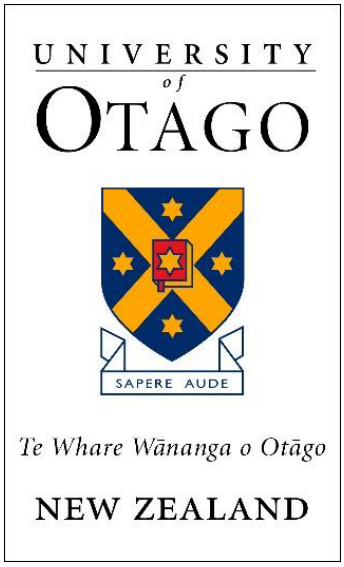


Regulatory pathways dependent on the RseP protease in *Pseudomonas aeruginosa*

Yida Jiang

A thesis submitted for the degree of

Master of Science in Biochemistry

University of Otago

Dunedin, New Zealand

March, 2017

Abstract

The opportunistic pathogen *P. aeruginosa* is responsible for a wide range of infections and is wide spread in many soils and human surroundings. In order to survive, the bacteria overcome limited iron bio-availability by releasing iron-chelating molecules known as siderophores. RseP, an intramembrane protease is critical for proteolysis of signal transduction proteins in siderophore uptake systems and Extra Cytoplasmic Function (ECF) sigma activity. RseP plays a crucial role in proteolytic cascades which allow complete degradation of anti-sigma factors and control the expression of sigma factor mediated genes in *P. aeruginosa*.

The aim of this study was to identify different genes or pathways that are mediated by RseP protease in *P. aeruginosa*, particularly new RseP-mediated genes in other pathways than siderophore uptake systems which have already been well characterized. An additional aim was to investigate whether the expression of siderophore receptor genes is altered by overexpressing RseP.

The expression of siderophore receptor genes was determined by RT-qPCR analysis in selected RseP mutant strains including overexpressed and knocked out strains. Similar expression of all receptor genes was observed in all the strains including an RseP overexpressing strain and wild-type. Findings of the present study rejected the hypothesis that overexpressed RseP may affect the transcription of siderophore receptor genes. We identified 69 genes which showed a significant differentially expression in overexpressed RseP and RseP knocked out strains by RNA-Seq analysis. Thus, these differentially expressed genes were thought to be RseP-dependent genes. Furthermore, minimal inhibitory concentration (MIC) testing demonstrated that the expression of *rseP* had no direct effect on the changes in resistance phenotypic of *P. aeruginosa*.

Acknowledgements

Thank you Iain Lamont, for giving me massive help and guidance academically throughout this project and making this research an enjoyable laboratory experience. Thanks for teaching me to design experiments independently and provide directions for troubleshooting when things go wrong. Thanks to Lois Martin for finding my way in the lab and providing technical guidance when I was starting to plan my own experiments and the supply of high quality baking for morning tea section. Thank you Sam for helping me through some hard parts of bioinformatics concepts in this project. Specifically, thank Iain, Attika and Charlotte for helpful comments on earlier version of this thesis. A big thank you to members of Lamont lab who were present throughout the whole duration of this project, thanks for helping to make Lamont lab an awesome place to work and study. Thank you to other members of the Lamont lab, past and present, who provided input into this project.

Table of Contents

Abstract		ii
Acknowledgements		iii
List of Tables		ix
List of Figures		x
List of abbreviations		xi
	Chapter 1 Introduction	1
1.1	Taxonomy and biology of *Pseudomonas aeruginosa*	1
1.2	Genomics of *Pseudomonas aeruginosa*	1
1.3	Pathogenesis and Disease	2
1.4	Antibiotics resistance in *P. aeruginosa*	2
1.5	Siderophores and iron condition in *P. aeruginosa*	3
1.5.1	Endogenous siderophores	3
1.5.2	Exogenous siderophores	4
1.6	Cell-Surface Signalling (CSS) system	4
1.6.1	Mechanisms of iron uptake	4
1.6.2	RseP-dependent proteolytic cascade in CSS	5
1.6.3	RseP protease in pyoverdine system	5
1.6.4	RseP protease in ferrichrome and desferroxiamine systems	7
1.7	RseP protease is responsible for envelop stress response and alginate synthesis pathways	8
1.8	Structure of RseP	11
1.9	Regulation of RseP activity	12
1.10	Aims of study	12
	Chapter 2 Materials and Methods	14
2.1	General	14
2.2	Media	15
2.3	Primers for PCR and RT-qPCR	17
2.4	Bacteria cell strains and plasmids	18

2.4.1	Bacterial strains used	18
2.4.2	Plasmids	19
2.4.3	Growth and maintenance of bacteria	19
2.5	DNA-Sequencing	19
2.6	General molecular biology methods	20
2.6.1	Purification of PCR product	20
2.6.2	Genomic DNA isolation	20
2.6.3	Plasmid DNA isolation	20
2.6.4	Restriction endonuclease (RE) enzyme digests	20
2.6.5	Agarose gel electrophoresis	21
2.6.6	Gel extraction	21
2.6.7	Butanol purification	22
2.6.8	DNA qualification	22
2.6.9	General cloning methods	22
2.6.9.1	Phosphatasing vectors prior to ligation	22
2.6.9.2	Ligation of DNA molecules	22
2.6.9.3	Transformation	23
2.6.9.4	Identification of recombinant clones	23
2.7	Manipulating involving Polymerase Chain Reaction (PCR)	24
2.7.1	PCR and Gradient PCR	24
2.7.2	PCR screening of bacterial colonies	25
2.8	Quantification of gene expression by Real-Time Polymerase Chain Reaction (RT-qPCR)	25
2.8.1	Induction of mRNA synthesis	25
2.8.2	RNA extraction and DNase treatment	25
2.8.3	RNA quantification	26
2.8.4	cDNA synthesis	26
2.8.5	RT-qPCR primer efficiency	26
2.8.6	RT-qPCR quantification	27
2.9	β-galactosidase assay	28
2.10	Minimum inhibitory concentration (MIC) test	29

2.11	Swimming and swarming motility assay	29
2.12	Whole Genome Sequencing(WGS) analysis	30
2.13	RNA-Sequencing analysis	30
2.14	Specialist database used in this study	31
	Chapter 3 Results	32
3.1	Research strategy of this study	32
3.2	β-galactosidase activity of pJN105 expression induced by arabinose	32
3.2.1	Introduction	32
3.2.2	Preparation of β-gal assay	33
3.2.3	Results of β-gal activity	34
3.3	DNA cloning of construct PAORseP^{++}	35
3.3.1	The aim of DNA cloning	35
3.3.2	The plasmid of pLUG-Prime vector and pJN105 vector	35
3.3.3	Cloned *rseP* sequence into the pJN105 vector	37
3.3.4	pJN105::*rseP* cloning selection and transformation	41
3.3.5	DNA sequencing to confirm cloning orientation	42
3.4	Sequence analysis of the intended PAO(*clpP*) strain	42
3.4.1	Introduction	42
3.4.2	Construction of intended PAO(*clpP*) strain	42
3.4.3	Analysis of PAO(*clpP*) by Whole Genome Sequencing	44
3.5	Sample collection for RNA-Seq	45
3.5.1	The growth of *P. aeruginosa* strains	45
3.5.2	Sampling preparation of *P. aeruginosa* strains for RNA-Seq	47
3.6	Quantifying the expression of *rseP* and siderophore receptor genes in *P. aeruginosa* strains	48
3.6.1	Introduction for RT-qPCR analysis	48
3.6.2	Relative quantification parameters	48
3.6.2.1	Relative quantification	48
3.6.2.2	Validation of housekeeping genes	49
2.6.2.3	Validation of primers	49

3.6.3	The result of RseP expression of *P. aeruginosa* strains	50
3.6.4	Quantification of FpvA, FoxA and FiuA expression	52
3.7	RNA-Seq analysis	55
3.7.1	Introduction	55
3.7.2	RNA-Seq sample preparation	55
3.7.3	Sample quality	57
3.7.4	Differential expression analysis	59
3.7.4.1	DESeq and EdgeR	60
3.7.4.2	Initial comparison between PAORseP^{++} and PAO *rseP*	60
3.7.4.3	Differential expressed comparison analysis between PAO1-WT, PAORseP^{++} and PAO *rseP*	61
3.7.4.4	The comparison of PAO(*clpP*) and PAO1-WT	63
3.7.4.5	Differential expressed comparison analysis between PAO(*clpP*), PAORseP^{++} and PAO *rseP*	63
3.7.5	RseP-dependent genes (69 DE genes) identified from PAORseP^{++} and PAO *rseP*	65
3.7.6	RT-qPCR validation of RNA-Seq results	68
3.7.7	Functions of differential expressed genes response to RseP	70
3.7.7.1	DE genes with unknown functions	70
3.7.7.2	DE genes with known functions	70
3.7.8	Conclusion	72
3.8	The Minimal inhibitory concentration (MIC) tests of *P. aeruginosa* strains	73
3.8.1	Introduction	73
3.8.2	MIC determination	73
3.8.2.1	Sample templates for MIC tests	73
3.8.2.2	Method of MIC test	73
3.8.3	MIC results of tobramycin and ciprofloxacin	74
3.8.4	Conclusion	76
3.9	Motility assay on *P. aeruginosa* strains	77
3.9.1	Introduction of the motility assay	77
3.9.2	Results of swimming motility	77

3.9.3	Results of swarming motility	80
	Chapter 4 Discussion	82
4.1	Summary of key findings	82
4.2	The growth temperature of *P. aeruginosa* strains	82
4.3	The correlation between RseP and siderophore receptor genes (*fpvA*, *foxA*, *fiuA*)	83
4.4	RNA-Seq analysis	85
4.4.1	The filtering parameters of differential expression analysis	85
4.4.2	The comparison analysis between the reference strains	85
4.4.3	RNA-Seq results may be affected by temperature	85
4.4.4	The established RseP-dependent pathways, ferrisiderophore systems and alginate synthesis pathway	86
4.4.5	RseP may play a role in the MexXY-OprM efflux transport system	87
4.4.6	RseP and the OpuC Osmolarity ABC transporter system	88
4.5	RseP and the antibiotic resistance of *P. aeruginosa*	89
4.6	Conclusion	91
4.7	Future directions	91
Bibliography		92
Appendix 1		100
Appendix 2		101
Appendix 3		103
Appendix 4		107

List of Tables

Table 2.1	Antibiotics for use with *P. aeruginosa* and *E. coli*	15
Table 2.3	Oligonucleotide primers used in this study	17
Table 2.4.1	Bacterial strains used in this study	18
Table 2.4.2	The plasmid used in this study	19
Table 2.6	Restriction digestion protocol	21
Table 3.5	The OD_{600} and growth hours of samples which were collected for RNA-Seq	47
Table 3.6	The T-test of the mean difference expression of *fpvA*, *foxA* and *fiuA* between PAORseP^{++}and PAO1-WT, PAO *rseP* and PAO1-WT strains	54
Table 3.7.1	The RNA concentration and A260/280 ratio of 12 RNA samples	56
Table 3.7.2	Total read counts summarised for each sample strain in biological triplicates, provided by NZGL	59
Table 3.7.3	The List of RseP-mediated genes (69) were identified from PAORseP^{++} vs PAO *rseP*, analysed by RNA-Seq	66
Table 4.4	The temperature-regulated genes of *P. aeruginosa* differentially expressed between PAORseP^{++} and PAO *rseP*	86

List of Figures

Figure 1.1	The chemical structure of pyoverdine from *Pseudomonas aeruginosa* strain PAO1	3
Figure 1.2	Siderophores systems in *P. aeruginosa*	4
Figure 1.3	Model for the regulation of ECF sigma-factor activity via proteolytic cascade	9
Figure 1.4	Topology of RseP in the cytoplasmic membrane	11
Figure 3.2.1	The map of pJN106 plasmid	33
Figure 3.2.2	β-galactosidase assay of PAOpJN106	34
Figure 3.3.1	Plasmids maps of pLUG Prime TA and pJN105 cloning vector	36
Figure 3.3.2	Construction of the pLUG::*rseP* construct	38
Figure 3.3.3	Construction of the final pJN105::*rse*P^{++}overexpressed strain	40
Figure 3.4.1	Construction of the intended PAO(*clpP*)	43
Figure 3.4.2	The alignment results visualized by IGV of PAO(*clpP*) strain	44
Figure 3.5.1	Growth analysis of *P. aeruginosa* strains	46
Figure 3.6.1	Validation of qPCR primer sets- example of Standard Curve (*rseP*)	49
Figure 3.6.2	RseP primers validated by conventional PCR	50
Figure 3.6.3	RseP expression level	52
Figure 3.6.4	The expression of *fpvA*, *foxA* and *fiuA*	53
Figure 3.7.1	Pipeline of RNA-Seq experiment	57
Figure 3.7.2	Quality score of samples of RNA-Sequencing	58
Figure 3.7.4.1	The differential expressed analysis between PAORseP^{++} and PAO *rseP*	61
Figure 3.7.4.2	The identified differential expressed genes between PAO1-WT, PAORseP^{++} and PAO *rseP*	62
Figure 3.7.4.3	The differential expressed genes identified from comparisons PAO(*clpP*) groups	64
Figure 3.7.6.1	PA5471 and PA2275 primers validated by gradient PCR from 55˚C to 64˚C	68
Figure 3.7.6.2	The expression in both RT-qPCR and RNA-Seq analysis between P. aeruginosa strains	70
Figure 3.8	Tobramycin and ciprofloxacin MIC results against four different strains	75
Figure 3.9.1	Swimming motility of *P. aeruginosa* strains	79
Figure 3.9.2	Swarming motility of *P. aeruginosa* strains	80
Figure 4.4.1	Regulation of the MexXY-OprM efflux in *P. aeruginosa*	87
Figure 4.4.2	OpuC ABC import system	89

List of Abbreviations

Unless otherwise stated, all abbreviations are described in the "Instructions to authors" of the Biochemical Journal (2005)

ABC	ATP-binding cassette
Amp	ampicillin
BCIG	5-Bromo-4-chloro-3-indolyl-β-D-galactopyranoside
β-gal	β-galactosidase
cDNA	complementary DNA
CF	Cystic Fibrosis
CIP	ciprofloxacin
CM	cytoplasmic membrane
CSS	Cell Surface Signalling
DE	differential expressed
ddH_2O	Double distilled water
dNTP	Deoxynucleotide
ECF	Extracellular Cytoplasmic Function
EDTA	ethylenediamminetetraacetic acid
EU	enzyme units
Fiu	ferrichrome
Fox	desferrioxamine
Fpv	ferripyoverdine
Gm	gentamycin
IGV	Integrative Genomics Viewer
IPTG	Isopropyl-β-D-thiogalactopyranoside
KB	King's B broth
lacZ	b-galactosidase gene
LB	Lysogeny Broth
MCS	multiple cloning site
MIC	minimal inhibitory concentration
mRNA	Messenger RNA
NZGL	New Zealand Genomics Limited
ODn	optical density (n=wavelength)
OM	outer membrane
ONPG	o-nitrophenyl-β-D-galactoside
PCR	polymerase chain reaction
R	resistance
RE	restriction enzyme
rep	replication initiator genes
RNA-Seq	RNA sequencing
RNAP	RNA polymerase
RND	resistance nodulation-division

RT	Reverse transcriptase
RT-qPCR	reverse transcription quantitative PCR
rpm	revolutions per minute
TAE	Tris-acetate-EDTA
TE	Tris-EDTA
TM	transmembrane
Tm	melting temperature
TOB	tobramycin
UV	Ultra violet
kv	kilovolts
w/v	weight per volume
WGS	whole genome sequencing
WT	wild type

Chapter 1 Introduction

1.1 Taxonomy and biology of *Pseudomonas aeruginosa*

Pseudomonas aeruginosa are Gram-negative motile rods with polar flagella (Bergey et al., 1923). *P. aeruginosa* are found in a wide range of environments including animals and aquatic environments (Lister et al., 2009). *P. aeruginosa* can utilize a range of energy sources to catabolize for growth such as organic acids and amino acids (Rojo, 2010).

To survive extreme conditions with low water activity, *P. aeruginosa* must adjust in response to osmotic changes. Thus, the uptake of osmoprotectant compounds (glycine betaine) allowed *P. aeruginosa* to maintain intracellular osmoprotection above a critical threshold to provide the osmotic homeostasis to cells (Fitzsimmons et al., 2012; Wargo, 2013). Bacteria encode a specific ATP binding cassette (ABC) transporter, the OpuC transporter which has been shown to be critical in the presence of osmotic stress with high concentration of salts (Fitzsimmons et al., 2012).

In addition to growing in a wide range of environments, *P. aeruginosa* is an opportunistic pathogen of clinical relevance and shows a high level of multi-drug resistance (Lister et al., 2009). This is due to the low permeability of their membranes and also the large number of efflux systems they have available (Stover et al., 2000). *P. aeruginosa* require flagella to allow bacteria to be able to cross tissue barriers and evade host immune defences (Murray et al., 2010).

1.2 Genomics of *Pseudomonas aeruginosa*

The genomes of pseudomonads are large and contain around 6.3 million base pairs genome (6,300,000bp) (Stover et al., 2000). The genome is predicted to contain 5570 open reading frames and *P. aeruginosa* is highly G+C rich (66.6% G+C). 8.4% of *P. aeruginosa* genes are involved in regulation and these include the Extra Cytoplasmic Function (ECF) transcriptional regulators (Stover et al., 2000). The complete *P. aeruginosa* genomic sequence data is available through the Pseudomonas Genome Database. *P. aeruginosa* PAO1 is widely used as a 'reference strain' for comparative genomic analyses, which allow genetic differences between mutated strains to be distinguished. Through genome comparisons, a conserved region has been identified with the genes that are present in all *P. aeruginosa* strains and a more variable

'accessory genome' which differs slightly between strains (Kung et al., 2010). This includes the regions of the flagella cluster and the pyoverdine locus (Spencer et al., 2003). To conclude, it has been reported that with a large conserved genome, *P. aeruginosa* can easily cause infection in different hosts (Kiewitz and Tümmler, 2000; Spencer et al., 2003).

1.3 Pathogenesis and Disease

P. aeruginosa is responsible for a wide range of chronic infections, such as cystic fibrosis (CF) (Reinhardt et al., 2007). Cystic fibrosis is a common autosomal recessive disorder that affects the lungs, which results in the regulation disorder of secretions and metabolism in the human body (Davies, 2002). Anaerobic respiration allows *P. aeruginosa* to form a thick biofilm, where a group of cells stick together and adhere to the surface and this can cause shortness of breath by blocking the bronchi. The ability of *P. aeruginosa* to cause infection is affected by virulence factors, involving the biofilm formation and the increased level of antibiotic resistance (Lister et al., 2009; Qiu et al., 2007).

1.4 Antibiotic resistance in *P. aeruginosa*

The treatment of *P. aeruginosa* infections is problematic due to the large number of antibiotic resistance genes which are chromosomally encoded (Lister et al., 2009). A variety of outer membrane proteins are encoded by the genome, which provide *P. aeruginosa* with the ability to import a large number of substrates and to export antibiotics and virulence factors (Stover et al., 2000). There is one pathway that has been well-studied recently, the MexXY-OprM efflux system of antibiotic resistance in *P. aeruginosa*. The MexXY-OprM efflux belongs to resistance-nodulation-division (RND) family transporters and is located in the cytoplasmic membrane that export antibiotics. The combination of MexY (cytoplasmic membrane transporter) and MexX (periplasmic membrane fusion protein) pumps out antimicrobials and results in increased antibiotic resistance of *P. aeruginosa*. A mutation in the *mexY* gene was found in cystic fibrosis lungs and this leads to an increase in multi-drug resistance (Vettoretti et al., 2009). This indicated that there is a need for MexXY in the hostile environment of cystic fibrosis lung. The MexXY efflux system has allowed *P. aeruginosa* to acquire higher levels of resistance after exposure to various antibiotics including aminoglycosides and fluoroquinolones. Aminoglycosides act as bacterial protein synthesis inhibitors, while quinolones interfere with DNA replication and transcription in bacteria (Lister et al., 2009). This information provides some evidence as to how this organism can cause infections in

different hosts. However, the molecular mechanisms and associated genomic events with a higher level of drug resistance of *P. aeruginosa* in CF patients remain uncertain.

1.5 Siderophores and iron condition in *P. aeruginosa*

Iron is an important element for the survival of bacteria because it is incorporated into proteins, which play roles in oxidative stress defences and metabolism (Bullen et al., 2005). The iron concentration in cystic fibrosis lungs may be higher than other environments but this is still too low to allow bacteria to survive and grow (Reid et al., 2009). High concentrations of iron appear to promote biofilm biosynthesis (Banin et al., 2005). Free ferric iron can interact with oxygen within the cell, which can cause damage to the cell. *P. aeruginosa* are found in oxygenic environments and are able to secrete iron chelating molecules known as siderophores in response to low iron levels within the cells (Andrews et al., 2003). Siderophores have a high affinity towards ferric ions (Fe^{3+}) forming hexadentate complexes, which allow binding and transport of iron in cells (Crosa, 1997).

1.5.1 Endogenous siderophores

Pyoverdines are endogenous siderophores that *P. aeruginosa* can produce themselves. Pyoverdines are required for virulence in animal models of disease (Lamont et al., 2002). Pyoverdine is a mixed catecholate-hyroxamate siderophore, which consists of three components, a quinoleinic chromophore which imparts the colour and fluorescence to the molecule, an acyl group (either dicarboxylic acid or monoamide) and an attached peptide chain (Figure 1.1) (Meyer and Abdallah, 1978). Secreted pyoverdine coordinates a hexadentate chelation of iron which is recognized by a pyoverdine receptor (FpvA) in the outer membrane (Meyer et al., 2002).

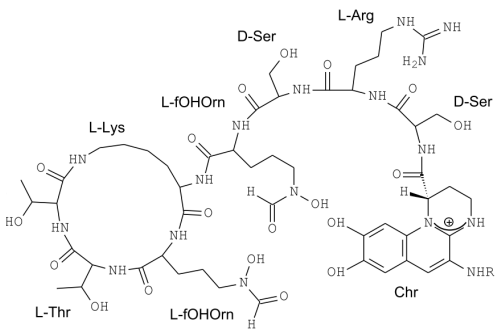


Figure 1.1 The chemical structure of pyoverdine from *Pseudomonas aeruginosa* strain PAO1 (Visca et al., 2007).

1.5.2 Exogenous siderophores

P. aeruginosa has ability to take up siderophores of both bacterial and fungal origin. These include exogenous siderophores such as ferrichrome (Fiu) and desferrioxamine (Fox). Ferrichrome, a cyclic trihydroxamate siderophore binds iron with high affinity and is taken up by *P. aeruginosa* through the FiuA receptor (Kuma et al., 2000; Llamas et al., 2006). Desferrioxamine is used therapeutically for the binding of excess blood iron in the treatment of thalassemia. It is transported into *P. aeruginosa* by the FoxA receptor (Llamas et al., 2006)

1.6 Cell-Surface Signalling (CSS) system

1.6.1 Mechanisms of iron uptake

Ferri-siderophores are actively transported across the outer membrane of Gram negative bacteria via binding to specific receptors. *P. aeruginosa* produces ferri-siderophores (siderophore-Fe^{3+}) including pyoverdine, ferrichrome and desferrioxamine which account for the high binding affinity ferric (Fe^{3+}) metal ion. In *Pseudomonas*, the uptake of ferri-siderophores requires active transport via a group of highly specific receptors and the pathways called Cell-Surface Signalling (CSS) cascade.

The activity of alternative sigma factors is controlled by CSS cascades that start at the outer membrane and end in the cytoplasm (Figure 1.2b). CSS enables self-uptake of iron via ferri-siderophores and signal transduction in the ferripyoverdine, ferrichrome and desferrioxamine siderophore systems were well-characterized in *P. aeruginosa*. The ferrisderophore complexes are recaptured by specific receptors. The CSS systems then transduce signals to the anti-sigma factor, leading to the activation of an ECF sigma factor in the cytoplasm that initiates the expression of receptor genes (Figure1.2b) (Llamas et al., 2014; Draper et al., 2011).

Systems		
Fpv	**Fox**	**Fiu**
FpvA	**FoxA**	**FiuA**
FpvR	**FoxR**	**FiuR**
FpvI	**FoxI**	**FiuI**
PvdS		

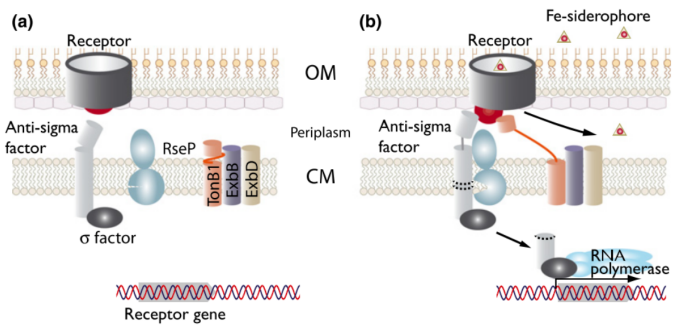


Figure 1.2 Siderophores systems in *P. aeruginosa* The siderophore systems include ferripyoverdine (Fpv), ferrichrome (Fiu) and desferrioxamine (Fox). Red bold represents the

ECF sigma factors. **(a)** In the inactive state, the anti-sigma factor sequesters the ECF sigma factor at the cytoplasmic membrane keeping it inactive. **(b)** When signals occur, receptors transmit the signals to the cytoplasm, and RseP protease mediates the partial degradation of anti-sigma factors allowing the expression of receptor genes involved in siderophore synthesis (Llamas et al., 2014). FpvA, FoxA and FiuA are the siderophore receptors of ferripyoverdine (Fpv), ferrichrome (Fiu) and desferrioxamine (Fox), respectively. FpvR, FoxR and FiuR are the anti-sigma factors and FpvI, PvdS, FoxI and FiuI are the sigma factors.

1.6.2 RseP-dependent proteolytic cascade in CSS

Intracellular proteases are involved in CSS pathways to control ECF-sigma-mediated gene expression via a regulatory intramembrane proteolysis cascade (Llamas et al., 2014; Koide et al., 2014). RseP, a periplasmic protease plays a role in the proteolysis cascade and it introduces the secondary cleavage within the transmembrane domain of the substrate in *P. aeruginosa*. This proteolytic cascade initiates degradation of the anti-sigma factors within the periplasmic domain (with site-1 proteases, DegS). The fragment of anti-sigma factor is susceptible to a second proteolytic event (site-2) (with RseP protease) after cleavage by 'site-1' proteolysis step. Site-2 proteolysis activates another proteolytic step, resulting in complete degradation of the remaining cytoplasmic anti-sigma domain (by ClpXP, site-3) (Kanehara et al., 2002). Signal transduction in the ferripyoverdine (Fpv), Ferrichrome (Fiu) and desferrioxamine (Fox) siderophore systems are required for RseP mediated proteolytic degradation of sigma regulators and release of the sigma ECF in the active form, allowing the expression of receptor genes for siderophore uptake (Llamas et al., 2014; Draper et al., 2011).

1.6.3 RseP protease in pyoverdine system

Pyoverdine is considered to be the major siderophore produced by *P. aeruginosa* because it has a high affinity for Fe^{3+} under conditions lacking ferric iron that allow bacteria to survive (Meyer and Abdallah, 1978). The ferripyoverdine system is the best characterized CSS system. In response to the environmental siderophore signal, the receptor (FpvA) transmits pyoverdine siderophore signals to the periplasm membrane and initiates the degradation of the anti-sigma factor (FpvR) by RseP protease. Consequently, ECF sigma factors (PvdS and FpvI) are released after degradation and increase the expression of receptor genes, in response to pyoverdine uptake in *P. aeruginosa* (Llamas et al., 2014). The signal of ferri-siderophore can cause conformational changes allowing FpvR to interact with receptor FpvA and initiate the signal transduction (Noinaj et al., 2010; Draper et al., 2011) (Figure1.2).

RseP participates in the regulation of pyoverdine synthesis through a sigma factor. The interaction of receptor and FpvR initiates a proteolytic cascade that involves a site-1 protease (DegS), and site-2 protease, RseP, which mediates the partial degradation of FpvR after initial cleavage and the anti-sigma factor is thought to be completely degraded and releases sigma factors in cells (Heinrich and Wiegert, 2009) (Figure 1.2b) whereas, in the absence of siderophore signals, FpvR is cleaved only by site-1 periplasmic proteases leading to incomplete cleavage with a stable 20 kDa fragment remaining in the cytoplasmic membrane that inhibits the target sigma factors, therefore preventing the transcription of receptor genes (Figure1.2a). Consequently, the proteolytic cascade is determined by ferripyoverdine levels (Draper et al., 2011).

Furthermore, a cytoplasmic protease (ClpXP) provides the third cleavage of FpvR, resulting in complete removal of the remaining fragments of FpvR at the cytoplasmic membrane and sigma factors are released (Figure1.2). Those sigma factors then bind to the RNA polymerase, facilitating promoter binding of target receptor genes. The sigma factor (PvdS) induces the transcription of pyoverdine synthesis genes (Lamont et al., 2002) and initiates the transcription of the receptor protein to increase the uptake of ferri-pyoverdine (Beare et al., 2003; Rédly and Poole, 2003) (Figure1.2).

To summarise, RseP protease is responsible for the activity of pyoverdine systems involved in degradation of anti-sigma factors. RseP is also required for complete proteolysis of the anti-sigma factor in ferrichrome and desferrioxamine siderophore systems of *P. aeruginosa.*

1.6.4 RseP protease in ferrichrome and desferrioxamine systems

Signal transduction in the ferrichrome (Fiu) and desferroxiamine (Fox) share many features with each other. For example, both systems control the activity of a single sigma factor by an anti-sigma factor and RseP-mediated proteolysis of the anti-sigma factor upon induction of the pathway as the mechanism for sigma factor activation (Llamas et al., 2014; Llamas et al., 2006). The model of the siderophore-induced CSS system is shown in Figure 1.2 and this process can be illustrated by Fox systems. As in pyoverdine CSS systems, iron-loaded ferrioxamine binds to the CSS receptor FoxA, which starts a signalling cascade via FoxR anti-sigma that results in the activation of the sigma factor (FoxI) through C-terminal interaction in the cytoplasm (Llamas et al., 2006) (Figure 1.2). Upon activation, sigma FoxI initiates the transcription of the gene encoding the FoxA receptor (Llamas et al., 2006). When siderophores are present, the level of anti-sigma factors is decreased, but sigma factor activity increases (FoxI) (Llamas et al., 2006; Mettrick and Lamont, 2009) (Figure 1.2). Similar to FpvR, FoxR and FiuR were degraded in the presence of RseP protease with siderophores and a stable fragment (18 kDa) of FoxR protein remained in the absence of siderophores, indicating siderophore signals were required for degradation of anti-sigma factors in these CSS systems (Draper et al., 2011). In addition, expressions of desferrioxamine and ferrichrome receptor genes were dramatically decreased in RseP mutants compared to wild-type in the presence of siderophore signal induction (Draper et al., 2011). Bastiaansen et al., (2015) also reported that RseP is directly involved in the FoxR system by secondary cleavage in intra-membrane regions. By using an HA-tag attached to the N terminal of FoxR protein, an N-terminal subfragment (12 kDa) was detected in the wild-type strain in the absence of a ferrioxamine signal in the *rseP* mutant. However, a larger fragment (15 kDa) was detected in the mutant (Bastiaansen et al., 2015). The presence of this fragment is associated with a significant increase of sigma FoxI activity, suggesting the cleavage of FoxR correlates with sigma FoxI activity (Bastiaansen et al., 2014).

To summarise, in the Fiu and Fox systems, the FiuA or FoxA receptor transmits a signal to FiuR or FoxR (antisigma protein). After sensing the presence of ferrichrome, RseP causes the secondary cleavage of anti-sigma factors (FiuR, FoxR) and releases sigma FiuI or FoxI. Similar to FpvR, two periplasmic proteases and a cytoplasmic protease are involved in the degradation of FoxR and FiuR in the proteolytic cascade. Siderophore signals are required to initiate the site-1 periplasmic protease activity and allowed for complete degradation of anti-sigma factors. The free sigma factor initiates the transcription of receptor genes in CSS systems. The involvement of RseP in CSS systems (particularly in pyoverdine, ferrichrome and

desferrioxamine siderophore systems) mediate a large class of anti-sigma factors that interact with receptors in *P. aeruginosa* (Mettrick and Lamont, 2009).

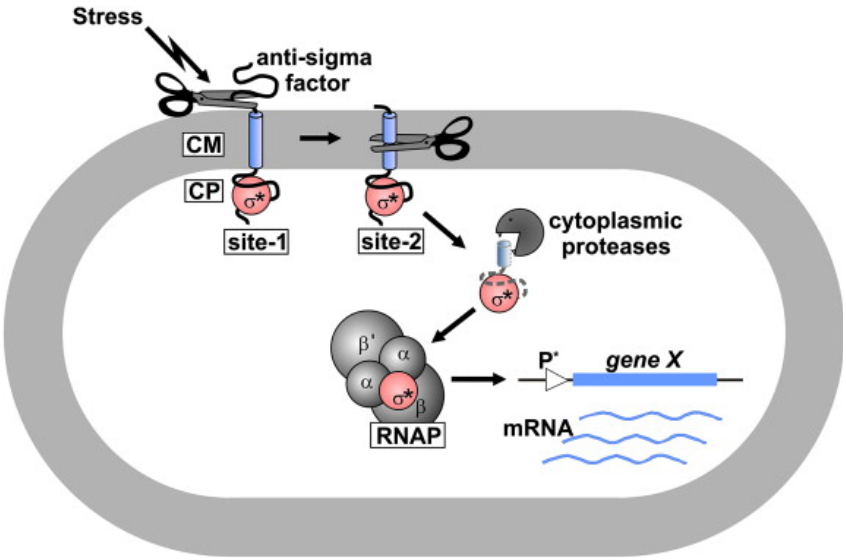
1.7 RseP protease is responsible for envelope stress response and alginate synthesis pathways

RseP is also responsible for the degradation of regulators in envelope stress response in *E. coli* and participate in alginate production in *P. aeruginosa*. The RseP protease belongs to the site-2 protease family of intramembrane cleaving proteases (Akiyama et al., 2004). Cell envelope stress responses maintain the integrity of the cell compartment and protect the envelope from damaging stresses. (MacRitchie et al., 2008; Rowley et al., 2006). The cell envelope consists of the inner membrane (IM), outer membrane (OM), and periplasmic space between membranes in cells. The σ^E-dependent envelope stress response in *E. coli* was found to be regulated by proteases in the proteolytic cascade. This σ^E-dependent pathway homologue to ECF sigma regulates the alginate synthesis pathway in *P. aeruginosa* where overproduction of alginate forms a biofilm that protects bacteria from antibiotic treatment. Alginate is a water-insoluble and cream-coloured substance. The alginate is involved in the pathogenesis of *P. aeruginosa* infection in CF patients by enhancing the biofilm formation (Govan and Deretic, 1996). The production of alginate promotes *P. aeruginosa* to grow in the lung against a range of stresses (Høiby et al., 2010).

MucA, anti-sigma factor and AlgU sigma factor in the alginate production pathway in *P. aeruginosa* are homologous to anti-sigma factors RseA, RseB and σ^E, in the envelope stress response system *E. coli* (Rowley et al., 2006; Schurr and Deretic, 1997) (Figure 1.3). The regulation of σ^E–dependent envelope stress response is shown in Figure 1.3. In *E. coli*, this is the best characterized example of envelope stress response that involves RseP protease. This protease had secondary cleavage at the membrane in *E. coli* (Kanehara et al., 2001). RseP catalyses proteolytic cleavage of the membrane-bound protein as an essential step in intra-membrane signal transduction in the envelope stress response system (Brachinger and Ades, 2013). In the presence of stress, RseP provides secondary cleavage of anti-sigma factors (RseA), resulting in the release of σ^E and initiating the transcription of the genes which are required for the stress response in *E. coli* cells (Figure 1.3). Consequently, the process results in maintaining the envelope integrity in the periplasm (Gruber and Gross, 2003). Thus, the RseP protease plays an important role in the envelope stress response in *E. coli*. RseA was

degraded by site-1 DegS protease in the presence of stress and activates the site-2 RseP protease in the intra-membrane space. Moreover, the RseA could be degraded by ClpXP protease (site-3) in the cytoplasmic membrane and initiate the transcription of genes in response to the stress (Flynn et al., 2003; Levchenko et al., 2005) (Figure 1.3).

A)



B)

Organism	σ/anti-σ	Site-1	Site-2	Cytoplasm
Escherichia coli	σ^E/RseA	DegS	RseP	ClpXP
Pseudomonas aeruginosa	AlgU/MucA	AlgW	RseP	ClpXP

Figure 1.3 Model for the regulation of ECF sigma-factor activity via proteolytic cascade **A)** The figure of ECF sigma-factor activity through transmembrane signalling that involves regulated intramembrane proteases. 'Site-1' and 'site-2' represent the periplasmic proteases involved in initial and secondary cleavage in proteolytic cascade. When stress signals occur extramembrane, the signal is transferred to the periplasmic space and activates the initial and secondary cleavage (RseP) of the anti-sigma factor and the remaining anti-sigma fragments are completed degraded by cytoplasmic protease. The released-sigma factor then activates the transcription of alginate production genes in *P. aeruginosa* OM: outer membrane, CM: cytoplasmic membrane, CP: cytoplasm; RNAP: RNA polymerase. **B)** The table represents the ECF sigma factors that are reported to be regulated by proteases in *E. coli* and *P. aeruginosa*. (Heinrich and Wigenet, 2009)

Studies demonstrated that signals are required to activate DegS protease in the proteolytic cascade. DegS provides the initial cleavage signal to RseP by releasing an activating peptide that initiates the secondary cleavage by RseP. Thus RseP protease activity must be preceded by a prior cleavage mediated by DegS (Gruber and Gross, 2003, Bohn et al., 2004). In the absence of signal/stress, DegS (site-1 protease) may exist in an inactive conformation and RseA could not be degraded without initiation by DegS (Brachinger and Ades, 2013). Consequently, RseA binds σ^E from proteolysis by DegS and prevents the σ activity. Schurr and Deretic, (1997) reported that AlgU controls alginate production in *P. aeruginosa*, similar to the *E. coli* sigma factor. In *P. aeruginosa*, MucA is an anti-sigma factor located in the inner membrane that binds to and inhibits the sigma factor AlgU. The AlgU directly regulates the production of alginate by directing transcription of the alginate-biosynthesis operon (first gene *algD*) from an AlgU-dependent promoter (Rowley et al., 2006). The transcription of the alginate biosynthetic enzymes from the *algD* promoter is activated by releasing the AlgU sigma factor in the cell (Figure 1.3). Induction of transcription of the alginate biosynthetic operon at the *algD* promoter appears to be the key event leading to the production of alginate (Deretic et al., 1994) (Figure 1.3).

By using an HA-Flag-tag attached to MucA, low expression of *algD* of *rseP* mutation strains were detected (Damron and Yu, 2011). A higher level of MucA and a lower level of AlgU were detected in the *rseP* mutant strain compared to wild-type. In addition, data have shown that *rseP* mutants could not have a mucoid phenotype (Qiu et al., 2007). These results suggest that the RseP is responsible for controlling the activity of MucA anti-sigma factor with a proteolytic process, which regulates the alginate synthesis pathway. Loss of *rseP* will decrease alginate production (by inhibiting sigma factor activity and decreasing the expression of sigma factors). Therefore, RseP protease is required for proteolysis of the anti-sigma factor in the different pathways at the intra-membrane region and mediates sigma factor activity.

To summarise, there are four known anti-sigma substrates (FpvR; FoxR; FpvR and MucA), of RseP. These anti-sigma factors are involved in ferrisiderophore CSS systems and alginate synthesis pathways in *P. aeruginosa* (Llamas et al., 2014).

1.8 Structure of RseP

RseP site-2 proteases have conserved amino acid sequence motifs, a typical zinc metalloprotease motif HE*XX*H, and a C-terminal-located LDG motif. RseP is located in the cytoplasmic membrane and consists of four transmembrane (TM) segments. The structure of RseP was well-determined in *E. coli* (Figure 1.4). The RseP active site is formed by HExxH motif in TM1 and the LDG motif in TM3 (Figure 1.4) (Kanehara et al., 2001). The metal ion is coordinated by histidine residues of HExxH motif in TM2 and the aspartate of the LDG motif in TM4 (Li et al., 2009). The PDZ domains are located in the periplasmic domain between TM2 and TM3 (Figure 1.4) (Kanehara et al., 2001). Mutations found in these motifs could abolish the proteolytic activities, thus these motifs are of vital importance to RseP protease activity (Kanehara et al., 2002).

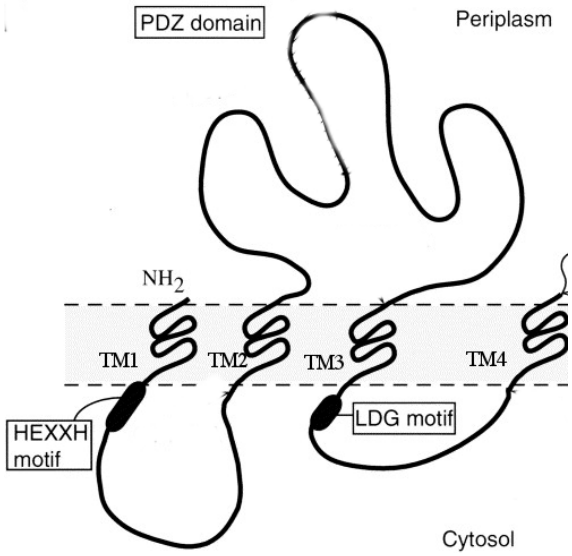


Figure 1.4 Topology of RseP in the cytoplasmic membrane RseP is shown by the thick line. The HExxH motif, the LDG motif and a PDZ-like domain and the transmembrane segments (TM1-4) are labelled. The middle dotted lines labelled the region called the cytoplasmic membrane (Kanehara et al., 2001).

Biochemical experiments provide insights into the active site of RseP, which was partially embedded into and stabilized by the lipid bilayer (Koide et al., 2007). The catalytic site of RseP is located within a pocket-like structure between the lipid and aqueous phases. The hydrophilic amino acid residues connect the RseP active site to the cytosolic side of the membrane which allows water molecules to access the active site (Koide et al., 2007). The major requirement for binding of substrates to RseP was the presence of helix-destabilizing residues in the

transmembrane region of substrates, which promote substrate degradation by activated RseP. Destabilization of a substrate helix may assist the substrate to enter into the RseP active site. RseP appears to have a broad substrate specificity with helix-destabilizing residues for specific substrate recognition (Koide et al., 2007).

1.9 Regulation of RseP activity

RseP has two activity PDZ domains with different roles. PDZ-1 is essential for the proteolytic activity and is involved in substrate recognition, whereas the PDZ-2 is required to maintain the stable structure in the periplasm (Inaba et al., 2008). PDZ domains have a substrate ligand binding groove that is formed by an α-helix and a β-strand. Previous study has found that a majority of mutations are mapped to the peptide-binding region of PDZ-1, suggesting a role in ligand binding (Inaba et al., 2008). Additionally, the PDZ domain regulates protease activity with conformational changes when peptide is bound to RseP and activated proteolysis (Inaba et al., 2008). In contrast, free RseP (without substrate peptide bind to its PDZ) had no degradation for an intra-membrane substrate, indicating the PDZ domain can inhibit the activity of RseP in the absence of substrate binding (Inaba et al., 2008; Li et al., 2009).

DegS provides the initial cleavage signal to RseP as the cleavage of substrate by RseP must be preceded by a prior cleavage mediated by DegS. Bohn et al., (2004) reported that RseP proteolysis is inhibited by a PDZ domain in the protein in the absence of proteolytic signals from site-1 protease DegS (Bohn et al., 2004). The mechanism by which these different factors act to restrain RseP is not fully understood. (Bohn et al., 2004).

Overall, the activation of RseP protease required initial signals from site-1 protease in the proteolytic cascade and the active site and PDZ domain play important roles in regulating the activity of RseP in cells. RseP protease had broad substrate specificity determined by helix-destabilizing residues, not only RseA-related sequence, but also target substrates (RseA, FpvR, FoxR and FiuR) in the CSS system in *P. aeruginosa*.

1.10 Aims of research

The overall aim of this study is to investigate RseP-dependent regulatory pathways in *P. aeruginosa*.

This first part of the study aimed to engineer *P. aeruginosa* to overexpress *rseP* and investigate the effects of overexpression, and absence of *rseP* on known RseP-dependent regulatory pathways. In addition, ClpXP protease is also mediated in *P. aeruginosa*. Experiments were also carried to investigate the effects of the absence of ClpXP protease.

The second part of the study aimed to identify additional possible RseP-dependent regulatory pathways by RNA-Seq analysis between wild-type, overexpressed RseP, RseP deleted strains and ClpP deleted strains. The last part of the study aimed to determine the correlation change in the expression of RseP-dependent genes (identified by RNA-Seq analysis) with changes in phenotype of *P. aeruginosa*.

Chapter 2 Materials and Methods

2.1 General

Centrifugation

All reactions were microcentrifuged using an Eppendorf microcentrifuge (Thermo Scientific).

Enzymes

All enzymes used for DNA based procedures such as restriction enzymes were sourced from Roche (Basel, Switzerland) and PCR enzymes were from Biolabs (New England).

L-arabinose

L-(+)-arabinose was supplied by Sigma and stored as a stock of 20 $\mu g.mL^{-1}$ and used at a final concentration of 0.5 $\mu g.mL^{-1}$ (0.5 %) in the experiment.

IPTG

Isopropyl-β-D-thiogalactopyranoside (IPTG) was supplied by Roche and stored as a stock of 24 $mg.mL^{-1}$ and used at a final concentration of 24 $\mu g.mL^{-1}$ on solid media.

BCIG

A stock solution of 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (BCIG) was supplied by Roche and kept in dimethyl formamide at 30 $mg.mL^{-1}$. The final concentration was 120 $\mu g.mL^{-1}$

Antibiotics

Antibiotic solutions were prepared with water and filter sterilized. Antibiotics were made up as stock solutions and stored at -20˚C. Gentamicin, ampicillin, tobramycin and ciprofloxacin were made in ddH_2O (double distilled water). Antibiotics used in this study are summarized in Table 2.1. Gentamicin was used for transformation, culture growth and experimental procedures for strain selection. Ampicillin was used for the selection of *E. coli* transformed with cloning on solid medium.

Antibiotic	Stock concentration	Final concentration	Strains
Tobramycin	40 mg.mL^{-1}	0.1 mg. mL^{-1}	*P. aeruginosa*
Ciprofloxacin	40 mg.mL^{-1}	0.1 mg. mL^{-1}	*P. aeruginosa*
Ampicillin	25 mg. mL^{-1}	0.1 mg. mL^{-1}	*E. coli*
Gentamycin	10 mg.mL^{-1}	0.024 mg.mL^{-1}	*P. aeruginosa*

Table 2.1 Antibiotics for use with bacterial strains

2.2 Media

Growth medium was prepared by using ddH_2O, adjusted to pH7.0 with NaOH as required and sterilized by autoclaving at 121˚C at 115 kilopascals pressure (15 pounds per square inch) for 20 minutes. Solid medium was prepared with the addition of 1.5% (w/v) agar.

Luria-Bertani (Lysogeny) broth (LB) (Sambrook and Russell, 2001)

This rich, iron replete growth medium was used for general bacterial cell culture.

10 g tryptone

5 g yeast extract

10 g NaCl

15 g agar (LB agar plates only)

1 L ddH_2O

Ingredients were dissolved in water and pH adjusted to 7.4 prior to autoclaving.

King's B (KB) broth and agar (King et al., 1954)

This is a rich, iron limited medium reported to be ideal for the growth of Pseudomonad species (King et al., 1954). The low levels of iron stimulate production of siderophores. King's B broth diluted one in ten was used for overnight bacterial growth.

20 g Bacto Peptone,

10 g K_2HPO_4,

10 mL glycerol

15 g of agar (KB agar plates).

After autoclaving, 6.1mL of 1 M $MgSO_4$ was added to 1 L broth to avoid precipitation of the magnesium.

Difco™ Mueller Hinton broth and agar (Haltiner et al., 1980)

This agar medium was used for antibiotic differentiation selection (MIC assay)

2.0 g Beef extract powder

17.5 g Acid Extract powder

1.5 g Starch

17.0 g Agar

2.3 Primers for PCR and RT-qPCR

Primer name	Sequence (5' -3')	Tm (˚C)
RseP wholeprimerF	GGGTCTAGAATCATCGAGGACGTGCTGAA	61
RseP wholeprimerR	CCCGAGCTCGGCGGCAGATTCGAATTCAA	60
RseP qPCR F	GGAAATGCTCCGCGAAGTGA	58
RseP qPCR R	GAAAATCCCCCACGCCAGAC	59
FpvA F	CCGCAACAACTACTATGCCC	57.5
FpvA R	TCGCCCCCTTGAGTACTTC	57
FoxA F	GCCCTACGTCAGCTACTCCGAG	61
FoxA R	GCCGACCGAGGTGTAGAAGTTGTC	61.3
FiuA F	TCACCCCCTTGAGTACTTC	56
FiuA R	CTGACGTCTTCGTCCACTGATG	57.4
PA5471 F	AAGCGCCTGGGCAATCTC	58.6
PA5471 R	CAGCCTGGTCGAACACCGT	60.2
PA2275 F	GAGCGCATCGTGGTCTCGG	61
PA2275 R	CGCCAGATAGGGATTGAGGTCG	59.3
MexX F	GGC CCT GGTCGCCCTATTC	60.8
MexX R	TCCTCGTACAGGCGACGG	59.3
PA3522 F	GGTGGTGCTGTTCCTGCGTA	60.1
PA3522 R	GAGAAAGGCCGAGGGGATGAA	59.1
OprL F	CCAACAGCGGTGCCGTTGA	61.3
OprL R	GCCATATTGTACTCGCGGGT	59
ClpX F	GTGGGCGAGGATGTCGAGAAC	60.1
ClpX R	CGGTACCCTCGATGAGCTTCAG	59.2
M13(-20) F	GTAAAACGACGGCCAG	55
M13(-20) R	CAGGAAACAGCTATGAC	54
pJN105 F	CGGATCCTACCTGACGCTTT	57
pJN105 R	TTGGGTAACGCCAGGGTTTT	57.4

Table 2.3 Oligonucleotide primers used in this study

All oligonucleotide primers used in this study were synthesized by IDT (Integrated DNA Technologies). The sequences of these primers are provided in Table 2.4. Each primer (20-30 nmoles) was resuspended in 150 µL of TE buffer (10mM Tris/ HCl (pH8.0) and 1mM EDTA). Working stocks were prepared to a concentration of 100 pmol.µL^{-1} and 10 pmol.µL^{-1}. The amplification product of RseP whole primers (Table 2.3) was around 1500 bp. The amplification products of other primer pairs were in the range of 200-300 bp with Tm (melting temperature) Min: 58˚C and Max: 62˚C, for RT-qPCR analysis. The sequencing primers were M13F&R, pJN105F&R for sequencing the cloning region of pLUG-Prime vector and pJN105 vector, respectively.

2.4 Bacterial cell strains and plasmids

2.4.1 Bacterial strains used

The strains of *E. coli* and *P. aeruginosa* used in this study with their corresponding genotypes are given in Table 2.4.1.

Strains	Genotype	Reference
E.coli JM83	Δ(lac-proAB) lacZΔM15	Messing et al., 1981
PAO1	Wild-type	Stover et al., 2000
PAORseP^{++}	PAO1*rseP*$^{++}$:Gm^{R}	This study
PAO*rseP*	PAO1Δ*rseP*	Draper et al., 2011
PAO(*clpP)*	PAO1(Δ*clpP)*	Lamont and Martin, unpublished
PAO1(pJN105)	Wild-type, Gm^{R}	This study
PAO*rseP–(pJN105)*	PAO1Δ*rseP*, Gm^{R}	This study
PAOrseP-(rseP^{++}*)*	PAO1Δ*rseP* with additional *rseP* plasmid (pNJ102)	This study

Table 2.4.1 Bacterial strains used in this study

2.4.2 Plasmids

The plasmids used in this study were isolated from laboratory glycerol stocks and are outlined in Table 2.4.2 with the description of the genotype of the plasmids.

Plasmid	Relevant Genotype and Phenotype	Reference
pLUG-Prime vector	*lacI*$^{\Phi}$*lacZ*(α-fragment),Amp^{R}	Lee et al., 2014
pLUG::*rseP*	1.5 kb XbaI-SacI PCR product containing the *rseP* genes from PAO1 cloned into pLUG-prime vector	This study
pJN105	araC-*pBAD* cassette cloned in pBBR1MCS-5,Gm^{R}	Newman and Fuqua, 1999
pJN105::*rseP*	1.5 kb XbaI-SacI PCR product containing the *rseP* genes from pLUG-prime vector cloned into pJN105 vector	This study
pJN106	*pBAD*::*lacZ* derivative of pJN105, Gm^{R}	Newman and Fuqua, 1999

Table 2.4.2 The plasmids used in this study

2.4.3 Growth and maintenance of bacteria

Bacterial strains were stored at -80°C in 80% (v/v) glycerol. King's B supplemented with antibiotics was used for culturing all *P. aeruginosa* strains. Culturing was performed in flasks with 10 mL of media. Cell density was estimated by measuring $OD_{600.}$ OD_{600} was recorded every 5-6 hours for each strain until the stationary phase.

2.5 DNA-Sequencing

DNA sequencing was sent to the Otago Genomics and Bioinformatics Facility with 150 ng.μL^{1} plasmid and 100 pmol.μL^{-1}of primers (e.g. M13F&R, pJN105F&R). Samples were submitted in 0.6 mL Eppendorf tubes. The sequence results were analysed by CLC Main-work Bench for determining the sequence quality.

2.6 General molecular biology methods

2.6.1 Purification of PCR product

PCR products were purified after PCR amplification and before DNA ligation. PCR products were purified using a high pure PCR product purification kit "Ultra-clean" (Roche). The procedure included mixing binding buffer with PCR solution and centrifuging three times and then centrifuged again with washing buffer before being eluted.

2.6.2 Genomic DNA isolation

Genomic DNA was isolated from 6 mL of an overnight bacterial culture using a UltraCleanTM DNA Isolation Kit (MO BIO) following the manufacturer's instructions. Cells were microcentrifuged at 10,000 rpm for 30 seconds and lysed by lysis solution. The genomic DNA was collected and stored at -20°C .

2.6.3 Plasmid DNA isolation

Plasmid DNA was isolated from a 4-6 mL overnight culture using a high pure Plasmid Isolation Kit (Roche). The pellet was resuspended in suspension buffer and RNase. The cells were then lysed by 250 μL lysis solution. The binding buffer (350 μL) was added and mixed by inversion. The supernatant was centrifuged at 14,000 rpm for 10 minutes then transferred to a high pure filter tube, followed by 1200 μL of column wash buffer, then transferred to a 1.5 mL microfuge tube and centrifuged twice. The column was transferred to a fresh tube and DNA was collected by the addition of 50 μL of elution buffer at full speed for 30 seconds.

2.6.4 Restriction endonuclease (RE) enzyme digests

Digest reactions were carried out according to the manufacturer's instructions in the recommended buffer. Reactions contained restriction enzymes at a maximum of 10% volume of appropriate restriction buffer and were made up to a final volume with sterile water. The standard reaction is profiled in Table 2.6.

For a ten μL reaction	Volume
Roche Buffer (10x)	1 μL
Deionized water	6 μL
DNA	2 μL
Enzyme	1 μL
Total	10 μL

Table 2.6 Restriction digestion protocol

2.6.5 Agarose gel electrophoresis

Reagents

- 5x sodium borate (SB) buffer stock:

  2 g NaOH pellets, pH adjusted to 8 with boric acid

  14 g Boric acid

  made up to 1 litre of buffer with ddH_2O

  To prepare 1xSB electrophoresis buffer (1L), ethidium bromide was added to a final concentration of 0.25 $mg.L^{-1}$

- 6x DNA loading buffer:

  30% (v/v) glycerol, 0.25% (w/v) bromophenol blue, 0.25% (w/v) xylene cyanol.

Agarose gels (0.85% (w/v)) were prepared in 1xSB buffer. The agarose was completely dissolved by heating in a microwave and then poured into a gel-forming apparatus. Once set, the gel was placed in the electrophoresis tank containing 1xSB buffer. DNA samples were prepared with 6x concentration of loading buffer and loaded with 2 μL 1 kb Plus DNA Ladder marker (Invitrogen) for size estimation of samples. Marker sizes (bp) were 12 000, 11 000, 10 000, 9 000, 8 000, 7 000, 6 000, 5 000, 4 000, 3 000, 2 000, 1 650, 1 000, 850, 650, 500, 400, 300, 200, 100. DNA was electrophoresed for 15 minutes at 200 v. The gel was exposed to UV light, and photographed using Bio-RAD Gel Doc™.

2.6.6 Gel extraction

DNA fragments produced by the action of DNA restriction enzymes on plasmids were separated according to size using agarose gel electrophoresis. Digested DNA samples were mixed with 6x DNA loading buffer at a ratio to achieve 1x loading buffer concentration and the samples were loaded onto 0.85% (w/v) agarose gels. Ethidium bromide stained DNA bands

were visualized using a table-top UV lamp and the DNA bands were excised using a sterile razor blade. The band of interest (at 1500 bp) was then processed using the QIAquick Gel Extraction Kit (QIAGEN).

2.6.7 Butanol purification

Butanol purification was carried out in order to purify DNA ligation reactions prior to electrotransformation. The ligated DNA was diluted to 50 µL Milli Q water and with 500 µL n-butanol. The reaction was microcentrifuged at 12,000 rpm for 5 minutes. After the supernatant was removed, the pellet was collected until it was completely dry (after 4 hours). The pellet was resuspended in 5 µL Milli Q water and electro-transformed to cells.

2.6.8 DNA quantification

Isolated genomic DNA, plasmid DNA and purified PCR were quantified using a Qubit Fluorimeter (Invitrogen) with the recommended buffer from kits.

2.6.9 General cloning methods

2.6.9.1 Phosphatasing vectors prior to ligation

The vector was treated with alkaline phosphatase to remove the 5' terminal phosphate group, which prevents the vector re-joining during ligation and the digested DNA mixed with 2 µL of 10x alkaline phosphatase buffer and 1 µL alkaline phosphatase (Boehringer Mannheim, USA). The mixture was incubated at 37˚C for 10 minutes and enzyme inactivated by heating for 2 minutes at 75˚C.

2.6.9.2 Ligation of DNA molecules

The 130 ng of PCR product, *rseP* (1500 bp) and 50 ng (2 µL) of pLUG prime were added to the ligation reaction for cloning the *rseP* insert into the pLUG prime vector (Intron, Korea).
The ligation reaction contained 50 ng of digested recipient vectors (pJN105) treated with appropriate restriction enzymes and then treated with alkaline phosphatase. Ligation was carried out by combining the pJN105 vector and target DNA (*rseP*) in a ratio of 1:3 mol (vector:target). The reaction contained 2 µL of 10x ligation buffer, 1 µL of T4 DNA ligase and made up the final volume of 20 µL with ddH_2O (Boehringer Mannheim, USA). Negative control tubes were set up with the vector, 1 µL buffer and 1 µL ligase but without insert genes.

All ligation reactions were incubated at 4˚C for 14 hours. The DNA ligation was followed by butanol purification (2.6.7) prior to transformation.

2.6.9.3 Transformation

E. coli

E. coli competent cells (JM83) stored as -80˚C laboratory stock and prepared by Lois Martin were used in transformation. Transformations were carried out by adding a 5 µL butanol-purified sample DNA (2.6.7) to 40 µL aliquots of JM83 *E. coli* cells, in 1.5 mL Eppendorf tubes. Negative control transformations were carried out with 40 µL *E. coli* cells with digested vector only. The mixtures were then applied to the Bio-Rad Gene Pulser with 1 pulse with each sample at 2.5 kv. Each transformation mixture was combined with 1 mL of LB and incubated for 2 hours at 37˚C on a shaking incubator. Aliquots (100 µL) of one tenth dilution from each transformation sample were spread onto LB-agar plates with appropriate selective antibiotics and incubated for 20 hours at 37˚C.

P. aeruginosa

Competent *P. aeruginosa* cells were prepared using the method described previously (ChuanChuen et al., 2002). A single bacterial *P. aeruginosa* colony was inoculated into 6 mL of LB broth and incubated for 18 hours at 37˚C. Aliquots 1 mL of overnight culture was harvested by spinning at maximum speed for 30 seconds. The supernatant was discarded and the pellet resuspended in 1 mL 300 mM sucrose solution sample three times in order to remove all the salts in samples. The cells were then ready for sucrose electrotransformation. Cells were transformed by adding 13.5 µL of 500 ng plasmid DNA to 100 µL aliquots of electrocomponent cells in 1.5 mL microcentrifuge tubes and transferring them into a 2 mm gap electroporation cuvette. The electroporation (2.5 kv) was carried out followed by the immediate addition of 1 mL of LB broth to the cells. Tubes were then incubated at 37˚C for 2 hours with shaking and one tenth dilution with (100 µL) of aliquots were spread onto LB agar plates containing the appropriate antibiotics.

2.6.9.4 Identification of recombinant clones

Ligated plasmid DNA (2.6.9.2) was transformed into *E. coli* and *P. aeruginosa* and spread onto specific selective agar plates. Bacterial colonies with pLUG-Prime insert recombinants showed white on BCIG/IPTG medium while colonies containing the vector with no insert were blue.

The recombinant colonies were screened using colony PCR (2.7.2) and plasmids were validated by restriction digestion (2.6.4).

2.7 Manipulation involving Polymerase Chain Reaction

2.7.1 PCR and Gradient PCR

All PCR reactions were performed using the BIO-RAD DNA Engine Peltier Thermal Cycler. Reactions were carried out in 20 μL volumes in 1.5 mL microcentrifuge tubes (20 μL), using of NEB 2x High-Fidelity PCR Master Mix and primers. Control reactions were always employed, in which the DNA template had been replaced with sterile ddH_2O. The temperature cycle of PCR and gradient PCR used in this study are shown below:

Basic PCR protocol

95˚C 5 minutes (initial denaturation step)

95˚C 10 seconds (denaturation) 58˚C 10 seconds (annealing) 72˚C 30 seconds (extension)	25 cycles

72˚C 7 minutes (final extension)

Then, held at 22˚C

Gradient PCR protocol

95˚C 5 minutes (initial denaturation step)

95˚C 10 seconds (denaturation) 55-65˚C 10 seconds (annealing) (Gradient temperature: 55-64˚C) 72˚C 30 seconds (extension)	25 cycles

72˚C 2 minutes (final extension)

Then, held at 22˚C

2.7.2 PCR screening of bacterial colonies

Bacterial colonies were inoculated up by tooth-pick from the agar plate and resuspended in 100 μL ddH_2O . The mixture was then incubated at 99˚C for 7 minutes and used as a template for PCR for screening with specific primers (amplified the gene of interest). PCR reactions used a NEB Next 2X PCR SuperMix which contained Taq polymerase, Taq DNA polymerase buffer and dNTP mixture. For each 10 μL reaction, the reagents were added to sterile 0.2 mL Eppendorf tubes: 5 μL of 2x SuperMix, 0.1 μL (150 pmol) of each primer stock solution, 0.8 μL of the colony template and ddH_2O to 10 μL. Negative control reactions were prepared containing identical components except that the template DNA was replaced with sterile ddH_2O. Template DNA was always added to the reaction last to avoid cross contamination. The temperature cycles used are shown below:

94˚C 5 minutes (denaturation)

94˚C 30 seconds (denaturation) 55˚C 30 seconds (annealing) 72˚C 30 seconds (extension)	25 cycles

72˚C 7 minutes (final extension)

Reactions were analysed by electrophoresis of a sample (2 μL) of each on an agarose gel (2.6.5).

2.8 Quantification of gene expression by Real-Time Polymerase Chain Reaction (RT-qPCR)

2.8.1 Induction of mRNA synthesis

The RNA samples were harvested from tubes containing cells with $1.3x10^9$ cells in 1.5 mL of bacterial cultures. (OD_{600} =1.6). The cells were then incubated on ice for 30 minutes to stop cell growth. A 2:1 ratio of RNA protection (QIAGEN): KB broth was mixed together in order to protect RNA samples for degradation (Assay Technologies, 2007). These mixtures were stored at -20˚C.

2.8.2 RNA extraction and DNase treatment

The protocol for RNA extraction using the GeneJet RNA purification kit (Thermo Scientific) is described in the manufacturer's manual. A volume of 0.5 mL bacterial culture mixtures (2.8.1) was used for RNA extractions. The amount of 0.4 mg of lysozyme (0.4 $mg.mL^{-1}$) was

needed to be fresh for each 1mL of RNA extraction. The extracted RNA (1 $\mu g.\mu L^{-1}$) was then estimated by the Nanodrop (2.8.3) and stored at -80˚C.

2.8.3 RNA quantification

The concentration of isolated RNA was quantified by Nanodrop (Thermo Scientific). The blank used was sample buffer and a 1 µl sample was used for analysis at 260 nm and 280 nm.

2.8.4 cDNA synthesis

cDNA conversion was carried out using qScript XLT cDNA SuperMix kit (Quanta) according to the manufacturer's instructions. First, RNA samples were treated with DNase I, (Quanta PerfeCTa DNase I) in order to eliminate genomic DNA from total RNA preparations for expression profiling by reverse transcription RT-qPCR. Two units per microliters (1 unit = 0.5 µL) of DNase I for 1 microgram of RNA. The proprietary reaction buffer and stop buffer was included to support a simple heat-kill step that permanently inactivates all trace levels of DNase activity. Complete inactivation of DNase I is critical before subsequent cDNA synthesis. The reaction mixture was then converted to cDNA by adding reverse transcriptase (+RT), which followed the protocol of qScript^{TM} XLT cDNA Supermix (with cDNA synthesis kits). Each RNA template (20 µL) was added with cDNA SuperMix (with reverse transcriptase +RT) in 4 µL and with a negative control (without adding reverse transcriptase (-RT)) respectively, for RT-qPCR analysis.

The DNase I provides all necessary components for first-strand synthesis including buffer, dNTPs, $MgCl_2$, primers, and RNase inhibitor protein. The unique blend of oligo and random primers in the qScript XLT cDNA SuperMix captures unbiased representation of all RNA sequences into the cDNA product. The reactions (with and without SuperMix) underwent PCR amplification with specific reaction protocol for 5 minutes at 25˚C, 60 minutes at 42˚C, (25 cycles) 5 minutes at 85˚C and held at 22˚C. One tenth of the first-strand product was used as a template for RT-qPCR amplification.

2.8.5 RT-PCR primer efficiency

Primer efficiency was calculated for each primer pair to validate them for use in RT-qPCR. Primer efficiency is a measure of the number of replications per PCR cycle and is calculated from the gradient of the line of best fit for a graph of crossing points vs log cDNA dilutions:

primer efficiency = $10^{(-1/gradient)}$. The maximum efficiency is two. Differences in primer efficiency can arise due to off target binding, low template binding efficiency or primer dimerization. A 10% variation from a primer efficiency of two (1.8-2.2) was considered acceptable for this study as primer efficiency was accounted for when carrying out relative quantification. To generate the efficiency of primers, a dilution series of $2x10^{-2}$, $1x10^{-2}$, $5x10^{-3}$, $1x10^{-3}$, $5x10^{-4}$, $2x10^{-4}$ ng.μL^{-1} of DNA was performed in duplicate to produce a standard curve (3.5).

2.8.6 RT-qPCR quantification

RT-PCR analyses were performed on a Roche Lightcycler ^{R} 480 (LC480, Platform GPL 8887) 96-well plates Real Time PCR instrument (Roche Applied science, Auckland) using Gene Scanning Software (Roche). After the primer efficiency had been calculated (2.8.5), the primers and SYBR green Master Mix (Invitrogen, USA) were combined and pipetted into the wells of a 96-well plate: 1 μL of 10 pM forward primers and 1 μl of 10 pM reverse primer for each qPCR primer (Table 2.3) and 5μL of SYBR Green Master Mix (contained Taq DNA polymerases and buffers). Next, 3 μL of 150 ng.μL^{-1} (1/10 dilution of 1500 ng) cDNA (with and without reverse transcriptase) mixed with 1.8 μL of nuclease-free water was added to the wells. The reactions without reverse transcriptase (-RT) of cDNA template were the negative control of the experiment. Each sample was assayed in quadruplicate and the experiments were in duplicate. The plates were centrifuged at 2,000 rpm for 2 minutes to ensure consistency of the reaction before reactions were carried out in the lightcycler for quantification. The PCR protocol included a cycle of denaturing/pre-incubation at 95˚C for 10 minutes, 40 cycles of denaturing at 95˚C for 15 seconds, annealing and extension at 60˚C for 1 minute and finally a cycle of extension/cooling at 40˚C for 30 seconds.

2.9 β-galactosidase assay (Miller, 1972)

pJN106 plasmid had reporter gene *lacZ* instead of gene of interest. *lacZ* is encoded to beta-galactosidase (β-gal). β-gal is assayed using an analogue of lactose, ONPG which is broken down by β-galactosidase to a yellow coloured product O – nitrophenol. The *lacZ* expression can be estimated from β-gal activity.

Reagents

- z buffer: 60 mM of Na_2HPO_4, 40 mM of NaH_2PO_4, 10 mM of KCl,1 mM of $MgSO_4$ and 50 mM of β-Gal were mixed in ddH_2O. The buffer was stored at 4˚C on the same day of assay
- ONPG: A 4 $mg.mL^{-1}$ solution of o-nitrophenyl-β-D-galactopyranoside (ONPG) was freshly prepared in phosphate buffer (34 mM K_2HPO_4, 22 mM KH_2PO_4 pH 7.5)

β-galactosidase assays were performed according the method of Miler, (1972). Overnight bacterial cultures were diluted 1:100 into King's B medium containing the appropriate supplements and were grown for 27 hours at 18˚C. The cells were chilled on ice for 20 minutes and the final OD_{600} of each culture (1.0) was recorded. Aliquots (50 µL) to 1.5 microcentrifuge tubes and Z buffer (950 µL) were added to a final volume of 1 mL. SDS (10 µL of 0.1%) and chloroform (20 µL) were added to each tube and these were then vortex-mixed for 15 seconds to permeabilise the cells. ONPG (200 µL) was added to each sample and each tube was mixed and incubated at 37˚C until a yellow coloration had developed. The reaction was stopped by the addition of 250 µL of 2 M Na_2CO_3 and the time noted. The cell debris was removed by centrifuging for 5 minutes at maximum speed. The A_{420} of the supernatant was determined and the units of β-galactosidase were calculated using the following equation (Miller, 1972):

$$\text{Units} = 1000 \times (A_{420}) / (t \times v \times OD_{600})$$

where A_{420} is the absorbance of the reaction mixture at 420 nm, OD_{600} indicates the bacterial culture cell density at 600 nm, "t" is the time (minutes) of ONPG incubation period and "v" is the volume of culture used in the assay (0.05 mL). Each sample was assayed in triplicate for each volume used.

2.10 Minimum inhibitory concentration (MIC) test

Media: Difco Mueller Hinton Broth (MH) (Becton and Dickson, France)

20 g Beef extract powder

17.5 g of Acid digest of Casein

1.5 g of solution starch

Agar: Mueller Hindon (MH) agar (38 g in 1 L of ddH_2O)

The agar concentration: 0.04% MH

Four to five bacterial colonies of each strain were inoculated by toothpicks of each strain and grown in Mueller Hinton broth for 24 hours at 22˚C. The bacterial culture was then diluted to $1x10^6$ cells/mL with 5 mL of MH broth for MIC plates assay. The control plate was carried out with strains in various cells contents (10^4,10^3,10^2cells/mL) without any antibiotics involved. Minimum inhibitory concentrations (MICs) for tobramycin (10 $\mu g.mL^{-1}$) and ciprofloxacin (10 $\mu g.mL^{-1}$) were determined by serial twofold dilution in MH broth by using dilution series (0.06-2 $\mu g.mL^{-1}$ per 25 mL MH agar plates). Results were determined after incubation at 22˚C for 48 hours.

2.11 Swimming and swarming motility assay

M8-agar (1 L batch) (Ha et al., 2014)

200 ml of M8-supplement solution (64 g Na_2HPO_4, 15 g KH_2PO_4, 2.5 g NaCl, ddH_2O bring the final volume to 1 L)

1 mL $MgSO_4$ (1 M)

10 mL glucose (0.2 $\mu g.mL^{-1}$)

25 mL casamino acid (0.5 $\mu g.mL^{-1}$)

3 g agar with 800 mL of ddH_2O (swimming plates)

6.5 g agar with 800 mL of ddH_2O (swarming plates)

0.3% LB agar

10 g peptone

5 g Yeast extract

10 g NaCl

3 g agar

Modified LB broth was used for growth adjustment of *P. aeruginosa* strains in the swimming motility assay in the study. The final LB adjusted agar concentration was 0.3%

Swimming and swarming motility assay:
P. aeruginosa uses flagella to swarm across a semi-solid surface and swim in a liquid or low-viscosity condition. Swarming motility was carried out on semi-solid agar surfaces (0.65% agar) in the presence of specific carbon and nutritional sources for the motility assay. Swimming motility occur on inner layer of 0.3% agar. *P. aeruginosa* strains were grown in LB from fresh streak plates for 40 hours at 18˚C before inoculation. The swimming M8-plates were solidified at room temperature for three hours whereas the swarming plates were dried under the laminar flow for 15 minutes after adding all the ingredients. Two bacteria colonies were inoculated by toothpicks for each strain and grown in LB broth for 40 hours at 18˚C. Bacterial cultures were then stabbed into the agar layer for the swimming motility assay and spotted on the surface centre (with 2.5 μL per inoculum) for the swarming assay (Ha et al., 2014). Swimming plates were incubated at 22˚C for 48 hours; swarming plates were incubated at 37˚C for 24 hours.

2.12 Whole Genome Sequencing (WGS) analysis

Genomic DNA samples were sent to New Zealand Genomics Limited (NZGL) for sequencing and raw sequence files were obtained from NZGL. The raw reads were then mapped to the reference genome (PAO1_IL_PAO1_107 (IL_PAO1) downloaded from Genbank by using Galaxy online alignment tools. The output files (bam files) were obtained from NZGL. NZGL job number, NZGL010191E (sample 14). The output results (bam files) were visualized by Integrative Genomics Viewer (IGV) which showed the overview coverage of nucleotide base of sequence throughout libraries (which mapped to IL_PAO1_ with identified genes ID).

2.13 RNA-Sequencing analysis

Twelve RNA samples were sent to NZGL for library preparation and sequencing. Samples were PAO1-WT, PAO1RseP^{++}, PAO *rseP* and PAO(*clpP*). DNA libraries were prepared from the RNA samples using the Illumina TruSeq^{TM} RNA library preparation. From this, 125 base pair (bp) paired-end reads were obtained. Raw sequence files following post-processing of raw reads included trimming and alignment to the reference genome PAO1-WT (NC_002516.2)

by NZGL. Each gene base on PAO1-WT was quantified in the sequence files, which were obtained from NZGL. (NZGL job number, NZGL010191E (sample 1-12)

Quantified transcripts were compared to each sample strain to identify differential expression of genes. The analysis was done using DESeq and EdgeR (R-Bioconductor packages) with R-studio (Anders and Huber, 2012). Comparisons were tested for significantly different expression between strains. A false discovery adjusted p values was used to determine the significance of the result. Any genes without a p-adjusted values ≤0.05 the significance from the final result.

2.14 Specialist database used in this study

- *P. aeruginosa* PAO1 genomic sequence information was obtained from the Pseudomonas Genome Database (available at: www.pseudomonas.com) .
- DNA sequences were analysed and compared using CLC Main-work Bench 7 and the online BLAST (Basic Local Alignment Search Tools) (available at: http://blast.ncbi.nlm.nih.gov/Blast.cgi). CLC Main-work Bench was also used to design primers for cloning.

Chapter 3 Results

3.1 Research strategy of this study

This study aimed to identify RseP-dependent genes and pathways in *P. aeruginosa*. Firstly, *P. aeruginosa* was engineered to overexpress RseP protease, by cloning into a plasmid (pJN105) for constructing PAORseP^{++}. The pJN105 expression vector is designed to facilitate arabinose responsive control of transcription of the insert gene and it consisted of a P_{BAD} promoter which is a strong inducible promoter in the presence of arabinose. This allow RNA polymerase to bind and recognize the sequence and transcribe *rseP*. Secondly, RT-qPCR was employed to quantify the expression level of siderophore receptor genes in CSS systems between PAO1-WT, PAORseP^{++}, PAO *rseP* strains. It was hypothesized that overproduced RseP may enhance the degradation of anti-sigma factors, which may result in the increased activity of receptor genes in *P. aeruginosa* and the results could be quantified by RT-qPCR. Furthermore, RNA-Seq was used to identify RseP-mediated genes by differential expression analysis between all samples including PAO1-WT, PAORseP^{++}, PAO1 *rseP* and PAO1(*clpP*). The aim of the last part of the study was to investigate the correlation change in the expression of selected RseP-dependent genes with changes in phenotype of *P. aeruginosa*. The minimal inhibitory concentration (MIC) test and motility assay were carried out for characterization of phenotypic changes, in terms of antibiotic resistance and motility of *P. aeruginosa,* respectively.

3.2 β-galactosidase activity of pJN105 expression induced by arabinose

3.2.1 Introduction

The pJN106 had similar genomic features to the pJN105 plasmid with the same P_{BAD} promoter and *araC* operon. L-arabinose is required for both plasmids in order to control the transcription of target genes. In the pJN106 plasmid, the *lacZ* gene which encodes β-galactosidase (β-gal) is involved in the breakdown of lactose to glucose and galactose. The β-gal assay involves o-nitrophenyl-β-galactosidase (ONPG) which is hydrolysed by β-gal to a yellow coloured product and the β-gal expression can be determined by quantifying the enzyme activity of ONPG (Miller, 1972). The β-gal assay was carried out to ensure the arabinose was sufficient to induce the transcription of *lacZ* prior induced the target gene in pJN105 plasmid (3.5.1). In the presence of arabinose, it can directly bind to the *araC* site activating the P_{BAD} promoter and

allowing for RNA polymerase recognition. The activated promoter initiates the transcription of the *lacZ* gene (Figure 3.2.1A) (Newman and Fuqua, 1999). In the absence of arabinose, the P_{BAD} promoter was repressed by a AraC-dependent mechanism which forms a DNA loop and prevents RNA polymerase from binding to the promoter. Thus the promoter was in the inactive form and it was not able to initiate the expression of *lacZ* (Figure 3.2.1B) (Newman and Fuqua, 1999).

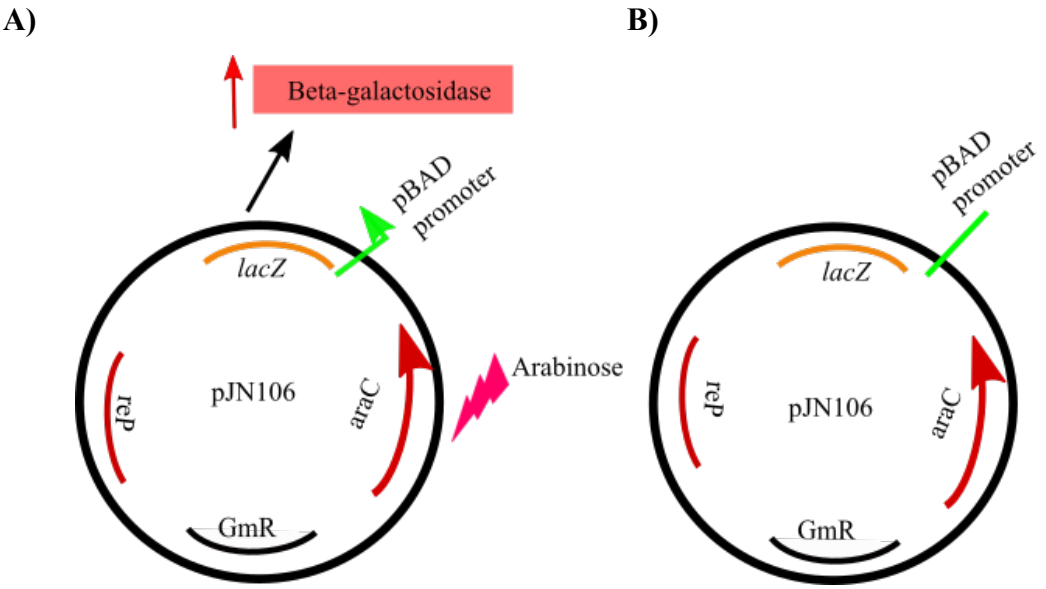


Figure 3.2.1 The map of pJN106 plasmid A) Arabinose binds to the *araC* site directly and activates the P_{BAD} promoter. The expression of β-galactosidase is highly transcribed in the cells. **B)** In the absence of arabinose, P_{BAD} promoter was inactive thus preventing RNA polymerase from binding to the promoter and repressing the expression of *lacZ*. The pJN106 plasmid included the Gentamicin resistance marker (Gm^R), *araC* site (red arc), *reP site (red* arc), *pBAD* (green arrow). The *lacZ* gene was labelled as orange (Newman and Fuqua, 1999).

3.2.2 Preparation of β-gal assay

The strains (PAOpJN106::*lacZ*) were incubated at 18˚C for 27 hours in LB broth before assay (2.4.3). The strains were grown at 18˚C to ensure arabinose was sufficient to transcribe the target gene in pJN105 plasmid with the growth condition to be used for bacterial sample preparation for RNA-Seq analysis (3.5). Arabinose with a final concentration of 10 µg. mL^{-1} was added to cultures when cells entered the growth phase (OD_{600} 0.3). During the growth phase, the gene of interest can be highly transcribed after adding arabinose when the cells are in fast replication.

3.2.3 Results of β-gal activity

The β-gal assay measured the activity of β-gal in both conditions, with and without arabinose induction (Figure 3.2.2). In the absence of arabinose, the expression was low with 300 enzyme units, indicating weak transcription compared to the strains with arabinose induction (Figure 3.2.2). By adding arabinose, a 47-fold increase in expression was consistently present in PAOpJN106 triplicates with a small error bar (Figure 3.2.2). This suggested that arabinose can successfully induce the transcription of *lacZ* gene in pJN106. Thus, the same batch of arabinose was used for further experimentation, so that we can be confident that the arabinose would be able to induce the transcription of target genes in pJN105 plasmid in subsequent experiments (3.5.1).

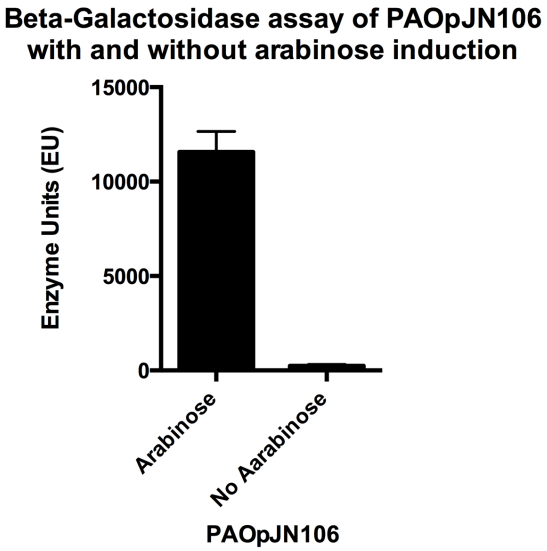


Figure 3.2.2 β-galactosidase assay of PAOpJN106 The assay carried out on the pJN106 strains with and without arabinose induction and the y-axis represented the enzyme activity (EU) of β-galactosidase. Data are representative of the average of three biological replicates ± standard error of the mean.

3.3 Construction of strain PAORseP^{++}

3.3.1 The aim of DNA cloning

The aim of this section was to construct a strain of *P. aeruginosa* where *rseP* was overexpressed, in order to identify RseP-dependent genes by comparing differential expression with other strains such as PAO1-WT and PAO *rseP* in RNA-Seq analysis. The *rseP* sequence (encoding the RseP protease) from *P. aeruginosa* strains was amplified and ligated into an expression vector (pJN105) in order to construct overexpressed strains in *P. aeruginosa*. The new construct was called PAOpJN105::*rseP*$^{++}$ (PAORseP^{++})

3.3.2 The plasmid of pLUG-Prime vector and pJN105 vector

The *rseP* gene was cloned into the pJN105 expression vector in a two-step cloning process. The *rseP* insert gene was first ligated into pLUG-Prime vector (pLUG-Prime^{R} TA-cloning manufacture instruction, Intron Biotechnology) and the *rseP* was digested by restriction enzymes and further cloned into the pJN105 expression vector. The reason why the *rseP* gene was cloned in two steps instead of clone into the pJN105 vector in one step was to ensure the integrity of *rseP* gene amplification and to further validate whether the *rseP* gene was able to be digested by restriction enzymes (RE) e.g. XbaI and SacI (the two RE found to have recognition sites on both of the vectors) in the correct location within the pLUG-Prime vector.

pLUG-Prime vector contains an ampicillin resistance marker (Amp^{R}), and *lacZ* genes encoding β-galactosidase allow for blue/white recombination selection. The vector was 2728 bp in size. For cloning ligation, DNA polymerase preferentially had a "3' end A" overhang to PCR products. The pLUG-Prime vector allowed direct insertion of such PCR products into this prelinearlized cloning vector, which has a T-overhang on each 3' end (Figure 3.3.1 left) (2.6.9.2). This eliminated the need for restriction enzyme digestion of the vector, and the insert sequence was cloned with the built-in restriction sites, resulting in a much more efficient and robust cloning procedure (Intron, Korea).

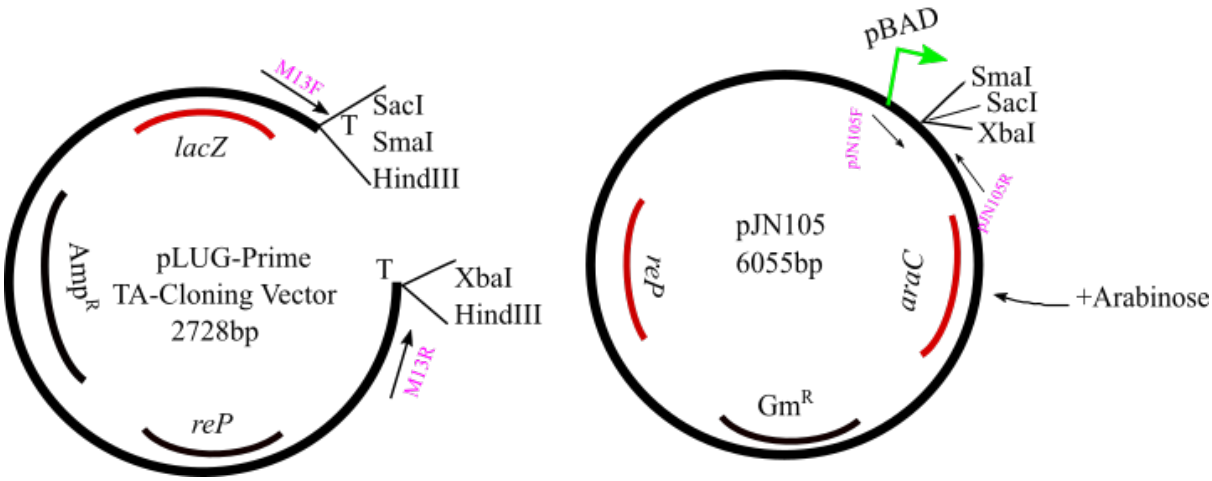


Figure 3.3.1 Plasmid maps of pLUG Prime TA (left) and pJN105 cloning vector (right) (without insert) The bold arcs above indicate the relative position of relevant genetic elements. The black bold arcs indicate the antibiotics resistance marker, ampicillin resistance (Amp^R) marker in pLUG-prime, and gentamicin selectable marker (Gm^R), replication initiator (*reP*) and *araC* shown in red. The pLUG Prime vector consisted of a *lacZ* gene that encodes the protein beta-galactosidase and *reP* genes of replication protein. The restriction enzymes used in the cloning steps were labelled (SacI, SmaI, HindIII, XbaI) in multiple cloning sites (MCS). The pJN105 vector includes trans-acting replication protein and the P_{BAD} promoter show by the green arrow is located close to cloning sites and the promoter can be targeted by arabinose induction.

The gene of interest in the pJN105 expression vector is induced by arabinose and its P_{BAD} promoter is the strongest induced promoter in the presence of L-arabinose. In the presence of arabinose, this allows RNA polymerase to bind and transcribe the *araC* operon and then initiate transcription of RseP. When arabinose is absent, the P_{BAD} promoter interacts with *araC* sites and forms a DNA loop located close to the promoter which prevents RNA polymerase from binding to the promoter of the *araC* operon and reduces the transcription rate, thus repressing P_{BAD} promoter activity (Newman and Fuqua, 1999). The pJN105 plasmid contains a selectable gentamycin marker (Gm^R) which allows for antibiotic selection (Figure 3.3.1 right). The pJN105 expression vector with the *rseP* inserted sequence was expressed from P_{BAD} promoter for an RseP overexpression construct (Figure 3.3.1 right).

3.3.3 Cloned *rseP* sequence into the pJN105 vector

i) Cloned *rseP* into the pLUG-Prime vector

To construct the RseP overexpressing plasmid, insertion of the *rseP* sequence was necessary. Oligonucleotide primers were designed (2.3) with an XbaI insertion site on the 5' end and a SacI site on the 3'end of the *rseP* sequence to amplify from the *P. aeruginosa* strain (PAO1-WT) by PCR (2.7.1). The XbaI and SacI restriction sites were chosen as these sites were present in both the pLUG-Prime and pJN105 vectors. The forward primers were located at 128 bp upstream of *rseP*, and the reverse primers were located 69 bp downstream of the *rseP* sequence. Primers were designed close to the *rseP* sequence in order to prevent amplifying the irrelevant sequence. These primers allowed amplification of the full length of the *rseP* gene to give a 1524 bp PCR amplification product. From Figure 3.3.2A, a 1500 bp fragment corresponding to the *rseP* PCR product can be identified and the fragment was extracted by agarose gel purification followed the method in 2.6.6. The purified *rseP* was then ligated to the pLUG-Prime TA cloning vector, which is a pre-linearized cloning vector and allows the direct insertion of the PCR product (Figure 3.3.2) (2.6.9.2). *E. coli* cells (JM83) were transformed with the ligation mixture and blue/white selection was performed (2.6.9.4). Bacteria were spread onto the LB-agar containing Amp, BCIG and IPTG. The transformants containing recombinant plasmids were selected as white colonies. The *rseP* was ligated into the plasmid vector which disrupts the *lacZα* and therefore the β-galactosidase would not be produced, resulting in a white colour.

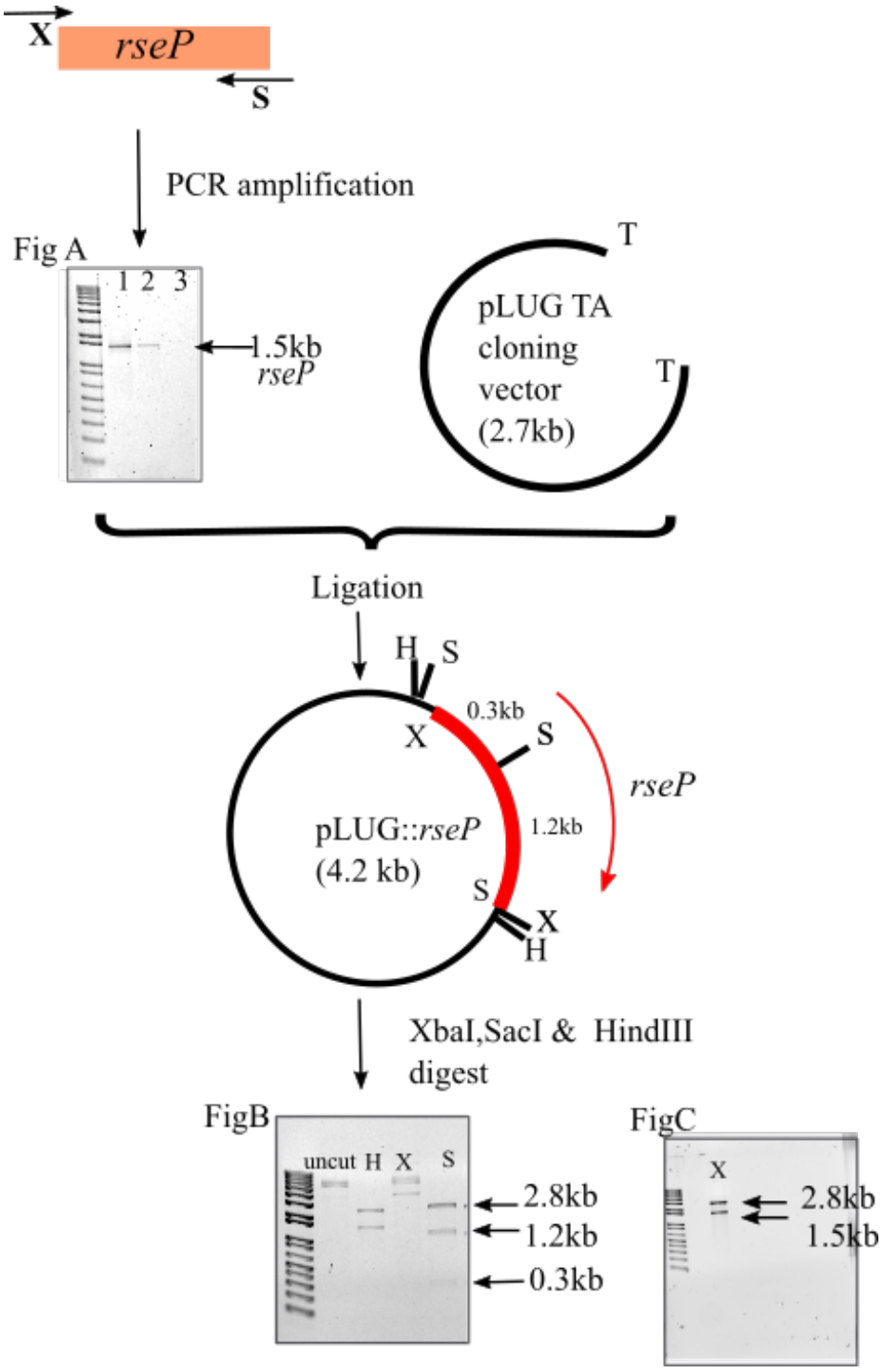


Figure 3.3.2 Construction of the pLUG::*rseP* construct A work flow from the *rseP* PCR product amplification with the primers designed with built-in restriction sites XbaI (X) and SacI (S) from genomic gDNA PAO1-WT through to the construction of pLUG::*rseP*. **Gel Figure A:** The PCR amplification from genomic PAO1 to obtain the *rseP* fragment (1524 bp) L1-2: *rseP* PCR product with the PAO1-WT gDNA template, L3: No template control. **Gel Figure B:** The restriction digest of pLUG::*rseP* to confirm the presence of cloned *rseP*. Restriction enzymes include HindIII (H), XbaI (X) and SacI (S).

ii) Results of restriction digestion of pLUG::*rseP*

Eighty white colonies were obtained, one of which was selected for plasmid purification (2.6.3) and the plasmid was validated by restriction digests using HindIII, XbaI and SacI enzymes (2.6.4), as shown on the gel in Figure 3.3.2B.

According to the Figure 3.3.2 B, the top bands of the gel (digested by HindIII) represent the pLUG-Primer vector which was 2800 bp. The insert gene was a 1600 bp fragment (Figure 3.3.2B) and shown in the lane which was digested by both HindIII and SacI. This indicated that the *rseP* insert sequence had been successfully ligated to the plasmid. However, at this step we realized that the SacI recognition site was also present in the middle of the *rseP* insert sequence, because a third band was obtained when partially digested with SacI (300 bp) (Figure3.3.2B).

Additionally, XbaI did not fully digest the insert sequence in Figure 3.3.2B, which had only one band (4000 bp) located near the top of the gel. It was found that the insert gene was fully digested from the vector by XbaI (Figure 3.3.2C) second time. From the restriction digestion validation, *rseP* was successfully ligated into the pLUG vector and the plasmid was sent for DNA sequencing with M13 universal primers sets (which allowed the entire cloned fragment to be sequenced) (Figure 3.3.1 left), which confirmed that no mutation had occurred.

iii) Cloned *rseP* into the pJN105 expression vector

The RseP fragment was then excised from the pLUG::*rse*P by agarose gel extraction with XbaI restriction enzyme (2.6.4). There were two XbaI recognition sites on pLUG::*rseP* (Figure 3.3.2) and therefore, according to the Figure 3.3.3 A, the top band was the pLUG-Prime vector with approximately 3000 bp and the second band (1600 bp) was the *rseP* sequence which was digested by XbaI. The pJN105 plasmid was initially digested with XbaI enzyme to linearize the vector. The linearized pJN105 plasmid was further treated with alkaline phosphatase to prevent the vector from re-ligating after digestion and allowed *rseP* insert genes to be ligated to vector (2.6.10.1) Full-length *rseP* was digested out of the plasmid and purified by gel extraction (2.6.6) as shown in Figure 3.3.3 C, from the excised second band. After the gel purification, 5 $ng.\mu L^{-1}$ of the purified donor DNA (150 ng) was ligated into the recipient pJN105 (50 ng) expression vector (Figure 3.3.3A) (2.6.9.2).

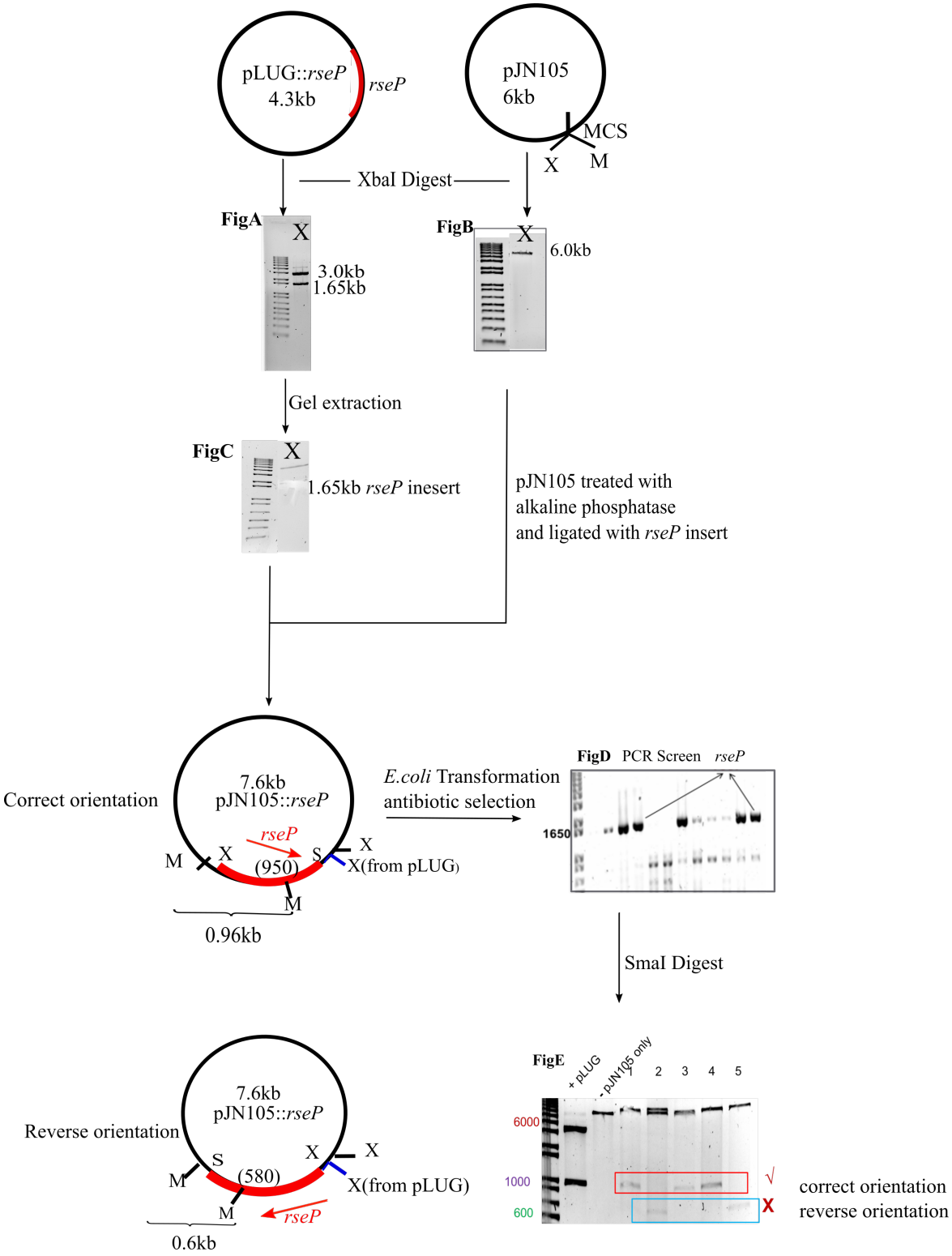


Figure 3.3.3 Construction of the final pJN105::*rseP*$^{++}$overexpressed strain The *rseP* sequence was excised from the pLUG Prime vector through gel extraction and was ligated into pJN105 to form the final pJN105::*rseP* construct (7.6 kb). The multiple cloning sites on the final construct (correct orientation) include SmaI (M), XbaI (X) on pJN105 vector with the correct position labelled. The XbaI sites in blue bold lines indicate the cuts from pLUG-prime vector which is 20 bp away from the *rseP* insert. XbaI and SacI labelled inward of the vector indicated the recognition sites on RseP primers (SmaI was labelled at 950 bp of insert sequence).

Gel Figure A: The pLUG::*rseP* plasmid was digested by XbaI (X) and the insert genes (1.65 kb) were excised off through gel extraction (**Gel Figure C**). **Gel figure B:** The pJN105 vector had first been fully digested by XbaI. **Gel Figure D**: Colony PCR screen of *E. coli* transformants, primers specific (pJN105 primers located at MCS sites) to the *rseP* gene were used to confirm the presence of the insert. **Gel Figure E:** The restriction digestion and orientation validation by SmaI (M). First two lanes, positive (pLUG-prime vector with *rseP* genes) and negative control (vector only). Lane 1-5, plasmids selected from PCR screening. Expected size 1 kb for correction orientation shown by red box and red tick and reverse orientation of insert shown in the blue box with 0.6 kb.

3.3.4 pJN105::*rseP* cloning selection and transformation

The purified *rseP* sequence was ligated to the pJN105 recipient vector and the ligation mixture was transformed into the JM83 *E. coli* cells (2.6.9.3). Since *E. coli* is commonly used for transformation, the plasmid was first transformed in *E. coli* to check whether the construct was integrated with the correct orientation, then further transformed into *P. aeruginosa* cells. Transformed bacteria were grown on the LB agar plates with gentamicin for the pJN105 plasmid selection. The transformants containing the recombinant plasmids were identified by colony PCR as shown in Figure 3.3.3.E (Colony PCR followed by method 2.7.2). Four colonies had a 1650 bp band on the gel (Figure 3.3.3E) and this was indicated as the *rseP* insert sequence. From the pJN105::*rseP*$^{++}$ final construct, the complete insert genes were digested from the pLUG:*rseP* and the result was 1524 bp with an additional 20 bp which is the distance between the XbaI sites to the end of *rseP* sequence (Figure 3.3.3).

In the final construct (pJN105::*rseP*$^{++}$) another restriction site, SmaI (M), was located at 950 bp of the insert sequence (when cloned in the desired orientation) with a second SmaI recognition site located at the pJN105 vector. Thus, the plasmid from PCR screening was digested by the SmaI enzyme resulting in a 900 bp fragment and the rest of the vector which was a 6000 bp fragment, thereby confirming of the correct orientation of the insert genes within the plasmid (Figure 3.3.3). According to Figure 3.3.3 E, plasmid 1, 3 and 4 had bands around 900 bp which indicated the partial fragment of the *rseP* inert gene and the *rseP* cloned in a correct orientation (forward orientation) whereas plasmid 2 and 5 had bands around 600 bp indicating the reverse orientation of *rseP* shown in Figure 3.3.3 (reverse orientation). One of the plasmids with the correct insert orientation was then transformed into *P. aeruginosa* (2.6.9.3) to construct the final RseP overexpressed strains: PAO(pJN105::*rseP*$^{++}$) (Figure 3.3.3E).

3.3.5 DNA sequencing to confirm cloning orientation

The selected pJN105 was sent for DNA sequencing with pJN105 primers (Table 2.3) which had been designed closer to the MCS along the pJN105 vector to allow sequencing of the entire inserted *rseP* fragment in the pJN105 vector. Forward and Reverse regions were 2000 bp in total and covered the whole *rseP* sequence (1524 bp) with no mutation occurring with a 99% identity mapped to PAO1-WT strains in a database (Appendix 2). Thus, we had successfully constructed the RseP overexpressed strains.

3.4 Sequence analysis of the intended PAO(*clpP*) strain

3.4.1 Introduction

ClpXP is a cytoplasmic protease which provides the third cleavage of substrate in the proteolytic cascade and it is also responsible for cleaving the remaining fragments in the cytoplasmic membrane of *P. aeruginosa* (Flynn et al., 2003,). No degradation of substrates occur as the ClpXP protease would not able to function if *clpP* genes had been knocked off (Baker and Sauer, 2012). Thus, the the PAO (*clpP*) mutant strain was to be used for determining gene expression levels when ClpXP protease was not functioning in cells. This section involved the process of constructing intended PAO(*clpP*) and analysis of its sequence by Whole Genome Sequencing (WGS).

3.4.2 Construction of intended PAO(*clpP*) strain

An intended PAO(*clpP*) had been generated in the laboratory (Lamont and Martin, unpublished). The steps of constructing the intended PAO(*clpP*) strain are shown in Figure 3.4.1 (Lamont and Martin, unpublished). There are two genes (*tig* and *clpX*) located upstream and downstream of *clpP*. Firstly, two sequences of *tig*-*clpP* and *clpP*-*clpX* were amplified from *P. aeruginosa* PAO1 by specific primers and ligated to the pEX18Gm plasmid without *clpP* genes. The plasmid was then transformed into *E. coli* cells for cloning validation followed by restriction digestion. The successful clone was transformed to *P. aeruginosa* cells through conjugation. These processes assisted the transference of the plasmid DNA in *E. coli* to the chromosome of *P. aeruginosa* by a bridge-like connection between two cells (Figure 3.4.1). Therefore, the *clpP* gene would be completely knocked off following by these steps. However, from the DNA sequencing results, the *clpP* gene somehow remained in the strain (3.4.3).

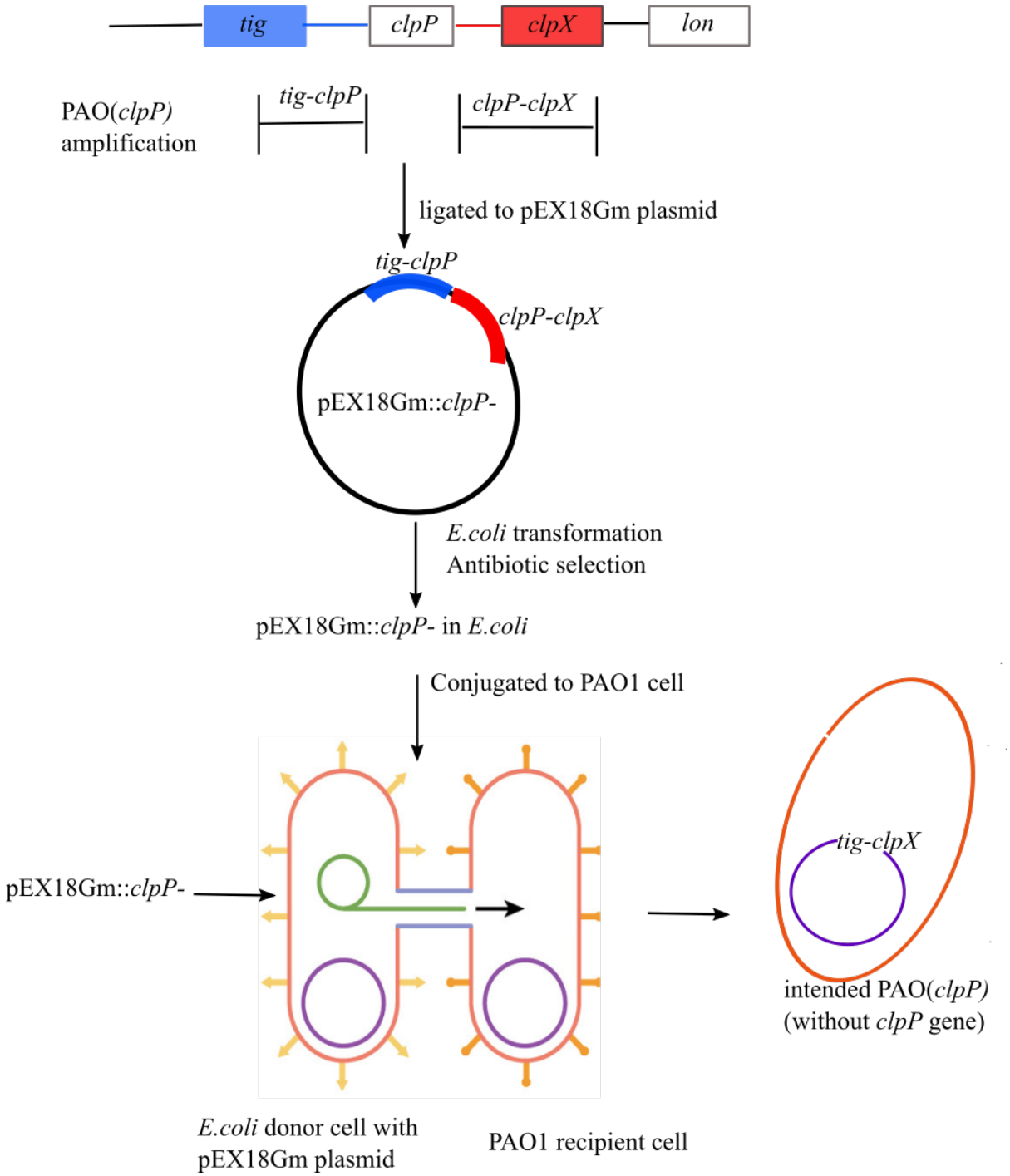


Figure 3.4.1 Construction of the intended PAO(*clpP*) Two designed primer sets amplified the region of *tig-clpP*, and *clpP-clpX* separately, leaving out the *clpP* entire sequence. The amplified sequence *tig-clpP* was first cloned to the pEX18Gm plasmid and then ligated the *clpP-clpX* to the plasmid to form the final construct, pEX18Gm::*clpP*-. Thus, two separate fragments were cloned into a plasmid (blue indicated the *tig-clpP* fragment and red indicated the *clpP-clpX* fragment). The plasmid was then transformed to the *E. coli* cell line and underwent antibiotic selection and the cloning result was verified by restriction digestion appropriate restriction enzymes. The plasmid (shown in green) was transformed to *P. aeruginosa* cells by conjugation. The intended PAO1(*clpP*) strain were constructed by these steps (Lamont and Martin, unpublished).

3.4.3 Analysis of PAO(*clpP*) by Whole Genome Sequencing

Whole Genome Sequencing (WGS) was carried out to identify the accuracy of the intended PAO(c*lpP*) strain and to ensure the *clpP* gene had been deleted with no mutation occurring in this strain. The process of WGS was performed by NZGL. The WGS reads were aligned to the reference genome (*P. aeruginosa* PAO1) by Galaxy alignment tools and the output results (bam files) were obtained by NZGL (2.12). The alignment output across the entire genome was visualized using the Integrative Genomics Viewer (IGV) (Figure 3.4.2).

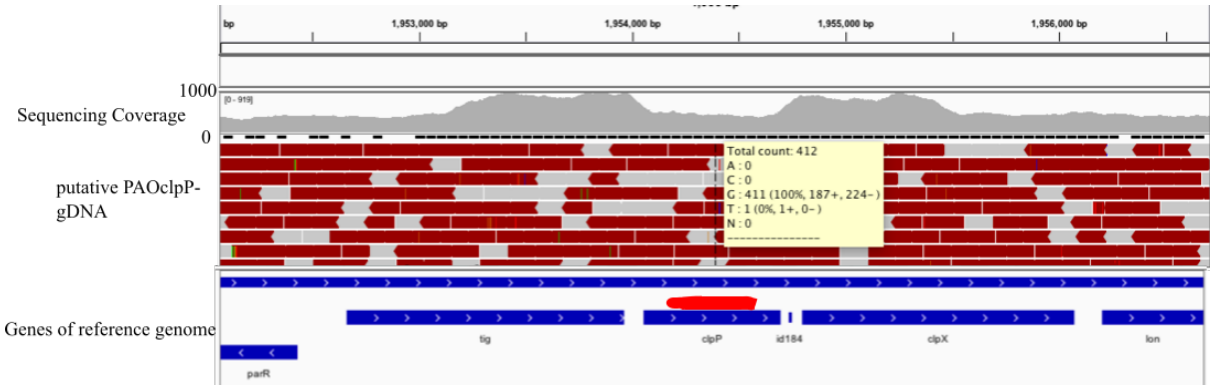


Figure 3.4.2 The alignment results visualized by IGV of the intended PAO(*clpP*) strain The strains were aligned to the reference genome. The top grey shading indicates the coverage of the reads of each individual gene. Red bars indicate the alignment of the base to the reference genome and a base call on the particular region indicated by a vertical dotted line. An example shown by a yellow box, G was the base on the region and it had 100% identity mapped to reference strains at the same region. The total read counts (412) were shown in a box for the *clpP* region. Blue bars (at the bottom of the figure) show the corresponding gene names from the reference genome (Robinson et al., 2011).

The partial reads coverage of intended PAO(*clpP*) sequences (from the region of *tig* to *clpX*) are displayed in the grey shade area (in the scale of 0-1000) of Figure 3.4.2. The sequencing coverage describes the average number of reads that mapped to reference bases of the PAO1 genome at each locus. For instance, the read count (412) of particular base G was shown in box (Figure 3.4.2), which was 100% matched to the reference base. The average of read-coverage of each base across the genome was approximately 500 (Figure 3.4.2) except for *tig* and *clpX*, which had 1000 reads coverage (Figure 3.4.2). These findings indicated that the read count of *tig* and *clpX* had duplicate read counts in contrast to other regions that had single read counts. These findings suggested that the plasmid with both *tig-clpP* and *clpP-clpX* sequences may have been conjugated to *P. aeruginosa* in which resulted in an increase of the number of reads of both *tig* and *clpX* genes in cells. However, the *clpP* gene had not been knocked out from the

strain and the reason why the *clpP* gene remained in this strain was unclear. Overall, the PAO(*clpP*) had no mutations with all the genes present including the *clpP* gene. Since no mutations occurred, PAO(*clpP*) could be considered as a reference strain as PAO1-WT.

3.5 Sample collection for RNA-Seq

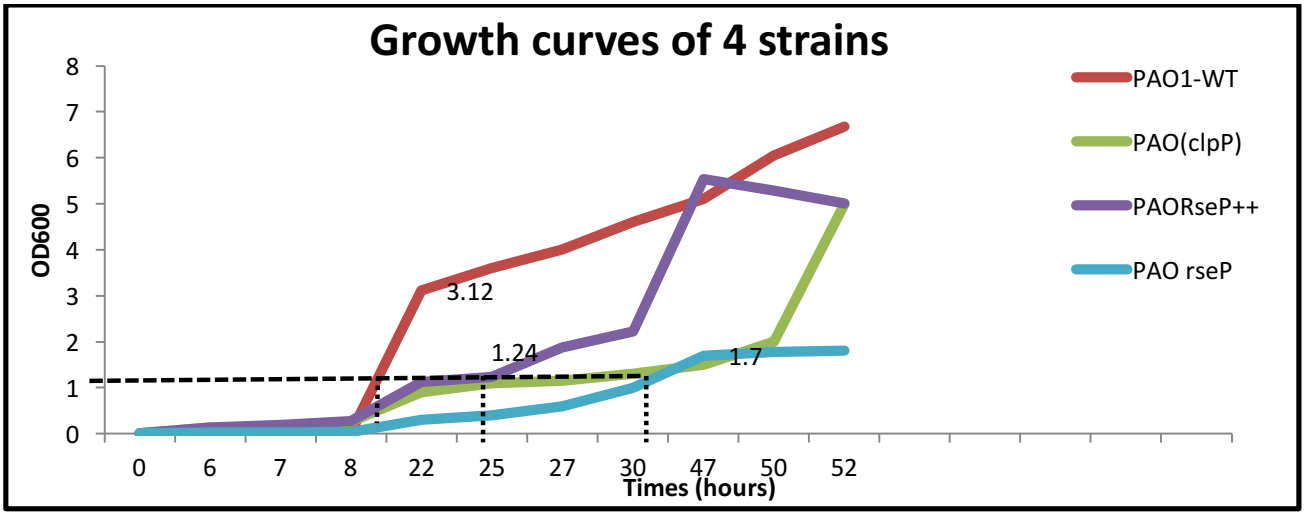
The *P. aeruginosa* samples including PAO1-WT, PAO1RseP^{++}, PAO1 *rseP* and PAO(*clpP*) were grown and harvested for RNA-extraction and were used as templates for RNA-Seq.

3.5.1 The growth of *P. aeruginosa* strains

The main goal of the growth experiment was to determine the growth rate and growth stages with corresponding time points for all strains. All the strains were cultured in King's B broth, a siderophore depletion broth, at 18˚C for 40 hours in order to determine the cell content at each time point. PAO *rseP* strain was temperature sensitive and it could not survive with the temperature above 22˚C so that all the strains were grown at the same temperature (22˚C) with a longer growth period.

The growth curves of four strains were plotted as incubation time against OD_{600} (Figure 3.5.1). The expected growth phases for all strains were shown (Figure 3.5.1): an initial lag phase followed by an exponential growth phase and then stationary phase. However, PAO1-(*clpP*), PAORseP^{++}and PAO *rseP* were slower in reaching an OD_{600} at 1.2 than PAO1-WT. After growing for 8 hours, PAO1-WT reached at OD_{600}=1.2 and the same OD_{600} reading for both PAORseP^{++} and PAO(*clpP*) was observed after growing for 25 hours (Figure 3.5.1). Arabinose was added to PAORseP^{++} after passing the lag phase (OD_{600}=0.6). PAO *rseP* reached OD_{600}=1 after 30 hours of growth (Figure 3.5.1). These findings indicated that the growth of PAORseP^{++}, PAO *rseP* and PAO(*clpP*) were similar but slower than PAO1-WT.

A)



B)

	OD_{600}			
Time （hours）	**PAO1-WT**	**PAO(*clpP*)**	**PAORseP^{++}**	**PAO *rseP***
0	0.01	0.01	0.01	0.015
6	0.02	0.13	0.13	0.03
7	0.02	0.017	0.19	0.03
8	0.03	0.26	0.28	0.035
22	3.12	0.9	1.12	0.3
25	3.6	1.1	1.24	0.4
27	4.01	1.15	1.87	0.6
30	4.6	1.3	2.23	1
47	5.1	1.5	5.53	1.7
50	6.05	2	5.29	1.78
52	6.68	5	5.01	1.8

Figure 3.5.1 Growth analysis of *P. aeruginosa* strains A) The growth curves of strains included PAO1-WT, PAORseP^{++}, PAO *rseP* and PAO(*clpP*) plotted in growth hours against to the corresponding cell density at OD600. The black dotted lines indicate the cells' concentration at OD_{600} 1.2 of each strain, for instance, PAO1-WT (10 hours), PAORseP^{++} (27 hours), PAO1(*clpP*) (27 hours) and PAO *rseP* (30 hours) **B)** The table of growth time (0-52 hours) and cell density at OD at 600 nm raw data of all four strains.

3.5.2 Sample preparation of *P. aeruginosa* strains for RNA-Seq

The strains were harvested when OD_{600} was around 1.5, because the cell density of cultures is required to meet the quality standard of RNA extraction, cut-off OD values are 1.1-2 (Assay Technologies, 2007). According to Table 3.5, all the strains were harvested when OD_{600} at 1.5-1.6 in triplicate which occurred at different time periods. For example, PAO1-WT strains were harvested after 13 hours of growing while other strains were harvested after 30-38 hours (Table 3.5).

To conclude, it is clear that PAO1-WT had a faster growth rate than other strains and it may be a problem for further analysis. The growth rate between wild-type and other strains may affect the expression level of these strains in the RNA-Seq results (3.7.4.2). Since all the sample strains were collected in triplicate at the same OD_{600}, all samples were at the same bacterial concentration to ensure consistency of results.

Strains/ culture	**OD_{600}**	**Hours**
PAO1-WT-1	1.4	13
PAO1-WT-2	1.62	13
PAO1-WT-3	1.46	13
PAORseP++1	1.63	25
PAORseP++ 2	1.53	31
PAORseP++ 3	1.6	31
PAO *rseP* -1	1.4	36
PAO *rseP* - 2	1.36	40
PAO *rseP* - 3	1.62	40
PAO(*clpP*) -1	1.47	31
PAO(*clpP*) -2	1.61	32
PAO(*clpP*) -3	1.72	30

Table 3.5 The OD_{600} and growth hours of samples which were collected for RNA-Seq

3.6 Quantifying the expression of *rseP* and siderophore receptor genes in *P. aeruginosa* strains

3.6.1 Introduction for RT-qPCR analysis

The initial aim of the RT-qPCR experiment was to determine the RseP expression between PAO1-WT, PAORseP^{++}, PAO *rseP* in order to determine if the overexpressed RseP strains worked successfully by increasing the RseP expression levels. The second goal of the RT-qPCR experiment was to investigate the effects of overexpression, and the absence of RseP on known RseP-dependent regulatory pathways (CSS systems). This aimed to determine whether the changes in RseP expression would alter the expression of siderophore receptor genes including *fpvA*, *foxA* and *fiuA* in CSS systems.

In the presence of ferri-siderophore, the siderophore binds to the outer membrane receptor and transmits the signal to the periplasm. The anti-sigma factors are then partially degraded by RseP site-2 protease in the periplasmic membrane and the anti-sigma factors are fully degraded by the cytoplasmic protease which allows release of the sigma factor and initiates the expression of receptor genes (1.6.2). Therefore, it was suggested that by overexpressing RseP protease in *P. aeruginosa*, the proteolytic cascade efficiency might be enhanced, which might cause a high level of transcription of the receptor genes (*fpvA*, *foxA* and *fiuA*) in CSS systems of *P. aeruginosa*.

3.6.2 Relative quantification parameters

3.6.2.1 Relative quantification

The RNA was harvested (2.8.1) and reverse transcribed into cDNA (2.8.4) for all samples and RT-qPCR was done by lightcycler, with *rseP* qPCR primers. For the measurement of *rseP*, *fpvA*, *fiuA* and *foxA* expression, cDNA was amplified and monitored by RT-qPCR and then analysed by relative quantification to reference genes (2.8.6). RT-qPCR amplification was monitored using SYBR green dye. The SYBR green dye can be used to monitor product amplification by taking a single collection of fluorescence following each DNA extension. The crossing point is the amplification cycle at which the SYBR green fluorescence signal crosses a certain threshold (Roche Applied Science, 2007).

3.6.2.2 Validation of housekeeping genes

To accurately measure gene expression by relative quantification, the reference genes needed to be constitutively and stably expressed to a level that was similar to the target genes. The *clpX* and *oprL* genes were used as reference genes in previous studies (Konings, 2010) and were again used as reference genes in this study. ClpX is the ATP-dependent part of the ClpXP serine protease. OprL is an outer-membrane protein that maintains outer membrane integrity and is involved in efflux systems. The primer efficiency and melting temperature of *clpX* and *oprL* primer pairs were validated by Konings, (2010). The amplification efficiency of *clpX* and *oprL* were 1.99 and 2.01 respectively, which were within the acceptable range of 1.8-2.0 (Konings, 2010). The amounts of *clpX* and *oprL* expression were used as references for measuring the relative expression of target genes in the RT-qPCR experiment.

3.6.2.3 Validation of primers

To measure the expression of *rseP*, *fpvA*, *foxA* and *fiuA* by RT-qPCR, specific primer pairs were designed for target gene amplification (2.4). The primers had been validated by previous study and the amplification efficiency of *fpvA*, *fiuA* and *foxA* was 1.95, 1.92 and 1.93 respectively (Draper, 2011). The efficiency of the *rseP* primer pair was validated in this study and it had an amplification efficiency of 2.044 which is able to meet the standard of primer efficiency, cut-off 1.8-2.0 (2.8.5) (Figure3.6.1).

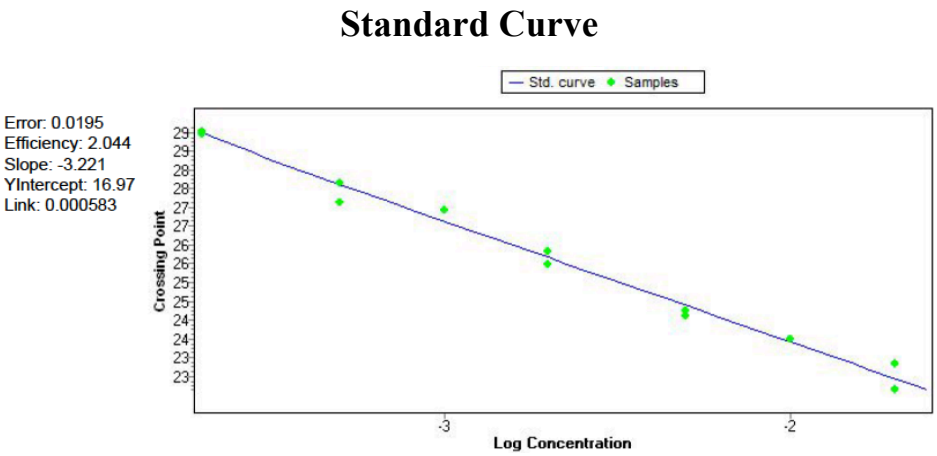


Figure 3.6.1 Validation of qPCR primer sets - example of Standard Curve (*rseP*) The RT-qPCR primers were validated for use with qPCR protocols by standard curve. The standard curve was generated by plotting the relative template concentration of duplicate dilutions against the quantification cycle as determined by the second derivative maximum method. A standard curve is shown for serial dilutions of genomic DNA template ($2x10^{-2}$, $1x10^{-2}$, $5x10^{-3}$, $1x10^{-3}$, $5x10^{-4}$, $2x10^{-4}$) and no template control reactions.

The *rseP*, *fpvA*, *foxA* and *fiuA* primer pairs were designed to target a 200 bp region of *rseP, fpvA, foxA* and *fiuA* (2.7.1). All the primers were tested by conventional PCR prior to RT-qPCR using the lightcycler, ensuring the primers amplified the correct region without any primer-dimer problems (2.7.1). According to Figure 3.6.2, RseP primers resulted in an expected amplification product with a clear band (200 bp). Therefore, the primer pair of RseP was shown to have acceptable amplification efficiency and the analysis of PCR products on the agarose gel showed there was a single expected product produced. The rest of the primers were validated in a previous study (Draper, 2011).

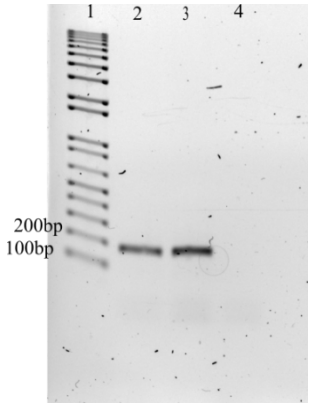


Figure 3.6.2 RseP primers validated by conventional PCR L1: ladder, L2-3: RseP PCR products with PAO1-WT gDNA template, L4: No template control with RseP primers only

3.6.3 The result of RseP expression of *P. aeruginosa* strains

The RseP expression of the *P. aeruginosa* strains including PAO1-WT, PAORseP^{++} and PAO *rseP* were quantified by RT-qPCR (Figure 3.6.3). The strains with normal relative expression of *rseP* were normalized to *clpX* and *oprL*, which have a consistent expression level in *P. aeruginosa* strains. The results showed that the *rseP* expression of PAO *rseP* was close to zero (Figure 3.6.3). When arabinose was added the bacteria containing the plasmid (PAORseP^{++}) had an average of a ten-fold increase in expression of *rseP* compared to PAO1-WT. These findings indicated that PAORseP^{++} strains worked successfully by elevating RseP expression levels (Figure 3.6.3).

Interestingly, the experiment showed that the bacteria which contained the pJN105 plasmid (PAORseP^{++}) had a four-fold increase in expression of *rseP* compared to PAO1-WT even in the absence of arabinose. The same trend was observed in the pJN106 plasmid which has the

same genomic features as pJN105, where 300 units of enzyme activity were recorded in *P. aeruginosa* in the absence of arabinose (3.2). In the absence of arabinose, the expression levels of *rseP* and *lacZ* genes were found to be low but measurable compared to the strains with arabinose induction in both the RT-qPCR and β-gal assays (3.2). These results matched the findings from previous studies where *lacZ* transcription of pJN106 had two enzyme units of activity which was detectable even in the absence of arabinose (Newman and Fuqua, 1999). Newman and Fuqua, (1999) reported that a combination of intact *araC* gene and P_{BAD} promoter in the plasmid may contribute to the low expression even in the absence of arabinose, but the reason why the *araC* gene can cause expression in the absence of arabinose was unknown. Overall, the results suggested that it was acceptable to have low expression quantified from the PAORseP^{++} strains in the absence of arabinose. However, most importantly, all these results indicated that arabinose had successfully induced the transcription of RseP in *P. aeruginosa*. Though the RseP protein expression of PAORseP^{++} was not measured in this study, a high level of *rseP* expression in the PAORseP^{++} construct is likely to have an increased level of RseP protein compared to wild-type. This is due to the positive correlation between mRNA and protein levels across a wide range organism including bacteria (Guo et al., 2008). To summarise, the RT-qPCR results suggested that *rseP* was successfully overexpressed in *P. aeruginosa*, being elevated ten-fold compared to the wild-type in the presence of arabinose.

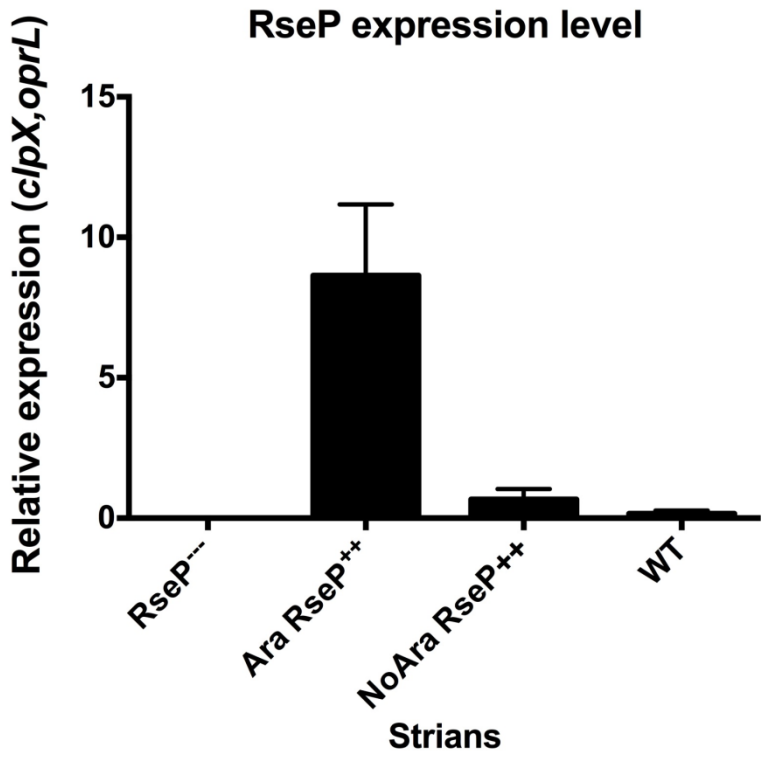


Figure 3.6.3 RseP expression level Quantification of the transcripts in PAO *rseP*, PAORseP^{++} (with and without arabinose induction) and PAO1-WT. The expression results were biological triplicates with standard error bars.

3.6.4 Quantification of FpvA, FoxA and FiuA expression

These experiments aimed to investigate the effect of overexpression, and absence of RseP on known RseP-dependent regulatory pathways including pyoverdine, desferrioxamine and ferrichrome CSS systems. This was carried out by using RT-qPCR to quantify the expression of siderophore receptor genes (*fpvA*, *foxA* and *fiuA*) in the strains PAO1-WT, PAORseP^{++} and PAO *rseP*.

Previous results suggested that the RseP protease mediated the expression of receptor genes in pyoverdine, desferrioxamine and ferrichrome systems (Draper et al., 2011). PAO *rseP* had significantly decreased in FpvA, FoxA and FiuA expression compared with PAO1-WT in the presence of siderophore (Draper et al., 2011). In this study, it was expected that the expression

of receptor genes (*fpvA*, *foxA* and *fiuA*) in CSS systems would be affected by differences in expression of *rseP* between PAORseP^{++} and PAO *rseP*.

The relative expression of these receptor genes (*fpvA, foxA* and *fiuA*) normalized to both *oprL* and *clpX* reference genes, was around 0.02-0.4 (Figure 3.6.4). Overexpressed RseP strains had a lower expression level of FpvA, FoxA, and FiuA than in wild–type strains (Figure 3.6.4). The PAORseP^{++}strain showed two-fold decreased transcription of *fpvA* than in the wild-type whereas the expression signals of both FoxA and FiuA were low and both had similar expression across all conditions including PAO1-WT, PAORseP^{++} and PAO *rseP* ($\pm$0.1) (Figure 3.6.4). All the strains were grown in KB broth, an iron limited medium without adding siderophores through the entire growth period in this study. The range of expression of *fpvA*, *fiuA* and *foxA* (0.1-0.3) agreed with the results of the expression of receptor genes in the absence siderophores from a previous study (Draper et al., 2011). These findings indicated that all the siderophore receptor genes had a weak expression (<0.1) in the PAORseP^{++}, PAO *rseP* and PAO1-WT strains when lacking siderophores.

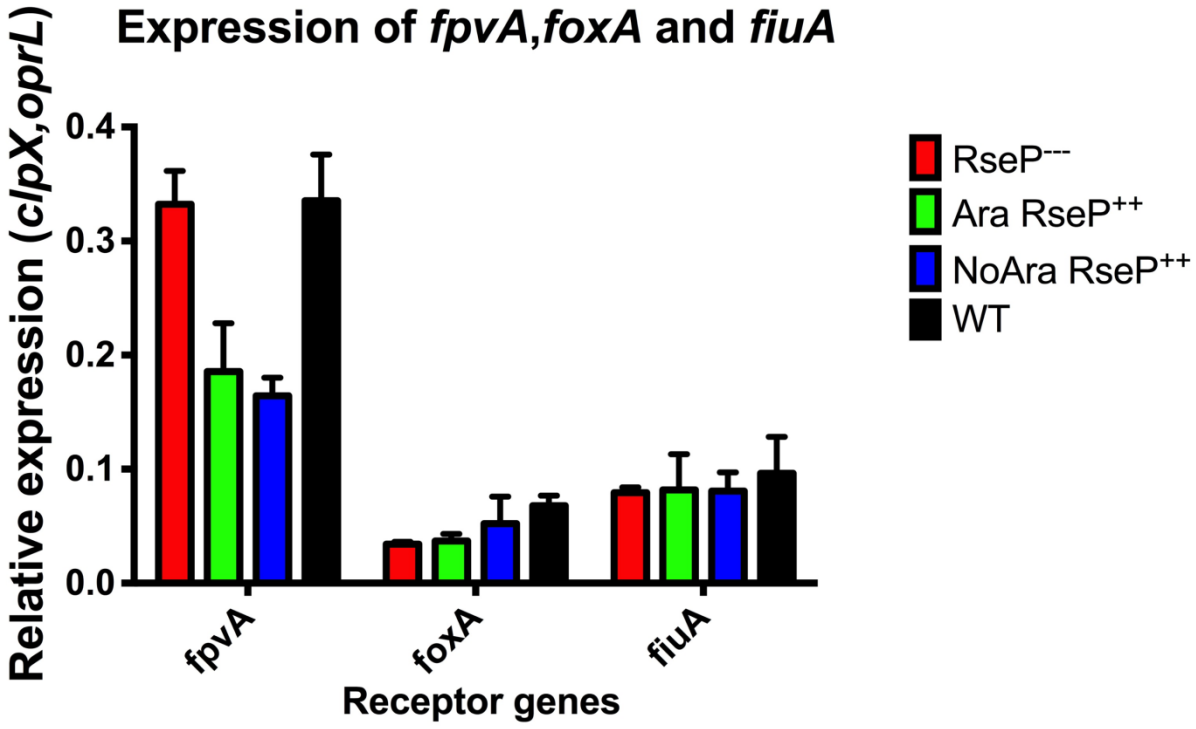


Figure 3.6.4 The expression of *fpvA*, *foxA* and *fiuA* Detection of the transcripts *fpvA*, *foxA* an *fiuA* in four different strains (PAO *rseP*, PAORseP^{++} (Arabinose/No Arabinose) and WT). Data are representative of three biological replicates with standard error. The expression was normalized to both *clpX* and *oprL* genes.

The results obtained through this RT-qPCR quantification were subjected to statistical analysis. An analysis of variance (T-test) was performed to compare the mean difference expression of receptor genes between PAO1-WT, PAORseP^{++} and PAO *rseP*. T-tests were done by using Prism 6 GraphPad. The p-values from these T-tests were adjusted for reducing the chance of false positive results and this was done by the Bonferroni-Holm method (Holm, 1979). A filter T-test with cut-off values of probability is equal to 0.05. The mean difference expression of all three receptor genes between PAO1-WT, PAORseP^{++} and PAO *rseP* is shown in Table 3.6. The results showed that the expression of all three receptor genes had no significant changes across all strains (PAORseP^{++}, PAO *rseP* and PAO1-WT) with adjusted p-values greater than 0.05.

Consequently, the results of the qPCR expression and statistical analysis suggested that altered expression of *rseP* may not alter the expression of FpvA, FoxA and FiuA in the Cell Surface Signalling (CSS) system in *P. aeruginosa* when lacking siderophores.

A)

Comparison: PAO1-WT vs PAORseP^{++}	Mean differences	Adjusted *p*-values
FpvA	0.15	0.13
FoxA	0.0313	0.05
FiuA	0.0145	0.76

B)

Comparison: PAO *rseP* vs PAO1-WT	Mean differences	Adjusted *p*-values
FpvA	0.00295	0.96
FoxA	0.0340	0.06
FiuA	0.017	0.7

Table 3.6 The T-test of the mean difference expression of FpvA, FoxA and FiuA between A) PAORseP^{++}and PAO1-WT, (B) PAO *rseP* and PAO1-WT strains

Chapter 3.7 RNA-Seq analysis

3.7.1 Introduction

The RNA-Seq analysis aimed to identify RseP-dependent genes that expressed differently in multiple comparison sets of reads from different *P. aeruginosa* strains. RNA-Sequencing was used for quantifying changes to the cellular transcriptome and the sample reads were quantified by counting the number of reads that mapped to each locus of the reference genome in the alignment and annotation step (Griffith et al., 2015). The differentially expressed genes were determined by using statistical tools (DESeq and EdgeR methods) to compare the datasets. The RNA-Sequencing technical support (library preparation and post-processing of raw reads) was provided by New Zealand Genomics Limited (NZGL).

3.7.2 RNA-Seq sample preparation

RNA samples including PAO1-WT, PAO *rseP*, PAORseP^{++} and PAO(*clpP*) in biological triplicate were extracted from bacterial cultures (2.8.1) and mixed with deoxyribonucleases (DNase) which were used to reduce the amount of genomic DNA. The concentration of each RNA sample was quantified by Nanodrop (2.8.3). The RNA concentration of all 12 samples had acceptable RNA concentration in the range of 800-1000 ng.μL^{-1} for cDNA library preparation (Table 3.7.1). All 12 RNA samples were measured at 260 nm and 280 nm which were used to assess RNA purity. Pure RNA has an A260/A280 ratio of 1.9-2.1 and 12 RNA samples had an A260/280 ratio approximately 2.1 which fulfilled the requirement of pure RNA with ratio between 1.9-2.1 (Assay Technologies, 2007) (Table 3.7.1). The RNA samples passed the quality measures, mRNA enrichment and cDNA library which were all carried out by NZGL.

Name of sample	Concentration ($ng.\mu L^{-1}$)	260/280 ratio
PAORseP^{++} 1	328.1	2.12
PAORseP^{++} 2	613.1	2.13
PAORseP^{++} 3	581.1	2.12
PAO1-WT 1	1416.5	2.17
PAO1-WT 2	620.9	2.19
PAO1-WT 3	1131	2.18
PAO *rseP* 1	948.1	2.18
PAO *rseP* 2	808.6	2.14
PAO *rseP* 3	1427.2	2.17
PAO(*clpP)* 1	631.7	2.18
PAO(*clpP)* 2	944.5	2.21
PAO(*clpP*) 3	683.1	2.22

Table 3.7.1 The RNA concentration and A260/280 ratio of 12 RNA samples

RNA sequencings and all the work described in following steps were done by NZGL except the last step of differential expression analysis (Figure 3.7.1). Full-length mRNAs were reverse transcribed to a library of cDNA fragments with adaptor primers attached to both ends (Figure 3.7.1) and the fragmentations were sent for the Illumina sequencing platform. Each molecule with amplification was then sequenced using the Illumina sequencing platform to obtain a short sequence (125 bp) from both ends. After sequencing, the fragments including adaptor sequences and reads with low quality (the quality score less than 30) were trimmed by NZGL. The resulting reads were mapped to a reference genome named PAO1-WT (PAO1 chromosome, complete genome NC_002516.2 from NCBI, consisting of 5570 genes in total), in order to produce a genome-scale transcription alignment that consisted of both the transcriptional structure and the level of expression for each gene (Figure 3.7.1).

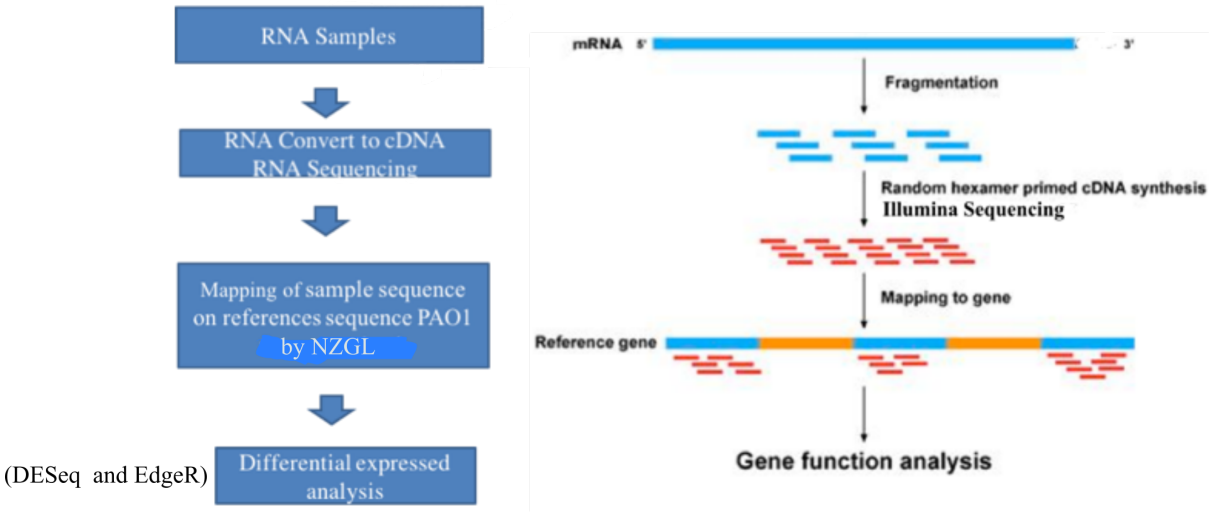


Figure 3.7.1 Pipeline of RNA-Seq experiment Total RNA template (mRNA) was converted into a library of cDNA fragments through fragmentation (blue bar). Sequencing adaptor primers were added to each cDNA fragment and a short sequence was obtained from each cDNA using an Illumina high-throughput sequencing platform. The resulting sequence reads (in red) were aligned with the reference genome to generate an expression profile for each gene annotated by NZGL. The statistical analysis was carried out with differential expressed analysis of each gene by DESeq and EdgeR methods (Bgisequence, 2012).

3.7.3 Sample quality

The quality score of 30 (Q30) to a base, is equivalent to the probability of an incorrect base call 1 in 1000 times (0.1%) and indicated that the base call accuracy is 99.9%. Q30 was considered as a benchmark for quality in RNA-Sequencing as a low base call accuracy will have higher error rate of genome and it can increase false-positive variants calls, which can result in inaccurate conclusions for experiments (Illumina, 2011). The quality per base of one RNA sample after the trimming process (which removed the adaptor DNA sequencing with a low quality score) is shown in Figure 3.7.2A. The major portion of base pairs lie in the green region indicating the mean quality score was above Q30. Overall, 93% of all sample genomes had mean quality score higher than Q30 with a mean quality score per bases equal to 35 (Figure 3.7.2B). The analysis from the current study demonstrated that a minimum number of 200 million reads or longer reads for all sequencing samples is sufficient to detect all annotated genes (Wang et al., 2011). The total of an approximate 400 million reads (with an overall depth of 55,272 million base) was obtained with 12 bacterial samples (Figure 3.7.2B) and indicated that it was a considerably large number of reads with high quality per base, as 98% average of quality greater than Q30 indicating sequencing consistency and accuracy.

A)



B)

Samples	Total number of reads (bp)	Overall depth of sequence (Mb)	Mean quality per base ≥ 30 (%)	Mean quality score
12	394,109,285	55,272	92.77%	35.41

Figure 3.7.2 Quality score of samples of RNA-Sequencing A) Typical sequence quality per base, as analysed by FastQC from RNA-Sequencing, of a portion of the reads from triplicate 3 of the PAORseP^{++} strains from this study. The x-axis represented the position of sequences in the read, with the read length of 125 bp, and the y-axis represented the quality score. The average of quality score of the sequence per base was greater than 30 **B)** A summary of RNA-Seq analysis showing total number of reads; average depth of sequence (million base); quality parameters and the cut–off values of quality per base is equal to or larger than 30.

The total read counts of each sample were summarised in Table 3.7.2. The number of reads for each strain were added up to measure these significant variations between triplicates. Therefore, the total read count in each sample showed no significant variation in the number of reads (with less than 6%) between each triplicate of PAORseP^{++}, PAO1-WT and PAO *rseP*. One strain PAO(*clpP*) had one biological replicate PAO(*clpP*)-2 with lower read counts but this difference

is not sufficient to influence the results of the analyses. Overall, the quality of sequence reads was high and a higher quality of sequences reads made the reads data more valid. The quality of samples fulfilled the standard and was acceptable for further analysis.

Strains	Total counts
PAORseP^{++}1	29622059
PAORseP^{++}2	31415444
PAORseP^{++}3	30930174
PAO1-WT 1	35872021
PAO1-WT 2	35137484
PAO1-WT 3	35080275
PAO *rseP* 1	33950759
PAO *rseP* 2	34724270
PAO *rseP* 3	33473458
PAO(*clpP*) 1	34199098
PAO(*clpP*) 2	25926068
PAO(*clpP*) 3	33778175

Table 3.7.2 Total read counts summarised for each sample strain in biological triplicates, provided by NZGL

3.7.4 Differential expression analysis

After validating the raw reads from RNA-Seq, quantified transcripts were then compared to each other to identify differential expression of genes. This was done by DESeq and EdgeR (two R Bioconductor packages). Comparisons were tested for significantly different expression between all samples (PAO1-WT, PAORseP^{++}, PAO *rseP* and PAO(*clpP*)). A false-discovery adjusted p-value (p-adj) was used to determine significance for reducing the chance of false positive results. Differential gene expression greater than 1 or less than -1 fold differences ($\log_2$Fold change) that were statistically significant (p-adj<0.05) were considered as RseP-affected genes.

Initial analysis focused on the comparison between PAORseP^{++} and PAO *rseP*. Further comparison was tested between reference strains (PAO1-WT and PAO(*clpP*)) and RseP mutated strains (PAORseP^{++} and PAO *rseP*).

3.7.4.1 DESeq and EdgeR

DESeq and EdgeR are two methods most commonly used for differential expression analysis. Both DESeq and EdgeR, two R Bioconductor packages used in R-studio, were used as a model based on the negative binomial distribution (Anders and Huber, 2010) and the two methods tended to rank the genes similarly (Soneson and Delorenzi, 2013).

The EdgeR method implies that the variance is equal to the mean expression but this method has no correction included for wide variations in read counts between triplicates (Nookaew et al., 2012). In this project, the variance of gene expression across biological triplicates of each strain could be larger than its mean expression values. Thus, the data accuracy obtained by the EdgeR method would be reduced (Rapaport et al., 2013). Whereas, DESeq breaks the variance estimate into mean expression of the gene and the biological expression variability between triplicates. The DESeq analysis with two parameters can reduce the probability of false discovery (Rapaport et al., 2013). However, DESeq has more stringent criteria for setting the threshold in the condition of the higher variance of gene expression across biological triplicates than the EdgeR models (Rapaport et al., 2013).

3.7.4.2 Initial comparison between PAORseP^{++} and PAO *rseP*

The initial differential expression (DE) analysis was performed between two extreme conditions, PAORseP^{++} and PAO *rseP*. This analysis was thought to provide the clearest cut differences among all comparison sets that identified those genes whose expression is affected only by RseP. The DE genes were identified between PAORseP^{++} and PAO *rseP* by both EdgeR and DESeq (Figure 3.7.4.1). There were 185 DE genes identified by the EdgeR method and 69 genes were found by DESeq, and 61 DE genes were found in common using both methods. EdgeR operates on a model with less restrictive statistical parameters than DESeq so that more genes were identified by the EdgeR method than DESeq (Figure 3.7.3). The results from EdgeR analysis showed increasing numbers of false positive being identified, primarily because genes which showed large variation in biological triplicate reads but no differences in each single repeat strains were identified, unlike DESeq analysis (Rapaport et al., 2013). Consequently, higher numbers of genes were identified by EdgeR than DESeq which has more

stringent criteria for filtering the genes. Additionally, the majority of DE genes were also identified by EdgeR which indicated that those genes were not identified by coincidence and ensured the consistency and robustness of DESeq method. DESeq analysis may be more suitable than EdgeR for samples in triplicates where it can provide accurate screening based on mean differences in each single strains, thus reducing the false discovery rate. Therefore, the EdgeR method was not suitable for this study and was deemed unsuitable for further analysis. Overall, 69 DE genes were identified by DESeq with the stringent filter (Table 3.7.3) and the changes which were observed are more likely to be the true transcriptional response.

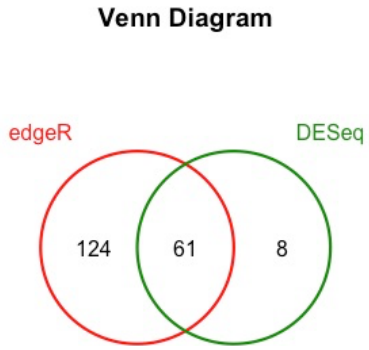


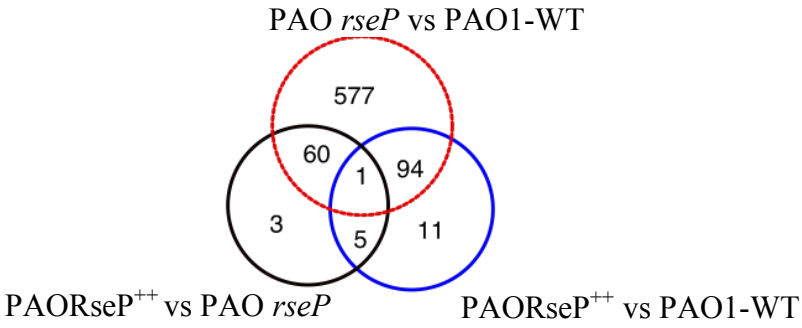
Figure 3.7.4.1 The differential expressed analysis between PAORseP^{++} and PAO *rseP*
The results are represented in the Venn diagram by both EdgeR (red) and DESeq (green). There were 185 DE genes identified by EdgeR and 69 DE genes by DESeq. 61 DE genes overlapped in the middle between the two methods.

3.7.4.3 Differential expressed comparison analysis between PAO1-WT, PAORseP^{++} and PAO *rseP*

Multiple comparisons were used for differential expressed (DE) analysis to strengthen the accuracy of the results by comparing multiple strains. Comparison were tested for significant differential expression between PAO1-WT, PAORseP^{++} and PAO *rseP* (Figure 3.7.4.2). The base-line expression of DE genes presented in PAO1-WT was expected to lie between extreme strains (PAORseP^{++} and PAO *rseP*), which could be identified in all three samples. However, the numbers of differential expressed genes (700) identified between PAO1-WT and PAO *rseP* were much higher than the numbers of DE genes identified in the initial analysis (PAORseP^{++} and PAO *rseP*) (69 DE genes) (Figure 3.7.4.2). The numbers of overlapping genes from the intersection of circles were summarised in Figure 3.7.4.2B. The genes (61 of 700 DE genes) which identified between PAO *rseP* and PAO1-WT were overlapped in PAORseP^{++} and PAO *rseP* analysis. Five DE genes which were identified in PAORseP^{++} vs PAO1-WT were also identified in PAORseP^{++} vs PAO *rseP*, analysis (Figure 3.7.4.2). The majority of DE genes in

PAORseP^{++} and PAO *rseP* were also found when compared to PAO1-WT groups. These findings suggested that some DE genes were identified in multiple comparison conditions (i.e. PAO1-WT, PAORseP^{++} and PAO *rseP*) as expected.

A)



B)

Comparison groups / Numbers of genes	PAORseP^{++} vs PAO *rseP* (69)
PAO *rseP* vs PAO1-WT (>700)	61
PAORseP^{++} vs PAO1-WT (111)	5

Figure 3.7.4.2 The identified differential expressed genes between PAO1-WT, PAORseP^{++} and PAO *rseP* A) The overall result is display by the Venn diagram. Only one gene overlapped in all three groups. **B)** The table summarised the number of common DE genes in the intersection of the circles of the Venn diagram. The numbers represented the DE genes. There were 61 genes overlapped in PAO *rseP* vs PAO1-WT and PAORseP^{++} vs PAO *rseP* groups, five genes were overlapped between PAORseP^{++} vs PAO *rseP* and PAORseP^{++} vs PAO1-WT groups. More than 700 DE genes in total were identified between PAO *rseP* and PAO1-WT and 111 DE genes were identified between samples PAORseP^{++} and PAO1-WT.

Additionally, some DE genes had the same expression level in PAORseP^{++} and PAO *rseP* strains while in PAO1-WT those genes had much higher or lower expression. The reason for the different gene expression in PAO1-WT may be that PAO1-WT had a faster growth rate than PAORseP^{++} and PAO *rseP* (3.5.1). The expression of PAO1-WT may be affected by the differences in growth rate compared to other strains (4.2). Therefore, those DE genes with

extreme expression could be due to false discovery from WT strains, which were not found in PAORseP^{++} and PAO *rseP* analysis.

3.7.4.4 The comparison of PAO(*clpP*) and PAO1-WT

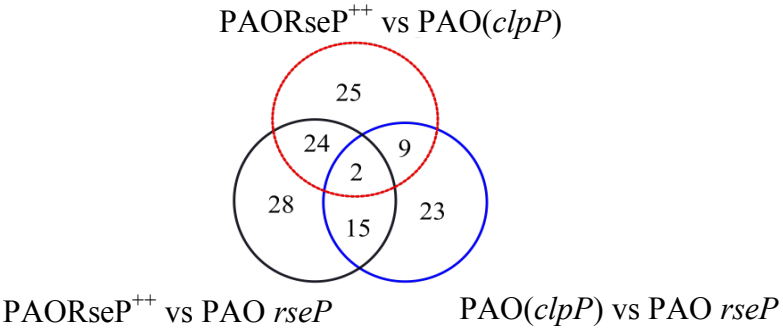
Since previous findings indicated PAO1-WT may not be a suitable reference genome, we used PAO(*clpP*) as a reference genome to make differential comparisons. The reason for choosing PAO(*clpP*) was its similarity in growth rate to PAORseP^{++} and PAO *rseP* which made it a suitable substitute reference genome. The comparison results showed 709 genes (a proportion of DE genes were shown in appendix 3) were identified that significantly (p-adj<0.05) differentially expressed between PAO(*clpP*) and PAO1-WT strains with same threshold as in other comparisons ($\log_2$Fold change ≤ -1 and ≥ 1).

The differentially expressed genes were compared in order to distinguish the similarity between PAO1-WT and PAO(*clpP*) while in fact we found that they were quite different considering 709 genes were chosen between two strains.

3.7.4.5 Differential expressed comparison analysis between PAO(*clpP*), PAORseP^{++} and PAO *rseP*

The alternative comparison analysis between PAO(*clpP*), PAORseP^{++} and PAO *rseP*, which had the same growth rates was performed. This comparison provided evidence that DE genes may be affected by RseP and not by environmental factors (i.e. growth rate). Comparison between three samples (PAO(*clpP*), PAORseP^{++} and PAO *rseP*) revealed 26 overlapping (out of 60) DE genes in all three comparison groups (Figure 3.7.4.3) (The genes were shown in Appendix 4). Additionally, there were 49 DE genes identified in PAO(*clpP*) vs PAO *rseP* and 17 of them were identified commonly in PAORseP^{++} vs PAO *rseP* groups (Figure 3.7.4.3). These indicated that half of the DE genes identified in PAO(*clpP*) comparison groups were commonly found in PAORseP^{++} and PAO *rseP*, initial analysis. These findings suggested that these overlapped genes may have a strong transcriptional response to RseP. Nevertheless, considering only the factor of growth rate, the 69 DE genes could not be neglected which use was identified by comparing two extreme condition (PAORseP^{++} and PAO *rseP*) which were important for further analysis. These findings ensured the consistency and accuracy of the results determined in initial analysis (between PAORseP^{++}and PAO *rseP*).

A)



B)

Comparison groups Numbers of genes	PAORseP^{++} vs PAO *rseP* (69)
PAO(*clpP*) vs PAO *rseP* (49)	17
PAORseP^{++} vs PAO(*clpP*) (60)	26

Figure 3.7.4.3 The differential expressed genes identified from comparisons PAO(*clpP*) groups A) The Venn diagram displays the overview of identified DE genes in each data set. Two genes overlapped in all three groups**. B)** The table summarises the number of common DE genes in the intersection of the circles of the Venn diagram.

3.7.5 RseP-dependent genes (69 DE genes) identified from PAORseP^{++} and PAO *rseP*

The final DE gene lists have 69 DE genes which were identified as strongly responsive to RseP in PAORseP^{++} and PAO *rseP* comparison samples (Table 3.7.3). This comparison provides the clearest differences among other sample sets. Therefore, the final 69 DE genes were carried forward in our analysis of gene function.

According to Table 3.7.3, 40 of the genes were shown upregulated in PAORseP^{++} and the remaining 29 genes were upregulated in PAO *rseP*. In addition, two DE genes were found in all three comparisons groups which are PA1592 (hypothetical protein) and PA3649 (RseP protease). RseP protease was found in a 965 fold (9-log2Fold) increase in PAORseP^{++} compare to a PAO *rseP* and these results matched the RseP expression of PAORseP^{++} quantified by RT-qPCR (3.6.3). Moreover, the PA1592 had a 52-fold increase when overexpression of RseP and these genes were overlapped in all comparison sets. These indicated the RseP protease had a strong association to PA1592 but with unknown function.

ID	Gene Names	Fold Change	log2Fold Change	p-adj	Functions
PA3649	*rseP*	964.77	9.91	5.63E-08	RseP protease
PA2294		694.08	9.44	3.26E-09	probable ATP-binding component of ABC transporter
PA3691		81.47	6.35	1.22E-02	hypothetical protein
PA1323		62.58	5.97	3.45E-03	hypothetical protein
PA1324		59.21	5.89	1.22E-02	hypothetical protein
PA1592		51.90	5.70	1.04E-08	hypothetical protein
PA3819		22.21	4.47	1.96E-05	conserved hypothetical protein
PA2414		21.10	4.40	2.25E-03	L-sorbosone dehydrogenase
PA3650	*dxr*	20.29	4.34	2.55E-03	1-deoxy-d-xylulose 5-phosphate reductoisomerase
PA4880		17.46	4.13	3.26E-09	probable bacterioferritin
PA3891	*opuCA*	14.46	3.85	8.95E-04	OpuC ABC transporter, ATP-binding protein, OpuCA
PA2415		14.26	3.83	5.63E-08	hypothetical protein
PA3044	*rocsS2*	11.93	3.58	3.40E-06	Two-component sensor RocS2
PA0737		11.05	3.47	3.81E-07	hypothetical protein
PA4154		10.20	3.35	1.93E-02	conserved hypothetical protein
PA0567		9.48	3.25	1.71E-05	conserved hypothetical protein
PA1921		8.93	3.16	4.92E-07	hypothetical protein
PA3889	*opuCC*	7.90	2.98	1.94E-05	OpuC ABC transporter, periplasmic substrate-binding protein, OpuCC
PA0062		7.30	2.87	1.71E-05	hypothetical protein
PA3888	*opuCD*	7.15	2.84	2.57E-05	OpuC ABC transporter, permease protein, OpuCD
PA5182		7.03	2.81	1.51E-04	hypothetical protein
PA0843	*plcR*	6.64	2.73	1.35E-05	phospholipase accessory protein PlcR precursor
PA2751		6.38	2.67	2.96E-04	conserved hypothetical protein
PA0763	*mucA*	6.00	2.58	1.87E-02	anti-sigma factor MucA
PA3890	*opuCB*	5.75	2.52	4.82E-04	OpuC ABC transporter, permease protein, OpuCB
PA2275		5.68	2.51	3.14E-02	probable alcohol dehydrogenase (Zn-dependent)
PA0833		5.20	2.38	1.93E-02	hypothetical protein
PA5297	*poxB*	5.13	2.36	2.76E-03	pyruvate dehydrogenase (cytochrome)
PA4204	*ppgL*	4.97	2.31	1.90E-03	periplasmic gluconolactonase, PpgL
PA5492.1	*ersA*	4.78	2.26	7.59E-03	ErsA
PA1471		4.67	2.22	3.34E-03	hypothetical protein
PA3902		4.40	2.14	7.10E-03	hypothetical protein
PA5483	*algB*	4.12	2.04	2.32E-02	two-component response regulator AlgB
PA1606		3.63	1.86	2.39E-02	hypothetical protein
PA3445		0.30	-1.76	2.32E-02	conserved hypothetical protein

PA2020	*mexZ*	0.27	-1.88	4.30E-02	MexZ
PA4160	*fepD*	0.27	-1.89	1.64E-02	ferric enterobactin transport protein FepD
PA3206		0.25	-2.02	2.39E-02	probable two-component sensor
PA3446		0.24	-2.08	7.59E-03	conserved hypothetical protein
PA3938		0.22	-2.16	2.56E-02	probable periplasmic taurine-binding protein precursor
PA2312		0.21	-2.25	7.10E-03	probable transcriptional regulator
PA3885	*tpbA*	0.21	-2.25	2.14E-03	protein tyrosine phosphatase TpbA
PA4517		0.17	-2.55	1.74E-03	conserved hypothetical protein
PA0201		0.17	-2.56	2.14E-03	hypothetical protein
PA2311		0.17	-2.59	8.88E-04	hypothetical protein
PA4228	*pchD*	0.16	-2.66	3.28E-02	pyochelin biosynthesis protein PchD
PA4218	*ampP*	0.16	-2.66	2.55E-02	AmpP
PA3449		0.15	-2.69	1.89E-04	conserved hypothetical protein
PA4443	*cysD*	0.15	-2.78	1.85E-02	ATP sulfurylase small subunit
PA4513		0.13	-2.97	1.78E-03	probable oxidoreductase
PA5471.1		0.12	-3.00	4.82E-02	PA5471 leader peptide
PA2786		0.11	-3.25	8.88E-04	hypothetical protein
PA2019	*mexX*	0.09	-3.42	9.75E-05	Resistance-Nodulation-Cell Division (RND) multidrug efflux membrane fusion protein MexX precursor
PA3205		0.08	-3.61	1.29E-05	hypothetical protein
PA4221	*fptA*	0.08	-3.62	1.10E-02	Fe(III)-pyochelin outer membrane receptor precursor
PA2018	*mexY*	0.08	-3.67	4.77E-04	Resistance-Nodulation-Cell Division (RND) multidrug efflux transporter MexY
PA4220		0.07	-3.83	2.97E-06	hypothetical protein
PA5471	*armZ*	0.07	-3.87	4.77E-04	ArmZ
PA4219	*ampO*	0.06	-3.99	1.40E-05	AmpO
PA5469		0.05	-4.37	1.94E-08	conserved hypothetical protein
PA5470		0.05	-4.41	4.92E-07	probable peptide chain release factor

Table 3.7.3 The list of RseP-mediated genes (69) were identified from PAORseP^{++} vs PAO *rseP*, analysed by RNA-Seq.

3.7.6 RT-qPCR validation of RNA-Seq results

The RT-qPCR was used to validate the RNA-Seq data. The experiments aimed to determine if the expression results in RNA-Seq had positive correlation of the expression in RT-qPCR. Three genes were selected for RT-qPCR validation including PA2019 (which was encoded by MexX protein), PA2275 and PA5471. These three genes had a significantly high fold change between PAORseP^{++} and PAO *rseP* in RNA-Seq analysis (Table 3.7.3). PA2275 was upregulated in PAORseP^{++} while both MexX and PA2275 were downregulated in PAORseP^{++}.

The cDNA of triplicate 2 (selected at random) of PAO1-WT, PAO(*clpP*), PAORseP^{++} and PAO *rseP* (Table 3.6.2), was isolated as templates of samples from RNA-Seq (2.8.4). The design primers of MexX, PA5471 and PA2275 were verified by conventional PCR with PAO1 genomic DNA for ensuring all the primers were amplifying the correct region without any contamination before conducting RT-qPCR. Both MexX and PA5471 amplified a 140 bp product and PA2275 had a product size around 380 bp (Figure 3.7.6.1). All the primers had clear bands with expected size (Figure 3.7.6.1) which made them likely right products of primers amplification.

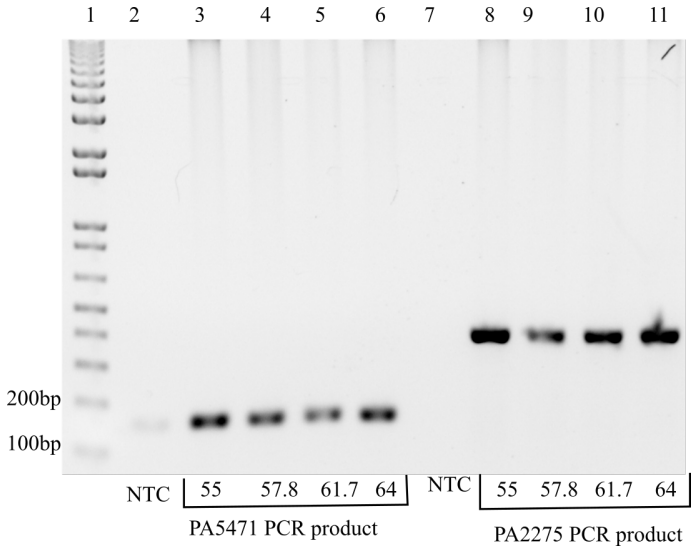


Figure 3.7.6.1 PA5471 and PA2275 primers validated by gradient PCR from 55˚C to 64˚C
L1:1 kb DNA ladder, L2 and 7: no template control with PA5471 primers and PA2275, respectively, L3-6: PA5471 PCR product from 55˚C to 64˚C of the range of gradient temperature. Expected size of 150 bp, L8-12: PA2275 PCR product with same temperature range.

All the primer pairs (PA5471, PA2275 and MexX) showed good amplification efficiency with low error between technical replicates. For instance, PA5471 had primer efficiency of 2.051,

PA2275 was 2.016 and MexX was 2.049 (Appendix 1). All these results meet the standard for primer efficiency; cut-off is 1.8-2.1. These results ensured the robustness and preciseness of RT-qPCR quantification (2.12.1). cDNA was amplified and monitored by RT-qPCR and then analysed by relative quantification to reference genes including *clpX* and *oprL* (3.6.2.2). The RT-qPCR experiments were repeated twice with the same cDNA template to ensure consistency of the results.

The relative fold change differences (in logarithm base of 2) expression of three candidate genes (*mexX*, PA5471 and PA2275) between PAORseP^{++}, PAO *rseP* and PAO1-WT samples were quantified by both RT-qPCR and RNA-Seq (Figure 3.7.6.2). The same trends of log2 fold differences expression between PAORseP^{++} and PAO *rseP* of all target genes were observed in both RT-qPCR and RNA-Seq analysis (Figure 3.7.6.2A). These findings indicated that the results of RNA-Seq differential expressed between PAORseP^{++} and PAO *rseP* had positive correlation to the expression monitored by RT-qPCR. However, the expression of PA2275 within the PAORseP^{++} and PAO1-WT samples (Figure 3.7.6.2B) had a negative correlation between RT-qPCR and RNA-Seq. These opposite findings of this candidate gene reflected that there was variability in expression of PAO1 performed by RT-qPCR and RNA-Seq. This inconsistency in PAO1 expression made these results unsuitable for further analysis. Thus, the genes identified from PAO1-WT comparison groups could not be relied on as being RseP-dependent genes and were not included in further analysis.

Overall, these results suggested that the small subset of DE genes from which expression was quantified by RT-qPCR was found reliably and consistently correlated to the results from RNA-Seq between PAORseP^{++} and PAO *rseP*.

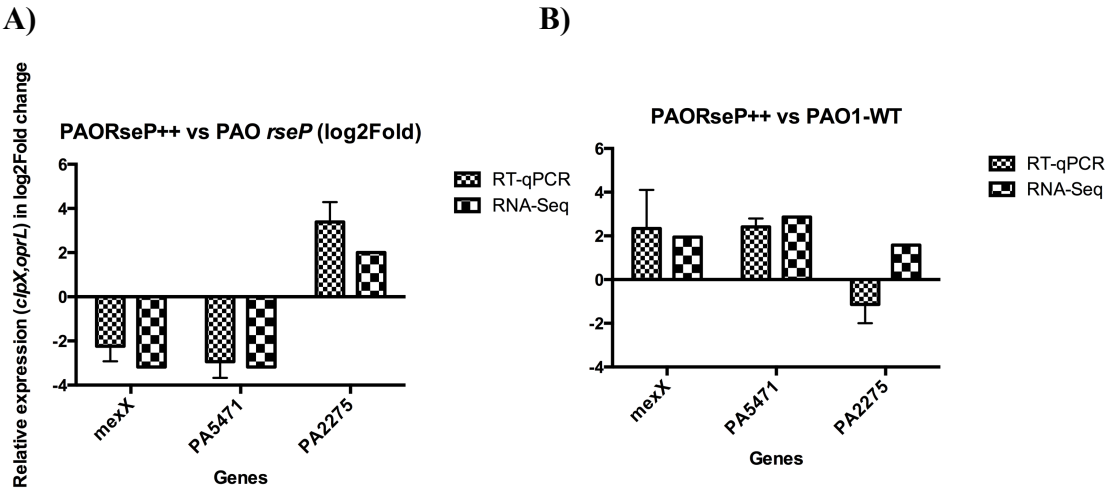


Figure 3.7.6.2 The expression in both RT-qPCR and RNA-Seq analysis between *P. aeruginosa* strains A) PAORseP^{++} and PAO *rseP* and **B)** PAORseP^{++} and PAO1-WT. The y-axis showed the log fold change expression in logarithm (base of 2) between data sets. The values are the average of two RT-qPCR technical replicates with standard error shown.

3.7.7 Functions of differential expressed genes response to RseP

All of the DE genes from Table 3.7.3 which showed a strong transcriptional response to RseP were determined by mapping to reference genome of PAO1-WT in Pseudomonas database (annotations for Genbank format for complete genomes, pseudomonas.com) for assessing the corresponding function names.

3.7.7.1 DE genes with unknown functions

Approximate 30 DE genes were encoded by hypothetical proteins with unknown function, which caused difficulty in identifying the pathways through which those genes are present. Some DE genes with unknown function were strongly altered by RseP, for example, PA3691 and PA1592 were highly upregulated in PAORseP^{++} (with 6.35 and 5.7-$\log_2$fold increase in expression, respectively) (Table 3.7.3). Future work on identifying the relevant pathways will be necessary to identify how RseP might regulate those genes.

3.7.7.2 DE genes with known functions

Of the 69 DE genes which were found to be differentially expressed between PAORseP^{++} and PAO *rseP* 39 had specific functions that had been identified. The majority of DE genes were found as neighbourhood genes which were closely related in clusters. For instance, PA0763 is encoded by MucA anti-sigma factor in alginate production and mucoid formation of *P. aeruginosa*. PA2414 and PA2415 are encoded to transmembrane transporters, PA3205 and PA3206 which are integral membrane proteins associated with signal transduction and

phosphorylation in cells. PA3445 and PA3446 are associated with membrane transporters for sulphur compounds. PA4218 and 4219 encoded AmpP and AmpO. PA4221 is a Fe^{3+} pyochelin membrane receptor (Table 3.7.3). The expression of these DE genes with known functions was significantly altered by RseP with significantly fold differences above the threshold. However, the functions of these DE genes have not yet been specifically characterized in the Pseudomonas database, thus we know little about these genes and their pathways.

Additionally, four genes which showed a strong transcriptional response to RseP in this study are involved in the MexXY efflux system, which had an increased expression (3-$\log_2$Fold) of PAO *rseP* (Table 3.7.3). This may suggest that a lesser amount of MexXY efflux transporter and regulated genes were present when the amount of RseP was increased and the genes were upregulated without RseP. In addition, a sub group of OpuC ABC osmolarity transporter systems involved four genes (*opuCA*, *opuCB*, *opuCC*, *OpuCD*), which had a four $\log_2$fold increase in expression of PAORseP^{++} than in PAO *rseP*. These findings in this study suggested that the genes which were involved in these two pathways are strongly associated with the expression of RseP (4.5.4). The MexXY efflux system had been well-studied in *P. aeruginosa*. The MexXY-efflux system was responsible for extruding antibiotics from outside bacterial cells. However, the mechanisms of OpuC osmolarity transporter are still not fully understood in Gram negative bacteria. OpuC transport is responsible for transport of osmoprotectant (glycine betaine) into cells in order to maintain osmotic homeostasis in cells but the intrinsic mechanisms remained unclear.

3.7.8 Conclusion

To summarise, the quality of sequence reads was high and showed the read counts data were valid for analysis. The differential expressed analysis comparing PAORseP^{++} and PAO *rseP* was only altered by RseP. The final DE genes partially overlapped in PAO(*clpP*) strains, PAORseP^{++} and PAO *rseP* strains showed strong evidence of accuracy of DE analysis. However, the expression of PAO1-WT may have been influenced by different growth rates and so comparisons with PAO1-WT data were not included in the final DE gene list (Table 3.7.3). Additionally, 69 genes were identified as differential expressed genes mediated by RseP in *P. aeruginosa*. Though 50% of the selected genes had unknown function in cells, some of these genes may have a strong correlation to RseP with significant differential expressed fold change altered by RseP. Interestingly, based on the genes with known function, RseP may play an important role in MexXY efflux systems by altering the expression of *mexX* and *mexY* genes (4.4.5) and RseP may also play a role in the regulation of OpuC transporter systems (4.4.6).

3.8 The Minimal inhibitory concentration (MIC) tests of *P. aeruginosa* strains

3.8.1 Introduction

The MexXY operon consisted of *mexX* and *mexY* genes, and MexXY efflux systems had natural resistance to aminoglycosides and fluoroquinolones in *P. aeruginosa* (Morita et al., 2012). The findings of RNA-Seq analysis suggested that increased expression of both *mexX* and *mexY* presented in the PAO *rseP* strains (Table 3.7.3). Recent studies reported that MexXY upregulation was the cause of increasing aminoglycosides and fluoroquinolones resistance (Hocquet et al., 2006). Thus, an increased level of antibiotic resistance in the PAO *rseP* strain was expected, given that *mexX* and *mexY* were upregulated in this strain compared to others. The Minimal Inhibitory Concentration (MIC) test was carried out to investigate whether the differences expression of *rseP* altered the antibiotic resistance level of the *P. aeruginosa* strains including PAO1-WT, PAORseP^{++} and PAO *rseP*. The MIC test measures the lowest concentration of antibiotic that prevents visible growth of bacterium strains after incubation. Two antibiotics (tobramycin and ciprofloxacin) were used in the MIC tests; these antibiotics belong to the aminoglycoside and fluoroquinolone antibiotic family respectively in *P. aeruginosa*.

3.8.2 MIC determination

3.8.2.1 Sample templates for MIC tests

A gentamicin resistance (Gm^{R}) marker was in the pJN105 vector and present in PAORseP^{++} strains only. This gentamicin resistance marker may affect the results of the resistance level in MIC tests. Thus, in order to ensure the accuracy of the results, the pJN105 vector was transformed into PAO1-WT and PAO *rseP* to give PAO1-WT(pJN105) and PAO*rseP(*pJN105) (2.6.10.3). Additionally, the pJN105 plasmid with cloned *rseP* genes was transformed into PAO *rseP* to form RseP complement strains, PAO *rseP*(pJN105::*rseP*$^{++}$). RseP complement strains were expected to have the same resistance phenotype as wild-type. The final strains including PAO1-WT(pJN105), PAO *rseP*(pJN105), PAOpJN105::*rseP*$^{++}$ and PAO*rseP* (pJN105::*rseP*$^{++}$) (RseP complement) strains were used as templates.

3.8.2.2 Method of the MIC test

Four bacterial strains were spotted onto Muller-Hinton (MH) agar plates for MIC tests (2.17). The MH agar is widely used for antibiotic sensitivity testing and it has a few properties have

benefit for antibiotics use. Firstly, MH agar is a non-differential medium and suitable for the growth of a wide range of microorganisms including *P. aeruginosa*. Secondly, MH media contains beef infusion, casamino acids, and starch. Starch acts as a colloid that protects bacteria from toxic conditions in the medium. Beef infusion and casamino acids provide nutrients for allowing bacteria to survive and grow (Haltiner et al., 1980). Thirdly, this medium was in low concentration (0.04% MH agar) which allows for better diffusion of the antibiotics than other media and all these properties of MH agar can lead to the accuracy of MIC results.

The range of antibiotic concentrations used for determining MIC is widely accepted to be in doubling dilution steps up and down from 1 $mg.L^{-1}$ as required. In this experiment, tobramycin (10 $\mu g.mL^{-1}$) was used in the range of 0.03 to 2 $\mu g.mL^{-1}$ per agar plate (25 mL per plate) whereas ciprofloxacin testing was carried out in the range of 0.03 to 0.5 $\mu g.mL^{-1}$ per plate (2.2, 2.17). The reason for the different concentration ranges between these two antibiotics tests was because previous studies have shown that the *P. aeruginosa* strains had lower resistance to ciprofloxacin than to tobramycin (Brazas and Hancock, 2005).

MIC tests can measure the lowest concentration of tobramycin and ciprofloxacin that will inhibit the growth of an organism. The inocula of 10^2,10^3,10^4 colony-forming units/spot of each strain were spotted to a control plate without spreading any antibiotics. This control plate allowed to find whether the growth of *P. aeruginosa* was being affected by antibiotics. The heavy inoculum could be adjusted so that 10^6 colony-forming units per spot were applied to the plates for MIC tests (Andrews, 2001). All the results were recorded after 48-hour incubation at 22˚C. Although the standard protocol required the plates to be incubated at 37˚C overnight (Andrews, 2001), the PAO *rseP* strain would not be able to grow at high temperature so the plates were incubated for a longer period at a low temperate (22˚C).

3.8.3 MIC results of tobramycin and ciprofloxacin

The MIC tests of tobramycin resistance are showed in Figure 3.8A, whereas the test of ciprofloxacin resistance was only done once (Figure 3.8B), which may lack reproducibility. According to the MIC results in Figure 3.7.1A, the RseP overexpressed strains and RseP complement strains had higher resistance to tobramycin than to PAO1-WT and PAO *rseP* (MIC=2 $\mu g.mL^{-1}$) and the MIC levels gave a consistent result in all triplicates of PAORseP^{++} and RseP complement strains (Figure 3.8A). Both PAO *rseP* and PAO1-WT had low MIC

resistance to tobramycin and the MIC levels fluctuated between 0.025 to 1 μg.mL^{-1} in both strains. These results indicated that the strains with additional RseP had higher tobramycin resistance than other PAO strains. Compared to the PAO1-WT, strains that overexpressed RseP showed a higher resistance to the aminoglycosides antibiotic tobramycin, however, the RseP^{++} and RseP complement strains still displayed a low level of resistance compared to the clinical isolates (MIC=4 μg.mL^{-1}) (Hocquet et al., 2008) .

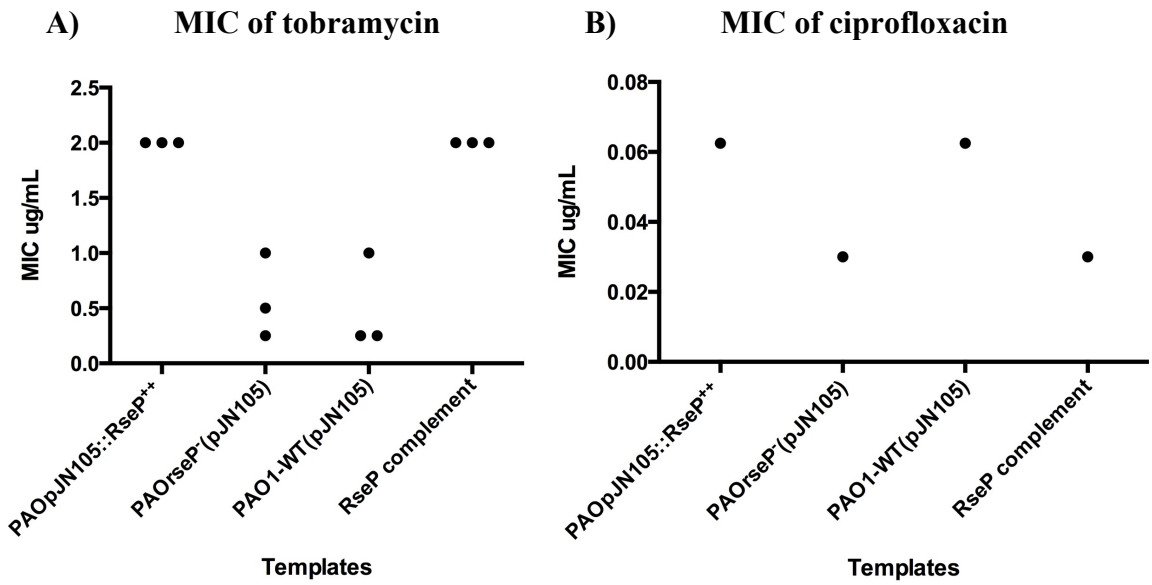


Figure 3.8 Tobramycin and ciprofloxacin MIC results against four different strains. The strains including PAORseP^{++}, PAO *rseP (*pJN105), PAO1-WT(pJN105) and RseP complement strains were carried out with **A)** tobramycin **B)** ciprofloxacin, respectively. The tobramycin MIC results were representative of three biological triplicates. MIC levels of ciprofloxacin antibiotics were from a single replicate.

Moreover, the MIC results for ciprofloxacin, as shown in Figure 3.8, had a maximum MIC (0.06 μg.mL^{-1}) of all strains. The MIC levels of both PAORseP^{++} and PAO1-WT were higher than the PAO *rseP* and RseP complement (MIC=0.03 μg.mL^{-1}) strains. This indicated that the strains that were examined in this experiment had lower resistance to ciprofloxacin than to tobramycin and PAO *rseP* was less susceptible to both tobramycin and ciprofloxacin with increased expression of the *mexXY* operon. The experiments on ciprofloxacin antibiotics were only done once, thus the MIC results lack reproducibility and consistency and may not be reliable.

3.8.4 Conclusion

As a result, the hypothesis that the higher resistance level of both aminoglycosides and fluoroquinolones in PAO *rseP* compared to PAO1-WT, because of increased expression of *mexXY,* was not confirmed in the MIC tests. PAORseP^{++} showed higher resistance to tobramycin in all these replicates compared to other strains (PAO *rseP* and PAO1-WT), though, this resistance of PAORseP^{++} was not clinically significant. The reason behind these findings remains unclear. Thus, the expression of *rseP* had no direct effect on the changes in resistance phenotypic of *P. aeruginosa*.

3.9 Motility assay on *P. aeruginosa* strains

3.9.1 Introduction of the motility assay

Recent study has shown that the differences in swarming motility occurred between PAO *rseP* and PAO1-WT in virto phenotypic analysis (Dubern et al., 2015). However, the study did not present conclusive results of the change in phenotype of PAO *rseP*.

In this project, the swimming and swarming motility were tested in RseP mutant strains (PAORseP^{++}, PAO *rseP* and RseP complement) in relation to the wild-type (PAO1-WT) to determine whether the changes in RseP amounts impacted on the motility of *P. aeruginosa* (2.11). *P. aeruginosa* has a polar flagellum that it uses to swim in liquid and swarm in a semi-solid environment. This surface organelle, flagellum has been identified as an important virulence factor in the chronic model of *P. aeruginosa* disease (Shapiro, 1995). Recent study has demonstrated that bacteria lacking flagella caused less inflammation and death than in the wild-type (Feldman et al., 1998).

Swimming and swarming motility both require a functional polar flagellum and determine the flagellum phenotypic behaviour of cells (Blair, 2003). The plate-based assays are carried out for both the swimming and swarming motility in *P. aeruginosa.* The differences between the swimming and swarming tests is the method of bacterial inoculation on agar plates. Swimming motility is defined as the movement in liquid (bacteria are inoculated in the middle of the agar layer), while swarming motility monitors the movement across the semisolid surface in cells (bacteria are inoculated on the surface of agar) (Ha et al., 2014).

The swimming and swarming motility assays were followed by method 2.12 (Ha et al., 2014). All plates were incubated upright at 22˚C for 40 hours. Because of the high level of sensitivity of the swarming assay, swarming could only occur at 37˚C. Thus, the swarming plates were incubated at 37˚C for all strains. However, the PAO *rseP* strain grew slowly at this temperature. Both swimming and swarming motility assays were performed in technical duplicates to ensure the accuracy and consistency of the results.

3.9.2 The results of swimming motility

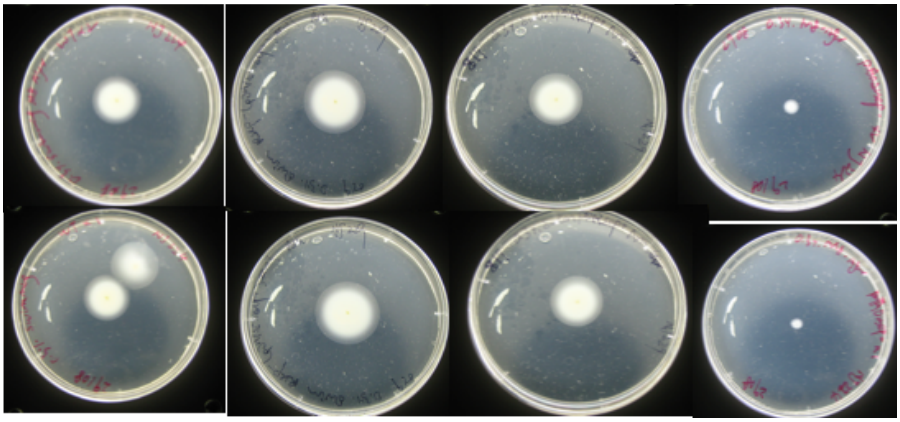
The results indicated that identical swimming phenotypes of the strains (PAO1-WT, PAORseP^{++} and RseP complement) had a similar inoculum size (as shown in Figure 3.9.1A).

PAO *rseP* had the smallest swimming diameters (0.6 cm), smaller than other strains (2.4 cm) (Figure 3.9.1B). Among isolates that were swim-positive, isolates with the largest swimming diameters tended to form larger swarming colonies than those with the smaller swim diameters (Murray et al., 2010). These findings suggested that PAO *rseP* might have a reduction in swimming motility but this phenotypic behaviour of PAO *rseP* could also be due to the slow growth of PAO *rseP* (slower than to PAO1-WT) that is mentioned in result 3.5.1.

To investigate whether the differences in motility between PAO *rseP* and PAO1-WT were affected by the differences in growth rate, both PAO1-WT and PAO *rseP* were inoculated on 0.3% LB agar plates with the same condition as for the swimming motility assay. The LB medium consisted of tryptone and NaCl, which provided a better growth condition for the growth of bacteria with enriched nutrient. It was expected that the growth of PAO *rseP* would be enhanced by plating in this media. However, the result showed that the inoculum size of PAO *rseP* was smaller than in PAO1-WT (Figure 3.9.1C), indicating the differences in growth between PAO *rseP* and PAO1-WT were still present. All these findings suggested that the slow growth rate of PAO *rseP* had not yet adapted to growing in the enriched media. Overall, we cannot conclude whether the reduction of swimming of PAO *rseP* was due to the strain having slow growth or if PAO *rseP* had an effect on the swimming motility of *P. aeruginosa*.

A)

PAO1-WT **RseP complement** **PAO RseP^{++}** **PAO *rseP***



B)

Strains	Diameters (cm)
PAO1-WT	2.1
RseP complement	2.9
PAORseP^{++}	2.4
PAO *rseP*	0.6

C)

PAO *rseP* **PAO1-WT**

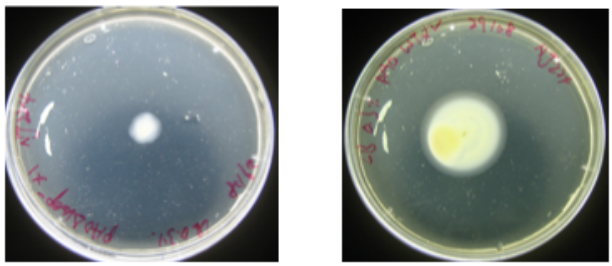


Figure 3.9.1 Swimming motility of *P. aeruginosa* strains A) The swimming plates with the strains included PAO1-WT, RseP complement, PAORseP^{++} (with arabinose) and PAO *rseP*. All the strains were in duplicate. The swimming plates (0.3% agar M8) were incubated for 40 hours at 22˚C. **B)** The table showed the averaged swimming diameters (in cm) of each strain in replicates. **C)** Swimming motility of both PAO1-WT and PAO *rseP* were plated on LB agar. Swimming plates (0.3% agar LB) were incubated for 40 hours at 22˚C.

3.9.3 Results of swarming motility

P. aeruginosa strains were associated with a circular pattern of swarming phenotypes (Figure 3.9.2A). Identical swarming inoculum size had been identified in both PAO1-WT, PAORseP^{++}, and RseP complement strains (Figure 3.9.2A). Among all strains, PAO *rseP* again had the smallest inoculum size (Figure 3.9.2A). However, PAO *rseP* could not grow sufficiently at 37˚C, thus the reduction in size of PAO *rseP* may have been due to the slow growth of this strain. Therefore, all these findings indicated that some differences in swarming motility were observed between PAO *rseP* and PAO1-WT but we cannot conclude whether the difference was due to slow growth or if the changes in RseP expression affected the swarming motility in *P. aeruginosa*.

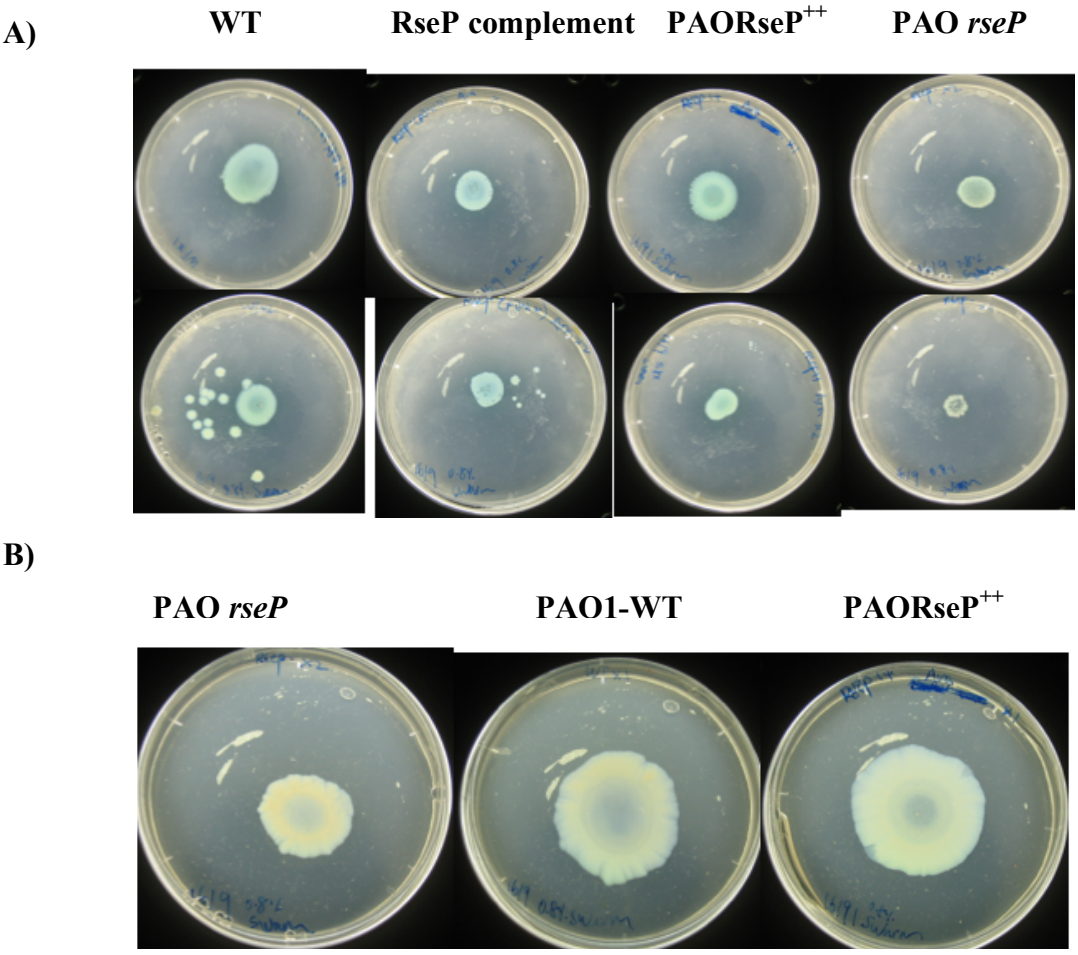


Figure 3.9.2 Swarming motility of *P. aeruginosa* strains A) The strains of PAO1-WT, RseP complement, PAORseP^{++} and PAO *rseP* were in technical duplicates shown vertically. The swarming plates (0.65% M8 agar) were incubated for 24 hours at 37˚C (2.11). **B)** Swarming motility of PAO *rseP*, PAO1-WT and PAORseP^{++} for a longer period**.** Swarming plates (0.65% agar M8) were first incubated for 24 hours at 37˚C and then at room temperature for 11 days.

Therefore, the swarming assay was further carried out by incubating the plates at a lower temperature for a longer period of time with PAO *rseP*, PAO1-WT and PAORseP^{++} in order to determine whether the motility was affected by the growth rate of each strain (Figure 3.9.2B). The swarming plates were incubated at 37˚C for the first 24 hours then at room temperature (22˚C) for 11 days (Figure 3.9.2B). After 11 days, PAO *rseP* had a similar phenotype as PAO1-WT, suggesting that a longer incubation time with a low temperature tended to enhance the size of the inoculum of PAO *rseP*. However, there is a possibility that prolonged incubation would have caused the portion of the bacteria species that caused the motile phenotype in the petri-dishes to be reduced, despite the large plaque observed (Figure 3.9.2B). Therefore, the differences growth rate between PAO1-WT and PAO *rseP* may be a major factor affecting the swarming motility of *P. aeruginosa*.

To summarise, the findings showed that similar patterns of swimming and swarming motility were presented in PAORseP^{++}, PAO1-WT and RseP complement strains. By applying overexpressed RseP in PAO *rseP* strains (RseP complement strains), the same motility phenotype was observed in PAO1-WT. The growth rate may have significant influence on the motility phenotype of PAO *rseP*. Moreover, the genes which encoded flagellum in *P. aeruginosa* were not found to have significant changes in expression between PAORseP^{++} and PAO *rseP* in RNA-Seq analysis (Table 3.7.3). These findings suggested that RseP may not play a role in changing the motility in *P. aeruginosa.* Therefore, all these findings implied that RseP may not be direct associated with the swimming and swarming motility of *P. aeruginosa*.

Chapter 4 Discussion

4.1 Summary of key findings

The overall aim of this research was to identify the differentially expressed genes mediated by RseP protease using RNA-Seq analysis. To achieve the aim, the overall approaches involved the use of *P. aeruginosa* strains including PAO1-WT, RseP overproduced strain (by cloning), *rseP* deleted strain and PAO(*clpP*). The fulfilment of this aim would provide a better understanding of the involvement of the RseP protease in other biological pathways and the generic background in bacteria.

To investigate the effect of the overexpression and absence of RseP on RseP-dependent pathways (e.g. CSS systems), RT-qPCR was used to quantify the expression of receptor genes in CSS systems. Gene expression data indicated that the higher amount of RseP protease does not increase the transcription of receptor genes in CSS systems when siderophores are lacking.

The RseP-mediated genes were identified by differential expression analysis from RNA-Seq and the results agreed with the expression results obtained by RT-qPCR. From the RNA-Seq analysis, 69 genes were identified with a significant response to RseP. However, the phenotype investigation demonstrated that RseP does not increase the aminoglycosides resistance, despite RseP affecting the expression of antibiotic efflux genes.

4.2 The growth temperature of *P. aeruginosa* strains

The growth rate of each of four strains PAO1-WT, PAORseP^{++}, PAO *rseP* and PAO(*clpP*) were examined (3.4), where all the strains were grown at 18˚C. The optimal growth temperature of *P. aeruginosa* was at 37˚C but it can survive at a wide range of temperatures from 4˚C to 42˚C (LaBauve and Wargo, 2012). Lower temperatures may cause slow growth of the *P. aeruginosa* cells (Bishop, 2012). Below 30˚C, some virulence pathways were not active in *P. aeruginosa* which was less susceptible to antibiotics and the cells may have decreased in accumulation at this stage (LaBauve and Wargo, 2012). Thus, the growth experiment was carried out at a low temperature in this study and this may have resulted in a slower growth rate of *P. aeruginosa* than the growth rate normally observed in 37˚C growth conditions (3.5.1). The changes in temperature may alter the expression of temperature-regulated genes in *P. aeruginosa* strains. PAO *rseP* was a temperature sensitive strain and it can only survive at low

temperatures, indicating there is a strong association between the RseP protease and the growth temperature of *P. aeruginosa*. PAORseP^{++} and PAO *rseP* had slower growth rates at 18˚C compared to PAO1-WT, which was a standard reference for growth. It was hypothesized that changes in the RseP expression may alter the expression of temperature-regulated genes. Thus, the altered expression of *rseP* may alter the growth rate of PAO *rseP* strain. Additionally, the double reads of *clpX* and *tig* genes present in the strains may contribute to slow growth rate of the PAO(*clpP*) strain (3.4.3). However, the growth analysis may lack reproducibility as the experiments were only been carried out once from each strain without any replication. Thus, the growth analysis (3.5) may not be reliable but this finding at least provides a sign that all mutant strains (PAORseP^{++}, PAO *rseP* and PAO(*clpP*)) had slower growth rates than in the PAO1-WT.

4.3 The correlation between RseP and siderophore receptor genes (*fpvA*, *foxA*, *fiuA*)

In CSS systems, the site-1 periplasmic protease (DegS) partially cleaves off the anti-sigma factor, leaving it more susceptible to degradation by RseP protease, which mediates the partial degradation of the sigma factor (1.6). Consequently, sigma factors were then released and initiated the expression of receptor genes in CSS systems (Draper et al., 2011; Llamas et al., 2014). It was hypothesized that overexpressed RseP may affect the transcription of the siderophore receptor genes in CSS systems by enhancing the degradation of anti-sigma factors, which leads to an increase in sigma factor activity.

Recent studies have implied that the RseP protease have mediates the degradation of anti-sigma factors (e.g. FpvR, FoxR and FiuR) and initiates the transcription of siderophore receptor genes (Draper et al., 2011; Bastiaansen et al., 2014). In the absence of RseP protease, the N-terminal fragments of FpvR were not degraded in the cytoplasmic region (Draper et al., 2011). RseP site-2 protease is essential in this signalling but it must be preceded by a prior cleavage mediated by the site-1 protease (DegS) (Saki et al., 1998). The proteolytic activity would not be enhanced by overexpression of RseP in cells as RseP plays a role only at one step (secondary cleavage) of the proteolytic cascade. The high level of proteolytic activity could be observed only when other proteases (DegS and ClpXP) in cascade were also expressed. Thus, this could be the possible explanation of the results where low expression of siderophore receptor genes

were found between PAORseP^{++} and PAO *rseP* (3.6.4). These suggested that the alter expression of *rseP* may not affect the expression of siderophores genes.

Additionally, a lack of siderophores may influence the sigma factor activity in *P. aeruginosa.* In CSS systems, ferrisiderophores would be transported into cells only when the siderophore signal was present This is the signal that activates the RseP-mediated cleavage of the intramembrane domain of anti-sigma factors and released sigma factors. The sigma factors bound to the polymerase and activated the transcription of the CSS receptor genes which facilitated the uptake of ferrisderophores (Llamas et al., 2014). In contrast, in the absence of siderophores, the siderophore uptake systems would not be activated (inactive state) and there would be no transcription of CSS receptor genes (Llamas et al., 2014). In this experiment, all the strains were cultured in King's B broth, with an insufficient amount of siderophores for signalling transduction in CSS systems (3.5.1). Consequently, the anti-sigma factor sequestered the sigma factor at the cytoplasmic membrane, keeping it inactive and reduced the transcription of siderophore receptor genes in cells. Hence, CSS systems may remain inactive because of a lack of siderophores in growth conditions, despite overexpressed RseP in *P. aeruginosa*.

In conclusion, we cannot determine whether overexpression of *rseP* had no effect on regulation of transcription of receptor genes in the CSS system as such, or if whether the very poor expression of receptor genes was due to these strains lacking siderophores.

4.4 RNA-Seq analysis

4.4.1 The filtering parameters of differential expression analysis

RNA-Seq was carried out to identify RseP-mediated genes and pathways in *P. aeruginosa* by comparing the differential expression analysis between PAORseP^{++} and PAO *rseP*. This approach provided the clearest cut differences among all comparison sets (3.7.3) and it ensured the changes which were a true transcriptional response to RseP. The differentially expressed genes were identified to transcript expression on the basis of significant (p-adj ≤ 0.05) increases or decreases of the logarithm (to base 2) of fold change cut-off (≥ 1 and ≤ -1) in expression between samples. This threshold was set in order to identify differentially expressed (DE) genes, whose expression was at least double or halved. A false-discovery adjusted p-value (p-adj) less than 0.05 was deemed suitable for control for false positives. Therefore, the combination of the fold change cut-off and the *p*-adjusted parameter in differential expressed analysis represents the right filter for RseP relevant changes.

4.4.2 The comparison analysis between the reference strains

PAO1-WT and PAO(*clpP*) strains were used as reference strains and both were involved in multiple comparison analysis for identifying the DE genes mediated by RseP. The analysis showed that more genes were identified as differentially expressed between PAO1-WT and RseP mutated strains and this may be due to the different growth rate between PAO1-WT and PAORseP^{++} and PAO *rseP* strains (3.5.1). Therefore, while PAO1-WT is involved for comparison analysis, it was unsuitable for further analysis due to the uncertainty of the transcripts of PAO1-WT which could lead to erroneous results.

4.4.3 RNA-Seq results may be affected by temperature

The expression of *P. aeruginosa* strains including PAO1-WT, PAO(*clpP*), PAORseP^{++} and PAO *rseP* may be altered by temperature. Temperature changes are responsible for the differential transcriptional regulation of 6.4% of *Pseudomonas aeruginosa* genes from 22˚C to 37˚C (Barbier et al., 2014). A Recent study reported that the temperatures significantly altered the expression of pyoverdine and pyochelin receptors (*fpvA* and *fptA*) which were observed to be down-regulated at 22˚C compared to 37˚C, which suggested the lower temperature may have influenced the expression of genes in *P. aeruginosa*. The *P. aeruginosa* strains in this study grew at 18˚C and the expression of genes in these strains may have be affected by temperature (Barbier et al., 2014). By comparing the list of temperature-regulated genes

(Barbier et al., 2014) and the final list of genes differentially expressed between PAORseP^{++} and PAO *rseP* (in this study), six genes overlapped in two lists (Table 4.4). This suggested that the expression of these six temperature-regulated genes may not be altered by RseP and these genes should not be considered as RseP-dependent genes in *P. aeruginosa*.

ID	Function
PA0567	Conserved hypothetical protein
PA4218	AmpP
PA4219	AmpO
PA4220	hypothetical protein
PA4221	Fe(III)-pyochelin outer membrane receptor precursor
PA4228	pyochelin biosynthesis protein PchD

Table 4.4 The temperature-regulated genes of *P. aeruginosa* differentially expressed between PAORseP^{++} and PAO *rseP* (Barbier et al., 2014).

4.4.4 The established RseP-dependent pathways, ferrisiderophore systems and alginate synthesis pathway

Recent study has shown that there are three substrates (FoxR; FiuR; FpvR) of anti-sigma factors, which control a range of biological functions such as uptake of iron components via ferrisiderophores and required RseP-mediated proteolysis in *Pseudomonas aeruginosa* in CSS systems (Draper et al., 2011). However, the genes which were encoded by anti-sigma factors and siderophore receptor proteins in CSS were not identified in the final DE gene-list (Table 3.7.3). These findings agreed with the results quantified by RT-qPCR, which showed no changes in expression of receptor genes between PAO1-WT, PAORseP^{++} and PAO *rseP*. These findings suggested that the difference expression of *rseP* may not affect the transcription of siderophore receptor genes in CSS systems (3.6.4, 4.4) when lacking siderophores. Additionally, MucA, an anti-sigma factor is degraded by RseP protease through proteolytic cascade and this increases the alginate synthesis in *P. aeruginosa* (Qiu et al., 2007). The MucA anti-sigma factor had a 2.85 $\log_2$Fold increase in expression in PAORseP^{++} compared to PAO *rseP* (Table 3.7.3). These findings suggested that the expression of *mucA* is regulated by differential expression of *rseP* in *P. aeruginosa*.

4.4.5 RseP may play a role in the MexXY-OprM efflux transport system

The resistance-nodulation-cell division (RND) family composes major efflux pumps in *P. aeruginosa* (Lister et al., 2009). One of the efflux systems, MexXY-OprM is constitutively expressed at a basal level in wild-type strains which is a major cause of the resistance to aminoglycosides and fluoroquinolones in clinical strains of *Pseudomonas aeruginosa* (Islam et al., 2004; Morita et al., 2012). The MexXY efflux transporter systems consists of a MexX (outer membrane fusion proteins), MexY (efflux transporter), ArmZ (ribosome-targeting regulator mediated expression of *mexY*) and PA5471.1 (a leader peptide of ArmZ). The MexX and MexY proteins are encoded by a single transcriptional unit, *mexXY*. The transporter MexY interacts with the outer membrane channel OprM and the periplasmic protein MexX to export aminoglycosides out of the cells (Figure 4.4.1) (Masuda et al., 2000) in order to contribute to the antibiotic resistance in *P. aeruginosa*. Additionally, the MexZ protein binds to and represses the promoter of the *mexXY* operon (Matsuo et al., 2004). The expression of *mexXY* is mediated by PA5471, a protein with unknown function that provides the functional link between the ribosome and MexXY-OprM (Morita et al., 2006) (Figure 4.4.1). Moreover, PA5471 is functions with PA5470 encoding an alternative peptide releasing factor in order to resume stalled ribosomes (Morita et al., 2006; Morita et al., 2009) (Figure 4.4.1).

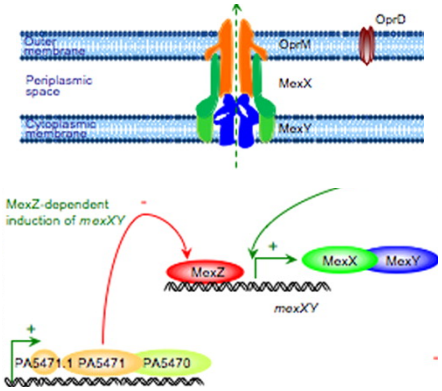


Figure 4.4.1 Regulation of the MexXY-OprM efflux system in *P. aeruginosa* The efflux pumps MexXY systems are driven by proton motive force, MexY in the cytoplasmic membrane, a periplasmic membrane fusion protein (MexX) and the outer membrane factors (OprM) (Monita et al., 2015). The regulation of *mexXY* operon was targeted by antimicrobials, the MexXY level is increased to provide resistance to the inducer, showing that the aminoglycoside is a substrate for the efflux pump. This induction process increased the levels of the PA5471 protein modulating MexZ repressor activity (Monita et al., 2015; Muller et al., 2011).

Both MexX and MexY were found to be upregulated in PAO *rseP* strains with a 3-$\log_2$fold increase greater than in RseP overexpressed strains (Table 3.7.3). These suggested that the membrane protein MexX and MexY may interact with RseP protease at the periplasmic

membrane. The MexX protein may be degraded by RseP. Thus, the expression level of both *mexX* and *mexY* may be lower in PAORseP^{++} compared to PAO *rseP*. In contrast, both MexX and MexY may not be completely degraded in the absence of RseP; a higher expression of both *mexX* and *mexY* would be observed in the cells (e.g. PAO *rseP*). All these findings provide a direction for future study on determining if there is some interaction between RseP protease and MexXY protein in the efflux system by western blotting assay and protein expression analysis in order to have a better understanding of the intrinsic mechanism between RseP and the MexXY efflux system.

4.4.6 RseP and the OpuC Osmolarity ABC transporter system

ATP binding cassette (ABC) function is transport of essential nutrients and export of toxic molecules that couple by ATP hydrolysis to transport process (Davidson et al., 2008). All ABC systems share a highly conserved ATP-hydrolysing domain. One of four osmoregulatory ABC transporters, OpuC, functions as the primary transporter for glycine betaine under high osmolarity in *P. aeruginosa* (Chen and Beattie, 2007). Glycine betaine and choline are natural osmolytes which are named as osmoprotectant. Glycine betaine was found as a byproduct of processing sucrose from sugar beets (Mäkelä*,* 2004*)*. Osmolytes play a role in maintaining cell volume and fluid balance. Bacteria accumulate osmolytes for protection against a high osmotic environment (Csonka., 1989).

The OpuC operon consists of four genes which are encoded by the ATP-binding protein (OpuCA) in the cytoplasmic membrane, two peripheral membrane permeases (OpuCB, OpuCD) and an extracellular substrate binding protein (OpuCC) allow targeting substrate from the outer membrane (Figure 4.4.2) (Angelidis et al., 2002). In the transport reaction, highly conserved ATP-binding domains (OpuCA) couple the binding of ATP to the translocation of particular substrates across the membrane, through interaction with membrane-spanning domains (OpuCB, OpuCD) of the transporter (Figure 4.4.2) (Davidson et al., 2008; Angelidis et al., 2002). The OpuC transporter was activated by osmotic upshift and allowed to transport osmoprotectant (glycine betaine) into *P. aeruginosa* cells in order to maintain the osmotic homeostasis of the cells. The signal transduction was to control the expression of osmoreglated genes.

In this study, arabinose was added to PAORseP^{++} when culturing the bacteria for initiating transcription of *rseP* genes (3.5.1). Arabinose is a monosaccharide containing five carbon

atoms and arabinose induction in strains which may cause an increased expression of osmoreglated genes in the ABC transporter system (Weast, 1981). The increased expression of *opuC* operon (including the expression of *opuCA*, *opuCB*, *opuCC* and *opuCD*) in PAORseP^{++} (Table 3.7.3) may have been affected by arabinose induction which leads to the imbalance of the homeostasis in cells but not by RseP. Consequently, the increased expression of *opuC* gene clusters may be caused by arabinose induction in PAORseP^{++} with increasing activity of uptake glycine betaine into cells transported by the OpuC transporter. On the other hand, the upregulation of *opuC* operon may be due to the high expression of *rseP*. We hypothesized that the RseP may target the cytoplasmic membrane protein (Figure 4.4.2) and facilitate the transportation of osmoprotectant compounds into cells against the concentration to drive the hydrolysis of ATP. However, the mechanisms of OpuC systems were still not fully understood. In conclusion, we cannot determine whether RseP protease plays a role in OpuC system, or whether the increased expression of *opuC* operon is due to arabinose induction in PAORseP^{++}.

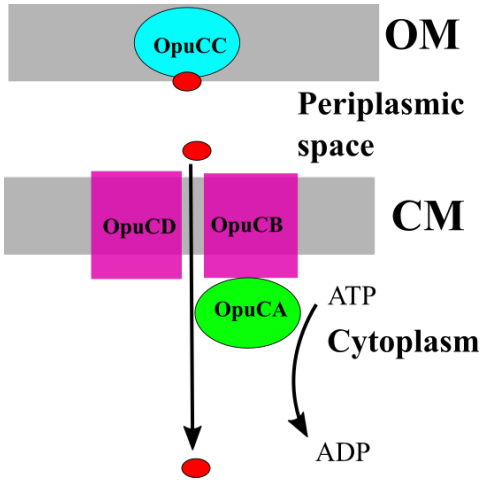


Figure4.4.2 OpuC ABC import system The compounds were exported through the outer membrane requires OpuCC. The compounds were captured by OpuCB and OpuCD in the periplasmic space and further transported to a cytoplasmic complex of OpuC which hydrolises ATP to ADP for energy coupling substrates uptake to cytoplasm (Davidson et al., 2008).

4.5 RseP and the antibiotic resistance of *P. aeruginosa*

Recent studies have reported that a highly conserved aminoglycoside resistance mechanism in clinical *P. aeruginosa* resistance strains was caused by the overproduction of *mexXY* (Hocquet et al., 2008). In *P. aeruginosa* isolates from a cystic fibrosis lung, a mutation in the *mexY* gene was found and it was observed that it led to an increase in multi-drug resistance, signifying that there is a need for MexXY in the hostile environmental of the cystic fibrosis lung (Vettoretti et al., 2009). Thus, the expression of *mexX* and *mexY* were important for developing a high

resistance level of *P. aeruginosa* to antibiotics. It was hypothesized that PAO *rseP* had high resistance to both tobramycin and ciprofloxacin with upregulated MexXY. Since more efflux transporters were presented in the cells, it is likely that an increased level of antibiotic would be exported out of the cells.

However, the findings of antibiotic resistance in this study suggested that PAO *rseP* strains had the lowest resistance to aminoglycosides and fluoroquinolones compared to other strains, despite a greater expression (13-fold) of *mexX* and *mexY* were being present in PAO *rseP* than in overexpressed RseP (3.8). Our findings in PAO *rseP* did not concur with the results of previous studies where it was found that a 48-fold increased expression of *mexXY* was associated with MIC (2-4 folds) higher than in PAO1-WT (Hocquet et al., 2008; Vettoretti et al., 2009). The low antibiotic resistance of *P. aeruginosa* strains observed in this study may be due to the small increase in expression of *mexX* and *mexY* in PAO *rseP*, which was not sufficient to lead resistance changes on *P. aeruginosa* to antibiotics. Furthermore, overexpressed RseP with the downregulation of *mexXY* increased MIC (2 $\mu g.mL^{-1}$) compared to other strains (MIC=0.05 $\mu g.mL^{-1}$) for tobramycin and ciprofloxacin, but was lower than in *P. aeruginosa* clinical isolates with MIC around 4 $\mu g.mL^{-1}$ (Hocquet et al., 2008; Brazas and Hancock., 2005).

To summarize, the higher expression of *mexXY* in PAO *rseP* does not appear to increase aminoglycosides or fluoroquinolones resistance. These findings suggested that the differences in expression of *rseP* may not cause the antibiotic resistance phenotypic changes of *P. aeruginosa.*

4.6 Conclusion

The overall aim of this research was to determine the RseP-dependent genes and pathways by differential expression analysis in *P. aeruginosa* and then analyse their biological function. The second aim was to investigate the differences in gene expression of receptor genes (*fpvA*, *foxA* and *fiuA*) altered by overexpressed RseP in *P.* aeruginosa. These experiments were shown that overproducing RseP does not increase in response to producing the transcription of receptor genes in CSS systems, especially under when siderophores are depleted. RseP is required for proteolysis of the anti-sigma factor in the pyoverdine, desferrioxamine and ferrichrome systems identified previously. In this study, 69 genes were identified with a strong transcription response to RseP from RNA-Seq analysis. Specifically, the expression of genes involved in MexXY efflux systems and OpuC ABC Osmolarity transporter systems were altered by RseP. This study reported that RseP does not increase the aminoglycosides resistance and RseP does not associate with flagellum motility phenotypes.

4.7 Future directions

The knowledge gained from this study has raised some questions about the involvement of RseP protease in both CSS pathways and other unknown regulatory pathways in *P. aeruginosa*. The results from RT-qPCR analysis suggested that the expression of receptor genes may be not altered by RseP protease. It is likely because RseP provides only one step (secondary cleavage) of the proteolytic cascade and so its overexpression may not have a major impact on enhancing the substrate protein degradation. Future research can focus on characterizing the expression of receptor genes with overexpressed DegS (site-1 protease) and ClpXP cytoplasmic protease in *P. aeruginosa respectively* by protein assay (i.e. western blots analysis). This will determine if other proteases in the proteolytic cascade would alter the receptor protein expression in CSS systems and it would be important to have a better understanding of these pathways. This study has provided a list of genes which had a strong transcriptional response to RseP. This would enable the researcher to determine which component(s) may interact with RseP and how RseP may play a role in the pathways through which those genes are presented.

Bibliography

1. Akiyama, Y., Kanehara, K. & Ito, K. RseP (YaeL), an *Escherichia coli* RIP protease, cleaves transmembrane sequences. *EMBO J.* **23,** 4434–42 (2004).

2. Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Nat. Preced.* 1–13 (2010). doi:10.1038/npre.2010.4282.1

3. Andrews, J. M. Determination of minimum inhibitory concentrations. *J. Antimic. Chemother.* **48,** Suppl. S I .5-16 (2001).

4. Andrews, S. C., Robinson, A. K. & Rodríguez-Quiñones, F. Bacterial iron homeostasis. *FEMS Microbiol. Rev.* **27,** 215–237 (2003).

5. Angelidis, A. S., Smith, L. T., Hoffman, L. M. & Smith, G. M. Identification of opuC as a chill-activated and osmotically activated carnitine transporter in Listeria monocytogenes. *Appl. Environ. Microbiol.* **68,** 2644–50 (2002).

6. Banin, E, Vasil, M.L. & Greenberg,E.P. Iron and *Pseudomonas aeruginosa* biofilm formation. *Proc Natl Acad Sci* USA **102** (31):11076-11083

7. Baker, T. A. & Sauer, R. T. ClpXP, an ATP-powered unfolding and protein-degradation machine. *Biochim. Biophys. Acta - Mol. Cell Res.* **1823,** 15–28 (2012).

8. Barbier, M. *et al.* From the environment to the host: Re-wiring of the transcriptome of *Pseudomonas aeruginosa* from 22°C to 37°C. *PLoS One* **9,** (2014).

9. Barchinger, S. E. & Ades, S. E. in 129–160 (Springer Netherlands, 2013). doi:10. 1007 /978–94–007-59404_6

10. Bastiaansen, K. C., Otero-Asman, J. R., Luirink, J., Bitter, W. & Llamas, M. A. Processing of cell-surface signalling anti-sigma factors prior to signal recognition is a conserved autoproteolytic mechanism that produces two functional domains. *Environ. Microbiol.* **17,** 3263–3277 (2015).

11. Bastiaansen, K. C., Ibañez, A., Ramos, J. L., Bitter, W. & Llamas, M. A. The Prc and RseP proteases control bacterial cell-surface signaling activity. *Environ. Microbiol.* **16,** 2433–2443 (2014).

12. Beare, P. A., For, R. J., Martin, L. W. & Lamont, I. L. Siderophore-mediated cell signaling in *Pseudomonas aeruginosa*: divergent pathways regulate virulence factor production and siderophore receptor synthesis. *Mol. Microbiol.* **47,** 195–207 (2002).

13. Bergey, D.H., F.C.Harrison, R.S.Breed, B. W. H. M. H. *Bregey's Manual of Determinative Bacteriology*. (Williams and Wilkins, 1923).

14. Bgisequence online. Basic RNA-seq processing: Unix tools and IGV. (2012). Available at: http://bio.lundberg.gu.se/courses/vt13/rnaseq.html.

15. Bishop, T. Identification of the Cytoplasmic Protease Involved in the Pyoverdine-Dependent Signal Transudction Pathway in *Pseudomonas aeruginosa (Honors)*. (University of Otago, 2012).

16. Blair, D.F. Flagellar movement driven by proton translocation. *FEBS Lett* **545**: 86–9 (2003).

17. Bohn, C., Collier, J. & Bouloc, P. Dispensable PDZ domain of Escherichia coli YaeL essential protease. *Mol. Microbiol.* **52,** 427–435 (2004).

18. Brazas, M. D. & Hancock, R. E. W. Ciprofloxacin induction of a susceptibility determinant in *Pseudomonas aeruginosa*. *Antimicrob. Agents Chemother.* **49,** 3222–7 (2005).

19. Bullen,J.J., Rogers,H.J., Spalding, P.B. & Ward,C.G. Iron and infection: the heart of the matter. FEMS Immunol Med Microbial **43**(3):325-330 (2005).

20. Chen, C. & Beattie, G. A. Characterization of the osmoprotectant transporter OpuC from *Pseudomonas syringae* and demonstration that cystathionine-beta-synthase domains are required for its osmoregulatory function. *J. Bacteriol.* **189,** 6901–12 (2007).

21. ChuanChuen, R., Narasaki, C. T. & Schweizer, H. P. *Biotechniques. BioTechniques* **33,** (2002).

22. Crosa,J.H. Signal transduction and transcriptional and posttranscriptional control of iron-regulated genes in bacteria. *Microbiol Mol Biol Rev* **61**(3):319-336 (1997)

23. Csonka, L. N. Physiological and genetic responses of bacteria to osmotic stress. *Microbiol. Rev.* **53,** 121–47 (1989).

24. Damron, F. H. & Yu, H. D. *Pseudomonas aeruginosa* MucD regulates the alginate pathway through activation of MucA degradation via MucP proteolytic activity. *J. Bacteriol.* **193,** 286–91 (2011).

25. Davidson, A. L., Dassa, E., Orelle, C. & Chen, J. Structure, function, and evolution of bacterial ATP-binding cassette systems. *Microbiol. Mol. Biol. Rev.* **72,** 317–64, table of contents (2008).

26. Davies, J. C. *Pseudomonas aeruginosa* in cystic fibrosis: pathogenesis and persistence. *Paediatr. Respir. Rev.* **3,** 128–134 (2002).

27. Draper, R. C., Martin, L. W., Beare, P. A. & Lamont, I. L. Differential proteolysis of sigma regulators controls cell-surface signalling in *Pseudomonas aeruginosa*. *Mol. Microbiol.* **82,** 1444–1453 (2011).

28. Draper, R. C. Divergent proteolysis of anti-sigma proteins controls cell-surface signalling in *Pseudomonas aeruginosa* (Doctorate). (University of Otago, 2011).

29. Dubern, J. F. *et al.* Integrated whole-genome screening for *Pseudomonas aeruginosa* virulence genes using multiple disease models reveals that pathogenicity is host specific. *Environ. Microbiol.* **17,** 4379–4393 (2015).

30. Feldman, M. *et al.* Role of flagella in pathogenesis of *Pseudomonas aeruginosa* pulmonary infection. *Infect. Immun.* **66,** 43–51 (1998).

31. Fitzsimmons, L. F., Hampel, K. J. & Wargo, M. J. Cellular choline and glycine betaine pools impact osmoprotection and phospholipase C production in *Pseudomonas aeruginosa*. *J. Bacteriol.* **194,** 4718–26 (2012).

32. Flynn, J. M., Neher, S. B., Kim, Y.-I., Sauer, R. T. & Baker, T. A. Proteomic Discovery of Cellular Substrates of the ClpXP Protease Reveals Five Classes of ClpX-Recognition Signals. *Mol. Cell* **11,** 671–683 (2003).

33. Govan, J. R. & Deretic, V. Microbial pathogenesis in cystic fibrosis: mucoid *Pseudomonas aeruginosa* and Burkholderia cepacia. Microbiol. Rev. 60, 539–574 (1996)

34. Griffith, M., Walker, J. R., Spies, N. C., Ainscough, B. J. & Griffith, O. L. Informatics for RNA Sequencing: A Web Resource for Analysis on the Cloud. *PLOS Comput. Biol.* **11,** e1004393 (2015).

34. Gruber, T. M. & Gross, C. A. Multiple Sigma Subunits and the Partitioning of Bacterial Transcription Space. *Annu. Rev. Microbiol.* **57,** 441–466 (2003).

35. Guo, Y. *et al.* How is mRNA expression predictive for protein expression? A correlation study on human circulating monocytes. *Acta Biochim. Biophys. Sin. (Shanghai).* **40,** 426–436 (2008).

36. Ha, D. G., Kuchma, S. L. & O'Toole, G. A. Plate-Based assay for swimming motility in *Pseudomonas aeruginosa*. *Methods Mol. Biol.* **1149,** 59–65 (2014).

37. Haltiner, R. C., Migneault, P. C. & Robertson, R. G. Incidence of thymidine-dependent enterococci detected on Mueller-Hinton agar with low thymidine content. *Antimicrob. Agents Chemother.* **18,** 365–8 (1980).

38. Heinrich, J. & Wiegert, T. Regulated intramembrane proteolysis in the control of extracytoplasmic function sigma factors. *Res. Microbiol.* **160,** 696–703 (2009).

39. Hocquet, D. *et al.* Relationship between antibiotic use and incidence of mexXY-OprM overproducers among clinical isolates of *Pseudomonas aeruginosa*. *Antimicrob. Agents Chemother.* **52,** 1173–1175 (2008).

40. Hocquet, D., Nordmann, P., El Garch, F., Cabanne, L. & Plésiat, P. Involvement of the MexXY-OprM efflux system in emergence of cefepime resistance in clinical strains of *Pseudomonas aeruginosa*. *Antimicrob. Agents Chemother.* **50,** 1347–51 (2006).

41. Høiby, N., Ciofu, O., & Bjarnsholt, T. Pseudomonas aeruginosa biofilms in cystic fibrosis. Future microbiology 5, 1663-1674. (2010).

42. Holm, S. Board of the Foundation of the Scandinavian Journal of Statistics A Simple Sequentially Rejective Multiple Test Procedure A Simple Sequentially Rejective Multiple Test Procedure. *Source Scand. J. Stat. Scand J Stat.* **6,** 65–70 (1979).

43. Illumina Online. Quality Scores for Next-Generation Sequencing. (2011). http:/www. Illumina.Com/Documents/Products/Technotes/Technote_Q-Scores.Pdf

44. Inaba, K. *et al.* A pair of circularly permutated PDZ domains control RseP, the S2P family intramembrane protease of *Escherichia coli*. *J. Biol. Chem.* **283,** 35042–52 (2008).

45. Islam, S., Jalal, S. & Wretlind, B. Expression of the MexXY efflux pump in amikacin-resistant isolates of *Pseudomonas aeruginosa*. *Clin. Microbiol. Infect.* **10,** 877–883 (2004).

46. Kanehara, K., Akiyama, Y. & Ito, K. Characterization of the yaeL gene product and its S2P-protease motifs in *Escherichia coli*. *Gene* **281,** 71–79 (2001).

47. Kanehara, K., Ito, K. & Akiyama, Y. YaeL (EcfE) activates the sigma(E) pathway of stress response through a site-2 cleavage of anti-sigma(E), RseA. *Genes Dev.* **16,** 2147–55 (2002).

48. Kiewitz, C. & Tümmler, B. Sequence diversity of *Pseudomonas aeruginosa*: impact on population structure and genome evolution. *J. Bacteriol.* **182,** 3125–35 (2000).

49. King, E. O., Ward, M. K. & Raney, D. E. Two simple media for the demonstration of pyocyanin and fluorescin. *J. Lab. Clin. Med.* **44,** 301–7 (1954).

50. Koide, K., Maegawa, S., Ito, K. & Akiyama, Y. Environment of the active site region of RseP, an *Escherichia coli* regulated intramembrane proteolysis protease, assessed by site-directed cysteine alkylation. *J. Biol. Chem.* **282,** 4553–4560 (2007).

51. Konings, A. F. *Pseudomonas aeruginosa* iron acquisition in cystic fibrosis (Doctorate). (University of Otago, 2010).

52. Kuma, K., Tanaka, J., Matsunaga, K. & Matsunaga, K. Effect of hydroxamate ferrisiderophore complex (ferrichrome) on iron uptake and growth of a coastal marine diatom, *Chaetoceros sociale*. *Limnol. Oceanogr.* **45,** 1235–1244 (2000).

53. Kung, V. L., Ozer, E. A. & Hauser, A. R. The accessory genome of *Pseudomonas aeruginosa*. *Microbiol. Mol. Biol. Rev.* **74,** 621–41 (2010).

54. LaBauve, A. E. & Wargo, M. J. Growth and laboratory maintenance of Pseudomonas aeruginosa. *Curr. Protoc. Microbiol.* 6E.1.1-6E.1.8 (2012). doi:10.1002/9780471729259.mc06e01s25

55. Lamont, I. L., Beare, P. A., Ochsner, U., Vasil, A. I. & Vasil, M. L. Siderophore-mediated signaling regulates virulence factor production in *Pseudomonas aeruginosa*. *Proc. Natl. Acad. Sci. U. S. A.* **99,** 7072–7 (2002).

56. Lee, J. G., Ji, M. J., Lee, O. M., Yoon., B.S., & Lim, H. Y. Nucleotide Sequence-Based Distinguishment of Sacbrood virus (SBV) using PCR Method. *J. Apic.* **29,** 43–49 (2014).

57. Levchenko, I., Grant, R. A., Flynn, J. M., Sauer, R. T. & Baker, T. A. Versatile modes of peptide recognition by the AAA+ adaptor protein SspB. *Nat. Struct. Mol. Biol.* **12,** (2005).

58. Li, X. *et al.* Cleavage of RseA by RseP requires a carboxyl-terminal hydrophobic amino acid following DegS cleavage. *Proc. Natl. Acad. Sci. U. S. A.* **106,** 14837–42 (2009).

59. Lister, P. D., Wolter, D. J. & Hanson, N. D. Antibacterial-resistant *Pseudomonas aeruginosa*: clinical impact and complex regulation of chromosomally encoded resistance mechanisms. *Clin. Microbiol. Rev.* **22,** 582–610 (2009).

60. Llamas, M. A. *et al.* The heterologous siderophores ferrioxamine B and ferrichrome activate signaling pathways in *Pseudomonas aeruginosa*. *J. Bacteriol.* **188,** 1882–91 (2006).

61. Llamas, M. A., Imperi, F., Visca, P. & Lamont, I. L. Cell-surface signaling in *Pseudomonas*: Stress responses, iron transport, and pathogenicity. *FEMS Microbiol. Rev.* **38,** 569–597 (2014).

62. MacRitchie, D. M., Buelow, D. R., Price, N. L. & Raivio, T. L. Two-component signaling and gram negative envelope stress response systems. *Advances in Experimental Medicine and Biology* **631,** 80–110 (2008).

63. Mäkelä, P. Agro-industrial uses of glycinebetaine. *Sugar Tech* **6,** 207–212 (2004).

64. Masuda, N. *et al.* Substrate specificities of MexAB-OprM, MexCD-OprJ, and MexXY-OprM efflux pumps in *Pseudomonas aeruginosa*. *Antimicrob. Agents Chemother.* **44,** 3322–7 (2000).

65. Messing, J., Crea, R. & Seeburg, P. H. A system for shotgun DNA sequencing. *Nucleic Acids Res.* **9,** 309–321 (1981).

66. Mettrick, K. A. & Lamont, I. L. Different roles for anti-sigma factors in siderophore signalling pathways of *Pseudomonas aeruginosa*. *Mol. Microbiol.* **74,** 1257–1271 (2009).

67. Meyer, J. & Abdallah, M. A. The Fluorescent Pigment of *Pseudomonas* fluorescens: Biosynthesis, Purification and Physicochemical Properties. *J. Gen. Microbiol.* **107,** 319–328 (1978).

68. Meyer, J.-M. *et al.* Siderophore typing, a powerful tool for the identification of fluorescent and nonfluorescent pseudomonads. *Appl. Environ. Microbiol.* **68,** 2745–53 (2002).

69. Miller, J. . Experiments in molecular genetics. *Cold Spring Jarbor, New York Cold Spring Harb. Lab.* (1972).

70. Morita, Y., Tomida, J. & Kawamura, Y. MexXY multidrug efflux system of *Pseudomonas aeruginosa*. *Front. Microbiol.* **3,** 408 (2012).

71. Morita, Y., Gilmour, C., Metcalf, D. & Poole, K. Translational control of the antibiotic inducibility of the PA5471 gene required for mexXY multidrug efflux gene expression in *Pseudomonas aeruginosa*. *J. Bacteriol.* **191,** 4966–75 (2009).

72. Morita, Y., Sobel, M. L. & Poole, K. Antibiotic inducibility of the MexXY multidrug efflux system of *Pseudomonas aeruginosa*: involvement of the antibiotic-inducible PA5471 gene product. *J. Bacteriol.* **188,** 1847–55 (2006).

73. Muller, C., Plésiat, P. & Jeannot, K. A two-component regulatory system interconnects resistance to polymyxins, aminoglycosides, fluoroquinolones, and β-lactams in *Pseudomonas aeruginosa*. *Antimicrob. Agents Chemother.* **55,** 1211–21 (2011).

74. Murray, T. S., Ledizet, M. & Kazmierczak, B. I. Swarming motility, secretion of type 3 effectors and biofilm formation phenotypes exhibited within a large cohort of *Pseudomonas aeruginosa* clinical isolates. *J. Med. Microbiol.* **59,** 511–20 (2010).

75. Newman, J. R. & Fuqua, C. Broad-host-range expression vectors that carry the L-arabinose-inducible *Escherichia coli araBAD* promoter and the *araC* regulator. *Gene* **227,** 197–203 (1999).

76. Noinaj, N., Guillier, M., Barnard, T. J. & Buchanan, S. K. TonB-Dependent Transporters: Regulation, Structure, and Function. *Annu. Rev. Microbiol.* **64,** 43–60 (2010).

77. Nookaew, I. *et al.* A comprehensive comparison of RNA-Seq-based transcriptome analysis from reads to differential gene expression and cross-comparison with microarrays: a case study in *Saccharomyces cerevisiae*. *Nucleic Acids Res.* **40,** 10084–10097 (2012).

78. Qiu, D., Eisinger, V. M., Rowen, D. W. & Yu, H. D. Regulated proteolysis controls mucoid conversion in *Pseudomonas aeruginosa*. *Proc. Natl. Acad. Sci. U. S. A.* **104,** 8107–12 (2007).

80. Rapaport, F. *et al.* Comprehensive evaluation of differential gene expression analysis methods for RNA-seq data. *Genome Biol.* **14,** R95 (2013).

81. Rédly, G. A. & Poole, K. FpvIR control of fpvA ferric pyoverdine receptor gene expression in *Pseudomonas aeruginosa*: demonstration of an interaction between FpvI and FpvR and identification of mutations in each compromising this interaction. *J. Bacteriol.* **187,** 5648–57 (2005).

82. Reid, D. W., O'May, C., Roddam, L. F. & Lamont, I. L. Chelated iron as an anti-*Pseudomonas aeruginosa* biofilm therapeutic strategy. *J. Appl. Microbiol.* **106,** 1058–1058 (2009).

83. Reinhardt, A. *et al.* Development and persistence of antimicrobial resistance in *Pseudomonas aeruginosa*: A longitudinal observation in mechanically ventilated patients. *Antimicrob. Agents Chemother.* **51,** 1341–1350 (2007).

84. Robinson, J. T. *et al.* Integrative genomics viewer. *Nat. Biotechnol.* **29,** 24–26 (2011).

85. Rojo, F. Carbon catabolite repression in *Pseudomonas*: Optimizing metabolic versatility and interactions with the environment. *FEMS Microbiology Reviews* **34,** 658–684 (2010).

86. Rowley, G., Spector, M., Kormanec, J. & Roberts, M. Pushing the envelope: extracytoplasmic stress responses in bacterial pathogens. *Nat. Rev. Microbiol.* **4,** 383–394 (2006).

87. Sambrook, J. & Russell, D. W. (David W. *Molecular cloning : a laboratory manual*. (Cold Spring Harbor Laboratory Press, 2001).

88. Schurr, M. J. & Deretic, V. Microbial pathogenesis in cystic fibrosis: co-ordinate regulation of heat-shock response and conversion to mucoidy in *Pseudomonas aeruginosa*. *Mol. Microbiol.* **24,** 411–420 (1997).

89. Roche Applied Science, . LightCycler® 480 Instrument Operator's Manual. *Roche Appl. Sci.* 1–395 (2007).

90. Shapiro, L. The bacterial flagellum: From genetic network to complex architecture. *Cell* **80,** 525–527 (1995).

91. Soneson, C. & Delorenzi, M. A comparison of methods for differential expression analysis of RNA-seq data. *BMC Bioinformatics* **14,** 91 (2013).

92. Spencer, D. H. *et al.* Whole-genome sequence variation among multiple isolates of *Pseudomonas aeruginosa*. *J. Bacteriol.* **185,** 1316–25 (2003).

93. Stover, C. K. *et al.* Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic pathogen. *Nature* **406,** 959–964 (2000).

94. Assay Technologies. RNAprotect® Cell Reagent Handbook. *Cell* (2007).

95. Vettoretti, L. *et al.* Efflux unbalance in *Pseudomonas aeruginosa* isolates from cystic fibrosis patients. *Antimicrob. Agents Chemother.* **53,** 1987–97 (2009). Visca P, Imperi F & Lamont IL. Pyoverdine siderophores: from biogenesis to biosignificance. *Trends Microbiol* **15**:22–30 (2007).

96. Visca P, Imperi F & Lamont IL. Pyoverdine siderophores: from biogenesis to biosignificance. *Trends Microbiol* **15**:22–30 (2007).

97. Wang, Y. *et al.* Evaluation of the coverage and depth of transcriptome by RNA-Seq in chickens. doi:10.1186/1471-2105-12-S10-S5

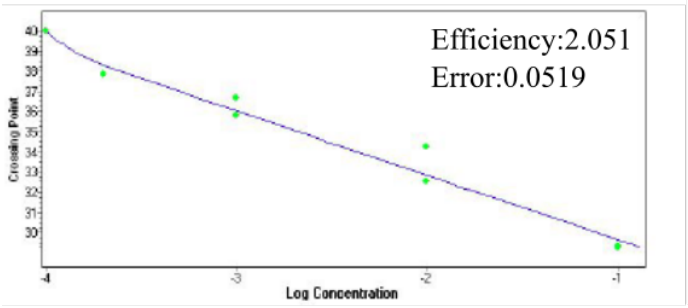
98. Wargo, M. J. Homeostasis and catabolism of choline and glycine betaine: lessons from *Pseudomonas aeruginosa*. *Appl. Environ. Microbiol.* **79,** 2112–20 (2013).

99. Weast, R. Physical constants of organic chemicals. In CRC Handbook of Chemistry and Physics. *CRC Press*. C51-C576 (1981).

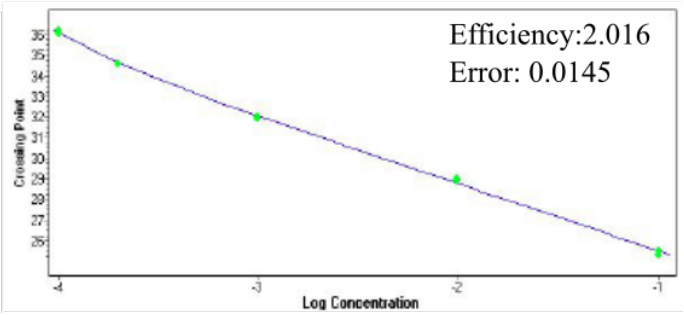
Appendixes

Appendix 1: qPCR Standard Curves Standard curves were produced using serial diluted genomic template as described previously (3.7.6).

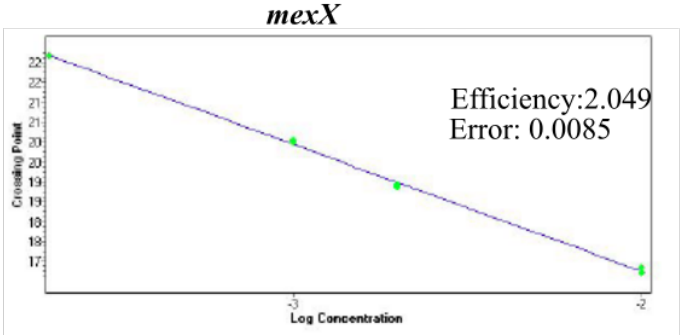
PA2275



PA5471



mexX



Appendix 2: Alignment for the pJN105::*rseP*$^{++}$ insert sequence and the *rseP* gene from NCBI database. The *rseP* insert gene from the l pJN105::*rseP*++ construct had 99% matched to the *rseP* sequence in NCBI database. The query is the insert sequence from the pJN105::*rseP*$^{++}$ plasmid. The subject stand for *rseP* sequence from database (with both forward and reverse alignment)

Forward sequence

Pseudomonas aeruginosa strain PAO1 MucP (mucP) gene, complete cds

Sequence ID: EF527878.1 Length: 1353 Number of Matches: 1

Range 1: 1 to 787 GenBank Graphics Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
1437 bits(778)	0.0	784/787(99%)	1/787(0%)	Plus/Plus

```
Query  156  ATGAGTGCGCTTTACATGATCGTCGGCACCCTGGTAGCCCTGGGTGTGCTGGTGACGTTC  215
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1    ATGAGTGCGCTTTACATGATCGTCGGCACCCTGGTAGCCCTGGGTGTGCTGGTGACGTTC  60

Query  216  CACGAATTCGGCCACTTCTGGGTCGCGCGACGCTGCGGGGTGAAGGTCCTGCGCTTCTCC  275
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  61   CACGAATTCGGCCACTTCTGGGTCGCGCGACGCTGCGGGGTGAAGGTCCTGCGCTTCTCC  120

Query  276  GTCGGCTTCGGCACGCCGCTGGTGCGCTGGCACGACCGGCACGGCACCGAGTTCGTGGTG  335
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  121  GTCGGCTTCGGCACGCCGCTGGTGCGCTGGCACGACCGGCACGGCACCGAGTTCGTGGTG  180

Query  336  GCGGCCATTCCGCTGGGCGGCTACGTCAAGATGCTCGACGAGCGCGAGGCGGAAGTACCG  395
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  181  GCGGCCATTCCGCTGGGCGGCTACGTCAAGATGCTCGACGAGCGCGAGGCGGAAGTACCG  240

Query  396  GCCCATCTGCTCGAACAGTCCTTCAATCGTAAGACGGTCCGTCAGCGGATCGCCATCGTC  455
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  241  GCCCATCTGCTCGAACAGTCCTTCAATCGTAAGACGGTCCGTCAGCGGATCGCCATCGTC  300

Query  456  GCGGCTGGTCCCATCGCCAATTTCCTACTGGCGATTCTGTTCTTCTGGGTCGTCGCCCTG  515
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  301  GCGGCTGGTCCCATCGCCAATTTCCTACTGGCGATTCTGTTCTTCTGGGTCGTCGCCCTG  360

Query  516  CTCGGTAGCCAGCAGGTACGCCCGGTCATCGGTTCCGTGGCGCCCGGAGAGCCTGGCCGCC  575
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  361  CTCGGTAGCCAGCAGGTACGCCCGGTCATCGGTTCCGTGGCGCCCGGAGAGCCTGGCCGCC  420

Query  576  CAGGCCGGCCTGGAGGCCGGGCAGGAGCTGCTGGCGGTGGATGGCGAGCCCGTCACCGGC  635
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  421  CAGGCCGGCCTGGAGGCCGGGCAGGAGCTGCTGGCGGTGGATGGCGAGCCCGTCACCGGC  480

Query  636  TGGAACGGGGTCAACCTGCAGCTGGTTCGCCGTCTCGGCGAGAGCGGTACGCTGGAAGTG  695
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  481  TGGAACGGGGTCAACCTGCAGCTGGTTCGCCGTCTCGGCGAGAGCGGTACGCTGGAAGTG  540

Query  696  CGGGTGCAGGAAAAAGGCTCGAATGTCGATTCCACCCAYCAGGTGCGTCTGGATGGCTGG  755
            |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Sbjct  541  CGGGTGCAGGAAAAAGGCTCGAATGTCGATTCCACCCACCAGGTGCGTCTGGATGGCTGG  600

Query  756  CTGAAGGGCGAGGACAATCCCGATCCCATCGCCTCGCTGGGCATCCGGCCCTGGCGGCCG  815
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  601  CTGAAGGGCGAGGACAATCCCGATCCCATCGCCTCGCTGGGCATCCGGCCCTGGCGGCCG  660

Query  816  GCCTTGCCGCCGGTGCTGGCCGAACTCGACCCCAAGGGGCCGGCGCA-GCCGCCGGGCTG  874
            |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Sbjct  661  GCCTTGCCGCCGGTGCTGGCCGAACTCGACCCCAAGGGGCCGGCGCAGGCCGCCGGGCTG  720

Query  875  AAGCTCGGCGACCGCTTGCAGAGCATCGATGGCATCGCCGTCGACGACTGGCAGCAGGTG  934
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  721  AAGCTCGGCGACCGCTTGCAGAGCATCGATGGCATCGCCGTCGACGACTGGCAGCAGGTG  780

Query  935  TTGGACA  941
            |||||
Sbjct  781  GTGGACA  787
```

Reverse sequence

Pseudomonas aeruginosa strain PAO1 MucP (mucP) gene, complete cds

Sequence ID: EF527878.1 **Length:** 1353 **Number of Matches:** 1

Range 1: 478 to 1353 GenBank Graphics Next Match Previous

Score	Expect	Identities	Gaps	Strand
1611 bits(872)	0.0	874/876(99%)	0/876(0%)	Plus/Minus

```
Query  55    CTACAGACGACTCAGATCGTTGACCAGAGCGAGCAACATCACGCCAACGACCAGACTGAT  114
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1353  CTACAGACGACTCAGATCGTTGACCAGAGCGAGCAACATCACGCCAACGACCAGACTGAT  1294

Query  115   GCCGATCTGCATCCCCCAAGCCTGTACCCGTTCGGATAATGGGCGCCCGCGCACCCACTC  174
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1293  GCCGATCTGCATCCCCCAAGCCTGTACCCGTTCGGATAATGGGCGCCCGCGCACCCACTC  1234

Query  175   AACCAGATAGAACAGCAGATGCCCGCCATCGAGGACAGGGATCGGCAGCAGGTTGAGAAC  234
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1233  AACCAGATAGAACAGCAGATGCCCGCCATCGAGGACAGGGATCGGCAGCAGGTTGAGAAC  1174

Query  235   CCCCAAGCTAATGCTCAGATAGGCCAGGAAATTGAGAAAATCCCCCACGCCAGACTGGGC  294
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1173  CCCCAAGCTAATGCTCAGATAGGCCAGGAAATTGAGAAAATCCCCCACGCCAGACTGGGC  1114

Query  295   TGAAGCGCCCGCCACTTTAGCAATGGTTATCGGCCCGCTCAAGTTTTTTACCGAGAGCTC  354
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1113  TGAAGCGCCCGCCACTTTAGCAATGGTTATCGGCCCGCTCAAGTTTTTTACCGAGAGCTC  1054

Query  355   GCCGAGCAACATTTTCTTTATGGAGTCGAGCGTTAGCAGGCTCATGGTCCAGGTTCGGGA  414
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1053  GCCGAGCAACATTTTCTTTATGGAGTCGAGCGTTAGCAGGCTCATGGTCCAGGTTCGGGA  994

Query  415   AAGCGCCTGACCCACCGCCTCCAGCGGTCCGTAGCTCACTTCGCGGAGCATTTCCGCCGG  474
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  993   AAGCGCCTGACCCACCGCCTCCAGCGGTCCGTAGCTCACTTCGCGGAGCATTTCCGCCGG  934

Query  475   CCATTCGGTACCGGCGACCCCCGCCCCCCATGTAGCCGCTCCGGGCCTTGCCCTCGCCGCG  534
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  933   CCATTCGGTACCGGCGACCCCCGCCCCCCATGTAGCCGCTCCGGGCCTTGCCCTCGCCGCG  874

Query  535   CACCGCCAGTTCCAGGGCAACATCCAGCACTTCGCCGTCGCGCAGCACCTTGAGCTGGAC  594
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  873   CACCGCCAGTTCCAGGGCAACATCCAGCACTTCGCCGTCGCGCAGCACCTTGAGCTGGAC  814

Query  595   GCGCTGCCCGGGACGGGCACGGACGCTGTCCACCACCTGCTGCCAGTCGTCGACGGCGAT  654
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  813   GCGCTGCCCGGGACGGGCACGGACGCTGTCCACCACCTGCTGCCAGTCGTCGACGGCGAT  754

Query  655   GCCATCGATGCTCTGCAAGCGGTCGCCGAGCTTCAGCCCGGCGGCCTGCGCCGGCCCCTT  714
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  753   GCCATCGATGCTCTGCAAGCGGTCGCCGAGCTTCAGCCCGGCGGCCTGCGCCGGCCCCTT  694

Query  715   GGGGTCGAGTTCGGCCAGCACCGGCGGCAAGGCCGGCCGCCAGGGCCGGATGCCCAGCGA  774
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  693   GGGGTCGAGTTCGGCCAGCACCGGCGGCAAGGCCGGCCGCCAGGGCCGGATGCCCAGCGA  634

Query  775   GGCGATGGGATCGGGATTGTCCTCGCCCTTCAGCCAGCCATCCAGACGCACCTGGTGGGT  834
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  633   GGCGATGGGATCGGGATTGTCCTCGCCCTTCAGCCAGCCATCCAGACGCACCTGGTGGGT  574

Query  835   GGAATCGACATTCGAGCCTTTTTCCTGCACCCGCACTTCCAGCGTACCGCTCTCRMCGAG  894
             ||||||||||||||||||||||||||||||||||||||||||||||||||||  ||||||
Sbjct  573   GGAATCGACATTCGAGCCTTTTTCCTGCACCCGCACTTCCAGCGTACCGCTCTCGCCGAG  514

Query  895   ACGGCGAACCAGCTGCAGGTTGACCCCGTTCCAGCC  930
             ||||||||||||||||||||||||||||||||||||
Sbjct  513   ACGGCGAACCAGCTGCAGGTTGACCCCGTTCCAGCC  478
```

Appendix 3: The genes identified between PAO(*clpP*) vs PAO1-WT

id	Fold Change	log2Fold Change	p-adj	Functions
PA3061	8.78	3.13	5.44E-42	PelD
PA1899	5.02	2.33	3.75E-03	probable phenazine biosynthesis protein
PA0547	4.41	2.14	2.40E-09	probable transcriptional regulator
PA1159	4.39	2.13	3.43E-03	probable cold-shock protein
PA3450	4.38	2.13	1.05E-03	probable antioxidant protein
PA2204	4.07	2.03	1.30E-16	probable binding protein component of ABC transporter
PA3601	0.25	-2.01	8.40E-04	conserved hypothetical protein
PA0525	0.25	-2.01	1.82E-04	probable dinitrification protein NorD
PA0459	0.24	-2.03	7.19E-11	probable ClpA/B protease ATP binding subunit
PA2153	0.24	-2.05	7.21E-05	1,4-alpha-glucan branching enzyme
PA5353	0.24	-2.07	1.02E-04	glycolate oxidase subunit GlcF
PA0836	0.24	-2.08	1.64E-06	acetate kinase
PA3573	0.24	-2.08	2.59E-05	probable major facilitator superfamily (MFS) transporter
PA3890	0.23	-2.13	7.67E-07	OpuC ABC transporter, permease protein, OpuCB
PA0460	0.23	-2.14	2.65E-14	hypothetical protein
PA0530	0.22	-2.20	4.31E-22	probable class III pyridoxal phosphate-dependent aminotransferase
PA1546	0.22	-2.22	3.29E-04	oxygen-independent coproporphyrinogen III oxidase
PA4067	0.21	-2.23	5.87E-08	Outer membrane protein OprG precursor
PA3415	0.21	-2.24	2.51E-04	probable dihydrolipoamide acetyltransferase
PA4204	0.21	-2.25	3.91E-15	periplasmic gluconolactonase, PpgL
PA3877	0.21	-2.25	1.36E-02	nitrite extrusion protein 1
PA5354	0.21	-2.25	8.44E-06	glycolate oxidase subunit GlcE
PA2321	0.21	-2.25	1.82E-05	gluconokinase
PA4172	0.21	-2.27	6.31E-11	probable nuclease
PA4919	0.21	-2.28	1.59E-19	nicotinate phosphoribosyltransferase
PA0527	0.21	-2.29	1.01E-04	transcriptional regulator Dnr
PA2127	0.20	-2.29	3.40E-04	cupA gene regulator A, CgrA
PA0836.1	0.20	-2.30	2.21E-03	P5
PA3879	0.20	-2.30	1.39E-03	two-component response regulator NarL
PA0197	0.20	-2.31	2.02E-11	TonB2

PA0198	0.19	-2.36	2.27E-04	transport protein ExbB
PA5483	0.19	-2.36	6.03E-19	two-component response regulator AlgB
PA5373	0.19	-2.36	1.95E-17	betaine aldehyde dehydrogenase
PA2500	0.19	-2.37	2.42E-03	probable major facilitator superfamily (MFS) transporter
PA5474	0.19	-2.38	3.21E-05	probable metalloprotease
PA5173	0.19	-2.39	4.93E-12	carbamate kinase
PA0119	0.19	-2.40	5.57E-10	probable dicarboxylate transporter
PA4837	0.19	-2.40	2.15E-04	probable outer membrane protein precursor
PA3972	0.19	-2.41	6.42E-23	probable acyl-CoA dehydrogenase
PA0105	0.19	-2.42	5.42E-04	cytochrome c oxidase, subunit II
PA4571	0.18	-2.45	2.56E-03	probable cytochrome c
PA3417	0.18	-2.45	1.06E-18	probable pyruvate dehydrogenase E1 component, alpha subunit
PA0106	0.18	-2.46	4.45E-03	cytochrome c oxidase, subunit I
PA3460	0.18	-2.51	5.15E-14	probable acetyltransferase
PA0519	0.17	-2.52	3.93E-03	nitrite reductase precursor
PA3416	0.17	-2.57	1.37E-10	probable pyruvate dehydrogenase E1 component, beta chain
PA2126	0.17	-2.59	1.76E-03	cupA gene regulator C, CgrC
PA3889	0.16	-2.62	2.02E-11	OpuC ABC transporter, periplasmic substrate-binding protein, OpuCC
PA0518	0.16	-2.62	5.33E-04	cytochrome c-551 precursor
PA3888	0.16	-2.64	9.91E-10	OpuC ABC transporter, permease protein, OpuCD
PA0506	0.16	-2.65	4.15E-17	probable acyl-CoA dehydrogenase
PA2138	0.16	-2.65	3.80E-13	Multifunctional non-homologous end joining protein LigD
PA4876	0.16	-2.69	4.84E-24	osmotically inducible lipoprotein OsmE
PA4596	0.15	-2.70	6.20E-03	EsrC
PA2663	0.15	-2.70	7.48E-04	psl and pyoverdine operon regulator, PpyR
PA3891	0.15	-2.71	4.64E-09	OpuC ABC transporter, ATP-binding protein, OpuCA
PA5172	0.15	-2.72	1.70E-12	ornithine carbamoyltransferase, catabolic
PA0843	0.15	-2.75	1.43E-17	phospholipase accessory protein PlcR precursor
PA1557	0.15	-2.75	2.62E-03	Cytochrome c oxidase, cbb3-type, CcoN subunit
PA3006	0.15	-2.75	1.43E-20	transcriptional regulator PsrA
PA1541	0.14	-2.79	4.04E-33	probable drug efflux transporter
PA1187	0.14	-2.79	1.40E-19	probable acyl-CoA dehydrogenase

PA2126.1	0.14	-2.86	4.14E-03	cupA gene regulator B, CgrB
PA5297	0.14	-2.86	7.84E-11	pyruvate dehydrogenase (cytochrome)
PA4587	0.14	-2.87	4.00E-03	cytochrome c551 peroxidase precursor
PA1555.1	0.13	-2.90	4.48E-03	Cytochrome c oxidase, cbb3-type, CcoQ subunit
PA3459	0.13	-2.95	3.34E-28	probable glutamine amidotransferase
PA5212	0.12	-3.08	4.87E-33	hypothetical protein
PA3458	0.12	-3.10	1.72E-05	probable transcriptional regulator
PA3880	0.12	-3.10	1.26E-06	conserved hypothetical protein
PA5170	0.11	-3.14	9.05E-09	arginine/ornithine antiporter
PA0059	0.11	-3.16	1.05E-36	osmotically inducible protein OsmC
PA1555	0.11	-3.17	2.49E-03	Cytochrome c oxidase, cbb3-type, CcoP subunit
PA1737	0.11	-3.17	1.92E-32	probable 3-hydroxyacyl-CoA dehydrogenase
PA3913	0.11	-3.24	2.35E-05	probable protease
PA3337	0.11	-3.25	6.69E-05	ADP-L-glycero-D-mannoheptose 6-epimerase
PA5171	0.10	-3.27	7.72E-17	arginine deiminase
PA1736	0.10	-3.35	3.09E-42	probable acyl-CoA thiolase
PA3515	0.10	-3.38	1.33E-04	hypothetical protein
PA0355	0.09	-3.51	1.16E-17	protease PfpI
PA4880	0.08	-3.57	8.74E-14	probable bacterioferritin
PA3692	0.08	-3.58	8.71E-14	Lipotoxon F, LptF
PA0526	0.08	-3.63	2.26E-08	hypothetical protein
PA2119	0.08	-3.67	7.07E-05	alcohol dehydrogenase (Zn-dependent)
PA3517	0.08	-3.67	1.27E-04	probable lyase
PA1196	0.08	-3.68	4.81E-04	probable transcriptional regulator
PA3305.1	0.08	-3.70	8.25E-06	PhrS
PA4358	0.07	-3.92	7.56E-07	probable ferrous iron transport protein
PA1920	0.07	-3.93	4.66E-03	class III (anaerobic) ribonucleoside-triphosphate reductase subunit, NrdD
PA1556	0.06	-3.95	7.49E-04	Cytochrome c oxidase, cbb3-type, CcoO subunit
PA1429	0.06	-4.01	1.90E-04	probable cation-transporting P-type ATPase
PA0199	0.06	-4.04	1.45E-05	transport protein ExbD
PA1137	0.05	-4.20	1.77E-03	probable oxidoreductase
PA5427	0.05	-4.24	1.98E-08	alcohol dehydrogenase
PA4878	0.05	-4.37	1.37E-12	BrlR
PA3574	0.04	-4.55	2.79E-13	NalD

PA4350	0.04	-4.62	1.89E-05	OlsB
PA0141	0.04	-4.62	2.57E-08	conserved hypothetical protein
PA3920	0.04	-4.72	1.95E-12	probable metal transporting P-type ATPase
PA3516	0.03	-4.84	1.08E-05	probable lyase
PA4351	0.02	-5.51	3.69E-08	OlsA
PA3521	0.02	-5.85	3.63E-05	probable outer membrane protein precursor
PA3574.1	0.02	-5.99	3.11E-19	Uncharacterized protein
PA3522	0.01	-6.51	4.67E-08	probable Resistance-Nodulation-Cell Division (RND) efflux transporter
PA3523	0.01	-7.14	4.16E-35	probable Resistance-Nodulation-Cell Division (RND) efflux membrane fusion protein precursor

Appendix 4: The genes identified between PAORseP^{++} vs PAO(*clpP*) and PAORseP^{++} vs PAO *rseP* groups

	PAORseP^{++} vs PAO(*clpP*)		PAORseP^{++}vs PAO *rseP*	
	log2Fold change	p-adj	log2Fold change	p-adj
PA2294	9.02	1.15E-08	9.44	3.26E-09
PA3649	7.0	2.56E-05	9.9	5.63E-08
PA2275	4.06	3.15E-04	2.51	3.10E-02
PA3650	4.06	2.76E-03	4.34	2.55E-03
PA3891	3.85	8.95E-04	2.79	2.43E-02
PA3044	3.62	3.61E-07	3.58	3.40E-06
PA3889	2.98	1.94E-05	2.35	6.31E-04
PA4880	2.95	8.89E-06	4.13	3.26E-09
PA1921	2.91	1.81E-06	3.16	4.92E-07
PA0567	2.79	4.60E-04	3.25	1.71E-05
PA1592	2.6	2.60E-02	5.7	1.04E-08
PA2751	2.41	5.68E-04	2.67	2.96E-04
PA1471	2.39	4.79E-04	2.22	3.33E-04
PA0737	2.37	5.65E-04	3.47	3.81E-07
PA3888	2.36	4.61E-04	2.84	2.57E-05
PA0843	2.28	3.19E-04	2.73	1.35E-05
PA1606	2.26	5.65E-04	1.86	4.92E-07
PA2415	2.15	2.20E-03	3.83	5.63E-08
PA3890	2.09	3.31E-03	2.52	4.82E-03
PA5297	1.99	7.00E-03	2.36	2.00E-03
PA4204	1.98	6.86E-03	2.31	1.90E-03
PA0738	1.75	8.36E-03	1.66	2.40E-02
PA5182	1.72	3.97E-02	2.81	1.51E-04
PA4220	-2.49	3.10E-04	-3.83	2.97E-06
PA4219	-2.84	2.96E-04	-3.99	1.40E-05
PA0062	2.87	1.71E-05	2.87	1.71E-05