

Characterisation of pathogenicity islands *in vivo* and *in vitro* and the *in vivo* virulence of *Pseudomonas aeruginosa*

Thesis submitted for the degree of
Doctor of Philosophy
University of Leicester

by

Melissa Elvira Koreen Carter BA (Hons)
Department of Infection, Immunity and Inflammation
University of Leicester

February 2009

Abstract

Pseudomonas aeruginosa is an opportunistic pathogen and usually targets immunocompromised patients such as burn victims and patients with AIDS, cancer or cystic fibrosis (CF). Acquisition of this organism is associated with high mortality and can cause death within 24 hours. The main themes covered within this thesis are pathogenicity island characterisation *in vitro* and *in vivo* as well as investigation of *in vivo* virulence of *P. aeruginosa*. The rationale behind this focus is that 10-20% of the *P. aeruginosa* genome is variable between strains and large variable regions such as genomic (pathogenicity) islands are considered more likely to contribute to the differences in disease-causing ability between strains.

The first project covers the development of a novel generic yeast-based genomic island capture method, which enables a complete genomic island to be present within a cloning vector. It was used in the characterisation of genomic islands in both *P. aeruginosa* and *Escherichia coli*. A novel genomic island in *E. coli* was captured and characterised.

The second project investigates the contribution of two pathogenicity islands, PAPI-1 and PAPI-2 to the *in vivo* virulence of *P. aeruginosa* PA14. Three pathogenicity island deletant isogenic mutants were tested for virulence in a murine acute respiratory model of infection developed for this project. The results showed that both pathogenicity islands contribute to virulence, but the presence of PAPI-2 is enough to maintain wild-type virulence.

The third project covers the exploration of the role of quorum-sensing in the virulence of *P. aeruginosa* LES; one of the transmissible epidemic strains and the most common strain recovered from CF patients across the UK. The project assessed whether over-expression of quorum-sensing products is a reliable indicator of increased virulence within a murine acute respiratory model of infection. The results showed in general, over-expressing mutants were more virulent than deficient mutants, but there was one exception, LESB58.

Statement of Originality

This accompanying thesis submitted for the degree of PhD entitled “Characterisation of pathogenicity islands *in vivo* and *in vitro* and the *in vivo* virulence of *Pseudomonas aeruginosa*” is based on work conducted by the author in the Department of Infection, Immunity and Inflammation at the University of Leicester mainly during the period between October 2005 and October 2008.

All the work recorded in this thesis is original unless otherwise acknowledged in the text or by references.

None of the work has been submitted for another degree in this or any other University.

Signed:_____

Date:_____

Acknowledgements

Dedicated completely to my loving husband Mark Haines

I would like to thank my supervisors Aras Kadioglu and Kumar Rajakumar. I also liked to thank those who I shared the good and bad days in the lab, including those in lab 136/212/227. Especially Sarah Glenn, Sarah Smeaton, Luke Richards, Saima Hussain, Hannah Brewin, Vanessa Terra, Luisa Crosatti, Barbara Rieck, Ewan Harrison, Jon van Aartsen, Eddie Cheah and James Lonnon. With an extra special thanks to Farah Shaikh, my extra supervisor and to Ken White 'because life upstairs would be a lot harder without you'.

Thanks for the love and support from my parents, my siblings and my best friend Rory O'Connor. Thanks to my other great friends who are also trying to achieve a PhD and have been a great support network, Sophie and Emily.

Abbreviations

%	Percent
°C	Degrees Centigrade
μg	Micrograms
μl	Microlitres
μM	Micromolar
μF	Microfaradays
ATP	Adenosine 5'-triphosphate
bp	Base pairs
DNA	Deoxyribonucleic acid
dNTPs	Deoxynucleosides
EDTA	Ethylenediamine tetraacetic acid
FCS	Fetal calf serum
g	Grams
<i>g</i>	Relative centrifugal force
GI	Genomic island
hr	Hours
kb	Kilobases
kV	Kilovolts
l	Litres
LA	Luria agar
LB	Luria broth
LD ₅₀	Lethal Dose, 50%
LT ₅₀	Lethal Time, 50%
M	Molar
mg	Milligrams
min	Minutes
ml	Millilitres
mm	Millimetres
mM	Millimolar
ng	Nanograms
NaCl	Sodium chloride
OD	Optical density

PAI	Pathogenicity island
PCR	Polymerase Chain Reaction
RNA	Ribonucleic acid
rpm	Revolutions per minute
RM	Restriction modification
sec	Seconds
TE	Tris-EDTA
tRIP	tRNA site interrogation for PAIs, prophages and other genomic islands
SDS	Sodium Dodecyl Sulfate
SGSP-PCR	Single genome-specific primer-PCR
TCS	Two component system
TSB	Tryptone soya broth
TTSS	Type three secretion system
v/v	Volume by volume
w/v	Weight by volume
X-gal	5-bromo-4-chloro-3-indolyl-b-D-galactopyranoside
YPD	Yeast extract-Dextrose and Peptone
Ω	Ohms

Table of contents

1	<u>INTRODUCTION</u>	1
1.1	<i>PSEUDOMONAS AERUGINOSA</i> AND ITS CLINICAL RELEVANCE	2
1.2	BACTERIAL GENOMES	4
1.2.1	PATHOGENCITY ISLANDS: A SUBGROUP OF GENOMIC ISLANDS	5
1.2.2	GENOMIC ISLANDS IN <i>PSEUDOMONAS AERUGINOSA</i>	6
1.2.3	BACTERIAL GENOME ANALYSIS	8
1.3	GENOMIC ISLAND CAPTURE USING YEAST RECOMBINATION TECHNOLOGY	10
1.3.1	TRANSFORMATION ASSOCIATED RECOMBINATION CLONING	10
1.3.2	PRINCIPLES OF YEAST GENOMIC ISLAND CAPTURE	12
1.3.3	AIMS	15
1.4	<i>IN VIVO</i> ANALYSIS OF PATHOGENCITY ISLAND DELETANT MUTANTS OF <i>PSEUDOMONAS AERUGINOSA</i> PA14	16
1.4.1	RESPIRATORY INFECTION MODELS	16
1.4.2	<i>PSEUDOMONAS AERUGINOSA</i> PA14	18
1.4.3	PAPI-1	20
1.4.4	PAPI-2	24
1.4.5	AIMS	25
1.5	<i>IN VIVO</i> ANALYSIS OF THE ROLE OF QUORUM-SENSING IN THE VIRULENCE OF <i>PSEUDOMONAS AERUGINOSA</i> LES	26
1.5.1	QUORUM-SENSING	26
1.5.2	<i>PSEUDOMONAS AERUGINOSA</i> LES	29
1.5.3	AIMS	30
1.6	PREAMBLE	30
2	<u>MATERIAL AND METHODS</u>	31
2.1	GENERAL INFORMATION	32
2.1.1	PREPARATION OF MEDIA	32
2.1.2	STRAINS USED	33
2.1.3	PREPARATION OF ELECTROCOMPETENT CELLS	35
2.1.4	PREPARATION OF CHEMICAL COMPETENT CELLS	36
2.1.5	BACTERIA PLASMID PREPARATION	36
2.1.6	YEAST PLASMID PREPARATION	36
2.1.7	RESTRICTION DIGESTION	37
2.1.8	GEL ELECTROPHORESIS	37
2.2	CONSTRUCTING A CAPTURE VECTOR	38
2.2.1	THE PRINCIPLES OF THE CREATION OF CAPTURE VECTOR	38
2.2.2	PCR PRIMER LIST	38
2.2.3	PRIMER DESIGN FOR GENOMIC WORK	40
2.2.4	AMPLIFICATION THE PLLX8 AMPLICON	41
2.2.5	AMPLIFICATION OF THE TARGETING SEQUENCES, TS1 AND TS2	41
2.2.6	LITHIUM ACETATE TRANSFORMATION	41
2.2.7	SELECTION OF CAPTURE VECTOR	42

2.3.1	CAPTURING A GENOMIC ISLAND	43
2.3.1	PREPARATION OF HIGH MOLECULAR WEIGHT GENOMIC DNA	43
2.3.2	PREPARATION OF SPHEROPLASTS	43
2.3.3	SCREENING OF SPHEROPLAST COLONIES: COLONY PCR	44
2.3.4	ELECTROPORATION: TRANSFER OF SUSPECT PLASMIDS TO <i>E. COLI</i>	45
2.3.5	RESTRICTION FRAGMENT LENGTH POLYMORPHISM (RFLP)	45
2.3.6	SUB-CLONING OF E106- <i>SERW</i>	45
2.4	PREPARATION OF PSEUDOMONAS AERUGINOSA FOR <i>IN VIVO</i> WORK	46
2.4.1	PREPARATION OF BACTERIAL STOCKS	46
2.4.2	INFECTION DOSE PREPARATION	47
2.5	<i>IN VIVO</i> MICE PROTOCOLS	47
2.5.1	ADMINISTRATION	47
2.5.2	ORGAN RETRIEVAL & CFU COUNTS	48
2.5.3	SYMPTOM SCORING	48
2.5.4	HISTOPATHOLOGY	48
2.5.5	LEUKOCYTE COUNTS	49
2.5.5	STATISTICS	50
3	<u>ANALYSIS OF GENOMIC ISLANDS CAPTURED USING YEAST-BASED TECHNOLOGY: THE GENOMIC ISLAND CAPTURE TECHNIQUE</u>	51
3.1	CAPTURE EXPERIMENTS INVOLVING GENOME SEQUENCED BACTERIAL STRAINS	53
3.1.1	K12 MG1655- <i>SERW</i>	54
3.1.2	K12 MG1655- <i>LEUX</i>	55
3.1.3	PAO1- <i>LYS10</i> AND PA14- <i>LYS10</i>	56
3.1.4	PA14- <i>LYS47</i>	56
3.2	CAPTURE EXPERIMENTS INVOLVING CLINICAL BACTERIAL STRAINS	58
3.2.1	KR115- <i>LYS10</i> AND KR159- <i>LYS10</i>	59
3.2.2	E105- <i>SERT</i>	59
3.3	E106-<i>SERW</i>	60
3.3.1	STRUCTURE OF THE E106- <i>SERW</i>	63
3.3.2	ANALYSIS OF E106- <i>SERW</i> CRISPR SYSTEM	67
3.4	E105-<i>LEUX</i>	70
3.4.1	DEFINING E105- <i>LEUX</i> AS A GENOMIC ISLAND (ISLET)	75
3.4.2	THE GENOMIC STRUCTURE OF THE E105- <i>LEUX</i> GENOMIC ISLET	78
3.4.3	EVOLUTION OF THE E105- <i>LEUX</i> GENOMIC ISLET	84
3.5	DISCUSSION	86
3.5.1	GENOMIC ISLAND CAPTURE AS A YEAST-BASED TECHNOLOGY	86
3.5.2	GENOMIC ISLAND CAPTURE AS A GENOMIC ISLAND DISCOVERY TOOL	90
3.6	CONCLUSION AND FURTHER WORK	92

<u>4 THE CONTRIBUTION OF PAPI-1 AND PAPI-2 TO THE VIRULENCE OF PSEUDOMONAS AERUGINOSA PA14 IN A MURINE ACUTE RESPIRATORY INFECTION MODEL</u>	94
4.1 CONSTRUCTION OF THE DELETANT MUTANTS AND <i>IN VITRO</i> DATA	95
4.2 PRELIMINARY STUDIES OF CHRONIC RESPIRATORY INFECTION MODEL	97
4.3 PRELIMINARY STUDIES OF ACUTE RESPIRATORY INFECTION MODEL	100
4.4 THE CONTRIBUTION OF PAPI-1 AND PAPI-2 TO THE VIRULENCE OF <i>PSEUDOMONAS AERUGINOSA</i> PA14	102
4.4.1 SURVIVAL TIMES POST-INFECTION	103
4.4.2 SYMPTOMS SCORES POST-INFECTION	103
4.4.3 BACTERIAL NUMBERS IN NASOPHARYNX, LUNGS AND BLOOD POST-INFECTION	104
4.4.4 INTRAVENOUS INFECTION	107
4.4.5 IMMUNE PROFILE	110
4.4.6 LUNG HISTOPATHOLOGY POST-INFECTION	112
4.4.7 PROGRESSION OF DISEASE POST-INFECTION WITH PA14 Δ PAPI-1 Δ PAPI-2	115
4.5 SUMMARY OF ACUTE RESPIRATORY OUTCOMES	118
4.6 DISCUSSION	119
4.6.1 THE ROLE OF PAPI-1 DURING ACUTE RESPIRATORY INFECTION	119
4.6.2 THE ROLE OF PAPI-2 DURING ACUTE RESPIRATORY INFECTION	124
4.6.3 PAO1 AS VIRULENT AS PA14 Δ PAPI-2 IN AN ACUTE RESPIRATORY MODEL	125
4.6.4 THE IMMUNE RESPONSE DURING ACUTE RESPIRATORY INFECTION	128
4.7 CONCLUSIONS AND FURTHER WORK	129
<u>5 THE ROLE OF QUORUM-SENSING IN THE VIRULENCE OF PSEUDOMONAS AERUGINOSA LES ISOLATES</u>	131
5.1 <i>IN VITRO</i> ANALYSIS OF THE QUORUM-SENSING EXPRESSION OF THE LES ISOLATES	132
5.2 ACUTE RESPIRATORY INFECTION OUTCOMES	135
5.2.1 SURVIVAL TIMES POST-INFECTION	135
5.2.2 SYMPTOMS SCORES POST-INFECTION	136
5.2.3 BACTERIAL NUMBERS WITHIN THE NASOPHARYNX, LUNGS AND BLOOD POST-INFECTION	140
5.2.4 LUNG HISTOPATHOLOGY POST-INFECTION	146
5.2.5 SUMMARY OF ACUTE RESPIRATORY INFECTION OUTCOMES	152
5.3 DISCUSSION	153
5.3.1 OVER-EXPRESSION OF QUORUM-SENSING PRODUCTS AS A MARKER FOR VIRULENCE IN AN ACUTE RESPIRATORY MODEL	154
5.3.2 THE DIFFERENCES IN VIRULENCE BETWEEN THE LES ISOLATES	156
5.3.3 THE LES ISOLATES VERSUS THE REFERENCE STRAINS, PAO1 & PA14	162
5.4 CONCLUSIONS AND FURTHER WORK	164
<u>6 THESIS CONCLUSION</u>	165

<u>7</u>	<u>APPENDIX</u>	<u>166</u>
7.1	RFLP COMPARISON BETWEEN APECO1-SERW AND E106-SERW	166
7.2	RESEARCH PUBLICATION	170
<u>8</u>	<u>BIBLIOGRAPHY</u>	<u>171</u>

1 Introduction

1 INTRODUCTION	1
1.1 <i>PSEUDOMONAS AERUGINOSA</i> AND ITS CLINICAL RELEVANCE	2
1.2 BACTERIAL GENOMES	4
1.2.1 PATHOGENICITY ISLANDS: A SUBGROUP OF GENOMIC ISLANDS	5
1.2.2 GENOMIC ISLANDS IN <i>PSEUDOMONAS AERUGINOSA</i>	6
1.2.3 BACTERIAL GENOME ANALYSIS	8
1.3 GENOMIC ISLAND CAPTURE USING YEAST RECOMBINATION TECHNOLOGY	10
1.3.1 TRANSFORMATION ASSOCIATED RECOMBINATION CLONING	10
1.3.2 PRINCIPLES OF YEAST GENOMIC ISLAND CAPTURE	12
1.3.3 AIMS	15
1.4 <i>IN VIVO</i> ANALYSIS OF PATHOGENICITY ISLAND DELETANT MUTANTS OF <i>PSEUDOMONAS AERUGINOSA</i> PA14	16
1.4.1 RESPIRATORY INFECTION MODELS	16
1.4.2 <i>PSEUDOMONAS AERUGINOSA</i> PA14	18
1.4.3 PAPI-1	20
1.4.4 PAPI-2	24
1.4.5 AIMS	25
1.5 <i>IN VIVO</i> ANALYSIS OF THE ROLE OF QUORUM-SENSING IN THE VIRULENCE OF <i>PSEUDOMONAS AERUGINOSA</i> LIVERPOOL EPIDEMIC STRAIN (LES)	26
1.5.1 QUORUM-SENSING	26
1.5.2 <i>PSEUDOMONAS AERUGINOSA</i> LES	29
1.5.3 AIMS	30
1.6 PREAMBLE	30

1.1 *Pseudomonas aeruginosa* and its clinical relevance

Pseudomonas aeruginosa is a Gram-negative, free-living, aerobic bacillus that is commonly isolated from soil and water. *P. aeruginosa* is an opportunistic pathogen to both humans and plants. It usually targets immunocompromised patients such as burn victims and patients with AIDS, cancer or cystic fibrosis (CF). It is one of the top three hospital-acquired pathogens (Spencer 1996) and is most commonly isolated from patients hospitalised for more than a week (Qarah S & Cunha BA). Infection can result in urinary tract infection, pneumonia and septicaemia. Acquisition of this organism is associated with high mortality and can cause fatality within 24 hours. The high mortality rate is partially explained by *P. aeruginosa* being intrinsically resistant to a broad range of antibiotics and its ability to rapidly acquire resistance genes via horizontal gene transfer. A restricted number of antibiotics are effective against *P. aeruginosa* these include fluoroquinolones, gentamicin and imipenem and susceptibility to these can vary between strains. Table 1-1 shows common virulence factors identified within *P. aeruginosa*.

	Virulence Factor	Role
<u>Cell-surface</u>	Pili	Adhesion to surfaces and motility (type IV)
	Flagella	Motility
	Type Three Secretion System (including toxins ExoS, ExoT, ExoU)	Translocates numerous effectors into host cells
	Alginate Biofilm	Adhesion to surfaces and defence against the environment
<u>Proteases</u>	LasA & LasB	Aid tissue invasion
	Alkaline Protease	
<u>Haemolysins</u>	Phospholipase C	Aid tissue invasion
	Rhamnolipid	
<u>Other factors</u>	Pyocyanin	Antimicrobial
	ExoA	Inhibits protein biosynthesis

Table 1-1 Summary of the common virulence factors within *P. aeruginosa*

The work described in this thesis investigates *P. aeruginosa* virulence in a murine respiratory model of infection. Therefore the focus of this introduction is to provide a background of *P. aeruginosa* as a respiratory pathogen. *P. aeruginosa* is associated with two types of respiratory infection within humans; acute and chronic. Acute infection usually occurs in hospitalised patients using a ventilator to aid breathing. *P. aeruginosa* is one of the leading causes of hospital-acquired pneumonia, causing significant morbidity and mortality. Ventilator-associated pneumonia caused by infection with *P. aeruginosa* is reported to have a high mortality rate of 70-80% (Chastre & Fagon 2002). Research into the virulence of *P. aeruginosa* in ventilator-associated pneumonia has focused on the type III secretion system (TTSS), which secretes four exotoxins known as ExoS, ExoT, ExoU and ExoY. ExoU is a potent effector of the TTSS and in ventilator-associated pneumonia it is correlated with more severe disease in terms of morbidity and mortality. It has also been reported that 40% of the isolates from these cases harbour

the *exoU* gene (Hauser *et al.* 2002). Clinical isolates from hospital-acquired pneumonia have shown a 4 fold lower 50% Lethal Dose (LD₅₀) of ExoU positive strains in comparison with ExoS positive strains (Schulert *et al.* 2003) when tested in a murine acute respiratory model.

P. aeruginosa is also associated with chronic infection, especially in patients with cystic fibrosis (CF). In CF patients, the leading cause of death is *P. aeruginosa* lung infection (Murray, Egan & Kazmierczak 2007). A recent phenomena noted within the CF community is the emergence of transmissible epidemic strains of *P. aeruginosa* (Scott, Pitt 2004). Prior to recent reports, CF patients were considered to acquire their own individual strain of *P. aeruginosa* from their local environment and not from other individuals. Research has shown that a strain of *P. aeruginosa* known as the Liverpool epidemic strain (LES) has become the most common isolate recovered from CF patients across the UK (Salunkhe *et al.* 2005, Scott & Pitt 2004). This strain has become a major focus of research as it has been reported to have been transmitted from a CF patient to both their non-CF parents, causing significant morbidity (McCallum *et al.* 2002). This was unusual as *P. aeruginosa* is considered an opportunistic pathogen and therefore unlikely to infect healthy individuals. There have also been reports of other epidemic strains of *P. aeruginosa*; listed in order of prevalent these are Midlands 1, Clone C and Manchester strains (Scott & Pitt 2004). These reports have prompted investigations into whether surveillance across UK CF centres is needed.

1.2 Bacterial genomes

This section proceeds to define what is meant by the terms genomic island and pathogenicity island. This is followed by a brief description of the large genomic islands identified in *P. aeruginosa*. The final subsection aims to give a brief overview of the use of bacterial genome sequences of pathogens to assess their virulence.

1.2.1 Pathogenicity islands: a subgroup of genomic islands

The genome of a bacterium consists of two components; the core genome and the mobile genome. The core genome is common to all bacteria within a species; it includes essential genes for the bacteria to prosper. In contrast the mobile genome, sometimes referred to as the mobilome (Ou *et al.* 2007), is defined as the parts of the genome that varies between strains within a species. The mobile genome consists of a range of genetic elements such as transposons, insertion sequences, plasmids, prophages and genomic islands. These elements are usually acquired via horizontal gene transfer (HGT) processes such as conjugation, transduction and transformation.

Clusters of genes within the mobile genome are known as genomic islands. Genomic islands encode gene products that can increase the fitness of a bacterium such as increasing host range, survivability in new environment and utilisation of new nutrients. There are a number of features used to define a genomic island that had been set out in review by Hacker and Kaper, 2000. The size of an island must exceed 10kb; a smaller insert is considered a genomic islet. The genomic island must have signs of being horizontally transferred. These include having a G+C content and codon usage that differ from the core genome and the presence of genes (or pseudogenes) that would enable the element to be transferred known as mobility genes. The genomic island is also usually flanked by direct repeats that were generated by the genomic island integrating into the core genome. These genomic islands are also often associated with the 3'-end of a tRNA gene. Some genomic islands have been shown to preferentially integrate at the 3'-end of a specific tRNA site.

Genomic islands that specifically encode virulence factors such as toxins that increase the ability of the bacterium to cause disease are termed pathogenicity islands. To be classified as a pathogenicity island, a genomic island must encode one or more virulence gene(s) as well having a higher incidence in pathogenic compared to non-pathogenic strains (Hacker & Kaper 2000). A schematic of the typical layout of a pathogenicity island is shown in Figure 1-1.

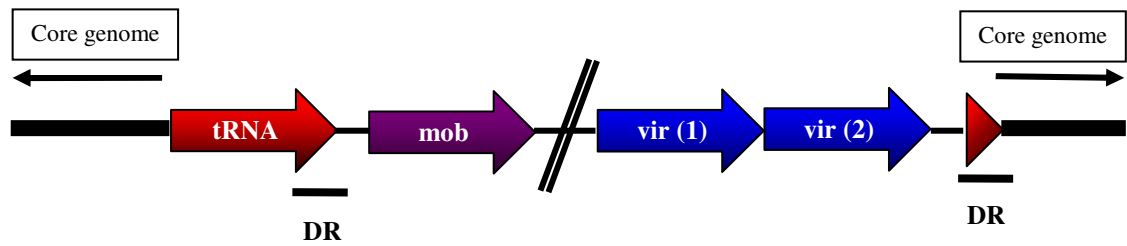


Figure 1-1 shows a schematic of the typical structure of a pathogenicity island

Key: mob (mobility genes), vir (virulence gene); DR (direct DNA sequence repeats)

1.2.2 Genomic islands in *Pseudomonas aeruginosa*

A number of large genomic islands have been identified within *P. aeruginosa* and are shown in Table 1-2. The abbreviation PAGI refers to *Pseudomonas aeruginosa* genomic island and PAPI refers to *Pseudomonas aeruginosa* pathogenicity island. All the large genomic islands are associated with tRNA sites with the exception of PAGI-1.

At present only PAPI-1 and PAPI-2 are explicitly named pathogenicity islands in the literature. *P. aeruginosa* is one of the more recently sequenced genomes, and there are only 4 complete genome sequences available, PAO1 (2001), PA14 (2005), PAS7 (2007) and LES (2008). It was reported that PAO1 had 44.2% of predicted ORFs in class 4 (Lee *et al.* 2006), which is defined as ORFs with unknown function. In order for genomic islands to be labelled pathogenicity islands they need to contain proven virulence genes, but if the functions of the genes present are unknown this is difficult.

The annotation process of *P. aeruginosa* genomes is still in its infancy, so tRNA sites are referenced by referring to the gene preceding the tRNA gene in the reference genome of *P. aeruginosa* PAO1. PAGI-1 is adjacent to the gene PA2217 and the full length island is found within isolate *P. aeruginosa* X24509. A number of analysed *P. aeruginosa* strains contain a truncated version of PAGI-1 (Salunkhe *et al.* 2005, Smart *et al.* 2006). The large genomic islands identified in the literature

appear to be related and there have been attempts to categorise them. Klockgether *et al*, 2007 defined two categories, PAGI-2-like islands and pKLC102-like islands. PAGI-2-like islands include PAGI-3; these islands are integrated into the glycine tRNA site that is adjacent to the PA2820. The pKLC102-like islands that include PAPI-1 are inserted in the lysine tRNA site adjacent to PA4541. A new family of genomic islands was discovered by Kulasekara *et al* (2006), called the ExoU islands. These genomic islands were discovered to integrate at the other lysine tRNA site within *P. aeruginosa*, adjacent to PA0976. They were demonstrated to be related to the previously identified islands within the literature; PAPI-2 and PAO1 genomic island associated with the same locus. PAPI-2 and the ExoU islands all contain the genes for the cytotoxin ExoU and its chaperone SpcU.

Genomic islands	Strain	Size	Insertion site	Function	Reference
PAPI-1	PA14	108kb	tRNA-Lys (PA4541)	Putative Type II effector, type IVB pilus, putative fimbrial chaperon-usher pathway cupD, two component regulatory systems (pvrR/PA14_59800, rcsB/rcsC), bacteriocin (pyocin S5)	(He <i>et al.</i> 2004)
PAPI-2	PA14	11kb	tRNA-Lys (PA0976)	Exotoxin and its chaperone (ExoU/SpcU)	(He <i>et al.</i> 2004)
PAGI-1	X24509	49kb	PA2217	Oxidative stress resistance	(Liang <i>et al.</i> 2001)
PAGI-2	C	158kb	tRNA-Gly (PA2830)	Transport of heavy metals and ions, cytochrome c biogenesis, mercury resistance	(Larbig <i>et al.</i> 2002)
PAGI-3	SG17M	128kb	tRNA-Gly (PA2830)	Transport of amino acids, coenzymes and porphyrins	(Larbig <i>et al.</i> 2002)
PAGI-4	C	34kb	tRNA-Lys (PA0976)		(Klockgether <i>et al.</i> 2004)
PAGI-5 to 11	Unknown	Various	-	-	Unpublished
Glycosylation	PAK		flgL/fliC	Flagellin glycosylation	(Arora <i>et al.</i> 2001)
ExoU island A	6077	81kb	tRNA-Lys (PA0976)	Exotoxin and its chaperone (ExoU/SpcU)	(Kulasekara <i>et al.</i> 2006)
ExoU island B	19660	30kb	tRNA-Lys (PA0976)	Exotoxin and its chaperone (ExoU/SpcU)	(Kulasekara <i>et al.</i> 2006)
ExoU island C	X13273	4kb	tRNA-Lys (PA0976)	Exotoxin and its chaperone (ExoU/SpcU)	(Kulasekara <i>et al.</i> 2006)

Table 1-2 lists the large genomic islands identified in *Pseudomonas aeruginosa*

PAGI = *Pseudomonas aeruginosa* genomic island and PAPI = *Pseudomonas aeruginosa* pathogenicity island

1.2.3 Bacterial genome analysis

In recent years, a new field of research called pathogenomics has developed, which aims to exploit genome-based knowledge as a rational for new vaccine and pharmacological targets. An example of the promise within the field is the development of new drugs that target drug-resistant *Mycobacterium tuberculosis* using microarray and proteomics data as a rational for design (Zhang, Amzel 2002). The branch of the field that will be explored within this introduction and thesis is depicted in Figure 1-2.

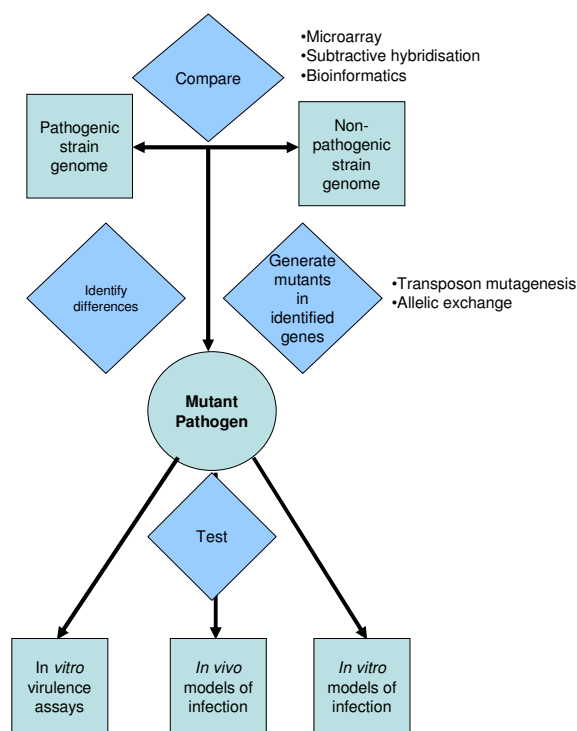


Figure 1-2 depicts the branch of pathogenomics explored within this thesis

The combination of high-throughput genome sequencing and comparing genome of strains using bioinformatics have highlighted areas of genomes that are variable between pathogenic and non-pathogenic strains. Bioinformatics tools are used to identify individual gene differences between pathogenic and non-pathogenic or less virulent strains. There has also been a trend towards analysing clusters of genes such as genomic (pathogenicity) islands that differ between strains using tools such as IslandFinder, PAIdatabase and MobilomeFinder. The individual genes are targeted for mutagenesis in the pathogenic strain and its effect on virulence is

assessed in number of assays and models of *in vivo* infection. There have been limited studies evaluating the effect on virulence of complete pathogenicity island within *P. aeruginosa*. Those reported made use of spontaneous pathogenicity island deletant mutants.

Pathogenomic techniques reported in the literature for *P. aeruginosa* at present do not strictly follow this model. As the complete genome of *P. aeruginosa* PAO1 only became available in 2001 and the other three completed strain genome sequences only became available in the last 3 years, the current research data available are based mainly on microarray and subtractive hybridisation, usually comparing the strains in question to the reference strain PAO1 (Choi *et al.* 2002, Smart *et al.* 2006). These genome differences are then targeted for further screening using targeted mutagenesis and virulence assays as well as in models of *in vivo* infection. In terms of *P. aeruginosa*, the current *in vivo* models of infection commonly used are based on infection of the nematode worm *Caenorhabditis elegans*, the plant *Arabidopsis thaliana*, the wax moth *Galleria mellonella* and the fruit fly *Drosophila melanogaster*. Murine infection models that replicate the human route of infection are less common (Choi *et al.* 2002, Rahme *et al.* 2000, Salunkhe *et al.* 2005, Tan, Mahajan-Miklos & Ausubel 1999).

To date, exploring virulence genes using these models has strictly used single gene knockouts. The work presented in this thesis has explored the role of whole genomic (pathogenicity) islands using *in vivo* models of infection. The focus of the project is a cluster of genes rather than single genes. The rationale behind this focus is that 10-20% of the *P. aeruginosa* genome is variable between strains. It could be suggested that these larger insertions of DNA such as genomic islands rather than single genes are more likely to contribute to the disease-causing ability of a specific strain. Genomic island genes were integrated as a unit into the core genome, been subjected to selection and have still been maintained as a unit. This warrants the study of genomic islands as a complete entity as further described in this thesis. A number of genomic islands have been identified in *P. aeruginosa* but a majority of the genes present within them have unknown function, performing studies such as those described in this thesis, will help define if any of these genomic islands are in fact pathogenicity islands.

The findings of recent pathogenicity island mutant studies have demonstrated the usefulness in determining their contribution to virulence. *E. coli* 536 contains five pathogenicity islands PAI I₅₃₆ – PAI V₅₃₆ and they were analysed for their contribution to *in vivo* virulence in a murine model of ascending urinary tract infection (Brzuszkiewicz *et al.* 2006). The authors reported the loss of PAI I₅₃₆ – PAI III₅₃₆ resulted in a reduction in virulence noted as an increase in LD₅₀ and therefore encode important virulence factors for urinary tract infection. There will always be a need for single genes knockout, but studies analysing complete pathogenicity island knockouts would be complementary.

1.3 Genomic island capture using yeast recombination technology

This section describes the genomic island capture technique. The technique is essential for the first project within this thesis, which aims to use the technique to capture and characterise genomic (pathogenicity) islands from *E. coli* and *P. aeruginosa*. It begins by providing the background of the technique followed by an outline of the principles and finishes with the aims and uses of the technique for the project reported in this thesis.

1.3.1 Transformation associated recombination cloning

Saccharomyces cerevisiae (baker's yeast) possesses an endogenous, efficient homologous recombination system. In molecular biology, this process is used to recombine DNA fragments with as little as 20bp of homology (Mezard *et al.* 1992). The process is known as transformation associated recombination (TAR) cloning. TAR cloning has been utilised in the laboratory for a variety of applications such as:

- Chromosomal gene replacement within yeast
- Creation of artificial chromosomes and yeast plasmids
- Insertion of genes into shuttle vectors
- ORF libraries
- Collections of mutant genes
- Producing tagged proteins

TAR cloning is a method designed by Larionov *et al* (1996) to exploit the yeast recombination system. TAR cloning involves designing a vector, containing two targeting sequences flanking the region of interest within the genome. Successful homologous recombination, after transformation of the yeast cells with the genomic DNA and the vector, results in specific cloning of the region of interest. The method was developed using eukaryotic DNA from mice and humans (Larionov *et al.* 1996).

When discussing the genomic island capture technique, the term capture vector refers to a shuttle vector which can replicate in both bacteria and yeast. Capturing refers to the process of successful homologous recombination between the capture vector and the genomic DNA. This results in the genomic island being present in the capture vector.

In terms of the history of the genomic island capture technique, the current application was piloted by Raymond *et al* (2002) when they first used this technique to capture variable genomic regions (O-antigen LPS region) from *P. aeruginosa* and were also the first to report using bacterial DNA in the capture process. The technique has also used to capture the pyoverdine region, another variable genomic region within *P. aeruginosa* (Smith *et al.* 2005). Wolfgang *et al* (2003) adapted this technique further by modifying the structure of the capture vector and used the method to target a variable region within the *P. aeruginosa* genome. They highlighted its use as a technique to capture genomic islands by constructing and using a capture vector to target the *lys10* tRNA site (p0975-0989capture) and captured genomic islands from a number of *P. aeruginosa* strains. The aim of the project described in this thesis is develop a generic genomic capture technique by further developing the published method and reporting the first use of the technique with bacterial species other than *P. aeruginosa*. The method described in this thesis streamlines the design and creation of capture vectors as well as providing an alternative protocol for effective screening of transformants. The project encompasses *E. coli* genomic island capture as well as *P. aeruginosa*. The technique has previously only been reported with *P. aeruginosa* and expansion to other bacterial species would make it a valuable molecular biology tool. In addition,

previous research data from within my research group (Thani A, University of Leicester, PhD Thesis) suggested there were two novel genomic islands to be captured within *E. coli*.

1.3.2 Principles of yeast genomic island capture

TAR cloning requires carefully constructed strains of yeast. In this thesis, the yeast strain used, *S. cerevisiae* CRY1-2, is uracil prototroph. This means the strain can not synthesis uracil and therefore it needs it to be present within the growth media. URA medium refers to a growth medium that is deficient in uracil and therefore the untransformed strain can not grow on this medium. *S. cerevisiae* CRY1-2 is also intrinsically resistant to cycloheximide, which is normally toxic to eukaryotic cells. Introduction of the dominant wild-type allele of the *cyh* gene into the yeast cell will render it susceptible to the toxic effects of cycloheximide. The principles described are shown in Figure 1-3 and summarised in Table 1-3. Three different colonies of *S. cerevisiae* CRY1-2 are shown. The first is an untransformed colony, as shown this can grow on a complete medium (YPD), it can also grow on complete medium supplemented with cycloheximide as it is resistant, but unable to grow on the URA medium. The second colony is transformed with a capture vector, this contains the dominant allele of the *cyh* gene, making the cells susceptible to the drug, but the plasmid contains the *ura* gene, so the strain can now grow on URA medium. The third colony is transformed with pLLX13, a plasmid that contains the *ura* gene, this enables the cell to grow on URA media and cycloheximide containing medium.

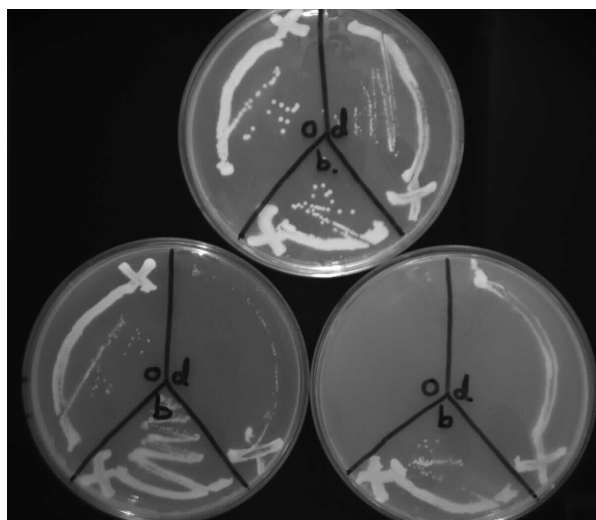


Figure 1-3 Top plate shows colony growth on YPD (non-selective medium); bottom left shows colony growth on YPD containing CYH (1µg/ml) media and bottom right shows colony growth on URA-deficient medium

	YPD	CYH	URA
<i>S. cerevisiae</i> CRY1-2	✓	✓	✗
+ capture vector (d)	✓	✗	✓
+ pLLX13 (b)	✓	✓	✓

Table 1-3 Summary of colonies growth on various media as depicted in Figure 1-3

The capture vectors are created from 4 DNA components (Figure 1-4). The first two are universal to all capture vectors. Linearised pLLX13 is used to provide all the genes essential for the capture vector to be maintained as a plasmid in both bacteria and yeast. The second component is the pLLX8 amplicon, which contains the dominant allele of the *cyh* gene and *bla* (ampicillin resistance) gene. The additional two components (TS1 and TS2) determine the target region of the recombination between the capture vector and the genomic DNA. The TS1 and TS2 correspond to a 1 kb region that can be found upstream and downstream of the tRNA site respectively. The TS1, TS2 and pLLX8 amplicon are all created by PCR and have a 40bp overlap at the ends to complement one another and pLLX13 (Figure 1-4). This provides a 40bp overlap which has been shown in the literature (Oldenburg *et al.* 1997) to be optimal length of homology for efficient recombination.

The *Xba*I-linearised pLLX13 plus the three amplicons TS1, TS2 and pLLX8 are co-transformed into lithium acetate treated *S. cerevisiae*. Lithium acetate transformation is based on the principle that alkali cations produce yeast cells that

are competent for DNA uptake. This results in homologous recombination at the site fixed by the amplicons within the yeast cell as shown in Figure 1-4. The transformed yeast cells are then plated onto URA medium; this does not differentiate between cells containing re-circularised pLLX13 and those containing the complete capture vector as explained in Figure 1-3.

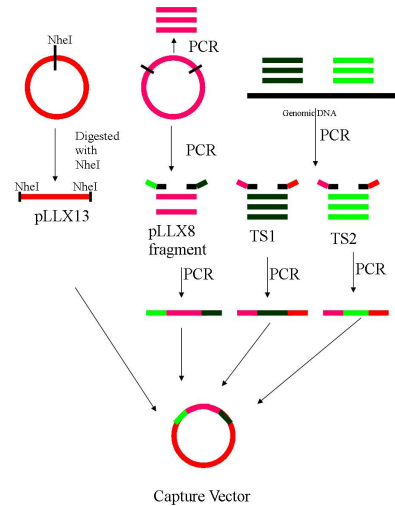


Figure 1-4 Schematic depicting how the four DNA components recombine to generate the complete capture vector

The capture vector is transferred to *E. coli*, its structure is verified and a DNA preparation is created. For genomic island capture to occur, the yeast cell wall is partially digested with a specialised enzyme known as Zymolyase (Zymo Research Corporation, USA), to allow exogenous DNA to enter. A yeast cell with a partially digested cell wall is known as a spheroplast. TAR cloning is stimulated by transforming the capture vector that has been linearised to stimulate DNA damage along with the genomic DNA of choice. A successful outcome results in the yeast ‘repairing’ the capture vector via homologous recombination between the vector and the co-transformed genomic DNA (Figure 1-5). This would result in the capture vector containing the targeted genomic island.

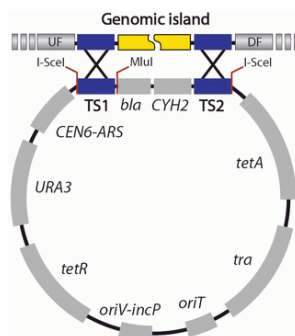


Figure 1-5 depicts the principles of genomic island capture method – Reproduced from the MobilomeFinder website (<http://mml.sjtu.edu.cn/MobilomeFINDER/ycv.htm>)

To increase the likelihood of recovery of a capture vector with a genomic island inserted, URA media is supplemented with cycloheximide. Any yeast cells that contain the capture vector only, will be rendered susceptible to cycloheximide, which is toxic to the cells, due to the dominant allele present on the capture vector. Successful recombination between the capture vector and the genomic DNA results in the loss of the *cyh* gene from the capture vector; therefore if the capture vector undergoes re-circularisation instead of homologous recombination, the presence of the cycloheximide within the media should eliminate them.

The resulting colonies are screened to ensure the correct recombination event has occurred by PCR. The capture vector: genomic island is then transferred to bacteria for further analysis. This capture vector plus genomic island can now be used for functional studies as well as sequencing projects. This molecular biology tool is very powerful as the targeting regions will be conserved among the majority of the members of the same species. This allows for comparative studies of these regions.

1.3.3 Aims

The aim of the first project was to create a generic genomic island capture technique to aid discovery and characterisation of novel genomic islands. To achieve this aim, the project will further develop the genomic island capture technique as described by Wolfgang *et al* (2003) by testing the parameters and capturing from bacterial species other than *P. aeruginosa*. The method described in this thesis streamlines the design and creation of capture vectors as well as providing an alternative protocol for effective screening of transformants. The project encompasses *E. coli*

genomic island capture as well as *P. aeruginosa*. The technique has previously only been reported with *P. aeruginosa* and expansion to other bacterial species would make it a valuable molecular biology tool. The project will also aim to characterise novel genomic islands captured using this technique.

1.4 In vivo analysis of pathogenicity island deletant mutants of *Pseudomonas aeruginosa* PA14

This section describes the respiratory models currently used to study *P. aeruginosa* virulence and provides a literature background to the strains used in this thesis project and finishes with the aims of the project.

1.4.1 Respiratory infection models

There two types of respiratory infection caused by *P. aeruginosa*; acute and chronic. In a clinical setting, there are two major patient types to reflect these two types of infection; those that have extended periods of ventilation (acute) and those who suffer from cystic fibrosis (chronic).

Through the use of acute murine respiratory models, a number of key virulence factors have been identified, using single gene knockout mutants of *P. aeruginosa*. The most studied of which is the TTSS proteins of which there are four, ExoS, ExoT, ExoU and ExoY. ExoS, ExoT and ExoU have been previously shown to contribute to mortality and morbidity in a murine acute respiratory model, demonstrated by intranasal infection of adult Balb/c females (Shaver & Hauser 2004). There is also a correlation in human clinical outcomes based on the presence or absence of components of the TTSS (Hauser *et al.* 2002). The production of pyocyanin has also been shown to cause significant damage to lungs during murine acute respiratory infection, demonstrated by intranasal infection of adult CD-1 mice (Lau *et al.* 2004). Pyocyanin is a blue pigment excreted by *P. aeruginosa* that has been shown to have anti-bacterial properties against other bacteria. The genes within the two *P. aeruginosa* quorum-sensing systems; LasR (Tang *et al.* 1996), LasI (Pearson *et al.* 2000) and RhlI (Pearson *et al.* 2000) has also been described as contributing to acute infection, all demonstrated by intranasal infection of neonatal

Balb/c mice. The models described by Allewelt *et al.* 2000, Laskowski, Kazmierczak 2006, Shaver, Hauser 2004 all use intranasal infection of adult female Balb/c mice, with analysis of bacterial loads in lungs, spleen and liver determined at 16-18 hours post-infection. The project described in this thesis will determine bacterial numbers in the lower respiratory tract and for the first time the upper respiratory tract by recovering nasopharyngeal tissue. The project will also investigate bacterial dissemination from the lungs into blood following respiratory challenge.

Current modelling of chronic lung infection with mice is primarily achieved through the use of agarose beads laden with *P. aeruginosa* (van Heeckeren & Schluchter 2002). Infection with free *P. aeruginosa* into the lungs results either in acute infection or clearance in healthy mice (van Heeckeren & Schluchter 2002). The use of chronic infection model is principally to mimic disease patterns of infection of CF patients with *P. aeruginosa* in conjunction with the use of CF mice. Chronic infection with *P. aeruginosa* was first achieved in rats by infecting intratracheally with agarose beads laden with *P. aeruginosa* (Cash *et al.* 1979). The model was adapted by Starke *et al* (1987) for use in mice. The mouse model has also been modified to model CF infection by using mucoid strains of *P. aeruginosa* as it closely mimics pathogenesis seen in lungs of CF patients. The mucoid phenotype of *P. aeruginosa* is due to over-production of alginate and is found in 70–80% of CF patients (van Heeckeren & Schluchter 2002). This phenotypic change correlates with lung-tissue destruction and pulmonary decline. A second model to study chronic infection in relation to biofilms was developed by Yanagihara *et al* (2000). This model involves placing a plastic tube within the bronchus that is pre-coated with *P. aeruginosa*. They used the model to evaluate the effectiveness of two antibiotics on the biofilms created during chronic infection of *P. aeruginosa*. Chronic infection with *P. aeruginosa* in transgenic CF mice has been achieved by infecting the drinking water with *P. aeruginosa* (Coleman *et al.* 2003). This model first shows chronic infection of the oropharyngeal with subsequent translocation to the lungs. Importantly, mucoid isolates of *P. aeruginosa* were recovered from the lungs of the CF mice.

To reflect the two patient types, the project outlined in this thesis will attempt to develop two models of respiratory infection (acute and chronic) to explore the role of pathogenicity islands in virulence of *P. aeruginosa*.

1.4.2 *Pseudomonas aeruginosa* PA14

Pseudomonas aeruginosa PA14 (UCBPP-PA14) is a highly virulent wild-type reference strain that has been used extensively to test the contribution of virulence factors to disease. The *P. aeruginosa* strain PAO1, another wild-type reference strain considered the laboratory strain of *P. aeruginosa*, has been shown to be less virulent than *P. aeruginosa* PA14 in a number of models of infection (He *et al.* 2004, Rahme *et al.* 2000, Tan, Mahajan-Miklos & Ausubel 1999). It has been shown in mouse and plant infection models to be less virulent than *P. aeruginosa* PA14 despite having the same host specificity (He *et al.* 2004). In a mouse burns-sepsis model, where a thermal injury is inflicted and *P. aeruginosa* is applied to the wound, PAO1 causes 25% mortality compared with 77% mortality caused by PA14 (Rahme *et al.* 2000). In the leaf infiltration model, where the leaf is infected with *P. aeruginosa* and the bacterial load at 5 days post-infection is determined, PAO1 was recovered at 8.0×10^5 cfu/cm² of leaf area in comparison to 2.6×10^7 cfu/cm² in PA14 (Rahme *et al.* 2000). The difference in virulence is also supported by the *C. elegans* model of infection, where the LT₅₀ in this slow killing model was 37.8 hours for PA14 and 68.2 hours for PAO1 (Tan, Mahajan-Miklos & Ausubel 1999).

This difference in virulence has prompted studies to determine the genomic differences between the two strains. There have been two previous studies to analyse the genomic difference between PAO1 and PA14. The first study was in 2002 and it used representational difference analysis (RDA), to compare the genome of these two strains to look for specific differences (Choi *et al.* 2002). This method led to the discovery of 4 genes that differed significantly between the strains. One of the genes highlighted is found within PA14 pathogenicity island PAPI-1, *uvrD*. The authors also discovered a gene that was a homologue of *ybtQ*, an ABC transporter found in *Yersinia pestis*. They produced a PA14 mutant in *ybtQ* and tested its virulence in a number of infection models, and it attenuated PA14 to levels comparable with PAO1 in both the wax moth *Galleria mellonella* and mouse

burns-sepsis model, but not in the nematode worm *Caenorhabditis elegans* model. The second study in 2006 involved sequencing the genome of PA14 and directly analysed the two genomes together (Lee *et al.* 2006). Their aim was to discover if the difference in virulence between PAO1 and PA14 was due to the presence of pathogenicity islands present in PA14 but absent in PAO1. They discovered that PA14 has a slightly larger genome than PAO1, 6.5Mb versus 6.3Mb and that PA14 has 58 additional gene clusters than PAO1. They tried to analyse these 58 clusters in two ways. Firstly, they screened 20 *P. aeruginosa* strains (including PA14 and PAO1) and compared their virulence in a *C. elegans* model of infection and used this data to rank the strains. They then proceeded to analyse the strains for the presence of PA14-specific gene regions with respect to PAO1. The results demonstrated no correlation between virulence and the presence of PA14-specific gene regions. A second approach taken was to create a transposon insertion mutant library for PA14. They screened the library using the *C. elegans* model of infection for those in the PA14-specific gene regions for attenuation. They discovered 9 genes that caused attenuation. The presence of these regions within the other *P. aeruginosa* strains, showed no set trend as to presence of these PA14-specific regions and the ranked virulence. This led the group to conclude that the presence of these gene regions within another *P. aeruginosa* strain was not an indicator of its virulence within the *C. elegans* model of infection. This study did not take into account the individual contribution to virulence each gene cluster makes and therefore the ranking of virulence may not give a true reflection of PA14 virulence.

Despite the extensive use of plants and invertebrate models of infection for *P. aeruginosa*, in order to evaluate how these virulence factors affect and interact in human infection they must be tested in a mammalian model of infection. The host immune response is more complex in mammals and therefore a murine model of infection such as described in this thesis will provide a more valuable insight into human disease.

A recent study (He *et al.* 2004) suggested that differences in virulence between PAO1 and PA14 could be due to the presence of two pathogenicity islands found in PA14 that are absent in PAO1, named *Pseudomonas aeruginosa* pathogenicity island 1 (PAPI-1) and *Pseudomonas aeruginosa* pathogenicity island 2 (PAPI-2).

The authors generated transposon mutants and created in-frame deletions within PAPI-1 and demonstrated a reduction in mortality compared to wild-type PA14 infected mice within a number of ORFs using a mouse burns sepsis model. The work presented in this thesis aims to directly assess the contribution of PAPI-1 and PAPI-2 to the virulence of *P. aeruginosa* PA14 within a murine respiratory model by using complete pathogenicity island deletant mutants.

1.4.3 PAPI-1

PAPI-1 is a 108kb pathogenicity island with 108 ORFs that is associated with the lysine tRNA adjacent to PA4541 and is a member of the pKLC102 island family (Klockgether *et al.* 2007). It contains a number of virulence factors, including a type IV B pili, pyocin S5, cupD system, as well as PvrR, a two component response regulator involved in biofilm synthesis. The large majority of encoded proteins within this island have no assigned function. It contains many features that are exhibited by genomic islands (He *et al.* 2004): (1) associated with a tRNA site (2) G+C content different from that of the core genome (59.7% compared with 66.6% of core genome) (3) bound by direct repeats (4) containing mobility gene i.e. integrases and transposases. An interesting property of PAPI-1 is its ability to excise from the genome and be maintained as a plasmid (Qiu, Gurkar & Lory 2006). Table 1-4 contains all the ORFs within PAPI-1 that have been assigned a function.

ORF ID	Gene name	Gene function
PA14_60140	<i>xerC</i>	Integrase
PA14_60120	<i>dcd_2</i>	Deoxycytidine deaminase
PA14_60100	<i>dtd</i>	Deoxycytidine triphosphate deaminase
PA14_60050	<i>parE</i>	Plasmid stabilization protein parE
PA14_59960	<i>dsbG</i>	Putative protein-disulfide isomerase
PA14_59790	<i>pvrR</i>	Regulator of two-component regulatory system; adhesion and antibiotic resistance
PA14_59780	<i>rscC</i>	Sensor of two-component regulatory system
PA14_59770	<i>rscB</i>	Regulator of two-component regulatory system
PA14_59760	<i>cupD5</i>	Probable pili assembly chaperone / adhesion and protein secretion
PA14_59750	<i>cupD4</i>	Adhesion and protein secretion
PA14_59740	<i>cupD3</i>	Probable fimbrial biogenesis usher / adhesion and protein secretion
PA14_59730	<i>cupD2</i>	Probable pili assembly chaperone / adhesion and protein secretion
PA14_59720	<i>cupD1</i>	Probable fimbrial precursor / adhesion and protein secretion
PA14_59390	<i>pilM2</i>	Type IV B pilus / adhesion and and protein secretion
PA14_59380	<i>pilV2</i>	Type IV B pilus / adhesion and and protein secretion
PA14_59370	<i>pilT2</i>	Type IV B pilus / putative peptidase / adhesion and and protein secretion
PA14_59360	<i>pilS2</i>	Type IV B pilus / adhesion and and protein secretion
PA14_59350	<i>pilR2</i>	Type IV B pilus / adhesion and and protein secretion
PA14_59340	<i>pilQ2</i>	ATPase / Type IV B pilus / adhesion and and protein secretion
PA14_59330	<i>pilP2</i>	Type IV B pilus / adhesion and and protein secretion
PA14_59320	<i>pilO2</i>	Type IV B pilus / adhesion and and protein secretion
PA14_59310	<i>pilN2</i>	Secretin / Type IV B pilus / adhesion and and protein secretion
PA14_59300	<i>pilL2</i>	Type IV B pilus / adhesion and and protein secretion
PA14_59290	Unassigned	Colicin immunity protein
PA14_59280	Unassigned	Colicin-like toxin (pyocin S5)
PA14_59270	Unassigned	DNA Helicase
PA14_59240	<i>topA</i>	Topoisomerase I
PA14_59210	<i>ssb</i>	Single-stranded DNA binding protein
PA14_59070	<i>dnaB</i>	DNA replication and recombination
PA14_59060	Unassigned	Putative phage protein
PA14_59010	<i>soj</i>	Chromosome partitioning

Table 1-4 Annotation of all ORFs that have assigned a function within the pathogenicity island, PAPI-1 in *Pseudomonas aeruginosa* PA14 (He *et al.* 2004) – Reproduced with modification by permission from PNAS 2004 101(8): 2530-2535. Copyright 2004 National Academy of Sciences, U.S.A.

Based on the literature available the loss of PAPI-1 from PA14 should result in a reduction in virulence (Choi *et al.* 2002, He *et al.* 2004). The remainder of this section describes genes that have been previously highlighted in the literature to be involved in a reduction in virulence and their ORFs have been assigned a function. To date 65% of the ORFs within PAPI-1 have no known function and are labelled

as hypothetical proteins (He *et al.* 2004). It is important to note that none of the ORFs have been previously investigated using *in vivo* respiratory infection models. The work presented in this thesis is the first reported analysis of the contribution of these ORFs to virulence within an acute respiratory model.

There are two, two-component regulatory systems (TCS) within PAPI-1 that have been highlighted through previous studies as important for general virulence of the strain (He *et al.* 2004, Lee *et al.* 2006). Two component regulatory systems act to change the gene expression within a cell to favour changes in the external environment. They consist of two parts; one is the sensor kinase, which senses the changes in the environment by being bound to the cellular membrane and two; the response regulator, which mediates the intracellular response to the changes in environment. These external changes manifest in response to changes in temperature, pH, presence of antibiotics and changes in nutrient availability.

The TCS involving RcsC and RcsB described in the literature is involved in the adaptation of the cell surface by changing its structure and integrity (Huang, Ferrieres & Clarke 2006). This TCS has been identified in a number of *Enterobacteriae* species, in which the system has the ability to activate capsule synthesis, suppress flagella biosynthesis and mobility as well as encouraging biofilm synthesis (Huang, Ferrieres & Clarke 2006). Literature suggests it is involved in a phenotypic change from mobility and attachment to cells during infection, to that of creating a stable biofilm for a more permanent environment. Interestingly, in *Yersinia* species, RcsB has been shown to be a positive regulator of Type III secretion system (TTSS) secretion (Huang, Ferrieres & Clarke 2006). In *E. coli* there are three components involved in this system, RcsC, the sensor kinase, which activates RcsD, an intermediate, which in turn activates RcsB, the response regulator. As of yet, no RcsD homologue has been discovered in *P. aeruginosa* PA14. Importantly, the *Yersinia* genome does not encode a functional RcsD, which suggests that the TCS has taken on a different role (Huang, Ferrieres & Clarke 2006). This is also the case in PA14 as mutagenesis of either RcsB or RcsC results in attenuation in a murine burns-sepsis model of infection (He *et al.* 2004), if they were redundant parts left over, mutagenesis should have no effect. In the circumstances that the *rsc* TCS system could still be involved in some of its usual

functional activities it could be suggested that loss of RcsB or RcsC from the pathway would result in cells that have a reduced ability to create a viable capsule and biofilms. The implication being the bacteria will be vulnerable to the immune system as the capsule can act to protect the bacteria from the host immune system.

The second TCS within PAPI-1 is PvrR (response regulator) and PA14_59800 (sensor kinase). Mutagenesis of either of these genes resulted in attenuation in a murine burns-sepsis model (He *et al.* 2004). Research by Drenkard *et al.*, 2002, which led to the discovery of PvrR in *P. aeruginosa* PA14 was based on a general observation that exposing *P. aeruginosa* to antibiotics on solid media resulted in small colony variants (SCV). These SCV produced biofilm in larger amounts and more rapidly than the 'normal' colonies as well being resistance to a large diversity of antibiotics. When these SCV were transferred to solid media that lacked antibiotics they could revert to 'normal' colonies. Drenkard *et al.*, 2002 suggested these results pointed to a change in phenotype and discovered the two component system response regulator, PvrR. The important outcome of their research was to show that PvrR acts to switch the phenotype from SCV to 'normal' and does not affect the opposite transition (Drenkard & Ausubel 2002). Another protein, WspR appears to control this function (Haussler 2004). Interestingly, the phenotype change is controlled by expression of *cup* genes. In *P. aeruginosa* PAO1, three gene clusters that encode chaperone-usher pathway components have been found known as cupA, cupB and cupC. The chaperone-usher pathway is conserved among Gram-negative bacteria for assembling fimbriae on the cell surface. In *P. aeruginosa* PA14, annotation by He *et al.* (2004) suggests that a cupD pathway exists within PAPI-1 and immediate adjacent to the PvrR TCS within the pathogenicity island.

Another protein encoded on PAPI-1 is UvrD, a DNA helicase deletion of this gene within PA14 (not present in PAO1) resulted in attenuation in the *G. mellonella* model of infection, but not notably in a *C. elegans* or mouse burns infection model (Choi *et al.* 2002). This protein has been shown to be important in *E. coli* as deletion of this gene as well as *rep* is lethal to the bacterial cell (Lestini & Michel 2008). The proposed function of UvrD is to aid nucleotide excision and mismatch repair and has been shown experimentally to be vital in *rep* mutants to offset deleterious RecA binding (Lestini, Michel 2008). The remaining protein encoded on

PAPI-1 which has been identified as being involved in attenuation is PA14_59150, a single stranded DNA binding protein. Proteins in this group function to protect ssDNA from damage, preventing mismatched annealing and to form complexes with other enzymes (Lu & Keck 2008).

1.4.4 PAPI-2

PAPI-2, an 11kb island with 15 ORFs, is a member of the family of ExoU islands (Kulasekara *et al.* 2006) and is associated with the lysine tRNA site adjacent to PA0976 (Figure 1-6). It contains a number of mobility genes, in addition to three ORFs with assigned function; ExoU, SpcU and PA14_51560. ExoU is a potent cytotoxin and SpcU is its chaperone and PA14_51560 is an acetyltransferase. ExoU mutants have been created within a number of different *P. aeruginosa* strains and have been shown in murine respiratory infection models to cause a reduction in virulence based on survival and recovered CFU at set timepoints as well as milder lung pathology (Allewelt *et al.* 2000, Schulert *et al.* 2003, Shaver & Hauser 2004). Allewelt *et al.* (2000) reported a clear difference in LD₅₀ for a panel of *P. aeruginosa* strains between those expressing ExoU and those strains that did not. Their paper also demonstrated that introducing the *exoU/spcU* genes into non-carrying *P. aeruginosa* increased its virulence as a respiratory pathogen; this was demonstrated in PAO1 by lowering the LD₅₀. Lee *et al.* (2005) also introduced ExoU into PAK and noted an increase in virulence within a murine acute respiratory model based on decreased survival times of the mice. The loss of ExoU from the *P. aeruginosa* strain, PA99, was also shown to be significantly attenuated in a murine acute respiratory model (Shaver, Hauser 2004). It has also been experimentally shown to contribute to the severity of human infection. In ventilator-associated pneumonia it is associated with more severe disease in terms of morbidity and mortality and 40% of the isolates from these cases harbour the *exoU* gene (Hauser *et al.* 2002). A 4 fold lower LD₅₀ of ExoU positive strains in comparison with ExoS positive strains was reported when a panel of clinical isolates from hospital-acquired pneumonia were tested in a murine acute respiratory model (Schulert *et al.* 2003).

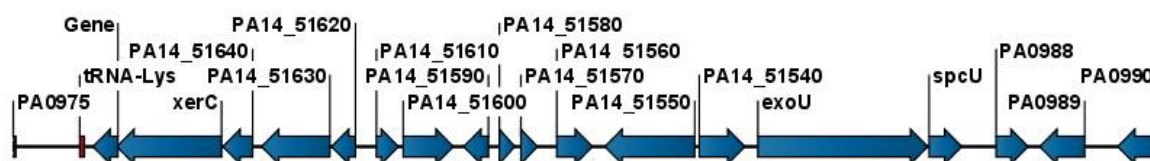


Figure 1-6 Schematic showing the ORFs found within PAPI-2

ORF ID	Gene name	Gene function
PA14_51650	xerC	Integrase
PA14_51630	Unassigned	Putative transposase
PA14_51620	Unassigned	Putative transposase
PA14_51560	Unassigned	Acetyltransferase
PA14_51550	Unassigned	Transposase
PA14_51540	Unassigned	Putative transposase
PA14_51530	exoU	Cytotoxin
PA14_51520	spcU	ExoU chaperone

Table 1-5 Annotation of all ORFs that have an assigned function located within the pathogenicity island PAPI-2 in *Pseudomonas aeruginosa* PA14 (He *et al.* 2004) – Reproduced with modification by permission from PNAS 2004 101(8): 2530-2535. Copyright 2004 National Academy of Sciences, U.S.A.

1.4.5 Aims

The overall aim of the second project is to compare the virulence of isogenic pathogenicity island deletant mutants using *in vivo* murine respiratory models of infection. To achieve this aim, the project will intend to develop both an acute and a chronic respiratory infection model to reproduce both infection types associated with *P. aeruginosa* infection. Pathogenicity islands are considered a distinguishing feature that separate non-pathogenic strains from pathogenic strains. The study of virulence within *P. aeruginosa* by removing pathogenicity islands has yet to be explored; based on published literature this is the first reported study to use pathogenicity island deletant mutants in models of respiratory infection. Research by He *et al* (2004) had previously suggested that increased virulence of *P. aeruginosa* PA14 over *P. aeruginosa* PAO1 was due to the presence of PAPI-1 and

PAPI-2. This project aims to investigate the contribution of PAPI-1 and PAPI-2 to the virulence of *P. aeruginosa* PA14 in murine respiratory infection models. To achieve this aim, a set of isogenic mutants will be used which consists of a mutant with PAPI-1 deleted, another with PAPI-2 deleted and a third mutant with both PAPI-1 and PAPI-2 deleted.

1.5 *In vivo* analysis of the role of quorum-sensing in the virulence of *Pseudomonas aeruginosa* Liverpool Epidemic Strain (LES)

This section will discuss quorum-sensing and its role in virulence within *P. aeruginosa*, followed by focussing on *P. aeruginosa* Liverpool Epidemic Strain (LES), a CF isolate that shows differential expression of the quorum-sensing system and finishes by explaining the aims of this project.

1.5.1 Quorum-sensing

Quorum-sensing is a control system intended to co-ordinate gene expression based on bacterial numbers. The system is controlled by secretion and detection of molecules known as autoinducers. When the level of an autoinducer reaches a threshold level, they bind to specific bacterial receptors, which then act to initiate gene transcription. The outcome of which is that the majority of the bacteria within a population now express the same phenotype. There are two quorum-sensing systems within *P. aeruginosa*, the Rhl system and the Las system. Figure 1-7 shows a simple schematic of the two quorum-sensing systems. Each system is controlled by two major components. First, the autoinducer synthase which synthesises the autoinducer for that specific system for example RhlI produces *N*-butyryl-L-homoserine lactone (C4-HSL). The second component is the receptor and transcriptional regulator which acts within the bacterial cell to activate the expression of the genes involved within that regulon.

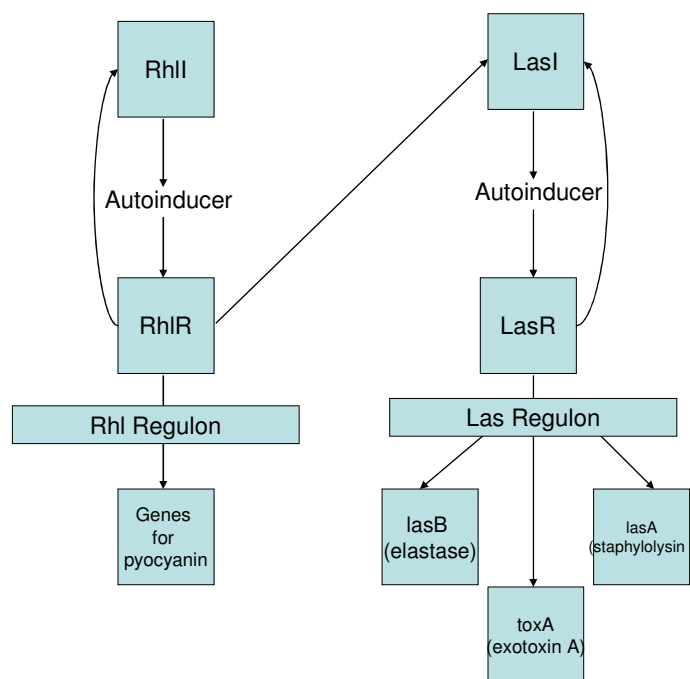


Figure 1-7 Schematic showing the two quorum-sensing systems within *Pseudomonas aeruginosa*; the genes used as examples have been highlighted in the published literature to be involved in *P. aeruginosa* virulence.

In *P. aeruginosa*, the expression of many virulent determinants is controlled by the quorum-sensing systems such as the secretion of toxins and the production of biofilm (Duan & Surette 2007). Previously published work has shown that a deficiency in the quorum-sensing systems can reduce virulence in a number of acute murine infection models. The use of quorum-sensing deficient mutants in a murine burns injury model led to a reduction in mortality of the mice and the ability of the bacteria to disseminate from the site of infection (Rumbaugh *et al.* 1999). In a murine interperitoneal infection model it was noted that clearance of bacteria was more rapid in quorum-sensing deficient mutants (Christensen *et al.* 2007). There have also been two previous studies (Pearson *et al.* 2000, Tang *et al.* 1996) that have shown in a neonatal murine acute respiratory model that the loss of quorum-sensing related genes results in a reduction in morbidity and mortality. These models involved inoculating 7 to 10 day old Balb/c mice with either 5×10^6 CFU (Tang *et al.* 1996) or 1×10^9 CFU (Pearson *et al.* 2000) bacteria per mice. Pearson *et al.* (2000) constructed a PAO1 mutant which lacked two quorum-sensing associated genes (*lasI* and *rhlI*), termed PAO_JP2. The mutant and the wild-type, PAO1 were tested in the neonatal respiratory model until 18 hours post-infection. The mutant had an increased survival rate (21% versus 5%), as well a reduced chance of

developing pneumonia (56% versus 10%) and bacteraemia (~50% versus ~15%). Tang *et al.*, 1996 performed experiments with a LasR mutant of PAO1 which resulted in similar findings. The authors also performed additional *in vitro* experiments with the LasR mutant, whereby a bacterial adherence assay on nasal polyp cells showed that the LasR mutant had 26% binding in contrast to wild-type PAO1 which had 58%. The authors suggested that the difference in binding was due to LasR-dependent exoproducts, which aid attachment to the cell surface. The previous published research described above has focussed on evaluating quorum-sensing deficient mutants; for this project I will evaluate how quorum-sensing *over-expression* affects virulence within a murine acute respiratory model.

The literature suggests that CF isolates of *P. aeruginosa* undergo changes in their genome to rid themselves of acute virulence determinants (D'Argenio *et al.* 2007, Smith *et al.* 2006). One of the most common genetic changes in CF isolates is the loss of LasR function (D'Argenio *et al.* 2007, Smith *et al.* 2006). A number of other acute infection models have shown that isolates deficient in quorum-sensing (including *lasR* mutants) are less virulent than their wild-type counterparts (Christensen *et al.* 2007, Pearson *et al.* 2000, Rumbaugh *et al.* 1999, Tang *et al.* 1996). It is suggested that this is due to selective pressures within the CF airway to select against acute virulence determinants and therefore allowing the bacteria to evade the host immune system. The current literature suggests that to be a successful chronic CF isolate, the isolate must lose its ability to be a successful acute infection isolate. The advantages of *lasR* mutants have been investigated. D'Argenio *et al.*, 2007 showed that the LasR mutants had an advantage over wild-type, for growth on selected carbon and nitrogen sources. Their findings showed that the loss of LasR also led to an increase in antibiotic resistance, noted as an increase in beta-lactamase activity. In this context, the loss of function of LasR would reduce the production of many virulence factors in a single mutation, in contrast to acquiring many small mutations in each of the virulence genes (Nguyen, Singh 2006).

1.5.2 *Pseudomonas aeruginosa* LES

The Liverpool epidemic strain (LES) is a lineage of *P. aeruginosa* clinical CF isolates. It is the most common strain of *P. aeruginosa* in the UK found among CF patients (Scott & Pitt 2004, Salunkhe *et al.* 2005). LES attracted research attention when identified as the first epidemic CF strain of *P. aeruginosa* in the UK (McCallum *et al.* 2001). It has also been reported to replace the existing *P. aeruginosa* strain infecting a CF patient (superinfection) (McCallum *et al.* 2001) and is also associated with higher morbidity in CF patients than other CF strains of *P. aeruginosa* (Al-Aloul *et al.* 2004). It was considered important enough to be one of the four *P. aeruginosa* to be sequenced to date. At present there has been limited research exploring LES isolates in *in vivo* models of infection. In a *Drosophila* fly infection model (Salunkhe *et al.* 2005), both LES431 and LES400 were shown to be more virulent than PAO1. In the *C. elegans* model of infection (Winstanley C, University of Liverpool, unpublished), all the LES isolates had a lower LT₅₀ than PAO1. In a chronic infection model within rats (Kukavica-Ibrulj *et al.* 2008), the growth curve of the bacteria within the lungs were shown to be similar for LESB58, PAO1 and PA14.

LES isolates have been shown to vary in their quorum-sensing ability (Fothergill *et al.* 2007). Research has shown those isolates that were defined as over-expressing quorum-sensing products were more virulent in a *C. elegans* slow killing assay (Winstanley C, University of Liverpool, unpublished). The project outlined in this thesis intends to add to the limited available data, by exploring the quorum-sensing phenotype and its implication on *in vivo* virulence within an acute respiratory model. This project aims to analyse the upper and lower respiratory tract in mice. This thesis will be the first reported analysis of the upper and lower respiratory tract within a *P. aeruginosa* murine acute respiratory model.

Four naturally occurring mutant isolates of LES were chosen based on availability of published literature, LES400, LES431, LESB58 and LESB65. LES400 is a quorum-sensing deficient mutant that was isolated from a CF patient. LES431 is an over-expressing quorum-sensing mutant, isolated from a pneumonia case from a non-CF parent of a CF child. A very unusual case and still today the only described

case of its kind (McCallum *et al.* 2002). LESB58 is an over-expressing quorum-sensing mutant from a CF patient; one of first isolated LES isolates and the genome sequenced isolate that represents the Liverpool Epidemic Strain. LESB65 is also an over-expressing quorum-sensing isolate that had came from a CF patient.

1.5.3 Aims

LES isolates have been shown to vary in their quorum-sensing ability (Fothergill *et al.* 2007). Research using a *C. elegans* slow killing assay has shown those isolates that were defined as over-expressing quorum-sensing products were more virulent (Winstanley C, University of Liverpool, unpublished). The aim of these studies was to determine in a relevant upper and lower respiratory tract infection model whether the over-expression of quorum-sensing products is a reliable marker for increased virulence exhibited by LES isolates. To analyse this view, a number of isolates displaying either a quorum-sensing deficient phenotype or an over-expressing quorum sensing phenotype were tested for virulence in an acute respiratory model. The over-expressing mutants are represented by LES431, LESB58 and LESB65. The deficient mutants are represented by LES400. This study is the first reported case of using LES isolates in an acute respiratory model.

1.6 Preamble

This thesis explores the role of pathogenicity islands and quorum-sensing in a murine respiratory model of infection as well as isolating pathogenicity islands using the genomic island capture technique. There are three results chapters and they cover the three projects outlined in this introduction. The first results chapter covers the development of a generic genomic island capture method and its uses in characterisation of novel genomic islands. The second results chapter covers the contribution of the two pathogenicity islands, PAPI-1 and PAPI-2, to the virulence of *P. aeruginosa* PA14 in murine respiratory models of infection. The third chapter results chapter covers the exploration of the role of quorum-sensing in virulence of *P. aeruginosa* LES in above mentioned *in vivo* model.

2 Material and Methods

2 MATERIAL AND METHODS 31

GENERAL INFORMATION	32
2.1.1 PREPARATION OF MEDIA	32
2.1.2 STRAINS USED	33
2.1.3 PREPARATION OF ELECTROCOMPETENT CELLS	35
2.1.4 PREPARATION OF CHEMICAL COMPETENT CELLS	36
2.1.5 BACTERIA PLASMID PREPARATION	36
2.1.6 YEAST PLASMID PREPARATION	36
2.1.7 RESTRICTION DIGESTION	37
2.1.8 GEL ELECTROPHORESIS	37
2.2 CONSTRUCTING A CAPTURE VECTOR	38
2.2.1 THE PRINCIPLES OF THE CREATION OF CAPTURE VECTOR	38
2.2.2 PCR PRIMER LIST	38
2.2.3 PRIMER DESIGN FOR GENOMIC WORK	40
2.2.4 AMPLIFICATION THE PLLX8 AMPLICON	41
2.2.5 AMPLIFICATION OF THE TARGETING SEQUENCES, TS1 AND TS2	41
2.2.6 LITHIUM ACETATE TRANSFORMATION	41
2.2.7 SELECTION OF CAPTURE VECTOR	42
2.3 CAPTURING A GENOMIC ISLAND	43
2.3.1 PREPARATION OF HIGH MOLECULAR WEIGHT GENOMIC DNA	43
2.3.2 PREPARATION OF SPHEROPLASTS	43
2.3.3 SCREENING OF SPHEROPLAST COLONIES: COLONY PCR	44
2.3.4 ELECTROPORATION: TRANSFER OF SUSPECT PLASMIDS TO <i>E. COLI</i>	45
2.3.5 RESTRICTION FRAGMENT LENGTH POLYMORPHISM (RFLP)	45
2.3.6 SUB-CLONING OF E106-SERW	45
2.4 PREPARATION OF <i>PSEUDOMONAS AERUGINOSA</i> FOR <i>IN VIVO</i> WORK	46
2.4.1 PREPARATION OF BACTERIAL STOCKS	46
2.4.2 INFECTION DOSE PREPARATION	47
2.5 <i>IN VIVO</i> MICE PROTOCOLS	47
2.5.1 ADMINISTRATION	47
2.5.2 ORGAN RETRIEVAL & CFU COUNTS	48
2.5.3 SYMPTOM SCORING	48
2.5.4 HISTOPATHOLOGY	48
2.5.5 LEUKOCYTE COUNTS	49
2.5.6 STATISTICS	50

General information

2.1.1 Preparation of media

Yeast media

Media	
Yeast extract/peptone/dextrose (YPD)	Yeast extract 1% (w/v) Bacto-peptone 2% (w/v) Dextrose 2% (w/v) Adenine 0.05% (w/v)
YPD agar	YPD plus Agar 2% (w/v)
Uracil-deficient media (URA) (MP Biomedicals Limited)	YNB 1.7g Ammonium sulphate 5g Dextrose 20g CSM-URA. 0.77g
URA Agar	URA plus Agar 2% (w/v)
0.5% Adenine	Distilled Water 475ml Adenine 2.5g 1M HCL 25ml

Table 2-1 Yeast media used during this study

Bacteria media

Media	
LB	Tryptone 1% (w/v) Yeast extract 0.5% (w/v) NaCl 86.2mM
LA	LB plus Agar 1.5% (w/v)
Tryptone soya broth (TSB)	Tryptone soya broth 3% (w/v)

Table 2-2 Bacteria media used during study

2.1.2 Strains used

Pseudomonas aeruginosa

Laboratory strain	Description	Source
PAO1 (KR124)	Reference laboratory strain of <i>P. aeruginosa</i>	Lab 212 Dept. of Infection, Immunity and Inflammation, University of Leicester
PA14 (KR125)	Reference laboratory strain; shown to be more virulent than <i>P. aeruginosa</i> PAO1	Lab 212 Dept. of Infection, Immunity and Inflammation, University of Leicester
KR115	Clinical isolate from Leicester Royal Infirmary	Lab 212 Dept. of Infection, Immunity and Inflammation, University of Leicester
KR159	Clinical isolate from Leicester Royal Infirmary	Lab 212 Dept. of Infection, Immunity and Inflammation, University of Leicester
LES400	Clinical respiratory isolate from CF patient	Craig Winstanley, Division of Medical Microbiology, University of Liverpool (Salunkhe <i>et al.</i> 2005)
LES431	Clinical respiratory isolate from non-CF patient	Craig Winstanley Division of Medical Microbiology, University of Liverpool (Salunkhe <i>et al.</i> 2005)
LESB58	Clinical respiratory isolate from CF patient	Craig Winstanley, Division of Medical Microbiology, University of Liverpool, Liverpool (Kukavica-Ibrulj I <i>et al.</i> 2008)
LESB65	Clinical respiratory isolate from CF patient	Craig Winstanley, Division of Medical Microbiology, University of Liverpool, Liverpool (Fothergill <i>et al.</i> 2007)

Table 2-3 shows the *Pseudomonas aeruginosa* strains used during this study

Escherichia coli

Laboratory strain	Description	Source
MG1655 K12 (E101)	Reference strain of <i>E. coli</i>	Lab 212 Dept. of Infection, Immunity and Inflammation, University of Leicester
E105	Clinical blood isolate from Leicester Royal Infirmary	Lab 212 Dept. of Infection, Immunity and Inflammation, University of Leicester
E106	Clinical blood isolate from Leicester Royal Infirmary	Lab 212 Dept. of Infection, Immunity and Inflammation, University of Leicester
DH10B	Laboratory strain of <i>E. coli</i> used for transformation of plasmids	Lab 212 Dept. of Infection, Immunity and Inflammation, University of Leicester
K12: <i>leuX</i> (KR507)		This study
E105: <i>leuX</i> (KR510)		This study
K12: <i>serW</i> (KR506)		This study
E106: <i>serW</i>		This study
PA01: <i>lys10</i> (KR508)		This study
PA14: <i>lys10</i> (KR509)		This study
KR115: <i>lys10</i> (KR559)		This study
KR159: <i>lys10</i> (KR560)		This study

Table 2-4 shows the *Escherichia coli* strains used and created during this study

Saccharomyces cerevisiae

Laboratory strain	Description	Source
<i>CRY1-2</i>	Mating type (<i>MATα</i>), cycloheximide resistant (<i>cyh2^R</i>) and uracil prototrophic (<i>ura3Δ</i>)	Stephen Lory Dept. of Microbiology and Molecular Genetics Harvard Medical School (Wolfgang <i>et al</i> 2003)

Table 2-5 shows the *Saccharomyces cerevisiae* strain used during this study

Plasmids

Plasmid	Description	Reference
pLLX13	(9.9kb) provides the backbone for the capture vector. It contains the <i>tetAR</i> genes that confer tetracycline resistance (10µg/ml) in <i>E. coli</i> . It has the <i>URA3</i> gene that will allow <i>S. cerevisiae</i> <i>CRY1-2</i> to grow on uracil –deficient media. The plasmid also contains genes to allow replication in both types of organisms. <i>CEN6</i> and <i>ARSH4</i> allow replication and segregation of the plasmid within yeast. There are two origin of replication; <i>oriV</i> for vegetative replication in <i>E. coli</i> and <i>oriT</i> which acts as an origin of transfer.	Stephen Lory Dept. of Microbiology and Molecular Genetics Harvard Medical School (Wolfgang <i>et al</i> 2003)
pLLX8	(4.6kb) is used to produce an amplicon of 2.9kb that contain the <i>cyh2</i> , which confers sensitivity to cycloheximide (2.5 µg/ml) in <i>S. cerevisiae</i> <i>CRY 1-2</i> and <i>bla</i> (b-lactamase) confer resistance to carbenicillin (50µg/ml) in <i>E. coli</i> .	Stephen Lory Dept. of Microbiology and Molecular Genetics Harvard Medical School (Wolfgang <i>et al</i> 2003)
pWSK29	Low copy vector used for the sub-cloning of E106- <i>serW</i> genomic island	Lab 212 Dept. of Infection, Immunity and Inflammation, University of Leicester (Wang, Kushner 1991)
<i>lys10</i> capture vector	Capture vector that targets the 10 th tRNA for lysine in <i>Pseudomonas aeruginosa</i> PAO1; adjacent to PA0976	This study
<i>lys47</i> capture vector	Capture vector that targets the 47 th tRNA for lysine in <i>Pseudomonas aeruginosa</i> PAO1; adjacent to PA4541	This study
<i>asnV</i> capture vector	Capture vector that targets <i>asnV</i> in <i>Escherichia coli</i> K12 MG1655	This study
<i>leuX</i> capture vector	Capture vector that targets <i>leuX</i> in <i>Escherichia coli</i> K12 MG1655	This study
<i>serT</i> capture vector	Capture vector that targets <i>serT</i> in <i>Escherichia coli</i> K12 MG1655	This study
<i>serW</i> capture vector	Capture vector that targets <i>serW</i> in <i>Escherichia coli</i> K12 MG1655	This study

Table 2-6 shows plasmids used and created during this study

2.1.3 Preparation of electrocompetent cells

A 5ml culture of *E. coli* DH10B was grown overnight 200rpm, 37°C in LB. The overnight culture was added to 495ml of LB and incubated at 200rpm, 37°C for approximately 2-3 hours until culture reached an OD₆₀₀ of 0.5. The culture was

transferred to a pre-chilled falcon tube and incubated on ice for 20 minutes. The culture was then centrifuged at 1,500g, 4°C for 15 minutes and the supernatant was discarded. The pellet was washed three times in 10% (v/v) ice-cold glycerol. After the final wash, the pellet was re-suspended in 5ml of 10% (v/v) ice-cold glycerol. The suspension was then separated into 100µl aliquots in pre-chilled microcentrifuge tubes. The aliquots were snap frozen in an ethanol/dry ice bath and stored at -70°C.

2.1.4 Preparation of chemical competent cells

A 5ml culture of *E. coli* DH10B was grown overnight 200rpm, 37°C in LB. 500µl of the overnight culture was added to 400ml of LB and incubated at 200rpm, 37°C until culture reached an OD₆₀₀ of 0.3. The culture was centrifuged at 1,500g, 10 minutes, 4°C. The supernatant was discarded and the pellet was washed once in 50ml 100mM magnesium chloride. The cells were then re-suspended in 50ml 100mM calcium chloride and incubated on ice for 10 minutes and then centrifuged at 1,500g, 4°C for 15 minutes. After the final wash, the pellet was re-suspended in 4ml of 100mM calcium chloride/15% glycerol solution. The suspension was then separated into 50µl aliquots in pre-chilled microcentrifuge tubes and stored at -70°C.

2.1.5 Bacteria plasmid preparation

All plasmid preparations were performed using the QIAprep Spin Mini-prep kit (Qiagen, UK) as described in the manual for mini-preps. For larger plasmid preparations, HiSpeed Plasmid Midi Kit (Qiagen, UK) or HiSpeed Plasmid Maxi Kit (Qiagen, UK) was used. The protocols followed were as stated by the manufacturer.

2.1.6 Yeast plasmid preparation

All plasmid preparations were performed using the Zymoprep™ Yeast Plasmid Mini-prep I kit (Zymoresearch, USA). The following modifications were made to the manufacturer's protocol, each preparation used 5ml of overnight yeast culture and an additional wash was performed in 1ml 70% ethanol after the stated isopropanol wash and the final re-suspension was in 10µl nanopure water.

In the case of capture vector construction, 1ml of nanopure water was added to a single plate and scraped with a sterile spreader to produce a suspension. 100µl of the suspension was added to a microcentrifuge tube and then the protocol provided with the kit was followed with the modification as stated above.

2.1.7 Restriction digestion

All restriction enzyme were performed using enzymes purchased from New England Biolabs (NEB, UK), using the Manufacturer's instructions. All restriction digests were incubated at 37°C overnight.

Digestion of pLLX13 was performed in a final volume of 20µl digestion using 200ng of pLLX13 and 2U of *Mlu*I enzyme for 3hrs. Digestion of capture vectors with *Mlu*I were performed by using a complete Qiagen Midi plasmid preparation in a final volume of 200µl, using 3µl of enzyme. Digestion of a capture vector with insert was generally performed using 5µl of Qiagen plasmid preparation in a final digestion volume of 30µl, using 0.5µl of enzyme.

2.1.8 Gel electrophoresis

Gel electrophoresis was used to estimate size and quantity of DNA. Agarose gels were produced using an appropriate w/v of agarose (Multipurpose agarose, Bioline, UK) in TAE buffer (40mM Tris-base, Glacial acetic acid, 1mM EDTA, pH 7). The mixture was heated until homogenised, cooled to 56°C. Ethidium bromide was added to a final concentration of 0.5 µg/ml and poured into an appropriate size gel tray. Two different DNA ladders were generally used; λ *Hind*III (Fermentas, UK) and GeneRuler® DNA ladder mix (Fermentas, UK), at a concentration of 500ng/6µl.

0.6% (w/v) agarose gels were used for capture clone RFLP digestion and quantifying capture vector after linearisation and 1% (w/v) agarose gels were used for colony PCR, otherwise a 0.8% (w/v) agarose gel was used.

2.2 Constructing a capture vector

2.2.1 The principles of the creation of capture vector

To create the capture vector, the lithium-acetate treated yeast cells are transformed with four fragments; the two 1kb targeting region amplicons (TS1 and TS2), the 2.9kb amplicon from pLLX8 and linearised pLLX13. The amplicons were created using tailed primers that carry homology to the other amplicons and pLLX13. The targeting regions PCR resulted in the targeting region amplicons having one flank with homology to pLLX13 and the other flank with homology to the pLLX8 amplicon. The pLLX8 amplicon was created with one flank homologous to TS1 and the other to TS2. The principles are depicted in Figure 2-1.

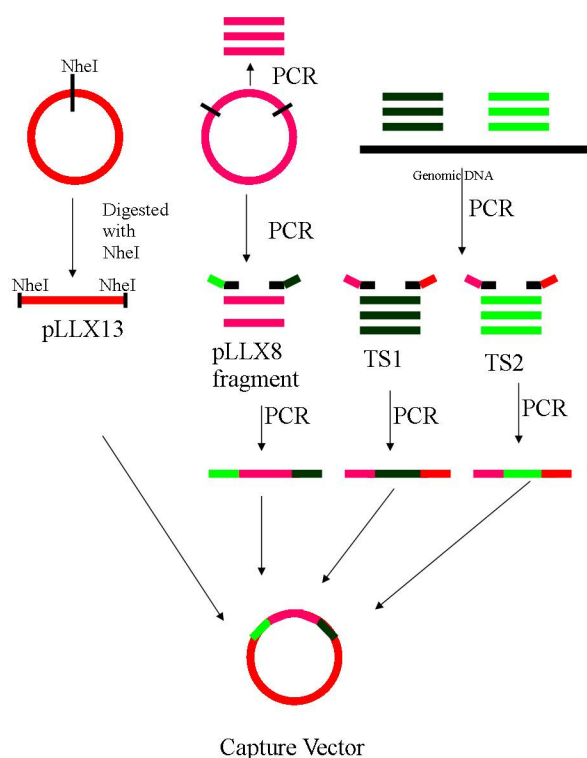


Figure 2-1 Schematic depicting how the four DNA components recombine to generate the complete capture vector

2.2.2 PCR primer list

All primers were synthesised by VH Bio Limited, Gateshead and had a T_m 60°C. Please note the pLLX8 primers were universal for the creation of all capture vectors.

Primer name		Sequence
Capture vector primers		
LeuX	LeuX TS1 U	5'-ATATTACCCTGTTATCCCTAGCGTAACTATCGATCTCGAG GCTGGATGGCGTCGTTAAT
	LeuX TS1 D	5'- GTTTACTCATATATACTTTAGATT ACGCGT AAAGCTATGCCGAAAAGAA
	LeuX TS2 U	5'- CTATCACCCATTTTCACCGTTTTTT AGTCAAATTCGCCTTGCATC
	LeuX TS2 D	5'- TAACAGGGTAATATAGAGATCTGGTACCCTGCAGGAGCTC ATAAGCTTGCTCCCGTCTCA
	LeuX Pos U	5'- GTGTTTCGTCAGTCCTCTTTC
	LeuX Pos D	5'-ATCGGGTGAATTATCGCAAG
SerW	SerW TS1 U	5'- ATATTACCCTGTTATCCCTAGCGTAACTATCGATCTCGAG CCGGATAAAGCCATTGATGT
	SerW TS1 D	5'- GTTTACTCATATATACTTTAGATT ACGCGT GTGCTTTTGTGACTCTGGA
	SerW TS2 U	5'- CTATCACCCATTTTCACCGTTTTTT GAATGGCGAGAAAGCTGAAC
	SerW TS2 D	5'- TAACAGGGTAATATAGAGATCTGGTACCCTGCAGGAGCTC GCGTTAATTGCCCTGTTTGT
	SerW Pos U	5'-TACCGCCATGACCGATAAAT
	SerW Pos D	5'-GTTTCAGCTTTTCTCGCCATTC
Lys10	Lys10 TS1 U	5'-ATATTACCCTGTTATCCCTAGCGTAACTATCGATCTCGAG ATGGTTTCAGCTCGGCTGATCTTC
	Lys10 TS1 D	5'-CATATATACTTTAGATTTTAATTAA ACGCGT TCTAGAAAATATCACCTTGTGCTCGATCACGC
	Lys10 TS2 U	5'-CATTTTCACCGTTTTTTGTTTAAACGTTAACTCTAGAGGG ATGCCGACCCCGGTTTTTCATAAGAGA
	Lys10 TS2 D	5'-TAACAGGGTAATATAGAGATCTGGTACCCTGCAGGAGCTC CCGGCGACTACTGGATCGTCTACGG
	Lys10 Pos U	5'-GGCTCGACCTCAATGGCATGGGCG
	Lys10 Pos D	5'-TCAGAAAATATGGCGTCGGGTCGGA
Lys47	Lys47 TS2 U	5'-ATATTACCCTGTTATCCCTAGCGTAACTATCGATCTCGAG ACATTCTCCTCCAGGTGCTC
	Lys47 TS2 D	5'-CATATATACTTTAGATTTTAATTAAACGCGTTCTAGAAAA GGTGGTCGGGGTAGAGAGAT
	Lys47 TS1 U	5'-CATTTTCACCGTTTTTTGTTTAAACGTTAACTCTAGAGGG AGCGGATCGAGGTATTTTCA
	Lys47 TS1 D	5'-TAACAGGGTAATATAGAGATCTGGTACCCTGCAGGAGCTC ACGCGT GAGGACTGGTCGGATACAGC
	Lys47 Pos U	5'- CACAGGTTGTAGTGGCGACCT
	Lys47 Pos D	5'- CGGTGAAAATACCTCGATCC
SerT	SerT TS1 U	5'-ATATTACCCTGTTATCCCTAGCGTAACTATCGATCTCGAG TTCTCTGCAATCACCACAGC
	SerT TS1 D	5'-CATATATACTTTAGATTTTAATTAA ACGCGT TCTAGAAAA TCCGCTGTAAACCTGTTC
	SerT TS2 U	5'-CATTTTCACCGTTTTTTGTTTAAACGTTAACTCTAGAGGG TTCGTTTTCAGCATCTCACG
	SerT TS2 D	5'-TAACAGGGTAATATAGAGATCTGGTACCCTGCAGGAGCTC CCAGACTGGATTGGGAAAGA
	SerT Pos U	5'-TAGCAACCGCTAAACATCC
	SerT Pos D	5'-AGCGGAGCGGATAAAAGATT
	SerT E105 Pos U	5'-GGGTGAAAACGCATATTGCT
	SerT E105 Pos D	5'-AGAGCTGGAAAAGCTGACCA
AsnV	AsnV TS1 U	5'-ATATTACCCTGTTATCCCTAGCGTAACTATCGATCTCGAG CAATGTCGCAATTTCTGTG
	AsnV TS1 D	5'-CATATATACTTTAGATTTTAATTAA ACGCGT TCTAGAAAA GGCGGTGATTATGAGCATT
	AsnV TS2 U	5'-CATTTTCACCGTTTTTTGTTTAAACGTTAACTCTAGAGGG CGCGAAAGCTATCCGTTAAG
	AsnV TS2 D	5'-TAACAGGGTAATATAGAGATCTGGTACCCTGCAGGAGCTC GTGCACAACGGCATATTGAC
	AsnV Pos U	5'- CAGTAAACGACTGCCCCACT
	AsnV Pos D	5'- CAGATGTCTCCTGCGATCAA
pLLX8	pLLX8 U	5'-TTTCTAGAACGCGTTTAATTAA AATCTAAAGTATATATGAGTAAAC
	pLLX8 D	5'-CCCTCTAGAGTTAACGTTTAAAC AAAAAACGGTGAAAATGGGTGATAG
Others	CYH2 U	5'-AGCACAGAGGTCACGTCTCA
	CYH2 D	5'-ATTCAACAACACCACCAGCA
	TetA U	5'-GCTACATCCTGCTTGCCTTC
	TetA D	5'-ATAGATCGCCGTGAAGAGGA
	T7	5'-GTAATCGACTCACTATAGGGC
	T3	5'-ATTAACCCTCACTAAAGGGA

Table 2-7 shows a list of all primers used in this study

2.2.3 Primer design for genomic work

The DNA sequence for the conserved flanks found upstream and downstream of the tRNA site were generated using the ExtractFlank program from Mobilome Finder (Ou *et al* 2006). These 2kb conserved flanks were used as a template to design the targeting regions for the capture vector. The primer3 program (Rozen S 2000) was used to create the targeting regions between 950bp-1050bp. The chosen primers were used in *in-silico* PCR against the reference genome to generate the amplicon DNA sequence using *insilico.ehu.es* (Bikandi *et al* 2004). For *E. coli*, the reference genome used was K12 MG1655 and the *P. aeruginosa* reference genome was PAO1. The amplicon DNA sequence was subjected to *in-silico* digestion using NEBCutter (Vincze, Posfai & Roberts 2003) to check for the presence of a *MluI* restriction site. If the restriction site was present the primer pair was discarded. A further consideration for the selection of the targeting sequences is to ensure that either TS1 or TS2 contains a partial gene. The residual part of the gene that is not present within the targeting sequence can be used to design the screening primers as shown in Figure 2-2.

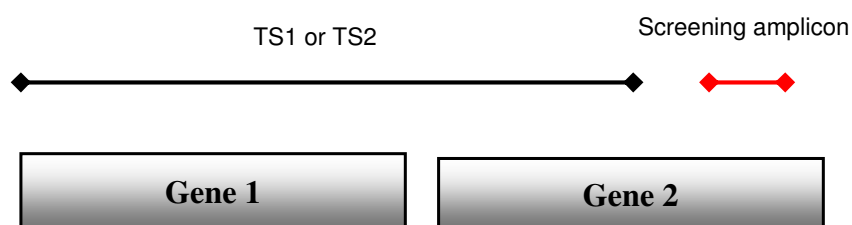


Figure 2-2 shows the methodology for designing the targeting sequences to allow the design of the screening primers.

To ensure the usefulness of the capture vector, the presence of both targeting sequences were checked against all available genomes within the chosen species. The genomes were checked using the targeting sequence primers to perform *in-silico* PCR. Any genomes that failed to produce an *in-silico* amplicon were investigated further. The reference genome DNA amplicon was blasted against the genome using NCBI Blast. In the few circumstances where this occurred, further investigation demonstrated a single nucleotide change within the core of the primer sequence.

The relevant homology sequence was added to the individual primers. The full-length primers are particularly long (40-60bp), so NetPrimer was used to check the secondary structures including hairpins, self-dimers, and cross-dimers (<http://www.premierbiosoft.com/netprimer/index.html>). If any of the secondary structures appeared to be obstructive, the individual primer was re-designed in the first instance.

2.2.4 Amplification the pLLX8 amplicon

Due to the importance of maintaining the functionality of *cyh* and *bla* gene a high-fidelity enzyme, Extensor Hi-Fidelity PCR enzyme mix (Abgene) was used to reduce the chance of mutations.

For amplification, a final volume of 50µl was used containing 5µl 10x Extensor Buffer 1 (22.5mM MgCl₂), 1µl of each primer (10pmol/µl), 1µl dNTPs (10mM each dNTP, Bioline), 1µl pLLX8 (10ng/µl), 0.25µl Extensor Hi-Fidelity PCR enzyme mix and 40.25µl nanopure water. The following cycling conditions were used; 94°C for 2 minutes for initial denaturation, followed by 30 cycles of 94°C for 30 seconds, 60°C for 30 seconds and 72°C for 3 minutes, proceeded by a final extension at 72°C for 7 minutes.

2.2.5 Amplification of the targeting sequences, TS1 and TS2

For amplification, a final volume of 20µl was used containing 4µl 5x Colorless Reaction Buffer (7.5mM MgCl₂), 2µl of each primer (10pmol/µl), 0.4µl dNTPs (10mM each dNTP, Bioline), 1µl reference genomic DNA (50ng/µl), 0.25µl GoTaq^R DNA Polymerase (Promega) and 10.35µl nanopure water. The following cycling conditions were used; 95°C for 5 minutes for initial denaturation, followed by 30 cycles of 95°C for 30 seconds, 60°C for 30 seconds and 72°C for 1 minutes, proceeded by a final extension at 72°C for 5 minutes.

2.2.6 Lithium acetate transformation

A single colony of *S. cerevisiae* CRY1-2 was grown overnight in 5ml YPD, 200rpm, 30°C. 500µl of the overnight culture was added to a fresh 5ml of YPD and incubated

at 200rpm, 30°C for 4 hours. The cells were centrifuged at 1,500g, 5 minutes. The cells were then washed twice in 1ml of sterile water in a microcentrifuge tube at 10,000g, 1 minute. An additional wash was performed in 0.1M lithium acetate after dividing into two microcentrifuge tubes. The supernatant was removed and the following was added, 240µl 50% (w/v) PEG 4,000, 36µl 1M lithium acetate, 50µl 2mg/ml salmon carrier DNA and 34µl DNA mixture (200ng TS1, 200ng TS2, 600ng pLLX8 amplicon and 200ng *Xba*I-linearised pLLX13 plus water). The cells were incubated at 30°C for 30 minutes and then 42°C for 40 minutes. One ml of sterile water was added to the yeast suspension and then centrifugation for 4,000g for 1 minute. The yeast cells were gently re-suspended in 300µl sterile water. 100µl of the yeast suspension was plated on URA plates. The plates were incubated at 30°C for 3-4 days.

2.2.7 Selection of Capture Vector

One ml of nanopure water was added to a single URA plate and scraped with spreader to produce a suspension and 100µl of this suspension was subjected to yeast plasmid preparation (Section 2.1.6). Electroporation was performed as mentioned in Section 2.3.4. The suspension is plated onto LB plates containing both ampicillin (100µg/ml) and tetracycline (10µg/ml), thereby selecting for the complete capture vector. Tetracycline resistance is from the pLLX13 backbone and ampicillin resistance from the pLLX8 amplicon.

Further verification was sought to confirm that the amplicons had recombined in the right order by performing a PCR extending from TS1 to TS2. A PCR was performed using the TS1 U primer and the TS2 D primer which should result in a ~5kb amplicon. For amplification, a final volume of 20µl was used containing 4µl 5x Colourless Reaction Buffer (7.5mM MgCl₂), 2µl each primer (10pmol/µl), 0.4µl dNTPs (10mM each dNTP, Bioline, UK), 1µl capture vector clone plasmid preparation (1:30 dilution), 0.25µl GoTaq^R DNA Polymerase (Promega) and 10.35µl nanopure water. The following cycling conditions were used; 95°C for 5 minutes for initial denaturation, followed by 30 cycles of 95°C for 30 seconds, 60°C for 30 seconds and 72°C for 5 minutes, proceeded by a final extension at 72°C for 5 minutes.

Important to the use of the capture vector in spheroplasting, is the introduction of the unique *Mlu*I restriction site by incorporation of its recognition site into the primer sequence. This is required to linearise the capture vector for the spheroplast transformation experiments. It is also important that the *I-Sce*I restriction sites are functional to enable estimation of the site of the insert after successful island capture. So restriction digestions using these enzymes were also carried out on the putative capture vector clones plasmid DNA preparation.

2.3 Capturing a genomic island

2.3.1 Preparation of high molecular weight genomic DNA

After optimisation of the genomic island capture technique, all genomic DNA preparations were prepared using the ArchivePure DNA Cell/Tissue Kit (Flowgen Biosciences, UK). For the initial yeast transformation experiments genomic DNA was prepared as described by Liang *et al* (2001) and the Caesium Chloride purification method was taken from Current Protocol in Molecular Biology (Wilson 2001).

2.3.2 Preparation of spheroplasts

A 5ml culture of *S. cerevisiae* CRY1-2 was grown overnight in YPD at 30°C and 200rpm. The overnight culture was added to 250ml flask containing 50ml of fresh YPD and incubated at 30°C, 200rpm for 3-4 hours until the culture reached an OD₆₆₀ 0.9-1.1. The cells were harvested by centrifugation at 1,500g for 5 minutes, 4°C. The cells were washed twice in 20ml chilled (4°C) sterile water. Re-suspended in 20ml of SPE (1M Sorbitol, 10mM sodium phosphate pH 7.5, 10mM EDTA) , added 200µl 1M DL-Dithiothreitol (DDT) and 50µl 10mg/ml Zymolyase; incubated at 30°C, 80rpm for approximately 60 minutes. The percentage of spheroplasting was monitored during the incubation by taking an initial 100µl of the cell prior to adding Zymolyase to act as the control. The OD₈₀₀ was compared of 900µl water plus 100µl cell suspension before digestion with that of 100µl after digestion; looking for approximate 10% decrease with a lower range ~0.07. Samples were taken approximately every 10 minutes after 40 minutes of digestion. When digestion was complete the cells were then centrifuged at 200g for 10 minutes, 4°C. All re-suspensions after this point were

done with gentle rocking motion, no pipetting or vortexing. Washed cells twice in 20ml chilled 1M STC (1M Sorbitol, 10mM CaCl₂, 10mM Tris-HCl pH 7.5). Re-suspended in 2ml STC and mixed 450µl aliquots of spheroplasts with 1µg *Mlu*I-linearised capture vector DNA plus 5µg genomic DNA in no more than a 45µl volume in 30ml sterlin tube; incubated for 10 minutes at room temperature. Added 4.5ml PEG reagent (20% PEG 8000, 10mM CaCl₂, 10mM Tris-HCl pH 7.5) mixed by gentle swirling and rocking motion; incubated at room temperature for 10 minutes. The cells were then centrifuged at 200g for 10 minutes at 4°C. The supernatant was discarded and the cells were re-suspended in 2ml SOS (25% YPD, 7mM CaCl₂, 1M sorbitol); incubate at 30°C for 40 minutes. 15ml of TOP agar (1M sorbitol, 2.5% agar, URA broth) which have been kept at 45°C was added and inverted quickly several times and plated immediately on pre-warmed (37°C) SORB plates (1M sorbitol, 2% agar, URA broth). Colonies were expected 5-7 days after incubation at 30°C.

2.3.3 Screening of spheroplast colonies: colony PCR

At best 100 colonies were patched onto fresh URA plates using sterile wooden toothpicks and left to grow overnight for at least 20hrs. The screening PCR reactions were set up in a total volume of 20µl; 2µl colony PCR buffer (45mM Tris-HCl pH 8.8, 11mM Ammonium sulphate, 45mM Magnesium chloride, 4.4mM EDTA pH 8.0, 1µM dNTP (each), 113µg/ml BSA and 7mM β-mercapoethanol), 2µl 10pmol µl⁻¹ of each positive screening primer, 1µl 10 pmol µl⁻¹ of each *tetA* primer, 0.4µl GoTaq^R DNA Polymerase (Promega, UK) and 13.6µl nanopure water. Each patch was touched gently with a yellow tip (20µl-200µl tip) and re-suspended by gentle pipetting into the PCR mixture. The following cycling conditions were used; 95°C for 5 minutes for initial denaturation, followed by 35 cycles of 95°C for 40 seconds, 60°C for 30 seconds and 72°C for 2 minutes, proceeded by a final extension at 75°C for 5 minutes.

Any positive patches were than grown overnight in 5ml URA broth and subjected to yeast plasmid preparation as stated in Section 2.1.6.

2.3.4 Electroporation: Transfer of suspect plasmids to *E. coli*

An electroporation cuvette (Gene Pulser Cuvette, 0.2cm, BioRad) and an aliquot of electro-competent cells were placed on ice. After the cells had thawed, 1µl of yeast plasmid preparation was added to the cuvette and 20µl *E. coli* DH10B was added to the DNA drop. The cuvette was incubated on ice for 5 minutes. The Gene Pulser electroporator (BioRad) was set to 1.6kV, 200Ω and 25µF. Excess moisture was removed from the cuvette and an electric pulse was given. The time constant was expected to be between 4.0 and 4.5 msec. 600µl SOC was added immediately and incubated at 37°C, 200rpm for 1 hour. 100µl of the bacterial suspension was plated onto LA containing tetracycline (10µg/ml) plates. The plates were incubated overnight at 37°C.

2.3.5 Restriction Fragment Length Polymorphism (RFLP)

A colony from each plate was chosen was subsequently used to produce a 5ml overnight for plasmid preparation and for re-streaking onto a fresh LA containing tetracycline (10µg/ml) plate. The following day the 5 ml overnight was harvested and subjected to bacterial plasmid preparation (Section 2.1.5). The restriction digests were performed as in Section 2.1.7, usually with the following enzymes, *I-SceI*, *BamHI* and *HindIII* and if appropriate compared with expected sequence fragment pattern generated using NEBCutter (Vincze, Posfai & Roberts 2003).

2.3.6 Sub-cloning of E106-*serW*

The RFLP of E106-*serW* plasmid suggested that in order to get the most from sample sequencing, *PstI* and *KpnI* were the best enzymes to use for sub-cloning.

In a total volume 60µl restriction digestion, 30µl of an E106-*serW* midi plasmid preparation and 3µl enzyme (either *PstI* or *KpnI*) were digested for 5 hours. A restriction digest was prepared of 100ng pWSK29 with either *PstI* or *KpnI*. All restriction digests were heat-inactivated at 65°C for 20 minutes, vacuum-dried, washed in 500µl 70% ethanol. Then centrifuged at 10,000g for 5 minutes, vacuum-dried and re-suspended in 10µl nanopure water. The ligation was set up a total volume

of 30µl, 10µl of digested pWSK29 and 10µl digested E106-*serW*, 3µl 10x ligase buffer, 1µl T4 ligase (Promega, UK) and 6µl nanopure water was added and left at 4°C overnight. All the ligation mixtures were heat-inactivated at 65°C for 20 minutes, vacuum-dried, washed in 500µl 70% ethanol, followed by centrifugation at 10,000g for 5 minutes, vacuum-dried and re-suspended in 10µl nanopure water.

A 50µl of aliquot of chemical competent cells were defrosted on ice. 2µl of the ligation mixture was added to the cells and was incubated on ice for 30 minutes. The mixture was incubated at 42°C, 15 seconds and then immediately incubated on ice for 2 minutes. 200µl LB was added to the mixture and left to incubate at 37°C for 1 hour. 100µl of the mixture plus 40µl X-Gal (20 mg/ml) was plated on LA plates containing ampicillin (100µg/ml) and incubated overnight at 37°C.

Forty µl X-Gal (20 mg/ml) plus 160µl nanopure water were spread onto a fresh plate of LA plates containing ampicillin (100µg/ml). Twenty white colonies were patched per enzyme onto this plate and the plate was incubated overnight at 37°C. Following incubation 10 white colonies were selected and overnight LB cultures were prepared. A mini plasmid preparation was performed on all 10 colonies and the appropriate restriction digest was performed (either *KpnI* or *PstI*) to check size of insert. A single clone for each insert size was selected and prepared for sequencing.

2.4 Preparation of *Pseudomonas aeruginosa* for *in vivo* work

2.4.1 Preparation of bacterial stocks

A single colony of *P. aeruginosa* was grown overnight in 5ml LB, 200rpm, 37°C. 600µl was removed and centrifuged at 13,000rpm for 1 minute. This pellet was added to 20ml of TSB (Sigma, UK) supplemented with 20% fetal calf serum (FCS - Sigma, UK). The culture was then incubated at 37°C, 200rpm for 3 hours. The culture was aliquoted into 1.5ml microcentrifuge tubes and stored at -70°C.

2.4.2 Infection dose preparation

The stock was left at -70°C for a minimum of 3 days and then one aliquot was thawed and centrifuged at 13,000rpm for 2 minutes. The supernatant was removed and the cells were re-suspended in 500µl of PBS (136mM sodium chloride, 2.68mM potassium chloride, 10.14mM di-sodium hydrogen phosphate, 1.76mM potassium di-hydrogen phosphate, pH 7.4).

Twenty µl of the re-suspension was serially diluted to 10^5 , in 180µl PBS. Six 20µl drops of each dilution were plated onto LA plates and incubated at 37°C overnight. The dilution with between 30 – 300 colonies was counted and the mean of the six spots was calculated, multiplied by 50 to calculate the number of colonies per ml and then multiplied by the dilution factor. The result is the colony forming unit per ml (CFU/ml). This figure was subsequently used to calculate whether a further dilution of the stock solution was needed before mouse infection.

2.5 In vivo mice protocols

All the mice used in this thesis were Balb/c females (Harlan, UK) aged between 8-12 weeks old that were specifically pathogen free.

2.5.1 Administration

The mice were anaesthetised with 5% (v/v) isoflurane over 1.8L/min oxygen and then each mouse was held in an upright position by the scruff of its neck. 50µl of bacteria re-suspended in PBS was carefully added to the nares to allow the fluid to be inhaled. The mice were placed on their backs in the cage to aid recovery. An acute infection was achieved by infecting the mice intranasally with *P. aeruginosa* 2×10^6 CFU in 50µl.

For intravenous infections, mice were then placed in a plastic cylinder with dimensions that restricted movement and closed leaving the tail exposed. An acute infection was achieved by infecting the mice via a lateral tail vein with *P. aeruginosa* 2×10^6 CFU in 100µl.

2.5.2 Organ retrieval & CFU counts

For acute infection experiments, mice were culled and samples were taken for analysis 18 hours post-infection, blood was taken via cardiac puncture, 20µl was immediately added to 180µl PBS, the nasopharynx and lungs were removed by dissection and homogenised in 10ml of pre-weighted PBS.

20µl of the homogenate or blood suspension was serially diluted to 10^3 , in 180µl PBS. Six 20µl drops of each dilution were plated onto LA plates, left to dry and incubated at 37°C overnight. The dilution with between 30 – 300 colonies was counted and the mean of the six spots was counted, multiplied by 50 to calculate the number of colonies per ml and then multiplied by the dilution factor. The result is the colony forming unit per ml (CFU/ml).

2.5.3 Symptom scoring

At the end point of 18 hours post-infection and frequently during survival experiments, the symptoms of each individual mouse were monitored. There were three signs to monitor; hunched, starry coat and lethargic each monitored on a scale of 0 = no signs, 1 = starting to show symptom and 2 = obviously exhibiting the symptom (Morton, Griffiths 1985). The three scores were added together to give a total symptom score out of 6. Mice that had a total symptom score of 6 were culled immediately as the experimental end point as defined by the home office project licence.

2.5.4 Histopathology

After the mice were culled, the lungs were removed and placed within 5ml PBS. Aluminium foil was made into a cylinder shape by firmly wrapping the foil around a 5ml Bijoux tube. A metal container was placed in dry-ice and iso-pentane was carefully poured into the metal container. Tissue Tek OCT tissue embedding compound (Miles) was poured into the cylinder aluminium mould to cover the bottom. The lungs were then added to the cylinder aluminium mould and small amount of OCT tissue embedding compound was added to just cover the lungs. The

mould was then added to the metal container with iso-pentane using tweezers and left until completely frozen. The mould was removed using tweezers and completely closed. After all lungs were frozen the metal container was removed from the ice and left to warm to room temperature before disposal. The lungs were stored at -70°C for long-term storage and were transferred to -20°C, 24 hours prior to sectioning. The lungs were cut into 15µm thick sections at -18°C using a Bright microtome.

To stain the sections, the slides were submerged in 90% ethanol for 5 minutes and then rinsed with tap water. The slides were then submerged in Harris' haematoxylin (BDH) for 30 seconds, rinsed with tap water and then immediately submerged in eosin Y-solution, 0.5% aqueous (BDH) and rinsed with tap water. The slides were dried. The slides were then hydrated for 30 seconds with 70% ethanol, then 90% ethanol and finally in 100% ethanol. The slides were then submerged in xylene. A coverslip was prepared by adding a drop of DPX mounting medium (Fisher Scientific, UK). A slide was removed from the xylene and pressed on the coverslip to mount.

Histopathology scores were determined from photographs taken using the x10 and x20 objective lens on a light microscope with a digital camera (NikonE4500) attached. The sections were scored on a scale of 1-5 to represent severity of lung pathology; starting from 1 minor pathology to 5 severe pathology based on a scale used by Magda Bortoni-Rodriguez (Bortoni Rodriguez 2006).

2.5.5 Leukocyte counts

Five ml of lung homogenate was filtered through Falcon 40µm nylon cell strainers (BD) into a 50ml Falcon tube. The homogenate was centrifuged at 300g, 5 minutes. The supernatant was discarded and the pellet re-suspended in 10ml Ammonium chloride lysing reagent diluted 1:10 with nanopure water. The suspension was left for 10 minutes at room temperature in the dark then centrifuged at 300g, 5 minutes. The cells were re-suspended in 8ml PBS. Three cytospin slides were prepared per individual mouse. 50µl of the re-suspension was centrifuged onto cytospin slide (Shandon) using a cytocentrifuge, Cytospin 2 (Shandon) at 1,500rpm, 3 minutes. The slides were allowed to dry overnight. The slides were stained using the Reastain^R Quik-Diff Kit (Reagen, UK), dipped 10 times in Reastain^R Quik-Diff FIX, rinsed

with distilled water, dipped 10 times in Reastain^R Quik-Diff RED, rinsed in distilled water and 10 times in Reastain^R Quik-Diff BLUE. The slides were finally rinsed with distilled water and left to air dry. They were counted using the x40 objective lens on a light microscope, identifying monocytes, macrophages, lymphocytes and polymorphonuclear leukocytes (neutrophils). The numbers were then used to calculate the total numbers of immune cells within the lung homogenate as well as the proportion of each cell type.

2.5.6 Statistics

Statistical analysis was performed with GraphPad Prism 5 software. Analysis of CFU was done by analysis of the variance (ANOVA) with Tukey-Kramer post-analysis. Analyse of symptom scores was done using the non-parametric Kruskal-Wallis test and Dunns post-analysis. Significantly difference results were $P < 0.05$. Survival data analysis was done using both the logrank test and the Gehan-Wilcoxon test.

3 Analysis of genomic islands captured using yeast-based technology: the genomic island capture technique

3 ANALYSIS OF GENOMIC ISLANDS CAPTURED USING YEAST-BASED TECHNOLOGY: THE GENOMIC ISLAND CAPTURE TECHNIQUE	51
3.1 CAPTURE EXPERIMENTS INVOLVING GENOME SEQUENCED BACTERIAL STRAINS	53
3.1.1 K12 MG1655– <i>SERW</i>	54
3.1.2 K12 MG1655– <i>LEUX</i>	55
3.1.3 PAO1- <i>LYS10</i> AND PA14- <i>LYS10</i>	56
3.1.4 PA14- <i>LYS47</i>	56
3.2 CAPTURE EXPERIMENTS INVOLVING CLINICAL BACTERIAL STRAINS	58
3.2.1 KR115- <i>LYS10</i> AND KR159- <i>LYS10</i>	59
3.2.2 E105- <i>SERT</i>	59
3.3 E106-<i>SERW</i>	60
3.3.1 STRUCTURE OF THE E106- <i>SERW</i>	63
3.3.2 ANALYSIS OF E106- <i>SERW</i> CRISPR SYSTEM	67
3.4 E105-<i>LEUX</i>	70
3.4.1 DEFINING E105- <i>LEUX</i> AS A GENOMIC ISLAND (ISLET)	75
3.4.2 THE GENOMIC STRUCTURE OF THE E105- <i>LEUX</i> GENOMIC ISLET	78
3.4.3 EVOLUTION OF THE E105- <i>LEUX</i> GENOMIC ISLET	84
3.5 DISCUSSION	86
3.5.1 GENOMIC ISLAND CAPTURE AS A YEAST-BASED TECHNOLOGY	86
3.5.2 GENOMIC ISLAND CAPTURE AS A GENOMIC ISLAND DISCOVERY TOOL	90
3.6 CONCLUSION AND FURTHER WORK	92

The aim of this project is to identify, capture and characterise genomic islands. This project explores the characterisation of pathogenicity islands *in vitro*. To target pathogenicity islands, we used known pathogenic strains of *E. coli* and *P. aeruginosa* (clinical isolates) from human infection. To increase the likelihood of any genomic islands found being pathogenicity islands.

To achieve the aim of the project the genomic island capture technique used by Wolfgang *et al*, 2003 was developed to produce a generic system. The main principle of this technique involves introducing the capture vector along with genomic DNA into yeast cell spheroplasts, which induces homologous recombination resulting in the capture vector now containing the genomic island. Genomic islands were captured from both *E. coli* and *P. aeruginosa*. This is the first report of the genomic island capture technique being used with a bacterial species other than *P. aeruginosa*. In order to achieve this aim optimisation of the efficiency of construction of tRNA loci capture vectors, the capturing of novel genomic islands, as well as detecting the colonies that contain the genomic island was essential. The methodology is described in Section 2.3.

This chapter explores the genomic islands captured using the genomic island technique. It proceeds by describing the genomic islands captured from sequenced strains to standardise the technique and is then followed by describing genomic islands captured from clinical strains of *P. aeruginosa* and *E. coli*. This is followed by detailed analysis of genomic islands captured from two clinical strains of *E. coli*. It concludes with the discussion of the data generated and an overall conclusion.

The captured inserts are referred to as the strain name preceding the tRNA loci, for example, PAO1-*lys10*, refers to the insert captured from *P. aeruginosa* PAO1 at the 10th tRNA lysine.

3.1 Capture experiments involving genome sequenced bacterial strains

This section will describe the initial capture experiments using reference genomes. The experiments were designed to develop the methodology associated with genomic island capture technique and to test the size limits of the captured insert. For *E. coli*, the reference genome was K12 MG1655 and for *P. aeruginosa* PAO1 and PA14.

The yeast cell spheroplasts are transformed with the capture vector plus the genomic DNA. Homologous recombination between the capture vector and the genomic DNA results in the genomic island now being present within the vector. The main principle of the yeast capture method is depicted in Figure 3-1. The term captured insert refers the DNA sequence running from TS1 to TS2, which consists of the genomic island with conserved flank genomic DNA.

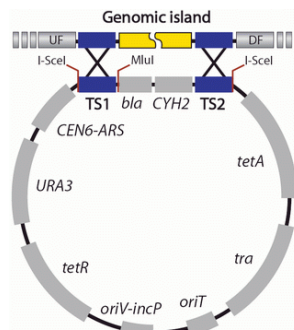


Figure 3-1 depicts the principles of genomic island capture technique. TS1 and TS2 are the conserved targeting regions found upstream and downstream of the genomic island. – Reproduced from the MobilomeFinder website (<http://mml.sjtu.edu.cn/MobilomeFINDER/ycv.htm>)

The targets for the initial capture experiments were chosen to vary in size based on the capture vectors available. The targets are listed in Table 3-1. The size of the captured insert was generated from the GenBank sequence of the reference genomes using the TS1 U and TS2 D primer sequences. The results also show the number of positive colonies determined by PCR discovered for each target. The results are supported by the literature which suggests that the larger the size of captured insert, the lower the recovery of positive colonies (Leem *et al.* 2003).

Number	Organism	Capture Vector	Size of insert	% positive clone
1	<i>E. coli</i> K12 MG1655	<i>serW</i>	4kb	16% (3/19)
2	<i>E. coli</i> K12 MG1655	<i>leuX</i>	43kb	6% (6/95)
3	<i>P. aeruginosa</i> PA01	<i>lys10</i>	11kb	8% (1/12)
4	<i>P. aeruginosa</i> PA14	<i>lys10</i>	13kb	8% (2/24)
5	<i>P. aeruginosa</i> PA14	<i>lys47</i>	110kb	Unsuccessful

Table 3-1 shows the targets of the initial capture experiments. The size of the captured insert was generated from the publicly available GenBank genome sequences.

The following subsections will describe the initial captured inserts from the reference genomes. The genomic island capture technique was designed to enable rapid estimation of the size of the genomic insert by having two *I-SceI* restriction sites flanking the captured insert, which are located immediately upstream of TS1 and immediately downstream TS2.

3.1.1 K12 MG1655–*serW*

The *serW* locus within *E. coli* K12 MG1655 is an empty site, no insertion events have occurred, the size of the insert is the distance between the two targeting sequences (TS1 and TS2). The first successful capture experiment resulted from homologous recombination between the *serW* capture vector and K12 MG1655-*serW* amplicon. The K12 MG1655-*serW* amplicon was generated from *E. coli* K12 MG1655 genomic DNA using the *serW* TS1 U and *serW* TS2 D primers.

The use of the K12 MG1655-*serW* amplicon provided a resolution of the initial hurdles to successful genomic island capture. The initial capture experiments involving genomic DNA resulted in no or very low recovery of colonies (Table 3-2). A number of experiments using the K12 MG1655-*serW* amplicon as a control demonstrated that the genomic DNA preparation was inhibiting the capture process (Table 3-3).

	DNA	No of colonies
Control	No DNA	0
Test	10µg K12 MG1655 - <i>serW</i>	0
Test	20ng E101 amplicon - <i>serW</i>	1744

Table 3-2 shows number of colonies retrieved during a capture experiment using 1µg *serW* capture vector as a control and testing *E. coli* K12 MG1655 genomic DNA as well as the E101 amplicon and *serW* capture vector

	DNA	No of colonies
Control	No DNA	0
Control	20ng pLLX13	173
Test	20ng pLLX13 + 10µg gDNA	0
Test	180ng K12 MG1655 - <i>serW</i>	178
Test	180ng K12 MG1655- <i>serW</i> + 10µg gDNA	0

Table 3-3 shows number of colonies retrieved during a capture experiment using 20ng pLLX13 as a control and testing K12 MG1655-*serW* amplicon and *serW* capture vector with replicas of 10µg *E. coli* K12 MG1655 genomic DNA (gDNA)

After the resolution of the genomic DNA problems, K12 MG1655-*serW* insert was also successfully captured from *E. coli* K12 MG1655 genomic DNA.

3.1.2 K12 MG1655-*leuX*

The *leuX*-associated genomic island within *E. coli* K12 MG1655 is 43kb in size and was the largest insert captured during this project. It is a genomic island described in the literature by Bishop *et al.* (2005). Figure 3-3 shows K12 MG1655-*leuX* digested with *Hind*III.

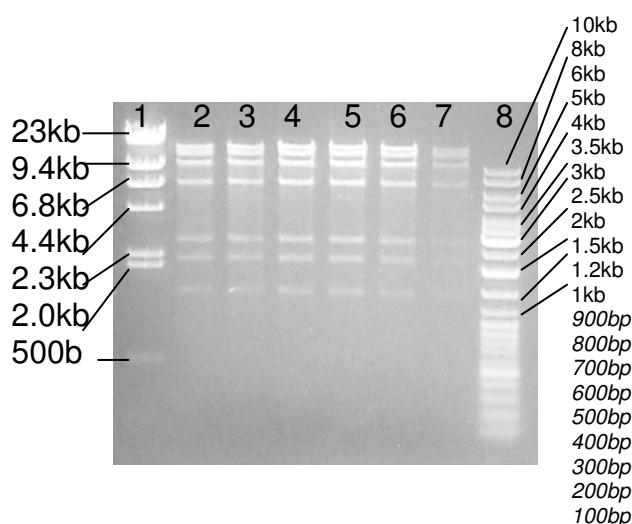


Figure 3-2 shows a 0.6% agarose gel with the restriction profile of K12 MG1655-*leuX* with *Hind*III Lane 1 λ*Hind*III 500ng; Lane 2-7 K12 MG1655-*leuX*; Lane 8 Generuler DNA ladder mix 500ng

3.1.3 PAO1-*lys10* and PA14-*lys10*

The genomic island at the *lys10* tRNA locus within *P. aeruginosa* PA14 is the pathogenicity island, PAPI-2 (He *et al.* 2004). The *P. aeruginosa* PAO1 has a genomic island present at this tRNA site locus defined in the literature (He *et al.* 2004), which has some sequence similarity with PAPI-2.

The capture vector containing the PAPI-2 island (PA14-*lys10*) was used subsequently in complementary studies performed by Ewan Harrison (Harrison E, PhD thesis, University of Leicester, unpublished) involving the *P. aeruginosa* pathogenicity island deletant mutants, PA14 Δ PAPI-2 and PA14 Δ PAPI-1 Δ PAPI-2. Figure 3-3 shows the restriction profile of both PAO1-*lys10* and PA14-*lys10* using *I-SceI*.

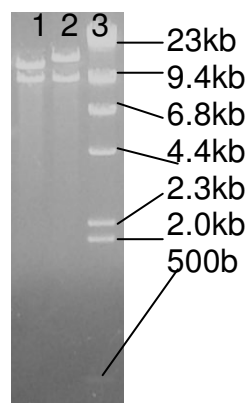


Figure 3-3 shows a 0.6% agarose gel with the restriction profile of PAO1-*lys10* and PA14-*lys10* with *I-SceI* Lane 1 PAO1-*lys10*; Lane 2 PA14-*lys10* Lane 3 λ HindIII 500ng

3.1.4 PA14-*lys47*

The genomic island present within *P. aeruginosa* PA14 at the *lys47* tRNA site is PAPI-1, the other pathogenicity island explored during the murine studies within this thesis. The successful capturing of this island would have been useful for complementary studies with the pathogenicity island deletant mutants. Unfortunately this island was never successfully captured. A set of experiments

were designed based on the assumption that the size of the island was approaching the limits of the genomic island capture technique; PAPI-1 is 108kb. The island is possibly too large for the technique in its current state; it had been previously used to capture an 80kb island (Wolfgang *et al.* 2003).

One approach was to increase the likelihood of the capture vector coming into contact with a genomic DNA fragment containing PAPI-1. The initial experimental design was to find restriction enzymes that would cut exclusively at the flanks of PAPI-1. Leem *et al.* (2003) support this method, they experimentally showed that cutting the target DNA at the flanks, increased recovery of positive clones in their example from 0.8% to 17%. Currently, no such restriction enzymes are commercially available. The experiment was redesigned to cut the genomic DNA with as many enzymes that that were non-cutters within PAPI-1 simultaneously. There were three restriction enzymes commercially available; *PmaI*, *SwaI* and *PacI*.

The second approach was to vary the amount of genomic DNA. In the DNA preparation the majority of DNA fragments are between 150 - 200kb. The larger the genomic island, the less likely the entire island is to be found on an individual genomic DNA fragment. This suggests that increasing the amount of genomic DNA would increase the number of positive colonies. Increasing the amount of genomic DNA conflicted with previous experiments that showed an excess of genomic DNA can be a limiting factor on the transformation efficiency. All experiments designed to capture PA14-*lys47* resulted in a low recovery of colonies and no PCR positive colonies.

In conclusion, a number of successful capture experiments were achieved using reference genome strains. The size of the captured insert varied from 4 – 43kb. These experiments showed promise of the development of a generic genomic island capture technique by successfully using the technique on bacterial genomic DNA from strains other than *P. aeruginosa*.

3.2 Capture experiments involving clinical bacterial strains

After standardising the technique, putative genomic islands within bacterial strains with un-sequenced genomes (clinical isolates) were targeted. This section proceeds by describing how the suspected genomic islands were identified. This is followed by a summary of the genomic islands targeted within the clinical isolates and brief descriptions of three targets. Two targets are excluded from this section (E106-*serW* and E105-*leuX*), as they are discussed and analysed within their own section.

The clinical isolate genomic island targets were determined by tRIP PCR (**t**RNA site **i**nterrogation for **p**athogenicity islands, prophages and other genomic islands) and SGSP PCR (**s**ingle **g**enome **s**pecific **p**rimers). The tRIP PCR and SGSP data were gathered by Ali Thani (Thani A, PhD thesis, University of Leicester) and Ewan Harrison (Harrison E, PhD thesis, University of Leicester, unpublished) for *E. coli* and *P. aeruginosa* respectively. The tRIP PCR technique is particularly usually for screening a large number of bacterial strains for genomic islands. The technique involves using a set of primers, one upstream and the other downstream of the tRNA site within the conserved flanks. If the site is empty or contains an islet of $\leq 5\text{kb}$, it will lead to a successful amplification of DNA, if there is an insert present then the PCR will fail. The PCR negative strains are then analysed further by SGSP PCR to get a snapshot of the structure of the genomic island and then analysed to determine whether there are similar islands already present within sequenced strains. The clinical isolate targets and the capture vectors used are listed in Table 3-4. The captured insert size is estimated from restriction digestion with *I-SceI* resolved on a 0.6% agarose gel.

Number	Organism	Capture Vector	Size of insert
1	<i>E. coli</i> E105	<i>leuX</i>	~11kb
2	<i>E. coli</i> E106	<i>serW</i>	~14kb
3	<i>E. coli</i> E105	<i>serT</i>	Unsuccessful
4	<i>P. aeruginosa</i> KR115	<i>lys10</i>	~9kb
5	<i>P. aeruginosa</i> KR159	<i>lys10</i>	~9kb

Table 3-4 shows clinical isolate strains targeted for genomic island capture. The sizes of the captured inserts were estimated from *I-SceI* restriction digest resolved on a 0.6% agarose gel.

3.2.1 KR115-*lys10* and KR159-*lys10*

Analysis of the *lys10* loci of the two clinical strains of *P. aeruginosa* KR115 and KR159 suggested a presence of a genomic island. Restriction digestion with *I-SceI* showed that the genomic islands were similar in size to one another and to that of the PAO1-*lys10* genomic island. A panel of restriction enzymes were selected to construct an RFLP for each genomic island. The results are shown in Figure 3-4. This demonstrates the ability of the technique to allow rapid comparison between genomic islands without the need for sequencing.

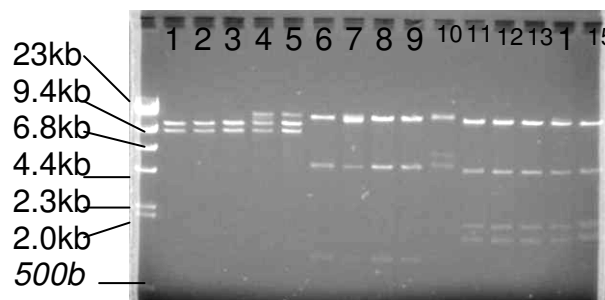


Figure 3-4 shows a 0.6% agarose gel with the restriction profile of PAO1-*lys10* and colonies of KR115-*lys10* and KR159-*lys10* with *I-SceI*, *HindIII* and *BamHI*

Lane 1 λ *HindIII* 500ng; Lane 2 -6 shows *I-SceI*, Lane 7 -11 shows *HindIII* and Lane 12 -16 shows *BamHI*, the clones represented are PAO1-*lys10*, three individual positive colonies of KR115-*lys10* and one colony of KR159-*lys10* respectively across the different digests

3.2.2 E105-*serT*

E105-*serT* positive clones were recovered, restriction digestion with *I-SceI* always resulted in a 2.4kb insert (Figure 3-5) even after multiple capture experiments. This result was unusual as tRIP PCR was negative, which suggests the insert was larger than 5kb. SGSP data provided by Ali Thani suggested an interesting genomic island was present. To investigate the problem, the 2.4kb insert was sequenced. The resulting sequence generated aligned to the expected sequence of an empty tRNA site as denoted by *E. coli* K12 MG1655. The data suggest that the genomic island was too unstable for the capture technique.

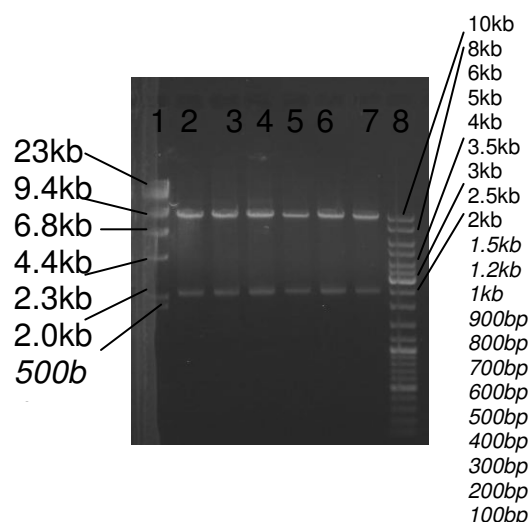


Figure 3-5 shows a 0.6% agarose gel with the restriction profile of E105-*serT* with *I-SceI*. Lane 1 λ HindIII 500ng; Lane 2-7 E105-*serT* *I-SceI*; Lane 8 Generuler DNA ladder mix 500ng

In conclusion, this section demonstrates that the genomic island capture technique can be used to compare genomic islands between strains. It can be used to generate an RFLP of the complete island to highlight similarities as well as differences between strains. It also shows that there are still possible limitations to the technique. The following two sections will describe two genomic islands also captured from clinical isolates explored in detail.

3.3 E106-*serW*

E. coli E106 is a clinical isolate recovered from a human septicaemia infection. The strain was originally part of a strain collection used to evaluate if a particular genome profile correlated with disease (Thani A, PhD thesis, University of Leicester). E106-*serW* was the second genomic island captured from a clinical isolate. This section will proceed by analysing the preliminary data generated for the genomic island. This is followed by experiments designed to determine the genomic structure of E106-*serW* and concludes with analysis of the genomic island.

Preliminary experiments were performed on the genomic island capture colony to verify the correct recombination event had occurred. Firstly, PCR was performed on the bacterial plasmid preparation to confirm that the plasmid contained TS1 and TS2 as well as the positive screening region. The plasmid was also subjected to

restriction digestion with *I-SceI* to get an approximation of the size of the captured insert (the genomic island plus the conserved flanks). As the captured insert was only approximately 14kb, it was within the range that could be amplified using long-range PCR. The captured insert was amplified from the genomic DNA of *E. coli* E106 using the *serW* TS1 U primer and the *serW* TS2 D primer. The size of the insert from long-range PCR and the *I-SceI* restriction digest are shown in Figure 3-6.

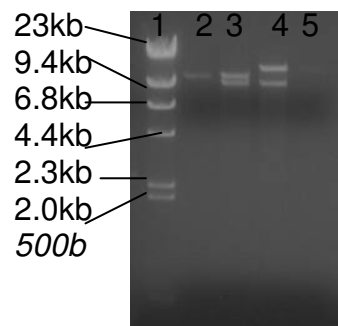


Figure 3-6 shows a 0.6% agarose gel with the restriction profile of E105-*leuX* and E106-*serW* with *I-SceI* and the long-range PCR product using the primer TS1 U and TS2 D
Lane 1 λ HindIII 500ng; Lane 2 E105-*leuX* long-range PCR product; Lane 3 E105-*leuX* *I-SceI*
Lane 4 E106-*serW* *I-SceI*; Lane 5 E106-*serW* long-range PCR product

The SGSP data provided by Ali Thani was re-analysed; construction of the SGSP library, screening and sequencing was performed by Ali Thani (Thani A, PhD thesis, University of Leicester). The following analysis was performed independently. The SGSP data provides a preview of the DNA sequence present at the genomic island flanks.

A 1.9kb SGSP amplicon was generated from the *EcoRV* genomic library of *E. coli* E106 using the *serW* U primer and the T3 universal primer. The amplicon was sent for sequencing with the *serW* U primer and 405bp of DNA sequence was produced. The most significant nucleotide alignment (*blastn*) was to the *E. coli* strains APEC O1 and UTI89 which produced identical score to one another (S: 719, E: 0, I: 99%). All other significant alignments were to the first 341bp of the amplicon sequence, only the *E. coli* strains APEC O1, B7A and UTI89 covered the first 401/405 nucleotides. The amplicon contained one gene, *infA* (58-277 nucleotides) and is part of the conserved U flank. The same SGSP amplicon was also sent for sequencing

from the opposite end using the KS universal primer and generated 364bp of DNA sequence. Based on *blastn* data APEC O1 and UTI89 (S: 567, E: 1e-158, I: 97%) gave full coverage of the DNA sequence 364/364 nucleotides. There were two additional bacterial species that aligned significantly *Citrobacter koseri* ATCC BAA 895 and *Enterobacter* spp 638. *Enterobacter* spp 536 give two partial alignments, one at nucleotides 1-69 (S: 68, E: 1e-06, I: 79%) and 221-255 (S: 44, E: 3.8, I: 85%) which come from the same ORF *Ent638_1406* (255bp). This indicated that the core DNA sequence does not significantly match. *Citrobacter koseri* matches from 107-257 (S: 78, E: 2e-09, I: 69%) ORF called *CKO_0219* (309bp). The nucleotide alignments identified the region as containing DNA sequence for the CRISPR-associated protein Csy4. Figure 3-7 summarises the significant nucleotide alignments of the SGSP amplicon.

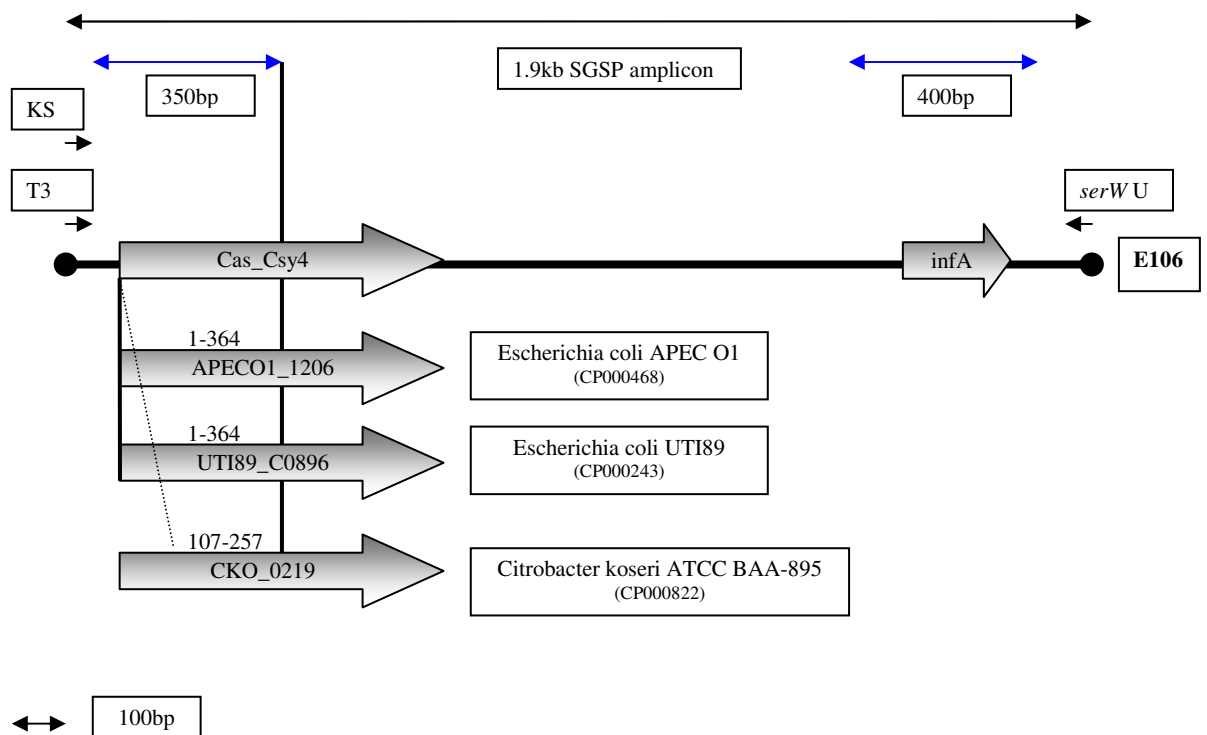


Figure 3-7 depicts the SGSP amplicon generated using the *serW U* and T3 primers from the *E. coli* E106 *EcoRV* genomic library. The most significant alignments are depicted adjacent to the SGSP amplicon.

A 2.4kb SGSP amplicon was generated from the *PstI* genomic library of E106 using the *serW D* primer and the T3 universal primer. The amplicon was sent for sequencing with the KS universal primer and 609bp of DNA sequence was

generated. The most significant nucleotide alignment (*blastn*) was to *E. coli* strains APEC O1 and UTI89 which produced an identical score to one another (S: 1054, E: 0, I: 98%). Another significant match was *Enterobacter* spp 638 (S: 255, E: 2e-64, I: 69%). All the alignments mentioned align for the full 609bp of the DNA sequence and aligns to genes predicted to encode the CRISPR-associated protein Cas1.

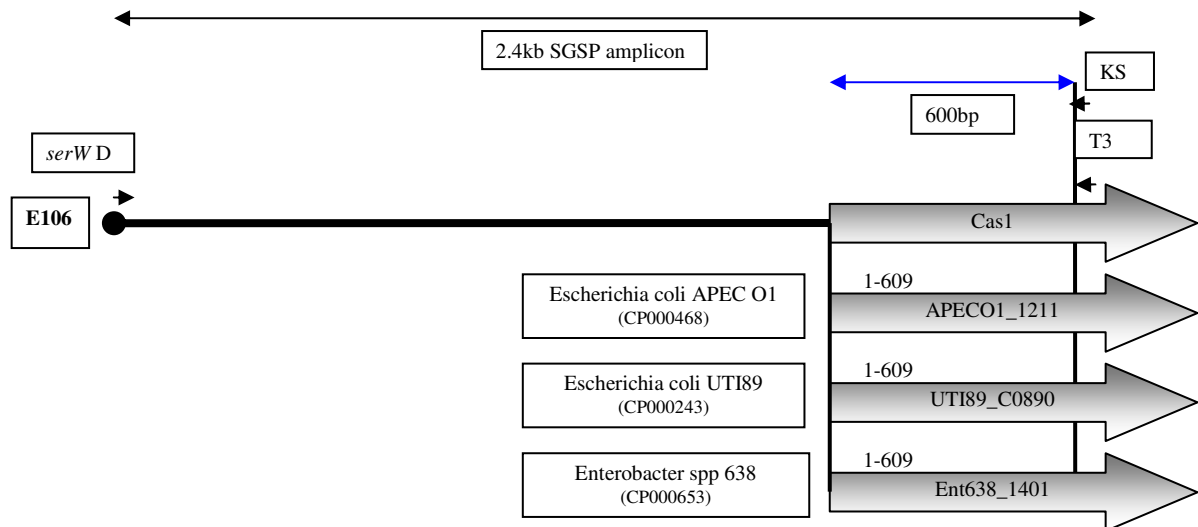


Figure 3-8 depicts the SGSP amplicon generated using the *serW* D and T3 primers from the *E. coli* E106 *Pst*I genomic library. The most significant alignments are depicted adjacent to the SGSP amplicon.

In conclusion, the preliminary data suggest that a similar genomic island has been found in previously sequenced strains of *E. coli*. The size and DNA sequence of the captured inserts for these *E. coli* strains were generated using the GenBank genome sequences. The results for the three *E. coli* strains were as follows: APEC O1 (13.3kb), B7A (16.6kb) and UTI89 (13.5kb). These captured insert sizes were similar to the estimated size of E106-*serW* (~14kb). The preliminary data also suggests that the genomic island, E106-*serW* contains a CRISPR system. The following subsection further investigates the similarity of E106-*serW* to those previously sequenced putative islands.

3.3.1 Structure of the E106-*serW*

This subsection proceeds by describing the structure of the previously identified genomic islands. It is then followed by the structure of E106-*serW* as determined by the experimental data.

Further investigation of the genomic islands present in the three *E. coli* strains APEC O1, B7A and UTI89 showed that all three contained CRISPR systems and a conserved set of genes as shown in Table 3-5. Analysis showed that the size discrepancy between strains APEC O1/UTI89 (13kb) and that of B7A (17kb) was primarily due to the addition of an insertion element (*IS100*) between the *serW* and *infA* genes.

	APEC O1	UTI89	B7A	Description
Genes				
clpA	clpA	clpA	EcB7A_2868	Part of conserved backbone ATP-dependent Clp protease ATP-binding subunit ClpA
#2	APECO1_1212	UTI89_C0889	EcB7A_2869	hypothetical protein
#3	APECO1_1211	UTI89_C0890	cas	CRISPR-associated protein Cas1
#4	APECO1_1210	UTI89_C0891	EcB7A_2871	CRISPR-associated helicase Cas3
#5	APECO1_1209	UTI89_C0892	csy1	CRISPR-associated protein (Cas_Csy1)
#6	APECO1_1208	UTI89_C0893	csy2	CRISPR-associated protein (Cas_Csy2)
#6a	N/A	UTI89_C0894	N/A	CRISPR-associated protein (Cas_Csy2)
#7	APECO1_1207	UTI89_C0895	csy3	CRISPR-associated protein (Cas_Csy3)
#8	APECO1_1206	UTI89_C0896	csy4	CRISPR-associated protein (Cas_Csy4)
serW	APECO1_t20	serW	EcB7A_2876	tRNA-Ser
#10	N/A	N/A	EcB7A_2877	IS100 transposase orfB
#11	N/A	N/A	EcB7A_2878	IS100 transposase orfA
#12	APECO1_1205	UTI89_C0898	EcB7A_2879	Hypothetical protein
infA	infA	infA	infA	Part of conserved backbone; translation initiation factor IF-1
aat	aat	aat	aat	Part of conserved backbone; leucyl/phenylalanyl-tRNA-protein transferase
cydC	cydC	cydC	cydC	Part of conserved backbone; cysteine/glutathione ABC transporter membrane/ATP-binding component

Table 3-5 The table shows and describes the genes found between *clpA* and *cydC* within the *E. coli* strains APEC O1, B7A and UTI89

To elucidate the structure of E106-*serW* two sets of experiments were performed. The first experiment was to subclone E106-*serW* to obtain sequence data for the interior of the genomic island. *KpnI* and *PstI* were used to generate the subclones, as

they both produced three fragments that varied in size. There were also a low number of cut sites for both enzymes within the capture vector backbone thereby increasing likelihood of getting putative island sequence. All DNA sequence generated aligned to the genomic islands previously identified in the three *E. coli* strains. The second experiment set of experiments were to identify if the RFLP (restriction fragment length polymorphism) corresponded to those expected from the previously identified genomic islands. RFLP is a technique that determines variation in a sequence of DNA using gel electrophoresis and restriction endonuclease enzymes. DNA sequence files of the capture vector with the captured insert were constructed using the GenBank genomic DNA sequences for APEC O1, B7A and UTI89. The *in-silico* restriction patterns were generated using NEBCutter (Vincze, Posfai & Roberts 2003) and were compared to those generated by E106-*serW* using an array of common restriction enzymes; agarose gel picture shown in Figure 3-9. The comparison between the RFLP and the expected RFLP can be found within the appendix (Section 7.1). The results showed that the E106-*serW* RFLP matched the pattern expected from the APEC O1-*serW* and UTI89-*serW*.

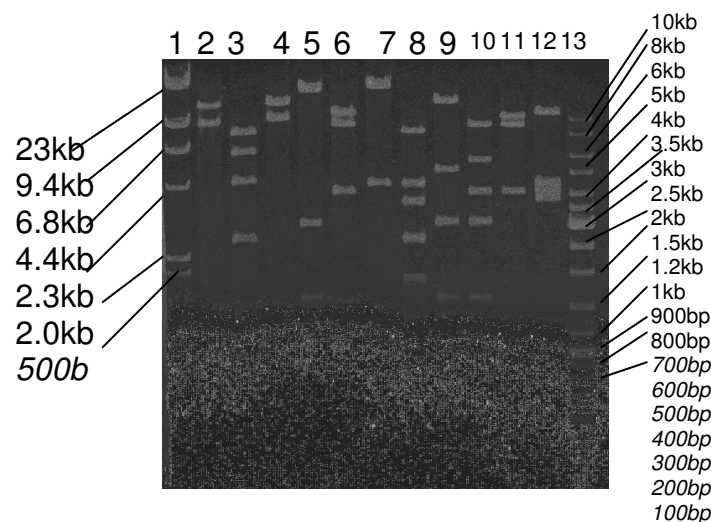


Figure 3-9 shows a 0.6% agarose gel of the RFLP of E106-*serW*

Lane 1 λ HindIII 500ng, Lane 2 *I-SceI*, Lane 3 *EcoRV*, Lane 4 *HindIII*, Lane 5 *KpnI*, Lane 6 *PstI*, Lane 7 *SalI*, Lane 8 *EcoRV-HindIII*, Lane 9 *KpnI-HindIII*, Lane 10 *KpnI-PstI*, Lane 11 *PstI-HindIII*, Lane 12 *PstI-SalI*, Lane 13 Generuler DNA ladder mix 500ng

The experimental data was analysed and compiled to generate the structure of E106-*serW* shown in Figure 3-10 using APEC O1-*serW* as a template. The DNA sequence

generated from the subcloning experiment does not verify the presence of *E106serW_1*. But the gene is found within the other three related genomic islands and the size of the captured insert is conserved thereby strongly supporting its presence. The DNA sequence data generated does not verify the presence of *E106_Cas-Csy1*. The *EcoRV* RFLP is consistent with the other related *serW*-associated regions and therefore the restriction site that is found within *Cas_Csy1* is maintained thereby supporting its presence.

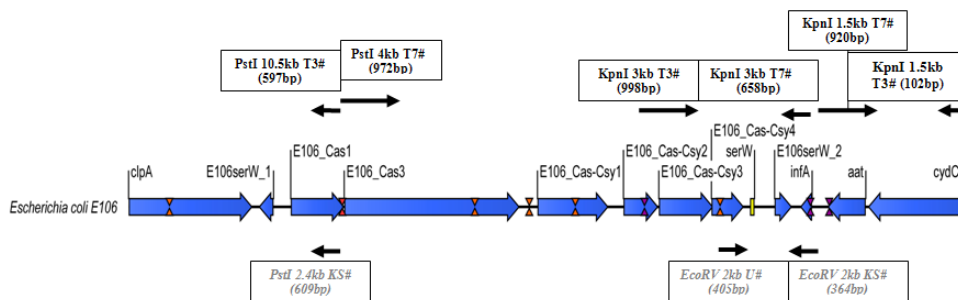


Figure 3-10 shows the DNA structure of E106-*serW* (~14kb). The text in grey represents data taken from SGSP data generated by Ali Thani. Restriction sites are represented by inverted triangles: *PstI* = red, *KpnI* = purple and *EcoRV* = orange.

In conclusion, this data shows that the insertion at the *serW* tRNA site within *E. coli* E106 is a CRISPR system. The results demonstrated that the E106-*serW* genomic island is closely related to APEC O1 and UTI89 *serW*-associated genomic islands. The E106-*serW* genomic island contains a number of CRISPR-associated genes and also two hypothetical genes. E106serW_1 and E106serW_2 are completely hypothetical proteins, containing no known conserved proteins domains but have counterparts in all three related *serW*-associated genomic islands.

The data highlights the limitations of SGSP as restriction sites assumed to be presence within the genomic island are in reality not. The data show that there is a single *PstI* site found within E106-*serW* (Figure 3-10). The SGSP data suggests there is an additional restriction site 2.4kb away. The subcloning data generated a 4kb fragment in the same direction. The SGSP data also suggests there is an additional *EcoRV* site that generates a 2kb fragment that includes *E106serW_Cas-Csy4* and *infA* (Figure 3-10). There is no *EcoRV* site proximal to *infA* in any of the

related genomic islands and this restriction fragment was not present in the E106-*serW* *EcoRV* RFLP.

3.3.2 Analysis of E106-*serW* CRISPR system

This subsection aims to provide a brief overview of CRISPR systems. This is followed by discussion and comparison of the E106-*serW* CRISPR system to other CRISPR systems identified in the literature and databases.

The CRISPR system is found in a wide range of bacterial species, it is currently found among 24 individual sequenced species ranging from *P. aeruginosa* PA14, *Enterobacter* spp 638 to *Yersinia pestis*. A CRISPR system consists of two components; the CRISPR array and the CRISPR-associated proteins. The CRISPR array consists of a number of palindromic repeat DNA sequences separated by DNA spacers of species-defined length. These CRISPR arrays are found in close proximity to a set of proteins known as CRISPR-associated proteins (Cas proteins). The CRISPR-associated proteins are divided into two categories. The first category is the core set of *cas* genes of which there are 6 (Peterson *et al.* 2001) named *cas1* to *cas6*. All genomes containing a CRISPR system contain *cas1* and one or more of the additional core *cas* genes (Jansen *et al.* 2002). The second category of genes associated with CRISPR systems define the subtype which are usually species related, there are 7 to 8 subtypes (Sorek, Kunin & Hugenholtz 2008).

The role of CRISPR systems has only recently been elucidated and is shown in Figure 3-11. The CRISPR system is suggested to work in a RNA interference dependent manner (Sorek, Kunin & Hugenholtz 2008). The DNA spacers between the CRISPR repeats have been identified as bacteriophage DNA. The CRISPR system acts to generate small RNAs that bind to complementary bacteriophage DNA which triggers degradation of the bacteriophage DNA. This acts to protect the bacterium from infection.

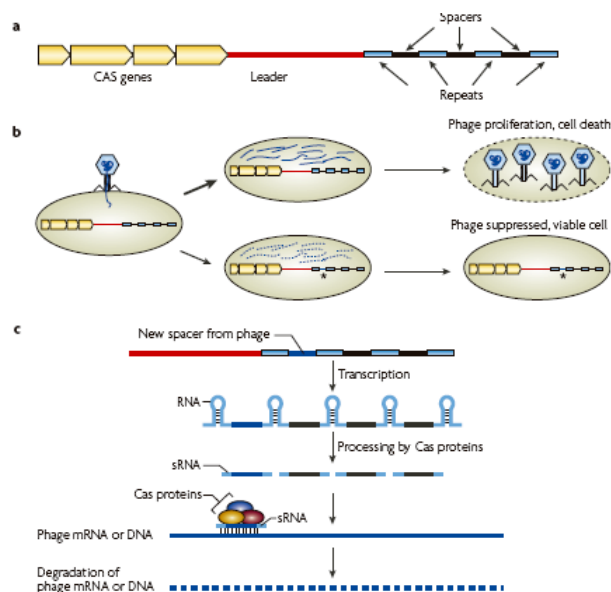


Figure 3-11 shows the structure and predicted function of a typical CRISPR system (Sorek, Kunin & Hugenholtz 2008). Reprinted by permission from Macmillan Publishers Ltd: Nature Reviews Microbiology 2008 Mar; 6(3):181-6, copyright 2008.

The CRISPR system has previously identified in other *E. coli* species (K12 MG1655, W3110, SMS3-5 O157:H7 VT2-sakai and O157:H7 EDL933) (Haft *et al.* 2005, Peterson *et al.* 2001). The CRISPR systems within these strains are structural related and contain the genes *cas1*, *cas2*, *cas3* and *cas5e* and the additional associated genes that encode the *E. coli* subtype proteins *cse1* - *cse4*. This is in contrast to the E106-*serW* and the related CRISPR systems in *E. coli* APEC O1, UTI89 and B7A, which contain *cas1*, *cas3* and the *Y. pestis* subtype proteins (*csy1* - *csy4*). The genomic structure of the CRISPR associated genes are shown in Figure 3-12.

E. coli CRISPR-associated protein gene structure (Ecoli subtype)



Yersinia CRISPR-associated protein gene structure (Ypest subtype)

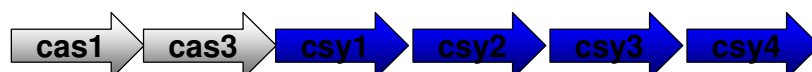


Figure 3-12 shows the genomic structure of the CRISPR-associated protein genes. The top diagram shows the structure of the CRISPR-associated protein genes within *E. coli* strains (K12 MG1655, W3110, SMS3-5 O157:H7 VT2-sakai and O157:H7 EDL933). The bottom structure shows those identified in *E. coli* E106, APEC O1, UTI89 and all sequenced *Yersinia* species.

All the *Yersinia* species available in the Comprehensive Microbial Resource (Peterson *et al.* 2001), contain only Cas1, Cas3 and the *Y. pestis* subtype proteins. The genomic structure of the CRISPR associated proteins are the same as depicted in Figure 3-12. The Cas1 and Cas3 proteins were analysed by amino acid sequence alignment (*blastp*). The Cas1 used APECO1_1211 as the template amino acid sequence and Cas3 used APECO1_1210. The Cas1 protein has showed a more significant protein identity with the *Yersinia* species (~69%) compared with the other *E. coli* (22%). Typical examples of protein alignments *blastp*, *Y. pestis* CO92 (S: 492, E: 2e-128 and I: 69%) and *E. coli* K12 MG1655 (S: 57.4, E: 2e-6 and I: 22%). Cas3 protein also more significantly aligned with the *Yersinia* homologue than the typical *E. coli* group. Typical examples of protein alignments *blastp*, *Y. pestis* CO92 (S: 1088, E: 0 and I: 52%) and *E. coli* O157:H7 EDL933 (S: 32.3, E: 7.6 and I: 39%). This data suggest that the E106-*serW* CRISPR is more closely related to the *Yersinia* systems than those of the typical *E. coli* systems. The CRISPR system within E106-*serW* contains all the genes found within the CRISPR system within the *Yersinia* system and the homologues appear to be more closely related to those of *Yersinia* than the other *E. coli* strains. The CRISPR system is likely to be functional as it contains all the genes required for a functional CRISPR system within *Yersinia*, which has been shown experimentally to active (Pourcel, Salvignol & Vergnaud 2005).

The CRISPR system within *E. coli* E106, APEC O1, B7A and UTI89 are all associated with a serine tRNA site. The literature has supported the notion that tRNA sites are hotspots for genomic island integration (Hacker, Kaper 2000). The data presented here suggest that the CRISPR system genomic island preferentially targeted the *serW* tRNA site. This is supported by analysis of *Enterobacter* spp 638 whose CRISPR system is similarity located and arranged as in *E. coli* E106 (APEC O1, B7A and UTI89); downstream of *clpA* and upstream of *infA*, adjacent to a serine tRNA. Analysing the genomic regions surrounding the CRISPR system within the *Yersinia* species showed that these systems are not associated with a tRNA site nor do they have similar flanking genes to those of E106-*serW*. The CRISPR systems in the other *E. coli* strains are also not associated with the *serW*

tRNA site. This suggests that two different CRISPR systems arose independently within *E. coli*.

The data presented does not differentiate between two scenarios for the island. The first scenario being that the genomic island was acquired via horizontal transfer by an ancestral *E. coli* to *E. coli* E106 and *E. coli* APEC O1, B7A and UTI89. The second scenario is that the genomic island is still mobile and has been horizontally transferred between these *E. coli* strains. The genomic region does contain any known mobility genes, but the literature suggests that the functional role of some CRISPR-associated proteins is to aid horizontal transfer of the CRISPR system (Haft *et al.* 2005). CRISPR systems have been suggested in the literature to be horizontally transferred, as Jansen *et al.* (2002) discovered in a larger scale study that there was no general correlation between the phylogenetic relationship of bacteria species and the CRISPR system present. Despite the method of acquiring the genomic island, the maintenance of a genomic island containing a CRISPR system would be beneficial to the bacteria as it plays a role in bacterial defence against bacteriophage (Sorek, Kunin & Hugenholtz 2008).

In conclusion, E106-*serW* genomic island has a functional CRISPR system that appears to target serine tRNA sites within *E. coli* strains (APEC O1, B7A and UTI89) and closely related species such as *Enterobacter* spp. 638. The system appears to be closely related to the CRISPR system subtype usually restricted to *Yersinia* species than those previously identified and highlighted in the literature for other *E. coli* strains. Analysis highlighted that the *Yersinia* CRISPR system is found within the core genome within the *Yersinia* strains, but when found in the other strains highlighted during this project it is found on a genomic island targeting a serine tRNA site. The data have also demonstrated the usefulness of the technique to compare the genomic islands to those previously identified.

3.4 E105-leuX

E. coli E105 is a clinical isolate recovered from a human septicaemia infection. The strain was originally part of a strain collection used to evaluate if a particular genome profile correlated with disease (Thani A, PhD thesis, University of

Leicester). E105-*leuX* was the first genomic island captured from a clinical isolate using the yeast-based genomic island capture technique. This section will proceed by analysing the preliminary data generated for the genomic island. This is followed by analysis to determine that the insertion at the *leuX* tRNA locus is a genomic island (islet). This is followed by analysis of the genomic structure of E105-*leuX* and concludes with the possible evolution of the genomic island.

Preliminary experiments were performed on the genomic island capture clone to verify the correct recombination event had occurred. Firstly, PCR was performed on the bacterial plasmid preparation to confirm that the plasmid contained TS1 and TS2 as well the positive screening region. The plasmid was also subjected to restriction digestion with *I-SceI* to get an approximation of the size of the captured insert (the genomic island plus the conserved flanks). As the captured insert was only approximately 11kb, it was within the range that could be amplified using long-range PCR. The capture insert was amplified from the genomic DNA of *E. coli* E105 using the *leuX* TS1 U primer and the *leuX* TS2 D primer. The size of the insert from long-range PCR and the *I-SceI* restriction digest are shown in Figure 3-6. The SGSP data provided by Ali Thani were re-analysed; construction of the SGSP library, screening and sequencing was performed by Ali Thani. The following analysis was performed independently.

A 1.4kb SGSP amplicon was generated from the *PstI* genomic library of *E. coli* E105 using the *leuX* U primer and the T3 universal primer. The amplicon was sent for sequencing with the KS universal primer. Results of the analysis are depicted in Figure 3-13.

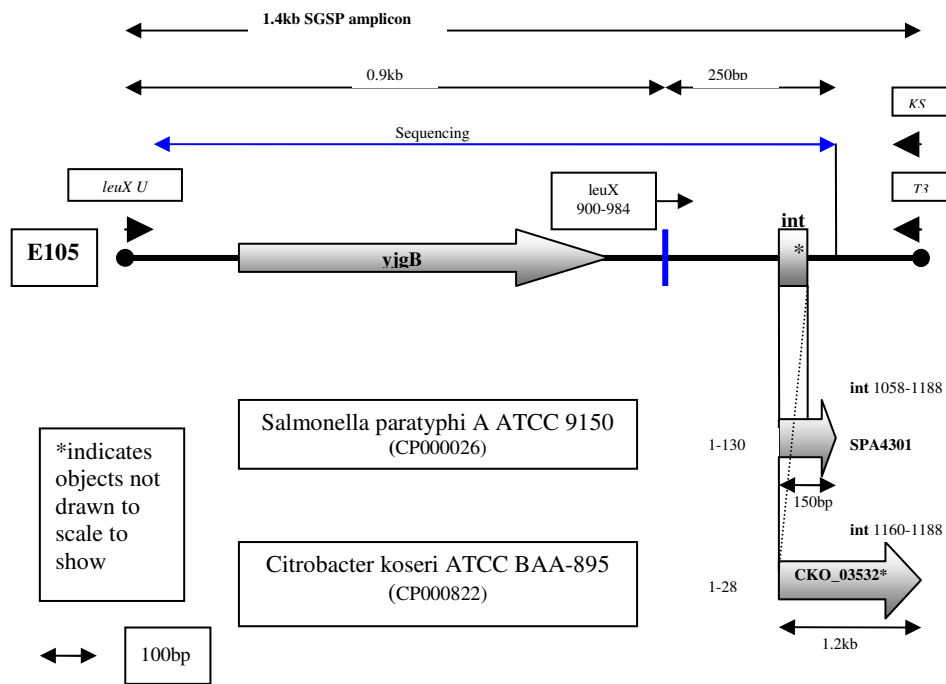


Figure 3-13 depicts the SGSP amplicon generated using the *leuX* U and T3 primers from the *E. coli* E105 *Pst*I genomic library. The most significant alignments are depicted adjacent to the SGSP amplicon.

The first kb of the amplicon corresponded to the conserved U flank for the *leuX* tRNA locus among *E. coli*; containing the gene *yigB* which is predicted to encode a zinc-dependent alcohol dehydrogenase; as well as the *leuX* gene itself. The residual 354bp of the DNA sequence do not align to any sequenced *E. coli* strains. The best nucleotide alignment across the whole SGSP amplicon DNA sequence was to the genome of *Citrobacter koseri* ATCC BAA-895 (S: 818, E: 0 and I: 76%) which encompasses the bases 126 - 1248 within the SGSP amplicon. The additional 354bp not found within the other sequenced *E. coli* strains contained an intergenic region found downstream of the *leuX* gene within *C. koseri* followed by 36bp corresponding to the first 36bp of a 1.2kb P4-like integrase found within *C. koseri*. The SGSP amplicon was also aligned significantly across the whole SGSP amplicon by three *Salmonella* genomes, *Salmonella enterica* subsp. *enterica* serovar *Paratyphi* A str. ATCC 9150 (S:725, E:0 and I:76%), *Salmonella enterica* subsp. *enterica* serovar *Typhi* CT18 and *Salmonella enterica* subsp. *enterica* serovar *Typhi* Ty2. These genomes had an additional 1kb of DNA between the *leuX* tRNA site and the integrase.

The strains highlighted with alignment in the residual 364bp were analysed to assess if the sequence was part of a genomic island. The genomic island in *C. koseri* has yet to be identified in the literature or by a database. The genomic island, if present, did not share a conserved D flank with *E. coli* after an extension of 5kb in any of the sequenced *E. coli* strains. The *Salmonella* strains all have defined and related pathogenicity islands associated with the *leuX* tRNA site (Table 3-6), functionally described as pathogenicity islands containing the *sefA-R* chaperone-usher fimbrial operon.

<i>Salmonella</i> strain	Size	GC%	Name
Paratyphi A	31.4kb	47.20%	-
CT18	32.9kb	46.60%	SPI-10
Ty2	32.7kb	46.60%	-

Table 3-6 Pathogenicity islands defined within *Salmonella* strains that are *leuX*-associated information gathered from Pathogenicity island database (Yoon *et al.* 2007)

A second SGSP amplicon was generated from the *EcoRV* genomic library of *E. coli* E105 using the *leuX* D primer and the T3 universal primer. Sequencing was performed using the KS universal primer and resulted in 750bp of DNA sequence data. The results are depicted in Figure 3-14. Alignment of the full length sequence (bases 1-750) to the D flank of a number of sequenced *E. coli* strains ($\geq 97\%$ identity) as shown in Table 3-7.

<i>E. coli</i> strains	Score	Expected value	Identity
UTI89	1315	0	98%
APEC O1	1315	0	98%
536	1315	0	98%
E24377A	1279	0	97%

Table 3-7 The highest scoring sequences from the nucleotide sequence database based on nucleotide alignment (*blastn*)

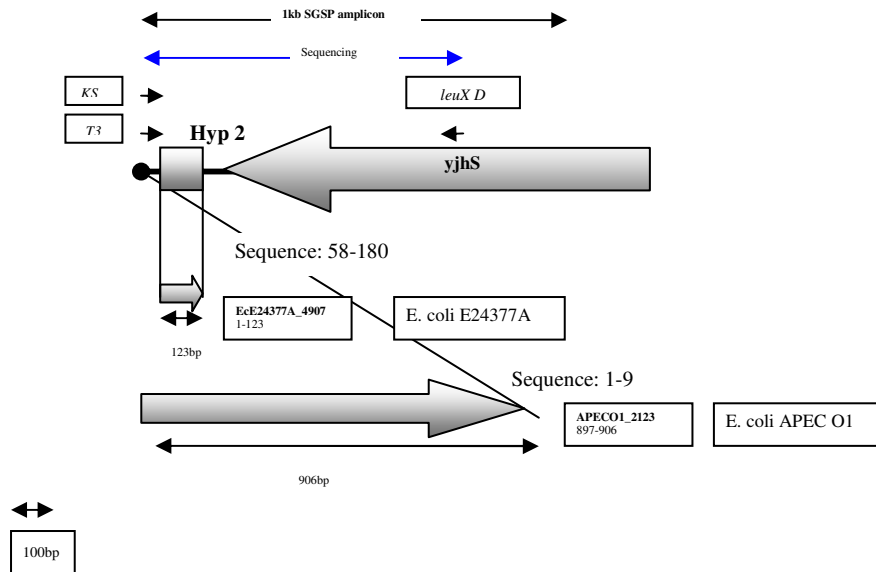


Figure 3-14 depicts the SGSP amplicon generated using the *leuX D* and T3 primers from the *E. coli* E105 *EcoRV* genomic library. The most significant alignments are depicted adjacent to the SGSP amplicon.

Blast alignment revealed that the D flank contains three hypothetical proteins. The first hypothetical protein is only slightly represented; the first 9bp of the SGSP amplicon represents the last 9bp of a conserved hypothetical protein, as suggested by the nucleotide alignment between *E. coli* APEC O1 (S: 326, E: 3e-86 and I:96%) and the SGSP amplicon. The predicted protein has the COG3440, which is putative conserved protein domain found in common with a Type II restriction endonuclease in *Neisseria* spp. The second hypothetical protein spans from nucleotide bases 58-180. The gene is predicted by glimmer to be 123bp long, the region is found in a number of *E. coli* and *Shigella* strains, but not necessarily defined as a gene. Based on *blastn* data for the first 200bp, a protein predicted in *E. coli* E24377A aligns significantly (S: 331, E: 6e-88 and I: 96%). The protein contains a putative conserved protein domain (COG3183) which predicts the protein to encode a restriction endonuclease. An interesting result was the region also partial aligned to a protein predicted by *blastx* analysis in *E. coli* B7A and SMS-3-5 (S: 150, I: 93%, E: 1e-33), which is predicted to be a protein with a HNH endonuclease domain. The alignment for the predicted protein within the latter two *E. coli* strains aligned for the last 55 amino acids of a 216 amino acid predicted length. The third hypothetical protein encoded within the DNA sequence is YjhS (191-748); protein comparisons

predict this protein to be a dnaJ-class molecular chaperone. This gene is part of the conserved D flank of the *leuX* tRNA site.

The preliminary data suggest that the genomic island, E105-*leuX* contains DNA sequence not previously found within *E. coli*. This suggests the possibility of a novel genomic island. The size of the genomic island does not correlate with those identified in *Salmonella* (32kb) in comparison to the E105-*leuX* (11kb). These preliminary results allowed the genomic island to be sent for full length sequencing at the Sanger Sequencing Centre, Cambridge.

The remainder of this section will proceed with analysis to determine that the insertion at the *leuX* tRNA locus is a genomic island (islet). This is followed by analysis of the genomic structure of E105-*leuX* and concludes with the possible evolution of the genomic island.

3.4.1 Defining E105-*leuX* as a genomic island (islet)

This subsection will explore the presence of markers within the DNA sequence of E105-*leuX* that will help define it as a genomic island (islet). The genomic island associated with the *leuX* tRNA site found within *E. coli* E105 is strictly defined as a genomic islet as it is smaller than 10kb at 8.8kb (8,775bp). The function of the genomic islet would be defined as a defence mechanism against bacteriophage infection, as it contains a type III restriction modification system. The genomic island contained a number of markers that are used to define genomic islands (islets).

The genomic islet is inserted at the 3'-end of a tRNA site, *leuX*. A large majority of genomic islands are inserted downstream of tRNA sites (Hacker, Kaper 2000). The literature suggests a number of reasons for this phenomenon. One reason states they are targeted because tRNA sites are a conserved set of genes among all bacterial species and therefore this preference facilitates the propagation of genomic islands between species. Another reason states that genomic islands are remnants of temperate bacteriophage that target the 3' end of tRNA sites for insertion such as the P4 family of bacteriophage (Campbell 2003), *Enterobacteria* phage P4 insertion site

is within the *leuX* tRNA gene. This suggests that genomic islands evolved from the bacteriophage DNA become defective and acquiring new functions (Campbell 2003).

Different species have different codon bias within their genome and this results in a definable G+C content. The codon usage differs between the E105-*leuX* and the genome of *E. coli* K12 MG1655 as shown in Figure 3-15. The G+C content of the genomic islet is 39%, which is lower than the typical 50% of the *E. coli* core genome. Amino acids are coded for by triplet of nucleotide consisting of adenine (A), thymine (T), guanine (G), and cytosine (C). This results in 64 different combinations for 20 different amino acids, known as codons. Multiple codons can code for the same amino acid, but within a particular organism, they will usually have preference for one codon for each specific amino acid. This is known as the codon usage bias and is used to compare segments of genomic DNA for evidence of horizontal transfer. A marked difference suggested the island was a result of a horizontal transfer event and that the event happened recently in the evolution of the strain. A well accepted dogma is that over time the G+C content and codon usage changes to that of the host chromosome as it become more stable and accepted, this is known as amelioration (Ermolaeva 2001).

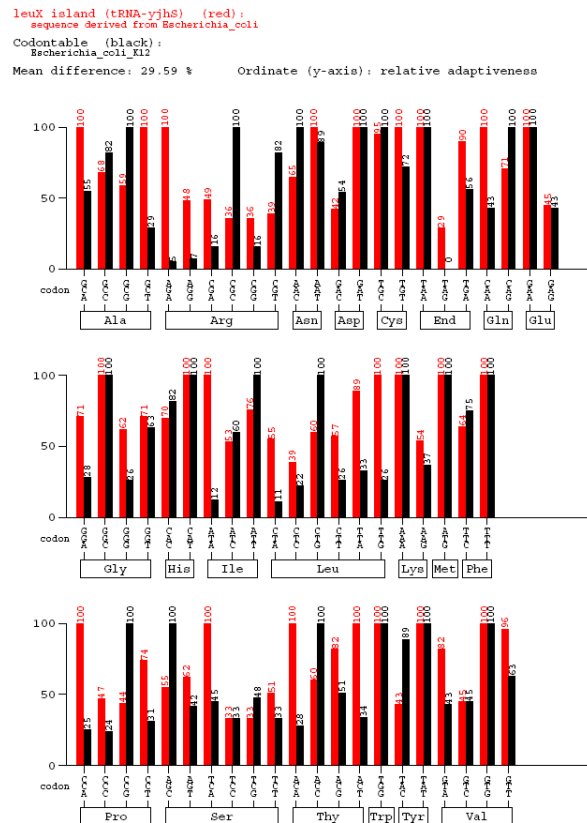


Figure 3-15 depicts the codon usage of the E105-*leuX* genomic islet versus the codon usage table of *E. coli* K12 MG1655. Graph generated by *graphical codon usage analyser* (Fuhrmann *et al.* 2004) and uses the relative adaptiveness as a scale.

The E105-*leuX* genomic islet contains a number of mobility genes, insertion sequence, *IS1* depicted by its two proteins, *insA* and *insB*. It also contains partial DNA sequence for *IS3*, the first 71bp out of total sequence length of 1258bp is found; neither of the two *IS3* ORFs are found within this small region. The E105-*leuX* genomic islet also contains a pseudogene for a P4-like integrase. These mobility genes can contribute to the ability of the genomic island to be propagated to other bacteria as well as contributing to the evolution and selection within the island itself, in the past and possibly in the future.

Another important marker for a genomic island, is direct repeats, usually manifesting as the 3'end of the tRNA which flanks upstream and downstream of the genomic island. No sequence repeat of this kind was found within E105-*leuX*.

In conclusion, E105-*leuX* is a genomic islet as contains a majority of the markers identified in the literature as essential. The next subsection will explore the genomic structure of the genomic islet.

3.4.2 The genomic structure of the E105-*leuX* genomic islet

Full length DNA sequence analysis led to the discovery of 6 genes (ORFs) within E105-*leuX*, they are depicted in Figure 3-16 and described in Table 3-8. A number of these genes were mobility genes, insertion elements and an integrase. The genomic islet is predicted to encode three restriction endonucleases. This subsection will give an overview of the genes present within E105-*leuX*, determining whether they are functional and their role within the genomic islet.



Figure 3-16 shows the genomic structure of the E105-*leuX* genomic islet; 11, 618bp sequence

ORF	Gene range	Blastx	Description
yjgB	Complement 79..1101	yjgB, <i>E. coli</i> DH10B	Putative zinc alcohol dehydrogenase oxidoreductase
leuX	1255..1336	-	<i>leuX</i> tRNA
int	1515..2709	CKO_03944, <i>C. koseri</i> (S: 350, E: 6e-110, I: 68%- 57%-60%)	Partial phage integrase, P4 type, pseudogene, two frameshifts
insA	2979..3746	IS1 InsA, <i>E. coli</i> W3110 (S: 1507, E: 0, I:99%)	Insertion element protein, IS1
insB	3228..3731	IS1 InsB, <i>E. coli</i> W3110 (S: 1507, E: 0, I:99%)	Insertion element protein, IS1
E105leuX _mod	3965..5982	EcoP15I, mod (S: 446, E: 3e-123, I:55%- 73%)	Type III restriction modification system, methylase, pseudogene
E105leuX _res	5985..8894	EcoP15I, res (S: 1816, E: 0, I: 90%)	Type III restriction modification system, restriction endonuclease
IS3	9354..9424	IS3, <i>E. coli</i> MG1655 K12 (S: 937, E: 2e-18, I: 91%)	IS3, partial
E105leuX _1	9452..9865	EcE24377A_4906, <i>E. coli</i> E24377A, (S: 249, E: 3e-65, I: 85%)	Putative restriction endonuclease, partial
E105leuX _2	Complement 9982..10086*	EcE24377A_4907, <i>E. coli</i> E24377A (S: 74, E: 1e-12 and I: 94%)	Putative restriction endonuclease, pseudogene
yjhS	10097..11104	yjhS, <i>E. coli</i> DH10B	Hypothetical protein
yjhT	Complement 11142..11534	yjhT, <i>E. coli</i> DH10B	Putative inner membrane protein

Table 3-8 The ORFs found within E105-*leuX* genomic islet.

Gene range defines where the genes are situated within the available E105-*leuX* sequence. The *blastx* scores were generated against the E105-*leuX* DNA sequence and include the most significant alignment S=Bit-score, E=Expected value and I=Identity percentage * Based on start of alignment and not a start codon.

The ORF marked as *int* was a P4-type integrase. The P4 family of bacteriophages integrate downstream of tRNA sites and this action is facilitated by the integrase protein. The integrase is rendered non-functional by two frameshifts. A frameshift is caused by insertion or deletion of nucleotides from a sequence that is not divisible by 3 (Figure 3-17). The earlier the disruption within the ORF occurs, the more

likely the resulting protein will be altered and in turn increases the likelihood of a non-functional protein.

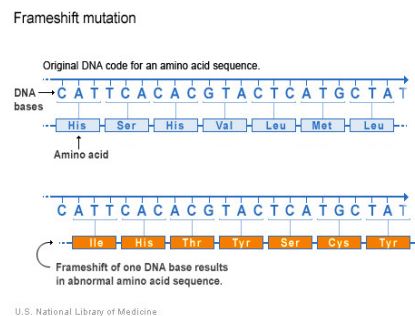


Figure 3-17 depicts the consequences of a frameshift mutation on the encoded protein- Reproduced with permission from US National Library of Medicine - Genetic Home Reference (<http://ghr.nlm.nih.gov/handbook/illustrations/frameshift>) – [Updated 2008 Oct 24; Cited 2008 Nov 3]

Two frameshifts occur within the integrase, the protein is originally encoded in the +1 reading frame (1-56), it shift to the +2 reading frame (58-94) and then shift again to the +3 reading frame (84-391). One of the most significant alignment based on *blastx* was *Citrobacter koseri* ATCC BAA-895 (S: 350, E: 6e-110, I: 68%-57%-60%). The first frameshift is caused by a single nucleotide deletion between nucleotide bases 1689 and 1690 and the second frameshift is caused by a single nucleotide deletion between nucleotide 1772 and 1773. The second frameshift could be an error due to sequencing or replication as it is situated in a string of adenine (A) nucleotides. A second explanation is this is an example of bacterial gene expression control. Ribosomal frameshifting is a method of control that required the ribosome to change reading frame whilst translating the mRNA in order to make a functional protein. This process is promoted by slippery sequences, which contain a string of adenine (A) residues as is found in the second frameshift within the integrase.

Restriction endonucleases are important weapons for bacteria against viral infection. Restriction endonucleases recognise specific nucleotide sequence (recognition sequence) and cut the DNA. These enzymes are maintained within the bacterial cell cytoplasm and if they come into contact with foreign DNA e.g. viral DNA that contains their recognition sequence they will restrict it. Importantly, if these

recognition sequences are found within the bacterial genome they are protected by enzymatic modification should as methylation. This results in restriction endonucleases being partnered with restriction methyltransferases within the genome, referred collectively as restriction modification (RM) system. There are four types of restriction endonuclease (Williams 2003). Type I restriction endonucleases cut DNA approximately 100 nucleotides after the recognition site and requires the use of ATP and S-adenosyl-L-methionine. This enzyme has both methylation and restriction function. Type II restriction endonucleases cut DNA within the recognition site, usually require magnesium and solely have the restriction function. Type III restriction endonucleases cut DNA approximately 20-30 base pairs after the recognition site, require the presences of ATP and exist as a complex with the restriction methyltransferases. Type IV restriction endonucleases target methylated DNA.

E105-*leuX* genomic islet contains a Type III RM system, the predicted mod (methylase) subunit protein consists of 672 amino acids and the predicted res (restriction) subunit consists of 969 amino acids. The most significant nucleotide alignment (*blastn*) for both the components is to the EcoP15I type III RM system from the *E. coli* plasmid p15 (Reiser, Yuan 1977). The two genes within this system are known as EcoP15I res (S: 1816, E: 0, I: 90%) and EcoP15I mod (S: 446, E: 3e-123, I: 55%-73%). There are similar type III RM systems found within *Neisseria meningitidis*, *Neisseria gonorrhoeae*, *Enterobacteria* phage P1 and *Enterobacteria* phage P7, which had similar putative conserved protein domains (COG). None of the currently sequenced *Neisseria* species have type III RM systems within a genomic island structure similar to E105-*leuX*.

The type III RM system within E105-*leuX* has a frameshift within the mod gene. The reading frame of E105leuX_mod changes from +1 to +2 after amino acid 437. Analysis of the conserved protein domains of E105leuX_mod was performed on a reconstituted protein sequence which rendered all the amino acids in the same original reading frame. The conserved protein domain, COG2189, a domain found within adenine specific DNA methylase (mod) aligned from amino acids 4 - 484 within E105leuX_mod (S: 122.94 E: 2e-28). The conserved protein domain spans 590 amino acids. The EcoP15I mod also only spans the first 484 amino acids and

has been proven experimentally to function (Redaschi, Bickle 1996). The frameshift within E105leuX_mod occurs within the conserved protein domain as denoted by the alignment between E105leuX_mod and EcoP15I mod, as it occurs at amino acid 437 which is within the first 484 amino acids. The residual 47 amino acids of the conserved domain could be important for the protein to function as a methylase. The residual amino acids outside of the mod conserved domain could also be essential by adding stability for the secondary and tertiary structure of the protein. The amino acids outside of the mod conserved domain could be also important because one of the functions of a type III RM system mod protein is to bind to the res subunit to enable the restriction activity of the type III RM system. This binding site is likely to be outside of the mod conserved domain. This analysis suggests that E105leuX_mod is non-functional.

The literature suggests that an RM system, with a functional res without a functional mod would be lethal to a bacterium cell. This is because in absence of the complementary mod, the res would restrict the bacterial DNA. In the current situation, the type III RM system, if closely related to EcoP15I, has a number of fail safes. Reaschi *et al.* (1996) reported that when the EcoP15I RM system is introduced into a bacterium, the *mod* gene is expressed immediately and expression of the *res* gene is suppressed until sufficient levels of mod is present within the bacterium. The same research paper also specified that the res subunit did not demonstrate any enzymatic activity unless it was complexed with the mod subunit, adding a further level of protection. The mutation of the last 47 residual amino acids within mod subunit, which are likely to be involved in the binding of res, suggests there would be no functional restriction activity. The most important evidence is *E. coli* E105 grows in culture despite being predicted to have a non-functional mod. Further work to demonstrate that the mod gene was non-functional would be to attempt to digest the genomic DNA from E105 with EcoP1I or EcoP15I but unfortunately none was commercially available, neither was there a type II or type III restriction enzyme with an identical recognition sequence commercially available.

The predicted ORF *E105leuX_1* translated to a protein that is 137 amino acids long. This protein is predicted to have the putative conserved protein domain, COG3440

(S: 150, E: 1e-37). This predicts the function of the protein to be a restriction endonuclease. The conserved domain is defined as 301 amino acids long, and the most significant *blastx* alignments to E105leuX_1 were longer proteins, approximately 300 amino acids compared with 137 of E105leuX_1. One of the most significant alignments was with a predicted protein in *E. coli* E24377A, EcE24377A_4906 (S: 249, E: 3e-65, I: 85%) which is 290 amino acids long and E105leuX_1 aligns from 157-290 amino acids. Interestingly, despite the length of the aligned protein, E105leuX_1 always aligned to the carboxyl terminal of the protein, suggesting the protein become truncated at the 5'-end during a recombination event and therefore it is predicted to be non-functional. This assumption is strengthened by the fact that the remnant of IS3 that is found immediately upstream of this ORF is truncated at the 3'-end, suggesting a deletion event led to the loss of DNA within the region.

E105leuX_2 is a predicted protein 40 amino acids in length and most significantly aligns to EcE24377A_4907 (S: 74, E: 1e-12 and I: 94%) based on *blastx* analysis. The putative conserved domain that aligns to this protein is COG3183 (S: 49 and E: 3e-07) the length of which is 272 amino acids and thereby predicts the function of this protein to also be a restriction endonuclease. This gene has a frameshift in the nucleotide sequence in the first 20bp, meaning that the first 6 amino acids are in the correct reading frame, but the residual amino acids are in a different reading frame and therefore no functional protein is being produced. The frameshift is due to a single deletion of a T nucleotide between 9978 and 9979 in E105-*leuX* sequence. It is possible this is a sequencing error, as this area within a string of 7 T's within the genome of *E. coli* E24377A. This also could be a replication error within the genome of the *E. coli* E105 or could also be a level of gene expression control.

In conclusion, we have explored the genes found within E105-*leuX*. Many appeared to be pseudogenes (producing non-functional gene products) or under some form of gene expression control. The genes found and the arrangement of the genes demonstrates E105-*leuX* to be a novel genomic islet.

3.4.3 Evolution of the E105-*leuX* genomic islet

This subsection aims to give an overview of the hypothetical genomic events that occurred to produce the structure of E105-*leuX* as described in this thesis. The *leuX* tRNA site is a recognised hot-spot for acquisition of horizontal acquired DNA and has been described as a hypervariable region (Bishop *et al.* 2005). The diversity of genomic islands found at this tRNA site within a single strain has been noted in the *Salmonella enterica* (Bishop *et al.* 2005). The *leuX* tRNA site has been noted as a hotspot for bacteriophage integration, especially bacteriophage within the P4 family (Campbell 2003), such as P4 that integrates at the *leuX* tRNA site. E105-*leuX* provides evidence for this theory as genes found within this genomic islet, even the conserved flank, seems to have significant counterparts in bacteriophage that target *E. coli*. The *int* gene is a pseudogene of a P4-like integrase. The area at the 3'-end of *leuX* seems to be under consistent pressure, as the *yjhS* gene has significant similarity to genes found in two bacteriophage, 86 (hosts: *Escherichia coli*) and VT2-Sa (hosts: *Escherichia coli* O157:H7 EDL933; *Escherichia coli* O157:H7) known to target *E. coli* specifically. This *E. coli* specific nature of these two bacteriophages might account for why *Citrobacter koseri* does not have the same D flank (TS2 region) as the *E. coli* that would aid identification of genomic island present. E105-*leuX* seems to have lost genes associated with this region in other *E. coli* strains such as *E. coli* K12 MG1655 has other *yjh* (p,q,x,z) genes immediately upstream of *yjhS*; whereas E105 has only maintained the *yjhS*. This suggests *yjhS* might code for an essential function that selected for its maintenance.

E105-*leuX* has a number of genes that are truncated, such as *E105leuX_1* and *E105leuX_2*, as well as remnants of insertion sequence (IS3), suggesting the *leuX* locus is an active area. The sequence data shows that the 3'-end of the IS3 element and the 5'-end of *E105leuX_1* has been deleted. This suggests that a recombination event occurred between these two genes which also led to their truncation. The E105-*leuX* contains a number of insertion sequences; these are known to aid in selective changes to genomic islands. The stability of the area is possibly improving due to the integrase now being non-functional and loss of the IS3, insertion sequence, but the IS1 is still intact. IS1 is found in its entirety, but there is only copy within the structure, it usually function by to aid changes in a genomic island by

having multiple copies of itself. An important feature of *IS1* is its ability to excise from the genomic DNA and circularise (Sekine *et al.* 1997). *IS3*, the first 71bp of total sequence length of 1258bp is found; none of the two ORFs are found within this small region; but *IS3* might have been important in the early changes in the genomic island structure. The 3'-end of *leuX* still has its core sequence (*attB*) is intact and therefore allows further acquisition of other P4 family prophages and could add to the current genomic islet in a modular fashion.

The evolution of the genomic islet, E105-*leuX* seems to have progressed originally from a lysogenic bacteriophage that integrated into the *leuX* tRNA site. This bacteriophage acquired a mutation in its excision gene; this left the bacteriophage unable to propagate. A research paper by Campbell (2003) supports this notion. It suggests that a genomic island started as a bacteriophage that originally integrated into a tRNA site, but became defective. The genome of the bacteriophage is then subjected to selective pressure to lose bacteriophage genes non-essential to the bacterial cell, but the site continued to collect other foreign DNA. The only remnant defined as a bacteriophage is the integrase (*int*) gene but lacks any other markers to enable identification of the original bacteriophage (Campbell 2003). In the case of E105-*leuX* this would suggest that the defective prophage underwent selection to maintain the type III RM system but promoted the loss of any essential genes for viral propagation such as those required to produce a viral capsid. Interestingly, the genes for *IS1* and the type III RM system are found in the same orientation to each other in the genome of *Enterobacteria* phage P1; separated by essential viral genes such as for the capsid, which would not be needed by the bacterium. This adds strength to the assumption that the DNA sequence of E105-*leuX* is a remnant of similar events. The integrase has also become non-functional as it is no longer needed by the prophage. Remnants of insertion sequences that have aided the evolution are left behind also such as *IS3*.

E105-*leuX* is a novel genomic islet discovered within clinical isolate *E. coli* E105 that was characterised within this section. Analysis of the DNA sequence suggests that the genomic island is a remnant of a larger genomic island that appeared to encode a number of functional restriction endonucleases. The genomic island functioned to protect the bacteria from bacteriophage infection. This work also

shows the success of the genomic island capture technique for the use in characterisation of novel genomic islands. In conjunction with other data generated within this project, the technique has been successfully used with bacterial species other than *P. aeruginosa*.

3.5 Discussion

This project explored the characterisation of pathogenicity islands *in vitro*. The aim of the project was to identify, capture and characterise genomic islands. To achieve this aim involved development of the genomic island capture technique used by Wolfgang *et al.* (2003) and the production of a generic system by targeting bacterial species other than *P. aeruginosa*. To achieve these aims, the development of the methodology was required to expand the use of the technique and to streamline the recovery of the genomic islands.

3.5.1 Genomic island capture as a yeast-based technology

In terms of history of the genomic island capture technique, the current application was piloted by Raymond *et al.* (2002a). Their first contribution to this application was the introduction of a counterselectable marker, cycloheximide into the capture vector (Raymond, Sims & Olson 2002b). They suggested (Raymond, Sims & Olson 2002b) and proved (Raymond *et al.* 2002a) that the use of a counterselectable markers allowed capturing directly from genomic DNA by increasing the percentage of clones that harboured a capture vector that had undergone homologous recombination as opposed to those that harboured only the capture vector. Raymond *et al.* (2002a) also piloted the first use of the technique to capture variable genomic regions (O-antigen LPS region) from *P. aeruginosa* and were the first to report using bacterial DNA in the capture process. Wolfgang *et al.* (2003) adapted this technique further by modifying the structure of the capture vector. Wolfgang *et al.* (2003) also used the method to target a variable region within the *P. aeruginosa* genome and highlighted its use as a technique to capture genomic islands. They constructed and used a capture vector to target the *lys10* tRNA site (p0975-0989capture) and captured genomic islands from a number of *P. aeruginosa* strains. The technique as proposed by Wolfgang *et al.*, 2003 is the basis of this project.

The capture vectors used during this project were constructed using the Wolfgang *et al.*, 2003 methodology. Six capture vectors were constructed targeting tRNA sites in both *P. aeruginosa* and *E. coli*. Five of which were designed during this project, the additional capture vector *lys10* is described in Wolfgang *et al.* (2003). Wolfgang *et al.*, 2003 designed their capture vectors by only specifying the requirement for TS1 and TS2 to be 1kb DNA regions upstream and downstream of the target region. During this project, the process was streamlined by utilising the MobilomeFinder bioinformatics tools (Ou *et al.* 2007). MobilomeFinder was able to define a 2kb conserved region flanking upstream and downstream of the tRNA site, of all (majority) of strains within a species with a GenBank available genome. This ensures the capture vector usefulness in a generic system, as it could be used against all strains within the target species and not solely the reference genome strain. Admittedly, this would have been difficult at the time for Wolfgang *et al.*, as there was only one *P. aeruginosa* genome sequence publicly available, *P. aeruginosa* PAO1. This could be considered as a limitation of the technique. The designing of the capture vectors requires at least one genome sequence being available and further genomes to increase the usefulness of the capture vector. The capture vector design protocol for this project also explicitly states that either TS1 or TS2 should contain a partial gene as this aids the design of the screening primers (Section 2.2.3). The assembly of the capture vector was also streamlined, as the recovery of the capture vector was not explicitly explained. A protocol was devised based on a research paper by Raymond *et al.* (2002b) to pool the colonies and perform a single yeast plasmid preparation (Section 2.2.7). Another technique that was intended to streamline the recovery of the capture vector was direct electroporation. This involved transferring the capture vector directly from the yeast colonies to *E. coli* by placing them together within an electroporation cuvette. These experiments were not as efficient as published in the literature (Marcil & Higgins 1992) and the technique was not included as part of the final methodology. The current structure of the capture vector, as a low copy yeast vector seems to impede this. This could be investigated further in future work. The methods and protocols in this thesis should streamline the process of design, construction and recovery of capture vectors.

Recent literature has suggested techniques to improve the efficiency of transformation-associated recombination (TAR) cloning by re-structuring the

capture vector. Noskov *et al.* (2003a) suggested that having a large region of non homology tagging either of the targeting sequences can impede the capture process by decreasing the likelihood of homologous recombination. It is debatable whether this research is relevant to the technique described in this project. They examined a variant of the TAR cloning technique that was used on eukaryotic DNA. They were also using two targeting sequences 60bp in length. They found tagging these targeting sequences with non-homologous DNA (such as the counterselectable marker) reduced efficiency of recovering positive clones. They noticed a 5 fold decrease after adding 37bp of non-homology DNA sequence. The technique described in this thesis, uses targeting sequence 1kb in length and the TS2 has 2.9kb of non-homologous DNA tagging it. Based on these observations the technique described in this thesis should not recover any positive clones; but as demonstrated by this work and results of Wolfgang *et al* the technique does result in the recovery of positive colonies. This discrepancy could be due to the use of a larger targeting sequence, which is approximately 10 times larger. Further work could be used to assess whether making changes to the structure of the capture vector as stated by Noskov *et al*, 2003a would increase the efficiency of recovering positive colonies. An increase in capture efficiency could make it possible to capture E105-*serT* and PA14-*lys47*.

The technique has previously implied for use across a number of bacterial species (Noskov *et al.* 2003a). Previous literature has only designed capture vectors against *P. aeruginosa* (Raymond *et al.* 2002, Smith *et al.* 2005, Wolfgang *et al.* 2003), this project has now expanded its use to *E. coli* and aided the construction of a capture vector for *Acinetobacter baumannii* (Shaikh, F, PhD thesis, University of Leicester, unpublished). The available literature suggests this thesis describes the first experimental account of the technique being used on bacterial species other than *P. aeruginosa* and first recorded use of this technique outside the USA. This supports the project aim to create a generic genomic island capture technique, by expanding the technique to other bacterial species.

The project has also highlighted obstacles with spheroplast transformation such as the quality and quantity of genomic DNA and the percentage of spheroplasting. A recent publication by Leem *et al.* (2003) analysed optimum conditions for TAR

cloning and their findings was incorporated into the methodology. During a controlled experiment to determine the optimal amount of genomic DNA. A trend was noted that increasing the amount of DNA, led to an increased recovery of colonies which then led to a plateau. Leem *et al* also saw this trend as increasing the amount of genomic DNA, increased numbers of colonies up to 2µg, then a plateau, finally leading to decrease in colony numbers. The paper also highlighted an idea that capture vectors recombine preferentially with the ends of the targeted genomic DNA fragments. This idea formed the basis for the experimental attempt to capture PA14-*lys47* (PAPI-1). The aim was to create a DNA fragment that contained PAPI-1, to achieve this aim restriction enzymes that cut at the flanks of the pathogenicity island were needed. Unfortunately, this was not possible as there were no commercially available enzymes that would cut exclusively at the flanks. The experiment was re-designed to include restriction enzymes that did not cut within PAPI-1. If they became available in the future, it would be possible to show that the island could be captured using this method. Another finding from the paper stated that shearing the genomic DNA to make smaller fragments increased the transformation efficiency. They reported an increase in efficiency of 15 times, when the average DNA fragment was reduced from 1Mb to 200kb. The data generated during this project supports this finding. The experiments aimed at analysing the genomic preparation methods used two genomic DNA preparation kits. Use of the *Illustra* genomic DNA preparation (≥20kb) preparation led to a larger number of colonies than that of the *ArchivePure* kit (150-200kb) in capture experiments involving PAO1-*lys10* (11kb). The findings within the paper suggest that the fragment sizes within the genomic DNA preparation, *ArchivePure* are within the optimum size range for TAR cloning which is used incorporated within the final methodology.

The project also refined the protocol for screening yeast colonies, by reviewing the literature and also incorporating techniques from other fields. In molecular biology, it is a well-established technique to create a multiplex PCR for difficult PCRs to reduce the chances of fail-negatives. The use of a second set of primers to highlight false-negatives allowed the method of directly adding a portion of the yeast colony to the PCR mixture to be reinstated, which in turn streamlined the methodology.

In summary, this project has further developed the genomic island capture technique by improving the methodology. This involves use of bioinformatics tools to rapidly design the targeting regions within the capture vector. The targeting regions are designed taking into account all GenBank available genomes ensures the usefulness of the resultant capture vector. Additional protocols were developed and incorporated to streamline the recovery of the capture vector after construction. A multiplex PCR was designed to streamline the screening PCR process and to decrease the likelihood of false negatives.

3.5.2 Genomic island capture as a genomic island discovery tool

Our research group works on the discovery of genomic islands. The research direction has been to work towards a high-throughput system. The system requires a fusion between bioinformatics and wet-bench science. MobilomeFinder (Ou *et al.* 2007) is a set of online bioinformatics tools that aid the discovery of genomic islands *in silico* and *in vitro*. Use of these tools has helped streamline the process of capture vector construction. These tools have also aided in the past the selection of tRNA sites to investigate using the tRIP PCR methodology by identifying tRNA sites that are more likely to have an insertion within a specific bacterial species.

The tRIP PCR screening method is a negative PCR system to screen tRNA sites for the presence of genomic islands or other insertion events. The basic principles are that the primers are designed within the conserved region upstream and downstream of the tRNA site. If the tRNA site is empty a PCR amplicon is generated. If an insertion has taken place then the product will be too large to be generated within the PCR conditions and therefore resulting in no amplicon. Those tRNA sites that are highlighted as tRIP negative are investigated further by a chromosomal walking technique known as SGSP PCR. This involves creating a genomic library and using the primers designed against the conserved region surrounding the tRNA site along with universal primers to generate amplicons that are sequenced. This generates sequence data of the putative island. In conjugation these techniques used together have been shown to be an efficient system for genomic island discovery (Ou *et al.* 2007).

The genomic island capture method can be added as a complementary tool to MobilomeFinder system, along with tRIP PCR and SGSP-PCR. SGSP-PCR is a method for gathering preview information of a genomic island, but it is restricted to analysing the islands starting at the flanking regions and working its way inwards. This can be a slow and repetitive process. The genomic island capture method has advantages over this method as it can be immediately apparent how large the island is by restriction digestion with *I-SceI*. The island is also immediately available for analysis such as RFLP and subcloning, as well as allowing functional studies of the genomic island. With this technique now available future researchers will be able to go straight from detecting a novel genomic island to having the island sequenced and available for functional studies.

Another great feature is that a single capture vector can be constructed for a single locus and used on a panel of strains of the same bacterial species. In some cases the capture vector can even be used cross-species for example, *Shigella*, *Salmonella* and *E. coli*. The capture vectors, *leuX* and *serW* which were designed for *E. coli* could also be used with *Salmonella* and *Shigella* as they share conserved regions surrounding these tRNA sites. A limitation to the technique could be allele variation between bacterial strains but this is diminished by using 1kb of targeting DNA sequence and successful recombination requires 75% DNA homology (Noskov *et al.* 2003b). The technique can not compensate for a flank that undergone DNA inversion and research has also shown that the U flank is well conserved among bacterial strains within the same species, but the D flank has been shown to be more flexible (Ou *et al.* 2006).

This technique as described still has limitations as attempts to capture two targets, PA14-*lys47* and E105-*serT* was unsuccessful. Literature suggests that TAR cloning can capture chromosomal regions up to 300kb (Noskov *et al.* 2003a). My current technique has only successfully captured 42kb (E101-*leuX*). PA14-*lys47* was 108kb pathogenicity island, PAPI-1. It is possible the island was approaching the limits of the technique, Wolfgang *et al.*, 2003 reported capturing a 80kb genomic island. It must also be taken into consideration the island has been reported to be unstable within the bacterial chromosome (Qiu, Gurkar & Lory 2006). Further experiments could be performed using my current technique with known stable islands that

exceed 42kb. Time constraints restricted this work being achieved during this project; as the current selection of capture vectors and strains available were not comparable with this aim.

The most important outcome of this project is the novel genomic island discovery of E105-*leuX*. This is the first record of the use of the genomic island capture technique to capture a genomic island from a bacterial species other than *P. aeruginosa*. The technique was successfully used to compare genomic islands to previously defined island and determine similarity, for example PAO1-*lys10* versus KR115-*lys10* and KR159-*lys10*. The PA14-*lys10* has been introduced into PA14 deletant mutants in experiments performed by Ewan Harrison (Harrison E, PhD thesis, University of Leicester, unpublished) in a complementary experiment. This molecular biology tool is very important as the targeting regions will in most cases be conserved among majority of the members of the same species. This allows for comparative studies of these regions. The technique has now opened doorways for allowing the roles of genomic islands to be elucidated via functional studies. The genomic islands can be introduced into strains that lack them and experiments to examine the increased fitness of the bacteria could be performed. These kinds of studies could be important for pathogenicity islands, to examine how they contribute to the virulence of a pathogen. These experiments could be used to support the central dogma, pathogenicity island cause non-pathogens to become pathogens. Functional experiments could be performed with E105-*leuX* to determine if the type III RM system is functional. A number of approaches could be considered such as purifying the mod and res subunit proteins and performing *in vitro* assays. Another approach is preparing the genomic DNA of *E. coli* E105 and attempting to restrict the DNA with enzymes that such as EcoP15I to see if the methylation has protected the genomic DNA.

3.6 Conclusion and further work

The aim of the project was to produce a generic genomic island capture method. This has been achieved by capturing genomic islands from a number of bacterial species *P. aeruginosa* and *E. coli* as well supporting work to capture genomic islands from *A. baumannii* (Shaikh F, PhD thesis, University of Leicester,

unpublished). This was achieved by developing the methodology associated with the genomic island capture technique. The project has continued the exploration of the use of this yeast-based technology on bacterial genomes and variable regions of genomic DNA. In previous application, the method have been used to capture variable regions from *P. aeruginosa*, O-antigen region (Raymond *et al.* 2002), pyoverdine region (Smith *et al.* 2005) and the *lys10* tRNA site (Wolfgang *et al.* 2003). Having a generic genomic island capture technique will be a useful tool for exploring the mobilome within a variety of bacterial species. Future work could include constructing a complete tRNA set of capture vectors to explore the whole genome. The use of the genomic island technique allowed the capture and characterisation of a novel genomic island E105-*leuX*.

The next step would be to use the captured islands in functional studies. This could help elucidate the true role of the genomic (pathogenicity) islands in the survival of bacteria. Further experiments could be done to improve my generic genomic island capture method such as Noskov *et al.*, 2003 (Noskov *et al.* 2003a) suggested that having a large region of non homology tagging the targeting sequence can impede the process and decrease the likelihood of homologous recombination. Further experimental work could be done to compare the capture efficiency of a redesigned capture vectors to the ones constructed during this project. An increase in capture efficiency could make it possible to capture E105-*serT* and PA14-*lys47*. Also further experiments could be performed to elucidate the true reasons for the obstacles to capturing islands exceeding >43kb.

4 The contribution of PAPI-1 and PAPI-2 to the virulence of *Pseudomonas aeruginosa* PA14 in a murine acute respiratory infection model

<u>4 THE CONTRIBUTION OF PAPI-1 AND PAPI-2 TO THE VIRULENCE OF <i>PSEUDOMONAS AERUGINOSA</i> PA14 IN A MURINE ACUTE RESPIRATORY INFECTION MODEL</u>	94
4.1 CONSTRUCTION OF THE DELETANT MUTANTS AND <i>IN VITRO</i> DATA	95
4.2 PRELIMINARY STUDIES OF CHRONIC RESPIRATORY INFECTION MODEL	97
4.3 PRELIMINARY STUDIES OF ACUTE RESPIRATORY INFECTION MODEL	100
4.4 THE CONTRIBUTION OF PAPI-1 AND PAPI-2 TO THE VIRULENCE OF <i>PSEUDOMONAS AERUGINOSA</i> PA14	102
4.4.1 SURVIVAL TIMES POST-INFECTION	103
4.4.2 SYMPTOMS SCORES POST-INFECTION	103
4.4.3 BACTERIAL NUMBERS IN NASOPHARYNX, LUNGS AND BLOOD POST-INFECTION	104
4.4.4 INTRAVENOUS INFECTION	107
4.4.5 IMMUNE PROFILE	110
4.4.6 LUNG HISTOPATHOLOGY POST-INFECTION	112
4.4.7 PROGRESSION OF DISEASE POST-INFECTION WITH PA14 Δ PAPI-1 Δ PAPI-2	115
4.5 SUMMARY OF ACUTE RESPIRATORY OUTCOMES	118
4.6 DISCUSSION	119
4.6.1 THE ROLE OF PAPI-1 DURING ACUTE RESPIRATORY INFECTION	119
4.6.2 THE ROLE OF PAPI-2 DURING ACUTE RESPIRATORY INFECTION	124
4.6.3 PAO1 AS VIRULENT AS PA14 Δ PAPI-2 IN AN ACUTE RESPIRATORY MODEL	125
4.6.4 THE IMMUNE RESPONSE DURING ACUTE RESPIRATORY INFECTION	128
4.7 CONCLUSIONS AND FURTHER WORK	129

Pseudomonas aeruginosa PA14 (UCBPP-PA14) is a wild-type reference strain that has been used extensively to test the contribution of virulence factors to disease. The *P. aeruginosa* strain PAO1, another wild-type reference strain, has been noted to be less virulent than *P. aeruginosa* PA14 in a number of models. A recent study (He *et al.* 2004) suggested the differences in virulence was due to the presence of two pathogenicity islands found in PA14 that were absent in PAO1, named *Pseudomonas aeruginosa* pathogenicity island 1 (PAPI-1) and *Pseudomonas aeruginosa* pathogenicity island 2 (PAPI-2).

The aim of the project is to compare the contribution of PAPI-1 and PAPI-2 to the virulence of *P. aeruginosa* PA14 within a murine acute pneumonia model. Importantly, this is the first project to look at the effects of pathogenicity islands on *in vivo* respiratory infection. To achieve this aim, three isogenic pathogenicity island deletant mutants were created in the strain *P. aeruginosa* PA14, a mutant with PAPI-1 deleted, another with PAPI-2 deleted and a third mutant with both PAPI-1 and PAPI-2 deleted. The PA14 mutants and wild-type PA14 were also compared against PAO1. Herein, the mutants will be referred to as Δ PAPI-1, Δ PAPI-2 and Δ PAPI-1 Δ PAPI-2.

This results chapter proceeds by describing the mutants used, followed by results generated whilst developing the murine infection models, the next section describes the results generated using an acute respiratory model to compare virulence and ends with the discussion and conclusions generated from this project.

4.1 Construction of the deletant mutants and *in vitro* data

Three isogenic pathogenicity island mutants were created by Ewan Harrison (University of Leicester, PhD, unpublished); Δ PAPI-1, Δ PAPI-2 and Δ PAPI-1 Δ PAPI-2. The first mutants of its kind reported in *P. aeruginosa*. They were created using a mutagenesis system developed by Choi and Schweizer (2005) specifically targeting *P. aeruginosa*. A number of *in vitro* test were performed on PA14 and its isogenic mutants by Ewan Harrison, to assess swarming mobility, twitching mobility, pyocyanin production, biofilm production and growth characteristics. There were no significant differences reported between PA14 and

any of its isogenic mutants under the conditions used within our laboratory. There data is in keeping with He *et al.* (2004) who reported no differences between wild-type PA14 in growth, production of pyocyanin, *in vitro* adhesion, swimming, twitching and swarming motilities and any of the PAPI-1 ORF mutants.

Growth curves were generated for PA14 and its isogenic mutants by inoculating fresh liquid media with a set starting OD and then the OD was measured hourly for 24 hours time period. No differences in growth characteristics suggest that the resulting isogenic mutants are equally as fit as the wild-type, PA14. Swarming and twitching are two types of mobility used by bacteria. Swarming uses the bacterial flagella to move across solid surfaces. The test involved inoculating bacterial cells in the centre of a minimal media plate that has a low agar content (~0.5%), using a toothpick. The variable parameter is the distance the bacteria have moved across the plate after growth overnight. Twitching is a type of movement that uses the type IV pili. The test is similar to that of swarming except the use of LB media plates that has an agar content of 1%. PAPI-1 has genes that have been annotated as encoding a putative type IV B pilus. The results do not show a difference between PA14 and the two mutants without PAPI-1; Δ PAPI-1 and Δ PAPI-1 Δ PAPI-2. This result could be due to the presence of the additional type IV gene cluster within the core genome. Pyocyanin is a virulence factor of *P. aeruginosa* that is excreted into the extracellular environment. The genes involved in pyocyanin production and its regulatory factors are found within the core genome, therefore no difference is expected. Biofilm is an extracellular matrix of polysaccharide and bacteria that aids attachment to solid surfaces. The test to determine biofilm formation involves growing a set density of cells in liquid media within a 96-well PVC plate and analysing biofilm growth via straining with crystal violet and quantification by OD reading. PAPI-1 contains the gene PvrR that is involved in the phenotypic switching of colony morphology types that correlates with biofilm formation. There are two basic types of *P. aeruginosa* colony, normal and small-colony variant (SCV). Under stressful condition such as the presence of antibiotics, there is an increase in colonies exhibiting the SCV compared with normal colony morphology. SCV produce larger amount of biofilm and produce it more rapidly than the normal colonies. The protein, PvrR acts to switch the SCV to the normal phenotype. Conclusions are that the conditions used during this test could be considered 'un-

stressful' to the bacteria and therefore PA14 and the isogenic mutants are all growing as normal colonies.

4.2 Preliminary studies of chronic respiratory infection model

An aim of the project was to develop a chronic infection model that more closely mimicked a natural infection and does not involve invasive procedures as previously published models. Previous models have involved using *P. aeruginosa* embedded in agarose beads and being infected intratracheally (Stotland, Radzioch & Stevenson 2000) or infecting drinking water with *P. aeruginosa* (Coleman *et al.* 2003).

The first challenge dose attempted was 2×10^4 CFU/50 μ l with 4 mice (Figure 4-2). The mice exhibited no signs of illness 48 hours post-infection and two mice were culled and their tissues recovered. *P. aeruginosa* was recovered from the nasopharynx and lungs of both mice, so the residual mice were monitored till 7 days post-infection. The tissues were recovered and *P. aeruginosa* was detected. The results appeared to indicate that there was an increase in CFU within the nasopharynx and lungs from 48 hours to 7 days post-infection. Repeat experiments were conducted but similar results were not achieved (Figure 4-2, Figure 4-3). The first experiment used mice that were 8 weeks old. The second experiment used mice that were 12 weeks old. To determine if this was the variable factor, the third experiment was performed using 8 week old mice.

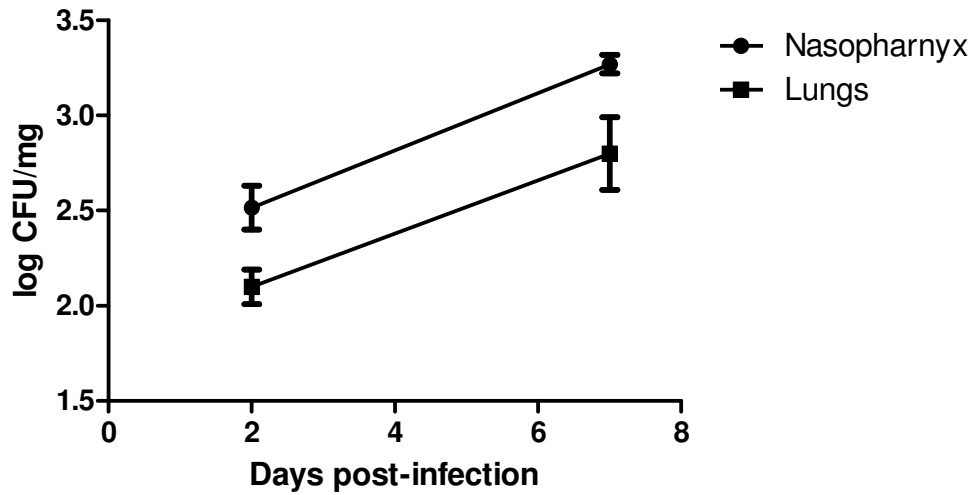


Figure 4-1 shows the CFU numbers recovered from the nasopharynx and lungs at 2 and 7 days post-infection with *Pseudomonas aeruginosa* PAO1 at 2×10^4 CFU/50 μ l; $n=2$ mice for each timepoint. The dot represents the mean log CFU/mg and the error bar represents the standard error of the mean.

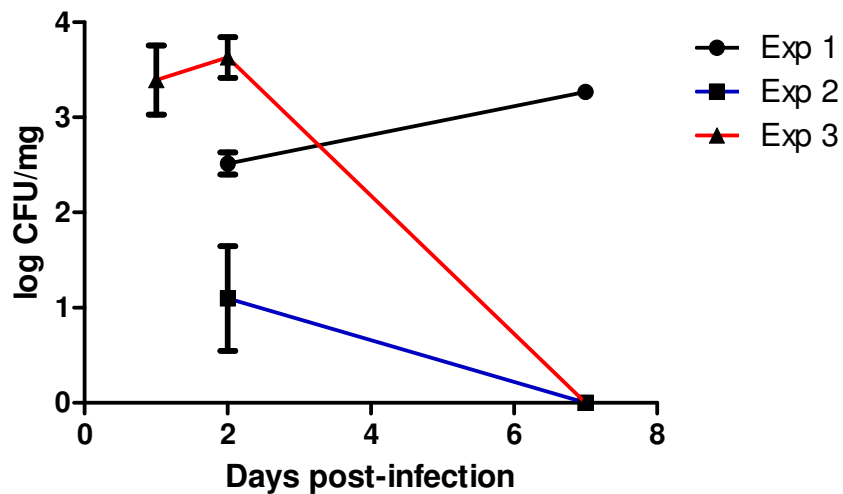


Figure 4-2 shows the CFU numbers recovered from the nasopharynx at 2 and 7 days post-infection with *Pseudomonas aeruginosa* PAO1 at 2×10^4 CFU/50 μ l. Experiment 1 $n=2$ for each timepoint. Experiment 2 $n=3$ for each timepoint. Experiment 3 $n=5$ for day 1 and $n=3$ for all other timepoints. The symbol represents the mean log CFU/mg and the error bar represents the standard error of the mean. Experiment 1 and 3 mice are 8 weeks old commencing experiment and experiment 2 mice that were 12 weeks old commencing experiment.

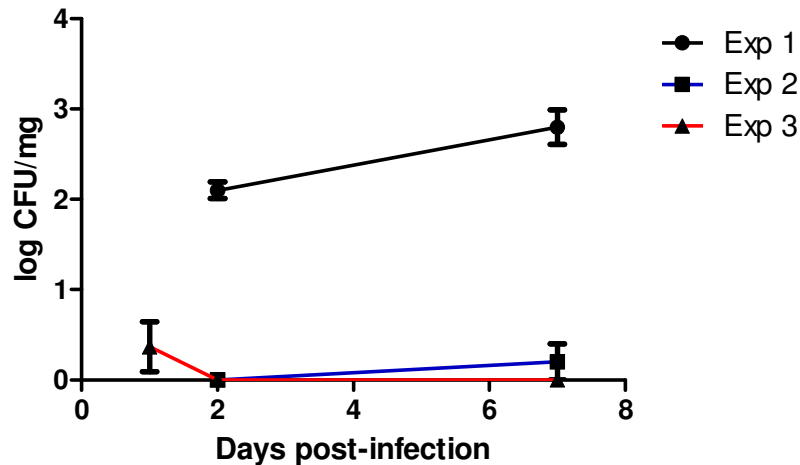


Figure 4-3 shows the CFU numbers recovered from the lungs at 2 and 7 days post-infection with *Pseudomonas aeruginosa* PAO1 at 2×10^4 CFU/50 μ l. Experiment 1 $n=2$ for each timepoint. Experiment 2 $n=3$ for each timepoint. Experiment 3 $n=5$ for day 1 and $n=3$ for all other timepoints. The symbol represents the mean log CFU/mg and the error bar represents the standard error of the mean. Experiment 1 and 3 mice are 8 weeks old commencing experiment and experiment 2 mice that were 12 weeks old commencing experiment.

A challenge dose of 2×10^4 CFU/50 μ l was also attempted with PA14 and Δ PAPI-1 Δ PAPI-2. Mice infected with PA14 had undetectable levels of bacteria within the nasopharynx and lungs at any of the set timepoint. The results of mice infected with Δ PAPI-1 Δ PAPI-2 are shown in Figure 4-4. CFU numbers were only detected in the nasopharynx up to 48 hours post-infection.

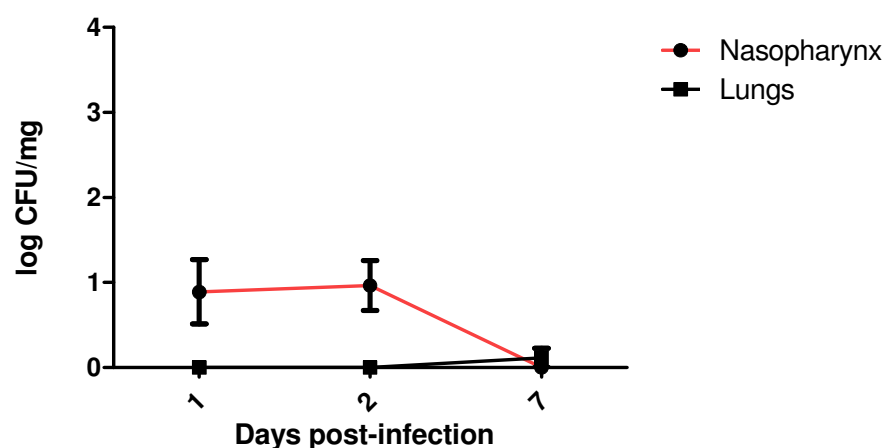


Figure 4-4 shows the CFU numbers recovered from the nasopharynx and lungs at 1, 2 and 7 days post-infection with *Pseudomonas aeruginosa* Δ PAPI-1 Δ PAPI-2 at 2×10^4 CFU/50 μ l; $n=5$ mice for each timepoint. The symbol represents the mean log CFU/mg and the error bar represents the standard error of the mean.

A challenge dose of 2×10^5 CFU/50 μ l was attempted with PA14 and the results are shown in Figure 4-5. CFU numbers in the nasopharynx and lungs were detected up to 48 hours post-infection.

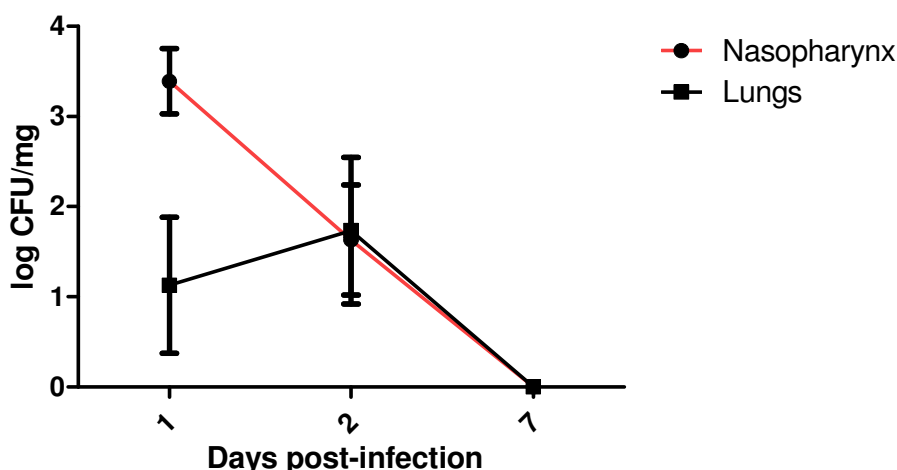


Figure 4-5 shows the CFU numbers recovered from the nasopharynx and lungs at 1, 2 and 7 days post-infection for *Pseudomonas aeruginosa* PA14 at 2×10^5 CFU/50 μ l ; $n=5$ mice for each timepoint. Each symbol the mean log CFU/mg and the error bar represents the standard error of the mean.

In summary, a pilot experiment with PAO1 at a challenge dose of 2×10^4 CFU/50 μ l allowed recovery of bacteria 7 days post-infection. Subsequent repeat experiments did not achieve this initial success. Similar experiments were attempted with PA14 and Δ PAPI-1 Δ PAPI-2. The chronic infection model was shown not to be robust and the subsequent experiments focussed solely on the acute respiratory model of infection.

4.3 Preliminary studies of acute respiratory infection model

The acute respiratory model used during this project is modified from Shaver *et al.* (2004). To develop a murine acute respiratory model, preliminary experiments were conducted using *P. aeruginosa* PAO1. The pilot studies involved determining the dose and the organs to recover for further analysis. The mice were symptom scored and the organs were recovered at 18 hours post-infection. The challenge dose was 2×10^6 CFU/50 μ l. The results of the pilot experiment are shown in Figure 4-6. The

mice exhibited visible symptoms at 18 hours post-infection with an average symptom score of 4. The maximum symptom score a mouse can achieve is 6. A small survival experiment resulted in the mice surviving till approximately 24 hours post-infection. The recovery of the bacteria from the nasopharynx, lungs and blood were selected for assessment of virulence.

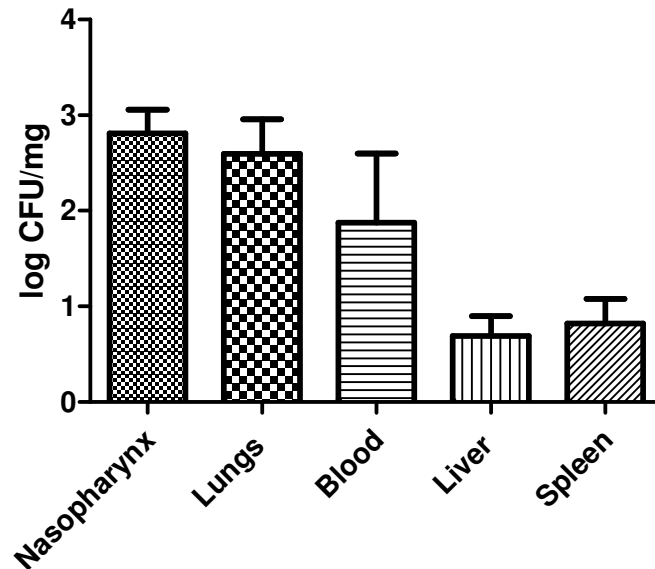


Figure 4-6 shows the CFU numbers recovered from the nasopharynx, lungs, blood, liver and spleen at 18 hours post-infection with *Pseudomonas aeruginosa* PAO1 at 2×10^6 CFU; $n=8$ mice. The bars represent the mean log CFU and the error bar represents the standard error of the mean. Scale for blood is log CFU/ml. All data was generated over 2 independent experiments for each strain.

The acute respiratory model was then tested with *P. aeruginosa* PA14. As previously stated, this strain has been described as being more virulent than PAO1 in a number of infection models. Figure 4-7 shows the difference in CFU number recovered from the chosen tissues and clearly demonstrates that the chosen tissues can differentiate virulence at this dose as there was significant differences ($P < 0.05$) between PAO1 and PA14 ($P < 0.05$) across all the chosen tissues. The symptom scores were an average of 6 for PA14 compared to an average of 4 for PAO1. The differences would allow for the testing of the contribution of PAPI-1 and PAPI-2 to differences in virulence between the two strains.

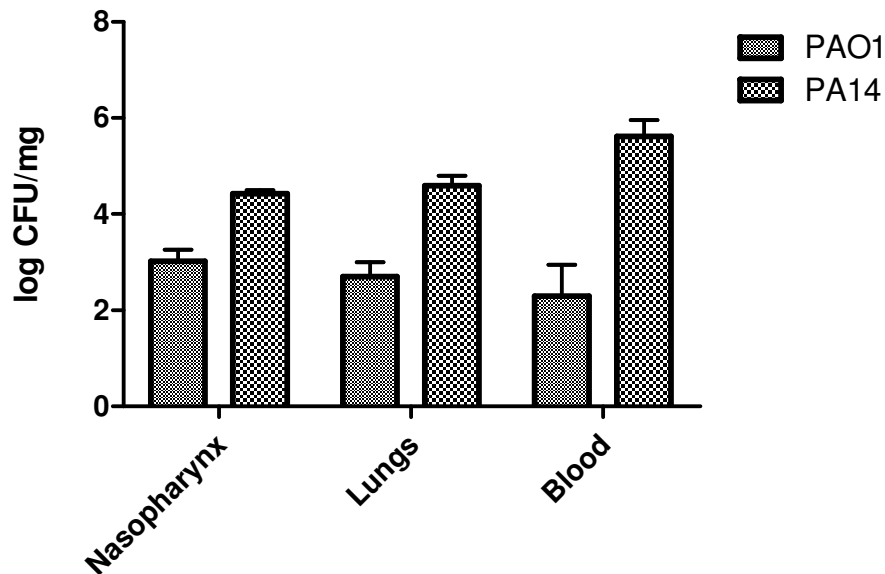


Figure 4-7 shows the CFU numbers recovered from the nasopharynx, lungs, blood at 18 hours post-infection for *Pseudomonas aeruginosa* PAO1 and PA14 at 2×10^6 CFU; $n=10$ mice for both strains. The bars represent the mean log CFU and the error bar represents the standard error of the mean. Scale for blood is log CFU/ml. All data was generated over 2 independent experiments for each strain.

4.4 The contribution of PAPI-1 and PAPI-2 to the virulence of *Pseudomonas aeruginosa* PA14

The experimental endpoint for the acute respiratory model was set at 18 hours. At this timepoint the mice were subjected to symptom scoring and tissues recovered. The nasopharynx and lungs were homogenised and serially diluted to estimate CFU numbers. Afterwards the lung homogenate was also used for leukocyte counts. For histopathology, additional mice were infected and lungs taken and fixed to be prepared for sectioning. For the survival experiments, mice were infected with the acute challenge dose and culled when they reached a maximum symptom score of 6.

This section proceeds by describing the virulence differences between the *P. aeruginosa* strains in terms of survival of the mice post-infection and then in terms of symptom scores at 18 hours post-infection. This is followed by CFU numbers recovered from the mice 18 hours post-infection, cellular infiltration into the lungs and the histopathology caused by the bacteria. The final sub-section analyses the progression of infection with the isogenic mutant, Δ PAPI-1 Δ PAPI-2 in more detail.

4.4.1 Survival times post-infection

For survival experiment, mice were culled when their symptom score reached 6, 10 mice were used for each strain. Survival was monitored for 96 hours (Figure 4-8). All 10 mice infected with either wild-type PA14 or Δ PAPI-1, survived 18 hours post-infection. The singular loss of either PAPI-1 or PAPI-2 from PA14 leads to death within 24 hours post-infection. This is in stark comparison with double loss of both PAPI-1 and PAPI-2 which results in 20% of the mice surviving the experiment and 80% surviving till 72 hours post-infection. Wild-type PAO1 infected mice survive 28 hours post-infection making it in terms of survival more virulent than Δ PAPI-1 Δ PAPI-2 and on par virulent with Δ PAPI-2.

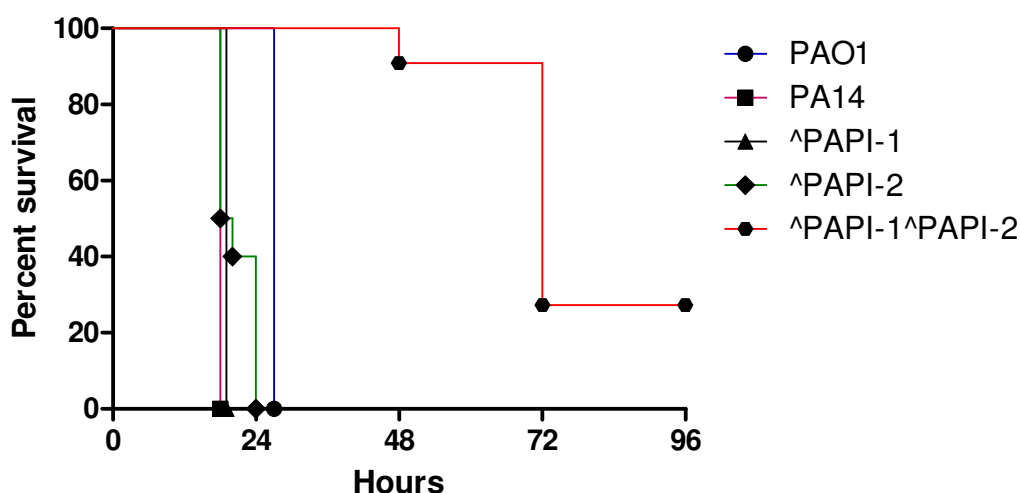


Figure 4-8 Survival graph showing percentage survival monitored for 96 hours post-infection with 2×10^6 CFU/50 μ l; $n=10$ for each strain. All data were generated over 2 independent experiments for each strain.

4.4.2 Symptoms scores post-infection

At 18 hours post-infection, the visible symptoms of each individual mouse were monitored. There were three signs to monitor, hunched, starry coat and lethargic each monitored on a scale of 0 = no signs, 1 = starting to show symptom and 2 = obviously exhibiting the symptom. Mice with a total symptom score of 6 were culled.

The signs of illness in the mice infected with Δ PAPI-1, Δ PAPI-2 and the wild-type, PA14 were not too dissimilar at the 18 hours post-infection. The only significant difference between PA14 and one its isogenic mutants was seen with the double knockout, Δ PAPI-1 Δ PAPI-2, where mice were demonstrating little or no signs of illness at 18 hours.

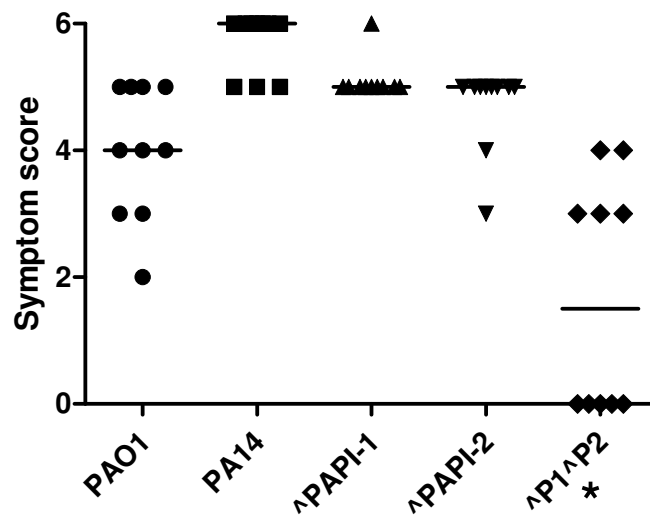


Figure 4-9 Symptom scores at 18 hours post-infection with 2×10^6 CFU/50 μ l; $n=10$ for each strain. Each dot represents a single mouse. All data were generated over 2 independent experiments for each strain. *denotes significant difference when compared with wild-type PA14 $P < 0.05$. The bar represents the median.

4.4.3 Bacterial numbers in nasopharynx, lungs and blood post-infection

The comparison of recovered CFU from the nasopharynx, lungs and blood at 18 hours post-infection can be found in Figure 4-10. The comparison of significant differences between PAO1, PA14 and its isogenic mutants are shown in the following tables, Table 4-1, Table 4-2 and Table 4-3. As discussed in the previous section PA14 is more virulent than PAO1; this is demonstrated by significantly higher ($P < 0.05$) numbers in the nasopharynx, lungs and blood at 18 hours post-infection. PA14 and Δ PAPI-1 show no significant difference ($P < 0.05$) in recovered CFU from all the tissues suggesting no change in virulence between the two strains. Δ PAPI-2 is attenuated when compared with PA14, based on a significant difference

($P < 0.05$) in CFU across all the tissues. Interestingly despite loss of PAPI-1 on its own showing no significant difference in virulence with wild-type PA14, when lost in conjunction with PAPI-2 (Δ PAPI-1 Δ PAPI-2), there is a drop in recovered CFU ($P < 0.05$) in comparison to Δ PAPI-2. This suggests PAPI-1 is important for virulence of PA14. PAO1 and Δ PAPI-2 have no significant difference in CFU recovered from nasopharynx, lungs and blood; suggesting a similar virulence between the two strains.

	PA14	Δ PAPI-1	Δ PAPI-2	Δ PAPI-1 Δ PAPI-2
PAO1	✓	✓	✗	✗
PA14		✗	✓	✓
Δ PAPI-1			✓	✓
Δ PAPI-2				✓

Table 4-1 Significant differences in CFU recovered from the nasopharynx at 18 hours post-infection with 2×10^6 CFU/50 μ l. $P < 0.05$ ✓=significant difference ✗=no significant difference.

	PA14	Δ PAPI-1	Δ PAPI-2	Δ PAPI-1 Δ PAPI-2
PAO1	✓	✓	✗	✓
PA14		✗	✓	✓
Δ PAPI-1			✓	✓
Δ PAPI-2				✓

Table 4-2 Significant differences in CFU recovered from the lungs at 18 hours post-infection with 2×10^6 CFU/50 μ l. $P < 0.05$ ✓=significant difference ✗=no significant difference.

	PA14	Δ PAPI-1	Δ PAPI-2	Δ PAPI-1 Δ PAPI-2
PAO1	✓	✓	✗	✓
PA14		✗	✓	✓
Δ PAPI-1			✓	✓
Δ PAPI-2				✓

Table 4-3 Significant differences in CFU recovered from the blood at 18 hours post-infection with 2×10^6 CFU/50 μ l. $P < 0.05$ ✓=significant difference ✗=no significant difference.

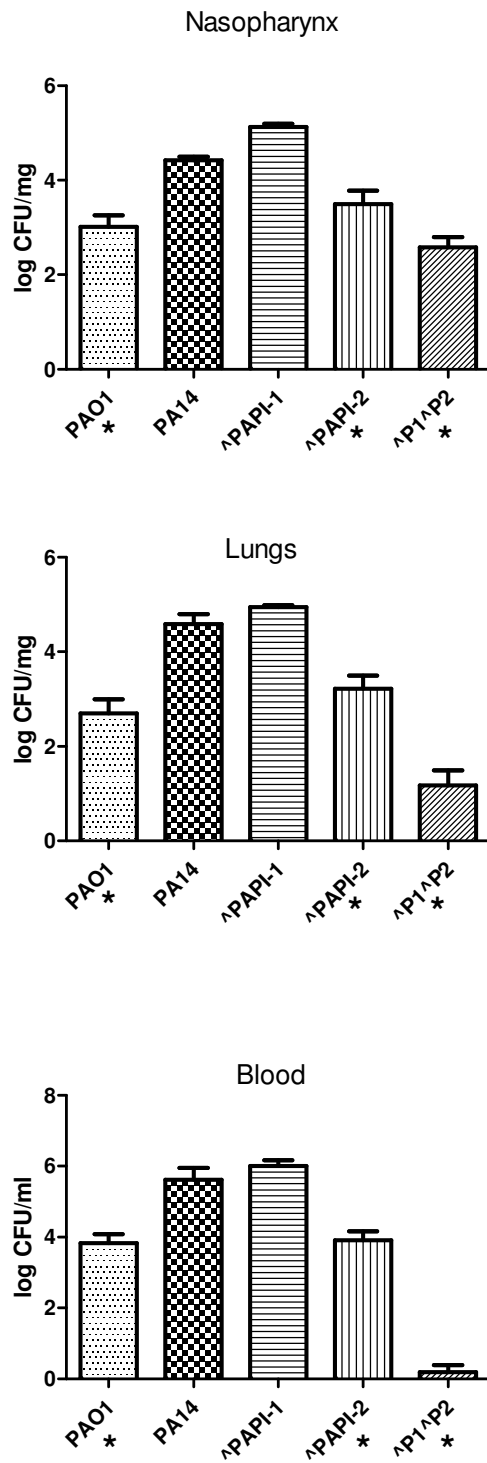


Figure 4-10 shows the CFU numbers recovered from the nasopharynx, lungs and blood at 18 hours post-infection 2×10^6 CFU/50 μ l; $n=10$ for each strain. Each bar represents the mean log CFU/mg except blood represents mean log CFU/ml and the error bar represents the standard error of the mean. All data was generated over 2 independent experiments for each strain. *denotes significant difference when compared with wild-type PA14 $P < 0.05$.

4.4.4 Intravenous infection

A question was raised by previous CFU data generated for the isogenic mutants. Why was there a significant decrease ($P < 0.05$) in blood CFU at 18 hours post-infection in two isogenic mutants namely, Δ PAPI-2 and Δ PAPI-1 Δ PAPI-2 compared to wild-type PA14? To address this question a blood infection model was developed to discriminate between two scenarios. The difference was caused by a decreased ability of the mutants to disseminate from the lungs or a result of a decreased ability to survive within the blood.

The first challenge dose was 2×10^5 CFU/100 μ l (Figure 4-11). At this dose the mice did not exhibit any signs of illness over the 24 hour period and the recovery of bacteria was very low at 24 hours post-infection. The second challenge dose was 2×10^6 CFU/100 μ l and the CFU numbers results are shown in Figure 4-12. The only significant difference for this data based on CFU numbers is between Δ PAPI-1 Δ PAPI-2, and PA14, Δ PAPI-1 and Δ PAPI-2 at the 2 hour timepoint. After this timepoint, there is no significant difference between any of the strains. The symptom scores are shown in Figure 4-13 for 2, 6 and 24 hours post-infection. Based on symptom scores, the only significant differences were between wild-type PA14 and all other strains used. At 2 hours, all the strains are exhibiting no symptoms. At 6 hours, all the mice infected with isogenic mutants and PAO1 are still exhibiting no symptoms except Δ PAPI-1. Δ PAPI-1 infected mice showed an average of symptom score of 1 and PA14 infected mice have an average of symptom score of 6 and were culled at this timepoint. By 24 hours post-infection all the isogenic mutants and PAO1 had a high symptom scores of 5 - 6. These data suggests that despite having similar CFU numbers within the blood, all the isogenic mutants had a decreased ability to cause severe disease in mice compared with PA14. The data recovered from the intravenous model suggests that Δ PAPI-2 and Δ PAPI-1 Δ PAPI-2 are deficient in some aspect that allows dissemination from the lungs to the blood. This may or may not be due to lower CFU numbers recovered from the lungs.

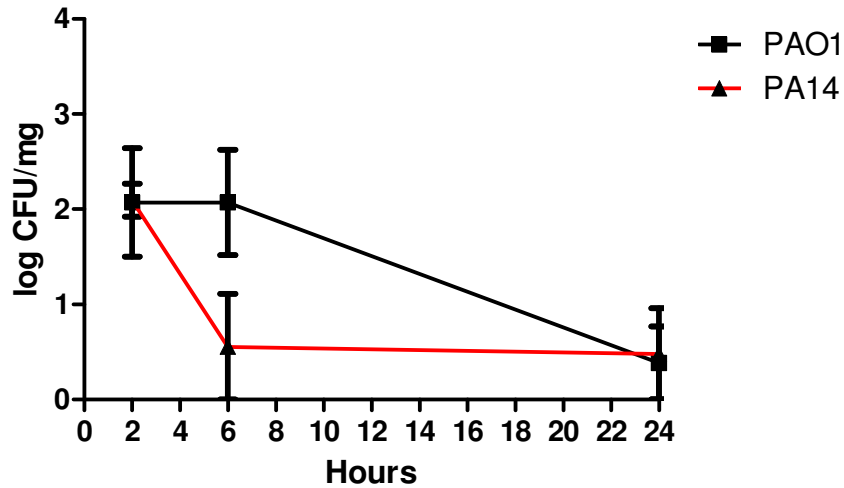


Figure 4-11 Blood CFU recovered from intravenous infection with 2×10^5 CFU/100 μ l taken at 2, 6 and 24 hours post-infection; $n=5$ for each strain. The same 5 mice are used for each timepoint. Each symbol represents the mean CFU/ml of blood. All data were generated over a single experiment for each strain. The error bar represents the standard error of the mean.

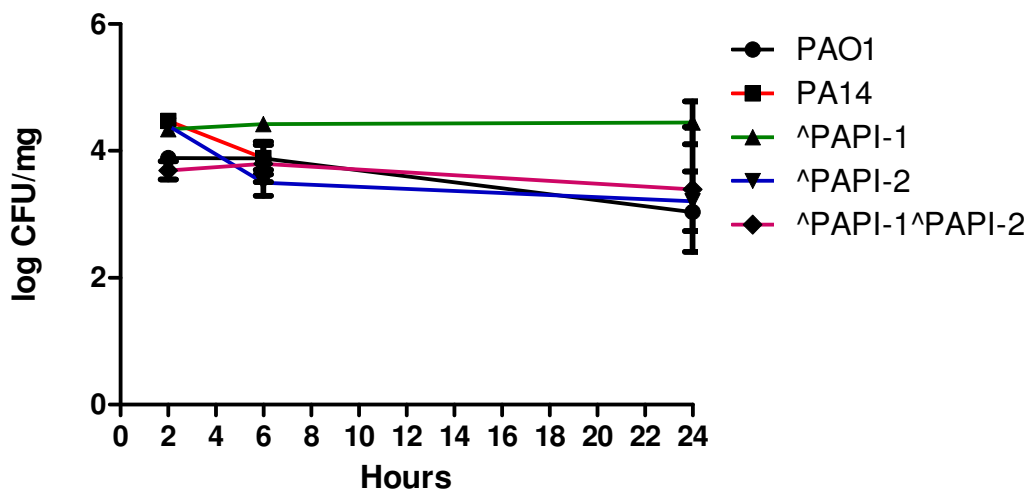


Figure 4-12 Blood CFU recovered from intravenous infection with 2×10^6 CFU/100 μ l taken at 2, 6 and 24 hours post-infection; $n=5$ for each strain. The same 5 mice are used for each timepoint. Each symbol represents the mean CFU/ml of blood. All data were generated over a single experiment for each strain. The error bar represents the standard error of the mean. Note that PA14 infected mice did not survive to generate data at the 24 hour timepoint.

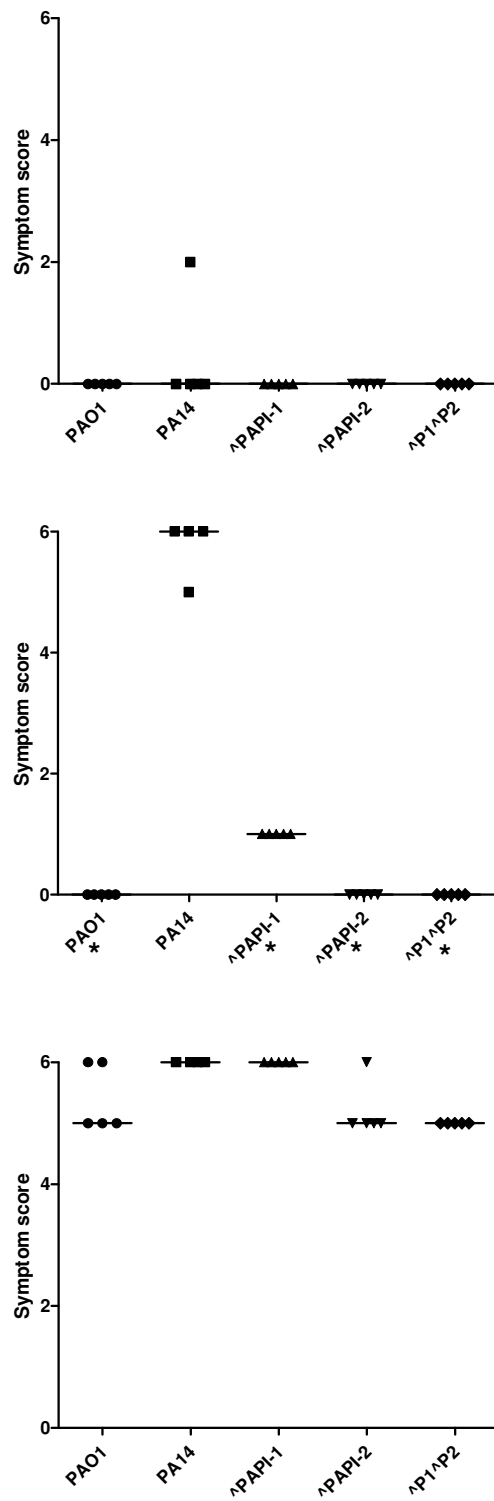


Figure 4-13 Symptom scores at 2, 6 and 24 hours post-infection with 2×10^6 CFU/50 μ l; $n=5$ for each strain. Each symbol represents a single mouse and the bar represents the median. All data was generated from a single independent experiment for each strain.

4.4.5 Immune profile

All the infected lungs had significant increase ($P < 0.05$) in all immune cell types in comparison to uninfected lungs. The leukocyte infiltration for all strains at 18 hours post-infection showed no significant differences between wild-types PAO1 and PA14, as well as the PA14 isogenic mutants (Figure 4-14). The proportion of each immune cell type was also insignificantly different for neutrophils and lymphocytes (Figure 4-14). The data suggest a significantly higher number of monocytes and macrophages in PAO1 and PA14 than Δ PAPI-1 Δ PAPI-2 (Table 4-4, Table 4-5). The data also suggest a significantly higher number of macrophages in Δ PAPI-1 and Δ PAPI-2 than Δ PAPI-1 Δ PAPI-2. (Table 4-5)

	PA14	Δ PAPI-1	Δ PAPI-2	Δ PAPI-1 Δ PAPI-2
PAO1	✕	✕	✕	✓
PA14		✕	✕	✓
Δ PAPI-1			✕	✕
Δ PAPI-2				✕

Table 4-4 Significant differences in monocytes recovered from the lung homogenate at 18 hours post-infection with 2×10^6 CFU/50 μ l. $P < 0.05$ ✓=significant difference ✕=no significant difference.

	PA14	Δ PAPI-1	Δ PAPI-2	Δ PAPI-1 Δ PAPI-2
PAO1	✕	✕	✕	✓
PA14		✕	✕	✓
Δ PAPI-1			✕	✓
Δ PAPI-2				✓

Table 4-5 Significant differences in macrophages recovered from the lung homogenate at 18 hours post-infection with 2×10^6 CFU/50 μ l. $P < 0.05$ ✓=significant difference ✕=no significant difference.

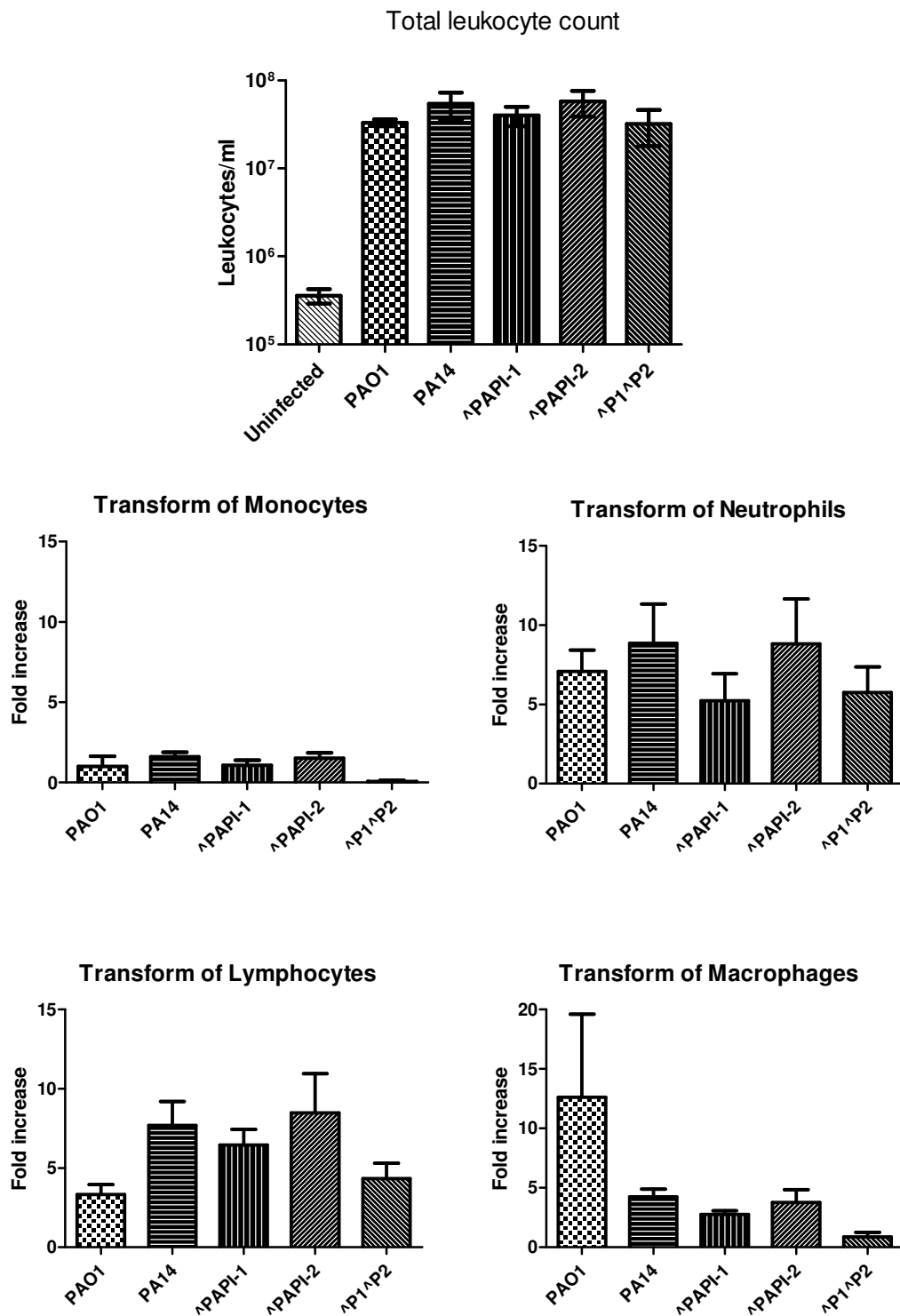


Figure 4-14 Top graph shows total leukocyte counts at 18 hours post-infection with 2×10^6 CFU/50 μ l, $n=4$ for each strain. The bar represents the mean total number of leukocytes per ml of lung homogenate. The subsequent graphs show the separate immune cell types that compose the total leukocyte count and the y-axis represents the fold-increase in comparison to the uninfected leukocyte counts. The error bars represents the standard error of the mean. $P < 0.05$

4.4.6 Lung histopathology post-infection

Lungs were recovered at 18 hours post-infection for each strain, 2 mice were culled and scored. The histopathology was scored on a number of different parameters, hypertrophy of the bronchial wall, general consolidation of the lung tissue and cellular infiltration of the immune cells. Hypertrophy of bronchial cell wall refers to presents of inflamed bronchi. Consolidation of lung tissue refers to visible loss of air spaces with the lung tissue. Cellular infiltration refers to the present of immune cells within the lung tissue, which can be seen moving from the blood vessels towards inflamed bronchi. Damage to the lungs 18 hours post-infection is summarised in Table 4-6 and photographs are shown in Figure 4-16. PA14 histopathology damage level (4.5) is on average higher than all other strains. The mean histopathology score for Δ PAPI-1 (3.5) was slightly higher than PAO1 (3), Δ PAPI-2 (3) and Δ PAPI-1 Δ PAPI-2 (3). The level of damage within the lungs between PAO1, Δ PAPI-2 and Δ PAPI-1 Δ PAPI-2 appeared to be the same. The data suggests that PAPI-2 contributes to lung damage.

		Hypertrophy of bronchial walls	General consolidation	Cellular infiltration	Score
PAO1	Mouse 1	✓	Medium	Medium	3
	Mouse 2	✓	Medium	Medium	3
PA14	Mouse 1	✓	High	Medium	4
	Mouse 2	✓	High	High	5
Δ PAPI-1	Mouse 1	✓	High	Medium	3
	Mouse 2	✓	High	High	4
Δ PAPI-2	Mouse 1	✓	Medium	Medium	3
	Mouse 2	✓	Medium	Medium	3
Δ PAPI-1 Δ PAPI-2	Mouse 1	✓	Medium	Medium	3
	Mouse 2	✓	Medium	Medium	3

Table 4-6 shows the histopathology scores for lungs recovered at 18 hours post-infection with 2×10^6 CFU/50 μ l; $n=2$ for each strain

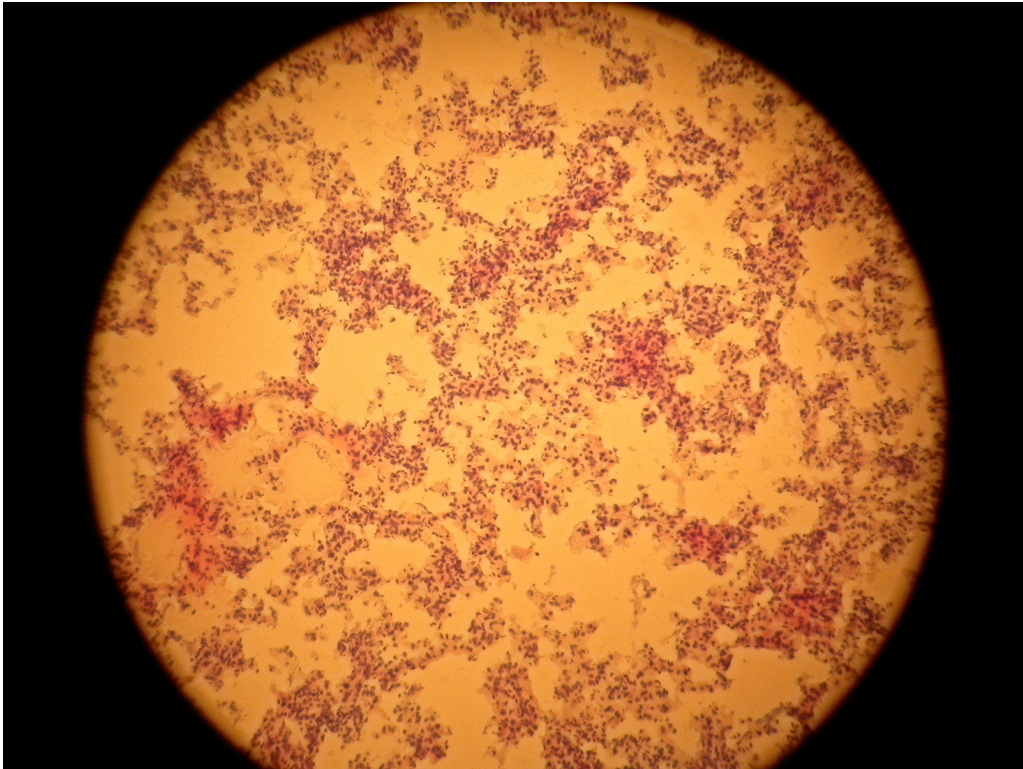


Figure 4-15 Histopathology lung section created from uninfected control mice

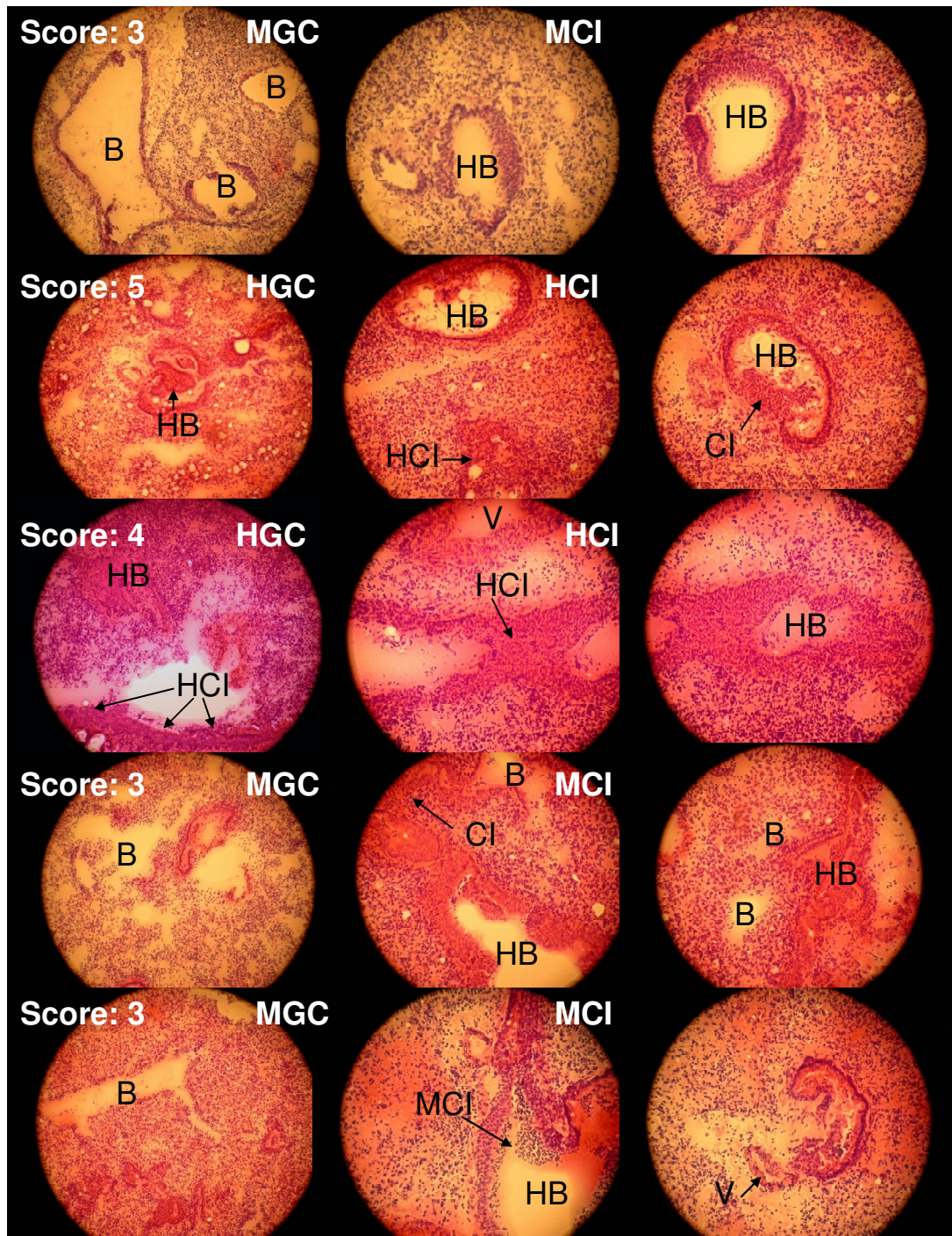


Figure 4-16 Histopathology lung sections created from mice infected with 2×10^6 CFU/50 μ l
 Left column represents from top to bottom a) PAO1 b) PA14 c) Δ PAPI-1 d) Δ PAPI-2
 e) Δ PAPI-1 Δ PAPI-2 at magnification x100; Middle and Right column represents the same
 stains respectively from top to bottom at magnification x200.

Key: LCI-MCI-HCI represents low/medium/high cellular infiltration; LGC-MGC-HGC
 represents low/medium/high general consolidation; B represents bronchi; V represents blood
 vessel

4.4.7 Progression of disease post-infection with PA14 Δ PAPI-1 Δ PAPI-2

An investigation into Δ PAPI-1 Δ PAPI-2 infected mice was conducted to verify the mice were succumbing to *P. aeruginosa* infection. To achieve this the mice were monitored over a 96 hour period, mice were culled at set timepoints during the experiment to monitor CFU numbers within the nasopharynx, lungs and blood, as well as monitoring the immune response via histopathology and leukocyte counts. All the mice in the experiment had to be culled at 72 hours, previous data generated during this project suggested to expect 70% to be culled at this timepoint.

The CFU recovered from the nasopharynx, lungs and blood significantly ($P > 0.05$) increased from 18 hours to 72 hours as shown in Figure 4-17. The numbers found within the nasopharynx and the lungs were not significantly dissimilar ($P > 0.05$) to the numbers found within mice infected with wild-type PA14 and Δ PAPI-1 at 18 hours post-infection. This suggests that these numbers represent a critical mass resulting in severe morbidity. The symptom scores also correlated with an increase in bacterial load within the tissues and blood (Figure 4-18).

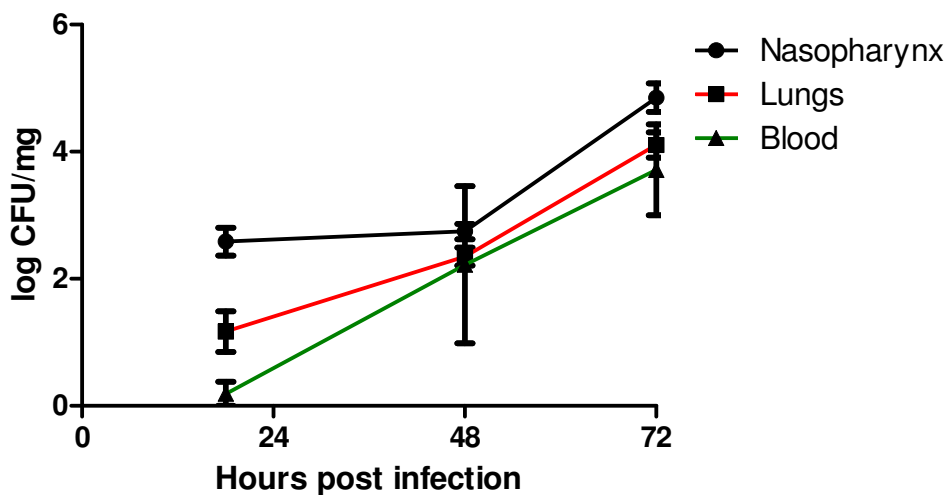


Figure 4-17 shows the CFU numbers recovered from the nasopharynx, lung and blood at 18, 48 and 72 hours post-infection with Δ PAPI-1 Δ PAPI-2 at 2×10^6 CFU/50 μ l; $n=10$ for 18 and 72 hours post-infection and $n=3$ for 48 hours post-infection. Each symbol represents the mean log CFU/mg for nasopharynx and lungs and the mean log CFU/ml of blood. The 18 hour timepoint data was generated over two independent experiments and the 48 and 72 hour data was generated over a single experiment. The error bar represents the standard error of the mean.

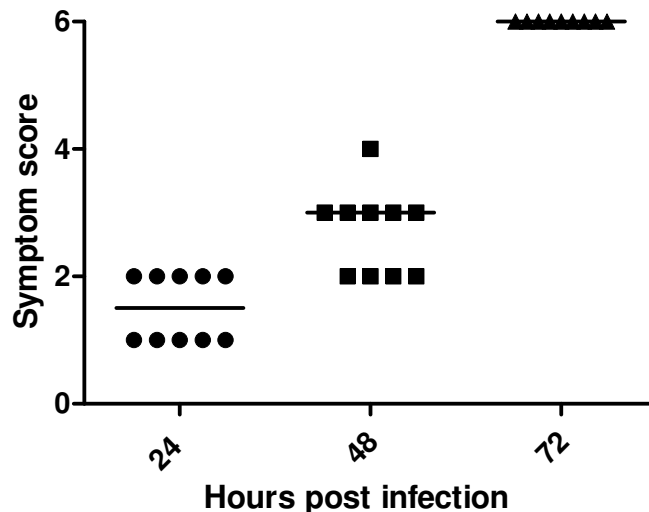


Figure 4-18 Symptom scores for mice infected with Δ PAPI-1 Δ PAPI-2 monitored over 72 hours with 2×10^6 CFU/50 μ l; $n=10$. Each dot represents a single mouse. All data was generated over a single experiment. The bar represents the median.

The immune response was also monitored for the 72 hours post-infection by leukocyte counts (Figure 4-19) and histopathology. These results also show a correlation between increase bacterial load within the tissues, increased cellular infiltration in the lungs and severity of histopathology as the infection progresses. Histopathology scores are summarised in Table 4-7 and photographs shown in Figure 4-20. The photograph of the lungs at 72 hours represents mice that were severely morbid with a symptom score of 6. The photograph of the lungs taken at 96 hours post-infection represents the 30% of mice that survived 96 hours post-infection with Δ PAPI-1 Δ PAPI-2.

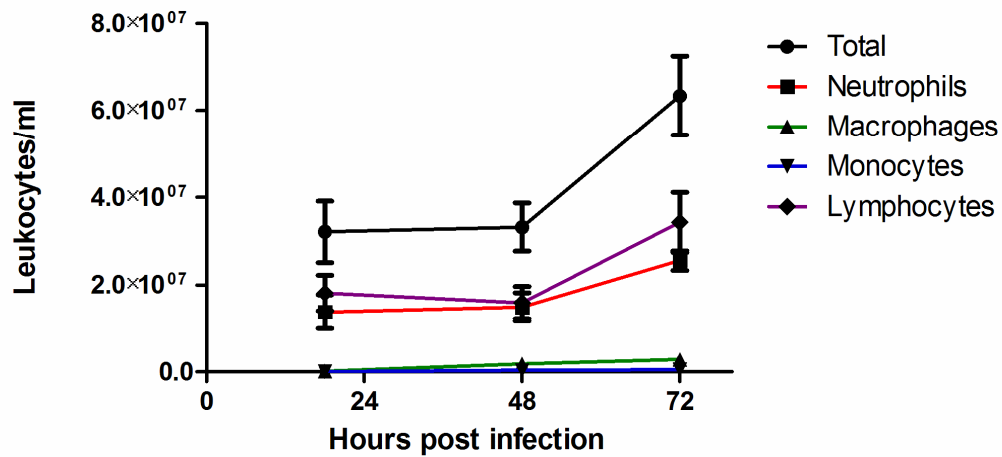


Figure 4-19 Leukocytes numbers monitored over 72 hours post-infection with Δ PAPI-1 Δ PAPI-2 at 2×10^6 CFU/50 μ l; $n=4$ for 18 hours post-infection and $n=3$ for 48 and 72 hours post-infection. The top line represents the total leukocyte count and the subsequent lines show the number of the individual immune cell types. The symbol represents the mean total number of leukocytes per ml of lung homogenate and the error bar represents the standard error of the mean.

		Hypertrophy of bronchial walls	General consolidation	Cellular infiltration	Score
PA14 Δ PAPI-1 Δ PAPI-2	Mouse 1	✓	Medium	Medium	3
18 hours	Mouse 2	✓	Medium	Medium	3
PA14 Δ PAPI-1 Δ PAPI-2	Mouse 1	✓	High	High	4
72 hours	Mouse 2	✓	High	High	5
PA14 Δ PAPI-1 Δ PAPI-2	Mouse 1	✓	Medium	Medium	3
96 hours					

Table 4-7 showing histopathology scores for lungs infected with Δ PAPI-1 Δ PAPI-2 monitored for 96 hours post-infection with 2×10^6 CFU/50 μ l

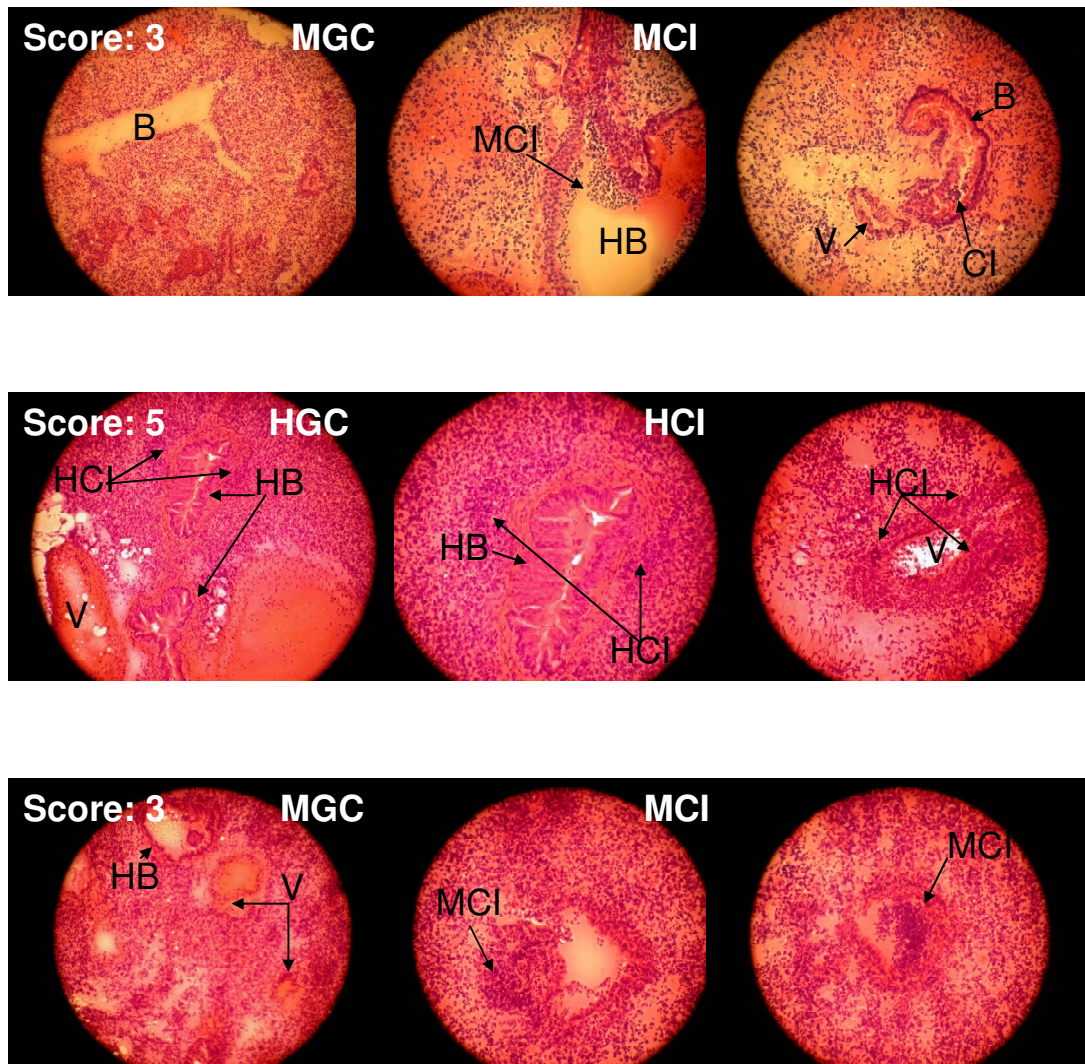


Figure 4-20 Figure 4-21 Histopathology lung sections produced from mice infected with Δ PAPI-1 Δ PAPI-2 at 2×10^6 CFU/50 μ l

Left column represents from top to bottom a) 18 hours b) 72 hours c) 96 hours at magnification x100; Middle and Right column represents the same stains respectively from top to bottom at magnification x200.

Key: LCI-MCI-HCI represents low/medium/high cellular infiltration; LGC-MGC-HGC represents low/medium/high general consolidation; B represents bronchi; V represents blood vessel

4.5 Summary of acute respiratory outcomes

The data would indicate that wild-type PA14 and Δ PAPI-1 have no significant differences in virulence within a murine acute respiratory model. Δ PAPI-2 is less virulent than wild-type, PA14. The double deletant mutant, Δ PAPI-1 Δ PAPI-2 is severely attenuated compared to the wild-type and both single mutants. An

interesting result is that the Δ PAPI-1 Δ PAPI-2 is less virulent than PAO1, a strain that lacks both PAPI-1 and PAPI-2. In fact in terms of virulence PAO1 is similar to Δ PAPI-2. The differences in virulence are reflected in recovered CFU numbers from tissue, symptom scores, survival times and histopathology, but not readily apparent from the immune profile.

4.6 Discussion

This project was to investigate the contribution of PAPI-1 and PAPI-2 to the virulence of *Pseudomonas aeruginosa* PA14 in a murine acute respiratory model by comparing wild-type PA14 and three isogenic mutants. The project shows that both pathogenicity islands contribute significantly to the virulence of PA14 and the loss of both pathogenicity islands results in a severely attenuated strain. Mice infected with the attenuated strain rather than the wild-type have an increased survival time from 18 hours to 72 hours (70%) to surviving the experiment (30%). Eighteen hours post-infection the CFU numbers recovered are drastically reduced in the lungs, nasopharynx and blood compared to wild-type, PA14, also a reduction in symptom score was only noted when both PAPI-1 and PAPI-2 are absent from PA14.

This section is going to discuss the role PAPI-1 and PAPI-2 play in virulence with respect to the project data and the available literature. This is followed by evaluating the hypothesis raised by He *et al.* (2004) regarding the differences in virulence between PAO1 and PA14. The final discussion point will be about the immune response during *P. aeruginosa* infection.

4.6.1 The role of PAPI-1 during acute respiratory infection

The data generated during project demonstrated that PAPI-1 has a role in acute respiratory infection. Δ PAPI-1 did not show a reduction in virulence compared to PA14. The bacterial numbers recovered from the nasopharynx, lungs and blood were insignificantly different ($P < 0.05$). The survival curves were identical. On average there was a slightly more lung damage in PA14 lungs in comparison to Δ PAPI-1. But the mutant Δ PAPI-1 Δ PAPI-2 did show a further reduction in virulence compared to Δ PAPI-2, demonstrating that PAPI-1 did have a role.

The role of PAPI-1 genes in virulence has been evaluated in other models of infection. Table 4-8 lists these highlighted in the literature as responsible for a reduction in virulence. To date only the model used during this project assesses the contribution of PAPI-1 in an acute respiratory model and therefore cannot be considered directly comparable.

Gene	Other names	Function	<i>C. elegans</i>	Mouse burn-sepsis
PA14_60130	RL003	Hypothetical protein		✓
<i>uvrD</i> (PA14_60080)	RL008	DNA helicase		✓
PA14_60090	RL009	Hypothetical protein		✓
PA14_59960	RL020	Putative protein-disulfide isomerase		✓
PA14_59870	RL029	Hypothetical protein		✓
PA14_59840	RL033	Hypothetical protein		✓
PA14_59800	RL036	Kinase sensor protein		✓
<i>pvrR</i> (PA14_59790)	RL037	Two component response regulator		✓
<i>rscC</i> (PA14_59780)	RL038	Putative kinase sensor protein		✓
<i>rscB</i> (PA14_59770)	RL039	Putative two component response regulator		✓
PA14_59490	RL068	Hypothetical protein		✓
PA14_59150	RL095	Single-strand DNA-binding protein		✓
PA14_59090	RL101	Hypothetical protein		✓
PA14_59070	RL102	Hypothetical protein	✓	
PA14_59010	RL107	Hypothetical protein	✓	
PA14_58950	RL112	Hypothetical protein		✓

Table 4-8 The genes that have identified in other models within PAPI-1 that contribute to virulence of PA14. *C. elegans* data generated from Lee *et al*, 2006 (Lee *et al.* 2006) Mouse burns-sepsis generated by He *et al*, 2004 (He *et al.* 2004)

Unfortunately, 65% of the ORFs within PAPI-1 have no known function and are labelled as hypothetical genes (He *et al.* 2004). These that have been assigned function and have sufficient literature are evaluated in the subsection.

There are two, two component regulatory systems (TCS) within PAPI-1 that have been highlighted through previous infection studies as important for general virulence of the strain. Two component regulatory systems act to change the gene expression within a cell to favour changes in the external environment. They consists of two parts, one the sensor kinase, which senses the changes in the environment by being bound to the cellular membrane and two, the response regulator, which mediates the intracellular response to the changes in environment. The published data in conjugation with the data generated during this project, suggests that the two TCS are important for mammalian infection, as the differences in virulence are currently only noted in mice models of infection.

The TCS involving RcsC and RcsB described in the literature is involved in the adaptation of the cell surface by changing its structure and integrity (Huang, Ferrieres & Clarke 2006). In *E. coli* there are three components involved in this system, RcsC, the sensor kinase, which activates RcsD, an intermediate, which in turn activates RcsB, the response regulator. A RcsD homologue has not been discovered in *P. aeruginosa* PA14, but the literature suggest that the *Yersinia* spp. genome does not encode a functional RcsD and the TCS has taken on a different role (Huang, Ferrieres & Clarke 2006). As this TCS, is now found on a pathogenicity island rather than the core genome and does not encode the intermediate protein RcsD, suggests the TCS is now performing some auxiliary role crucial to PAPI-1. This is supported by the fact that mutagenesis of either RcsB or RcsC results in attenuation in other models of infection (He *et al.* 2004), demonstrating that they are functional.

In the circumstances that the *rsc* TCS system maintained its usual functional activities the loss of RcsB or RcsC would result in cells that have a reduced ability to create a capsule and to form biofilms. A capsule is a highly noted protection of bacteria from the host innate immunity. This phenomenon could contribute to the

further reduction in virulence seen between Δ PAPI-1 and Δ PAPI-1 Δ PAPI-2. PAPI-2 encodes ExoU, a potent cytotoxin which has been shown lethal to immune cells (Finck-Barbancon *et al.* 1997). In the presence of ExoU, as in the case of Δ PAPI-1, the loss of the protection afforded from the capsule and biofilm formation is negligible in terms of reduction of virulence. Δ PAPI-1 Δ PAPI-2 cells do not produce ExoU and have compromised capsule and biofilm formation, leaving the cells more susceptible to the immune system. This could be used as a partial explanation for the reduction in bacterial numbers in the lungs at 18 hours post-infection. The *in vitro* data provided by Ewan Harrison (University of Leicester, PhD, unpublished) does not support these assumptions as there was no significant reduction in the ability of Δ PAPI-1 Δ PAPI-2 to form biofilms in comparison to Δ PAPI-1.

The second TCS within PAPI-1 is PvrR, the response regulator and PA14_59800, the sensor kinase. This TCS controls colony form and biofilm formation. There are two colony forms, normal and small colony variants (SCV). SCVs produced biofilm in larger amounts and more rapidly than the 'normal' colonies as well being resistance to a large diversity of antibiotics. The change in colony type to SCV is noted when the cells are subjected to stressful conditions such as exposure to antibiotics. When conditions become more favourable they revert back to normal colony morphology. PvrR acts to switch the phenotype from SCV to 'normal' and does not affect the opposite transition (Drenkard, Ausubel 2002), another protein, WspR appears to control this function (Haussler 2004). The phenotype change is controlled by expression of *cup* gene clusters. The putative *cupD* gene cluster is found with PAPI-1 and immediately adjacent to the PvrR TCS. There are three other *cup* clusters found within the core genome, *cupA*, *cupB* and *cupC*. The *cupA* system was experimentally shown to be more important in biofilm formation, than either *cupB* and *cupC* (Ruer *et al.* 2007). This data suggests that as long as *cupA* is present then 'normal' cell activity is retained. This suggests that *cupD* loss would have no effect as *cupA* could compensate for the loss.

This data could be relevant to the data generated in this project in relation to the seeding of bacteria from the lungs to the blood. The Δ PAPI-1 mutant contains PAPI-2 and therefore ExoU. This reduces the likelihood of having stressful growth conditions within the lungs and therefore the phenotypic switch to the SCV

morphology. Δ PAPI-1 Δ PAPI-2 leaves the cells susceptible to immune attack and acts as a trigger to switch to SCV morphology. In this state, the colonies are small and attached tightly in a biofilm, therefore less mobile and in turn cause a reduction in dissemination from the lungs to the blood. This explanation could be used to partially explain, the severely reduced numbers found within the blood and also the lower numbers found within the lungs, as these SCV colonies are less likely to reproduce. The cells cannot revert back to normal morphology due to the loss of PvrR. This could also be used to explain the mice death at 72 hours post-infection, as SCV colonies persist within the lungs and in the later stages of infection gain an advantage. It would be interesting further work to analyse cells that have changed to the SCV phenotype and unable to revert and the subsequent impact on virulence.

He *et al.* (2004) reported that disrupting genes within PAPI-1 caused a reduction in mortality compared with wild-type PA14 in a mouse burns-sepsis model. In the acute respiratory model, the mortality of mice infected with either wild-type PA14 or Δ PAPI-1 was identical. The results generated from the murine intravenous sepsis model supports He *et al.* (2004) data. PAO1 and all the PA14 isogenic mutants infected mice survived till the endpoint of the experiment (24 hours) but PA14 infected mice had to be culled 6 hours post-infection.

In summary, despite the loss of PAPI-1 on its own having no noticeable effect on virulence, there is a role for PAPI-1 in acute respiratory infection. This is demonstrated by further reduction of virulence between Δ PAPI-2 and Δ PAPI-1 Δ PAPI-2. The discussion in this subsection suggests there is no notable reduction in virulence due to the potency of ExoU and its importance in respiratory infection. The masking effects of ExoU have been noted in other models, the true role of the loss of ExoT could not be seen until ExoU had also been deleted (Cowell *et al.* 2000). This work therefore can highlight the need for deletion of ExoU to truly determine the role of any other island or suspected virulence factors. The role of PAPI-1 is further established with my intravenous model that shows that the loss of PAPI-1 attenuates the strain allowing the mice to survive 24 hours in comparison to the 6 hours of the wild-type parent. This could also show that PA14 has adapted to a role as burns-sepsis clinical strain PA14 was originally isolated from burns patient.

4.6.2 The role of PAPI-2 during acute respiratory infection

A reduction in virulence in PA14 was only seen when PAPI-2 was deleted, mutants Δ PAPI-2 and Δ PAPI-1 Δ PAPI-2. The only genes highlighted in PAPI-2 within the literature as contributing to virulence are ExoU which is a potent cytotoxin plus SpcU its chaperone. The literature appears to suggest that the majority of the reduction in virulence could be due to the loss of ExoU.

ExoU has been previously been shown to be an important factor in terms of acute respiratory illness within murine acute respiratory models (Allewelt *et al.* 2000, Schulert *et al.* 2003, Shaver, Hauser 2004). Allewelt *et al.*, 2000 reported a clear difference in LD₅₀ for a panel of *P. aeruginosa* strains between those expressing ExoU and those strains that did not and that introducing the *exoU* gene with its chaperone *spcU* into a non-carrying strain could significantly reduce the LD₅₀. The loss of ExoU from the *P. aeruginosa* strain, PA99, was also shown to be significantly attenuated in a murine acute respiratory model (Shaver, Hauser 2004). This model is very similar to the one described in this thesis, they used a similar dose and culling timepoint (18 hours). They evaluated differences in virulence by performing CFU counts on lung, liver and spleen homogenate. They did not evaluate the nasopharynx or directly assess dissemination from the lungs to the blood by cardiac puncture.

The data generated during this project support these observations as they show a loss of PAPI-2 results in an attenuated strain in comparison with PA14. The data suggest that PAPI-2 plays a number of important roles in acute respiratory infection. Infection with Δ PAPI-2 in comparison with wild-type PA14 led to a loss of one log in nasopharynx as well as the lungs and a two log in blood at 18 hours post-infection. Consideration of the data generated for Δ PAPI-1, suggests that the presence of PAPI-2 solely is enough to ensure persistence in nasopharynx and lungs as well as successful dissemination to the blood. This data is the first to recognise that PAPI-2 is important for upper respiratory tract (nasopharynx) infection.

The intravenous sepsis model data seemed to demonstrate that all the isogenic mutants can survive equally well within the blood; it is merely the ability to get into the bloodstream where the mutants differ. This data is supported by Shaver *et al.* (2004) who observed that the presence of ExoU was important for dissemination from the lungs. They constructed mutants in PA99, so the strain solely excreted ExoS, ExoT or ExoU. They observed dissemination by evaluating CFU recovered from the spleen and liver. They observed that in 60% of the mice showed dissemination when secreting ExoU solely in comparison to ExoS (40%) and ExoT (29%).

In addition to having a similar ability to survive within the blood, both the mutants lacking PAPI-2 showed a decrease in CFU within the blood at 18 hours post-infection in comparison to PA14 in the acute respiratory model. Δ PAPI-1 Δ PAPI-2 has lower CFU numbers than Δ PAPI-2, but this could be explained by increased susceptibility to the immune system decreasing the number of bacteria within the lungs as previously discussed. An alternative is that the CFU within the lungs correlates with the CFU recovered within the blood. An additional explanation is PAPI-1 encodes additional factors that allows successful dissemination from the lungs to occur, still allowing Δ PAPI-2 to dissemination, but not Δ PAPI-1 Δ PAPI-2. This factor would be less potent than ExoU, which would explain the lack of difference blood dissemination seen with Δ PAPI-1.

In summary, PAPI-2 contributes significantly to the virulence of PA14. Δ PAPI-2 was the only single mutant to be attenuated in comparison to wild-type, PA14. The importance of this pathogenicity island is solidified as the role of PAPI-1 in acute respiratory infection is masked till PAPI-2 is deleted. The data generated from this project suggest that PAPI-2 is important for upper respiratory infection and for the ability of PA14 to disseminate from the lungs. The data generated during this project is the first to highlight that PAPI-2 (ExoU) contributes to the ability of PA14 to get into the bloodstream, but not essential for survival in the bloodstream.

4.6.3 PAO1 as virulent as PA14 Δ PAPI-2 in an acute respiratory model

This subsection aims to address the hypothesis raised by He *et al.* (2004) that suggested the major differences in virulence between PAO1 and PA14 is the presence of PAPI-1 and PAPI-2. PAO1 lacks both PAPI-1 and PAPI-2; based on this hypothesis PAO1 should have similar virulence to Δ PAPI-1 Δ PAPI-2.

The data generated during this project shows that PAO1 is more virulent than Δ PAPI-1 Δ PAPI-2. PAO1 has a lower survival time than Δ PAPI-1 Δ PAPI-2. PAO1 has a significantly higher CFU ($P < 0.05$) in the nasopharynx, lungs and blood 18 hours post-infection. Interestingly, in terms of virulence, PAO1 is similar to Δ PAPI-2, showing no significant difference ($P < 0.05$) across the tissues at 18 hours post-infection. Wild-type PAO1 infected mice survive 28 hours post-infection making it in terms of survival more virulent than Δ PAPI-1 Δ PAPI-2 (70% (72hrs), 30% survive) and on par virulent with Δ PAPI-2 (100% by 22hrs). The data therefore suggests that PAO1 is as virulent as Δ PAPI-2 and not Δ PAPI-1 Δ PAPI-2 as expected. The data strengthens the argument that the true differences between PAO1 and PA14 are more complex than the presence or absence of pathogenicity islands (Lee *et al.* 2006).

There have been two previous studies to analyse the genomic differences between PAO1 and PA14 that could contribute to the difference in virulence. The first study in 2003, used representational difference analysis (RDA), this technique is to compare the genome of the two strains and looks for specific differences (Choi *et al.* 2002). This method lead to the discovery of 4 genes reported that differed significantly from PAO1. One of the genes highlighted is found within PAPI-1, *uvrD*. They also discovered a gene that was a homologue of *ybtQ*, which is not found within either PAPI-1 or PAPI-2. They produced a PA14 mutant in *ybtQ*, and tested the effect on virulence in a number of infection models. The mutant was attenuated to virulence levels comparable with PAO1 in both a *G. mellonella* and mouse burns-sepsis model, but not *C. elegans*.

The second study in 2006, involved sequencing the complete genome of PA14 (Lee *et al.* 2006). This allowed the group to analyse the genomes of the two strains directly. The aim of the study was to discover if the difference in virulence in PAO1 and PA14 was due to presence of pathogenicity islands present in PA14 but absent

in PAO1(Lee *et al.* 2006). They showed that PA14 had 58 additional gene clusters than PAO1. They tried to analyse these 58 gene clusters in two ways. Firstly, they screened 20 *P. aeruginosa* strains (including PA14 and PAO1) and compared them for virulence in a *C. elegans* model of infection and ranked their virulence. They then analysed the strains for the presence of the identified PA14-specific gene regions. The results showed no correlation between virulence and the presence of the PA14-specific gene regions. A second approach taken was to create a transposon insertion mutant library for PA14. They screened the library using the *C. elegans* model of infection for those in the PA14-specific gene regions for attenuation. They discovered 9 genes, 2 of which are present in PAPI-1 that caused attenuation. The present of these regions within the other *P. aeruginosa* strains, showed no set trend for correlation for virulence within *C. elegans* model of infection. This led the group to conclude that the presence of these gene regions within another *P. aeruginosa* strain was not an indicator of its virulence within a *C. elegans* model of infection.

The publication by Lee *et al.* (2006) justify the need to create *P. aeruginosa* pathogenicity islands mutants. To evaluate the true impact of a pathogenicity island, it is essential to remove the whole island. Pathogenicity islands have evolved over time to become single elements that encode all the genes that are required to increase virulence. The island has evolved and continued to function as a single element because it functions better as a whole. Lee *et al.* (2006) discovered 9 genes that were important for virulence in a *C. elegans* model of infection, 2 of which were in PAPI-1. The data could represent ‘kinks in the armour’ of the pathogenicity island that causes the whole island to be rendered non-functional. The availability of the technique to create unmarked mutants within *P. aeruginosa* is available and our laboratory has proven that 100kb of genome can be successfully removed. This could be used to explore the roles of the 58 clusters of PA14-specific regions within the strain by removing them directly, especially as so few genes within the PA14-specific regions have an assigned function; 63% have unknown function (Lee *et al.* 2006). The use of PA14-specific region mutants could be used as a cost-effective method to decide on individual genes to target for further virulence characterisation.

In conclusion the data shows that PAPI-1 and PAPI-2 are not the only differences between PAO1 and PA14 and that these differences may be more complex than a straight forward absent of the pathogenicity islands, PAPI-1 and PAPI-2 (Choi *et al.* 2002, Lee *et al.* 2006). For example, previous studies have also shown that in acute respiratory models, strains harbouring ExoU are more virulent than strains harbouring ExoS (Hauser *et al.* 2002, Schulert *et al.* 2003); PA14 has ExoU and PAO1 has ExoS. A study has demonstrated that there is a 4 fold lower LD₅₀ of ExoU positive strains in comparison with ExoS positive strains (Schulert *et al.* 2003) when tested in a murine acute respiratory model, on a panel of clinical isolates from hospital-acquired pneumonia.

4.6.4 The immune response during acute respiratory infection

This subsection is to discuss the data generated based on lung histopathology and the cellular infiltration into the lungs. The total amount of cellular infiltration into the lungs did not differ significantly ($P < 0.05$) between all strains 18 hours post-infection. The proportion of each immune cell type was also insignificantly different between strains for neutrophils and lymphocytes. The data suggests a significantly ($P < 0.05$) higher number of monocytes in PAO1 and PA14 than Δ PAPI-1 Δ PAPI-2. The data also suggests a significantly ($P < 0.05$) higher number of macrophages in PAO1, PA14, Δ PAPI-1 and Δ PAPI-2 than Δ PAPI-1 Δ PAPI-2. For the levels of histopathology damage in the lungs, PA14 damage level (4.5) is on average higher than all other strains. Δ PAPI-1 (3.5) on average appears to have slightly higher damage in the lungs than PAO1 (3), Δ PAPI-2 (3) and Δ PAPI-1 Δ PAPI-2 (3). The level of damage found within the lungs infected with PAO1, Δ PAPI-2 and Δ PAPI-1 Δ PAPI-2 appeared to be equivalent.

The data appears to show a correlation between the histopathology and the cellular infiltration. The isogenic mutants show no significant differences in total numbers of cells that infiltrate into the lungs. This could signify that the histopathology score of 3 is caused principally by the cellular infiltration and the extracellular products common to PA14 and its isogenic mutants. Additional damage appears to be caused by factors present on PAPI-2, as PA14 and Δ PAPI-1 on average had higher histopathology scores than Δ PAPI-2 than Δ PAPI-1 Δ PAPI-2. Previous studies have

shown that deletion of the *exoU* gene can lead to a decrease in histopathology (Finck-Barbancon *et al.* 1997). The group measured levels of the oedema within the lungs caused by exoproducts. The PA103 mutant in the *exoU* gene had significantly lower amount of damage in comparison to the wild-type as measured 4 hour post-infection. The presence of PAPI-2 and therefore ExoU is likely the cause of the additional histopathology damage seen with the isogenic mutants, PA14 and Δ PAPI-1.

4.7 Conclusions and further work

The aim of this project was to explore the role of pathogenicity islands on the virulence of *Pseudomonas aeruginosa* within a murine respiratory model. This is first study to use *P. aeruginosa* pathogenicity island deletant mutants *in vivo*. The project involved developing an acute respiratory model and its use demonstrated that PAPI-1 and PAPI-2 contribute to the virulence of PA14 in a murine acute respiratory model. This project reports a role for PAPI-1 in an acute respiratory model of infection. This is the first acute respiratory model developed for *P. aeruginosa* to evaluate the upper respiratory tract. The data generated highlighted that PAPI-2 plays an important role in infection of the nasopharynx. The reduction in dissemination from lungs to blood in mutants lacking PAPI-2 led to the development of an intravenous sepsis model. The sepsis model demonstrated that PAPI-2 (ExoU) was essential for dissemination from the lungs but was not essential for survival within the blood. The results also support the literature suggesting that the difference in virulence between PAO1 and PA14 is more complex than the presence or absence of pathogenicity islands. The virulence of PAO1 was more closely mimicked by Δ PAPI-2 than Δ PAPI-1 Δ PAPI-2. This is despite PAO1 itself not containing either PAPI-1 or PAPI-2.

The project has shown that the use of pathogenicity island mutants in virulence models is feasible. The use of this model in the future could aid the labelling of known genomic islands as pathogenicity islands. The data generated could also be used as a starting point to pin point virulence genes. Further experiments based on this premise, could be to evaluate the 58 PA14-specific regions highlighted by Lee *et al.* (2006). This could produce interesting results especially as our laboratory has

shown that ~100kb of genome can be successful deleted. An additional investigation that could be useful is the creation of a PAPI-1 deletant mutant with ExoU also deleted. This data could be used to evaluate if the hypothesis generated during this project that ExoU is masking the true effects of the loss of PAPI-1 on virulence within a murine acute respiratory model. This mutant was intended to be created during this study but due to time constraints this was not achieved. The loss of ExoU could also be used to determine whether any of the other genes on PAPI-2 contribute to virulence, as a large proportion have unknown function.

5 The role of quorum-sensing in the virulence of *Pseudomonas aeruginosa* Liverpool Epidemic Strain (LES) isolates

<u>5 THE ROLE OF QUORUM-SENSING IN THE VIRULENCE OF <i>PSEUDOMONAS AERUGINOSA</i> LES ISOLATES</u>	131
5.1 <i>IN VITRO</i> ANALYSIS OF THE QUORUM-SENSING EXPRESSION OF THE LES ISOLATES	132
5.2 ACUTE RESPIRATORY INFECTION OUTCOMES	135
5.2.1 SURVIVAL TIMES POST-INFECTION	135
5.2.2 SYMPTOMS SCORES POST-INFECTION	136
5.2.3 BACTERIAL NUMBERS WITHIN THE NASOPHARYNX, LUNGS AND BLOOD POST-INFECTION	140
5.2.4 LUNG HISTOPATHOLOGY POST-INFECTION	146
5.2.5 SUMMARY OF ACUTE RESPIRATORY INFECTION OUTCOMES	152
5.3 DISCUSSION	153
5.3.1 OVER-EXPRESSION OF QUORUM-SENSING PRODUCTS AS A MARKER FOR VIRULENCE IN AN ACUTE RESPIRATORY MODEL	154
5.3.2 THE DIFFERENCES IN VIRULENCE BETWEEN THE LES ISOLATES	156
5.3.3 THE LES ISOLATES VERSUS THE REFERENCE STRAINS, PAO1 & PA14	162
5.4 CONCLUSIONS AND FURTHER WORK	164

This project explores *in vivo* virulence of *Pseudomonas aeruginosa*. The Liverpool epidemic strain (LES) is a *P. aeruginosa* clinical strain that causes respiratory infection within humans. It is the most common strain isolated from CF patients across the UK (Salunkhe *et al.* 2005, Scott, Pitt 2004). It was considered important to test the acute respiratory model developed during this thesis with isolates specifically recovered from human respiratory infection cases. The aim of this project is to determine whether the over-expression of quorum-sensing products is a reliable marker for the increased virulence exhibited by some LES isolates. This hypothesis was tested in an acute respiratory infection model, by analysing the viable bacterial numbers post-infection in the upper respiratory tract (nasopharynx), the lower respiratory tract (lungs) and septicaemia secondary to pneumonia, as well as comparing lung damage (histopathology). The over-expressing isolates are represented by LES431, LESB58 and LESB65. The deficient isolates are represented by LES400. This study is the first reported case of using LES isolates in an acute respiratory model.

5.1 In vitro analysis of the quorum-sensing expression of the LES isolates

This section provides evidence that the LES isolates used in this project represent the correct quorum-sensing expression phenotype. To verify that the LES isolates did represent the correct quorum-sensing phenotype, a number of standard tests were performed. The quorum-sensing over-expressing phenotype can be represented by over-production of pyocyanin (Fothergill *et al.* 2007). Pyocyanin is a soluble blue pigment that is secreted extracellularly by *P. aeruginosa* and can be seen clearly within growth media. The photograph (Figure 5-1) of the overnight growth cultures clearly shows that the over-expressing isolates (LES431, LESB58 and LESB65) produce excesses of pyocyanin, which turns the LB growth media green. This contrasts with the deficient isolate, LES400 where the growth media remains yellow.



Figure 5-1 shows overnight growth of the LES isolates in LB growth media

From left to right: LES400, LES431, LESB58 and LESB65

Another standard test for pyocyanin is to obtain an OD_{695} reading. The OD_{695} readings of the supernatant for each of the LES isolates are shown in Figure 5-2, after 4 hour re-growth in TSB supplemented with FCS from an overnight starter culture. The readings were taken after standardisation of the culture to OD_{600} of 1. The OD readings are consistent with the data given in Fothergill *et al.* (2007) for defining over-expression of quorum-sensing products. The authors defined over-expression as an OD_{695} reading > 0.1 at a culture reading at OD_{600} 1.2. The photograph (Figure 5-3) shows the 4 hour re-growth of each of the LES isolates.

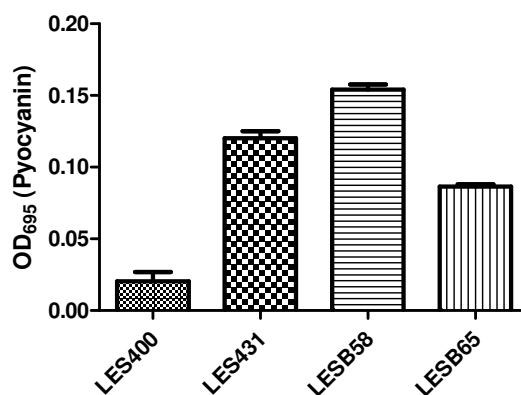


Figure 5-2 shows the pyocyanin OD_{695} reading of growth supernatant from the 4 hours re-growth for the LES isolates. All growth cultures were standardised to an OD_{600} reading of 1.

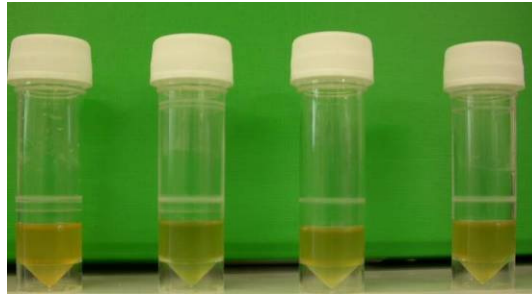


Figure 5-3 shows 4 hour growth cultures of the LES isolates in TSB growth media supplemented with FCS From left to right: LES400, LES431, LESB58 and LESB65

The production of pyocyanin was also noted after growth on solid media. The following photograph (Figure 5-4) shows the LA growth media plates with the dose serially diluted after overnight incubation. The green pigmentation is present on the growth media plates of the over-expressing isolates (LES431, LESB58 and LES65) and is not visible on the quorum-sensing deficient isolate plate (LES400).

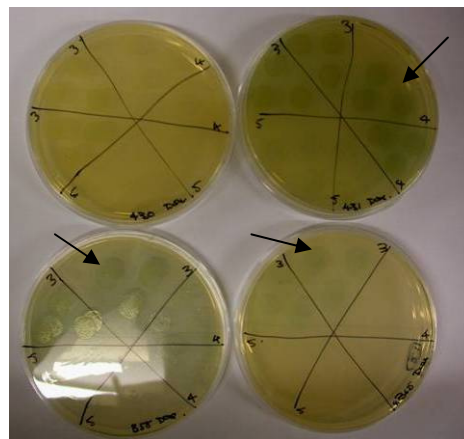


Figure 5-4 shows the growth of the LES isolates on LA growth media after overnight incubation

Top left LES400, Top right LES431, Bottom left LESB58 and Bottom right LESB65

The arrows indicates pyocyanin evident areas.

The photograph (Figure 5-5) shows the LA plates with homogenised nasopharynx or lungs as well as blood serially diluted from mice culled 18 hours post-infection after being challenged with a dose of 2×10^6 CFU; that had been incubated overnight for each of the LES isolates. The green pigmentation is present on the LA plates representing the lung CFU of the over-expressing isolates, LES431 and LESB65. The colouring is not present on lung plate for the deficient isolate, LES400 or the over-expressing isolate LESB58. The most probable explanation for the lack of pyocyanin on the LESB58 media plates is the low recovery of bacteria.

Quorum-sensing is bacterial cell density dependent and the CFU recovered has failed to reach the critical threshold for expression. The LA growth plates representing the blood CFU shows that CFU was only recovered for LES431.

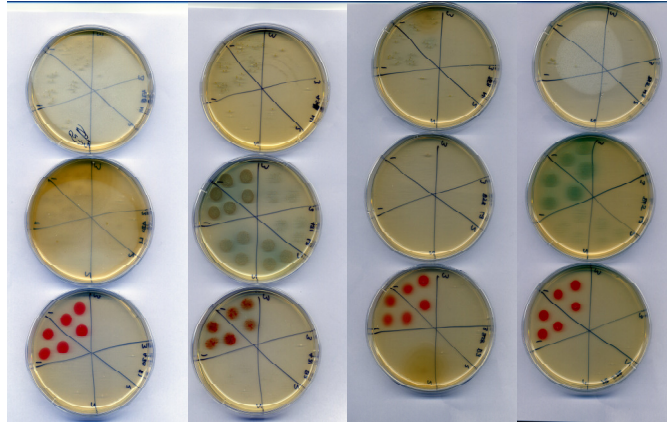


Figure 5-5 shows growth of LES isolates after recovery from nasopharynx, lungs and blood 18 hours post-infection with 2×10^6 CFU on LA growth media

From left to right: LES400, LES431, LESB58 and LESB65

From top to bottom: nasopharynx, lungs and blood

In summary, based on pyocyanin production data presented during this project and the published literature; LES400 is a quorum-sensing deficient isolate and LES431, LESB58 and LESB65 are quorum-sensing over-expression isolates (Fothergill *et al.* 2007, Kukavica-Ibrulj *et al.* 2008, Salunkhe *et al.* 2005).

5.2 Acute respiratory infection outcomes

The mice were infected with the challenge dose and were monitored for 96 hours post-infection. The mice were subjected to symptom scoring and tissue was recovered at set timepoints throughout the 96 hours. For histopathology, additional mice were infected and their lungs were recovered at set timepoints. For the survival experiments, mice were infected with the challenge dose and culled when their symptom score reached the maximum of 6.

5.2.1 Survival times post-infection

10 mice were used to evaluate the survival times for each LES isolate for 96 hours (Figure 5-6); a mouse would be culled when its total symptom score reached 6.

The results show that mice infected with LES431 did not survive beyond 24 hours post-infection. Mice infected with either LES400 or LESB58 survived the 96 hours of the experiment. 60% of the mice infected with LESB65 survived the experiment. These results suggest that LES431, an over-expressing quorum-sensing mutant is significantly ($P < 0.05$) more virulent than LES400, LESB58 and LESB65.

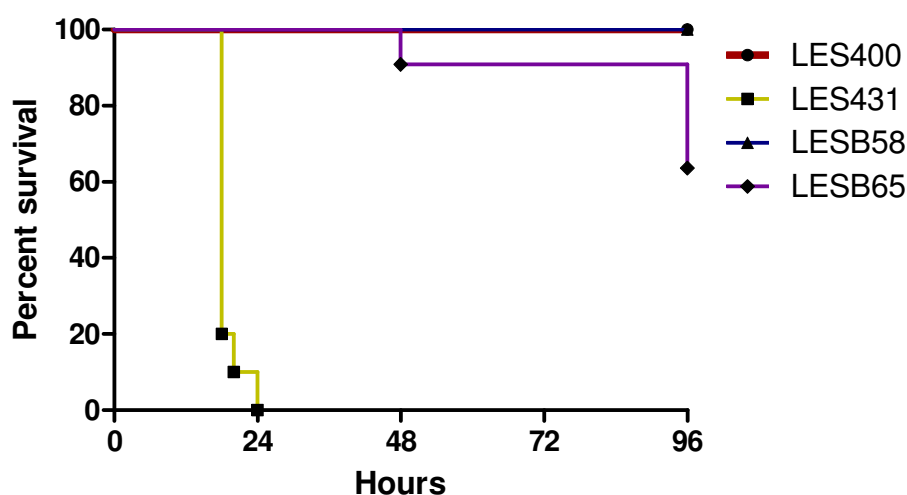


Figure 5-6 Survival graph showing percentage survival monitored for 96 hours post-infection with 2×10^6 CFU, $n=10$ for each isolate. All data were generated over 2 independent experiments for each isolate, with exception of LESB65 that was generated over a single experiment.

5.2.2 Symptoms scores post-infection

The symptom scores of each individual mouse were monitored continuously for 96 hours post-infection. There were three symptom signs to monitor; hunched, starry coat and lethargic, each monitored on a scale of 0 = no signs, 1 = initial symptom signs and 2 = fully exhibiting symptom signs. Mice with a total symptom sign score of 6 were culled.

The results show at 18 hours post-infection (Figure 5-7), there were no significant ($P < 0.05$) differences based on symptom signs between LES400, LESB58 and LESB65. The mice infected with LES431 had significantly ($P < 0.05$) higher symptom scores than all other isolates, which correlates with a lower survival time.

Interestingly, the quorum-sensing deficient isolate, LES400 had a similar symptom score to the two over-expressing isolates, LESB58 and LESB65.

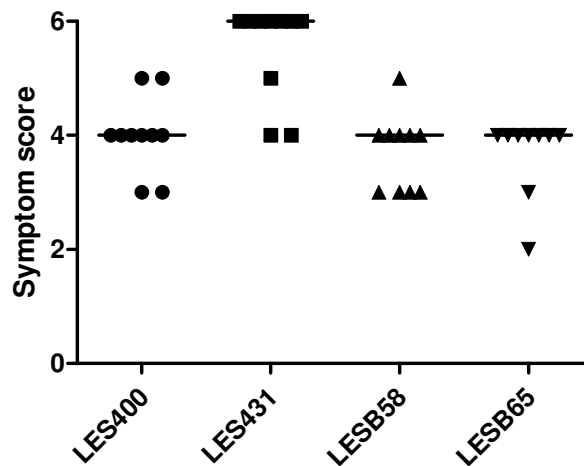


Figure 5-7 Symptom scores at 18 hours post-infection with 2×10^6 CFU; $n=10$ for each isolate. Each dot represents a single mouse. All data were generated over 2 independent experiments for each isolate. The bar represents the median.

The symptom scores for all isolates from 18 hours to 96 hours post-infection are shown in Figure 5-8. The graphs post 24 hours exclude LES431-infected mice as they do not survive beyond 24 hours post-infection. The two mice represented on the 24 hours post-infection graph represent those that survived beyond 18 hours.

The symptom scores of LES400-infected mice peaked at 48 hours and showed improvement beyond this timepoint. The symptom scores of LESB58-infected mice were at their highest from 24 hours to 48 hours post-infection and appeared completely devoid of visible symptoms 72 hours onwards. This correlates with the survival data for both LES400 and LESB58 as 100% of the mice survive the experiment (Figure 5-6). LESB65 median symptom scores (4) were consistent throughout the 96 hours. Interestingly, at the 96 hour timepoint, two groups were evident; those that appeared to be improving (30%), and those that had reached the maximum symptom score of 6 (30%).

Comparing LES400 to LESB65, the only significant ($P < 0.05$) difference in symptom score was at 24 hours post-infection where LESB65 was higher. LESB65 was consistently significantly ($P < 0.05$) higher than LESB58 from 18 hours to 72

hours post-infection. Interestingly, LESB58 (over-expressing isolate) consistently had lower symptom scores than LES400 (deficient isolate) throughout the 96 hours post-infection.

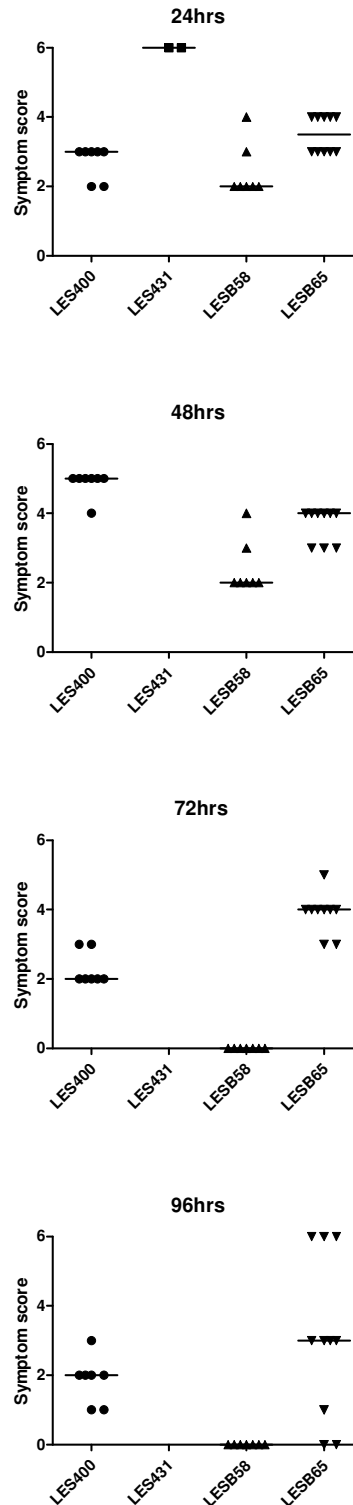


Figure 5-8 Symptom scores at 24, 48, 72 and 96 hours post-infection with 2×10^6 CFU; $n=10$ for each isolate. Each dot represents a single mouse. All data were generated over 2 independent experiments for each isolate, with the exception of LESB65 that generated in a single experiment. The two mice representing LES431 at 24 hours survived beyond 18 hours post-infection. The bar represents the median.

5.2.3 Bacterial numbers within the nasopharynx, lungs and blood post-infection

The bacterial numbers within the nasopharynx, lungs and blood were monitored for 96 hours post-infection at 24 hour intervals. Comparisons of the CFU recovered from the nasopharynx, lungs and blood at 18 hours post-infection are shown in Figure 5-9. The significant differences between the LES isolates are shown in Table 5-1, Table 5-2 and Table 5-3. LES431 is more virulent than either LES400 or LESB58, as it has significantly ($P < 0.05$) higher CFU numbers recovered from the nasopharynx, lungs and blood. This is in keeping with the survival data as LES431-infected mice do not survive beyond 24 hours whereas those infected with either LES400 or LESB58 all survive the 96 hours survival experiment. LES431 has a similar bacterial burden in the respiratory tract as LESB65 based on insignificant differences ($P < 0.05$) in CFU recovered from nasopharynx and lungs. It must be noted that LES431 has significantly higher ($P < 0.05$) CFU recovered from the blood. LESB65 is more virulent than either LES400 or LESB58, as demonstrated by higher CFU numbers recovered from the nasopharynx and lungs at 18 hours post-infection. At 18 hours post-infection there were no significant differences ($P < 0.05$) between LES400 and LESB58 in any of the recovered tissues suggesting a similar virulence. Interestingly, there were no significant differences ($P < 0.05$) between LES400, LESB58 and LESB65 in the blood at 18 hours post-infection.

	LES431	LESB58	LESB65
LES400	✓	✗	✓
LES431		✓	✗
LESB58			✓

Table 5-1 Significant differences in CFU recovered from the nasopharynx at 18 hours post-infection. $P < 0.05$ ✓=significant difference ✗=no significant difference.

	LES431	LESB58	LESB65
LES400	✓	✗	✓
LES431		✓	✗
LESB58			✓

Table 5-2 Significant differences in CFU recovered from the lungs at 18 hours post-infection.
P < 0.05 ✓=significant difference ✗=no significant difference.

	LES431	LESB58	LESB65
LES400	✓	✗	✗
LES431		✓	✓
LESB58			✗

Table 5-3 Significant differences in CFU recovered from the blood at 18 hours post-infection.
P < 0.05 ✓=significant difference ✗=no significant difference.

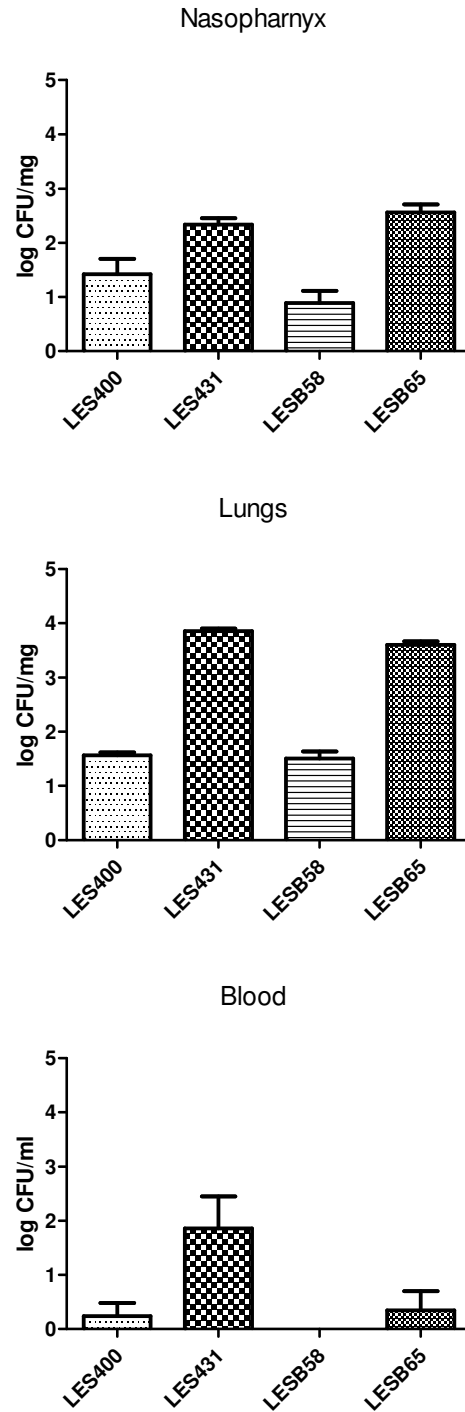


Figure 5-9 shows the CFU numbers recovered from the respective tissue, nasopharynx, lungs and blood at 18 hours post-infection with 2×10^6 CFU; $n=10$ for each isolate. Each bar represents the mean log CFU/mg except blood which represents mean log CFU/ml and the error bar represents the standard error of the mean. All data were generated over 2 independent experiments for each isolate.

Additional comparisons of the CFU recovered from the nasopharynx, lungs and blood were performed at 24 hour intervals for the 96 hour duration of the experiment (Figure 5-10) for all LES isolates. The exception is LES431 as the infected mice did not survive beyond 24 hours post-infection.

The significant differences ($P < 0.05$) between the LES isolates in CFU recovered from the nasopharynx are shown in Table 5-4. LES400 and LESB58 CFU numbers decrease to virtually undetectable levels by 24 hours post-infection within the nasopharynx. LESB65 has a significant ($P < 0.05$) increase in bacterial numbers from 18 hours to 96 hours, suggesting it has a greater ability to infect the upper respiratory tract than either LESB58 or LES400.

	LES431	LESB58	LESB65
LES400	N/A	✖	✓
LES431		N/A	N/A
LESB58			✓

Table 5-4 Significant differences in CFU recovered from the nasopharynx at 48, 72 and 96 hours post-infection. $P < 0.05$ ✓=significant difference ✖=no significant difference.

The comparison of the significant differences ($P < 0.05$) between the LES isolates in CFU recovered from lungs is shown in Table 5-5 (48 hours) and Table 5-6 (72 and 96 hours). For mice infected with the LESB58, the lung CFU numbers decreased steadily between 18 and 52 hours post-infection and were undetectable beyond the latter timepoint. LES400 lung CFU levels increased between 18 hours and 52 hours post-infection. The lung CFU numbers peak at 52 hours and then decrease steadily to 96 hours post-infection. LESB65 infected mice show a significant ($P < 0.05$) increase in CFU numbers recovered from the lungs between 18 and 96 hours post-infection. Based on these numbers the order of virulence for the isolates would be LESB65, LES400 and followed by LESB58.

	LES431	LESB58	LESB65
LES400	N/A	✓	✓
LES431		N/A	N/A
LESB58			✓

Table 5-5 Significant differences in CFU recovered from the lungs at 48 hours post-infection. $P < 0.05$ ✓=significant difference ✖=no significant difference.

	LES431	LESB58	LESB65
LES400	N/A	✖	✓
LES431		N/A	N/A
LESB58			✓

Table 5-6 Significant differences in CFU recovered from the lungs at 72 and 96 hours post-infection. $P < 0.05$ ✓=significant difference ✖=no significant difference.

Comparison of the blood CFU numbers show that the LES400 and LESB58 infected mice never developed septicaemia secondary to pneumonia. The majority of LESB65 infected mice (90%) also did not develop septicaemia. Interestingly, this suggests that these LES isolates are unable to disseminate / seed from the lungs to the blood.

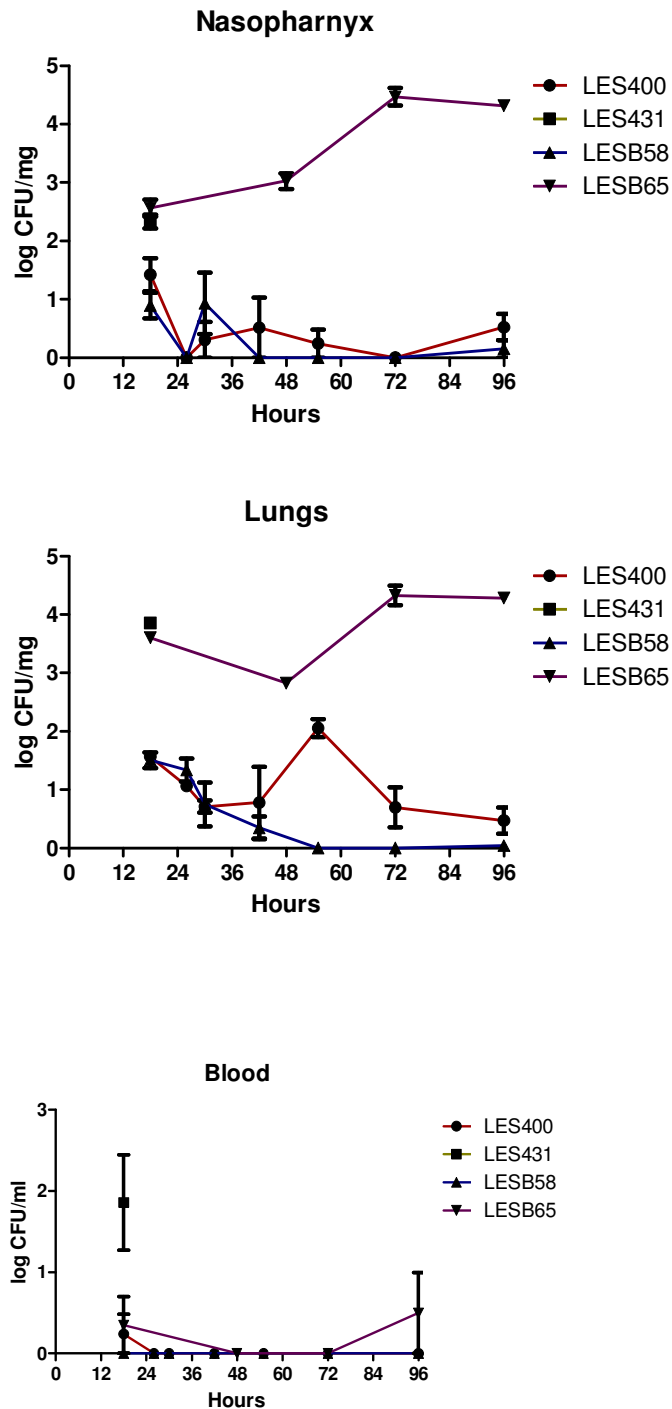


Figure 5-10 Nasopharynx, lung and blood CFU recovered at 18, 48, 72 and 96 hours post-infection with 2×10^6 CFU; $n=10$ for 18 and 96 hours post-infection and $n=4$ for 48 and 72 hours post-infection. Each dot represents the mean log CFU/mg for nasopharynx and lungs and the mean log CFU/ml of blood. The 18 and 96 hour timepoint data were generated over two independent experiments and the 48 and 72 hour data were generated over a single experiment. With the exception of LESB65 96 hour data that were generated in a single experiment. The error bar represents the standard error of the mean.

5.2.4 Lung histopathology post-infection

Lungs were taken at 18, 55 and 96 hours post-infection for each LES isolate experiment, the exception being LES431, as the mice do not survive beyond 24 hours post-infection. Two mice were culled and scored for each timepoint. The histopathology was scored on a number of different parameters, such as hypertrophy of the bronchial wall, general consolidation of the lung parenchymal tissue and cellular infiltration of host immune cells. The results are shown in Table 5-8 and in Figure 5-11 (18 hours), Figure 5-12 (55 hours) and Figure 5-13 (96 hours). Hypertrophy of bronchial cell wall refers to presence of an enlarged bronchial cell wall with a visible increase in the size and number of cells. Consolidation of lung parenchymal tissue refers to visible loss of air spaces within the lung tissue. Cellular infiltration refers to the increased presence of host immune cells within the lung tissue, which can be observed to be migrating from the blood vessels towards inflamed bronchi. Based on these observations the lungs are given a histopathology score between 1 (minor signs of histopathology) – 5 (severe signs of histopathology). Table 5-7 summarises all the data presented to show the relationship between the bacterial burden within the tissues, survival times, symptom scores and histopathology scores.

At 18 hours post-infection, LES400 and LESB58 have a significantly ($P < 0.05$) similar bacterial lung burden and median symptom score of 4. Infection with LESB58 resulted in higher level of lung damage than LES400. These data suggest that the over-expression of quorum-sensing products correlates with more severe lung damage at 18 hours post-infection.

By 55 hours post-infection, the lung damage caused by LES400 has increased to its peak score of 4. This coincides with a peak in bacterial load with the lungs and the highest median symptom score of 4 recorded for the mice. By 96 hours post-infection, the bacterial numbers within the lungs are undetectable and the symptom score has lowered dramatically to 2 and the histopathology score has lowered to 3. The recovery of the mice is represented by 100% survival by 96 hours post-infection.

For LESB58-infected mice the histopathology score of 4 is maintained between 18 and 55 hours post-infection. During this time the bacterial numbers within the lungs have gone from low to undetectable and the symptom scores has decreased considerably from 4 to 2. By 96 hours post-infection, the histopathology has decreased, bacterial numbers are still undetectable and the mice appear healthy with a symptom score of 0 and 100% surviving till 96 hours post-infection.

At 18 hours post-infection, LES431 and LESB65 also have a significantly ($P < 0.05$) similar bacterial lung burden. LES431 causes more lung damage than LESB65, adding further evidence to LES431 being more virulent than LESB65. This coincides with LES431 (0% by 24 hours) having a lower survival time than LESB65 (60% by 96 hours). This also coincides with the LES431-infected mice having a median symptom score of 6, the maximum in comparison to the LESB65 which has a score of 4.

At 55 hours post-infection, the mean histopathology score of LESB65-infected mice has increased to 4.5 coinciding with an increase in bacterial load within the lungs. The median symptom score is maintained at 4 and 90% of the mice has survived till this timepoint. At 96 hours post-infection, the level of lung damage is maintained and the median symptom score has lowered to 3. These data show that there are two groups (Figure 5-8); those that were improving (30%) and those that have declined (30%). Those that declined have reached the maximum symptom score of 6. Only 60% of the mice survived until 96 hours post-infection.

In conclusion, at 18 hours post-infection, mice infected with LES431 had the highest level of histopathology (4.5), which coincided with the survival end-point. Mice infected with LESB58 (4) and LESB65 (4) have the next highest level of histopathology. Mice infected with LES400, the deficient quorum-sensing deficient isolate have a lower level of histopathology than all the quorum-sensing over-expressing isolates. At 96 hours post-infection, the LES400 and LESB58-infected lung histopathology improved alongside their symptom scores. The lung histopathology of LESB65-infected mice had increased by this timepoint and was the survival end-point for 30% of mice.

		LES400	LES431	LESB58	LESB65
18 hrs	Survival	100%	20%	100%	100%
	Symptom Score/6	4	6	4	4
	Histopathology/5	3	4.5	4*	4
	Lung CFU	Low	Medium	Low	Medium
55hrs	Survival	100%	0%	100%	80%
	Symptom Score/6	5	N/A	2	4
	Histopathology/5	4	N/A	4*	4.5
	Lung CFU	Low	N/A	Undetectable	Medium
96 hrs	Survival	100%	0%	100%	60%
	Symptom Score/6	2	N/A	0	3
	Histopathology/5	3	N/A	3	4.5
	Lung CFU	Undetectable	N/A	Undetectable	High

Table 5-7 Summary table correlating the survival data, symptom scores and CFU at 18, 55 and 96 hours post-infection with 2×10^6 CFU. Survival refers to the percentage of mice alive. Symptom score is the median symptom score for the mice. Histopathology score refers to the mean histopathology score. Lung CFU refers to a grading system of log CFU <1 = undetectable, 1-2 = low, 2-4 = medium and 4-5 = high.

***Please note that LESB58 still had a relatively high histopathology score despite having a low CFU; this could be partially explained by the over-expression of virulence factors especially pyocyanin which has been shown to cause lung inflammation when instilled into mice lungs (Lau *et al.* 2004)**

Strain	Mouse	Timepoint	HB	CI	GC	Score
LES400	1	18hr	✓	Heavy	Medium	3
	2	18hr	✓	Heavy	Medium	3
	1	52hr	✓	Medium	Heavy	4
	2	52hr	✓	Medium	Heavy	4
	1	96hr	✓	Heavy	Medium	3
	2	96hr	✓	Medium	Medium	3
LES431	1	96hr	✓	Heavy	Heavy	5
	2	96hr	✓	Heavy	Medium	4
LESB58	1	18hr	✓	Heavy	Heavy	4
	2	18hr	✓	Heavy	Heavy	4
	1	52hr	✓	Heavy	Heavy	4
	2	52hr	✓	Heavy	Heavy	4
	1	96hr	✓	Medium	Low	3
	2	96hr	✓	Medium	Low	3
LESB65	1	18hr	✓	Heavy	Medium	4
	2	18hr	✓	Heavy	Medium	4
	1	55hr	✓	Heavy	Heavy	4
	2	55hr	✓	Heavy	Heavy	5
	1	96hr	✓	Heavy	Heavy	4
	2	96hr	✓	Heavy	Heavy	5

Table 5-8 shows the histopathology scores for lungs recovered at 18, 52 (55) and 96 hours post-infection with 2×10^6 CFU/50 μ l; $n=2$ for each strain

Key: HB = hypertrophy of bronchi, CI = cellular infiltration and GC = general consolidation of lung parenchyma

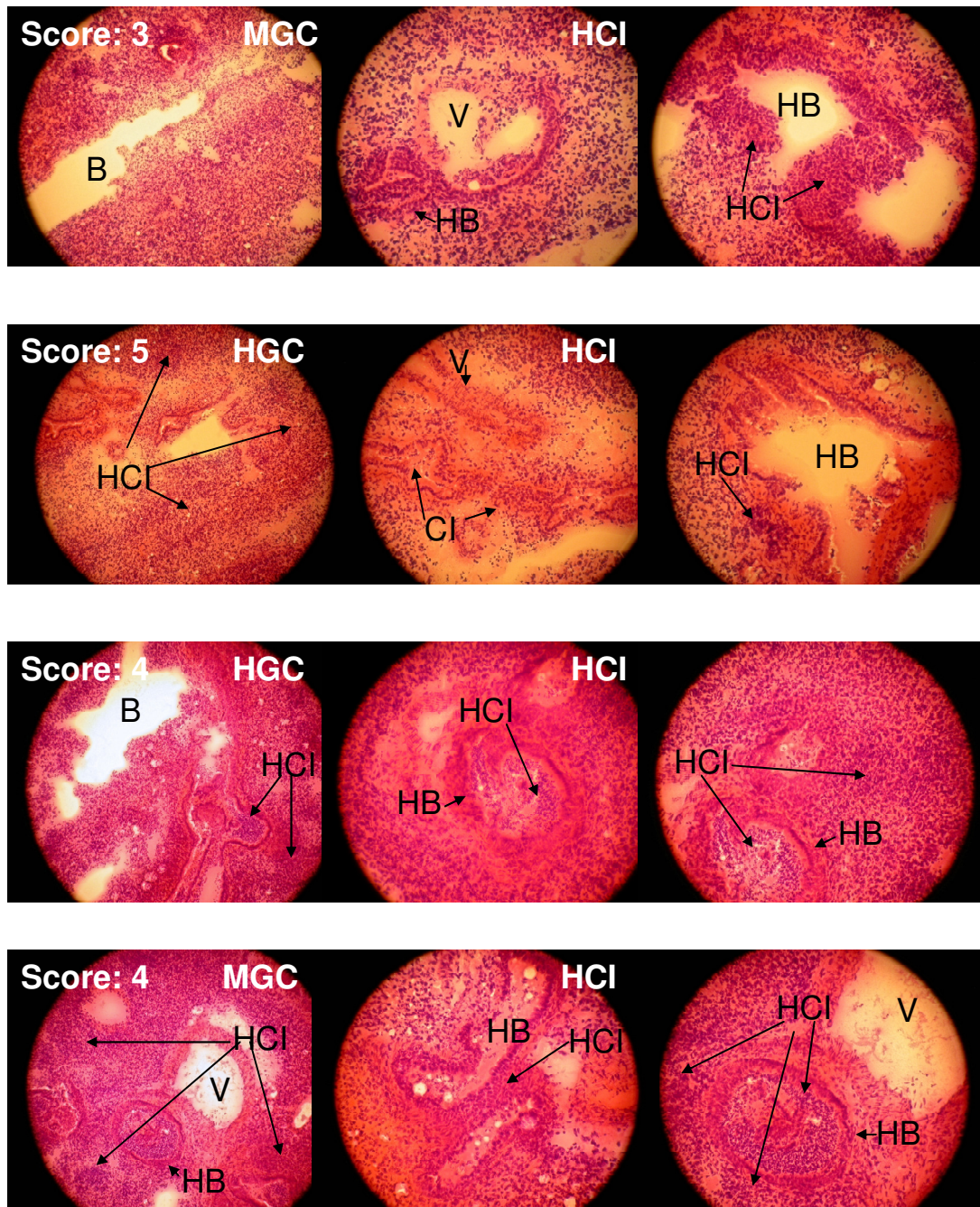


Figure 5-11 Histopathology lung sections produced from mice infected with 2×10^6 CFU/50 μ l at 18 hours post-infection

Left column represents from top to bottom a) LES400 b) LES431 c) LESB58 d) LESB65 at magnification x100; Middle and Right column represents the same stains respectively from top to bottom at magnification x200.

Key: LCI-MCI-HCI represents low/medium/high cellular infiltration; LGC-MGC-HGC represents low/medium/high general consolidation; B represents bronchi; V represents blood vessel

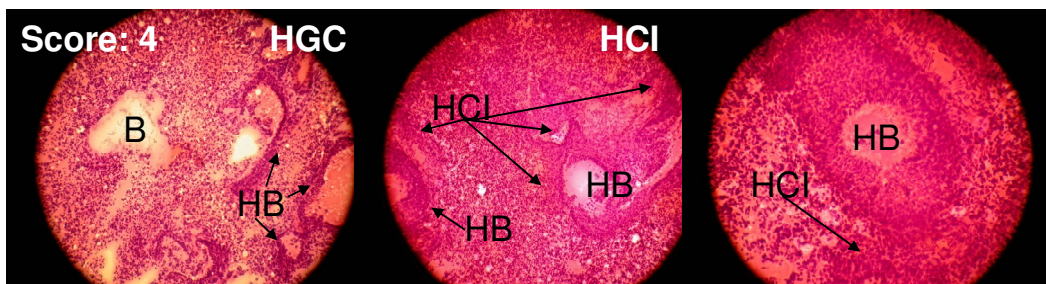
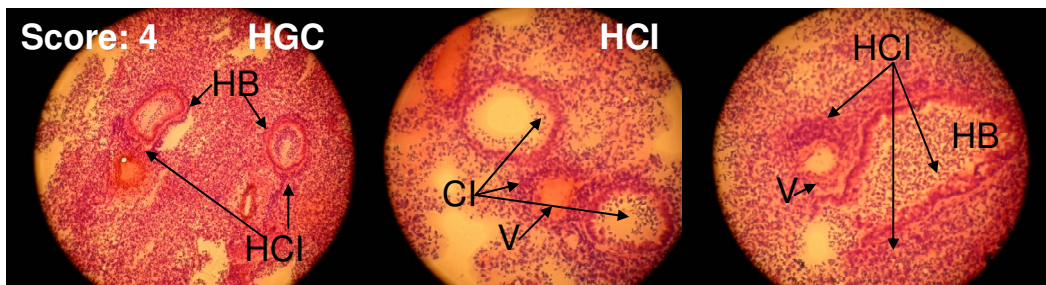
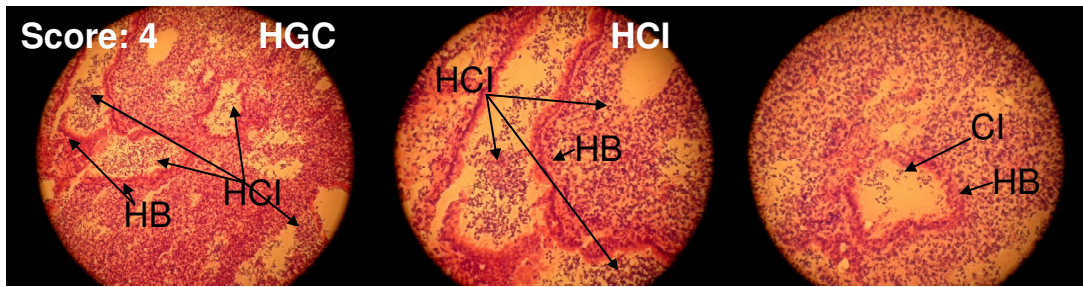


Figure 5-12 Histopathology lung sections produced from mice infected with 2×10^6 CFU/50 μ l at 55 hours post-infection

Left column represents from top to bottom a) LES400 b) LESB58 c) LESB65 at magnification x100; Middle and Right column represents the same stains respectively from top to bottom at magnification x200.

Key: LCI-MCI-HCI represents low/medium/high cellular infiltration; LGC-MGC-HGC represents low/medium/high general consolidation; B represents bronchi; V represents blood vessel

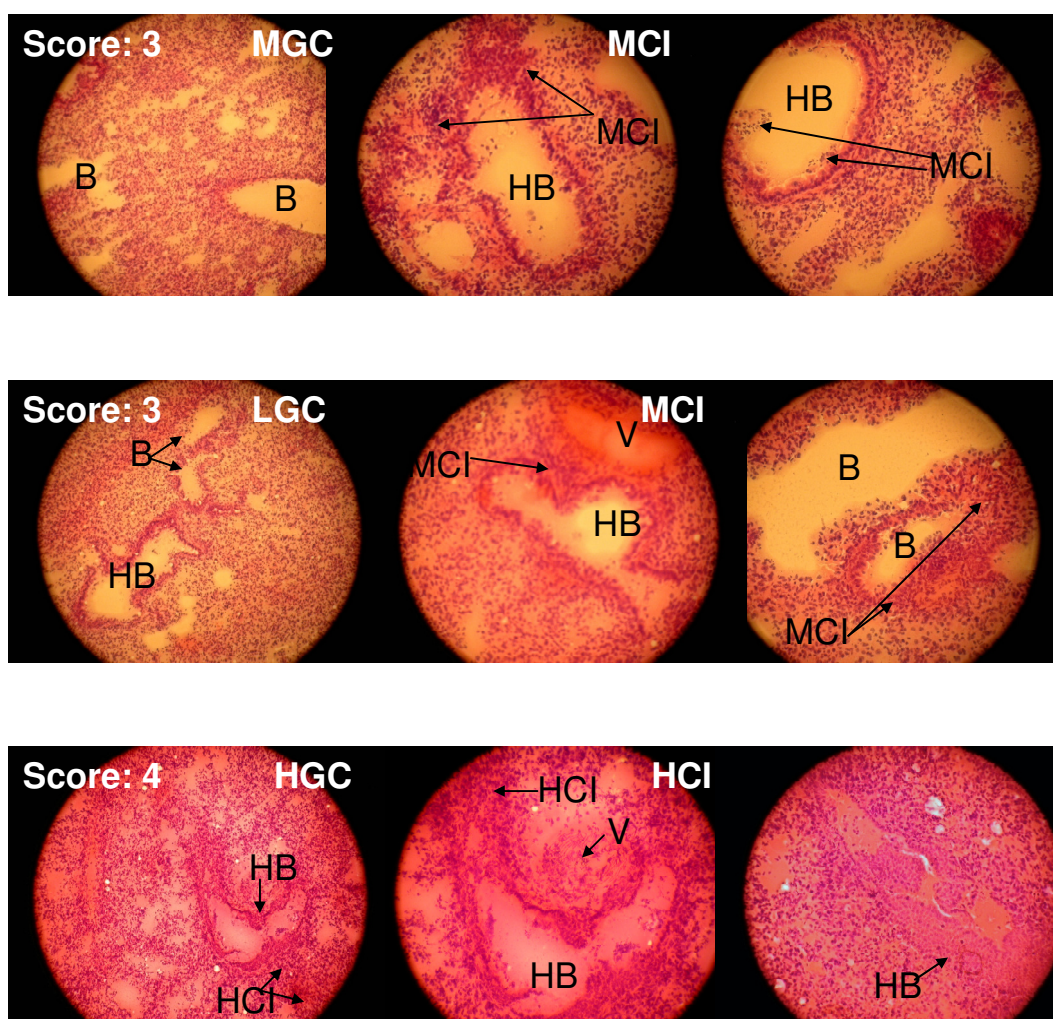


Figure 5-13 Histopathology lung sections produced from mice infected with 2×10^6 CFU/50 μ l at 96 hours post-infection

Left column represents from top to bottom a) LES400 b) LESB58 c) LESB65 at magnification x100; Middle and Right column represents the same stains respectively from top to bottom at magnification x200.

Key: LCI-MCI-HCI represents low/medium/high cellular infiltration; LGC-MGC-HGC represents low/medium/high general consolidation; B represents bronchi; V represents blood vessel

5.2.5 Summary of acute respiratory infection outcomes

In the survival experiments, LES431-infected mice did not survive beyond 24 hours post-infection. Of the mice infected with LESB65, 60% survived the experiment, whereas all the mice infected with LES400 or LESB58 survived the experiment. The symptom scores of LESB58-infected mice improved continually from 18 hours

post-infection to that of healthy mice (symptom score = 0) by 96 hours post-infection. The symptoms scores coincides with the lung CFU that peak at 18 hours and decreased to undetectable levels by 48 hours post-infection. This coincides with the lung damage that is consistent through 18 to 55 hours post-infection and improves by 96 hours post-infection. In contrast LES400-infected mice had their symptom scores peak at 48 hours post-infection and improved to 96 hours post-infection; this also correlates with the lung CFU which peaked at 48 hours post-infection and decreased steadily till 96 hours post-infection. This also coincides with the lung histopathology that peaks at 55 hours post-infection and improved until 96 hours post-infection. The LESB65-infected mice maintained a consistent median symptom score throughout the experiment. The CFU numbers recovered from the nasopharynx and lungs increased steadily from 18 hours to 96 hours post-infection. This correlates with the increase in histopathology score between 18 and 55 hours post-infection ($4 > 4.5$) and this high score was maintained until 96 hours post-infection. Based on the data generated during this project, the most virulent to the least virulent LES isolate is LES431, to LESB65, to LES400 and finally LESB58.

5.3 Discussion

The aim of this study was to determine whether the over-expression of quorum-sensing products is a reliable marker for increased virulence exhibited by LES isolates. This hypothesis was tested in an acute respiratory infection model. The over-expressing isolates were represented by LES431, LESB58 and LESB65. The deficient isolates were represented by LES400.

This section proceeds by discussing quorum-sensing over-expression as a marker of virulence of *P. aeruginosa* within a murine acute respiratory model. This is followed by discussions on the differences in virulence between LES isolates. The final discussion point is to compare the virulence of the LES isolates and the reference strains, PAO1 and PA14, within a murine acute infection model. This is followed by the conclusions and future work.

5.3.1 Over-expression of quorum-sensing products as a marker for virulence in an acute respiratory model

Several LES isolates have been highlighted to exhibit an over-expression of quorum-sensing products labelled as the hypervirulence phenotype (Fothergill *et al.* 2007, Salunkhe *et al.* 2005). An increased virulence of LES isolates with the hypervirulence phenotype was first noted in a *Drosophila* fly model of infection (Salunkhe *et al.* 2005); when comparing over-expressing quorum-sensing products isolate LES431 to that of a quorum-sensing deficient isolate LES400. These findings were later replicated in a *C. elegans* model of infection (Winstanley C, University of Liverpool, unpublished), where isolates with the hypervirulence phenotype (including LESB58 and LESB65) had a lower LT₅₀ in the slow killing assay.

Previously published work has shown that a deficiency in the quorum-sensing system can reduce virulence in a number of acute murine infection models. The use of quorum-sensing deficient mutants in a murine burns injury model led to a reduction in mortality of the mice and the ability of the bacteria to disseminate (Rumbaugh *et al.* 1999). In a murine interperitoneal infection model they noted that clearance of the bacteria was more rapid in quorum-sensing deficient mutants (Christensen *et al.* 2007). There have also been two previous studies (Pearson *et al.* 2000, Tang *et al.* 1996) that have shown that in a neonatal murine acute respiratory model the loss of quorum-sensing related genes resulted in a reduction in morbidity and mortality. The models involved using 7 to 10 day old Balb/c mice inoculating them with either 5×10^6 CFU (Tang *et al.* 1996) or 1×10^9 CFU (Pearson *et al.* 2000) per mice. Pearson *et al.*, constructed a PAO1 mutant which lacked two quorum-sensing associated genes (*lasI* and *rhII*), termed PAO_JP2. Using this mutant in the neonatal respiratory model increased the survival of the mice (21% versus 5%), as well reducing the chance of developing pneumonia (56% versus 10%) and bacteraemia (~50% versus ~15%) 18 hours post-infection when compared with the wild-type, PAO1. Tang *et al.*, presented similar results using a *lasR* mutant of PAO1.

Tang *et al.* (1996) performed additional *in vitro* experiments with the *lasR* mutant. One experiment involved an adherence assay on nasal polyp cells and the results showed that the *LasR* mutant had 26% binding in contrast to wild-type PAO1 which had 58%. The authors suggested that the difference in binding was due to *LasR*-dependent exoproducts which aid attachment to the cell surface. The data presented in this thesis supports this finding. LES400, which is a quorum-sensing deficient due to a mutation in the *lasR* gene, exhibited a poor recovery of CFU from the nasopharynx. In contrast, over-expressing quorum-sensing isolates, LES431 and LESB65 had significantly higher ($P < 0.05$) CFU recovered from the nasopharynx suggesting they were better at persisting within the nasopharynx. The anomaly to this trend is LESB58, which was cleared from the nasopharynx more rapidly than LES400. Further analysis could elucidate the reasons for this. It is possible that LESB58 could have secondary mutation(s) in a gene or gene cluster or transcriptional regulator of a pathway involved in attachment of *P. aeruginosa* to host cells within the nasopharynx.

The literature provides data to demonstrate that quorum-sensing is an important *P. aeruginosa* virulence factor; as quorum-sensing deficient mutants are less virulent than the wild-type strain. The limited data provided for quorum-sensing over-expression mutants suggests they are more virulent than wild-type or deficient expression strains.

In general, this trend correlates with the data provided within this thesis. The over-expressing isolates, LES431 and LESB65 are more virulent than the deficient isolate, LES400. This is reflected in significantly ($P < 0.05$) higher CFU recovered from the nasopharynx and lungs at 18 hours post-infection in the over-expressing isolates. The histopathology scores at 18 hours post-infection were also higher and the survival rates were also significantly ($P < 0.05$) different. LES431-infected mice did not survive beyond 24 hours, 60% of LESB65-infected mice survived the experiment, in contrast to LES400-infected mice which all survived the experiment. The exception to the trend is LESB58 which is less virulent than LES400 in this model. This is reflected in lower CFU numbers recovered from the nasopharynx and lungs at 18 hours post-infection. LES400-infected mice had an increase in symptom scores between 18 and 48 hours post-infection, which coincided with a peak of CFU

numbers within the lungs and then proceed to recover. In contrast, LESB58-infected mice had their peak in symptom scores at 18 hours post-infection and recovered beyond this timepoint. This coincided with a decrease in CFU recovered from the lungs. This anomaly suggests that quorum-sensing over-expression is not always an over-riding virulence determinant in terms of acute respiratory infection.

In conclusion, quorum-sensing over-expression is a reliable marker for increased virulence of *P. aeruginosa* within a murine acute respiratory model, but it is possible that the effects of quorum-sensing can be over-ridden in some cases. The quorum-sensing over-expressing LES isolates, LES431 and LESB65 were more virulent than the deficient isolate, LES400. LESB58, an over-expressing isolate appears to be an anomaly to this trend as it is less virulent than LES400. Further analysis of the differences in virulence between the LES isolates are explored in the following subsection.

5.3.2 The differences in virulence between the LES isolates

All of the isolates used during this project are natural mutants of *P. aeruginosa* LES. The literature shows that there is a conserved genomic backbone between *P. aeruginosa* strains of approximately 90% (Wolfgang *et al.* 2003). As these isolates are from the same strain, the genome similarity should be closer to 100%. The differences between isolates will be most likely attributed to an additive effect of a small number of gene mutations and changes in gene expression profile. A research project that was targeted at determining the differences between the genomes of these isolates could determine the virulence factors that contribute to respiratory disease and pinpoint the differences in virulence. Genomic variation between LES isolates has been previously reported (Fothergill *et al.* 2007). The small genome differences are exemplified by Figure 5-14 showing the PFGE profile of LES400 against that of LES431 and PAO1; there is a two band difference between LES400 and LES431.

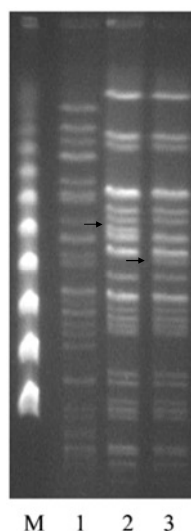


Figure 5-14 PFGE comparison of the LES isolates and strain PAO1 (Salunkhe *et al.* 2005). **M**, pulse marker, 50 to 1,000 kb (Sigma-Aldrich); lane 1, PAO1; lane 2, LES400; lane 3, LES431. Reproduced by permission from American Society for Microbiology: *Journal of Bacteriology* 2005 Jul;187(14):4908-20.

The genome of LESB58, the earliest available isolate of the LES, has been sequenced (Winstanley *et al.* 2009). The authors reported the mobile genome (variable regions) of LESB65 consists primarily of six prophage gene clusters and five genomic islands. Table 5-9 shows a comparison of the LES isolates based on available data. The table summarises the quorum-sensing ability, difference in genomic islands and other variable DNA regions, the presence of Type III secretion system (TTSS) exotoxins and colony morphology.

	LES400	LES431	LESB58	LESB65
Quorum-sensing	Deficient	Over-expression	Over-expression	Over-expression
PAGI-1	✓	✓	✓	✓
PAGI-2	✓	✓	✓	✓
LES Genomic Island -5	✓	✓	✓	✓
Prophage 2	✓	✗	✓	✓
Prophage 3	Unconfirmed	Unconfirmed	✓	✓
Prophage 4	✓	✓	✓	✓
Prophage 5	Unconfirmed	Unconfirmed	✓	✗
TTSS	ExoS	ExoS	ExoS	ExoS
Colony morphology	SCV	SCV	SCV	NM
Infection source	CF	Non-CF	CF	CF

Table 5-9 shows summary of the genomic, phenotype and source of each of the LES isolates
Genomic information gathered by written communication with Winstanley, C, University of
Liverpool. Key: SCV (Small colony variant) and NM (Normal morphology).

Further work could elucidate the genomic differences between the over-expressing quorum-sensing isolates. It would be valuable to compare the two extremes, LES431 and LESB58. LES431-infected mice do not survive beyond 24 hours, the mice have high bacterial numbers within the lungs and the majority have bacteraemia. This is in contrast to LESB58-infected mice, which clear the bacteria within 48 hours and 100% of the mice survive the 96 hour survival experiment. The *Caenorhabditis elegans* model of infection does not differentiate between the isolates (Winstanley C, University of Liverpool, unpublished), implying that the

isolates have a similar virulence in the model tested. The work described in this thesis shows there is a larger continuum, supporting the notion that *in vivo* work within mammalian models of infection are important for assessing virulence and for detecting subtle differences in virulence. The acute respiratory model of infection has shown that over-expression of quorum-sensing products is not an over-riding factor and that other factors must also have a vital role in virulence. The results of further investigation between these two isolates could show the mutation or acquisitions that enable LES431 to cross from a CF individual to a non-CF individual. LES431 was isolated from a non-CF parent of a CF patient (McCallum *et al.* 2002). The isolate was originally labelled as an acute infection isolate (Salunkhe *et al.* 2005) but the non-CF patient was later re-admitted to hospital with respiratory symptoms in 2007 and LES was recovered from the sputum (Winstanley C, University of Liverpool, unpublished); suggesting a chronic infection.

LESB65 represents a strain amidst the two extremes in terms of survival post-infection, as 60% of the mice challenged with LESB65 survived the 96 hour experiment. However, the lung and nasopharynx CFU numbers increased steadily over the same 96 hour period. At 18 hours post-infection, the LES431 and LESB65 have no significant differences ($P < 0.05$) in lung CFU. At this timepoint, the LES431 were culled as they reached the maximum symptom score of 6, whereas the lung CFU within LESB65-infected mice continued to rise over the 96 hours and resulted in a 10% loss of mice at 48 hours and further 30% loss at 96 hours post-infection. Further research would be important to analyse the differences between the two isolates at 18 hours post-infection to determine why LES431 cause more severe morbidity compared to LESB65 despite having a similar CFU burden within the lungs. A possible explanation is that LES431-infected mice develop septicaemia secondary to pneumonia resulting in systemic multi-organ failure. LESB65-infected mice in comparison rarely (10%) developed septicaemia. LESB65 persists in the lungs throughout the 96 hour experiment; in contrast LESB58 was cleared within 48 hours. Comparing these two isolates for gene mutations and changes in expression profile could highlight genes that are essential for persistence within the lungs. The work was not attempted due to time constraints and considered beyond the scope of this project. This aim of the project was to evaluate the effect of quorum-sensing on the *in vivo* virulence of the LES isolates and not secondary gene mutations.

Comparing strains from the same type of infection could also provide valuable data. In the current set of isolates LES431 was recovered from an acute respiratory infection and therefore it is expected to perform well in an acute respiratory infection model. LES431 is therefore expected to out-perform the CF isolates which were recovered from a chronic infection. The data provided in this thesis supports this notion. Each CF patient has a population of *P. aeruginosa*; this population usually consists of a single strain that is a mixture of isolates that vary by genotype and phenotype (Fothergill *et al.* 2007). The data presented in this thesis examines single pure isolates of a strain, LES. The literature suggests that CF isolates of *P. aeruginosa* undergo changes in their genome to rid themselves of acute virulence determinants (D'Argenio *et al.* 2007, Smith *et al.* 2006). LES431 had not undergone these changes which enables the isolate to be virulent in an acute infection model. An interesting hypothesis is the change from wild-type expression to quorum-sensing over-expression, which would allow LES431 to be more successful acute infection isolate. This change was needed to increase the production of acute virulence determinants and make the jump from CF isolate to non-CF isolate.

One of the most common genetic changes in CF isolates is the loss of LasR function as exemplified by LES400 (D'Argenio *et al.* 2007, Smith *et al.* 2006), the quorum-sensing transcriptional regulator. A number of acute infection models have shown that isolates deficient in quorum-sensing (including *lasR* mutants) are less virulent than their wild-type counterparts (Christensen *et al.* 2007, Pearson *et al.* 2000, Rumbaugh *et al.* 1999, Tang *et al.* 1996). It is suggested this is due to selective pressures within the CF airway to select against acute virulence determinants. The advantages of LasR mutants have been investigated. D'Argenio *et al.*, 2007 showed that the LasR mutants had an advantage over wild-type strains, for growth on selected carbon and nitrogen sources. Their findings also showed that the loss of function of LasR led to an increase in antibiotics resistance, noted as an increase in beta-lactamase activity. In this context, the loss of function of LasR would reduce the production of many virulence factors due to a single mutation, in contrast to acquiring many small mutations in each of the virulence genes (Nguyen, Singh 2006). The current literature would suggest that to succeed as a CF isolate, the isolate must lose its ability to cause acute infection.

The data presented in this thesis supports this notion. LES400 and LESB58-infected mice survive and clear the infection within 96 hours. LES400 is a LasR mutant, so has lost function of many of its acute virulence determinants. This would result in an isolate that is poor at causing acute infection and in turn a more successful chronic isolate. LESB58 is also poor at causing acute infection, suggesting it has acquired mutations that decrease the expression of acute virulence determinants. The cause of the decreased ability of the LES isolate to cause acute infection does not appear to have affected its ability to cause chronic infection. Kukavica-Ibrulj *et al.* 2008 demonstrated that LESB58 has the ability to cause infection in a rat respiratory model of chronic infection. A generalisation from the data generated during this project is that CF isolates are cleared within 96 hours post-infection, as these isolates had lost many of their acute virulent determinants. The exception being LESB65, as 60% of the mice infected survive the experiment and the CFU numbers recovered from the nasopharynx and lungs increases steadily over the 96 hours. Further research between LESB65 and the other CF isolates could clarify the differences in virulence.

There are published data available that oppose the theory that there is selection for a decrease in acute virulent determinants over the course of the chronic infection. The data suggest the *P. aeruginosa* population of a CF patient maintains a mixture of phenotypes and genotypes, with no skew over the period of chronic infection. For example, Fothergill *et al.* (2007) suggest that CF patients continue to have a mixture of quorum-sensing phenotypes (over-expression, wild-type and deficient) within their *P. aeruginosa* population throughout their infection.

In conclusion, the differences in virulence between the LES isolates could be a reflection of two factors; the quorum-sensing ability and the type of infection where the isolate originated. The data provided in this thesis could be valuable for future research, as the isolates used are from the same Liverpool Epidemic Strain. This makes the genomic differences between the isolates small. Detection of these genomic differences as well as changes in gene expression profile could highlight genes that play important roles in acute respiratory infection. The data could

determine genes that are important for acute and chronic respiratory infection as well as genes that are essential for persistence in the nasopharynx and lungs.

5.3.3 The LES isolates versus the reference strains, PAO1 & PA14

The LES isolates have been compared in several research papers to the *P. aeruginosa* reference strains, PAO1 and PA14. In a *Drosophila* fly infection model (Salunkhe *et al.* 2005), both LES431 and LES400 were shown to be more virulent than PAO1. In the *C. elegans* model of infection (Winstanley C, University of Liverpool, unpublished), all the LES isolates had a lower LT₅₀ than PAO1. In a chronic infection model within rats (Kukavica-Ibrulj *et al.* 2008), the growth curve of the bacteria within the lungs and thereby the lung colonisation capability were shown to be similar for LESB58, PAO1 and PA14.

In this thesis the LES isolates and the reference strains have been analysed using a newly developed acute respiratory model as described in methods section. Mice infected with PAO1 did not survive beyond 24 hours and PA14-infected mice did not survive beyond 18 hours. LES431 infected mice had a similar survival curve to PA14. This coincides with significantly ($P < 0.05$) similar levels of bacteria found within the lungs at 18 hours post-infection. This suggests LES431 and PA14 have a similar virulence within the acute respiratory model. Further research could be performed to analyse the two genomes and gene expression profile to highlight the common features. These common features could be compared to the other LES isolates used during this project; especially LESB58 as it is the least virulent of the LES isolates. The results could generate a number of virulence factors that are important for acute respiratory infection within *P. aeruginosa*. PA14 and LESB65 also had significantly ($P < 0.05$) similar CFU numbers within the lungs at 18 hours post-infection. But 60% of the LESB65-infected mice survived the experiment, and even had a steady increase in the number of bacteria within the lungs over the 96 hour period. Uncovering the genomic differences between PA14 and LESB65 could provide further interesting data.

PAO1 and PA14-infected mice developed septicaemia by 18 hours post-infection. The levels of bacteria within the blood were significantly higher ($P < 0.05$) in PA14

than PAO1. Despite the bacterial lung CFU being insignificantly different ($P < 0.05$) between PA14, LES431 and LESB65 at 18 hours post-infection, they differed significantly ($P < 0.05$) from one another in blood CFU levels. 100% of PA14-infected mice developed septicaemia, 60% for LES431 and 10% for LESB65. The blood CFU in PA14 was significantly higher ($P < 0.05$) than LES431 and LESB65. In turn, LES431 was significantly higher ($P < 0.05$) than LESB65.

These results are supported in the literature. Acute respiratory infection usually leads to dissemination of *P. aeruginosa* from lungs to the blood, followed closely by death (Furukawa, Kuchma & O'Toole 2006). In chronic infection, dissemination from the lungs to the blood is rare (Furukawa, Kuchma & O'Toole 2006). PAO1, PA14 and LES431 were originally isolated from acute infections, so respiratory infection leads to septicaemia. LES400, LESB58 and LESB65 were from chronic infections, thereby less likely to cause septicaemia. This is supported by the data presented in this thesis. LES400 and LESB58-infected mice never developed septicaemia over the 96 hours post-infection and only 10% of LESB65-infected mice develop septicaemia. Interestingly, LES431 and PA14 have similar lung CFU but significantly differ in blood CFU. They were both isolated from acute infection. An important point to note is that PA14 harbours ExoU, which has been demonstrated as an important virulence factor for dissemination from the lungs to the blood in acute respiratory infection (Shaver, Hauser 2004). In contrast LES431 has ExoS, which has been demonstrated in literature as less potent in this role (Shaver, Hauser 2004).

In conclusion, the results generated from the *in vivo* virulence of *P. aeruginosa* projects demonstrate that *P. aeruginosa* adapts to their infection lifestyle. *P. aeruginosa* isolated from acute infections are the most virulent in an acute respiratory model, exemplified by PAO1 (wound isolate), PA14 (burns isolate) and LES431. These acute infection isolates cause death within 24 hours post-infection as the mice develop septicaemia. This is in contrast to the chronic infection isolates, LES400 and LESB58 that are cleared from the lungs by 96 hours post-infection.

5.4 Conclusions and further work

This project explored *in vivo* virulence of *Pseudomonas aeruginosa*. The project utilised the murine acute respiratory model developed in this thesis to assess whether over-expression of quorum-sensing products is a marker of virulence. This project demonstrated that over-expression of quorum-sensing products is a marker of increased virulence within an acute respiratory model. The data shows the over-expressing isolates, LES431 and LESB65 were more virulent than the deficient isolate, LES400. The exception to this generalisation was LESB58, which was less virulent than LES400. This anomaly could be used to support the hypothesis that CF isolates adapt to the CF airway environment by reducing the expression of virulence determinants involved in acute infection. This is the first report of quorum-sensing over-expressing isolates causing increased virulence in an acute infection model. This study is also the first reported case of using LES isolates in an acute respiratory model as the LES isolates are principally considered members of an aggressive chronic CF strain and are usually reported in this context.

The data suggested a broader continuum of virulence for the quorum sensing over-expressing LES isolates in an acute respiratory infection model. Previous studies have been unable to distinguish between the over-expressing isolates, such as the *C. elegans* model of infection (Winstanley C, University of Liverpool, unpublished). Further research that aimed to discover the genomic differences between the LES isolates could clarify the differences in virulence and provide data for genes involved in persistence within the lungs and nasopharynx, as well highlighting acute virulence determinants for acute respiratory infection.

6 **Thesis conclusion**

The underlying themes of this thesis are pathogenicity island characterisation *in vitro* and *in vivo* as well as *in vivo* virulence of *P. aeruginosa*. The rationale behind this focus is that 10-20% of the *P. aeruginosa* genome is variable between strains. Large variable regions such as genomic (pathogenicity) islands are considered more likely to contribute to the differences in disease-causing ability between strains.

The first project focused on pathogenicity island characterisation *in vitro*. The aim of the project was to create a generic genomic island capture technique to aid discovery and characterisation of novel genomic islands. This aim was achieved by developing the genomic island capture technique as described by Wolfgang *et al*, 2003, by testing the parameters in three stages. This included the DNA preparation used in transformation, the process of preparing yeast spheroplasts and finally the screening process to detect clones that have undergone homologous recombination and therefore contained the genomic island. The technique was used to capture genomic islands from bacterial species other than *P. aeruginosa*, which included *E. coli* and supported work with *Acinetobacter baumannii* (Shaikh F, PhD thesis, University of Leicester, unpublished). The use of the technique led to the characterisation of a novel genomic island in *E. coli*, E105-*leuX*. The project also used the technique to compare captured genomic islands to those that had previously been identified (E106-*serW*, KR115-*lys10* and KR159-*lys10*). The project has opened up a new technique to explore the mobile genome across a number of bacterial species. It has also highlighted the usefulness of the technique to discover novel genomic islands.

The second project focused on *in vivo* characterisation of pathogenicity islands. The project relied on the development of a novel murine acute infection model and is the first reported use of *P. aeruginosa* in a murine acute infection model within the UK. The aim of the project was to analyse the contribution of the two pathogenicity islands, PAPI-1 and PAPI-2, to the virulence of *P. aeruginosa* PA14 in murine acute respiratory model of infection. To achieve this aim, a set of isogenic mutants were used, a mutant with PAPI-1 deleted, another with PAPI-2 deleted and a third

mutant with both PAPI-1 and PAPI-2 deleted. The results show that both pathogenicity islands contribute to virulence in a murine acute respiratory model of infection. The results demonstrated that the presence of PAPI-2, possibly ExoU, is enough to maintain wild-type virulence. The data presented suggested that recovery of the nasopharynx can provide an important variable for assessing virulence of *P. aeruginosa*. The results from this project suggest that PAPI-2 have a role in infection within the upper respiratory tract. The data generated also suggests that PAPI-2 is important for dissemination of PA14 from the lungs to the blood, but not for survival within the blood as demonstrated by testing in an intravenous sepsis model of infection. The results supported the literature by showing the difference in virulence between PAO1 and PA14 is more complex than the presence or absence of pathogenicity islands alone. The virulence of PAO1 was more closely mimicked by PA14 Δ PAPI-2 than PA14 Δ PAPI-1 Δ PAPI-2. This is despite PAO1 itself not containing either PAPI-1 or PAPI-2. The project has shown that the use of pathogenicity island mutants in virulence models is feasible. The use of this model in the future could aid the labelling of known genomic islands as pathogenicity islands. The data generated could also be used as a starting point to pin point virulence genes within PAPI-1.

The third project focused on *in vivo* virulence of *P. aeruginosa*. The aim of the project was to determine if quorum-sensing over-expression was a reliable marker of virulence within *P. aeruginosa* Liverpool Epidemic Strain (LES) isolates in a murine acute respiratory model. The results suggest that this theory is generally true in an acute respiratory model of infection. The over-expressing isolates (LES431 and LESB65) were more virulent than the deficient isolate (LES400), the exception being the over-expressing isolate LESB58. The results also showed that LES isolates (as well as PAO1 and PA14) from acute infection were more virulent than those from chronic infection (CF patients). The project has highlighted a need to determine the genomic differences between the isolates to explain the continuum of virulence. Further work that could expand the data generated during this thesis is to evaluate LES isolate, LES416 in the acute respiratory model. This LES isolate was taken from the CF patient who parents also developed a respiratory infection, to compare its virulence with LES 431, which was taken from one of the CF patient's parents (McCallum *et al.* 2002).

Hopefully the methodologies described during this thesis will enable future project aimed at evaluating pathogenicity islands to link *in vivo* and *in vitro* data analysis. The *in vitro* work could enable rapid sequencing of genomic islands as well as allowing detailed analysis of the functional role of the genes present. When combined with the *in vivo* work, known and newly discovered genomic islands could be described as pathogenicity islands. This is especially important for *P. aeruginosa* where for example in PAO1, 44% of the ORFs within the genome have no known function (Class 4).

P. aeruginosa is currently the largest of all sequenced bacterial genomes. It contains a large portion of genes with no known homologues within other bacterial species. Both the *in vivo* projects described in this thesis could contribute to elucidating the role of virulence genes, at least in an acute respiratory model. A more detailed comparison between the LES isolates used during this thesis could highlight key genomic differences between the most virulent isolate (LES431) and least virulent (LESB58), as well as highlight new markers of virulence within *P. aeruginosa*.

7 Appendix

7.1 RFLP comparison between APECO1-serW and E106-serW

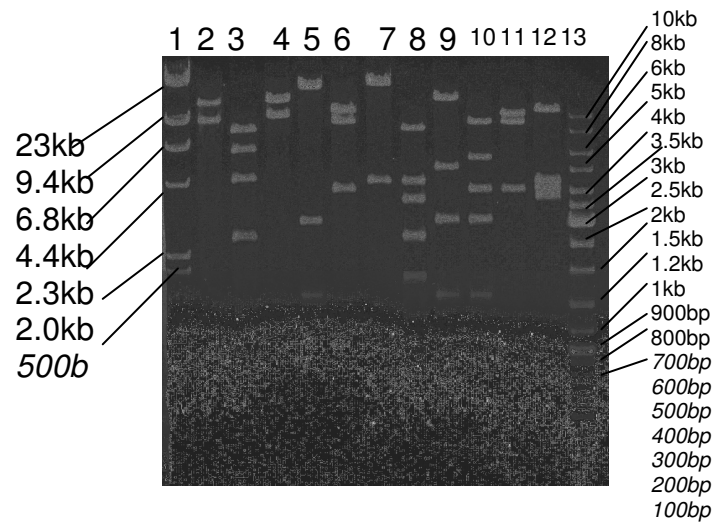


Figure 7-1 shows restriction profile of E106-serW, please see table 7-1 for detailed comparison

1 λ HindIII ladder	2 <i>I-SceI</i>		3 <i>EcoRV</i>		4 <i>HindIII</i>		5 <i>KpnI</i>		6 <i>PstI</i>		7 <i>Sall</i>		8 <i>EcoRV</i> - <i>HindIII</i>	9 <i>KpnI</i> - <i>HindIII</i>	10 <i>KpnI</i> - <i>PstI</i>	11 <i>PstI</i> - <i>HindIII</i>	12 <i>PstI</i> - <i>Sall</i>	13 Generuler
500ng	APEC O1	E106	APEC O1	E106	APEC O1	E106	APEC O1	E106	APEC O1	E106	APEC O1	E106	APEC O1	APEC O1	APEC O1	APEC O1	APEC O1	500ng
	13335	~14kb	8135	~8kb	13167	~14kb	18231	~19kb	10593	~11kb	18964	~19kb	8135	13167	8819	10046	10953	
	9922	~10kb	6075	~6kb	10090	~10kb	3057	~3kb	8825	~9kb	4293	~4.5kb	4180	5026	5573	8781	4293	
			4526	~4.5kb			1621	~1.5kb	3839	~4kb			3674	3057	3839	3839	3839	
			2668	~2.8kb			348						2668	1621	3057	547	3749	
			1003	~1 kb									1895	348	1615	44	783	
			850	~900bp									1003	38	348			
													852		6			
													850					

Table 7-1 shows the comparison between E106-*serW* restriction profile and the *in-silico* profile generated for APECO1-*serW*. The lane numbers within the table are in reference to the lanes from left to right in Figure 7-1

7.2 Research Publication

Prior to using *P. aeruginosa* in a respiratory model of infection, I generated data using *Streptococcus pneumoniae* and the following publication uses this data.

Sialic Acid: A Preventable Signal for Pneumococcal Biofilm Formation, Colonization, and Invasion of the Host

Claudia Trappetti,^{1,a} Aras Kadioglu,^{3,a} Melissa Carter,³ Jasvinder Hayre,¹ Francesco Iannelli,¹ Gianni Pozzi,^{1,2} Peter W. Andrew,³ and Marco R. Oggioni^{1,2}

¹Laboratorio di Microbiologia Molecolare e Biotecnologia, Dipartimento di Biologia Molecolare, Università di Siena and ²UOC Batteriologia, Azienda Ospedaliera Universitaria Senese, Siena, Italy; ³Department of Infection, Immunity, and Inflammation, University of Leicester, Leicester, United Kingdom

The correlation between carbohydrate availability, pneumococcal biofilm, nasopharyngeal colonization, and invasion of the host has been investigated. Of a series of sugars, only sialic acid (i.e., *N*-acetylneuraminic acid) enhanced pneumococcal biofilm formation in vitro, at concentrations similar to those of free sialic acid in human saliva. In a murine model of pneumococcal carriage, intranasal inoculation of sialic acid significantly increased pneumococcal counts in the nasopharynx and instigated translocation of pneumococci to the lungs. Competition of both sialic acid–dependent phenotypes was found to be successful when evaluated using the neuraminidase inhibitors DANA (i.e., 2,3-didehydro-2-deoxy-*N*-acetylneuraminic acid), zanamivir, and oseltamivir. The association between levels of free sialic acid on mucosae, pneumococcal colonization, and development of invasive disease shows how a host-derived molecule can influence a colonizing microbe and also highlights a molecular mechanism that explains the epidemiologic correlation between respiratory infections due to neuraminidase-bearing viruses and bacterial pneumonia. The data provide a new paradigm for the role of a host compound in infectious diseases and point to new treatment strategies.

Recurrent colonization of the nasopharynx is a common characteristic of respiratory pathogens. Although colonization is normally asymptomatic, it is believed to be a forerunner of the invasive diseases caused by these pathogens; however, the events that bring about this key change from colonization to invasion are unknown. To address this conundrum, we present new data derived from in vitro and in vivo experiments involving *Streptococcus pneumoniae*.

S. pneumoniae is a gram-positive, anaerobic, aerotolerant bacterium responsible for a variety of acute dis-

eases in humans, including otitis media, pneumonia, sepsis, and meningitis [1–3]. The burden of invasive pneumococcal disease is great worldwide, in both developed and developing countries. Colonization of the human nasopharynx by *S. pneumoniae* is a natural process that occurs during the first few months of life and is thought to be a prerequisite for invasive pneumococcal infection. Although most colonized individuals are asymptomatic, progression from colonization to disease has been reported to occur soon after acquisition of a new colonizing strain [3–5]. Successive episodes of colonization are common, and these episodes, along with the incidence of pneumococcal invasive disease, have a clear seasonal variation (with the peak incidence occurring in winter or early spring [4, 6]), indicating the influence of environmental factors on host-pathogen interactions.

The human nasopharynx is the principal ecological niche of many bacterial pathogens. Although this environment offers bacteria a plethora of signals, sugars present in the nasopharynx primarily have been regarded as nutrients or as foci for adhesion, although they

Received 25 August 2008; accepted 4 December 2008; electronically published XX March 2009.

Potential conflicts of interest: none reported.

Financial support: European Union (grants LSHM-CT-2005–512099 [to M.R.O. and P.W.A.] and LSHB-CT-2005–512061 [to G.P. and M.R.O.]).

^a C.T. and A.K. contributed equally to this work.

Reprints or correspondence: Dr. Marco R. Oggioni, LAMMB (Laboratorio di Microbiologia Molecolare e Biotecnologia), Policlinico Le Scotte (lotto 5, piano 1), 53100 Siena, Italy (oggioni@unisi.it).

The Journal of Infectious Diseases 2009; 199:xxx

© 2009 by the Infectious Diseases Society of America. All rights reserved.

0022-1899/2009/19910-00XX\$15.00

DOI: 10.1086/598483

also may be signals that lead to enhanced virulence. In the pneumococcus, the interconnection between environmental signals and the capacity to colonize or cause disease was first indicated by genomic screening for virulence factors, which identified uptake systems (including carbohydrate uptake systems [7]) as a requirement for the development of full virulence in animal models [8–10]. In other species, sugar-utilizing systems specific for host carbohydrates have been associated with virulence, including heparin in *Staphylococcus* organisms, fucose in *Bacteroides* organisms, and sialic acid in nontypeable *Haemophilus influenzae* and *Escherichia coli* [11–15]. To our knowledge, at this time, the only example demonstrated to have a significant influence in humans is sucrose-induced colonization and cariogenicity of *Streptococcus mutans*, a model in which the dual role of the dietary sugar (as fermentable carbohydrate and as substrate for extracellular biofilm-matrix production) determines the increased virulence of the bacterium [16].

MATERIALS AND METHODS

Strains and growth conditions. The pneumococcal strains used in the present study were the serotype 2 strain D39, the serotype 4 strain TIGR4, and their respective rough derivatives DP1004 and FP23. Bacteria routinely were grown in tryptic soy broth (TSB; Becton Dickinson) or tryptic soy agar supplemented with 3% vol/vol horse blood at 37°C in an atmosphere enriched with CO₂. Sialic acid-free liquid media were each prepared with 10 g/L yeast extract (Becton Dickinson), 5 g/L NaCl₂, and 0.5 mol/L 3% wt/vol K₂HPO₄.

Mutant construction. Isogenic mutants were constructed by means of gene splicing by overlap extension (i.e., SOEing), as described elsewhere [17]. The primers used for deletion of the *nanA* and *nanB* locus were based on the sequences of upstream segments NanA1 (TG TAGCCGTCATTTTATTGCTAC) and NanA2 (TCCACTAGTTCTAGAGCGATTTTCTGCCTGATGTTGGTAT), downstream segments NanA3 (ATCGCTCTTGAAGGGAATGCTATTTACACCATACTTCCT) and NanA4 (CAGCTTCGCCTTGCCGTAGGT), and spectinomycin cassettes IF100 (GCTCTAGAACTAGTGA) and IF101 (TTCCTTCAAGAGCGAT). The size of the locus deletion was 23 kb and included open-reading frames SP1674–SP1693 in TIGR4 and open-reading frames spr1518–spr1536 in DP1004 (R6 genome). The mutant in TIGR4 was named “FP285,” that in FP23 (rough TIGR4) was named “FP236,” and that in DP1004 (rough D39) was named “FP240.”

Microtiter biofilm methodologic approach. The methodologic approach used for the static model of *S. pneumoniae* biofilm has been described elsewhere [18]. In brief, bacteria were grown in 96-well flat-bottom polystyrene plates (Sarstedt) with inocula of 1:100. CSP1 (30 ng/mL) was used for D39 and its derivatives, and CSP2 (100 ng/mL) was used for TIGR4 and its derivatives. In all cases, microtiter plates were incubated at 37°C in an atmosphere

enriched with CO₂. To detach biofilm bacteria, plates were washed 3 times, sealed, floated on a sonicator water bath (a Transonic 460/H ultrasonic bath [Elma]), and sonicated for 2 s at 35 kHz. Sugars assayed for the enhancement of biofilm formation were prepared as filter-sterilized 20% wt/vol stocks and assayed at 0.2% wt/vol. The sugars used were adonitol (Sigma 45502), arabinose (Sigma A3131), cellobiose (Sigma C7252), fructose (Fluka 47740), fucose (Sigma F8150), galactose (Carlo Erba 453125), glucose (Baker 0115), glycerol (Baker 7044), inositol (Sigma I5125), lactose (Carlo Erba 457552), maltose (Carlo Erba 459865), mannitol (Sigma M4125), mannose (Sigma M6020), melibiose (Sigma M5500), *N*-acetyl-D-galactosamine (Sigma A2795), *N*-acetyl-D-glucosamine (Sigma A8625), *N*-acetylneuraminic (sialic) acid (Sigma A9646), *N*-glycolylneuraminic acid (Sigma G9793), raffinose (Fluka 83400), rhamnose (Sigma R3875), sorbitol (Sigma S1876), stachyose (Sigma S4001), starch (Carlo Erba 417585), sucrose (Baker 0334), trehalose (Fluka P0210), xylitol (Fluka 95649), and xylose (Sigma X1500). Solutions of carbohydrates had a neutral pH in PBS, except for *N*-acetylneuraminic (sialic) acid (2% wt/vol [pH 2], 0.6% [pH 4], 0.2% [pH 6.8], and 0.06% [pH 7.0]) and *N*-glycolylneuraminic acid (2% wt/vol [pH 2], 0.6% [pH 3], 0.2% [pH 5.5], 0.06% [pH 6.0], 0.02% [pH 6.5], and 0.006 [pH 6.8]), which had acidic pH. To evaluate competition for a sialic acid-dependent effect, the neuraminidase inhibitors DANA (2,3-didehydro-2-deoxy-*N*-acetylneuraminic acid; D9050 Sigma), zanamivir (4-guanidino-2,4-dideoxy-2,3-dehydro-*N*-acetylneuramic acid (Relenza; Glaxo SmithKline), and oseltamivir (i.e., oseltamivir phosphate [Tamiflu; Roche]) were used at concentrations of 0.01–300 µg/mL.

Surface protein detection. For flow cytometric analysis, bacteria were grown either in TSB to the early stationary phase or in biofilm, and ~10⁶ cfu were used for the assay. Bacterial cells were washed, resuspended in PBS 1% bovine serum albumin, and incubated with rabbit anti-NanA and anti-NanB polyclonal antibodies prepared against cloned neuraminidases (data not shown). Cells were washed twice in PBS and then were resuspended in 0.5 mL of PBS and analyzed by flow cytometry (FACScan; Becton Dickinson). Data analysis was performed using CellQuest software (Becton Dickinson).

Induction of gene expression by sialic acid. Pneumococci were grown in yeast extract medium. Midexponential pneumococcal cultures were aliquoted into vials containing 0.025 % wt/vol sialic acid, and samples were obtained at 5, 10, and 20 min. After being chilled on ice, samples were directly processed for RNA extraction. Experiments were performed in triplicate for TIGR and in quintuplicate for DP1004. Real-time polymerase chain reaction (PCR) analysis was performed as described elsewhere [18, 19]. Relative gene expression was analyzed using the 2^{−ΔΔCT} method [20]. The reference gene was *gyrB*, and the reference condition was the exponential phase of growth in TSB. Statistical analysis was performed using a 2-tailed Student’s *t* test, as described elsewhere [20], for quantitative PCR.

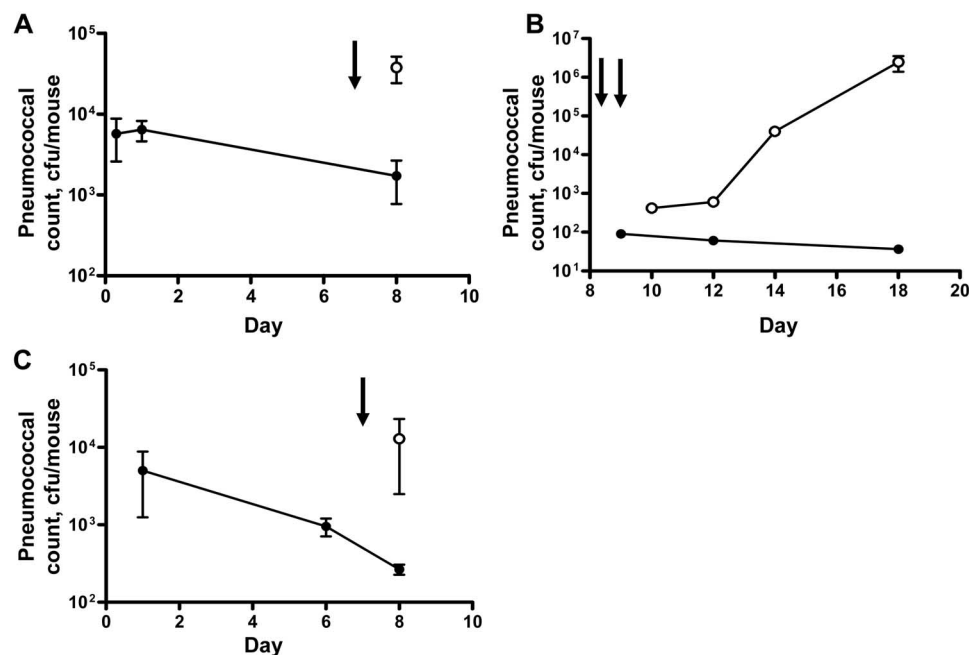


Figure 1. Influence of sialic acid on pneumococcal colonization of mice. Outbred MF1 mice were intranasally challenged with 10^4 cfu of pneumococci, and colonization was monitored by nasopharyngeal lavage (black circles). Pneumococcal colonization in mice intranasally inoculated with sialic acid (1 mg/mouse) (arrow) is denoted by white circles (with the mean count \pm SD shown). *A*, Colonization with strain TIGR4 and the effect of a single intranasal dose of sialic acid on day 7 after infection. *B*, Colonization with TIGR4 and the effect of 2 intranasal doses of sialic acid on days 8 and 9 after infection and monitoring of colonization up to 18 days after infection. *C*, Colonization with the serotype 2 strain D39 and the effect of a single intranasal dose of sialic acid on day 7 after infection. The increase in the number of colonizing bacteria was statistically significant for mice given sialic acid ($P < .05$ for all panels, by 2-tailed Student's *t* test)

Colonization model. Outbred MF1 mice were purchased from Harlan Nossan and Harlan Olac. Experiments were performed at the University of Siena, Siena, Italy (for TIGR4), and at the University of Leicester, Leicester, United Kingdom (for D39), in accordance with respective national and institutional guidelines. Five different time course experiments were performed to determine (1) the influence of sialic acid on colonization by D39, (2) the influence of a 2-dose challenge with sialic acid on colonization by TIGR4, (3) the influence of a single-dose challenge with sialic acid on colonization by TIGR4 and on its spread to the lungs and blood, (4) the influence of different sugars on colonization by TIGR4, and (5) the competition in vivo of sialic acid-dependent effects by neuraminidase inhibitors.

In all experiments, the samples used for microbiological analysis were nasopharyngeal and lung lavage fluid and blood samples. For all mice, bacterial counts are expressed as the total number of colony-forming units per sample. In experiment 1, a total of 12 mice (4 mice/data point) were intranasally infected with 2×10^4 cfu of D39 in 50 μ L. On day 7 after infection, each mouse in 1 additional group of 4 mice received 1 mg of sialic acid in 30 μ L (40 mg/kg/day) [21, 22]; samples were obtained from these mice 24 h later (figure 1C). In experiment 2, a total of 21 mice (3 mice/data point) were intranasally infected with 3×10^4 cfu of TIGR4 in 30 μ L while under anesthesia. In this experiment, 2 doses of sialic acid (1 mg/mouse) were given on days 7

and 8 after infection, and colonization in nasopharyngeal lavage fluid was assayed (figure 1B). In experiment 3, a total of 28 mice were intranasally infected with 1×10^4 cfu of TIGR4 in 30 μ L, and a single dose of sialic acid (1 mg/mouse) was given on day 7 after infection. On days 6 and 19 after infection, samples were obtained from groups of mice that had not been given sialic acid, whereas on days 8, 9, 13, 16, and 19, samples were obtained from groups of mice that had been given sialic acid. In addition to nasopharyngeal lavage fluid, lung homogenates and blood samples were also analyzed (figure 2). In experiment 4, a total of 42 mice were intranasally infected with 2×10^4 cfu of TIGR4 in 30 μ L. In this experiment, samples were obtained at 6 and 24 h after infection and on day 8 after infection (figure 1A). Additional groups of mice were challenged with different sugars on day 7 after infection, and samples for analysis were collected on day 8. The sugars (1 mg/mouse) were given intranasally. In experiment 5, a total of 44 mice were intranasally infected with 2×10^4 cfu of TIGR4 in 30 μ L. The groups of mice in experiment 5 included 10 control mice that were killed either on day 7 or day 11; ten mice that received sialic acid (1 mg/mouse) on day 7 after infection and had samples obtained on days 9 and 11; three groups of 4 mice each that were treated with either DANA, zanamivir, or oseltamivir phosphate (1 mg/mouse) on days 7 and 8; and 3 groups of 4 mice each that received a single dose of sialic acid on day 7 and subsequently were treated with either DANA, zanamivir,

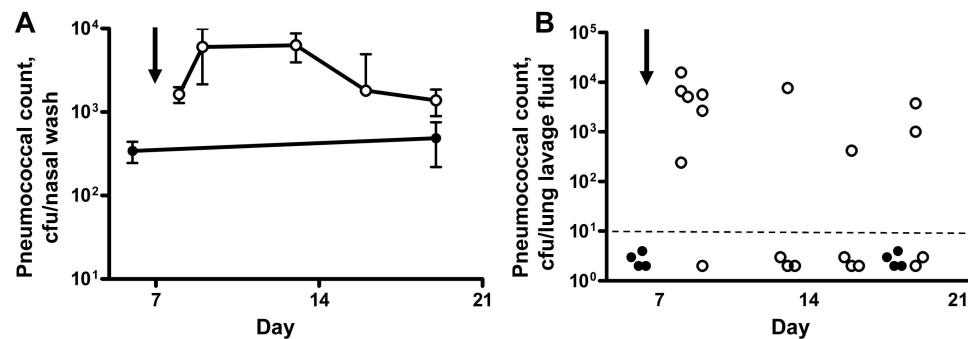


Figure 2. Influence of sialic acid on the spread of pneumococci to the lungs of colonized mice. Outbred MF1 mice were intranasally challenged with 10⁴ cfu of TIGR4. Two groups of 4 mice each were colonized and received PBS intranasally as mock treatment on day 7 after infection (*filled circles*), and 5 groups of 4 mice each received sialic acid intranasally (1 mg/mouse on day 7 after infection) (*empty circles*). Bacterial counts in nasopharyngeal lavage fluid (*A*) are expressed as the mean value \pm SD, whereas lung homogenate counts are expressed as individual counts (*B*; *dashed line* denotes the cutoff value). Blood samples plated in parallel had negative results at all time points. Nasopharyngeal counts were significantly greater in mice treated with sialic acid than in mice in the control group, at all time points (as determined using Student's *t* test) (*A*). Detection of pneumococci in lung samples was analyzed with Fisher's exact test; results were statistically significant at early time points (*B*).

vir, or oseltamivir (1 mg/mouse) on days 9 and 10. Statistical analysis was performed using a 2-tailed Student's *t* test.

RESULTS

Pneumococcal biofilm is induced by sialic acid. We recently demonstrated that pneumococci grown in static biofilms *in vitro* exhibited the same gene expression pattern as did pneumococci from the lungs of infected mice and, hence, that this *in vitro* biofilm model could be used to develop hypotheses regarding events occurring during invasive pneumococcal disease [18]. In this model, adherent pneumococcal cells can be recovered after 24 h if a competence-stimulating peptide is added to the pneumococcal culture [14].

In the present study, we used this biofilm model to investigate the possible effects of carbohydrates on the pneumococcal physiologic profile during infection. Twenty-seven sugars were individually added to standard TSB medium at 0.2% wt/vol. Under these conditions, only the aminosugar sialic acid significantly increased the number of bacteria attached to the wells after 24 h of incubation (figure 3A); this increase in bacterial levels was comparable to that noted after stimulation of the competence system by a competence-stimulating peptide (data not shown) [18]. Additional experiments were performed using medium based on soy or yeast extract only, because both constituents are devoid of animal products and, hence, are without sialic acid. In both media, effects on biofilm were also statistically significant at 6 h after infection and were specific to sialic acid among all sugars tested, including other monosaccharides that constitute the O-glycans of mucin (figure 3B). A time course experiment in yeast medium, supplemented with 0.2% wt/vol sialic acid, showed that the effect of sialic acid on the attachment of pneumococci to surfaces occurred after a few hours of incubation, corresponding to the exponential phase of growth (figure 3C).

The sialic acid concentration found to be sufficient to induce biofilm formation was between 1 and 3 μ g/mL (figure 3D), a concentration comparable to that of free sialic acid in human saliva [23–25]. Pneumococci carry 2–3 genes that code for sialidases/neuraminidase [26, 27]. An \sim 100-fold increase in the sialic acid concentration was necessary for biofilm formation by a *nanA-nanB* mutant (figure 3D). To exclude strain-specific factors, almost all experiments in this study were performed in parallel with the use of the serotype 2 strain D39 and the serotype 4 strain TIGR4. TIGR4 carries a third neuraminidase gene (*nanC*) in a separate chromosomal locus, but this gene was found to be irrelevant to biofilm formation (figure 3D) [28]. Because of the genomic organization of *nanA* and *nanB* within a single metabolic operon, the *nanAB* mutant is also deleted for cotranscribed transport systems and metabolic enzymes. In the *nanAB* mutant, recovery of the biofilm phenotype only when there is a non-physiologic excess of sialic acid (figure 3D) could result from complementation of the lost high-affinity uptake by additional low-specificity transport system(s) [29]. Alternatively, a signaling event triggered by binding of sialic acid to other receptors could explain the data.

Neuraminidase production and expression. Increased expression of neuraminidase genes in pneumococcal biofilm and in animal models of infection has been reported elsewhere [18, 30, 31]. By performing cytofluorimetric analysis of pneumococci grown in liquid culture, we detected NanA in 29% of cells and noted almost no reactivity with serum with anti-NanB serum. An increase in the detection rate to 74% for NanA and to 38% for NanB in bacteria recovered from biofilms indicates more-abundant neuraminidase production in sessile cells. The association between biofilm, neuraminidases, and sialic acid was further emphasized by the observation that sialic acid induced expression of both *nanA* and *nanB* (figure 4). The increase in

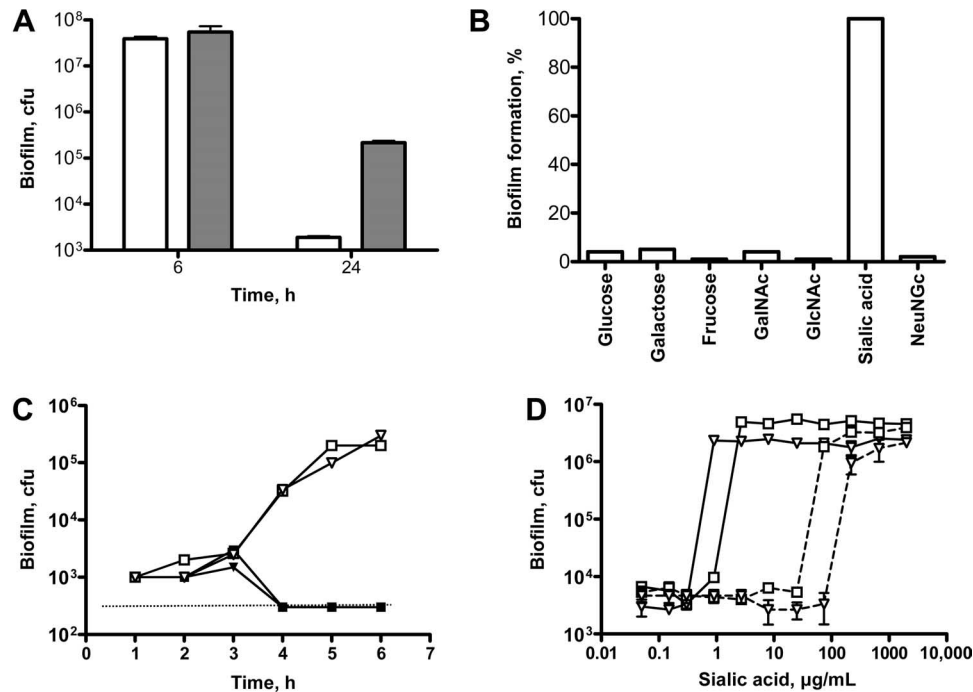


Figure 3. Biofilm formation by *Streptococcus pneumoniae* is dependent on sialic acid in the growth medium. *A*, Formation of biofilm DP1004 (a rough D39 derivative) was assayed in a static microtiter model after 6 and 24 h of incubation in unsupplemented tryptic soy broth (TSB) medium (white bars) and in TSB medium supplemented with 0.2% wt/vol of sialic acid (black bars). The effect of *N*-acetylneuraminic (sialic) acid on biofilm is significant at 24 h ($P < .001$). *B*, Biofilm formation in sialic acid-free yeast medium supplemented with 0.2% wt/vol monosaccharide after 6 h of incubation. *C*, A time course experiment performed in yeast medium and showing similar dynamics of attachment to surfaces in TIGR4 (white triangles and solid line) and DP1004 (white squares and solid line). Biofilm formation of both TIGR4 (black triangles and solid line) and DP1004 (black squares and solid line, respectively) in the absence of sialic supplementation. *D*, The effects of different sialic acid concentrations in yeast medium on biofilm formation by DP1004 (white squares and solid line), its *nanAB* mutant (white squares and dashed line), TIGR4 (open triangles and solid line), and the respective *nanAB* mutant (white triangles and dashed line) at 6 h after infection (3 replicates; the SD is shown in the graph).

expression of both *nanA* and *nanB* is steady and significant at 10 and 20 min for *nanA* in DP1004 and at 20 min for both *nanA* and *nanB* in TIGR4. This association between sialic acid and its hydrolases is typical of a substrate-inducible regulatory system [32]. In addition to the neuraminidase genes, the virulence regulator *mgrA*, which is up-regulated in biofilms [18] and is important for nasopharyngeal carriage and pneumonia [33], was induced by sialic acid (figure 4E). The absence of up-regulation of *nanC* (data not shown) is in accordance with the findings of a report published elsewhere [28], as well as with the observation that there is no significant variation in the capacity to form biofilm between the *nanAB*[−]/*nanC*⁺ TIGR4 mutant and the *nanAB*[−] D39 derivative (figure 3D). These data demonstrate that sialic acid functions as a regulatory signal for its own cleavage from oligosaccharides and has additional influence on more-complex phenotypes, such as biofilm and virulence.

Influence of sialic acid on carriage and invasion. To investigate the path of respiratory bacterial pathogens from colonization to invasion, we established a model of long-term nasopharyngeal carriage [3, 29, 34, 35]. Figure 1 shows that, in this model, pneumococci remained in the nasopharynx for at least 8–18 days without infection of the lower respiratory tract or

bacteremia developing. However, in the present study, we made the exciting observation that intranasal administration of sialic acid after 1 week of carriage resulted in a 10- to 100-fold increase in the number of colonizing pneumococci in the nasopharynx (figure 1). The increase in the number of colonizing pneumococci was even more pronounced if 2 doses of sialic acid were administered (figure 1B). Experiments performed in a different laboratory with strain D39 (figure 1C), instead of TIGR4 (figure 1A), produced an identical sialic acid-dependent increase in colonization. In both cases, this sialic acid-dependent increase in nasopharyngeal colonization was accompanied by pneumococcal spread to the lungs (figure 2). It is noteworthy that no such invasion of the lower respiratory tract by bacteria can be detected in any of the colonized mice without an intranasally administered sialic acid boost. However, pneumococcal counts in the lungs of these mice were ~100 times lower than those in models of acute pneumonia [36], and blood culture results were negative. In nearly all mice without signs of acute disease, translocation of bacteria to the lung is quite reminiscent of the situation noted in humans during the seasonal increase in carriage, when very few subjects develop invasive disease. Importantly, the booster effect on pneumococcal colonization of the naso-

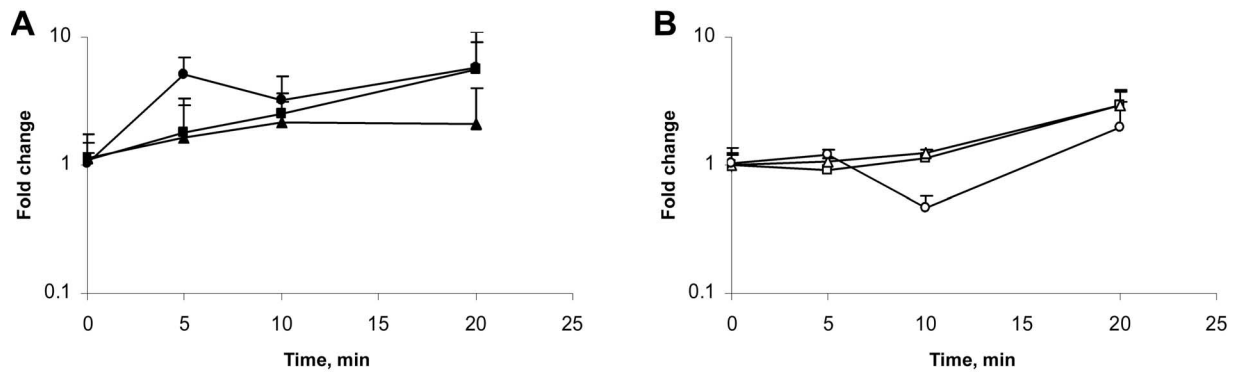


Figure 4. Changes in the expression of pneumococcal neuraminidases. Quantitative real-time polymerase chain reaction analysis was used to evaluate changes in gene expression of *nanA* (squares), *nanB* (triangles), and *mgrA* (circles) in response to sialic acid. Sialic acid (25 μ g/mL) was added to exponentially growing pneumococci in yeast medium. A significant increase in gene expression was detected in the D39 derivative DP1004 (A), for *nanA* (at 10 min, a 2.5-fold increase was noted [$P < .01$]; at 20 min, a 5.6-fold increase was noted [$P < .05$]) and for *mgrA* (at 5 min, a 5.1-fold increase was noted [$P < .01$]; at 10 min, a 3.2-fold increase was noted [$P < .05$]; and at 20 min, a 5.7-fold increase was noted [$P < .05$]), as well as in the strain TIGR4 (B), for *nanA* (at 20 min, a 2.9-fold increase was noted [$P < .05$]) and *nanB* (at 20 min, a 3-fold increase was noted [$P < .05$]). Statistics are for ≥ 3 nonparallel biological replicates and were obtained using a 2-tailed Student's *t* test [20].

pharynx and subsequent spread to the lower respiratory tract was specific to sialic acid; it did not occur after intranasal administration of other amino sugars (figure 5).

Competition of sialic acid-dependent phenotypes. To confirm the specificity of the 2 sialic acid-dependent phenotypes, we performed competition experiments with use of the transition state analogue of sialic acid DANA, its commercial neuraminidase inhibitor drug derivative zanamivir, and the cyclohexene-derived neuraminidase inhibitor oseltamivir [37]. Using all 3 compounds, it was possible to reduce by 1000-fold the capacity of pneumococci to form sialic acid-dependent biofilm in vitro (figure 6). In mice, intranasal delivery of DANA and of both anti-influenza drugs enabled a significant reduction in pneumococcal carriage (previously boosted by sialic acid)

to levels lower than those detected in uninduced carriage (figure 7). Among mice with uninduced carriage, the reduction in pneumococcal counts was detected in all mice that were treated, although this reduction was not statistically significant (figure 7). The successful competition, both in vitro and in vivo, of the sialic acid-dependent phenotypes confirms the specificity of the observed phenomenon. Nevertheless, we want to emphasize that, although pneumococcal biofilms have been demonstrated in the host in otitis media [38], and although colonization images are suggestive of biofilms [3], the present study does not claim that a biofilm is involved in pneumococcal carriage, but only that the 2 models share important analogies and that the in vitro system may serve as a suitable model for analysis of carriage phenotypes.

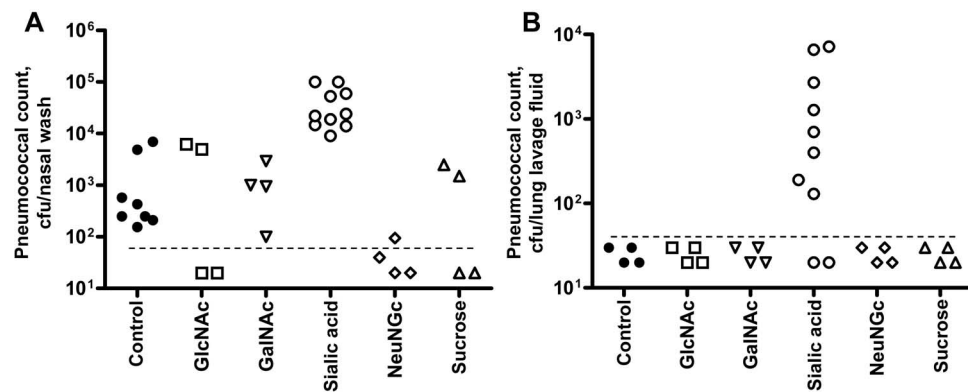


Figure 5. Modulation of pneumococcal colonization of mice by distinct sugars. Six days after intranasal infection with 10⁴ cfu of TIGR4, mice received different sugars intranasally. Pneumococcal colony-forming units in the nasopharynx (A) and lungs (B) are reported for individual mice. Dashed line, the cutoff for detection of pneumococci. The difference in nasal carriage in mice that received sialic acid is significant with respect to the control group ($P < .01$, by 2-tailed Student's *t* test), whereas there were no significant differences ($P > .05$, by 2-tailed Student's *t* test) after administration of the other sugars. The presence of bacteria in the lungs of mice treated with sialic acid is significant ($P < .05$, by Fishers exact test). In addition to sugars used in previous experiments, *N*-glycolylsialic acid, a sialic acid derivative missing in humans, was also used in the present study.

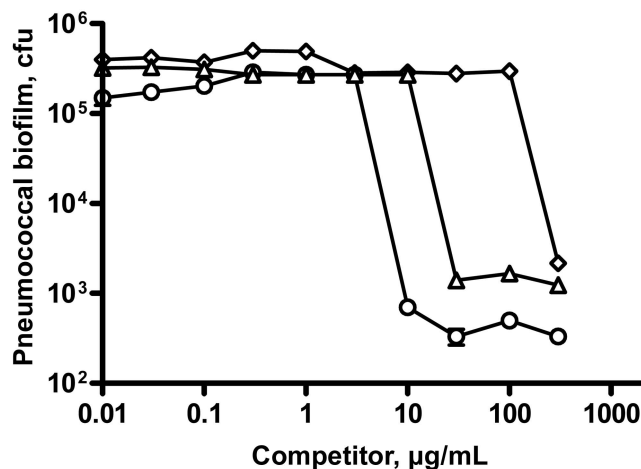


Figure 6. Competition for sialic acid-dependent biofilm formation in vitro. The effect of serial dilutions of sialic acid transition-state analogue DANA (2,3-didehydro-2-deoxy-*N*-acetylneuraminic acid) (triangles), its commercial drug derivative zanamivir (diamonds), and the cyclohexene-derived neuraminidase inhibitor oseltamivir (circles) on the capacity of pneumococci to form sialic acid-dependent biofilm was evaluated in yeast medium containing 8 µg/mL free sialic acid.

DISCUSSION

This study is, to our knowledge, the first to demonstrate (1) a causal association between sialic acid and pneumococcal biofilm formation in vitro and (2) a significant increase in nasopharyngeal colonization and spread to the lungs in vivo. These observations are highly relevant to our understanding of the mechanism of the virulence of this important human pathogen, and they could be more widely applicable to other bacterial pathogens that have both colonizing and invasive phenotypes. In humans, the epidemiologic profile of both carriage and invasive bacterial diseases shows clear seasonal variation (with peaks occurring in late winter and early spring) that parallels that of influenza virus, parainfluenza virus, and respiratory syncytial virus [2, 6, 39]. This correlation indicates that viral infection of the upper respiratory tract contributes to peaks in the incidence of invasive pneumococcal disease in humans [39]. Sialic acid has long been believed to play a central role in this interplay, although only because its removal was important. In humans, sialic acid is present as a terminal carbohydrate of the O-glycan chains of mucins and cell surface glycoproteins [23–25]. Binding to sialoglycoconjugates on host cells via hemagglutinin and release of viral particles by neuraminidases contributes to host range, tissue tropism, and the pathogenesis of the influenza virus and other viruses. Because of their central role in the pathogenesis of influenza A virus infection, neuraminidase inhibitors are among the best-known antiviral drugs [40]. Although it has not been found to be directly effective against pneumococci in vivo, treatment with neuraminidase inhibitors has been shown to reduce the severity of postinfluenza pneumococcal infection

[22, 41]. These observations have been explained by cleavage of sialic acids revealing new receptors for the bacteria and/or mitigating direct damage to host cells [42, 43]. The data presented in the current study suggest a new paradigm—namely, that sialic acid itself is a signaling molecule that results in fundamental alterations in the physiologic profile of bacteria, resulting in an enhanced capacity to adhere to surfaces and/or survive within the biofilm environment.

An important observation was that the sialic acid concentrations normally found in human saliva signal to the bacterium for increased production of neuraminidases, which is known to be associated with concomitant increased production of the co-transcribed transporters and metabolic enzymes [29]. The consequence of this signaling would be expected to be a local increase in the concentration of sialic acid and a more efficient import and metabolism. Desialylation of cell-bound and secreted glycoproteins in the nasopharynx and the lungs caused by the neuraminidases of respiratory viruses, including influenza A virus, also will increase in the concentration of free sialic acid [44, 45]. Pneumococcal neuraminidases and their role in infection have been a subject of interest for some time and have been

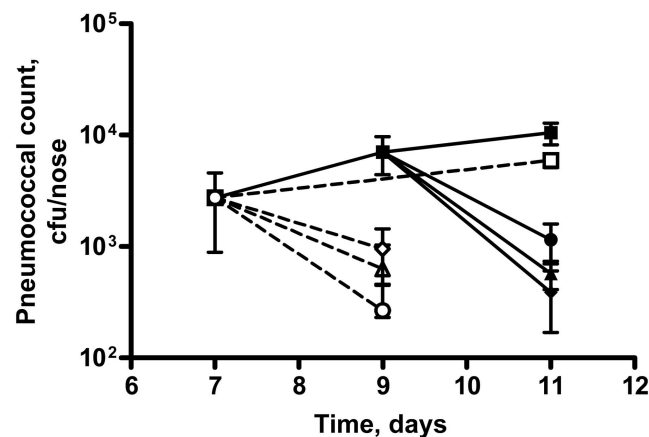


Figure 7. Competition for sialic acid-dependent colonization increases in vivo. Mice intranasally infected with pneumococci had stable nasopharyngeal colonization for 11 days (white squares and dashed line) (no significant variation was noted between days 7 and 11 after infection). After intranasal administration of sialic acid (1 mg given twice daily on days 7 and 8), pneumococcal counts increased, albeit not significantly (black squares). Intranasal inoculation (1 mg given twice daily on days 7 and 8 after infection) of either DANA (white triangle and dashed line), zanamivir (white diamond and dashed line), or oseltamivir (white circle and dashed line) reduced pneumococcal carriage ($P < .05$ for all 3 molecules, with respect to mice treated with sialic acid on day 9; reduction was nonsignificant until day 7). The reduction in colonization was even more pronounced if sialic analogues (straight lines and black triangle [for DANA], black diamond [for zanamivir], and black circle [for oseltamivir]) were administered to mice (1 mg given twice daily on days 9 and 11) that had received sialic acid on day 7 and 8 ($P < .05$, for all 3 molecules with respect to mice treated with sialic acid on day 9 and mice treated with on sialic acid day 11). Statistical analysis was performed using a 2-tailed Student's *t* test.

the focus of studies describing their desialylation of host cells and involvement in adhesion *in vitro* and *in vivo* [3, 26, 27, 42, 43, 46]. Deletion of either *nanA* or *nanB* was shown to have significant effects on nasopharyngeal colonization and pneumonia [34, 47], a finding that strongly matches the proposed role of these enzymes in the novel sialic acid–inducible carriage phenotype described in the present study.

We now propose a new model for the pathogenesis of pneumococcal pneumonia in which a local, virus-induced increase in the free sialic acid concentration in the upper respiratory tract would serve as a signal for colonizing pneumococci to increase in numbers, possibly by promoting the formation of biofilm. In this context, the up-regulation of the pneumococcal neuraminidases would provide a positive feedback loop. The higher colonization density in the nasopharynx would eventually lead to passive shedding of pneumococci to the lower respiratory tract, which, worsened by virus-induced damage to bronchoalveolar cells [48], would finally initiate development of an acute pulmonary infection.

In conclusion, the data from the present study demonstrate that free sialic acid (1) is essential for biofilm formation *in vitro* at concentrations with physiologic relevance, (2) is able to induce the expression of neuraminidase genes and a virulence regulator, and (3) leads to an increase in nasopharyngeal pneumococcal loads and instigates the spread of pneumococci to the lower respiratory tract. These data support the hypothesis that free sialic acid is a trigger that converts a harmless colonizing pneumococcus to an invasive pathogen. This finding, in addition to adding substantial knowledge to the pathogenesis of pneumococcal disease, points to novel strategies for the prevention of invasive disease, similar to those proposed for the prevention of disease due to nontypeable *H. influenzae* [15], by treating carriage or even overt disease through interventions affecting free sialic acid on human mucosae and, thus, preventing sialic acid from acting as an environmental cue to the pneumococcus.

Acknowledgments

We thank Velia Braione, Tiziana Braccini, and Anna Cuppone for help with laboratory procedures.

References

1. Ispahani P, Slack RCB, Donald FE, Weston WC, Rutter N. Twenty year surveillance of invasive pneumococcal disease in Nottingham: serogroups responsible and implications for immunisation. *Arch Dis Child* **2004**; 89:757–62.
2. Melegaro A, Edmunds WJ, Pebody R, Miller E, George R. The current burden of pneumococcal disease in England and Wales. *J Infect* **2006**; 52:37–48.
3. Kadioglu A, Weiser JN, Paton JC, Andrew PW. The role of *Streptococcus pneumoniae* virulence factors in host respiratory colonization and disease. *Nat Rev Microbiol* **2008**; 6:288–301.
4. Gray BM, Converse GM 3rd, Dillon HC Jr. Epidemiologic studies of *Streptococcus pneumoniae* in infants: acquisition, carriage, and infection during the first 24 months of life. *J Infect Dis* **1980**; 142:923–33.
5. Hogberg L, Geli P, Ringberg H, Melander E, Lipsitch M, Ekdahl K. Age- and serogroup-related differences in observed durations of nasopharyngeal carriage of penicillin-resistant pneumococci. *J Clin Microbiol* **2007**; 45:948–52.
6. Mühlemann K, Uehlinger DE, Büchi W, Gorgievski M, Aebi C. The prevalence of penicillin-non-susceptible *Streptococcus pneumoniae* among children aged <5 years correlates with the biannual epidemic activity of respiratory syncytial virus. *Clin Microbiol Infect* **2006**; 12: 873–9.
7. Iyer R, Camilli A. Sucrose metabolism contributes to *in vivo* fitness of *Streptococcus pneumoniae*. *Mol Microbiol* **2007**; 66:1–13.
8. Polissi A, Pontiggia A, Feger G, et al. Large-scale identification of virulence genes from *Streptococcus pneumoniae*. *Infect Immun* **1998**; 66:5620–9.
9. Lau GW, Haataja S, Lonetto M, et al. A functional genomic analysis of type 3 *Streptococcus pneumoniae* virulence. *Mol Microbiol* **2001**; 4:555–71.
10. Hava DL, Camilli A. Large-scale identification of serotype 4 *Streptococcus pneumoniae* virulence factors. *Mol Microbiol* **2002**; 45:1389–406.
11. Shanks RM, Donegan NP, Graber ML, et al. Heparin stimulates *Staphylococcus aureus* biofilm formation. *Infect Immun* **2005**; 73:4596–606.
12. Hooper LV, Xu J, Falk PG, Midtvedt T, Gordon JL. A molecular sensor that allows a gut commensal to control its nutrient foundation in a competitive ecosystem. *Proc Natl Acad Sci USA* **1999**; 96:9833–8.
13. Sohanpal BK, El-Labany S, Lahooti M, Plumbridge JA, Blomfield IC. Integrated regulatory responses of *fimB* to *N*-acetylneuraminic (sialic) acid and GlcNAc in *Escherichia coli* K-12. *Proc Natl Acad Sci USA* **2004**; 101:16322–7.
14. Greiner LL, Watanabe H, Phillips NJ, et al. Nontypeable *Haemophilus influenzae* strain 1919 produces a biofilm containing *N*-acetylneuraminic acid that may mimic sialylated O-linked glycans. *Infect Immun* **2004**; 72:4249–60.
15. Johnston JW, Apicella MA. Sialic acid metabolism and regulation by *Haemophilus influenzae*: potential novel antimicrobial therapies. *Curr Infect Dis Rep* **2008**; 10:83–4.
16. Kilian M, Röllä G. Initial colonization of teeth in monkeys as related to diet. *Infect Immun* **1976**; 14:1022–7.
17. Iannelli F, Pozzi G. Method for introducing specific and unmarked mutations into the chromosome of *Streptococcus pneumoniae*. *Mol Biotechnol* **2004**; 26:81–6.
18. Oggioni MR, Trappetti C, Kadioglu A, et al. Switch from planktonic to sessile life: a major event in pneumococcal pathogenesis. *Mol Microbiol* **2006**; 61:1196–210.
19. Oggioni MR, Iannelli F, Ricci S, et al. Antibacterial activity of a competence-stimulating peptide in experimental sepsis caused by *Streptococcus pneumoniae*. *Antimicrob Agents Chemother* **2004**; 48:4725–32.
20. Schmittgen TD, Livak KJ. Analyzing real-time PCR data by the comparative C(T) method. *Nat Protoc* **2008**; 3:1101–8.
21. Alymova IV, Portner A, Takimoto T, Boyd KL, Sudhakara Babu Y, McCullers JA. The novel parainfluenza virus hemagglutinin-neuraminidase inhibitor BCX 2798 prevents lethal synergism between a paramyxovirus and *Streptococcus pneumoniae*. *Antimicrob Agents Chemother* **2005**; 49:398–405.
22. McCullers JA. Effect of antiviral treatment on the outcome of secondary bacterial pneumonia after influenza. *J Infect Dis* **2004**; 190:519–26.
23. Tram TH, Brand Miller JC, McNeil Y, McVeagh P. Sialic acid content of infant saliva: comparison of breast fed with formula fed infants. *Arch Dis Child* **1997**; 77:315–8.
24. Lee JY, Chung JW, Kim YK, Chung SC, Kho HS. Comparison of the composition of oral mucosal residual saliva with whole saliva. *Oral Dis* **2007**; 13:550–4.
25. Siqueira WL, Siquiera MF, Mustacchi Z, de Oliveira E, Nicolau J. Salivary parameters in infants aged 12 to 60 months with Down syndrome. *Spec Care Dentist* **2007**; 27:203–6.
26. Berry AM, Paton JC, Glare EM, Hansman D, Catcheside DE. Cloning and expression of the pneumococcal neuraminidase gene in *Escherichia coli*. *Gene* **1988**; 71:299–305.

27. Camara M, Mitchell TJ, Andrew PW, Boulnois GJ. *Streptococcus pneumoniae* produces at least two distinct enzymes with neuraminidase activity: cloning and expression of a second neuraminidase gene in *Escherichia coli*. *Infect Immun* **1991**; 59:2856–8.
28. Burnaugh AM, Frantz LJ, King SJ. Growth of *Streptococcus pneumoniae* on human glycoconjugates is dependent upon the sequential activity of bacterial exoglycosidases. *J Bacteriol* **2008**; 190:221–30.
29. King SJ, Hippe KR, Gould JM, et al. Phase variable desialylation of host proteins that bind to *Streptococcus pneumoniae* *in vivo* and protect the airway. *Mol Microbiol* **2004**; 54:159–71.
30. Orihuela CJ, Radin JN, Sublett JE, Gao G, Kaushal D, Tuomanen EI. Microarray analysis of pneumococcal gene expression during invasive disease. *Infect Immun* **2004**; 72:5582–96.
31. LeMessurier KS, Ogunniyi DA, Paton JC. Differential expression of key pneumococcal virulence genes *in vivo*. *Microbiology* **2006**; 152:305–11.
- AQ: 48 32. Deutscher J, Francke C, Pot B, Postma PW. How phosphotransferase system-related protein phosphorylation regulates carbohydrate metabolism in bacteria. *Microbiol Mol Biol Rev* **2006**; 70:939–1031.
33. Hemsley CJ, Joyce E, Hava DL, Kawale A, Camilli A. MgrA, an orthologue of Mga, acts as a transcriptional repressor of the genes within the *rlrA* pathogenicity islet in *Streptococcus pneumoniae*. *J Bacteriol* **2003**; 185:6640–7.
34. Manco S, Hernon F, Yesilkaya H, Paton JC, Andrew PW, Kadioglu A. Pneumococcal neuraminidases A and B both have essential roles during infection of the respiratory tract and sepsis. *Infect Immun* **2006**; 74:4014–20.
35. Kadioglu A, Taylor S, Iannelli F, Pozzi G, Mitchell TJ, Andrew PW. Upper and lower respiratory tract infection by *Streptococcus pneumoniae* is affected by deficiency of pneumolysin and by differences in serotype. *Infect Immun* **2002**; 70:2886–90.
36. Kadioglu A, Gingles NA, Grattan K, Kerr A, Mitchell TJ, Andrew PW. Host cellular immune response to pneumococcal lung infection in mice. *Infect Immun* **2000**; 68:492–501.
37. von Itzstein M. The war against influenza: discovery and development of sialidase inhibitors. *Nat Rev Drug Discov* **2007**; 6:967–74.
38. Hall-Stoodley L, Hu FZ, Gieseke A, et al. Direct detection of bacterial biofilms on the middle-ear mucosa of children with chronic otitis media. *JAMA* **2006**; 296:202–11.
39. Watson M, Gilmour R, Menzies R, Ferson M, McIntyre P. The association of respiratory viruses, temperature and other climatic parameters with the incidence of invasive pneumococcal disease in Sydney, Australia. *Clin Infect Dis* **2006**; 42:211–5.
40. Streicher H. Inhibition of microbial sialidases—What happened beyond the influenza virus? *Curr Med Chem- Anti-infective Agents* **2004**; 3:149–61.
41. McCullers JA, Rehg JE. Lethal synergism between influenza virus and *Streptococcus pneumoniae*: characterization of a mouse model and the role of platelet-activating factor receptor. *J Infect Dis* **2002**; 186:341–50.
42. Tong HH, Blue LE, James MA, De Maria TF. Evaluation of the virulence of a *Streptococcus pneumoniae* neuraminidase-deficient mutant in nasopharyngeal colonization and development of otitis media in the chin-chilla model. *Infect Immun* **2000**; 68:921–4.
43. Grewal PK, Uchiyama S, Ditto D, et al. The Ashwell receptor mitigates the lethal coagulopathy of sepsis. *Nat Med* **2008**; 14:648–55.
44. Tess BR, Kempf JE. Decrease of bound sialic acid and inhibitor in chorioallantoic membranes infected with influenza virus. *J Bacteriol* **1963**; 86:239–45.
45. McCullers JA, Bartmess KC. Role of neuraminidase in lethal synergism between influenza virus and *Streptococcus pneumoniae*. *J Infect Dis* **2003**; 187:1000–9.
46. King SJ, Hippe KR, Weiser JN. Deglycosylation of human glycoconjugates by the sequential activities of exoglycosidases expressed by *Streptococcus pneumoniae*. *Mol Microbiol* **2006**; 59:961–74.
47. Orihuela CJ, Gao G, Francis KP, Yu J, Tuomanen EI. Tissue-specific contributions of pneumococcal virulence factors to pathogenesis. *J Infect Dis* **2004**; 190:1661–9.
48. Matrosovich MN, Matrosovich TY, Gray T, Roberts NA, Klenk H. Human and avian influenza viruses target different cell types in cultures of human airway epithelium. *Proc Natl Acad Sci USA* **2004**; 101:4620–4.

QUERIES TO THE AUTHOR

1. Au: Your article has been edited for grammar, clarity, consistency, and adherence to journal style. To expedite publication, we no longer ask authors for approval of routine grammatical and style changes. Please read the article to make sure your meaning has been retained; any layout problems (including table and figure placement) will be addressed after we have incorporated corrections. Note that we may be unable to make changes that conflict with journal style, obscure meaning, or create grammatical or other problems. If you are writing corrections by hand, please print clearly, and be aware that corrections written too close to the edges of the paper may not transmit by fax. Finally, please note that a delayed, incomplete, or illegible response may delay publication of your article. Thank you!

2. Au: The following queries are for sentence 1 (“The correlation. . .”): (a) Should “formation” be inserted after “biofilm”? (b) Okay that I inserted “of the host” after “invasion”?

3. Au: Please carefully review and verify edits to the sentence “In a murine model. . .”

4. Au: (a) Okay that I inserted “Formation” after “Biofilm” in the title of your article? (b) Please expand “UOC” in “UOC Bacteriologia” in author affiliation 2. (c) Please verify that the author affiliations are correct as shown and that they are correctly matched with the appropriate author names.

5. Au: In the sentence “Competition both. . .” please (a) verify my expansion of “DANA” and (b) indicate whether “oseltamivir” should be changed to “oseltamivir phosphate” at first mention in the abstract and text. (c) Please verify all other edits to the sentence.

6. Au: In the sentence “The association. . .” okay that I inserted “pneumococcal” before “colonization”?

7. Au: Please review edits to the sentence “To address. . .” and either verify that they are acceptable as shown or provide clarification of your intended meaning.

8. Au: In the sentence “The burden of. . .” okay that I changed “high” to “great”?

9. Au: In the sentence “Successive. . .” okay that I inserted “incidence” after “peak”?

10. Au: (a) Please note, that because of edits made to the sentence “In the pneumococcus. . .” I renumbered original references 7–10 (and their citations) so they would appear in sequential order. Please verify that no errors were introduced during

renumbering of your references/citations. (b) Please verify all edits to the sentence “In the pneumococcus. . .”

11. Au: Okay that I inserted “To our knowledge” before “at this time” at the beginning of the sentence?

12. Au: (a) Please carefully review and verify all edits to the sentence “Sialic acid-free. . .” (b) Please note that, as per JID style, I changed all text mentions of “M” (to denote molarity) to “mol/L.” Please carefully review all such changes, and verify that they are acceptable as shown.

13. Au: Please carefully review and verify edits to the information in the 2 sentences “The primers used. . .” and “The size of the. . .”

14. Au: In the sentence “To detach. . .” okay that I changed “Transonic 460, 35 kHz” to “sonicator water bath (a Transonic 460/H ultrasonic bath [Elma])”?

15. Au: (a) In the sentence “The sugars used. . . here and in the sentence “Solutions of carbohydrates. . .” below, okay that I inserted “(sialic)” between “N-acetylneuraminic” and “acid”? (b) In the sentence “Solutions of carbohydrates. . .” okay that I placed each pH in brackets?

16. Au: Please specify the version number of the CellQuest software used.

17. Au: In the sentence “Statistical analysis. . .” and elsewhere in your text, okay that I changed mentions of “*t* test” to “Student’s *t* test”?

18. Au: Okay that, here in the sentence “In all experiments. . .” and elsewhere in your text, I specified “nasopharyngeal lavage fluid” and “lung lavage fluid”?

19. Au: Your tables and/or figures have been edited in accordance with journal style. Please check carefully to ensure that all edits are acceptable and that the integrity of the data has been maintained. Please also confirm, where applicable, that units of measure are correct, that table column heads accurately reflect the information in the columns below, and that all material contained in figure legends and table footnotes (including definitions of symbols and abbreviations) is correct.

20. Au: Because the figures in your original article your cited out of sequential order, I renumbered original figures 1–4 and their text citations. Please carefully review all renumbered figures and figure citations, and verify that no errors were introduced during renumbering.

21. Au: Please verify DOSE information here and elsewhere in this article; make sure that both the values and the units have been typeset correctly.
22. Au: In the sentence “In this experiment. . .,” okay that I changed “at 8 days” to “on day 8 after infection”?
23. Au: In the first subhead in the “Results” section, should “formation” be inserted after “biofilm”?
24. Au: Okay that I started a new paragraph with the sentence “In the present study. . .”?
25. Au: (a) I deleted “(N-acetylneuraminic acid; NeuNAc)” from the sentence “Under these conditions).” Okay? (b) Please carefully review and verify all other edits to the sentence.
26. Au: (a) For figures 3 and 5, please expand “GalNAc,” “GlcNAc,” and “NeuNGc.” (b) Where exactly in figure 3D does the SD appear? (The last line of the legend says that the SD is shown in the graph.)
27. Au: In the sentence “In both media. . .,” okay that I inserted “statistically” before “significant”?
28. Au: Please verify edits to the sentence “In the *nanAB*. . . mutant. . .” or clarify your intended meaning.
29. Au: (a) Should “antibodies be inserted after mentions of “NanA” and “NanB” in the sentences “By performing. . .” and “An increase. . .”? (b) Also, in the same sentence, please clarify what is meant by “no reactivity with serum with anti-NanB serum.” (c) In the sentence “An increase. . .,” okay that I changed “positivity” to “detection rate”?
30. Au: In the figure 4 legend, please verify all edits to the sentence “A significant increase. . .”
31. Au: Please verify edits to the sentence “The increase in expression. . .”
32. Au: (a) In the sentence “The absence. . .,” I changed “previous reports” to “a report published elsewhere,” because only 1 report was cited (reference 28). Okay? (b) Please verify edits to the sentence “These data. . .”
33. Au: Please verify edits to the sentence “To investigate. . .”
34. Au: (a) Please verify edits to the sentence “Figure 1 shows. . .” or provide clarification of your intended meaning. (b) In the sentence “However, in the present study. . .,” okay that I changed “nasal” to “intranasal,” to match other text mentions?
35. Au: In the sentence “However, pneumococcal counts. . .,” okay that I changed “lung counts” to “pneumococcal counts in the lungs”?
36. Au: The following queries are for the sentence “To confirm. . .”: (a) Okay that I changed “analog” to “analog ue”? (b) Okay that I deleted “Neu5Ac2en”?
37. Au: (a) In the sentence “In mice. . .,” okay that I changed “nasal” to “intranasal,” to match other text mentions? (b) Please carefully review and verify all other edits to the sentence.
38. Au: (a) Please carefully review all symbol definitions in the figure 7 legend, and verify that they are correct as shown. (b) In the y -axis title of figure 7, should “nose” be changed to “nasal wash”?
39. Au: In the sentence “Nevertheless. . .,” okay that I changed “the present data,” to “the present study”? Please verify this and all other edits to the sentence.
40. Au: (a) Okay that I inserted “to our knowledge” in the sentence “This study. . .”? (b) If possible, please provide a paragraph break(s) for this somewhat lengthy paragraph.
41. Au: In the sentence “Binding. . .,” please clarify what is meant by “host range.” Are you perhaps referring to the range of hosts in which pneumococcal invasion occurs?
42. Au: (a) Okay that I specified “The consequence of this signaling”? (b) Please verify edits to the sentence “Desialylation. . .”
43. Au: The following queries are for the sentence “Pneumococcal. . .”: (a) Please verify edits to the sentence, or provide clarification of your intended meaning. (b) In the string of reference citations appearing in the original typescript, there were 2 mentions of “[43].” Okay that I deleted one of these mentions, or was another reference supposed to be cited here?
44. Au: Please verify edits to the sentence “This finding. . .,” or provide clarification of your intended meaning.
45. Au: I changed the year of publication of reference 3 from “2005” to “2006,” to match the Medline entry for this article (PubMed ID no. 16368459). Okay?
46. Au: In reference 10, I changed the last number in the page range from “405” to “406,” to match the Medline entry for this article (PubMed ID no. 12207705). Okay?

47. Au: In reference 25, okay that I changed the page range from “203–206” to “202–205,” to match the Medline entry for this article (PubMed ID no. 17990480).

48. Au: Okay that I deleted “Pot B” from the author list for reference 32, because he/she does not appear in the Medline entry for this article (PubMed ID no. 17158705)?

49. Au: **(a)** Reference 40 could not be matched in a search of the Medline database. Please verify that it is correct as shown, and, if possible, provide its PMID no. (not UI) in that database. **(b)** Please expand the journal title.

8 Bibliography

- AL-ALOUL, M., CRAWLEY, J., WINSTANLEY, C., HART, C.A., LEDSON, M.J. and WALSHAW, M.J., 2004. Increased morbidity associated with chronic infection by an epidemic *Pseudomonas aeruginosa* strain in CF patients. *Thorax*, **59**(4), pp. 334-336.
- ALLEWELT, M., COLEMAN, F.T., GROUT, M., PRIEBE, G.P. and PIER, G.B., 2000. Acquisition of expression of the *Pseudomonas aeruginosa* ExoU cytotoxin leads to increased bacterial virulence in a murine model of acute pneumonia and systemic spread. *Infection and immunity*, **68**(7), pp. 3998-4004.
- ARORA, S.K., BANGERA, M., LORY, S. and RAMPHAL, R., 2001. A genomic island in *Pseudomonas aeruginosa* carries the determinants of flagellin glycosylation. *Proceedings of the National Academy of Sciences of the United States of America*, **98**(16), pp. 9342-9347.
- BARRANGOU, R., FREMAUX, C., DEVEAU, H., RICHARDS, M., BOYAVAL, P., MOINEAU, S., ROMERO, D.A. and HORVATH, P., 2007. CRISPR provides acquired resistance against viruses in prokaryotes. *Science (New York, N.Y.)*, **315**(5819), pp. 1709-1712.
- BECKER, D.M. and LUNDBLAD, V., 2001. Introduction of DNA into yeast cells. *Current protocols in molecular biology / edited by Frederick M. Ausubel ...[et al.]*, **Chapter 13**, pp. Unit13.7.
- BIKANDI, J., SAN MILLAN, R., REMENTERIA, A. and GARAIZAR, J., 2004. In silico analysis of complete bacterial genomes: PCR, AFLP-PCR and endonuclease restriction. *Bioinformatics*, **20**(5), pp. 798-9.
- BISHOP, A.L., BAKER, S., JENKS, S., FOOKES, M., GAORA, P.O., PICKARD, D., ANJUM, M., FARRAR, J., HIEN, T.T., IVENS, A. and DOUGAN, G., 2005. Analysis of the hypervariable region of the *Salmonella enterica* genome associated with tRNA(LeuX). *Journal of Bacteriology*, **187**(7), pp. 2469-2482.
- BORTONI RODRIGUEZ, M.E., 2006. Studies on pneumolysin and on pneumococcal gene expression under aerobiosis. PhD Leicester: University of Leicester.
- BRACHMANN, C.B., DAVIES, A., COST, G.J., CAPUTO, E., LI, J., HIETER, P. and BOEKE, J.D., 1998. Designer deletion strains derived from *Saccharomyces cerevisiae* S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. *Yeast*, **14**(2), pp. 115-32.
- BRZUSZKIEWICZ, E., BRUGGEMANN, H., LIESEGANG, H., EMMERTH, M., OLSCHLAGER, T., NAGY, G., ALBERMANN, K., WAGNER, C., BUCHRIESER, C., EMODY, L., GOTTSCHALK, G., HACKER, J. and DOBRINDT, U., 2006. How to become a uropathogen: comparative genomic analysis of extraintestinal pathogenic *Escherichia coli* strains. *Proceedings of the*

National Academy of Sciences of the United States of America, **103**(34), pp. 12879-12884.

BURGERS, P.M. and PERCIVAL, K.J., 1987. Transformation of yeast spheroplasts without cell fusion. *Analytical Biochemistry*, **163**(2), pp. 391-397.

CALI, S., SPOLDI, E., PIAZZOLLA, D., DODD, I.B., FORTI, F., DEHO, G. and GHISOTTI, D., 2004. Bacteriophage P4 Vis protein is needed for prophage excision. *Virology*, **322**(1), pp. 82-92.

CAMPBELL, A., 2003. Prophage insertion sites. *Research in microbiology*, **154**(4), pp. 277-282.

CASH, H.A., WOODS, D.E., MCCULLOUGH, B., JOHANSON, W.G., JR and BASS, J.A., 1979. A rat model of chronic respiratory infection with *Pseudomonas aeruginosa*. *The American Review of Respiratory Disease*, **119**(3), pp. 453-459.

CHASTRE, J. and FAGON, J.Y., 2002. Ventilator-associated pneumonia. *American journal of respiratory and critical care medicine*, **165**(7), pp. 867-903.

CHOI, J.Y., SIFRI, C.D., GOUMNEROV, B.C., RAHME, L.G., AUSUBEL, F.M. and CALDERWOOD, S.B., 2002. Identification of virulence genes in a pathogenic strain of *Pseudomonas aeruginosa* by representational difference analysis. *Journal of Bacteriology*, **184**(4), pp. 952-961.

CHOI, K.H. and SCHWEIZER, H.P., 2005. An improved method for rapid generation of unmarked *Pseudomonas aeruginosa* deletion mutants. *BMC microbiology*, **5**(1), pp. 30.

CHRISTENSEN, L.D., MOSER, C., JENSEN, P.O., RASMUSSEN, T.B., CHRISTOPHERSEN, L., KJELLEBERG, S., KUMAR, N., HOIBY, N., GIVSKOV, M. and BJARNSHOLT, T., 2007. Impact of *Pseudomonas aeruginosa* quorum sensing on biofilm persistence in an in vivo intraperitoneal foreign-body infection model. *Microbiology (Reading, England)*, **153**(Pt 7), pp. 2312-2320.

COLEMAN, F.T., MUESCHENBORN, S., MELULENI, G., RAY, C., CAREY, V.J., VARGAS, S.O., CANNON, C.L., AUSUBEL, F.M. and PIER, G.B., 2003. Hypersusceptibility of cystic fibrosis mice to chronic *Pseudomonas aeruginosa* oropharyngeal colonization and lung infection. *Proceedings of the National Academy of Sciences of the United States of America*, **100**(4), pp. 1949-1954.

COMOLLI, J.C., HAUSER, A.R., WAITE, L., WHITCHURCH, C.B., MATTICK, J.S. and ENGEL, J.N., 1999. *Pseudomonas aeruginosa* gene products PilT and PilU are required for cytotoxicity in vitro and virulence in a mouse model of acute pneumonia. *Infection and immunity*, **67**(7), pp. 3625-3630.

COWELL, B.A., CHEN, D.Y., FRANK, D.W., VALLIS, A.J. and FLEISZIG, S.M., 2000. ExoT of cytotoxic *Pseudomonas aeruginosa* prevents uptake by corneal epithelial cells. *Infection and immunity*, **68**(1), pp. 403-406.

D'ARGENIO, D.A., WU, M., HOFFMAN, L.R., KULASEKARA, H.D., DEZIEL, E., SMITH, E.E., NGUYEN, H., ERNST, R.K., LARSON FREEMAN, T.J., SPENCER, D.H., BRITTNACHER, M., HAYDEN, H.S., SELGRADE, S., KLAUSEN, M., GOODLETT, D.R., BURNS, J.L., RAMSEY, B.W. and MILLER, S.I., 2007. Growth phenotypes of *Pseudomonas aeruginosa* lasR mutants adapted to the airways of cystic fibrosis patients. *Molecular microbiology*, **64**(2), pp. 512-533.

DRENKARD, E. and AUSUBEL, F.M., 2002. *Pseudomonas* biofilm formation and antibiotic resistance are linked to phenotypic variation. *Nature*, **416**(6882), pp. 740-743.

DUAN, K. and SURETTE, M.G., 2007. Environmental regulation of *Pseudomonas aeruginosa* PAO1 Las and Rhl quorum-sensing systems. *Journal of Bacteriology*, **189**(13), pp. 4827-4836.

ERMOLAEVA, M.D., 2001. Synonymous codon usage in bacteria. *Current Issues in Molecular Biology*, **3**(4), pp. 91-97.

FINCK-BARBANCON, V., GORANSON, J., ZHU, L., SAWA, T., WIENER-KRONISH, J.P., FLEISZIG, S.M., WU, C., MENDE-MUELLER, L. and FRANK, D.W., 1997. ExoU expression by *Pseudomonas aeruginosa* correlates with acute cytotoxicity and epithelial injury. *Molecular microbiology*, **25**(3), pp. 547-557.

FINCK-BARBANCON, V., GORANSON, J., ZHU, L., SAWA, T., WIENER-KRONISH, J.P., FLEISZIG, S.M., WU, C., MENDE-MUELLER, L. and FRANK, D.W., 1997. ExoU expression by *Pseudomonas aeruginosa* correlates with acute cytotoxicity and epithelial injury. *Molecular microbiology*, **25**(3), pp. 547-557.

FOTHERGILL, J.L., PANAGEA, S., HART, C.A., WALSHAW, M.J., PITT, T.L. and WINSTANLEY, C., 2007. Widespread pyocyanin over-production among isolates of a cystic fibrosis epidemic strain. *BMC microbiology*, **7**, pp. 45.

FUHRMANN, M., HAUSHERR, A., FERBITZ, L., SCHODL, T., HEITZER, M. and HEGEMANN, P., 2004. Monitoring dynamic expression of nuclear genes in *Chlamydomonas reinhardtii* by using a synthetic luciferase reporter gene. *Plant Molecular Biology*, **55**(6), pp. 869-881.

FURUKAWA, S., KUCHMA, S.L. and O'TOOLE, G.A., 2006. Keeping their options open: acute versus persistent infections. *Journal of Bacteriology*, **188**(4), pp. 1211-1217.

GARAU, J. and GOMEZ, L., 2003. *Pseudomonas aeruginosa* pneumonia. *Current opinion in infectious diseases*, **16**(2), pp. 135-143.

GIETZ, R.D. and WOODS, R.A., 2001. Genetic transformation of yeast. *Biotechniques*, **30**(4), pp. 816-20, 822-6, 828 passim.

GIETZ, R.D. and WOODS, R.A., 2002. Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. *Methods in enzymology*, **350**, pp. 87-96.

- HACKER, J. and KAPER, J.B., 2000. Pathogenicity islands and the evolution of microbes. *Annual Review of Microbiology*, **54**, pp. 641-679.
- HAFT, D.H., SELENGUT, J., MONGODIN, E.F. and NELSON, K.E., 2005. A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. *PLoS computational biology*, **1**(6), pp. e60.
- HAUSER, A.R., COBB, E., BODI, M., MARISCAL, D., VALLES, J., ENGEL, J.N. and RELLO, J., 2002. Type III protein secretion is associated with poor clinical outcomes in patients with ventilator-associated pneumonia caused by *Pseudomonas aeruginosa*. *Critical care medicine*, **30**(3), pp. 521-528.
- HAUSER, A.R., KANG, P.J. and ENGEL, J.N., 1998. PepA, a secreted protein of *Pseudomonas aeruginosa*, is necessary for cytotoxicity and virulence. *Molecular microbiology*, **27**(4), pp. 807-818.
- HAUSSLER, S., 2004. Biofilm formation by the small colony variant phenotype of *Pseudomonas aeruginosa*. *Environmental microbiology*, **6**(6), pp. 546-551.
- HE, J., BALDINI, R.L., DEZIEL, E., SAUCIER, M., ZHANG, Q., LIBERATI, N.T., LEE, D., URBACH, J., GOODMAN, H.M. and RAHME, L.G., 2004. The broad host range pathogen *Pseudomonas aeruginosa* strain PA14 carries two pathogenicity islands harboring plant and animal virulence genes. *Proceedings of the National Academy of Sciences of the United States of America*, **101**(8), pp. 2530-5.
- HUANG, Y.H., FERRIERES, L. and CLARKE, D.J., 2006. The role of the Rcs phosphorelay in Enterobacteriaceae. *Research in microbiology*, **157**(3), pp. 206-212.
- IMAMURA, Y., YANAGIHARA, K., TOMONO, K., OHNO, H., HIGASHIYAMA, Y., MIYAZAKI, Y., HIRAKATA, Y., MIZUTA, Y., KADOTA, J., TSUKAMOTO, K. and KOHNO, S., 2005. Role of *Pseudomonas aeruginosa* quorum-sensing systems in a mouse model of chronic respiratory infection. *Journal of medical microbiology*, **54**(Pt 6), pp. 515-518.
- JANDER, G., RAHME, L.G. and AUSUBEL, F.M., 2000. Positive correlation between virulence of *Pseudomonas aeruginosa* mutants in mice and insects. *Journal of Bacteriology*, **182**(13), pp. 3843-3845.
- JANSEN, R., EMBDEN, J.D., GAASTRA, W. and SCHOULS, L.M., 2002. Identification of genes that are associated with DNA repeats in prokaryotes. *Molecular microbiology*, **43**(6), pp. 1565-1575.
- KLOCKGETHER, J., REVA, O., LARBIG, K. and TUMMLER, B., 2004. Sequence analysis of the mobile genome island pKLC102 of *Pseudomonas aeruginosa* C. *Journal of Bacteriology*, **186**(2), pp. 518-534.
- KLOCKGETHER, J., WURDEMAN, D., REVA, O., WIEHLMANN, L. and TUMMLER, B., 2007. Diversity of the abundant pKLC102/PAGI-2 family of

genomic islands in *Pseudomonas aeruginosa*. *Journal of Bacteriology*, **189**(6), pp. 2443-2459.

KOUPRINA, N. and LARIONOV, V., 2006. Selective isolation of mammalian genes by TAR cloning. *Current protocols in human genetics / editorial board, Jonathan L.Haines ...[et al.]*, **Chapter 5**, pp. Unit 5.17.

KUKAVICA-IBRULJ, I., BRAGONZI, A., PARONI, M., WINSTANLEY, C., SANSCHAGRIN, F., O'TOOLE, G.A. and LEVESQUE, R.C., 2008. In vivo growth of *Pseudomonas aeruginosa* strains PAO1 and PA14 and the hypervirulent strain LESB58 in a rat model of chronic lung infection. *Journal of Bacteriology*, **190**(8), pp. 2804-2813.

KULASEKARA, B.R., KULASEKARA, H.D., WOLFGANG, M.C., STEVENS, L., FRANK, D.W. and LORY, S., 2006. Acquisition and evolution of the *exoU* locus in *Pseudomonas aeruginosa*. *Journal of Bacteriology*, **188**(11), pp. 4037-4050.

LARBIG, K.D., CHRISTMANN, A., JOHANN, A., KLOCKGETHER, J., HARTSCH, T., MERKL, R., WIEHLMANN, L., FRITZ, H.J. and TUMMLER, B., 2002. Gene islands integrated into tRNA(Gly) genes confer genome diversity on a *Pseudomonas aeruginosa* clone. *Journal of Bacteriology*, **184**(23), pp. 6665-6680.

LARIONOV, V., KOUPRINA, N., GRAVES, J. and RESNICK, M.A., 1996. Highly selective isolation of human DNAs from rodent-human hybrid cells as circular yeast artificial chromosomes by transformation-associated recombination cloning. *Proceedings of the National Academy of Sciences of the United States of America*, **93**(24), pp. 13925-30.

LASKOWSKI, M.A. and KAZMIERCZAK, B.I., 2006. Mutational analysis of RetS, an unusual sensor kinase-response regulator hybrid required for *Pseudomonas aeruginosa* virulence. *Infection and immunity*, **74**(8), pp. 4462-4473.

LAU, G.W., RAN, H., KONG, F., HASSETT, D.J. and MAVRODI, D., 2004. *Pseudomonas aeruginosa* pyocyanin is critical for lung infection in mice. *Infection and immunity*, **72**(7), pp. 4275-4278.

LEE, D.G., URBACH, J.M., WU, G., LIBERATI, N.T., FEINBAUM, R.L., MIYATA, S., DIGGINS, L.T., HE, J., SAUCIER, M., DEZIEL, E., FRIEDMAN, L., LI, L., GRILLS, G., MONTGOMERY, K., KUCHERLAPATI, R., RAHME, L.G. and AUSUBEL, F.M., 2006. Genomic analysis reveals that *Pseudomonas aeruginosa* virulence is combinatorial. *Genome biology*, **7**(10), pp. R90.

LEE, V.T., SMITH, R.S., TUMMLER, B. and LORY, S., 2005. Activities of *Pseudomonas aeruginosa* effectors secreted by the Type III secretion system in vitro and during infection. *Infection and immunity*, **73**(3), pp. 1695-1705.

LEEM, S.H., NOSKOV, V.N., PARK, J.E., KIM, S.I., LARIONOV, V. and KOUPRINA, N., 2003. Optimum conditions for selective isolation of genes from complex genomes by transformation-associated recombination cloning. *Nucleic acids research*, **31**(6), pp. e29.

- LESTINI, R. and MICHEL, B., 2008. UvrD and UvrD252 counteract RecQ, RecJ and RecFOR in the rep mutant. *Journal of Bacteriology*, **190**(17), pp. 5995-6001.
- LIANG, X., PHAM, X.Q., OLSON, M.V. and LORY, S., 2001. Identification of a genomic island present in the majority of pathogenic isolates of *Pseudomonas aeruginosa*. *Journal of Bacteriology*, **183**(3), pp. 843-53.
- LIN, H.H., HUANG, S.P., TENG, H.C., JI, D.D., CHEN, Y.S. and CHEN, Y.L., 2006. Presence of the *exoU* gene of *Pseudomonas aeruginosa* is correlated with cytotoxicity in MDCK cells but not with colonization in BALB/c mice. *Journal of clinical microbiology*, **44**(12), pp. 4596-4597.
- LOBOCKA, M.B., ROSE, D.J., PLUNKETT, G., 3RD, RUSIN, M., SAMOJEDNY, A., LEHNHERR, H., YARMOLINSKY, M.B. and BLATTNER, F.R., 2004. Genome of bacteriophage P1. *Journal of Bacteriology*, **186**(21), pp. 7032-7068.
- LU, D. and KECK, J.L., 2008. Structural basis of *Escherichia coli* single-stranded DNA-binding protein stimulation of exonuclease I. *Proceedings of the National Academy of Sciences of the United States of America*, **105**(27), pp. 9169-9174.
- MARCIL, R. and HIGGINS, D.R., 1992. Direct transfer of plasmid DNA from yeast to *E.coli* by electroporation. *Nucleic acids research*, **20**(4), pp. 917.
- MCCALLUM, S.J., CORKILL, J., GALLAGHER, M., LEDSON, M.J., HART, C.A. and WALSHAW, M.J., 2001. Superinfection with a transmissible strain of *Pseudomonas aeruginosa* in adults with cystic fibrosis chronically colonised by *P. aeruginosa*. *Lancet*, **358**(9281), pp. 558-560.
- MCCALLUM, S.J., GALLAGHER, M.J., CORKILL, J.E., HART, C.A., LEDSON, M.J. and WALSHAW, M.J., 2002. Spread of an epidemic *Pseudomonas aeruginosa* strain from a patient with cystic fibrosis (CF) to non-CF relatives. *Thorax*, **57**(6), pp. 559-560.
- MEISSNER, A., WILD, V., SIMM, R., ROHDE, M., ERCK, C., BREDENBRUCH, F., MORR, M., ROMLING, U. and HAUSSLER, S., 2007. *Pseudomonas aeruginosa* *cupA*-encoded fimbriae expression is regulated by a GGDEF and EAL domain-dependent modulation of the intracellular level of cyclic diguanylate. *Environmental microbiology*, **9**(10), pp. 2475-2485.
- MEZARD, C., POMPON, D. and NICOLAS, A., 1992. Recombination between similar but not identical DNA sequences during yeast transformation occurs within short stretches of identity. *Cell*, **70**(4), pp. 659-70.
- MICHEL-BRIAND, Y. and BAYSSE, C., 2002. The pyocins of *Pseudomonas aeruginosa*. *Biochimie*, **84**(5-6), pp. 499-510.
- MORTON, D.B. and GRIFFITHS, P.H., 1985. Guidelines on the recognition of pain, distress and discomfort in experimental animals and an hypothesis for assessment. *The Veterinary record*, **116**(16), pp. 431-436.

- MURRAY, A.W. and SZOSTAK, J.W., 1983. Construction of artificial chromosomes in yeast. *Nature*, **305**(5931), pp. 189-93.
- MURRAY, T.S., EGAN, M. and KAZMIERCZAK, B.I., 2007. *Pseudomonas aeruginosa* chronic colonization in cystic fibrosis patients. *Current opinion in pediatrics*, **19**(1), pp. 83-88.
- NGUYEN, D. and SINGH, P.K., 2006. Evolving stealth: genetic adaptation of *Pseudomonas aeruginosa* during cystic fibrosis infections. *Proceedings of the National Academy of Sciences of the United States of America*, **103**(22), pp. 8305-8306.
- NOSKOV, V.N., KOUPRINA, N., LEEM, S.H., OUSPENSKI, I., BARRETT, J.C. and LARIONOV, V., 2003. A general cloning system to selectively isolate any eukaryotic or prokaryotic genomic region in yeast. *BMC genomics*, **4**(1), pp. 16.
- NOSKOV, V.N., LEEM, S.H., SOLOMON, G., MULLOKANDOV, M., CHAE, J.Y., YOON, Y.H., SHIN, Y.S., KOUPRINA, N. and LARIONOV, V., 2003. A novel strategy for analysis of gene homologues and segmental genome duplications. *Journal of Molecular Evolution*, **56**(6), pp. 702-710.
- OLDENBURG, K.R., VO, K.T., MICHAELIS, S. and PADDON, C., 1997. Recombination-mediated PCR-directed plasmid construction in vivo in yeast. *Nucleic Acids Res*, **25**(2), pp. 451-2.
- ORR-WEAVER, T.L., SZOSTAK, J.W. and ROTHSTEIN, R.J., 1981. Yeast transformation: a model system for the study of recombination. *Proceedings of the National Academy of Sciences of the United States of America*, **78**(10), pp. 6354-8.
- OU, H.Y., CHEN, L.L., LONNEN, J., CHAUDHURI, R.R., THANI, A.B., SMITH, R., GARTON, N.J., HINTON, J., PALLAN, M., BARER, M.R. and RAJAKUMAR, K., 2006. A novel strategy for the identification of genomic islands by comparative analysis of the contents and contexts of tRNA sites in closely related bacteria. *Nucleic acids research*, **34**(1), pp. e3.
- OU, H.Y., HE, X., HARRISON, E.M., KULASEKARA, B.R., THANI, A.B., KADIOGLU, A., LORY, S., HINTON, J.C., BARER, M.R., DENG, Z. and RAJAKUMAR, K., 2007. MobilomeFINDER: web-based tools for in silico and experimental discovery of bacterial genomic islands. *Nucleic acids research*, **35**(Web Server issue), pp. W97-W104.
- PEARSON, J.P., FELDMAN, M., IGLEWSKI, B.H. and PRINCE, A., 2000. *Pseudomonas aeruginosa* cell-to-cell signaling is required for virulence in a model of acute pulmonary infection. *Infection and immunity*, **68**(7), pp. 4331-4334.
- PETERSON, J.D., UMayAM, L.A., DICKINSON, T., HICKEY, E.K. and WHITE, O., 2001. The Comprehensive Microbial Resource. *Nucleic acids research*, **29**(1), pp. 123-125.

POURCEL, C., SALVIGNOL, G. and VERGNAUD, G., 2005. CRISPR elements in *Yersinia pestis* acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology (Reading, England)*, **151**(Pt 3), pp. 653-663.

QARAH S, C.B., *Pseudomonas aeruginosa* infections.
<http://www.emedicine.com/med/topic1943.htm#section~AuthorsandEditors> edn.

QIU, X., GURKAR, A.U. and LORY, S., 2006. Interstrain transfer of the large pathogenicity island (PAPI-1) of *Pseudomonas aeruginosa*. *Proceedings of the National Academy of Sciences of the United States of America*, **103**(52), pp. 19830-19835.

RAHME, L.G., AUSUBEL, F.M., CAO, H., DRENKARD, E., GOUMNEROV, B.C., LAU, G.W., MAHAJAN-MIKLOS, S., PLOTNIKOVA, J., TAN, M.W., TSONGALIS, J., WALENDZIEWICZ, C.L. and TOMPKINS, R.G., 2000. Plants and animals share functionally common bacterial virulence factors. *Proceedings of the National Academy of Sciences of the United States of America*, **97**(16), pp. 8815-8821.

RAHME, L.G., STEVENS, E.J., WOLFORT, S.F., SHAO, J., TOMPKINS, R.G. and AUSUBEL, F.M., 1995. Common virulence factors for bacterial pathogenicity in plants and animals. *Science (New York, N.Y.)*, **268**(5219), pp. 1899-1902.

RAYMOND, C.K., SIMS, E.H., KAS, A., SPENCER, D.H., KUTYAVIN, T.V., IVEY, R.G., ZHOU, Y., KAUL, R., CLENDENNING, J.B. and OLSON, M.V., 2002a. Genetic variation at the O-antigen biosynthetic locus in *Pseudomonas aeruginosa*. *Journal of Bacteriology*, **184**(13), pp. 3614-22.

RAYMOND, C.K., SIMS, E.H. and OLSON, M.V., 2002b. Linker-mediated recombinational subcloning of large DNA fragments using yeast. *Genome Research*, **12**(1), pp. 190-7.

RAYMOND, C., 2001-last update, how to gapture.

<http://www.genome.washington.edu/UWGC/protocols/Gapture.cfm>

REDASCHI, N. and BICKLE, T.A., 1996. Posttranscriptional regulation of EcoP1I and EcoP15I restriction activity. *Journal of Molecular Biology*, **257**(4), pp. 790-803.

REISER, J. and YUAN, R., 1977. Purification and properties of the P15 specific restriction endonuclease from *Escherichia coli*. *The Journal of biological chemistry*, **252**(2), pp. 451-456.

ROZEN, S., and SKALETISKY, H., 2000. Primer3 on the WWW for general users and for biologist programmers. *Methods in Molecular Biology*, **132**, pp. 365-386

RUER, S., STENDER, S., FILLOUX, A. and DE BENTZMANN, S., 2007. Assembly of fimbrial structures in *Pseudomonas aeruginosa*: functionality and

specificity of chaperone-usher machineries. *Journal of Bacteriology*, **189**(9), pp. 3547-3555.

RUMBAUGH, K.P., GRISWOLD, J.A., IGLEWSKI, B.H. and HAMOOD, A.N., 1999. Contribution of quorum sensing to the virulence of *Pseudomonas aeruginosa* in burn wound infections. *Infection and immunity*, **67**(11), pp. 5854-5862.

SALUNKHE, P., SMART, C.H., MORGAN, J.A., PANAGEA, S., WALSHAW, M.J., HART, C.A., GEFFERS, R., TUMMLER, B. and WINSTANLEY, C., 2005. A cystic fibrosis epidemic strain of *Pseudomonas aeruginosa* displays enhanced virulence and antimicrobial resistance. *Journal of Bacteriology*, **187**(14), pp. 4908-4920.

SCHULERT, G.S., FELTMAN, H., RABIN, S.D., MARTIN, C.G., BATTLE, S.E., RELLO, J. and HAUSER, A.R., 2003. Secretion of the toxin ExoU is a marker for highly virulent *Pseudomonas aeruginosa* isolates obtained from patients with hospital-acquired pneumonia. *The Journal of infectious diseases*, **188**(11), pp. 1695-1706.

SCOTT, F.W. and PITT, T.L., 2004. Identification and characterization of transmissible *Pseudomonas aeruginosa* strains in cystic fibrosis patients in England and Wales. *Journal of medical microbiology*, **53**(Pt 7), pp. 609-615.

SEIDMAN, C.E., STRUHL, K., SHEEN, J. and JESSEN, T., 2001. Introduction of plasmid DNA into cells. *Current protocols in molecular biology / edited by Frederick M. Ausubel ...[et al.]*, **Chapter 1**, pp. Unit 1.8.

SEKINE, Y., EISAKI, N., KOBAYASHI, K. and OHTSUBO, E., 1997. Isolation and characterization of IS1 circles. *Gene*, **191**(2), pp. 183-190.

SHAIKH, F., 2005. *Cloning and analysis of specific Trypanosoma brucei TREU 927 telomeres by TAR cloning strategy*, University of Leicester.

SHANKS, R.M., CAIAZZA, N.C., HINSA, S.M., TOUTAIN, C.M. and O'TOOLE, G.A., 2006. *Saccharomyces cerevisiae*-based molecular tool kit for manipulation of genes from gram-negative bacteria. *Applied and Environmental Microbiology*, **72**(7), pp. 5027-36.

SHAVER, C.M. and HAUSER, A.R., 2004. Relative contributions of *Pseudomonas aeruginosa* ExoU, ExoS, and ExoT to virulence in the lung. *Infection and immunity*, **72**(12), pp. 6969-6977.

SIKORSKI, R.S. and HIETER, P., 1989. A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. *Genetics*, **122**(1), pp. 19-27.

SMART, C.H., WALSHAW, M.J., HART, C.A. and WINSTANLEY, C., 2006. Use of suppression subtractive hybridization to examine the accessory genome of the Liverpool cystic fibrosis epidemic strain of *Pseudomonas aeruginosa*. *Journal of medical microbiology*, **55**(Pt 6), pp. 677-688.

SMITH, E.E., SIMS, E.H., SPENCER, D.H., KAUL, R. and OLSON, M.V., 2005. Evidence for diversifying selection at the pyoverdine locus of *Pseudomonas aeruginosa*. *Journal of Bacteriology*, **187**(6), pp. 2138-2147.

SMITH, E.E., BUCKLEY, D.G., WU, Z., SAENPHIMMACHAK, C., HOFFMAN, L.R., D'ARGENIO, D.A., MILLER, S.I., RAMSEY, B.W., SPEERT, D.P., MOSKOWITZ, S.M., BURNS, J.L., KAUL, R. and OLSON, M.V., 2006. Genetic adaptation by *Pseudomonas aeruginosa* to the airways of cystic fibrosis patients. *Proceedings of the National Academy of Sciences of the United States of America*, **103**(22), pp. 8487-8492.

SMITH, R.S., WOLFGANG, M.C. and LORY, S., 2004. An adenylate cyclase-controlled signaling network regulates *Pseudomonas aeruginosa* virulence in a mouse model of acute pneumonia. *Infection and immunity*, **72**(3), pp. 1677-1684.

SOREK, R., KUNIN, V. and HUGENHOLTZ, P., 2008. CRISPR--a widespread system that provides acquired resistance against phages in bacteria and archaea. *Nature reviews.Microbiology*, **6**(3), pp. 181-186.

SPENCER, R.C., 1996. Predominant pathogens found in the European Prevalence of Infection in Intensive Care Study. *European journal of clinical microbiology & infectious diseases : official publication of the European Society of Clinical Microbiology*, **15**(4), pp. 281-285.

STARKE, J.R., EDWARDS, M.S., LANGSTON, C. and BAKER, C.J., 1987. A mouse model of chronic pulmonary infection with *Pseudomonas aeruginosa* and *Pseudomonas cepacia*. *Pediatric research*, **22**(6), pp. 698-702.

STOTLAND, P.K., RADZIOCH, D. and STEVENSON, M.M., 2000. Mouse models of chronic lung infection with *Pseudomonas aeruginosa*: models for the study of cystic fibrosis. *Pediatric pulmonology*, **30**(5), pp. 413-424.

STOVER, C.K., PHAM, X.Q., ERWIN, A.L., MIZOGUCHI, S.D., WARRENER, P., HICKEY, M.J., BRINKMAN, F.S., HUFNAGLE, W.O., KOWALIK, D.J., LAGROU, M., GARBER, R.L., GOLTRY, L., TOLENTINO, E., WESTBROCK-WADMAN, S., YUAN, Y., BRODY, L.L., COULTER, S.N., FOLGER, K.R., KAS, A., LARBIG, K., LIM, R., SMITH, K., SPENCER, D., WONG, G.K., WU, Z., PAULSEN, I.T., REIZER, J., SAIER, M.H., HANCOCK, R.E., LORY, S. and OLSON, M.V., 2000. Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen. *Nature*, **406**(6799), pp. 959-964.

TAN, M.W., MAHAJAN-MIKLOS, S. and AUSUBEL, F.M., 1999. Killing of *Caenorhabditis elegans* by *Pseudomonas aeruginosa* used to model mammalian bacterial pathogenesis. *Proceedings of the National Academy of Sciences of the United States of America*, **96**(2), pp. 715-720.

TANG, H.B., DIMANGO, E., BRYAN, R., GAMBELLO, M., IGLEWSKI, B.H., GOLDBERG, J.B. and PRINCE, A., 1996. Contribution of specific *Pseudomonas aeruginosa* virulence factors to pathogenesis of pneumonia in a neonatal mouse model of infection. *Infection and immunity*, **64**(1), pp. 37-43.

- VAN HEECKEREN, A.M. and SCHLUCHTER, M.D., 2002. Murine models of chronic *Pseudomonas aeruginosa* lung infection. *Laboratory animals*, **36**(3), pp. 291-312.
- VINCZE, T., POSFAI, J. and ROBERTS, R.J., 2003. NEBcutter: A program to cleave DNA with restriction enzymes. *Nucleic acids research*, **31**(13), pp. 3688-3691.
- WALKER, G.M., 1998. Yeast physiology and biotechnology. Chichester: J. Wiley.
- WANG, R.F. and KUSHNER, S.R., 1991. Construction of versatile low-copy-number vectors for cloning, sequencing and gene expression in *Escherichia coli*. *Gene*, **100**, pp. 195-199.
- WILLIAMS, R.J., 2003. Restriction endonucleases: classification, properties, and applications. *Molecular biotechnology*, **23**(3), pp. 225-243.
- WILSON, K., 2001. Preparation of genomic DNA from bacteria. *Current protocols in molecular biology / edited by Frederick M. Ausubel ... [et al.]*, **Chapter 2**, pp. Unit 2.4.
- WINSTANLEY, C., LANGILLE, M.G., FOTHERGILL, J.L., KUKAVICA-IBRULJ, I., PARADIS-BLEAU, C., SANSCHAGRIN, F., THOMSON, N.R., WINSOR, G.L., QUAIL, M.A., LENNARD, N., BIGNELL, A., CLARKE, L., SEEGER, K., SAUNDERS, D., HARRIS, D., PARKHILL, J., HANCOCK, R.E., BRINKMAN, F.S. and LEVESQUE, R.C., 2009. Newly introduced genomic prophage islands are critical determinants of in vivo competitiveness in the Liverpool Epidemic Strain of *Pseudomonas aeruginosa*. *Genome research*, **19**(1), pp. 12-23.
- WOLFGANG, M.C. and ET AL, 27 April 2007-last update, **protocol** for yeast recombinational cloning. Available: http://mml.sjtu.edu.cn/MobilomeFINDER/YCV_data/Protocol_for_Yeast_Recombinational_cloning.pdf.
- WOLFGANG, M.C., KULASEKARA, B.R., LIANG, X., BOYD, D., WU, K., YANG, Q., MIYADA, C.G. and LORY, S., 2003. Conservation of genome content and virulence determinants among clinical and environmental isolates of *Pseudomonas aeruginosa*. *Proceedings of the National Academy of Sciences of the United States of America*, **100**(14), pp. 8484-9.
- YANAGIHARA, K., TOMONO, K., SAWAI, T., KUROKI, M., KANEKO, Y., OHNO, H., HIGASHIYAMA, Y., MIYAZAKI, Y., HIRAKATA, Y., MAESAKI, S., KADOTA, J., TASHIRO, T. and KOHNO, S., 2000. Combination therapy for chronic *Pseudomonas aeruginosa* respiratory infection associated with biofilm formation. *The Journal of antimicrobial chemotherapy*, **46**(1), pp. 69-72.
- YOON, S.H., PARK, Y.K., LEE, S., CHOI, D., OH, T.K., HUR, C.G. and KIM, J.F., 2007. Towards pathogenomics: a web-based resource for pathogenicity islands. *Nucleic acids research*, **35**(Database issue), pp. D395-400.

ZHANG, Y. and AMZEL, L.M., 2002. Tuberculosis drug targets. *Current Drug Targets*, **3**(2), pp. 131-154.