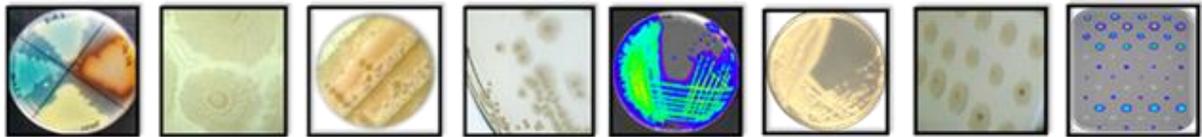




**Investigating the genetic basis of
preservative resistance in an industrial
Pseudomonas aeruginosa strain**



Thesis presented for the Degree of Philosophiae Doctor by

**Angharad Ellen Green, BSc (Hons) Equine Science,
MSc Biomedical Science (Medical Microbiology)**

In candidature for the Degree of Philosophiae Doctor

Organisms and the Environment Division
Cardiff School of Biosciences
Cardiff University

September 2017



“If you want something you've never had,
you must be willing to do something
you've never done.” — Thomas Jefferson.

Declaration

This work has not been submitted in substance for any other degree or award at this or any other university or place of learning, nor is being submitted concurrently in candidature for any degree or other award.

Signed AEGreen (candidate) Date: 27th September 2017

STATEMENT 1

This thesis is being submitted in partial fulfillment of the requirements for the degree of PhD.

Signed AEGreen (candidate) Date: 27th September 2017

STATEMENT 2

This thesis is the result of my own independent work/investigation, except where otherwise stated, and the thesis has not been edited by a third party beyond what is permitted by Cardiff University's Policy on the Use of Third Party Editors by Research Degree Students. Other sources are acknowledged by explicit references. The views expressed are my own.

Signed AEGreen (candidate) Date: 27th September 2017

STATEMENT 3

I hereby give consent for my thesis, if accepted, to be available online in the University's Open Access repository and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed AEGreen (candidate) Date: 27th September 2017

STATEMENT 4: PREVIOUSLY APPROVED BAR ON ACCESS

I hereby give consent for my thesis, if accepted, to be available online in the University's Open Access repository and for inter-library loans **after expiry of a bar on access previously approved by the Academic Standards & Quality Committee.**

Signed AEGreen (candidate) Date: 27th September 2017

Acknowledgments

I would like to express my deepest thanks to the following people:

- My primary supervisor Professor. Eshwar Mahenthiralingam for his encouragement, direction and much appreciated enthusiasm throughout my PhD project. It has been a pleasure and a great opportunity to be a part of the Mahenthiralingam lab, and the experience and knowledge I have gained will take me forward into my future career.
- My secondary Cardiff supervisor Dr. Thomas Connor, for his much appreciated advice on bioinformatics and providing the Cloud Infrastructure for Medical Microbial Bioinformatics (CLIMB).
- My industrial supervisors Dr. Alejandro Amézquita and Dr. Yvan Le Marc for their guidance on industry procedures and support whilst on placement at Unilever Safety and Environment Assurance Centre (SEAC), Colworth. I would like to express my gratitude to everyone else at SEAC for their support during my PhD and for being very welcoming during my placements.
- Professor Jason Papin for hosting me on a 3-month placement at the University of Virginia and providing training in genome-scale metabolic modelling. I would also like to express my thanks to Ms. Anna Blazier for her modelling expertise and much appreciated advice. Also thanks to everyone in the Papin lab for providing a friendly and enjoyable environment whilst on placement.
- The Biotechnology and Biological Sciences Research Council (BBSRC) and industrial CASE sponsor Unilever SEAC, Colworth as my PhD funding bodies.
- Dr. Matthew Bull for his much valued advice and extensive expertise during the genome sequencing and RNA-Seq components of this project. A special thanks also to Dr. Angela Marchbank of the Cardiff Genome Research Hub, for her guidance whilst planning and carrying out the RNA-Seq experiments.
- My microbiology buddies, Dr. Rebecca Weiser, Dr. Laura Osbourne and Dr. Laura Ruston for providing such a fun and supportive work environment, and for sharing their expert knowledge throughout my PhD. I have enjoyed all the laughs and adventures along the way. My time at Cardiff University would not have been the same without you guys. I look forward to continuing our friendships long into the future.
- My family for their love and encouragement; a special thanks goes to my Mum, Annette, and Dad, Chris, for your unwavering support. Also thanks for delivering all the inspiring quotes such as “life is an adventure, take on the challenge!” and “dare to dream”. I would not have got this far in my career without you!
- My best friend Sophie Carnihan (aka Greeny) for your invaluable friendship and for putting up with my science talk. Also, thanks for providing all the fun times when I needed to take a break from PhD life.
- Everyone in the Microbiomes, Microbes and Informatics Group at Cardiff University who have made my time in the lab lots of fun with a special thanks to Dr. Cerith Jones who has offered much needed guidance and advice throughout my PhD.
- My MSc supervisor Dr Sarah Maddocks for your help in developing my skills during the early stages of my research career and for her encouragement to publish my project findings.
- My Brynteg Comprehensive School A-Level biology teacher, Mrs Phillips who established my enthusiasm for science.

Scientific Conferences and Awards

Poster presentation: Using bacterial biosensors to understand the genetic basis for antimicrobial resistance. Speaking of Science, Student led Conference, 2014, Cardiff, Wales (**Awarded 1st prize**).

Poster presentation: Using bacterial biosensors to understand the genetic basis for antimicrobial resistance. Microbiology and Infection Translational Research Group (MITReG), Annual Event, 2015, Cardiff, Wales.

Oral presentation: Using bacterial biosensors to understand the genetic basis for antimicrobial resistance. Society for General Microbiology (SGM), Annual Conference, 2015, Birmingham, UK.

Oral presentation: Using bacterial biosensors to understand the genetic basis for antimicrobial resistance. Cardiff School of Biosciences Organisms and Environment Division Away Day, 2015 (**Awarded 1st prize**).

Oral presentation: Using bacterial biosensors to understand the genetic basis for antimicrobial resistance. Speaking of Science, Student led Conference, 2015, Cardiff, Wales.

Poster presentation: Using bacterial biosensors to understand the genetic basis for antimicrobial resistance. The Learned Society Partnership on Antimicrobial Resistance, Networking workshop for researchers, 2015, London, UK.

Poster presentation: Using bacterial biosensors to understand the genetic basis for antimicrobial resistance. American Society for Microbiology, *Pseudomonas* Conference, 2015, Washington DC, USA.

Oral presentation: Using bacterial biosensors to understand the genetic basis for antimicrobial resistance." British Society for Antimicrobial Chemotherapy (BSAC), Antibiotic Resistance Mechanisms Workshop for researchers, 2015, Birmingham, UK.

<http://www.bsac.org.uk/meetings/arm-workshop-for-researchers-arm-2015/>.

Poster presentation: Harnessing the power of bacterial bioluminescence to shine a light on antimicrobial resistance. Junior Award for Microbiology (JAM) Talk (sponsored by the Institute of Microbiology and Infection (IMI) and the Microbiology Society), 2016, University of Birmingham, UK. <https://pbs.twimg.com/media/CdwwuuZW0AE8OHD.jpg:large>.

Poster presentation: Harnessing the power of bacterial bioluminescence to shine a light on antimicrobial resistance. Microbiology Society Annual Conference, 2016, Birmingham, UK (**Awarded society conference travel grant**).

Travel Grant: Awarded the Cardiff School of Biosciences Gillian Powell Memorial Travel Scholarship Grant (2017) to undertake a 3-month international placement at the University of Virginia (UVA) Biomedical Engineering department, Charlottesville USA (January 2017-March 2017).

Poster presentation: Mapping the resistance and metabolic pathways associated with preservative exposure in an industrial *Pseudomonas aeruginosa* strain. *Pseudomonas* Conference, 2017, Liverpool, UK (**Awarded society conference travel grant**).

Summary

Pseudomonas aeruginosa is a common industrial contaminant associated with costly recalls of home and personal care (HPC) products. Preservation systems are used to prevent bacterial contamination and protect consumers, but little is known about the mechanisms of preservative resistance in *P. aeruginosa*. The aim of this research was to map genetic and metabolic pathways associated with preservative resistance and bacterial growth in HPC products.

The genome of the industrial strain *P. aeruginosa* RW109 was sequenced, functionally annotated, and compared to other strains of the species. This revealed the first complete genome of a *P. aeruginosa* isolate from the HPC industry. Comparative analysis with 102 *P. aeruginosa* strains from various sources, showed industrial strains' genomes to be significantly larger than clinical and environmental strains, and RW109's genome was the largest of the species (7.8 Mbp) and included two plasmids.

Identification of differentially expressed genes by RNA-Seq (more informative than mini-Tn5-*luxCDABE* mutagenesis), revealed complex genetic networks utilised by RW109 when exposed to benzisothiazolone (BIT), phenoxyethanol (POE) and a laundry detergent formulation. Differential expression of five sets of genes was consistently observed in response to these industry relevant conditions - MexPQ-OpmE efflux pump, sialic acid transporter and isoprenoid biosynthesis (*gnyRDBHAL*) genes were frequently up-regulated; whereas *phnBA* and *pqsEDCBA* genes encoding PQS production and quorum-sensing, respectively, were consistently down-regulated. Genome-scale metabolic network reconstruction of RW109, the first with a *P. aeruginosa* industrial strain, along with integration of transcriptomic data, predicted essential pathways for RW109's preservative resistance (e.g. cell membrane phospholipid biosynthesis as a key pathway for POE resistance).

This study highlights the utility of integrating genomic, transcriptomic and metabolic modelling approaches to uncover the basis of industrial bacterial resistance to preservative and product formulations. The ability to predict the metabolic basis of *P. aeruginosa* preservative resistance will inform the development of targeted industrial preservation systems, enhancing product safety and minimising future resistance development.

Contents

Declaration.....	I
Acknowledgments.....	II
Scientific Conferences and Awards.....	III
Summary.....	IV
List of Tables.....	XIII
List of Figures	XV
Abbreviations	XVIII
1. Introduction	1
1.1 <i>Pseudomonas aeruginosa</i>	1
1.1.1 The environmental versatility of <i>P. aeruginosa</i>	2
1.1.2 Linking genome size with the complex biological functionalities of <i>P. aeruginosa</i>	3
1.2 <i>P. aeruginosa</i> as a multidrug resistant microorganism.....	4
1.2.1 Antimicrobial resistance mechanisms of <i>P. aeruginosa</i>	4
1.2.1.1 Obstructed entry into the cell.....	5
1.2.1.2 Reducing intracellular accumulation of antimicrobials	7
1.2.1.2.1 Efflux pumps of <i>P. aeruginosa</i>	7
1.2.1.2.2 Modification of the Gram-negative cell wall	10
1.2.1.3 Preventing antimicrobials within the cell reaching target sites	10
1.2.1.3.1 Inactivating antimicrobials with enzymes	10
1.2.1.3.2 Modification of antimicrobial target sites	11
1.2.1.4 <i>P. aeruginosa</i> biofilm formation.....	13
1.3 The challenge of microbial contamination in home and personal care industry.....	17
1.3.1 Microbial contamination in home and personal care products	17
1.3.2 Preventing microbial contamination in the HPC industry	18
1.3.2.1 Isothiazolinone preservatives	19
1.3.2.2 Alcohol-based preservatives	20
1.3.3 <i>P. aeruginosa</i> as a contaminant of the HPC industry.....	22
1.4 Mechanisms of resistance to industrial antimicrobial agents.....	23
1.4.1 Alteration of cell wall structures in response to industrial antimicrobials.....	23
1.4.2 The use of efflux pumps in response to industrial antimicrobials	24
1.4.3 Alteration of bacterial targets sites and inhibition of industrial antimicrobial agents	25
1.4.4 Biofilm resistance and industrial antimicrobial agents.....	25
1.4.5 Cross-resistance between industrial antimicrobial agents and antibiotics.....	26
1.5 Project Aims	27

1.5.1	Research project rationale	27
2.	Materials and methods.....	30
2.1	Chemicals and growth media.....	30
2.2	Bacterial Strains.....	30
2.2.1	Storage of bacterial isolates.....	30
2.2.2	Routine growth conditions of an industrial <i>P. aeruginosa</i> strain	30
2.2.3	Enumeration of viable bacteria.....	30
2.3	Antimicrobial agents.....	31
2.3.1	Preservatives and formulations	31
2.3.2	Antibiotics	31
2.4	DNA Extraction from bacterial cells	31
2.4.1	Rapid DNA extraction using Chelex®100 resin	31
2.4.2	Automated Maxwell® 16 system.....	31
2.4.3	Quantitation and quality assessment of Genomic DNA	32
2.5	Whole genome sequencing and comparative genomics of <i>P. aeruginosa</i> strain RW109.....	33
2.5.1	Single Molecule, Real-Time (Pacific Biosciences) sequencing of the RW109 strain.....	33
2.5.2	Sequence assembly and annotation	33
2.5.3	Identifying the basic features of the RW109 whole genome sequence	34
2.5.4	Identifying plasmids in the assembled genome	34
2.5.4.1	Nucleotide BLAST comparison.....	34
2.5.4.2	ACT: the Artemis Comparison Tool.....	34
2.5.5	Assigning functional groups to the RW109 whole genome sequence.....	34
2.5.5.1	Cluster of Orthologous Group (COG) functional annotation.....	34
2.5.5.2	Determining the interaction of RW109 CDS	35
2.5.5.3	Kyoto Encyclopedia of Genes and Genomes (KEGG) functional module assignment and determining the metabolic and physiological potential	35
2.5.5.4	Identifying putative antimicrobial resistance and virulence genes in the RW109 genome.....	36
2.5.5.5	Identifying antibacterial biocide and metal resistance genes in the RW109 genome	36
2.5.6	Identifying RW109 genomic islands	36
2.5.7	Identifying prophages in RW109 genome.....	36
2.5.8	Visualising the RW109 whole genome sequence.....	36
2.5.9	Comparative genomics of the RW109 whole genome sequence.....	37
2.5.9.1	Comparing RW109 annotated functional categories with PA14 and PAO1	37
2.5.9.2	Roary: Rapid large-scale prokaryote pan genome analysis and core-gene alignment	37

2.5.9.3	Comparison of antimicrobial resistance genes across 103 <i>P. aeruginosa</i> genome sequences.....	37
2.5.9.4	Identifying the unique genes of RW109 when compared to the 102 <i>P. aeruginosa</i> genome sequences.....	38
2.6	Transposon mutagenesis experiments.....	42
2.6.1	Construction of a mini-Tn5- <i>luxCDABE</i> transposon library in RW109 <i>P. aeruginosa</i> strain.....	42
2.6.1.1	Optimising transconjugant selective media: Minimum Inhibition Concentration (MIC) testing of antibiotics using broth and agar assays	42
2.6.1.2	Conjugal transfer of mini-Tn5	42
2.6.2	Identifying the mini-Tn5- <i>luxCDABE</i> transposon in low light emitters.....	43
2.6.3	Random Amplified Polymorphic DNA (RAPD) typing of PA RW109 <i>lux</i> mutants	44
2.6.4	MIC testing of industrial preservatives using a broth doubling dilution method	45
2.6.5	Screening the transposon library to identify preservative responding mutants	45
2.6.5.1	Master plate screening	45
2.6.5.2	Isolated mutant screening	46
2.6.5.3	Statistical analysis of RLU measurements.....	46
2.6.6	Sequencing of transposon flanking DNA and bioinformatics analysis	48
2.7	RNA-Sequencing (RNA-Seq) experiments.....	49
2.7.1	Growth curve analysis of RW109 when exposed to individual preservatives.....	49
2.7.2	Growth curve analysis of RW109 when exposed to BIT and POE in combination	49
2.7.3	Growth curve analysis of RW109 when exposed to a product formulation with and without the presence of a preservative.....	49
2.7.4	RNA-Seq exposure experiments	50
2.7.5	Enumeration of viable bacteria from exposure experiments	52
2.7.6	Total RNA (toRNA) extraction.....	52
2.7.7	Total RNA (toRNA) quantification and quality assessment	52
2.7.8	Messenger RNA (mRNA) enrichment, quantification and quality assessment.....	53
2.7.9	cDNA library preparations and sequencing	53
2.7.10	RNA-Seq bioinformatics analysis	54
2.7.10.1	Quality control and adaptor trimming	54
2.7.10.2	Mapping RNA-Seq reads to the RW109 genome.....	54
2.7.10.3	Counting the number of RNA-Seq reads which map to the RW109 genome.....	54
2.7.10.4	Differential gene expression analysis	55
2.7.11	Generating heat maps of expression data	55
2.7.12	Grouping differentially expressed genes into functional categories.....	55

2.8	Metabolic modelling reconstruction experiments	56
2.8.1	Updating the PA14 model	56
2.8.2	Assigning new reactions to the RW109 metabolic model reconstruction	57
2.8.3	Loading the model into MATLAB	57
2.8.4	Removing PA14 genes with no similarities to RW109 genes	59
2.8.5	Integrating RNA-Seq gene expression data into the RW109 model	59
2.8.6	Gene and reaction essentiality predictions from the condition specific MADE models	60
3.	Whole genome sequencing and comparative genomics of the industrial <i>P. aeruginosa</i> strain RW109	62
3.1	Introduction	62
3.1.1	Aim and objectives	64
3.2	Results	65
3.2.1	Sequence assembly, annotation and basic feature identification	65
3.2.2	Identifying plasmids in the assembled genome	67
3.2.2.1	Nucleotide BLAST	67
3.2.2.2	Comparison of contig synteny	69
3.2.3	Assigning functional groups to the RW109 whole genome sequence	71
3.2.3.1	COG functional annotation	71
3.2.3.1.1	Whole genome COG annotation	71
3.2.3.1.2	RW109 main chromosome COG annotation	71
3.2.3.1.3	RW109 plasmid 1 COG annotation	72
3.2.3.1.4	RW109 plasmid 2 COG annotation	73
3.2.3.2	RW109 KEGG functional module assignment and metabolic and physiological potential	77
3.2.3.2.1	Biologically feasible KEGG modules identified within the main chromosome	77
3.2.3.2.2	Biologically feasible KEGG modules identified within plasmid 1	80
3.2.3.2.3	Biologically feasible KEGG modules identified within plasmid 2	82
3.2.3.3	Identifying the antimicrobial resistance and virulence genes in the RW109 genome	83
3.2.3.4	Identifying biocide and metal resistance genes in the RW109 genome	85
3.2.4	Identifying RW109 genomic islands	86
3.2.4.1	GIs encoded on the main chromosome	86
3.2.4.2	Genomic islands encoded on RW109 plasmid 1	89
3.2.4.3	Genomic islands encoded on plasmid 2	89
3.2.5	Identifying phages in the RW109 genome	91
3.2.6	Comparative genomics of the RW109 whole genome sequence	93
3.2.6.1	Comparing RW109 annotated functional categories with PA14 and PAO1	93
3.2.6.1.1	COG functional annotation comparisons	93

3.2.6.1.2	KEGG functional modules and MAPLE MRCs comparisons	96
3.2.6.3	Antimicrobial and virulence functional annotation comparisons	98
3.2.6.3.3.1	ABRicate analysis of PA14 and PAO1	98
3.2.6.3.3.2	BacMet analysis of PA14 and PAO1	98
3.2.6.2	Core-genome comparison of RW109 with 102 <i>P. aeruginosa</i> genome sequences.....	100
3.2.6.3	Presence and absence of CARD predicted antibiotic resistance genes in RW109 and the 102 <i>P. aeruginosa</i> genome sequences.....	102
3.2.6.4	The unique genes of RW109.....	105
3.2.6.5	Genome size comparison of RW109 with the 102 <i>P. aeruginosa</i> genome sequences.....	107
3.3	Discussion	109
3.3.1	Obtaining a complete sequence of the industrial <i>P. aeruginosa</i> strain RW109	109
3.3.2	Functional annotation of a <i>P. aeruginosa</i> strain isolated from an industrial setting	110
3.3.3	Notable functional annotations of the RW109 strain in comparison to the clinical reference strains PA14 and PAO1	114
3.3.4	Core genome analysis of 103 <i>P. aeruginosa</i> genome sequences isolated from clinical, environmental and industrial sources	115
3.4	Conclusions.....	117
4.	Use of an industrial <i>P. aeruginosa</i> mini-Tn5-<i>luxCDABE</i> mutant library to investigate the genetic basis of preservative resistance	118
4.1	Introduction.....	118
4.1.1	Aim and objectives.....	122
4.2	Results.....	123
4.2.1	Construction of a mini-Tn5- <i>luxCDABE</i> mutant library using an industrial <i>P. aeruginosa</i> strain	123
4.2.1.1	Transconjugant selective media	123
4.2.1.2	The RW109 <i>lux</i> tagged bioluminescent mutant bank	123
4.2.2	Optimising the preservative screening method to identify responding mutants	126
4.2.2.1	Screening with BIT	126
4.2.2.2	Screening with POE.....	126
4.2.3	Screening of the constructed mini-Tn5- <i>luxCDABE</i> mutant library to identify preservative responding mutants.....	128
4.2.4	Identifying transposon insertion sites in isolated mutants of interest	132
4.2.5	The COG category distributions of identified transposon interrupted genes	136
4.2.6	Distribution of the distinct transposon insertion sites throughout the RW109 whole genome.....	138
4.2.7	Key insertion sites of BIT positive responding mutants.....	140
4.2.8	Key insertion sites of BIT negative mutants and mutants with minimal RLU change in response to BIT.....	149

4.3	Discussion	153
4.3.1	The industrial RW109 mini-Tn5- <i>luxCDABE</i> mutant library.....	153
4.3.2	The key genetic determinants associated with BIT positive mutants	155
4.3.3	The key genetic determinants associated with BIT negative mutants	156
4.4	Conclusions.....	158
5.	RNA-Seq gene expression analysis of RW109 in response to preservatives and product formulations.....	159
5.1	Introduction.....	159
5.1.1	Aims and objectives	160
5.2	Results	162
5.2.1	Growth of <i>P. aeruginosa</i> RW109 in the presence of preservatives BIT and POE at various concentrations	162
5.2.2	Initial total RNA extractions and mRNA enrichment	164
5.2.3	Growth of <i>P. aeruginosa</i> RW109 when exposed to BIT and POE in combination	166
5.2.4	Growth of <i>P. aeruginosa</i> RW109 when exposed to unpreserved laundry detergent and laundry detergent with the presence of the preservative BIT.....	168
5.2.5	RNA-Seq exposure conditions	170
5.2.6	toRNA and mRNA quantification, quality assessment and sequencing outputs	172
5.2.7	RNA-Seq data analysis.....	175
5.2.7.1	Overview of gene expression analysis	175
5.2.7.2	Classification of DEGs based on KEGG pathway annotations.....	177
5.2.7.3	Key differentially regulated genes identified following exposure of <i>P. aeruginosa</i> RW109 to single preservatives	179
5.2.7.3.1	Gene expression analysis following exposure to BIT at 20% of the MIC	179
5.2.7.3.2	Gene expression analysis following exposure to BIT at 50% of the MIC	181
5.2.7.3.3	Gene expression analysis following exposure to POE at 20% of the MIC	185
5.2.7.3.4	Gene expression analysis following exposure to POE at 50% of the MIC	189
5.2.7.4	Key differentially regulated genes identified following exposure of RW109 to a combination of preservatives.....	193
5.2.7.4.1	Gene expression analysis following exposure to a combination of BIT and POE both at 20% of the MIC.....	193
5.2.7.5	Key differentially regulated genes identified following exposure of RW109 to a laundry detergent with and without the addition of the preservative BIT	197
5.2.7.5.1	Gene expression analysis following exposure to laundry detergent at 1:100 dilution with no preservative added	197
5.2.7.5.2	Gene expression analysis following exposure to a combination of laundry detergent at 1:100 dilution and BIT at 20% of the MIC	202
5.2.7.6	Gene expression analysis of RW109 antimicrobial predicted genes	212

5.2.7.7	Differentially regulated genes commonly identified during exposure to the industry relevant conditions.....	216
5.3	Discussion	219
5.3.1	RNA-Seq analysis of the RW109 following exposure to industry relevant conditions.....	219
5.3.2	Overview of the transcriptomic analysis when RW109 was exposed to industry relevant conditions	220
5.3.3	Prevalent RW109 genes differentially expressed in response to multiple industry relevant conditions	222
5.4	Conclusions.....	226
6.	Genome-scale metabolic network reconstruction of an industrial <i>P. aeruginosa</i> strain, RW109.....	227
6.1	Introduction.....	227
6.1.1	Aim and objectives.....	229
6.2	Results	230
6.2.1	Updating the PA14 model	230
6.2.2	Classifying the additional reactions of the RW109 metabolic model	230
6.2.3	The RW109 model in COBRA format	234
6.2.4	Removing PA14 genes with no similarities to RW109 genes	234
6.2.5	The overall key features of the RW109 metabolic model	236
6.2.6	Transcriptomic data integration to constrain the RW109 metabolic model and predict reaction essentiality during exposure to industry relevant conditions	238
6.2.6.1	Predicted essential reactions involved in lipid metabolism.....	243
6.2.6.2	Predicted essential reactions involved in nucleotide metabolism	246
6.2.6.3	Predicted essential reactions involved in membrane transport	248
6.2.6.4	Predicted essential reactions involved in xenobiotics biodegradation and metabolism	253
6.2.6.5	Predicted essential reactions involved in carbohydrate metabolism	255
6.2.6.6	Predicted essential reactions involved in amino acid metabolism	258
6.2.6.7	Predicted essential reactions involved in metabolism of cofactors and vitamins.....	263
6.2.6.8	Predicted essential reactions involved in fatty acid metabolism and energy metabolism	265
6.2.7	Comparison of the predicted essential reactions between POE20 and POE50	265
6.2.8	Linking up-regulated genes with those predicted as essential.....	268
6.2.9	Additional modelling work	270
6.2.9.1	Finding an alternative pathway in RW109 for the essential rxn00770 reaction	270
6.2.9.2	Growth of RW109 on various carbon sources	270
6.3	Discussion	271

6.3.1	The outcomes of reconstructing a genome-scale metabolic model specific to an industrial <i>P. aeruginosa</i> strain and the key characteristics of RW109	271
6.3.2	Important points to consider when using the RW109 model reconstruction for computational predictions of biological functionality	274
6.3.3	The successful integration of gene expression data to produce models specific to industry related conditions	275
6.3.4	The key predicted essential reactions and their potential function in the ability of RW109 to survive exposure to industry relevant conditions	277
6.3.5	Preservative concentration greatly influences essentiality predictions.....	284
6.4	Conclusions.....	285
7.	General conclusions, discussion and future research.....	286
7.1	Conclusions.....	286
7.2	Discussion and future research	290
7.2.1	The use of whole genome sequencing to extensively characterise an industrial <i>P. aeruginosa</i> strain.....	290
7.2.2	The efficacy of an industrial <i>P. aeruginosa</i> mini-Tn5- <i>luxCDABE</i> mutant library when investigating the genetic basis of industrial resistance	292
7.2.3	The use of RNA-Seq to characterise the global gene expression response of an industrial <i>P. aeruginosa</i> strain in reaction to industry relevant conditions.....	293
7.2.4	The construction of an industrial <i>P. aeruginosa</i> genome-scale metabolic model for predicting the essentially of reactions during exposure to industry relevant conditions.....	294
References	296

List of Tables

Table 1. RND Efflux pumps of <i>P. aeruginosa</i>	9
Table 2. Important isothiazolinones and alcohol-based preservatives used in the HPC industry.....	21
Table 3. <i>P. aeruginosa</i> genome sequences used in the Roary pan genome analysis	39
Table 4. The basic genome features of the industrial <i>P. aeruginosa</i> strain RW109 in comparison to the reference strains PA14 and PAO1	65
Table 5. <i>P. aeruginosa</i> plasmid sequence information	68
Table 6. The Number of RW109 genes associated with general COG functional categories.....	74
Table 7. Biologically feasible two-component regulatory systems on the RW109 main chromosome	79
Table 8. The biologically feasible KEGG functional modules of RW109 Plasmid 1	81
Table 9. Antimicrobial resistance genes identified in the RW109 genome	84
Table 10. Genomic Islands within the <i>P. aeruginosa</i> RW109 main chromosome	87
Table 11. Genomic Islands encoded on the RW109 plasmids.....	90
Table 12. Phages within the RW109 genome	92
Table 13. KEGG functional modules which differ when MCRs and Q-values of RW109 were compared to PA14 and PAO1.....	97
Table 14. Industrial isolates with the BpeEF-OprC efflux pump KEGG module (M00698).....	106
Table 15. The use of transposon mutagenesis to identify the genetic basis of functional characteristics in <i>P. aeruginosa</i>	119
Table 16. Transposon-interrupted genes found in BIT positive mutants	133
Table 17. Transposon-interrupted genes found in BIT negative mutants.....	134
Table 18. Transposon-interrupted genes found in mutants with minimal RLU change in response to BIT ...	135
Table 19. Total RNA and messenger RNA quantification and quality assessment	173
Table 20. KEGG pathway classifications of the up and down-regulated DEGs for each test exposure condition.....	178
Table 21. The number of genes differentially regulated in the main chromosome and two plasmids after exposure to the test exposure conditions	180
Table 22. Top 10 up-regulated genes in response to BIT at 50% of the MIC	183
Table 23. Top 10 down-regulated genes in response to BIT at 50% of the MIC	184
Table 24. Top 10 up-regulated genes in response to POE at 20% of the MIC.....	187
Table 25. Top 10 down-regulated genes in response to POE at 20% of the MIC	188
Table 26. Top 10 up-regulated genes in response to POE at 50% of the MIC.....	191
Table 27. Top 10 down-regulated genes in response to POE at 50% of the MIC.....	192
Table 28. Top 10 up-regulated genes in response to BIT and POE in combination at 20% of the MIC	195
Table 29. The top ten down-regulated genes in response to BIT and POE in combination at 20% of the MIC	196
Table 30. The top ten up-regulated genes in response to LD with no preservative added.....	200
Table 31. The top ten down-regulated genes in response to LD with no preservative added	201

Table 32. The top ten up-regulated genes in response to laundry detergent in combination with BIT at 20% of the MIC	208
Table 33. The top ten genes down-regulated in response to laundry detergent in combination with BIT at 20% of the MIC	211
Table 34. Differentially expressed antimicrobial resistance genes identified when RW109 was exposed to test conditions 2-7	214
Table 35. The key features of the RW109 and PA14 metabolic models	236
Table 36. Results of transcriptomic gene expression data integration into the RW109 model reconstruction	241
Table 37. Predicted essential reactions for the lipid metabolism KEGG functional category	245
Table 38. Predicted essential reactions for the nucleotide metabolism KEGG functional category	247
Table 39. Predicted essential reactions for the membrane transport category for test conditions 1, 2, 3, 4 and 5.....	250
Table 40. Predicted essential reactions for the membrane transport category for test conditions 1, 2, 6 and 7	252
Table 41. Predicted essential reactions for the xenobiotics biodegradation and metabolism category for test conditions 3 and 5	254
Table 42. Predicted essential reactions for the carbohydrate metabolism category for test conditions 3, 4, 5, 6 and 7.....	257
Table 43. Predicted essential reactions for the amino acid metabolism category for test conditions 2, 3, 4 and 5	260
Table 44. Predicted essential reactions for the amino acid metabolism category for test conditions 2, 6 and 7	262
Table 45. Predicted essential reactions for the metabolism of cofactors and vitamins KEGG functional category	264
Table 46. Predicted essential reactions when RW109 was exposed to POE20 and POE50.....	267
Table 47. Significantly up-regulated essential genes for each preservative exposure condition.....	269

List of Figures

Figure 1. Representation of the <i>P. aeruginosa</i> Gram-negative cell wall	6
Figure 2. Structure of a typical Gram-negative resistance-nodulation division (RND) family efflux pump	8
Figure 3. Resistance mechanisms used to prevent antimicrobials within the cell from reaching target sites	12
Figure 4. Stages in the formation of a biofilm.....	15
Figure 5. Overview of the ubiquitous microorganism <i>Pseudomonas aeruginosa</i>	16
Figure 6. The main experimental approaches used to investigate the genetic basis of preservative resistance in an industrial <i>P. aeruginosa</i> strain	29
Figure 7. The organisation of luminescent plates when screening master plates and individual mutants	47
Figure 8. The arrangement of a Bioscreen microplate during RNA-Seq exposure experiments	51
Figure 9. Read length distributions from PacBio Sequencing of RW109 using SMRT cell technology	66
Figure 10. Comparative synteny line plots of the RW109 contigs 1 and 2 Vs. plasmid sequences and the RW109 whole genome sequence Vs. PA14	70
Figure 11. Circular maps of the RW109 main chromosome and plasmids 1 and 2.....	75
Figure 12. STRING network analyses of RW109 Plasmid 1 CDS within the E COG category	76
Figure 13. Biologically feasible KEGG functional modules of the RW109 main chromosome	78
Figure 14. The COG categories of the RW109 main chromosome Genomic Island's CDS.....	88
Figure 15. Comparison of the number of CDS within the well characterised COG categories for RW109, PA14 and PAO1	94
Figure 16. Percentage number of CDS within each COG category for RW109, PA14 and PAO1	95
Figure 17. Virulence and BacMet predicted genes which differed between RW109, PA14 and PAO1.....	99
Figure 18. Core-gene phylogenetic tree of 103 <i>P. aeruginosa</i> and RW109 genome sequences	101
Figure 19. The presence and absence of CARD antimicrobial genes across 103 <i>P. aeruginosa</i> sequences	104
Figure 20. Genome size distributions of 103 <i>P. aeruginosa</i> sequences of strains isolated from clinical, environmental and industrial sources.....	108
Figure 21. A bioluminescent mini-Tn5- <i>luxCDABE</i> -Tc transposon.....	121
Figure 22. RAPD PCR gel for selected RW109 transposon mutants	125
Figure 23. Effect of BIT concentration on the RLU response of three mutants	127
Figure 24. Box plots of mutants with a significant RLU fold increase when exposed to BIT for 24hours	129
Figure 25. Box plots of mutants with a significant RLU fold decrease when exposed to BIT for 24hours.....	130
Figure 26. Box plots of mutants with no significant RLU fold change when exposed to BIT for 24hours.....	131
Figure 27. The distribution of identified transposon interrupted genes based on COG categories	137
Figure 28. Distribution of the distinct <i>lux</i> transposon insertion sites throughout the main chromosome of RW109.....	139
Figure 29. The <i>luxCDABE</i> transposon mutation of genes encoding a bacterial extracellular solute-binding transport system operon and their RLU responses when exposed to BIT for 24 hours.....	141
Figure 30. The overproduction of a brown/red pigment by the BIT positive mutant 14:C4	143
Figure 31. The <i>luxCDABE</i> transposon mutation of the gene encoding a KdgR transcriptional regulator	144
Figure 32. The <i>luxCDABE</i> transposon mutation of the gene within the <i>napEFDABC</i> operon.....	145
Figure 33. The <i>luxCDABE</i> transposon mutation of a gene encoding a ferrous-iron efflux pump.....	147

Figure 34. The <i>luxCDABE</i> transposon mutation of gene encoding a bifunctional enzyme CysN	148
Figure 35. The <i>luxCDABE</i> transposon mutation of a gene encoding for an acid-activated urea channel	150
Figure 36. The <i>luxCDABE</i> transposon mutation of a gene encoding a tripartite ATP-independent periplasmic (TRAP) transporters DctQ component	151
Figure 37. The <i>luxCDABE</i> transposon mutation of a gene encoding for the multidrug efflux pump subunit AcrB	152
Figure 38. Overview of stages involved RNA-Seq transcriptomic analysis	161
Figure 39. The effect of the individual preservatives BIT and POE at different concentrations on the growth of RW109	163
Figure 40. Bioanalyzer electropherograms of extracted total RNA and enriched mRNA	165
Figure 41. The effect of BIT and POE in combination on the growth of RW109	167
Figure 42. The effect of unpreserved laundry detergent and laundry detergent with the preservative BIT on the growth of RW109	169
Figure 43. Growth dynamics of RW109 when exposed to the chosen RNA-Seq control and test conditions ...	171
Figure 44. Bioanalyzer electropherograms of total RNA extractions and enriched messenger RNA samples	174
Figure 45. Overview of RNA-Seq gene expression data	176
Figure 46. The differentially regulated genes identified on the main chromosome in response to laundry detergent in combination with BIT	205
Figure 47. The differentially regulated genes identified on plasmid 1 in response to laundry detergent and BIT in combination	206
Figure 48. The differentially regulated genes identified on plasmid 2 in response to laundry detergent and BIT in combination	207
Figure 49. Gene expression analysis the RW109_03936-03940 predicted membrane transport operon and a transcriptional regulator GI <i>t</i> R when exposed to test conditions 2-7	209
Figure 50. Gene expression analysis of the <i>pilGHI</i> and <i>pilJK-chpA-E</i> operons when exposed to test conditions 5, 6 and 7	210
Figure 51. Differentially regulated RW109 antimicrobial genes when exposed to test conditions 2-7	213
Figure 52. The MexPQ-OpmE RND efflux system identified as up-regulated when exposed to POE- and LD-associated test conditions (3-7)	215
Figure 53. The up-regulated response of the <i>gnyRDBHAL</i> gene cluster and KdgR transcriptional regulator during exposure to multiple test exposure conditions	217
Figure 54. The down-regulated response of the <i>phnAB</i> and <i>pqsABCDE</i> operon during exposure to multiple test exposure conditions	218
Figure 55. KEGG functional categories of the additional RW109 model reactions	232
Figure 56. Breakdown of the key reactions in the two major functional categories associated with the RW109 additional model reactions	233
Figure 57. Flow diagram of reactions removed from the RW109 metabolic model	235
Figure 58. KEGG functional categories of the RW109 metabolic model reconstruction	237
Figure 59. Computational pipeline of transcriptomic data integration into the RW109 metabolic model	240

Figure 60. KEGG functional categories of the predicted essential reactions for each RNA-Seq exposure condition.....	242
Figure 61. Grouping of the lipid metabolism predicted essential reactions	244
Figure 62. Grouping of the membrane transport predicted essential reactions for the test conditions 1, 2, 3, 4 and 5	249
Figure 63. Grouping of the membrane transport predicted essential reactions for test conditions 1, 2, 6 and 7.....	251
Figure 64. Grouping of the carbohydrate metabolism predicted essential reactions	256
Figure 65. Grouping of the amino acid metabolism predicted essential reactions for test conditions 2, 3, 4 and 5.....	259
Figure 66. Grouping of the amino acid metabolism predicted essential reactions for test conditions 2, 6 and 7	261
Figure 67. Comparison of the predicted essential reactions between POE20 and POE50.....	266
Figure 68. Key enzymes and pathways linked to reactions projected to be essential during exposure of RW109 to industry relevant conditions.....	283
Figure 69. Key features of the industrial <i>P. aeruginosa</i> strain RW109.....	289

Abbreviations

ABC	ATP binding cassette
ACT	Artemis Comparison Tool
ATP	Adenosine triphosphate
BIT	Benzisothiazolinone
BLAST	Basic Local Alignment Search Tool algorithm
BLASTn	Nucleotide BLAST
BLASTp	Protein-Protein BLAST
bp	Base pairs
BWA	Burrows-Wheeler Aligner
BWA-MEM	BWA-maximal exact matches
cDNA	Complementary Deoxyribonucleic acid
CDP-DAG	CDP-diacylglycerol
CDS	Coding sequences
CF	Cystic Fibrosis
CFU / ml	Colony forming units per millilitre
CFU/g	Colony forming units per gram
CGView	Circular Genome Viewer
CLIMB	Cloud Infrastructure for Microbial Bioinformatics
CMIT	Chloro-N-methylisothiazolinone
CITMIT	Chloro-N-methylisothiazolinone and N-methylisothiazolinone
COBRA	COntstraint-Based Reconstruction and Analysis
COG	Cluster of Orthologous Group
DEGs	Differentially expressed genes
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic acid
dNTPs	deoxyribonucleotide triphosphates
EggNOG	Evolutionary Genealogy of Genes: Non-supervised Orthologous Groups
EPI	Efflux Pump Inhibitor
EPS	Exopolysaccharide
FBA	FLUX Balance Analysis
FVA	FLUX Variability Analysis
GIs	Genomic Islands
GMPs	Good Manufacturing Practices

GPR	Gene-Protein-Reaction
grRateKO	deletion strain growth rates
grRateWT	wild type growth rate
GTR	Generalized Time-Reversible
HPC	Home and Personal Care
KASS	KEGG Automatic Annotation Server
Kbp	Kilobase pairs
KEGG	Kyoto Encyclopedia of Genes and Genomes
KO	KEGG Orthology
LB	Luria bertain
LD	Laundry Detergent
LPS	Lipopolysaccharide
MADE	Metabolic Adjustment by Differential Expression
MAPLE	Metabolic and Physiological Potential Evaluator
Mbp	Mega base pairs
MIC	Minimum inhibitory concentration
MIT	N-methylisothiazolone
MLST	Multi locus sequence typing
mRNA	Messenger RNA
NCBI	National Centre of Biotechnology and Information
NO	Nitric oxide
OD	Optical density
PA	<i>Pseudomonas aeruginosa</i>
PacBio	Pacific Biosciences
PATRIC	PathoSystems Resource Integration Center
PAβN	Phenylalanine-Arginine Beta-Naphthylamide
PBS	Phosphate buffer solution
PCR	Polymerase chain reaction
PHASTER	PHAge Search Tool Enhanced Release
Pmx	Polymixin B
POE	Phenoxyethanol
PQS	Pseudomonas quinolone signal
qRT-PCR	Quantitative real-time PCR
QS	Quorum sensing
QUAST	Quality Assessment Tool for Genome Assemblies
RAPD	Random amplified polymorphic DNA

RAST	Rapid Annotation Subsystems Technology
RIN	RNA Integrity Number
RLU	Relative Light Units
RNA	Ribonucleic acid
RNA-Seq	RNA Sequencing
RND	Resistance Nodulating Division
rpm	Revolutions per minute
rRNA	Ribosomal RNA
RT-PCR	Reverse transcriptase PCR
RT-PCR	Reverse transcriptase PCR
SMRT	Single Molecule, Real-Time
Taq	<i>Thermus aquaticus</i>
Tc	Tetracycline
TCS	Two-component regulatory systems
TIGER	Toolbox for Integrating Genome-scale Metabolism, Expression, and Regulation
tmRNA	Transfer Messenger RNA
toRNA	Total RNA
TraDIS	Transposon Directed Insertion Sequencing
tRNA	Transfer RNA
TSA	Tryptone soya agar
TSB	Tryptone soya broth
UV	Ultra violet
WGS	Whole genome sequencing
WHO	World Health Organisation
WT	Wild-Type

1. Introduction

1.1 *Pseudomonas aeruginosa*

Pseudomonas aeruginosa (*P. aeruginosa*) is a Gram-negative, non-fermentative, aerobic rod-shaped bacterium, with a cell diameter of 0.5-1.0 µm and length of 1.5-3.0 µm. Optimum growth occurs at 37 °C, although the bacterium has the ability to grow in a wide range of temperatures, reported between 4 °C and 42 °C (Morrison and Wenzel, 1984, LaBauve and Wargo, 2012). *P. aeruginosa* demonstrates proficiency at surviving with minimal nutritional levels and exhibits multiple metabolic capabilities, allowing the microorganism to grow and proliferate within a variety of environments (Stover et al., 2000, Rossolini and Mantengoli, 2005, LaBauve and Wargo, 2012).

The pathogenic potential of *P. aeruginosa* renders it an important clinical microorganism, commonly causing disease in the immunocompromised, especially in transplant recipients, individuals with abnormally low neutrophil levels (neutropenia), those receiving chemotherapy and patients with HIV (Sadikot et al., 2005, Bodey, 2001, Vidal et al., 1999). This pathogen is the causative agent of a broad spectrum of infections associated with the respiratory system, urinary tract, inner ear, blood stream and wounds (Kerr and Snelling, 2009, Rossolini and Mantengoli, 2005, Sadikot et al., 2005, Wagenlehner and Naber, 2006). *P. aeruginosa* rarely causes disease in healthy individuals (Lyczak et al., 2000, Sadikot et al., 2005), which characterises this bacterium as an opportunistic pathogen. In cystic fibrosis (CF) patients, *P. aeruginosa* is a dominant pathogen and results in chronic, severe pulmonary infections which are largely responsible for the morbidity and mortality observed with CF (Govan and Deretic, 1996, Oliver et al., 2000). *P. aeruginosa* infections in the CF lung are extremely difficult to eradicate (Gómez and Prince, 2007), especially with the increased occurrence of multidrug resistance reported with this pathogen, which is of great concern with in the research field of *P. aeruginosa* (Langan et al., 2015).

1.1.1 The environmental versatility of *P. aeruginosa*

The ubiquitous nature of *P. aeruginosa* results in its ability to colonise a diverse range of habitats. This microorganism is commonly isolated from a variety of settings including marine habitats (Khan et al., 2007), freshwater systems (Tripathy et al., 2007), soil (Marques et al., 1979), plants and vegetables (Green et al., 1974, Kominos et al., 1972) and hospital locations (Whitby and Rampling, 1972). The ability to colonise different environments results in the bacterium being able to infect and cause disease in numerous hosts other than humans, such as animals (Haenni et al., 2015), insects (Miyata et al., 2003, Apidianakis and Rahme, 2009) and plants (Elrod and Braun, 1942). *P. aeruginosa* also has the ability to grow in extreme environments which have high and low pH levels, excesses of hot and cold, elevated amounts of toxic compounds, increased and decreased salinity and high-pressure levels (Pikuta et al., 2007, Montgomery et al., 2013). Examples of harsh settings where *P. aeruginosa* has been reported growing include crude oil and petroleum hydrocarbons (Guo-liang et al., 2005, Song et al., 2006), the Antarctic (Villeret et al., 1997), hazardous chemicals used in conventional leather-making (Lama et al., 2012), disinfectants (Dantas et al., 2008, Kung et al., 2010) and pharmaceutical and cosmetic products (Jimenez, 2007). The degradative abilities of *Pseudomonas* spp. also results in their common incorporation into bioremediation of environmental pollutants such as heavy metals pesticides and phenols (Wasi et al., 2013), with *P. aeruginosa* being reported as functional in the removal of cadmium (Wang et al., 1997).

P. aeruginosa has the ability to exist in diverse environments as either planktonic unicellular cells which are free floating or they can form multicellular aggregates of communicating sessile cells, which attach to surfaces and are recognised as biofilms (Spoering and Lewis, 2001, Moradali et al., 2017). The ability to survive as planktonic cells allows the organism to move around its environment, initiate colonisation and cause acute infection (Chua et al., 2014). Existing as a multicellular community enables prolonged persistence and established infections to occur, with 65-80% of all human chronic microbial infections being associated with biofilms (Spoering and Lewis, 2001, Donlan, 2002).

Successful adaptation to a wide-range of environmental niches, each presenting unique challenges, demands exceptional functional versatility from a microorganism. *P. aeruginosa* exhibits a broad metabolic capacity demonstrated by the ability to grow on 75 organic compounds (Stanier et al., 1966) and can survive in low-nutrient conditions, exhibited by the aptitude for growing in distilled and bottled water (Favero et al., 1971, Eckmanns et al., 2008). The metabolic potential of *P. aeruginosa* is also emphasised by the capacity to proliferate in very low oxygen levels (Alvarez-Ortega and Harwood, 2007) where it can utilise nitrogen as a terminal electron acceptor to respire anaerobically (Sharma et al., 2006). The multiple secondary metabolites and virulence factors produced by this bacterium are also indicative of its extensive metabolic proficiencies (Latifi et al., 1995, Gallagher and Manoil, 2001, Balasubramanian et al., 2012).

The metabolic versatility of *P. aeruginosa* has also been demonstrated via the modelling of its biochemical networks (Oberhardt et al., 2008, Oberhardt et al., 2011, Bartell et al., 2017), which represent the gene-protein-reaction relationships and metabolic potential. In the clinical reference strain PA14, a

number of distinct biological functional groups (14) represented the reactions, genes and metabolites within a genome-scale metabolic model; the majority of biochemical pathways were associated with central metabolism such as lipid, amino acid and carbohydrate metabolism as well as nutrient transport (Bartell et al., 2017). Although metabolic models represent a small proportion of a microorganisms metabolic potential, in combination with whole genome sequencing, they have demonstrated the assortment of biochemical pathways available to *P. aeruginosa*.

1.1.2 Linking genome size with the complex biological functionalities of *P. aeruginosa*

The genome size of *P. aeruginosa* ranges between ~5.2–7.0 Mbp, which is among the largest recognised in Gram-negative bacteria and encodes around 5,570 open reading frames (Schmidt et al., 1996, Stover et al., 2000, Winsor et al., 2016). This extensive genomic size is indicative of the genetic complexity of *P. aeruginosa*, providing the microorganism with an extensive availability of multipurpose genes. A high number of genes (8.4%) are associated with regulatory networks and two component systems, as well as those linked to the utilisation of various carbon sources and nutrient transporters (Stover et al., 2000, Arai, 2011, Nishijyo et al., 2001). This observation suggests a reason for the ability of *P. aeruginosa* to metabolise and utilise various compounds for nutritional purposes, and would permit environmental adaptability. It has previously been characterised that large bacterial genomes are excessively enriched in genes associated with regulation, secondary metabolism and to a lesser extent energy conversion-related genes, whereas those linked to protein translation, DNA replication, cell division, and nucleotide metabolism have less representation (Konstantinidis and Tiedje, 2004, van Nimwegen, 2006). Bacteria with larger genome sizes are better adapted at persisting in environments with insufficient but varied resources, due to their portrayed wide-ranging metabolic diversity and regulatory functions (Konstantinidis and Tiedje, 2004).

The *P. aeruginosa* genome is comprises a single, supercoiled circular chromosome which resides in the cytoplasm (Klockgether et al., 2011, Vallet-Gely and Bocard, 2013). The presence of additional plasmids which incorporate antimicrobial resistance genes have also been reported (Shahid and Malik, 2003). Investigations of the large genome have revealed 4000 “core genes” which are generally well conserved and invariable between *P. aeruginosa* isolates (Klockgether et al., 2011, Freschi et al., 2015) and another 20% of genes representing the “accessory genome”, commonly identified on bacterial genomic islands and indicative of horizontal gene transfer (Rumbaugh, 2014, Juhas et al., 2009). The “accessory genome” is strain specific and contains genes involved in metabolism, virulence and antibiotic resistance (Rumbaugh, 2014, Kung et al., 2010). The variability observed with the “accessory genome” accounts for the strain diversity of *P. aeruginosa* and contributes to the wide ranging environmental niches where the bacterium can be isolated (Kung et al., 2010). Defining and characterising the genome of *P. aeruginosa* is central to understanding the diverse functionality of this microorganism.

1.2 *P. aeruginosa* as a multidrug resistant microorganism

The International Nosocomial Infection Control Consortium reported that *P. aeruginosa* hospital acquired infections have become a worldwide healthcare issue (Rosenthal et al., 2016). The World Health Organisation (WHO) has also recognised *P. aeruginosa* within the most critical group of “antibiotic resistant priority pathogens” which pose the greatest threat to human health (WHO, 2017). This microorganism has substantial antimicrobial resistance capabilities and is notoriously difficult to control with antibiotics and disinfectants (Hancock, 1998, Lambert, 2002). Antibiotic resistance is classified as the capability of a bacterium to survive in antibiotic concentrations that normally inhibits growth of most other bacteria (Russell, 2000). *P. aeruginosa* displays low susceptibility to antimicrobials due to its intrinsic genetically encoded resistance mechanisms (Livermore, 2002). It also has the ability to acquire resistance via temporary or permanent adaptive mutations and through exogenous genetic materials such as plasmids (Flores et al., 1997, Hancock and Speert, 2000).

The increased occurrence of reported resistance in *P. aeruginosa* is a result of de novo mechanisms, which develop during prolonged and frequent exposure to antimicrobials in hospital settings, along with the spread of resistant organisms between patients (Carmeli et al., 1999, Fridkin and Gaynes, 1999, Aloush et al., 2006). Medical settings however are not the only environment where antimicrobial resistance can develop. Locations which are subjected to contamination with human pollution include sewage works (Amos et al., 2014), industrial pharmaceutical settings (Ahmad et al., 2017), aquaculture facilities (Buschmann et al., 2012) and agricultural environments (Witte, 1998) and usually have elevated bacterial loads (Berendonk et al., 2015). These areas are also frequently exposed to sub-inhibitory concentrations of antimicrobials, resulting in bacteria further developing and acquiring resistant mechanisms, which can then spread throughout the environment (Angulo et al., 2004, Di Cesare et al., 2013, Ahmad et al., 2017). As discussed previously, *P. aeruginosa* is capable of living within a wide range of diverse settings and as a result undoubtedly encounters antimicrobials on a recurrent basis.

1.2.1 Antimicrobial resistance mechanisms of *P. aeruginosa*

The wide-ranging antimicrobial aptitudes of *P. aeruginosa* result from a combination of elements such as the low permeability of its cell wall, the genetic capacity to encode numerous resistance mechanisms such as efflux pumps and degradative enzymes, and ability to form multicellular aggregations known as biofilms (Aloush et al., 2006, Hancock, 1998, Stewart and William Costerton, 2001). These resistant mechanisms may already be innate to the bacterium and can also become accessible via mutations of chromosome resistance regulation genes (Oliver et al., 2004, Lister et al., 2009) and through obtaining plasmids and transposons which encode additional resistance genes (Livermore, 2002, Lambert, 2002, Shahid and Malik, 2003). Three widely acknowledged modes of resistance actions are used by *P. aeruginosa* to reduce its susceptibility to antimicrobials, which include obstructing antimicrobial penetration, decreasing antimicrobial accumulation within the cell and preventing antimicrobials from reaching target sites. (Aloush et al., 2006, Lambert, 2002). Resistance mechanisms are frequently used simultaneously, especially when pathogens display multi-antimicrobial resistance characteristics (Higgins,

2007, Livermore, 2002). Another feature of *P. aeruginosa* is the ability to form surface-attached microbial biofilm communities which when compared with planktonic equivalents can be 10–1,000 times more resist to antimicrobials (Hoyle and Costerton, 1991, Hancock, 1998, Mah et al., 2003).

1.2.1.1 Obstructed entry into the cell

An important factor behind the intrinsic resistance of *P. aeruginosa* to antimicrobial agents is associated with the low permeability of its cell envelope (Chapman, 1998, Lambert, 2002) which restricts antimicrobial uptake. The highly structured multi-layered bilayer arrangement of the Gram-negative cell wall includes an outer membrane of phospholipids and lipopolysaccharides (LPS), a thin peptidoglycan layer and a phospholipid inner cytoplasmic membrane (Figure 1) (Chapman, 1998, Brown et al., 2015, Zgurskaya et al., 2015). *P. aeruginosa* has a high cell wall lipid content (Schweizer, 2001) and an outer membrane permeability identified as 10–100 fold lower when compared to other Gram-negative bacteria, such as *Escherichia coli* (*E. coli*) (Hancock and Speert, 2000, Schweizer, 2001). The narrow nonspecific porins and specific uptake channels of the outer membrane restrict the penetration of antimicrobials by their size (Hancock, 1998, Zgurskaya et al., 2015). LPS consists of three elements, a highly conserved lipid A glycolipid hydrophobic outer membrane attachment protein, central core polysaccharide, with two well-conserved 3-deoxy-d-mannooctulosonic acid residues, a more changeable oligosaccharide region and the architecturally varied O-antigen (Wang and Quinn, 2010). The *P. aeruginosa* LPS phospholipid bilayers are more inflexible compared to standard bilayers, which reduces the passive diffusion of hydrophobic compounds (Zgurskaya et al., 2015). Strong links between the LPS molecules are also facilitated by a high Mg⁺ outer membrane content (McDonnell and Russell, 1999).

A large portion of the *P. aeruginosa* genome encodes for surface molecules and two extremely long open reading frames, identified in PAO1 (PA2462 and PA41) are predicted to be involved in the synthesis of LPS and extracellular polysaccharides (Stover et al., 2000). The proteins associated with the outer membranes interact with the surrounding environment and transmit signals to the cytoplasm, which enable the bacterium to adapt appropriately to changes in the surrounding conditions (Koebnik et al., 2000). A large amount of proteins are exported out of the cytoplasm for use in the cell wall and include those associated with nutrient uptake, pili and flagella biogenesis, immunogenicity, antimicrobial resistance, phospholipid assembly and environmental sensing (Molloy et al., 2000, Lewenza et al., 2005). The Gram-negative *P. aeruginosa* cell wall provides the bacterium with an intrinsic first line of defence when faced with antimicrobials and has an imperative role in obstructing entry into the cell.

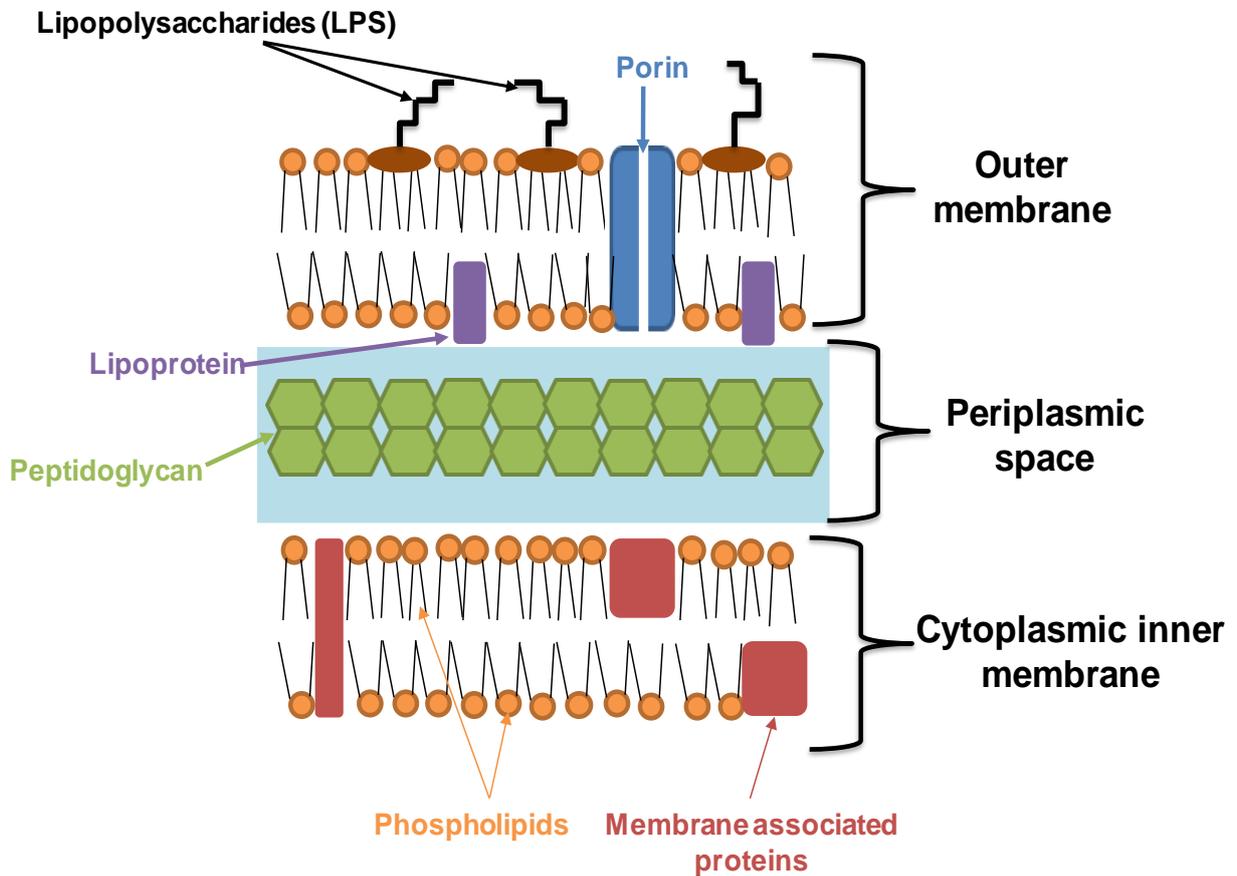


Figure 1. Representation of the *P. aeruginosa* Gram-negative cell wall

The *P. aeruginosa* Gram-negative cell wall is made up of a cytoplasmic inner membrane, a thin peptidoglycan layer within the hydrophilic periplasmic space, periplasm and an outer membrane. The hydrophobic outer membrane consists of phospholipid molecules as well as lipopolysaccharide (LPS) molecules, lipoproteins and also porins which permit the entry and exit of hydrophilic low-molecular weight substances and may be specific or non-specific. The hydrophobic cytoplasmic membrane is a phospholipid bilayer and contains associated membrane proteins. Adapted from (Brown et al., 2015).

1.2.1.2 Reducing intracellular accumulation of antimicrobials

1.2.1.2.1 Efflux pumps of *P. aeruginosa*

The cell wall of *P. aeruginosa* is also important due to the numerous active efflux pump systems it encompasses (McArthur et al., 2013, Askoura et al., 2011); these systems work to prevent the accumulation of toxic agents within the cell by actively removing them (Lambert, 2002, Poole, 2005b). The synergy between the low-permeability of the cell wall and active efflux systems are considered the major mechanisms of antimicrobial resistance, which have allowed *P. aeruginosa* to become a prevalent pathogen in hospital acquired infections (Poole, 2001, Poole, 2005b, Schweizer, 2003). Efflux systems are composed of three protein structures; an energy-reliant cytoplasmic membrane transporter, a porin outer membrane protein and linker proteins which couple the structures together within the periplasm (Figure 2) (Poole, 2001, Paulsen, 2003, Piddock, 2006, Lister et al., 2009).

Efflux pumps actively expel chemicals from within the bacterial cell and a variety of these systems are used by *P. aeruginosa* (Table 1), with the MexAB-OprM system acting as the primary efflux pump which can export a wide range of antimicrobials out of the cell (Askoura et al., 2011, Poole, 2001). The major efflux pumps of *P. aeruginosa* belong to the Resistance-Nodulation-Division (RND) (Poole, 2011, Soto, 2013, Avrain et al., 2013) (Table 1). The structure of these efflux pump systems (Figure 2) allows the direct expelling of toxic substances from the inner membrane directly to the extracellular environment (Avrain et al., 2013). The MexAB-OprM and MeXY-OprM systems are continuously expressed at a base line level in *P. aeruginosa*, with the expression of MexAB-OprM being the most notable (Lister et al., 2009). The genes encoding these two pumps can also be up-regulated in response to antimicrobials (Riou et al., 2010), along with the other efflux systems shown in Table 1; the genes associated with the other pumps are not constantly 'on' in *P. aeruginosa* (Avrain et al., 2013). The RND efflux pumps are associated with multi-drug resistant strains of *P. aeruginosa* as they are capable of transporting numerous types of antimicrobials (Schweizer, 2003, Piddock, 2006).

The importance of efflux pumps in *P. aeruginosa* biofilms have also been identified, with genes encoding these systems demonstrating increased expression in biofilms when compared to planktonic bacteria (Zhang and Mah, 2008). Disruption of genes associated with efflux pump systems, renders the *P. aeruginosa* biofilms more susceptible to antibiotics (Wright, 2005), highlighting efflux as an important mechanism behind biofilm associated antimicrobial resistance.

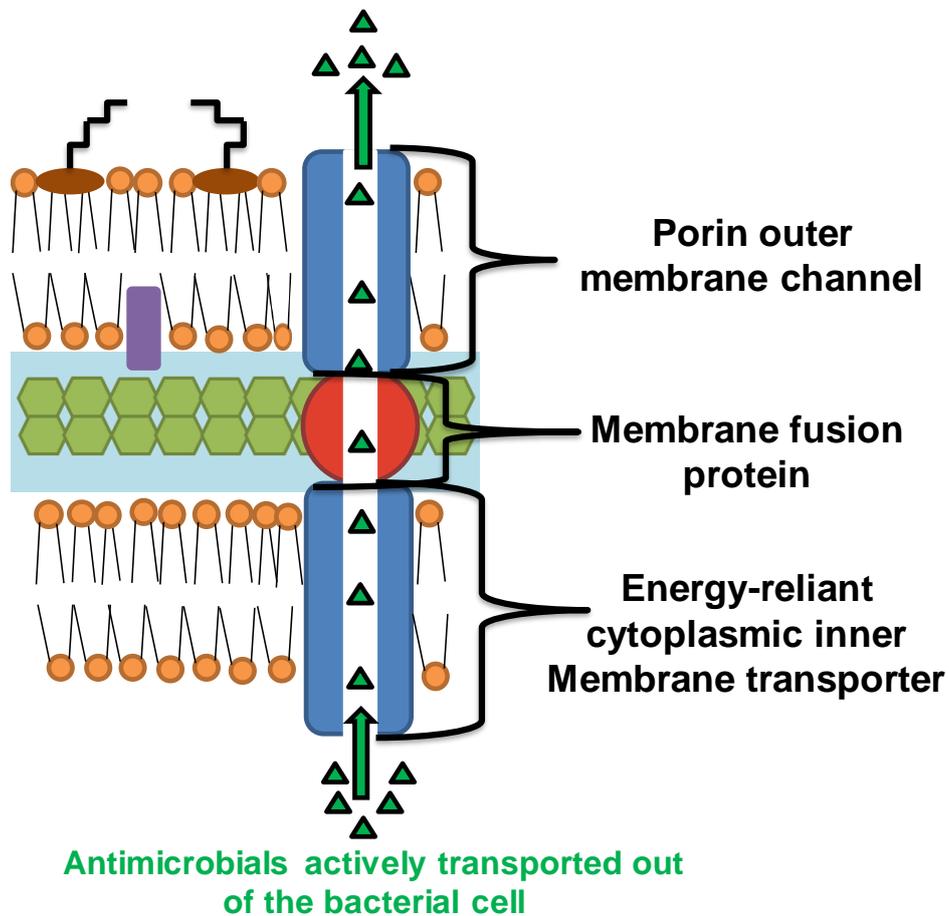


Figure 2. Structure of a typical Gram-negative resistance-nodulation division (RND) family efflux pump

The *P. aeruginosa* resistance-nodulation division (RND) family efflux pumps comprise an energy-reliant transporter protein located in the cytoplasmic membrane, a periplasmic membrane fusion-linker protein and an outer membrane porin protein channel. Efflux pumps can be specific for one substrate or may transport a range of structurally dissimilar compounds. The resistance from efflux pumps can be intrinsic to the bacteria and may be acquired via mutations resulting in the up-regulation of efflux systems. Adapted from (Pidcock, 2006, Lister et al., 2009, Askoura et al., 2011).

Table 1. RND Efflux pumps of *P. aeruginosa*

Efflux pump system	Types of substrates exported	References
MexAB-OprM	β -lactam, SDS, fluoroquinolones, tetracycline, novobiocin, chloramphenicol, macrolides, trimethoprim, triclosan, ethidium bromide, aromatic hydrocarbons, thiolactomycin, cerulenin, acylated homoserine lactones, colistin.	(Nehme and Poole, 2007, Sugimura et al., 2008, Pamp et al., 2008, Masuda et al., 2000)
MexCD-OprJ	β -lactam, fluoroquinolones, tetracycline, chloramphenicol, novobiocin, macrolides, trimethoprim, triclosan, ethidium bromide, SDS, aromatic hydrocarbons, crystal violet, acriflavin, chlorheximide, benzalkonium chloride	(Srikumar et al., 1998, Schweizer, 2003, Masuda et al., 2000, Fraud et al., 2008, Morita et al., 2003)
MexEF-OprN	Fluoroquinolones, chloramphenicol, trimethoprim, triclosan, aromatic hydrocarbons, ciprofloxacin	(Köhler et al., 1999, Llanes et al., 2011, Schweizer, 2003)
MexGHI-OprD	Vanadium, acylated homoserine lactones	(Aendekerk et al., 2002, Aendekerk et al., 2005)
MexPQ-OpmE	Fluoroquinolones, tetracycline, chloramphenicol, erythromycin, macrolides, metal ions	(McArthur et al., 2013, Mima et al., 2005, Lister et al., 2009, Thaden et al., 2010)
MexXY (OprA and OprM)	Tetracycline, erythromycin, aminoglycosides, fluoroquinolones	(Morita et al., 2012, Hocquet et al., 2003, Mine et al., 1999, Masuda et al., 2000)
MexJK (OprM and OpmH)	Tetracycline, erythromycin, triclosan	(Chuanchuen et al., 2002, Chuanchuen et al., 2005)
MexVW-OprM	Fluoroquinolones, tetracycline, chloramphenicol, erythromycin, acriflavin	(Li et al., 2003)
MexMN-OprM	Chloramphenicol, thiamphenicol	(Mima et al., 2009)
CzrAB-OpmN	Cadmium, Zinc	(Schweizer, 2003)
TriABC-OpmH	Triclosan	(Mima et al., 2007)

1.2.1.2.2 Modification of the Gram-negative cell wall

Certain antimicrobial agents are able to cross the Gram-negative cell wall barrier and reach target sites. Hydrophobic antimicrobials utilise a lipid-mediated pathway and hydrophilic agents diffuse through porins within the cell wall (Delcour, 2009). *P. aeruginosa* is able to become resistant to these types of antimicrobials by altering the pathways used to cross the cell envelope barrier in order to reduce intracellular accumulation (Tumah, 2009). The organism can adapt by structurally altering the hydrophobic properties of the membrane, replacement of porins with more-selective channels, completely removing a particular porin via null mutations and through the modification of individual lipopolysaccharide binding sites (Delcour, 2009, Blair et al., 2015). Decreased permeability due to alterations in the *P. aeruginosa* LPS outer membrane, is well documented in aminoglycoside-resistant clinical isolates, especially in strains from CF patients (MacLeod et al., 2000, Poole, 2005a). A decreased availability of the OprD *P. aeruginosa* porin protein, has been reported to reduce the entry of the carbapenem antibiotic imipenem into the cell (Fukuoka et al., 1993). The modification of the cell wall composition can also occur within the phospholipid bilayer and increased content of important cell wall components such as phosphatidylglycerol (Arendt et al., 2012) and cardiolipins (Bernal et al., 2007), have been shown to reduce antibiotic permeability across the cell membranes. This highlights the importance of modifying the cell wall as a resistance mechanism of *P. aeruginosa*.

1.2.1.3 Preventing antimicrobials within the cell reaching target sites

If an antimicrobial is successful at crossing the Gram-negative cell wall barrier and is not expelled by efflux pumps, other resistance mechanisms can be employed which prevent the agent from reaching its target within the cell. These resistance traits are usually acquired either through genes encoded on mobile genetic elements such as plasmids and via mutation or the up-regulation of chromosomally encoded resistance genes (Livermore, 2002, Strateva and Yordanov, 2009)

1.2.1.3.1 Inactivating antimicrobials with enzymes

The inactivation of antimicrobials by enzymes is a mechanism used by Gram-negative bacteria to reduce susceptibility to these agents and prevent them from reaching target sites within the cell (Lambert, 2002). *P. aeruginosa* can adapt to antimicrobial exposure via the production of enzymes which degrade or modify the agent (Figure 3A). Multiple bacterial antibiotic degrading enzymes have been identified in *P. aeruginosa* (Wright, 2005). It has been well established that genes encoding beta-lactamases are up-regulated in response to the presence of beta-lactam antibiotics, and in particular an overexpression of a chromosomally encoded regulatory *ampR* gene is observed (Lambert, 2002). Beta-lactamase enzymes break down antibiotics through hydrolysing the β -lactam ring which is present in penicillin, cephalosporin, monobactam, and carbapenem antibiotics (Santajit and Indrawattana, 2016). Modification of aminoglycoside antibiotics has been characterised in *P. aeruginosa* and as commonly carried out by aminoglycoside-acetyltransferases, aminoglycoside-adenyltransferases and aminoglycoside-

phosphotransferases, with genes encoding these enzymes frequently found on mobile genetic elements (Ramirez and Tolmasky, 2010, Poole, 2005a).

1.2.1.3.2 Modification of antimicrobial target sites

The alteration, protection and overproduction of bacterial target sites have also been identified as mechanisms used by *P. aeruginosa* to reduce susceptibility to those antimicrobials with the ability to cross the Gram-negative cell wall (Figure 3B) (Lambert, 2005, Kohanski et al., 2010). Alteration of target sites usually occur due to spontaneous bacterial chromosome gene mutations and antibiotic selection pressures (Lambert, 2005). The alteration of the fluoroquinolone DNA gyrase target in *P. aeruginosa* has also been reported as a common resistance mechanism in response to this antibiotic (Mouneimné et al., 1999, Mesaros et al., 2007). In *P. aeruginosa* the binding of the aminoglycoside antibiotics to bacterial ribosomes can be compromised by the ability of the bacterium to modify the ribosomal RNA (rRNA) binding sites of the antibiotic through methylation of 16S rRNA (Gutierrez et al., 2013).

Protection of target sites can be implemented to prevent antimicrobials from binding and causing disruption in the cell (Figure 3B). This technique can be achieved via the production of ribosomal protection proteins which confer resistance to antibiotics such as tetracycline, macrolides, chloramphenicol erythromycin and phenicols (Thaker et al., 2010, Zhou et al., 2015a). Another mechanism involved in the modification of target sites, is an over production of a protein which can also be bound by the antimicrobial (Figure 3B). The up-regulation of bacterial genes which synthesise antimicrobial-insensitive sites of action, prevents the agent from efficiently binding to its target, rendering it inactive within the cell (Dever and Dermody, 1991). The alteration of metabolic pathways to induce overproduction of a specific metabolite which competes for the active target site of an antimicrobial, is also a reported mechanism of resistance (Zhou et al., 2015a).

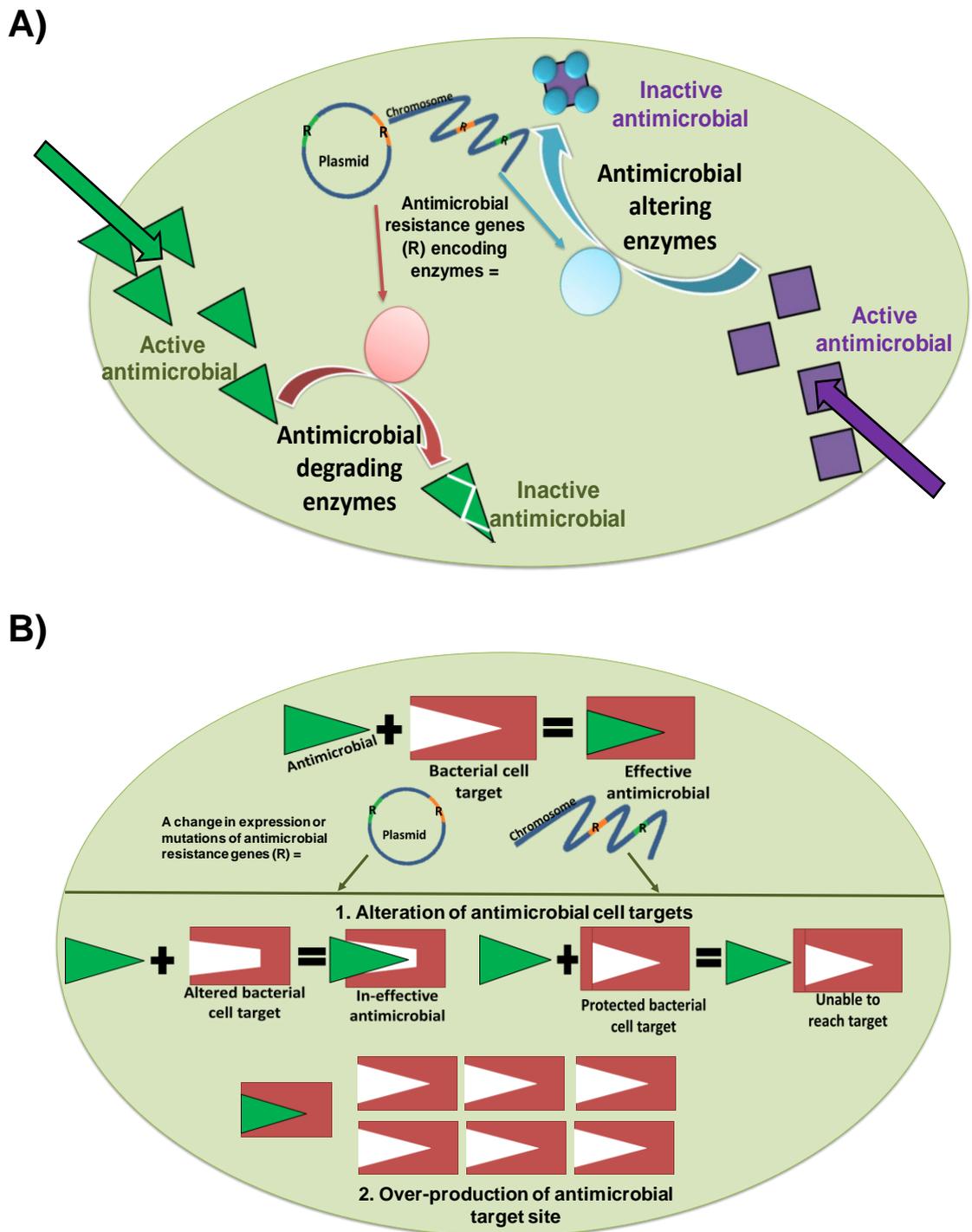


Figure 3. Resistance mechanisms used to prevent antimicrobials within the cell from reaching target sites

In order to prevent antimicrobials which are able to enter the cell from reaching their intended target sites, antimicrobial inhibiting enzymes can be encoded by antimicrobial resistance genes (A). These enzymes can break up the antimicrobial or alter its structure, ensuring it's not able to reach its intended target from within the cell. The antimicrobial target sites can also be modified and protected so the antimicrobial cannot interact with it effectively or the target sites can be overproduced (B), which minimises the impact of the antimicrobial.

1.2.1.4 *P. aeruginosa* biofilm formation

P. aeruginosa is able to produce a polysaccharide alginate layer which can provide an additional barrier when exposed to antimicrobials and environmental stresses (Orgad et al., 2011, Taylor et al., 2014). The alginate polysaccharide and other secreted moieties enable the bacteria to aggregate and grow as a biofilm structure (Figure 4). This is a complex accumulation of bacteria embedded within self-produced, hydrated polymer exopolysaccharide (EPS) matrix (Whiteley et al., 2001, Flemming and Wingender, 2010); the EPS represents 85% of the total biofilm biomass (Flemming et al., 2007). *P. aeruginosa* chronic biofilm infections are commonly reported in patient tissues and body surfaces as well as on medical devices including indwelling venous and urinary catheters and joint replacements (Mulcahy et al., 2014). Biofilms of *P. aeruginosa* have also been found in waste water treatment systems and are difficult to eradicate (Breathnach et al., 2012).

The formation of a biofilm is an endless cycle and involves planktonic cells reversibly attaching to a biotic or abiotic surface with the adhesion becoming irreversible with the production of an EPS matrix (Figure 4, stages 1 and 2) (Ma et al., 2009). Multicellular biofilm communities then become more densely structured with non-colonised space quickly becoming filled with bacteria and a three-dimensional community is observed (Figure 4, stages 3 and 4) (Flemming and Wingender, 2010). Lastly, bacterial cells detach from a mature sessile biofilm structure and revert back to a planktonic state to spread and restart colonisation and initiation of biofilm formation on other surfaces (Figure 4, stage 5 and 1) (Chua et al., 2014, Rasamiravaka et al., 2015). A swarming distribution technique has also been described in non-mucoid *P. aeruginosa* strains, where after initial biofilm growth and the development of an outer layer of sessile bacteria, motile cells of planktonic phenotype are observed within the inner regions and can swarm out of the multicellular community (Tolker-Nielsen et al., 2000, Sauer et al., 2002). A hollow mound in the centre of a biofilm is a characteristic of this swarming dispersal technique (Hall-Stoodley et al., 2004)

Exploring the regulatory networks involved in biofilm formation has recognised cyclic diguanosine monophosphate (ci-di-GMP) as a common initiator of biofilm formation (Karatan and Watnick, 2009). Increased intracellular levels of ci-di-GMP have been shown to promote the adjustment from a planktonic to sessile lifestyle in Gram-negative bacteria (Simm et al., 2004, Karatan and Watnick, 2009). The importance of two-component regulatory systems has also been recognised in regulating *P. aeruginosa* biofilm development, with the highly conserved GacS-GacA system being involved in EPS production via the use of small regulatory RNAs (Wei and Ma, 2013, Mikkelsen et al., 2011b). Biofilm communities have been shown to communicate via quorum sensing (QS) molecules which enable cells to identify each other's presence and regulate gene expression based on the extent of the population (Camilli and Bassler, 2006, Wei and Ma, 2013). Quorum sensing involves the production, secretion, and detection of Auto-Inducer (AIs) molecules such as acyl-homoserine-lactone (AHL), which can alter bacterial community responses within a biofilm (Camilli and Bassler, 2006). Null mutations in genes associated with QS, such as the *P. aeruginosa* PQS gene clusters *pqsABCD*, *phnAB* and *pqsH*, impact on the development and structural integrity of a biofilm (Lee and Zhang, 2015, Vital-Lopez et al., 2015).

In addition to exopolysaccharides, the production of large quantities of extracellular DNA has been reported during *P. aeruginosa* biofilm formation (Muto and Goto, 1986, Whitchurch et al., 2002, Allesen-Holm et al., 2006). This extracellular DNA is organised into distinctive grid-like patterns within the biofilm matrix (Allesen-Holm et al., 2006). High concentrations are identified on the outer parts of a mature biofilm, suggesting a cause for the build-up of migrating bacteria frequently observed in this area (Allesen-Holm et al., 2006, Flemming and Wingender, 2010). The origin of extracellular DNA within a biofilm appears to be from indiscriminate chromosomal DNA (Allesen-Holm et al., 2006) and the autolysis of cells, controlled by the QS system, within multicellular communities (Webb et al., 2003, Montanaro et al., 2011, Zemke and Bomberger, 2016). The presence of extracellular DNA has also been shown to contribute to antimicrobial resistance in *P. aeruginosa* biofilms (Mulcahy et al., 2008, Wilton et al., 2016).

Growing as a biofilm community provides the bacteria with considerably diminished susceptibility to antimicrobials when compared to planktonic cells (Stewart and William Costerton, 2001, Singh et al., 2017). This increased resistant characteristic of biofilms is a result of reduced or incomplete diffusion of antimicrobials; the EPS matrix restricts the amount of antimicrobials which can reach the bacteria below the outer layers of the biofilm, causing limited target binding (Stewart and William Costerton, 2001, Walters et al., 2003). The interaction of EPS matrix glucose polymers with antimicrobial agents has also been suggested as a mechanism of preventing agents from reaching target sites. *P. aeruginosa* with a mutation in an *ndvB* gene was unable to produce periplasmic glycans, and biofilms were more susceptible to treatment with the tobramycin antibiotic compared to the wild-type strain (Mah et al., 2003). It is proposed that glucose polymers are able to isolate antimicrobial agents within the periplasm, preventing their diffusion throughout the biofilm. Antibiotics which are able to penetrate and diffuse *P. aeruginosa* biofilms, such as tobramycin and ciprofloxacin, are still not effective at completely eradicating cells (Walters et al., 2003). Within a biofilm structure a number of cells, termed persisters undergo physiological changes due to the limited access to oxygen and nutrients, resulting in reduced metabolic activity and limited growth rate (Whiteley et al., 2001). As antimicrobials are more effective against metabolically active bacteria, persister cells demonstrate increased tolerance levels and are able to remain viable within a biofilm after exposure to antimicrobial treatments (Tumah, 2009, Walters et al., 2003).

The wide-ranging characteristics of *P. aeruginosa* (Figure 5) contribute to its success as an important opportunistic pathogen. Its metabolic versatility, low nutritional requirements, proficiency at forming biofilms and antimicrobial resistance mechanisms, also enable this microorganism to contaminate pharmaceutical and home and personal (HPC) products, which present harsh environmental conditions to survive in. The next section will consider the problem of HPC microbial contamination, the use of preservation systems in product formulations and known antimicrobial resistance mechanisms used in response to industrial antimicrobials.

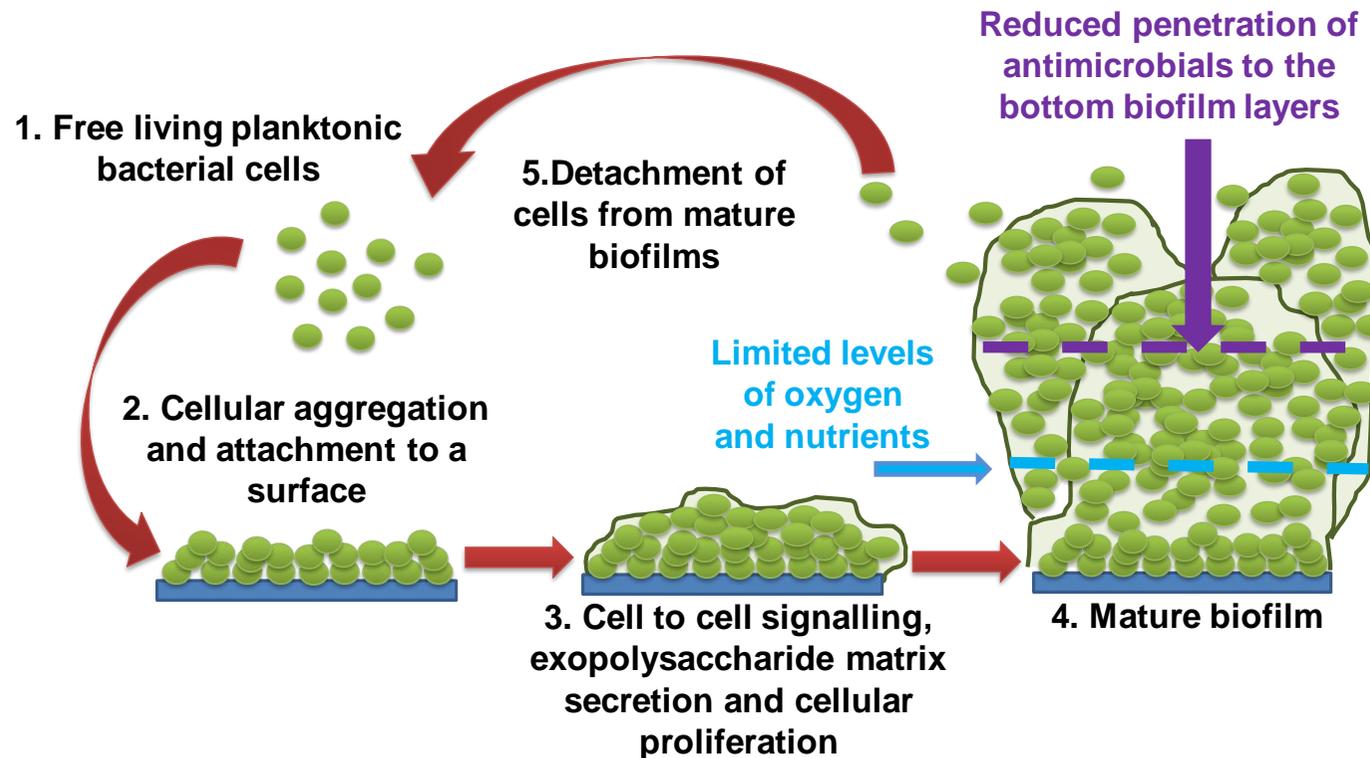


Figure 4. Stages in the formation of a biofilm

P. aeruginosa is able to grow as free living bacterial cells or as a multicellular accumulation of sessile cells known as a biofilm. The stages of biofilm formation involve planktonic cells attaching to a surface and forming aggregates followed by the development of an exopolysaccharide matrix and increased cell to cell communication. The cells then proliferate to form a mature biofilm which has limited oxygen levels and reduced antimicrobial penetration. Planktonic cells are able to detach from a mature biofilm and start the formation of a new biofilm elsewhere. Adapted from (Ma et al., 2009, Flemming et al., 2007, Taylor et al., 2014).

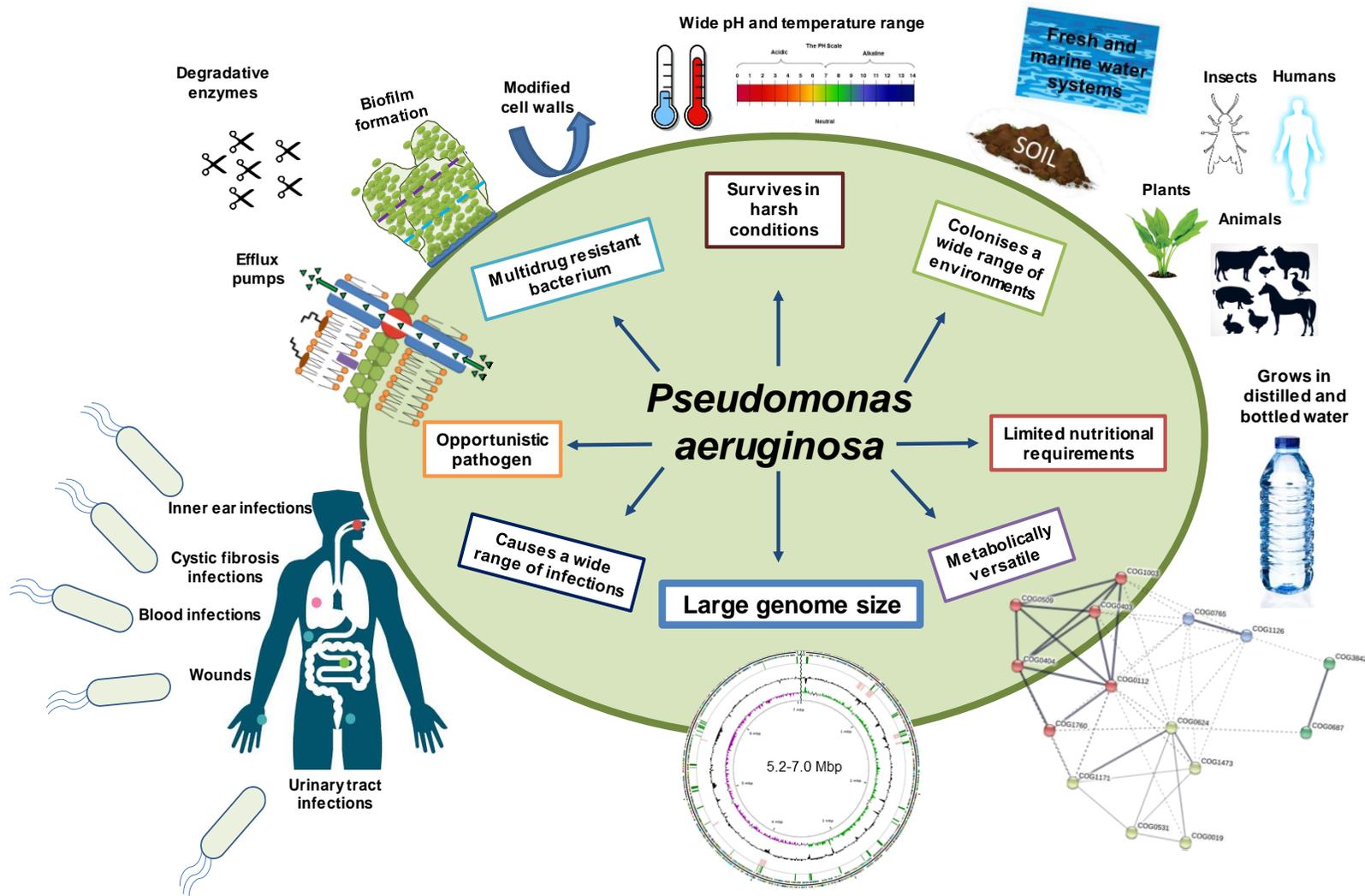


Figure 5. Overview of the ubiquitous microorganism *Pseudomonas aeruginosa*

The key functional characteristics of *P. aeruginosa* enable the bacterium to survive in a wide range of environments owing to its limited nutritional requirements and metabolic versatility. The microorganism is an opportunistic pathogen which can cause a wide range of infections and a major concern is the growing problem of multi-drug resistance. The overall versatility observed with *P. aeruginosa* is linked to the large genome size which ranges between 5.2-7.0 Mbp.

1.3 The challenge of microbial contamination in home and personal care industry

1.3.1 Microbial contamination in home and personal care products

The home and personal care (HPC) industry manufactures a broad range of products from household cleaners and laundry detergents through to shampoos, soaps, body creams and cosmetics (Orth et al., 2006, Orus and Leranoz, 2005). The abundance of HPC product types results in high quantities being used globally on a daily basis (Daughton and Ternes, 1999, Boxall et al., 2012). There is no requirement for HPC products to be sold sterile (Lundov et al., 2009), however there is a legal responsibility for manufacturers to ensure product safety for consumer use (Brannan, 1995, Orus and Leranoz, 2005, FDA, 2017b). Good Manufacturing Practices (GMPs) are incorporated into industrial manufacturing procedures to reduce the occurrence of unacceptable levels or types of microorganisms, from contaminating the unsterile raw materials and water to be incorporated into products (Brannan, 1997, Jimenez, 2001, El-Housseiny et al., 2013).

The microbial contamination of HPC products can occur, especially in those with a high water content (Siegert et al., 2011), with reports of Gram-negative contamination in consumer products such as body and hand creams (Álvarez-Lerma et al., 2008, Becks and Lorenzoni, 1995), mouthwashes (Stephenson et al., 1984, Prevention, 1998, Kutty et al., 2007) and hand washes (Zapka et al., 2011). The proliferation of bacteria within a HPC product can result in financial and reputational losses for the manufacturer due to product recalls, may present a risk to consumer health and could be problematic for immunocompromised individuals (Russell, 2003b, Orth et al., 2006, Jimenez, 2007). Microbial contaminants can also breakdown the product which can result in alterations to the pH value, colour, odour, and viscosity, which can all impact on the product performance (Smart and Spooner, 1972); packaging of HPC products may also swell and become damaged with the production of microbial gases. Even though HPC products are not natural environments for microorganisms to grow, they do contain a variety of nutrients suitable for metabolism such as water, lipids, polysaccharides, amino acids, glycosides, peptides and vitamins (Herrera, 2004, Neza and Centini, 2016)

Contamination of HPC products can occur either through the manufacturing process or by the consumer themselves (Farrington et al., 1994, Siegert et al., 2011). Products may already have a high microbial burden when sold to consumers, resulting from poor production procedures, storage and transport of goods. Raw materials and water containing bacteria can become incorporated into the product manufacturing process, and non-sterile industrial environments can introduce microbes (Jimenez, 2001, Kulakov et al., 2002, Siegert et al., 2011, El-Housseiny et al., 2013). Furthermore, non-sterile products can acquire microorganisms from the consumer and the domestic environment after purchase, especially if kept after the recommended 'best used before' and 'period after opening' guidelines as implemented by the European Union (Jimenez, 2001, Baird, 2007, Lundov et al., 2009, Orus and Leranoz, 2005).

The Gram-negatives *P. aeruginosa* and *Burkholderia cepacia* are considered the most common bacteria related to recalls of non-sterile HPC products (Wong et al., 2000, Lundov and Zachariae, 2008, Chen et al., 2010, Neza and Centini, 2016). Other frequently isolated microbial contaminants include *Salmonella*

spp., *Staphylococcus aureus*, *Enterobacter gergoviae*, *Enterococcus spp.*, *Klebsiella oxytoca*, *Serratia marcescens* and *Escherichia coli* (Jimenez, 2001, Jimenez, 2007, Lundov and Zachariae, 2008, Neza and Centini, 2016) as well as *Candida albicans* and *Rhodotorula* (Dadashi and Dehghanzadeh, 2016). There is also a high number of unidentified microorganisms which are recovered from recalled products (Jimenez, 2007, Sutton and Jimenez, 2012). Currently there is no requirement for HPC manufacturers to disclose and publish incidences of product contamination, which results in a reduced appreciation of the different bacterial species capable of proliferating within consumer products. This also causes consumers to be unaware of the problem of microbial contamination as well as prevention and protection measures.

1.3.2 Preventing microbial contamination in the HPC industry

Preservative formulations in the HPC industry are key strategies, which aim to achieve microbial sterility of products after being sold to a consumer. Preservatives are a class of biocide antimicrobials with a broad spectrum of activity and are added to products with the intent to inhibit the growth of microorganisms and prevent the breakdown of industrial formulations (Reinhard et al., 2001, Ikarashi et al., 2009, Mikami et al., 2002, Gilbert and McBain, 2003). Globally by weight, preservatives are the most common type of antimicrobial which microorganisms are exposed to, due to the high occurrence of HPC product usage (Gilbert and McBain, 2003, Ikarashi et al., 2009, Alvarez-Rivera et al., 2012). Currently, isothiazolinones and alcohol-based preservatives are widely used in industry to avoid the proliferation of bacteria within aqueous-based home and personal care products (Reinhard et al., 2001, Alvarez-Rivera et al., 2012, Brannan, 1997). These two forms of preservatives are of great interest to the HPC industry, and there is a requirement to increase the understanding of their mechanisms of action and how bacteria respond when exposed to them. This form of investigation can enhance the development of preservation formulations, which incorporate isothiazolinones and alcohols.

Other types of preservatives not as commonly utilised in HPC products include parabens, the biguanide biocide chlorhexidine, formaldehyde and formaldehyde-releasing agents. The reduced use of these preservatives is due to concerns over their safety with reports of parabens being linked to tumour development (Harvey and Everett, 2004) and a rise in the incidences of contact allergies when formaldehyde preservatives are incorporated into personal care products (Lundov et al., 2009, Sasseville, 2004). Rare but severe allergic reactions have also been reported with increased contact of chlorhexidine on the skin (FDA, 2017a).

Preservative enhancers are also of interest to the HPC industry as these can amplify the antimicrobial efficacy of product preservation formulations without the need to increase concentrations of the antimicrobials (Orth et al., 2006). Examples of enhancers which can be incorporated into HPC products include chelating agents, such as ethylenediamine tetracetic acid (EDTA), which increases cell membrane permeability by interfering with levels of Mg^{2+} and Ca^{2+} (Brannan, 1997). Compounds which can interrupt the phospholipid bilayer such as caprylyl glycol, ethylhexylglycerin and sorbitan caprylate, are also used to enhance the performance of preservation systems (Papageorgiou et al., 2010). Agents which can restore the susceptibility of microorganisms to an antimicrobial, which have become resistant due to

efflux pumps, are also of interest to the HPC industry (Askoura et al., 2011). Phenyl-arginine-beta-naphthylamide (PA β N) can act as an efflux pump inhibitor and also increases the penetrability of the Gram-negative cell membrane when used in combination with several antibiotics (Lamers et al., 2013).

When compared to antibiotics, preservative mechanisms of action are not as well established and appear to have multiple cellular targets. The activity and application of two widely used HPC preservative categories will now be considered.

1.3.2.1 Isothiazolinone preservatives

The isothiazolinone preservatives, also known as isothiazolones, are heterocyclic chemical compounds that include benzisothiazolone (BIT), N-methylisothiazolone (MIT), 5-chloro-N-methylisothiazolone (CMIT) and a blend of chloro-N-methylisothiazolinone and N-methylisothiazolinone (CITMIT). These agents are commonly used to prevent the proliferation of microorganisms in settings such as cooling towers, paint products, the manufacturing of paper and textiles as well as HPC products (Laopaiboon et al., 2001, Xu et al., 2009, Lundov et al., 2014). Isothiazolinones have a wide application range against Gram-positive and Gram-negative bacteria and also fungi (Pucci et al., 2011, Lundov et al., 2014) which makes them suitable as preservatives for HPC products. These types of preservatives have also been associated with contact allergies (Lundov et al., 2009, Chomiczewska-Skóra et al., 2014), with only MIT and CITMIT approved for use in personal care products at concentrations 0.01% and 0.0015% active respectively. However, their use is limited in these types of products, as industrial companies seek to uphold their reputation at delivering safe products. BIT is widely incorporated into preservation systems of household cleaning products such as washing up liquids and laundry detergents in order to prevent microbial contamination (Table 2) (Alvarez-Rivera et al., 2012). The recommended concentration usage levels of BIT in these products is 0.05 – 0.15% active.

The main mechanism of action for isothiazolinone preservatives is the oxidation of the thiols groups found frequently on bacterial enzymes and proteins; this results in the inhibition of growth and metabolic activity of microorganisms (Collier et al., 1990a, Collier et al., 1990b, Williams, 2007). The oxidation effects of these preservatives can also result in the generation of free radicals causing oxidative stress and further irreversible damage to the cell (Williams, 2007). Isothiazolinones have been shown to have direct inhibitory effects on metabolic enzymes such as dehydrogenases (Collier et al., 1991), isomerases (DNA gyrase and topoisomerase IV) (Cheng et al., 2007, Pucci et al., 2011), oxidoreductases (Collier et al., 1991), acetyltransferases (Stimson et al., 2005, Dekker et al., 2009, Furdas et al., 2011), polymerases (Yan et al., 2007) kinases (Lippa et al., 2006, Reddy et al., 2012), lipases (King et al., 2009), synthases (Gedi et al., 2011). This therefore highlights the multifactorial isothiazolinone impact on bacterial metabolism. Investigations into mechanisms of action of the preservative BIT and its derivatives, have identified its ability target and inhibit the sortase-A enzyme which is involved in Gram-positive peptidoglycan cell wall biosynthesis (Zhulenkovs et al., 2014). In *P. aeruginosa* the regulatory c-di-GMP phosphodiesterase RocR enzyme was inhibited by a BIT derivative and a reduced swarming ability was also noted (Zheng et al., 2016). It was suggested the BIT derivative was not suitable for growth inhibition purposes, but would be

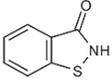
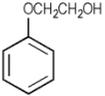
useful to attenuate the virulence of *P. aeruginosa*, as mutants deficient in c-di-GMP phosphodiesterase remained viable but had reduced virulence.

1.3.2.2 Alcohol-based preservatives

Alcohol-based preservatives can also be incorporated into HPC preservation formulations and types used include phenoxyethanol (POE) (Ikarashi et al., 2009, Krowka et al., 2014), benzyl alcohol, (Nair, 2000, Mikami et al., 2002) and phenethyl alcohol (Ziosi et al., 2013). The use of POE in personal care products has increased due the low occurrence of allergy contact sensitisation reported with this preservative (Poudrier, 1990, Steinberg, 2008, Johnsson et al., 2011, Krowka et al., 2014). Under European Union regulations, POE is recommended to be used in concentrations between 0.5-1% in personal care products (Table 2) and this preservative is also highly compatible with numerous product formulations (Krowka et al., 2014).

POE demonstrates antimicrobial properties against Gram-negative bacteria such as *P. aeruginosa*, (Orth et al., 2006) and the proposed mechanism of action is the disruption of bacterial cell membranes, cell lysis and leakage of cellular protein material (Fitzgerald et al., 1992, Maillard, 2002, Langsrud et al., 2016). However, the exact mechanism of POE antimicrobial activity is not fully comprehended. Gilbert et al. carried out numerous studies on the effect of POE with *E. coli*, in the 1970's and multiple mechanisms, which caused inhibitory and bactericidal actions, were noted. The antimicrobial activity of POE was associated with disruptions and inhibitions of malate dehydrogenase enzyme-systems involved in the Krebs cycle (Gilbert et al., 1976a, Gilbert et al., 1976b) and RNA, DNA and protein-biosynthesis pathways (Gilbert et al., 1979) as well as leakage of membrane protons which decreased cell wall permeability (Gilbert et al., 1977). These investigations into the mechanisms of action of POE demonstrate the multiple potential targets of the preservative. It has also been suggested that the antimicrobial activity is concentration dependant and bactericidal effects result from the preservative interfering with several of the proposed targets, causing irreversible cell damage (Langsrud et al., 2016). The effectiveness of POE is also enhanced when combined with other preservatives such as chlorhexidine (Fitzgerald et al., 1992) isothiazolinones (Lundov et al., 2011) and ethylhexylglycerin (Beilfuß et al., 2005, Langsrud et al., 2016).

Table 2. Important isothiazolinones and alcohol-based preservatives used in the HPC industry

Preservative Category	Preservative trade name, active ingredient, solvent, supplier and recommended usage levels.	Chemical structure	Preservative applications	Targets and mechanisms of action	References
Isothiazolinones	<p>PROXEL GXL</p> <ul style="list-style-type: none"> Active ingredient: 1,2-benzisothiazolinon-3-one BIT supplied as 20% active. Soluble in dH₂O. Yellow brown Solution Supplied by Arch UK Biocides Ltd LONZA, UK. Recommended usage levels: 0.05 – 0.15% active. 		<p>House hold products:</p> <ul style="list-style-type: none"> Liquid laundry detergents and fabric softeners. Dishwasher detergents. All purpose cleaners. Polishes/waxes. Incorporated into raw materials and surfactants used in the Household products. Not used in personal care products due to sensitising capabilities. 	<ul style="list-style-type: none"> Nonspecific mode of action/multiple cellular targets. Bacteriostatic effects. Electrophilic agents which react with nucleophilic cell entities. Oxidation of thiol groups, found on enzymes and other proteins, which are essential for function. Loss of proteins as thiols disrupts pathways critical for cell metabolism. Direct inhibitory effects on following metabolic enzymes: dehydrogenases, oxidoreductases, acetyltransferases, kinases, DNA gyrase and topoisomerases, lipases, kinases, polymerase, and synthases. Oxidation generates free radicals which further damage the cell. 	<p>(Collier et al., 1990a, Collier et al., 1990b, Collier et al., 1991, Lippa et al., Stimson et al., 2005, 2006, Williams 2007, Cheng et al., 2007, Yan et al., 2007, Dekker et al., 2009, King et al., 2009, Gedi et al., 2011, Furdas et al., 2011, Pucci et al., 2011, Alvarez-Rivera et al. 2012, Reddy et al., 2012, Zhou et al., 2014, Zhou et al., 2015b)</p>
Alcohols	<p>PHENOXYETOL</p> <ul style="list-style-type: none"> Active ingredient: 2-Phenoxyethanol POE supplied as >99 % active. Soluble in dH₂O. Supplied by Clariant Produkte GmbH, Germany. Recommended usage levels: 0.5 – 1.0 % active. 		<p>Personal care products:</p> <ul style="list-style-type: none"> Personal skin care preparations due to its low sensitising capabilities. Aqueous based home and personal care products such as shampoos, foam baths, shower gels or liquid detergents. Cosmetic makeup products. 	<ul style="list-style-type: none"> Good activity against Gram-negative bacteria. Exhibits a micro-biocidal activity. Membrane active lytic agent, disrupting cellular membranes and causing the leakage of intracellular components. Increases permeability of the cell membrane due to proton leakage Denaturing of essential proteins and enzymes. Disrupts RNA, DNA and protein synthesis. Inhibits malate dehydrogenase which is involved in the Krebs cycle. 	<p>(Gilbert et al., 1976a, 1976b, 1977, 1979), (Poudrier 1990, (Fitzgerald et al., 1992 Maillard 2002, Orth et al., 2006 Ikarashi et al., 2009, Krowka et al., 2014, Langsrud et al., 2016)</p>

1.3.3 *P. aeruginosa* as a contaminant of the HPC industry

Due to its dominance as a health care pathogen, research has mainly focused on the epidemiology, pathogenesis and treatment of clinical *P. aeruginosa* isolates. In comparison, there is an incomplete understanding of *P. aeruginosa* strains isolated from the natural environment (Chavarría et al., 2013) and much less is known about isolates which are successful contaminants of non-sterile cosmetic, personal and household products (Jimenez, 2007, Lundov and Zachariae, 2008, Siegert et al., 2011, Siegert, 2012). As previously noted, *P. aeruginosa* can be recovered from numerous environmental niches (Figure 5), and due to its metabolic versatility, has the ability to proliferate in harsh chemicals and disinfectants (Dantas et al., 2008, Kung et al., 2010, Lama et al., 2012). Consequently its aptitude to grow within the unnatural setting of a HPC product is not surprising, and a recent analysis of recalled contaminated products, identified *P. aeruginosa* as the most frequently isolated microorganism (35.48%) (Neza and Centini, 2016).

Due to the opportunistic characteristics of this pathogen, high levels of the bacterium within a product could be harmful to those with weakened immune systems. The action of some personal care products can cause damage to skin and mucous membranes, which could promote infection if microbial contaminants are present (Neza and Centini, 2016). *P. aeruginosa* is commonly found in cosmetics (Tan et al., 2013), shampoo products (Neza and Centini, 2016), contact lens solutions and eye cosmetic products, where it has been reported to cause a number of infections of the eye and surrounding tissues (Reid and Wood, 1979, Blumenfeld et al., 2005, Hazlett, 2004, Manuj et al., 2006). This highlights the requirement to minimise the occurrence of HPC product contamination with this pathogen through the improvement of manufacturing practices, and with enhanced product preservation systems. *P. aeruginosa* strains which are resistant to isothiazolinone preservatives are regularly identified from incidences of product contamination where spoilage has occurred (Chen et al., 2010). This suggests the bacterium is able to overcome preservation strategies used in industry. In order to understand *P. aeruginosa* in an industrial context, those isolated from contaminated products should be analysed in depth and resistant mechanisms used in response to preservative formulations require characterisation.

1.4 Mechanisms of resistance to industrial antimicrobial agents

The majority of research investigating *P. aeruginosa* antimicrobial resistance has focussed on the mechanisms employed to reduce susceptibility to antibiotics. There is limited knowledge of the resistance mechanisms used by *P. aeruginosa* when exposed to industrial antimicrobials such as preservatives (Chapman, 1998). Industrial preservatives have been observed to target multiple sites within bacteria and cause universal damage with efficient bactericidal effects. However, the presence of preservatives at sub-lethal concentrations can result in incidences of adaptive resistance in *P. aeruginosa* (Brözel and Cloete, 1994, Winder et al., 2000, Mc Cay et al., 2010, Abdel-Malek and Badran, 2010). This section will consider characterised bacterial mechanisms of resistance in response to industrial antimicrobial agents.

1.4.1 Alteration of cell wall structures in response to industrial antimicrobials

A principal resistance mechanism of Gram-negative bacteria in response to preservatives is the reduction of cellular permeability, by altering or removing cell envelope components such as proteins, phospholipids and fatty acids (Chapman, 2003, Tumah, 2009). A characterised *P. aeruginosa* resistance mechanism in response to the isothiazolinone CITMIT, was associated with an alteration in the outer membrane, via the removal of the OprD protein (Brözel and Cloete, 1994). The protein was readily detected in non-resistant strains, however in those with decreased susceptibility to isothiazolinones, OprD could not be detected (Brözel and Cloete, 1994). This suggests OprD was the mode of entry for isothiazolinones and *P. aeruginosa* had adapted to exposure via the modification of the cell wall structure. Winder et al., (2000) also identified the removal of the OprD in the outer membrane in a PAO1 strain with induced resistance to BIT, MIT, and CMIT. Furthermore this study found OprD reappeared when resistant strains were passaged without a preservative, however induced resistance was not fully reversed, suggesting the occurrence of additional adaptations (Winder et al., 2000). In addition, induced resistance towards isothiazolinone preservatives resulted in decreased susceptibility to other thiol-interactive agents.

A *P. aeruginosa* strain with adapted resistance to the preservative POE also demonstrated alterations in the outer membrane protein profile, with decreases in protein densities observed (Abdel-Malek and Badran, 2010). An interesting observation was the cross-resistance demonstrated with unrelated biocides and with penicillin and aminoglycoside antibiotics. In the presence of the preservative BIT, a three-fold reduction in susceptibility was observed with the POE adapted strain, suggesting these two preservatives have a common point of entry in *P. aeruginosa*. The reduced susceptibility observed was not suggested to be linked to an efflux pump as the resistance did not reduce in the presence on an efflux pump inhibitor (Abdel-Malek and Badran, 2010). These investigations highlight modification of cell wall structures as a fundamental *P. aeruginosa* resistance mechanism in response to HPC preservatives.

1.4.2 The use of efflux pumps in response to industrial antimicrobials

In order to reduce intracellular concentrations of antimicrobials, which are able to cross the cell wall, Gram-negative bacteria overexpress efflux systems. However, there is limited research identifying the specific pump systems used by *P. aeruginosa* in response to industrially used antimicrobials. Efflux pumps associated with multidrug resistance are able to transport a range of structural substrates as demonstrated by Table 1; these established systems within *P. aeruginosa* could also be used to expel industrial preservatives and prevent their accumulation within the cell.

The use of efflux as a primary resistance response to the broad-spectrum industrial biocide triclosan was identified in the up-regulation of multiple efflux pump systems in *P. aeruginosa* due to mutations in the *nfxB* and *mexL* regulatory genes (Schweizer, 2001, Chuanchuen et al., 2001, Chuanchuen et al., 2002). The MexJK efflux system (Table 1) functions as a two-component RND pump to expel triclosan, however the presence of an outer membrane channel protein is necessary to form a tripartite antibiotic efflux system (Chuanchuen et al., 2002, Chuanchuen et al., 2005). A *P. aeruginosa* triclosan specific efflux pump has also been identified TriABC-OpmH (Table 1) (Mima et al., 2007). These studies all demonstrate the importance of efflux as a principal *P. aeruginosa* resistance mechanism in response to triclosan.

Another example of efflux as a biocide resistance mechanism in *Pseudomonas spp.*, was identified with glutaraldehyde, a widely used antimicrobial in hospitals and water treatment (Vikram et al., 2015). Transcriptomic analysis identified up-regulation of genes encoding an efflux pump in *Pseudomonas fluorescens* which had high homology to the *P. aeruginosa* MexEF-OprN efflux pump (Table 1) (Vikram et al., 2015). When PAO1 was treated with both glutaraldehyde and an efflux pump inhibitor, it demonstrated increased susceptibility compared to treatment with glutaraldehyde only (Vikram et al., 2015). This study proposed efflux as a *P. aeruginosa* resistance mechanism in response to the glutaraldehyde biocide. The MexCD-OprJ efflux pump system in *P. aeruginosa* (Table 1) is also inducible during exposure to the disinfectants chlorhexidine and benzalkonium chloride as well as other membrane lytic agents (Morita et al., 2003, Fraud et al., 2008).

Transcriptomic research has also identified an RND efflux pump as a key resistance mechanism when a *Burkholderia lata* (*B. lata*) strain with induced resistance to isothiazolinone preservatives was exposed to CMIT (Rushton et al., 2013). *B. lata* are also common Gram-negative bacterial contaminants of the HPC industry (Wong et al., 2000). This suggests efflux as a possible resistance mechanism in response to preservative formulations, which could also be used by *P. aeruginosa* HPC product contaminants.

1.4.3 Alteration of bacterial targets sites and inhibition of industrial antimicrobial agents

The alteration of antimicrobial target sites within the cell has also been identified as a resistance mechanism of *P. aeruginosa* in response to triclosan. This biocide targets the bacterial enoyl-acyl carrier protein (ACP) FabI reductase within the fatty acid biosynthetic pathway (Chuanchuen et al., 2001). *P. aeruginosa* is able to confer resistance when exposed to triclosan by alteration of the FabI enzyme, rendering it insensitive to the biocide (Chuanchuen et al., 2001, Schweizer, 2001). There is limited research to suggest the modification of target sites is utilised by *P. aeruginosa* to reduce susceptibility to industrial preservatives, due to the broad spectrum activity demonstrated by these antimicrobials (Williams, 2007, Alvarez-Rivera et al., 2012). The inhibition of industrial antimicrobial agents is demonstrated by the ability of *P. aeruginosa* to detoxify formaldehyde via a NAD⁺ glutathione-dependent dehydrogenase enzyme, which is predicted to be plasmid-encoded (Cloete, 2003).

Recent transcriptomic investigations of an isothiazolinone-resistant *P. aeruginosa* strain when exposed to CITMIT, identified a number of differentially expressed genes associated with nitrogen metabolism and oxidative phosphorylation pathways (Zhou et al., 2016). It was proposed that the upregulation of genes encoding nitric oxide synthases was to increase the availability of nitric oxide (NO) in the cell (Zhou et al., 2016). NO can structurally alter and inhibit toxic compounds and alleviate the oxidative stress, which increases tolerance to a wide-range of antibiotics such as lactams, aminoglycosides, and quinolones (Gusarov et al., 2009, McCollister et al., 2011). The isothiazolinone-resistant strain also generated increased levels of NO when compared to the wild-type strain, (Zhou et al., 2015b, Zhou et al., 2016), thus highlighting the importance of increased NO production in the resistance of *P. aeruginosa* to CITMIT.

1.4.4 Biofilm resistance and industrial antimicrobial agents

A delay in the ability of industrial antimicrobials to penetrate *P. aeruginosa* biofilm structures is also a characterised resistance mechanism. When the biocidal activity of peracetic acid and benzalkonium chloride was assessed, the presence of the EPS matrix obstructed the ability of the biocides to diffuse throughout the *P. aeruginosa* biofilm (Bridier et al., 2011). A reduced susceptibility to the oxidative biocides monochloramine and hydrogen peroxide when *P. aeruginosa* formed thin biofilms, has also been reported in comparison to planktonic cells (Cochran et al., 2000). In *B. lata* strains which demonstrated decreased susceptibility to CMIT isothiazolinone preservatives and the biocide benzethonium chloride, an increase in biofilm biomass formation was also observed (Rushton et al., 2013). This suggests biofilm development is a possible resistance mechanism for strains, which are able to contaminate HPC products such as *B. lata* and *P. aeruginosa*.

1.4.5 Cross-resistance between industrial antimicrobial agents and antibiotics

The overuse of industrial antimicrobial agents may contribute to the emergence of highly resistant microorganisms, and increased tolerance to preservatives could select for strains which are also less susceptible to clinically-relevant antibiotics (McBain et al., 2002, Poole, 2002, Russell, 2003a). Resistance mechanisms used by *P. aeruginosa* in response to triclosan have been shown to also convey resistance to numerous antibiotics (Carey and McNamara, 2014) especially when sub-inhibitory concentrations are utilised (Schweizer, 2001). Cross-resistance between triclosan and antibiotics such as ciprofloxacin in *P. aeruginosa* is facilitated by the overexpression a multidrug efflux pump MexCD-OprJ (Table 1) as a result of mutations in its regulatory gene *nfxB*, selected for by triclosan (Chuanchuen et al., 2001). Preservative cross-resistance and altered antibiotic susceptibility was also reported in a *B. lata* strain with induced resistance to benzethonium chloride and isothiazolinones including BIT (Rushton et al., 2013). Another interesting link between preservative and antibiotic resistance was observed with the reduced susceptibility of a POE adapted *P. aeruginosa* strain to clinically relevant antibiotics and other unrelated preservatives and biocides (Abdel-Malek and Badran, 2010). A *P. aeruginosa* strain with increased tolerance to the isothiazolinone CITMIT was also found to display cross-resistance to triclosan and ciprofloxacin (Zhou et al., 2016). This suggests the preservatives used in HPC products such as POE and isothiazolinones, are capable of initiating cross-resistance to antibiotics in *P. aeruginosa*.

The high quantities of HPC products being used on a daily basis results in their continuous release into the environment (Daughton and Ternes, 1999). Consequently bacteria will be exposed to the preservative formulations, which are likely to be diluted and at sub-lethal levels. Strains, which are successful contaminants of HPC products, are also proficient at overcoming the preservation system. Therefore, the occurrence of reduced susceptibility towards industrially used preservatives and the evidence of cross-resistance, may result in the appearance of hyper-resistant antibiotic *P. aeruginosa* strains.

1.5 Project Aims

The Biotechnology and Biological Sciences Research Council (BBSRC) funded this PhD studentship in collaboration with the industrial CASE sponsor Unilever Safety and Environment Assurance Centre (SEAC). The interests of Unilever, with the intention of informing industrial practices, guided the research questions addressed. The overall aim of this work was to identify important genetic pathways, which contribute to preservative resistance in an industrial *P. aeruginosa* strain. The main approaches used to achieve the aims of this PhD project are summarised in Figure 6.

1.5.1 Research project rationale

Industry utilises preservative systems for non-sterile cosmetic and household items to inhibit microbial contamination, these agents are essential to prevent product spoilage and potential consumer infection. However, *P. aeruginosa* has the ability to overcome preservation formulations and the resistance mechanisms utilised to survive the presence of industrial antimicrobials are not well established. A better understanding of these mechanisms can contribute to optimal use of HPC preservation strategies via the development of targeted preservative formulations. The adaptive resistance mechanisms employed by *P. aeruginosa* in response to industrially used antimicrobials, may also be sufficient to confer multidrug resistance in a clinical environment. Therefore, there is concern that excessive industrial use of preservatives, may contribute to the emergence of highly antimicrobial resistant strains. This further highlights the requirement for industry to minimise the use of antimicrobials through the optimisation of preservation strategies.

The purpose of this project was to characterise key genetic pathways involved in resistance to industry relevant antimicrobials, to provide a better understanding of the success of *P. aeruginosa* as a HPC product contaminant. The identification of key genes and regulatory pathways is an invaluable approach when investigating microbial resistance mechanisms. An industrial *P. aeruginosa* strain RW109, isolated from a contamination incident, was chosen for this project. RW109 served as a model organism for investigating industrial antimicrobial resistance due to its previous exposure to HPC preservatives and product formulations. The use of the RW109 strain in this project also demonstrates the genetic and metabolic determinants, which enable *P. aeruginosa* to become a successful contaminant of the home and personal care industry. The antimicrobial agents selected for investigation included BIT to represent a commonly used preservative in household products for example, laundry detergents and washing-up liquids and POE, which is widely incorporated into personal care products such as shampoos, shower gels and cosmetics. The results of this work will inform the future development of targeted preservation systems to enhance product sterility whilst also minimising future resistance development.

The specific aims of the project were as follows:

- 1) Characterise the complete genome of an industrial *P. aeruginosa* strain (Chapter 3).** A complete genome of the industrial strain RW109 was obtained using PacBio Single Molecule, Real-Time (SMRT) technology. This enabled a comprehensive annotation of the functional characteristics of an industrial strain. A comparison of RW109 with a panel of *P. aeruginosa* genome sequences representing strains isolated from clinical, environmental and industrial sources was also completed.
- 2) Investigate the genetic basis of industrial *P. aeruginosa* preservative resistance (Chapter 4).** A mini-Tn5-*luxCDABE* transposon library was constructed using the RW109 industrial *P. aeruginosa* strain. A screening procedure was developed to identify mutants with altered light emission responses when exposed to industry relevant preservatives. Mutants of interest were linked to the putative gene pathways of the RW109 *P. aeruginosa* genome to identify genetic determinants of preservative resistance.
- 3) Investigate the global gene expression response of *P. aeruginosa* RW109 when exposed to industry-relevant conditions (Chapter 5).** Appropriate industrial exposure and growth conditions were determined which enabled the extraction of good quality RNA for subsequent transcriptomic investigations of the RW109 strain. An RNA-Seq experiment was implemented following exposure of RW109 to individual preservatives, preservative combinations and a HPC product with and without the addition of a preservative. The differentially regulated genes with significant increases and decreases in expression when exposed to the industry relevant conditions were analysed.
- 4) Reconstruct and utilise a genome-scale metabolic model specific to the industrial *P. aeruginosa* RW109 strain (Chapter 6).** A recently published PA14 genome-scale metabolic model was updated to represent the industrial RW109 strain. RNA-Seq transcriptomic data was integrated to generate condition-specific models that represent RW109 when exposed to industrial conditions. These models were used to predict the essentiality of metabolic reactions when RW109 was exposed to HPC preservatives and product formulations.

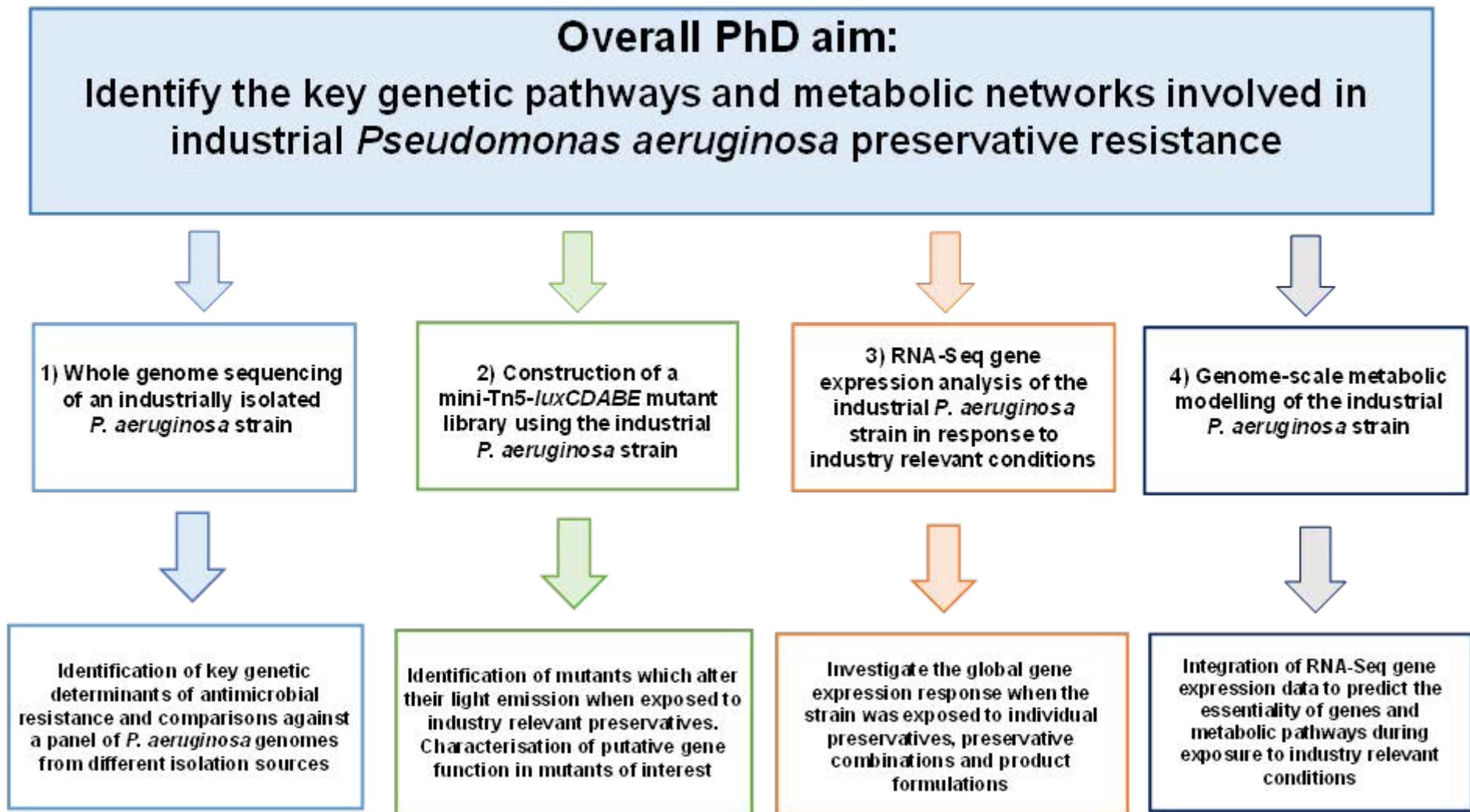


Figure 6. The main experimental approaches used to investigate the genetic basis of preservative resistance in an industrial *P. aeruginosa* strain

Four key strategies were used to identify genetic pathways and metabolic networks involved in industrial *P. aeruginosa* preservative resistance. These included: 1) Whole genome sequencing of an important industrial strain isolated from a source of contamination. 2) The use of a mini-Tn5-*luxCDABE* mutant library to identify mutants of interest when screened with preservatives. 3) An investigation of the global gene expression response when exposed to industry relevant conditions, with the use of RNA-Seq. 4) Genome-scale metabolic modelling of an industrial strain and integration of transcriptomic data to predict the essentiality of genes and reactions when the industrial strain was exposed to HPC preservatives and product formulations.

2. Materials and methods

2.1 Chemicals and growth media

The chemicals used throughout this study were supplied by Sigma-Aldrich (Dorset, UK), and Fisher Scientific (Loughborough, UK), unless otherwise stated. All media was prepared according to the manufacturer's guidelines using double-deionised water and sterilised by autoclaving at 121°C for 15 minutes. Tryptone soya medium (TSA agar and TSB broth) and Luria bertani (LB) medium were obtained from Oxoid Ltd (Basingstoke, UK) and were routinely used for growing *P. aeruginosa* and *Escherichia coli* (*E.coli*) strains respectively.

2.2 Bacterial Strains

2.2.1 Storage of bacterial isolates

Bacterial stocks used in this study were stored at -80°C in TSB or LB containing 8% Dimethylsulfoxide (DMSO) (Fisher Scientific).

2.2.2 Routine growth conditions of an industrial *P. aeruginosa* strain

Throughout this project, the industrial *P. aeruginosa* challenge test strain, RW109, was used. This strain was isolated from a contamination incident and was routinely incorporated into the industrial sponsor's challenge testing processes. These tests were used to determine the susceptibility of a product to microbial contamination, and the efficacy of the preservative system. RW109 had previously been exposed to industrial preservatives and product formulations, and served as a model organism to investigate industrial antimicrobial resistance. RW109 was routinely revived from a freezer stock onto TSA, and incubated static for 24 hours at 30°C. Overnight liquid cultures were prepared by inoculating 3 ml of TSB with confluent growth from a 24-hour plate culture, and incubated for 18 hours shaking with aeration at 30°C. Throughout this project, the lower incubation temperature of 30°C was used to be more representative of industrial product manufacture and storage temperatures, but still enabling sufficient growth for 24 and 18-hours incubation times.

2.2.3 Enumeration of viable bacteria

Bacterial cultures were enumerated using the Miles and Misra surface viable plate count method (Miles et al., 1938). Serial dilutions were performed in TSB broth unless otherwise stated; 100 µl drops were spread onto triplicate TSA plates and incubated at 30°C for 24 hours. Individual colonies were counted and the number of viable cells calculated and expressed as colony forming units per ml (CFU/ml).

2.3 Antimicrobial agents

2.3.1 Preservatives and formulations

Throughout this study, the preservative benzisothiazolone (BIT) was used to represent the isothiazolinones group and phenoxyethanol (POE) was used to represent the alcohol preservative category. Information about the preservatives used in this study and their source information can be found in Chapter 1, Table 2. Stock concentrations of preservatives were prepared in sterilised double-deionised water on day of use to 0.05% activity for BIT and 2% activity for POE. Stock concentrations were then further diluted in growth media to obtain the required test concentrations. In addition, an unpreserved laundry detergent formulation containing C12-15 Pareth-7 and Sodium Laureth Sulfate, was prepared by the industrial sponsor and used for bacterial RNA expression analysis (see Section 2.7)

2.3.2 Antibiotics

Antibiotic stock solutions were prepared for tetracycline (Tc) at 50 mg/ml on day of use in DMSO, and -20°C freezer stocks solutions of polymyxin B (Pmx) were prepared in autoclaved double-deionised water at 600,000 units/ml, filter sterilised and used within 2 months. Antibiotic stock solutions were further diluted to the required concentration in growth media.

2.4 DNA Extraction from bacterial cells

2.4.1 Rapid DNA extraction using Chelex®100 resin

The following procedure was used for rapid bacterial DNA extraction for screening techniques such as PCR. A sterile 200 µl plastic pipette tip was lightly dipped into confluent growth of a strain grown for 24 hour on agar and material transferred into 50 µl of 5% Chelex® 100 resin solution (Bio-rad, Hertfordshire, UK; autoclaved prior to use). A c1000™ Thermal Cycler (Bio-Rad) was used to run the following conditions: samples were heated to 98°C for 5 minutes, then placed at 4°C for 5 minutes, process was repeated and sample then centrifuged at 800 x g for 1 minute to deposit the Chelex® resin and cellular debris. Crude DNA from the supernatant was used in stated PCR procedures, with 2-5 µl added to each reaction. Storage of the DNA was at 4°C and DNA was used on the same day as the Chelex® preparation was performed.

2.4.2 Automated Maxwell® 16 system

The Automated Maxwell® 16 instrument nucleic acid purification system, along with the Maxwell® 16 Tissue DNA purification kit (Promega) were used to extract genomic DNA from bacterial cultures, following the manufacturer's instructions. Cells from a 3 ml overnight culture were centrifuged (1,400 x g for 10 minutes), the pellet re-suspended in 300 µl TE buffer (10 mM Tris, 1 mM EDTA, pH 8) (Sambrook et al., 1989) and added into the DNA purification kit cartridges. The instrument was cycled through its normal DNA extraction protocol and the final DNA extracts (approximately 250 µl) were stored short-term at 4 °C and at -20 °C for long-term storage.

2.4.3 Quantitation and quality assessment of Genomic DNA

Gel electrophoresis was used to analyse bacterial genomic DNA; a 1.5% high resolution (w/v) agarose gel (molecular grade agarose; Severn Biotech Ltd.), stained with SafeView (NBS Biologicals Ltd., Cambridgeshire, UK; 10 µl SafeView per 100 ml of agarose gel) was run at 80V for 8 cm (approximately 3-3.5 hours). A Qubit™ fluorometer system (Invitrogen, Massachusetts, USA) was used to quantify the genomic DNA, according to the manufacturer's instructions and nuclease free water (Severn Biotech Ltd.) was used if DNA dilutions were required. Quality assessment of DNA was carried out by measuring the absorbance ratios 260/280 nm and 260/230 nm using a NanoDrop Spectrophotometer (Thermo Fisher Scientific, Massachusetts, USA); ratios of ≥ 1.8 were used as an indication of good quality DNA.

2.5 Whole genome sequencing and comparative genomics of *P. aeruginosa* strain RW109

2.5.1 Single Molecule, Real-Time (Pacific Biosciences) sequencing of the RW109 strain

Genomic DNA from RW109 was extracted using the automated Maxwell® 16 system and quantitation and quality assessment was carried out as described in Section 2.4.3. A 1.5 ml aliquot of the overnight culture used for the DNA extraction was prepared as a freezer stock via the addition of 8% DMSO; purity of the culture was checked through the preparation of a streak plate from the freezer stock. The extracted DNA was sent as 100 µl at 236 ng/µl to the Centre for Genomic Research, University of Liverpool. Sequencing was carried out using two Single Molecule, Real-Time (SMRT) cells with P6/C4 chemistry on a Pacific Biosciences (PacBio) RSII. All subsequent bioinformatics analysis was carried out using a virtual machine, hosted by the Cloud Infrastructure for Microbial Bioinformatics (CLIMB) consortium (Connor et al., 2016).

2.5.2 Sequence assembly and annotation

The following protocol was used to create FASTQ DNA sequence files from the raw data PacBio files. From each sequencing run, three bax.h5 files and one bas.h5 file (a pointer file to the three bax.h5 files) resulted from each SMRT cell. The bax.h5 files contained the base call information from the sequencing run, and both sets from the two SMRT cells were converted into two separate binary format (bam) files using the bax2bam tool (v0.0.8, PacBio, California). The two resulting bam files were merged and a single FASTQ was extracted using the BamTools toolkit (v2.4.0, <https://github.com/pezmaster31/bamtools>). Assembly of the FASTQ sequence data was carried out as follows. The contigs were created from FASTQ files using the Canu Assembler (v1.3) (Koren et al., 2017). Assembled contigs were checked for overlapping ends and trimmed where necessary using Circulator (v1.2.1) (Hunt et al., 2015). The resulting assembly was polished using the GenomicConsensus Package (v2.1.0, PacBio) The FASTA sequence file of the RW109 genome was run through the Quality Assessment Tool for Genome Assemblies (QUAST), (v4.5.5) (Gurevich et al., 2013). Annotation of the assembly was carried out using Prokka (v1.11) (Seemann, 2014), which analyses a variety of databases to associate the genome's CDS with predicted biological functions. The annotation tool first uses BLAST+ to search against a panel of well-characterised bacterial proteins obtained from UniProtKB and databases specific to the genus of the bacteria (Seemann, 2014). A HMMER3 based search is then carried out to examine bacteria specific Prokka-HMM libraries for finalisation of the CDS functional predictions (Seemann, 2014).

2.5.3 Identifying the basic features of the RW109 whole genome sequence

Basic genome features of RW109 such as the total number of bases, numbers of coding sequences (CDS), transfer RNA (tRNA), ribosomal RNA (rRNA) and transfer-messenger (tmRNA), were obtained from the Prokka annotation summary files, and the DNA G + C content taken from the QUASt analysis. The complete genome sequence FASTA files of the clinical *P. aeruginosa* reference strains UCBPP-PA14 (PA14) (Accession: CP000438) and PAO1 (Accession: AE004091) were obtained from the *Pseudomonas* Genome Database (Winsor et al., 2016), annotated using Prokka and run through QUASt as described in Section 2.5.2. The genome features of these two strains after annotation with Prokka were used as a comparison against the RW109 strain.

2.5.4 Identifying plasmids in the assembled genome

2.5.4.1 Nucleotide BLAST comparison

To identify if the RW109 genome contained plasmids, six circular *P. aeruginosa* complete plasmid sequences were downloaded as a single multi-FASTA file from the PathoSystems Resource Integration Center (PATRIC) database (v3.3.10) (Wattam et al., 2014, Wattam et al., 2017). A BLAST database was created from the plasmid sequences using a command line standalone NCBI BLAST (BLAST v2.2.29+) (Camacho et al., 2009). The assembled un-annotated amino acid sequence of RW109 was compared to the plasmid BLAST database using a command line Nucleotide (BLASTn) search.

2.5.4.2 ACT: the Artemis Comparison Tool

The Artemis ACT comparison tool (Carver et al., 2005) was used to construct comparative synteny line plots between the assembled annotated RW109 genome and various plasmid sequences and the *P. aeruginosa* reference strain PA14. Comparison crunch files for input into ACT were generated using BLASTall (v2.2.26).

2.5.5 Assigning functional groups to the RW109 whole genome sequence

2.5.5.1 Cluster of Orthologous Group (COG) functional annotation

The COG functional annotation (Tatusov et al., 2000) of the RW109 Prokka predicted coding sequences (CDS; specifically translated amino acid sequences) was carried out with the command line EggNOG-mapper downloaded from the evolutionary genealogy of genes: Non-supervised Orthologous Groups (EggNOG) database (v4.5.1) (Huerta-Cepas et al., 2016a, Huerta-Cepas et al., 2016b). Functional orthologs were assigned using the EggNOG HMMER3 based homology search to the optimised bacterial database and the COG categories and accession numbers were extracted. COG categories were divided into three well characterised functional classes' information storage and processing, cellular processes and signalling and metabolism. A poorly characterised functional class was also used where the COG category was unknown.

2.5.5.2 Determining the interaction of RW109 CDS

The STRING database (v10.5) (Szklarczyk et al., 2016) was used to predict interactions between the RW109 CDS via their assigned COG accession numbers. A multiple list of the COG accession numbers were entered into the database, the organism *Pseudomonas* was selected and interactions were predicted (the following active interaction sources were applied: neighbourhood, gene fusion, co-occurrence, co-expression, experiments, and databases). Networks between CDS were determined at a medium confidence level score of 0.400 and were grouped via the STRING k-means clustering data-mining algorithm, which divides data into a fixed number of groups, without any prior knowledge of their relationships. In the network output, each node represents a COG accession number and the edge demonstrates the interaction confidence scores.

2.5.5.3 Kyoto Encyclopedia of Genes and Genomes (KEGG) functional module assignment and determining the metabolic and physiological potential

KEGG Orthology (KO) terms were assigned to the RW109 translated amino acid sequences from Prokka predicted CDS, using the KEGG Automatic Annotation Server (KAAS) (Moriya et al., 2007) through the NCBI BLAST single-directional best hit search method, against the prokaryotes organism list. The Metabolic and Physiological Potential Evaluator (MAPLE) tool (v2.1.0) (Takami et al., 2016), was subsequently used to map the groups of KO-assigned CDS to KEGG defined modules. These modules are a collection of functional units linked to specific metabolic abilities and phenotypic features, and identified with M numbers. KEGG modules are grouped into pathway modules, structural complexes, functional sets and signature modules (Takami et al., 2016). The percentage of a module's completeness was determined by calculating the module completion ratio (MCR), which evaluated how many KO components of the module were present. If all KO assigned CDS within each module were present, the MCR was equal to 100%, according to a Boolean algebra-like equation (Takami et al., 2016). For each MCR, a Q-value was also calculated which indicated the significance of the module completion. The MAPLE tool inferred that reaction modules with Q-values of less than 0.5 were biologically feasible even if the MCR is less than 100% (Takami et al., 2016).

2.5.5.4 Identifying putative antimicrobial resistance and virulence genes in the RW109 genome

The ABRicate tool (v0.5-dev, <https://github.com/tseemann/abricate.git>) via the command line was used to screen the Prokka annotated nucleotide sequence of RW109 to identify antimicrobial and virulence genes. The following databases downloaded with ABRicate were searched against; Resfinder (Zankari et al., 2012), Antibiotic Resistance Gene-ANNOTation (ARG-ANNOT) (Gupta et al., 2014), the Comprehensive Antibiotic Resistance Database (CARD) (McArthur et al., 2013), and the Virulence Factor Database (VFDB) (Chen et al., 2016). A $\geq 80\%$ cut off was used for both coverage and identity ABRicate scores.

2.5.5.5 Identifying antibacterial biocide and metal resistance genes in the RW109 genome

The BacMet database (v1.1) (Pal et al., 2013) of antibacterial biocide and metal resistance genes was downloaded and the Prokka annotated amino acid sequence of RW109 was compared against it using command line BLASTp. Similarity matches were determined using an Expected-value (E-value) cut-off of $1E-40$, a BIT score cut off of 400 and setting the best-hit filtering algorithm overhang parameter to 0.25. CDS identified by the BacMet database were termed BacMet resistance genes.

2.5.6 Identifying RW109 genomic islands

RW109 genomic islands (GIs) were predicted using Islandviewer (v4.0) (Bertelli et al., 2017) with the Prokka generated GeneBank file with default settings applied. The results from IslandViewer prediction methods SIGI-HMM and IslandPath-DIMOB was used to identify the total number of GIs within the RW109 whole genome sequence.

2.5.7 Identifying prophages in RW109 genome

Prophage sequences within the RW109 genome were predicted using PHAge Search Tool Enhanced Release (PHASTER) (v1.0) (Arndt et al., 2016, Zhou et al., 2011). Comparisons against the PHASTER databases and feature identifications were carried out with the Prokka generated GeneBank file. Phage sequence regions were given a PHASTER score and were identified as being complete if the score was > 90 , questionable with a score of 70-90 and incomplete with a score < 70 .

2.5.8 Visualising the RW109 whole genome sequence

The Circular Genome Viewer (CGView) application and CGView source code (Stothard and Wishart, 2005) were used via the command line to convert Prokka generated GeneBank files into CGView XML files for generation of a RW109 genome circular map. The CDS were coloured according to the assigned COG categories, and the RW109 GIs and prophage sequences were also illustrated on the map.

2.5.9 Comparative genomics of the RW109 whole genome sequence

2.5.9.1 Comparing RW109 annotated functional categories with PA14 and PAO1

The Prokka translated amino acid sequences of *P. aeruginosa* PA14 and PAO1 (see Section 2.5.3) were assigned COG functional annotations, KEGG functional modules and MAPLE MCRs, as described in Sections 2.5.5.1, 2.5.5.2 and 2.5.5.3. The antimicrobial resistance and virulence genes were also identified in PA14 and PAO1, as described in Sections 2.5.5.4 and 2.5.5.5. The results from these annotations were used to compare against the functional annotations of the industrial *P. aeruginosa* RW109 strain.

2.5.9.2 Roary: Rapid large-scale prokaryote pan genome analysis and core-gene alignment

The Roary pipeline (v3.7.0) (Page et al., 2015) was used for pan-genome comparisons of the RW109 strain with genome sequences of *P. aeruginosa* strains isolated from clinical, environmental and industrial sources (Table 3). In total 102 genomes were obtained; 70 complete genome FASTA files of *P. aeruginosa* strains were downloaded from the *Pseudomonas* Genome Database (Winsor et al., 2016) and 9 were obtained from GenBank. Also included in the analysis were 23 unpublished *P. aeruginosa* strain genome sequences, previously used as a panel by a Cardiff University PhD student (Weiser, 2015). All the FASTA sequences of the strains were annotated using Prokka, as described in Section 2.5.2. The resulting GFF files of the 102 *P. aeruginosa* genome sequences were used in the Roary analysis along with the GFF file of the RW109 strain. The `-e` parameter was applied to create a multi-FASTA codon aware alignment of all the core genes using PRANK (Löytynoja, 2014). The resulting core-gene multi-FASTA alignment file was used to build a phylogenetic maximum-likelihood tree based on 1,000 bootstrap resampling replicates with FastTree (v2.1.8), (Double precision, No-SSE3, OpenMP) (Price et al., 2010). The Generalized Time-Reversible (GTR) model with default settings was applied. The tree was visualised with FigTree (v1.4.3, <http://tree.bio.ed.ac.uk/software/figtree/>).

2.5.9.3 Comparison of antimicrobial resistance genes across 103 *P. aeruginosa* genome sequences

The FASTA sequences of the 102 *P. aeruginosa* genomes from Section 2.5.9.2 were input into the ABRicate tool to identify CARD antimicrobial resistance genes. A $\geq 80\%$ cut off was used for both coverage and identity ABRicate scores. The ABRicate summary reports from each genome sequence and the RW109 genome were combined into a single spreadsheet to show the presence and absence of antimicrobial resistance genes. This spreadsheet along with the phylogenetic tree built with FastTree in Section 2.5.9.2 were uploaded to the online Interactive Tree of Life (iTOL) (v3.5.4) (Letunic and Bork, 2016) to construct a tree annotated with a heat map illustrating the presence/absence of the identified antimicrobial resistance genes for each of the *P. aeruginosa* sequences.

2.5.9.4 Identifying the unique genes of RW109 when compared to the 102 *P. aeruginosa* genome sequences

From the Roary pan-genome analysis (Section 2.5.9.2), the gene presence and absence output was used to determine the genes specific to the *P. aeruginosa* RW109 strain which were not identified in the 102 *P. aeruginosa* genome sequences. The position of these genes in the RW109 genome and their functional annotations such COG, KEGG, antimicrobial and virulence identifications were analysed.

Table 3. *P. aeruginosa* genome sequences used in the Roary pan genome analysis

Strain Name	Strain Type	Isolation Source	Reference	GenBank Accession Number
12-4-4.59	Clinical	Blood culture of a burn patient	(Karna et al., 2016)	CP013696
19BR	Clinical	Collected as part of a Brazilian Surveillance study between 2002 and 2004	(Boyle et al., 2012)	AFXJ00000000
213BR	Clinical	Collected as part of a Brazilian surveillance study between 2002 and 2004	(Boyle et al., 2012)	AFXK00000000
8380-3922	Clinical	Human gut	(Ichise et al., 2015)	AP014839
AES-1R-2482	Clinical	Sputum of a 14 month old infant with cystic fibrosis	(Naughton et al., 2011)	CP013680
ATCC-15692	Clinical	Infected wound	Not published	CP017149
ATCC-27853	Clinical	Unknown	(Feng et al., 2016)	CP015117
B136-33	Clinical	Infant with community acquired diarrhoea	Not Published	CP004061
BAMC-07-48	Clinical	Combat injury wound	(Sanjar et al., 2016)	CP015377
Carb01-63	Clinical	Unknown	Not published	CP011317
DHS01	Clinical	Nose of a patient	(Valot et al., 2014)	CP013993
DK2	Clinical	CF sputum	(Rau et al., 2012)	CP003149
DN1	Environmental	Soil, China	Not published	CP017099
DSM-50071	Clinical	Hospital, Japan	(Nakano et al., 2015)	CP012001
F9676	Environmental	Diseased rice, China	Not published	CP012066
F22031	Clinical	Pubic bone	Not published	CP007399
FA-HZ1	Environmental	Wastewater, dibenzofuran_degrading bacterium	Not published	CP017353
FRD1	Clinical	CF sputum	Not published	CP010555
IOMTU-133	Clinical	Female participant in the dbGaP microbiome study	Not published	AP017302
F9670	Clinical	Unknown	Not published	CP008873
F23197	Clinical	Unknown	Not published	CP008856
F30658	Clinical	Unknown	Not published	CP008857
F63912	Clinical	Unknown	Not published	CP008858
H5708	Clinical	Unknown	Not published	CP008859
H27930	Clinical	Unknown	Not published	CP008860
H47921	Clinical	Unknown	Not published	CP008861
M1608	Clinical	Unknown	Not published	CP008862
M37351	Clinical	Unknown	Not published	CP008863
S86968	Clinical	Unknown	Not published	CP008865
T38079	Clinical	Unknown	Not published	CP008866
T52373	Clinical	Unknown	Not published	CP008867
T63266	Clinical	Unknown	Not published	CP008868
W16407	Clinical	Unknown	Not published	CP008869
W36662	Clinical	Unknown	Not published	CP008870
W45909	Clinical	Unknown	Not published	CP008871
W60856	Clinical	Unknown	Not published	CP008864
X78812	Clinical	Unknown	Not published	CP008872

Table 3. Continued

LES431	Clinical	CF sputum, Liverpool epidemic strain	(Jeukens et al., 2014)	CP006937
LESB58	Clinical	CF sputum, Liverpool epidemic strain	(Winstanley et al., 2009)	FM209186
M18	Environmental	Plant isolate, China	(Wu et al., 2011)	CP002496
MTB-1	Environmental	Hexachlorocyclohexane contaminated soil	(Ohtsubo et al., 2014)	CP006853
N17-1	Environmental	Soil	Not published	CP014948
NCGM2.S1	Clinical	Isolated from a hospital in Japan	(Miyoshi-Akiyama et al., 2011)	AP012280
NCGM257	Clinical	Urine, Japan	Not published	AP014651
NCGM1900	Clinical	Urinary catheter	Not published	AP014622
NCGM1984	Clinical	Urinary catheter	(Tada et al., 2016)	AP014646
NCTC10332	Clinical	Unknown	Not published	LN831024
D1	Clinical	Clinical isolate from ventilator associated pneumonia patient	Not published	CP012585
D2	Clinical	Clinical isolate from ventilator associated pneumonia patient	Not published	CP012578
D5	Clinical	Clinical isolate from ventilator associated pneumonia patient	Not published	CP012579
D9	Clinical	Clinical isolate from ventilator associated pneumonia patient	Not published	CP012580
D16	Clinical	Clinical isolate from ventilator associated pneumonia patient	Not published	CP012581
D21	Clinical	Clinical isolate from ventilator associated pneumonia patient	Not published	CP012582
D22	Clinical	Clinical isolate from ventilator associated pneumonia patient	Not published	CP012583
D25	Clinical	Clinical isolate from ventilator associated pneumonia patient	Not published	CP012584
PA1	Clinical	Respiratory tract infection	(Lu et al., 2015)	CP004054
PA1R	Clinical	Respiratory tract infection	(Lu et al., 2015)	CP004055
PA1RG	Clinical/environmental	Hospital sewage	(Li et al., 2016)	CP012679
PA7	Clinical	Non-respiratory clinical isolate	(Roy et al., 2010)	CP000744
PA121617	Clinical	CF Sputum	Not published	CP016214
PACS2	Clinical	CF Sputum	Not published	AAQW01000001
PAO1	Clinical	Wound isolate	(Stover et al., 2000)	AE004091
RP73	Clinical	CF Sputum	(Jeukens et al., 2013)	CP006245
SCV20265	Clinical	CF Lung	(Eckweiler et al., 2014)	CP006931
UCBPP-PA14	Clinical	Burn patient	(Lee et al., 2006a)	CP000438
USDA-ARS-USMARC-41639	Environmental	Nasopharynx of a cow in Kansas, USA	Not published	CP013989
VA-134	Clinical	Burn wound	(Miller et al., 2016)	CP013245
VRFPA04	Clinical	Corneal button from patient with corneal keratitis	(Murugan et al., 2016)	CP008739
YL84	Environmental	Compost	(Chan et al., 2014)	CP007147
PA96	Clinical	Clinical isolate from Guangzhou, China	(Déraspe et al., 2014)	CP007224
S04-90	Environmental	Microbial mat material	Not published	NZ_CP011369
N002	Environmental	Crude oil contaminated soil	(Roy et al., 2013)	ALBV00000000

Table 3. Continued

SJTD-1	Environmental	Soil	(Liu et al., 2012)	CP015877
ATCC-700888	Environmental	Industrial water system	(Chugani et al., 2012)	AKZF00000000
E2oS	Environmental	Soil	(Stewart et al., 2014)	ASQV00000000
MSH-3	Environmental	Environmental, Mount St. Helens	(Stewart et al., 2014)	ASQU00000000
ATCC-14886	Environmental	Soil	(Chugani et al., 2012)	AKZD00000000
MSH-10	Environmental	Environmental	(Stewart et al., 2014)	ASWW00000000
XMG	Environmental	Soil, China	(Gao et al., 2012)	AJXX00000000
NCTC-12903*	Clinical	Antibiotic efficacy testing reference strain; originally isolated from blood	Not published	Not available
RW18*	Clinical	Chronic prostatitis isolate	Not published	Not available
RW27*	Clinical	CF sputum	(Mahenthiralingam et al., 1996)	Not available
RW30*	Clinical	CF sputum	(Mahenthiralingam et al., 1996)	Not available
RW99*	Environmental	Domestic isolate, washing machine drawer biofilm	Not published	Not available
RW110*	Industrial	Strain used in Unilever preservative efficacy testing	Not published	Not available
RW130*	Industrial	Dish washer liquid; Italy; isolated 2010	Not published	Not available
RW131*	Industrial	Dish washer liquid; Italy; isolated 2010	Not published	Not available
RW138*	Industrial	Liquid abrasive cleaner; Italy; isolated 2001	Not published	Not available
RW146*	Industrial	Dish washer liquid; Italy; isolated 2004	Not published	Not available
RW149*	Industrial	Personal care product; origin location unknown; isolated 2003	Not published	Not available
ATCC-9027*	Clinical	Reference strain used in industrial testing; isolated from an outer ear infection	Not published	Not available
RW168.2*	Industrial	Unknown	Not published	Not available
RW172*	Industrial	Dish washer liquid; suspected origin is Thailand; isolated 2009	Not published	Not available
RW176*	Industrial	Dish washer liquid; Indonesia; isolated 2010	Not published	Not available
RW184*	Industrial	Dish washer liquid; UK; isolated 2006	Not published	Not available
RW192*	Industrial	Surface cleaner; isolated from a contaminated Unilever product in 2012	Not published	Not available
ATCC-13388*	Unknown	Origin unknown; reference strain used in industrial testing (ISO 846C)	Not published	Not available
ATCC-15442*	Environmental	Reference strain used in industrial testing; isolated from an animal room water bottle	Not published	Not available
RW199*	Industrial	Metal working fluid product	Not published	Not available
RW200*	Industrial	Timber care product	Not published	Not available
RW202*	Industrial	Liquid abrasive cleaner; EU; isolated 2012	Not published	Not available
RW204*	Industrial	Dish washer liquid; isolated 2012	Not published	Not available

Footnotes = * Panel of 23 isolates used by a previous Cardiff University PhD Student (Weiser, 2015).

2.6 Transposon mutagenesis experiments

2.6.1 Construction of a mini-Tn5-*luxCDABE* transposon library in RW109 *P. aeruginosa* strain

Random mini-Tn5 transposon mutagenesis was initiated by performing a bi-parental conjugal mating with the recipient *P. aeruginosa* RW109 strain and an *E. coli* S17-1 λ pir donor strain carrying pUTmini-Tn5-*luxCDABE*-Tc (Winson et al., 1998). The *E. coli* was revived on LB agar containing 10 μ g/ml Tc at 37°C. Overnight 3 ml liquid LB and 10 μ g/ml Tc cultures of the *E. coli* strain were prepared by inoculating with confluent growth from a 24-hour plate culture and incubated for 18-hours with aeration at 37°C.

2.6.1.1 Optimising transconjugant selective media: Minimum Inhibition Concentration (MIC) testing of antibiotics using broth and agar assays

A micro-dilution broth assay was used to determine the MIC of the donor and recipient strains in the presence of the antibiotic Tc. This form of assay allowed a wide range of Tc concentrations to be assessed before carrying out an agar dilution assay. Twelve individual Tc test concentrations at 0, 15, 20, 25, 30, 35, 40, 45, 50, 100, 150 and 200 μ g/ml were assessed and prepared at double the concentration with 100 μ l added to wells of a clear flat bottom 96 well plate (Fisher Scientific). The concentrations were halved to the required concentration when inoculated with 100 μ l of an overnight culture at approximately 10^6 cfu/ml. The 96 well plates were incubated shaking (150 rpm) at 30°C, and after 24 hours the optical density (OD) at 600 nm of each well was recorded using the absorbance function of a Tecan Infinite® M200 PRO plate reader (Labtech International LTD, UK). The MIC was taken as the lowest concentration of the Tc antibiotic, at which there was an 80% or more reduction in OD when compared to growth in control wells containing TSB-only. Experiments were performed in triplicate using different starting cultures and Tc stock solutions to obtain biological replicates, with each repeat having four technical replicates.

An agar dilution assay was used to investigate the susceptibility of donor and recipient strains to Tc and Pmx antibiotics. TSA agar was cooled to 50°C after autoclaving, and an appropriate volume of antibiotic stock added to the 25 ml of agar to achieve final concentration ranges of 0, 100, 125 and 150 μ g/ml for Tc, and 0, 100 and 120 units/ml for Pmx. Concentrations were assessed separately and in combination for Tc and Pmx. Strains were cultured as described previously and 50 μ l of approximately 10^6 cfu/ml of culture was transferred to the surface of triplicate agar plates and spread evenly. Visual assessments and colony counts were performed following 24 – 48 hours of incubation at 30°C.

2.6.1.2 Conjugal transfer of mini-Tn5

Overnight cultures of recipient and donor cells were harvested by centrifugation (1,400 x g for 10 minutes) and washed with LB + 10 mM MgSO₄. As a control for background growth, 50 μ l of approximately 10^6 cfu/ml of both recipient and donor cultures were plated onto the transconjugant selective media. Recipient cells were heat shocked for 10 minutes at 42°C prior to being mixed with the donor at a ratio of 1:1. 100 μ l of the mating mixture was transferred onto a 25 mm sterile nitrocellulose

filter membrane (0.22 µm pore size, Sigma-Aldrich) which was placed centrally on the surface of a warm dry LB+10 mM MgSO₄ plate. After 18-hour incubation at 37°C, the mating mixture was recovered from the membrane and re-suspended in TSB.

To selectively kill the *E. coli* donor and the RW109 recipient but select for transconjugants carrying the mini-Tn5-*luxCDABE* transposon, 50 µl of re-suspended mating mixture at 10⁻¹ dilution was transferred to selective TSA plates. All plates were incubated at 30°C for approximately 48 hours. Using sterile wooden tooth picks, transconjugants were individually picked into flat-bottom white 96-well micro-titre plates (Greiner, BioOne) containing 200 µl TSB and 50 µg/ml Tc, with wells H:6 and H:12 containing TSB-only, acting as blank controls.

Plates were cultured in the dark for 24 hours at 30°C and mutant relative light units (RLU), were measured in triplicate using the luminescence function of a Tecan plate reader with a 4-second integration time. The mean RLU of the blank wells was subtracted from each mutant well to normalise measurements. Mutants were individually identified based on their well position within each master plate and were categorised according to the baseline bioluminescence measurement as follows: ≥ 1.00E+06 RLU (Category 1), ≥ 1.00E+05 RLU (Category 2), ≥ 1.00E+04 RLU (Category 3), ≥ 1.00E+03 RLU (Category 4), ≥ 5.00E+02 RLU (Category 5) and ≤ 4.99E+02 RLU (Category 6). After reading the baseline light emissions, 8% DMSO was added to each well for long-term storage as master plates at -80°C.

2.6.2 Identifying the mini-Tn5-*luxCDABE* transposon in low light emitters

Twenty mutants in light emission Category 6, the lowest emission category as assigned above, were checked to determine if they contained the mini-Tn5-*luxCDABE* transposon using a PCR that amplifies a region of the *luxA* gene. Twenty category 6 mutants from across the constructed mutant bank were isolated and grown on TSA agar. DNA was extracted using Chelex®100 resin and used as template for a PCR reaction with a *luxA*_forward primer (5'-CGGTTTGGTATTTGCCGAGG-3') and *luxA*_reverse primer (5'GTAGGCCAAATTGAGCA GCC-3'). The reaction took place in a c1000™ Thermal Cycler according to the following cycling conditions: 95°C for 2 minutes, 35 cycles of 94°C for 30 seconds, 60°C for 30 seconds and 72°C for 1 minute and 30 seconds and a final extension step of 72°C for 5 minutes. Each PCR reaction mixture was prepared with an overall volume of 25µl, with reagents obtained from Qiagen, UK. The mixtures contained sterile H₂O, 1X PCR buffer, 1X Q-solution, 200 µM of each dNTP, 0.4 pmol/µl of each primer, 1U of Taq DNA polymerase and 2 µl of the template DNA (approximately 20 ng). The PCR reaction products were run along with a 2-Log DNA ladder (New England Biolabs, NEB UK; 0.1-10.0 kb) on 1.5% high resolution (w/v) agarose-TAE gels stained with Safeview and run at 80V for 8 cm (approximately 3-3.5hr). The wild-type RW109 strain was used a negative control and a high light emitter along with the *E.coli* donor strain were used as positive controls; the expected size of the PCR product was 262 bp to confirm the presence of the *luxA* gene.

2.6.3 Random Amplified Polymorphic DNA (RAPD) typing of PA RW109 *lux* mutants

Eighteen mutants of varying light emissions were randomly selected from across the library and analysed using a RAPD PCR as described by (Mahenthiralingam et al., 1996). This was carried out to ensure the picked transconjugants had the same genotype as the recipient RW109 strain. The majority of mutants selected were from the lowest light emission category; the PCR also included the recipient, donor and a negative control. Mutants were isolated, grown on TSA agar and DNA was extracted using Chelex®100 resin and used as template for a PCR reaction with a RAPD 272 primer (5'-AGC GGG CCA A-3'). Using a c1000™ Thermal Cycler the following cycling conditions were run: 94°C for 5 minutes, 4 cycles of 36°C for 5 minutes, 72°C for 5 minutes, 30 cycles of 94°C for 1 minute, 36°C for 1 minute 72°C for 2 minutes followed by a final extension step of 72°C for 6 minutes. PCR reaction mixtures for each mutant sample were prepared with an overall volume of 25µl, with reagents supplied by Qiagen. The mixtures contained Sterile H₂O, 10x Coraload buffer, 1 x Q-solution, 25 mM MgCl₂, 200 µM (dNTPs), 1.6 µM RAPD 272 primer, 1 U Taq DNA polymerase and 2µl of the template DNA (approximately 20 ng). 12 µl of the PCR products and 10 µl of a 2-Log DNA ladder (0.1-10.0 kb) were loaded into a 1.5% high resolution (w/v) agarose gel stained with Safeview with and run at 80V for 8 cm.

2.6.4 MIC testing of industrial preservatives using a broth doubling dilution method

To obtain approximate MIC values for the preservatives BIT and POE, serial doubling dilutions of aqueous stock solutions were prepared in TSB to achieve final concentrations ranging from 0 – 0.075% for BIT and 0 – 2% for POE. Preservatives were prepared at double the concentrations and 96 well plates were set up for MIC testing and carried out as described in Section 2.6.1.1. Preservative MIC values were taken as the lowest concentration at which there was an 80% or more reduction in OD when compared to the growth control wells containing TSB-only. Experiments were performed in triplicate using different starting cultures and preservative stock solutions to obtain biological replicates with each repeat having four technical replicates.

2.6.5 Screening the transposon library to identify preservative responding mutants

2.6.5.1 Master plate screening

To identify mutants with altered light emission responses when exposed to industrial preservatives, master plates were replicated and screened with the preservative at sub-MIC concentrations; the MIC was an approximate calculation for the wild type RW109 strain. A 96-pin replicator was used to replicate master plates into 2.0 ml deep-well plates (Nunc, Thermo Scientific) containing 1.5 ml of TSB which were incubated shaking (150 rpm) for 24 hours at 30°C. The replicated master plate was then sub-cultured by transferring 100 µl of the left side (columns 1-6) to both the left and right side of triplicate white 96-well micro-titre plates (Figure 7). This was also repeated for the right side of the master plate (columns 7-12) (Figure 7). Mutants on the left side of sub-cultured plates were exposed to 100 µl of TSB-only (control condition) and the right side was exposed to 100 µl of screening preservative (test condition) (Figure 7). Plates were incubated at 30°C in the dark and after 24-hours following the addition of TSB and preservative, the RLU was measured using the luminescence setting of a Tecan plate reader. Triplicate RLU readings were normalised to the mean RLU of the blank wells and individual mutants of interest were isolated from the replicated master plate revived on TSA and freezer stocks prepared. Mutants were selected for isolation depending on the RLU fold change between the control and test conditions after 24 hours; those which demonstrated a ≥ 0.4 fold RLU increase (positive responding mutants) and a ≥ 0.3 fold RLU decrease (negative responding mutants) in the master plate screen were isolated. The fold RLU selection criteria used may not represent true preservative responding mutants, but it allowed for the isolation of a wide range of transconjugants that demonstrated various responses to BIT exposure. Mutants which displayed minimal light emission change when exposed to BIT (RLU fold change between +0.05 to -0.05; minimal change mutants), were also isolated from master plates.

2.6.5.2 Isolated mutant screening

Isolated mutants of interest were revived onto TSA from individual freezer stocks and overnight cultures diluted to approximately 10^6 cfu/ml. 100 μ l of the diluted overnight culture was transferred in a staggered format to four wells in both columns 1 and 7 of a white 96 well plate; this was carried for six mutants per plate (Figure 7). The wells in between the mutants contained 100 μ l of TSB-only for normalisation of RLU data. To columns 1-6, 100 μ l of TSB was added to all wells (control condition) and 100 μ l of the screening preservative was added to all wells of columns 7-12 (test condition) (Figure 7). After incubation in the dark at 30°C for 24-hours, luminescence was measured. The replicate RLU readings were normalised to the mean RLU of the blank wells.

2.6.5.3 Statistical analysis of RLU measurements

Using R statistical software (v3.3.3), mutant RLU data was analysed by comparing the two means for the control and test conditions using a paired two-sample T-test for means; RLU differences with a p-value of ≤ 0.05 were considered to be significant. Boxplots were generated using BoxPlotR (Spitzer et al., 2014) to examine the difference in RLU measurements between the control and test conditions and observe the distribution of quantified light emission data.

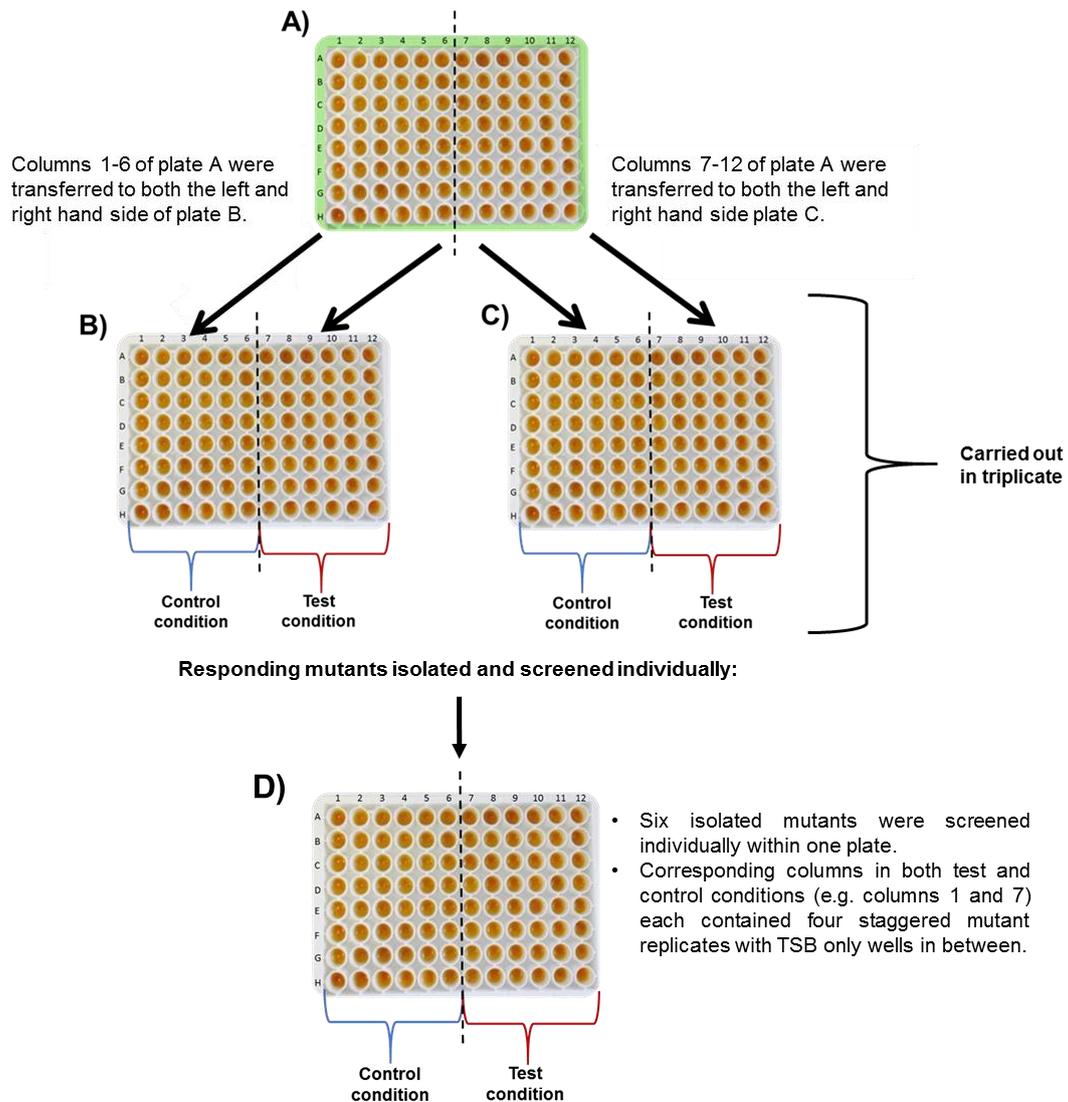


Figure 7. The organisation of luminescent plates when screening master plates and individual mutants

A master plate was replicated into a deep-well plate (A) and sub-cultured into two plates, one containing columns 1-6 in both the left and right hand side (B) and the other consisting of columns 7-12 in both the left and right hand side (C). Screening of a master plate was carried out in triplicate resulting in three sets of plates B and C. Mutants of interest were isolated from plate A and six individual mutants were screened with both control and test conditions within one plate (D).

2.6.6 Sequencing of transposon flanking DNA and bioinformatics analysis

An inverse PCR technique was used to identify the transposon insertion sites in mutants of interest as described by (Lewenza et al., 2005). Genomic DNA was extracted using the automated Maxwell® 16 system. DNA flanking the left side of the transposon was digested and ligated using a *SacII* restriction enzyme and T4 DNA ligase (NEB, UK) following the manufacturer's instructions. The ligated product was used as template for an inverse PCR reaction using Tn5-SacII primer 1 (5'-GTCAAAGGACGATTTTCGGTTTGG-3') and Tn5-SacII primer 2 (5'-GATCCCCGGGTACCGAGCTCGAATTC-3') in a c1000™ Thermal Cycler. The following cycling conditions were used: 95°C for 3 minutes, 5 cycles of 95°C for 30 seconds, 60°C for 1 minute and 72°C for 3 minutes, followed by 30 cycles of 95°C for 30 seconds, 60°C for 1 minute and 72°C for 3 minutes and finishing with a single step of 72°C for 5 minutes.

Each PCR reaction mixture was prepared with an overall volume of 25 µl, with reagents obtained from Qiagen, UK. The mixtures contained sterile H₂O, 1X PCR buffer, 1X Q-solution, 0.5 mM MgCl₂, 200 µM of each dNTP, 0.4 pmol/µl of each primer 1U of *Taq* DNA polymerase and 2.5 µl of the ligated DNA as a template. The PCR reaction products were run along with a 2-Log DNA ladder on 1.5% high resolution (w/v) agarose-TAE gels stained with Safeview and run at 80V for 8 cm. All samples yielding a single major band were purified with the QIAquick PCR purification kit (Qiagen, UK), according to the manufacturer's instructions and quantified using a Qubit. The purified PCR products were forwarded to the MWG Eurofins DNA Sanger sequencing service following the sample submission guidelines for 'MWG Eurofins Value Read Service in Tubes'. A sequencing primer Tn5-SacII primer 3 (5'-CCGGGTACCGAGCTCGAATTCG-3') was sent along with the PCR products to identify the DNA flanking the left side of the transposon. Insertion sites were identified using a command line standalone NCBI Nucleotide BLAST (BLASTn, v2.2.29+) (Camacho et al., 2009) with a BLAST database of the RW109 PacBio genome sequence and an E-Value cut-off of 1E-6 was used. The COG functional associations of the transposon insertion sites were identified from the annotation of RW109 in Section 2.5.5.1. The distribution of the mutant transposon insertion sites within the RW109 genome were mapped using CGView, as described in Section 2.5.8.

2.7 RNA-Sequencing (RNA-Seq) experiments

All RNA-Sequencing (RNA-Seq) experiments were carried out using the same *P. aeruginosa* RW109 freezer stock prepared in Section 2.5.1 which was used for PacBio sequencing.

2.7.1 Growth curve analysis of RW109 when exposed to individual preservatives

To determine the growth dynamics of RW109 with and without preservatives, a Bioscreen C Microbiological Growth Analyser (Labsystems, Finland) was used. Within a Bioscreen C microplate, four blank control wells contained 200 µl TSB-only and test wells contained 200 µl TSB supplemented with preservatives and inoculated with approximately 10⁶ cfu/ml of an overnight culture. The preservatives BIT and POE were analysed with a range of 0 – 115% of the MIC value (calculated in Section 2.6.4), going up in 5% increments. Growth was monitored for 48 hours at 30°C with turbidity measurements taken at 15-minute intervals using a wide band filter (450-580 nm), after shaking the microplates for 10 seconds at an intermediate intensity. Each experiment contained four technical replicates for the control and test conditions and was performed in triplicate with different starting overnight cultures and test condition preparations. The data was pooled and analysed in Microsoft Excel with the mean value of the control blank wells subtracted from all the test OD data, which was averaged and logarithmically transformed and used to produce line graphs. After this initial assessment, another Bioscreen microplate was set up where RW109 was exposed to BIT and POE at only 20 and 50% of the MIC for 24 hours for further analysis of the growth dynamics at these concentrations.

2.7.2 Growth curve analysis of RW109 when exposed to BIT and POE in combination

The growth dynamics of RW109 when exposed to a combination of the preservatives BIT and POE were also assessed. The Bioscreen microplate was set up and run as described in Section 2.7.1 and RW109 was grown in the presence of the preservatives both in combination at 5, 10, 15 and 20% of the calculated MIC values.

2.7.3 Growth curve analysis of RW109 when exposed to a product formulation with and without the presence of a preservative

The growth dynamics of RW109 when exposed to a product formulation with and without the presence of a preservative was also assessed. A liquid laundry detergent product was supplied by Unilever, which did not have a preservation system (Section 2.3.1); the laundry detergent was colourless and soluble in H₂O. The Bioscreen microplate was set up and run as described in Section 2.7.1 and RW109 was grown in the presence of the laundry detergent at the following dilutions: 1:10, 1:20, 1:50 and 1:100. The dilutions of the laundry detergent were prepared in TSB. After the initial analysis of RW109 grown with only the laundry detergent, a Bioscreen experiment was also set up with exposure to the laundry detergent and the preservative BIT in combination.

2.7.4 RNA-Seq exposure experiments

The RW109 strain was exposed to eight conditions, which included:

Control condition: TSB-only,

- 1) BIT at 20% of the MIC (0.01875% active),
- 2) BIT at 50% of the MIC (0.00469%),
- 3) POE at 20% of the MIC (0.1% active)
- 4) POE at 50% of the MIC (0.25% active),
- 5) BIT and POE in combination, both at 20% of the MIC (see above for active concentrations),
- 6) 1:100 dilution of unpreserved laundry detergent composed of C12-15 Pareth-7 and Sodium Laureth Sulfate (Section 2.3.1), and
- 7) 1:100 dilution of the laundry detergent in combination with BIT at 20% of the MIC.

These chosen concentrations were lower than product recommended usage levels of 0.05-0.15% active for BIT and 0.5-1.0% active for POE.

The exposure experiments were carried out using the Bioscreen C instrument and the microplate was set up as described in Section 2.7.1 with the control condition and seven test conditions each represented by four technical replicates within the plate (Figure 8). RW109 was exposed to the conditions for 24 hours at 30°C and turbidity measurements were taken every 15 minutes as previously described. After exposure for 24 hours the cultures in each of the four technical replicate wells for the control and exposure conditions were promptly pooled into separate micro-centrifuge tubes, and immediately snap-cooled in an ethanol and dry ice bath, before centrifuging at 20,000 x g at 4°C for 1 minute.

The supernatant was removed and pellets were immediately snap-frozen and stored at -80°C until required. Exposure experiments were repeated with different starting overnight cultures, and exposure condition preparations to obtain four biological replicates for the control and exposure conditions. The mean Log₁₀ OD value (450-580 nm) of the four biological replicates at 24 hours for each RNA-Seq exposure condition were analysed by comparing the two means for the control and test exposure conditions using a paired two-sample t-test for means; Log₁₀ OD value differences with a p-value of ≤ 0.05 were considered to be significant.

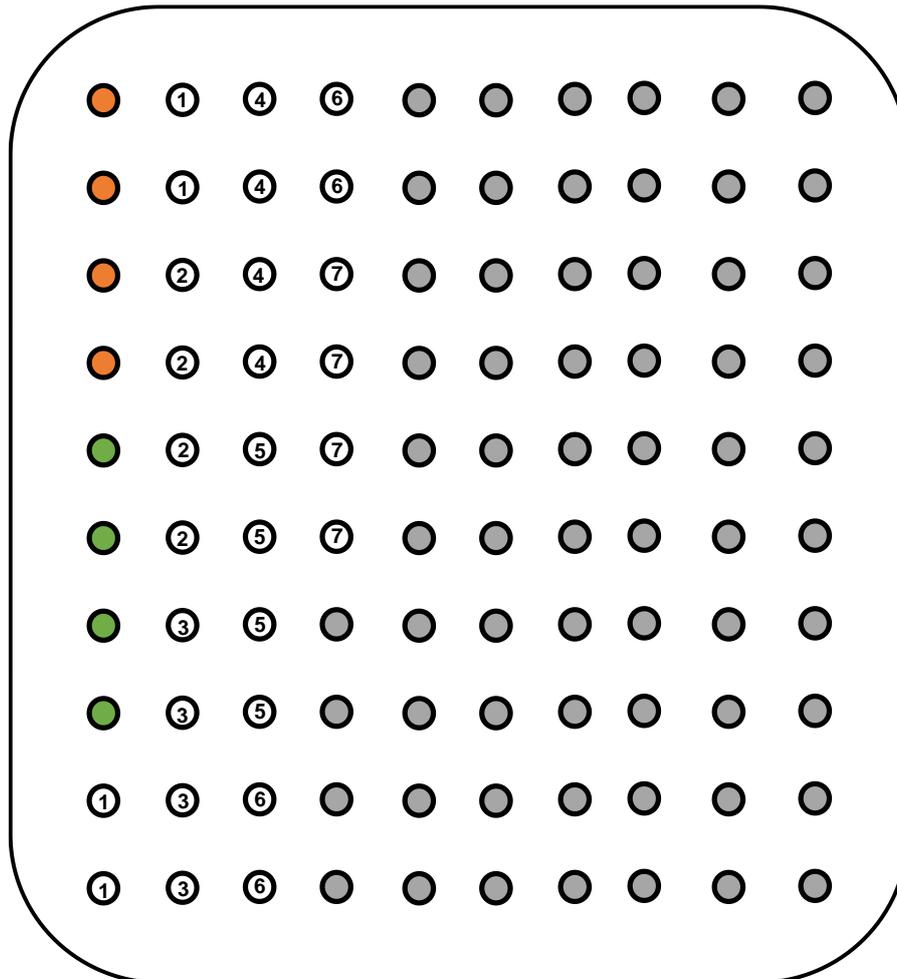


Figure 8. The arrangement of a Bioscreen microplate during RNA-Seq exposure experiments

A Bioscreen microplate was arranged as follows for each exposure experiment; orange wells were the blanks and contained 200 μ l of TSB and were not inoculated with RW109, wells coloured in green indicate the control condition and contained 200 μ l of TSB inoculated with approximately 10^6 cfu/ml from an overnight culture. The white wells numbered 1-7 represent the test exposure conditions which were as follows: 1) BIT at 20% of the MIC, 2) BIT at 50% of the MIC, 3) POE at 20% of the MIC, 4) POE at 50% of the MIC, 5) BIT and POE in combination, both at 20% of the MIC, 6) 1:100 dilution of unpreserved laundry detergent, and 7) 1:100 dilution of laundry detergent in combination with BIT at 20% of the MIC. Grey wells were empty and not used during the exposure experiments.

2.7.5 Enumeration of viable bacteria from exposure experiments

After exposure to the RNA-Seq conditions, the enumeration of viable cells was carried out as described in Section 2.2.3. For each serial dilution of the exposure conditions, 100 µl of was spread onto 3 replicate TSA agar plates and incubated for 24 hours at 30°C. This was repeated with different starting overnight cultures and exposure condition preparations, to obtain three biological replicates for the control and exposure conditions.

2.7.6 Total RNA (toRNA) extraction

Total RNA was extracted within one week of harvesting cells using the RiboPure™ RNA Purification Bacteria Kit (Ambion, Life Technologies Ltd, UK) according to the manufacturer's instructions. The kit uses zirconia beads for cell wall disruption and organic extraction of the lysate, and RNA purification with silica spin columns. After the extraction and wash steps, the RNA bound to a silica filter was eluted with 2 x 30 µl of elution buffer. Each sample was treated to remove any contaminating genomic DNA using the Ambion DNA-free™ reagents following the manufacturer's instructions.

2.7.7 Total RNA (toRNA) quantification and quality assessment

A Qubit™ fluorometer system with the broad range RNA kit (Invitrogen, USA) was used to quantify the total RNA (toRNA) samples according to the manufacturer's instructions. An additional RNA precipitation step was incorporated for samples with low RNA concentrations (< 100 ng/µl) via addition of 0.1 volumes of 3 M sodium acetate, 5 µg of glycogen and 2.5 volumes of 100% ethanol. Mixtures were incubated for 30 minutes at -80°C, centrifuged at $\geq 12,000 \times g$ for 30 minutes at 4°C, and RNA pellets washed twice with 70% ethanol and re-suspended in 17 µl of the elution buffer from the RiboPure™ RNA Purification Bacteria Kit and re-quantified. The quality of toRNA was assessed with a Bioanalyzer using the RNA 6000 Nano kit (Agilent Technologies Ltd, UK), following the manufactures protocol. If the RNA Integrity Number (RIN) was ≥ 8 and the ratio of 23S:16S rRNAs was ≥ 1.5 , the toRNA samples were used for further applications. If the toRNA was found to be of low quality the exposure, experiment and toRNA extraction was repeated.

2.7.8 Messenger RNA (mRNA) enrichment, quantification and quality assessment

A MICROBExpress™ bacterial mRNA enrichment kit (Ambion), was used to enrich for messenger RNA (mRNA) from the total RNA; samples were added within a concentration range of 1.5 – 3 µg in a maximum volume of 15 µl. The kit works by mixing the total RNA with an optimized set of capture oligonucleotides, which bind to the bacterial 16S and 23S ribosomal RNA (rRNA). Derivatized magnetic microbeads remove the rRNA, which have hybridised to the oligonucleotide hybrids, but leave the mRNA in the supernatant, which is retrieved by ethanol precipitation. The resulting enriched mRNA pellets were re-suspended in 10 µl of nuclease free water and quantified using the Qubit™ fluorometer system, with the broad range RNA kit to ensure the concentration was greater than 2 ng/µl for input into the complementary DNA (cDNA) library preparation kit. The Bioanalyzer with the RNA 6000 Nano kit was used to evaluate the depletion of rRNA by comparing RNA samples before and after MICROBExpress mRNA enrichment. Noticeable reduction in the 16S and 23S rRNA peaks were observed on the bioanalyzer electropherograms; samples were allowed to continue forward to RNA sequencing even if traces of the ribosomal RNA were still present. RIN values and ratios of 23S:16S rRNAs cannot be calculated when quality assessing mRNA samples with the Bioanalyzer.

2.7.9 cDNA library preparations and sequencing

Complementary DNA (cDNA) library preparations and sequencing were carried out in collaboration with the Genomics Research Hub at Cardiff School of Biosciences:

<https://www.cardiff.ac.uk/biosciences/research/technology-research-hubs/genomics-research>.

The Illumina® TruSeq® Stranded mRNA Sample Preparation Kit was used to prepare sequencing libraries with the low sample protocol. The manufacturer's instructions describing the use of the protocol with previously isolated mRNA was followed; 5 µl of enriched mRNA within the Illumina recommended concentration range of 10 – 400 ng was added to 13 µl of the fragment, prime, finish mix to prepare sequencing libraries for each sample. Sample library concentrations were checked with the Qubit™ fluorometer system using the broad range DNA kit (Invitrogen) and were between the ranges of 20.0 – 83.3 ng/µl. The fragment size of each library was evaluated using the Agilent Tape-System 2200, with the standard D1000 Screen-Tape and sizes varied between 290-320 bp. The libraries were then normalised to approximately 2 nM and pooled in equimolar concentrations with 16 samples pooled for the sequencing runs; in total 2 sequencing runs were carried out for the 32 samples. Pooled libraries were checked with the Qubit™ fluorometer, using the broad range DNA kit and the Agilent Tape-System 2200, with the High Sensitivity D100 Screen-Tape. The overall concentration of the first pooled library was 0.306 ng/µl at 297 bp, equating to 1.58 nM. For the second library, the overall concentration was 0.317 ng/µl at 276 bp, equating to 1.34 nM. The Illumina NextSeq-500 was used to carry out the sequencing. The pooled cDNA libraries were loaded into a mid-output cartridge to produce paired-end reads of 75 bp in length, following the manufacturer's instructions.

2.7.10 RNA-Seq bioinformatics analysis

The relevant files and scripts used throughout the bioinformatics analysis of the RNA-Sequencing data are provided on within the relevant Chapter folder on the CD-ROM, which accompanies this thesis (folder: RNA-Seq_BioinformaticAnalysis_Scripts). All subsequent bioinformatics analysis was carried out using a virtual machine, hosted by the CLIMB consortium (Connor et al., 2016).

2.7.10.1 Quality control and adaptor trimming

Quality control and adaptor trimming of the sequencing sample raw reads was carried out using Trim Galore for paired-end reads. Default settings were used to remove short low quality reads (< 20 bp), predict and trim Illumina adaptor sequences, and eliminate poor-quality bases from the sequences (< Q20). FastQC confirmed that the resulting trimmed reads were of sufficient quality for subsequent analysis.

2.7.10.2 Mapping RNA-Seq reads to the RW109 genome

Alignment of the RNA-Seq reads to the reference RW109 PacBio genome (Section 2.5) was carried out via a Burrows-Wheeler Aligner transformation (BWA) with the BWA-MEM algorithm (v0.7.13-r1126) (Li, 2013). The algorithm works by seeding alignments with maximal exact matches (MEMs) and then extending seeds with the affine-gap Smith-Waterman algorithm (SW), to map low-divergent sequences against a large reference genome. The programme is highly recommended for high-quality alignments for RNA-Seq experiments due to its accuracy and speed. An index database of the reference genome of RW109 nucleotide FASTA file was generated using the BWA index command, and the BWA-MEM command was used to run with default settings for the paired-end trimmed RNA-Seq reads. The output was a SAM file of aligned sequence reads which were sorted into BAM files using the Sequence Alignment/Map (SAM) Tools toolkit (v1.3) (Li et al., 2009).

2.7.10.3 Counting the number of RNA-Seq reads which map to the RW109 genome

The Python programme HTSeq-count (v0.6.0) (Anders et al., 2015) was used to count the number of sorted aligned RNA-Seq reads which map to the gene features of the RW109 reference genome. The input was the sorted alignment BAM files and the Prokka generated RW109 GFF file (parameters set: --stranded = reverse, --order = pos, --feature type = gene, --id attribute = locus_tag and --mode = interSection-strict). The output was a count matrix file of RNA-Seq reads, which mapped to the RW109 gene features for each biological replicate for the test and control conditions (4 files for control condition and 4 for each test condition = 32 count files in total).

2.7.10.4 Differential gene expression analysis

The Bioconductor programme DESeq2 (v1.14.1) (Love et al., 2014) was used in R to normalise the count data and determine differential gene expression between the control and test samples. The HTSeq-count matrix files were input into DESeq2, using the `DESeqDataSetFromHTSeqCount` command and differential gene expression calculated using the `DESeq` command. An example of a DESeq2 script used for differential gene expression analysis between the control and exposure to POE at 20% of the MIC can be found on CD-ROM (`Example_DESeq2_POE20RScript.r`).

DESeq2 has a normalization method built in which uses negative binomial distribution and estimates size factors dispersions, to control for differences in the library size of the sequencing experiments. The programme then estimates the dispersion for each gene and fits a generalized linear model. Fold change is reported on a logarithmic scale to base 2, and indicates how much the expression of a gene changed in the test condition when compared to the control. The p-values were adjusted via the `p.adjust` function with the Benjamini and Hochberg (BH) method. Differentially expressed genes (DEGs) were defined as exhibiting a log₂-fold change of ≥ 1.5 and an adjusted p-value of ≤ 0.05 , when each test condition was compared against the control condition.

2.7.11 Generating heat maps of expression data

The `heatmap.2` function from the R package `gplots` (v2.3.2) was used to build heat maps. A distance matrix was generated with log₂-fold changes using the hierarchical cluster analysis ‘`hclust`’ ‘`dist`’ functions to produce a dendrogram tree. The hierarchical clustered distance matrix was used as an input for the `heatmap.2` function. To obtain a preliminary overview of the gene expression data, all log₂ fold changes generated by DESeq2 for the test conditions when compared to the control were input into `heatmap.2` regardless of the adjusted p-value.

2.7.12 Grouping differentially expressed genes into functional categories

DEGs were grouped based on the KEGG KO number assignment (Section 2.5.5). The KEGG Mapper reconstruct pathway tool (v2.8) was used to characterise the pathways of the up and down regulated DEGs (Kanehisa et al., 2017). Gene expression analysis of the predicted antimicrobial genes of RW109 (Section 2.5.5) was carried out in response to the test exposure conditions. The distributions of DEGs across the RW109 genome were visualised with some test conditions using CGView as described in Section 2.5.8.

2.8 Metabolic modelling reconstruction experiments

Relevant programming scripts used throughout this section are provided within the relevant Chapter folder on the CD-ROM, which accompanies this thesis.

2.8.1 Updating the PA14 model

A recently published metabolic model for the *P. aeruginosa* PA14 strain (Bartell et al., 2017) was used as a basis to reconstruct a biochemical network for the RW109 industrial strain. The PA14 model was provided in an Excel format, which had two tabs labelled ‘reactions’ and ‘metabolites’.

The reactions tab had the following headers:

Column 1: Abbreviation = rxn03253
 Column 2: Name = acyl-CoA dehydrogenase (decanoyl-CoA)
 Column 3: Reaction = cpd00015[c] + cpd03128[c] -> cpd00982[c] + cpd03129[c]
 Column 4: Reactions with MetNames = FAD + Decanoyl-CoA -> FADH2 + (2E)-Decenoyl-CoA
 Column 5: Gene Protein reaction (GPR) = ((PA14_52900) or (PA14_06600) or (PA14_31580))
 Column 6: Genes = PA14_52900 PA14_06600 PA14_31580
 Column 7: Protein = (optional)
 Column 8: Subsystem = (optional)
 Column 9: KEGG Subsystem = Fatty acid metabolism
 Column 10: Reversible = 0
 Column 11: Lower bound = 0
 Column 12: Upper bound = 1000
 Column 13: Objective = 0 (optional)
 Column 14: Confidence Score = (optional)
 Column 15: Enzyme Commission (EC) Number = 1.1.1.1 (optional)
 Column 16: Notes = N/A (optional)
 Column 17: References = Supporting the inclusion of the reaction (optional)

The metabolites tab had the following headers:

Column 1: Abbreviation = cpd00015[c]
 Column 2: Name = FAD
 Column 3: Formula (neutral) = (optional)
 Column 4: Formula (charged) = C27H31N9O15P2
 Column 5: Charge = - 2
 Column 6: Compartment = (optional)
 Column 7: KEGG ID = C00016
 Column 8: PubChem ID = (optional)
 Column 9: ChEBI ID = (optional)
 Column 10: InChI string = (optional)
 Column 11: Smiles = (optional)

For the PA14 metabolic reconstruction described by (Bartell et al., 2017), the *Pseudomonas_aeruginosa_PA14_109.faa* annotated amino acid FASTA sequence from the *Pseudomonas* database was used. This file was downloaded and compared against the Prokka annotated amino acid sequence of the RW109 strain, using the command line standalone NCBI Protein-Protein BLAST (BLASTp, v2.2.29+) (Camacho et al., 2009). Matches between genes in the two strains were determined by applying an E-value cut-off of 1E-40, and setting the best-hit filtering algorithm overhang parameter to 0.25.

The BLASTp results were used to replace the PA14 locus tags with the matching RW109 locus tags, within columns 5 and 6 of the model reactions tab in the Excel format. This was achieved using the Excel multi-find and replace function. Column 5 within the reactions tab of the model represents the gene-protein-reaction (GPR) associations; a reaction can be associated with just one gene, or more than one gene can be involved. A protein complex occurs when all the genes are required for a reaction to take place, and within the Excel model in column 5 these genes were separated by the word 'and'. Isozymes occur when a reaction is associated with more than one gene, which encodes for the reaction enzyme, but not all are required for that reaction to take place and in the model, these genes were separated by the word 'or'. PA14 genes that matched with more than one RW109 gene in the BLASTp search were included in column 5, following the rule for isozymes.

Any PA14 model genes that did not match with genes in the RW109 strain, were manually evaluated by reducing the strict BLASTp filtering parameters. The PA14 biomass reaction was also updated to represent the industrial strain, by incorporating the RW109 strain specific compositions of DNA, RNA and amino acids (see *RW109_Biomass Parameters.xlsx* on CD-ROM). During the PA14 model reconstruction the approach outlined by (Chavali et al., 2008) was used to generate the biomass formulation.

2.8.2 Assigning new reactions to the RW109 metabolic model reconstruction

RW109 FASTA sequences were run through the modelSEED (Henry et al., 2010) and KEGG (Kanehisa et al., 2017) databases to identify additional reactions for the RW109 strain. New reactions were added to the model in Excel format following the order for the reactions tab (Section 2.8.1). Reactions were classified via their KEGG functional categories and assigned enzyme commission (EC) numbers from the ExPASy:SIB bioinformatics resource portal (Artimo et al., 2012). The metabolites of the new reactions, not already in the model were added to Excel model format following the order for the metabolites tab (Section 2.8.1).

2.8.3 Loading the model into MATLAB

MATLAB (v9.1.0.441655, R2016b) was used throughout metabolic modelling experimental analysis, along with the COntstraint-Based Reconstruction and Analysis (COBRA) Toolbox (v2.0) (Schellenberger et al., 2011). The COBRA toolbox was initiated in MATLAB using the `initCobraToolbox` function:

(<https://opencobra.github.io/cobratoolbox/deprecated/docs/cobra/initCobraToolbox.html>). The Gurobi 6 solver was used for all COBRA analysis. The RW109 metabolic model in Excel format was loaded into MATLAB using a `d_xls2model_JAB.m` script (CD-ROM). This script writes in a model from an Excel spreadsheet and converts it into a COBRA Toolbox model format. Once in MATLAB the following steps were always carried out before any model analysis and simulations:

- 1) Set objective function: `RW109model = changeObjective(RW109model,'RW109_Biomass')`
- 2) Change media to LB: `RW109_LB = changeMedia_SEED(RW109model,1)`
- 3) Optimise model: `optimizeCbModel(RW109_LB)`

The `changeObjective` function was used to designate the `RW109_Biomass` reaction as the objective function for the constraint-based model:

(<https://opencobra.github.io/cobratoolbox/latest/modules/FBA.html>). The `changeMedia_SEED.m` script (CD-ROM) worked by changing the lower and upper bounds of the RW109 model's exchange reactions to alter the *in-silico* media. LB media was always chosen, and the substrates associated with this media are provided within the `changeMedia_SEED.m` script. LB was chosen as it represented a complete *in-silico* medium. This script is an approximation of LB medium as LB is not well-defined, and the lower and upper bounds set within the script allow for a degree of flexibility, in what substrates the model would uptake.

The `optimizeCbModel` function was used to perform flux balance analysis (FBA) on the model: (<https://opencobra.github.io/cobratoolbox/latest/modules/FBA.html>).

FBA is regularly used as a constraint-based reconstruction and analysis tool to examine the metabolite flow through genome-scale metabolic network reconstructions (Orth et al., 2010). FBA enables the prediction of an organism's metabolic capabilities, such as growth, via a mathematical representation of the metabolic reactions displayed as stoichiometric matrix, together with a biologically relevant objective function (Lee et al., 2006b). A fundamental assumption of FBA is the optimum performance of an organism in relation to a specified metabolic reaction, known as the objective function; one commonly used objective function is the biomass reaction, which serves as an approximation for growth (Orth et al., 2010). Upon setting an objective function, subsequent analysis of the network is based on the conservation of mass through solving a system of equations. As well as identifying the resulting distribution of metabolic fluxes, that lead to optimal flux through an objective function (Raman and Chandra, 2009).

The output of `optimizeCbModel` was a solution structure that contained the maximal flux through the chosen `RW109_Biomass` objective function (f , units = hr^{-1}) and was used as a prediction of model growth in LB complete media. Additionally, this solution structure included the primal solution (x = the flux carried by each reaction within the model), the dual solution (y = metabolite shadow prices; indicates how much the addition of a metabolite will increase or decrease the objective value), the reduced cost (w = how much each reaction affects the objective) and a universal status flag (`stat` = solver status which

outputs 1 for an optimal solution, 2 for an unbounded solution, 0 if solution is infeasible and -1 if no solution is reported).

2.8.4 Removing PA14 genes with no similarities to RW109 genes

Once loaded into MATLAB the PA14 model genes with no similarities to any genes within the RW109 genome sequence, as identified from the BLASTp search, were individually deleted from the RW109 model using the SingleGeneDeletion COBRA Toolbox function:

(<https://opencobra.github.io/cobratoolbox/latest/modules/deletion.html>).

The function outputs the computed growth-rate ratio between the deletion strain and wild type strain (grRatio), the deletion strain growth rate (grRateKO- hr⁻¹), and the wild type growth rate (grRateWT- hr⁻¹). The function also demonstrates if the gene deletion results in a reaction being removed from the model (hasEffect) and will produce a list of deleted reactions for each gene grRateKO (delRxns) and a list of FBA fluxes for KO strains (fluxSolution).

These outputs were used to determine if the PA14 model genes, which had no similarities to any in the RW109 genome sequence were essential for model growth. If they could be deleted without any effect, they were removed from Excel version of the model, along with the reactions and metabolites associated with them (metabolites were only removed if they were not associated with any reactions remaining in the model). Genes that could not be deleted due to their effect on model growth were kept in the model and further analysis was carried out using BLAST searches, to determine an alternative RW109 gene associated with the essential reaction.

2.8.5 Integrating RNA-Seq gene expression data into the RW109 model

Transcriptomic data which profiled gene expression changes when RW109 was exposed to relevant industrial conditions (Section 2.7), was integrated into the model using the Metabolic Adjustment by Differential Expression (MADE) algorithm (Jensen and Papin, 2011), within the Toolbox for Integrating Genome-scale metabolic models, Expression data, and transcriptional Regulatory networks (TIGER) (Jensen et al., 2011). The MADE_Model_integration.m MATLAB script was used for the integration of the RNA-Seq data into the RW109 model (CD-ROM).

To predict which model genes were switched on or off in RW109, the log₂-fold changes and corresponding adjusted p-values of the test exposure conditions when compared to the control condition, were input into the MADE algorithm. A gene was considered switched ‘off’ by MADE in the control condition and switched ‘on’ in the test condition, if expression increased significantly between the two. A gene would also remain ‘on’ in a model if it had no change in expression, and was considered to be necessary for the model to function, as defined by a necessity for the minimal flow of flux through the specified objective function.

When a gene was found to have a significant decrease in expression between the control and test conditions, it was considered switched ‘off’ in the test condition if flux was still able to flow through the

defined objective function, meaning it was not required for metabolism. In a situation where the integrated expression data prevents flux through the objective function, MADE would ensure that the related gene remains 'on' in the model. The MADE algorithm, with the calculation of a penalty score, maintained functionality within condition specific models. This score was sought to be kept at a minimum to limit discrepancies with expression data; statistical significance of the log₂-fold changes, was also used to prioritise inconsistencies in gene states. Integration of gene expression data into the model, results in a metabolic representation more consistent with the physiological status of a bacterium when exposed to a particular condition.

The output of the algorithm was a series of eight condition-specific MADE models in the TIGER Toolbox format with a constrained set of reactions. These reactions were considered to be either required for a functioning model (as specified by the objective function), and/or were verified by the transcriptomic integrated data. These MADE models represented the functional gene states and metabolic adjustments of the control exposure condition (control model) and the 7-test exposure conditions (models 1-7). The models differed in that some of their reactions were 'turned on' or 'turned off' according to the gene expression levels of their corresponding genes. The TIGER MADE models were converted into a COBRA format, via the deletion of the model genes from the base RW109 model in COBRA, according to the MADE functional gene states output (See MADE_Model_integration.m MATLAB script for more details). A COBRA format was necessary for further analysis and simulations of the condition-specific models.

2.8.6 Gene and reaction essentiality predictions from the condition specific MADE models

The COBRA Toolbox was used to predict gene and reaction essentiality from the 8 MADE models using the MATLAB script COBRA_Essentiality_Predictions.m (CD-ROM). This script worked by simulating the individual removal of each remaining gene within the condition specific models in COBRA format, using the singleGeneDeletion function (<https://opencobra.github.io/cobratoolbox/latest/modules/deletion.html>). A reaction that requires a deleted gene to function will have inhibited flux, and if the reaction was necessary for the condition specific model to produce biomass, the gene was predicted as essential. These genes were identified as resulting in a grRateKO of < 0.0001 hr⁻¹.

The reactions associated with predicted essential genes for each condition specific model, were individually removed using the singleRxnDeletion function

(<https://opencobra.github.io/cobratoolbox/latest/modules/deletion.html>).

(See MATLAB script COBRA_Essentiality_Predictions.m). The reactions required for the condition specific model to produce biomass were identified as those resulting in a grRateKO of < 0.0001hr⁻¹. The essential genes and reactions predicted for the control condition were compared to those for each of the 7-test exposure conditions; essential reactions for each test condition that were not found within the control condition, were grouped according to their KEGG functional category assignment. The JVENN

online tool was used to create Edwards-Venn diagrams to determine relationships between the predicted essential reactions for each test exposure condition (Bardou et al., 2014).

3. Whole genome sequencing and comparative genomics of the industrial *P. aeruginosa* strain RW109

3.1 Introduction

The first complete whole genome sequence of *P. aeruginosa* was acquired for the clinical strain PAO1 by Stover et al., (2000) and at the time was the largest bacterial genome ever to be sequenced using whole-genome-shotgun-sampling. Analysing the genome in such detail, provided valuable insights into the genetic diversity and vast functional characteristics of PAO1, whilst also providing a perspective on the association between genome size and genetic complexity (Stover et al., 2000). The PAO1 sequencing project also supported the development of the *Pseudomonas* Genome database (Winsor et al., 2016) which currently contains 3,348 annotated genome sequences and provides valuable information on drug targets, antibiotic resistance genes and vaccine candidates. The technology of whole genome sequencing has advanced considerably since the year 2000 and has evolved at an exponential pace with development of new methods which are increasingly accurate and cost-effective (Levy and Myers, 2016, Klemm and Dougan, 2016).

Complete genome sequencing of bacteria is an important analysis tool, which provides the framework for understanding the characteristics of a particular bacterial strain. There have also been advancements in the development of databases for annotation of genome coding sequences (CDS) such as the Clusters of Orthologous Groups (COGs) (Tatusov et al., 2000) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa et al., 2017) which assign information about the strain's biological functions using a complete genome sequence. Antimicrobial resistance and virulence genes can also be characterised using databases such as The Comprehensive Antimicrobial Resistance Database (CARD) (McArthur et al., 2013), BacMet: the biocide and metal resistance database (Pal et al., 2013) and the Virulence Factor Data Base (VFDB) (Chen et al., 2016b). Functional annotations are important for assessing the pathogenic potential and the environmental adaptability of strains (Stover et al., 2000). For accurate transcriptomic profiling analysis techniques such as RNA-Seq, a detailed understanding of biological characteristics of a reference strain is also crucial (Conesa et al., 2016).

There are currently numerous methods available to obtain a bacterial whole genome, and Pacific Biosciences (PacBio) single-molecule real-time (SMRT) sequencing technology is one of the most accurate methods for an error free complete genome (Rhoads and Au, 2015, Nakano et al., 2017, Quail et al., 2012). It is classified as a third generation sequencing technique because unlike second generation methods such as Illumina, PacBio is able to provide information on regions difficult to sequence by having an unbiased G+C coverage and it generates extra-long multi-kilobase reads (Shin et al., 2013, Chin et al., 2013). The technology works by attaining the sequence data during the replication process of the DNA in real time (Rhoads and Au, 2015). The equipment uses a SMRTbell template which is loaded into the numerous sequencing units of a SMRT cell and the template acts as a closed, single-stranded circular DNA which is generated by ligating hairpin adaptors to both ends of the target double-stranded DNA molecule to be sequenced (McCarthy, 2010). Each sequencing unit contains an immobilised single

polymerase which binds to the hairpin adaptor of the SMRTbell template resulting in a DNA polymerase/template complex; this binding initiates the replication of the target DNA molecule (Rhoads and Au, 2015). Within the sequencing units are four fluorescently labelled phospholinked-nucleotides which each produce unique emission spectrums related to the colour of the fluorophore, when a base is detected during the replication of the target DNA (Berlin et al., 2015). The sequencing units are made up of zero-mode waveguides (ZMWs) which only permit light to illuminate at the bottom of the sequencing unit where the DNA polymerase/template complex is immobilized (Rhoads and Au, 2015, McCarthy, 2010). Specific light pulses related to the colour of the four different fluorophores are recorded and interpreted as a sequence of bases. The ZMWs result in a small detection volume, which provides 1000-fold enhancement in limiting the amount of background noise during sequencing (McCarthy, 2010, English et al., 2012, Rhoads and Au, 2015). This real-time sequencing process occurs in parallel within the thousands of SMRT cell sequencing units containing ZMWs (Chin et al., 2013, Rhoads and Au, 2015). The end result is a collection of long PacBio reads which can be combined using *de novo* assembly to form a high-quality complete whole genome sequence (Chin et al., 2013, Rhoads and Au, 2015); an integral component for informative bacterial genomic investigations.

PacBio sequencing technology was first used to obtain a complete *P. aeruginosa* genome in 2015 with the DSM-50071^T strain (Nakano et al., 2015). This analysis resulted in a single circular contig, which represented the main *P. aeruginosa* chromosome. The completeness of the sequence provided a reliable reference genome due to the higher-resolution and accuracy of the base order. For bacteria with large genomes and high GC content such as *Pseudomonas*, the reported accuracy of using PacBio SMRT technology is very useful for detailed characterisation of the antimicrobial and pathogenic potential of different strains. The re-sequencing of the reference PA1 strain using SMRT technology also revealed a number of errors in the genome previously obtained using Illumina sequencing by synthesis (SBS) technology; errors included nucleotide sequence mutations, inverted regions and contig arrangement error (Li et al., 2016). The complete PacBio PA1 genome was used for accurate classification of the strain's virulence factors, regulatory proteins, secretion system proteins, type II toxin-antitoxin pairs and genomic islands. The sequence also enabled detailed whole genome comparisons which demonstrated the similarity of PA1 to other *P. aeruginosa* strains and also highlighted the variations in horizontal gene transfer characteristics within the genomic islands and prophage regions of the strain (Li et al., 2016).

Whole genome comparative genomics is an important research tool to determine similarities and differences between bacterial strains; it provides detailed analysis on the relatedness of strains and how their genomes have evolved (Mathee et al., 2008, Winsor et al., 2016). The *P. aeruginosa* strains PAO1 and UCBPP-PA14 (PA14) are classified as universal reference strains in both laboratory experiments and genome analysis investigations (Stover et al., 2000, He et al., 2004, Klockgether et al., 2010, Mikkelsen et al., 2011a). PAO1 is employed as a reference as it was the first *P. aeruginosa* strain to be associated with a complete whole genome sequence (Stover et al., 2000) and the highly virulent PA14 strain provides an informative model for pathogenesis and biofilm formation investigations (Mikkelsen et al., 2011a). These reference strains are both clinical isolates and can be used for comparison against *P. aeruginosa* strains

obtained from the environment and industry to determine genomic characteristics, which enable survival within niche settings. Another advantageous component of comparative genomics is classifying the pan-genome of a collection of bacterial sequences to identify the core and accessory genes (Page et al., 2015). This analysis determines the network of genes, which are shared between collections of isolates and is used to establish the functional features specific to a certain strain type.

There are currently over 100 complete *P. aeruginosa* genomes available in the *Pseudomonas* database (and 3000+ draft genomes) which represent clinical and natural environments (Winsor et al., 2016). A previous study determined the draft genome sequences for eighteen industrial strains (Weiser, 2015), but as far as we are aware, there are no complete genome sequences available for a *P. aeruginosa* strain isolated from industrial contamination.

3.1.1 Aim and objectives

The overall aim of this chapter was to characterise the complete genome of an industrial *P. aeruginosa* strain. This was achieved with the following objectives:

- 1) Obtain a complete genome of the industrial strain RW109 using PacBio Single Molecule, Real-Time (SMRT) technology.
- 2) Identify the basic features of the RW109 genome.
- 3) Characterise the functional categories of the RW109 coding sequences (CDS).
- 4) Identify distinctive functional characteristics of the RW109 strain when compared to the clinical reference strains PA14 and PAO1.
- 5) Comparison of RW109 with a panel of *P. aeruginosa* genome sequences representing strains isolated from clinical, environmental and industrial sources.

3.2 Results

3.2.1 Sequence assembly, annotation and basic feature identification

RW109 was sequenced using PacBio SMRT technology and the FASTQ files were assembled with Canu as described in Section 2.5. The total output bases resulting from the two SMRT sequencing cells was 3,583 Mb with a length distribution that peaked at 10 kb (Figure 9). The assembly generated three contigs and annotation with Prokka identified the function of 7,303 CDS with 2,199 of these being hypothetical proteins (Table 4). The genome sequence of RW109 revealed a G+C content of 65.11% and a genome size of 7.8 Mb with twelve ribosomal RNA (rRNA), seventy-six transfer RNA (tRNA) and one transfer-messenger RNA (tmRNA) molecules. General features of the Prokka annotated reference strains PA14 and PAO1 were also shown in Table 4. Both reference strains have one contig and in comparison to RW109, their genomes were smaller; PA14 genome was 15.7% smaller and the PAO1 genome 19.2% smaller. RW109 had 1,402 more CDS when compared to PA14 and 1,632 more CDS than PAO1. The annotated amino acid, nucleotide and general feature format (GFF) files for the RW109 sequence can be found in the CD-ROM (RW109.FAA, RW109.FFN and RW109.GFF). The annotated and assembled sequence was uploaded onto the European Nucleotide Archive (ENA) under the accession number SAMEA104432335.

Table 4. The basic genome features of the industrial *P. aeruginosa* strain RW109 in comparison to the reference strains PA14 and PAO1

Genome Features	<i>P. aeruginosa</i> Strain		
	RW109	PA14	PAO1
Genome size (bp)	7,756,224	6,537,648	6,264,404
Contig numbers	3	1	1
DNA G + C content (%)	65.11	66.29	66.56
Coding sequences (CDS)	7,303	5,901	5,671
Number of hypothetical proteins	2,199	2,026	1,858
rRNA genes	12	12	12
tRNA genes	76	69	73
tmRNA	1	1	1

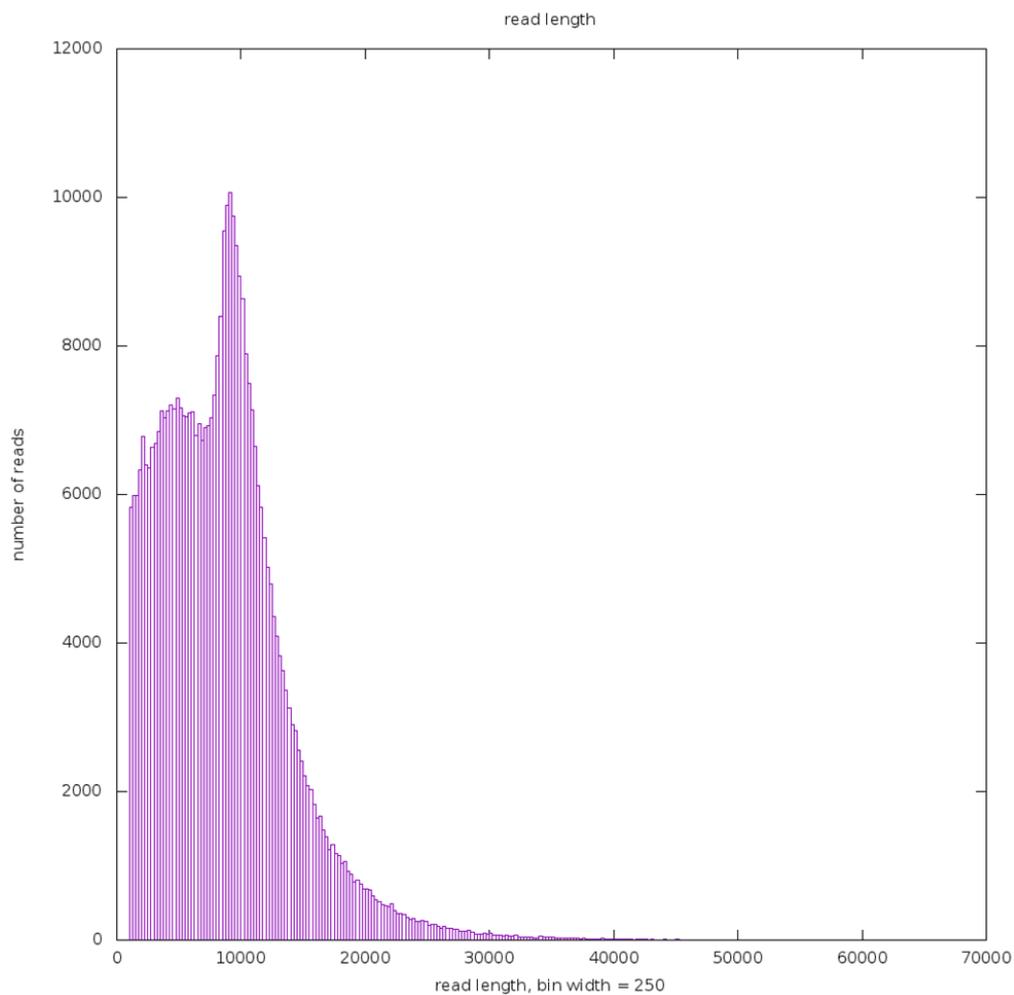


Figure 9. Read length distributions from PacBio Sequencing of RW109 using SMRT cell technology

Bar chart showing the overview of the length distribution of the reads generated from PacBio sequencing of RW109. The total output size of the bases was 3,583 Mb with a length distribution that peaked at ~10 kb.

3.2.2 Identifying plasmids in the assembled genome

3.2.2.1 Nucleotide BLAST

The assembly of the PacBio sequence of RW109 resulted in three contigs with contig 2 being the largest with 7,049,347 bp. Contig 1 had 555,265 bp and contig 3 had 151,612 bp. QUAST analysis of the FASTA sequence of each RW109 contig was obtained as described in Section 2.5.2. The G+C content of each contig was as follows, contig 1 = 58.09%, contig 2 = 65.84% and contig 3 = 57.28%. Due to the smaller size of contigs 1 and 3 and the difference observed in G+C content when compared to contig 2, comparative analysis was carried out to determine if these contigs were plasmids.

A BLASTn comparison of the assembled un-annotated RW109 genome was carried out against six circular *P. aeruginosa* complete plasmid sequences (Table 5). Contig 1 was found to have high alignment similarities to plasmid pOZ176 with 76 BLASTn hits over the sequence (mean values for percentage ID= 94.0%, alignment length= 5046.46, E-value= 7.49e-09 and BIT-Score= 8605.57). Contig 1 was also found to have high alignment similarities to plasmid pBM413 with 52 BLASTn hits over the sequence (mean values for percentage ID= 95.54%, alignment length= 6494.37, E-value= 1.46e-09 and BIT-Score= 11182.15). The third contig had high alignment similarities to a plasmid from the *P. aeruginosa* strain S04 90 with 14 BLASTn hits over the sequence (mean values for percentage ID= 94.72%, alignment length= 8561.29, E-value= 1.14286e-67 and BIT-Score= 11182.15).

Table 5. *P. aeruginosa* plasmid sequence information

<i>P. aeruginosa</i> strain		Plasmid			
Name	Isolation source	Name	GC content (%)	Length (bp)	Accession number
S04 90	Microbial mat material	S04 90 (plasmid un-named)	57.73	159,187	CP011370
BH6	Tissue of patient with lung disease	pBH6	55.86	3,652	CM003767
121617	Sputum	pBM413	56.41	423,017	CP016215
3448	Bloodstream	pPA3448	58.9	49,094	CM007350
7790	Unknown	pPA7790	58.9	49,021	CP015000
96	Clinical isolate from Guangzhou, China	pOZ176	57.00	500,839	KC543497

3.2.2.2 Comparison of contig synteny

ACT comparisons were carried out between contig 1 and plasmids pOZ176 and pBM413 and between contig 2 and the plasmid from the *P. aeruginosa* strain S04 90 (Figure 10 A and B). An ACT comparison was also carried out between the whole RW109 genome sequence (all contigs) and the PA14 reference strain (Figure 10C). The comparative synteny plot in Figure 10A shows contig 1 of RW109 in the middle with pOZ176 at the top and pBM413 at the bottom of the plot. The plot indicates that contig 1 has a high degree of similarity with these two plasmids as shown by the red bars, which indicate conserved sequences and gene order. A small number of blue lines were present between pOZ176 and contig 1 suggesting regions of sequence similarity that have been inverted. The gaps between the red bars indicate no sequence similarity and these gaps occur in approximately the same positions on the plot for both the plasmids when compared to contig 1, suggesting that these regions were only found in RW109 contig 1 when compared to these plasmids. There was also a high degree of similarity between contig 3 of RW109 and the plasmid from the *P. aeruginosa* strain SO4 90 (Figure 10B; see the red bars). However, RW109 contig 3 had no sequence homology to the end of the SO4 90 plasmid (Figure 10B). Figure 10C illustrates a comparative synteny plot for the whole genome sequence of RW109 (including all contigs) and the reference *P. aeruginosa* PA14 strain. Only contig 2 had a high degree of sequence similarity with PA14 whereas contigs 1 and 3 had limited sequence homology to PA14.

The comparative analysis of contigs 1 and 3 with the *P. aeruginosa* plasmid sequences suggest that these were plasmids acquired by RW109 and that contig 2 was the main *P. aeruginosa* chromosome of the industrial strain. Throughout the rest of this thesis, contig 1 (554 kb) was referred to as RW109 plasmid 1, contig 2 (7,049 kb) as RW109 main chromosome and contig 3 (161 kb) as RW109 plasmid 2.

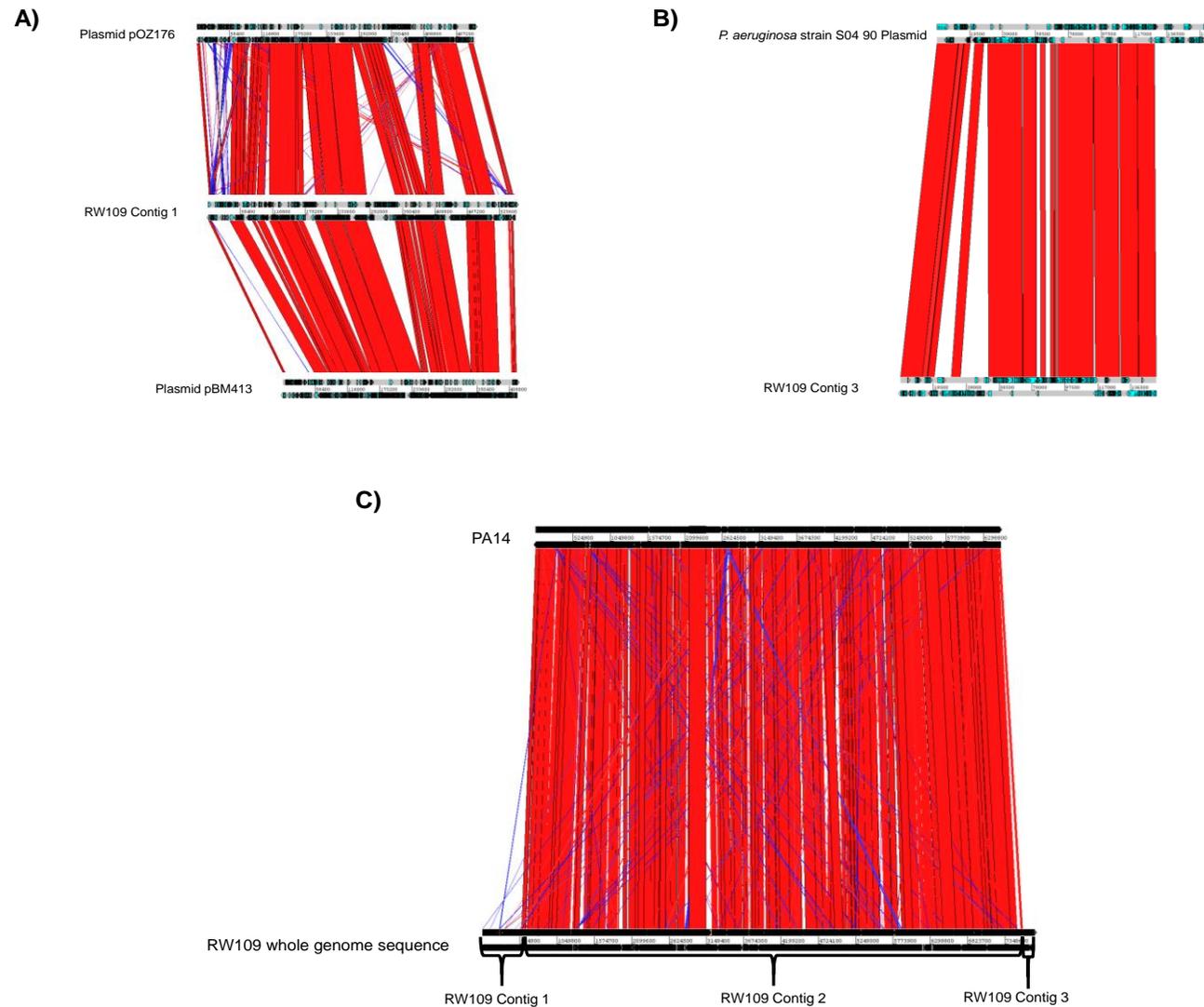


Figure 10. Comparative synteny line plots of the RW109 contigs 1 and 2 Vs. plasmid sequences and the RW109 whole genome sequence Vs. PA14

The analysis was carried out using the Artemis Comparison Tool (ACT) with the comparative crunch files generated using BLASTall. The red bars indicate regions of similarity, the blue bars show regions oriented in opposite directions and the gaps between bars were regions of non-similarity between the sequences. ACT comparisons were carried out between contig 1 and plasmids pOZ176 and pBM413 (A), contig 2 and the plasmid from the *P. aeruginosa* strain S04 90 (B) and between the whole RW109 genome sequence (all contigs) and the PA14 reference strain (C).

3.2.3 Assigning functional groups to the RW109 whole genome sequence

3.2.3.1 COG functional annotation

3.2.3.1.1 Whole genome COG annotation

The whole genome sequence of RW109 was categorised using COG functional annotation as described in Section 2.5.5.1. Out of the 7,303 CDS in the RW109 genome, 6,220 (85.17%) were assigned to a category (Table 6) with 123 CDS being annotated with two or more COG annotations. The percentage of each COG category was based on the total number of CDS and COG categories were divided into three well-characterised functional classes with metabolism having the highest number of CDS (46.01%), followed by cellular processes and signalling (22.28%), and information storage and processing (15.34%). A high number of CDS in RW109 were assigned to the poorly characterised COG categories with 25.59% being termed as function unknown (S). The CDS that were not assigned to a COG category were termed as a protein not being similar to any COG in the database and 1,083 RW109 CDS (14.83%) were allocated to this group.

Across the well-characterised functional classes, the top three COG categories with the highest percentage CDS totals throughout the whole genome sequence were transcription (K= 7.38%), amino acid transport and metabolism (E= 6.82%) and replication, recombination and repair (L= 5.13%). No CDS were assigned to the extracellular structures (W), nuclear structure (Y), cytoskeleton (Z) and general function prediction only (R) COG categories. The COG categories of the whole genome sequence were also divided up to represent the assigned functional annotations for each of the three contigs (Table 6). Circular maps of the main chromosome and plasmids 1 and 2 were generated using command line CGView as described in Section 2.5.8 (Figure 11). The first two outermost rings on the maps illustrate the predicted CDS on the reverse and forward strands respectively and were coloured according to the assigned COG categories.

3.2.3.1.2 RW109 main chromosome COG annotation

Within in the well-characterised COG functional annotations of the main chromosome the categories with the largest percentage of CDS were transcription (K= 7.75%), amino acid transport and metabolism (E= 7.04%), energy production and conversion (C= 5.44%) and inorganic ion transport and metabolism (P= 5.31%). Of the poorly characterised COG categories, the main chromosome had 26.46% CDS assigned to function unknown (S) and 10.89% which could not be assigned to any COG category (unknown). The COG categories which had the least number of CDS within the main chromosome (disregarding those with no CDS) were RNA processing and modification (A= 0.05%), chromatin structure and dynamics (B= 0.05%), cell cycle control, cell division, chromosome partitioning (D= 0.55%) cell motility (N= 1.31%), defence mechanisms (V= 1.37%) and nucleotide transport and metabolism (F= 1.71%).

3.2.3.1.3 RW109 plasmid 1 COG annotation

When focusing on the well characterised COG functional annotations of plasmid 1, the replication, recombination and repair (L) category represented the largest percentage of CDS (9.00%) and when the annotations of these CDS were analysed, 35.85% encoded for transposases. Other COG categories with large representations on plasmid 1 were amino acid transport and metabolism (E= 6.11%) and transcription (K= 4.92%). When analysing the CDS on plasmid 1 which were assigned to the E COG category, seventeen CDS were found to be in close proximity to each other within the genome, and when the COG accession numbers were analysed in the STRING protein-protein interaction database, 4 clusters were predicted (Figure 12). The cluster represented in red (Figure 12) comprised five CDS which was the biggest number of CDS with a high confidence interaction score and was predicted to be a cleavage system involved in glycine/serine degradation (RW109_00289, RW109_00290, RW109_00291, RW109_00292 and RW109_00293). The cluster with the greatest interaction confidence score (coloured as blue in Figure 12) contained 3 CDS (RW109_00305, RW109_00306 and RW109_00307) which were predicted to be involved in an ABC-type amino acid transport system. Two threonine efflux proteins were also identified within the E COG category of plasmid 1 (RW109_00569 and RW109_00570).

When analysing the CDS on plasmid 1 assigned to the K COG category, 58.62% were found to be regulatory proteins with 37.93% with a predicted transcriptional regulator function. Out of the total number of CDS on plasmid 1, 65.70% were poorly characterised whereas only 37.34% of the CDS on the main chromosome fell into this category. The COG categories with the lowest numbers of CDS on plasmid 1 were lipid transport and metabolism (I= 0.17%), cell cycle control, cell division, chromosome partitioning (D= 0.17%), RNA processing and modification (A= 0.17%), nucleotide transport and metabolism (F= 0.34%), secondary metabolites biosynthesis, transport, and catabolism (Q= 0.34%), cell motility (N= 0.51%) and defence mechanisms (V= 0.51%).

3.2.3.1.4 RW109 plasmid 2 COG annotation

The replication, recombination and repair (L) COG category also represented the highest number of well-characterised CDS on plasmid 2 with a percentage total of 14.10%, which was higher than the percentage of CDS represented in this category for both plasmid 1 and the main chromosome. These replication, recombination and repair CDS were spread out throughout the plasmid sequence (Figure 11; see deep pink coloured COG category). When the L COG category CDS of plasmid 2 were analysed, 27.27% were found to encode for transposases. The intracellular trafficking, secretion, and vesicular transport (U) COG category was found to have the second largest representation of CDS in plasmid 2 with of 5.77% of the total CDS. Plasmid 2 has a greater percentage of CDS represented in COG-category U when compared to the main chromosome and plasmid 1 where 0.68% and 2.73% of CDS were represented within this COG functional group.

When the annotated CDS from the plasmid 2 U-COG category were analysed, 8 of the genes were all in close proximity to each other and encoded 2 twitching mobility proteins (RW109_07269 and RW109_07279), a toxin co-regulated pilus biosynthesis protein E (RW109_07271), a type II secretion system protein E (RW109_07272), a pilin accessory protein (*pilO*) (RW109_07274), an outer membrane lipoprotein BfpB precursor (RW109_07275) and 2 hypothetical proteins (RW109_07268 and RW109_07270). A toxin co-regulated pilus biosynthesis protein Q (RW109_07276) which was assigned to COG functional group unknown (S), was also found in close proximity to these 8 U-COG category CDS. The amino acid sequences of these 9 CDS and the RW109_07276 CDS, were formatted into a multi-FASTA file and used to conduct a *Pseudomonas* database DIAMOND BLASTP search (Winsor et al., 2016). The resulting output for 8 out of the 9 CDS (excluding RW109_07276) were found to have a high degree of sequence similarity to a Section of the *P. aeruginosa* S04 90 plasmid encoding for a Type II secretion system (percentage ID= 100%, alignment length= 3307.00, E-value= 0.0 and BIT-Score= 6218.20). This further confirms the similarity of the third RW109 contig with the plasmid of the S04 90 strain. The RW109_07276 CDS was found to have sequence similarity to a plasmid (pND6-2) from the *Pseudomonas putida* strain ND6 (percentage ID= 88.0%, alignment length= 208.00, E-value= 0.0 and BIT-Score= 379.00). On the circular map of plasmid 2 in Figure 11, to the right on the second track this group of CDS were coloured in green/yellow representing the U COG category and have been labelled as an operon with a number of determinants for a Type II secretion system.

74.36% of CDS on plasmid 2 were poorly characterized which was a higher proportion when compared to plasmid 1 and the main chromosome. Plasmid 2 also has limited CDS within the metabolism COG functional class (1.28%) with only 2 CDS being associated with Lipid transport and metabolism (I), whereas the metabolism associated COGs of Plasmid 1 and the main chromosome represent 11.04% and 25.40% respectively of their total numbers of CDS.

Table 6. The Number of RW109 genes associated with general COG functional categories

COG Category		Description	<i>P. aeruginosa</i> strain RW109							
			Whole Genome		Plasmid 1		Main Chromosome		Plasmid 2	
			Value	Percentage (%) [*]	Value	Percentage (%) [*]	Value	Percentage (%) [*]	Value	Percentage (%) [*]
Well characterised categories										
Information storage and processing										
A	RNA processing and modification	4	0.05	1	0.17	3	0.05	0	0.00	
B	Chromatin structure and dynamics	3	0.04	0	0.00	3	0.05	0	0.00	
J	Translation, ribosomal structure and biogenesis	199	2.72	0	0.00	199	3.03	0	0.00	
K	Transcription	539	7.38	29	4.92	508	7.75	2	1.28	
L	Replication, recombination and repair	375	5.13	53	9.00	300	4.57	22	14.10	
Cellular processes and signalling										
D	Cell cycle control, cell division, chromosome partitioning	38	0.52	1	0.17	36	0.55	1	0.64	
O	Post-translational modification, protein turnover, and chaperones	222	3.04	7	1.19	214	3.26	1	0.64	
M	Cell wall/membrane/envelope biogenesis	323	4.42	10	1.70	312	4.76	1	0.64	
N	Cell motility	91	1.25	3	0.51	86	1.31	2	1.28	
P	Inorganic ion transport and metabolism	363	4.97	15	2.55	348	5.31	0	0.00	
T	Signal transduction mechanisms	305	4.18	18	3.06	285	4.35	2	1.28	
U	Intracellular trafficking, secretion, and vesicular transport	192	2.63	4	0.68	179	2.73	9	5.77	
V	Defence mechanisms	93	1.27	3	0.51	90	1.37	0	0.00	
W	Extracellular structures	0	0.00	0	0.00	0	0.00	0	0.00	
Y	Nuclear structure	0	0.00	0	0.00	0	0.00	0	0.00	
Z	Cytoskeleton	0	0.00	0	0.00	0	0.00	0	0.00	
Metabolism										
C	Energy production and conversion	371	5.08	14	2.38	357	5.44	0	0.00	
G	Carbohydrate transport and metabolism	225	3.08	5	0.85	220	3.35	0	0.00	
E	Amino acid transport and metabolism	498	6.82	36	6.11	462	7.04	0	0.00	
F	Nucleotide transport and metabolism	114	1.56	2	0.34	112	1.71	0	0.00	
H	Coenzyme transport and metabolism	172	2.36	5	0.85	167	2.55	0	0.00	
I	Lipid transport and metabolism	203	2.78	1	0.17	200	3.05	2	1.28	
Q	Secondary metabolites biosynthesis, transport, and catabolism	150	2.05	2	0.34	148	2.26	0	0.00	
Poorly characterised										
R	General function prediction only	0	0.00	0	0.00	0	0.00	0	0.00	
S	Function unknown	1869	25.59	101	17.15	1735	26.46	33	21.15	
Unknown	Protein is not similar to any COG	1083	14.83	286	48.56	714	10.89	83	53.21	
Total Number of CDS =		7,303		589		6,558		156		

Footnotes: *The percentage of COG categories is based on the total number of CDS.

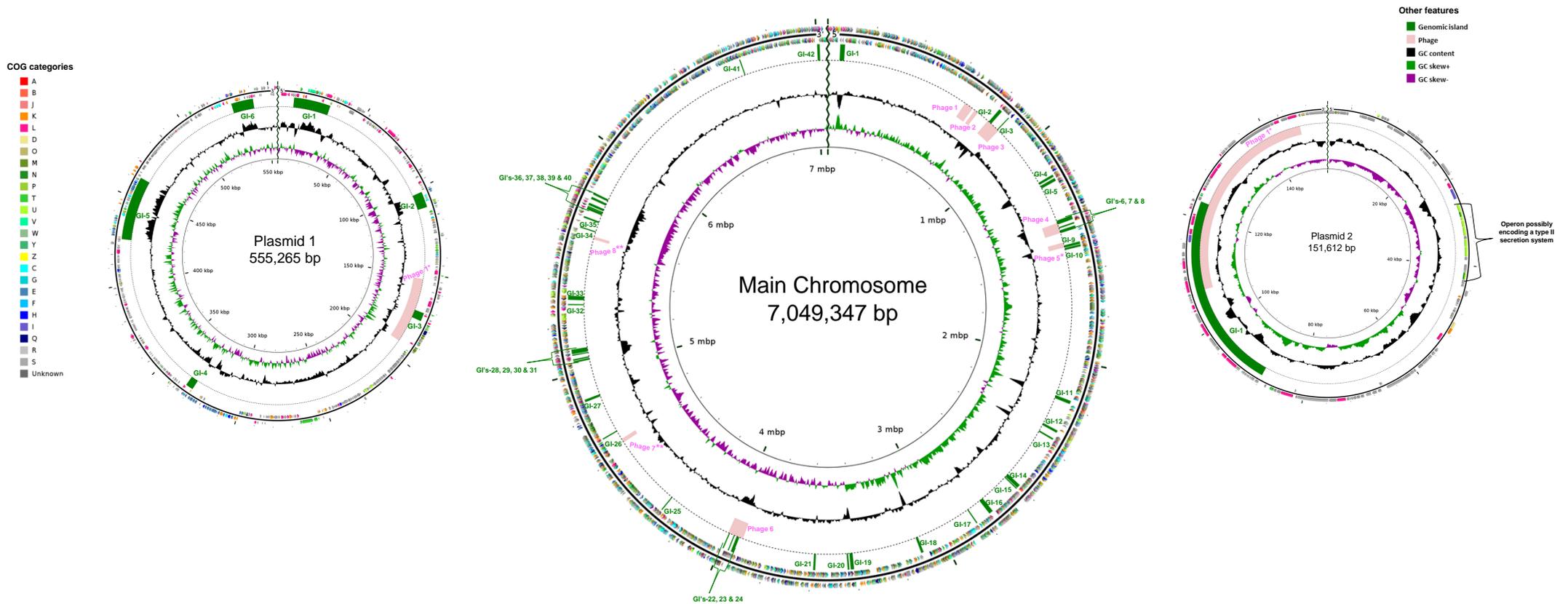


Figure 11. Circular maps of the RW109 main chromosome and plasmids 1 and 2

Starting from the outermost circle for each map and moving inwards the following tracks were identified: (1) predicted CDS on forward strand coloured according to COG categories, (2) predicted CDS on reverse strand coloured according to COG categories, (3) genomic islands (GIs) coloured green, (4) phages coloured light pink (those labelled with * indicate an incomplete phage and those labelled with ** indicate a questionable phage), (5) GC content (black), (6) positive and negative GC skew (green and purple, respectively) and (7) genome region by mbp for the main chromosome and kbp for plasmids 1 and 2. The size of the main chromosome and the two plasmids were not drawn to scale.

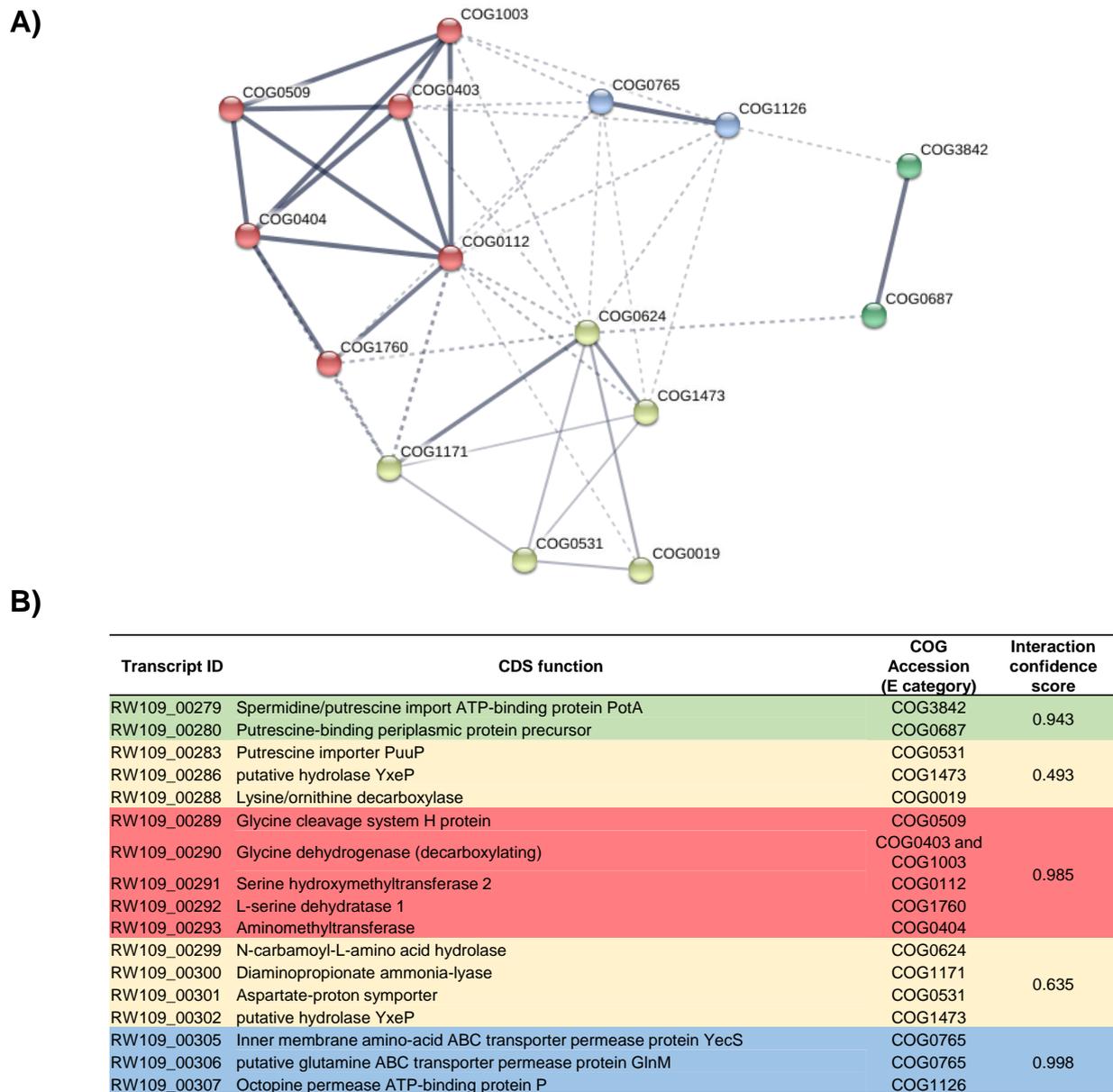


Figure 12. STRING network analyses of RW109 Plasmid 1 CDS within the E COG category

The interaction of seventeen CDS within the E COG category were analysed using the STRING database. A network of COG accession numbers which were split into 4 groups (A) via kmeans clustering as represented with the colour coding. The nodes represent the different COG accession numbers and the line thickness between the nodes indicates the strength of data support with thicker lines indicating a high confidence level score. B) The table represents the CDS and COG accession numbers and were colour coded in relation to A) to represent the clusters they fall into.

3.2.3.2 RW109 KEGG functional module assignment and metabolic and physiological potential

KEGG Orthology (KO) terms were assigned to the main chromosome and the two plasmid amino acid sequences using KASS as described in Section 2.5.5.3. Out of the 6,558 CDS in the main chromosome, 3,381 were annotated with a KO number (51.56%) with six CDS (RW109_01016, RW109_02053, RW109_02258, RW109_02260, RW109_04953 and RW109_06543) having two or more KO number associations. Out of the 589 plasmid 1 CDS, 146 were assigned with a KO number (24.79%) with 1 CDS (RW109_00245) having 2 more KO number associations and of the 156 CDS of plasmid 2, 28 were annotated with a KO number (17.95%). Using the KO-assigned CDS of the RW109 genome, KEGG defined modules were mapped using MAPLE, which searched against 279 pathways, 291 complexes, 155 functions and 40 signatures. The module completion ratios (MCR) were calculated to determine the completeness of the assigned modules. The whole genome of RW109, including the main chromosome and two plasmids, encoded for 171 biologically feasible KEGG modules.

3.2.3.2.1 Biologically feasible KEGG modules identified within the main chromosome

When KEGG functional modules were mapped to the KO-assigned CDS of the main chromosome 68 pathways, 59 complexes, 31 functional sets and 7 signatures were identified as biologically feasible (MCRs with Q-values of <0.5) (Figure 13). The two-component regulatory system category within the functional set group had the greatest number of modules which were biologically feasible across the four categories (24 MCRs (all 100%) with Q-values of <0.5) (Table 7). The cofactor and vitamins biosynthesis category within the pathways group had the second largest number of modules which were biologically feasible (13 MCRs with Q-values of <0.5), and the mineral and organic ion transport system category within the complexes group had the third biggest number of biologically feasible modules (11 MCRs with Q-values of <0.5).

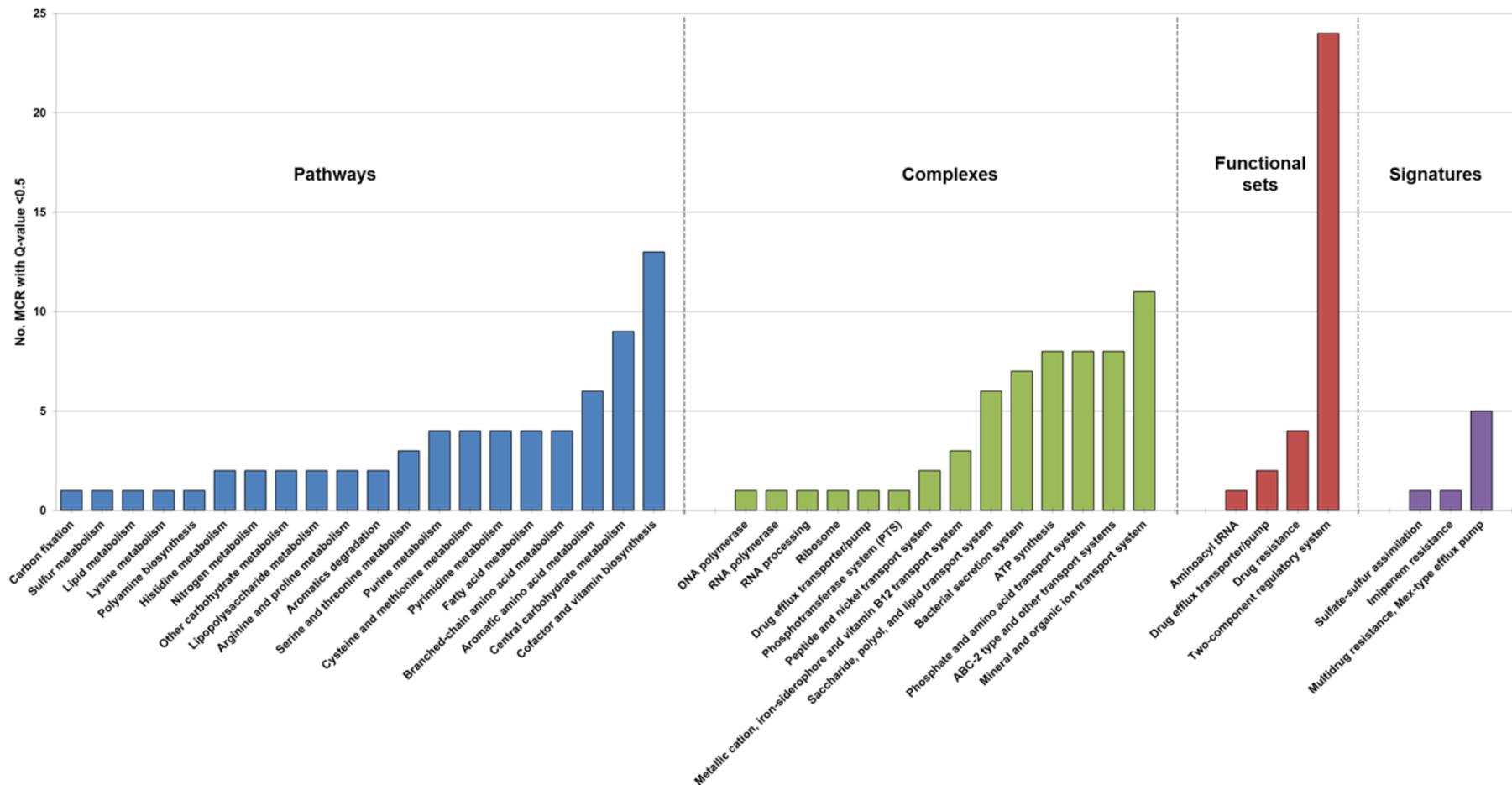


Figure 13. Biologically feasible KEGG functional modules of the RW109 main chromosome

Bar chart represents the number of KEGG modules mapped to KO-assigned CDS of the RW109 main chromosome which had MCRs with Q -values < 0.5 . Modules were grouped into four categories with pathways in blue, complexes in green, functional sets in red and signatures in purple. The MRC of each module and associated Q -value were calculated using the MAPLE system.

Table 7. Biologically feasible two-component regulatory systems on the RW109 main chromosome

Module ID	Two-component regulatory system	RW109 main chromosome KO Terms	RW109 main chromosome CDS*
M00434	PhoR-PhoB (phosphate starvation response)	K07636 K07657	RW109_06969 RW109_06968
M00444	PhoQ-PhoP (magnesium transport)	K07637 K07660	RW109_04990 RW109_04991
M00445	EnvZ-OmpR (osmotic stress response)	K07638 K07659	RW109_06800 RW109_06801
M00446	RstB-RstA	K07639 K07661	RW109_05012 RW109_04258
M00449	CreC-CreB (phosphate regulation)	K07641 K07663	RW109_01067 RW109_01066
M00452	CusS-CusR (copper tolerance)	K07644 K07665	RW109_03165 RW109_03166
M00453	QseC-QseB (quorum sensing)	K07645 K07666	RW109_06365 RW109_06364
M00454	KdpD-KdpE (potassium transport)	K07646 K07667	RW109_04526 RW109_04525
M00457	TctE-TctD (tricarboxylic acid transport)	K07649 K07774	RW109_05484 RW109_05485
M00471	NarX-NarL (nitrate respiration)	K07673 K07684	RW109_01839 RW109_01838
M00473	UhpB-UhpA (hexose phosphates uptake)	K07675 K07686	RW109_04072 RW109_04071
M00475	BarA-UvrY (central carbon metabolism)	K07678 K07689	RW109_05256 RW109_03415
M00477	EvgS-EvgA (acid and drug tolerance)	K07679 K07690	RW109_01768 RW109_00625
M00493	AlgZ-AlgR (alginate production)	K08082 K08083	RW109_06867 RW109_04341
M00497	GlnL-GlnG (nitrogen regulation)	K07708 K07712	RW109_06720 RW109_06721
M00501	PilS-PilR (type 4 fimbriae synthesis)	K02668 K02667	RW109_06121 RW109_06122
M00504	DctB-DctD (C4-dicarboxylate transport)	K10125 K10126	RW109_06765 RW109_06766
M00505	KinB-AlgB (alginate production)	K11383 K11384	RW109_07101 RW109_07100
M00507	ChpA-ChpB/PilGH (chemosensory)	K06596 K02487 K02657 K02658 K06597	RW109_01016 RW109_03121 RW109_01011 RW109_01012 RW109_01017
M00509	WspE-WspRF (chemosensory)	K13490 K11444 K13491	RW109_02111 RW109_02113 RW109_02112
M00515	FliB-FliC (polar flagellar synthesis)	K10942 K10943	RW109_05079 RW109_05078
M00523	RegB-RegA (redox response)	K15011 K15012	RW109_05834 RW109_05833
M00653	AauS-AauR (acidic amino acids utilization)	K17060 K17061	RW109_04685 RW109_04686
M00770	PfeS-PfeR (enterobactin-dependent iron acquisition)	K19609 K19610	RW109_03310 RW109_03311

Footnotes: * The order of the CDS matches with the order of the associated KO terms in column 2

3.2.3.2.2 Biologically feasible KEGG modules identified within plasmid 1

When KEGG functional modules were mapped to the KO-assigned CDS of plasmid 1, 3 pathways, 2 complexes, 0 functional sets and 1 signatures were identified as biologically feasible (MCRs with Q-values of <0.5) (Table 8). Within the pathways group, the 3 mapped modules were pyrimidine metabolism (Pyrimidine degradation: MRC=100%, Q-value <0.5), serine and threonine metabolism (Betaine biosynthesis: MRC=100%, Q-value <0.5) and cofactor and vitamin biosynthesis (C1-unit interconversion: MRC= 66.7%, Q-value <0.5). A plasmid 1 CDS (RW109_00291= K00600) which is thought to be involved in a glycine/serine cleavage system in Section 3.2.3.1.3 was identified as being part of the cofactor and vitamin biosynthesis module (M00140). The complexes group had two mapped modules both involved in transport; phosphate and amino acid transport system (Putative polar amino acid transport system: MRC=100%, Q-value <0.5) and ABC-2 type transport systems (lipopolysaccharide transport system: MRC=100%, Q-value <0.5). The CDS involved in the phosphate and amino acid transport system module (M00236) were also predicted by STRING protein-protein network analysis in Section 3.2.3.1.3 as being part of an ABC-type amino acid transport system. The complete module within the signature group was involved in multidrug resistance (efflux pump BpeEF-OprC: MRC=100%, Q-value <0.5). All of the pathway and complex modules identified in plasmid 1 were also found within the main chromosome, except for the multidrug resistance efflux pump BpeEF-OprC signature module.

Table 8. The biologically feasible KEGG functional modules of RW109 Plasmid 1

Module	Module ID	Functional Process	Plasmid 1 KO Number	Plasmid 1 CDS*
Pathways				
Pyrimidine metabolism	M00046	Pyrimidine degradation, uracil => beta-alanine, thymine => 3-aminoisobutanoate	K06016	RW109_00299
Serine and threonine metabolism	M00555	Betaine biosynthesis, choline => betaine	K00108 K00130	RW109_00565 RW109_00449
Cofactor and vitamin biosynthesis	M00140	C1-unit interconversion	K00600 K01491	RW109_00291 RW109_00468
Complexes				
Phosphate and amino acid transport system	M00236	Putative polar amino acid transport system	K02030 K02029 K02028	RW109_00304 RW109_00306 RW109_00307
ABC-2 type and other transport systems	M00250	Lipopolysaccharide transport system	K09690 K09691	RW109_00168 RW109_00169
Signatures				
Multidrug resistance, efflux	M00698	Multidrug resistance, efflux pump BpeEF-OprC	K18900 K18901 K18902 K18903	RW109_00010 and RW109_00017 RW109_00012 RW109_00013 RW109_00014

Footnotes: * The order of the CDS matches with the order of the associated KO terms in column 2

3.2.3.2.3 Biologically feasible KEGG modules identified within plasmid 2

When KEGG functional modules were mapped to the KO-assigned CDS of plasmid 2, no biologically feasible modules were identified and across all the module groups only one had an MCR of greater than zero but was associated with a Q-value of greater than 0.5. This module was found within the functional set group and was identified as a ChpA-ChpB/PilGH (chemosensory) two-component regulatory system (M00507) (MCR = 20% and Q-value =0.968). Plasmid 2 had a CDS associated with 1 of the 5 KO-terms required for the M00507 module to be complete; this KO-term was K02487, which was assigned to toxin co-regulated pilus biosynthesis protein Q (RW109_07276). This same CDS was found to be in close proximity to eight U-COG category genes, which encoded a putative type II secretion system (described in Section 3.2.3.1.4). Two out of the eight U-COG category CDS had KO numbers assigned (RW109_07269= K02669 and RW109_07279= K12203); when the KO terms were searched on the KEGG database, they correlated to a KEGG type II secretion system twitching motility protein PilT and a Icm/Dot secretion system protein, respectively.

3.2.3.3 Identifying the antimicrobial resistance and virulence genes in the RW109 genome

The ABRicate tool was used to identify antimicrobial and virulence genes within the RW109 genome. Table 9 shows the antimicrobial genes identified when searched against the CARD, ARG-ANNOT and Resfinder databases. Thirty-eight RW109 CDS were classified as having an antimicrobial function with three being recognised by all databases and two found by both CARD and Resfinder. One CDS (RW109_04031) was found to have percentage coverage and identity scores of greater than 98% for two different CARD antimicrobial resistance genes, AmrA and MexX. All predicted antimicrobial resistance genes were located on the main chromosome of RW109 and not on the two plasmids. Thirty-one of the identified resistance CDS were associated with efflux pump systems, two were beta-lactamase encoding CDS, three were antimicrobial target modification proteins, one encoded for an aminoglycoside phosphotransferase and there was one chloramphenicol acetyltransferase associated CDS. The assigned COG functional annotations of the RW109 antimicrobial CDS were also examined and the three categories with the greatest number of CDS were defence mechanisms (V= 36.84%), cell wall/membrane/envelope biogenesis (M= 28.95%) and the inorganic ion transport and metabolism (P= 10.53%). Out of the other CDS identified as being antimicrobial, 5.26% were found within the translation, ribosomal structure and biogenesis (J), transcription (K) and function unknown (S) categories and 2.63% were situated in signal transduction mechanisms (T), energy production and conversion (C) and amino acid transport and metabolism (E) categories.

When the ABRicate tool was used to identify the virulence genes of the RW109 genome by searching against the VFDB database, 239 CDS were classified as being associated with virulence (3.27% of the RW109 genome) (Gene annotations provided on the CD-ROM, Predicted virulence and BacMet RW109 genes.xlsx). Again, all predicted virulence CDS were found on the main chromosome and not on the two plasmids. Prokka annotated 25 of these CDS as hypothetical proteins, whereas the VFDB database identified a putative virulence function for them. The well-characterised COG functional classes for the virulent CDS were examined and the cellular processes and signalling class contained the greatest number of CDS (67.78%). Within this functional class 22.59% were assigned to the cell motility (N) category, 9.67% to the intracellular trafficking, secretion, and vesicular transport (U) category, 10.46% to the cell wall/membrane/envelope biogenesis (M) category and 9.62% to the signal transduction mechanisms (T) category. The poorly characterised class had 21.34% of the virulence CDS and the metabolism functional class had 15.48% and within this class the secondary metabolites biosynthesis, transport, and catabolism (Q) category had 7.11% of CDS.

Table 9. Antimicrobial resistance genes identified in the RW109 genome

RW109 CDS ID	ABRicate Antimicrobial Gene	ABRicate Antimicrobial Gene Function	Database	Accession Number
RW109_00759	TriA	Membrane protein that is fused to TriB and within the <i>P. aeruginosa</i> efflux complex TriABC-OpmH	Card	ARO:3003679
RW109_00760	TriB	Membrane protein that is fused to TriA and within the <i>P. aeruginosa</i> efflux complex TriABC-OpmH	Card	ARO:3003680
RW109_00761	TriC	RND transporter that is a part of TriABC-OpmH, a triclosan-specific efflux protein	Card	ARO:3003681
RW109_01029	MexA	Membrane fusion protein of the MexAB-OprM multidrug efflux complex	Card	ARO:3000377
RW109_01030	MexB	Inner membrane multidrug exporter of the efflux complex MexAB-OprM	Card	ARO:3000378
RW109_01031	OprM	Outer membrane factor protein part of the MexAB-OprM, MexVW-OprM, MexXY-OprM and the AmrAB-OprM efflux pump complexes	Card	ARO:3000379
RW109_01466	OpmD	Outer membrane channel protein of the efflux complex MexGHI-OpmD	Card	ARO:3000809
RW109_01467	MexI	Inner membrane transporter of the efflux complex MexGHI-OpmD	Card	ARO:3000808
RW109_01468	MexH	Membrane fusion protein of the efflux complex MexGHI-OpmD	Card	ARO:3000807
RW109_01469	MexG	Membrane protein required for MexGHI-OpmD efflux activity	Card	ARO:3000806
RW109_01555	APH(3')-IIb	Chromosomal-encoded aminoglycoside phosphotransferase in <i>P. aeruginosa</i>	Card, Resfinder and ARG-ANNOT	ARO:3002645
RW109_01564	PDC-1	Beta-lactamase found in <i>P. aeruginosa</i>	Card and Resfinder	ARO:3002497
RW109_02137	MexL	Specific repressor of mexJK efflux transcription	Card	ARO:3003710
RW109_02138	MexJ	Membrane fusion protein of the MexJK multidrug efflux protein	Card	ARO:3003692
RW109_02139	MexK	Inner membrane RND transporter in the MexJK multidrug efflux protein	Card	ARO:3003693
RW109_02266	ArnA	Modifies lipid A with 4-amino-4-deoxy-L-arabinose (Ara4N) allowing gram-negative bacteria to resist cationic antimicrobial peptides and polymyxin	Card	ARO:3002985
RW109_02297	MexP	Membrane fusion protein of the MexPQ-OpmE multidrug efflux complex	Card	ARO:3003698
RW109_02298	MexQ	Inner membrane transporter of the multidrug efflux pump MexPQ-OpmE	Card	ARO:3003699
RW109_02299	OpmE	Outer membrane factor protein that is part of the multidrug efflux pump MexPQ-OpmE	Card	ARO:3003700
RW109_03516	OprN	Outer membrane channel component of the MexEF-OprN multidrug efflux complex	Card	ARO:3000805
RW109_03517	MexF	Multidrug inner membrane transporter of the MexEF-OprN efflux complex	Card	ARO:3000804
RW109_03518	MexE	Membrane fusion protein of the MexEF-OprN multidrug efflux complex	Card	ARO:3000803
RW109_03520	MexS	Suppressor of MexT, which is an activator of the multidrug efflux pump MexEF-OprN	Card	ARO:3000813
RW109_04031	AmrA	Efflux pump subunit of the AmrAB-OprM multidrug efflux complex	Card	ARO:3002982
	MexX	Membrane fusion protein of the MexXY-OprM multidrug efflux complex.	Card	ARO:3003034
RW109_04032	AmrB	Membrane fusion protein of the AmrAB-OprM multidrug efflux complex	Card	ARO:3002983
RW109_04814	MexN	Inner membrane transporter of the MexMN-OprM multidrug efflux complex	Card	ARO:3003705
RW109_04815	MexM	Membrane fusion protein of the MexMN-OprM multidrug efflux complex	Card	ARO:3003704
RW109_05044	FosA	Confers resistance to fosfomycin via catalyzing the conjugation of glutathione to carbon-1 of fosfomycin	Card and Resfinder	ARO:3000149
RW109_05523	CatB7	Chromosome-encoded variant of the cat gene (chloramphenicol acetyltransferase gene)	Card, Resfinder and ARG-ANNOT	ARO:3002679
RW109_05712	MexV	Membrane fusion protein of the MexVW-OprM multidrug efflux complex	Card	ARO:3003030
RW109_05713	MexW	RND-type membrane protein of the efflux complex MexVW-OprM	Card	ARO:3003031
RW109_06173	OprJ	Outer membrane channel component of the MexCD-OprJ multidrug efflux complex	Card	ARO:3000802
RW109_06174	MexD	Multidrug inner membrane transporter of the MexCD-OprJ efflux complex	Card	ARO:3000801
RW109_06175	MexC	Membrane fusion protein of the MexCD-OprJ multidrug efflux complex	Card	ARO:3000800
RW109_06176	NfxB	Repressor of the efflux pump mexCD-oprJ	Card	ARO:3000820
RW109_06365	PmrB	Histidine protein kinase sensor Lipid A modification gene; part of a two-component system involved in polymyxin resistance that senses high extracellular Fe(2+)	Card	ARO:3003583
RW109_06568	OpmH	Outer membrane efflux protein required for triclosan-specific efflux pump function	Card	ARO:3003682
RW109_07131	OXA-50	Beta-lactamase found in <i>P. aeruginosa</i>	Card, Resfinder and ARG-ANNOT	ARO:3001796

3.2.3.4 Identifying biocide and metal resistance genes in the RW109 genome

The BacMet database was used to identify antibacterial biocide and metal resistance genes in the RW109 genome and a total number of sixty-three CDS were classified (Gene annotations provided on the CD-ROM = Predicted virulence and BacMet RW109 genes.xlsx). Four CDS from plasmid 1 were identified as BacMet resistance genes, 59 CDS identified were from the main chromosome and there were no CDS identified by BacMet from plasmid 2. Two plasmid 1 CDS identified by BacMet (RW109_00013 and RW109_00014) were previously identified as part of the multidrug resistance efflux pump BpeEF-OprC KEGG signature module (Table 8). The other plasmid 1 CDS identified with BacMet were annotated with functions in mercury (RW109_00034) and tellurium (RW109_00121) resistance.

The three COG categories with the greatest number of BacMet resistance CDS were inorganic ion transport and metabolism (P= 33.33%), defence mechanisms (V= 23.81%) and cell wall/membrane/envelope biogenesis (M= 12.70%). Nineteen of the BacMet resistance CDS were previously identified as antimicrobial genes using ABRicate in Section 3.2.3.3 and were associated with efflux pump systems. One BacMet gene was also classified as virulence CDS (RW109_01453) and was annotated as a Fe (3+)-pyochelin receptor-OS by both the BacMet and VFDB databases. The BacMet predicted CDS for RW109 are provided on the CD-ROM.

3.2.4 Identifying RW109 genomic islands

Genomic islands (GIs) for the main chromosome and two plasmids were predicted using Island Viewer as described in Section 2.5.6. The predicted genomic islands were visualised in the circular replicon maps (see Figure 11; highlighted in green on the third track in from the outside of the map). A total of forty-two possible GIs were identified within the main chromosome (Table 10), six were found in plasmid 1 and one in plasmid 2 (Table 11).

3.2.4.1 GIs encoded on the main chromosome

The mean length of the forty-two GIs identified in the main chromosome was 9810 bp and the mean number of CDS per GI was ten. GI-8 had the greatest number of CDS with eighteen, and GI-19 and GI-33 were the second largest with each having 17 genes (Table 10). Every GI contained hypothetical proteins (GI-18 and GI-41 only contained hypothetical proteins), 15 had transposase proteins and 9 had phage-associated proteins. The number of assigned COG categories for the GIs CDS of the main chromosome were analysed and the poorly characterised COG categories S and unknown represented the highest number of GI CDS (Figure 14). Within the well-characterised COG classes the three categories with the greatest number of GI CDS were replication, recombination and repair (L = 70), transcription (K = 26), cell wall/membrane/envelope biogenesis (M = 14) and defence mechanisms (V = 11). The CDS assigned to COG category L and K were well distributed throughout the 42 GIs. Nine out of the fourteen CDS assigned to the M category were found within GI-11 and five out of the eleven V category CDS were found in GI-16. There were no antimicrobial and/or virulence CDS identified in Section 3.2.4 within the forty-two main chromosome GIs. Two main chromosome GI CDS were classified as putative BacMet resistance genes (Section 3.2.3.4); a cobalt-zinc-cadmium resistance protein CzcA (RW109_02795) in GI-12 and a nickel and cobalt resistance protein CnrA (RW109_05607) found in GI-33.

Table 10. Genomic Islands within the *P. aeruginosa* RW109 main chromosome

Genomic Island	Start	End	Length (bp)	Number of CDS	Annotation
GI-1	53593	74190	20597	11	Hypothetical Protein (x5), Transposase IS200 like protein (x3), ParB-like nuclease domain protein, Adenosine monophosphate-protein transferase and cysteine protease IbpA precursor, Transposase DDE domain protein
GI-2	807810	822504	14694	12	Site-specific tyrosine recombinase XerC, Phage integrase family protein, Hypothetical Protein (x5), helix-turn-helix protein, recombination protein F, Transposase, Modification methylase DpnIIB, Type III restriction enzyme, res subunit
GI-3	860197	864197	4000	5	Hypothetical Protein (x4), Mor transcription activator family protein
GI-4	1169714	1179963	10249	12	Hypothetical Protein (x8), ATP-dependent RNA helicase HrpB (x2), helix-turn-helix protein (x2)
GI-5	1186579	1199486	12907	11	Hypothetical Protein (x3), LexA repressor, site-specific tyrosine recombinase XerC, Phage integrase family protein, SdiA-regulated, RNA polymerase-binding transcription factor DksA, Universal stress protein family protein, Bicarbonate transporter BicA, DNA polymerase IV
GI-6	1361129	1378932	17803	13	Phage integrase family protein (x2), Hypothetical Protein (x8), ATP-dependent RecD-like DNA helicase, Transposase DDE domain protein, Tyrosine recombinase XerC
GI-7	1392112	1398449	6337	13	Replicative DNA helicase, Hypothetical Protein (x6), Serine/threonine-protein phosphatase 1, Phage regulatory protein Rha (Phage_pRha), Phage holin family (Lysis protein S), Chitinase class I, Phage DNA packaging protein Nu1, Phage terminase large subunit (GpA)
GI-8	1407149	1420149	13000	18	Hypothetical Protein (x17), methionine aminopeptidase
GI-9	1478291	1488819	10528	10	Hypothetical Protein (x6), Prophage CP4-57 regulatory protein (AlpA), Chaperone protein HtpG, Competence protein CoiA-like family protein, Sporulation initiation inhibitor protein Soj
GI-10	1490712	1499857	9145	13	IS66 Orf2 like protein, putative Helix-turn-helix domain of transposase IS66, Hypothetical Protein (x6), Antiseptic resistance protein, HTH-type transcriptional regulator YfmP, PAAR motif protein, DNA-binding transcriptional activator AIs, Transposase
GI-11	2170714	2184863	14149	14	Hypothetical Protein (x5), lipopolysaccharide biosynthesis protein WzzE, Polysaccharide biosynthesis protein, UDP-glucose 4-epimerase, NAD dependent epimerase/dehydratase family protein, UDP-2, 3-diacetamido-2, 3-dideoxy-D-glucuronate 2-epimerase, putative glycosyl transferase, 3 beta-hydroxysteroid dehydrogenase/Delta 5->4-isomerase, Transposase DDE domain protein, putative undecaprenyl-phosphate N-acetylglucosaminyl 1-phosphate transferase
GI-12	2309559	2319897	10338	10	Penicillinase repressor, heat shock protein HtpX, Cobalt-zinc-cadmium resistance protein CzC precursor, Cobalt-zinc-cadmium resistance protein CzCB, Cobalt-zinc-cadmium resistance protein CzCA, Hypothetical Protein (x4), Hca operon transcriptional activator
GI-13	2347603	2354444	6841	6	Hypothetical Protein (x4), Replicase family protein, HTH-type transcriptional regulator SinR
GI-14	2593796	2609655	15859	14	putative deferrocetolase/peroxidase YfeX, DNA-binding transcriptional regulator CynR, Hca operon transcriptional activator, Formaldehyde dismutase, Hypothetical Protein (x3), Integrase core domain protein, Transposase DDE domain protein, UvrABC system protein B, HTH-type transcriptional regulator DmiR, SnaoL-like domain protein, 2-oxoglutarate amidase, Transcriptional repressor FrmR
GI-15	2611449	2616363	4914	4	Hypothetical Protein, S-formylglutathione hydrolase FrmB, NADH dehydrogenase-like protein YjID, Replication-associated recombination protein A
GI-16	2754265	2772704	18439	12	putative type I restriction enzyme P M protein, EcoKI restriction-modification system protein HsdS, Hypothetical Protein (x3), ATP-dependent helicase HepA, 5-methylcytosine-specific restriction enzyme B, 5-methylcytosine-specific restriction enzyme subunit McrC, Type-1 restriction enzyme R protein, Integrase core domain protein, Transposase, Toxic protein SymE
GI-17	2837405	2842473	5068	8	Type II secretion system protein G precursor, Hypothetical Protein (x5), General secretion pathway protein K, General secretion pathway protein M
GI-18	3101647	3112481	10834	5	Hypothetical Protein (x5)
GI-19	3411526	3425833	14307	17	6-aminohexanoate-dimer hydrolase, HTH-type transcriptional regulator VirS, Hypothetical Protein (x13), Phytanoyl-CoA dioxygenase (PhyH), Ycf48-like protein precursor
GI-20	3431903	3438789	6886	6	Hypothetical Protein, Gamma-glutamylputrescine oxidoreductase, Phytanoyl-CoA dioxygenase (PhyH), putative DNA-binding transcriptional regulator, FAD-containing monooxygenase EthA, Putative aminoacylate hydrolase RutD
GI-21	3579713	3588705	8992	10	Alanine--tRNA ligase, Serine acetyltransferase, Cysteine synthase, putative adenyltransferase/sulfurtransferase MoeZ, Hypothetical Protein, putative inner membrane transporter YedA, Transposase DDE domain protein, putative HTH-type transcriptional regulator YjIR, putative oxidoreductase, Putative acetyl-hydrolase LipR precursor
GI-22	3941539	3954604	13065	13	Hypothetical Protein (x10), Lysozyme RrrD, O-acetyltransferase OatA, Prophage tail length tape measure protein
GI-23	3964377	3968753	4376	11	Phage portal protein, Hypothetical Protein (x7), Phage Terminase, small subunit, HNH endonuclease
GI-24	3988289	3993220	4931	13	Hypothetical Protein (x11), Siphovirus Gp157, LytTr DNA-binding domain protein
GI-25	4300185	4304881	4696	5	Hypothetical Protein (x3), thiamine biosynthesis protein ThiF, Helix-turn-helix domain protein
GI-26	4695760	4700811	5051	6	Pyocin-S2 immunity protein, Pyocin-S1, Hypothetical Protein (x2), Integrase core domain protein, Transposase
GI-27	4873196	4883713	10517	12	Virulence sensor protein BvgS precursor, NADH pyrophosphatase, Hypothetical Protein (x3), 1, 4-dihydroxy-2-naphthoyl-CoA hydrolase, IS2 transposase TnpB, Transposase (x3), Colicin-Ia, DSBA-like thioredoxin domain protein
GI-28	5047600	5055035	7435	6	Prophage CP4-57 integrase, Hypothetical Protein (x4), N-6 DNA Methylase
GI-29	5057139	5063286	6147	6	Hypothetical Protein (x4), T5orf172 domain protein, helix-turn-helix protein
GI-30	5078468	5083193	4725	6	Hypothetical Protein (x3), Transposase, Integrase core domain protein, Histidinol-phosphatase
GI-31	5085496	5105213	19717	14	Hypothetical Protein (x10), Chromosome partition protein Smc, TM2 domain protein, PAAR motif protein, VRR-NUC domain protein
GI-32	5296079	5300880	4801	3	ATP-dependent DNA helicase PcrA, Hypothetical Protein, RNA polymerase sigma factor
GI-33	5318785	5336645	17860	17	Hypothetical Protein (x3), mce related protein, Methionine import ATP-binding protein MetN, putative phospholipid ABC transporter permease protein MlaE, Nickel and cobalt resistance protein CnrA, Multidrug resistance protein MdtE precursor, HTH-type transcriptional regulator AcrR, putative efflux pump outer membrane protein TtgC precursor, Hca operon transcriptional activator, Quaternary ammonium compound-resistance protein SugE, HTH-type transcriptional regulator SyrM 1, Conjugal transfer protein TraG, Ribbon-helix-helix protein, copG family, Type IV secretion system protein PilH
GI-34	5632149	5637115	4966	8	Hypothetical Protein (x7), Replicative DNA helicase
GI-35	5677209	5684343	7134	10	Hypothetical Protein (x9), Putative SOS response-associated peptidase YedK
GI-36	5714944	5726091	11147	13	Hypothetical Protein (x6), Antitoxin ParD1, Toxin ParE1, Integrase core domain protein, Transposase, Putative helicase, site-specific tyrosine recombinase XerC, Sporulation initiation inhibitor protein Soj
GI-37	5726582	5731142	4560	7	Hypothetical Protein (x6), Replicative DNA helicase
GI-38	5735127	5745199	10072	13	Nucleoid-associated protein YejK, DNA adenine methylase, Arc-like DNA binding domain protein, ParB-like nuclease domain protein, Hypothetical Protein (x6), Phage regulatory protein Rha (Phage_pRha), Single-stranded DNA-binding protein, Helix-turn-helix domain protein
GI-39	5763025	5772960	9935	10	Hypothetical Protein (x3), Twitching motility protein, PilM, Integrase core domain protein (x2), Transposase, IS66 Orf2 like protein, Transposase IS66 family protein
GI-40	5784090	5793274	9184	9	Hypothetical Protein (x8), Transposase DDE domain protein
GI-41	6658315	6662965	4650	3	Hypothetical Protein (x3)
GI-42	7002470	7013629	11159	12	Hypothetical Protein (4), Putative transposase (4), Hydroxyacylglutathione hydrolase (x3), 2-hydroxyhexa-2,4-dienoate hydratase

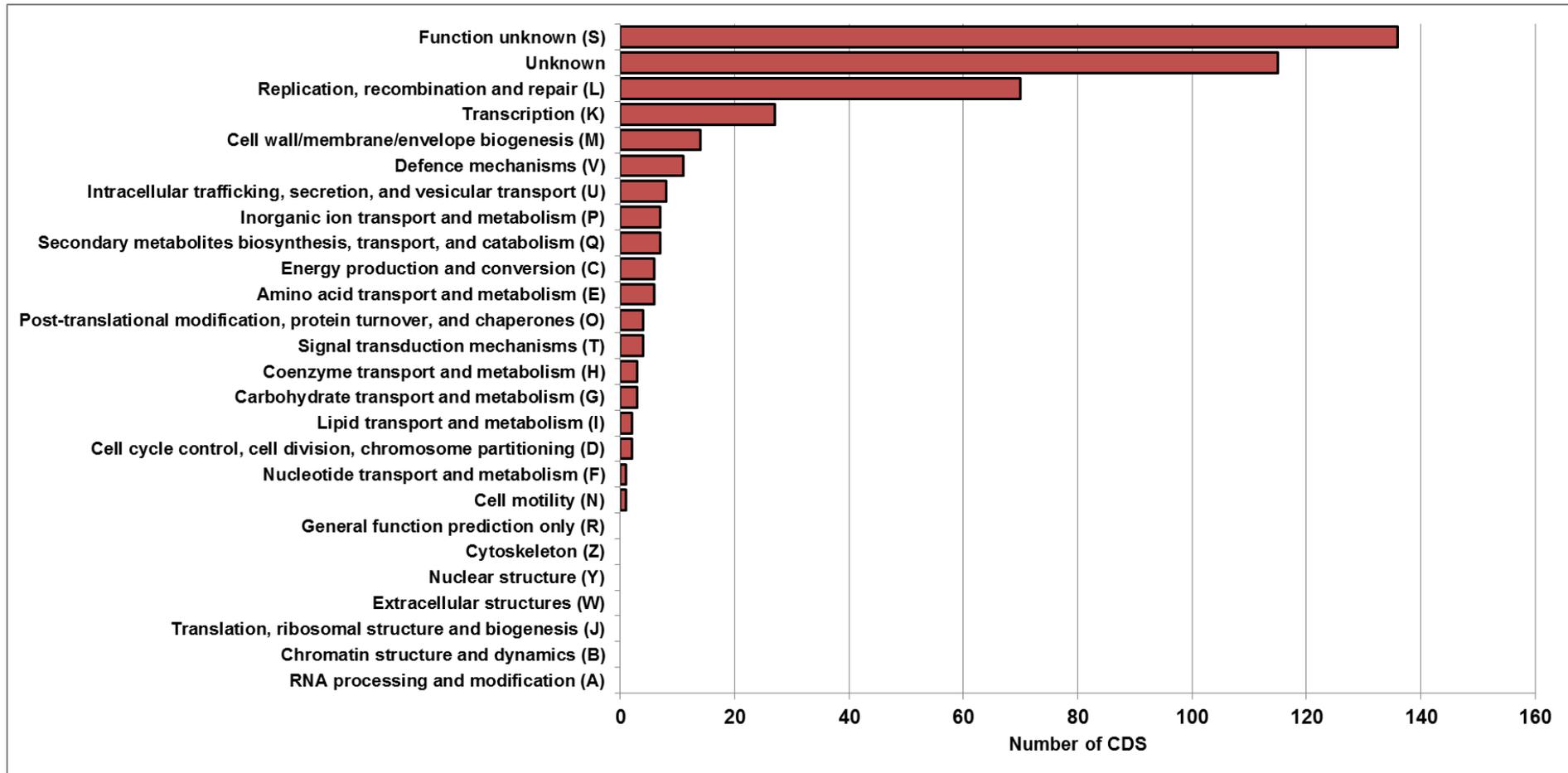


Figure 14. The COG categories of the RW109 main chromosome Genomic Island's CDS

Bar chart represents the numbers of the GI CDS within each COG category for the main chromosome. The poorly characterised COG categories S and unknown represented the highest proportion of GI CDS. Within the well characterised COG classes the four categories with the greatest number of GI CDS were L (70 CDS), K (26 CDS), M (14 CDS) and V (11 CDS). No GI CDS were found in categories R, Z, Y, W, J, B and A.

3.2.4.2 Genomic islands encoded on RW109 plasmid 1

The mean length of the six GIs identified in plasmid 1 was 14,574 bp and the mean number of CDS per GI was 15. GI-5 had a large number of CDS with the other GIs having between 5-19 CDS (Table 11). The number of assigned COG functional groups for the GI CDS of plasmid 1 were analysed and the amino acid transport and metabolism (E) had the highest number with 19 CDS. The E-COG category CDS were found in GIs 2, 4, 5 and 6; GI-4 had four out of its six CDS assigned to this category and GI-5 had the highest number of E-COG CDS (9). The poorly categorised COG group S had the second largest number with 17 CDS and K category had the third largest number with 13 CDS; both S and K COG CDS were distributed throughout the plasmid 1 GIs. Two CDS with efflux pump functions were found in GI-1 and these were identified as predicted BacMet resistance genes encoding an efflux pump membrane transporter BepE (RW109_00013) and putative efflux pump outer membrane protein TtgC precursor (RW109_00014). These two genes were also associated with the KEGG module M00698, the multidrug resistance efflux pump BpeEF-OprC (Section 3.2.3.2.2). The other three genes in the M00698 module (RW109_00010, RW109_00012 and RW109_00017) (Table 8) were also identified within GI-1 of plasmid 1 suggesting this efflux pump is part of a genomic island. No other BacMet resistance genes were found in the GI's of plasmid 1. GI-6 also had two threonine efflux proteins (RW109_00569 and RW109_00570) and there were seven CDS with transporter functions were also identified throughout the GIs with one in GI-1, and two in GI-2, GI-4 and GI-5.

3.2.4.3 Genomic islands encoded on plasmid 2

Plasmid 2 was found to have one GI with a length of 35,019 bp with forty CDS; 55% of the CDS were annotated as hypothetical proteins (Table 11). When the COG categories of the GI were analysed the majority were within the poorly characterised group (67.50%). 30% were assigned to the replication, recombination and repair (L) category and these CDS were annotated as transposase and integrase proteins. One CDS was assigned to the lipid transport and metabolism (I) category (2.5% of total number of CDS) and was an outer membrane protein P1 precursor.

Table 11. Genomic Islands encoded on the RW109 plasmids

Genomic Island	Start	End	Length (bp)	Number of CDS	Annotation
Plasmid 1					
GI-1	9313	30288	20975	19	Tniq, Bacterial TniB Protein, Transposon Tn7 Transposition Protein Tnsb, Aerotaxis Receptor, HTH-Type Transcriptional Regulator Dmlr(X2), Transposase DDE Domain Protein, Multidrug Resistance Protein MdtA Precursor, Efflux Pump Membrane Transporter BepE, Putative Efflux Pump Outer Membrane Protein TtgC Precursor, NADH Oxidase, HTH-Type Transcriptional Regulator LeuO, Glutathione S-Transferase GST-6.0, DNA-Invertase Hin, Hypothetical Protein (X2), PIN Domain Protein, Tn3 Transposase DDE Domain Protein, Mercuric Transport Protein Periplasmic Component Precursor
GI-2	102388	111297	8909	11	Putative M18 Family Aminopeptidase 2, Phosphoserine Phosphatase, Cytochrome C-552 Precursor (X2), Aralkylamine Dehydrogenase Light Chain Precursor, Methylamine Utilization Protein MauD, Methylamine Utilization Protein MauE, Aralkylamine Dehydrogenase Heavy Chain Precursor, Hypothetical Protein, Sulfate Transport System Permease Protein CysW, Putative 2-Aminoethylphosphonate Transport System Permease Protein PhnV
GI-3	171629	176324	4695	5	Hypothetical Protein(X3), Transposase DDE Domain Protein (X2)
GI-4	328383	332688	4305	6	Leucine-Responsive Regulatory Protein, Cystine-Binding Periplasmic Protein Precursor, Inner Membrane Amino-Acid ABC Transporter Permease Protein YecS, Putative Glutamine ABC Transporter Permease Protein GlnM, Octopine Permease ATP-Binding Protein P, Hypothetical Protein
GI-5	425648	461677	36029	36	Hypothetical Protein (X7), HTH-Type Transcriptional Regulator PuuR, Glutamine Synthetase 3, Amidophosphoribosyltransferase Precursor, GXGXG Motif Protein, Glutamate Synthase [NADPH] Large Chain, Ammonia Channel Precursor, Phenoxybenzoate Dioxygenase Subunit Beta, Aminomethyltransferase (X2), Quaternary Ammonium Compound-Resistance Protein SugE, GMP Synthase [Glutamine-Hydrolyzing], Betaine Aldehyde Dehydrogenase, Ammonia Channel, Methyl-Accepting Chemotaxis Protein PctB, HTH-Type Transcriptional Regulator LutR, Eama-Like Transporter Family Protein, Sialic Acid TRAP Transporter Permease Protein SiatT, LysR Substrate Binding Domain Protein, Imelysin, Cytochrome C, Variant SH3 Domain Protein, Pca Regulon Regulatory Protein, Sarcosine Oxidase, Gamma Subunit Family, Delta Subunit Family, Sarcosine Oxidase Subunit Beta, Bifunctional Protein Fold Protein, Formyltetrahydrofolate Deformylase, 4-Methylaminobutanoate Oxidase (Formaldehyde-Forming), Acetate Operon Repressor
GI-6	528619	541150	12531	11	Transcriptional Regulator AcuR, Alcohol Dehydrogenase, Chromosome-Partitioning Protein Spo0J, Hypothetical Protein (X3), Bifunctional Transcriptional Activator/DNA Repair Enzyme Ada, Threonine Efflux Protein (X2), Glycine Cleavage System Transcriptional Activator, Putative Prophage CPS-53 Integrase
Plasmid 2					
GI-1	122949	35019	35019	40	Site-Specific Tyrosine Recombinase XerC, Abi-Like Protein, Hypothetical Protein (X22), Transposon Tn7 Transposition Protein TnsA, Transposon Tn7 Transposition Protein TnsB, Transposase DDE Domain Protein, Transposon Tn7 Transposition Protein TnsC, Integrase Core Domain Protein(X3), Transposase (X2), ATP-Dependent Helicase HepA, Replicative DNA Helicase, Outer Membrane Protein P1 Precursor, Transposase IS66 Family Protein, IS66 Orf2 Like Protein, Antitoxin DinJ, mRNA Interferase YafQ

3.2.5 Identifying phages in the RW109 genome

The PHASTER tool was used to identify phage sequences in the RW109 genome (Table 12). Eight phages were identified in the main chromosome; five were classified as intact, one was incomplete and two were questionable. One phage was identified in each of the two plasmids and these were both characterised as incomplete. On the circular replicon maps (Figure 11), the phage sequences were highlighted in light pink on the forth track in from the outside of the map. Phage's 3, 4 and 6 had overlaps with the GIs (see Figure 11). A similar overlap with annotated GIs was observed for the phage's identified in plasmids 1 and 2 (Figure 11). Out of the phages on the main chromosome, six were characterised as *Pseudomonas* phages, phage 5 was classed as a *Salmonella* phage and phage 8 was an *Enterobacteria*. The phage sequence in plasmid 1 was identified as a *Shigella* phage and the phage in plasmid 2 was classified as a *Pseudomonas* phage.

Table 12. Phages within the RW109 genome

Phage Number	Region Length (bp)	Region position	Total number CDS in region	GC %	PHASTER Score	Phage Completeness	Phage sequence most similar to	Phage NC Number
Plasmid 1								
1	39600	152692-192335	15	57.25%	60	Incomplete	<i>Shigella_Sf6</i>	NC_005344(2)
Main Chromosome								
1	41900	670851-712805	60	63.91%	130	Intact	<i>Pseudomonas_vB_PM105</i>	NC_028667(40)
2	18400	724969-743443	25	65.19%	100	Intact	<i>Pseudomonas_JBD44</i>	NC_030929(10)
3	66200	802521-868726	65	61.05%	150	Intact	<i>Pseudomonas_vB_PaeS</i>	NC_028667(42)
4	47800	1371811-1419673	63	61.12%	109	Intact	<i>Pseudomonas_F10</i>	NC_007805(46)
5	22800	1464864-1487750	7	61.55%	20	Incomplete	<i>Salmonella_SPN1S</i>	NC_016761(2)
6	72900	3923988-3996972	96	58.92%	150	Intact	<i>Pseudomonas_PMG1</i>	NC_016765(50)
7	18000	4637833-4655921	11	61.20%	85	Questionable	<i>Pseudomonas_Pf1</i>	NC_001331(8)
8	12900	5599645-5612594	16	60.32%	70	Questionable	<i>Enterobacteria_HK022</i>	NC_002166(3)
Plasmid 2								
1	39300	106946-146332	19	58.27%	50	Incomplete	<i>Pseudomonas_F116</i>	NC_006552(2)

3.2.6 Comparative genomics of the RW109 whole genome sequence

3.2.6.1 Comparing RW109 annotated functional categories with PA14 and PAO1

3.2.6.1.1 COG functional annotation comparisons

The COG functional annotation of predicted CDS of PA14 and PAO1 was carried out and compared to the COG assignment of the RW109 whole genome. Figure 15 shows a bar chart of the number of CDS assigned to the well characterised COG categories for the three strains. The poorly characterised categories were not shown as a high number of CDS for each strain were assigned to this group when compared to the well-characterised COG categories. Within the poorly characterised COG functional class, RW109 had 2,952 CDS (40.42% of its genome), PA14 1,999 CDS (33.86% of its genome) and PAO1 1845 CDS (32.53% of its genome).

Out of the twenty well characterised COG categories (ignoring categories W, Y and Z which had no CDS assigned for the three strains), RW109 had a greater number of CDS in seventeen categories when compared to PA14 and PAO1. In the replication, recombination and repair category (L) RW109 had 96.34% more CDS than PA14 and 123.21% more than PAO1. RW109 was also found to have 22.29% extra CDS than PA14 and 42.22% more CDS than PAO1 in the intracellular trafficking, secretion, and vesicular transport (U). The other categories where RW109 was found to have 10% or more CDS when compared to PA14 and PAO1 included cell cycle control, cell division, chromosome partitioning (D), defence mechanisms (V) and transcription (K). There was a less than 1% difference in the number of CDS for the nucleotide transport and metabolism (F), chromatin structure and dynamics (B) and translation ribosomal structure and biogenesis (J) categories when RW109 was compared to PA14 and PAO1.

The percentage of CDS within each COG category for the RW109, PA14 and PAO1 strains were also compared and were found to be similar throughout the three genomes (Figure 16); the category where there was a noticeable difference in the percentage of CDS was the replication, recombination and repair (L) category. RW109 had 5.13% of its total CDS assigned to the L COG category, whereas PA14 had 3.24% CDS and PAO1 2.96% CDS with L COG annotation.

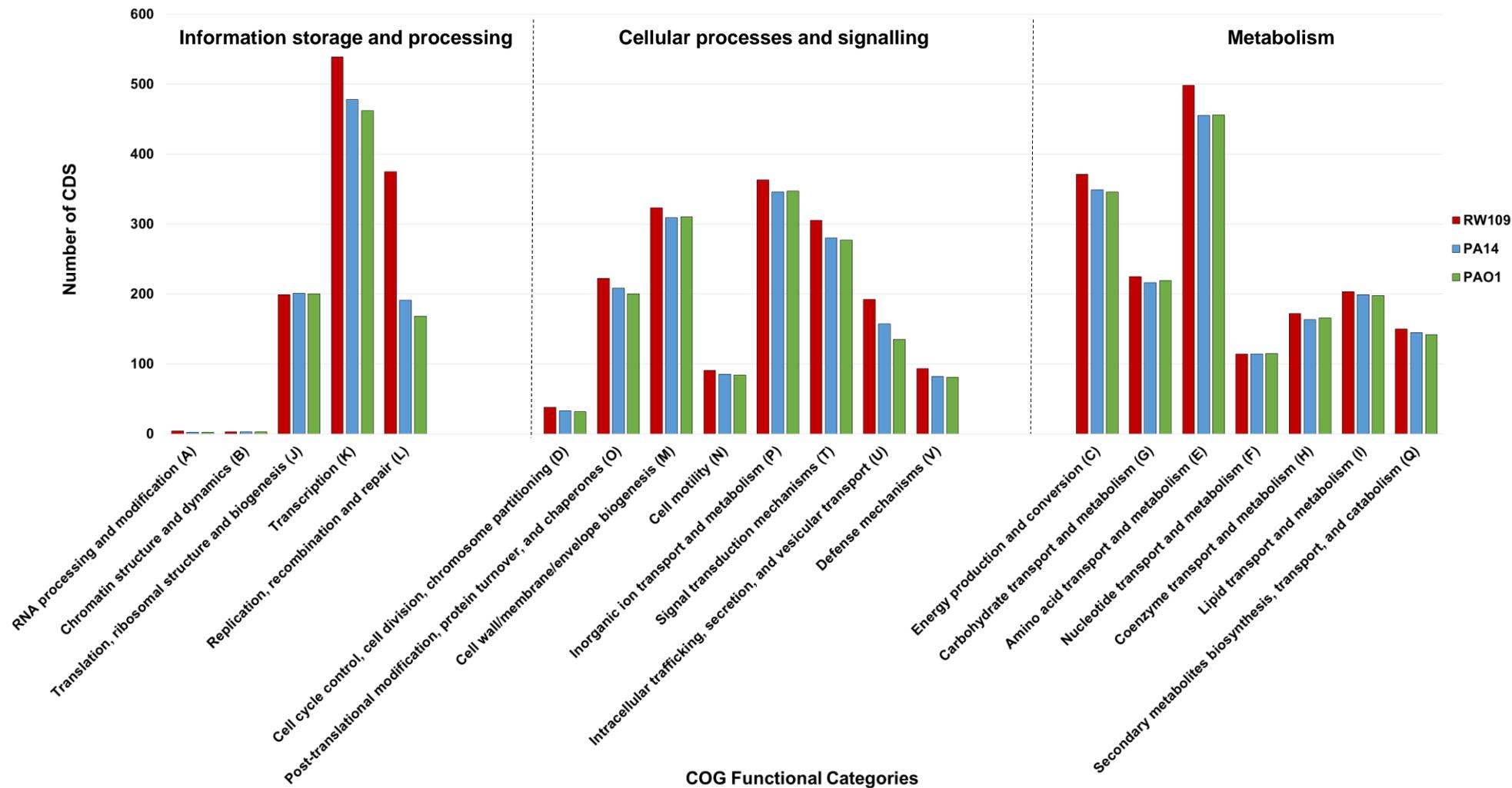


Figure 15. Comparison of the number of CDS within the well characterised COG categories for RW109, PA14 and PAO1

The bar chart shows the number of CDS which were assigned to the well characterised COG categories for RW109 (red), PA14 (blue) and PAO1 (green). The poorly characterised COG categories were not shown as a high number of CDS were assigned to this group.

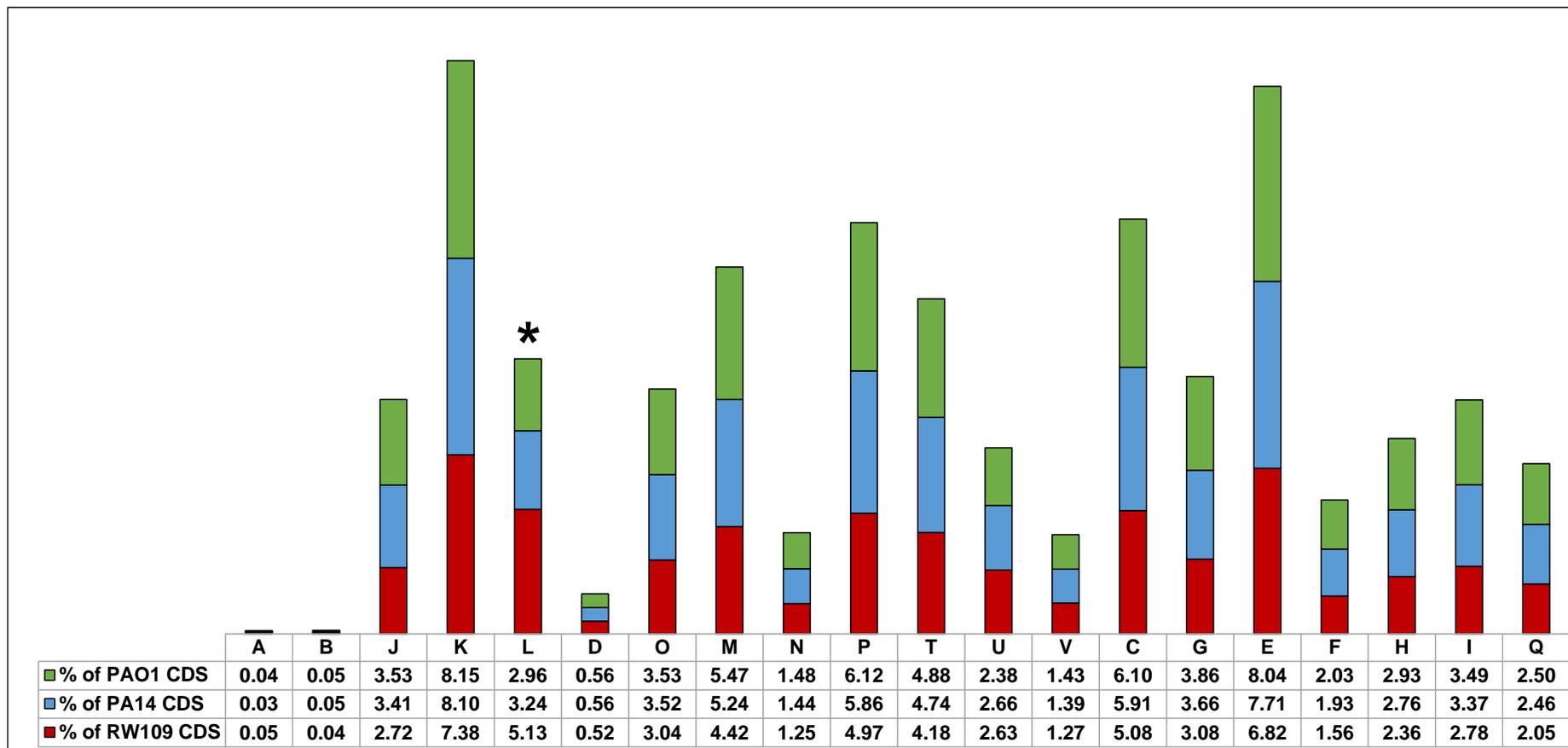


Figure 16. Percentage number of CDS within each COG category for RW109, PA14 and PAO1

The percentage number of CDS within each COG category when compared to the size of the genome was calculated for RW109 (red), UCBPP-PA14 (blue) and PAO1 (green). The stacked column graph demonstrates the percentage of CDS for each of the COG categories across the three strains. The asterisk above the column for L indicates the COG category where RW109 was found to have a noticeably greater percentage number of CDS when compared to PA14 and PAO1.

3.2.6.1.2 KEGG functional modules and MAPLE MRCs comparisons

The KEGG MAPLE MRCs and Q-values for PAO1 and PA14 were calculated and compared against those identified as biologically achievable across the whole genome sequence of RW109 (combining the annotations of the main chromosome and plasmids 1 and 2) from Section 3.2.3.2. Table 13 demonstrates the KEGG functional modules, which differed when RW109 was compared to PA14 and PAO1. RW109 contained two modules which were not biologically feasible in both PA14 and PAO1 (MCRs with Q-values of >0.5). These were a UhpB-UhpA (hexose phosphates uptake) two-component regulatory system (M00473) and a multidrug resistance, efflux pump BpeEF-OprC (M00698) which was found on GI-1 of plasmid 1. Another two RW109 modules, which were not biologically feasible in PA14 but were in PAO1, included a lipopolysaccharide metabolism module (M00064- ADP-L-glycero-D-manno-heptose biosynthesis) and an AauS-AauR (acidic amino acids utilisation) two-component regulatory system (M00653). In PA14, two modules were found to be not biologically feasible in RW109 and PAO1; these were involved in aromatic anthranilate degradation (M00637) and nitrate assimilation (M00615). A central carbohydrate metabolism module involved in 5-phospho- α -ribosyl-1-pyrophosphate (PRPP) biosynthesis (M00005) was found to be biologically achievable in both PA14 and PAO1 but not in RW109.

Table 13. KEGG functional modules which differ when MCRs and Q-values of RW109 were compared to PA14 and PAO1

KEGG Functional Module	Module Type	Module number	Module description	RW109		PA14		PAO1	
				MRC (%)	Q-Value	MRC (%)	Q-Value	MRC (%)	Q-Value
Two-component regulatory system	Functional	M00473	UhpB-UhpA (hexose phosphates uptake) two-component regulatory system	100	0	50	0.667	50	0.667
Drug resistance	Signature	M00698	Multidrug resistance, efflux pump BpeEF-OprC	100	0	50	0.933	50	0.8
Two-component regulatory system	Functional	M00653	AauS-AauR (acidic amino acids utilization) two-component regulatory system	100	0	50	0.667	100	0
Lipopolysaccharide metabolism	Pathway	M00064	ADP-L-glycero-D-manno-heptose biosynthesis	100	0	75	0.533	100	0
Central carbohydrate metabolism	Pathway	M00005	PRPP biosynthesis, ribose 5P => PRPP	0	1	100	0	100	0
Aromatics degradation	Pathway	M00637	Anthranilate degradation, anthranilate => catechol	0	0.571	100	0	0	0.571
Metabolic capacity	Signature	M00615	Nitrate assimilation	33.3	0.545	66.7	0.396	33.3	0.545

Footnotes: KEGG functional modules were considered biologically achievable when they were assigned a MCR with a Q-value of < 0.05.

3.2.6.3.3 Antimicrobial and virulence functional annotation comparisons

3.2.6.3.3.1 ABRicate analysis of PA14 and PAO1

The antimicrobial resistance genes of PA14 and PAO1 were identified using ABRicate; thirty-eight CDS were associated with antimicrobial resistance in both PA14 and PAO1 which was the same number as identified in RW109 (see Section 3.2.3.3). When the resistance associated CDS for the three strains were compared, RW109 and PAO1 were found to obtain the same set of ABRicate CDS whereas minor differences were seen with PA14. A resistance gene called *PDC-1*, which is a *P. aeruginosa* beta-lactamase (Table 9) was found in RW109 and PAO1 but not in PA14. A resistance gene called *PDC-9* was identified in PA14 but not in RW109 and PAO1, this gene is an extended-spectrum beta-lactamase found in *P. aeruginosa* (McArthur et al., 2013).

PA14 was found to have less predicted virulence genes, 235 CDS, compared to 239 in RW109 (Section 3.2.3.3). However, more predicted virulence genes were observed in PAO1 246 CDS (Figure 17A). RW109 and PAO1 were found to have eleven virulence genes that were not found in PA14 with eight of these being associated with Type IV fimbrial biogenesis. Five virulence genes were found in PA14 and PAO1 but not in RW109 and their function was related to pyoverdine biosynthesis. Two type III secretion system virulence genes were found in PA14, which were not identified in both PAO1 and RW109 (Figure 17A). In PAO1, lipopolysaccharide biosynthesis virulence genes were observed that were not found in RW109 and PA14 (Figure 17A).

3.2.6.3.3.2 BacMet analysis of PA14 and PAO1

The amino acid sequences of PA14 and PAO1 were searched against the BacMet database; sixty PA14 CDS and fifty-six PAO1 CDS were identified as being antibacterial biocide and metal resistance genes. Figure 17B shows the BacMet resistance genes which differed when the sixty-three predicted RW109 BacMet CDS were compared to those identified in PA14 and PAO1. RW109 had nine BacMet associated genes, which were not found in PA14 and PAO1; four were part of efflux pump systems, two were involved in mercury resistance, two had functions in resistance to nickel, zinc, cadmium and cobalt and one was a tellurium resistance protein. There were six BacMet CDS which were identified in both RW109 and PA14 but not in PAO1 and two found in both RW109 and PAO1 but not in PA14 (Figure 17B). PA14 and PAO1 also both contained a BacMet resistance gene (BAC0513) that was not found in RW109, which functioned as multidrug resistance protein (vmeF). Two BacMet resistance genes were also found in PA14 (BAC0067 and BAC0549) but not in RW109 or PAO1; one was involved in the tolerance to copper and the other was a nickel, cobalt and cadmium resistance protein.

A)

VFDB ID	Gene name	Virulence information from VFDB	RW109	PA14	PAO1
NP_248771	<i>fha1</i>	Fha domain-containing protein; type VI secretion system protein, protein secretion/export apparatus			
NP_252530	<i>exoS</i>	Exoenzyme S; type III Secretion system protein, secreted Factors (toxins, enzymes, alginate)			
NP_253217	<i>pilC</i>	Pilin biogenesis protein; protein secretion by the type II secretion system, motility and attachment			
NP_253239	<i>fimT</i>	Type IV fimbrial biogenesis protein; protein secretion, motility and attachment			
NP_253240	<i>fimU</i>	Type IV fimbrial biogenesis protein; protein secretion, motility and attachment			
NP_253241	<i>pilV</i>	Type IV fimbrial biogenesis protein; pilus biogenesis, motility and attachment			
NP_253242	<i>pilW</i>	Type IV fimbrial biogenesis protein; pilus biogenesis, motility and attachment			
NP_253243	<i>pilX</i>	Type IV fimbrial biogenesis protein; type IV pilus-dependent motility, motility and attachment			
NP_253244	<i>pilY1</i>	Type IV fimbrial biogenesis protein; pilus biogenesis, motility and attachment			
NP_253245	<i>pilY2</i>	Type IV fimbrial biogenesis protein; pilus biogenesis, motility and attachment			
NP_253246	<i>pilE</i>	Type IV fimbrial biogenesis protein; type IV pilus-dependent motility, motility and attachment			
NP_251092	<i>pvdI</i>	Pyoverdine biosynthetic process; pathogenesis, regulation of other virulence factors and biofilm formation			
NP_251090	<i>pvdJ</i>	Pyoverdine biosynthetic process; pathogenesis, regulation of other virulence factors and biofilm formation			
NP_251089	<i>pvdD</i>	Pyoverdine biosynthetic process; pathogenesis, regulation of other virulence factors and biofilm formation			
NP_251088	<i>fpvA</i>	Ferripyoverdine receptor; pyoverdine biosynthetic process; pathogenesis, regulation of other virulence factors and biofilm formation			
NP_251087	<i>pvdE</i>	Pyoverdine biosynthetic process; pathogenesis, regulation of other virulence factors and biofilm formation			
NP_250386	<i>pscP</i>	Translocation protein in type III secretion; protein secretion/export apparatus			
AAC16023	<i>exoU</i>	Type III secretory toxin; secreted Factors (toxins, enzymes, alginate)			
NP_251844	<i>wzy</i>	Lipopolysaccharide biosynthetic process; membrane protein; adherence, endotoxin;			
NP_251850	<i>wzz</i>	Lipopolysaccharide biosynthetic process; cell wall, LPS, capsule			

Absent

Present

B)

BacMet ID	Gene name	Experimentally verified BacMet gene information	RW109	PA14	PAO1
BAC0015*	<i>adeG</i>	Part of the AdeFGH efflux system; contributes to resistance of antibacterial biocides such as dyes, anionic detergents			
BAC0016*	<i>adeH</i>	Part of the AdeFGH efflux system; contributes to resistance of antibacterial biocides such as dyes, anionic detergents			
BAC0046*	<i>bepG</i>	Efflux pump membrane transporter BepG; Contributes to resistance of Deoxycholate, Sodium dodecyl sulphate and Nalidixic acid			
BAC0119	<i>czcA</i>	Cobalt-zinc-cadmium resistance protein			
BAC0203	<i>cnrA</i>	Nickel and cobalt resistance protein			
BAC0388*	<i>terC</i>	Tellurium resistance protein; membrane protein which may provide a barrier to Tellurium ion entry			
BAC0425	<i>vexD</i>	Transporter, AcrBDF efflux system; contributes to resistance of Ethidium bromide and Deoxycholate			
BAC0650*	<i>merA</i>	Mercuric reductase; contributes to resistance of Mercury and Phenylmercury acetate			
BAC0653*	<i>merA</i>	Mercuric ion reductase; contributes to resistance of Mercury			
BAC0263	<i>ncrA</i>	Part of ncrABC nickel resistance chromosomal determinant system; encodes a membrane protein that forms an efflux with ncrC			
BAC0267	<i>nczA</i>	Nickel-cobalt-zinc resistance protein; part of the nczCBA operon			
BAC0526	<i>vmeV</i>	Transporter, AcrBDF efflux system; contributes to resistance of Sodium dodecyl sulphate, Sodium cholate, Sodium deoxycholate, Sodium glycocholate, Ethidium bromide, Benzylkonium chloride, Chlorhexidine, Acriflavine and Tetraphenylphosphonium			
BAC0570	<i>actP</i>	Acetate Permease; contributes to resistance to Tellurium			
BAC0652*	<i>merA</i>	Mercury reductase enzyme; contributes to resistance to Mercury			
BAC0698	<i>ncrA</i>	Part of the ncrABC determinant system; contributes to Nickel and Cobalt resistance			
BAC0044	<i>bepE</i>	Efflux pump membrane transporter; involved in resistance to Sodium deoxycholate, Ethidium bromide and Crystal violet			
BAC0130	<i>czrA</i>	Part of the czrSRCBA resistance operon; contributes to resistance to Zinc and Cadmium			
BAC0513	<i>vmeF</i>	Putative multidrug resistance protein; contributes to the resistance of Ethidium bromide, Sodium dodecyl sulphate, Sodium cholate, Sodium deoxycholate, Sodium glycocholate and Acriflavine			
BAC0067	<i>cinA</i>	Copper-containing azurin-like protein; involved in Copper tolerance			
BAC0549	<i>nccA</i>	Nickel-cobalt-cadmium resistance protein			

Figure 17. Virulence and BacMet predicted genes which differed between RW109, PA14 and PAO1

Comparison of RW109, UCBPP-PA14 and PAO1 showed differences in the identification of genes associated with virulence (A) and biocide and metal resistance (B). The figure shows the presence (dark blue) and absence (light blue) of these genes when the three strains were compared. The database accession number, gene name and information about the gene were shown. In panel B) the * illustrates the BacMet associated RW109 CDS which were found on plasmid 1.

3.2.6.2 Core-genome comparison of RW109 with 102 *P. aeruginosa* genome sequences

Pan-genome (Roary) analysis of 102 *P. aeruginosa* selected genomes (Chapter 2, Table 3) along with the RW109 genome sequence identified 4,009 core genes; a 95% identity cut-off was applied and core genes were defined as those in 99-100% of the genome sequences. A further 222 genes were shared by 95-99% of isolates and were termed soft-core genes. The genome sequences used in this analysis represented 66 strains from a clinical setting, 19 environmental strains, 16 industrial strains, one strain classified as both clinical and environmental (PA1RG) and one strain with an unknown isolation source (ATCC-13388) (Chapter 2, Table 3). The alignment of the core-genes was used to construct a maximum-likelihood tree of the 103 *P. aeruginosa* sequences as shown in Figure 18.

The tree was rooted with the PA7 genome sequence as this strain was highly divergent when compared to the other sequences (Figure 18). The *P. aeruginosa* sequences were colour coded depending on whether they were clinical, environmental, industrial strains or unknown and genome sequences for each type of strain were distributed relatively equally throughout the tree. The genome sequences of the *P. aeruginosa* were split into 2 major clades, with 13 industrial positioned in clade 1 and 3 located within clade 2 (Figure 18). The sub-clade which contained RW109 was highlighted by a light purple box; within this sub-clade were six industrial strains (including RW109) and six clinical strains; there were no environmental type strains identified within this sub-clade. RW109 was found to be most closely related to the industrial RW192 strain, which was isolated from a contaminated Unilever surface cleaner product in 2012 (Chapter 2, Table 3). The other industrial strains in this sub-clade included RW138 (isolated from liquid abrasive cleaner product), RW149 (isolated from personal care product), RW176 (isolated from dishwasher liquid) and RW200 (isolated from a timber care product); this sub-clade had the highest number of industrially isolated strains when compared to others in the tree. The clinical strains within this sub-clade included Carb01-63 (isolation unknown), F30658 (isolation unknown), FRD1 (isolated from CF sputum), F23197, (isolation unknown), RW27 (isolated from CF sputum) and RW30 (isolated from CF sputum) (Figure 18).

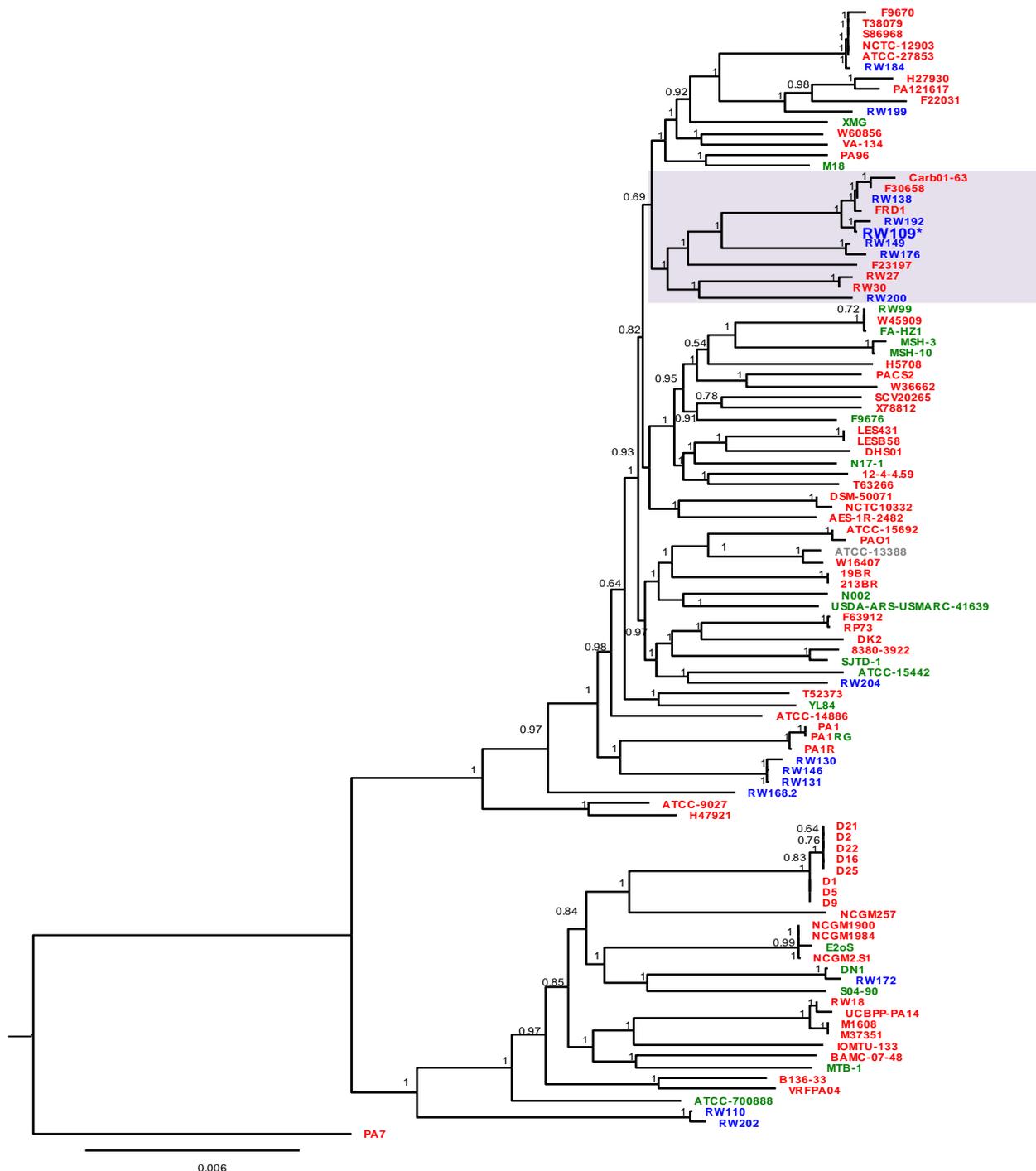


Figure 18. Core-gene phylogenetic tree of 103 *P. aeruginosa* and RW109 genome sequences

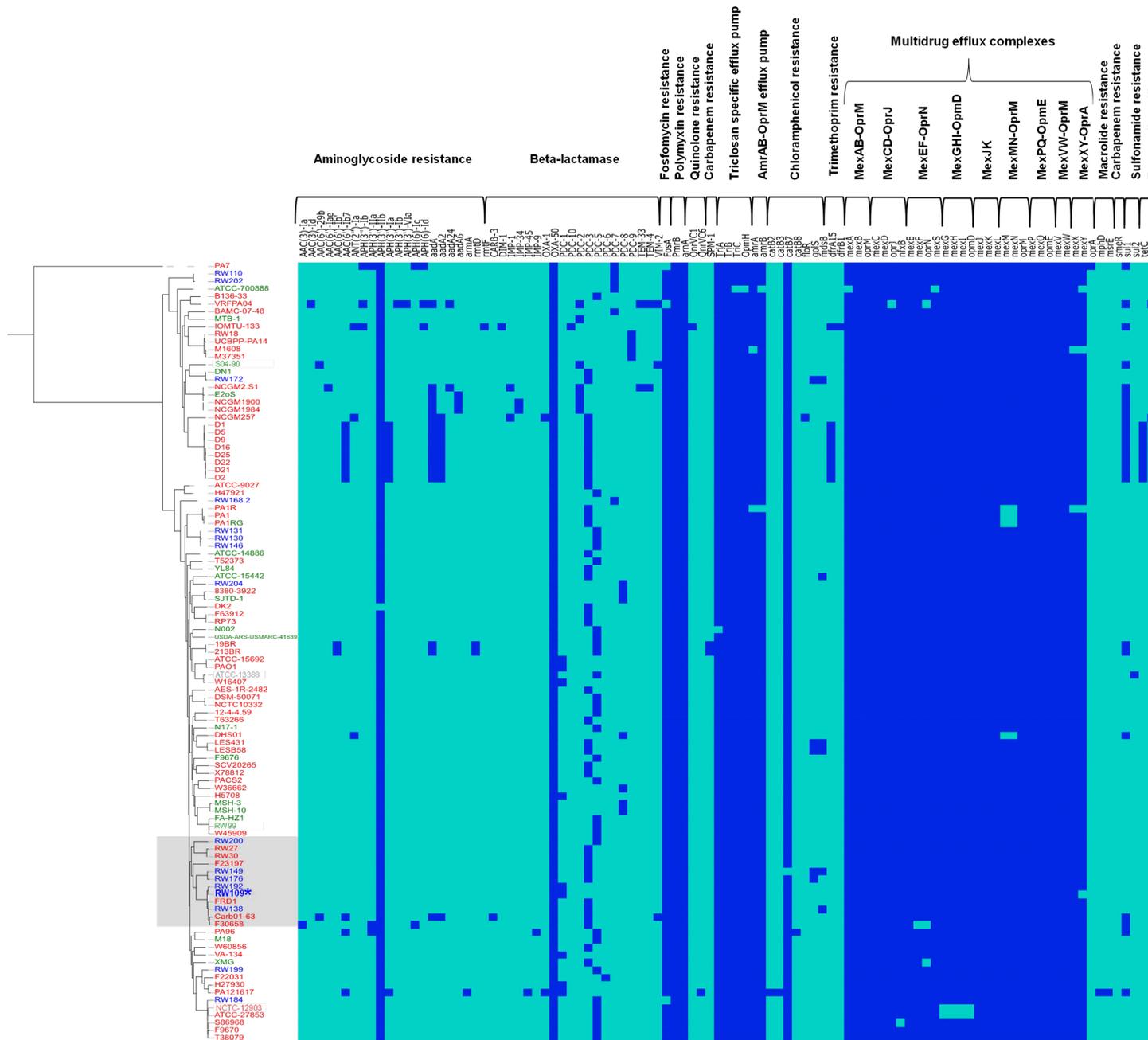
The phylogenetic tree was generated based on the Roary core-genome generated from the 103 *P. aeruginosa* genome sequences and RW109. The strain names of the genome sequences were colour coded to indicate if the strain was clinical (red), environmental (green) or industrial (blue) (the PA1RG strain was classified as both clinical and environmental so was coloured red and green). The strain with an unknown source of isolation is shown in grey (ATCC-13388). The numbers on the nodes represent the bootstrap values with 1 being 100% and the tree was drawn to scale with genetic distances given as the number of base substitutions per site as shown by the scale bar. The tree was rooted with the PA7 genome sequence as it was highly divergent and was ordered based on the increasing nodes. The sub-clade highlighted by a light purple square towards the top of the tree contains the RW109 strain which is in bold and has an asterisks next to it.

3.2.6.3 Presence and absence of CARD predicted antibiotic resistance genes in RW109 and the 102 *P. aeruginosa* genome sequences

The antibiotic resistance genes present in the 103 *P. aeruginosa* sequences were determined by CARD and visualised as a heat map presence/absence annotation aligned to a core-genome alignment phylogenetic tree (Figure 19; the core-genome alignment was not ordered based on increasing nodes as with the tree in Figure 18). The antimicrobial resistance genes were grouped according to their predicted function. Overall, the clinical, environmental and industrial strain sequence types had similar CARD resistance gene profiles. The mean number of CARD resistance genes was determined as 43 for clinical strains, 41 for environmental strains, and 41 for industrial strains (the strain classified as both clinical and environmental had 39 resistance genes and the strain with an unknown isolation source had 42). The highest numbers of antimicrobial resistance genes were found in the two clinical strains PA121617 (isolated from CF sputum) and VRFPA04 (isolated from a corneal button of a patient with corneal keratitis); both strains had 50 CARD predicted resistance genes. Strains PA1R (a clinical strain isolated from a respiratory tract infection) and E2oS (a soil environmental strain) had the lowest number of predicted resistance genes with 35 identified.

All 103 genome sequences had the polymixin resistance genes *PmrB* and *ArnA* and the aminoglycoside resistance gene *OXA-50*. All but one strain, RW149 (isolated from a personal care product), had the chloramphenicol resistance gene *catB7* and all strains except for DK2 (isolated from CF sputum) had the aminoglycoside *APH(3')-Iib* gene. The four genes involved in the triclosan specific efflux pump were present in all strains except for the environmental strain N002 (isolated from crude oil contaminated soil) which did not have *TriA* and the environmental strain ATCC-700888 (isolated from an industrial water system) where the *TriC* and *OpmH* genes were not present. An industrial strain RW184 (isolated from dishwasher liquid) and two clinical strains VRFPA04 and PA7 (non-respiratory clinical isolate) were the only strains not to have the fosfomycin resistance gene *FosA*. The majority (101 of 103) *P. aeruginosa* sequences encoded a number of genes predicted to be Mex-type multidrug efflux complexes. Two clinical strains NCTC-12903 (an antibiotic efficacy testing reference strain) and ATCC-15692 (isolated from an infected wound) were the only two, which did not have the genes for the MexGHI-OpmD efflux pump. The clinical strain PA7 was the only strain to encode a predicted *oprA* gene from the MexXY-OprA efflux pump.

Within the industrial genome sequences, RW109 and RW192 were the only two with the beta-lactamase *PDC-1* gene; this gene was also found within seven clinical strains (ATCC-15692, H27930, H5708, PA121617, PAO1, VA-134, and W16407). The majority of the genes, which encode the Mex-type efflux pump systems, were found in all 16 industrial strains. The exceptions were *oprA* which was not predicted in any of the industrial strain genomes, and the *mexY* gene which was absent in RW109. The sulfonamide resistance gene *sul2* was found only in the ATCC-13388 genome sequence, which had an unknown isolation source and is used in industrial testing. For the industrial strain sequences, CARD did not identify any antimicrobial genes within the quinolone, carbapenem, trimethoprim, macrolide and tetracycline efflux resistance functional groups, and this was true for the environmental strains.



←SEE PREVIOUS PAGE

Figure 19. The presence and absence of CARD antimicrobial genes across 103 *P. aeruginosa* sequences

The phylogenetic tree on the left of the figure was annotated with a heat map to illustrate the presence/absence of antimicrobial resistance genes from the CARD database for a panel of 103 *P. aeruginosa* genome sequences of clinical (red), environmental (green) and industrial (blue) strains. The strain with an unknown isolation source is shown in grey. The CARD antimicrobial genes were grouped via their resistance function on the top of the heat map; light blue shows its absence of an antimicrobial genes and dark blue indicates the presence. The sub-clade, which contains RW109, is highlighted with a light purple box and RW109 is indicated in bold with an asterisk next to it.

3.2.6.4 The unique genes of RW109

The gene presence and absence output from the pan-genome analysis of the 103 *P. aeruginosa* sequences was used to determine genes, which were unique to RW109. In total, 329 genes were found to be specific to RW109 with 105 encoded on plasmid 1, 202 on the main chromosome and 22 encoded on plasmid 2. No predicted antimicrobial or virulence genes (Sections 3.2.3.3) were found within the unique subset of RW109 genes. The COG annotation assessment identified 51.06% of the RW109 unique genes as poorly characterised.

The well-characterised COG categories with the highest number of unique RW109 genes were replication, recombination and repair (L) with 23.10%, followed by amino acid transport and metabolism (E) with 7.90%, transcription (K) with 4.86% and energy production and conversion (C) with 4.56%. The KEGG modules identified as biologically feasible (Section 3.2.3.2) were analysed and the CDS required for a phosphate and amino acid transport system module (M00236: RW109_00304, RW109_00306 and RW109_00307) and a pyrimidine metabolism module (M0046: RW109_00299) from plasmid 1 were found within the unique genes of RW109 (Table 8). One out of the two genes required for a plasmid 1 encoded cofactor and vitamin biosynthesis module (M00140- RW109_00291) was also found within the unique genes.

Genomic islands (GI's) genes (Section 3.2.4) were examined and more than 70% of the genes in GI's 9, 10, 13, 28, 29 and 42 of the main chromosome (Table 10) were associated with the unique RW109 genes; all of the genes in GI-13 and GI-42 were identified as unique. For plasmid 1, all genes found in GI's 2 (RW109_00089 - RW109_00099) and 4 (RW109_00303 - RW109_00308) were also within the unique RW109 gene set (Table 11). 16 out of the 40 CDS, which make up the one GI in plasmid 3, were also identified as unique RW109 genes. The RW109 PHASTER sequences predicted in Section 3.2.5 (Table 12) were also examined and seventy-four phage CDS were identified as unique genes; within the main chromosome four were identified in phage 1, 28 in phage 3, three in phage 4, two in phage 5, 27 in phage 6 and one in phage 8. Two unique genes were identified in the one phage sequence identified on plasmid 1 phage and seven were found within the plasmid 2 phage.

When analysing the presence and absence output from pan-genome analysis (Roary), the five genes associated with the KEGG BpeEF-OprC efflux pump (M00698), were also found in the genome sequences for RW130, RW146, RW172, RW176 and RW184, but were not identified in the remaining 97 *P. aeruginosa* genome sequences. These five genome sequences were all industrial isolates obtained from contaminated dishwasher liquids at different time points and in various locations (Chapter 2 Table 3). The KEGG BpeEF-OprC efflux pump (M00698) module was considered biologically feasible in all of these 5 industrial strains when checked with the KEGG MAPLE tool (MCR of 100% and a Q-value of 0). The genes, which matched to the M00698 KO numbers within each strain along with the RW109 strain, are shown in Table 14. These 6 industrial strains were spread throughout the core genome phylogenetic tree (Figure 18) and did not cluster within one sub-clade.

Table 14. Industrial isolates with the BpeEF-OprC efflux pump KEGG module (M00698)

Industrial Isolates	M00698 BpeEF-OprC efflux pump (KO number)			
	K18900	K18901	K18902	K18903
RW130	RW130_06334 and RW130_06328	RW130_06329	RW130_06330	RW130_06331
RW146	RW146_05850 and RW146_05844	RW146_05845	RW146_05846	RW146_05847
RW172	RW172_05145 and RW172_05139	RW172_05140	RW172_05141	RW172_05142
RW176	RW176_03984 and RW176_03978	RW176_03979	RW176_03980	RW176_03981
RW184	RW184_02738 and RW184_02744	RW184_02743	RW184_02742	RW184_02741
RW109	RW109_00010 and RW109_00017	RW109_00012	RW109_00013	RW109_00014

Footnotes: Each strain gene ID, which correspond to the KO numbers are shown in the columns.

3.2.6.5 Genome size comparison of RW109 with the 102 *P. aeruginosa* genome sequences

The number of base pairs for each of the 103 *P. aeruginosa* sequences (including RW109) were obtained from the Prokka summary annotation files and the genome size distribution for all the strains were examined (Figure 20A). The median genome size for all 103 sequences was 6,643,410 bp and six strains (RW109, RW146, RW130, RW131, Carb01-63 and RW172) were identified as outliers with larger genome sizes (Figure 20A). RW109 was found to have the largest genome size with 7,756,224 bp and was identified as the top outlier on Figure 20A. Five of the outliers were identified as industrial strains and Carb01-63 was a clinical strain with an unknown isolation source. Four of these outlying industrial strains (RW109, RW130, RW146 and RW172) had previously been identified as obtaining the 5 genes associated with a BpeEF-OprC efflux pump KEGG module (M00698) (Sections 3.2.3.2.2 and 3.2.6.4).

The distribution of genome sizes were examined for the three different strain isolation sources (Figure 20B); this comparative analysis excluded PAIRG, which was classified as both clinical and environmental, and the ATCC-13388 strain with an unknown source of isolation. The genome sizes of those isolated from an industrial setting were found to be larger than strains from a clinical and an environmental source (Figure 20 B and C). Strains isolated from an environmental setting were found to have the lowest genome size when compared to the other strain types. The mean sizes of the genomes were calculated (Figure 20C) and statistical analysis using a two-sample T-test (assuming unequal variances, as decided by the F-test), was performed to determine if the mean genome size of strains isolated from industry was significantly different to those obtained from a clinical and environmental setting. The industrial strains were found to have a significantly larger mean genome size when compared to the clinical strains (7,086,444 bp compared to 6,660,367 bp, p-value = 0.0020) and those isolated from an environmental setting (7,086,444 bp compared to 6,527,492 bp, p-value = 0.0002).

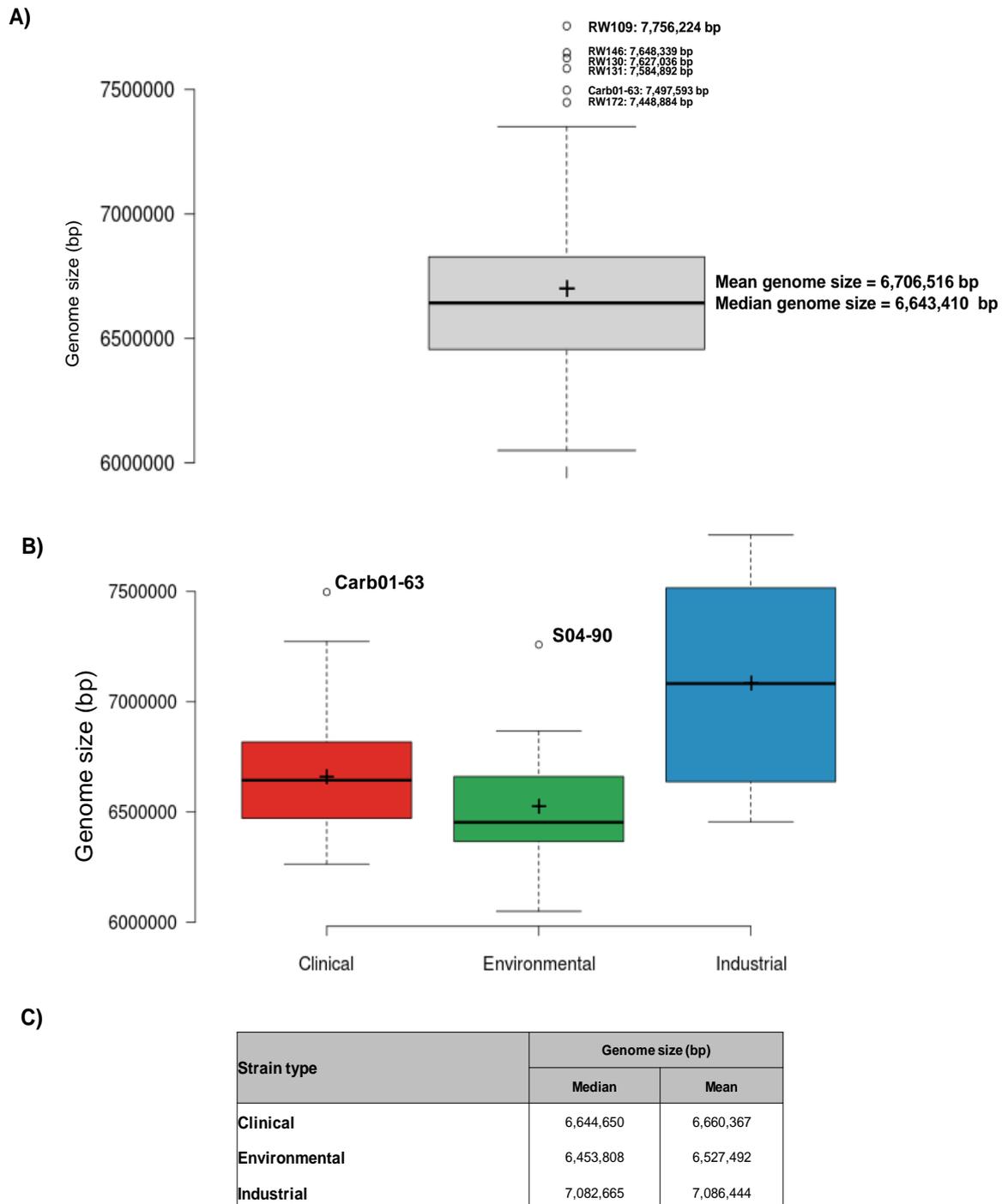


Figure 20. Genome size distributions of 103 *P. aeruginosa* sequences of strains isolated from clinical, environmental and industrial sources

The distribution of the genome sizes for the 103 *P. aeruginosa* sequences were visualised using boxplots which display the median (black horizontal line), upper quartile, lower quartile, maximum and minimum values. The mean values are also represented with a cross. Outliers are shown as open circles outside of the boxplots with the strain names and genome sizes indicated. A boxplot to represent the genome size distribution of the 103 sequences combined are shown in A). The sequences were also divided into three boxplots (B) to indicate the genome sizes of the clinical strains (red, n = 66), environmental strains (green, n = 19) and industrial (blue, n = 16). The values for the median and mean genome sizes (bp) for the clinical, environmental and industrial strains are shown in panel C).

3.3 Discussion

3.3.1 Obtaining a complete sequence of the industrial *P. aeruginosa* strain RW109

When searching the literature with a combination of the key words ‘*Pseudomonas aeruginosa*’, ‘industrial contamination’, ‘home and personal care products’ and ‘whole genome sequencing’, no studies appear to have obtained a whole genome sequence of an industrial *P. aeruginosa* strain specifically isolated from a home and personal care product. In this project PacBio SMRT sequencing technology, was successfully used to acquire a complete genome of the industrial RW109 strain, which assembled into three contigs. The genome sequence provided a valuable insight into the characteristics of an industrial strain and detailed functional annotation was carried out using COG and KEGG. The antimicrobial resistance and virulence genes were classified and the genomic islands and phage sequences of the strain were identified. Determining these important features of a genome sequence is essential for in-depth transcriptomic investigations, which were also carried out with the industrial strain as part of this PhD project.

RW109 was found to have a genome size of 7.8 Mbp, which is larger than the reported *P. aeruginosa* genome size range of 5.2 to 7.0 Mbp (Schmidt et al., 1996, Stover et al., 2000, Juhas et al., 2005, Winsor et al., 2016). The RW109 genome comprised a 7,049,347 bp main chromosome and two plasmids; 555,265 bp were found in plasmid 1 and plasmid 2 had 151,612 bp. The large genome size and the two additional plasmids may contribute to the ability of the RW109 strain to survive in industrial HPC products by encoding for additional genes. The genome size of the 103 *P. aeruginosa* sequences were compared and RW109 was found to have the largest number of bp. The mean genome size of the 16 industrial strains was also found to be significantly higher when compared to those strains obtained from a clinical and environmental setting, suggesting larger genome size is indicative of a strain’s ability to reside within an industry setting. The numbers of genome sequences examined in this study were not however equal for the three different isolation sources; a greater number of industrial and environmental genome sequences need to be included in future analysis.

Plasmid 1 of RW109 was found to have similar sequence similarity with two previously identified *P. aeruginosa* plasmids pOZ176 (Xiong et al., 2013) and pBM413 (Chen et al., 2016a, Botelho et al., 2017). The pOZ176 plasmid was identified in the clinical *P. aeruginosa* strain 96 isolated from a hospital setting in China. The pOZ176 plasmid was found to encode genes, which mediate carbapenem resistance, and was thought to have originated from an environmental source. The pBM413 plasmid was found in the clinical PA121617 strain isolated from CF sputum in China. When the pOZ176 and pBM413 plasmids were sequenced, they were described as the largest plasmids which had been isolated from *P. aeruginosa* (Xiong et al., 2013, Botelho et al., 2017). As both these plasmids were from strains isolated in clinical settings, it suggests that RW109 may have been acquired from a clinical source, providing the strain with plasmid associated antimicrobial resistance genes.

Contig 3 had similar sequence homology to the un-named plasmid from the environmental *P. aeruginosa* strain SO4 90 which was isolated from microbial mat material in Rotterdam, the Netherlands (Kraak and van der Zee, 2015). This suggests plasmid 3 may have been acquired by RW109 from an environmental source. The RW109 plasmids were only compared against *P. aeruginosa* plasmid sequences in this study; these plasmids may have been acquired from other bacterial species. For further classification of the plasmids, sequence comparisons need to be carried out against a larger collection of plasmid sequences across bacterial genera isolated from various settings. Searching databases such as the plasmid MLST collection may also help with characterisation of these industrial strain plasmids (<https://pubmlst.org/plasmid>). Additional work to increase the understanding of these plasmids and their importance in RW109's capabilities as an industrial strain would be to experimentally remove them. This would determine the effect the plasmids have on the strain's ability to grow in preservative formulations and HPC products. Additional plasmids in environmental *P. aeruginosa* isolates are related to heavy metal resistance (Raja and Selvam, 2009, Durve et al., 2013) and plasmids also have associations with increased tolerance to the industrial antimicrobials chlorhexidine, quaternary ammonium compounds, triclosan and formaldehyde (McDonnell and Russell, 1999). However when compared to antibiotic resistance, the role of plasmids in delivering increased resistance to industrial preservatives is not well established.

3.3.2 Functional annotation of a *P. aeruginosa* strain isolated from an industrial setting

The CDS of the RW109 genome were annotated with COG categories which provided the ability to investigate the functional classifications of the main chromosome and the two plasmids. The highest percentage of CDS within the well characterised COG categories of the main chromosome was in transcription (K= 7.75%) and the replication, recombination and repair (L) category was found to represent 9.00% and 14.00% of the CDS of plasmid 1 and 2 respectively. *P. aeruginosa* has previously been identified as having one of the largest transcriptional regulatory networks of those reported in bacteria (Galán-Vásquez et al., 2011, Balasubramanian et al., 2012, Coggan and Wolfgang, 2012); this defines *P. aeruginosa* capabilities as a metabolically versatile and virulent bacterium. The highest percentage of PA14 and PAO1 CDS were also found within the K COG category in this study and a large portion of genes associated with regulation was reported when the first genome *P. aeruginosa* genome (PAO1) was sequenced by (Stover et al., 2000); highlighting the importance of the regulatory functional characteristic of *P. aeruginosa*. A substantial number of the L category CDS on plasmids 1 and 2 were found to be encoding transposases, a common characteristic of plasmids; presence of these mobile gene elements suggests these plasmids are suitable for transposition providing RW109 with the ability to adapt to various environments. The poorly characterised COG categories represented the biggest percentage of CDS within the main chromosome and the two plasmids. Overall 40.42% of the RW109 CDS were within the poorly characterised class whereas PA14 had 33.86% and PAO1 32.53% had CDS within this class; this suggests that more of the industrial strain CDS are of unknown function and these uncharacterised proteins may have a novel mechanism of action.

Due to the large genome size of RW109, the strain was found to have a greater number of CDS in seventeen of the twenty well-characterised COG categories when compared to PA14 and PAO1. This suggests that RW109 has a broader functional potential than the reference strains. RW109 had over 95% more CDS in the L category than the reference strains, which is owing to the large number of CDS on the two additional plasmids within this category. RW109 was found to have 20% more CDS in the intracellular trafficking, secretion, and vesicular transport (U) category compared to the reference strains. The high numbers of CDS found within the L and U COG categories highlights their importance in the lifestyle of RW109 industrial strain. A recent study carried out COG comparisons between clinical and environmental isolates of *Burkholderia* by counting the number of CDS per COG category (Peeters et al., 2017). It demonstrated that isolates obtained from the environment were significantly enriched with the L and U COG categories when compared to clinical isolates, suggesting that the isolation source of bacteria has an effect on functional characteristics of a genome.

The percentage distribution of the well characterised COG categories when compared to the genome size was also analysed and found to be similar in the three strains, with the K and amino acid transport and metabolism (E) categories having greatest percentage of CDS (Figure 16). This illustrates the functional importance of the K and E categories in the lifestyle of *P. aeruginosa*. The energy production and conversion (C) category had the third largest percentage of CDS for PA14 and PAO1 and this category was ranked fourth for RW109. The L category had the third largest percentage of CDS for RW109 whereas for the reference strains, this category was ranked eleventh in the percentage distributions of the well-characterised COG categories. This difference between the two clinical strains and RW109 suggests that the biological functions of the CDS within the L category are important for a strain to adapt and endure industrial conditions.

As a preliminary analysis of the RW109 genome, identifying the COG functional annotations provided a valuable insight into the strain's biological functions and was important for comparative analysis with reference strains. The database is a universal tool which is regularly used for genome-scale analysis of protein functions and provides a consistent classification of bacterial genomes based on orthologous groups (Galperin et al., 2014). Assigning COG categories also provides a classification method of grouping genes of interest identified in transcriptomic analysis. However due to the large size of the COG database, it is not regularly updated (Huerta-Cepas et al., 2016) which might explain the high percentage of CDS which were placed within the poorly characterised COG classes.

The KEGG database offers a larger actively maintained reference system for classifying biological systems (Kanehisa et al., 2017). This method of functional annotation was used to interpret the genome sequence of RW109 by determining biologically feasible KEGG modules within the strain. Across the whole genome, including the main chromosome and two plasmids, 171 biologically achievable KEGG modules were identified and found to be involved in a variety of functions; this suggests the strain has a high metabolic and physiological potential. The two-component regulatory system group was found to contain the greatest number of complete modules, which were all on the main chromosome of RW109.

CHAPTER 3 - Whole genome sequencing and comparative genomics of the industrial *P. aeruginosa* strain RW109

The *P. aeruginosa* species has previously been reported to encompass a high number of two component regulatory systems which facilitate their ability to live in a variety of habitats (Gooderham and Hancock, 2009, Stover et al., 2000, Rodrigue et al., 2000) and KEGG annotation identified 24 in RW109 (Table 7). Two component regulatory systems are used to sense and mediate responses to environmental stimuli by regulating target genes which have roles in antimicrobial resistance, virulence, quorum sensing and biofilm formation (Gooderham and Hancock, 2009, Stock et al., 2000, Mikkelsen et al., 2011b). When the RW109 plasmids were evaluated with KEGG, five out of the six modules identified on plasmid 1 were also found on the main chromosome, suggesting the plasmid contains duplicate copies of main chromosome genes. A multidrug resistance efflux pump BpeEF-OprC module (M00698) was found only in plasmid 1 of RW109 within GI-1 and plasmid 2 was found to have any biologically feasible KEGG modules.

Using the ABRicate tool and BacMet database, RW109 was found to encode a high number of genes with predicted associations to antimicrobial resistance and virulence. This is supported by previous research demonstrating the ability of *P. aeruginosa* in producing a wide spectrum of virulence factors and antibiotic resistance capabilities (Jácome et al., 2012, Beceiro et al., 2013, Breidenstein et al., 2011). However, across all the 103 *P. aeruginosa* isolates analysed using the CARD database, the number of antibiotic genes was broadly equivalent. This suggests the industrial strains do not have additional predicted genomic capacity for antibiotic resistance and isolation source has limited impact on antimicrobial proficiency. These findings need to be assessed phenotypically with antibiotic susceptibility testing.

Island Viewer predicted that the RW109 strain had 42 GIs on the main chromosome (Table 10), six on plasmid 1 and one in plasmid 2 (Table 11). GIs are important regions of interest within strains as they frequently contain virulence and antimicrobial resistance genes as well as contributing metabolic fitness (Bertelli et al., 2017). A high number of GIs is important for *P. aeruginosa* to survive and adapt to niche environments (Battle et al., 2008, Schmidt et al., 1996); they belong to the accessory genome and up to 10% of an individual isolate's genetic material can be made up of GIs (Spencer et al., 2003, Shen et al., 2006). The high number of GIs identified in RW109 is consistent with other *P. aeruginosa* strains such as PSE9 which has been described as a hyper-virulent strain (Battle et al., 2008).

The RW109 GIs contained a large amount of hypothetical proteins and the poorly characterised COG classes represented the highest number of GI CDS suggesting novel functions. GI-1 within plasmid 1 was of interest as it encoded for the KEGG BpeEF-OprC efflux pump module (M00698); this module was also biologically feasible in five other industrial contamination strains but was not identified in any of the other 97 *P. aeruginosa* sequences examined. Genomic islands and phage sequences may also have a role as mobile genetic elements within the accessory genome of *P. aeruginosa* and are indicative of horizontal gene transfer during bacterial evolution (Li et al., 2016, Kung et al., 2010). RW109 was predicted to have five intact phage sequence regions which is comparable to other strains such as the clinical PA1 strain (Li et al., 2016) and the environmental isolates M18 (Wu et al., 2011) and YL84 (Chan et al., 2014).

**CHAPTER 3 - Whole genome sequencing and comparative genomics of the industrial
P. aeruginosa strain RW109**

Characterising the CDS of RW109 associated with antimicrobial resistance, virulence, GIs and phages, provides a valuable resource for further comparative genomics and transcriptomics work to distinguish genes important for survival within a niche industrial setting.

CHAPTER 3 - Whole genome sequencing and comparative genomics of the industrial *P. aeruginosa* strain RW109

3.3.3 Notable functional annotations of the RW109 strain in comparison to the clinical reference strains PA14 and PAO1

The RW109 whole genome sequence was compared to the annotated genome sequences of PA14 and PAO1 which represented well studied *P. aeruginosa* reference strains. The differences in COG functional annotations were previously discussed in Section 3.3.2. The KEGG module MCRs were evaluated, RW109 was found to contain 2 modules, which were not biologically feasible in PA14 and PAO1. These included the UhpB-UhpA two-component regulatory system (M00473) from the main chromosome and the multidrug resistance, efflux pump BpeEF-OprC (M00698) found on GI-1 of plasmid 1. The UhpB-UhpA system is involved in the UhpT transport system of hexose phosphates and is commonly found in *E. coli* and *Vibrio cholera* (*V. cholera*) (Moisi et al., 2013, Yamamoto et al., 2005). When searching the literature the UhpB-UhpA system was not evidently classified as being a *P. aeruginosa* two-component system. UhpT transports fosfomycin into the bacterial cell and inactivation of the regulatory genes *uhpA* and *uhpB* results in the down regulation of *uhpT* contributing to fosfomycin antibiotic resistance (Kadner and Shattuck-Eidens, 1983, Castañeda-García et al., 2013). A study which investigated the transport of fosfomycin in *P. aeruginosa* using a PA14 insertion mutant library identified a permease GlpT protein as the only mutational target involved in fosfomycin resistance (Castañeda-García et al., 2009). This suggests the UhpB-UhpA two-component regulatory system is not commonly found in *P. aeruginosa* strains. Another study which carried out a pan-genome wide analysis of 181 *P. aeruginosa* genomes identified one gene of the UhpB-UhpA system within the accessory genome (Mosquera-Rendón et al., 2016) further confirming the absence of this two component regulatory system in *P. aeruginosa*. As UhpB-UhpA system was identified in RW109, it suggests this two component regulatory system may provide a functional role for survival in diverse environments such as those encountered in industry.

The RW109 BpeEF-OprC efflux pump identified within GI-1 of plasmid 1, is a multi-drug efflux pump found in *Burkholderia pseudomallei* (*B. pseudomallei*) (Chan et al., 2004). It is linked to the resistance of clinical and environmental *B. pseudomallei* isolates to chloramphenicol and trimethoprim (Kumar et al., 2006, Podnecky et al., 2013). This efflux pump was not identified in PA14 and PAO1 and when further analysis was carried out on the 102 *P. aeruginosa* genomes from Section 3.2.6, the genes encoding this efflux module were identified in only industrial strains isolated from contaminated dishwasher liquids. This suggests the efflux pump may provide protection against the industrial antimicrobials present within dishwasher liquids. These industrial *P. aeruginosa* strains may have acquired the efflux pump genes from a *Burkholderia* species which are also commonly found to contaminate HPC products (Russell, 2003c, Jimenez, 2007). Four out of the five industrial strains, which obtained this efflux pump, were also identified as having a larger genome sizes when compared to the panel of 103 *P. aeruginosa* sequences (Figure 20). Further work is required to determine if these industrial strains are carrying the BpeEF-OprC efflux pump genes on an additional plasmid similar to the one in RW109.

One KEGG module was found to be biologically achievable in both PA14 and PAO1 but not in RW109 and this was a central carbohydrate metabolism module involved in 5-phospho- α -ribosyl-1-pyrophosphate

CHAPTER 3 - Whole genome sequencing and comparative genomics of the industrial *P. aeruginosa* strain RW109

(PRPP) biosynthesis (M00005). This metabolic pathway involves the ribose-phosphate diphosphokinase enzyme (K00948) which converts ribose 5-phosphate to PRPP, a ribose sugar required for purines and pyrimidines synthesis (Hove-Jensen et al., 2003, Kanehisa et al., 2017). This module is important for bacterial metabolism, suggesting that RW109 would have an alternative means of encoding for the ribose-phosphate diphosphokinase enzyme; further interpretation of the genome is required to identify CDS in RW109 associated with this enzyme.

The antimicrobial resistance genes of RW109 were predicted using the ABRicate tool and the BacMet database. ABRicate compares genome sequences against four well maintained antimicrobial resistance gene databases whereas BacMet has not been updated since 2014 but provides experimentally confirmed information about biocide and metal resistance genes. Using ABRicate, little differences were identified in the antimicrobial resistance genes between RW109 and the clinical reference strains. The only variation observed was with two beta lactamase encoding genes; RW109 and PAO1 were both found to have *PCD-1* and PA14 had the *PCD-9* gene. The *PCD-9* gene encodes for a beta lactamase which is capable of degrading an extended spectrum of antibiotics when compared to *PCD-1* (Rodríguez-Martínez et al., 2009). The BacMet database was used to compare RW109 with PA14 and PAO1 and the industrial strain was found to have nine additional genes associated with biocide and metal resistance with multiple antimicrobial functions; six of these genes were on plasmid 1. This suggests that the RW109 strain encompasses a clinical antimicrobial resistance profile similar to PA14 and PAO1 but also has additional genes, which are linked to biocide antimicrobial resistance.

Comparative genomic analysis to determine differences in functional annotations between strains was a useful research method to identify candidate genes for further examination during transcriptomic analysis. For example, it would be interesting to establish if the genes encoding for the BpeEF-OprC efflux pump or the UhpB-UhpA two component regulatory system were differentially regulated in response to industrially relevant conditions.

3.3.4 Core genome analysis of 103 *P. aeruginosa* genome sequences isolated from clinical, environmental and industrial sources

The core genome comparison of the 103 *P. aeruginosa* sequences identified that isolation source had limited influence on the distributions of strains throughout the constructed phylogenetic tree as all three types of strains were found within the tree clades (Figure 18). This was previously shown in a whole genome multi locus sequence typing (MLST) profile analysis where strains isolated from clinical, environmental and industrial settings were also dispersed evenly throughout the network (Weiser, 2015). These results highlight that at a genomic level, isolation source does not result in specific niche sub-groups in *P. aeruginosa*. The core-genome alignment did however identify the highest number of industrially isolated strains to fall within the sub-clade, which contained RW109 when compared to the others in the tree. This indicates these 6 industrial strains share a comparable core genome and may have originated from a similar origin. However, there were also 6 clinical strains within the sub-clade suggesting that a related core genome does not necessarily determine strain type. More industrial strains

CHAPTER 3 - Whole genome sequencing and comparative genomics of the industrial *P. aeruginosa* strain RW109

were also identified within clade 1 of the tree when compared to clade 2 suggesting found in clade 1 share a similar core genome.

The accessory genome of these strains may provide the essential genes required for living within either a clinical or industrial environment. It has previously been identified that *P. aeruginosa* strains obtain highly diverse genomes, even amongst closely related isolates (Mathee et al., 2008, Jun et al., 2015). The 5 strains within the RW109 sub-clade may all be capable of living within a clinical or an industrial environment; their strain type was determined by their isolation source, but this may not be where they are only capable of residing. The numbers of genome sequences from the three isolation sources were not equal for the core genome analysis in this study; a higher number of the sequences were from strains isolated from a clinical setting (n = 66). This was due to the greater availability of sequenced clinical strains when compared to environmental (n = 19) and industrial isolates (n= 16). Potentially, if more environmental and industrial sequences were included, clustering within these particular strain types may be observed, as strains from the same source are likely to share a similar source of origin.

The core genome alignment identified 2 major tree clades with the panel of 103 *P. aeruginosa* genome sequences. The tree was rooted with the PA7 genome sequence as its core genome was highly divergent when compared to the other sequences. The identification of two major groupings of *P. aeruginosa* whole genome sequences with the contrasting PA7 strain solely being placed into a separate third group, has previously been confirmed (Freschi et al., 2015).

The CARD database was used to predict the presence and absence of antimicrobial genes across the 103 *P. aeruginosa* genomes; small differences were observed between the genome sequences of the strains isolated from the different sources (Figure 19). When compared to the clinical strains, CARD did not identify any antimicrobial resistance genes within the quinolone, carbapenem, trimethoprim, macrolide and tetracycline efflux resistance functional groups for the industrial and environmental strain sequences. This suggests that these resistance genes may be more specific to survival in clinical settings rather than in the environment and industry. This analysis only included one antimicrobial resistance gene database and incorporating others may increase the accuracy of the predicted resistance profiles of the genome sequences. Further relevant analysis would be to search the genes of the 103 genome sequences using the BacMet database to determine if the industrial strains have a different biocide and metal gene resistance profile when compared to the clinical and environmental strains. In addition, functional screening for resistance traits associated with the predicted genes would be needed for a complete analysis of antimicrobial resistance. Interestingly the six industrial strains with the biologically achievable antimicrobial BpeEF-OprC efflux pump (M00698) did not all cluster within one sub-clade and were dispersed throughout the core-genome phylogenetic tree suggesting different sources of origins for these strains.

3.4 Conclusions

The main conclusions from this detailed investigation of an industrial *P. aeruginosa* strain genome were as follows:

- 1) A complete genome of the industrial RW109 strain was successfully obtained using PacBio SMRT sequencing technology.
- 2) RW109 had a larger than average sized genome when compared to a panel of 102 *P. aeruginosa* genome sequences which were isolated from clinical, environmental and industrial sources; the 7.8 Mbp genome comprised a main chromosome and two plasmids.
- 3) A thorough functional annotation of RW109 was achieved using COG, KEGG, ABRicate and BacMet databases providing an accurate reference sequence for further transcriptomic analysis of the industrial strain (see Chapter 5).
- 4) Comparative analysis with the clinical reference strains along with the pan-genome analysis of 102 *P. aeruginosa* genome sequences identified important RW109 genes. Notable genes included those from a genomic island on plasmid 1 encoding for a BpeEF-OprC efflux pump only identified as biologically achievable in industrial strains isolated from contaminated dishwasher liquids.

4. Use of an industrial *P. aeruginosa* mini-Tn5-*luxCDABE* mutant library to investigate the genetic basis of preservative resistance

4.1 Introduction

Important research tools such as transposon mutant libraries, have been used to characterise the genetic basis of functional traits in *P. aeruginosa*, such as quorum sensing (Whiteley et al., 1999), swarming (Overhage et al., 2007) and antimicrobial resistance (Aires et al., 1999, McPhee et al., 2003, Lewenza et al., 2005) (Table 15). Transposition can be achieved by incorporating the reporter gene into transposon vectors such as mini-Tn5, and delivering on a pUT suicide plasmid by conjugation; due to the loss of the *tnp* gene after insertion, the transposon is stably integrated into the chromosome and inherited (Winson et al., 1998, De Lorenzo et al., 1990, Dehio and Meyer, 1997). The transposon results in knockout polar mutations of downstream genes in the same operon, generating a transcriptional fusion with a random bacterial gene (Weitz et al., 2001, Lewenza et al., 2005). Mutant libraries can be analysed to identify those with altered functional attributes in comparison to the wild-type strain. Key mutants can then be characterised to determine the genetic basis for their altered phenotype. A variety of promoterless reporters (*lux*, *phoA*, *lacZ*) can also be integrated into the transposon to generate measurable transcriptional or translational signals (Winson et al., 1998, Whiteley et al., 1999, Aendekerk et al., 2002).

The *luxCDABE* operon developed from *Photobacterium luminescens*, incorporates all the necessary genes for luminescence and requires only ATP from metabolically active cells for the encoded luciferase to emit light (Figure 21) (Winson et al., 1998, Meighen, 1994). Bioluminescent *lux* reporters provide a method of quantifying gene expression through the detection of light emission alterations in response to a particular stimuli (Su et al., 2011, Winson et al., 1998). A mini-Tn5-*luxCDABE* transposon ideally generates individual mutants which have the *lux* reporter infused within different genes (Lewenza et al., 2005). The resulting mutants can function as novel whole cell biosensors, allowing the measurement of luminescence change in response to environmental changes. This research tool provides a high-throughput and highly sensitive way of reporting gene expression. A multi-functional mini-Tn5-*luxCDABE* mutant library using the PAO1 *P. aeruginosa* strain was successfully constructed, with the resulting transposon insertion sites being evenly distributed throughout the genome (Lewenza et al., 2005). *P. aeruginosa* mini-Tn5-*luxCDABE* mutant libraries have effectively been used to identify genes involved in the ability of the microorganism to swarm (Overhage et al., 2007) and were exploited in the classification of a PmrA-PmrB regulatory system relating to polymyxin resistance (McPhee et al., 2003) (Table 15).

Table 15. The use of transposon mutagenesis to identify the genetic basis of functional characteristics in *P. aeruginosa*

Transposon mutagenesis	Screening using transposon mutagenesis	Identification of transposon insertion site	Major research findings	Reference
Tn501 random insertion mutagenesis of <i>P. aeruginosa</i> PAO1.	Random Tn501 tagging of the <i>P. aeruginosa</i> PAO1 strain was carried out using a temperature sensitive pMT1000 plasmid derived from R68::Tn501. Mutants screened for resistance or hyper-susceptibility to aminoglycosides by individual replication on Mueller Hinton agar plates containing 0.5 or 16 µg of amikacin ml ⁻¹ .	A 7-kbp genomic fragment from strain PAO1 which contained the Tn501 insertion site from a mutant of interest was cloned into <i>E. coli</i> and the fragment was sequenced to identify insertion site.	Identification of a mutant hyper-susceptible to aminoglycosides, tetracycline and erythromycin. Mutant of interest found to be deficient in the MexXY active efflux pump system.	(Aires et al., 1999)
Tn5-B22 promoterless <i>lacZ</i> random transposon insertion mutagenesis in the chromosome of a <i>P. aeruginosa</i> (PAO-MW1) quorum sensing acyl-acyl-homoserine lactone (HSL) synthesis mutant.	Library constructed was screened to identify mutants showing acyl-HSL-dependent β-galactosidase expression using the <i>lacZ</i> reporter. Transcription of the Tn5-B22 <i>lacZ</i> within a quorum-sensing gene responds to an acyl-HSL signal as the transposon contains a promoterless <i>lacZ</i> and PAO-MW1 is a <i>lasI</i> , <i>rhlI</i> null mutant that does not make acyl-HSL signals.	Arbitrary PCR was used to identify DNA sequences flanking Tn5-B22 insertions. Sequences that could not be identified by using this method were cloned into <i>E. coli</i> and sequenced.	Identification of 39 quorum sensing-regulated genes, which code for putative virulence factors or production of secondary metabolites.	(Whiteley et al., 1999)
Random Mini-Tn-<i>phoA3</i> transposon mutagenesis of <i>P. aeruginosa</i> strain PA59.20, a type III pyoverdine-producing strain.	Mutants were screened by replica-plating on Casamino acid medium containing VOSO ₄ for the detection of Vanadium-sensitive mutants.	An inverse PCR was used to identify insertion sites, the method involved: <ul style="list-style-type: none"> • The isolation of genomic DNA, • Digestion with restriction enzymes (<i>SmaI</i> digestion) that cut once within the transposon but numerous times within the <i>Pseudomonas</i> genome. • Ligation to circularise all linear genomic fragments. • Ligated products amplified via PCR, which was followed by a nested PCR reaction. • Products of inverse PCR were sequenced. 	Identification of one mutant, which could not grow in the presence of VOSO ₄ and the transposon had inserted in a non-coding region upstream of a 4 gene cluster: <i>mexGHI-opmD</i> . This cluster is associated with a <i>P. aeruginosa</i> antibiotic efflux pump systems.	(Aendekerk et al., 2002)

<p>Bank of bioluminescent Mini-Tn5 <i>luxCDABE</i> fusion strains in various <i>P. aeruginosa</i> strains.</p>	<p>Constructed bioluminescent bank was screened for mutants with differential light emission expressions in response to high or low MgSO₄ concentrations. This was used to identify regulatory systems in <i>P. aeruginosa</i>, which respond to varying MgSO₄ levels.</p>	<p>Interrupted genes were identified by two step arbitrarily primed PCR. In first round used a primer specific for the mini-Tn5 <i>luxCDABE</i> transposon and an ARB1 primer. Products from this PCR reaction were then used as the template in second round PCR, which used an ARB2 primer and same mini-Tn5 <i>luxCDABE</i> primer from first round PCR.</p>	<p>Screening for Mg²⁺ responsive promoters using the library of mini-Tn5 <i>luxCDABE</i> mutants identified a two-component regulatory system, PmrA-PmrB. This system regulates resistance to polymyxin B and cationic antimicrobial peptides.</p>	<p>(McPhee et al., 2003)</p>
<p>pBTK30 <i>mariner</i> C9 mini-transposon insertions in <i>P. aeruginosa retS</i> null mutants.</p>	<p>Mutants were used to screen for suppressors of distinct RetS-dependent phenotypes such as decreased hyper-adhesion and repression of the type III secretion system (TTSS) when grown on minimal and TTSS-inducing culture medium.</p>	<p>Insertion sites were identified using two round arbitrary PCR and the amplified transposon-chromosome junctions were sequenced using BTK30-Seq.</p>	<p>In suppressors of distinct RetS-dependent phenotypes, a high number of transposon insertions were found in the genes of the GacS/GacA/rsmZ regulatory pathway. This pathway functions as a multicomponent switch for the transition from acute colonisation to chronic infection.</p>	<p>(Goodman et al., 2004)</p>
<p>Whole genome random mutagenesis of PA01 using a bioluminescent Mini-Tn5 <i>luxCDABE</i> reporter.</p>	<p>The mini Tn5-<i>luxCDABE</i>-Tc transposon was exploited to create a multi-functional mutant library using the PA01 <i>P. aeruginosa</i> strain. Differentially regulated novel genes were identified when exposed to magnesium and phosphate-limiting conditions.</p>	<p>Insertion sites were mapped via a high-throughput inverse PCR similar to the one used by Aendekerck et al., 2002. Differences included using <i>NarI</i> restriction enzymes which cut once within the transposon but numerous times within the PA01 genome. Products of ligation were amplified using a PCR, which incorporated two outward-facing, transposon-specific primers.</p>	<p>The genes identified through screening were mapped to wild-type genome and found to be involved in bacterial resistance to cationic antimicrobial peptides.</p>	<p>(Lewenza et al., 2005)</p>
<p>Mini-Tn5-<i>luxCDABE</i> transposon PA01 mutant Library</p>	<p>Use of library constructed by Lewenza et al., (2005) to screen for alterations in swarming motility. 5,000 mutants were tested by growing on brain heart infusion agar plates containing 0.5% (wt/vol) Difco agar. Phenotypic screening was carried out on 36 swarming-deficient mutants by analysing the swimming and twitching motility.</p>	<p>Same inverse PCR used by Lewenza et al., (2005).</p>	<p>Tn5 insertions identified a variety of genes involved in <i>P. aeruginosa</i> swarming. The genes were found to be involved in the synthesis or function of flagellin, type IV pilus, in genes for the Xcp-related type II secretion system, and in regulatory, metabolic, chemosensory, and hypothetical genes with unknown functions. A high number of swarming-negative mutants also showed impairment in biofilm formation.</p>	<p>(Overhage et al., 2007)</p>

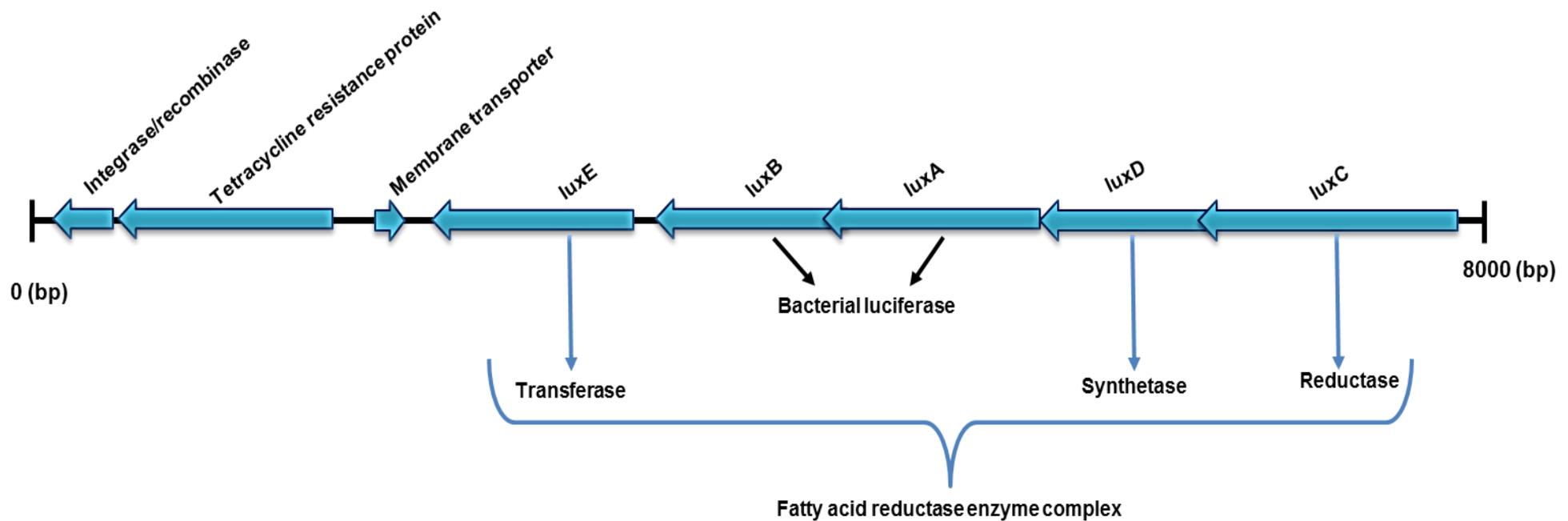


Figure 21. A bioluminescent mini-Tn5-*luxCDABE*-Tc transposon

The promoterless mini-Tn5-*luxCDABE*-Tc transposon can be exploited to identify mutants which respond to preservative exposure by monitoring changes in light emission over a set time period. The transposon used in this study contains a tetracycline (Tc) resistance cassette flanked by transcriptional and translational termination signals and a promoterless *luxCDABE* operon. All the genes required for bacterial bioluminescence are found within the operon so there is no requirement to add an exogenous substrate (Winson et al., 1998, Meighen, 1994). The *luxAB* genes encode the heterodimeric subunits of bacterial luciferase, whilst *luxCDE* encode proteins that make up a fatty acid reductase enzyme complex for synthesis of the long-chain fatty aldehyde luciferin substrate (Meighen, 1994). Bacterial luciferases emit light at 490 nm by catalysing the oxidation of reduced flavin mononucleotide (FMNH₂) and the long-chain aldehyde.

CHAPTER 4 – Use of an industrial *P. aeruginosa* mini-Tn5-*luxCDABE* mutant library to investigate the genetic basis of preservative resistance

4.1.1 Aim and objectives

The overall aim of this chapter was to investigate the genetic basis of industrial *P. aeruginosa* preservative resistance. This was achieved with the following objectives:

- 1) Construct a mini-Tn5-*luxCDABE* mutant library within the industrially isolated RW109 *P. aeruginosa* strain.
- 2) Develop a screening method to identify mutants with altered light emission responses when exposed to sub-lethal concentrations of industry relevant preservatives.
- 3) Determine the putative gene function in mutants of interest.

CHAPTER 4 – Use of an industrial *P. aeruginosa* mini-Tn5-*luxCDABE* mutant library to investigate the genetic basis of preservative resistance

4.2 Results

4.2.1 Construction of a mini-Tn5-*luxCDABE* mutant library using an industrial *P. aeruginosa* strain

4.2.1.1 Transconjugant selective media

A broth assay was used to determine the MIC of the donor and recipient with the antibiotic Tc before further assessment with an agar dilution assay. The lowest Tc concentration which resulted in a >80% reduction when compared to growth in the control wells, was 25 µg/ml for the donor and 100 µg/ml for the recipient strain. The wells, which corresponded to the MIC values, demonstrated no signs of visible growth in the 96 well plates, supporting the MICs determined by the absorbance readings.

An agar assay was used to evaluate concentrations and combinations of Tc and Pmx within TSA, in order to prevent growth of the donor and recipient strains, whilst allowing growth of the transconjugants. Selective plates, which contained Tc only at 150 µg/ml and Tc at 125 µg/ml, along with Pmx at 120 units/ml, prevented the growth of both the donor and recipient. These selective TSA plates were then further assessed by plating 50 µl of re-suspended mating mixture at $10^0 - 10^{-3}$ dilution onto triplicate plates. The plates containing Tc only at 150 µg/ml with 10^{-1} dilution of the mating mixture resulted in around 50 and to 150 individual, evenly spread colonies, allowing for ease of transconjugant collection. As Tc only at 150 µg/ml prevents growth of both donor and recipient, this concentration was used within selective media during mutant bank construction.

4.2.1.2 The RW109 *lux* tagged bioluminescent mutant bank

An industrial *P. aeruginosa* strain RW109, was used to construct a mini-Tn5-*luxCDABE* mutant library by performing a bi-parental mating with an *E. coli* donor carrying the mini-Tn5 on the S17-1 λ pir suicide plasmid pUT. In total 3,760 mutants were generated and stored in 96-well master plates (n=40 plates) at -80 °C. Each mutant was categorised based on the RLU measurements and mutants with light emission readings above 499 RLU represented 52.82% of the library. The transposon mutant library bioluminescent light emission categories, as described in Section 2.6.1.2, were represented as follows: category 1 ($\geq 1.00E+06$ RLU) = 0.03% (1 mutant), category 2 ($\geq 1.00E+05$ RLU) = 11.81%, category 3 ($\geq 1.00E+04$ RLU) = 15.05%, category 4 ($\geq 1.00E+03$ RLU) = 18.94%, category 5 ($\geq 5.00E+02$ RLU) = 6.99% and category 6 ($\leq 4.99E+02$ RLU) = 47.18%. A *luxA* PCR, as described in Section 2.6.2, was used to determine if mutants within the lowest light emission category 6 contained the *luxCDABE* transposon. Out of the 20 mutants tested across the library, 13 had a positive result and exhibited a light emission reading of ≥ 227 RLU. There were 2,272 mutants identified which had a light emission of ≥ 227 RLU with the following statistical characteristics, a mean of 45,184 RLU, a lower quartile of 1,092 and an upper quartile of 63,106 RLU.

CHAPTER 4 – Use of an industrial *P. aeruginosa* mini-Tn5-*luxCDABE* mutant library to investigate the genetic basis of preservative resistance

RAPD PCR typing was also used to ensure the picked mutants from the transconjugant selective media, had the same genotype as the recipient RW109 strain and not the donor. Eighteen of the 2,272 mutants from various light emission categories were selected at random and run through a RAPD typing PCR. RAPD PCR banding patterns for the mutants and were compared against the lanes representing the recipient, donor and a negative control (Figure 22). All mutants analysed demonstrated a comparable banding pattern to the recipient strain, except for the one in lane 6, where no bands were visible (Figure 22). This suggests that the selected mutants, with a comparable banding pattern to the positive control, had the same genotype as the wild type RW109 strain. No mutants evaluated had a banding pattern similar to donor strain, indicating the mutants did not have the genotype of the *E.coli lux* strain used in the bi-parental mating procedure.

CHAPTER 4 – Use of an industrial *P. aeruginosa* mini-Tn5-*luxCDABE* mutant library to investigate the genetic basis of preservative resistance

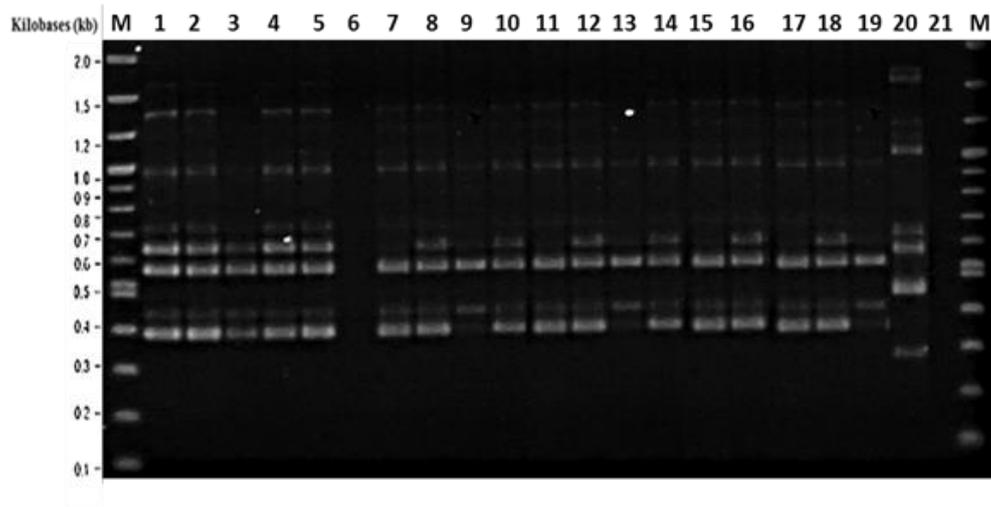


Figure 22. RAPD PCR gel for selected RW109 transposon mutants

The lanes were arranged as follows; M represents the 2-Log DNA ladder (0.1-10.0 kb), 1 contains the wild type RW109 (recipient strain), 2-18 show RW109 *lux* mutants arranged in order of increasing light emission category, lane 20 contains the *E. coli* (*lux* tagged donor) and lane 21 the negative control. No bands were observed in lane 6.

CHAPTER 4 – Use of an industrial *P. aeruginosa* mini-Tn5-*luxCDABE* mutant library to investigate the genetic basis of preservative resistance

4.2.2 Optimising the preservative screening method to identify responding mutants

Twelve mutants from across the mutant bank from light emission categories 1 – 5 were randomly selected, to choose an appropriate preservative concentration for identifying the maximum number of responding mutants during screening experiments. The preservatives BIT and POE were used to screen the mutant bank and the MICs for wild-type RW109 strain were determined as 0.00938% active for BIT and 0.5% active for POE using a broth doubling dilution assay, as described in Section 2.6.4. Each mutant was isolated from the master plates and screened as previously described in Section 2.6.5.2 with BIT and POE at 10%, 20%, 30%, 40% and 50% of the calculated MIC. Screening experiments were carried out twice using different mutant starting cultures and preservative stock solutions to obtain biological replicates, with each repeat having four technical replicates.

4.2.2.1 Screening with BIT

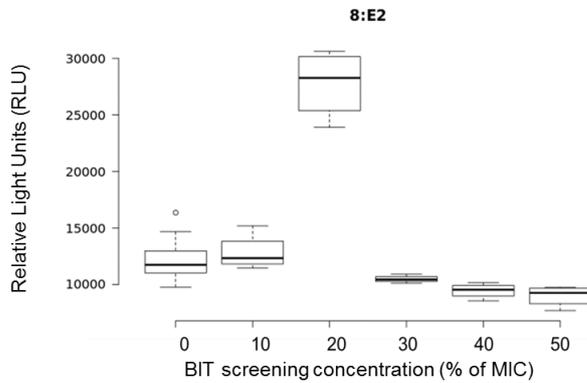
Screening with BIT at 20% of the calculated MIC resulted in the greatest number of mutants with a RLU-fold increase or decrease in response to BIT. Figure 23 demonstrates mutants 8:E2, 15:D4 and 30:C5 which had an increase (+1.29 fold), decrease (-0.59 fold) and minimal (-0.13 fold) RLU change respectively, when screened with 20% of the calculated BIT MIC. Screening with 10% of the calculated MIC resulted in the least change in RLU, whereas concentrations of 30%, 40% and 50% of the MIC resulted in an overall RLU-fold decrease. BIT at 20% of the MIC was chosen as an appropriate concentration to identify the maximum number of responding mutants during screening of the constructed master plates.

4.2.2.2 Screening with POE

POE exposure with all five screening concentrations resulted in an overall RLU fold decrease for the twelve mutants investigated. The mean RLU fold values for POE at 10%, 20%, 30%, 40% and 50% of the calculated MIC were -0.63, -0.68, -0.80, -0.83 and -0.87 respectively. It may have been possible that the 12 mutants selected to optimise the screening concentration, did not represent genes up-regulated in response to POE exposure. It was decided to include POE in master plate screens at 20% of the calculated MIC to determine if a varied RLU response was observed when more mutants were exposed to POE. The concentration of 20% of the MIC was chosen for comparison with the screening concentration selected for BIT exposure. Eight master plates were screened with POE, (approximately 752 mutants) and a mean -0.75-fold decrease was observed (data not shown), with not one mutant demonstrating an increase in RLU in response to POE exposure. It was decided to not include POE in further screening experiments as the results implied that the preservative was impairing the activity of the *luxCDABE* operon.

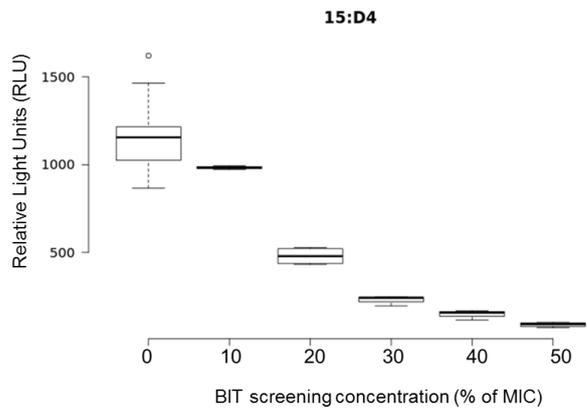
CHAPTER 4 – Use of an industrial *P. aeruginosa* mini-Tn5-*luxCDABE* mutant library to investigate the genetic basis of preservative resistance

A)



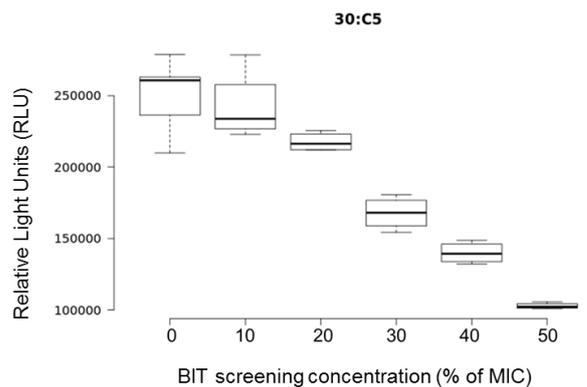
BIT screening concentration (% of MIC)	8:E2 RLU fold change
10	+0.06
20	+1.29
30	-0.14
40	-0.18
50	-0.26

B)



BIT screening concentration (% of MIC)	15:D4 RLU fold change
10	-0.15
20	-0.59
30	-0.80
40	-0.87
50	-0.92

C)



BIT screening concentration (% of MIC)	30:C5 RLU fold change
10	-0.04
20	-0.13
30	-0.33
40	-0.44
50	-0.59

Figure 23. Effect of BIT concentration on the RLU response of three mutants

The box plots illustrate the light emission response of mutants 8:E2 (A), 15:D4 (B) and 30:C5 (C) when exposed to BIT at 0%, 10%, 20%, 30%, 40% and 50% of the calculated MIC for 24 hour. Centre lines show the medians; box limits indicate the 25th and 75th percentiles as determined by R software; whiskers extend 1.5 times the interquartile range from the 25th and 75th percentiles, outliers are represented by dots. n = 8 sample points. The RLU fold change between readings taken for control (0% BIT) and BIT screening concentrations (% MIC) are also shown for each mutant.

CHAPTER 4 – Use of an industrial *P. aeruginosa* mini-Tn5-*luxCDABE* mutant library to investigate the genetic basis of preservative resistance

4.2.3 Screening of the constructed mini-Tn5-*luxCDABE* mutant library to identify preservative responding mutants

The 40 master plates in the transposon library were screened in triplicate with the preservative BIT at 20% of the calculated MIC for 24 hours, as previously described in section 2.6.5.1. The genes interrupted by the *lux* transposon were then identified in mutants isolated from the master plates to test for randomness of transposon insertion sites before embarking on individual mutant screens. In total 404 mutants of interest were isolated and -80°C freezer stocks prepared. This included 204 with increased RLU measurements (≥ 0.4 -fold RLU increase with a p -value ≤ 0.05 ; BIT positive responding mutants), 97 with decreased light emission (≥ 0.3 -fold RLU decrease with a p -value ≤ 0.05 ; BIT negative responding mutants) and 103 with minimal change in light emission (fold change between -0.05 to +0.05 with a p -value ≥ 0.05 ; no significant RLU change mutants) when exposed to BIT for 24 hours. The fold RLU selection criteria used in this study may not represent true preservative responding mutants. However, it allowed isolation of a wide range of transconjugants that demonstrated various responses to BIT exposure.

Out of the 204 BIT positive responding mutants, Figure 24 indicates the top ten with a significant RLU fold increase (p -value ≤ 0.05) during master plate screening. Mutant 33:F4 had the greatest RLU fold increase of +2.53 (p -value=0.0209) when exposed to BIT at 20% of the MIC for 24 hours. This suggests the transposon-disrupted gene was up-regulated in response to BIT exposure. BIT negative responding mutants represented 24.0% of those isolated from master plate screening, whereas 50.5% of the mutants isolated were BIT positive responders. Therefore, when identifying BIT negative mutants from the libraries, the threshold was lowered from ≥ 0.4 -fold to a ≥ 0.3 -fold RLU decrease, in order to isolate more with a decreased RLU response with BIT. Figure 25 illustrates the top ten mutants selected from the BIT master plate screen with a significant RLU fold decrease (p -value ≤ 0.05). Mutant 28:H11 had the greatest fold decrease of -0.98 (p =0.0088) when exposed to BIT for 24 hours. This suggests that the transposon-disrupted gene was down-regulated in response to BIT exposure. When compared to the BIT positive mutants (Figure 24), the overall fold difference for BIT negative responding mutants was lower (Figure 25).

Mutants demonstrating minimal RLU fold changes when exposed to BIT for 24 hours were also isolated from master plate screening. These mutants were of interest to provide information regarding the RW109 genes not differently regulated when exposed to the preservative. Minimal change mutants had a RLU fold changes between -0.05 to +0.05 and represented 25.5% of those isolated. Figure 26 illustrates the mutants with the minimal amount of RLU fold change when exposed to BIT for 24 hours (p -value ≥ 0.05 – no significant RLU change). Mutant 9:B7 had the least change in light emission with a RLU fold change of -0.0002 (p =0.9921), suggesting the transposon insertion gene did not have an altered gene expression in response to the preservative BIT.

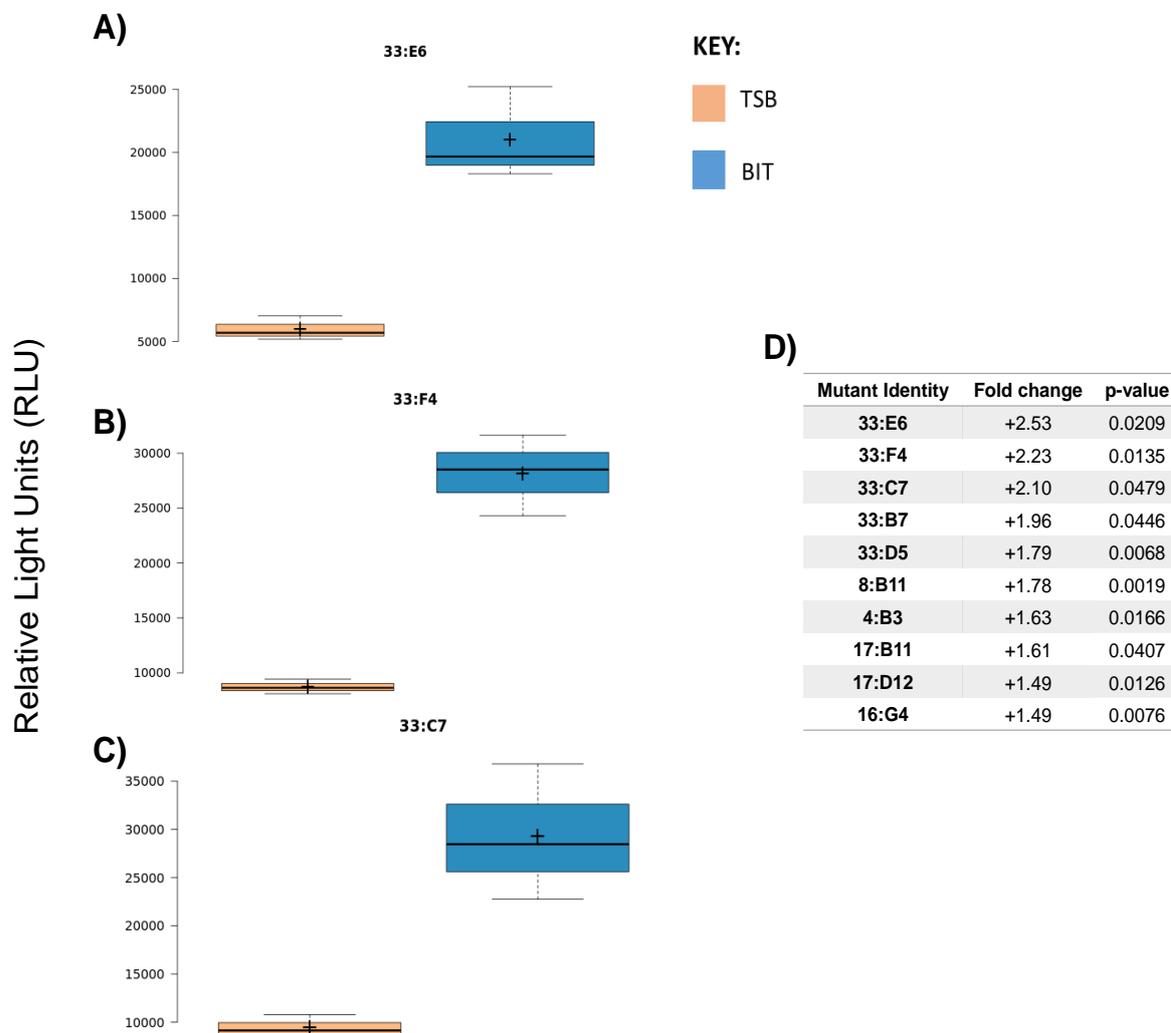


Figure 24. Box plots of mutants with a significant RLU fold increase when exposed to BIT for 24hours

Mutants were exposed to BIT at 20% of the calculated MIC within a master plate screen for 24 hours and RLU measurements taken. The top 3 mutants with a significant RLU fold increase were 33:E6 (A), 33:F4 (B) and 33:C7 (C). On the boxplots the centre lines show the medians and crosses represent the sample means; box limits indicate the 25th and 75th percentiles as determined by R software; whiskers extend 1.5 times the interquartile range from the 25th and 75th percentiles. n = 3 sample points. Panel D demonstrates the RLU fold change and p-values for the top ten BIT positive responding mutants.

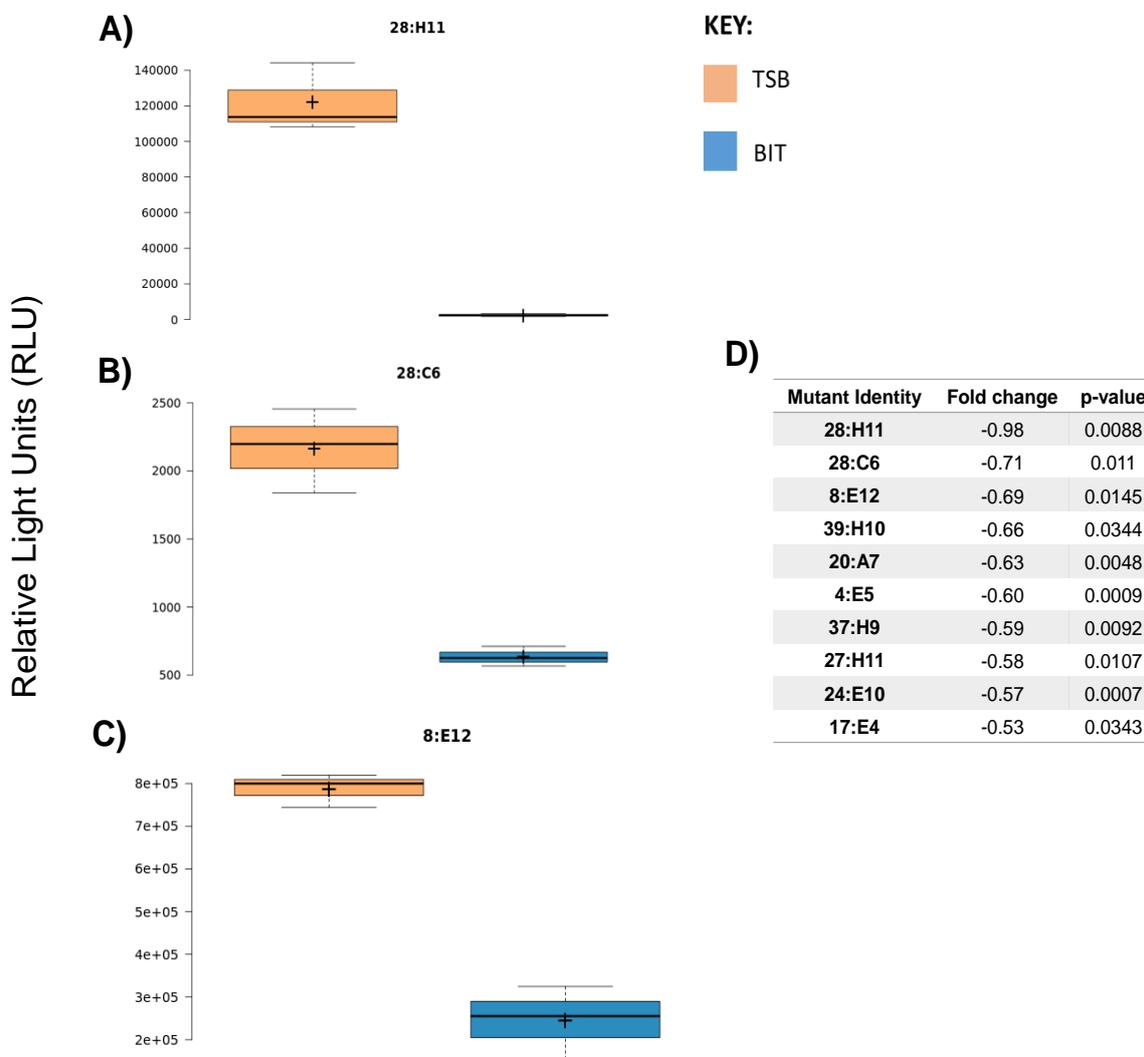


Figure 25. Box plots of mutants with a significant RLU fold decrease when exposed to BIT for 24hours

Mutants were exposed to BIT at 20% of the calculated MIC within a master plate screen for 24 hours and RLU measurements taken. The top 3 mutants with a significant RLU fold decrease were 28:H11 (A), 28:C6 (B) and 8:E12 (C). On the box plots the centre lines show the medians and crosses represent the sample means; box limits indicate the 25th and 75th percentiles as determined by R software; whiskers extend 1.5 times the interquartile range from the 25th and 75th percentiles. n = 3 sample points. Panel D demonstrates the RLU fold change and p-values for the top ten BIT negative responding mutants.

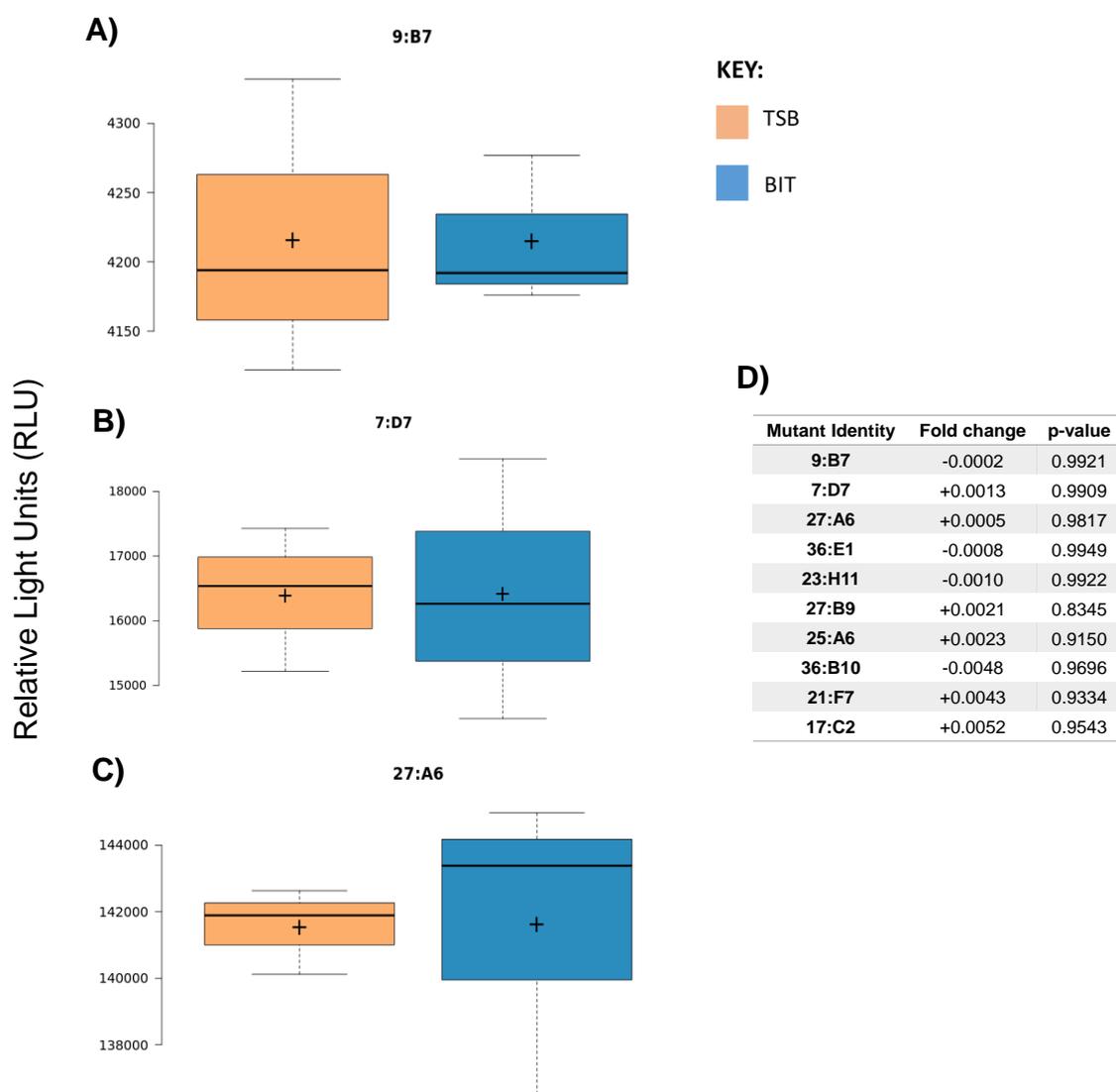


Figure 26. Box plots of mutants with no significant RLU fold change when exposed to BIT for 24hours

Mutants were exposed to BIT at 20% of the calculated MIC within a master plate screen for 24 hours and RLU measurements taken. Three mutants with the least amount of RLU fold change were 9:B7 (A), 7:D7 (B) and 27:A6 (C). On the box plots the centre lines show the medians and crosses represent the sample means; box limits indicate the 25th and 75th percentiles as determined by R software; whiskers extend 1.5 times the interquartile range from the 25th and 75th percentiles. n = 3 sample points. Panel D demonstrates the RLU fold change and p-values for mutants with the least amount of light emission alterations when exposed to BIT.

CHAPTER 4 – Use of an industrial *P. aeruginosa* mini-Tn5-*luxCDABE* mutant library to investigate the genetic basis of preservative resistance

4.2.4 Identifying transposon insertion sites in isolated mutants of interest

The most successful method for identifying transposon insertion sites in isolated mutants was the amplification and sequencing of the DNA flanking the left side of the transposon, as described in Section 2.6.6. Lewenza et al. (2005) illustrates a method for amplifying and sequencing from the right side of the transposon when the left side fails to amplify and was described in Table 15. However when this PCR was adapted for this study, it proved unsuccessful and inconsistent. An arbitrary primed PCR as described by McPhee et al., (2003) and shown in Table 15, was also modified and tested with the RW109 *lux* mutants, but failed to work correctly.

Out of the 404 mutants isolated, 157 demonstrated successful amplification with the transposon insertion site identification method (Section 2.6.6). The inverse PCR method, which focuses on the left side of the transposon, failed to work with the remaining 247 isolated mutants. The PCR conditions were further evaluated by altering the annealing temperatures and magnesium ion concentrations. However, no PCR products were identified, resulting in the insertion sites not being identified for these mutants.

The resulting PCR products, which successfully amplified using the insertion identification method, were purified and sent away for sequencing. These represented 85 BIT positive mutants, 36 BIT negative mutants and 36 mutants with minimal RLU change in response to BIT. In total 53 distinct insertion sites were identified (Table 16, Table 17 and Table 18). Out of these mutants, 19 were classified only within BIT positive responding mutants, 16 represented only BIT negative responders and 14 were only identified within mutants with minimal RLU change when exposed to BIT.

Four RW109 genes were identified as insertion sites in mutants from different BIT response categories and all encoded for hypothetical proteins. RW109_06287 was identified in 3 BIT positive mutants and 1 minimal RLU change mutant (Table 16 and Table 18), RW109_04396 was found in 1 BIT positive mutant and 1 minimal RLU change mutant (Table 16 and Table 18) and RW109_05920 was identified in 1 BIT negative and 3 minimal RLU change mutants (Table 17 and Table 18). The RW109_07253 gene was identified in all 3 BIT response categories; in BIT positive and negative mutants it was found twice and in the minimal RLU change mutants it was identified nine times (Table 16, Table 17 and Table 18). Interestingly, the sequence obtained for the BIT negative mutant 25:H3 returned no significant hits when compared against the RW109 genome (Table 17).

Table 16. Transposon-interrupted genes found in BIT positive mutants

COG Category	Mutant ID	Bioluminescence category	Master plate screening RLU Fold Change	Gene ID	Gene annotation
Cell wall/membrane/envelope biogenesis (M)	31:B3	3	+0.46	RW109_06517	Small-conductance mechanosensitive channel
	8:E2	3	+0.93		
Energy production and conversion (C)	8:B11*	3	+1.78	RW109_04996	Periplasmic nitrate reductase precursor
	32:G7	2	+1.23	RW109_05781	Bifunctional enzyme CysN/CysC
Inorganic ion transport and metabolism (P)	37:C5	3	+0.73	RW109_01716	Ferrous-iron efflux pump FieF
	28:B12	3	+0.49	RW109_00596	D-glycero-beta-D-manno-heptose-1,7-bisphosphate 7-phosphatase
Amino acid transport and metabolism (E)	4:B3* (40 mutants in total)	All category 4	+1.63 (Mean RLU fold change = +0.86)	RW109_01055 [£]	Bacterial extracellular solute-binding proteins, family 3
	3:B3	4	+0.63		
Post-translational modification, protein turnover, and chaperones (O)	3:B4	4	+0.8	RW109_01054	Modulator of FtsH protease HflK
	28:G6	4	+0.52		
	32:C10	2	+0.72	RW109_04951	Quercetin 2,3-dioxygenase
	33:F2	3	+0.65		
Transcription (K)	14:C4	2	+0.64	RW109_04041	Transcriptional regulator KdgR
Replication, recombination and repair (L)	37:D10	3	+0.77	RW109_06956	DNA-binding protein HU-beta
	26:F11	2	+0.58		
Function unknown (S)	27:C9	2	+0.48	RW109_06287	Hypothetical protein
	28:A12	2	+0.53		
	31:C5	2	+0.45		
	37:A7	2	+0.6	RW109_07276	Toxin co-regulated pilus biosynthesis protein Q
	27:E4	2	+0.58		
	11:E12	5	+0.99		
	13:C9	4	+0.86		
	13:G9	4	+0.9		
	13:F12	5	+0.87		
	16:F4	5	+1.03		
	16:G4	5	+1.49		
	16:F5	5	+1.15		
	16:G6	5	+1.12		
	37:D1	2	+0.83		
	37:B3	2	+0.7		
	37:E9	4	+1.00	RW109_05921	Hypothetical protein
	37:G9	2	+0.47		
	38:G6	3	+0.47		
	38:C2	5	+0.6		
	38:E2	5	+0.45		
	38:G2	5	+0.55	RW109_07026	Hypothetical protein
	38:A6	3	+0.56		
	10:F4	4	+0.64	RW109_04396#	Hypothetical protein
	24:H2	4	+0.45		
	29:A10	4	+0.53	RW109_00597	Hypothetical protein
	8:D3	4	+0.75	RW109_05512	Hypothetical protein
	9:C10	4	+0.71	RW109_06265	Hypothetical protein
	12:A2	3	+0.42	RW109_07105	Endonuclease/Exonuclease/phosphatase family protein
	22:C6	4	+0.45	RW109_00433	Fic/DOC family protein
	Not similar to any COG	28:D10	2	+0.67	RW109_07253~
28:F12		2	+0.68		

Footnotes: *Mutants with the highest RLU fold increase with identified transposon insertion sites; [£] Insertion site RW109_01055 was found in 40 BIT positive mutants; # indicates mutant insertion sites identified in two different BIT response categories; ~ RW109_07253 encodes a hypothetical protein which was identified as an insertion site in BIT positive, negative and no change mutants.

Table 17. Transposon-interrupted genes found in BIT negative mutants

COG Category	Mutant ID	Bioluminescence category	Master plate screening RLU Fold Change	Gene ID	Gene annotation
Carbohydrate transport and metabolism (G)	12:D11	3	-0.5	RW109_02629	Phosphogluconate dehydratase
	14:C2	4	-0.33		
	8:E12*	2	-0.69	RW109_06768	Tripartite ATP-independent periplasmic transporters, DctQ component
Amino acid transport and metabolism (E)	12:B3	3	-0.41	RW109_06692	Glycine betaine transporter periplasmic subunit
Intracellular trafficking, secretion, and vesicular transport (U)	13:D11	3	-0.38	RW109_04175	Type I secretion system membrane fusion protein PrsE
Cell wall/membrane/envelope biogenesis (M)	13:A12	2	-0.48	RW109_05246	Cryptic autophosphorylating protein tyrosine kinase Etk
Inorganic ion transport and metabolism (P)	27:B2	3	-0.39	RW109_02452	Acid-activated urea channel
	27:B7	3	-0.36		
	27:H8	3	-0.37		
	27:E11	3	-0.36		
	28:D2	3	-0.41		
	31:F2	3	-0.39		
	33:B8	3	-0.35		
37:H9*	3	-0.59			
Energy production and conversion (C)	18:A3	3	-0.47	RW109_04704	Cytochrome bo(3) ubiquinol oxidase subunit 1
Transcription (K)	38:A8	2	-0.48	RW109_06908	HTH-type transcriptional regulator PuvR
Function unknown (S)	22:C7	2	-0.34	RW109_05170	Putative metallo-hydrolase YycJ
	27:F9	4	-0.54	RW109_00750	2-oxoglutarate-dependent ethylene/succinate-forming enzyme
	27:H11*	3	-0.58		
	31:H4	4	-0.47		
	24:E11	3	-0.42	RW109_00047	PhoD-like phosphatase
	24:F11	3	-0.42		
	24:F12	2	-0.39		
	15:H1	2	-0.44	RW109_04388	Phosphoenolpyruvate synthase regulatory protein
	16:C10	2	-0.5		
No COG	15:D4	4	-0.45	RW109_03550	Hemolysin precursor
	28:H12	2	-0.98	RW109_07253~	Hypothetical protein
	36:D12	2	-0.33	RW109_05920#	Hypothetical protein
	30:A11	4	-0.38	RW109_00153	Cyclic AMP receptor-like protein
	4:E5*	3	-0.6	RW109_00117	Hypothetical protein
	37:D12	3	-0.52		
	38:A2	3	-0.48		
	38:H3	4	-0.53	RW109_02284	Hypothetical protein
	38:D7	3	-0.37		
38:C9	3	-0.43	No Significant hits	No Significant hits	
25:H3	4	-0.48			
N/A	25:H3	4	-0.48	No Significant hits	No Significant hits

Footnotes: *Mutants with the highest RLU fold decrease with identified transposon insertion sites; # indicates mutant insertion sites identified in two different BIT response categories; ~ RW109_07253 encodes a hypothetical protein which was identified as an insertion site in BIT positive, negative and no change mutants.

Table 18. Transposon-interrupted genes found in mutants with minimal RLU change in response to BIT

COG Category	Mutant ID	Bioluminescence category	Master plate screening RLU Fold Change	Gene ID	Gene annotation
Amino acid transport and metabolism (E)	37:F6	1	+0.0539	RW109_05316	Phenylalanine-4-hydroxylase
Intracellular trafficking, secretion, and vesicular transport (U)	2:A2	5	+0.018	RW109_07271	Toxin coregulated pilus biosynthesis protein E
Defense mechanisms (V)	1:E6	4	+0.0077	RW109_04032	Multidrug efflux pump subunit AcrB
Nucleotide transport and metabolism (F)	26:D11	3	+0.0318	RW109_06702	8-oxoguanine deaminase
	28:D5	3	+0.0229		
	31:F3	3	+0.0299		
Replication, recombination and repair (L)	29:F7	4	-0.0127	RW109_06040	DNA topoisomerase 1
	34:A11	4	+0.0093		
Post-translational modification, protein turnover, and chaperones (O)	39:C12	3	+0.0108	RW109_00874	Metalloprotease LoiP precursor
Transcription (K)	37:F8	4	-0.0119	RW109_05312	HTH-type transcriptional regulator DmlR
	7:E7	2	+0.02	RW109_06238	HTH-type transcriptional repressor YcgE
	38:B3	2	+0.0251	RW109_02216	Oxygen regulatory protein NreC
Function unknown (S)	27:G6	4	+0.0383	RW109_03466	Translocation and assembly module TamB
	28:D4	4	+0.0103		
	28:H9	4	+0.0213		
	29:A11	4	+0.024		
	32:D8	4	+0.055		
	28:C11	3	+0.0201	RW109_02100	Hypothetical protein
	29:E11	3	+0.0315		
	30:A11	4	+0.0422	RW109_05920#	Hypothetical protein
	31:C4	3	+0.0232		
	32:G12	4	+0.0495	RW109_04396#	Hypothetical protein
	9:B7	4	-0.0002		
	28:G5	4	+0.0379		
31:B8	4	+0.0307			
39:D8	3	-0.0142			
27:C7	4	+0.0161			
No COG	23:H11	2	-0.001	RW109_07253~	Hypothetical protein
	24:A11	2	+0.0243		
	27:B9	2	+0.0021		
	29:C11	2	+0.0501		
	29:G6	2	-0.0372		
	30:G5	2	-0.0222		
	31:B6	2	+0.0417		
	31:D11	2	+0.047		
	37:C7	2	+0.0274		

Footnotes: # indicates mutant insertion sites identified in two different BIT response categories; ~ RW109_07253 encodes a hypothetical protein which was identified as an insertion site in BIT positive, negative and no change mutants.

CHAPTER 4 – Use of an industrial *P. aeruginosa* mini-Tn5-luxCDABE mutant library to investigate the genetic basis of preservative resistance

4.2.5 The COG category distributions of identified transposon interrupted genes

The transposon insertion site genes were categorised based on the Clusters of Orthologous Groups (COGs) annotation of the RW109 strain (Chapter 3, Section 3.2.3.1) and 13 COG clusters were determined (Figure 27). Out of the 53 distinct insertion sites, 49.06% were represented by the poorly characterised COG category (S and No COG), with 16 out of 26 encoding for hypothetical proteins. The other 10 poorly characterised insertion sites had predicted Prokka annotations as illustrated in Table 16, Table 17 and Table 18. COG category distribution comparisons of mutant insertion sites, did not include the 4 hypothetical protein encoding genes identified in more than one BIT response group (Section 4.2.4). BIT positive responding mutants represented the highest number of poorly characterised (S and No COG) insertions sites (n=9), followed by BIT negative mutants (n=8) and mutants with minimal RLU change (n=5).

Within the well-characterised COG groups, BIT positive mutants were associated with 7 categories (Figure 27). These included, inorganic ion transport and metabolism (n=2), amino acid transport and metabolism (n=2), post-translational modification, protein turnover, and chaperones (n=2), cell wall/membrane/envelope biogenesis (n=1), energy production and conversion (n=1), transcription (n=1), replication, recombination and repair (n=1). Seven COG categories were also identified in BIT negative mutants (Figure 27) and these included, carbohydrate transport and metabolism (n=2), intracellular trafficking, secretion, and vesicular transport (n=1), cell wall/membrane/envelope biogenesis (n=1), energy production and conversion (n=1), transcription (n=1), inorganic ion transport and metabolism (n=1) and amino acid transport and metabolism (n=1). Mutants with minimal RLU change when exposed to BIT were also linked to 7 COG categories (Figure 27). These included transcription (n=3), defence mechanisms (n=1), nucleotide transport and metabolism (n=1), replication, recombination and repair (n=1), post-translational modification, protein turnover, and chaperones (n=1), intracellular trafficking, secretion, and vesicular transport (n=1) and amino acid transport and metabolism (n=1).

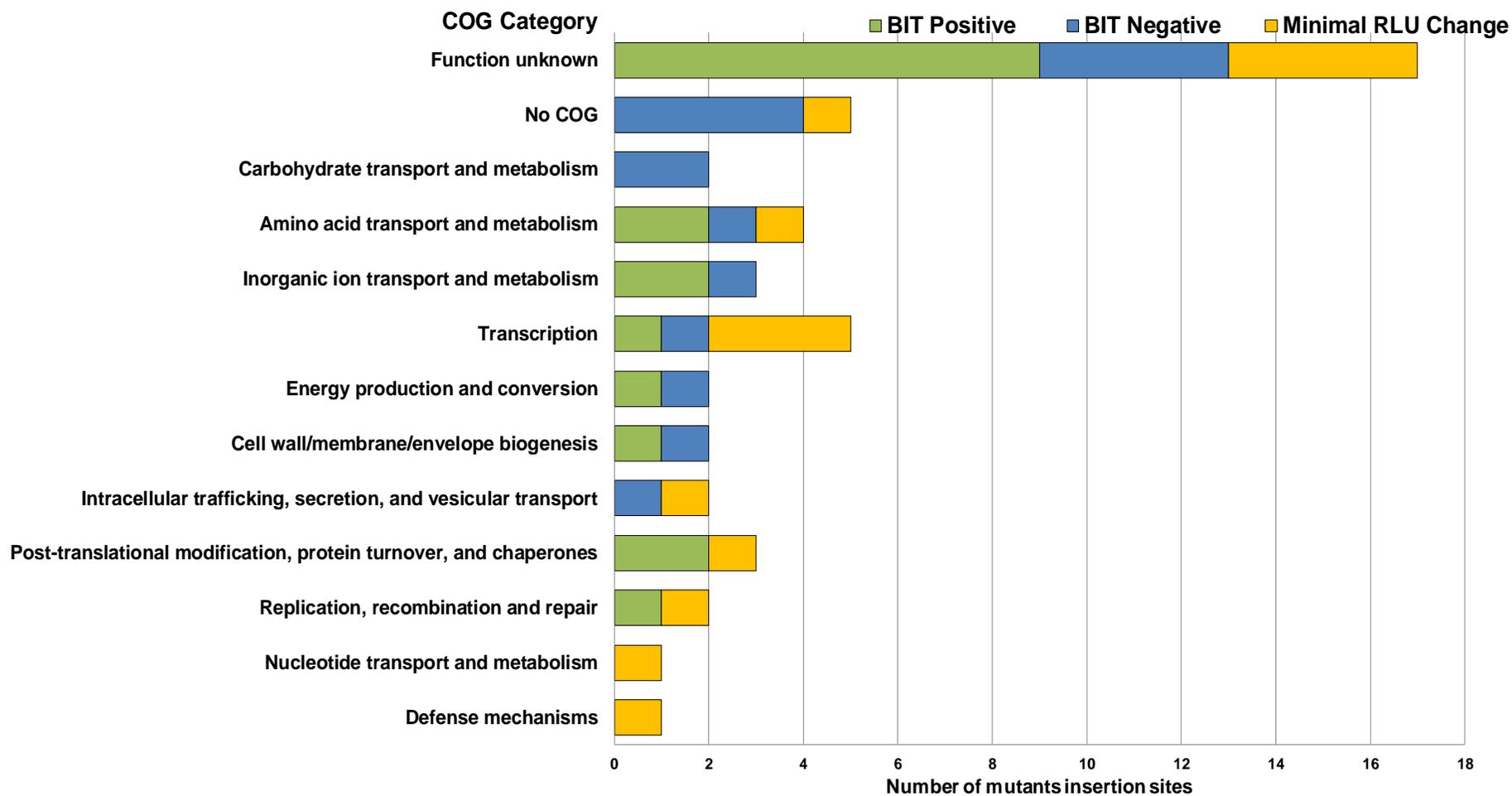


Figure 27. The distribution of identified transposon interrupted genes based on COG categories

The green bars represent transposon interrupted genes from mutants with a RLU fold increase (BIT positive), blue bars correspond to genes where a RLU fold decrease was observed (BIT negative) and yellow bars show genes from mutants with minimal RLU change in response to BIT. This analysis does not include the 4 insertion sites highlighted in Section 4.2.4 which were classified in mutants from different BIT response categories.

CHAPTER 4 – Use of an industrial *P. aeruginosa* mini-Tn5-luxCDABE mutant library to investigate the genetic basis of preservative resistance

4.2.6 Distribution of the distinct transposon insertion sites throughout the RW109 whole genome

Out of the 53 distinct insertion sites, 45 were identified on the main chromosome. With the exception of a large region between ~2.2 and 2.9 Mbp, the 45 insertion sites were evenly distributed throughout the RW109 main chromosome (Figure 28). Mutant insertion sites were also found on the plasmids, with 5 on plasmid 1 and 3 on plasmid 2.

On plasmid 1, the 5 insertion sites genes encoded 2 hypothetical proteins (RW109_00038 and RW109_00117), a PhoD-like phosphatase (RW109_00047), a Cyclic AMP receptor-like protein (RW109_00153) and a Fic/DOC family protein (RW109_00433). All plasmid 1 transposon insertion genes were classified within the poorly characterised COG categories. The RW109_00047 and RW109_00117 genes were both identified within 3 BIT negative mutants and RW109_00153 was found within 1 BIT negative mutant (Table 17). RW109_00433 was classified within 1 BIT positive mutant (Table 16) and RW109_00038 was found in a mutant with minimal RLU change in response to BIT (Table 18).

The 3 insertion sites identified on plasmid 2, encoded for a hypothetical protein (RW109_07253) and the toxin co-regulated pilus biosynthesis proteins E and Q (RW109_07271 and RW109_07276, respectively). RW109_07253 and RW109_07276 were classified within the poorly characterised COG category and RW109_07271 was annotated within the intracellular trafficking, secretion and vesicular transport (U) COG category. The RW109_07253 gene, was a previously noted insertion site, identified in mutants from three different BIT response categories (Section 4.2.4). The genes encoding for the two toxin co-regulated pilus biosynthesis proteins were not classified within the same BIT response group. The RW109_07271 gene was identified in a mutant with minimal RLU change in response to BIT (Table 18) and 3 BIT positive mutants were identified with insertion sites in the RW109_07276 gene (Table 16). The RW109_07271 gene was identified within a predicted operon on RW109 plasmid 2. This operon had high sequence similarity to a *P. aeruginosa* S04 90 plasmid type II secretion system and RW109_07276 was also found in close proximity to this operon (Chapter 3, Section 3.2.3.1.4).

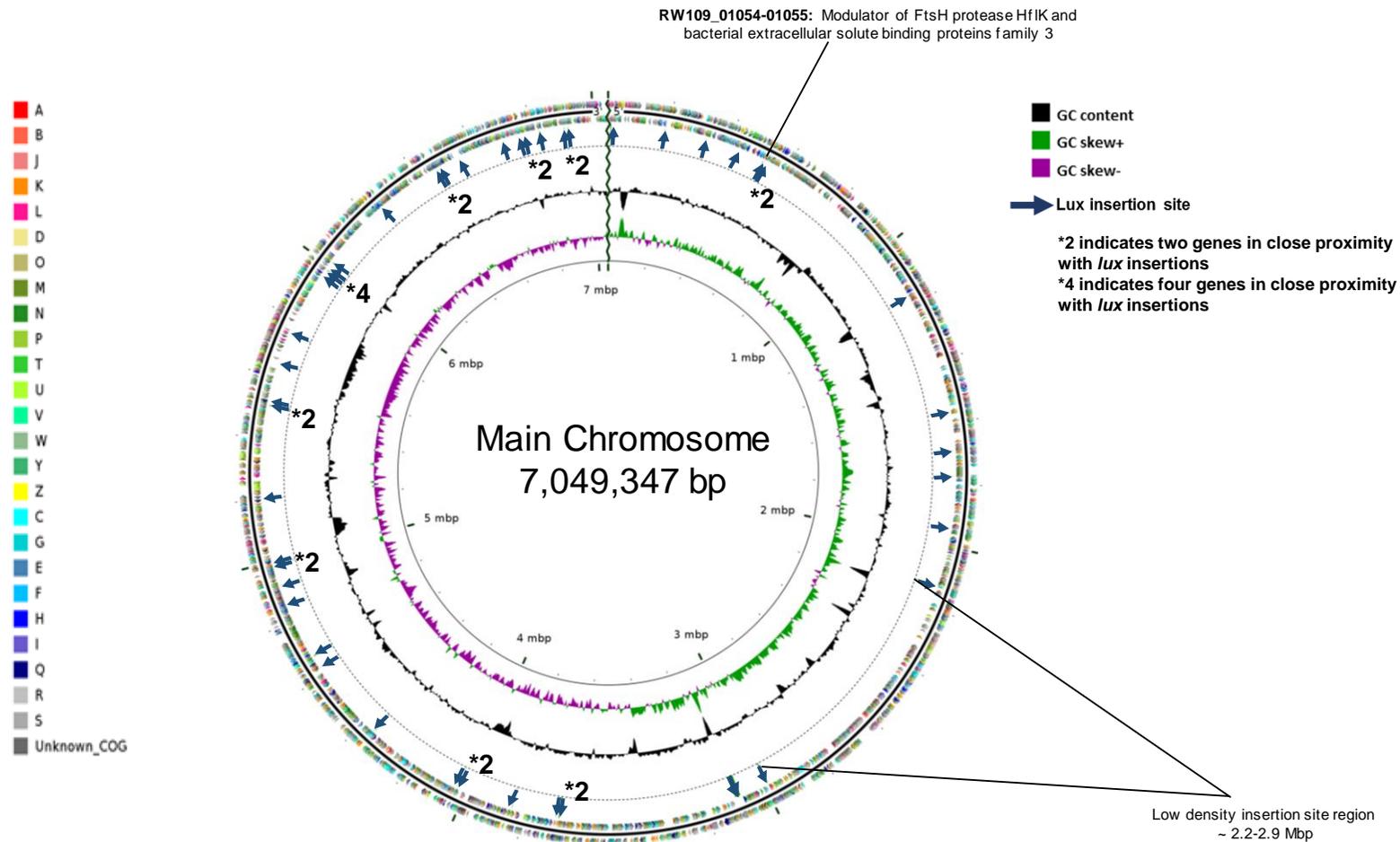


Figure 28. Distribution of the distinct *lux* transposon insertion sites throughout the main chromosome of RW109

On the main chromosome of RW109, 45 distinct transposon insertion sites were identified. On the map from the outermost circle and moving inwards the following tracks were identified: (1) predicted CDS on forward strand coloured according to COG categories, (2) predicted CDS on reverse strand coloured according to COG categories, (3) transposon insertion sites indicated by blue arrows, (4) GC content (black), (5) positive and negative GC skew (green and purple, respectively) and (6) genome region by mbp. The size of the main chromosome was not drawn to scale.

4.2.7 Key insertion sites of BIT positive responding mutants

Out of the 85 BIT positive mutants which had successful inverse PCR amplification, 40 had transposon insertion sites in a gene encoding for the bacterial extracellular solute binding proteins family 3 (gene locus: RW109_01055). This protein has multiple functions which include chemoreceptors, components of molecular transport systems and initiators of signal transduction pathways (Tam and Saier, 1993, Sauer and Camper, 2001). The location of this insertion site in RW109 was also indicated in Figure 28. A mean RLU fold increase of +0.86 was observed for the 40 mutants, which were all from bioluminescence category 4. Of the mutants identified with this insertion site, 4:B3 had the greatest RLU increase of +1.63-fold (p-value ≤ 0.05) (Table 16). Mutant 4:B3 was also screened with BIT individually, as previously described in Section 2.6.5.2, to ensure a reproducible RLU response and a +1.79-RLU fold increase was observed (p-value ≤ 0.05) (Figure 29). In this study gene operons were predicted using the *Pseudomonas* database DIAMOND BLASTP tool (Winsor et al., 2016).

Two BIT positive mutants 3:B3 and 3:B4, were found to have transposon insertion sites in the gene locus RW109_01054, which encodes a modulator of FtsH protease HflK. This gene was located within the same three-gene operon as the bacterial extracellular solute binding proteins family 3 (Figure 29). The position for this gene within RW109 was also indicated in Figure 28. When mutants 3:B3 and 3:B4 were exposed to BIT for 24 hours in a master plate, a +0.63 and +0.80 fold RLU increase was observed respectively (p-value ≤ 0.05) (Table 16). The RLU response of 3:B4 when exposed to BIT was found to be reproducible when the mutant was screened individually (Figure 29). The gene product HflK has previously been associated with tobramycin antibiotic resistance in *P. aeruginosa* (Hinz et al., 2011). Results observed from the mutants suggest the genes encoding for the bacterial extracellular solute binding proteins family 3 and modulator of FtsH protease (HflK) were up-regulated when exposed to BIT at 20% of the MIC.

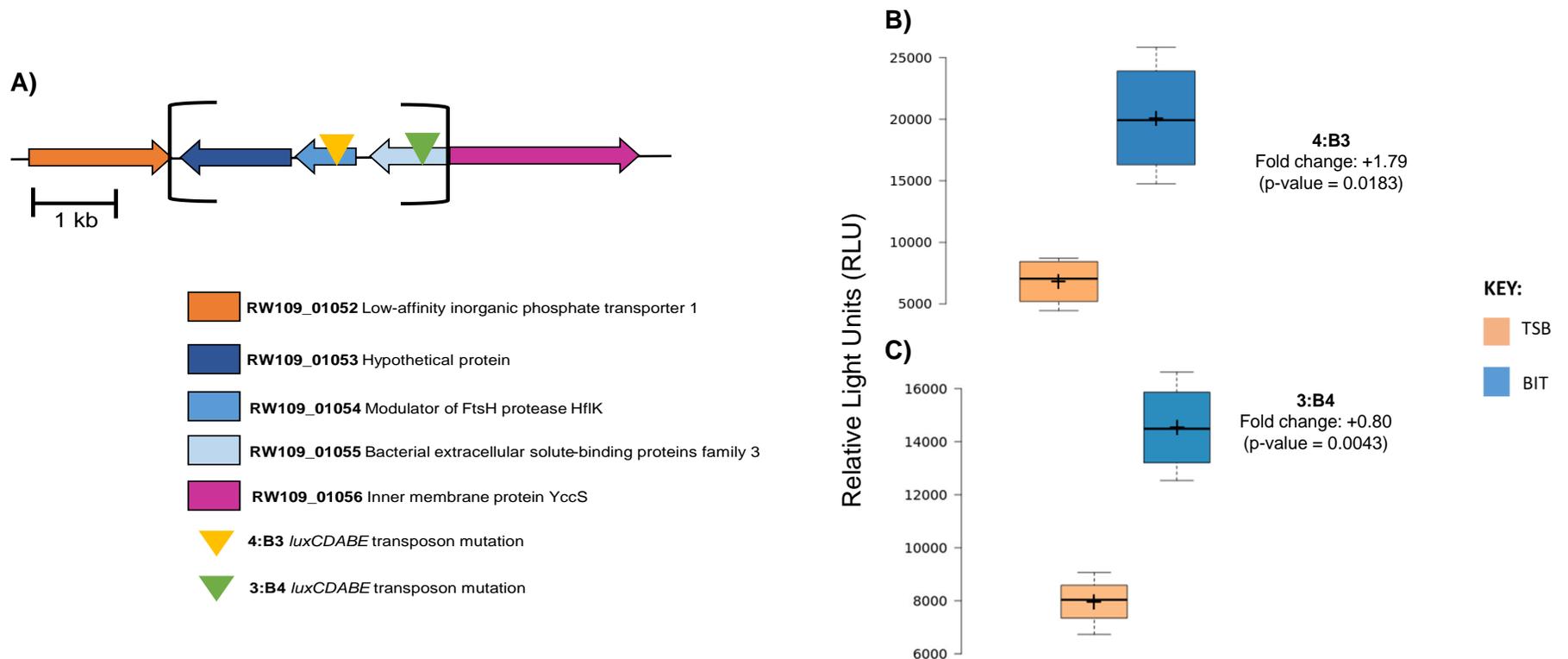


Figure 29. The *luxCDABE* transposon mutation of genes encoding a bacterial extracellular solute-binding transport system operon and their RLU responses when exposed to BIT for 24 hours

The schematic (A) shows the arrangement of three genes within an operon encoding proteins of a bacterial extracellular solute-binding transport system (loci: RW109_01052-01056). Each gene is drawn to the scale given by the 1 kb bar. The putative products of the genes are shown by the colour coded key, and the site of transposon-insertion for mutants 4:B4 and 3:B4 are labelled; both these mutants were found to be BIT positive.

Box plots of mutants 4:B4 (B) and 3:B4 (C) indicates the change in light emission when exposed to BIT for 24 hours. Centre lines show the medians and crosses represent the sample means; box limits indicate the 25th and 75th percentiles as determined by R software; whiskers extend 1.5 times the interquartile range from the 25th and 75th percentiles (n= 4 sample points). The RLU fold changes and corresponding p-values are indicated next to the box plots.

CHAPTER 4 – Use of an industrial *P. aeruginosa* mini-Tn5-*luxCDABE* mutant library to investigate the genetic basis of preservative resistance

Mutant 14:C4 when exposed to BIT in a master plate screen demonstrated a +0.64-fold RLU increase (p-value ≤ 0.05) (Table 16). When 14:C4 was isolated and revived on TSA an overproduction of a red/brown pigment was observed which was not detected with the wild-type strain and other isolated mutants (Figure 30). *P. aeruginosa* produces a red/brown pigment called pyorubin when under oxidative stress conditions (Ferguson *et al.*, 2007) and this pigment is produced by highly virulent clinical strains (Kaleli *et al.*, 2007, Rajyaguru *et al.*, 2014). The transposon insertion site was identified in a gene encoding for a transcriptional regulator (KdgR) (gene locus: RW109_04041) and was directly located next to the *gnyRDBHAL* gene operon (loci: RW109_04035-04040) (Figure 31). This gene operon is involved in acyclic isoprenoid degradation (Diaz-Perez *et al.*, 2004, Förster-Fromme *et al.*, 2006) and isoprenoids have previously been associated with bacterial pigment production (Cañizares-Villanueva *et al.*, 1997, Saha *et al.*, 2008). When 14:C4 was screened with BIT as an individual mutant, a greater RLU fold increase of +1.78-fold RLU was observed (p-value ≤ 0.05) when compared to the result obtained from master plate screening (Figure 31 and Table 16). The increased light emission demonstrated with this mutant, suggests the transposon insertion gene was up-regulated in response to BIT at 20% of the MIC.

Another mutant of interest identified from the BIT master plate screening was 8:B11, where a +1.78-fold RLU was observed (p-value ≤ 0.05). This was the largest RLU fold increase observed during master plate screening with a BIT positive mutant (Table 16). The transposon insertion site was identified in the periplasmic nitrate reductase precursor, *napA* gene (RW109_04996) of the *napEFDABC* operon (loci: RW109_04993-04998) (Figure 32). This operon functions in the anaerobic regulatory network for *P. aeruginosa* nitrate respiration under anoxic conditions (Schreiber *et al.*, 2007, Van Alst *et al.*, 2009). Growth conditions within a biofilm are characterised as anoxic especially within the cystic fibrosis lung (Van Alst *et al.*, 2007). When this mutant was screened individually with BIT, a +2.33-fold RLU increase was observed (p-value ≤ 0.05), which was a greater fold change when compared to the master plate screening result (Figure 32 and Table 16). The RLU-fold increase observed with 8:B11 suggests the *napA* gene, within the *napEFDABC* operon, was up-regulated during exposure to the preservative BIT at 20% of the MIC.

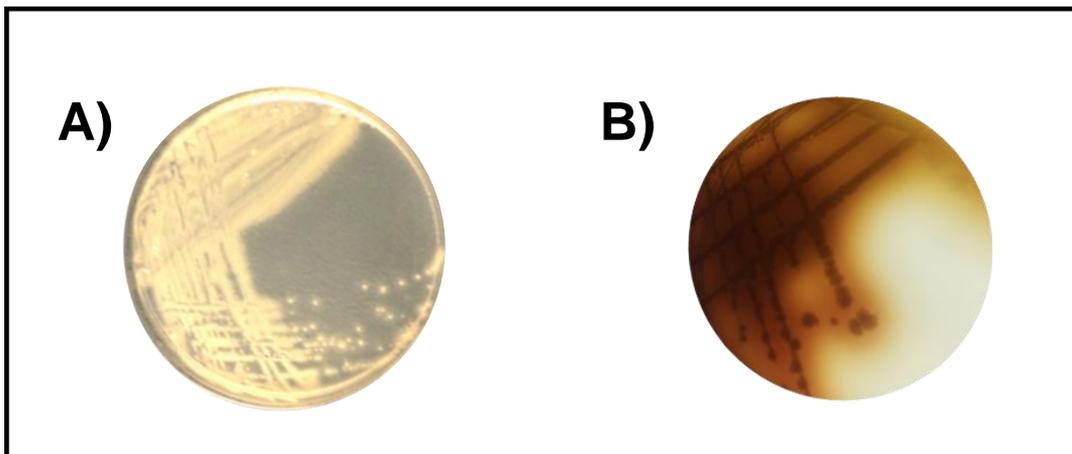


Figure 30. The overproduction of a brown/red pigment by the BIT positive mutant
The wild-type RW109 strain (A) and a BIT positive transposon mutant 14:C4 (B) were both grown for 24 hours at 30°C on TSA. When compared to A, the mutant shown in B was found to overproduce a brown/red pigment.

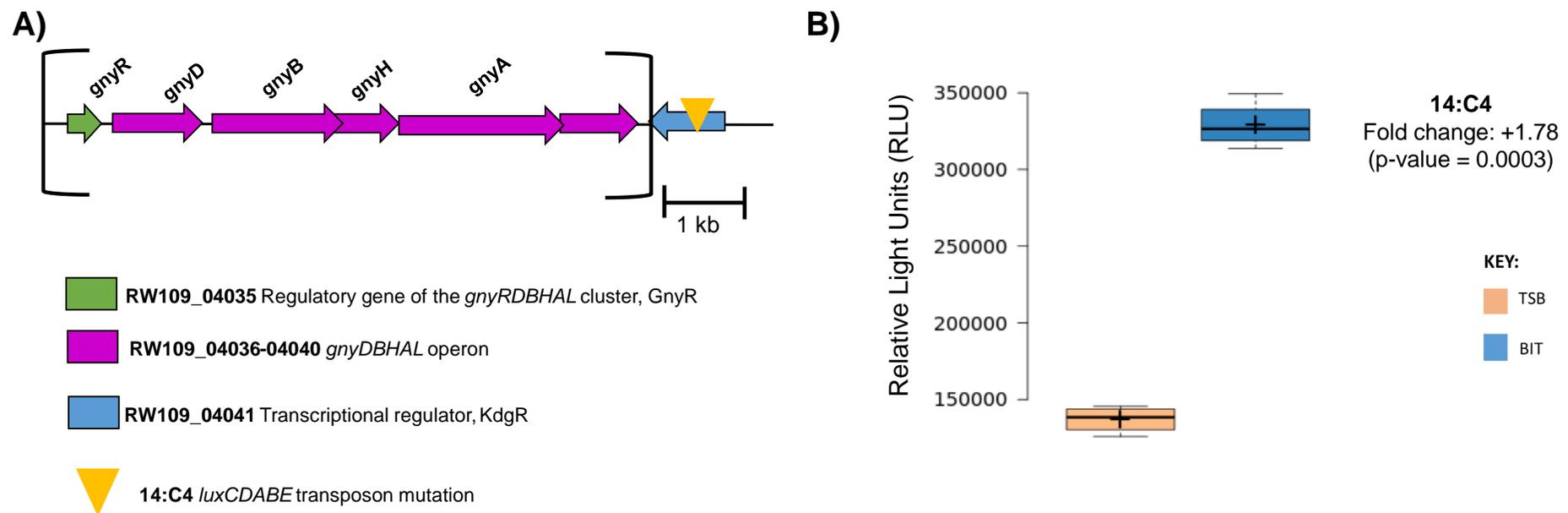


Figure 31. The *luxCDABE* transposon mutation of the gene encoding a KdgR transcriptional regulator

The schematic shows the site of transposon-insertion for the BIT positive mutant 14:C4 (A). The transposon mutation gene is located directly next to the *gnyRDBHAL* gene operon (loci: RW109_04035-04040) which is involved in acyclic isoprenoid degradation. Each gene is drawn to the scale given by the 1 kb bar and the putative products are shown by the colour coded key. A box plot of 14:C4 (B) indicates the change in light emission when exposed to BIT for 24 hours. Centre lines show the medians and crosses represent the sample means; box limits indicate the 25th and 75th percentiles as determined by R software; whiskers extend 1.5 times the interquartile range from the 25th and 75th percentiles (n= 4 sample points). The RLU fold change and corresponding p-value is shown next to the box plot.

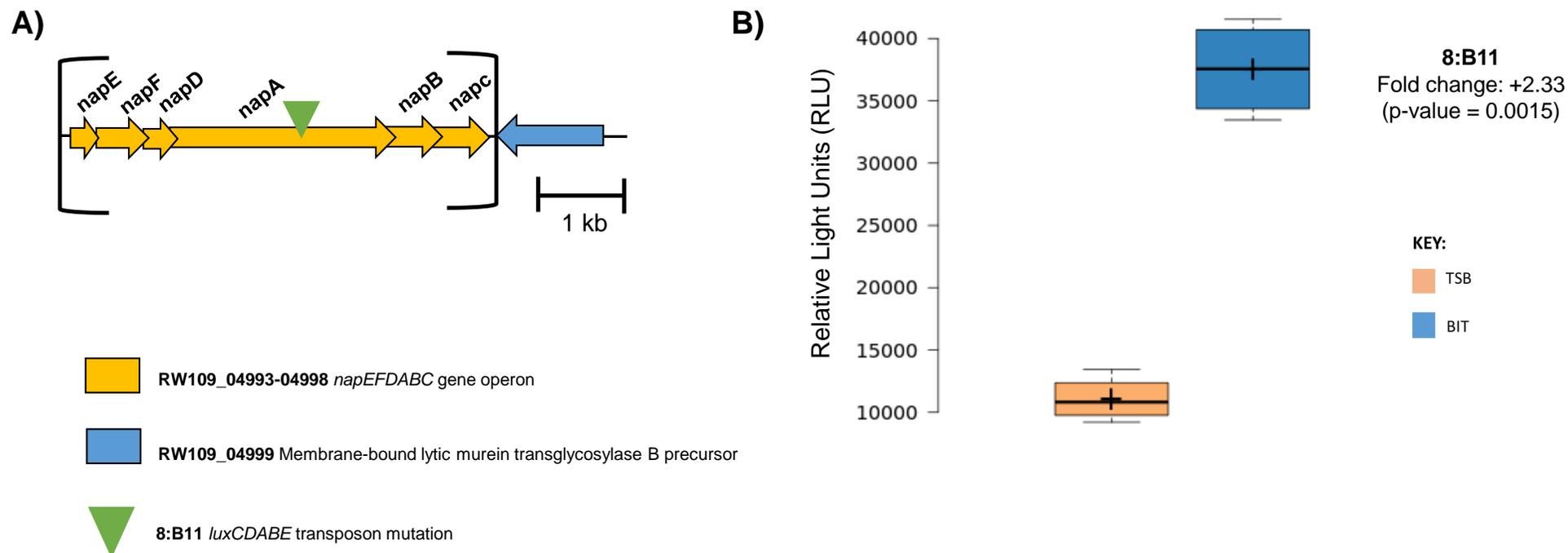


Figure 32. The *luxCDABE* transposon mutation of the gene within the *napEFDABC* operon

The schematic shows the site of transposon-insertion for mutant 8:B11 (A) which was found to be BIT positive. The transposon mutation gene is located in the *napA* gene of the *napEFDABC* operon (loci: RW109_04993-04998) which is part of the anaerobic regulatory network for *P. aeruginosa* nitrate respiration. Each gene is drawn to the scale given by the 1 kb bar and the putative products are shown by the colour coded key. A box plot of 8:B11 (B) indicates the change in light emission when exposed to BIT for 24 hours. Centre lines show the medians and crosses represent the sample means; box limits indicate the 25th and 75th percentiles as determined by R software; whiskers extend 1.5 times the interquartile range from the 25th and 75th percentiles (n= 4 sample points). The RLU fold change and corresponding p-value is shown next to the box plot.

CHAPTER 4 – Use of an industrial *P. aeruginosa* mini-Tn5-luxCDABE mutant library to investigate the genetic basis of preservative resistance

A transposon insertion site in a gene encoding for an efflux pump was also identified during master plate screening with BIT. The mutant 37:C5 was found to increase RLU by +0.73-fold (p-value ≤ 0.05) (Table 16) and its insertion site was found within a ferrous-iron efflux pump, FieF (loci RW109_01716) (Figure 33). FieF is an iron-efflux transporter, which is responsible for iron detoxification (Grass *et al.*, 2005). When 37:C5 was screened individually with BIT, a +2.09-fold RLU increase was observed (p-value ≤ 0.05), which was a greater fold change when compared to the master plate screening result (Figure 33 and Table 16). The RLU-fold induction observed with this mutant, suggests the FieF efflux pump was up-regulated during exposure to the preservative BIT at 20% of the MIC.

A gene encoding for a bi-functional enzyme CysN, (RW109_05781) was identified as the insertion site for the BIT positive mutant 32:G7. A RLU fold increase of +1.23 (p-value ≤ 0.05) was observed when the mutant was screened within a master plate (Table 16). The enzyme encoded by this gene functions in the metabolism of sulphur in *P. aeruginosa* within the *cysND* operon (Hummerjohann *et al.*, 1998). When screened individually, a +1.70-fold RLU increase was observed (p-value ≤ 0.05) (Figure 34), suggesting that this insertion site gene was up-regulated in response to the preservative BIT at 20% of the MIC.

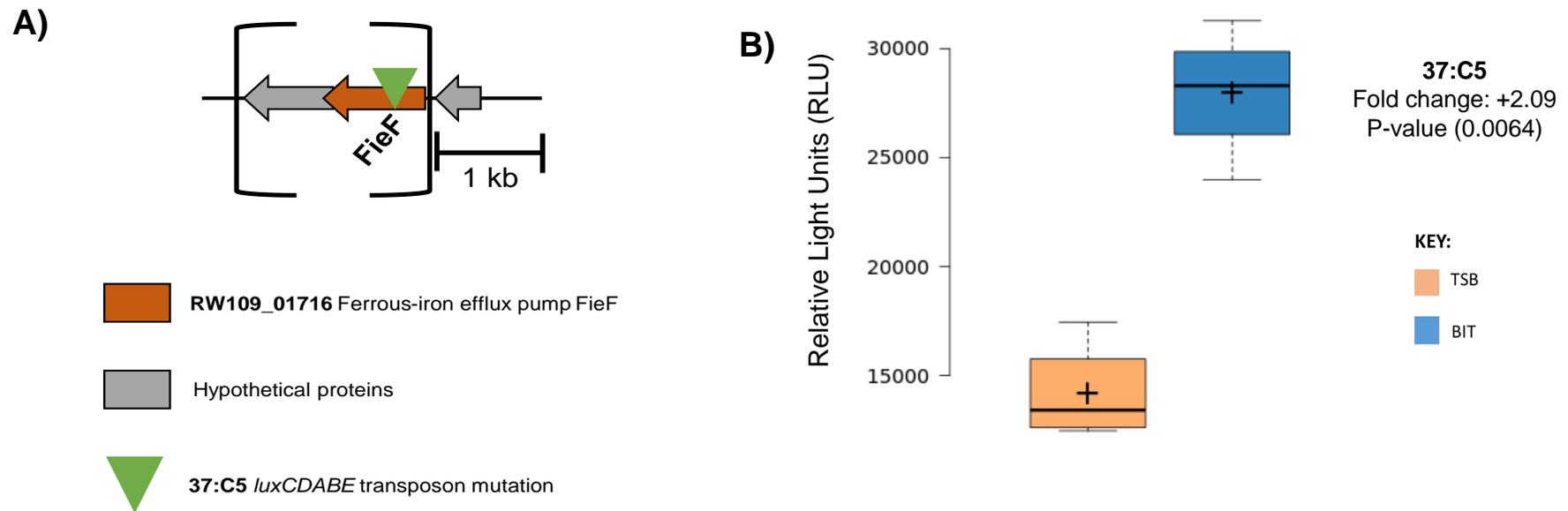


Figure 33. The *luxCDABE* transposon mutation of a gene encoding a ferrous-iron efflux pump

The schematic shows the site of transposon-insertion for mutant 37:C5 (A) which was found to be BIT positive. The transposon mutation is located in a gene encoding a ferrous-iron efflux pump FieF (loci: RW109_01716) which is involved in metal cation transportation. Each gene is drawn to the scale given by the 1 kb bar and the putative products are shown by the colour coded key. A box plot of 37:C5 (B) indicates the change in light emission when exposed to BIT for 24 hours. Centre lines show the medians and crosses represent the sample means; box limits indicate the 25th and 75th percentiles as determined by R software; whiskers extend 1.5 times the interquartile range from the 25th and 75th percentiles (n= 4 sample points). The RLU fold change and corresponding p-value is shown next to the box plot.

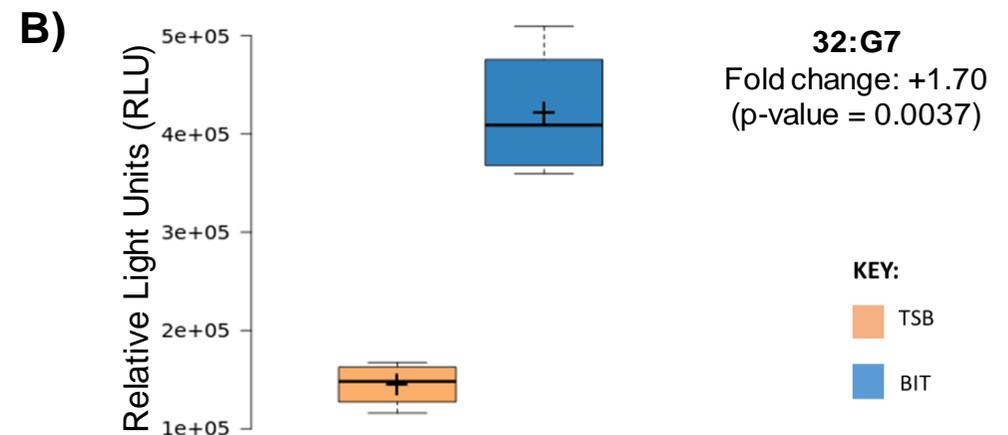
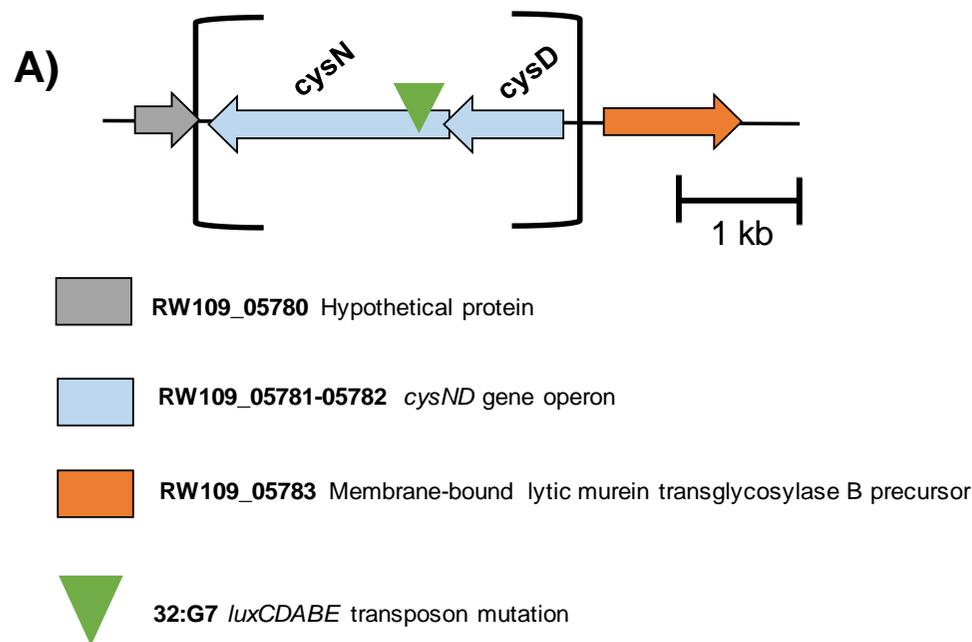


Figure 34. The *luxCDABE* transposon mutation of gene encoding a bifunctional enzyme CysN

The schematic shows the site of transposon-insertion for mutant 32:G7 (A) which was found to be BIT positive. The transposon mutation is located in a gene encoding a bifunctional enzyme CysN within a *cysND* gene operon (loci: RW109_05781-05782), involved in sulphur metabolism. Each gene is drawn to the scale given by the 1 kb bar and the putative products are shown by the colour coded key. A box plot of 32:G7 (B) indicates the change in light emission when exposed to BIT for 24 hours. Centre lines show the medians and crosses represent the sample means; box limits indicate the 25th and 75th percentiles as determined by R software; whiskers extend 1.5 times the interquartile range from the 25th and 75th percentiles (n= 4 sample points). The RLU fold change and corresponding p-value is shown next to the box plot.

CHAPTER 4 – Use of an industrial *P. aeruginosa* mini-Tn5-*luxCDABE* mutant library to investigate the genetic basis of preservative resistance

4.2.8 Key insertion sites of BIT negative mutants and mutants with minimal RLU change in response to BIT

Eight BIT negative mutants had a transposon insertion site identified within a gene encoding an acid activated urea channel (gene locus RW109_02452) which was located next to the *amiEBCR* gene operon (Figure 35). Acid activated urea channels are involved in the transport of urea across the cytoplasmic membrane (Sangari et al., 2010) and the *amiEBCR* gene operon produces the amidase enzyme in response to short chain aliphatic amides, such as acetamide (Kelly and Clarke, 1962, Wilson et al., 1995). A mean RLU fold decrease of -0.40 (p-value ≤ 0.05) was identified for the 8 mutants which were all from bioluminescence category 3. Mutant 37:H9 had the greatest RLU fold decrease of -0.59 (p-value ≤ 0.05) (Table 17); when this mutant was screened with BIT individually, a -0.80-fold decrease in RLU was observed (p-value ≤ 0.05) (Figure 35). This suggests the gene encoding for the acid activated urea channel, was down-regulated in response to exposure with the preservative BIT at 20% of the MIC.

Mutant 8:E12 had the third greatest RLU-fold decrease when exposed to BIT during master plate screening, with a -0.69-fold RLU reduction observed (p-value ≤ 0.05) (Table 17). The transposon insertion site gene within this mutant encoded for a tripartite ATP-independent periplasmic (TRAP) transporters DctQ component (gene locus RW109_06768) (Figure 36). TRAP transporters are substrate-binding, protein-dependent secondary transporters (Mulligan et al., 2011) and DctQ component is essential for transport of C4-dicarboxylates such as malate, succinate, and fumarate (Kelly and Thomas, 2001, Forward et al., 1997). When 8:E12 was screened individually with BIT, a -0.62 RLU fold decrease was observed (p-value ≤ 0.05) (Figure 36), suggesting the gene encoding the DctQ component of the TRAP operon was down-regulated in response to BIT exposure.

Mutant 1:E6 had no significant RLU fold change when exposed to BIT within a master plate screen (+0.0077 fold change, p-value ≥ 0.05) (Table 18) and the transposon insertion site gene encoded for an AcrB efflux pump subunit (gene locus: RW109_04032) of a multidrug efflux pump (Figure 37). The gene was identified within the multidrug efflux pump *acrEB* operon (gene loci RW109_04031-04032) which is part of the Resistance-Nodulation-Division (RND) superfamily of efflux pumps (Nikaido and Takatsuka, 2009). The AcrB subunit acts as an efflux transporter located within the inner membrane (Poole, 2004, Poole, 2005b). When this mutant was screened individually, no significant RLU fold change was observed (+0.0379 RLU fold, p-value ≥ 0.05) (Figure 37). This suggests that the gene insertion site within the AcrB efflux pump sub-unit had no change in expression during exposure to the preservative BIT at 20% of the MIC.

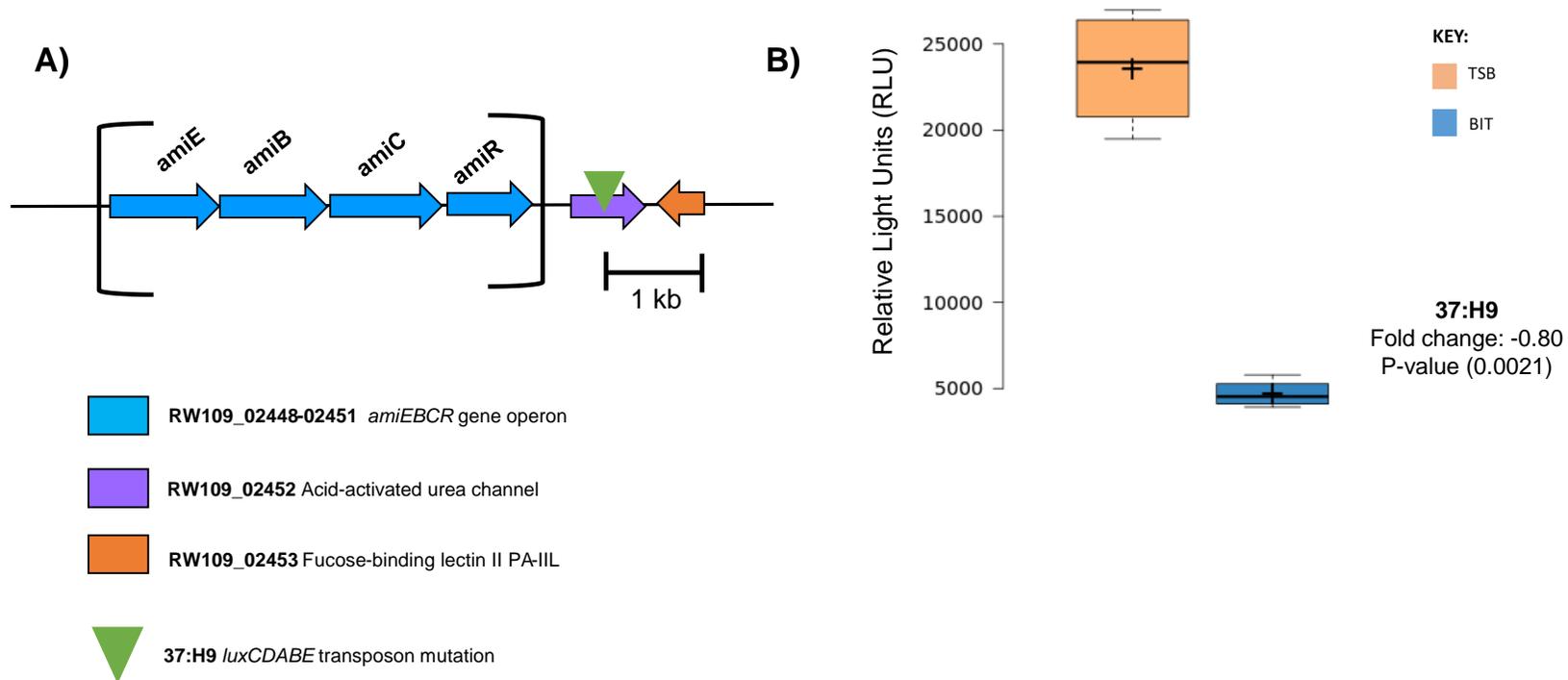


Figure 35. The *luxCDABE* transposon mutation of a gene encoding for an acid-activated urea channel

The schematic shows the site of transposon-insertion for mutant 37:H9 (A) which was found to be BIT negative. The transposon mutation gene was annotated as an acid active urea channel (loci: RW109_02452). This gene was located next to the *amiEBCR* gene operon (loci: RW109_02448-02451). Each gene is drawn to the scale given by the 1 kb bar and the putative products are shown by the colour coded key. A box plot of 37:H9 (B) indicates the change in light emission when exposed to BIT for 24 hours. Centre lines show the medians and crosses represent the sample means; box limits indicate the 25th and 75th percentiles as determined by R software; whiskers extend 1.5 times the interquartile range from the 25th and 75th percentiles (n= 4 sample points). The RLU fold change and corresponding p-value is shown next to the box plot.

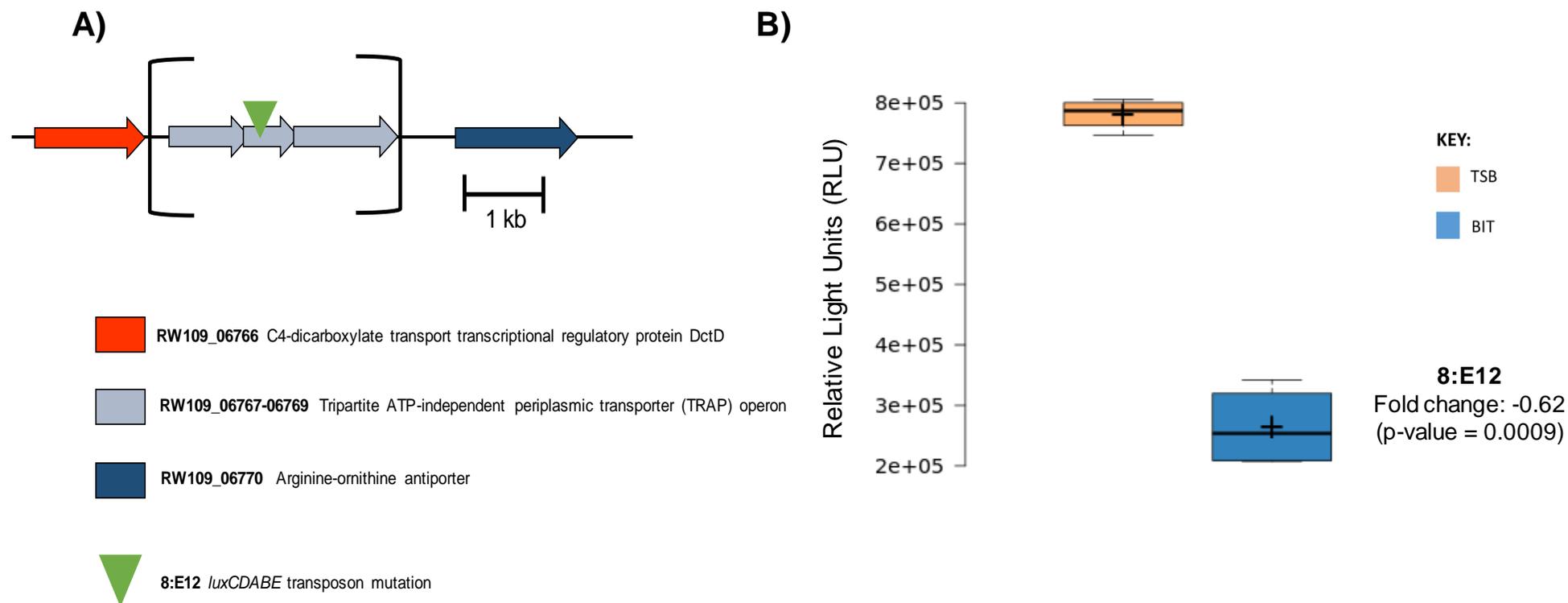


Figure 36. The *luxCDABE* transposon mutation of a gene encoding a tripartite ATP-independent periplasmic (TRAP) transporters DctQ component
 The schematic shows the site of transposon-insertion for mutant 8:E12 (A) which was found to be BIT negative. The transposon mutation gene is located within the tripartite ATP-independent periplasmic transporters DctQ component (loci: RW109_06768) of the TRAP operon (loci: RW109_06767-06769). Each gene is drawn to the scale given by the 1 kb bar and the putative products are shown by the colour coded key. Box plots of 8:E12 (B) indicates the change in light emission when exposed to BIT for 24 hours. Centre lines show the medians and crosses represent the sample means; box limits indicate the 25th and 75th percentiles as determined by R software; whiskers extend 1.5 times the interquartile range from the 25th and 75th percentiles (n= 4 sample points). The RLU fold change and corresponding p-value is shown next to the box plot.

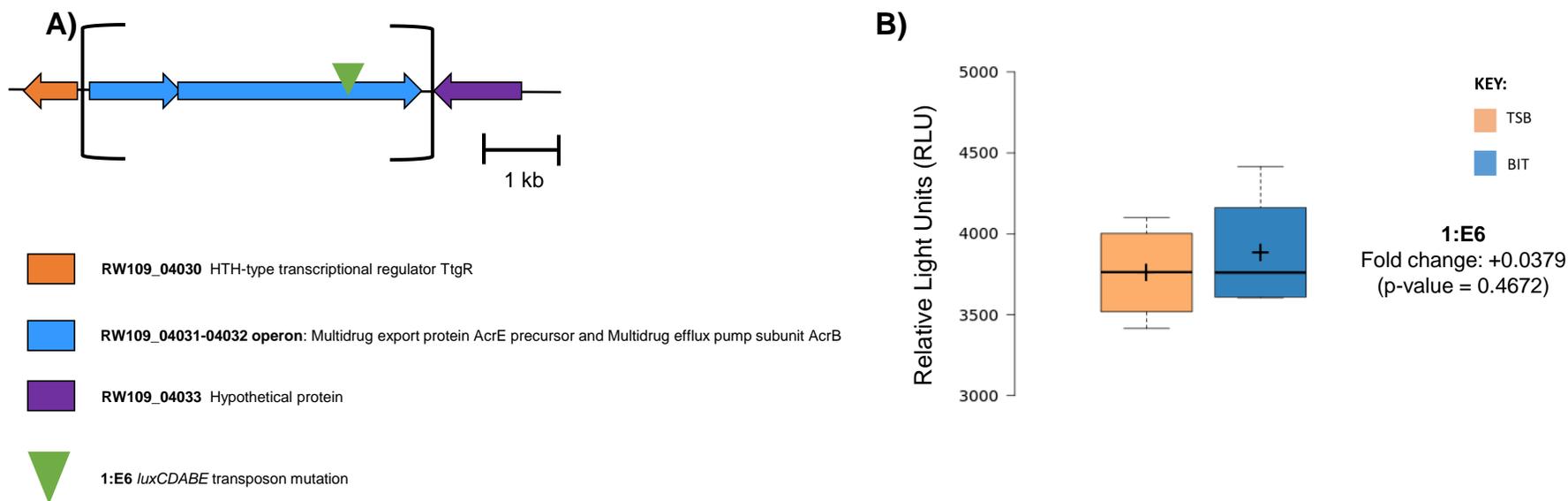


Figure 37. The *luxCDABE* transposon mutation of a gene encoding for the multidrug efflux pump subunit AcrB

The schematic shows the site of transposon-insertion for mutant 1:E6 (A) which was found to have minimal RLU change when exposed to BIT. The transposon mutation gene is located within the multidrug efflux pump subunit AcrB (loci: RW109_04032). Each gene is drawn to the scale given by the 1 kb bar and the putative products are shown by the colour coded key. A box plot of 1:E6 (B) indicates the change in light emission when exposed to BIT for 24 hours. Centre lines show the medians and crosses represent the sample means; box limits indicate the 25th and 75th percentiles as determined by R software; whiskers extend 1.5 times the interquartile range from the 25th and 75th percentiles (n= 4 sample points). The RLU fold change and corresponding p-value is shown next to the box plot.

4.3 Discussion

4.3.1 The industrial RW109 mini-Tn5-*luxCDABE* mutant library

This study demonstrated the successful construction of a *P. aeruginosa* mini-Tn5-*luxCDABE* transposon mutant library using the industrial RW109 strain. The mutants with distinct insertion sites were evenly dispersed throughout chromosome of the RW109 strain, with the exception of the region between ~2.2 and 2.9 Mbp (Figure 28). There were also 5 mutants with insertions on plasmid 1 and 3 mutant insertions on plasmid 2; this uniform distribution indicates the transposon was not likely to insert into only one area of the genome. Even though a low number of insertion sites were identified, their dissemination throughout the RW109 genome enabled the representation of genes from different areas. The mutants with distinct insertion sites were also characterised with a wide range of well characterised COG categories; this suggests the genes were involved in a variety of biological functions within the RW109 strain. A large number were also assigned to the poorly characterised COG categories. This may be due to these categories representing biggest percentage of genes within RW109 genome (Chapter 3, Section 3.2.3.1.1); increasing the likelihood of a transposon mutation occurring within poorly characterised genes.

The low insertion site density region identified between ~2.2-2.9 Mbp, was previously reported in the mini-Tn5-*lux* PAO1 transposon mutant bank (Lewenza et al., 2005). The region was predicted to be the terminus of replication in the *P. aeruginosa* genome (Song et al., 2003, Hallin and Ussery, 2004). This led to (Lewenza et al., 2005) suggesting that the low number of insertion sites identified was due to the reduced number of target-gene replicates within this region. The predicted essential genes of *P. aeruginosa* are also recognised as spread randomly throughout the genome (Jacobs et al., 2003, Lewenza et al., 2005, Stover et al., 2000, Liberati et al., 2006). Thus, indicating the region between 2.2-2.9 Mbp would not necessary result in a number of lethal mutations.

The constructed library was used to classify key mutants with an altered light emission response when screened with the isothiazolinone preservative BIT, at 20% of the calculated MIC. Mutants of interest identified from master plate screening were isolated and the genetic basis of the mutation was determined via comparison against the RW109 PacBio reference genome. The characterisation of insertion site genes within the BIT responding mutants identified key putative determinants, which may be involved in the resistance of RW109 to the isothiazolinone preservative. A number of mutants of interest where the transposon insertion site was successfully identified, also had reproducible RLU responses when screened individually. The mutants from the master plate screen where the transposon insertion sites was known (Table 16, Table 17 and Table 18), should also be screened individually with BIT to ensure reproducible RLU responses in comparison to master plate screening. The transposon insertion sites identified in key mutants within this study were only predictions with results requiring validation by quantitative real-time PCR (qRT-PCR) and via the generation of site-directed/non-polar mutants. The RW109 transposon mutant bank could also be used to identify mutants, which were susceptible or more tolerant to preservatives in comparison to the wild-type strain. This analyse could be used to further classify the genetic basis for industrial preservative resistance.

CHAPTER 4 – Use of an industrial *P. aeruginosa* mini-Tn5-*luxCDABE* mutant library to investigate the genetic basis of preservative resistance

More insertion sites also need to be classified in the mutants identified from mutant bank screening to further identify key genetic determinants. This could be carried out with further optimisation of the inverse PCR technique used or by the whole genome sequencing of selected mutants. Transposon Directed Insertion Sequencing (TraDIS) could also be used for the generation of a greater density of RW109 mutants with characterised insertion sites. TraDIS involves the construction of a large mutant bank followed by the fragmentation of mutant DNA and selective enrichment for transposon-flanking sequences via PCR amplification (Langridge et al., 2009, Barquist et al., 2015). The technique enables the classification of the frequency of mutations within the population and can be applied to Tn5-based mutant libraries (Barquist et al., 2015). TraDIS has successfully been used to construct characterised mutation libraries with a wide-ranging genome coverage in numerous bacteria, including in *E. coli* (Dziva et al., 2013), *Salmonella Typhi* (Langridge et al., 2009), *Clostridium difficile* (Dembek et al., 2015), *Mycobacterium marinum* (Weerdenburg et al., 2015) and *Pseudomonas syringae* (Mesarich et al., 2017).

This study also recognised that the mini-Tn5-*luxCDABE* transposon was not an appropriate reporter for gene expression with all preservative groups. When master plates were screened with POE, mutants demonstrated an overall decrease in light emission, suggesting this type of preservative was impairing *luxCDABE* operon activity. Higher concentrations of BIT also had the effect of diminishing light emission. Therefore, to fully characterise genetic pathways involved in resistance to the preservative POE, as well as increased concentrations of BIT, there is a requirement to use alternative forms of transcriptomic analysis, such as RNA-Seq. This method would enable the investigation of the whole RW109 transcriptome in response to industry relevant conditions (Wang et al., 2009). Applying RNA-Seq would also remove the requirement for construction, characterisation and screening of a large mutant library, which is necessary for sufficient genome coverage with transposon mutagenesis. Consequently, RNA-Seq would be a more efficient method to determine the genetic basis of preservative resistance.

CHAPTER 4 – Use of an industrial *P. aeruginosa* mini-Tn5-*luxCDABE* mutant library to investigate the genetic basis of preservative resistance

4.3.2 The key genetic determinants associated with BIT positive mutants

Screening the mini-Tn5-*luxCDABE* RW109 mutant bank with BIT, identified 204 mutants with a RLU-fold increase and transposon insertion sites were successfully identified in 85 mutants. Out of these mutants, 40 were found to have insertion sites within the bacterial extracellular solute binding proteins family 3 (gene locus: RW109_01055). The high number of mutants with this insertion site suggests it was a possible hotspot for the Tn5-*luxCDABE* transposon in RW109. However, the Tn5 inserts at a 9 bp degenerate and uneven sequence (Kang et al., 2004) and past studies using this transposon have not reported an incidence of hotspots. Lewenza et al., (2005) found the 9 bp Tn5 insertion site sequence to be randomly distributed throughout the PAO1 genome; indicating regions of hotspots were not likely in *P. aeruginosa*. The large number of insertions within the RW109_01055 gene may have occurred due to the method used for picking individual mutants during library construction.

All 40 mutants with an insertion in the gene encoding for bacterial extracellular solute binding proteins family 3, demonstrated a RLU-fold increase in response to BIT and were from the same bioluminescence category (Table 16). The gene product is involved in the active transport of solutes across the cytoplasmic membrane (Tam and Saier, 1993, Sauer et al., 2002) and the role of transport systems for removing antimicrobials from bacterial cells is a well characterised mechanism of resistance. Class 3 solute binding proteins have also been linked with roles in adhesion (Sauer et al., 2002) and the ability of *P. aeruginosa* to adhere has been associated with biofilm-mediated industrial antimicrobial resistance (Brown et al., 1995, Cochran et al., 2000). The mutant RLU result suggest the gene encoding bacterial extracellular solute binding proteins family 3, was up-regulated when exposed to BIT. Increased expression could be to actively remove the preservative from the cell or to initiate increased adhesion for a biofilm-mediated resistance response. Two additional BIT positive mutants had transposon insertion sites in RW109_01054, which encodes an HflK protein. This gene was located within the same operon as the bacterial extracellular solute binding proteins family 3 (Figure 29). The increased RLU observation, suggests the insertion site operon genes were important to RW109 in response to the preservative BIT. HflK regulates the specificity of FtsH membrane protease (Ito and Akiyama, 2005) and inactivation of this protein results in increased sensitivity to the aminoglycoside antibiotic tobramycin in *P. aeruginosa* (Hinz et al., 2011). Increased expression of the gene encoding HflK has also been observed in response to aminoglycoside treatment (Kohanski et al., 2008). The suggested up-regulation of the gene encoding HflK in this study could be associated with decreasing RW109 susceptibility to BIT.

The BIT positive mutant 14:C4 was phenotypically different from the wild type RW109 strain as it overproduced a red/brown pigment, possibly pyorubin, when grown on TSA only (Figure 30). The transposon insertion site was identified within the RW109_04041 gene that encodes the transcriptional regulator KdgR and was located directly next to the *gnyRDBHAL* gene operon (loci RW109_04035-04040) (Figure 31). This operon is involved in acyclic isoprenoid degradation (Diaz-Perez et al., 2004, Förster-Fromme et al., 2006) and isoprenoids are implicated in bacterial pigment production (Cañizares-Villanueva et al., 1997, Saha et al., 2008). The overproduction of a pigment in this mutant suggests the insertion site gene

CHAPTER 4 – Use of an industrial *P. aeruginosa* mini-Tn5-luxCDABE mutant library to investigate the genetic basis of preservative resistance

encoding KdgR, may have a role in the regulation of the *gnyRDBHAL* gene operon. The KdgR is a member of the IclR family of bacterial transcriptional regulators and functions as a repressor of gene expression (Molina-Henares et al., 2006). KdgR regulates the RsmA-rsmB system, which is involved in quorum sensing and in the synthesis of numerous secondary metabolites and virulence factors such as protease, elastase, staphylolytic enzymes, hydrogen cyanide and pyocyanin (Hyytiäinen et al., 2001, Pessi et al., 2001, Burrowes et al., 2005). The increased RLU observed when mutant 14:C4 was exposed to BIT, suggests the transcriptional regulator KdgR was important when RW109 in response to the industrial preservative.

A transposon insertion site was also identified in the periplasmic nitrate reductase *napA* gene of the *napEFDABC* operon (loci RW109_04993-04998) in mutant 8:B11 (Figure 32). The *napEFDABC* operon is associated with biofilm-mediated antimicrobial resistance and increased expression of the operon allows *P. aeruginosa* to proliferate under aerobic and anaerobic conditions (Schreiber et al., 2007, Van Alst et al., 2009, Van Alst et al., 2007). The result from mutant screening with BIT at 20% of the MIC, suggests the up-regulation of the *napA* gene within the *napEFDABC* operon and indicates its importance in RW109's response to the preservative. The BIT positive 37:C5 had an insertion site gene RW109_01716 that encoded for a ferrous-iron efflux pump, FieF (Figure 33). Ferrous iron transport systems are associated with increased virulence and biofilm development in *P. aeruginosa* (Marshall et al., 2009, Seyedmohammad et al., 2014). The screening results of mutant 37:C5 suggest the ferrous-iron efflux pump, FieF is up-regulated when the RW109 strain was exposed to the preservative BIT. This response could be to remove BIT from within the cell or to initiate biofilm formation.

Another BIT positive mutant of interest was 32:G7 where the insertion site gene encoded for the bi-functional enzyme CysN within a *cysND* gene operon (loci: RW109_05781-05782) (Figure 34). The gene operon has previously been observed as up-regulated in *P. aeruginosa* during conditions of oxidative stress (Goldová et al., 2011) and has a role in the metabolism of sulphur in *P. aeruginosa* (Hummerjohann et al., 1998). The preservative BIT causes an oxidative response (Collier et al., 1990) as it's able to act on thiol containing cytoplasmic and membrane bound enzymes (Denyer, 1995). As the mutant had increased RLU when exposed to BIT, it suggests *cysN* was up-regulated to combat adverse oxidative stress effects. It may also have an increased expression in order to expand sulphur availability, which is required for repairing disrupted thiol groups (Ezraty et al., 2017).

4.3.3 The key genetic determinants associated with BIT negative mutants

Screening the mini-Tn5-luxCDABE mutant bank with BIT identified 97 mutants, which had an RLU-fold decrease, and transposon insertion sites were successfully identified in 36 of these mutants. Eight key BIT negative mutants had transposon insertion sites, which encoded an acid, activated urea channel (gene locus RW109_02452). Acid activated urea channels are involved in the transport of urea across the cytoplasmic membrane (Sangari et al., 2010) (Figure 35). Gram-negative membrane channels and porins can be used to influx various chemicals including antimicrobials, and adaptations to reduce entry via this process decreases antimicrobial susceptibility (Pagès et al., 2008, Low et al., 2001). Potentially, acid activated

CHAPTER 4 – Use of an industrial *P. aeruginosa* mini-Tn5-luxCDABE mutant library to investigate the genetic basis of preservative resistance

urea channels also allow BIT to enter the cell. RW109 could be down-regulating the gene encoding the activated urea channel to prevent the resulting adverse effects of preservative build-up. Another BIT negative mutant 8:E12, with a reproducible RLU response, also had a transposon insertion site involved in membrane transport (Figure 36). The gene encoded for the DctQ component of a TRAP transporter, which act as periplasmic importers of a diverse range of solutes into the cytoplasm (Brautigam et al., 2012, Mulligan et al., 2011, Kelly and Thomas, 2001). Again, this could be used as a possible method of BIT entry and the gene encoding the import transporter protein may be down-regulated by RW109 as a resistance mechanism.

From master plate screening, 103 mutants were isolated, which demonstrated minimal change when exposed to BIT and transposon insertion sites were identified in 36. The genes represented by this response category suggest there was no altered expression when RW109 was exposed to the preservative BIT. An example of a minimal RLU change mutant which had a reproducible RLU response when screened individually was 1:E6. The transposon insertion site was identified in a gene encoding for an AcrB efflux pump subunit, which formed an operon with an AcrE export protein (Figure 37). These two proteins have reported similarities with those involved in the *E.coli* AcrAB–TolC efflux pump (Smith and Blair, 2013) which is able to export multiple antibiotics including tetracycline, chloramphenicol and fluoroquinolones (Edward et al., 2003). The results from the mutant screening suggest this efflux system did not have an altered expression in response to BIT, indicating it may not be important the resistance to this preservative.

4.4 Conclusions

The main conclusions from this chapter were as follows:

- 1) The industrially isolated strain RW109 strain was successfully used to construct a mini-Tn5-*luxCDABE* mutant library. The distinct transposon insertion sites were evenly distributed throughout the whole genome and identified on the two plasmids.
- 2) The mutant bank was screened with the industry relevant preservative BIT at 20% of the MIC and mutants with either an increased, decreased or an unchanged RLU response, were identified and isolated for further analysis. The RLU light emission response observed within a master plate was also reproducible when selected mutants were screened individually.
- 3) The mutant library was not suitable for screening with the preservative POE and increased concentrations of BIT, as an overall fold decrease in light emission was observed. This suggests the activity of the *luxCDABE* operon was being diminished.
- 4) The putative gene function in mutants with a reproducible RLU response were determined and linked to gene pathways of the RW109 PacBio reference genome. The mutants of interest had insertions in genes encoding solute binding proteins, transcriptional regulators, efflux pumps and proteins associated with biofilm formation and oxidative stress responses. These findings provide an insight into the genetic basis of BIT preservative resistance in the industrial RW109 strain.
- 5) Out of the 7,303 predicted CDS of the RW109 genome, 53 distinct transposon insertion sites were identified using the mini-Tn5-*luxCDABE* mutant library preservative screening approach. This indicates the requirement for an alternative and more enhanced method to identify the genetic basis of preservative resistance in the RW109 strain.

5. RNA-Seq gene expression analysis of RW109 in response to preservatives and product formulations

5.1 Introduction

The transcriptome profiling technique RNA-Seq, utilises high-throughput next generation sequencing to map and quantify gene expression across a genome (Wang et al., 2009). This technique can be used to investigate the frequency with which genes are activated or repressed when a bacterium is subjected to different conditions. RNA-Seq is the successor of microarray technology and in comparison, provides a more holistic method of analysing gene expression without the reliance on hybridisation techniques (Wang et al., 2009, Mäder et al., 2011). RNA-Seq demonstrates an extensive dynamic range of expression levels, whereas hybridisation techniques are not as efficient at detecting and quantifying expression of genes at low or very high levels. Hybridisation techniques are also limited by the occurrence of cross hybridisation signal interference and have a limited ability to generate in-depth, high quality detection of transcripts (Wang et al., 2009, Mantione et al., 2014, Croucher and Thomson, 2010). RNA-Seq is also not dependent on the range of target gene probes present on a hybridisation array, and has been particularly useful for the detection of intergenic non-coding regulatory RNAs (Miyakoshi et al., 2015, Saliba et al., 2017).

Overall RNA-Seq enables a wide-ranging analysis of transcriptomic changes and produces high-resolution results, thus enhancing the reliability and accuracy when quantifying gene expression in response to a specified condition. RNA-Seq works via the extraction of RNA from bacterial cells which were exposed to either control conditions or test conditions and the enrichment of mRNA via the removal of rRNA (Figure 38) (Wang et al., 2009, Levin et al., 2010). The mRNA is converted into cDNA libraries, which are sequenced and resulting reads aligned to a reference genome. The number of reads which map to each gene feature of the genome are quantified and differential gene expression determined by comparing test conditions with a control (Figure 38) (Wang et al., 2009, Levin et al., 2010).

This transcriptomic sequencing technology is advantageous in understanding how bacteria interact with their environment. RNA-Seq has been successfully implemented to determine genetic pathways involved in multiple bacterial functions such as antimicrobial resistance, biofilm development and pathogenicity (Khaledi et al., 2016, Jones et al., 2014, Maura et al., 2016). The use of RN-Seq to investigate antimicrobial gene expression responses, identified dominant genetic resistance traits when 135 clinical *P. aeruginosa* isolates were exposed to different antibiotics (Khaledi et al., 2016); an example was the up-regulation of a sensor kinase-encoding gene *cbrA* in response to polymyxins, tobramycin, and ciprofloxacin (Khaledi et al., 2016). These types of findings can be used to develop targeted treatment approaches. RNA-Seq will enable the investigation of global gene expression of the industrial *P. aeruginosa* RW109 when exposed to industry-relevant conditions, thus enabling the detailed analysis of genetic pathways, which allow the industrial strain to survive the presence of preservatives and HPC

products. A better understanding of common genetic resistant traits used in response to industrial antimicrobials, could also enable the development of targeted HPC preservation systems.

5.1.1 Aims and objectives

The overall aim of this chapter was to investigate the global gene expression response of *P. aeruginosa* RW109 when it was exposed to industry-relevant conditions. This was achieved by carrying out the following objectives:

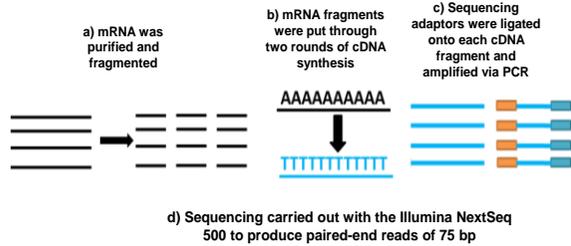
- 1) Identify appropriate industrial exposure and growth conditions, which enable the extraction of good quality RNA for subsequent transcriptomic investigations of the RW109 strain.
- 2) Carry out RNA-Seq analysis following exposure of RW109 to the individual preservatives BIT and POE, a combination of these two preservatives and a laundry detergent with and without the addition of the preservative BIT.
- 3) Analyse differentially regulated genes with significant increases and decreases in expression when exposed to the industry relevant conditions.

1) Industrial *P. aeruginosa* strain was grown in the presence of test and control conditions for 24-hours:



Four biological replicates obtained for each exposure condition
 tRNA extraction and mRNA enrichment
 Quantification and QC assessment with Qubit fluorometer and Agilent Bioanalyzer

2) mRNA converted into cDNA and prepared into sequencing libraries



Bioinformatic analysis

3) Differential gene expression

a) Quality control and adaptor trimming followed by aligning the sequence reads to the reference genome.



b) Quantify the number of reads mapped to each reference genome CDS



3) Determining log₂-fold changes and adjusted p-values when test conditions were compared to the control

Figure 38. Overview of stages involved RNA-Seq transcriptomic analysis

The schematic gives an overview of the RNA-Seq protocol used in this study. The process can be divided into three stages which involve 1) exposing the bacterial strain to relevant conditions, extracting the toRNA, mRNA enrichment and quality assurance checking. 2) Converting the mRNA to cDNA and preparing libraries to carry out sequencing. 3) Bioinformatics analysis to determine differentially expressed genes. This step includes aligning the sequence reads to a reference genome, counting the number of reads which map to the CDS features and determining significant log₂-fold changes when comparing test conditions with a control.

5.2 Results

5.2.1 Growth of *P. aeruginosa* RW109 in the presence of preservatives BIT and POE at various concentrations

To analyse the growth dynamics of RW109 in the presence of increasing concentrations of BIT and POE, the strain was grown for 48 hours in TSB supplemented with the preservatives at concentrations between 0 – 115% of the MIC. The MIC values were previously calculated as 0.00938% active for BIT and 0.5% active for POE (Chapter 4, Section 4.2.2). When deciding on RNA-Seq exposure conditions, BIT at 20% of the MIC was of interest to compare with the screening condition for identifying mini-Tn5-*luxCDABE* responding mutants (Chapter 4, Section 4.2.2). An exposure to POE at 20% of the MIC was relevant to relate gene expression results with those obtained with BIT at 20% of the MIC. Test conditions with increased concentrations for both BIT and POE were also sought, and half the MIC (50%) was seen as a possible concentration of interest.

The growth dynamics of RW109 in the presence of BIT (A) and POE (B) at increasing concentrations altered over 48-hours (Figure 39). An extension of the lag phase was observed when RW109 was exposed to increasing concentrations of BIT when compared to growth in TSB-only (Figure 39A). RW109 was able to grow in concentrations of up to 80% of the MIC with BIT, and reached an OD close to the maximum observed in TSB-only, at approximately 46-hours. Growth was also observed with BIT up to 90% of the calculated MIC, with the OD starting to increase at approximately 38-40 hours at 85% and 90% of the MIC. When exposed to increasing concentrations of POE for 48-hours, the effect was an overall reduction of growth; RW109 was not able to achieve an OD close to the maximum reached in TSB-only (Figure 39B). Growth was observed up to 90% of the MIC, with the OD starting to increase at around 30-35 hours for 85% and 90% of the MIC (Figure 39B).

It was decided to repeat the growth of RW109 in the presence of BIT and POE both at 20% and 50% of the MIC, and carry out OD measurements for 24-hours (Figure 39C). The mean Log₁₀ OD was 0.421 (450-480 nm) at 24-hours when grown in just TSB; this value was used as a comparison to calculate the percentage difference in mean Log₁₀ OD for the other exposure conditions (Figure 39D). BIT at 20% and 50% of the MIC had a percentage reduction of -1.21% and -11.41% respectively at 24-hours. POE at 20% and 50% of the MIC had a greater effect on growth and resulted in a percentage decrease of -27.16% and -67.96% respectively. RW109 was at different growth stages after 24-hours exposure to these preservative conditions. However, rather than concentrate on harvesting cells at a particular growth stage, the initial RNA-Seq experiment aimed to investigate the gene expression response after 24-hours of exposure to draw greater correlation to the transposon mutagenesis and *lux*-reported gene expression (Chapter 4). The next step in deciding on appropriate test conditions was to assess the quality of extracted RNA from RW109 after exposure to these conditions.

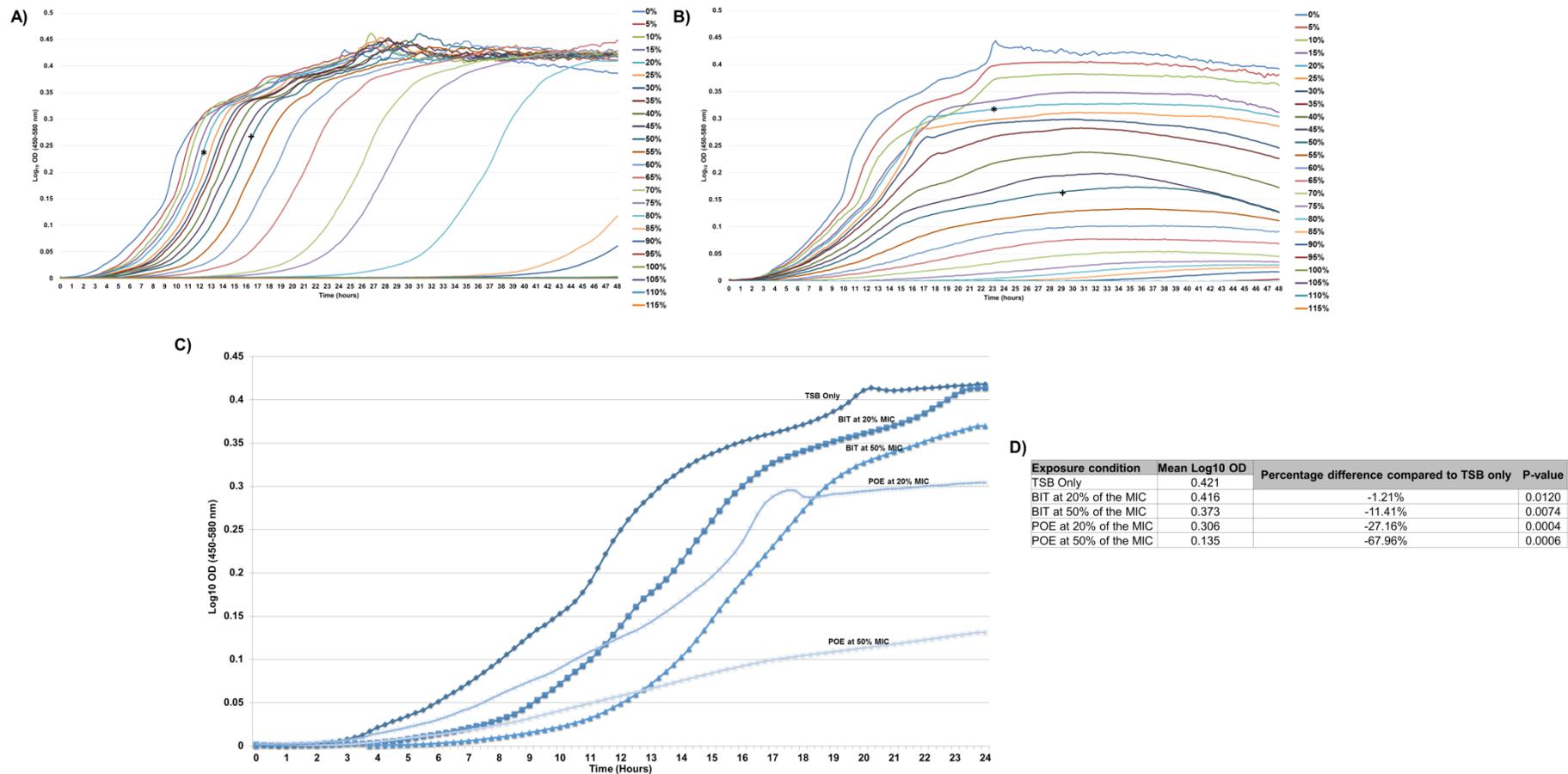


Figure 39. The effect of the individual preservatives BIT and POE at different concentrations on the growth of RW109

Line graphs A and B demonstrate the growth dynamics of the RW109 strain when exposed to the individual preservatives BIT (A) and POE (B) at the concentration range of 0 – 115% of the calculated MICs for 48 hours at 30°C. On the line graphs A and B the * highlights growth at 20% of the MIC and the + shows growth at 50% of the MIC. Growth analysis of RW109 at 20% and 50% of the MIC for both BIT and POE was repeated and shown in C). Panel D demonstrates the Log₁₀ OD values at 24 hours and the percentage difference in OD when compared to TSB-only for each of the exposure conditions shown in C).

5.2.2 Initial total RNA extractions and mRNA enrichment

Total RNA (toRNA) extractions from RW109 were carried out after growth for 24-hours at 30°C in TSB-only and POE at 50% of the MIC. These conditions were chosen for assessment of RNA quality as they represented the highest and the lowest OD values after 24 hours exposure (Figure 39D). Both toRNA samples passed quality analysis with RNA integrity Numbers (RINs) of ≥ 8 and 23S:16S rRNAs ratios of ≥ 1.5 (Figure 40). The concentrations were checked with the Qubit and were both above the required value of 100 ng/ μ l (Figure 40). This suggests good quality RNA can be extracted when the OD at 24-hours is reduced by more than 50% when compared to RW109 grown in TSB-only. Messenger RNA (mRNA) enrichment was then carried out and quality assessment showed a noticeable reduction in the 16S and 23S rRNA bands for both samples when compared to the toRNA (Figure 40). The concentration of the mRNA-enriched samples was 22 ng/ μ l for the TSB-only treated sample and 28 ng/ μ l when exposed to POE at 50% MIC. These concentrations were within the required range of 10 – 400 ng per 5 μ l, for input into the Illumina cDNA library preparation kit.

When choosing further RNA-Seq test conditions, it was decided the mean \log_{10} OD (450-580 nm) should not result in a percentage reduction of more than -67.96% when compared to RW109 grown in TSB for 24 hours. This parameter was used in an attempt to ensure the extraction of good quality RNA that had previously been achieved when this percentage reduction in OD was observed. For the RNA-Seq experiment, BIT and POE both at 20% and 50% of the MIC were chosen as test conditions to represent individual preservatives.

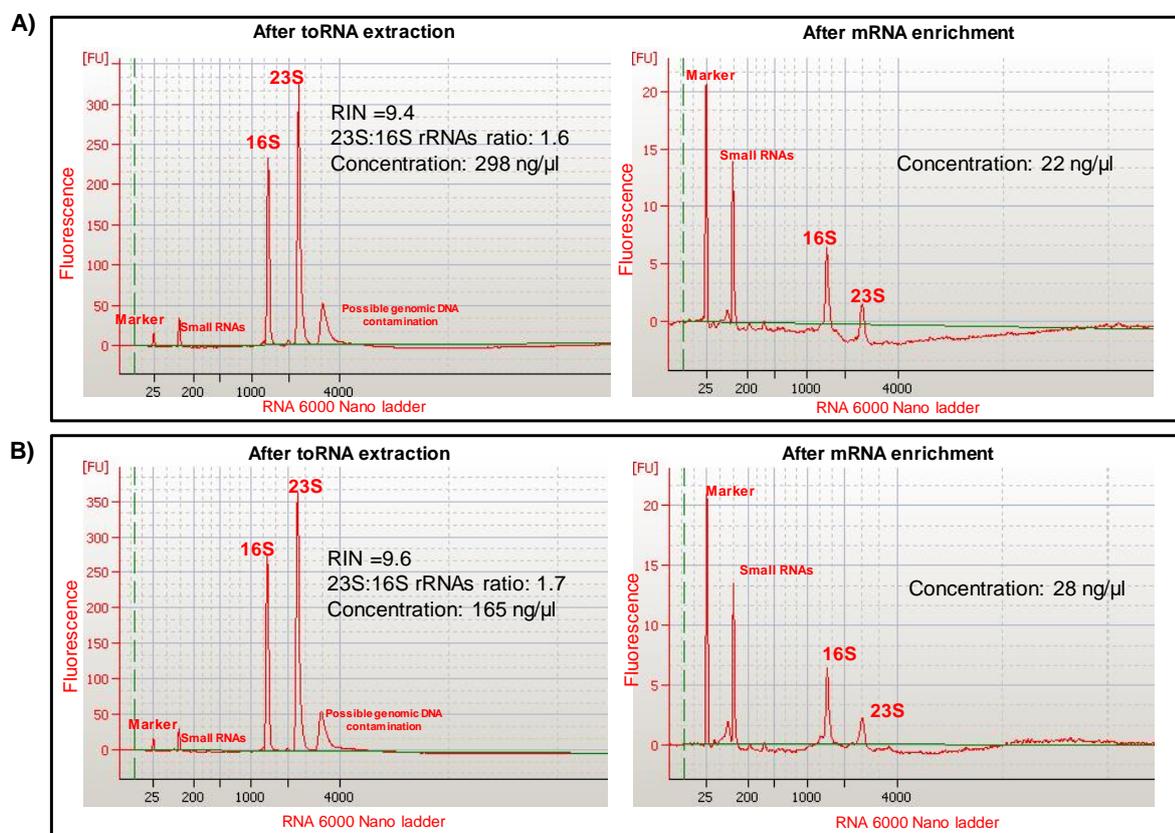


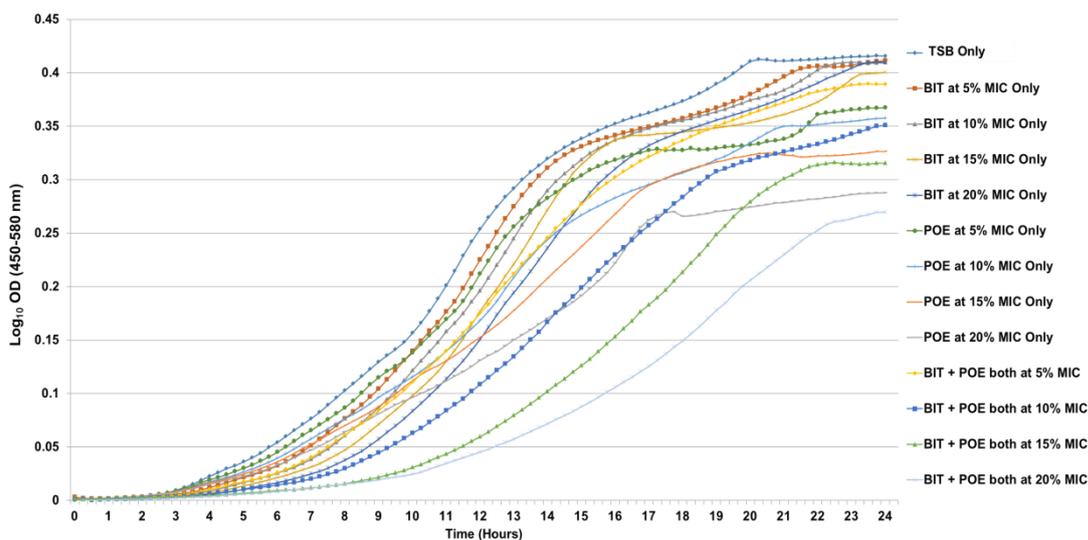
Figure 40. Bioanalyzer electropherograms of extracted total RNA and enriched mRNA

RNA extracted from RW109 grown in TSB-only (A) and in the presence of POE at 50% of the MIC (B) were analysed with the Bioanalyzer to produce electropherograms. Total RNA (toRNA) was quality assessed using the RIN value and 23S:16S ribosomal RNA (rRNA) ratio and concentrations analysed with the Qubit. Messenger RNA (mRNA) which had been enriched from the toRNA was also checked with the Bioanalyzer to determine if the 16S and 23S rRNA peaks had been reduced and concentration was again analysed with the Qubit. Both RNA samples shown in A and B passed quality assessment checks and were above required concentration values.

5.2.3 Growth of *P. aeruginosa* RW109 when exposed to BIT and POE in combination

RW109 was grown in the presence of BIT and POE both in combination at 5%, 10%, 15% and 20% of the MIC. This analysis was used for growth dynamics and to calculate the percentage reduction in OD when compared to TSB-only. Exposure to the individual preservatives BIT and POE at 20% of the MIC were chosen as test conditions, and consequently it would be appropriate to expose the industrial strain to a combination of preservatives both at 20% of the MIC. A combination of both preservatives at 50% of the MIC was not desirable, due to the large growth reduction observed with POE at this concentration (Figure 39). Lower combination concentrations were also assessed in the event that percentage reduction in OD was more than 67.96%. Increasing the concentration of the preservatives in combination had the effect of extending the lag phase, when compared to growth in TSB-only (Figure 41A). A percentage reduction of -35.15% when compared to growth in TSB-only was observed following exposure of RW109 to BIT and POE in combination both at 20% for 24 hours (Figure 41B). Therefore, it was decided the preservatives together at this concentration would be suitable for good quality RNA extractions.

A)



B)

Exposure condition	Mean Log ₁₀ OD at 24 hours	Percentage difference compared to TSB only	P-value
TSB Only	0.416		
BIT + POE both at 5% of the MIC	0.389	-6.42%	0.0029
BIT + POE both at 10% of the MIC	0.351	-15.59%	0.0069
BIT + POE both at 15% of the MIC	0.315	-24.14%	0.0054
BIT + POE both at 20% of the MIC	0.270	-35.15%	0.0034

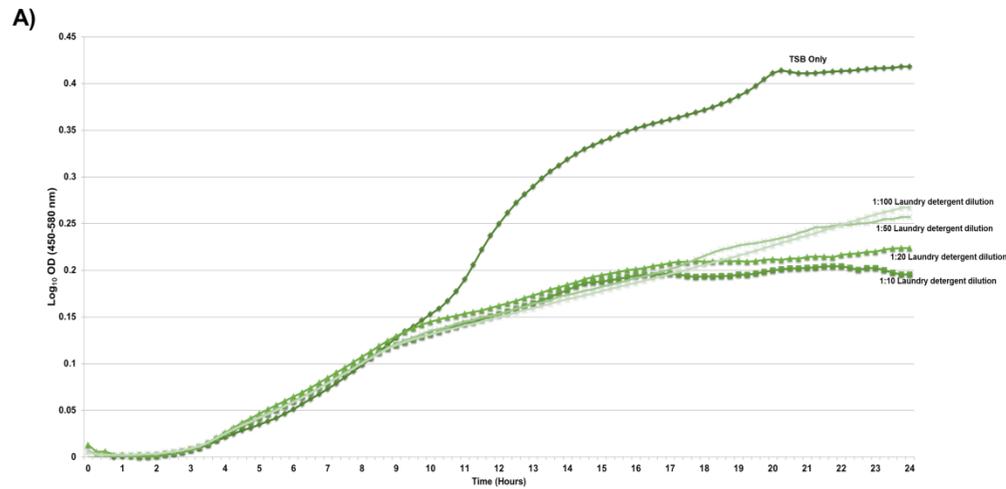
Figure 41. The effect of BIT and POE in combination on the growth of RW109

Line graphs demonstrate the growth of the RW109 strain in the presence of BIT and POE in combination both at 5, 10, 15 and 20% of their calculated MIC values for 24 hours at 30°C (A). Growth was also observed for the preservatives individually at these concentrations within the same graph (A). The Log₁₀ OD values at 24 hours for each exposure condition and the percentage difference in OD when compared to TSB-only are shown in B).

5.2.4 Growth of *P. aeruginosa* RW109 when exposed to unpreserved laundry detergent and laundry detergent with the presence of the preservative BIT

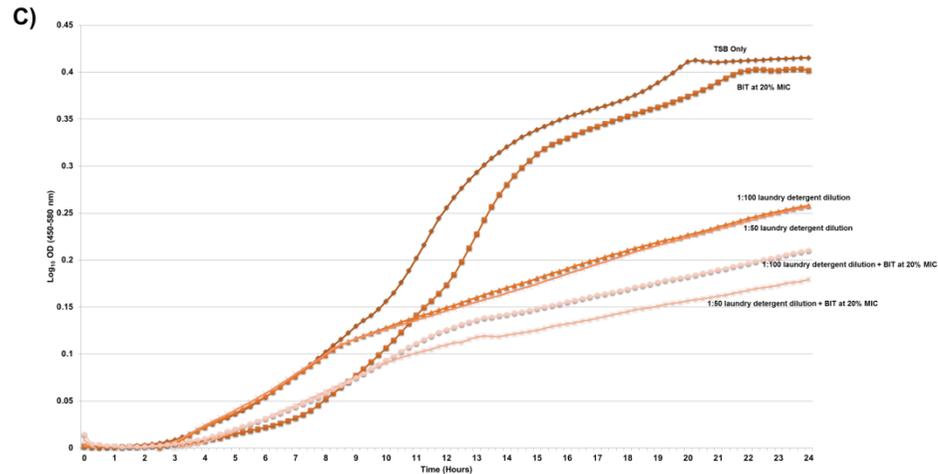
The growth dynamics of RW109 when exposed to unpreserved laundry detergent and laundry detergent in combination with the preservative BIT were assessed for 24-hours. The different dilutions of laundry detergent analysed included 1:10, 1:20, 1:50 and 1:100 and when compared to growth of RW109 in TSB, similar lag and exponential phases were observed up to the 8-10 hour point (Figure 42A). A reduction in growth was seen after 8-10 hours with the laundry detergent dilutions. The 1:10 and 1:20 dilutions resulted in an early stationary phase, whereas growth increased slowly with dilutions 1:50 and 1:100. Unpreserved laundry detergents at these concentrations for 24-hours, did not result in a mean OD percentage reduction of above 67.96% (Figure 42B); dilutions 1:10 and 1:20 were however difficult to prepare due to the thickness of the laundry detergent and presence of bubbles during mixing. A small amount of foaming was observed during mixing of 1:50 dilutions, but no foam was noted with the 1:100 dilution. It was decided that dilutions 1:50 and 1:100 would be used to assess the growth dynamics of laundry detergent in combination with BIT at 20% of the MIC.

When RW109 was grown with the laundry detergent dilutions in combination with BIT at 20% of the MIC, the effect was an extension in the lag phase and overall slower growth (Figure 42C). A reduction in the mean OD was observed in comparison to growth in TSB-only (Figure 42D), however the percentage differences observed did not fall below the chosen -67.96% cut off. It was decided to proceed with the laundry detergent dilution of 1:100 due to small amount of bubbling observed when mixing with dilutions 1:50.



B)

Exposure condition	Mean Log ₁₀ OD at 24 hours	Percentage difference compared to TSB only	P-value
TSB Only	0.418		
1:10 laundry detergent dilution	0.196	-53.15%	0.0007
1:20 laundry detergent dilution	0.224	-46.47%	0.0024
1:50 laundry detergent dilution	0.256	-38.71%	0.0056
1:100 laundry detergent dilution	0.267	-36.12%	0.0011



D)

Exposure condition	Mean Log ₁₀ OD at 24 hours	Percentage difference compared to TSB only	P-value
TSB Only	0.415		
1:50 laundry detergent dilution + BIT at 20% MIC	0.180	-56.55%	0.0004
1:100 laundry detergent dilution + BIT at 20% MIC	0.210	-49.40%	0.0006

Figure 42. The effect of unpreserved laundry detergent and laundry detergent with the preservative BIT on the growth of RW109

Line graphs demonstrate the growth of RW109 for 24 hours at 30°C in the presence of an unpreserved laundry detergent at different dilutions (A) and laundry detergent with the preservative BIT (C). The Log₁₀ OD values at 24 hours for each exposure condition and the percentage difference in OD when compared to TSB-only are shown in B) for laundry detergent only and in D) for laundry detergent in combination with the preservative BIT.

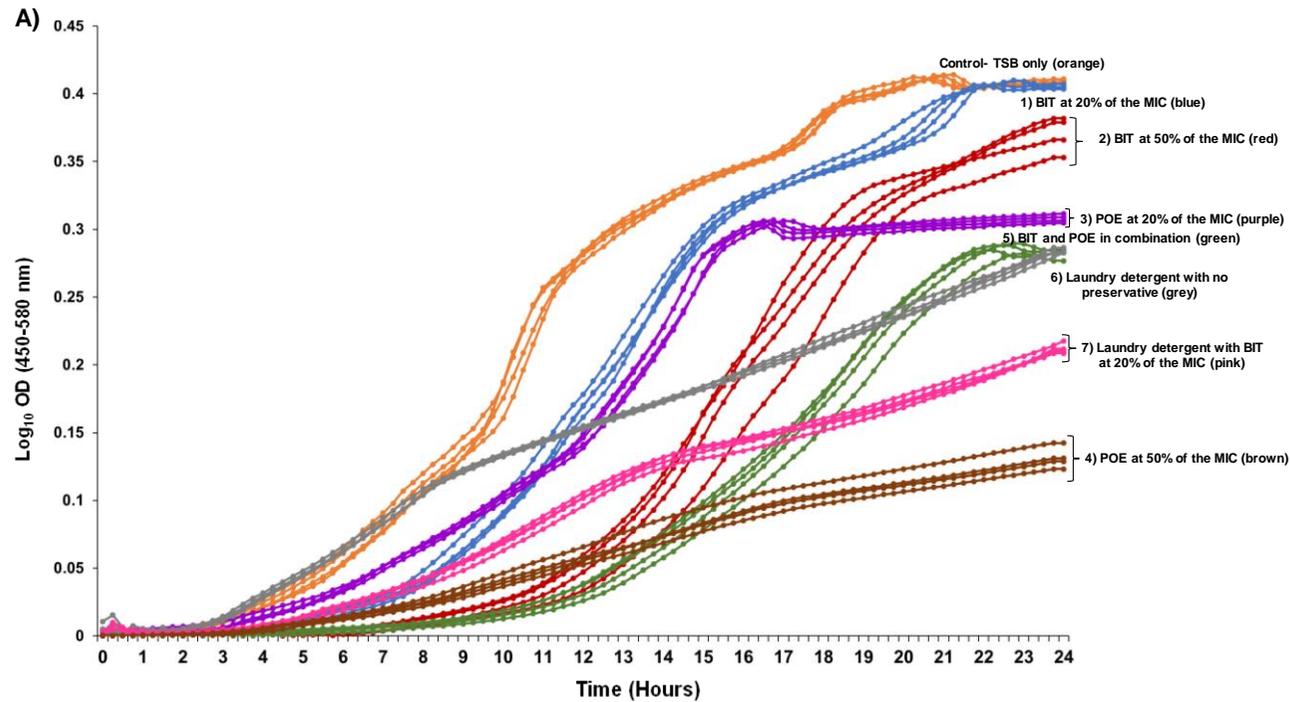
5.2.5 RNA-Seq exposure conditions

From the growth analysis investigations, the following exposure conditions were selected to investigate gene expression changes in RW109 using RNA-Seq:

Control condition: TSB-only

- 1) BIT at 20% of the MIC (0.01875% active)
- 2) BIT at 50% of the MIC (0.00469%)
- 3) POE at 20% of the MIC (0.1% active)
- 4) POE at 50% of the MIC (0.25% active)
- 5) POE and BIT in combination at 20% of the MIC
- 6) Unpreserved laundry detergent at 1:100 dilution only
- 7) Laundry at 1:100 dilution in combination with BIT at 20% of the MIC

RW109 was grown in the presence of the control condition and test exposure conditions for 24 hours at 30°C and tRNA was extracted. Experiments were repeated with different starting overnight cultures and exposure condition preparations, to obtain four biological replicates for the control and test exposure conditions. Figure 43 demonstrates the growth of RW109 in response to the control and test conditions, with growth curves for the 4 biological replicates of each condition shown. The Log₁₀ OD values were consistent during the 4 replications of each exposure condition, as demonstrated by the tight groupings of the replicate data points (Figure 43). BIT at 20% of the MIC (1) resulted in the smallest difference in mean Log₁₀ OD when compared to the control with a percentage reduction of -2.73% observed. A minimum difference in mean viable cell counts (CFU/ml) when compared to the control was also observed with test condition 1 with a percentage reduction of -21.50% observed (Figure 43 B and C). POE at 50% of the MIC had the greatest effect of the RW109 growth when compared to the control condition. A -68.46% reduction in mean Log₁₀ OD and a -100% reduction in mean viable cell counts (CFU/ml) were observed in comparison to the control condition (Figure 43 B and C).



B)

Exposure condition	Mean Log ₁₀ OD	Percentage difference when compared to condition 1	P-value
Control- TSB Only	0.417		
1) BIT at 20% MIC	0.405	-2.73%	0.0070
2) BIT at 50% MIC	0.370	-11.17%	0.0088
3) POE at 20% MIC	0.308	-26.06%	6.58E-07
4) POE at 50% MIC	0.131	-68.46%	3.78E-06
5) BIT and POE in combination	0.282	-32.23%	5.64E-06
6) Laundry detergent only	0.284	-31.86%	5.30E-07
7) Laundry detergent and BIT	0.212	-49.08%	1.22E-06

C)

Exposure condition	Mean viable cell count (CFU/mL)	Percentage difference when compared to condition 1	P-value
Control- TSB only	1.19E+10		
1) BIT at 20% MIC	9.32E+09	-21.50%	0.0004
2) BIT at 50% MIC	4.94E+09	-58.39%	0.0049
3) POE at 20% MIC	3.84E+09	-67.69%	0.00686
4) POE at 50% MIC	7.93E+08	-100.00%	5.96E-05
5) BIT and POE in combination	2.60E+09	-77.78%	0.01846
6) Laundry detergent only	1.81E+09	-84.77%	0.00026
7) Laundry detergent and BIT	9.59E+08	-91.93%	6.10E-05

Figure 43. Growth dynamics of RW109 when exposed to the chosen RNA-Seq control and test conditions

The growth dynamics of RW109 when exposed to the RNA-Seq control and test conditions (1-7) for 24 hours at 30°C are represented by the line graph in panel A). The four biological replicates of each exposure condition are plotted individually on the graph and the numbers represent the test exposure conditions which are described in the key. The mean Log₁₀ OD (450-580 nm) at 24 hours for each condition is shown in panel B) along with the percentage difference in OD when test conditions were compared to the control condition. The mean viable cell count (CFU/mL) for each condition and the percentage difference when the test conditions were compared the control are shown in panel C).

5.2.6 toRNA and mRNA quantification, quality assessment and sequencing outputs

The 32 toRNA extractions and mRNA enrichments, obtained after RW109 was exposed to the control and test conditions, were quantified, and quality assessed. The toRNA extractions used in this study were analysed and the RINs were all ≥ 8 and the rRNAs (23S:16S) ratios were all ≥ 1.5 (Table 19). The concentrations were checked with the Qubit to ensure they were above 100 ng/ μ l; 2 samples were below this value after an additional RNA precipitation step as indicated in Table 19. These 2 samples were allowed to continue to the mRNA stage, as they were considered close enough to 100 ng/ μ l and passed quality assurance checks; the lower concentrations also did not result in the enriched mRNA having a low concentration. The enriched mRNA concentrations were all above the required concentration of 2 ng/ μ l for subsequent cDNA sequence library preparations (Table 19). The mRNA samples were also assessed to ensure the sufficient depletion of the 23S and 16S Ribosomal RNA peaks (Figure 44).

After sequencing the resulting number of paired-end reads for all of the 32 samples ranged from 6.07E+06 to 2.62E+07. The reads were aligned to the sequenced RW109 genome and a mean percentage number of 97.87% were found to align (range of 91.32% - 99.03%). From these aligned reads, the mean percentage number which mapped to a gene feature, was calculated as 98.56% (range of 96.74% - 99.23%). The full details of the number of sequence reads, the number of reads aligned with BWA-MEM and the feature mapping with HTSeq-count are provided on the CD-ROM in the corresponding Chapter and Section folder (SequenceReads_BWA_HTSeq_Stats.xlsx). The raw RNA-Seq reads are available in the ArrayExpress database (<http://www.ebi.ac.uk/arrayexpress>) under accession number E-MTAB-6344. Overall, the sequencing analysis demonstrated the achievement of representative global profiling of gene expression across all the industrially relevant growth conditions assessed (Figure 43).

Table 19. Total RNA and messenger RNA quantification and quality assessment

Biological Replicate Number	Exposure Condition	toRNA Extractions			mRNA Enrichment
		Qubit toRNA Concentration (ng/ul)	toRNA rRNA Ratio (23s / 16s)	toRNA RNA Integrity Number (RIN)	Qubit mRNA Concentration (ng/ul)
1	TSB Only (Control)	357.0	1.7	9.9	27.4
	BIT at 20% MIC (1)	343.0	1.6	10.0	23.2
	BIT at 50% MIC (2)	328.0	1.6	10.0	21.6
	POE at 20% MIC (3)	100.0	1.6	10.0	23.8
	POE at 50% MIC (4)	113.0	1.9	9.6	28.4
	BIT at 20% MIC + POE at 20% MIC (5)	234.0	1.6	10.0	23.0
	1:100 dilution of laundry detergent (6)	172.0	1.5	10.0	27.2
	1:100 dilution of laundry detergent + BIT at 20% MIC (7)	103.0	1.5	10.0	20.8
2	TSB Only (Control)	323.0	1.7	10.0	27.0
	BIT at 20% MIC (1)	401.0	1.6	10.0	24.8
	BIT at 50% MIC (2)	400.0	1.5	10.0	22.6
	POE at 20% MIC (3)	141.0	1.5	10.0	18.3
	POE at 50% MIC (4)	408.0	1.6	10.0	21.8
	BIT at 20% MIC + POE at 20% MIC (5)	112.0	1.5	10.0	18.6
	1:100 dilution of laundry detergent (6)	88.6*	1.5	10.0	38.6
	1:100 dilution of laundry detergent + BIT at 20% MIC (7)	532.0	1.6	9.9	35.2
3	TSB Only (Control)	233.0	1.6	10.0	23.8
	BIT at 20% MIC (1)	374.0	1.5	9.7	19.4
	BIT at 50% MIC (2)	376.0	1.7	9.8	18.4
	POE at 20% MIC (3)	132.0	1.6	10.0	17.4
	POE at 50% MIC (4)	180.0	1.6	10.0	30.2
	BIT at 20% MIC + POE at 20% MIC (5)	125.0	1.7	10.0	22.0
	1:100 dilution of laundry detergent (6)	249.0	1.7	10.0	19.0
	1:100 dilution of laundry detergent + BIT at 20% MIC (7)	97.7*	2.3	10.0	17.0
4	TSB Only (Control)	178.0	1.5	10.0	26.4
	BIT at 20% MIC (1)	323.0	1.5	9.9	19.0
	BIT at 50% MIC (2)	240.0	1.6	10.0	17.8
	POE at 20% MIC (3)	153.0	1.7	10.0	25.6
	POE at 50% MIC (4)	196.2	1.6	9.0	23.2
	BIT at 20% MIC + POE at 20% MIC (5)	106.0	1.6	9.8	25.4
	1:100 dilution of laundry detergent (6)	190.0	1.7	10.0	33.4
	1:100 dilution of laundry detergent + BIT at 20% MIC (7)	103.0	1.7	10.0	20.8

Footnotes: The numbers in brackets next to the exposure conditions indicates the test exposure condition number. * indicates toRNA concentrations below the required concentration of 100 ng/μl. These 2 toRNA samples were allowed to continue forward to mRNA enrichment as they were of high quality and considered close enough to the required concentration. The lower concentration did not result in the enriched mRNA having a low concentration.

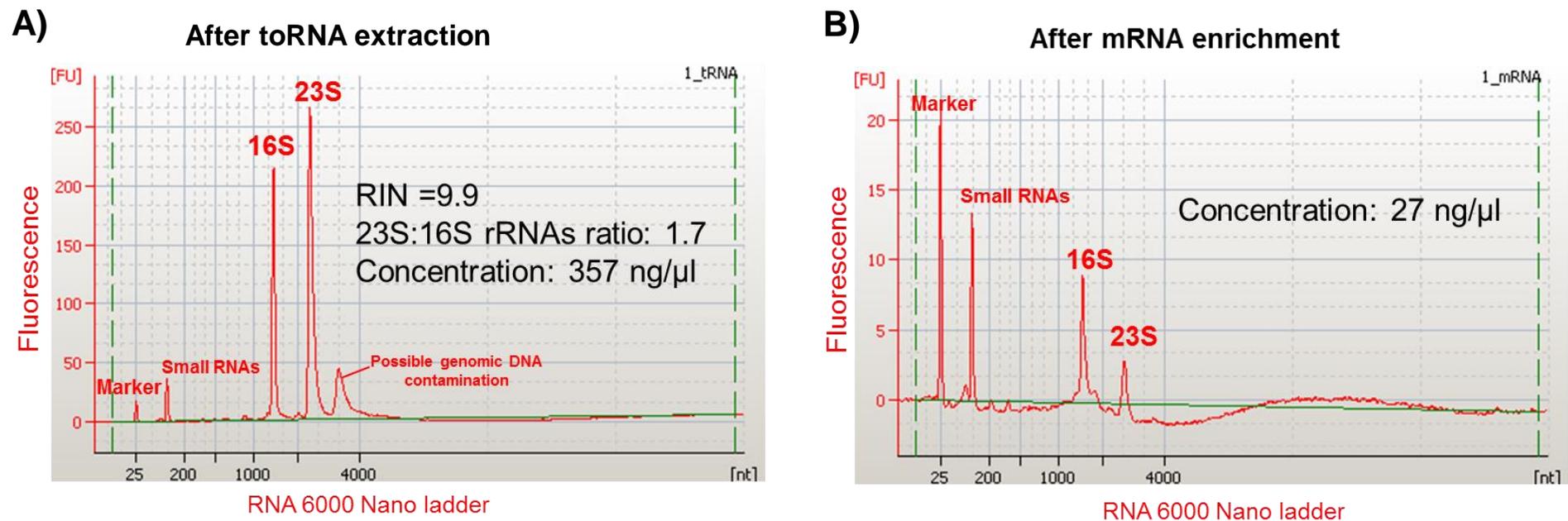


Figure 44. Bioanalyzer electropherograms of total RNA extractions and enriched messenger RNA samples

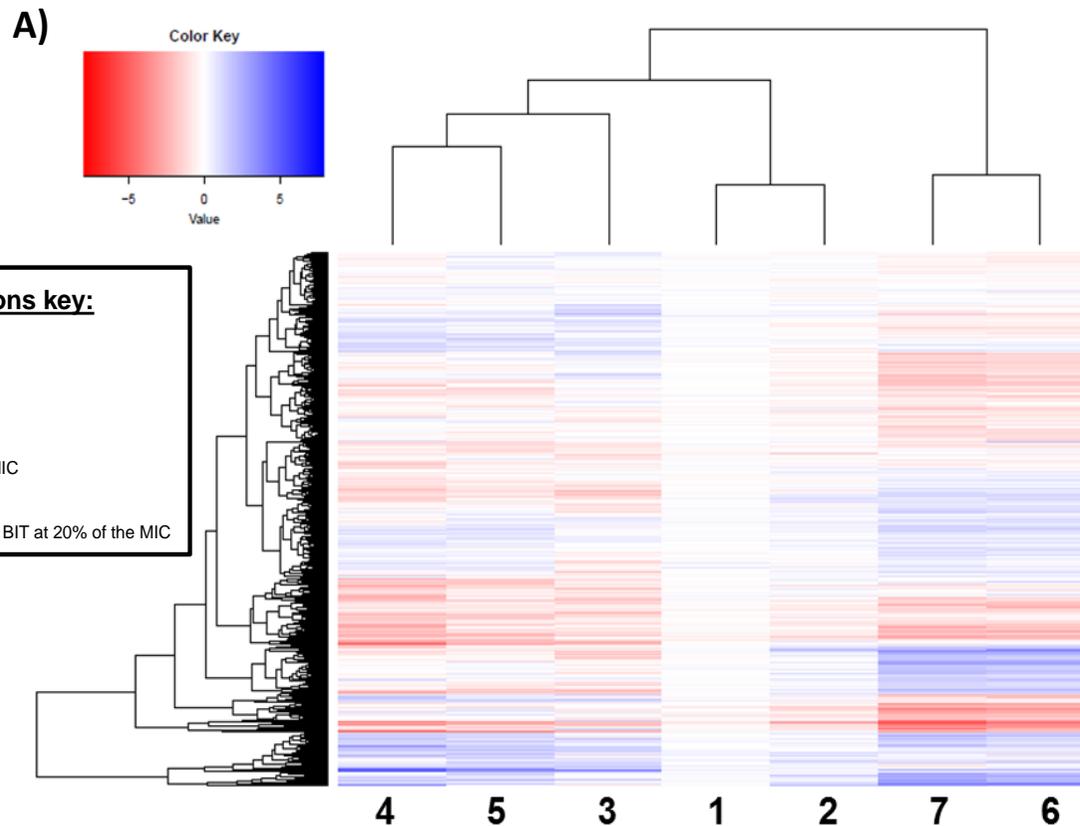
The toRNA extracted when RW109 was exposed to the control condition (TSB only) for 24 hours (replicate 1) was quality assessed with the Bioanalyzer to produce an electropherogram (A). The RIN and 23S:16S rRNAs ratio for this toRNA sample are shown in A) and both were above the recommended values. The concentration of toRNA was checked with the qubit to ensure it was above 100 ng/μl. The mRNA was enriched from this toRNA sample and was also analysed with the Bioanalyzer to produce an electropherogram (B). This analysis was used to ensure sufficient depletion of the 23S and 16S ribosomal peaks when compared to the toRNA sample and this was observed in B). The Qubit was used to ensure the mRNA concentration was above 2 ng/μl. This analysis was carried out for all the 32 RNA extractions used in this study.

5.2.7 RNA-Seq data analysis

5.2.7.1 Overview of gene expression analysis

To summarise the gene expression data of RW109 when exposed to the test exposure conditions, all log₂-fold changes obtained in comparison to the control, were used to generate a heat map (Figure 45A). The preservative BIT at 20% of the MIC (test condition 1) resulted in the least amount of gene expression changes when compared to the other conditions. Laundry detergent in combination with BIT at 20% of the MIC (test condition 7) had the greatest number of gene expression alterations. Within the heat map, the test exposure conditions were grouped according to their gene expression similarities via the top dendrogram. This resulted in three global gene expression groups comprising of test conditions 1 and 2, associated with the presence of BIT, test conditions 3, 4 and 5, linked to the presence POE and test conditions 6 and 7, correlating to the presence of laundry detergent (Figure 45A).

A differentially expressed gene (DEG) was defined as exhibiting a log₂-fold change of ≥ 1.5 and an adjusted p-value of ≤ 0.05 when the test conditions 1-7 were compared against the control condition. Figure 45B illustrates the numbers of genes for each condition, which were either up or down regulated in response to the test conditions. Again the test exposure condition with the least amount of DEGs was BIT at 20% of the MIC; 1 gene was up-regulated and 1 gene was down-regulated. Condition 7, laundry detergent in combination with BIT at 20% of the MIC, had the greatest number of DEGs with 714 up-regulated and 773 down-regulated. Excel spreadsheets of all the identified DEGs with each tab representing a test condition can be found on the CD-ROM provided with this thesis in the corresponding folder (Significantly_Down_Regulated_DEGs.xlsx and Significantly_Up_Regulated_DEGs.xlsx).



B)

Test exposure Condition	No. of up-regulated genes	No. of down-regulated genes
1	1	1
2	63	50
3	183	310
4	346	514
5	279	307
6	594	670
7	714	773

Cut off values: ≥ 1.5 log₂ fold change and adjusted p-value ≤ 0.05
(All test exposure conditions compared to the control condition)

Figure 45. Overview of RNA-Seq gene expression data

A heat map was generated (A) to show the overview of gene expression data for all the RNA-Seq test exposure conditions when compared to the control. The numbers below the heat map represent the test conditions as shown by the test exposure conditions, blue indicates genes which are up-regulated and red represents genes which are down-regulated and the more intense the colours the greater the gene expression (see color key). The top dendrogram displays the test exposure conditions which have been grouped together via hierarchical cluster analysis and the dendrogram to the left of the heat map represents the gene clusters which are grouped according to their log₂-fold change values across all the conditions. The Table (B) represents the number of genes which were differentially regulated for each test exposure condition for the whole genome sequence of RW109 (main chromosome and 2 plasmids). Differentially expressed genes were defined as exhibiting a log₂-fold change of ≥ 1.5 and an adjusted p-value of ≤ 0.05 when the test conditions were compared against the control condition.

5.2.7.2 Classification of DEGs based on KEGG pathway annotations

The DEGs from each test exposure condition were grouped based on KEGG Orthology (KO) numbers and the KEGG pathways characterised. The percentage of DEGs for each KEGG pathway was based on the total number of up or down regulated DEGs with an assigned KO number for each of the test exposure conditions (Table 20). DEGs identified with BIT at 20% of the MIC were excluded due to the low number identified. DEGs without a KO number were also not included in the analysis. A high number of DEGs were assigned to the carbohydrate metabolism pathway for each test condition. A greater percentage were up-regulated for test conditions 2-5 whereas more carbohydrate pathways were down-regulated in conditions 6 and 7 associated with the presence of laundry detergent.

Amino acid metabolism pathways also had a high percentage of DEGs for each test condition. Condition 2, BIT at 50% MIC, had a greater percentage of down-regulated DEGs whereas a higher percentage of up-regulated genes was observed with condition 3. Limited differences were observed between the percentage numbers of up and down regulated DEGs within the amino acid metabolism category with other test conditions. A prominent difference in the percentage of up-regulated DEGs was observed for all test conditions in the energy metabolism pathway group, and this was especially true for test conditions 2 (BIT-associated), 6 and 7 (laundry detergent associated). A greater percentage of membrane transport pathways were identified as down-regulated in conditions 3, 4 and 5 (POE associated); contrastingly more were up-regulated in conditions 2 (BIT-associated) and 6 (laundry detergent associated). An equal percentage of up and down regulated DEGs were observed in condition 7 (laundry detergent associated) for membrane transport (Table 20).

Within biofilm pathways a greater percentage of down-regulated DEGs were observed in all test conditions, except for test condition 3 (POE at 20% MIC) and prominent differences between up and down regulated DEGs were identified in conditions 6 and 7 (laundry detergent associated) (Table 20). All test exposure conditions had a higher percentage of down-regulated DEGs within the quorum-sensing pathway, which was especially true for test condition, 2 (BIT at 50% MIC). The xenobiotics and metabolism pathway had a higher percentage of up-regulated DEGs with test conditions 2, 3, 4 and 5, associated with presence of either BIT or/and POE, whereas a limited difference was observed with conditions 6 and 7 where laundry detergent was present. Test exposure condition 5 had the greatest percentage of up-regulated DEGs within the xenobiotics pathway group followed by condition 4, both linked to the presence of POE. Other notable observations with KEGG pathway analysis included test condition 3, POE at 20% MIC, having the highest percentage of up-regulated genes assigned to the cell motility and signal transduction: two-component system pathways (Table 2). The genes of two RW109 KEGG modules of interest, the BpeEF-OprC efflux pump (M00698) (Chapter 3, Section 3.2.3.2.2) and the UhpB-UhpA two-component regulatory system (M00473) (Chapter 3, Section 3.2.3.2.1), were not identified as significantly up or down regulated in response to the test exposure conditions used within this study.

Table 20. KEGG pathway classifications of the up and down-regulated DEGs for each test exposure condition

KEGG Pathways	Test condition 2: BIT at 50% of the MIC		Test condition 3: POE at 20% of the MIC		Test condition 4: POE at 50% of the MIC		Test condition 5: BIT and POE at 20% of the MIC		Test condition 6: LD Only		Test condition 7: LD and BIT at 20% of the MIC		KEGG Pathways
	Up-regulated DEGs (%)	Down-regulated DEGs (%)	Up-regulated DEGs (%)	Down-regulated DEGs (%)	Up-regulated DEGs (%)	Down-regulated DEGs (%)	Up-regulated DEGs (%)	Down-regulated DEGs (%)	Up-regulated DEGs (%)	Down-regulated DEGs (%)	Up-regulated DEGs (%)	Down-regulated DEGs (%)	
Carbohydrate metabolism	22	12	17	15	25	11	19	8	15	20	16	20	Carbohydrate metabolism
Energy metabolism	10	0	8	6	12	9	6	4	13	7	12	6	Energy metabolism
Amino acid metabolism	17	24	21	16	13	17	19	16	11	16	12	15	Amino acid metabolism
Membrane transport	11	4	0	4	5	8	4	8	10	6	9	9	Membrane transport
Translation	1	4	1	14	1	6	1	8	8	1	8	1	Translation
Metabolism of cofactors and vitamins	3	0	3	4	2	6	2	6	5	4	6	5	Metabolism of cofactors and vitamins
Nucleotide metabolism	6	0	0	7	1	7	1	7	5	3	6	3	Nucleotide metabolism
Signal transduction: Two-component system	4	8	12	2	5	4	6	2	7	7	6	7	Signal transduction-Two-component system
Lipid metabolism	7	4	9	2	6	2	8	3	4	3	4	3	Lipid metabolism
Glycan biosynthesis and metabolism	0	0	1	0	0	1	1	1	3	0	4	0	Glycan biosynthesis and metabolism
Cellular community: Biofilm formation	1	12	5	4	4	5	5	7	4	10	3	9	Cellular community - Biofilm formation
Folding, sorting and degradation	0	0	0	2	0	2	0	3	3	0	3	0	Folding, sorting and degradation
Replication and repair	0	4	0	0	0	0	0	2	2	1	2	1	Replication and repair
Metabolism of other amino acids	4	0	3	2	4	3	5	3	2	3	2	3	Metabolism of other amino acids
Cellular community: Quorum sensing	0	20	0	4	0	3	0	9	2	7	2	7	Cellular community - Quorum sensing
Biosynthesis of other secondary metabolites	2	8	0	5	1	6	1	8	1	3	2	3	Biosynthesis of other secondary metabolites
Metabolism of terpenoids and polyketides	2	0	1	4	2	3	3	3	1	1	1	1	Metabolism of terpenoids and polyketides
Xenobiotics biodegradation and metabolism	8	0	9	4	13	4	18	4	2	3	1	2	Xenobiotics biodegradation and metabolism
Cell motility	1	0	9	2	3	2	2	1	1	6	1	5	Cell motility
Transcription	0	0	0	1	0	1	0	0	1	0	1	0	Transcription

Footnotes: The numbers in each column correspond to the percentage of DEGs for each KEGG pathway in relation to the total number of up or down regulated DEGs with an assigned KO number. The colour intensities decrease with the decreasing percentage of DEGs.

CHAPTER 5 – RNA-Seq gene expression analysis of RW109 in response to preservatives and product formulations

5.2.7.3 Key differentially regulated genes identified following exposure of *P. aeruginosa* RW109 to single preservatives

5.2.7.3.1 Gene expression analysis following exposure to BIT at 20% of the MIC

Gene expression analysis of the RW109 strain when exposed to BIT at 20% of the MIC (test condition 1), identified two significant DEGs. One gene was up-regulated and the other down-regulated and both were on the main chromosome (Table 21). No differentially regulated genes were observed on the two plasmids in response to this test condition. The up-regulated gene was RW109_02248, which was annotated as 3-hydroxyisobutyrate dehydrogenase I and a 1.5 log₂ fold change increase in expression was observed (adjusted p-value ≤ 0.05). RW109_02248 was also up-regulated in response to BIT at 50% of the MIC with a 1.8 log₂-fold increase (adjusted p-value ≤ 0.05). This gene was characterised with a KO number involved in valine degradation (K00020) within the amino acid metabolism KEGG pathway.

The one down-regulated gene identified after exposure to BIT at 20% of the MIC encoded a hypothetical protein (RW109_05945) and expression decreased with a -1.6 log₂-fold change (adjusted p-value ≤ 0.05). This hypothetical annotated gene was also down-regulated in response to BIT at 50% of the MIC, where a -2.6 log₂-fold change was observed (adjusted p-value ≤ 0.05). No KO numbers were assigned to this hypothetical protein. In response to BIT at 50% of the MIC, 5 other hypothetical protein encoding genes within close proximity to RW109_05945 included RW109_05942-05943 and RW109_05947-05949 were all down-regulated ≥ 1.7 log₂-fold in response to 50% of the MIC.

Table 21. The number of genes differentially regulated in the main chromosome and two plasmids after exposure to the test exposure conditions

Test exposure condition	Up-regulated genes								Down-regulated genes								Overall proportions of DEGs*	
	Plasmid 1		Main Chromosome		Plasmid 2		Total		Plasmid 1		Main Chromosome		Plasmid 2		Total			
	Number	Percentage (%)	Number	Percentage (%)	Number	Percentage (%)	Number	Percentage (%)	Number	Percentage (%)	Number	Percentage (%)	Number	Percentage (%)	Number	Percentage (%)	Number	Percentage (%)
1) BIT at 20% of the MIC	0	0	1	0.02	0	0	1	0.01	0	0	1	0.02	0	0	1	0.01	2	0.03
2) BIT at 50% of the MIC	0	0	63	0.96	0	0	63	0.086	0	0	50	0.76	0	0	50	0.68	113	1.55
3) POE at 20% of the MIC	13	2.21	160	2.44	10	6.41	183	2.51	4	0.68	306	4.67	0	0	310	4.24	493	6.75
4) POE at 50% of the MIC	10	1.70	320	4.88	16	10.26	346	4.74	6	1.02	507	7.73	1	0.64	514	7.04	860	11.78
5) BIT and POE in combination	20	3.40	248	3.78	11	7.05	279	3.82	5	0.85	301	4.59	1	0.64	307	4.20	586	8.02
6) LD with no preservative	7	1.19	581	8.86	6	3.85	594	8.13	32	5.43	627	9.56	11	7.05	670	9.17	1264	17.31
7) LD with BIT	23	3.90	678	10.34	13	8.33	714	9.78	35	5.94	724	11.04	14	8.97	773	10.58	1487	20.36

Footnotes: The percentages represent the proportion of DEGs in relation to the size of each replicon. In the total columns the percentages represent the proportion of DEGs in relation and to the genome as a whole. * total number of up and down regulated DEGs

5.2.7.3.2 Gene expression analysis following exposure to BIT at 50% of the MIC

Gene expression analysis of the RW109 strain when exposed to BIT at 50% of the MIC (test condition 2), identified 113 significant DEGs (Table 21); 63 were up-regulated and 50 were down-regulated. All the DEGs identified in response to condition 2 were on the main chromosome.

Within the top 10 up-regulated DEGs, an anthranilate 1,2-dioxygenase large subunit, *antA* (RW109_03497) and an anthranilate 1,2-dioxygenase small subunit *antB* (RW109_03496) were identified (Table 22). These genes are predicted to be involved in the first step of a sub-pathway, which synthesises catechol, during the degradation of anthranilate via hydroxylation. Another gene associated with this sub-pathway, anthranilate 1,2-dioxygenase electron transfer component, *antC* (RW109_03495), was also significantly up-regulated in response to BIT at 50% of the MIC, with a 2.0 log₂-fold change. These three genes were assigned KO numbers, which mapped to the KEGG xenobiotic biodegradation and metabolism pathways. RW109_03496 was associated with benzoate, fluorobenzoate and xylene degradation, and RW109_03497 and RW109_03495 were linked to aminobenzoate degradation.

Genes RW109_02633, RW109_02635 and RW109_02636 were also within the top ten up-regulated genes (Table 22) and all had KO numbers associated with a glucose/mannose ABC transport system. The RW109_02634 gene was also associated with this transport system and in response to BIT at 50% of the MIC, a 2.4 log₂-fold increase was observed. Furthermore, genes encoding for a putative sialic acid transporter (RW109_03936) and a porin B precursor (RW109_02637) were identified in the top ten up-regulated genes (Table 22).

Other up-regulated DEGs of interest included RW109_04035-04040, which were all co-located on the main chromosome. The genes were involved in leucine degradation and all had an increased expression of between 1.5 - 2.6 log₂-fold. The amino acid sequences were analysed using the *Pseudomonas* database DIAMOND BLASTP search (Winsor et al., 2016) and a high degree of sequence similarity was identified with the PA14 *gnyRDBHAL* gene cluster (mean BLASTP scores: percentage ID= 98.27%, alignment length= 379, E-value= 0.0 and BIT Score= 734). Other genes also involved in leucine degradation encoded for an acetyl-CoA acetyltransferase (RW109_04050, 1.5log₂-fold), and subunits A and B of a putative succinyl-CoA:3-ketoacid coenzyme A transferase (RW109_04051, 2.2 log₂-fold and RW109_04052, 2.0 log₂-fold).

In response to BIT at 50% of the MIC, 5 up-regulated genes involved in sulphur metabolism were identified. These included a sulfite reductase [ferredoxin] (RW109_04218, 1.9 log₂-fold), a bi-functional enzyme CysN/CysC (RW109_05781, 1.7 log₂-fold), a sulfate adenylyl-transferase subunit II (RW109_05782, 1.6 log₂-fold), a cytochrome c-551 precursor (RW109_01122, 1.7log₂-fold) and a sulfate-binding protein precursor, which was identified within the top 10 up-regulated DEGs (RW109_00880,) (Table 22). Genes involved in nitrogen metabolism were also up-regulated and encoded for a carbamate kinase 1 (RW109_06773, 1.9 log₂-fold) and a nitrite reductase precursor (RW109_01123, 2.0 log₂-fold). A high number of up-regulated DEGs in response to BIT at 50% of the MIC were

characterised within the KEGG energy metabolism category (Table 20) which includes sulphur and nitrogen metabolism sub-pathways.

Within the top ten down-regulated genes in response to BIT at 50% of the MIC, included 6 within close proximity to each other in the RW109 genome, (Table 23). A gene annotated as a PQB biosynthetic 3-oxoacyl-[acyl-carrier-protein] synthase III, (RW109_05184) had decreased expression with a 2.0 log₂-fold change and was in close proximity to those identified in Table 23. From the *Pseudomonas* database analysis, the RW109_05181-05187 had a high degree of sequence similarity was identified with the PA14 quorum sensing operons *phnAB* and *pqsABCDE* (mean BLASTP scores: percentage ID= 99.51%, alignment length= 358, E-value= 0.0 and BIT Score= 723).

Out of the DEGs down-regulated in response to BIT at 50% of the MIC, 24% were classified within the amino acid metabolism pathway (Table 20). Further analysis revealed these genes were involved with leucine biosynthesis, inversely correlating to the up-regulation of leucine degradation genes, (RW109_04035-04040) as noted above. These encoded a 3-isopropylmalate dehydrogenase (RW109_02705, -1.7 log₂-fold), a 2-isopropylmalate synthase (RW109_01995, -1.5 log₂-fold) and a 3-isopropylmalate dehydratase small subunit I (RW109_02703, -1.5 log₂-fold). An asparagine synthase annotated gene involved in aspartate metabolism was also down-regulated in response to BIT at 50% of the MIC (RW109_00655, -1.6 log₂-fold).

Table 22. Top 10 up-regulated genes in response to BIT at 50% of the MIC

Transcript ID	Prokka putative gene function	Cog category	Log2 fold change	Adjusted P-value
Main Chromosome				
RW109_03497	Anthranilate 1,2-dioxygenase large subunit	Q	3.9	1.24E-13
RW109_02636	sn-glycerol-3-phosphate import ATP-binding protein UgpC	G	3.5	7.26E-16
RW109_03936	Putative sialic acid transporter	G	3.4	7.10E-10
RW109_02633	Putative sugar-binding periplasmic protein precursor	G	3.4	2.60E-20
RW109_02637	Porin B precursor	M	3.2	2.49E-13
RW109_00880	Sulfate-binding protein precursor	P	2.8	4.25E-07
RW109_03496	Anthranilate 1,2-dioxygenase small subunit	Q	2.7	0.000032
RW109_02628	D-erythrose-4-phosphate dehydrogenase	G	2.7	2.89E-08
RW109_02635	L-arabinose transport system permease protein AraQ	P	2.7	0.0000105
RW109_04038	1,4-Dihydroxy-2-naphthoyl-CoA synthase	I	2.6	1.27E-06

Table 23. Top 10 down-regulated genes in response to BIT at 50% of the MIC

Transcript ID	Prokka putative gene function	Cog category	Log2 fold change	Adjusted P-value
Main chromosome				
RW109_05182	Anthranilate synthase component II, pyocyanine specific*	E	-3.0	5.87E-09
RW109_05185	3-oxoacyl-(acyl carrier protein) synthase III*	I	-2.8	1.81E-09
RW109_05945	Hypothetical protein	S	-2.6	1.77E-20
RW109_05186	3-oxoacyl-(acyl carrier protein) synthase III*	Unknown	-2.6	6.08E-09
RW109_05183	Putative diflavin flavoprotein-A III*	S	-2.4	0.0000341
RW109_05181	Anthranilate synthase component II, pyocyanine specific*	E	-2.4	0.0000971
RW109_05187	Anthranilate-CoA ligase*	I	-2.3	5.02E-08
RW109_07078	Hypothetical protein	Unknown	-2.3	2.82E-10
RW109_05948	Hypothetical protein	S	-2.3	0.0000448
RW109_04992	Hypothetical protein	M	-2.2	0.0000338

Footnotes: *indicates genes in close proximity to each other within the RW109 genome

5.2.7.3.3 Gene expression analysis following exposure to POE at 20% of the MIC

Gene expression analysis of the RW109 strain when exposed to POE at 20% of the MIC (test condition 3), identified 493 significant DEGs (Table 21). On the main chromosome, 160 genes were up-regulated and 306 were down-regulated. On plasmid 1, 13 genes were up-regulated and 4 genes were down-regulated. Within plasmid 2, 10 genes were up-regulated and no genes were identified as down-regulated (Table 21).

Within the top 10 up-regulated main chromosome genes, RW109_02297 and RW109_02298 were identified and formed part of a MexPQ-OprM efflux pump system (Table 24). The gene annotated as the outer membrane OprM protein (RW109_02299) was not included in the top ten up-regulated genes, but did have an increased expression of 3.1 log₂-fold. Also recognised were 3 genes in close proximity to each other on the main chromosome with one having the greatest log₂-fold change and encoded for a hypothetical protein (RW109_01137). The other 2 genes were annotated as a gamma-glutamylputrescine oxidoreductase, PuuB (RW109_01139) and a HTH-type transcriptional regulator PuuR (RW109_01140) and both had a log₂-fold increase of ≥ 4.7 . These genes were involved in Gamma-aminobutyric acid (GABA) biosynthesis from putrescine within the amino acid metabolism KEGG pathway. PuuB functioned in the breakdown of putrescine via the oxidation of L-glutamylputrescin and PuuR regulates the putrescine utilisation pathway.

Out of the up-regulated genes characterised in the KEGG signal transduction, 12% were assigned to the two-component system pathway (Table 20). Two genes, RW109_05645 and RW109_05646, both encoded for the PctA and PctB methyl-accepting chemotaxis proteins and were up-regulated by 1.7 and 1.8 log₂-fold respectively. The genes, RW109_01838 and RW109_01839, were previously identified as encoding for a NarX-NarL (nitrate respiration) two-component regulatory system (Chapter 3, Section 3.2.3.2). These had an increased expression of 2.1 and 1.7 log₂-fold respectively in response to POE at 20% of the MIC. Two other up-regulated genes identified were RW109_04786 and RW109_04787, which both encode for the chemotaxis protein CheW.

Up-regulated genes of interested classified within the cell motility pathway, had functions in flagella assembly and encoded for an FlgN protein (RW109_02463, 1.7log₂-fold) and an anti-sigma-28 factor, FlgM (RW109_02462, 2.3log₂-fold). Two other notable genes were annotated as Flagellar protein FliS (RW109_05082, 1.8log₂-fold) and a B-type flagellin (RW109_05085, 2.2log₂-fold). Within the xenobiotics and biodegradation pathway, noticeable up-regulated genes included RW109_01049 (2.2 log₂-fold) and RW109_04050 (2.4 log₂-fold) which encoded for acyl-CoA dehydrogenase and acetyl-CoA acetyltransferase respectively; both enzymes were involved in benzoate degradation. Two genes involved in caprolactam degradation encoded for the alkane 1-monooxygenase I and II enzymes (RW109_03429, 1.8 log₂-fold and RW109_04639, 2.0 log₂-fold). A gene annotated as an alcohol dehydrogenase, RW109_07041, was up-regulated in response to POE at 20% of the MIC with a 2.5 log₂-fold. The RW109_04035-04039 genes, identified as part of the *gnyRDBHAL* gene cluster, also had an increased expression of between 1.8 – 4.2 log₂-fold with this test condition.

CHAPTER 5 – RNA-Seq gene expression analysis of RW109 in response to preservatives and product formulations

Within the top 10 genes up-regulated on plasmid 1 in response to POE, a mercuric transport protein periplasmic component precursor (RW109_00024) had the greatest increase in expression (Table 24). This was also the only up-regulated plasmid 1 gene assigned with a KO number and was characterised as a pores ion channel. Seven out of the 10 up-regulated plasmid 2 genes were annotated as hypothetical proteins. Three genes of interest encoded for an antitoxin DinJ (RW109_07361), an mRNA interferase YafQ (RW109_07362) and an antitoxin ChpS (RW109_07367). The genes encoding DinJ and YafQ, work together as a toxin-antitoxin (TA) system. The YafQ toxin is an RNase enzyme which interferes with translation and the DinJ is an antitoxin which counteracts the effects of YafQ toxicity (Armalytè et al., 2012). ChpS was also part of a TA-System involving ChpB and RW109 was found to encode for ChpB on plasmid 2. However, no gene expression changes in response to the industrial conditions were observed for this gene.

The main chromosome genes within the top 10 POE down-regulated genes are shown in Table 25. Of these, RW109_03624, which encoded an L-ornithine 5-monooxygenase, had the largest fold change (Table 25). The RW109_03625 and RW109_03626 (not included in the top 10 down-regulated genes) were of interest due to their close proximity to RW109_03624. These encoded for an acyl-homoserine lactone acylase PvdQ (RW109_03625) and a ferric uptake regulation protein (RW109_03626) and were down-regulated ≥ 2.9 log₂-fold. The enzymes encoded by RW109_03624 and RW109_03625 were involved in quorum sensing and the RW109_03626 was involved in the regulation of iron-uptake. Other genes of note in the top 10 down-regulated genes were RW109_05337 encoding a GlcNAc-binding protein A precursor, and RW109_02092 which encodes for a pseudolysin precursor. A RW109_03982 gene encoding for a decarbamoylnovobiocin carbamoyltransferase (K00612) was also identified. Siderophore biosynthesis genes RW109_01443-01450 which had decreased expression between the ranges of 1.6 – 3.6 log₂-fold in response to POE at 20% of the MIC, were also notable. The genes RW109_05182-05187, previously identified as part of *phnAB* and *pqsABCDE* operons, were also down-regulated in response to test condition 3 and a decrease of between 1.6 – 2.2 log₂-fold was observed. The plasmid 1 gene with the greatest decrease in expression was RW109_00284 and encoded for a taurine-pyruvate aminotransferase (Table 25).

Table 24. Top 10 up-regulated genes in response to POE at 20% of the MIC

Transcript ID	Prokka putative gene function	Cog category	Log2 fold change	Adjusted P-value
Plasmid 1				
RW109_00024	Mercuric transport protein periplasmic component precursor	P	2.5	0.033044
RW109_00588	Chromosome partition protein Smc	S	1.8	6.79E-15
RW109_00303	Leucine-responsive regulatory protein	K	1.7	5.24E-08
RW109_00516	Hypothetical protein	Unknown	1.7	0.02571
RW109_00367	Arc-like DNA binding domain protein	Unknown	1.6	0.010028
RW109_00019	DNA-invertase hin	L	1.6	3.24E-11
RW109_00048	Hypothetical protein	S	1.6	0.00000119
RW109_00101	Hypothetical protein	S	1.6	0.00000873
RW109_00562	HTH-type transcriptional regulator DmlR	K	1.6	0.00000831
RW109_00020	Hypothetical protein	S	1.6	0.0000403
Main chromosome				
RW109_01137	Hypothetical protein	Unknown	6.7	7.00E-34
RW109_02297	Membrane fusion protein MexP	V	5.5	6.94E-15
RW109_02298	Efflux pump membrane transporter MexQ	V	5.0	3.65E-14
RW109_01140	HTH-type transcriptional regulator PuuR	K	4.9	1.82E-14
RW109_01139	Gamma-glutamylputrescine oxidoreductase	E	4.7	5.28E-69
RW109_02242	Copper chaperone CopZ	P	4.5	8.99E-23
RW109_04035	HTH-type transcriptional regulator CueR	K	4.2	8.05E-28
RW109_05611	Hca operon transcriptional activator	K	4.0	5.01E-52
RW109_05610	Putative efflux pump outer membrane protein TtgC precursor	M	3.9	1.90E-75
RW109_06470	Multidrug-efflux transporter 1 regulator	K	3.9	1.05E-32
Plasmid 2				
RW109_07319	Hypothetical protein	Unknown	2.3	4.37E-21
RW109_07282	Hypothetical protein	S	2.3	2.67E-34
RW109_07367	Antitoxin ChpS	Unknown	2.2	0.0000634
RW109_07253	Hypothetical protein	Unknown	2.1	2.14E-18
RW109_07362	mRNA interferase YafQ	S	2.1	1.12E-09
RW109_07255	Hypothetical protein	Unknown	2.0	5.85E-19
RW109_07254	Hypothetical protein	Unknown	2.0	8.10E-09
RW109_07330	Hypothetical protein	Unknown	1.7	0.00614
RW109_07361	Antitoxin DinJ	L	1.7	1.35E-08
RW109_07256	Hypothetical protein	Unknown	1.6	0.00000989

Footnotes = Top 10 genes from the main chromosome and the two plasmids are shown

Table 25. Top 10 down-regulated genes in response to POE at 20% of the MIC

Transcript ID	Prokka putative gene function	Cog category	Log2 fold change	Adjusted P-value
Plasmid 1				
RW109_00284	Taurine-pyruvate aminotransferase	H	-2.2	0.012463
RW109_00047	PhoD-like phosphatase	S	-1.7	1.58E-09
RW109_00460	Imelysin	S	-1.7	0.016447
RW109_00125	Hypothetical protein	E	-1.6	0.006475
Main chromosome				
RW109_03624	L-ornithine 5-monooxygenase	Q	-5.2	1.56E-16
RW109_05337	GlcNAc-binding protein A precursor	G	-4.8	4.75E-37
RW109_03982	Decarbamoylnovobiocin carbamoyltransferase	O	-4.7	5.28E-13
RW109_03587	Linear gramicidin synthase subunit B	Q	-4.6	3.90E-09
RW109_07078	Hypothetical protein	Unknown	-4.6	1.20E-10
RW109_04157	Hypothetical protein	S	-4.5	0.00000301
RW109_03598	Diaminobutyrate--2-oxoglutarate aminotransferase	E	-4.2	0.00000365
RW109_02092	Pseudolysin precursor	E	-4.2	1.26E-101
RW109_02394	Hemophore HasA	S	-4.1	0.00000159
RW109_03645	Hypothetical protein	S	-4.0	7.88E-13

Footnotes = Top 10 genes from the main chromosome are shown as well as and the total 4 down-regulated from plasmid 1.

5.2.7.3.4 Gene expression analysis following exposure to POE at 50% of the MIC

Gene expression analysis of the RW109 strain when exposed to POE at 50% of the MIC (test condition 4) identified 860 significant DEGs (Table 21). On the main chromosome, 320 were up-regulated and 507 were down-regulated. Ten genes were up-regulated on plasmid 1 and 6 genes were down-regulated and on plasmid 2, 16 were up-regulated and one gene was down-regulated.

Within the top 10 up-regulated on the main chromosome, the gene with the greatest fold change encoded for a hypothetical protein RW109_01137 (Table 26), which was also had the highest increase in gene expression in response to POE at 20% of the MIC (5.2.7.3.3). A putative sialic acid transporter associated with RW109_03936 had the second greatest log₂-fold increase. The *antA* RW109_03497 gene was identified as the third gene most up-regulated on the main chromosome. The *antB* RW109_03496 gene and *antC* RW109_03495 gene also both had increases in expression with 4.7 log₂-fold and 3.3 log₂-fold changes observed respectively (Table 26). These three genes were previously identified with BIT at 50% of the MIC (Section 5.2.7.3.2).

The efflux pump membrane transporter MexQ (RW109_02298), was also identified within the top up-regulated main chromosome genes (Table 26). The membrane fusion protein MexP (RW109_02297, 4.9log₂-fold) and outer membrane protein OpmE precursor (RW109_02299, 3.2log₂-fold) were significantly up-regulated but not included within the top ten. Other genes of interest were the putative tartrate transporter (RW109_03749) and a bacterioferritin encoding gene (RW109_02289) (Table 26). The RW109_04035-04040 genes which form the *gnyRDBHAL* gene cluster were also up-regulated in response to POE at 50% of the MIC with increases observed between 3.6 – 5.5 log₂-fold. A transcriptional regulator, KdgR (RW109_04041) was also significantly up-regulated with a 2.6 log₂-fold change. The RW109_01139 and RW109_01140 genes involved in GABA biosynthesis from putrescine were up-regulated by 2.3 and 3.8 log₂-fold respectively, and were identified with increased expression in response to POE at 20% of the MIC (Section 5.2.7.3.3).

Within the xenobiotics biodegradation and metabolism KEGG pathway, 4 genes were identified as up-regulated in response to POE at 50% of the MIC and were involved in benzoate degradation. These were annotated as a muconolactone delta-isomerase (RW109_03501, 3.8 log₂-fold), a catechol 1,2-dioxygenase (RW109_03502, 4.8 log₂-fold), an acyl-CoA dehydrogenase (RW109_01049, 1.9 log₂-fold) and an acetyl-CoA acetyltransferase (RW109_04050, 1.8 log₂-fold). The RW100_07041 and RW109_02105 genes were up-regulated in response to POE at 50% of the MIC by 2.0 and 1.7 log₂-fold respectively and encoded for alcohol dehydrogenases. The RW100_07041 gene was also up-regulated with POE at 20% of the MIC by 2.5 log₂-fold. POE is an alcohol based preservative making genes encoding putative alcohol dehydrogenases of interest.

The gene with the highest increase in expression on plasmid 1 in response to POE at 50% of the MIC encoded a glutathione S-transferase GST (RW109_00018) (Table 26). This gene was characterised as part of the KEGG xenobiotics by cytochrome P450 pathway. The majority of the genes with increased

CHAPTER 5 – RNA-Seq gene expression analysis of RW109 in response to preservatives and product formulations

expression on plasmid 2 encoded for hypothetical proteins (Table 26). The genes encoding for the TA-System antitoxin ChpS (RW109_07367) and mRNA interferase YafQ (RW109_07362) that increased expression in response to POE at 20% of the MIC (Section 5.2.7.3.3) were also up-regulated in response to POE at 50% of the MIC.

On the main chromosome, a hypothetical protein encoded gene (RW109_05811) had the greatest log₂-fold change within the top ten down-regulated genes (Table 27). The RW109_05337 gene involved in chitin-binding, a pseudolysin precursor gene RW109_02092 and a decarbamoylnovobiocin carbamoyltransferase encoding RW109_03982 gene were down-regulated in response to POE at 50% of the MIC and were identified in Table 27. These three genes were also down-regulated in response to POE at 20% of the MIC. Other notable genes identified on the main chromosome were RW109_01443- RW109_01450, which had a decrease in expression within the range of 2.7 – 6.0 log₂-fold. These genes were also down-regulated in response to POE at 20% of the MIC and functioned in the biosynthesis of siderophores (Section 5.2.7.3.3). The down-regulation of these 8 genes together was only observed with POE at 20% and 50% of the MIC. The RW109_05181-05187 genes previously identified as part of the *phnAB* and *pqsABCDE* operons (Section 5.2.7.3.2), were also down-regulated in response to POE at 50% of the MIC, with a decrease of between 1.8 – 4.4 log₂-fold.

A Cytochrome c encoding gene RW109_00461 had the greatest decrease in expression on plasmid 1. The taurine-pyruvate aminotransferase encoding gene RW109_00284 previously identified with POE at 20% of the MIC was also down-regulated with 50% of the MIC (Table 27). Other plasmid 1 genes of interest were an acrylyl-CoA reductase AcuI (RW109_00560) and a formyltetrahydrofolate deformylase (RW109_00469). RW109_07290 encoding for a hypothetical protein was the only plasmid 2 down-regulated gene identified.

Table 26. Top 10 up-regulated genes in response to POE at 50% of the MIC

Transcript ID	Prokka putative gene function	Cog category	Log2 fold change	Adjusted P-value
Plasmid 1				
RW109_00018	Glutathione S-transferase GST-6.0	O	2.3	0.000578
RW109_00576	Hypothetical protein	Unknown	2.0	7.96E-11
RW109_00005	DNA-invertase hin	L	1.9	4.72E-13
RW109_00180	Hypothetical protein	Unknown	1.9	0.014388
RW109_00516	Hypothetical protein	Unknown	1.9	0.001518
RW109_00386	Hypothetical protein	Unknown	1.7	0.031553
RW109_00068	Bifunctional (p)ppGpp synthase/hydrolase RelA	S	1.7	0.00093
RW109_00589	Site-specific tyrosine recombinase XerD	L	1.5	4.66E-13
RW109_00170	Putative S-adenosylmethionine-dependent methyltransferase/MSMEI_2290	Q	1.5	2.92E-08
RW109_00588	Chromosome partition protein Smc	S	1.5	6.24E-08
Main chromosome				
RW109_01137	Hypothetical protein	Unknown	6.3	5.87E-21
RW109_03936	Putative sialic acid transporter	G	6.0	7.97E-15
RW109_03497	Anthranilate 1,2-dioxygenase large subunit	Q	5.7	9.42E-13
RW109_02298	Efflux pump membrane transporter MexQ	V	5.5	2.02E-14
RW109_03749	Putative tartrate transporter	G	5.5	4.79E-12
RW109_04035	HTH-type transcriptional regulator CueR	K	5.5	5.00E-49
RW109_02628	D-erythrose-4-phosphate dehydrogenase	G	5.5	1.74E-29
RW109_02289	Bacterioferritin	P	5.3	3.76E-70
RW109_04036	Acyl-CoA dehydrogenase	I	5.3	6.68E-70
RW109_03934	Hypothetical protein	S	5.2	9.79E-09
Plasmid 2				
RW109_07256	Hypothetical protein	Unknown	2.5	5.22E-18
RW109_07367	Antitoxin ChpS	Unknown	2.2	0.00000611
RW109_07352	Hypothetical protein	Unknown	2.0	0.021039
RW109_07312	Hypothetical protein	Unknown	1.9	0.006913
RW109_07362	mRNA interferase YafQ	S	1.9	0.00000105
RW109_07382	Hypothetical protein	Unknown	1.9	1.69E-08
RW109_07266	Transposase DDE domain protein	L	1.9	9.24E-11
RW109_07365	Hypothetical protein	Unknown	1.8	0.0000792
RW109_07254	Hypothetical protein	Unknown	1.8	0.00000758
RW109_07250	Hypothetical protein	Unknown	1.8	5.11E-18

Footnotes = Top 10 genes from the main chromosome and the two plasmids are shown

Table 27. Top 10 down-regulated genes in response to POE at 50% of the MIC

Transcript ID	Prokka putative gene function	Cog category	Log2 fold change	Adjusted P-value
Plasmid 1				
RW109_00461	Cytochrome c	C	-2.7	0.01304
RW109_00269	Hypothetical protein	S	-1.8	0.0000424
RW109_00284	Taurine-pyruvate aminotransferase	H	-1.7	0.009095
RW109_00135	Hypothetical protein	S	-1.6	2.98E-08
RW109_00560	Acrylyl-CoA reductase Acul	C	-1.6	0.000279
RW109_00469	Formyltetrahydrofolate deformylase	F	-1.6	0.000918
Main chromosome				
RW109_05811	Hypothetical protein	S	-7.9	5.85E-23
RW109_03585	Putative RNA polymerase sigma factor FecI	K	-7.7	1.73E-17
RW109_02092	Pseudolysin precursor	E	-7.2	1.48E-226
RW109_05337	GlcNAc-binding protein A precursor	G	-6.9	2.41E-51
RW109_01518	Pesticin receptor precursor	P	-6.8	1.82E-23
RW109_05808	Superoxide dismutase [Mn/Fe]	P	-6.8	2.65E-27
RW109_03711	Chitinase D precursor	G	-6.7	3.84E-18
RW109_03982	Decarbamoylnovobiocin carbamoyltransferase	O	-6.6	3.32E-18
RW109_05809	Hypothetical protein	S	-6.5	2.14E-17
RW109_05840	Periplasmic dipeptide transport protein precursor	E	-6.5	3.20E-37
Plasmid 2				
RW109_07290	Hypothetical protein	Unknown	-2.0	0.007861

Footnotes = Top 10 genes from the main chromosome are shown as well as the total 6 down-regulated on plasmid 1 and the one from plasmid 2

5.2.7.4 Key differentially regulated genes identified following exposure of RW109 to a combination of preservatives

5.2.7.4.1 Gene expression analysis following exposure to a combination of BIT and POE both at 20% of the MIC

Gene expression analysis of the RW109 strain when exposed to BIT and POE in combination at 20% of the MIC (test condition 5) identified 586 significant DEGs (Table 21). On the main chromosome, 248 were up-regulated and 301 genes were down-regulated and on plasmid 1, 20 genes were up-regulated and 5 genes were down-regulated. Eleven genes had increased in expression on plasmid 2, and one gene was down-regulated.

Within the top ten up-regulated genes on the main chromosome, the *antA* RW109_03497 gene had the greatest log₂-fold change (Table 28). The *antB* RW109_03496 gene had the fourth greatest increase in expression and the *antC* RW109_03495 gene was also identified as up-regulated in response to POE and BIT in combination with a 3.4 log₂-fold increase (not included within the top ten unregulated). These three genes were involved in the degradation of anthranilate via hydroxylation and previously observed had an increased expression with test conditions 2 and 4 (Sections 5.2.7.3.2 and 5.2.7.3.4). The genes encoding MexQ (RW109_02298) and MexP (RW109_02298) had the second and third biggest increases in expression when exposed to BIT and POE in combination (Table 28). Also up-regulated by 3.1 log₂-fold was the RW109_02299 gene annotated as the OprM outer membrane protein was encoded in close proximity to the latter efflux genes, although its fold increase was not within the top 10 up-regulated genes. Other genes of interest up-regulated in response to BIT and POE in combination, were a putative sialic acid transporter encoding gene (RW109_03936), a gene annotated as a copper chaperone CopZ (RW109_02242) and a citrate transporter encoding gene (RW109_04047) (Table 28).

Out of the genes up-regulated in response to BIT and POE in combination, 18% were characterised as part of the xenobiotics metabolism and degradation KEGG pathway (Table 20). Within this pathway three genes involved in benzoate degradation were found in close proximity to each other on the RW109 genome and encoded for a muconate cycloisomerase 1 (RW109_03500, 2.5 log₂-fold), a muconolactone delta-isomerase Q (RW109_03501, 3.1 log₂-fold) and a catechol 1,2-dioxygenase Q (RW109_03502, 4.3 log₂-fold). Other up-regulated genes also involved in benzoate degradation encoded an acetyl-CoA acetyltransferase (RW109_04050, 2.5 log₂-fold) and a carboxymuconolactone decarboxylase family protein (RW109_00826, 1.6 log₂-fold). A gene on plasmid 1, RW109_00018 encoded a glutathione S-transferase, was up-regulated by 2.2 log₂-fold (Table 28) and characterised as part of the metabolism of xenobiotics by cytochrome P450. RW109_00018 was also up-regulated with POE at 50% of the MIC. Three genes encoding for alcohol dehydrogenases were also identified with increased expression in response to BIT and POE in combination and these included RW109_07041 (2.6 log₂-fold), RW109_02105 (1.5 log₂-fold) and RW109_06146 (2.2 log₂-fold). The RW109_04035-04040 genes, which form the *gnyRDBHAL* gene cluster, were also up-regulated in response to BIT and POE with

CHAPTER 5 – RNA-Seq gene expression analysis of RW109 in response to preservatives and product formulations

increased expression observed between 3.0 – 4.2 log₂-fold. The RW109_04041 gene, which encodes for the transcriptional regulator, KdgR was also significantly up-regulated with a 2.2 log₂-fold change.

On plasmid 1 in response to BIT and POE, a gene encoding for a hypothetical protein had the greatest increase in expression with a 2.3 log₂-fold change. Five other genes annotated as hypothetical proteins were also identified within the top ten up-regulated on plasmid 1 (Table 28). This was also the case with plasmid 2 where 8 hypothetical encoding genes had increased expression in response to BIT and POE in combination (Table 28).

The gene with the greatest fold decrease on the main chromosome in response to the combined preservatives encoded a pesticin receptor precursor (RW109_01518) (Table 29). Four genes within close proximity on the main chromosome encoded a superoxide dismutase (RW109_05808), two hypothetical proteins (RW109_05809 and RW109_05811) and a fumarate hydratase class II (RW109_05810) (Table 29). The RW109_05181-05187 genes, which encode for the quorum sensing *phnAB* and *pqsABCDE* operons, were also all down-regulated after combined preservative exposure, with a fold decrease of 2.4 – 3.6 log₂-fold observed. The genes RW109_01443-01450 which functioned in the biosynthesis of siderophores were all down-regulated in response to the BIT-POE combination with a decrease of ≥ 2.2 log₂-fold. These genes were also down-regulated in response to POE at 20% and 50% of the MIC.

Four down-regulated genes encoded in close proximity to each other on plasmid 1 were identified (RW109_00464-00469; Table 29) and 3 had been assigned to KEGG pathways. RW109_00469 was characterised in the glyoxylate and dicarboxylate metabolism pathway and RW109_00464-00465 were involved in glycine, serine and threonine metabolism. The one gene down-regulated on plasmid 2 was annotated as a hypothetical protein (RW109_07290) and had decreased expression in response to POE at 50% of the MIC.

Table 28. Top 10 up-regulated genes in response to BIT and POE in combination at 20% of the MIC

Transcript ID	Prokka putative gene function	Cog category	Log2 fold change	Adjusted P-value
Plasmid 1				
RW109_00266	Hypothetical protein	Unknown	2.3	0.0000241
RW109_00018	Glutathione S-transferase GST-6.0	O	2.2	0.000127
RW109_00576	Hypothetical protein	Unknown	2.2	2.23E-12
RW109_00324	Recombination-associated protein RdgC	L	2.1	4.51E-11
RW109_00180	Hypothetical protein	Unknown	1.9	0.005873
RW109_00375	Hypothetical protein	S	1.9	0.017366
RW109_00516	Hypothetical protein	Unknown	1.9	0.001297
RW109_00194	Hypothetical protein	Unknown	1.8	0.027186
RW109_00049	Hypothetical protein	S	1.7	0.003275
RW109_00367	Arc-like DNA binding domain protein	Unknown	1.7	0.000387
Main chromosome				
RW109_03497	Anthranilate 1,2-dioxygenase large subunit	Q	6.1	2.25E-20
RW109_02298	Efflux pump membrane transporter MexQ	V	5.3	5.05E-17
RW109_02297	Membrane fusion protein MexP	V	5.2	1.54E-13
RW109_03496	Anthranilate 1,2-dioxygenase small subunit	Q	4.9	7.83E-09
RW109_04047	Citrate transporter	E, G	4.5	9.93E-09
RW109_02242	Copper chaperone CopZ	P	4.5	9.54E-26
RW109_03936	Putative sialic acid transporter	G	4.4	2.97E-08
RW109_03502	Catechol 1,2-dioxygenase	Q	4.3	0.00000209
RW109_04037	Methylmalonyl-CoA carboxyltransferase 12S subunit	I	4.2	9.88E-30
RW109_01137	Hypothetical protein	Unknown	4.2	9.46E-10
Plasmid 2				
RW109_07266	Transposase DDE domain protein	L	2.1	1.33E-15
RW109_07345	Hypothetical protein	Unknown	2.0	0.024347
RW109_07352	Hypothetical protein	Unknown	2.0	0.005691
RW109_07250	Hypothetical protein	Unknown	2.0	9.41E-18
RW109_07342	Transposase	L	1.7	0.009392
RW109_07282	Hypothetical protein	S	1.7	1.60E-17
RW109_07376	DNA-binding protein HRL53	L	1.7	1.70E-13
RW109_07297	Hypothetical protein	S	1.7	0.000398
RW109_07382	Hypothetical protein	Unknown	1.6	0.00000133
RW109_07256	Hypothetical protein	Unknown	1.6	0.00000144
RW109_07312	Hypothetical protein	Unknown	1.5	0.027199

Footnotes = Top 10 genes from the main chromosome and plasmid 1 are shown, as well as the total 11 genes down regulated from plasmid 2

Table 29. The top ten down-regulated genes in response to BIT and POE in combination at 20% of the MIC

Transcript ID	Prokka putative gene function	Cog category	Log2 fold change	Adjusted P-value
Plasmid 1				
RW109_00464	Sarcosine oxidase, gamma subunit family~	E	-2.9	0.010672
RW109_00461	Cytochrome c~	C	-2.2	0.00854
RW109_00469	Formyltetrahydrofolate deformylase~	F	-1.9	0.00000628
RW109_00465	Aminomethyltransferase~	E	-1.8	0.017953
RW109_00423	Hypothetical protein	Unknown	-1.5	0.046256
Main chromosome				
RW109_01518	Pesticin receptor precursor	P	-6.4	3.44E-36
RW109_05811	Hypothetical protein*	S	-6.3	3.40E-45
RW109_03585	Putative RNA polymerase sigma factor FecI	K	-5.8	1.82E-31
RW109_02394	Hemophore HasA	S	-5.7	3.52E-14
RW109_03624	L-ornithine 5-monooxygenase	Q	-5.5	4.82E-41
RW109_05808	Superoxide dismutase [Mn/Fe]*	P	-5.3	1.53E-48
RW109_05809	Hypothetical protein*	S	-5.2	1.09E-28
RW109_03982	Decarbamoylnovobiocin carbamoyltransferase	O	-4.9	3.70E-31
RW109_06145	Hypothetical protein	Unknown	-4.9	6.94E-21
RW109_05810	Fumarate hydratase class II*	C	-4.8	1.13E-84
Plasmid 2				
RW109_07290	Hypothetical protein	Unknown	-1.5	0.00024

Footnotes = Top 10 genes from the main chromosome are shown as well as the total 5 genes from plasmid 1 and the one gene from plasmid 2. The* indicates those genes within close proximity on the main chromosome. The ~ illustrates plasmid 1 genes in close proximity.

5.2.7.5 Key differentially regulated genes identified following exposure of RW109 to a laundry detergent with and without the addition of the preservative BIT

5.2.7.5.1 Gene expression analysis following exposure to laundry detergent at 1:100 dilution with no preservative added

Gene expression analysis of the RW109 strain when exposed to laundry detergent with no preservative added (LD-only, test condition 6), identified 1,264 significant DEGs (Table 21). On the main chromosome, 581 were up-regulated and 627 genes were down-regulated and on plasmid 1, 7 genes were up-regulated and 32 genes were down-regulated. Six genes had increased expression on plasmid 2 and 11 were down-regulated.

A gene encoding a putative sialic acid transporter RW109_03936, had the greatest fold increase within the top ten up-regulated genes on the main chromosome (Table 30). Four other genes within close proximity to RW109_03936 were also identified in Table 30. These genes encoded a hypothetical protein (RW109_03934), a porin-like protein NicP precursor (RW109_03937), a LamB/YcsF family protein (RW109_03938) and a kinase-A inhibitor (RW109_03939). The putative KipI antagonist (RW109_03940) and a HTH-type transcriptional regulator GltR (RW109_03935), were also up-regulated in response to LD-only by 4.4 and 1.5 log₂-fold respectively. The amino acid sequences of these 7 genes were searched using DIAMOND BLASTP on the *Pseudomonas* database. The genes RW109_03936-03940 had high sequence homology to four genes within a predicted PA14 membrane transport operon (mean BLASTP scores: percentage ID= 99.2%, alignment length= 329, E-value= 0.0 and BIT Score= 650). Another gene of interest, up-regulated in response to the laundry detergent was RW109_05501 (Table 30), which encoded a metallo-beta-lactamase superfamily protein and was not observed in conditions 2-5.

Outside of the top 10 up-regulated genes, a notable observation was the increased expression of the syntenic genes RW109_01012 – 01017 ranging between 1.8 – 3.9 log₂-fold in response to LD-only. The genes had high homology to those within the PA14 pilin biogenesis operons *pilGHI* and *pilJK-chpA-E* when analysed using the *Pseudomonas* database (mean BLASTP scores: percentage ID= 99.6%, alignment length= 682, E-value= 0.0 and BIT Score= 1171). These genes encode chemotaxis-like chemosensory signal transduction systems for regulation of type IV pili production and twitching motility in *P. aeruginosa* (Mattick, 2002).

Out of the genes up-regulated in response to LD-only, 12% were characterised within the energy metabolism KEGG pathways (Table 20). Notable genes RW109_07219-07226 were associated with oxidative phosphorylation and had increased expression between 1.5 – 2.6 log₂ fold; the genes were part of the *atpCDGAHFEBI* operon. The genes RW109_00877-00880 functioned in sulphur metabolism and increased their expression between 3.0 – 4.2 log₂-fold. RW109_00880 encoded for a sulphate-binding protein, and the RW109_00877-00879 genes were associated with a sulphate transport operon *cysAWT*. A large cluster of genes (RW109_01333-01370), were up-regulated between 1.9 – 3.7 log₂-fold in response to LD-only. These genes encoded for 32 ribosomal proteins, 3 DNA-directed RNA polymerases, 2 elongation factors and 1 preprotein translocase.

CHAPTER 5 – RNA-Seq gene expression analysis of RW109 in response to preservatives and product formulations

Genes up-regulated in response to LD-only were also represented by membrane transport pathways (10%) (Table 20). RW109_02634-02637 encoded for an ABC sugar transporter operon and RW109_02635 and RW109_02636 were within the top ten up-regulated main chromosome genes (Table 30). RW109_02634 encoded for a glycerol-3-phosphate transport system permease protein and was up-regulated by 4.1 log₂-fold. RW109_02637 was annotated as a porin B precursor and had a 4.6 log₂-fold increase in expression. The genes encoding the MexP and MexQ proteins of the MexPQ-OpmE efflux pump were up-regulated by 4.2 and 4.5 log₂-fold respectively in response to LD-only; these genes had previously been identified with test conditions 3-5.

The genes within the *gnyDBHAL* operon (RW109_04036-04040) and the KdgR transcriptional regulator, encoding gene (RW109_04041) had log₂-fold increases of ≥ 2.0 in response to LD-only. These genes were also up-regulated with test conditions 2-5. Seven plasmid 1 genes were up-regulated in response to LD-only and a hypothetical protein encoding gene RW109_00524 had the greatest increase with 3.0 log₂-fold (Table 30). A hypothetical protein encoding gene RW108_07297 also had the greatest fold increase out of the 6 genes up-regulated on plasmid 2.

Within the top ten down-regulated chromosome genes identified in response to LD-only, RW109_05642 annotated as an Flp/Fap pilin component had the greatest fold decrease (Table 31). RW109_05181 was the second-most downregulated chromosomal gene (Table 31) and was associated with the quorum sensing operon *phnAB*. The RW109_05182 gene associated with the same quorum sensing operon *phnAB* was also down-regulated in response to LD-only by 4.2 log₂-fold. The genes, which form the pqsABCDE operon RW109_05183-05187, also had a decreased expression with this test condition of between 4.0 – 5.3 log₂-fold (Table 31). Genes RW109_06227 and RW109_06228 (Table 31) encoded two spore Coat Protein U domain proteins and were both highly down-regulated by 6.0 log₂-fold. An outer membrane channel protein, LapE (RW109_04177) and a *cna* B-type domain protein (RW109_04178) were identified in the top ten down-regulated DEGs. Two other genes in close proximity were also down-regulated and encoded a type I secretion system membrane fusion protein LapC (RW109_04175, -4.3 log₂-fold) and a toxin RTX-I translocation ATP-binding protein LapB (RW109_04176, -4.5 log₂-fold). RW109_04176-04177 were characterised by KEGG as a *lapBCE* adhesin protein transport system operon. Two genes of interest with decreased expression with LD-only encoded for two bacterial type II secretion system proteins (RW109_05636, -3.6 log₂-fold and RW109_05637, -2.4 log₂-fold). Out of the genes down-regulated in response to LD-only, 20% were associated with carbohydrate metabolism (Table 20). Notable genes RW109_02383-02385 within this KEGG pathway formed a pyruvate dehydrogenase complex and had a decreased expression of between 2.5 – 3.6 log₂-fold.

The gene with the greatest fold decrease on plasmid 1 in response to LD-only was annotated as a hypothetical protein (RW109_00135) (Table 31). Other notable genes down-regulated on the plasmid included RW109_00469 and RW109_00464 which also had decreased expression in response to BIT and POE in combination (Section 5.2.7.4.1). A plasmid 1 aerotaxis receptor annotated gene (RW109_00009) had the third greatest log₂-fold decrease in response to LD-only and was part of a KEGG bacterial

CHAPTER 5 – RNA-Seq gene expression analysis of RW109 in response to preservatives and product formulations

chemotaxis two-component system. Two other notable down-regulated plasmid 1 genes encoded transposition TniQ and TniB proteins (RW109_00006, -1.9 log₂-fold and RW109_00007, -2.1 log₂-fold). On plasmid 2, a down-regulated gene of interest encoded for an antitoxin DinJ protein (RW109_07361); the majority of plasmid 2 genes with decreased expression were annotated as hypothetical proteins (Table 31).

Table 30. The top ten up-regulated genes in response to LD with no preservative added

Transcript ID	Prokka putative gene function	Cog category	Log2 fold change	Adjusted P-value
Plasmid 1				
RW109_00524	Hypothetical protein	Unknown	3.0	0.000877
RW109_00459	Hypothetical protein	S	2.6	0.00000509
RW109_00174	N-acetylglucosaminyl-diphospho-decaprenol L-rhamnosyltransferase	M	1.9	1.15E-12
RW109_00057	Site-specific tyrosine recombinase XerC	L	1.8	2.73E-08
RW109_00460	Imelysin	S	1.6	1.77E-09
RW109_00225	Hypothetical protein	Unknown	1.5	0.00000256
RW109_00198	Anaerobic sulfatase-maturing enzyme	S	1.5	0.002315
Main chromosome				
RW109_03936	Putative sialic acid transporter*	G	6.9	1.20E-21
RW109_03934	Hypothetical protein*	S	5.3	9.89E-10
RW109_05501	Metallo-beta-lactamase superfamily protein	Q	5.1	4.70E-11
RW109_03937	Porin-like protein NicP precursor*	M	5.0	1.12E-12
RW109_02635	L-arabinose transport system permease protein AraQ	P	5.0	1.44E-12
RW109_03939	Kinase A inhibitor*	E	4.9	3.11E-08
RW109_03690	High-affinity gluconate transporter	E, G	4.8	5.26E-08
RW109_06963	Putative FAD-linked oxidoreductase	C	4.7	5.82E-09
RW109_03938	LamB/YcsF family protein*	E	4.6	0.00000202
RW109_02636	sn-glycerol-3-phosphate import ATP-binding protein UgpC	G	4.6	1.73E-24
Plasmid 2				
RW109_07297	Hypothetical protein	S	2.0	0.0000024
RW109_07302	Hypothetical protein	Unknown	1.9	0.000268
RW109_07275	Outer membrane lipoprotein BfpB precursor	U	1.8	9.94E-10
RW109_07312	Hypothetical protein	Unknown	1.7	0.005522
RW109_07299	Hypothetical protein	Unknown	1.6	0.004189
RW109_07267	O-acetyltransferase OatA	I	1.5	0.0000075

Footnotes = Top 10 genes from the main chromosome as well as the 7 genes from plasmid 1 and the 6 genes from plasmid 2. The * indicates genes within close proximity on the RW109 genome.

Table 31. The top ten down-regulated genes in response to LD with no preservative added

Transcript ID	Prokka putative gene function	Cog category	Log2 fold change	Adjusted P-value
Plasmid 1				
RW109_00135	Hypothetical protein	S	-3.7	3.52E-34
RW109_00469	Formyltetrahydrofolate deformylase	F	-3.6	3.08E-22
RW109_00009	Aerotaxis receptor	T	-3.4	4.13E-34
RW109_00303	Leucine-responsive regulatory protein	K	-3.3	1.34E-21
RW109_00081	Hypothetical protein	Unknown	-3.2	0.017561
RW109_00269	Hypothetical protein	S	-2.9	6.38E-13
RW109_00231	Hypothetical protein	Unknown	-2.8	0.00016
RW109_00464	Sarcosine oxidase, gamma subunit family	E	-2.7	0.008721
RW109_00446	Aminomethyltransferase	E	-2.7	0.002159
RW109_00566	Chromosome-partitioning protein Spo0J	K	-2.6	0.015754
Main chromosome				
RW109_05642	Flp/Fap pilin component	Unknown	-6.9	3.96E-38
RW109_05181	Anthranilate synthase component 2, pyocyanine specific	E	-6.4	4.68E-13
RW109_04178	Cna protein B-type domain protein	O	-6.1	2.22E-163
RW109_06227	Spore Coat Protein U domain protein	S	-6.0	2.78E-45
RW109_06228	Spore Coat Protein U domain protein	S	-6.0	4.32E-11
RW109_04177	outer membrane channel protein	M	-5.9	4.22E-48
RW109_05641	SAF domain protein	U	-5.8	2.36E-15
RW109_03434	PA-I galactophilic lectin	S	-5.7	4.60E-32
RW109_00725	Aegerolysin	Unknown	-5.5	1.29E-126
RW109_01791	Hypothetical protein	S	-5.4	1.23E-197
Plasmid 2				
RW109_07281	Hypothetical protein	S	-2.3	0.028305
RW109_07290	Hypothetical protein	Unknown	-2.3	0.000000173
RW109_07361	Antitoxin DinJ	L	-1.8	7.81E-10
RW109_07316	Hypothetical protein	Unknown	-1.8	9.4E-13
RW109_07384	Hypothetical protein	Unknown	-1.7	0.0000139
RW109_07253	Hypothetical protein	Unknown	-1.6	2.69E-08
RW109_07238	Transglycosylase SLT domain protein	M	-1.6	0.02264
RW109_07257	Hypothetical protein	Unknown	-1.5	0.00149
RW109_07280	Hypothetical protein	S	-1.5	0.035025
RW109_07268	Hypothetical protein	U	-1.5	2.97E-08
RW109_07293	Initiator Replication protein	L	-1.5	0.000000098

Footnotes = Top 10 genes from the main chromosome and plasmid 1 are shown as well as the total 11 genes from plasmid 2

5.2.7.5.2 Gene expression analysis following exposure to a combination of laundry detergent at 1:100 dilution and BIT at 20% of the MIC

Gene expression analysis of the RW109 strain when exposed to a combination of laundry detergent at a 1:100 dilution and BIT at 20% of the MIC (test condition 7) identified 1,487 significant DEGs (Table 21). On the main chromosome, 678 were up-regulated and 724 genes were down-regulated. On plasmid 1, 23 genes were up-regulated and 35 genes were down-regulated and on plasmid 2, 13 genes had increased expression and 14 were down-regulated.

Due to the large number of DEGs identified with this test condition, it was decided to illustrate those with increased and decreased expression on circular maps representing the main chromosome and two plasmids (Figure 46, Figure 47 and Figure 48). Both the up-regulated and down-regulated genes were dispersed throughout the main chromosome. A large grouping of the up-regulated RW109_01333-01370 genes were observed on main chromosome which encoded for ribosomal proteins (Figure 46) and were also identified as up-regulated during exposure to laundry detergent only (Section 5.2.7.5.1). An increase in ribosomal protein gene expression suggests the repair and replacement of ribosomes is required by the RW109 strain when exposed to LD with and without BIT. On plasmid 1, greater co-localisation of up-regulated genes was observed, whereas the down-regulated genes were spread throughout the plasmid (Figure 47). Interestingly, a similar split and co-localisation of up-regulated genes on plasmid 2 was also observed (Figure 48). The down-regulated genes were also more dispersed throughout plasmid 2 when compared to the up-regulated genes (Figure 48).

On the main chromosome, the most up-regulated gene in response to LD with BIT was the putative sialic acid transporter encoding gene RW109_03936 (Table 32); this gene was also up-regulated in response to multiple test conditions (2, 4, 5 and 6) as indicated by Figure 49. RW109_03936 was characterised within a predicted membrane transport operon, which included RW109_03937-03940 and a transcriptional regulator GltR, encoded by RW109_03935 (Figure 49). All genes associated with this operon were up-regulated in response to LD with BIT, and RW109_03937 and RW109_03939 were within the top 10 up-regulated main chromosome DEGs (Table 32).

The RW109_05501 gene encoding a metallo-beta-lactamase superfamily protein had the second highest log₂-fold change in response to LD with BIT (Table 32); this gene was also up-regulated with LD-only. The MexPQ-OpmE RND efflux system encoding genes RW109_02297-02299 were up-regulated with LD with BIT, and the gene RW109_02298 was the third gene within the top ten up-regulated main chromosome DEGs (Table 32). The RW109_04035-04040 *gnyRDBHAL* operon and the RW109_04041 gene encoding the KdgR transcriptional regulator, previously identified with conditions 2-6, were also up-regulated with LD and BIT with an expression increase of ≥ 1.9 log₂-fold.

A notable observation was the up-regulation of the genes RW109_01012 – 01017, that ranged between 1.7 – 3.7 log₂-fold. These genes were associated with the pilin biogenesis operons *pilGHI* and *pilJK-chpA-E* (Section 5.2.7.5.1) and increased expression with LD-only (Figure 50). All the genes of the *pilGHI* operon

CHAPTER 5 – RNA-Seq gene expression analysis of RW109 in response to preservatives and product formulations

were identified as up-regulated in response to test condition 5 as indicated in Figure 50. In response to LD with BIT, 4 up-regulated genes were identified which were not observed together in any other test condition. These included RW109_04463 (*pscN*, 2.4 log₂-fold), RW109_04466 (*pscQ*, 2.9 log₂-fold), RW109_04467 (*pscR* 3.0 log₂-fold) and RW109_04470 (*pscU*, 3.0 log₂-fold). These genes formed part of the *pscNOPQRTU* operon of a KEGG type III section system. Four genes associated with a glucose/mannose membrane ABC transport system (RW109_02633-02636) were up-regulated in response to LD and BIT with a ≥ 3.4 log₂-fold increase. These genes were also all up-regulated together in response to BIT at 50% of the MIC.

On plasmid 1, within the top ten up-regulated genes, 7 encoded for hypothetical proteins which was also observed with plasmid 2 (Table 32). A gene of interest on plasmid 1 was RW109_00174, which encoded an N-acetylglucosaminyl-diphospho-decaprenol L-rhamnosyltransferase WbbL, and functioned in cell wall peptidoglycan biosynthesis (Figure 47). On plasmid 2 a notable gene, RW109_07269, encoded for a twitching mobility protein as indicated in Figure 48.

On the main chromosome, a gene encoding for decarbamoylnovobiocin carbamoyltransferase (RW109_03982) had the greatest fold decrease in response to LD with BIT (Table 33). This gene was previously identified as down-regulated in response to POE at 20% and 50% of the MIC. Also within the top ten down-regulated genes were RW109_05186 and RW109_05187, which formed part of the RW109_05183-05187 quorum sensing *pqsABCDE* operon. The other genes associated with this operon and the RW109_05181-05182 *phnAB* operon all had decreased expression in response to LD with BIT. The genes encoding a Flp/Fap pilin component (RW109_05642), an outer membrane channel protein (RW109_04177) and a Cna B-type domain protein (RW109_04178) were identified in Table 33 and also had a decreased expression with LD-only. Two down-regulated DEGs previously identified with LD-only encoded for bacterial type II secretion system proteins (RW109_05636, -4.1 log₂-fold and RW109_05367, -2.2 log₂-fold); these genes were only observed after exposure to LD with and without BIT. The genes encoding for a KEGG iron complex transport system were only observed in response to LD with BIT and these included RW109_01515 and RW109_01516 which had 3.5 and 3.2 log₂-fold decreases respectively. A KEGG glycine betaine/proline membrane transport system was also only down-regulated with this test exposure condition and was encoded for by RW109_06690 (-2.1 log₂-fold) and RW109_06691 (-1.8 log₂-fold).

Hypothetical protein encoded genes were repressed on both plasmid 1 and 2, with 7 being within the top 10 down-regulated genes on each replicon (Table 33). On plasmid 1, the greatest fold decrease was observed with the hypothetical protein encoding gene RW109_00135 (Table 33). The RW109_00464 and RW109_00469 plasmid 2 genes (Table 33) were also identified with BIT-POE and LD-only test conditions and encoded for a sarcosine oxidase and formyltetrahydrofolate deformylase respectively. The aerotaxis receptor encoding gene (RW109_00009) which had decreased expression with LD-only, was also observed within the top ten down-regulated plasmid 1 genes, in response to LD with BIT (Table 33). A down-regulated gene of interest on plasmid 2 encoded for an antitoxin DinJ protein (RW109_07361)

CHAPTER 5 – RNA-Seq gene expression analysis of RW109 in response to preservatives and product formulations

which had previously been identified as up-regulated in response to POE at 20% of the MIC (Section 5.2.7.3.3). A large number of genes with decreased expression were annotated as hypothetical proteins (Table 33) and this was true during exposure to LD-only.

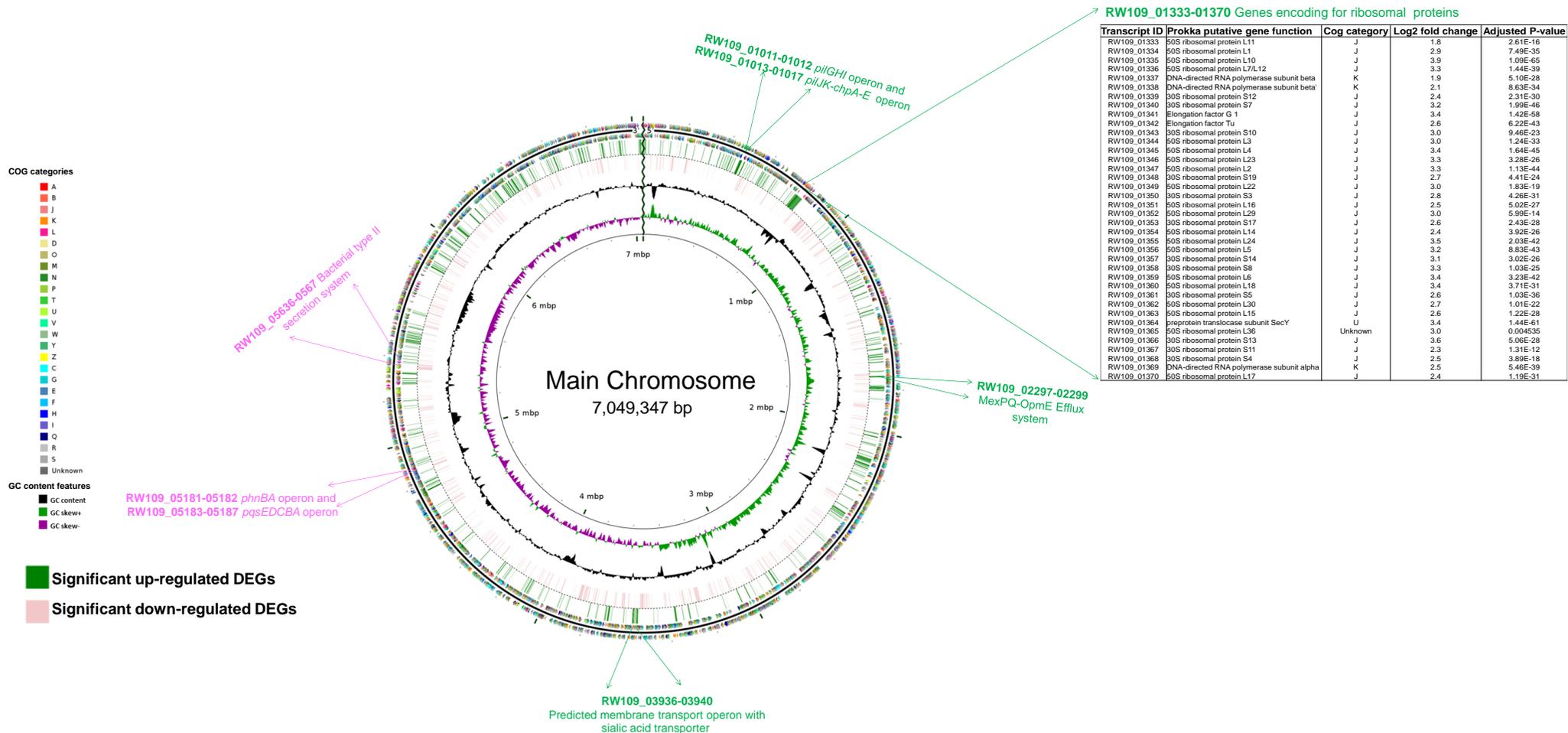


Figure 46. The differentially regulated genes identified on the main chromosome in response to laundry detergent in combination with BIT

Starting from the outermost circle for each map and moving inwards the following tracks were identified: (1) predicted CDS on forward strand coloured according to COG categories, (2) predicted CDS on reverse strand coloured according to COG categories, (3) significantly up-regulated DEGs coloured green, (4) significantly down-regulated DEGs coloured light pink, (5) GC content (black), (6) positive and negative GC skew (green and purple, respectively) and (7) genome region by Mbp. The size of the main chromosome was not drawn to scale. Notable up-regulated genes were labelled in green and down-regulated genes of interest were labelled in pink.

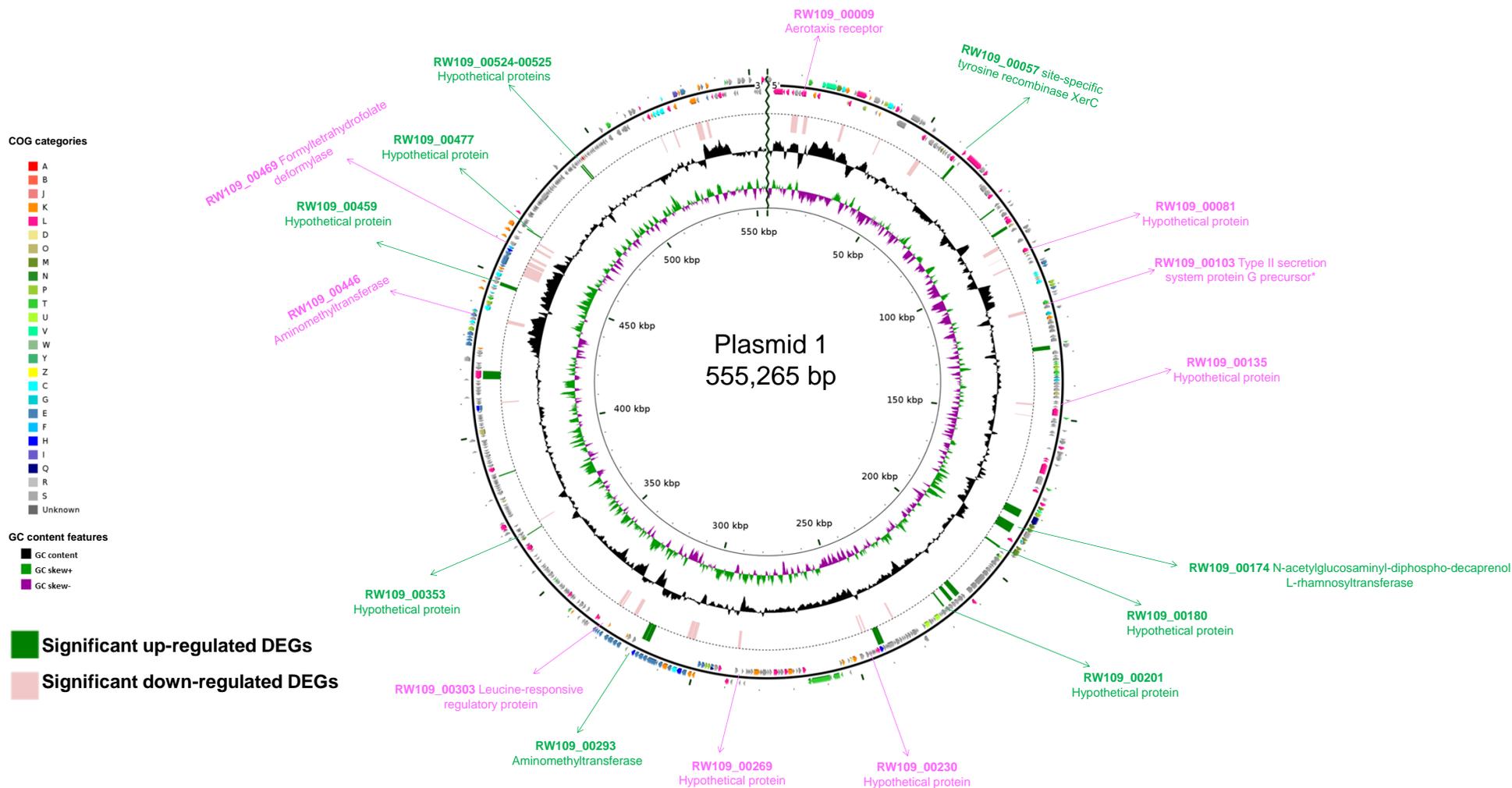


Figure 47. The differentially regulated genes identified on plasmid 1 in response to laundry detergent and BIT in combination

Starting from the outermost circle for each map and moving inwards the following tracks were identified: (1) predicted CDS on forward strand coloured according to COG categories, (2) predicted CDS on reverse strand coloured according to COG categories, (3) significantly up-regulated DEGs coloured green, (4) significantly down-regulated DEGs coloured light pink, (5) GC content (black), (6) positive and negative GC skew (green and purple, respectively) and (7) genome region by kbp. The size of plasmid 1 was not drawn to scale. The top ten up-regulated genes were labelled in green and the top ten down-regulated were labelled in pink (the * on RW109_00103, indicates a gene of interest not included within the top 10 down-regulated).

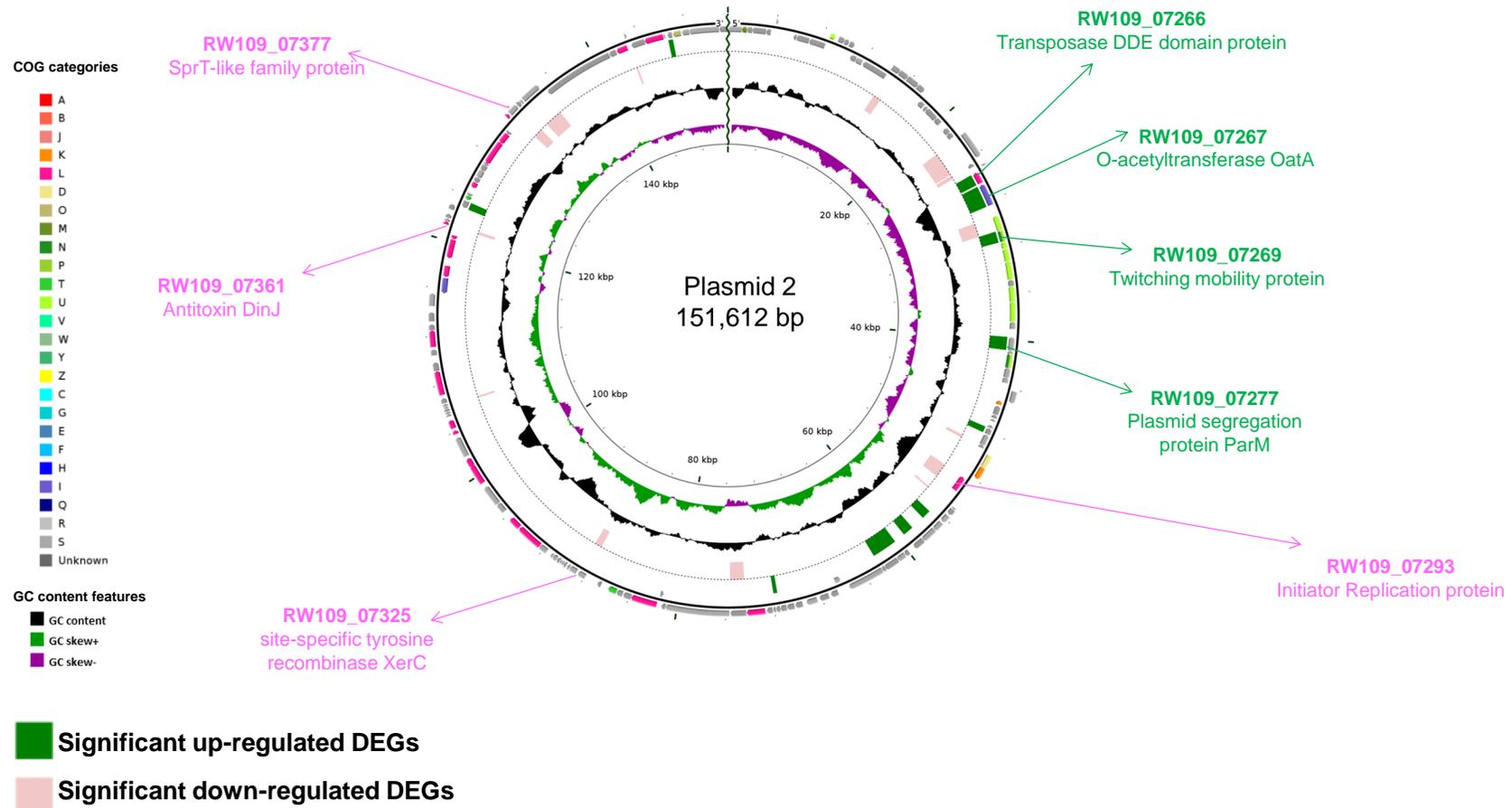


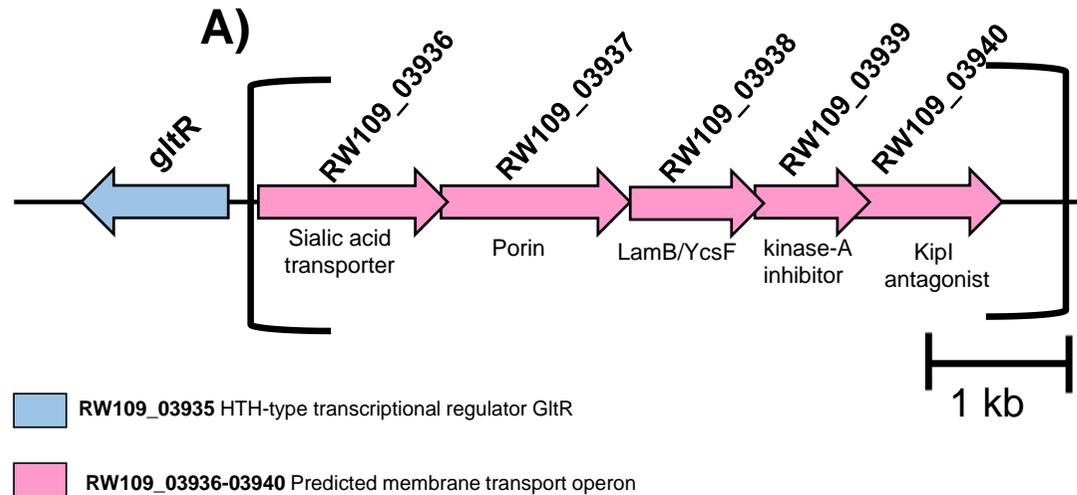
Figure 48. The differentially regulated genes identified on plasmid 2 in response to laundry detergent and BIT in combination

Starting from the outermost circle for each map and moving inwards the following tracks were identified: (1) predicted CDS on forward strand coloured according to COG categories, (2) predicted CDS on reverse strand coloured according to COG categories, (3) significantly up-regulated DEGs coloured green, (4) significantly down-regulated DEGs coloured light pink, (5) GC content (black), (6) positive and negative GC skew (green and purple, respectively) and (7) genome region by kbp. The size of plasmid 2 was not drawn to scale. All the up-regulated genes not encoding hypothetical proteins were labelled in green and all the down-regulated genes without a hypothetical protein annotation were labelled in pink.

Table 32. The top ten up-regulated genes in response to laundry detergent in combination with BIT at 20% of the MIC

Transcript ID	Prokka putative gene function	Cog category	Log2 fold change	Adjusted P-value
Plasmid 1				
RW109_00524	Hypothetical protein	Unknown	3.4	0.000187
RW109_00174	N-acetylglucosaminyl-diphospho-decaprenol L-rhamnosyltransferase	M	2.3	4.26E-21
RW109_00477	Hypothetical protein	Unknown	2.3	0.028076
RW109_00057	Site-specific tyrosine recombinase XerC	L	2.2	7.27E-14
RW109_00459	Hypothetical protein	S	2.2	0.000157
RW109_00353	Hypothetical protein	Unknown	2.0	0.030427
RW109_00201	Hypothetical protein	Unknown	2.0	0.035608
RW109_00293	Aminomethyltransferase	E	1.9	0.003284
RW109_00180	Hypothetical protein	Unknown	1.7	0.008859
RW109_00525	Hypothetical protein	Unknown	1.7	0.0000212
Main chromosome				
RW109_03936	Putative sialic acid transporter	G	6.8	4.72E-20
RW109_05501	Metallo-beta-lactamase superfamily protein	Q	5.9	3.78E-14
RW109_02298	Efflux pump membrane transporter MexQ	V	5.5	1.19E-14
RW109_03934	Hypothetical protein	S	5.3	1.58E-09
RW109_03690	High-affinity gluconate transporter	E, G	5.2	7.79E-09
RW109_06963	Putative FAD-linked oxidoreductase	C	5.1	2.65E-10
RW109_03937	Porin-like protein NicP precursor	M	5.0	8.07E-12
RW109_02635	L-arabinose transport system permease protein AraQ	P	5.0	1.14E-10
RW109_03939	Kinase A inhibitor	E	4.9	8.43E-08
RW109_02640	Glucose-6-phosphate 1-dehydrogenase	G	4.9	1.33E-53
Plasmid 2				
RW109_07303	Hypothetical protein	Unknown	2.9	0.014245
RW109_07302	Hypothetical protein	Unknown	2.3	0.00000496
RW109_07312	Hypothetical protein	Unknown	2.0	0.0000876
RW109_07267	O-acetyltransferase OatA	I	1.9	0.000000281
RW109_07299	Hypothetical protein	Unknown	1.9	0.0000628
RW109_07288	Hypothetical protein	Unknown	1.7	2.08E-11
RW109_07297	Hypothetical protein	S	1.6	0.0000369
RW109_07277	Plasmid segregation protein ParM	S	1.5	0.042658
RW109_07388	Hypothetical protein	Unknown	1.5	0.000000259
RW109_07269	Twitching mobility protein	N, U	1.5	0.001901

Footnotes = Top 10 genes from the main chromosome and the two plasmids are shown.

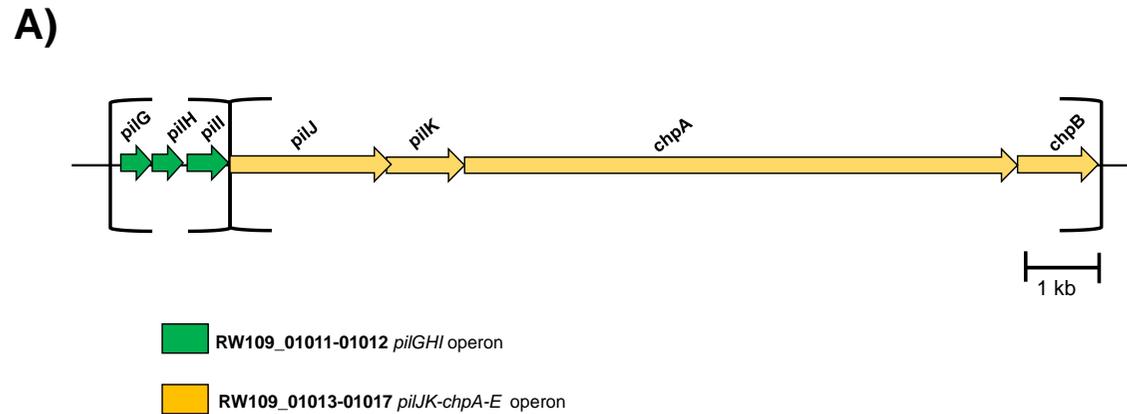


B)

Gene ID	Test exposure conditions									
	BIT at 50% of the MIC (2)		POE at 50% of the MIC (4)		BIT and POE in combination (5)		LD only (6)		LD and BIT at 20% of the MIC (7)	
	Fold change	Adjusted p-value	Fold change	Adjusted p-value	Fold change	Adjusted p-value	Fold change	Adjusted p-value	Fold change	Adjusted p-value
RW109_03935	NA	NA	NA	NA	NA	NA	1.5	0.034038	1.8	0.005044
RW109_03936	3.4	7.10E-10	6.0	7.97E-15	4.4	2.97E-08	6.9	1.20E-21	6.8	4.72E-20
RW109_03937	NA	NA	3.0	0.002298	2.8	0.000815	5.0	1.12E-12	5.0	8.07E-12
RW109_03938	NA	NA	NA	NA	NA	NA	4.6	2.02E-06	4.6	3.55E-06
RW109_03939	NA	NA	NA	NA	2.6	0.014728	4.9	3.11E-08	4.9	8.43E-08
RW109_03940	NA	NA	NA	NA	2.3	0.04825	4.4	1.26E-05	4.6	3.05E-06

Figure 49. Gene expression analysis the RW109_03936-03940 predicted membrane transport operon and a transcriptional regulator GltR when exposed to test conditions 2-7

Panel A shows the transcriptional regulator GltR in blue (RW109_03935) and the predicted membrane transport operon in pink (RW109_03936-03940). The log₂-fold changes and adjusted p-values for these genes are shown in panel B for conditions 2-7. All the genes of the predicted operon were up-regulated with conditions 6 and 7 as well as the gltR annotated gene. The RW109_03936 gene encoding for a sialic acid transporter had increased expression in conditions 2, 4, 5, 6 and 7. In condition 4 and 5 the porin annotated gene RW109_03937 was up-regulated and the genes RW109_03939-03940 also had increased expression in condition 5.



B)

Gene ID	Test exposure conditions					
	BIT and POE in combination (5)		LD only (6)		LD and BIT at 20% of the MIC (7)	
	Fold change	Adjusted p-value	Fold change	Adjusted p-value	Fold change	Adjusted p-value
RW109_01011	1.8	2.03E-19	NA	NA	NA	NA
RW109_01012	2.1	1.39E-11	2.3	1.95E-18	1.9	1.34E-10
RW109_01013	2.8	1.96E-05	3.9	4.08E-10	3.7	9.08E-10
RW109_01014	NA	NA	2.8	1.60E-34	2.7	7.43E-38
RW109_01015	NA	NA	1.8	1.31E-05	1.7	3.78E-06
RW109_01016	NA	NA	2.2	1.64E-16	NA	NA
RW109_01017	NA	NA	2.5	4.49E-05	2.3	1.25E-04

Figure 50. Gene expression analysis of the *pilGHI* and *pilJK-chpA-E* operons when exposed to test conditions 5, 6 and 7

Panel A shows the *pilGHI* operon in green (RW109_01011-01012) and the *pilJK-chpA-E* operon in yellow (RW109_01013-01017). The log₂-fold changes and adjusted p-values for these genes are shown in panel B. All three genes of the *pilGHI* operon were up-regulated after exposure to condition 5 and the *pilH* and *pilI* genes had increased expression with 6 and 7. All genes in the *pilJK-chpA-E* operon were up-regulated in response to conditions 6 and 7.

Table 33. The top ten genes down-regulated in response to laundry detergent in combination with BIT at 20% of the MIC

Transcript ID	Prokka putative gene function	Cog category	Log2 fold change	Adjusted P-value
Plasmid 1				
RW109_00135	Hypothetical protein	S	-3.9	1.88E-52
RW109_00469	Formyltetrahydrofolate deformylase	F	-3.7	2.48E-22
RW109_00269	Hypothetical protein	S	-3.6	6.34E-25
RW109_00303	Leucine-responsive regulatory protein	K	-3.6	1.15E-26
RW109_00009	Aerotaxis receptor	T	-3.6	2.18E-50
RW109_00081	Hypothetical protein	Unknown	-2.9	0.033649
RW109_00230	Hypothetical protein	Unknown	-2.7	0.006339
RW109_00446	Aminomethyltransferase	E	-2.7	0.001958
RW109_00085	Hypothetical protein	Unknown	-2.6	9.60E-12
RW109_00464	Sarcosine oxidase, gamma subunit family	E	-2.4	0.009834
Main chromosome				
RW109_03982	Decarbamoylnovobiocin carbamoyltransferase	O	-7.6	3.32E-42
RW109_06227	Spore Coat Protein U domain protein	S	-6.9	5.42E-42
RW109_03434	PA-I galactophilic lectin	S	-6.8	2.63E-29
RW109_05642	Flp/Fap pilin component	Unknown	-6.7	4.01E-41
RW109_05186	3-oxoacyl-(acyl carrier protein) synthase III	Unknown	-6.6	1.63E-29
RW109_00725	Aegerolysin	Unknown	-6.5	1.43E-178
RW109_05187	Anthranilate--CoA ligase	I	-6.3	9.72E-27
RW109_04177	outer membrane channel protein	M	-6.3	6.99E-53
RW109_05811	Hypothetical protein	S	-6.2	3.66E-71
RW109_04178	Cna protein B-type domain protein	O	-6.2	1.00E-171
Plasmid 2				
RW109_07290	Hypothetical protein	Unknown	-2.8	4.97E-15
RW109_07253	Hypothetical protein	Unknown	-2.8	1.80E-25
RW109_07384	Hypothetical protein	Unknown	-2.4	2.41E-19
RW109_07263	Hypothetical protein	S	-1.8	1.29E-14
RW109_07316	Hypothetical protein	Unknown	-1.8	1.04E-14
RW109_07264	Hypothetical protein	Unknown	-1.7	1.85E-11
RW109_07347	Hypothetical protein	Unknown	-1.6	0.015387
RW109_07380	Hypothetical protein	S	-1.6	7.71E-13
RW109_07361	Antitoxin DinJ	L	-1.6	7.52E-09
RW109_07293	Initiator Replication protein	L	-1.5	3.90E-10

Footnotes = Top 10 genes from the main chromosome and plasmid 1 are shown as well as the total 11 down-regulated genes from plasmid 2

5.2.7.6 Gene expression analysis of RW109 antimicrobial predicted genes

The output from ABRicate and BacMet (Chapter 3, Sections 3.2.3.3 and 3.2.3.4), were used to investigate the expression of known and predicted antimicrobial resistance genes in response to the industry relevant conditions. The analysis revealed 34 genes differentially regulated in at least one of the test exposure conditions (not including BIT at 20% of the MIC due to the minimal genes regulated) (Figure 51 and Table 34). The highest number of differentially regulated antimicrobial DEGs was identified after exposure to LD in combination with BIT at 20% of the MIC, with 18 up-regulated and 7 down-regulated (Table 34). The least number of antimicrobial DEGs were identified with BIT at 50%, where RW109_06416 encoding for magnesium-transport protein MgtA, was up-regulated. The *mgtA* gene also had increased expression after exposure to conditions 6, LD-only, and 7, LD with BIT, whereas it was down-regulated in response to condition 4, POE at 50% MIC (Table 34).

A notable observation was the increased expression of the genes encoding for a MexPQ-OpmE efflux pump system (RW109_02297-02299) in response to a broad range of POE and LD associated conditions (3-7; *opmE* gene RW109_02299 not identified with condition 6) (Figure 52). The genes linked to the MexPQ proteins of this efflux pump had the greatest expression increases within the predicted antimicrobial genes (Table 34). A gene encoding for a copper transporter, CueA, RW109_01794 was up-regulated in response to multiple POE and LD associated conditions (3-7; Table 34). In contrast, the RW109_05808 gene annotated as a superoxide dismutase, SodA protein, was down-regulated after exposure to the same five conditions (Table 34).

The RW109_02137 gene encoding for a MexL protein was classified as a repressor of the MexJK efflux pump and was up-regulated in response to multiple POE and LD-associated conditions (4, 5, 6 and 7; Table 34). The RW109_02139 gene associated with MexK was up-regulated only in conditions associated with laundry detergent (6 and 7; Table 34). The RW109_05607 gene associated with a transporter protein VmeV, also had increased expression after exposure to multiple conditions (4, 5, 6 and 7; Table 34). The genes encoding for the MexA and MexB proteins of a MexAB-OprM efflux pump (RW109_01029 and RW109_01030), were identified as up-regulated in response to LD-associated conditions (6 and 7; Table 34). The RW109_06268 gene, which encoded an FbpB2, Fe-transport system, also demonstrated increased expression in response to these LD associated conditions (Figure 51 and Table 34).

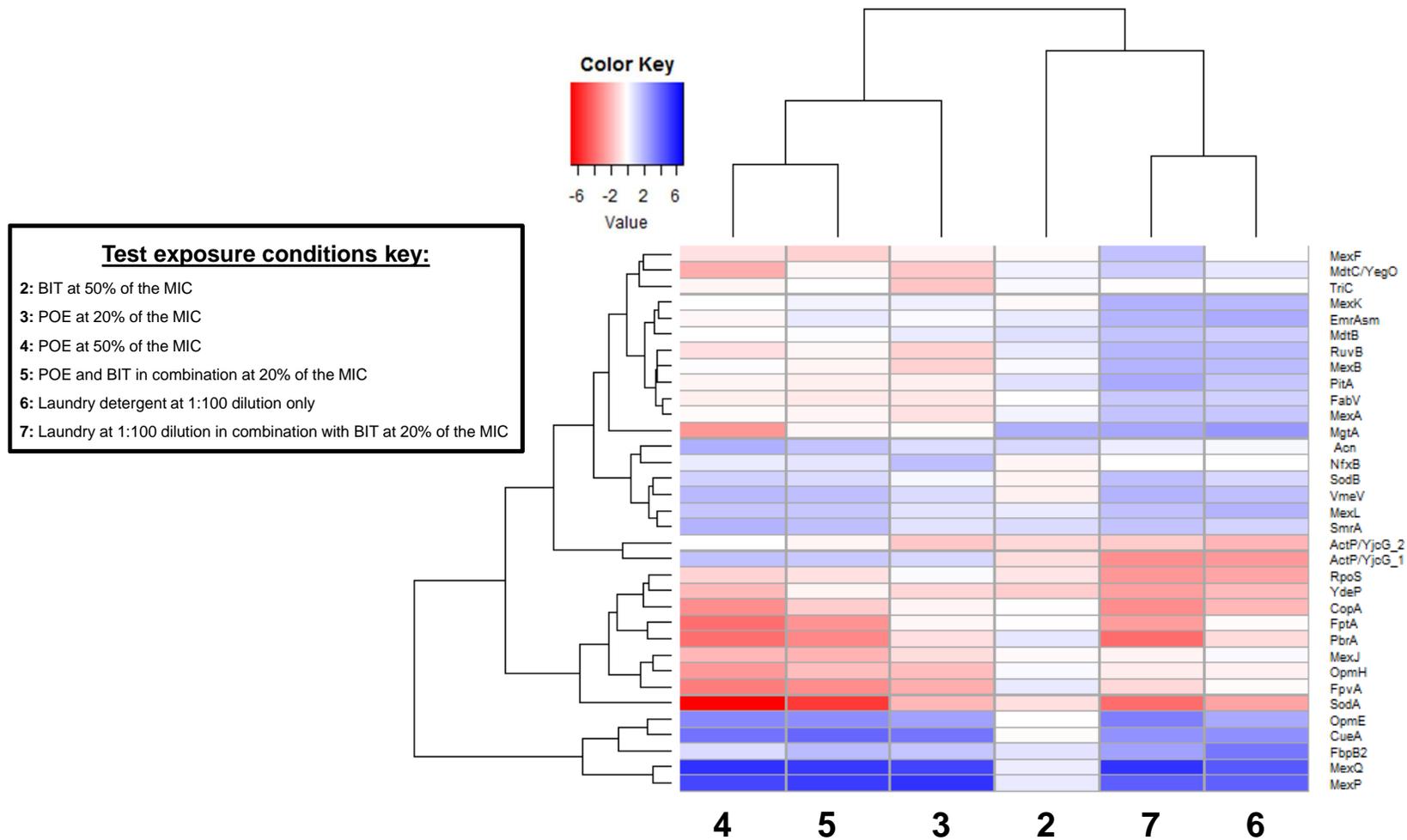


Figure 51. Differentially regulated RW109 antimicrobial genes when exposed to test conditions 2-7

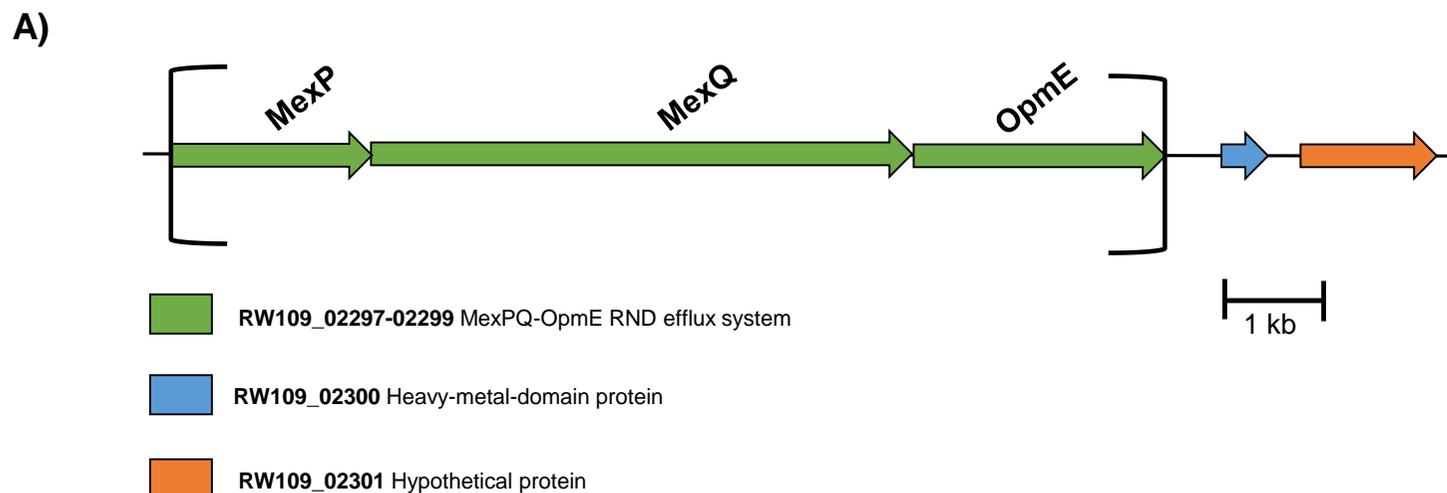
Heat map representing the hierarchical distance cluster analysis of the log₂ fold changes of 34 RW109 antimicrobial genes. The numbers below the heat map represent the test conditions as shown by the test exposure conditions key, blue indicates genes which are up-regulated and red represents genes which are down-regulated and the more intense the colours the greater the gene expression (see color key). The top dendrogram displays the test exposure conditions which have been grouped together via hierarchical cluster analysis and the dendrogram to the left of the heat map represents the gene clusters which are grouped according to their log₂-fold change values across all the conditions.

CHAPTER 5 – RNA-Seq gene expression analysis of RW109 in response to preservatives and product formulations

Table 34. Differentially expressed antimicrobial resistance genes identified when RW109 was exposed to test conditions 2-7

Test exposure	Antimicrobial gene database	CDS ID	CDS Function	Log-2 Fold Change	Adjusted P-value
Condition 2: BIT at 50% of the MIC					
	BacMet	RW109_06416	MgtA,Magnesium-transport	2.2	4.17E-07
Condition 3: POE at 20% of the MIC					
	ABRicate	RW109_02297	MexP,MexPQ-OpmE efflux	5.5	6.94E-15
	ABRicate and BacMet	RW109_02298	MexQ,MexPQ-OpmE efflux	5.0	3.65E-14
	BacMet	RW109_01794	Copper transporter,CueA	3.7	6.91E-12
	ABRicate	RW109_02299	OpmE,MexPQ-OpmE efflux	2.5	0.040597
	ABRicate	RW109_06176	NfxB,Repressor ofMexCD-OprJ efflux	1.8	0.00231
	BacMet	RW109_01506	FpvA,Ferripyoverdine receptor	-2.2	0.007294
	BacMet	RW109_05808	SodA,Superoxide dismutase	-1.9	1.67E-10
	ABRicate	RW109_06568	OpmH,Triclosan efflux	-1.8	0.000729
	ABRicate and BacMet	RW109_00761	TriC,Triclosan efflux	-1.6	0.002666
Condition 4: POE at 50% of the MIC					
	ABRicate and BacMet	RW109_02298	MexQ,MexPQ-OpmE efflux	5.5	2.02E-14
	ABRicate	RW109_02297	MexP,MexPQ-OpmE efflux	4.9	4.60E-09
	BacMet	RW109_01794	Copper transporter,CueA	3.8	1.50E-28
	ABRicate	RW109_02299	OpmE,MexPQ-OpmE efflux	3.2	0.011837
	BacMet	RW109_05444	Acn,Aconitate hydratase	2.2	6.67E-08
	BacMet	RW109_02594	SmrA,ABC-type efflux	2.0	7.38E-06
	BacMet	RW109_05607	VmeV,Transporter AcrB/D/F	1.9	1.39E-09
	BacMet	RW109_02587	ActP/YjcG_1,Cation/acetate symporter	1.7	8.15E-15
	ABRicate	RW109_02137	MexL,Repressor of MexJK efflux	1.6	0.002156
	BacMet	RW109_05808	SodA,Superoxide dismutase	-6.8	2.65E-27
	BacMet	RW109_02125	PbrA,Pb(II) resistance	-3.9	1.91E-21
	BacMet and VFDB	RW109_01453	FptA,Fe-pyochelin receptor	-3.9	1.78E-55
	BacMet	RW109_01506	FpvA,Ferripyoverdine receptor	-3.5	0.00000245
	BacMet	RW109_03986	CopA,Copper resistance protein A	-3.1	0.027726
	BacMet	RW109_06416	MgtA,Magnesium-transport	-2.8	0.002518
	ABRicate	RW109_06568	OpmH,Triclosan efflux	-2.7	6.06E-09
	BacMet	RW109_03482	MdtC/YegO,Multidrug resistance protein	-2.2	0.045269
	BacMet	RW109_06781	YdeP Protein	-1.9	0.0000216
Condition 5: BIT and POE in combination at 20% of the MIC					
	ABRicate and BacMet	RW109_02298	MexQ,MexPQ-OpmE efflux	5.3	5.05E-17
	ABRicate	RW109_02297	MexP,MexPQ-OpmE efflux	5.2	1.54E-13
	BacMet	RW109_01794	Copper transporter,CueA	4.1	2.12E-30
	ABRicate	RW109_02299	OpmE,MexPQ-OpmE efflux	3.1	0.002978
	BacMet	RW109_02594	SmrA,ABC-type efflux	1.8	0.0000441
	BacMet	RW109_05607	VmeV,Transporter AcrB/D/F	1.7	3.12E-10
	BacMet	RW109_05444	Acn,Aconitate hydratase	1.6	0.0000499
	ABRicate	RW109_02137	MexL,Repressor of MexJK efflux	1.5	0.00128
	BacMet	RW109_02587	ActP/YjcG_1,Cation/acetate symporter	1.5	2.37E-11
	BacMet	RW109_05808	SodA,Superoxide dismutase	-5.3	1.53E-48
	BacMet	RW109_02125	PbrA,Pb(II) resistance	-3.2	1.88E-23
	BacMet	RW109_01506	FpvA,Ferripyoverdine receptor	-3.1	0.00000224
	BacMet and VFDB	RW109_01453	FptA,Fe-pyochelin receptor	-2.9	7.73E-47
	ABRicate and BacMet	RW109_02138	MexJ,MexJK efflux	-2.1	0.014979
	ABRicate	RW109_06568	OpmH,Triclosan efflux	-1.8	0.000000735
Condition 6: Laundry detergent unpreserved					
	ABRicate and BacMet	RW109_02298	MexQ,MexPQ-OpmE efflux	4.5	1.77E-09
	ABRicate	RW109_02297	MexP,MexPQ-OpmE efflux	4.2	0.000000267
	BacMet	RW109_06268	FbpB2,Fe-transport system	3.7	0.0000027
	BacMet	RW109_01794	Copper transporter,CueA	3.0	1E-16
	BacMet	RW109_06416	MgtA,Magnesium-transport	2.8	4.55E-18
	BacMet	RW109_06758	EmrAsm,Multidrug resistance protein A	2.3	0.027519
	ABRicate	RW109_02137	MexL,Repressor of MexJK efflux	2.1	0.000000179
	ABRicate and BacMet	RW109_02139	MexK,MexJK efflux	1.9	0.000000611
	ABRicate and BacMet	RW109_01030	MexB,MexAB-OprM efflux	1.8	1.82E-32
	BacMet	RW109_06642	RuvB,Malic protein	1.8	4.48E-10
	BacMet	RW109_05607	VmeV,Transporter AcrB/D/F	1.7	3.96E-12
	BacMet	RW109_05577	PitA,Phosphate transporter	1.6	0.00059
	ABRicate and BacMet	RW109_01029	MexA,MexAB-OprM efflux	1.5	5.01E-11
	BacMet	RW109_02587	ActP/YjcG_1,Cation/acetate symporter	-2.7	4.78E-32
	BacMet	RW109_05808	SodA,Superoxide dismutase	-2.5	2.03E-34
	BacMet	RW109_02193	RpoS,RNA polymerase sigma factor	-2.4	4.76E-30
	BacMet	RW109_03816	ActP/YjcG_2,Cation/acetate symporter	-2.0	0.032927
	BacMet	RW109_03986	CopA,Copper resistance protein A	-2.0	0.049462
	BacMet	RW109_06781	YdeP Protein	-1.8	5.92E-10
Condition 7: Laundry detergent with BIT at 20% of the MIC					
	ABRicate and BacMet	RW109_02298	MexQ,MexPQ-OpmE efflux	5.5	1.19E-14
	ABRicate	RW109_02297	MexP,MexPQ-OpmE efflux	4.3	0.000000572
	ABRicate	RW109_02299	OpmE,MexPQ-OpmE efflux	3.4	0.001503
	BacMet	RW109_01794	Copper transporter,CueA	3.0	6.27E-24
	BacMet	RW109_06268	FbpB2,Fe-transport system	2.5	0.003991
	BacMet	RW109_06416	MgtA,Magnesium-transport	2.4	4.11E-11
	BacMet	RW109_05577	PitA,Phosphate transporter	2.3	0.000000529
	ABRicate and BacMet	RW109_02139	MexK,MexJK efflux	2.1	2.65E-08
	BacMet	RW109_05607	VmeV,Transporter AcrB/D/F	2.1	4.67E-18
	ABRicate and BacMet	RW109_01030	MexB,MexAB-OprM efflux	2.0	9.15E-39
	BacMet	RW109_06642	RuvB,Malic protein	2.0	6.67E-13
	BacMet	RW109_05704	SodB,Superoxide dismutase	1.8	2.93E-16
	ABRicate	RW109_02137	MexL,Repressor of MexJK efflux	1.7	0.0000926
	ABRicate and BacMet	RW109_03517	MexF,MexEF-OprN efflux pump	1.7	0.037303
	ABRicate and BacMet	RW109_01029	MexA,MexAB-OprM efflux	1.6	4.81E-15
	BacMet	RW109_02594	SmrA,ABC-type efflux	1.6	0.00000853
	BacMet	RW109_03481	MdtB,Multidrug resistance protein	1.6	0.002274
	BacMet	RW109_02932	FabV,Putative reductase	1.5	1.37E-08
	BacMet	RW109_02125	PbrA,Pb(II) resistance	-3.9	1.64E-45
	BacMet	RW109_05808	SodA,Superoxide dismutase	-3.9	3.57E-35
	BacMet	RW109_03986	CopA,Copper resistance protein A	-3.0	0.001242
	BacMet	RW109_02587	ActP/YjcG_1,Cation/acetate symporter	-3.0	2.01E-27
	BacMet	RW109_02193	RpoS,RNA polymerase sigma factor	-2.8	8.33E-59
	BacMet and VFDB	RW109_01453	FptA,Fe-pyochelin receptor	-2.7	3.04E-32
	BacMet	RW109_06781	YdeP Protein	-2.6	1.56E-17

Footnotes: Genes indicated in blue were up-regulated and those in red were down-regulated.



B)

Gene ID and annotation	Test exposure conditions									
	POE at 20% of the MIC (3)		POE at 50% of the MIC (4)		BIT and POE in combination (5)		LD Only (6)		LD and BIT at 20% of the MIC (7)	
	Fold Change	Adjusted p-value	Fold Change	Adjusted p-value	Fold Change	Adjusted p-value	Fold Change	Adjusted p-value	Fold Change	Adjusted p-value
RW109_02297	5.5	6.94E-15	5.0	4.60E-09	5.2	1.54E-13	4.2	2.67E-07	4.3	5.72E-07
RW109_02298	5.0	3.65E-14	5.5	2.02E-14	5.3	5.05E-17	4.5	1.77E-09	5.5	1.19E-14
RW109_02299	2.5	4.06E-02	3.2	1.18E-02	3.1	2.98E-03	NA	NA	3.4	1.50E-03
RW109_02300	2.8	1.68E-10	2.4	1.90E-07	3.1	3.14E-18	NA	NA	NA	NA
RW109_02301	3.3	6.85E-13	2.0	2.04E-04	2.8	9.56E-12	2.2	1.23E-06	2.5	3.41E-07

Figure 52. The MexPQ-OpmE RND efflux system identified as up-regulated when exposed to POE- and LD-associated test conditions (3-7)

Panel A shows the MexPQ-OpmE RNA efflux pump operon genes coloured in green (RW109_02297-02299) along with an associated heavy metal domain protein in blue (RW109_02300) and a hypothetical protein in orange (RW109_02301). The log₂-fold changes and adjusted p-values for these genes are shown in panel B. The efflux pump operon genes and the gene encoding the hypothetical protein next to the operon were all up-regulated in conditions 3-7 and the heavy metal domain protein gene was up-regulated in conditions 3, 4 and 5.

5.2.7.7 Differentially regulated genes commonly identified during exposure to the industry relevant conditions

Several DEGs were frequently identified when RW109 was exposed to different industrial test conditions. A prominent observation was the up-regulation of all the genes encoding the MexPQ-*OpmE* efflux pump system in response to test conditions where POE or laundry detergent was present (3-5 and 7; Figure 52). The genes encoding MexP and MexQ were up-regulated with condition 6, laundry detergent with BIT, but there was no increased expression of the *opmE* gene (Figure 52). Both the MexP and MexQ encoding genes were identified within the top 10 up-regulated with conditions 3 and 5, both POE-associated (Table 24 and Table 28) and the MexQ gene was observed within the top ten up-regulated in conditions 4 and 7, POE and LD-associated respectively (Table 26 and Table 32). Interestingly, the hypothetical protein encoded by RW109_02301 next to the efflux system, increased in expression in multiple conditions 3-7. A heavy metal domain protein annotated gene RW109_02300, was up-regulated in conditions 3, 4 and 5 (Figure 52).

A notable gene, which encoded a putative sialic acid transporter, RW109_03936, was always present in the top 10 up-regulated genes in response to multiple test conditions 2, 4, 5, 6 and 7. This gene was identified within a predicted membrane transporter operon (RW109_03936-03940) and when exposed to LD with and without the addition of BIT, all the operon genes were up-regulated (Figure 49). The RW109_04035-04041 genes were also of interest due to the increased expression observed in response to the industrial exposure conditions 2-7 (Figure 53). The RW109_04040 gene was not up-regulated in response to condition 3 and the RW109_04041 gene did not have increased expression with conditions 2 and 3. These genes RW109_04035-04040, encoded for a *gnyRDBHAL* gene cluster and RW109_04041 was annotated as a transcriptional regulator, KdgR (Figure 53). Another gene of interest annotated as a copper transporter, CueA (RW109_01794), was up-regulated in response to test conditions 3-7 (Table 34).

An anthranilate 1,2-dioxygenase small subunit *antB* gene (RW109_03496), an anthranilate 1,2-dioxygenase large subunit, gene (RW109_03497) and an anthranilate 1,2-dioxygenase electron transfer component, *antC* gene (RW109_03495) were of relevance due to the up-regulation observed in response to test conditions 2, 4 and 5. Within the up-regulated genes of condition 6, the *antA* had an increased expression but *antB* and *antC* were not observed. In response to the industry relevant conditions 2-7, an interesting observation was the common down-regulation of the RW109_05181-05187 genes (RW109_05181 did not decrease expression with condition 3). These were identified as the quorum sensing operons *phnAB* and *pqsABCDE* involved in the production of the *Pseudomonas* Quinolone Signal (PQS) (Wade et al., 2005) (Figure 54).

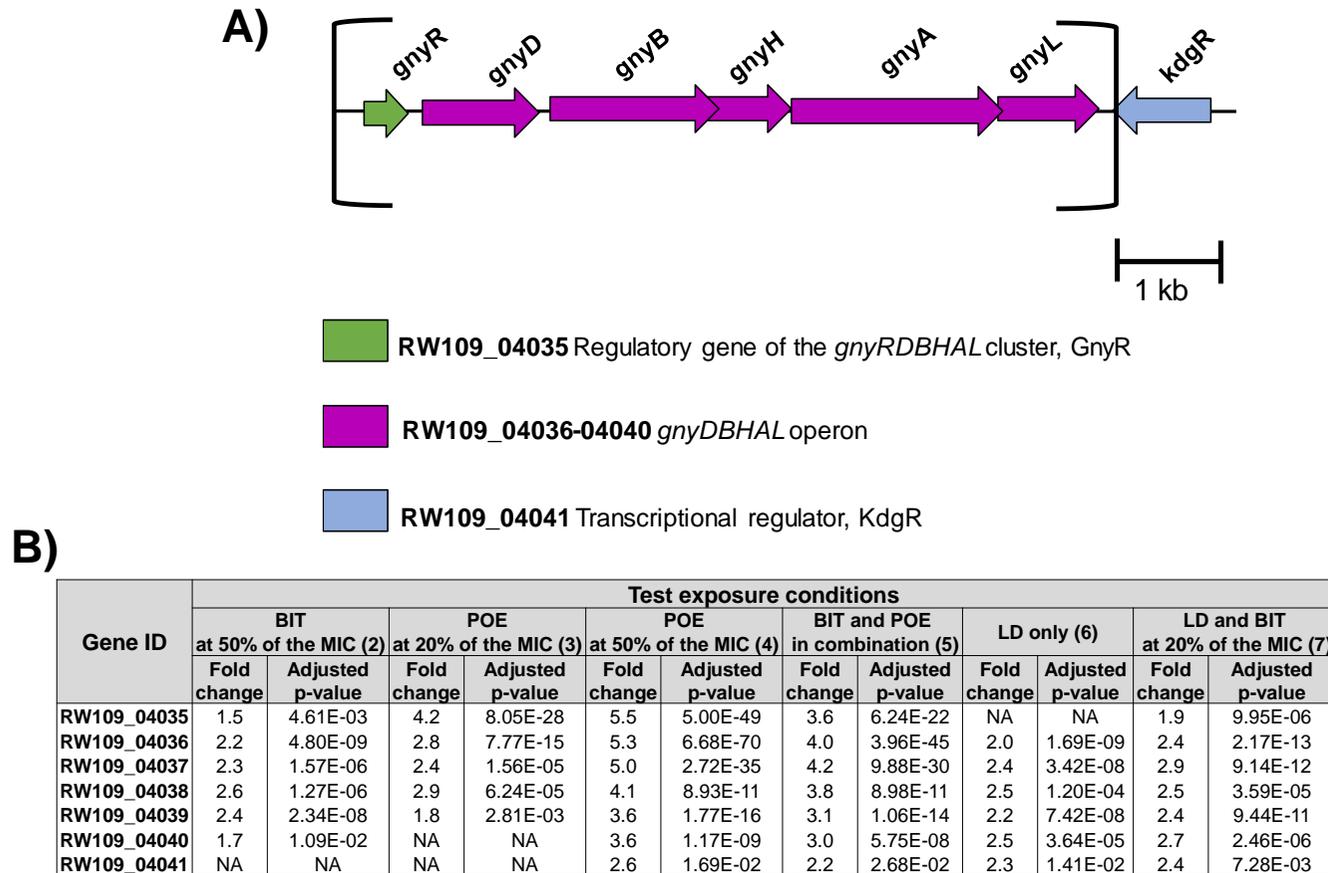


Figure 53. The up-regulated response of the *gnyRDBHAL* gene cluster and KdgR transcriptional regulator during exposure to multiple test exposure conditions

Panel A shows the *gnyDBHAL* operon (RW109_04036-04040) in purple, the gene encoding GnyR (RW109_04035) in green and the KdgR transcriptional regulator gene (RW109_04041) in blue. The log₂-fold changes and adjusted p-values for these genes are shown in panel B. The *gnyRDBHAL* gene cluster and *kdgR* gene were up-regulated in response to conditions 4-7. In condition 2 the operon and *gnyR* gene were up-regulated but the *kdgR* gene was not and in condition 3 the expression of the RW109_04035-04039 genes were increased.

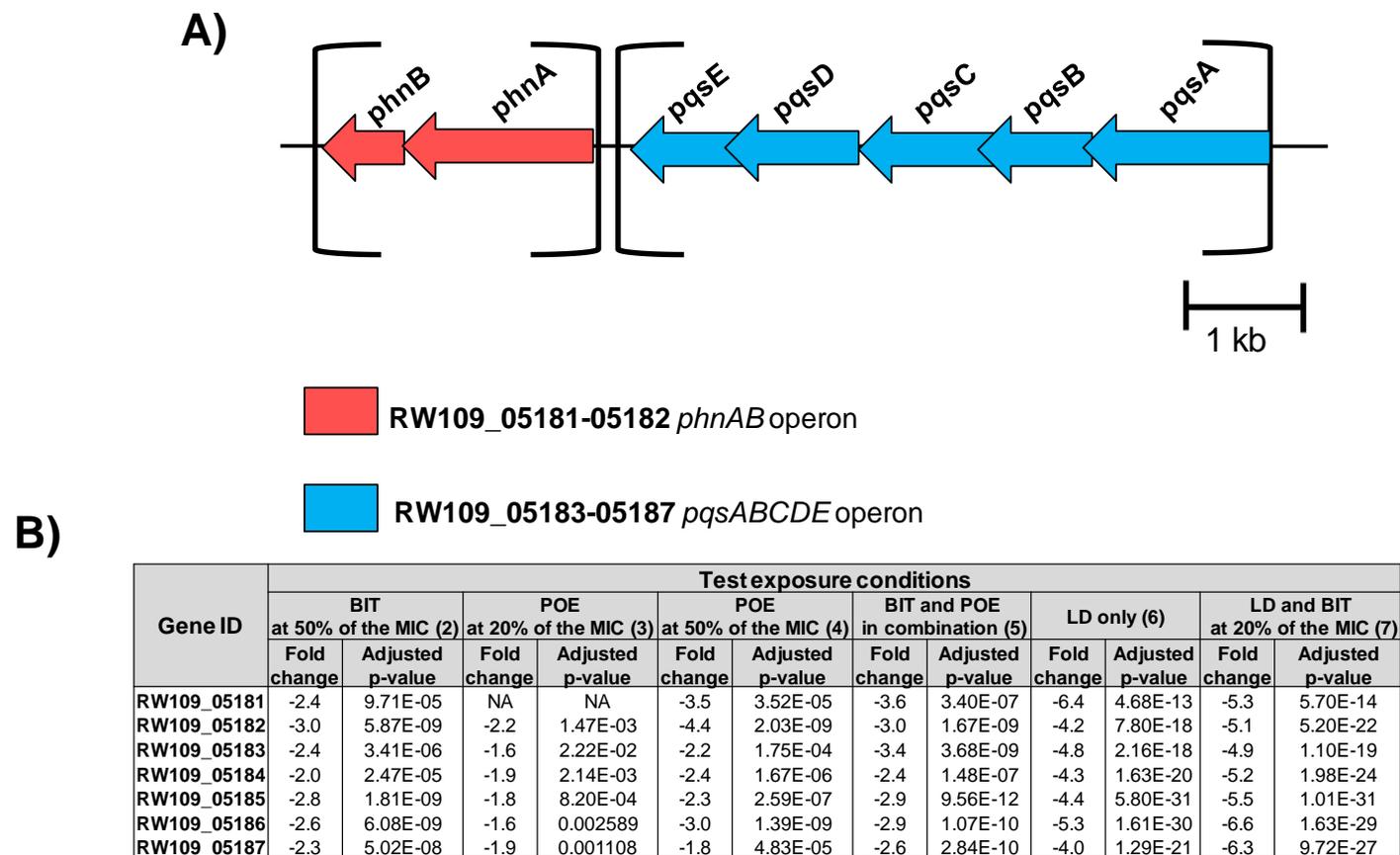


Figure 54. The down-regulated response of the *phnAB* and *pqsABCDE* operon during exposure to multiple test exposure conditions

Panel A shows the *phnAB* operon (RW109_05181-05182) in red and the *pqsABCDE* operon (RW109_05183-05187) in blue. The log₂-fold changes and adjusted p-values for these genes are shown in panel B. In conditions 2-7 both operons were down-regulated, except for the RW109_05181 gene in the *phnAB* operon which was not differentially regulated in response to condition 3.

5.3 Discussion

5.3.1 RNA-Seq analysis of the RW109 following exposure to industry relevant conditions

RNA-Seq analysis was successfully used to investigate the transcriptome of the RW109 strain in response to industry relevant conditions. Analysis of the gene expression data revealed numerous genetic pathways putatively utilised by the industrial strain, for survival in the presence of HPC preservatives and product formulations. The test conditions involved exposure to the preservatives BIT and POE individually and in combination with each other. Also investigated was the response to laundry detergent with and without the addition of the preservative BIT. The selected conditions were used as a primarily investigation of the gene expression response after 24 hours of exposure, regardless of the growth stage of RW109. This was due to the differences in growth dynamics observed when the strain was in the presence of the industrial conditions. This investigation provides a novel insight into the genetic pathways associated with exposure to preservative formulations and HPC products. However, the results from this study require further validation with quantitative RT-PCR on select genes to confirm the transcriptomic responses observed.

The exposure experiments and RNA extractions procedures generated good quality toRNA and mRNA for incorporation into cDNA library preparations and sequencing protocols. The bioinformatics pipeline resulted in a high number of reads aligning and mapping to the gene features of the RW109 PacBio sequence. This study utilised 4 biological replicates to determine log₂-fold changes in gene expression when the test exposure conditions were compared to the control. There is no absolute consensus for classifying a differentially expressed gene and in this study DEGs were defined as exhibiting a log₂-fold change of ≥ 1.5 as described previously (Rushton et al., 2013, Sass et al., 2011) along with an adjusted p-value of ≤ 0.05 . These cut off values were used for initial analysis and the log₂-fold threshold could be increased to 2.0 due to the high number of DEGs identified with conditions 2-7. However the strict cut-off for the adjusted p-value, controls for a false discovery rate when identifying DEGs (Love et al., 2014).

The number of biological replicates necessary for a successful RNA-Seq experiment remains unclear in the literature. A study which investigated the best RNA-Seq analysis tools using 48 biological replicates concluded at least 12 are required (Schurch et al., 2016). However, it was acknowledged that less than 12 replicates were commonly used in RNA-Seq experiments, due to the cost and time incurred in preparing samples. A consensus on the leading tools for accurate differential gene expression analysis with lower biological replicates, includes the DESeq2 and edgeR R tools (Schurch et al., 2016, Tang et al., 2015, Seyednasrollah et al., 2013). With the 4 biological replicates generated for each exposure condition in this study, DESeq2 was successfully implemented to identify the DEGs log₂-fold changes and adjusted p-values. For further validation of the results, another analysis tool could also be applied. The edgeR tool, which is part of the R Bioconductor package (Robinson et al., 2010), was also used to compare the BIT at 50% of the MIC test condition with the control condition. A comparable number of genes were up and down regulated as identified with DESeq2 (data not shown). However, the tool should also be used to validate the gene expression results obtained with the other test exposure conditions. Another

consideration for future work would be to select less exposure conditions and include more replicates to enable increased accuracy and more detailed analysis of the gene expression data.

5.3.2 Overview of the transcriptomic analysis when RW109 was exposed to industry relevant conditions

When exposed to BIT at 20% of the MIC the lowest number of DEGs were identified (0.03% of the whole genome) whereas laundry detergent in combination with BIT at 20% of the MIC, resulted in the greatest number of up and down regulated genes observed in this study (Figure 45). The up and down regulated genes identified in response to LD with BIT were also distributed evenly throughout the RW109 genome (Figure 46, Figure 47 and Figure 48). This indicates a large global gene expression effort was necessary for RW109 to survive, with 20.36% of the whole genome differently expressed in the presence this HPC product in combination with BIT (Table 21).

The RW109 main chromosome had the greatest number of DEGs in response to the test conditions (Table 21). This is understandable due to the higher number of genes on the main chromosome when compared to the two plasmids. Up and down regulated genes were also observed on plasmid 1 and 2 in all test conditions except for BIT at 20% and 50% of the MIC; the number of identified DEGs on plasmids 1 and 2 were generally comparable. Multiple plasmid DEGs encoding for hypothetical proteins were identified, suggesting the uncharacterised plasmid genes play an important role in RW109's provenance as an industrial strain. On plasmid 1, the highest number of DEGs were observed in response to LD with BIT (up-regulated $n = 23$ and down-regulated $n = 35$; Table 21). Within the top ten up-regulated plasmid 1 genes, a DEG of interest encoded for N-acetylglucosaminyl-diphospho-decaprenol L-rhamnosyltransferase WbbL, and was proposed to be involved in the biosynthesis of peptidoglycan and LPS O-antigen production (Crick et al., 2001, McNeil, 1999). The upregulation of this gene suggests a role in cell wall maintenance of plasmid 1 when RW109 was exposed to laundry detergent in combination with BIT. The greatest number of DEGs identified on plasmid 2 were in response to POE at 50% of the MIC and two up-regulated genes of interest encoded for a YafQ and DinJ. These function as a Toxin-antitoxin (TA) system, which are mechanisms for responding appropriately to harsh environments and enabling the multidrug tolerance of biofilms (Harrison et al., 2009). TA-systems also function in stabilising bacterial plasmids by destroying daughter cells not encompassing the plasmid (Gerdes et al., 1986, Pal et al., 2015). The up-regulation of the TA-system in response to POE indicates its role for survival of RW109 in the presence of the preservative and signifies the value of plasmid 2 in response to the industrial conditions.

The limited number of significant up and down regulated genes identified following exposure to BIT at 20% of the MIC, suggests that the RW109 *P. aeruginosa* strain was well adapted to survival in the presence of low concentrations of this preservative. This condition was chosen to determine if genes of interest identified from screening the mini-Tn5-*luxCDABE* mutant library (Chapter 4), were differentially expressed during RNA-Seq analysis. However, only a minimal number of genes could be compared across

the mutagenesis and global gene expression approaches, with the two genes identified from RNA-Seq not being observed in the mutant library analysis.

On the other hand, the same genes recognised with 20% of the MIC in this chapter were also found with BIT at 50% of the MIC, suggesting their importance when RW109 was exposed to this preservative. The up-regulated gene encoded for 3-hydroxyisobutyrate dehydrogenase I and in the Gram-negative *Ralstonia pickettii* bacterium, a gene mutation encoding for this enzyme revealed an inability to neutralise the adverse effects of oxidative stress (Takanashi et al., 2009). A BIT mechanism of action involves the oxidation of thiol groups, resulting in the generation of free radicals and damage to the cell (Collier et al., 1990). The up-regulation of a gene, which encodes an enzyme linked to neutralising oxidative stress, suggests its use in minimising the adverse effects of the preservative BIT. The down-regulated gene identified with exposure to both 20% and 50% of the BIT MIC was annotated as a hypothetical protein.

Interestingly a gene involved in sulphur metabolism was up-regulated in response to BIT at 50% of the MIC and was also identified as an insertion site of a BIT positive *lux* transposon mutant (Chapter 4, Table 2). This gene encoded for a bi-functional enzyme CysN/CysC (RW109_05781) and an adjacent gene annotated as a sulfate adenylyl-transferase subunit II CysD (RW109_05782), also had increased expression in response to BIT at 50% of the MIC. The *cysND* operon has been associated with an increased expression in *P. aeruginosa* in response to oxidative stress (Goldová et al., 2011). This suggests the oxidation effect of the preservative BIT results in RW109 up-regulating the *cysND* operon. Another interesting observation with BIT at 50% of the MIC was the increased expression of genes encoding carbamate kinase 1 and a nitrite reductase precursor, which are involved in nitrogen metabolism. This suggest that up-regulation of nitrogen metabolism-related enzymes enhances the ability of RW109 to survive the presence of BIT. A previous study on a *P. aeruginosa* strain resistant to isothiazolinones demonstrated a significant number of DEGs involved in nitrogen metabolism when exposed to a Kathon preservative formulation (Zhou et al., 2016). This illustrates the importance of nitrogen metabolism when *P. aeruginosa* strains are exposed to isothiazolinone preservatives.

This investigation also identified 34 differentially expressed RW109 predicted antimicrobial resistance genes (Figure 51). This illustrates the importance of several previously characterised antimicrobial resistance mechanisms for the survival in the presence of industry relevant conditions. Numerous reports indicate that increased resistance to biocides also decreases susceptibility to antibiotics (McBain et al., 2002, Pal et al., 2015, Russell, 2003a). However, it may also be the case that a strain encompassing antibiotic resistance capabilities such as RW109, already has the inherent structural and functional characteristics to be resistant to industrial biocides. Proving direct associations between biocides and antibiotic resistance is a point of much debate and only a few specific mechanisms of co-resistance have been defined. For example with *Burkholderia lata*, exposure to isothiazolone preservatives resulted in the up-regulation of an efflux pump which was also directly involved in fluoroquinolone antibiotic resistance (Rushton et al., 2013). The *Burkholderia* strain involved did not gain multiple antibiotic resistances and

overall was less tolerant of antibiotics, but just more resistant to fluoroquinolones as a direct correlation to the efflux pump up-regulation (Rushton et al., 2013).

5.3.3 Prevalent RW109 genes differentially expressed in response to multiple industry relevant conditions

An interesting observation of this study was the frequent up-regulation of the same genetic pathways in response to different test conditions. This identified common resistance mechanisms used by RW109 when exposed to industrial preservatives and product formulations. The genes associated with the predicted antimicrobial resistance MexPQ-OpmE RND efflux pump system, were up-regulated in response to POE at both 20% and 50% of the MIC, BIT and POE in combination and laundry detergent with BIT (Figure 52). The genes encoding MexP and MexQ were up-regulated with the LD-only condition but an increased expression with the OpmE gene was not observed. The MexP and MexQ were also commonly identified within the top ten up-regulated in response to these industrial conditions. This efflux pump system is used by *P. aeruginosa* to export a variety of antibiotics such as fluoroquinolones, tetracycline, chloramphenicol and macrolides (Lister et al., 2009, Mima et al., 2005). This efflux system also functions as a metal ion exporter and knocking out the *mexP* or *mexQ* genes elevates *P. aeruginosa* susceptibility to copper (Thaden et al., 2010).

The common up-regulation of the efflux pump in response to industry conditions, suggests it could also be used to pump out preservatives. In particular, the efflux pump genes had increased expression when exposed to POE at 20% and 50% of the MIC, signifying its use as a resistance mechanism in the presence of this preservative. When exposed to BIT at these concentrations, the efflux pump genes were not up-regulated indicating it was not used with this preservative. In response to BIT and POE in combination both at 20% of the MIC, all the efflux pump genes demonstrated an increased expression. This suggests the efflux pump may be used as a resistance mechanism in response to preservative combinations. The efflux pump was linked with the presence of POE within a formulation, as a limited number of genes were up-regulated when exposed to just BIT at 20% of the MIC. Consequently, the increased expression of all the efflux pump genes was observed with LD in combination with BIT, whereas in response to LD without BIT, the OpmE encoding gene was not up-regulated. This indicates the addition of BIT to the laundry detergent results in RW109 utilising the MexPQ-OpmE RND efflux pump system as a resistance mechanism.

Follow-up research aimed at knocking out the efflux pump system genes in the RW109 strain and investigating changes in susceptibility to industry relevant conditions should be carried out. As noted above the role of efflux as an important industrial preservative resistance mechanism was previously demonstrated with the up-regulation of a *B. Lata* RND-type efflux system in response to isothiazolone preservatives (Rushton et al., 2013). Examples of *P. aeruginosa* efflux pump-mediated biocide resistance include the MexJK and MexAB-OprM efflux systems which export the biocide triclosan (Chuanchuen et al., 2002, Chuanchuen et al., 2001, Schweizer, 1998) and the MexCD-OprJ efflux pump which is induced by disinfectants such as benzalkonium chloride, chlorhexidine and ethidium bromide (Morita et al., 2003).

This highlights the versatility of multidrug efflux pumps in response to agents other than antibiotics. The consistent up-regulation of the *P. aeruginosa* MexPQ-OpmE RND efflux system genes in response to the industry relevant conditions used in this study, emphasises the approach to target efflux pumps in preservation strategies. Chemicals which act as efflux pump inhibitors (Stavri et al., 2006, Lomovskaya et al., 2001) could be used in combination with industrial preservatives to potentiate their activity and enhance HPC preservation systems.

Another gene, which frequently had increased expression when RW109 was exposed to the industry relevant conditions, was a putative sialic acid transporter. This gene was continuously present within the top ten up-regulated genes in response to multiple test conditions associated with all preservatives as well as laundry detergent, 2, 4, 5, 6 and 7 (Figure 49), signifying its importance as a potential mechanism of survival in the presence of industry preservatives and HPC products. The sialic acid transporter encoding gene was identified within a predicted membrane transporter operon (RW109_03936-03940) and in response to LD with and without the addition of BIT, all the genes within this operon were up-regulated (Figure 49). Sialic acids incorporate more than 50 naturally occurring nine-carbon amino sugars, which can be used by bacteria as sole carbon sources and are imported via sialic acid transporters (Vimr et al., 2004, Almagro-Moreno and Boyd, 2009). Sialic acid is commonly found in mucus rich environments (Vimr, 2013) and has been linked to *P. aeruginosa* pathogenicity in immunocompromised patients (Khatua et al., 2010). It is proposed acquiring and assembling sialic acids on the surface mediates enhanced binding to host cells, confers resistance to complement-mediated killing and facilitates the establishment of persistent infection (Khatua et al., 2010, Greiner et al., 2004).

Sialic acid uptake has also been linked to biofilm formation in the Gram-negative bacterium *Haemophilus influenzae* (*H. influenzae*); sialic acid was identified as an important biofilm matrix component during persistent respiratory tract infections (Jurcisek et al., 2005, Greiner et al., 2004). The *P. aeruginosa* RW109 strain may increase the expression of genes involved in the sialic acid transport system in order to uptake the molecule for enhanced survival in industrial conditions. Sialic acid may be required for the RW109 to adhere and form a biofilm to permit viability when exposed to preservative formulations and HPC products. The increased expression of the transport system could also be for use as an antimicrobial export mechanism of the preservative formulations and HPC products. Since the global gene expression was performed under planktonic conditions with limited potential for biofilm formation, this functional role would appear more likely. The sialic acid transport system is clearly important during exposure to industry related conditions and targeting its components in preservation strategies could be beneficial. A patent for an agent which specifically inhibits sialic acid transporters in numerous pathogens, including *P. aeruginosa*, has been filed and reports a reduction in bacterial growth upon exposure (Gibson et al., 2006). Agents, which specifically target common resistance mechanisms, could be incorporated into HPC preservative formulations to minimise proliferation of contaminants.

Genes within the isoprenoid degradation associated *gnyRDBHAL* operon were commonly up-regulated when exposed to the industrial test conditions (not including BIT at 20% of the MIC) (Figure 53). The

adjacent transcriptional regulator KdgR, had increased expression in response to conditions 4-7 and was characterised in a transposon mutant with an increased light emission in response to BIT at 20% of the MIC (BIT positive) (Chapter 4, Table 16). However, during RNA-Seq analysis the associated gene was not up-regulated when exposed to the preservative BIT individually. The *gnyRDBHAL* gene cluster encodes enzymes which function in *P. aeruginosa* acyclic isoprenoid degradation (Diaz-Perez et al., 2004). The frequent up-regulation of these genes suggests a potential role of the *gnyRDBHAL* operon in the survival of RW109 when exposed to industry relevant conditions. The gene cluster has previously been identified as up-regulated during transcriptomic analysis of developing *P. aeruginosa* biofilms (Waite et al., 2006). Increased transcript levels of the *gny* genes were also observed during growth of *P. aeruginosa* in mucus when compared to cultivation in minimal media (Cattoir et al., 2012). Therefore the *gnyRDBHAL* operon may have a role in the biofilm and virulence capabilities of *P. aeruginosa* and the up-regulation of the operon, suggests its use as a survival mechanism when RW109 was exposed to industrial conditions.

A gene characterised via the BacMet biocide resistance database encoded for a copper transporter, CueA and was up-regulated in response to the industrial conditions (not including BIT at 20% and 50% of the MIC) (Table 34). In *P. aeruginosa*, the *cueA* gene has a role in copper tolerance, with mutations in the gene resulting in increased sensitivity to the heavy metal (Thaden et al., 2010). *P. aeruginosa* strains with a *cueA* mutation also demonstrated a 20-fold attenuation in the spleens of mice when compared to the wild-type (Schwan et al., 2005), highlighting the involvement of the gene in virulence. There are various reports of bacterial strains resistant to heavy metals that also have decreased susceptibility to antibiotics (Yazdankhah et al., 2014, Seiler and Berendonk, 2012, Berg et al., 2010), and it may be the case that mechanisms of copper tolerance could also be used for survival in the presence of industrial preservative formulations. This is supported by the up-regulation of the *cueA* gene in response to the majority of preservative formulations and HPC products used in this study (excluding BIT).

The *antABC* genes had increased expression when RW109 was exposed to BIT and POE individually at 50% of the MIC and in response to BIT in combination with POE both at 20% of the MIC. The *antA* and *antB* genes were also observed within ten up-regulated in response to these conditions. The *antABC* operon functions in the first stage of anthranilate degradation and the ability of *P. aeruginosa* to form biofilms has been identified as partly reliant on this degradative pathway (Costaglioli et al., 2012). Transcriptomic analysis of *P. aeruginosa* biofilm sessile cells identified an increased expression of the *antA*, *antB*, and *antC* genes and mutations in these genes, significantly reduced adhesion and biofilm forming abilities (Costaglioli et al., 2012). The formation of biofilms especially with sessile cells, is a recognised resistance mechanism, often associated with a reduced susceptibility to antimicrobials (Donlan and Costerton, 2002). This link between anthranilate degradation and biofilm formation, suggests the RW109 strain may be transitioning to a biofilm state in order to survive exposure to the preservatives BIT and POE at 50% of the MIC and when in combination with each other at 20% of the MIC.

The recurrent down-regulation of the same genes in response to different test conditions signifies commonly suppressed mechanisms. Two operons *phnAB* and *pqsABCDE* were frequently identified with decreased expression in response to the industrial test conditions (not including BIT at 20% of the MIC) (Figure 54). These operons were adjacent to each other within the RW109 genome and encode enzymes necessary for the biosynthesis of PQS (Wade et al., 2005). The *pqsABCDE* operon encodes enzymes that catalyse production of five different types of the PQS molecule 4-hydroxy-2-alkylquinolines (HAQs) (Déziel et al., 2004). The operon *phnAB* is involved in the conversion of the central metabolite chorismate to anthranilate by anthranilate synthase, which is an essential precursor pathway in the biosynthesis of HAQs (Essar et al., 1990, Déziel et al., 2004).

PQS are cell to cell signalling molecules involved in quorum sensing enabling the development of multicellular processes for survival purposes such as biofilm formation (Diggle et al., 2007). These molecules are generally produced in response to harsh environments and involve substantial investment of bacterial reserves (Häussler and Becker, 2008). After an extensive exposure to rate limiting conditions such as preservative formulations and HPC products, PQS production may not be sustainable. The *phnAB* and *pqsABCDE* operons were possibly beneficial during initial exposure to the industrial test conditions, however after 24 hours PQS biosynthesis may not have been viable for RW109. Further work could involve exposing the strain to the same test conditions for a shorter amount of time to determine alterations in the expression of these operons. It is also proposed that under increased exogenous stresses, the overproduction of PQS becomes harmful to *P. aeruginosa* (D'Argenio et al., 2002, Heurlier et al., 2005).

Down-regulation of the PQS genes in response to industrial preservatives may also be a mechanism, which allows the RW109 strain to persist in the presence of these antimicrobials and harsh conditions. Excessive PQS disrupts electron flow through the respiratory chain which releases reactive oxygen species, this impairs membrane integrity inducing bacterial cell autolysis and the release of DNA (Hazan et al., 2016). A *P. aeruginosa* mutant deficient in the *pqsA* gene was grown in the presence of the antibiotics ciprofloxacin, imipenem and gentamicin, and killing was found to be significantly delayed in comparison to the wild-type and a mutant which over produced PQS (Häussler and Becker, 2008). The *pqsA* mutant also demonstrated an enhanced resistance when exposed to hydrogen peroxide in relation to the wild-type and the PQS overproducing mutant (Häussler and Becker, 2008). Again, the operons may have been advantageous as an initial response to the test conditions used in this study. However after prolonged exposure, PQS biosynthesis may have become detrimental to RW109, resulting in the down-regulation of *phnAB* and *pqsABCDE*.

The transcriptomic analysis performed in this study generated a vast amount of data. Consequently, it was decided to mainly focus the discussion on the genetic pathways commonly identified as differently expressed when exposed to industry relevant conditions. This interpretation has highlighted central resistance mechanisms utilised by RW109, with many of the genes and operons discussed being found within the top ten up and down regulated DEGs in response to each test condition. This signifies the

importance of these recognised mechanisms in the ability of RW109 to survive the presence of preservative formulations and HPC products. This study has also highlighted a variety of genetic pathways with differential regulation in response to preservatives and product formulations. This indicates the resistance mechanisms used by *P. aeruginosa* to survive in an industrial environment are complex and may involve more than one genetic pathway. Further interpretation of the findings from this study would enable a comprehensive characterisation of the resistance mechanisms utilised in response to each of the industry relevant conditions.

5.4 Conclusions

The main conclusions from this chapter were as follows:

- 1) Good quality RNA for use in RNA-Seq analysis was extracted from *P. aeruginosa* RW109 following exposure to conditions which resulted in a OD percentage reduction of up to 68%, when compared to growth in TSB for 24 hours.
- 2) When exposed to industry relevant conditions, RNA-Seq transcriptomic analysis was successfully applied to investigate the gene expression response of *P. aeruginosa* RW109. The exposure conditions included the individual preservatives BIT and POE at different concentrations, a combination of these two preservatives and laundry detergent with and without the addition of the preservative BIT.
- 3) Exposure of *P. aeruginosa* RW109 to BIT at 20% of the calculated MIC resulted in the least number of differentially regulated genes, with 0.03% of DEGs in relation to the whole genome. In response to laundry detergent in combination with BIT, the greatest proportion of up and down regulated genes (20.36%) within the RW109 genome was observed.
- 4) The consistent up-regulation of the same genetic pathways when exposed to the different test conditions identified common resistance mechanisms used by *P. aeruginosa* RW109 in response to industrial preservatives and product formulations. These included a MexPQ-OpmE RND efflux pump system, a putative sialic acid transporter and a *gnyRDBHAL* operon. These frequently used resistant mechanisms can be targeted during the development of HPC preservation strategies.
- 5) The repeated decreased expression of certain genetic pathways in response to different test conditions signifies commonly suppressed *P. aeruginosa* mechanisms. Two operons *phnBA* and *pqsEDCBA* involved in PQS production and quorum-sensing, were repeatedly down-regulated when exposed to industry relevant conditions after 24 hours.

6. Genome-scale metabolic network reconstruction of an industrial *P. aeruginosa* strain, RW109

6.1 Introduction

Genome-scale metabolic network reconstructions provide a comprehensive overview of bacterial biochemical networks and can be analysed to predict the essentiality of genes and reactions when a bacterium is exposed to various environmental conditions (Thiele and Palsson, 2010). Information from a well-annotated whole-genome sequence, together with databases and relevant literature searching, enables the gene-protein-reaction relationship to be determined for characterisation of the network of reactions involved in bacterial metabolism. Updating a previously developed metabolic model with the use of genetic information is a useful starting point during model reconstruction to better represent a specific bacterial strain (Rocha et al., 2008). The metabolic potential of an organism under various conditions can be simulated using a constraint-based modelling approach called flux balance analysis (FBA) to further understand the biological system of interest (Thiele and Palsson, 2010). Constraint-based reconstruction and analysis (COBRA) methods are widely used; they are built on three fundamental concepts which include the constraints of a microorganism based on its genome and biochemical networks, the objective function which accounts for the necessary components for growth and the association of gene-protein-reaction relationships (Lewis et al., 2012). A COBRA model is represented mathematically enabling the computational interpretation of genome-encoded reactions and a better understanding of the metabolism of a microorganism. FBA is commonly used in conjunction with COBRA when investigating the metabolite flow through genome-scale metabolic network reconstructions (Orth et al., 2010). The rapid increase in the availability of high-throughput genomic, proteomic and metabolomics data has advanced the capabilities of genome-scale metabolic models as investigative tools (Lewis et al., 2012). As an advanced research technique, metabolic models are valuable when examining the virulence or antimicrobial resistance capabilities of a microorganism.

The recently published metabolic model for the *P. aeruginosa* strain PA14 evaluated the interrelationships between growth and virulence-linked pathways using an *in silico* media to mimic the lung of a CF patient (Bartell et al., 2017). The results enabled the prediction of 46 essential genes, which were involved in growth inhibition and virulence reactions, with a high representation of fatty acid and phospholipid metabolism pathways. The findings provide information to enhance the design of effective CF antimicrobial treatments which have a wide-impact on *P. aeruginosa* metabolism (Bartell et al., 2017). The reconstruction and biological interpretation of the PA14 metabolic model demonstrates how computational screening can be a useful tool in understanding how bacteria respond to their environments.

There is currently motivation in the field of modelling for the development of condition-specific models, which represent the metabolic state of a bacterium when exposed to a particular environmental condition (Schmidt et al., 2013, Hyduke et al., 2013, Ebrahim et al., 2016). Various methods have been developed which enable transcriptomic data to be integrated into a genome-scale model (Blazier and Papin, 2012,

Lewis et al., 2012, Vivek-Ananth and Samal, 2016). The integration applies further constraints on a model, allowing for an accurate representation of gene expression dynamics and resulting condition-specific models can be used to predict the essentiality of genes and reactions. Hypotheses can then be formulated regarding the importance of specific metabolic pathways during exposure to relevant conditions, which can then be validated. This offers a streamlined hypothesis-driven approach for experimentally investigating the genetic and metabolic determinants of antimicrobial resistance, by initially focusing on predicted genes and enzymes of interest. Metabolic modelling can examine the genes deemed necessary for a bacterium to survive within harsh environments and draw attention to potential novel antimicrobial targets to combat the increasing problem of resistance development. Putative targets to combat the resistance with the antibiotic colistin have recently been identified using a constraint-based approach with an *Acinetobacter baumannii* (*A. baumannii*) genome-scale metabolic model, combined with gene expression data (Presta et al., 2017). A set of novel genes were identified using *in silico* gene essentiality predictions when the bacteria was exposed to colistin. The study identified essential reaction targets for development of new drugs against resistant *A. baumannii*. This work highlights the use of genome-scale models along with transcriptomic data in understanding the response of bacteria to stressful conditions.

There are over 100 published and experimentally validated metabolic models which represent various microorganisms (Broddrick, 2017). Genome-scale metabolic models serve as prominent resources in understanding how bacteria respond to various environmental conditions, and enable the testing and generation of relevant hypotheses. As far as we know, no genome-scale metabolic models exist, which represent an industrial *P. aeruginosa* strain; this would be useful to determine the metabolic requirements for survival in home and personal care products (HPC) products and in the presence of industrial preservatives.

6.1.2 Aim and objectives

The overall aim of this chapter was to reconstruct and utilise a genome-scale metabolic model specific to an industrial *P. aeruginosa* strain and this was achieved by carrying out the following objectives:

- 1) Up-date a recently published PA14 metabolic model to represent the industrial RW109 strain using the PacBio genome sequence (Chapter 3).
- 2) Integrate the RNA-Seq transcriptomic gene expression data (Chapter 5) with the reconstructed model and generate condition-specific models that represent RW109 when exposed to industrial conditions.
- 3) Predict the essentiality of model reactions when RW109 was exposed to industry relevant conditions.

6.2 Results

6.2.2 Updating the PA14 model

The recently published and validated PA14 model was updated to represent an industrial *P. aeruginosa* strain by using the complete RW109 genome sequence (Chapter 3). Updating was initially carried out by replacing the PA14 model gene locus tags with RW109 orthologs identified from a BLASTp search between the two strains. Out of the 1,129 genes in the PA14 model, 7 did not match to any RW109 genes even when the BLASTp filtering parameters were reduced. The remaining 1,122 PA14 model genes matched to 1,185 RW109 genes using the set BLASTp filtering parameters.

6.2.3 Classifying the additional reactions of the RW109 metabolic model

When the FASTA sequences were run through the modelSEED (Henry et al., 2010) and KEGG databases (Kanehisa et al., 2017), 166 additional reactions were identified and included in the updated model. 114 genes were associated with these additional reactions as follows: 38 were unique to RW109 when compared to PA14 (there was no match during a BLASTp search to any PA14 genome genes); 54 were RW109 genes which had similarity to PA14 genes not included in the original model; and 22 were RW109 genes which corresponded to PA14 genes already in the original model and associated with other reactions. Running the RW109 FASTA sequences through the modelSEED database also identified existing reactions within the PA14 model, which were associated with 7 unique RW109 genes and an additional 45 RW109 genes, which had similarity to PA14 genes not in the original model.

KEGG functional categories were assigned to the 166 additional reactions of the RW109 model and the xenobiotic biodegradation and metabolism category was found to be the most represented, with 33 reactions (Figure 55). Out of the 33 reactions in this category, 31 were linked to genes from plasmid 1. Within this functional category, 60% of the reactions were associated with the metabolism of xenobiotics by cytochrome P450 sub-category (Figure 56A). The original PA14 model had 11 reactions within the xenobiotic biodegradation and metabolism category; 1 reaction was associated with aminobenzoate degradation and 10 with benzoate degradation; the additional reactions of the RW109 model were not associated with the benzoate degradation group. The second functional category with the greatest number of reactions was amino acid metabolism, which had 30 reactions (Figure 55). Within this functional category, 43% of the reactions were associated with the glycine, serine and threonine metabolism sub-category (Figure 56B). When analysing the genes associated with this sub-category, two genes (RW109_00449 and RW109_00565) from plasmid 1 were previously described as being part of the serine and threonine metabolism pathway KEGG module (M00555) (Chapter 3, Table 8; Section 3.2.3.2.2). Within the glycine, serine and threonine metabolism sub-category in the RW109 model, RW109_00449 was associated with 7 reactions and RW109_00565 was found within 2 reactions.

CHAPTER 6 - Genome-scale metabolic network reconstruction of an industrial
***P. aeruginosa* strain, RW109**

When examining the location of the genes associated with the 166 additional reactions of the RW109 model, 33 genes were from plasmid 1, 78 from the main chromosome and 3 genes were from plasmid 2. Plasmid 1 genes were associated with 97 additional reactions and the xenobiotics biodegradation and metabolism functional category represented the highest number of reactions (31.96%) linked to plasmid 1 genes, followed by amino acid metabolism (18.56%), nucleotide metabolism (10.31%), glycan biosynthesis and metabolism (9.28%), and carbohydrate metabolism (9.28 %). The 3 genes from plasmid 2 were linked to 5 additional reactions; 3 were involved in membrane transport, 1 with replication and repair and the other in nucleotide metabolism.

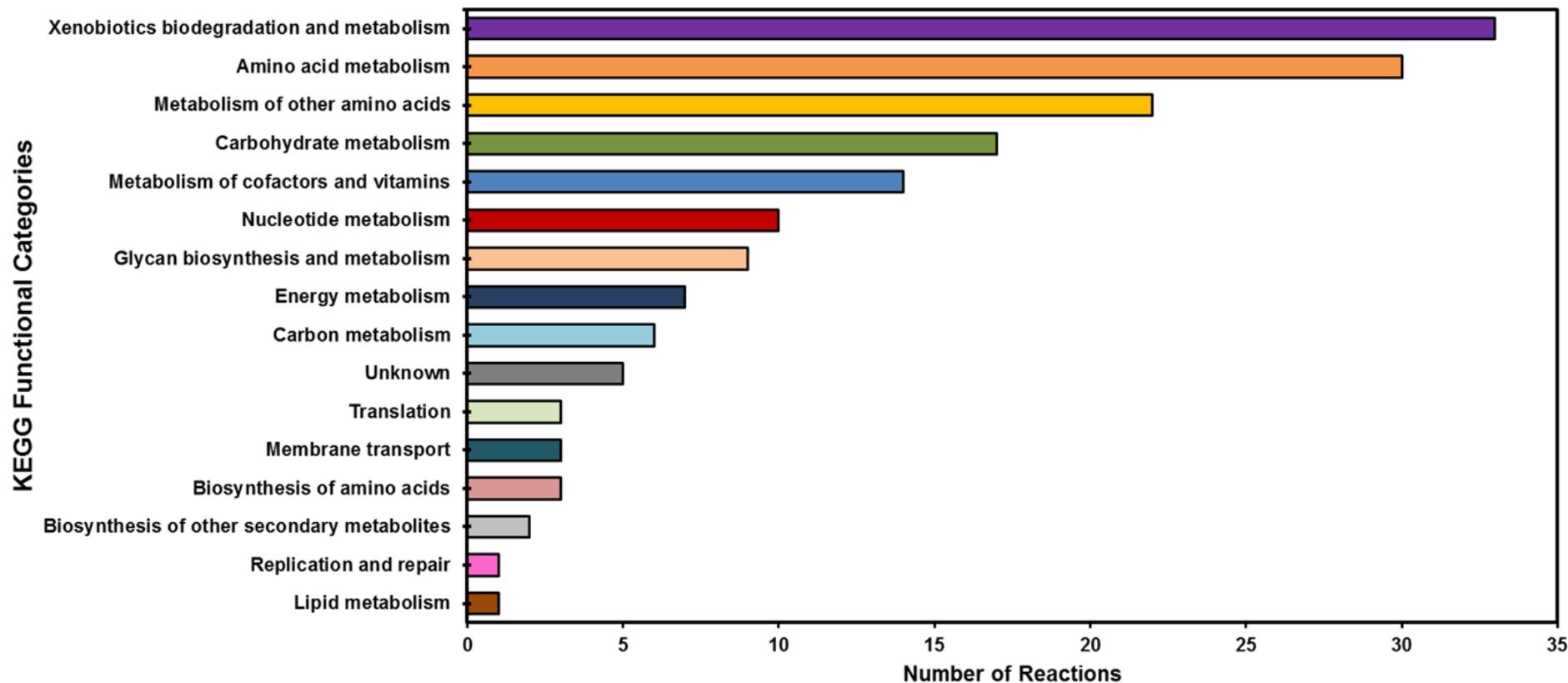


Figure 55. KEGG functional categories of the additional RW109 model reactions

The bar-chart shows the functional categories of the 166 additional reactions included within the RW109 metabolic model. The functional categories were assigned to the reactions via the KEGG database and the xenobiotics biodegradation and metabolism category had the greatest number of the additional reactions associated with it, followed by amino acid and metabolism.

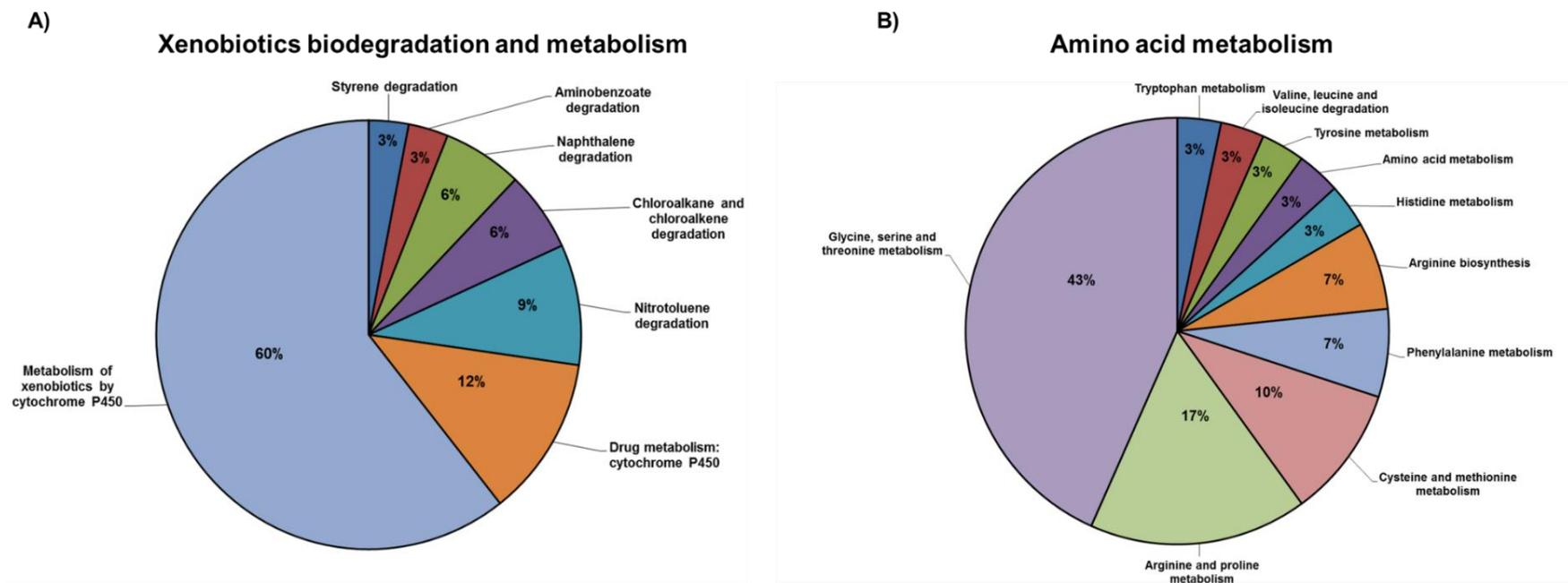


Figure 56. Breakdown of the key reactions in the two major functional categories associated with the RW109 additional model reactions

The pie charts represent the percentage of reactions assigned to the sub-categories of A) xenobiotics biodegradation and metabolism and B) amino acids metabolism KEGG functional categories. Within A) the metabolism of xenobiotics by cytochrome P450 was found to have the greatest number of reactions assigned to it with 60% and in B) the glycine, serine and threonine metabolism sub-category had the greatest number of reactions (43%) within the amino acid metabolism category.

6.2.4 The RW109 model in COBRA format

The RW109 model in Excel format was loaded into MATLAB and converted into a COBRA toolbox model as described in Section 2.8.3. The model was optimised with FBA, once the model objective was set to the RW109_Biomass reaction, which serves as an approximation of growth and the *in silico* media was changed to LB complete media. This media was chosen as it represented the complete *in silico* media available on the change media script (see changeMedia_SEED.m script on CD-ROM Chapter 2 folder). The output of FBA was an objective function value of 17.45 hr⁻¹, indicating that the model was able to grow and an optimal flux balance analysis solution for the model was found *in silico*.

6.2.5 Removing PA14 genes with no similarities to RW109 genes

The 7 genes from the PA14 model with no similarity to any RW109 genes were individually deleted from the model and this resulted in the removal of 5 non-essential reactions (Figure 57). Out of these non-essential reactions, 4 were within the glycan biosynthesis and metabolism category and functioned in lipopolysaccharide biosynthesis of B-band O antigen; within the model, 38 remaining reactions were associated with lipopolysaccharide biosynthesis. The other non-essential reaction removed from the model was associated with lipid metabolism and had a function in glycerophospholipid metabolism; however, 56 model reactions remained which were also involved in glycerophospholipid metabolism. The PA14_60120 gene, which did not match to any within RW109, was also deleted from the model and its removal did not result in any loss of reactions. The PA14_60120 gene was associated with the rxn01672 reaction and the enzyme associated could be encoded for by more than one gene within the PA14 model. The other gene associated with the rxn01672 reaction was PA14_19090 and this was matched to the RW109_02319 gene allowing the reaction to remain within the RW109 model reconstruction.

When the PA14_61770 gene was removed from the model, it resulted in a biomass objective function value of 0, indicating that this gene is essential for biomass production. Indeed, this gene was associated with the essential reaction ATP: D-ribose-5-phosphate pyrophospho-transferase (rxn00770), which produces 5-Phospho-alpha-D-ribose 1-diphosphate. Reducing the strict BLASTp filtering parameters did not identify any match for the PA14_61770 gene within the RW109 genome. Within the KEGG database this reaction is termed R01049 and the functional units of this reaction (KEGG module M00005; Chapter 3, Section 3.2.6.1.2) were identified as biologically achievable for the PA14 and PAO1 strains, but was not for RW109. The enzyme involved in the reaction rxn00770, ATP:ribose-1,5-bisphosphate phosphotransferase, was searched on the Uni-Prot database (Consortium, 2017) and 203 amino acid sequences from different bacterial species encoding the enzyme retrieved. These were compiled into a multi-FASTA file and a local BLASTp search was carried out against the RW109 genome. This comparison loosely matched the RW109_04621 gene with a *Sulfolobus solfataricus* gene encoding the ribose-phosphate pyrophosphokinase enzyme (percentage identity= 38.03%, e-Value: 3e-07, BIT-Score = 47.4). This indicates that the RW109_04621 gene may be linked to the rxn00770 reaction. The RW109_04621 gene had already been found in the model and was associated with the adenine phosphoribosyltransferase rxn00139 reaction. Overall, the rxn00770 reaction was kept in the updated RW109 model and a note made about the more weakly associated RW109_04621 gene.

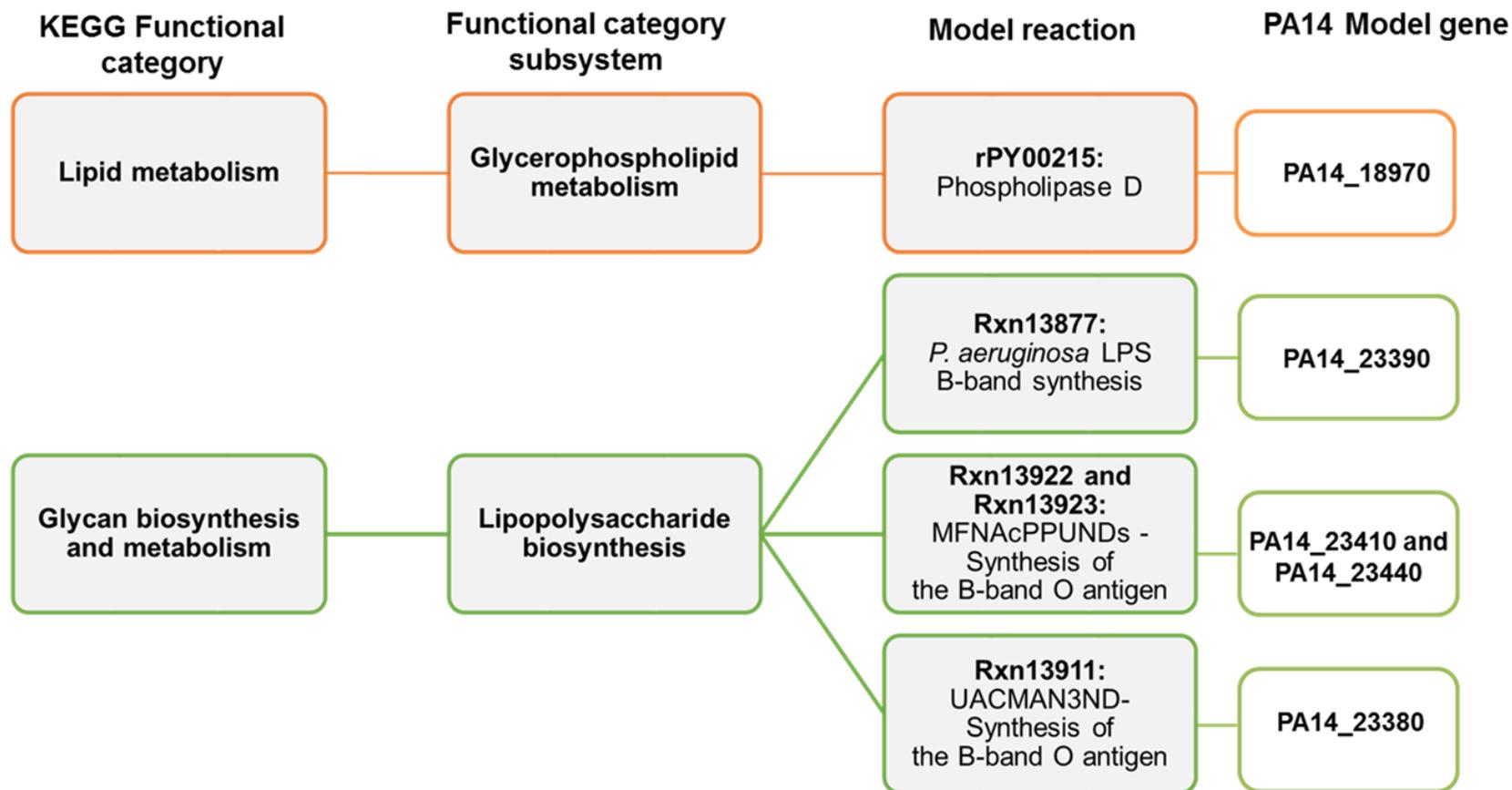


Figure 57. Flow diagram of reactions removed from the RW109 metabolic model

The flow diagram demonstrates the 5 PA14 model reactions which were removed from the model as they were found to be non-essential when the five PA14 genes associated with them were deleted. These genes were deleted as no similarities were found when compared to the RW109 strain during a BLASTp search. Within the flow diagrams the first box represents the KEGG functional category, the second the sub-system of the functional category, the third is the model reaction and the fourth column corresponds to the PA14 gene that was deleted from the model.

6.2.6 The overall key features of the RW109 metabolic model

The reconstructed model of *P. aeruginosa* RW109 accounts for the predicted function of 1,329 genes, 1,656 reactions and 1,465 metabolites (Table 35) (See relevant folder on the CD-ROM: RW109_Model_23March_UPtoDate.xlsx). The RW109 model had more predicted reactions, genes and metabolites when compared to the original PA14 metabolic model (Table 35), consistent with having a larger genome. The biomass production was predicted to be higher for the RW109 model when simulating growth in LB complete media, compared to simulations with the PA14 model, as indicated by the objective function value (Table 35). The locations of all the RW109 model genes on the genome were examined and 1,266 were from the main chromosome, 59 from plasmid 1 and 4 genes were from plasmid 2.

A diverse distribution of biological functions for the RW109 model reconstruction was observed when the reactions were grouped using KEGG functional categories; a total number of 19 KEGG categories were represented in the model (Figure 58). The lipid metabolism category had the greatest number of reactions associated with it (18.36%), followed by membrane transport (14.55%) and amino acid metabolism (13.41%). The KEGG functional categories, which had the least representation within the model, were replication and repair (0.06%), biosynthesis of amino acids (0.18%) and biosynthesis of other secondary metabolites (0.36%).

Table 35. The key features of the RW109 and PA14 metabolic models

Features	RW109	PA14
Number of genes	1,329 (18.2% of the genome=7,304 genes)	1,129 (18.9% of the genome= 5,978 genes)
Number of reactions	1,656 (1,223 metabolic, 241 membrane transport, 172 exchange and 20 unknown)	1,495 (1,070 metabolic, 238 membrane transport, 172 exchange and 15 unknown)
Number of metabolites	1,465	1,286
Objective function value*	17.45 hr ⁻¹	15.73 hr ⁻¹

Footnotes: *When simulating growth in LB complete media using the FBA optimised model solution

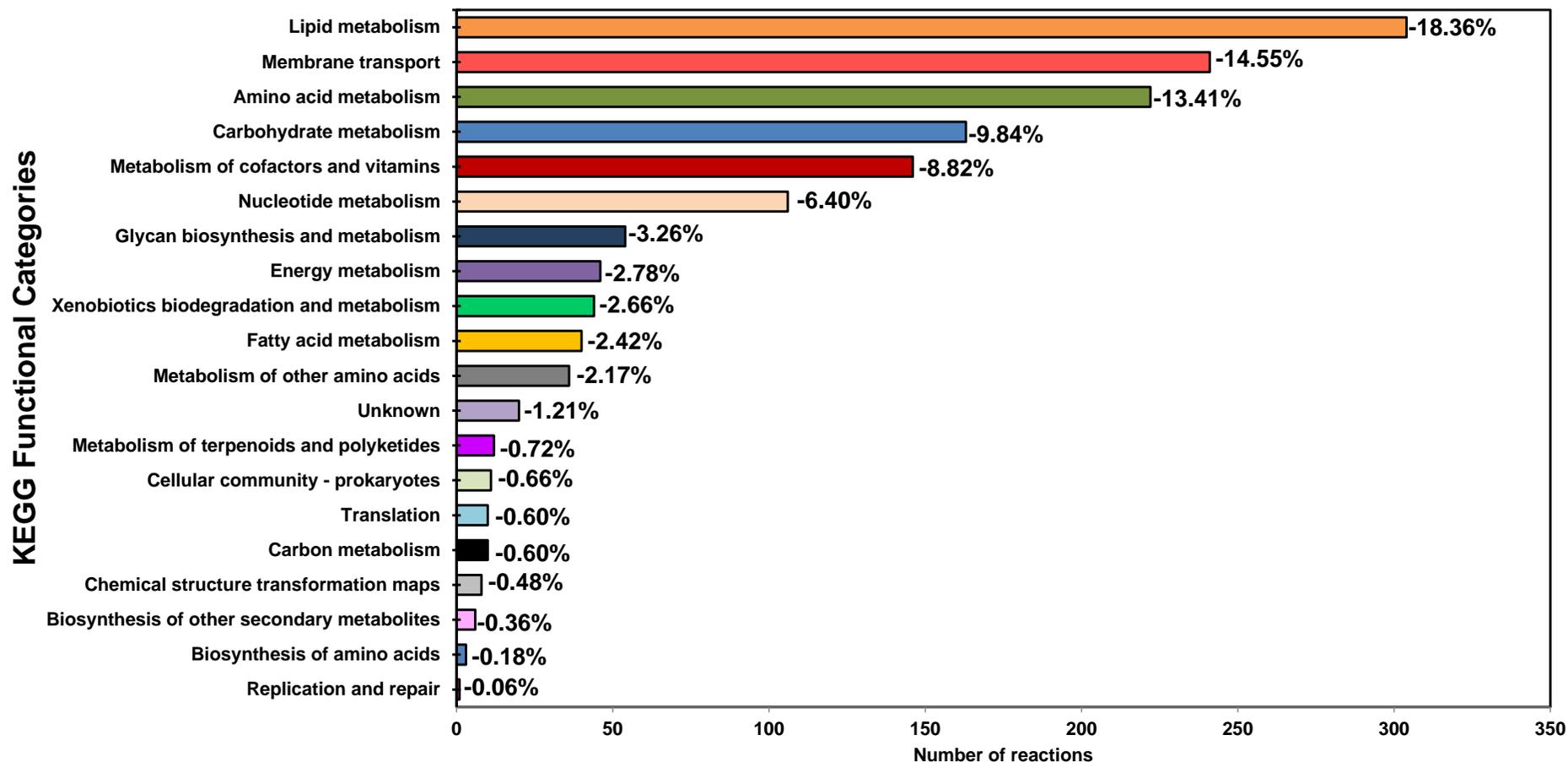


Figure 58. KEGG functional categories of the RW109 metabolic model reconstruction

The bar-chart shows the numbers of reactions within each KEGG functional category for the RW109 metabolic model's 1,656 the reactions. The functional categories were assigned via the KEGG database. The percentage representations of each functional category are indicated next to the bars and are based on the total number of reactions within the model. The lipid metabolism category had the greatest number of reactions associated with it, followed by membrane transport and amino acid metabolism.

6.2.7 Transcriptomic data integration to constrain the RW109 metabolic model and predict reaction essentiality during exposure to industry relevant conditions

To generate condition-specific models of RW109 metabolism, gene expression data from exposure of the strain to industry relevant conditions (Chapter 5; log₂-fold changes and adjusted p-values) were integrated following the computational pipeline as shown in Figure 59. The exposure conditions transcriptomically evaluated were (abbreviations used throughout this analysis are shown in brackets after the test condition number): TSB-Only (control condition), BIT at 20% of the MIC (test condition 1, BIT20), BIT at 50% of the MIC (test condition 2, BIT50), POE at 20% of the MIC (test condition 3, POE20), POE at 50% of the MIC (test condition 4, POE50), BIT and POE in combination at 20% of the MIC (test condition 5, BIT+POE), unpreserved laundry detergent only (test condition 6, LDOnly) and unpreserved laundry detergent in combination with BIT at 20% of the MIC (test condition 7, LD+BIT). The MADE algorithm (as described in detail in Section 2.8.5) was used to integrate statistically significant gene expression changes into the RW109 metabolic model. This method does not require a predefined threshold to be set and generates functioning condition-specific metabolic models, which accurately represent gene expression dynamics.

The outputs from the integration produced 8 separate models, which represented the functional gene states and metabolic adjustments of the control exposure condition (TSB-only, control model), and each of the 7 test exposure conditions (models 1-7). The MATLAB versions of these 8 models can be found in the relevant folder on the CD-ROM supplied with this thesis. The initial results from the integration were as follows (Table 36); the models, which represented the test conditions, each had different objective flux values, except for the BIT20 model, which had the same value as the control model. The greatest objective flux value of 7.46 hr⁻¹ was observed for the control model and the BIT20 model. The lowest objective flux value of 5.28 hr⁻¹ was observed for the BIT+POE model. The objective flux indicates the maximum amount of flux permitted to flow through the defined objective biomass function after the models were constrained with the gene expression data. When no constraints were placed on the RW109 base model, the objective flux value was 17.45 hr⁻¹ (Table 35). The number of functional genes (i.e., those not removed by the MADE algorithm) within each constrained model were also identified in Table 36.

The essentiality of the remaining functional genes for each model were predicted by performing computational knock-outs; if the resulting flux through the objective biomass function was < 0.0001 hr⁻¹ after performing the knock-out, that gene was predicted as essential. The reactions associated with the essential genes were also deleted and predicted to be essential if the flux through the objective biomass function was < 0.0001 hr⁻¹ after a deletion. The 140 genes and 165 reactions predicted to be essential in the control condition model were used as a comparison against those identified with the test conditions. For each test condition, the predicted essential reactions not found within the control condition were recognised (Table 36) and grouped according to their KEGG functional category assignment (Figure 60). This resulted in the classification of reactions for each test condition, which were projected to be essential for survival of RW109 within each industrial exposure condition. Spreadsheets of the model essential

reactions and genes for each test condition can be found in the relevant folder on the CD-ROM supplied with this thesis (TestConditions_Essential_Genes_and_Reactions.xlsx).

The BIT20 model had 1 gene and 1 reaction predicted to be essential, which was the least number of essential genes and reactions identified. The POE50 model had the highest quantities predicted with 38 essential genes and 53 essential reactions. KEGG categories were used to group the essential reactions of the test conditions to understand their biological functions and 9 categories were found to represent the essential reactions of the test condition models (Figure 60). Lipid metabolism had the greatest number of overall essential reactions with 49, followed by nucleotide metabolism with 42 and membrane transport with 39. The fatty acid metabolism and energy metabolism were the least represented functional categories for the test conditions (Figure 60). There were also KEGG functional categories of the RW109 model reconstruction, which had no representation within the predicted essential reactions. These included metabolism of other amino acids, glycan biosynthesis and metabolism, carbon metabolism, translation, biosynthesis of amino acids, biosynthesis of secondary metabolites and replication and repair.

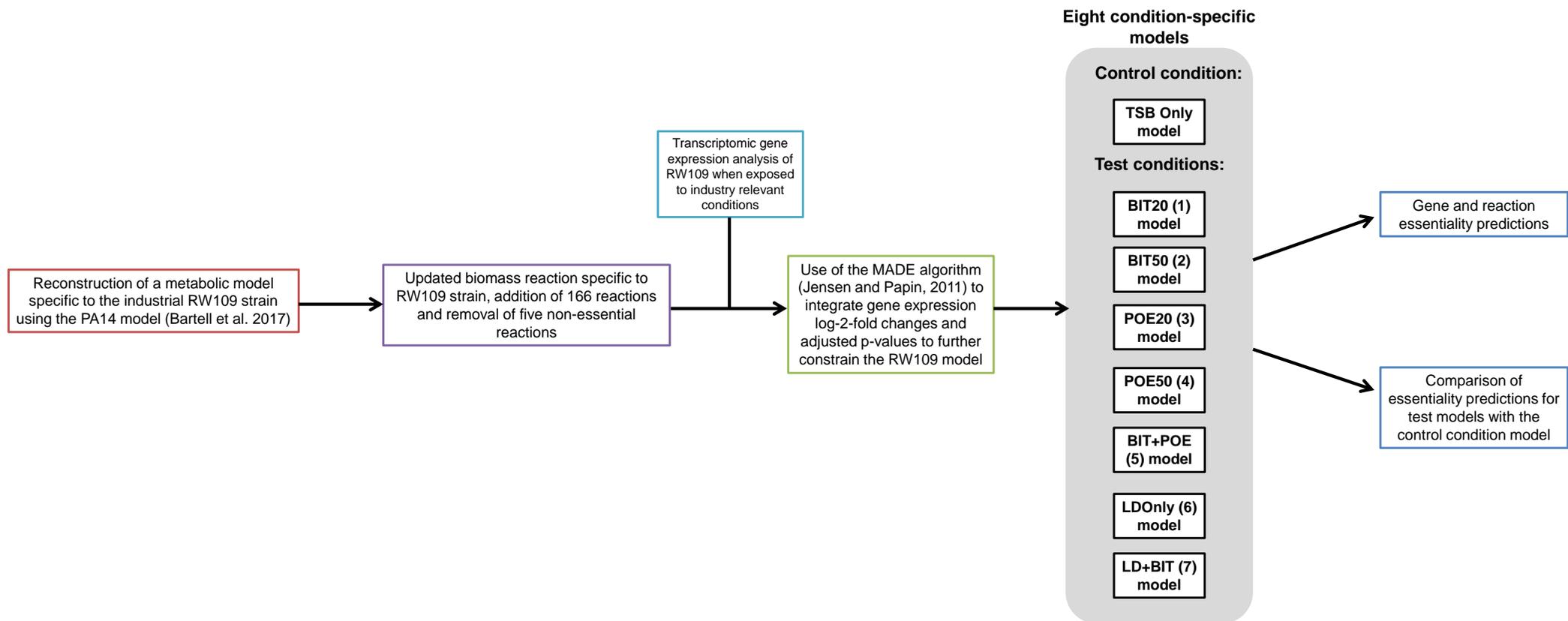


Figure 59. Computational pipeline of transcriptomic data integration into the RW109 metabolic model

The pipeline illustrates the steps used to integrate the transcriptomic data from Chapter 5 into the RW109 metabolic model reconstruction. The output from the MADE algorithm was a series of 8 condition-specific models which represent the functional gene states and metabolic adjustments of the control condition and the 7 test exposure conditions. The model names correspond to the exposure conditions: TSB-Only (control condition), BIT20(1) (BIT at 20% of the MIC), BIT50(2) (BIT at 50% of the MIC), POE20(3) (POE at 20% of the MIC), POE50(4) (POE at 50% of the MIC), BIT+POE(5) (BIT and POE in combination at 20% of the MIC), LDOnly(6) (unpreserved laundry detergent only) and LD+BIT(7) (unpreserved laundry detergent in combination with BIT at 20% of the MIC). These models were then used to predict the essentiality of genes and reactions when RW109 was exposed to industry relevant conditions.

Table 36. Results of transcriptomic gene expression data integration into the RW109 model reconstruction

Condition	Objective flux(hr-1)	Number of functional genes	Number of essential genes	Number of essential reactions
Control: TSB only	7.46	739	140	165
Test conditions:				
(1) BIT20	7.46	742	1*	1*
(2) BIT50	6.51	723	8*	10*
(3) POE20	5.74	563	35*	45*
(4) POE50	5.89	552	38*	53*
(5) BIT+POE	5.28	590	33*	37*
(6) LD Only	5.57	621	26*	34*
(7) LD+BIT	5.59	629	29*	47*

Footnotes: The * in columns 5 and 6 indicates the number of essential genes and reactions which were not identified in the control condition.

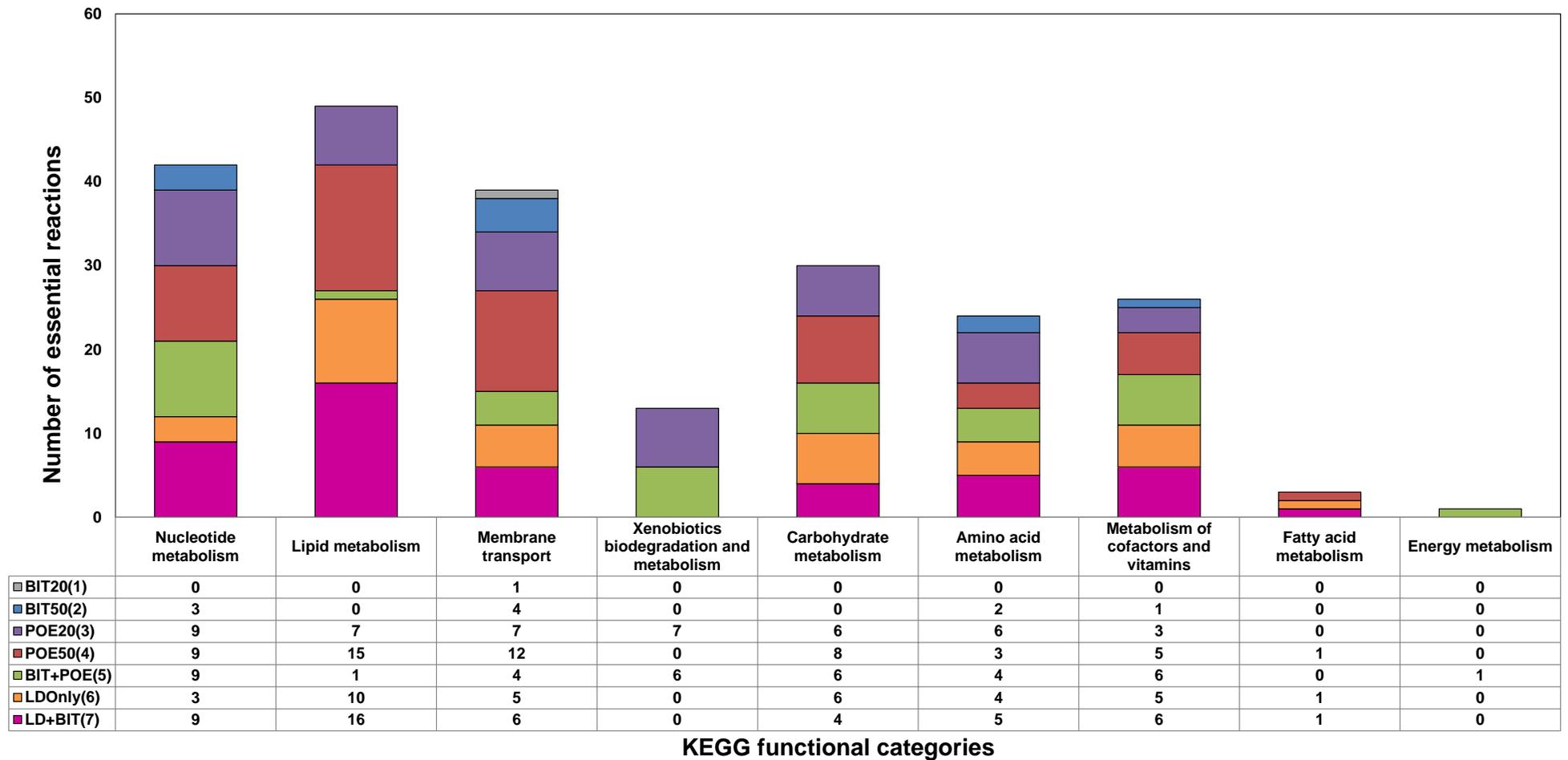


Figure 60. KEGG functional categories of the predicted essential reactions for each RNA-Seq exposure condition

The stacked bar chart demonstrates the number of predicted essential reactions within each test condition grouped via the KEGG functional category. BIT20(1) is shown in grey, BIT50 (2) is in blue, POE20 (3) is in purple, POE50 (4) is in red, BIT+POE (5) is shown in green, LDOnly (6) is in orange and LD+BIT (7) is shown in pink. The lipid metabolism functional category had the greatest number of genes associated with predicted essential reactions.

6.2.7.1 Predicted essential reactions involved in lipid metabolism

The greatest numbers of predicted essential reactions, 49, were associated with lipid metabolism when compared to the other KEGG functional categories (Figure 60). All test exposure conditions except for BIT20 and BIT50, were predicted to have essential reactions within this category; LD+BIT and POE50 had the most with 16 and 15 predicted respectively. The groupings between the exposure conditions of the lipid metabolism essential reactions are shown in Figure 61 and described in detail in Table 37. A notable observation were the 5 predicted reactions only identified as essential after exposure to POE50 (Figure 61); these reactions were all found to be associated with CDP-diacylglycerol synthetase (CDP-DAG-synthetase) and were encoded by the same gene RW109_02164 within the model (Table 37). Six cardiolipin synthase reactions were predicted to be essential when RW109 was exposed to POE20, POE50 and LD+BIT (Figure 61 and Table 37); these reactions were all associated with the essential model gene RW109_03891. During exposure to LDOnly and LD+BIT, 6 reactions were predicted as essential which were all linked to (3R)-3-Hydroxypalmitoyl-[acyl-carrier-protein]:NADP⁺ oxidoreductase and were encoded for by the RW109_02915 FabG model gene (Table 37). Two reactions related to glycerolipid metabolism were also found to only be essential during exposure to POE20 (rxn00763) and BIT+POE (rxn01102).

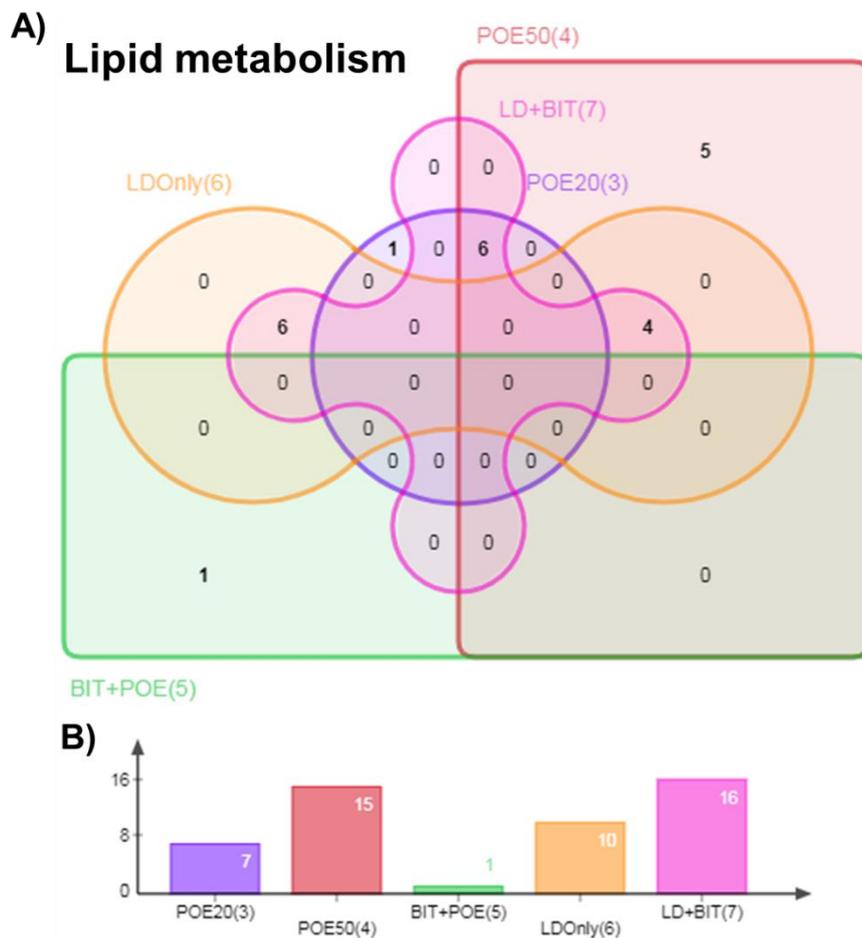


Figure 61. Grouping of the lipid metabolism predicted essential reactions

The Edwards-Venn diagram (A) represents the number of lipid metabolism essential reactions (numbers in black) which either overlapped or were found only in the test conditions POE20(3) in purple, POE50(4) in red, BITandPOE(5) in green, LDOnly(6) in orange and LD and BIT(7) in pink. The bar chart (B) is used to clearly show the numbers of essential reactions within each of the test conditions which are coloured the same as in A).

Table 37. Predicted essential reactions for the lipid metabolism KEGG functional category

Test exposure condition	Essential Reaction	Reaction Name	Reaction Function	Reaction formula	EC number	Essential genes
POE20(3)	rxn00763	Alcohol dehydrogenase (glycerol)	Glycerolipid metabolism	NAD + Glycerol <=> NADH + H+ + D-Glyceraldehyde	1.1.1.1,1.1.1.2,1.1.1.21,1.1.1.72	RW109_03810
POE50(4)	rPY00163	CDP-diacylglycerol synthetase (n-C16:0 16:1)	Fatty acid biosynthesis	H+ + pa160161 + CTP -> PPI + CDP-pa160161	2.7.7.41	RW109_02164
	rPY00168	CDP-diacylglycerol synthetase (n-C18:0 18:1)	Fatty acid biosynthesis	H+ + CTP + pa180181 + PPI + CDP-pa180181	2.7.7.41	RW109_02164
	rxn08310	CDP-diacylglycerol synthetase (n-C16:1)	Glycerolipid and Glycerophospholipid Metabolism in Bacteria	CTP + 1,2-dihexadec-9-enoyl-sn-glycerol_3-phosphate <=> PPI + CDP-1,2-dihexadec-9-enoylglycerol	2.7.7.41	RW109_02164
	rxn08311	CDP-diacylglycerol synthetase (n-C18:0)	Glycerolipid and Glycerophospholipid Metabolism in Bacteria	CTP + 1,2-dioctadecanoyl-sn-glycerol_3-phosphate <=> PPI + CDP-1,2-dioctadecanoylglycerol	2.7.7.41	RW109_02164
	rxn08312	CDP-diacylglycerol synthetase (n-C18:1)	Glycerolipid and Glycerophospholipid Metabolism in Bacteria	CTP + 1,2-dioctadec-11-enoyl-sn-glycerol_3-phosphate <=> PPI + CDP-1,2-dioctadec-11-enoylglycerol	2.7.7.41	RW109_02164
BIT+POE(5)	rxn01102	glycerate kinase	Glycerolipid metabolism	ATP + Glycerate <=> ADP + 3-Phosphoglycerate	2.7.1.31	RW109_05126
LDOOnly(6) & LD+BIT(7)	rxn05336	(3R)-3-Hydroxypalmitoyl-[acyl-carrier-protein]:NADP+ oxidoreductase	Fatty acid biosynthesis	NADP + R-3-hydroxypalmitoyl-acyl-carrierprotein- <=> H+ + NADPH + 3-oxohexadecanoyl-acp	1.1.1.100,2.3.1.85,2.3.1.86	RW109_02915
	rxn05337	(3R)-3-Hydroxyhexanoyl-[acyl-carrier-protein]:NADP+ oxidoreductase	Fatty acid biosynthesis	NADP + D-3-Hydroxyhexanoyl-[acp] <=> H+ + NADPH + 3-Oxohexanoyl-[acp]	1.1.1.100,2.3.1.85,2.3.1.86	RW109_02915
	rxn05339	(3R)-3-Hydroxybutanoyl-[acyl-carrier protein]:NADP+ oxidoreductase	Fatty acid biosynthesis	NADP + (R)-3-Hydroxybutanoyl-[acyl-carrier protein] <=> H+ + NADPH + Acetoacetyl-ACP	1.1.1.100,2.3.1.85,2.3.1.86	RW109_02915
	rxn05340	3R)-3-Hydroxydodecanoyl-[acyl-carrier-protein]:NADP+ oxidoreductase	Fatty acid biosynthesis	NADP + D-3-Hydroxydodecanoyl-[acp] <=> H+ + NADPH + 3-oxododecanoyl-acp	1.1.1.100,2.3.1.85,2.3.1.86	RW109_02915
	rxn05341	(3R)-3-Hydroxyoctanoyl-[acyl-carrier-protein]:NADP+ oxidoreductase	Fatty acid biosynthesis	NADP + (R)-3-Hydroxyoctanoyl-[acyl-carrier protein] <=> H+ + NADPH + 3-oxooctanoyl-acp	1.1.1.100,2.3.1.85,2.3.1.86	RW109_02915
	rxn05342	3R)-3-Hydroxytetradecanoyl-[acyl-carrier-protein]:NADP+ oxidoreductase	Fatty acid biosynthesis	NADP + HMA <=> H+ + NADPH + 3-oxotetradecanoyl-acp	1.1.1.100,2.3.1.85,2.3.1.86	RW109_02915
POE20(3) & POE50(4) & LD+BIT(7)	rPY00183	Cardiolipin synthase (n-C18:0 18:1)	Glycerophospholipid metabolism	2 pg180181 -> Glycerol + clpn180181	2.7.8.-	RW109_03891
	rPY00184	Cardiolipin synthase (n-C16:0 16:1)	Glycerophospholipid metabolism	2 pg160161 -> Glycerol + clpn160161	2.7.8.-	RW109_03891
	rxn08229	Cardiolipin synthase (n-C16:0)	Glycerolipid and Glycerophospholipid Metabolism in Bacteria	2 Phosphatidylglycerol_dihexadecanoyl <=> Glycerol + clpn160	2.7.8.-	RW109_03891
	rxn08230	Cardiolipin synthase (n-C16:1)	Glycerolipid and Glycerophospholipid Metabolism in Bacteria	2 Phosphatidylglycerol_dihexadec-9-enoyl <=> Glycerol + clpn161	2.7.8.-	RW109_03891
	rxn08231	Cardiolipin synthase (n-C18:0)	Glycerolipid and Glycerophospholipid Metabolism in Bacteria	2 Phosphatidylglycerol_dioctadecanoyl <=> Glycerol + clpn180	2.7.8.-	RW109_03891
rxn08232	Cardiolipin synthase (n-C18:1)	Glycerolipid and Glycerophospholipid Metabolism in Bacteria	2 Phosphatidylglycerol_dioctadec-11-enoyl <=> Glycerol + clpn181	2.7.8.-	RW109_03891	
POE50(4) & LDOOnly(6) & LD+BIT(7)	rPY00162	PA160161S	Fatty acid biosynthesis	Hexadecanoyl-ACP + 1-hexadecanoyl-sn-glycerol_3-phosphate -> ACP + pa160161	2.3.1.51	RW109_00595(4), RW109_03471(6 and 7)
	rPY00167	PA180181S	Fatty acid biosynthesis	Octadecanoyl-ACP + 1-octadecanoyl-sn-glycerol_3-phosphate -> ACP + pa180181	2.3.1.51	RW109_00595(4), RW109_03471(6 and 7)
	rxn08087	1-hexadec-7-enoyl-sn-glycerol 3-phosphate O-acyltransferase (n-C16:1)	Phosphate metabolism Glycerolipid and Glycerophospholipid Metabolism in Bacteria	Hexadecanoyl-ACP + 1-hexadec-9-enoyl-sn-glycerol_3-phosphate -> ACP + 1,2-dihexadec-9-enoyl-sn-glycerol_3-phosphate	2.3.1.51	RW109_00595(4), RW109_03471(6 and 7)

Footnotes:

- 1) The test exposure conditions in column 1 are indicated by their abbreviation and test condition number e.g. POE20(3) is test condition 3: POE at 20% of the MIC.
- 2) The genes in the last column without numbers in brackets demonstrate a gene which was essential in all the test conditions within the corresponding row in column, e.g. RW109_02915 was essential for both LDOOnly(6) and LD+BIT(7).
- 3) The numbers in brackets next to the genes in the last column indicates the test exposure condition where the gene was identified as essential e.g. RW109_00595(4) was found to be only essential in test condition POE50(4).

6.2.7.2 Predicted essential reactions involved in nucleotide metabolism

Two groupings of essential reactions within the nucleotide metabolism category were observed (Table 38). All exposure conditions were predicted to have essential reactions within this KEGG functional category except for BIT20. The greatest numbers of nucleotide metabolism essential reactions were predicted in POE20, POE50 and BIT+POE with 9 predicted in each test condition. Three essential reactions identified in the 6 exposure conditions were all involved in the salvage pathway nucleoside-diphosphate kinase (Ndk) and encoded by the RW109_01980 model gene (Table 38). Six essential reactions were predicted when RW109 was exposed to POE20, POE50, BIT+POE and LD+BIT (not identified with BIT50); two were involved in the salvage pathways cytidylate kinase (rxn00364) and another with the Ndk enzyme (rxn00409). The other 4 reactions had a function in purine and pyrimidine metabolism and biosynthesis and 3 were identified as reactions which were additional to the RW109 model (Section 6.2.3); rxn05816, rxn05817 and rxn05818.

Table 38. Predicted essential reactions for the nucleotide metabolism KEGG functional category

Test exposure condition	Essential Reaction	Reaction name	Reaction function	Reaction formula	EC number	Essential genes
POE20(3) & POE50(4) & BIT+POE(5) & LD+BIT(7)	rxn00364	Cytidylate kinase (CMP)	Salvage Pathways	ATP + CMP <=> ADP + CDP	2.7.4.14	RW109_02660
	rxn00409	Nucleoside-diphosphate kinase (ATP:CDP)	Salvage Pathways	ATP + CDP <=> ADP + CTP	2.7.4.6	RW109_01980
	rxn00917	GMP synthase (glutamine-hydrolysing)	Purine and Pyrimidine Biosynthesis	H ₂ O + ATP + L-Glutamine + XMP -> H ⁺ + L-Glutamate + P _{Pi} + AMP + GMP	6.3.5.2	RW109_00448
	rxn05816	Deoxyadenosine 5'-triphosphate:DNA deoxynucleotidyltransferase (DNA-directed)	Purine metabolism	dATP <=> P _{Pi}	2.7.7.7	RW109_00252 (3,4,5), RW109_00592 (7)
	rxn05817	Deoxyguanosine 5'-triphosphate:DNA deoxynucleotidyltransferase (DNA-directed)	Purine metabolism	dGTP <=> P _{Pi}	2.7.7.7	RW109_00252 (3,4,5), RW109_00592 (7)
	rxn05818	Deoxycytidine triphosphate:DNA deoxynucleotidyltransferase (DNA-directed)	Pyrimidine metabolism	dCTP <=> P _{Pi}	2.7.7.7	RW109_00252 (3,4,5), RW109_00592 (7)
BIT50(2) & POE20(3) & POE50(4) & BIT+POE(5) & LDOly(6) & LD+BIT(7)	rxn00839	Nucleoside-diphosphate kinase (ATP:dADP)	Salvage Pathways	ATP + dADP <=> ADP + dATP	2.7.4.6	RW109_01980
	rxn01353	Nucleoside-diphosphate kinase (ATP:dGDP)	Salvage Pathways	ATP + dGDP <=> ADP + dGTP	2.7.4.6	RW109_01980
	rxn01673	Nucleoside-diphosphate kinase (ATP:dCDP)	Salvage Pathways	ATP + dCDP <=> ADP + dCTP	2.7.4.6	RW109_01980

Footnotes:

- 1) The test exposure conditions in column 1 are indicated by their abbreviation and test condition number e.g. POE20(3) is test condition 3: POE at 20% of the MIC.
- 2) The genes in the last column without numbers in brackets demonstrate a gene which was essential in all the test conditions within the corresponding row in column 1.
- 3) The numbers in brackets next to the genes in the last column indicates the test exposure condition where the gene was identified as essential.

6.2.7.3 Predicted essential reactions involved in membrane transport

Numerous groupings of the essential reactions involved in membrane transport metabolism were noted. To enable display of the data, the analysis was divided into comparisons of the essential reactions between the test conditions: (i) BIT20(1), BIT50(2), POE20(3), POE50(4) and BIT+POE(5) (Figure 62 and Table 39) and (ii) BIT20(1), BIT50(2), LDOnly(6) and LD+BIT(7) (Figure 63 and Table 40). Test conditions 1, 2, 3, 4 and 5 were analysed together as these conditions represented the preservatives BIT and POE both individually and in combination; POE50 had the greatest number of essential reactions with twelve predicted. The test conditions 1, 2, 6 and 7 were examined collectively as these conditions represented the preservative BIT and laundry detergent separately and in combination; the greatest number of essential reactions identified was during LD+BIT exposure with six predicted.

Eight groupings of the essential reactions were observed for test conditions BIT20(1), BIT50(2), POE20(3), POE50(4) and BIT+POE(5) (Figure 62 and Table 39). An L-phenylalanine transport reaction (rxn05306) was only predicted to be essential with POE20 and 5 membrane transport reactions with various functions were identified during exposure to POE50. Three reactions were predicted as essential during exposure to the test conditions POE20 and POE50, and these had functions in transporting urea, L-histidine and L-proline (rxn05164, rxn05165 and rxn09653). A FeoB iron (II) transporter reaction (rxn10963) encoded by the RW109_05695 model gene was identified during exposure to the test conditions POE50 and BIT+POE and an L-lysine ABC system transport reaction (rxn05151) was predicted as essential in both POE50 and BIT+POE and was encoded by RW109_02958. An L-methionine transport via ABC system (rxn05219) was classified as essential in all of the test conditions and was the only reaction predicted for the BIT20 exposure condition. With BIT20 and BIT50, the essentiality of rxn05219 was associated with the RW109_07122 model gene and in conditions POE20, POE50 and BIT+POE, 3 genes RW109_03660, RW109_03661 and RW109_03662 were linked to the reaction.

Seven groupings of essential reactions were recognised when the test conditions BIT20(1), BIT50(2), LDOnly(6) and LD+BIT(7) were compared (Figure 63 and Table 40). A L-proline transport reaction (rxn05638) was predicted as essential only during exposure to LDOnly, and 2 reactions involved in the transport of L-arginine and acetate were only classified as essential with LD+BIT (rxn05154 and rxn05488). Exposure to the test conditions LDOnly and LD+BIT classified the essentiality of 2 reactions, which transported L-histidine and adenosine (rxn05164 and rxn05318). The L-methionine transport via ABC system (rxn05219) was predicted as essential during exposure to BIT20, BIT50 and LD+BIT and the FeoB iron (II) transporter was essential during exposure to BIT50, LDOnly and LD+BIT.

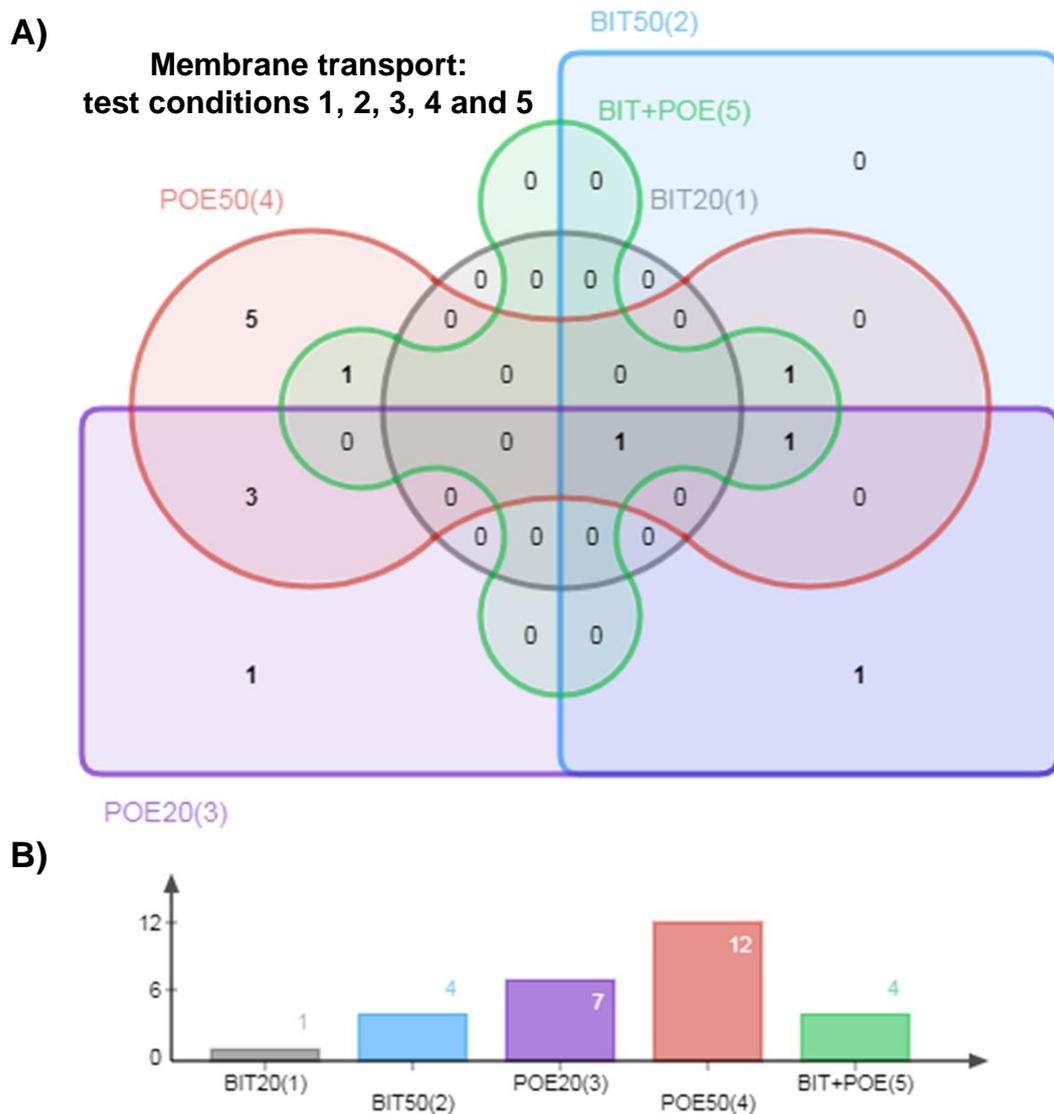


Figure 62. Grouping of the membrane transport predicted essential reactions for the test conditions 1, 2, 3, 4 and 5

The Edwards-Venn diagram (A) represents the number of membrane transport essential reactions (numbers in black) which either overlapped or were found only in the test conditions BIT20(1) in grey, BIT50(2) in blue, POE20(3) in purple, POE50(4) in red and BITandPOE(5) in green. The bar chart (B) is used to clearly show the numbers of essential reactions within each of the test conditions which are coloured the same as in A).

Table 39. Predicted essential reactions for the membrane transport category for test conditions 1, 2, 3, 4 and 5

Test exposure condition	Essential Reaction	Reaction function	Reaction formula	EC number	Essential genes
POE20(3)	rxn05306	L-phenylalanine transport in/out via proton symport	H+ + L-Phenylalanine <=> H+ + L-Phenylalanine	TC-2.A.3.1,2.A.3.1	RW109_02881
POE50(4)	rJB00225	D-Fructose transport via PEP:Pyr PTS	Phosphoenolpyruvate + D-Fructose -> Pyruvate + D-fructose-6-phosphate	-	RW109_02260 RW109_02259 RW109_02258
	rxn05161	L-leucine transport via ABC system	H2O + ATP + L-Leucine -> H+ + ADP + Phosphate + L-Leucine	-	RW109_05103 RW109_05104
	rxn05168	L-valine transport via ABC system	H2O + ATP + L-Valine -> H+ + ADP + Phosphate + L-Valine	-	RW109_05103 RW109_05104
	rxn05179	L-isoleucine transport via ABC system	H2O + ATP + L-Isoleucine -> H+ + ADP + Phosphate + L-Isoleucine	-	RW109_06504
	rxn05560	D-fructose transport via PEP:Pyr PTS	Phosphoenolpyruvate + D-Fructose <=> Pyruvate + D-fructose-1-phosphate	-	RW109_02260 RW109_02258
BIT50(2) & POE20(3)	rxn10963	FeoB iron (II) transporter	Fe2+ <=> Fe2+	-	RW109_05695
POE20(3) & POE50(4)	rxn05164	L-histidine transport via ABC system	H2O + ATP + L-Histidine -> H+ + ADP + Phosphate + L-Histidine	-	RW109_02958
	rxn05165	L-proline transport via ABC system	H2O + ATP + L-Proline -> H+ + ADP + Phosphate + L-Proline	-	RW109_01824
	rxn09653	urea reversible transport via proton symport (2 H+)	2 H+ + Urea <=> 2 H+ + Urea	-	RW109_04753
POE50(4) & BIT+POE(5)	rxn05151	L-lysine transport via ABC system	cpd00001 + cpd00002 + cpd00039[e] -> cpd00067 + cpd00008 + cpd00009 + cpd00039	-	RW109_02958
BIT50(2) & POE50(4) & BIT+POE(5)	rxn05663	L-tryptophan transport in via proton symport	H+ + L-Tryptophan <=> H+ + L-Tryptophan	TC-2.A.3.1,2.A.3.1	RW109_05322(2 and 5) RW109_02881(4)
BIT50(2) & POE20(3) & POE50(4) & BIT+POE(5)	rxn05169	L-threonine transport via ABC system	H2O + ATP + L-Threonine -> H+ + ADP + Phosphate + L-Threonine	-	RW109_06501(2,3 and 5) RW109_05103(4), RW109_05104(4)
BIT20(1) & BIT50(2) & POE20(3) & POE50(4) & BIT+POE(5)	rxn05219	L-methionine transport via ABC system	H2O + ATP + L-Methionine -> H+ + ADP + Phosphate + L-Methionine	-	RW109_07122(1 and 2), RW109_03660(3, 4 and 5), RW109_03661(3, 4, and 5), RW109_03662(3, 4 and 5)

Footnotes:

- 1) The test exposure conditions in column 1 are indicated by their abbreviation and test condition number e.g. POE20(3) is test condition 3: POE at 20% of the MIC.
- 2) The genes in the last column without numbers in brackets demonstrate a gene which was essential in all the test conditions within the corresponding row in column 1.
- 3) The numbers in brackets next to the genes in the last column indicates the test exposure condition where the gene was identified as essential.

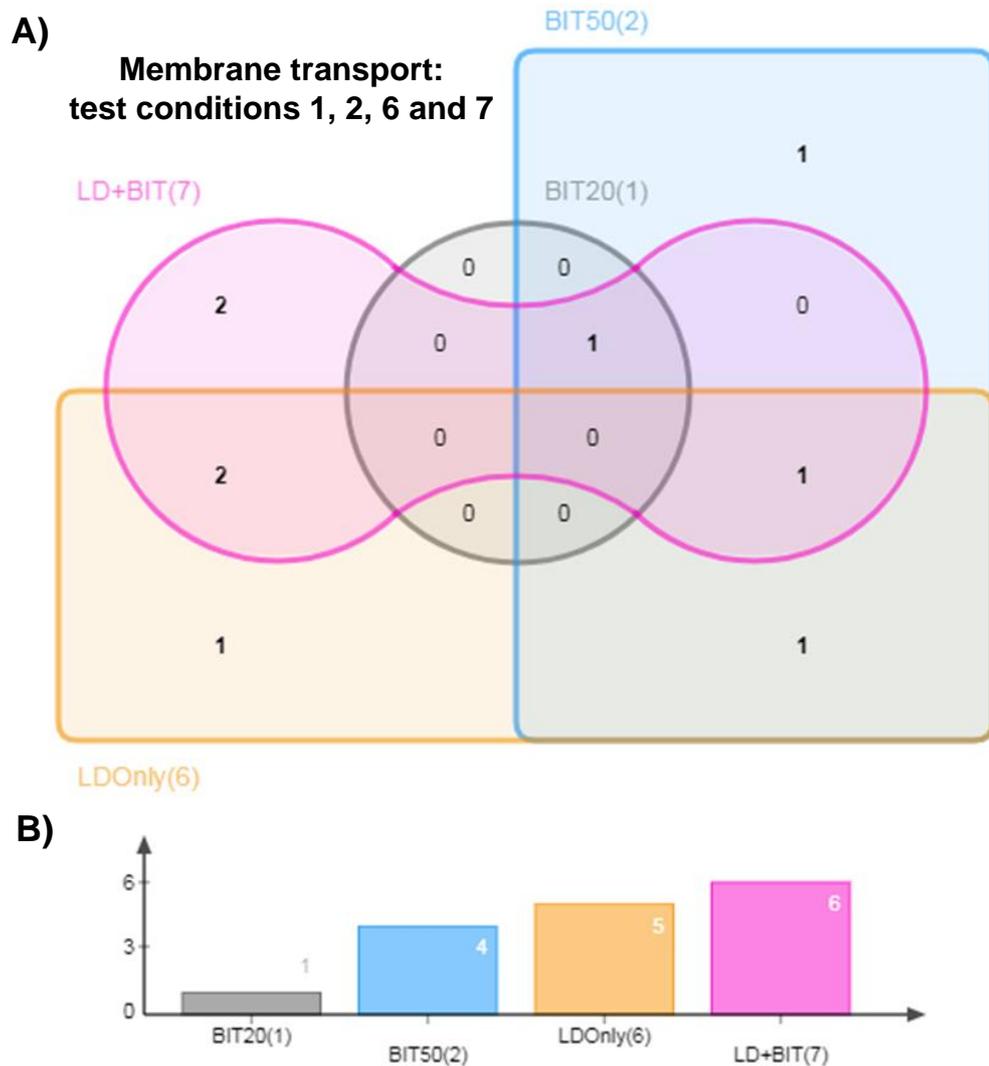


Figure 63. Grouping of the membrane transport predicted essential reactions for test conditions 1, 2, 6 and 7

The Edwards-Venn diagram (A) represents the number of membrane transport essential reactions (numbers in black) which either overlapped or were found only in the test conditions BIT20(1) in grey, BIT50(2) in blue, LDOnly(6) in orange and LDandBIT(7) in pink. The bar chart (B) is used to clearly show the numbers of essential reactions within each of the test conditions which are coloured the same as in A).

Table 40. Predicted essential reactions for the membrane transport category for test conditions 1, 2, 6 and 7

Test exposure condition	Essential Reaction	Reaction function	Reaction formula	EC number	Essential genes
BIT50(2)	rxn05663	L-tryptophan transport in via proton symport	H+ + L-Tryptophan <=> H+ + L-Tryptophan	TC-2.A.3.1,2.A.3.1	RW109_02881
LDOnly(6)	rxn05638	L-proline transport in via proton symport	H+ + L-Proline -> H+ + L-Proline	-	RW109_06983
LD+BIT(7)	rxn05154	L-arginine transport via ABC system	H2O + ATP + L-Arginine -> H+ + ADP + Phosphate + L-Arginine	-	RW109_05296
	rxn05488	Acetate transport in/out via proton symport	H+ + Acetate <=> H+ + Acetate	-	RW109_03816
BIT50(2) & LDOnly(6)	rxn05169	L-threonine transport via ABC system	H2O + ATP + L-Threonine -> H+ + ADP + Phosphate + L-Threonine	-	RW109_06501(2), RW109_05105(6), RW109_05106(6), RW109_05107(6), RW109_06504(6)
LDOnly(6) & LD+BIT(7)	rxn05164	L-histidine transport via ABC system	H2O + ATP + L-Histidine -> H+ + ADP + Phosphate + L-Histidine	-	RW109_05296
	rxn05318	ADNt2r	H+ + Adenosine <=> H+ + Adenosine	-	RW109_04831
BIT20(1) & BIT50(2) & LD+BIT(7)	rxn05219	L-methionine transport via ABC system	H2O + ATP + L-Methionine -> H+ + ADP + Phosphate + L-Methionine	-	RW109_071229(1 and 2), RW109_03660(7), RW109_03661(7), RW109_03662(7)
BIT50(2) & LDOnly(6) & LD+BIT(7)	rxn10963	FeoB iron (II) transporter	Fe2+ <=> Fe2+	-	RW109_05695

Footnotes:

- 1) The test exposure conditions in column 1 are indicated by their abbreviation and test condition number e.g. POE20(3) is test condition 3: POE at 20% of the MIC.
- 2) The genes in the last column without numbers in brackets demonstrate a gene which was essential in all the test conditions within the corresponding row in column 1.
- 3) The numbers in brackets next to the genes in the last column indicates the test exposure condition where the gene was identified as essential.

6.2.7.4 Predicted essential reactions involved in xenobiotics biodegradation and metabolism

Essential reactions involved in xenobiotics biodegradation and metabolism were only found within the test conditions POE20(3) and BIT+POE(5). The most essential reactions were identified during exposure to POE20, with seven predicted (Figure 60). Table 41 illustrates the groupings of these reactions and all were found to have a function in benzoate degradation via hydroxylation. A reaction involved in 4-carboxymuconolactone decarboxylase (rxn02483) was predicted only to be essential in the POE20(3) test condition. Six reactions with various enzymes involved in benzoate degradation were identified as essential in both the POE20(3) and BIT+POE(5) test conditions (Table 41).

Table 41. Predicted essential reactions for the xenobiotics biodegradation and metabolism category for test conditions 3 and 5

Test exposure condition	Essential Reaction	Reaction name	Reaction function	Reaction formula	EC number	Essential genes
POE20(3)	rxn02483	4-carboxymuconolactone decarboxylase	Benzoate Degradation	H+ + 4-Carboxymuconolactone -> CO2 + 3-oxoadipate-enol-lactone	4.1.1.44	RW109_05826
POE20(3) & BIT+POE(5)	rxn00598	3-oxoadipyl-CoA thiolase	Benzoate Degradation	Acetyl-CoA + Succinyl-CoA <=> CoA + 3-Oxoadipyl-CoA	2.3.1.16,2.3.1.174	RW109_02226(3), RW109_00825(5)
	rxn00962	4-hydroxybenzoate 3-monooxygenase	Benzoate Degradation	O2 + H+ + NADPH + 4-Hydroxybenzoate -> H2O + NADP + Protocatechuate	1.14.13.2, 1.14.13.33	RW109_00844
	rxn01192	Protocatechuate 3,4-dioxygenase	Benzoate Degradation	O2 + Protocatechuate -> 2 H+ + 3-Carboxy-cis,cis-muconate	1.13.11.3	RW109_00756
	rxn02143	3-oxoadipate CoA-transferase	Benzoate Degradation	Succinyl-CoA + 3-Oxoadipate <=> 3-Oxoadipyl-CoA + Succinate	2.8.3.6	RW109_04051
	rxn02144	3-oxoadipate enol-lactonase	Benzoate Degradation	H2O + 3-oxoadipate-enol-lactone -> H+ + 3-Oxoadipate	3.1.1.24	RW109_01084
	rxn02369	3-carboxy-cis,cis-muconate cycloisomerase	Benzoate Degradation	4-Carboxymuconolactone <=> H+ + 3-Carboxy-cis,cis-muconate	5.5.1.2	RW109_00827

Footnotes:

- 1) The test exposure conditions in column 1 are indicated by their abbreviation and test condition number e.g. POE20(3) is test condition 3: POE at 20% of the MIC.
- 2) The genes in the last column without numbers in brackets demonstrate a gene which was essential in all the test conditions within the corresponding row in column 1.
- 3) The numbers in brackets next to the genes in the last column indicates the test exposure condition where the gene was identified as essential.

6.2.7.5 Predicted essential reactions involved in carbohydrate metabolism

All conditions, except test conditions BIT20 and BIT50, were found to have predicted essential reactions involved in carbohydrate metabolism (Figure 60). The test exposure condition POE50 was found to have the greatest number of reactions with 8 predicted as essential. The essential reactions were compared and 7 test conditions groupings were observed (Figure 64 and Table 42). A reaction involved in butanoate metabolism (rxn00988) was found to only be essential during exposure to POE20. Two glycolysis/gluconeogenesis/Entner-Doudoroff reactions (rxn00216 and rxn00786) and one alternate carbon metabolism reaction (rxn01492) had predicted essentiality with only the POE50 test condition. The essentialities of 5 reactions with functions in glycolysis/gluconeogenesis were predicted during exposure to BIT+POE (Table 42). Four of these pathways were also predicted to be involved in alternative Entner-Doudoroff pathways. Two citrate (TCA) cycle reactions were identified with the LDOnly test condition (Table 42) and one of these reactions (rxn00265) was specifically an additional RW109 model reaction (Section 6.2.3). A leucine degradation reaction (rxn00290) was predicted to be essential during exposure to the test conditions POE20, POE50 and BIT+POE. Three pentose phosphate cycle reactions were identified in exposure to POE20, POE50, LDOnly and LD+BIT. The RW109_02629 gene involved in the phosphogluconate dehydratase reaction (rxn01477) was previously characterised in two transposon mutants that reduced their light emission in response to BIT (BIT negative) in Chapter 4 (Table 17). A common reaction found to be essential in the 5 test conditions was a glyoxylate and dicarboxylate metabolism reaction, which breaks down glycoaldehyde to glycolate (rAB00001); this reaction was associated with the same RW109_00559 *aldA* gene in all the conditions.

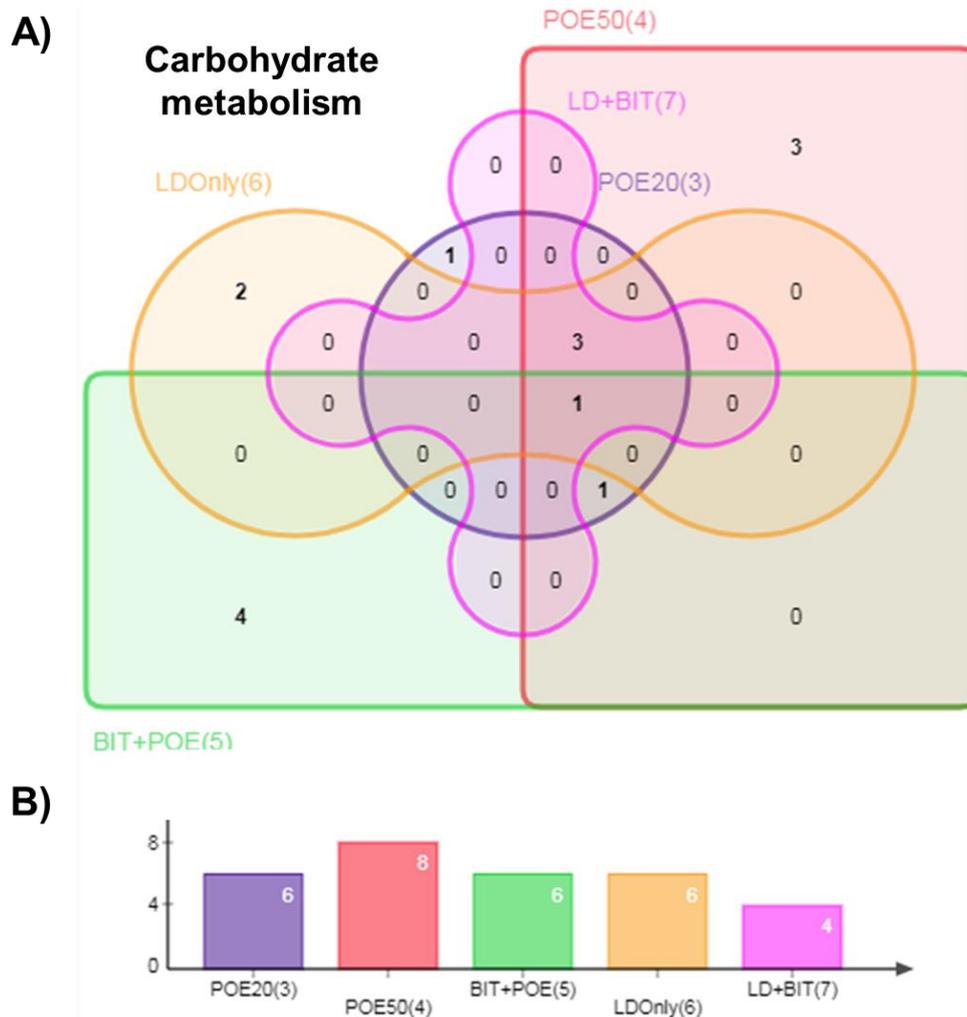


Figure 64. Grouping of the carbohydrate metabolism predicted essential reactions

The Edwards-Venn diagram (A) represents the number of carbohydrate metabolism essential reactions (numbers in black) which either overlapped or were found only in the test conditions POE20(3) in purple, POE50(4) in red, BIT and POE(5) in green, LD Only(6) in orange and LD and BIT(7) in pink. The bar chart (B) is used to clearly show the numbers of essential reactions within each of the test conditions which are coloured the same as in A).

Table 42. Predicted essential reactions for the carbohydrate metabolism category for test conditions 3, 4, 5, 6 and 7

Test exposure condition	Essential Reaction	Reaction name	Reaction function	Reaction formula	EC number	Essential genes
POE20(3)	rxn00988	Acetoacetyl-CoAsynthetase	Butanoate metabolism	H+ + CoA + ATP + Acetoacetate -> PPi + AMP + Acetoacetyl-CoA	6.2.1.16	RW109_04054
POE50(4)	rxn00216	Hexokinase (D-glucose:ATP)	Glycolysis/Gluconeogenesis/Entner-Doudoroff Pathway	ATP + D-Glucose -> ADP + D-glucose-6-phosphate	2.7.1.1,2.7.1.2	RW109_02630
	rxn00786	Fructose-bisphosphate aldolase	Glycolysis/Gluconeogenesis	D-fructose-1,6-bisphosphate <=> Glycerone-phosphate + Glyceraldehyde3-phosphate	4.1.2.13	RW109_01160
	rxn01492	Fructose-1-phosphate kinase	Alternate Carbon Metabolism	ATP + D-fructose-1-phosphate <=> ADP + D-fructose-1,6-bisphosphate	2.7.1.56	RW109_02259
BIT+POE(5)	rxn00459	Enolase	Glycolysis/Gluconeogenesis/Entner-Doudoroff Pathway	2-Phospho-D-glycerate <=> H2O + Phosphoenolpyruvate	4.2.1.11	RW109_04052
	rxn00747	Triose-phosphate isomerase	Glycolysis/Gluconeogenesis	Glyceraldehyde3-phosphate <=> Glycerone-phosphate	5.3.1.1	RW109_06336
	rxn00781	Glyceraldehyde-3-phosphate dehydrogenase (NAD)	Glycolysis/Gluconeogenesis/Entner-Doudoroff Pathway	NAD + Phosphate + Glyceraldehyde3-phosphate <=> NADH + 1,3-Bisphospho-D-glycerate	1.2.1.12,1.2.1.13,1.2.1.59	RW109_02628
LDOnly(6)	rxn01100	Phosphoglycerate kinase	Glycolysis/Gluconeogenesis/Entner-Doudoroff Pathway	H+ + ATP + 3-Phosphoglycerate <=> ADP + 1,3-Bisphospho-D-glycerate	2.7.2.3	RW109_01157
	rxn00256	Citrate synthase	Citrate cycle (TCA cycle)	H+ + CoA + Citrate <=> H2O + Acetyl-CoA + Oxaloacetate	2.3.3.1,2.3.3.3,4.1.3.7	RW109_05443
	rxn00265	Citrate oxaloacetate-lyase (forming acetate from the pro-S carboxymethyl group of citrate)	Citrate cycle (TCA cycle)	Citrate <=> Acetate + Oxaloacetate	4.1.3.6	RW109_00126
POE20(3) & POE50(4) & BIT+POE(5)	rxn00290	3-oxoacid CoA-transferase (Succinyl-CoA: acetoacetate)	Leucine degradation	Acetoacetate + Succinyl-CoA <=> Acetoacetyl-CoA + Succinate	2.8.3.5	RW109_04052
POE20(3) & POE50(4) & LDOnly(6) & LD+BIT(7)	rxn01476	6-phosphogluconolactonase	Pentose Phosphate Cycle/Entner-Doudoroff pathway	H2O + 6-phospho-D-glucono-1-5-lactone -> H+ + 6-Phospho-D-gluconate	3.1.1.31	RW109_02641
	rxn01477	Phosphogluconate dehydratase	Pentose Phosphate Cycle/Entner-Doudoroff pathway	6-Phospho-D-gluconate -> H2O + 2-Keto-3-deoxy-6-phosphogluconate	4.2.1.12	RW109_02629
	rxn03884	2-dehydro-3-deoxy-phosphogluconate aldolase	Pentose Phosphate Cycle/Entner-Doudoroff Pathway	2-Keto-3-deoxy-6-phosphogluconate <=> Pyruvate + Glyceraldehyde3-phosphate	4.1.2.14	RW109_02692(3), RW109_02642(4,6,7)
POE20(3) & POE50(4) & BIT+POE(5) & LDOnly(6) & LD+BIT(7)	rAB00001	Glycoaldehyde:NAD+ oxidoreductase	Glyoxylate and dicarboxylate metabolism	Glycoaldehyde + NAD+ + H2O <=> Glycolate + NADH + H+	1.2.1.21	RW109_00559

Footnotes:

- 1) The test exposure conditions in column 1 are indicated by their abbreviation and test condition number e.g. POE20(3) is test condition 3: POE at 20% of the MIC.
- 2) The genes in the last column without numbers in brackets demonstrate a gene which was essential in all the test conditions within the corresponding row in column 1.
- 3) The numbers in brackets next to the genes in the last column indicates the test exposure condition where the gene was identified as essential.

6.2.7.6 Predicted essential reactions involved in amino acid metabolism

Numerous groupings of the essential reactions involved in amino acid metabolism were noted and hence the analysis was divided to compare the essential reactions between the test conditions: (i) BIT50(2), POE20(3), POE50(4) and BIT+POE(5) (Figure 65 and Table 43) and, (ii) BIT50(2), LDOnly(6) and LD+BIT(7) (Figure 66 and Table 44). The test conditions 2, 3, 4 and 5 were analysed together as these conditions signified the preservatives BIT+POE individually and in combination; POE20 had the most essential reactions with 6 being predicted. The test conditions 2, 6 and 7 were examined collectively as these conditions represented the preservative BIT and laundry detergent separately and in combination with each other; LD+BIT resulted in the highest number of essential reactions with 5 predicted.

When the predicated essential amino acid metabolism reactions for the test conditions BIT50, POE20, POE50 and BIT+POE were compared (Table 65 and Table 43), three reactions were identified only during POE20 exposure; 2 were involved in arginine putriscine and spermidine metabolism (rxn00405 and rxn00858) and one functioned in the biosynthesis of aromatic amino acids (rxn00727). An asparagine metabolism reaction (rxn00260) was only predicted essential in POE50. Two reactions were observed solely in the test condition BIT+POE; one had a function in the biosynthesis of aromatic amino acids (rxn01332) and the other was involved in glycine, serine and threonine metabolism (rxn08647) and was an additional reaction of the RW109 metabolic model (Section 6.2.3). Exposure to the test conditions POE20 and POE50 resulted in the essentiality of another glycine, serine and threonine metabolism reaction rxn00692 and the test conditions BIT50, POE20 and BIT+POE all had predicted essentiality with an asparagine metabolism reaction (rxn00416). A 3-dehydroquinatase reaction (rxn02213) involved in the biosynthesis of aromatic amino acids was identified as an essential reaction with all the test exposure conditions in this comparison (Table 43).

The predicted essential amino acid metabolism reactions were compared when RW109 was exposed to the BIT50, LDOnly and LD+BIT test conditions (Figure 66 and Table 44). A methionine metabolism reaction (rxn00141) was identified only with exposure to LDOnly and the essentialities of 2 threonine and lysine metabolism reactions (rxn01069 and rxn01302) were recognised with LD+BIT. The reaction involved in the biosynthesis of aromatic amino acids (rxn01332) was predicted as essential when RW109 was exposed to the test conditions LDOnly and LD+BIT. An additional biosynthesis of aromatic amino acids reaction (rxn02213) was also found to be essential for the all 3 test conditions BIT50, LDOnly and LD+BIT.

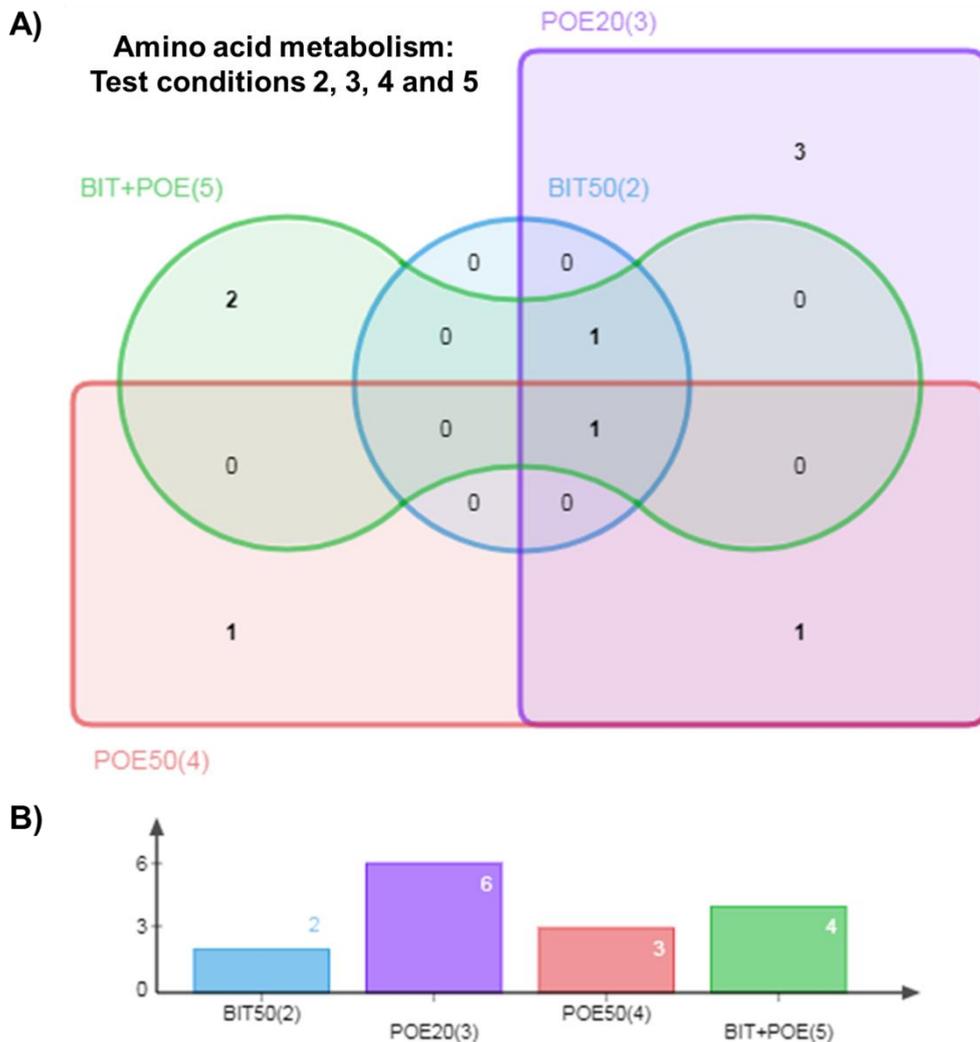


Figure 65. Grouping of the amino acid metabolism predicted essential reactions for test conditions 2, 3, 4 and 5

The Edwards-Venn diagram (A) represents the number of amino acid metabolism essential reactions (numbers in black) which either overlapped or were found only in the test conditions BIT50(2) in blue, POE20(3) in purple, POE50(4) in red and BIT and POE(5) in green. The bar chart (B) is used to clearly show the numbers of essential reactions within each of the test conditions which are coloured the same as in A).

Table 43. Predicted essential reactions for the amino acid metabolism category for test conditions 2, 3, 4 and 5

Test exposure condition	Essential Reaction	Reaction name	Reaction function	Reaction formula	EC number	Essential genes
POE20(3)	rxn00405	Arginine decarboxylase	Arginine Putrescine and Spermidine Metabolism	H+ + L-Arginine -> CO2 + Agmatine	4.1.1.19	RW109_06430
	rxn00727	Anthranilate synthase	Biosynthesis of aromatic amino acids	L-Glutamine + Chorismate -> H+ + L-Glutamate + Pyruvate + Anthranilate	4.1.3.27	RW109_01274
	rxn00858	Agmatinase	Arginine Putrescine and Spermidine Metabolism	H2O + Agmatine -> Urea + Putrescine	3.5.3.11	RW109_00885
POE50(4)	rxn00260	Aspartate transaminase	Asparagine metabolism	2-Oxoglutarate + L-Aspartate <=> L-Glutamate + Oxaloacetate	2.6.1.1	RW109_06570
BIT+POE(5)	rxn01332	3-deoxy-D-arabino-heptulosonate 7-phosphate synthetase	Biosynthesis of aromatic amino acids	H2O + Phosphoenolpyruvate + D-Erythrose4-phosphate -> H+ + Phosphate + DAHP	2.5.1.54,4.1.2.15	RW109_04409
	rxn08647	ATP:(R)-glycerate 2-phosphotransferase	Glycine, serine and threonine metabolism	ATP + Glycerate <=> ADP + H+ + 2-Phospho-D-glycerate	2.7.1.165	RW109_04751
POE20(3) & POE50(4)	rxn00692	Glycine hydroxymethyltransferase	Glycine, serine and threonine metabolism	H2O + Glycine + 5-10-Methylenetetrahydrofolate <=> Tetrahydrofolate + L-Serine	2.1.2.1	RW109_00291
BIT50(2) & POE20(3) & BIT+POE(5)	rxn00416	Asparagine synthase (glutamine-hydrolysing)	Asparagine metabolism	H2O + ATP + L-Glutamine + L-Aspartate -> L-Glutamate + PPi + AMP + L-Asparagine	6.3.5.4	RW109_03968
BIT50(2) & POE20(3) & POE50(4) & BIT+POE(5)	rxn02213	3-dehydroquininate dehydratase	Biosynthesis of aromatic amino acids	5-Dehydroquininate -> H2O + 3-Dehydroshikimate	4.2.1.10	RW109_00842

Footnotes:

- 1) The test exposure conditions in column 1 are indicated by their abbreviation and test condition number e.g. POE20(3) is test condition 3: POE at 20% of the MIC.
- 2) The genes in the last column without numbers in brackets demonstrate a gene which was essential in all the test conditions within the corresponding row in column 1.
- 3) The numbers in brackets next to the genes in the last column indicates the test exposure condition where the gene was identified as essential.

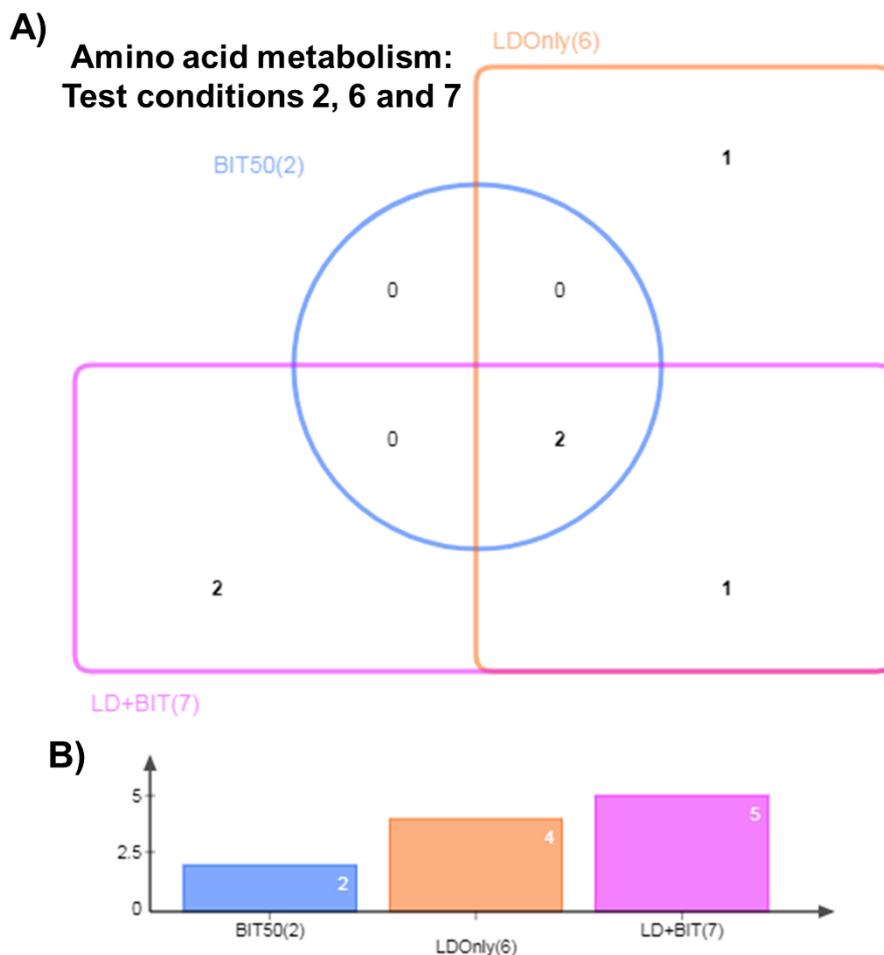


Figure 66. Grouping of the amino acid metabolism predicted essential reactions for test conditions 2, 6 and 7

The Edwards-Venn diagram (A) represents the number of amino acid metabolism essential reactions (numbers in black) which either overlapped or were found only in the test conditions BIT50 (2) in blue, LD Only(6) in orange and LD and BIT(7) in pink. The bar chart (B) is used to clearly show the numbers of essential reactions within each of the test conditions which are coloured the same as in A).

Table 44. Predicted essential reactions for the amino acid metabolism category for test conditions 2, 6 and 7

Test exposure condition	Essential Reaction	Reaction name	Reaction function	Reaction formula	EC number	Essential genes
LDOnly(6)	rxn00141	Adenosylhomocysteinase	Methionine Metabolism	H ₂ O + S-Adenosyl-homocysteine <=> Adenosine + Homocysteine	3.3.1.1	RW109_02760
LD+BIT(7)	rxn01069	Threonine synthase	Threonine and Lysine Metabolism	H ₂ O + O-Phospho-L-homoserine -> H ⁺ + Phosphate + L-Threonine	4.2.3.1,4.2.99.2	RW109_02080
	rxn01302	Homoserine dehydrogenase (NADPH)	Threonine and Lysine Metabolism	NADP + L-Homoserine <=> H ⁺ + L-Aspartate4-semialdehyde + NADPH	1.1.1.3	RW109_02079
LDOnly(6) & LD+BIT(7)	rxn01332	3-deoxy-D-arabino-heptulosonate 7-phosphate synthetase	Biosynthesis of aromatic amino acids	H ₂ O + Phosphoenolpyruvate + D-Erythrose4-phosphate -> H ⁺ + Phosphate + DAHP	2.5.1.54,4.1.2.15	RW109_03043 (6,7)
BIT50(2) & LDOnly(6) & LD+BIT(7)	rxn00416	Asparagine synthase (glutamine-hydrolysing)	Asparagine metabolism	H ₂ O + ATP + L-Glutamine + L-Aspartate -> L-Glutamate + PPi + AMP + L-Asparagine	6.3.5.4	RW109_03968 (2,6) RW109_02340 (7)
	rxn02213	3-dehydroquinate dehydratase	Biosynthesis of aromatic amino acids	5-Dehydroquinate -> H ₂ O + 3-Dehydroshikimate	4.2.1.10	RW109_00842 (2,6) RW109_06438 (7)

Footnotes:

- 1) The test exposure conditions in column 1 are indicated by their abbreviation and test condition number e.g. POE20(3) is test condition 3: POE at 20% of the MIC.
- 2) The genes in the last column without numbers in brackets demonstrate a gene which was essential in all the test conditions within the corresponding row in column 1.
- 3) The numbers in brackets next to the genes in the last column indicates the test exposure condition where the gene was identified as essential.

6.2.7.7 Predicted essential reactions involved in metabolism of cofactors and vitamins

Six reactions were predicted to be essential in the metabolism of cofactors and vitamins KEGG functional category with no reactions being associated with only one test exposure condition (Table 45). The test exposure conditions BIT+POE and LD+BIT had the highest numbers of essential reactions with 6 predicted for each. The coproporphyrinogen oxidase reaction (rxn02303) involved in tetrapyrrole biosynthesis was predicted as essential in all test exposure conditions except for BIT20. Two reactions involved in one carbon metabolism, termed as methylene-tetrahydrofolate dehydrogenase (NADP) (rxn00907) and methenyltetrahydrofolate cyclohydrolase (rxn01211) were predicted as essential in the test conditions POE50, BIT+POE, LDOnly and LD+BIT.

Table 45. Predicted essential reactions for the metabolism of cofactors and vitamins KEGG functional category

Test exposure condition	Essential Reaction	Reaction name	Reaction function	Reaction formula	EC number	Essential genes
POE50(4) & BIT+POE(5) & LDOonly(6) & LD+BIT(7)	rxn00907	Methylenetetrahydrofolate dehydrogenase (NADP)	One Carbon Metabolism	NADP + 5-10-Methylenetetrahydrofolate <=> NADPH + 5-10-Methenyltetrahydrofolate	1.5.1.5	RW109_00468 (4,5) RW109_04361 (6,7)
	rxn01211	Methenyltetrahydrofolate cyclohydrolase	One Carbon Metabolism	H2O + 5-10-Methenyltetrahydrofolate <=> H+ + 10-Formyltetrahydrofolate	3.5.4.9,6.3.4.3	RW109_00468 (4,5) RW109_04361 (6,7)
	rxn03841	4-aminobenzoate synthase	Folate Metabolism	ADC -> H+ + Pyruvate + ABEE	4.1.3.38	RW109_02038 (4,5) RW109_02918 (6,7)
POE20(3) & BIT+POE(5) & LD+BIT(7)	rxn01629	Glutamate-1-semialdehyde aminotransferase	Tetrapyrrole Biosynthesis	5-Aminolevulinate <=> L-Glutamate1-semialdehyde	5.4.3.8	RW109_07141 (3) RW109_01591 (5,7)
POE20(3) & POE50(4) & BIT+POE(5) & LDOonly(6) & LD+BIT(7)	rxn00299	GTP cyclohydrolase I	Folate Metabolism	H2O + GTP -> H+ + 7,8-Dihydroneopterin 3'-triphosphate + Formate	3.5.4.16	RW109_02362(3,4,7) RW109_04487 (5,6)
BIT50(2) & POE20(3) & POE50(4) & BIT+POE(5) & LDOonly(6) & LD+BIT(7)	rxn02303	Coproporphyrinogen oxidase	Tetrapyrrole Biosynthesis	O2 + 2 H+ + CoproporphyrinogenIII -> 2 H2O + 2 CO2 + ProtoporphyrinogenIX	1.3.3.3	RW109_00614

Footnotes:

- 1) The test exposure conditions in column 1 are indicated by their abbreviation and test condition number e.g. POE20(3) is test condition 3: POE at 20% of the MIC.
- 2) The genes in the last column without numbers in brackets demonstrate a gene which was essential in all the test conditions within the corresponding row in column 1.
- 3) The numbers in brackets next to the genes in the last column indicates the test exposure condition where the gene was identified as essential.

6.2.7.8 Predicted essential reactions involved in fatty acid metabolism and energy metabolism

Limited essential reactions were found within the KEGG functional categories for fatty acid metabolism and energy metabolism. One essential reaction was identified in the fatty acid metabolism category which involved the was Enoyl-[acyl-carrier-protein] reductase (NADH) enzyme (rxn08398) and it was predicted to be essential for the 3 test exposure conditions POE50(4), LDOnly(6) and LD+BIT(7). In the energy metabolism functional category, the carbonic acid hydrolase reaction (rxn00102) was found to be essential in only the test condition BIT+POE(5).

6.2.8 Comparison of the predicted essential reactions between POE20 and POE50

The reactions predicted to be essential were examined when RW109 was exposed to the test conditions POE20 and POE50 (Figure 67 and Table 46). The essentialities of 29 reactions were predicted in both POE20 and POE50 and out of these 9 were found to be involved in nucleotide metabolism. When exposed to POE20 a notable observation were the 6 benzoate degradation reactions predicted to be essential, which were not identified with POE50. A higher number of lipid metabolism reactions involved in fatty acid and glycerolipid metabolism were identified with POE50 and a greater number of membrane transport reactions. This illustrates the differences in the essentiality of reactions during exposure to the same preservative at different concentrations.

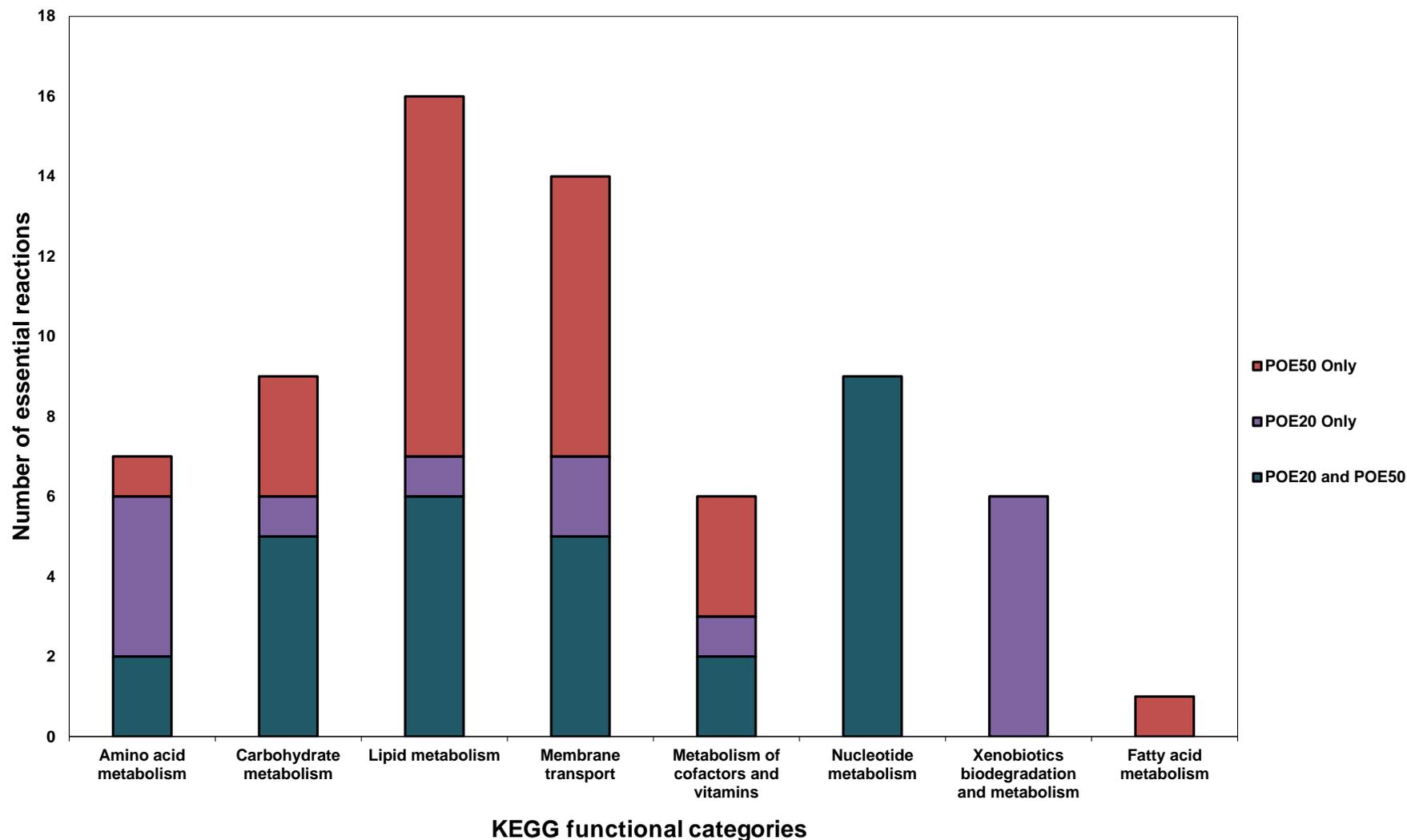


Figure 67. Comparison of the predicted essential reactions between POE20 and POE50

The stacked bar-chart demonstrates the predicted essential reactions found in both POE20 and POE50 and those found only in each test condition; the reactions are grouped via their KEGG functional categories. Reactions only found in POE50 are shown in red, those only identified in POE20 are in purple and reactions predicted to be essential in both POE20 and POE50 are shown in turquoise.

Table 46. Predicted essential reactions when RW109 was exposed to POE20 and POE50

Exposure condition(s)	Essential reaction	Reaction name	Reaction function	KEGG functional category	
Both POE20 and POE50	rxn00692	Glycine hydroxymethyltransferase	Glycine, serine and threonine metabolism	Amino acid metabolism	
	rxn02213	3-dehydroquininate dehydratase	Phenylalanine, tyrosine and tryptophan biosynthesis	Amino acid metabolism	
	rAB00001	Glycoaldehyde:NAD+ oxidoreductase	Glyoxylate and dicarboxylate metabolism	Carbohydrate metabolism	
	rxn00290	3-oxoacid CoA-transferase (Succinyl-CoA: acetoacetate)	Butanoate metabolism	Carbohydrate metabolism	
	rxn01476	6-phosphogluconolactonase	Pentose phosphate pathway	Carbohydrate metabolism	
	rxn01477	Phosphogluconate dehydratase	Pentose phosphate pathway	Carbohydrate metabolism	
	rxn03884	2-dehydro-3-deoxy-phosphogluconate aldolase	Pentose phosphate pathway	Carbohydrate metabolism	
	rPY00183	Cardiolipin synthase (n-C18:0 18:1)	Glycerophospholipid metabolism	Lipid metabolism	
	rPY00184	Cardiolipin synthase (n-C16:0 16:1)	Glycerophospholipid metabolism	Lipid metabolism	
	rxn08229	Cardiolipin synthase (n-C16:0)	Glycerophospholipid metabolism	Lipid metabolism	
	rxn08230	Cardiolipin synthase (n-C16:1)	Glycerophospholipid metabolism	Lipid metabolism	
	rxn08231	Cardiolipin synthase (n-C18:0)	Glycerophospholipid metabolism	Lipid metabolism	
	rxn08232	Cardiolipin synthase (n-C18:1)	Glycerophospholipid metabolism	Lipid metabolism	
	rxn05164	L-histidine transport via ABC system	Transport	Membrane transport	
	rxn05165	L-proline transport via ABC system	Transport	Membrane transport	
	rxn05169	L-threonine transport via ABC system	Transport	Membrane transport	
	rxn05219	L-methionine transport via ABC system	Transport	Membrane transport	
	rxn09653	Urea reversible transport via proton symport (2 H+)	Transport	Membrane transport	
	rxn00299	GTP cyclohydrolase I	Folate biosynthesis	Metabolism of cofactors and vitamins	
	rxn02303	Coproporphyrinogen oxidase	Porphyrin and chlorophyll metabolism	Metabolism of cofactors and vitamins	
	rxn00364	Cytidylate kinase (CMP)	Pyrimidine metabolism	Nucleotide metabolism	
	rxn00409	Nucleoside-diphosphate kinase (ATP:CDP)	Pyrimidine metabolism	Nucleotide metabolism	
	rxn00839	Nucleoside-diphosphate kinase (ATP:dADP)	Purine metabolism	Nucleotide metabolism	
	rxn00917	GMP synthase (glutamine-hydrolysing)	Purine metabolism	Nucleotide metabolism	
	rxn01353	Nucleoside-diphosphate kinase (ATP:dGDP)	Purine metabolism	Nucleotide metabolism	
	rxn01673	Nucleoside-diphosphate kinase (ATP:dCDP)	Pyrimidine metabolism	Nucleotide metabolism	
	rxn05816	Deoxyadenosine 5'-triphosphate:DNA deoxynucleotidyltransferase (DNA-directed)	Purine metabolism	Nucleotide metabolism	
	rxn05817	Deoxyguanosine 5'-triphosphate:DNA deoxynucleotidyltransferase (DNA-directed)	Purine metabolism	Nucleotide metabolism	
	rxn05818	Deoxycytidine triphosphate:DNA deoxynucleotidyltransferase (DNA-directed)	Pyrimidine metabolism	Nucleotide metabolism	
	POE20 Only	rxn00405	Arginine decarboxylase	Arginine and proline metabolism	Amino acid metabolism
		rxn00416	Asparagine synthase (glutamine-hydrolysing)	Alanine, aspartate and glutamate metabolism	Amino acid metabolism
		rxn00727	Anthranylase synthase	Phenylalanine, tyrosine and tryptophan biosynthesis	Amino acid metabolism
		rxn00858	Agmatinase	Arginine and proline metabolism	Amino acid metabolism
rxn00988		Acetoacetyl-CoAsynthetase	Butanoate metabolism	Carbohydrate metabolism	
rxn00763		Alcohol dehydrogenase (glycerol)	Glycerolipid metabolism	Lipid metabolism	
rxn05306		L-phenylalanine transport in/out via proton symport	Transport	Membrane transport	
rxn10963		FeoB iron (II) transporter	Transport	Membrane transport	
rxn01629		Glutamate-1-semialdehyde aminotransferase	Porphyrin and chlorophyll metabolism	Metabolism of cofactors and vitamins	
rxn00598		3-oxoadipyl-CoA thiolase	Benzoate degradation	Xenobiotics biodegradation and metabolism	
rxn00962		4-hydroxybenzoate 3-monooxygenase	Benzoate degradation	Xenobiotics biodegradation and metabolism	
rxn01192		Protocatechuate 3,4-dioxygenase	Benzoate degradation	Xenobiotics biodegradation and metabolism	
rxn02143		3-oxoadipate CoA-transferase	Benzoate degradation	Xenobiotics biodegradation and metabolism	
rxn02144		3-oxoadipate enol-lactonase	Benzoate degradation	Xenobiotics biodegradation and metabolism	
rxn02369		3-carboxy-cis,cis-muconate cycloisomerase	Benzoate degradation	Xenobiotics biodegradation and metabolism	
rxn02483	4-carboxymuconolactone decarboxylase	Benzoate degradation	Xenobiotics biodegradation and metabolism		
POE50 Only	rxn00260	Aspartate transaminase	Alanine, aspartate and glutamate metabolism	Amino acid metabolism	
	rxn00216	Hexokinase (D-glucose:ATP)	Glycolysis / Gluconeogenesis	Carbohydrate metabolism	
	rxn00786	Fructose-bisphosphate aldolase	Glycolysis / Gluconeogenesis	Carbohydrate metabolism	
	rxn01492	Fructose-1-phosphate kinase	Fructose and mannose metabolism	Carbohydrate metabolism	
	rxn08398	Fnoyl-[acyl-carrier-protein] reductase (NADH) (n-C18:1)	Fatty acid metabolism	Fatty acid metabolism	
	rPY00162	FA160161S	Fatty acid biosynthesis	Lipid metabolism	
	rPY00163	CDP-diacylglycerol synthetase (n-C16:0 16:1)	Fatty acid biosynthesis	Lipid metabolism	
	rPY00167	FA180181S	Fatty acid biosynthesis	Lipid metabolism	
	rPY00168	CDP-diacylglycerol synthetase (n-C18:0 18:1)	Fatty acid biosynthesis	Lipid metabolism	
	rxn08087	1-hexadec-7-enoyl-sn-glycerol 3-phosphate O-acyltransferase (n-C161)	Glycerolipid metabolism	Lipid metabolism	
	rxn08089	1-octadec-7-enoyl-sn-glycerol 3-phosphate O-acyltransferase (n-C181)	Glycerolipid metabolism	Lipid metabolism	
	rxn08310	CDP-diacylglycerol synthetase (n-C161)	Glycerolipid metabolism	Lipid metabolism	
	rxn08311	CDP-diacylglycerol synthetase (n-C180)	Glycerolipid metabolism	Lipid metabolism	
	rxn08312	CDP-diacylglycerol synthetase (n-C181)	Glycerolipid metabolism	Lipid metabolism	
	rJB00225	D-Fructose transport via PEP:Pyr PTS	Transport	Membrane transport	
	rxn05151	L-lysine transport via ABC system	Transport	Membrane transport	
	rxn05161	L-leucine transport via ABC system	Transport	Membrane transport	
	rxn05168	L-valine transport via ABC system	Transport	Membrane transport	
	rxn05179	L-isoleucine transport via ABC system	Transport	Membrane transport	
	rxn05560	D-fructose transport via PEP:Pyr PTS	Transport	Membrane transport	
rxn05663	L-tryptophan transport in via proton symport	Transport	Membrane transport		
rxn00907	Methylenetetrahydrofolate dehydrogenase (NADP)	One carbon pool by folate	Metabolism of cofactors and vitamins		
rxn01211	Methylenetetrahydrofolate cyclohydrolase	One carbon pool by folate	Metabolism of cofactors and vitamins		
rxn03841	4-aminobenzoate synthase	Folate biosynthesis	Metabolism of cofactors and vitamins		

6.2.9 Linking up-regulated genes with those predicted as essential

A number of essential genes from each constrained model (except for BIT20) were also significantly up-regulated when exposed to the test condition in which they were predicted to be essential (Table 47). LDOnly had the greatest number of predicted essential up-regulated genes with 10 and BIT50 had the least with 1. A few genes were found to be up-regulated in more than one condition. The RW109_00614 was identified in BIT50, LDOnly and LD+BIT and encoded for a coproporphyrinogen oxidase involved in the rxn02303 reaction. In POE20, POE50 and BIT+POE the RW109_04052 gene encoding a 3-oxoacid CoA-transferase (Succinyl-CoA: acetoacetate) increased in expression and) was involved in the rxn00290 reaction. The RW109_02629 and RW109_02642 genes, both involved in alternative pentose phosphate pathways, were significantly up-regulated in POE50, LDOnly and LD+BIT. The essential RW109_04051 gene associated with the benzoate degradation rxn02143 reaction in POE20 and BIT+POE was significantly up-regulated when exposed to these test conditions. In LD and LD+BIT the RW109_01980 gene (encoding a nucleoside-diphosphate kinase [Ndk] enzyme) was significantly up-regulated but was associated with different essential reactions in the two exposure conditions; it was linked to rxn00839 in LDOnly and rxn00409 in LD+BIT. This was also true for the RW109_05296 gene, which had up-regulated expression in these two, conditions; in LDOnly it was linked to the essential L-histidine transport reaction rxn05164 and was associated with the rxn05154 L-arginine transport reaction in LD+BIT.

Table 47. Significantly up-regulated essential genes for each preservative exposure condition

Test exposure condition	Essential gene	Essential Reaction	Reaction name	Reaction function	Log2Fold Change	Adjusted P-Value
BIT50	RW109_00614	rxn02303	coproporphyrinogen oxidase	Porphyrin and chlorophyll metabolism	1.6	0.009454
POE20	RW109_04052	rxn00290	3-oxoacid CoA-transferase (Succinyl-CoA: acetoacetate)	Butanoate metabolism	3.5	6.64E-16
	RW109_04051	rxn02143	3-oxoadipate CoA-transferase	Benzoate degradation	3.7	3.67E-10
POE50	RW109_04052	rxn00290	3-oxoacid CoA-transferase (Succinyl-CoA: acetoacetate)	Butanoate metabolism	2.9	2.16E-10
	RW109_02629	rxn01477	phosphogluconate dehydratase	Pentose phosphate pathway/Entner-Doudoroff pathway	3.3	2.58E-20
	RW109_02630	rxn00216	hexokinase (D-glucose:ATP)	Glycolysis / Gluconeogenesis/Entner-Doudoroff pathway	2.0	0.027894
	RW109_02259	rxn01492 and rJB00225	fructose-1-phosphate kinase and D-Fructose transport via PEP	Fructose and mannose metabolism and Membrane transport	2.7	0.040176
	RW109_02642	rxn03884	2-dehydro-3-deoxy-phosphogluconate aldolase	Pentose phosphate pathway/Entner-Doudoroff pathway	2.9	0.000884
	RW109_02258	rxn01492 and rxn0556	fructose-1-phosphate kinase and D-Fructose transport via PEP	Fructose and mannose metabolism/Membrane transport	3.0	0.000165
BIT+POE	RW109_04052	rxn00290	3-oxoacid CoA-transferase (Succinyl-CoA: acetoacetate)	Butanoate metabolism	3.4	1.33E-18
	RW109_04051	rxn02143	3-oxoadipate CoA-transferase	Benzoate degradation	4.0	3.35E-14
	RW109_00825	rxn00917	GMP synthase (glutamine-hydrolysing)	Purine metabolism	1.8	0.046409
LDOnly	RW109_05695	rxn10963	FeoB iron (II) transporter	Membrane transport	2.1	1.23E-15
	RW109_00614	rxn02303	coproporphyrinogen oxidase	Porphyrin and chlorophyll metabolism	2.8	7.3E-09
	RW109_01980	rxn00839	nucleoside-diphosphate kinase (ATP:dADP)	Purine metabolism	1.7	4.42E-12
	RW109_02641	rxn01476	6-phosphogluconolactonase	Pentose phosphate pathway/Entner-Doudoroff pathway	4.1	0.000044
	RW109_02629	rxn01477	phosphogluconate dehydratase	Pentose phosphate pathway/Entner-Doudoroff pathway	2.8	1.02E-18
	RW109_02642	rxn03884	2-dehydro-3-deoxy-phosphogluconate aldolase	Pentose phosphate pathway/Entner-Doudoroff pathway	4.2	2.31E-13
	RW109_05296	rxn05164	L-histidine transport via ABC system	Membrane transport	1.7	0.005361
	RW109_05105	rxn05169	L-threonine transport via ABC system	Membrane transport	2.2	0.038819
	RW109_05106	rxn05170	L-threonine transport via ABC system	Membrane transport	1.6	0.001321
	RW109_05107	rxn05171	L-threonine transport via ABC system	Membrane transport	1.5	0.000215
LD+BIT	RW109_00614	rxn02303	coproporphyrinogen oxidase	Porphyrin and chlorophyll metabolism	2.9	7.97E-09
	RW109_01980	rxn00409	nucleoside-diphosphate kinase (ATP:CDP)	Pyrimidine metabolism	1.9	1.6E-18
	RW109_02641	rxn01476	6-phosphogluconolactonase	Pentose phosphate pathway/Entner-Doudoroff pathway	4.7	0.0000223
	RW109_02629	rxn01477	phosphogluconate dehydratase	Pentose phosphate pathway/Entner-Doudoroff pathway	3.2	8.17E-27
	RW109_02660	rxn00364	cytidylate kinase (CMP)	Pyrimidine metabolism	2.5	0.000458
	RW109_02642	Rxn03884	2-dehydro-3-deoxy-phosphogluconate aldolase	Pentose phosphate pathway/Entner-Doudoroff pathway	4.4	4.42E-16
	RW109_05296	rxn05154	L-arginine transport via ABC system	MembraneTransport	2.0	0.000298

6.2.10 Additional modelling work

6.2.10.1 Finding an alternative pathway in RW109 for the essential rxn00770 reaction

A pathway encoding the essential reaction rxn00770 (identified in Section 6.2.5) involved in purine biosynthesis was not identified in RW109 using the standard computational modelling approaches as previously described. Putative genes and pathways that may encode this reaction were subsequently investigated as follows. Further searching on the KEGG database identified a *P. aeruginosa* R06836 reaction specific to the PAO1 strain, which uses the ATP: ribose-1, 5-bisphosphate phosphotransferase enzyme; this reaction was found to produce 5-Phospho-alpha-D-ribose1-diphosphate which was the same product of the essential rxn00770 reaction (Section 6.2.5). A BLASTp comparison of the PAO1 gene associated with this reaction was carried out against the RW109 genome and a match was identified with the RW109_02429 gene (percentage identity= 99.46%, E-Value: 3.00e-130, BIT-Score = 365). This indicates that RW109 may use the R06836 reaction as an alternative for producing 5-Phospho-alpha-D-ribose1-diphosphate instead of rxn00770. This R06836 reaction was identified after the simulation and transcriptomic integration work described to this point and hence was not included in the latter analysis.

6.2.10.2 Growth of RW109 on various carbon sources

The industrial sponsors carried out preliminary growth of the RW109 strain on various carbon sources using BIOLOG phenotype microarrays with the MicroPlates PM1 and PM2a (data not shown). The plates were inoculated with approximately 10⁶ cfu/ml of an overnight culture and turbidity was read using the OmniLog® ID System (BIOLOG) at 30°C for 48 hours; this was carried out in triplicate. The industrial sponsors report the carbon utilisation results for the RW109 strain were consistent with those obtained with the PA14 strain (Bartell et al., 2017).

6.3 Discussion

6.3.1 The outcomes of reconstructing a genome-scale metabolic model specific to an industrial *P. aeruginosa* strain and the key characteristics of RW109

In this project an existing PA14 genome-scale metabolic reconstruction (Bartell et al., 2017) was curated to represent the *P. aeruginosa* RW109 strain, enabling an extended understanding of the metabolic functionality of the industrial strain. It also provided a basis to simulate the essentiality of reactions during exposure to industrial conditions. The PA14 model was published in 2017 after a great deal of curation and validation using previous *P. aeruginosa* PAO1 metabolic models (iMO1056 (Oberhardt et al., 2008) and iMO1086 (Oberhardt et al., 2011), genome sequences, carbon source utilisation results and gene essentiality data sets (Bartell et al., 2017). Core genome alignment analysis had identified the large genome of the industrial RW109 strain as not being closely related to the PA14 strain (Chapter 3, Section 3.2.6.2, Figure 18), but as a basis to build an initial network, the PA14 model provided comprehensive and up-to-date information of the biochemical networks for *P. aeruginosa*. The limited differences in COG and KEGG annotations between the PA14 and RW109 strains (Chapter 3, Section 3.2.6.1) also indicated the PA14 model as a suitable base for the reconstruction of the industrial metabolic model.

Obtaining a complete well-annotated whole-genome for the RW109 strain was highly beneficial for accurate comparison of the PA14 model genes and for the assigning of additional model reactions using modelSEED and KEGG databases. The modelSEED is an informative resource regularly used for metabolic modelling providing a comprehensive database of experimentally and literature validated balanced reactions and metabolites (Henry et al., 2010, Devoid et al., 2013). The KEGG database offers a wide-ranging collection of pathway maps and biological functional annotations which are updated on a regular basis (Kanehisa et al., 2017, Thiele and Palsson, 2010). KEGG pathways are also referenced within the modelSEED (Bartell et al., 2017). The use of these resources during the RW109 model reconstruction along with the whole genome sequence provided a reliable method to establish reactions. There are a number of other databases available for genome-scale metabolic modelling which include Metacyc (Caspi et al., 2007) and Reactome (Fabregat et al., 2015) which again offer large sets of reactions and pathways and could be used to further expand and justify the reactions within the industrial model. However, because of the complexity of the analysis, these were not applied to RW109.

Throughout this chapter, the COBRA tool box (Schellenberger et al., 2011) was used to perform FBA simulations. One major assumption of this analysis was the implementation of a well-defined objective function for accurate predictions of metabolic capabilities. The RW109 biomass reaction was set as the objective function throughout this study to serve as an approximation of growth. The process of making the reaction specific to the industrial strain involved updating the PA14 model biomass which had previously been expanded to accurately represent the *P. aeruginosa* species (Bartell et al., 2017). Once converted into a COBRA model within MATLAB, the RW109 model was able to produce biomass and an optimal flux balance analysis solution was found, enabling further analysis to be carried out.

The PA14 model genes, which did not match with any in RW109, were removed using the COBRA toolbox and five non-essential reactions were deleted from the model (Figure 57). Four of these non-essential reactions functioned in glycan biosynthesis and metabolism, specifically in lipopolysaccharide (LPS) biosynthesis of B and O band antigens. *P. aeruginosa* is associated with a wide-range of genes involved in the biosynthesis of LPS (King et al., 2009, Lam et al., 2011). This suggests an alternative pathway is used by the RW109 strain when compared to PA14, which is supported by the numerous alternative LPS associated reactions remaining within the RW109 model. A lipid-associated reaction was also removed which had a function in glycerophospholipid metabolism, and again a variety of similar reactions were kept in the industrial strain model. Comparisons between *P. aeruginosa* isolates have also identified differences between strain-specific genes which have related functions but divergent DNA sequences, such as those encoding for lipopolysaccharide O antigens (Spencer et al., 2003, Lee et al., 2006a).

The KEGG functional annotations of the 1,656 reactions of the RW109 model revealed the wide-ranging metabolic capabilities of the industrial strain suggesting it can adapt to survive in a variety of different environments. The ability of bacteria to grow quickly and persist in adverse environments has been linked to metabolic flexibility with the development of an 'environmental scope index' (ESI) score which is defined as the fraction of environments in which the strain is viable (Freilich et al., 2009). *P. aeruginosa* had a high ESI score highlighting its ability to proliferate and thrive in a diverse range of environments (Freilich et al., 2009). Reactions involved in lipid metabolism, transport and amino acid metabolism represented the highest numbers in the RW109 model which was also observed in the PA14 model (Bartell et al., 2017). This also suggests the reactions characterised with these functional categories are prominent in the overall lifestyle of *P. aeruginosa*.

The RW109 metabolic model had 166 additional reactions when compared to PA14, and these represented 15 KEGG functional categories highlighting the surplus wide-ranging metabolic capabilities of the RW109 strain. The xenobiotics biodegradation and metabolism functional category characterised the highest number of these additional reactions with the majority being linked to genes on the large plasmid 1. The RW109 plasmid 1 was 555 kb in size and provided the RW109 strain with more additional genes when compared to the smaller plasmid 2, which was 152 kb (Chapter 3). Xenobiotics are defined as un-natural compounds released into an environment (Van Der Meer et al., 1992, Top and Springael, 2003) and HPC products are man-made un-natural environments for bacteria. The occasional growth and survival of bacteria within HPCs may be attributed to the ability to metabolise and degrade compounds found in these industrial products. Bacteria are used in bioremediation to degrade harmful chemicals accidentally released into the environment (Das and Chandran, 2010) and *P. aeruginosa* strains have been shown to breakdown crude oil and petroleum hydrocarbons (Song et al., 2006, Pasumarthi et al., 2013). Within the xenobiotic functional category the reactions were mainly linked to genes from plasmid 1, suggesting it may play a role in the ability of RW019 to survive in harsh environments containing toxic hydrocarbons. The metabolism of xenobiotics by the cytochrome P450 sub-category represented 60% of the additional reactions within the xenobiotic functional group (Figure 56). A *P. aeruginosa* strain isolated

from oil contaminated soil was found to obtain two P450 monooxygenase enzymes which had an important role in the ability of the strain to degrade hydrocarbons (Liu et al., 2012). This suggests the cytochrome P450 enzymes have a principal function in the ability of *P. aeruginosa* to survive in harsh and un-natural environments. Cytochrome P450 enzymes also function in the detoxification of un-natural substances (Cerniglia, 1993, Wang and Shao, 2013) and may also play a role in the degradation of compounds encountered by *P. aeruginosa* in HPC products.

Amino acid metabolism represented the second largest number of additional RW109 model reactions with 43% characterised within the glycine, serine and threonine metabolism subcategory (Figure 56). This amino acid metabolic pathway has been associated with biofilm formation in *P. aeruginosa* (Vital-Lopez et al., 2015). Metabolites of the glycine, serine, and threonine metabolism pathways were frequently depleted when a genome-scale kinetic model was used to measure metabolites during biofilm associated reactions (Vital-Lopez et al., 2015). RW109 has an increased number of reactions within this amino acid metabolic subcategory, suggesting the RW109 strain has the potential for an elevated survival during biofilm formation. This could be evaluated experimentally by comparing the ability of the RW109 and PA14 strains to establish biofilms.

6.3.2 Important points to consider when using the RW109 model reconstruction for computational predictions of biological functionality

A principal issue concerning the industrial *P. aeruginosa* model was the number of associated model genes, which only represented 18.2% of the total RW109 genome coding capacity. The significance of the >80% of genes not included in the model would be overlooked during analysis and the likelihood of the model missing strain specific reactions would be high. Therefore, any model simulations should be considered as approximations of the full potential of the RW109 strain. However, as a first-look, the predictions provide valuable insights into the biological functionality of an industrial *P. aeruginosa* strain. They were useful in highlighting genes and reactions of interest for experimental validation in further investigations, such as those encoding cell membrane phospholipid biosynthesis reactions. The RW109 model is also representative of the wide-ranging metabolic capabilities of *P. aeruginosa* as the reactions annotations are divided throughout a number of KEGG biological functional categories.

Gene essentiality predictions carried out using the PA14 model were found to be 91% accurate when compared to *in vitro* essentiality experiments (Bartell et al., 2017). In correlation with the RW109, the PA14 model also only represents a small percentage of the genes (18.9%) within its own genome, highlighting a modelling dilemma in the applicability of these research techniques when bacteria obtain versatile genomes. Since the RW109 model reconstruction was based on the PA14 model, the accuracy of the gene and reaction essentiality predictions should be high, however there is a requirement for further validation of fundamental pathways using mutagenesis *in vitro* experimentation.

Preliminary carbon source utilisation data for the RW109 strain, obtained from the industrial sponsors, was found to be comparable to PA14 (data not shown). This supports the inclusion of the PA14 reactions within the industrial strain model, however further work using carbon source utilisation is required to validate the addition of each RW109 metabolic reaction. Overall, this study presents a first, outline version of a *P. aeruginosa* model for an industrial strain. Further refinement should include gap-filling of reaction networks, examining the balancing of reactions, identifying any dead-end metabolites and minimal media simulations together with flux variability analysis as recommended by (Henry et al., 2010), (Thiele and Palsson, 2010, Opdam et al., 2017).

6.3.3 The successful integration of gene expression data to produce models specific to industry related conditions

The transcriptomic data obtained from *P. aeruginosa* preservative exposure experiments (Chapter 5) was successfully integrated into the RW109 metabolic model using the MADE algorithm (Jensen and Papin, 2011). Various methods for the integration of expression data are available, but MADE is advantageous as it does not rely on user-specified threshold values but focuses on the statistical significance of fold changes (Blazier and Papin, 2012, Jensen and Papin, 2011). Other methods which utilise constraint-based FBA methods to integrate gene expression data include but are not limited to, the Gene Inactivity Moderated by Metabolism and Expression (GIMME) algorithm (Becker and Palsson, 2008), the Integrative Metabolic Analysis Tool (iMAT) (Shlomi et al., 2008) and E-FLUX (Colijn et al., 2009). These three methods require the user to decide on a threshold for differentially regulated gene expression when compared to the MADE algorithm. The availability of multiple datasets for successful integration is required by the MADE algorithm (Jensen and Papin, 2011); this was not an issue in this project as the log₂-fold changes for 7 test exposure conditions compared to a control condition were available. To validate the results of this study, the gene expression data could be integrated using one of the indicated alternative methods and gene and reaction essentiality results compared. A similar computational pipeline to the one followed in this study was also used in a recent investigation into the metabolic network for the malaria *Plasmodium falciparum* parasite (Carey et al., 2017). Transcriptomic data was integrated using the MADE algorithm and gene and reaction essentiality predictions used to detect unique pathways in the resistant and sensitive parasite phenotypes to identify candidate drug targets (Carey et al., 2017).

Essential genes and reactions are required to maintain vital functions when bacteria are exposed to different environmental conditions (Liu et al., 2010, Zhang and Lin, 2008, Conesa et al., 2016, Schellenberger et al., 2011). Those identified are also considered promising targets for the development of novel targeted antimicrobials due to their importance in bacterial survival (Juhás et al., 2012). This study predicted the essentiality of shared and unique reactions for seven industrial exposure conditions (Figure 60). Out of all the industry relevant conditions, exposure to POE at 50% of the MIC resulted in the greatest number reactions predicted to be essential (n = 53). This suggests the preservative at this concentration had a more substantial effect on RW109. This test exposure condition did not however result in the largest percentage of DEGs in relation to the whole genome size (Chapter 5, Table 21). 11.78% of the RW109 genes were differentially regulated in response to POE at 50% of the MIC whereas 17.31% and 20.36% of RW109 genes had altered expression in response to LD Only and LD with BIT, respectively. After the expression data integration, the POE50 model had the least number of functional genes compared to the other exposure conditions (Table 36). This infers that a higher number of genes were considered inactive by the MADE algorithm, increasing the likelihood of the remaining functioning genes being predicted as essential when RW109 was exposed to POE at 50% of the MIC. BIT20 and BIT50 had the least number of predicted essential reactions with 1 and 10 respectively; this is consistent with the gene expression data where 0.03% and 1.55% of RW109 genes were differentially regulated in response to BIT at 20% and 50% of the MIC, respectively. This signifies the preservative BIT at these

concentrations had a limited effect on the RW109 strain. The growth dynamic analysis of RW109 in the presence of these test conditions supports these findings (Chapter 5, Figure 43). POE50 exposure after 24 hours resulted in the greatest reductions in culture optical density and viable cell counts when compared to the control condition, whereas BIT20 and BIT50 demonstrated minimal differences.

Overall, the modelling and gene expression integration suggests that POE at 50% of the MIC was more difficult for RW109 to grow in when compared to the other test conditions, resulting in the occurrence of more predicted essential genes and reactions in order to grow. The limited biological functional diversity of reactions predicted to be essential with BIT also suggests this preservative potentially has a more specific mode of action when compared to POE. The BIT20 essential reaction was involved in membrane transport and the BIT50 reactions had functions in amino acid metabolism, membrane transport, metabolism of cofactors and vitamins and nucleotide metabolism (Figure 60). There were also no essential reactions unique to the exposure of RW109 with BIT at both 20% and 50% of the MIC, as the predicted reactions were also identified with other exposure conditions. Overall, this indicates that fewer alterations in metabolism were required by RW109 when it was exposed to BIT.

Interestingly, a few genes predicted as essential for survival in a test condition, were also identified as significantly up-regulated during exposure to that same condition (Table 47). LD only had the greatest number of predicted essential genes, which also had increased expression, and half of these were linked to reactions involved in membrane transport pathways (Table 47). This association suggests that under certain conditions, there might be a benefit to increasing the expression of a gene, which encodes for a vital enzyme/protein in a reaction essential for survival. These identified genes are of interest as potential targets due to their predicted essentiality and the evidence of their utilisation as recognised by the up-regulation during exposure to the industry relevant conditions. This observation also demonstrates how transcriptomic profiling techniques alone may overlook the importance of essential genes in response to a particular condition. Not all predicted essential genes were identified as significantly up-regulated, suggesting the association is only relevant with certain gene types. The observed link could also be an increased likelihood of a gene, which is up-regulated also being predicted as essential, due to the fewer genes within the constrained MADE models.

6.3.4 The key predicted essential reactions and their potential function in the ability of RW109 to survive exposure to industry relevant conditions

A large number of POE50 predicted essential reactions were found within the lipid metabolism category, and of these, five reactions were specifically associated with CDP-DAG-synthetase cell membrane phospholipid biosynthesis. CDP-DAG-synthetase is an intra-membrane liponucleotide synthetase which has a crucial role in phospholipid biosynthesis via catalysing the formation of cytidine-diphosphate diacylglycerol (CDP-DAG) during cell membrane synthesis (Liu et al., 2014). A previous study which carried out whole genome sequencing of *Staphylococcus aureus* (*S. aureus*) strains exposed to daptomycin, identified mutations in the CDP-DAG-glycerol-3-phosphate 3-phosphatidyltransferase (*pgsA*) gene which resulted in a thicker cell wall and decreased susceptibility to the antibiotic (Peleg et al., 2012). This suggests the requirement of CDP-DAG- synthetase associated reactions for RW109 survival during exposure to POE at 50% of the MIC. It also corroborates that the cellular membrane is a key target for POE as a preservative, which has been noted in the literature (Gilbert et al., 1977, Fitzgerald et al., 1992, Langsrud et al., 2016).

When RW109 was exposed to POE at both 20% and 50% of the MIC, six cardiolipin synthase reactions were also predicted to be essential. Cardiolipin synthase is involved the synthesis of the bacterial inner membrane lipid cardiolipin (Zhang et al., 2014). The gene encoding for cardiolipin synthase was previously knocked out in a *Pseudomonas putida* (*P. putida*) mutant and the membrane was found to contain less cardiolipin (Bernal et al., 2007). Mutants, when compared to the parent strain, were more susceptible to several antibiotics and during exposure to ethidium bromide, an increased accumulation was identified within the mutant cells. This suggests that cardiolipin may play a role in impeding the preservative POE from crossing the cell membrane. These cardiolipin reactions were also predicted as essential when exposed to LD in combination with BIT at 20% of the MIC. Increased accumulation of cardiolipins have also been reported when bacteria are exposed to extreme environmental conditions (Czolkoss et al., 2016). Targeting cardiolipin components of the *P. aeruginosa* inner membrane with amphiphilic aminoglycoside derivatives, have proven successful in increasing the permeability and depolarization of the cell wall (El Khoury et al., 2017). Therefore, it would be beneficial to develop product preservation systems, which directly interact with cardiolipin-enriched areas to enhance the efficiency of POE preservatives.

Lipid metabolism reactions were also predicted to be essential when RW109 was exposed to LDOnly and LD+BIT. Six (3R)-3-Hydroxypalmitoyl-[acyl-carrier-protein]:NADP⁺ oxidoreductase reactions were identified which have a function in fatty acid biosynthesis elongation. These reactions were all associated with the FabG 3-oxoacyl-[acyl-carrier-protein] reductase gene in RW109 and mutations in this gene can confer resistance to the biocide triclosan which blocks the final step in fatty acid elongation by inhibiting FabG reductase enzymes (Khan et al., 2016). The essentiality of these reactions when exposed to laundry detergent with and without the preservative BIT suggests they are important targets to consider during preservation system development. Tea catechins and related plant polyphenols can act as effective

inhibitors of FabG reductase in a wide range of Gram-positive and Gram-negative bacterial species when used at relatively low concentrations (Zhang and Rock, 2004, Taylor et al., 2005). These additives could be applied as boosters to target the fatty acid biosynthesis reactions predicted to be essential for survival in BIT preservative systems of laundry detergents.

A high number of essential reactions in test exposure conditions were involved in nucleotide metabolism with the same reactions predicted as essential in numerous exposure conditions (Figure 60 and Table 38). Nucleoside-diphosphate kinase (Ndk) reactions involved in salvage pathways were commonly predicted as essential. These reactions function in balancing the nucleotide pool of the cell (Kamath et al., 2000) and have important functions in growth and virulence of *P. aeruginosa* under stress conditions (Chakrabarty, 1998). The Ndk enzyme supplies guanosine triphosphate (GTP) which is required for alginate biosynthesis. The overproduction of the alginate exopolysaccharide is associated with establishment of chronic infections in the CF lung by mucoid *P. aeruginosa* isolates (Kamath et al., 2000). Although not essential for biofilm formation, alginate has a notable effect on the structural integrity of biofilms produced by mucoid *P. aeruginosa* strains (Franklin and Ohman, 1993, Tielen et al., 2005, Flemming and Wingender, 2010). These important survival characteristics may also be used by RW109 to survive within the harsh conditions found in industrial products. The Ndk enzyme in bacteria is highly conserved (Spooner and Yilmaz, 2012) making it an ideal antimicrobial target. A preservative booster could be developed which disrupts the enzyme, enhancing the effectiveness of a preservative system. The RW109_01980 essential gene, which is linked to reactions involving Ndk, was also significantly up-regulated during exposure of RW109 to LD with and without BIT, highlighting the importance of the gene in surviving exposure to HPC preservatives and product formulations.

Reactions associated with membrane transport were also commonly predicted as essential across all test exposure conditions. Membrane transporters mediate the uptake of necessary substrates for growth and are also used to export metabolic end products and toxic substances out of the cell (Nikaïdo and Saier, 1992, Ma et al., 2003). The majority of the transporters identified by the analysis were characterised as ATP-binding cassette (ABC) protein exporters. ABC transporters, which transport proteins, can also have a secondary role in multidrug export across the cell membrane in Gram-negative bacteria (Ma et al., 2003, Venter et al., 2003, Rees et al., 2009). This suggests the ABC transporters predicted as essential could have a role in removing the preservatives from inside the cell.

A reaction involved in the transport of L-methionine via an ABC system (rxn05219) was identified as essential throughout all of the test conditions except for LD only (Table 39 and Table 40). An increased availability of methionine has been linked with an enhanced aptitude for reducing the effects of oxidative stress caused by the disinfectant hydrogen peroxide (Zhang et al., 2012, Romsang et al., 2013). The common essentiality prediction of a methionine transport system suggests a requirement for methionine during exposure of RW109 to industrial preservatives.

Xenobiotics biodegradation and metabolism reactions were only predicted as essential when RW109 was exposed to POE20 and BIT+POE, with all the implicated reactions functioning in benzoate degradation. It is probable the 5 reactions identified when BIT and POE were in combination both at 20% of the MIC, occurred due to the presence of POE within the formulation, as a limited number of reactions were predicted to be essential when exposed to BIT20 alone. POE is an aromatic ether alcohol which is a derivative of benzene (Fiege et al., 2000) suggesting the same bacterial pathways used to breakdown benzoate structures could be active on POE. Resistance to the antibiotic tigecycline (Tgc) which contains 4 aromatic six-carbon rings has been linked with an up-regulation of genes encoding for benzoate degradation enzymes in Gram-negative *Acinetobacter baumannii* clinical isolates (*A. baumannii*) (Liu et al., 2016). This suggests certain bacteria may use the benzene degradation pathways as a resistance mechanism which could be linked to surviving exposure to antimicrobial agents. The *P. aeruginosa* RW109_04051 gene associated with the essential rxn02143 benzoate degradation reaction was also significantly up-regulated in response to POE20 and BIT+POE, indicating its importance during exposure to these conditions.

Reactions involved in carbohydrate metabolism were predicted as essential in all conditions except with BIT20 and BIT50 (Table 42 and Figure 64). An important observation was the common essentiality of a glyoxylate and dicarboxylate metabolism reaction, involving the glycoaldehyde:NAD⁺ oxidoreductase enzyme, which breaks down glycoaldehyde into glycolate (rAB00001). The glyoxylate pathway has previously been identified as a principal metabolic pathway used by *P. aeruginosa* when maintaining infection in the CF lung (Hagins et al., 2011, Flynn et al., 2017). Transposon insertion mutants in *P. aeruginosa* genes involved in glyoxylate metabolism displayed decreased fitness when growing in media which represented the high level of mucins found in the CF lung (Flynn et al., 2017). The frequent identification of the glyoxylate metabolism reaction during the RW109 essentiality predictions suggests a possible role for this pathway in the ability to survive to exposure of the industrial conditions. Various essential reactions involved in alternative pathways for glycolysis and pentose phosphate were predicted during exposure to POE50, BIT+POE and LD (with and without BIT). This suggests the industrial growth conditions could be impairing the mainstream reactions of glycolysis and the pentose phosphate pathway which could result in RW109 having to use alternative pathways to survive in the presence of these conditions. Ethanol and isopropanol have been associated with inhibiting glycolysis enzymes (Russell, 2003c) and this could also be the mode of action used by industrial preservatives such as POE. These alternative reactions are also used in order to generate redox NADPH to counteract the effects of oxidative stress when exposed to harsh environments (Chavarría et al., 2013). A *P. aeruginosa* RW109 transposon mutant (Chapter 4) was characterised with an insertion in the essential RW109_02629 gene involved in the phosphogluconate dehydratase reaction (rxn01477). This reaction was identified when RW109 was exposed to POE at 20% and 50% of the MIC, and LD with and without BIT. Further analysis of this transposon mutant could be used to examine fitness in the presence of these conditions.

Amino acid metabolism reactions were predicted to be essential in all the test exposure conditions except for BIT20. An interesting observation was the predicted essentiality in response to BIT+POE of a glycine,

serine and threonine metabolism reaction (rxn08647), which was an additional reaction of the RW109 model in comparison to PA14. The glycerate 2-phosphotransferase enzyme involved in this reaction has functions in the detoxification of formaldehyde via a serine pathway (Chen et al., 2016, Kanehisa et al., 2017). BIT and POE preservatives are not formaldehyde associated agents, however a glutathione-dependent pathway with a possible role in formaldehyde detoxification was identified as a putative resistant determinant to the CMIT isothiazolinone preservative in *Burkholderia lata* (*B. lata*) (Thomas, 2011). Transposon-based interruption of a glutathione synthase gene resulted in increased isothiazalinone susceptibility (Thomas, 2011). This suggests formaldehyde detoxification systems may be used as a resistance mechanism of *P. aeruginosa* in response to the exposure of other industrial preservatives.

POE at 20% of the MIC had the greatest number of essential reactions (n= 6) within the amino acid category and one of the reactions uniquely essential was involved in the breakdown of agmatine into urea and putrescine with an agmatinase enzyme (rxn00858). It has previously been identified that putrescine reduces antibiotic-induced oxidative stress in *Burkholderia cenocepacia* (*B. cenocepacia*) (El-Halfawy and Valvano, 2014). A reduction in the ability to produce putrescine resulted in increased reactive oxygen species and also susceptibility to the antimicrobial peptide polymyxin B. Putrescine is able to protect less-resistant cells by competing with polymyxin B in binding to the surface of *B. cenocepacia* (El-Halfawy and Valvano, 2013). The RW109 strain may be producing putrescine in order to prevent POE from attaching to the cell membrane. One of the other reactions found to be essential with POE20 was the production of agmatine with the enzyme arginine decarboxylase (rxn00405). In a study by Chou et al. (2008) *P. aeruginosa* was grown in the presence of agmatine and demonstrated up to 16-fold increases in the MIC's of aminoglycosides and quinolones antibiotics along with the dyes ethidium bromide and acridine orange. It was proposed that this decreased susceptibility was due to the function of agmatines in the induction of expression levels of the *oprH-phoP-phoQ* operon of the PhoPQ two-component system which is involved in resistance to cationic antimicrobial peptides such as polymyxin B (Macfarlane et al., 1999, Chou et al., 2008). This two-component system was also identified as biologically feasible in the *P. aeruginosa* RW109 strain (Chapter 3, Section 3.2.3.2.1) and the predicted essentiality of the rxn00405 reaction when exposed to POE at 20% of the MIC supports the importance of agmatine production in resistance to the preservative.

A reaction involving the glycine hydroxymethyltransferase enzyme (rxn00692) was identified as essential when RW109 was exposed to POE at both 20 and 50% of the MIC. This enzyme has been associated with the assimilation of glycine, cell signalling processes and pyocyanin biosynthesis in *P. aeruginosa* PAO1 (Lundgren et al., 2013). The RW109_00291 gene, associated with the essential rxn00692 reaction, was encoded on the large plasmid 1 and was characterised as a component a glycine cleavage system (Chapter 3, Section 3.2.3.1.3). During environmental stresses and starvation of nutrients, glycine metabolism has been identified as a nutritional cue for induction of quorum-sensing phenotypes such as pyocyanin production, especially in the CF lung (Lundgren et al., 2013). This glycine utilisation virulence mechanism could also have a role in the survival of RW109 when exposed to POE.

Reactions encompassing the enzymes threonine synthase (rxn01069) and homoserine dehydrogenase (rxn01302), were predicted to be essential solely during exposure to LD with the preservative BIT. Both these reactions have associations with pyoverdine production as a key virulence characteristic of *P. aeruginosa* (Bartell et al., 2017, Calcott et al., 2014). It's known that pyoverdine production increases resistance of *P. aeruginosa* to the antibiotic tigecycline (Oglesby-Sherrouse et al., 2014) suggesting it may have a role during survival in the presence of industrial conditions. This resistance was associated with pyoverdine facilitating the uptake of iron which contributes to the ability *P. aeruginosa* to form biofilms (Banin et al., 2005). Interestingly a FeoB iron (II) transporter (rxn10963) was also deemed essential during exposure to LD in combination with BIT, adding a further link to iron scavenging. Overall, the essentiality of these reactions may be due to the requirement of RW109 to produce pyoverdine for biofilm development and iron acquisition during growth in industrial products and their preservation systems.

In all test conditions (excluding BIT20), a reaction involved in the biosynthesis of aromatic amino acids (rxn02213) via the 3-dehydroquinate dehydratase (DHQase) enzyme, was predicted as essential, implicating it as a possible enzyme to target for future preservations systems. A flavonoid polyketide has been identified which inhibits DHQase (Cheung et al., 2014). Therapeutics which target the DHQase in tuberculosis infections result in the successful inhibition of *Mycobacterium tuberculosis* (*M. tuberculosis*) (Frederickson et al., 1999). Inhibitors of this enzyme could be used as a potential preservative booster in industrial products.

A common reaction identified as essential in all test exposure conditions (except for BIT20) within the metabolism of cofactors and vitamins category was a tetrapyrrole biosynthesis pathway. The reaction requires a coproporphyrinogen oxidase enzyme (rxn02303) and mediates the oxidative decarboxylation of coproporphyrinogen III to form protoporphyrinogen IX in tetrapyrrole heme biosynthesis (Rompf et al., 1998). The inhibition of the heme biosynthesis by antimicrobials such as the metalloid tellurite leads to the accumulation of the pathway intermediates and the resulting build up is toxic to the cell (Morales et al., 2017). The enzyme involved in this reaction is a potential antimicrobial target of relevance for enhancing the effects of preservation systems and is also highly conserved in prokaryotic genomes (Dailey et al., 2015). The RW109_00614 essential gene linked to the rxn02303 reaction was also significantly up-regulated during exposure to POE at 20% and 50% of the MIC and BIT and POE in combination, corroborating the importance of the gene for survival in industrial environments.

The integration of transcriptomic data into the *P. aeruginosa* RW109 metabolic model provided a valuable insight to predicted essential metabolic reactions when the strain was exposed to industry relevant conditions. The key predicted essential pathways identified in this study are shown in Figure 68. The wide-ranging variation in the biological functions of essential reactions highlights the multifactorial nature of *P. aeruginosa*'s resistance and survival mechanisms. Common prediction of essential pathways such as those associated with the enzymes NdK and DHQase, also highlights important targets to inform the development of targeted preservation systems. These highlighted essentiality predictions and implicated

pathways require validation, using techniques such as the fitness evaluation of transposon insertion mutants when exposed to industry relevant conditions.

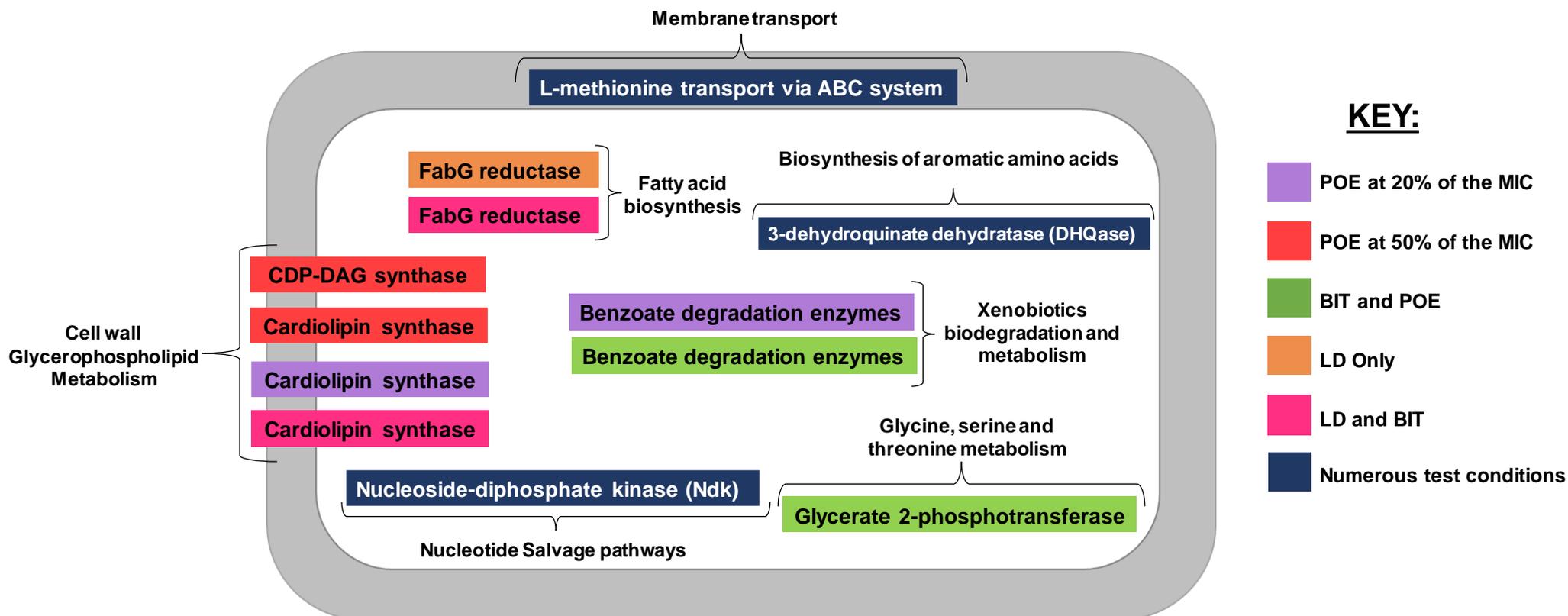


Figure 68. Key enzymes and pathways linked to reactions projected to be essential during exposure of RW109 to industry relevant conditions

Key enzymes and pathways associated with predicted essential reactions are colour coded according to the industrial test condition where they were identified (See Key). This includes POE at 20% (purple) and 50% of the MIC (red), BIT and POE (green), LD only (orange) and LD and BIT (pink). Also illustrated were reactions commonly predicted as essential in response to multiple test conditions (dark blue) such as the enzymes NdK and DHQase and the ABC transport pathway for L-methionine.

6.3.5 Preservative concentration greatly influences essentiality predictions

Interesting observations of this study were the differences in the prediction of essential reactions when the preservative POE was screened at 20% and 50% of the MIC (Figure 67 and Table 46). Reactions involved in benzoate degradation were predicted as essential at the lower concentration but were not identified during exposure to 50% of the MIC, where a greater number of lipid metabolism reactions were observed. At 20% of the MIC, the strain may be able to grow more and degrade POE, but when exposed to higher concentrations the reactions involved in cell membrane phospholipid biosynthesis become increasingly essential. This suggests that higher concentrations result in greater damage to the cell membrane, which requires the strain to focus its efforts on fixing the membrane rather than degrading the preservative. Membrane transport reactions were predicted to be essential in at both 20% and 50% of the MIC; however, more were identified at the higher concentration suggesting a greater requirement to move substances across the membrane.

The essentiality of alternative glycolysis reactions were observed when exposed to 50% of the MIC. A higher number of essential amino acid metabolism reactions were identified during exposure to POE at 20% of the MIC, whereas at 50% MIC a lower number of amino acid reactions were found to be essential. This variations observed in the essentiality of metabolic pathways at 20% and 50% of the MIC with POE, highlights the alteration in the response of a strain when exposed to the same preservative used at different concentrations. The concentration dependant activity of antibiotics is well characterised (Bernier and Surette, 2013) and results from this study indicate the mechanism of action of POE is also heavily influenced by its concentration. It has previously been noted that POE antimicrobial activity is concentration dependant and bactericidal effects result from the preservative interfering with several targets, causing irreversible cell damage (Langsrud et al., 2016). This signifies the importance of concentration when formulating preservative systems and demonstrates the necessity to predict bacterial responses to industrial preservatives. It also highlights a key challenge in designing research strategies based on microbial growth for evaluation of preservative action.

6.4 Conclusions

The main conclusions from this chapter were as follows:

- 1) The PacBio sequence of the RW109 strain was utilised to accurately update the PA14 metabolic model (Bartell et al., 2017), enabling a representation of the metabolic capabilities of an industrial *P. aeruginosa* strain for the first time.
- 2) The model provides important insight into the biological characteristics of the industrial *P. aeruginosa* RW109 strain. Reactions associated with lipid metabolism were found to be the most abundant and in comparison to the PA14 model. RW109 had additional reactions within the xenobiotics biodegradation and metabolism functional category.
- 3) Transcriptomic gene expression data was successfully integrated into the RW109 model reconstruction to produce condition specific models to predict the essentiality of reactions during exposure to industry relevant conditions.
- 4) Exposure to the preservative POE at 50% of the MIC resulted in the highest number of predicted essential reactions with the majority having a function lipid metabolism. Reactions involved in cell membrane phospholipid biosynthesis via the CDP-DAG-synthetase enzyme were identified as uniquely essential when RW109 was exposed to POE at 50% of the MIC.
- 5) The reactions predicted to be essential throughout the industrial exposure conditions had biologically diverse functions, suggesting a multifactorial antimicrobial response in *P. aeruginosa*. However, within this diversity the enzymes such as Ndk and DHQase were commonly predicted in essential reactions, and represent potential novel targets to enhance the effectiveness of preservative systems.
- 6) Preservative concentration heavily influences the predicted essentiality of different metabolic pathways. Dissimilarities were observed in the reactions predicted to be essential when exposed to different concentrations of POE; at a lower concentration reactions involved in degradation of xenobiotics were essential and with an increase in concentration, cell membrane phospholipid metabolism reactions were considered as crucial. This suggests *P. aeruginosa* is capable of responding to the concentration-dependant mechanism of action of POE.

7. General conclusions, discussion and future research

7.1 Conclusions

This study represents a novel and comprehensive investigation of an industrial *P. aeruginosa* strain, RW109. A variety of research techniques were implemented to examine the genetic and metabolic pathways, involved in the ability of RW109 to grow in the presence of industrial preservatives and HPC product formulations. A complete genome of RW109 was attained for functional annotation and for comparison with a panel of *P. aeruginosa* genome sequences representative of strains from clinical, environmental and industrial sources. The genetic basis of industrial preservative resistance was also investigated via the construction of a mini-Tn5-*luxCDABE* transposon mutant library, and in more depth with global RNA-Seq gene expression analysis. Finally, a metabolic model specific to RW109 was constructed by updating an existing *P. aeruginosa* PA14 genome-scale biochemical network representation. The RW109 RNA-Seq transcriptomic data was successfully integrated into this model, allowing for the prediction of reaction essentiality during exposure to HPC preservatives and product formulations. Overall, the results from this study highlight the utility of incorporating multiple research techniques when investigating the ability of industrial *P. aeruginosa* strains to resist preservatives and product formulations.

The main conclusions from the study were as follows:

- 1. The complete genome sequence analysis of the industrial strain RW109 revealed it's larger than average sized genome, when compared to a panel of 102 *P. aeruginosa* genome sequences isolated from clinical, environmental and industrial sources (Chapter 3).**

This was the first study to obtain a whole genome sequence of an industrial *P. aeruginosa* strain specifically isolated from a home and personal care product. PacBio SMRT sequencing technology was successfully used to acquire a complete genome of RW109, which comprised a main chromosome and two plasmids. This work also allowed accurate functional annotations, which provided an important insight into the biological characteristics of the strain. The 7.8 Mbp size of RW109, was the largest genome identified within the *P. aeruginosa* panel used in this study (Figure 69). This genome was also bigger than the reported genome size range of 5.2 to 7.0 Mbp for this microorganism (Schmidt et al., 1996, Stover et al., 2000, Juhas et al., 2005, Winsor et al., 2016). Another interesting observation was the significantly bigger mean genome size of the industrial strains in the panel when compared to those from clinical and environmental settings. The functional characteristics of the RW109 strain were also compared to PA14 (He et al., 2004) and PAO1 (Stover et al., 2000), the most studied clinical reference strains for *P. aeruginosa*. The comprehensive functional annotations and comparisons obtained with RW109 present this isolate as an appropriate reference strain for the HPC industry.

2. When investigating the genetic basis of preservative exposure with the RW109 strain, the use of a mini-Tn5-*luxCDABE* mutant library proved to be not as informative and was less reliable when compared to RNA-Seq analysis (Chapter 4 and 5).

The RW109 strain was successfully used to construct a mini-Tn5-*luxCDABE* mutant library and the distinct transposon insertion sites identified were evenly distributed throughout the whole genome sequence. Screening the mutant bank with the preservative BIT at 20% of the MIC for 24-hours, identified mutants of interest, which were isolated, and insertion sites identified. However, an overall fold decrease in light emission was observed when the mutant bank was screened with the preservative POE as well as increased concentrations of BIT. When the mutant bank screening condition was analysed with RNA-Seq, only two genes with significantly altered expression were identified. This indicates that BIT at 20% of the MIC was suitable at identifying mutants of interest during mutagenesis screening, but was too low to induce major changes in gene expression when used during global transcriptomic analysis. RNA-Seq was also useful for investigating gene expression with POE, higher BIT concentrations along with HPC product formulations. Furthermore, the difficulties with identifying the insertions sites in mutants of interest, was also a drawback with using the mini-Tn5-*luxCDABE* library in this study. The mutant bank was however not used to identify mutants, which were susceptible or more tolerant to preservatives which maybe a useful tool for this type of investigation.

3. Global gene expression analysis with RNA-Seq revealed both condition specific and consistent transcriptomic responses of RW109, following exposure to industry related conditions (Chapter 5).

Obtaining RNA from bacterial cells can be challenging, especially if chemically harsh compounds such as preservatives or HPC product formulations are present in the growth medium. In this study, good quality RNA was recovered following a 24-hour exposure of RW109 to conditions which reduced the percentage OD by up to 68%, in comparison to TSB only growth. This enabled a novel comparative RNA-Seq gene expression analysis to be carried out with a wide range of industrial conditions. The successful identification of DEGs provided an increased understanding of the complex genetic networks utilised by RW109, when exposed to HPC preservatives and product formulations. An interesting observation was the consistent up-regulation of the same operon genes in response to different test conditions, examples include those encoding for an efflux pump, a sialic acid transporter and acyclic isoprenoid degradation enzymes (Figure 69). These findings suggest the RW109 strain potentially exploits a consistent set of resistance mechanisms for survival in a variety of harsh industrial environments. In order to validate if the mechanisms identified represent strain specific mechanisms, transcriptomic comparisons with other industrial strains would need to be carried out. To enhance the sterility of HPC products preservation systems could be developed which specifically target commonly used resistance mechanisms. The transcriptomic changes observed in this study still require further validation with techniques such as quantitative RT-PCR.

- 4. A genome-scale metabolic model specific to an industrial *P. aeruginosa* was constructed for the first time and used to incorporate global gene expression data from growth under industrial conditions. This work provided valuable insights into the metabolic characteristics of RW109 and was used to predict reactions essential for the strain to survive in the presence of HPC preservatives and product formulations (Chapter 6).**

The preliminary construction of an industrial metabolic model was the first of its kind for a *P. aeruginosa* strain isolated from a HPC product. This form of analysis was only achievable following the successful whole genome sequencing and transcriptomic analysis of the RW109 strain, along with the availability of a recently published *P. aeruginosa* genome-scale metabolic model for the PA14 strain (Bartell et al., 2017). This work offers a good example of how results generated with different research techniques can be linked together to provide a powerful investigative tool. Building the RW109 metabolic model, increased the understanding of the biological characteristics of an industrial strain. There are over 100 published and experimentally validated metabolic models which represent various microorganisms (Broddrick, 2017). For the important *P. aeruginosa* clinical reference strains PAO1 and PA14, mapping of the biochemical networks provided a valuable insight into the metabolic potential of the strains and how they respond to various environmental conditions (Oberhardt et al., 2008, Oberhardt et al., 2011, Bartell et al., 2017). The integration of transcriptomic data into the RW109 metabolic model allowed the development of condition specific models. A valuable outcome of the study was the prediction of reactions classified essential for RW109 to survive in the industrially relevant conditions. POE at half the MIC resulted in the highest number of reactions expected to be vital for the strain to persist, with notable reactions involved in cell membrane phospholipid biosynthesis. This analysis also predicted the essentiality of common reactions, which highlighted enzymes for consideration as potential targets for preservation strategies (Figure 69). The results from this chapter require further validation of the predicted fundamental pathways via mutagenesis *in vitro* experimentation.

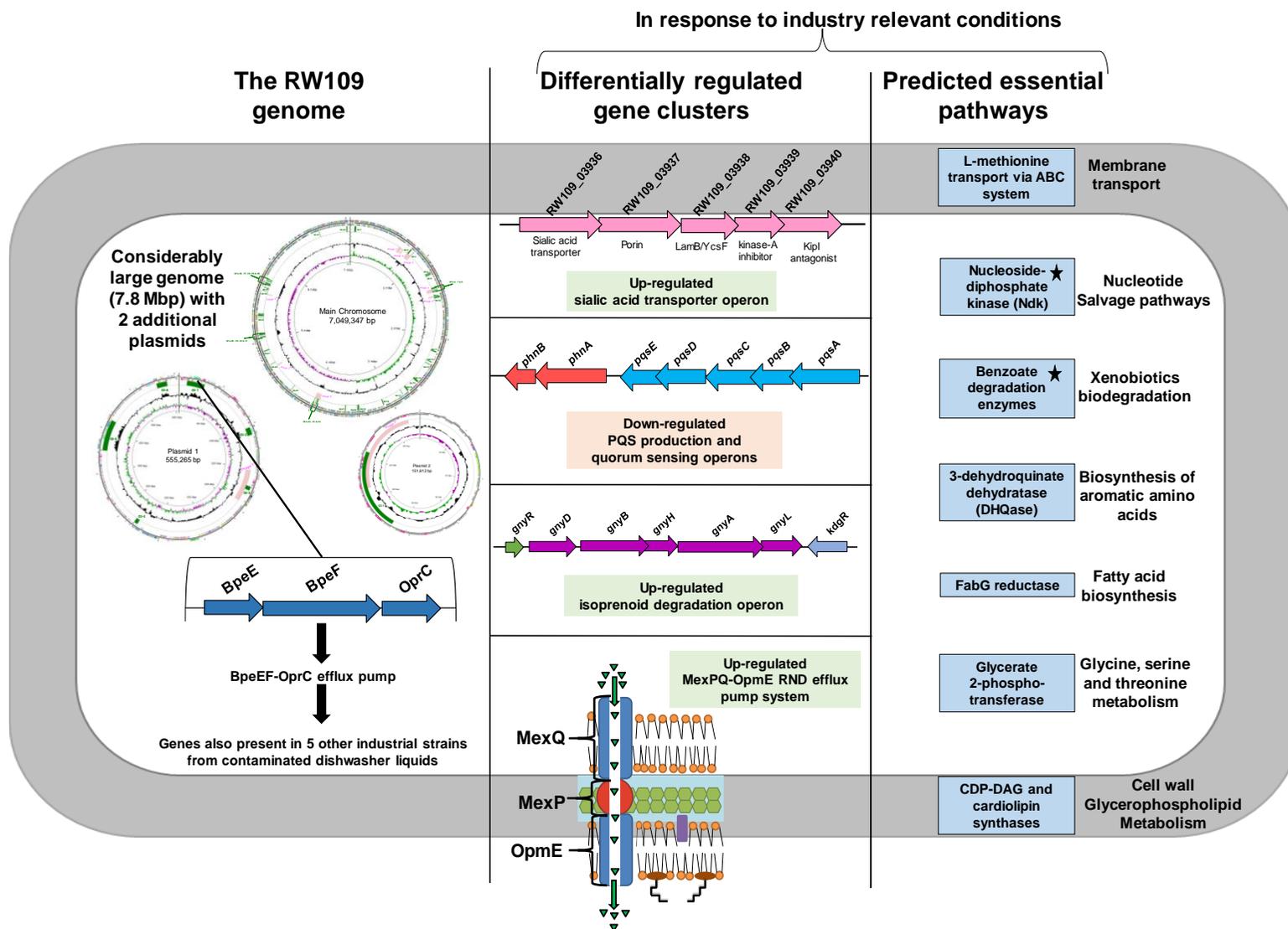


Figure 69. Key features of the industrial *P. aeruginosa* strain RW109

This schematic provides an overview of the genome traits, differentially regulated gene clusters and predicted essential pathways, which are key to the success of RW109 as an industrial contaminant. Those to note include the large genome size, additional plasmids, common up-regulation of the MexPQ-OpmE efflux system and frequently predicted essential pathways associated with nucleotide salvage and cell wall glycerophospholipid metabolism. The predicted essential pathways with an adjacent star, indicate those associated with a gene which had increased expression in response to industry relevant conditions.

7.2 Discussion and future research

The resistance of *P. aeruginosa* to antibiotics has been well studied and there is a comprehensive understanding of the mechanisms used by the bacterium to survive antibiotic exposure. This information has permitted the optimisation of clinical treatment strategies that minimise the occurrence of antimicrobial resistance, whilst also clearing infection effectively (Monogue et al., 2016, Paterson et al., 2016). Conversely, details about the genetic and metabolic pathways used by *P. aeruginosa* in response to HPC industrial preservative formulations are not well characterised. A few resistance mechanisms in response to isothiazolinones HPC preservatives have been characterised such as alterations in the *P. aeruginosa* outer membrane structure (Brözel and Cloete, 1994), upregulation of an RND efflux pump in *Burkholderia lata* (Rushton et al., 2013) and increased biofilm formation in *Burkholderia cepacia* (Costerton and Lashen, 1984). Recent transcriptomic investigations of a isothiazolone-resistant *P. aeruginosa* strain when exposed to the CMIT/MIT isothiazolone, identified a number of differentially expressed genes associated with nitrogen metabolism and oxidative phosphorylation pathways (Zhou et al., 2016).

This PhD project provided beneficial understandings of the functional characteristics of an industrial *P. aeruginosa* strain, and identified genetic pathways and metabolic networks following exposure to HPC preservatives and product formulations (Figure 69). The overall results of the study are discussed below with reference to how the research could be translated into industrial practices. Future directions for investigating the genetic basis of *P. aeruginosa* industrial preservative resistance are also suggested.

7.2.1 The use of whole genome sequencing to extensively characterise an industrial *P. aeruginosa* strain
Obtaining a complete genome for the RW109 strain facilitated the comprehensive functional classification and comparative analysis of an industrially prominent isolate. This study provides a well characterised industrial reference strain for the future whole genome sequencing of other industry isolates. The HPC industry would benefit from the development of a database containing complete whole genomes of key strains, which are sequenced, assembled and annotated using a parallel pipeline. During industrial contamination incidents, the routine sequencing of suspected *P. aeruginosa* strains could be used to produce a comprehensive genome collection of industry isolates. This would enable the generation of a genome database of industrial isolates, to assist in establishing the functional characteristics, which allow strains to contaminate HPC products.

A large *P. aeruginosa* genome is a possible indicator of a strain's ability to reside within an industry setting; this interesting observation was recognised for the first time by this study. The RW109 genome was considerably larger than all 102 strains analysed (Figure 69), and those isolated from an industrial setting had a significantly larger mean genome size when compared to clinical and environmental strains. This result suggests a large genomic size as a possible predictor of a *P. aeruginosa* strain's ability to cause industrial contamination. However this investigation only included a limited number of industrial strains (n=16); this further accentuates the requirement for incorporating genome sequencing into routine industrial practices when analysing a *P. aeruginosa* contamination incident. Understandably, the cost of

CHAPTER 7 – General conclusions, discussion and future research

sequencing is an obstacle, but with the current exponential development of new methods in this field, which are increasingly accurate and cost-effective (Levy and Myers, 2016, Klemm and Dougan, 2016), sequencing integration may become achievable. The application of whole genome sequencing to collections of *P. aeruginosa* isolates has the ability to provide vital information on resistance profiles, aiding in the prediction of antimicrobial resistance (Jeukens et al., 2017, Winsor et al., 2016, Freschi et al., 2015).

Future research should also examine the plasmids within *P. aeruginosa* strains isolated from HPC products. The RW109 strain was unusual in having a very large plasmid (555 kb) as well as a smaller plasmid (152 kb) (Figure 69) and pathways were differentially expressed on both these replicons when exposed to industry relevant conditions (Chapter 5). The RW130, RW146, RW172, RW176 and RW184 isolates from dishwasher liquid products all encoded the KEGG BpeEF-OprC efflux pump module (M00698). The 5 genes associated with this KEGG module were identified on plasmid 1 of the RW109 strain (Figure 69) suggesting these 5 industrial isolates may also incorporate an additional plasmid carrying this locus. The functionality of the efflux pump should also be investigated via susceptibility analysis of the *P. aeruginosa* isolates to the trimethoprim and chloramphenicol, the substrates for the BpeEF-OprC system in *B. pseudomallei* (Kumar et al., 2006, Podnecky et al., 2013). Transcriptomic analysis of the efflux pump operon did not reveal an altered gene expression when exposed to the preservatives and product formulations investigated in this study. However, it may be used in response to alternative industrial antimicrobials and if the pump was knocked out, the susceptibility of the mutants to preservatives could be investigated. The removal of the two plasmids from RW109 may also form a strategy, to determine their importance for the survival of the strain in the presence of industry relevant conditions. The curing of the plasmids could be carried out as described by (Raja and Selvam, 2009, Durve et al., 2013), where plasmid removal resulted in increased susceptibility to antibiotics and heavy metals in *P. aeruginosa* strains isolated from the environment.

7.2.2 The efficacy of an industrial *P. aeruginosa* mini-Tn5-*luxCDABE* mutant library when investigating the genetic basis of industrial resistance

An aim of Chapter 4 was to develop a transposon mutant bank in the industrial *P. aeruginosa* strain that could be used to understand differentially regulated genes when exposed to industrial preservatives. However, the use of a mini-Tn5-*luxCDABE* mutant library as a luminescent reporter did not prove an appropriate research tool for investigating the genetic basis of preservative resistance in this way. The *lux* reporter strategy had successfully been used as a direct method of identifying differentially regulated genes in the *P. aeruginosa* strain PAO1. Mutants with differential light emission responses when exposed to MgSO₄, magnesium and phosphate-limiting conditions were recognised and linked to genes with altered expression levels (McPhee et al., 2003, Lewenza et al., 2005). However, in this study the RW109 *lux* mutants could not be screened with the preservative POE as well as high concentrations of BIT and this limited screening to one condition involving BIT at 20% of the MIC. Even though mutants with altered light emission responses could be identified, difficulties were encountered with identification of transposon insertion sites.

As applied, the *lux*-reporter transposon method did not provide a genome-scale understanding of the genetic basis of preservative resistance. Out of 404 mutants of interest, the putative gene function was successfully characterised in 157, which resulted in only 53 distinct insertion sites being recognised. Key mutants with increased light emission responses when screened with BIT had insertions in genes encoding bacterial solute binding proteins, a ferrous-iron efflux pump, a KdgR transcriptional regulator and the *napA* gene within a *napEFDABC* operon. Those with decreased light emission responses were associated with an acid active urea channel and a tripartite ATP-independent periplasmic transporter. For a complete validation of functional knockouts from this study, complementation can be carried out to account for polar effects or targeted gene knockout mutants could be generated.

To achieve the characterisation of the small number of insertion sites identified in this study was a time consuming process. The initial attractiveness of the approach was the possibility of building a mutant panel representing key genes, where the expression responses to a variety of industry relevant conditions were characterised. This panel could then be passed onto industry for use as a direct measurement of gene expression during preservative formulation development. However, the results of the study show that the *luxCDABE* operon was possibly impaired by the activity of preservatives and not enough mutant insertion sites were characterised to identify key genetic functions.

Alternatively, as mutant insertions were evenly dispersed throughout the RW109 genome, they could be used to assess functionality by incorporating industry relevant conditions into agar-based mutant screens. Furthermore, there is the requirement to characterise the insertion sites in the remaining mutants of interest. This could be achieved via whole genome sequencing or via an alternative approach such as TraDIS could be implemented to develop a large well-characterised mutant library (Langridge et al., 2009, Barquist et al., 2015). This future work could also enhance the gene essentiality predictions when RW109 was exposed to industry relevant conditions (Chapters 6).

7.2.3 The use of RNA-Seq to characterise the global gene expression response of an industrial *P. aeruginosa* strain in reaction to industry relevant conditions

The RNA-Seq gene expression analysis identified several key genetic pathways when RW109 was exposed to HPC preservatives and product formulations (Figure 69). The results provided a better understanding of the strains success as an industrial contaminant and pinpointed common resistance mechanisms used when exposed to different industrial conditions. In order to enhance the antimicrobial capabilities of product formulations, knowledge of frequently used resistance mechanisms would enable the selection of appropriate preservative enhancing agents. The consistent up-regulation of genes encoding the MexPQ-OpmE RND efflux pump was observed during transcriptomic analysis (Chapter 5) highlighting it as a possible target for the HPC industry during preservative system development (Figure 69). The success of an enhancing agent at increasing the susceptibility of *B. lata* to the isothiazolinone CITMIT, was demonstrated when the bacterium was exposed to the Efflux Pump Inhibitor (EPI) Phenylalanine-Arginine Beta-Naphthylamide (PAβN), together with the preservative (Rushton et al., 2013). The use of EPIs such as PAβN in antibiotic treatment strategies have proven successful against *P. aeruginosa* (Liu et al., 2010, Lamers et al., 2013). Targeting efflux systems may therefore be an appropriate method to enhance the effectiveness of HPC antimicrobial formulations. The growth dynamics and gene expression alterations of RW109 in the presence of the RNA-Seq industrial conditions combined with an EPI should be investigated in future research.

Implementing transcriptomic analysis techniques such as RNA-Seq, into industrial practices should be considered to further the understanding of resistance mechanisms used by *P. aeruginosa* in response to preservation formulations. During the development of an antimicrobial formulation, global gene expression investigations would enable the selection of preservatives and enhancing agents, which work well in combination. This study also demonstrated resistance mechanisms used when RW109 was exposed to an un-preserved laundry product; these identified genes may also require consideration during preservation development. Understandably, the cost and complexity of transcriptomic analysis techniques (Ozsolak and Milos, 2011, Conesa et al., 2016) may be a hindrance when incorporating such methods into HPC industrial practices. Nonetheless, the wide-ranging practicality of the results generated should be taken into consideration. The development of a well-defined experimental pipeline comparable to the one used in this study, together with bioinformatics computational platforms such as CLIMB (Connor et al., 2016), would enable the use of RNA-Seq in industry.

7.2.4 The construction of an industrial *P. aeruginosa* genome-scale metabolic model for predicting the essentially of reactions during exposure to industry relevant conditions

An initial genome-scale network reconstruction for an industrial *P. aeruginosa* strain was constructed (Chapter 6), expanding the understanding of RW109 metabolic functionality. Up-dating the PA14 metabolic model (Bartell et al., 2017) to better represent an industrial strain was greatly enhanced by the high-quality complete RW109 genome sequence (Chapter 3). This work also benefited from the extensive transcriptomic analysis (Chapter 5), which permitted *in silico* gene essentiality simulations of RW109, when exposed to industry relevant conditions. The resulting condition specific models were productively analysed to propose essential genes and reactions for use as candidate HPC preservation system targets (Figure 69).

Metabolic modelling is highly applicable to HPC industry practices and the study provides a preliminary genome-scale metabolic model incorporating gene-protein-reaction associations for an important industrial *P. aeruginosa* strain. In the food industry, metabolic modelling for predicting pathogen behaviour in response to environmental conditions has proven beneficial for characterising essential genes and novel antimicrobial targets. In a newly constructed model for *Clostridium difficile* (*C. difficile*), transcriptomic data was integrated to represent the gene expression alterations when exposed to a variety of environmental conditions. (Kashaf et al., 2017). Gene essentiality predictions identified novel therapeutic targets in *C. difficile*, such as phosphoglycerate kinase, glycerate kinase and bi-functional carbon monoxide dehydrogenase enzymes (Kashaf et al., 2017).

The industrial partner will take on the RW109 genome-scale and condition-specific metabolic models for further expansion and validation. Incorporating transcriptomic analysis into the industrial development of different HPC preservation formulations, could allow the production of additional condition specific models. This permits further predictions of novel genes and reactions considered essential for the design of targeted industrial antimicrobial formulations. Furthermore, exploiting whole genome sequencing of industrially relevant *P. aeruginosa* strains, could potentially allow the development of a collection of metabolic models for different isolates. This has the capacity to inform the design of preservation systems with increased relevance to a wide range of strains, which may vary in response to HPC formulations.

Even though the metabolic model only represented 18.2% of the RW109 genome, the genes and reactions predicted to be essential can be used to generate hypotheses for focused driven validation investigations. For example, future research to assess the essentiality of genes associated with the enzymes Ndk and DHQase (Figure 69), which were commonly predicted as essential during exposure to a wide range of industrial conditions should be carried out. Also necessary is to confirm the importance of cell membrane phospholipid biosynthesis enzymes such as CDP-DAG-synthetase (Figure 69) in response to POE at half the MIC. Essential genes and reactions are considered promising targets for the development of novel directed antimicrobials, due to their importance in bacterial survival (Juhás et al., 2012).

CHAPTER 7 – General conclusions, discussion and future research

To validate key essentiality predictions and implicated pathways, the fitness evaluation of transposon insertion mutants during exposure to an industry relevant condition could be examined. This verification approach of genome-scale metabolic modelling has proven worthwhile in the characterisation of essential pathogenic genes when subjected to environmental conditions, which represent the host. Transposon mutagenesis and genome-scale metabolic modelling were successfully used to confirm the essentiality of genes involved in leucine, histidine, glycine and proline amino acid production when *Neisseria meningitidis* was grown in conditions to mimic growth in blood (Mendum et al., 2011). The PA14 transposon mutant bank database (Liberati et al., 2006) could also be exploited for further validation and investigation of the key genetic pathways identified in this study. This database provides a genome-wide set of *P. aeruginosa* mutations with fully characterised insertions. Since the PA14 metabolic model was used as a basis to map the biochemical networks of RW109, the library will be useful to assess the essentiality of the mutants when exposed to industry related conditions.

In summary, this PhD project has for the first time developed a holistic genomic and metabolic understanding of an industrial *P. aeruginosa* strain. In addition to these findings, this undertaking also enabled broad training related to genome sequencing, bioinformatics, transcriptomic gene expression analysis and metabolic modelling, which are key skills for a successful career in molecular microbiology.

References

- ABDEL-MALEK, S. & BADRAN, Y. 2010. *Pseudomonas aeruginosa* PAO1 adapted to 2-phenoxyethanol shows cross-resistance to dissimilar biocides and increased susceptibility to antibiotics. *Folia Microbiologica*, **55**, 588-592.
- AENDEKERK, S., DIGGLE, S. P., SONG, Z., HØIBY, N., CORNELIS, P., WILLIAMS, P. & CAMARA, M. 2005. The MexGHI-OpmD multidrug efflux pump controls growth, antibiotic susceptibility and virulence in *Pseudomonas aeruginosa* via 4-quinolone-dependent cell-to-cell communication. *Microbiology*, **151**, 1113-1125.
- AENDEKERK, S., GHYSELS, B., CORNELIS, P. & BAYSSE, C. 2002. Characterization of a new efflux pump, MexGHI-OpmD, from *Pseudomonas aeruginosa* that confers resistance to vanadium. *Microbiology*, **148**, 2371-2381.
- AHMAD, A., PATEL, I. & KHAN, M. U. 2017. Pharmaceutical waste and antimicrobial resistance. *The Lancet Infectious Diseases*, **17**, 578-579.
- AIRES, J. R., KÖHLER, T., NIKAIDO, H. & PLÉSIAT, P. 1999. Involvement of an active efflux system in the natural resistance of *Pseudomonas aeruginosa* to aminoglycosides. *Antimicrobial Agents and Chemotherapy*, **43**, 2624-2628.
- ALLESEN-HOLM, M., BARKEN, K. B., YANG, L., KLAUSEN, M., WEBB, J. S., KJELLEBERG, S., MOLIN, S., GIVSKOV, M. & TOLKER-NIELSEN, T. 2006. A characterization of DNA release in *Pseudomonas aeruginosa* cultures and biofilms. *Molecular Microbiology*, **59**, 1114-1128.
- ALMAGRO-MORENO, S. & BOYD, E. F. 2009. Insights into the evolution of sialic acid catabolism among bacteria. *BMC Evolutionary Biology*, **9**, 118.
- ALLOUSH, V., NAVON-VENEZIA, S., SEIGMAN-IGRA, Y., CABILI, S. & CARMELI, Y. 2006. Multidrug-resistant *Pseudomonas aeruginosa*: risk factors and clinical impact. *Antimicrobial Agents and Chemotherapy*, **50**, 43-48.
- ÁLVAREZ-LERMA, F., MAULL, E., TERRADAS, R., SEGURA, C., PLANELLS, I., COLL, P., KNOBEL, H. & VÁZQUEZ, A. 2008. Moisturizing body milk as a reservoir of *Burkholderia cepacia*: outbreak of nosocomial infection in a multidisciplinary intensive care unit. *Critical Care*, **12**, R10.
- ALVAREZ-ORTEGA, C. & HARWOOD, C. S. 2007. Responses of *Pseudomonas aeruginosa* to low oxygen indicate that growth in the cystic fibrosis lung is by aerobic respiration. *Molecular Microbiology*, **65**, 153-165.
- ALVAREZ-RIVERA, G., DAGNAC, T., LORES, M., GARCIA-JARES, C., SANCHEZ-PRADO, L., LAMAS, J. P. & LLOMPART, M. 2012. Determination of isothiazolinone preservatives in cosmetics and household products by matrix solid-phase dispersion followed by high-performance liquid chromatography–tandem mass spectrometry. *Journal of Chromatography A*, **1270**, 41-50.
- AMOS, G. C., HAWKEY, P., GAZE, W. H. & WELLINGTON, E. 2014. Waste water effluent contributes to the dissemination of CTX-M-15 in the natural environment. *Journal of Antimicrobial Chemotherapy*, **69**, 1785-1791.
- ANDERS, S., PYL, P. T. & HUBER, W. 2015. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics*, **31**, 166-169.
- ANGULO, F. J., BAKER, N. L., OLSEN, S. J., ANDERSON, A. & BARRETT, T. J. 2004. Antimicrobial use in agriculture: controlling the transfer of antimicrobial resistance to humans. *Seminars in Pediatric Infectious Diseases*, **15**, 78-85.
- APIDIANAKIS, Y. & RAHME, L. G. 2009. *Drosophila melanogaster* as a model host for studying *Pseudomonas aeruginosa* infection. *Nature Protocols*, **4**, 1285.
- ARAI, H. 2011. Regulation and function of versatile aerobic and anaerobic respiratory metabolism in *Pseudomonas aeruginosa*. *Frontiers In Microbiology*, **2**.
- ARENDET, W., HEBECKER, S., JÄGER, S., NIMTZ, M. & MOSER, J. 2012. Resistance phenotypes mediated by aminoacyl-phosphatidylglycerol synthases. *Journal of Bacteriology*, **194**, 1401-1416.
- ARMALYTĚ, J., JURĚNAITĚ, M., BEINORAVIČIŪTĚ, G., TEIŠERSKAS, J. & SUŽIEDĚLIENĚ, E. 2012. Characterization of *Escherichia coli* dinJ-yafQ toxin-antitoxin system using insights from mutagenesis data. *Journal of Bacteriology*, **194**, 1523-1532.
- ARNDT, D., GRANT, J. R., MARCU, A., SAJED, T., PON, A., LIANG, Y. & WISHART, D. S. 2016. PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Research*, **44**, W16-W21.
- ARTIMO, P., JONNALAGEDDA, M., ARNOLD, K., BARATIN, D., CSARDI, G., DE CASTRO, E., DUVAUD, S., FLEGEL, V., FORTIER, A. & GASTEIGER, E. 2012. ExpASY: SIB bioinformatics resource portal. *Nucleic Acids Research*, **40**, W597-W603.
- ASKOURA, M., MOTTAWEA, W., ABUJAMEL, T. & TAHER, I. 2011. Efflux pump inhibitors (EPIs) as new antimicrobial agents against *Pseudomonas aeruginosa*. *Libyan Journal of Medicine*, **6**, 5870.
- AVRAIN, L., MERTENS, P. & VAN BAMBEKE, F. 2013. RND efflux pumps in *P. aeruginosa*: an underestimated resistance mechanism. *Antibiotic susceptibility*, **26321**, 26-28.
- BAIRD, R. 2007. Microbial spoilage, infection risk and contamination control. *Hugo and Russell's: Pharmaceutical Microbiology, Seventh Edition*, 263-284.
- BALASUBRAMANIAN, D., SCHNEPER, L., KUMARI, H. & MATHEE, K. 2012. A dynamic and intricate regulatory network determines *Pseudomonas aeruginosa* virulence. *Nucleic Acids Research*, **41**, 1-20.

REFERENCES

- BANIN, E., VASIL, M. L. & GREENBERG, E. P. 2005. Iron and *Pseudomonas aeruginosa* biofilm formation. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 11076-11081.
- BARDOU, P., MARIETTE, J., ESCUDIÉ, F., DJEMIEL, C. & KLOPP, C. 2014. Jvenn: an interactive Venn diagram viewer. *BMC Bioinformatics*, **15**, 293.
- BARQUIST, L., MAYHO, M., CUMMINS, C., CAIN, A. K., BOINETT, C. J., PAGE, A. J., LANGRIDGE, G. C., QUAIL, M. A., KEANE, J. A. & PARKHILL, J. 2015. The TraDIS toolkit: sequencing and analysis for dense transposon mutant libraries. *Bioinformatics*, **32**, 1109-1111.
- BARTELL, J. A., BLAZIER, A. S., YEN, P., THØGENSEN, J. C., JELSBÆK, L., GOLDBERG, J. B. & PAPIN, J. A. 2017. Reconstruction of the metabolic network of *Pseudomonas aeruginosa* to interrogate virulence factor synthesis. *Nature Communications*, **8**.
- BATTLE, S. E., RELLO, J. & HAUSER, A. R. 2008. Genomic islands of *Pseudomonas aeruginosa*. *FEMS Microbiology Letters*, **290**, 70-78.
- BECEIRO, A., TOMÁS, M. & BOU, G. 2013. Antimicrobial resistance and virulence: a successful or deleterious association in the bacterial world? *Clinical Microbiology Reviews*, **26**, 185-230.
- BECKER, S. A. & PALSSON, B. O. 2008. Context-specific metabolic networks are consistent with experiments. *Plos Computational Biology*, **4**, e1000082.
- BECKS, V. E. & LORENZONI, N. M. 1995. *Pseudomonas aeruginosa* outbreak in a neonatal intensive care unit: a possible link to contaminated hand lotion. *American Journal of Infection Control*, **23**, 396-398.
- BEILFUß, W., LESCHKE, M. & WEBER, K. 2005. A new concept to boost the preservative efficacy of phenoxyethanol. *SÖFW-Journal*, **131**, 30-36.
- BERENDONK, T. U., MANAIA, C. M., MERLIN, C., FATTA-KASSINOS, D., CYTRYN, E., WALSH, F., BÜRGMANN, H., SØRUM, H., NORSTRÖM, M. & PONS, M.-N. 2015. Tackling antibiotic resistance: the environmental framework. *Nature Reviews Microbiology*, **13**, 310-317.
- BERG, J., THORSEN, M. K., HOLM, P. E., JENSEN, J., NYBROE, O. & BRANDT, K. K. 2010. Cu exposure under field conditions co-selects for antibiotic resistance as determined by a novel cultivation-independent bacterial community tolerance assay. *Environmental Science & Technology*, **44**, 8724-8728.
- BERLIN, K., KOREN, S., CHIN, C.-S., DRAKE, J. P., LANDOLIN, J. M. & PHILLIPPY, A. M. 2015. Assembling large genomes with single-molecule sequencing and locality-sensitive hashing. *Nature Biotechnology*, **33**, 623-630.
- BERNAL, P., MUÑOZ-ROJAS, J., HURTADO, A., RAMOS, J. L. & SEGURA, A. 2007. A *Pseudomonas putida* cardiolipin synthesis mutant exhibits increased sensitivity to drugs related to transport functionality. *Environmental Microbiology*, **9**, 1135-1145.
- BERNIER, S. P. & SURETTE, M. G. 2013. Concentration-dependent activity of antibiotics in natural environments. *Frontiers in Microbiology*, **4**.
- BERTELLI, C., LAIRD, M. R., WILLIAMS, K. P., LAU, B. Y., HOAD, G., WINSOR, G. L. & BRINKMAN, F. S. 2017. IslandViewer 4: expanded prediction of genomic islands for larger-scale datasets. *Nucleic Acids Research*, **45**, W30-W35.
- BLAIR, J. M., WEBBER, M. A., BAYLAY, A. J., OGBOLU, D. O. & PIDDOCK, L. J. 2015. Molecular mechanisms of antibiotic resistance. *Nature Reviews Microbiology*, **13**, 42.
- BLAZIER, A. S. & PAPIN, J. A. 2012. Integration of expression data in genome-scale metabolic network reconstructions. *Frontiers in Physiology*, **3**.
- BLUMENFELD, O., NATHANSOHN, N., YESHURUN, I. & ASHKENAZI, I. 2005. Eye cosmetics--the beauty and the beast. *Harefuah*, **144**, 357-362.
- BODEY, G. P. 2001. *Pseudomonas aeruginosa* infections in cancer patients: have they gone away? *Current Opinion in Infectious Diseases*, **14**, 403-407.
- BOTELHO, J., GROSSO, F., QUINTEIRA, S., MABROUK, A. & PEIXE, L. 2017. The complete nucleotide sequence of an IncP-2 megaplasmid unveils a mosaic architecture comprising a putative novel blaVIM-2-harboring transposon in *Pseudomonas aeruginosa*. *Journal of Antimicrobial Chemotherapy*.
- BOXALL, A. B., RUDD, M. A., BROOKS, B. W., CALDWELL, D. J., CHOI, K., HICKMANN, S., INNES, E., OSTAPYK, K., STAVELEY, J. P. & VERSLYCKE, T. 2012. Pharmaceuticals and personal care products in the environment: what are the big questions? *Environmental Health Perspectives*, **120**, 1221.
- BOYLE, B., FERNANDEZ, L., LAROCHE, J., KUKAVICA-IBRULJ, I., MENDES, C. M., HANCOCK, R. W. & LEVESQUE, R. C. 2012. Complete genome sequences of three *Pseudomonas aeruginosa* isolates with phenotypes of polymyxin B adaptation and inducible resistance. *Journal of Bacteriology*, **194**, 529-530.
- BRANNAN, D. K. 1995. Cosmetic preservation. *Journal of the Society of Cosmetic Chemists*, **46**, 199-220.
- BRANNAN, D. K. 1997. *Cosmetic microbiology: a practical handbook*, Florida: CRC Press.
- BRAUTIGAM, C. A., DEKA, R. K., SCHUCK, P., TOMCHICK, D. R. & NORGARD, M. V. 2012. Structural and thermodynamic characterization of the interaction between two periplasmic *Treponema pallidum* lipoproteins that are components of a TPR-protein-associated TRAP transporter (TPAT). *Journal of Molecular Biology*, **420**, 70-86.
- BREATHNACH, A., CUBBON, M., KARUNAHARAN, R., POPE, C. & PLANCHE, T. 2012. Multidrug-resistant *Pseudomonas aeruginosa* outbreaks in two hospitals: association with contaminated hospital waste-water systems. *Journal of Hospital Infection*, **82**, 19-24.

- BREIDENSTEIN, E. B., DE LA FUENTE-NÚÑEZ, C. & HANCOCK, R. E. 2011. *Pseudomonas aeruginosa*: all roads lead to resistance. *Trends in Microbiology*, **19**, 419-426.
- BRIDIER, A., DUBOIS-BRISSENET, F., GREUB, G., THOMAS, V. & BRIANDET, R. 2011. Dynamics of the action of biocides in *Pseudomonas aeruginosa* biofilms. *Antimicrobial Agents and Chemotherapy*, **55**, 2648-2654.
- BRODDRICK, J. 2017. Available predictive genome-scale metabolic network reconstructions [Online]. Available: <http://sbrg.ucsd.edu/InSilicoOrganisms/OtherOrganisms> [Accessed 15 August 2017].
- BROWN, L., WOLF, J. M., PRADOS-ROSALES, R. & CASADEVALL, A. 2015. Through the wall: extracellular vesicles in Gram-positive bacteria, mycobacteria and fungi. *Nature Reviews Microbiology*, **13**, 620.
- BROWN, M. L., ALDRICH, H. C. & GAUTHIER, J. J. 1995. Relationship between glycocalyx and povidone-iodine resistance in *Pseudomonas aeruginosa* (ATCC 27853) biofilms. *Applied and Environmental Microbiology*, **61**, 187-193.
- BRÖZEL, V. & CLOETE, T. 1994. Resistance of *Pseudomonas aeruginosa* to isothiazolone. *Journal of Applied Microbiology*, **76**, 576-582.
- BURROWES, E., ABBAS, A., O'NEILL, A., ADAMS, C. & O'GARA, F. 2005. Characterisation of the regulatory RNA RsmB from *Pseudomonas aeruginosa* PAO1. *Research in Microbiology*, **156**, 7-16.
- BUSCHMANN, A. H., TOMOVA, A., LÓPEZ, A., MALDONADO, M. A., HENRÍQUEZ, L. A., IVANOVA, L., MOY, F., GODFREY, H. P. & CABELLO, F. C. 2012. Salmon aquaculture and antimicrobial resistance in the marine environment. *Plos One*, **7**, e42724.
- CALCOTT, M. J., OWEN, J. G., LAMONT, I. L. & ACKERLEY, D. F. 2014. Biosynthesis of novel Pyoverdines by domain substitution in a nonribosomal peptide synthetase of *Pseudomonas aeruginosa*. *Applied and Environmental Microbiology*, **80**, 5723-5731.
- CAMACHO, C., COULOURIS, G., AVAGYAN, V., MA, N., PAPADOPOULOS, J., BEALER, K. & MADDEN, T. L. 2009. BLAST+: architecture and applications. *BMC Bioinformatics*, **10**, 421.
- CAMILLI, A. & BASSLER, B. L. 2006. Bacterial small-molecule signalling pathways. *Science*, **311**, 1113-1116.
- CAÑIZARES-VILLANUEVA, R., RIOS-LEAL, E., OLVERA, R. R., PONCE, N. T. & MÁRQUEZ, R. F. 1997. Microbial sources of pigments. *Revista Latinoamericana de Microbiología*, **40**, 87-107.
- CAREY, D. E. & MCNAMARA, P. J. 2014. The impact of triclosan on the spread of antibiotic resistance in the environment. *Frontiers in Microbiology*, **5**.
- CAREY, M. A., PAPIN, J. A. & GULER, J. 2017. Novel Plasmodium falciparum metabolic network reconstruction identifies shifts associated with clinical antimalarial resistance. *bioRxiv*, **119941**.
- CARMELI, Y., TROILLET, N., ELIOPOULOS, G. M. & SAMORE, M. H. 1999. Emergence of antibiotic-resistant *Pseudomonas aeruginosa*: comparison of risks associated with different antipseudomonal agents. *Antimicrobial Agents and Chemotherapy*, **43**, 1379-1382.
- CARVER, T. J., RUTHERFORD, K. M., BERRIMAN, M., RAJANDREAM, M.-A., BARRELL, B. G. & PARKHILL, J. 2005. ACT: the Artemis comparison tool. *Bioinformatics*, **21**, 3422-3423.
- CASPI, R., FOERSTER, H., FULCHER, C. A., KAIPA, P., KRUMMENACKER, M., LATENDRESSE, M., PALEY, S., RHEE, S. Y., SHEARER, A. G. & TISSIER, C. 2007. The MetaCyc Database of metabolic pathways and enzymes and the BioCyc collection of Pathway/Genome Databases. *Nucleic Acids Research*, **36**, D623-D631.
- CASTAÑEDA-GARCÍA, A., BLÁZQUEZ, J. & RODRÍGUEZ-ROJAS, A. 2013. Molecular mechanisms and clinical impact of acquired and intrinsic fosfomicin resistance. *Antibiotics*, **2**, 217-236.
- CASTAÑEDA-GARCÍA, A., RODRÍGUEZ-ROJAS, A., GUELFO, J. R. & BLÁZQUEZ, J. 2009. The glycerol-3-phosphate permease GlpT is the only fosfomicin transporter in *Pseudomonas aeruginosa*. *Journal of Bacteriology*, **191**, 6968-6974.
- CATTOIR, V., NARASIMHAN, G., SKURNIK, D., ASCHARD, H., ROUX, D., RAMPHAL, R., JYOT, J. & LORY, S. 2012. Transcriptional response of mucoid *Pseudomonas aeruginosa* to human respiratory mucus. *MBio*, **3**, e00410-12.
- CERNIGLIA, C. E. 1993. Biodegradation of polycyclic aromatic hydrocarbons. *Current Opinion in Biotechnology*, **4**, 331-338.
- CHAKRABARTY, A. 1998. Nucleoside diphosphate kinase: role in bacterial growth, virulence, cell signalling and polysaccharide synthesis. *Molecular Microbiology*, **28**, 875-882.
- CHAN, K.-G., YIN, W.-F. & LIM, Y. L. 2014. Complete genome sequence of *Pseudomonas aeruginosa* strain YL84, a quorum-sensing strain isolated from compost. *Genome Announcements*, **2**.
- CHAN, Y., TAN, T., ONG, Y. & CHUA, K. 2004. BpeAB-OprB, a multidrug efflux pump in *Burkholderia pseudomallei*. *Antimicrobial Agents and Chemotherapy*, **48**, 1128-1135.
- CHAPMAN, J. S. 1998. Characterizing bacterial resistance to preservatives and disinfectants. *International Biodeterioration & Biodegradation*, **41**, 241-245.
- CHAPMAN, J. S. 2003. Biocide resistance mechanisms. *International Biodeterioration & Biodegradation*, **51**, 133-138.
- CHAVALI, A. K., WHITTEMORE, J. D., EDDY, J. A., WILLIAMS, K. T. & PAPIN, J. A. 2008. Systems analysis of metabolism in the pathogenic trypanosomatid *Leishmania major*. *Molecular Systems Biology*, **4**, 177.

- CHAVARRÍA, M., NIKEL, P. I., PÉREZ-PANTOJA, D. & LORENZO, V. 2013. The Entner–Doudoroff pathway empowers *Pseudomonas putida* KT2440 with a high tolerance to oxidative stress. *Environmental Microbiology*, **15**, 1772-1785.
- CHEN, D., XU, Z. & LIU, J. 2016a. Complete Sequence of pBM413, a 423-Kilobase Plasmid with qnrVC gene [Online]. Available: https://www.ncbi.nlm.nih.gov/nucore/NZ_CP016215.1 [Accessed 03 March 2017].
- CHEN, L., ZHENG, D., LIU, B., YANG, J. & JIN, Q. 2016. VFDB 2016: hierarchical and refined dataset for big data analysis—10 years on. *Nucleic Acids Research*, **44**, D694-D697.
- CHEN, N. H., DJOKO, K. Y., VEYRIER, F. J. & MCEWAN, A. G. 2016. Formaldehyde stress responses in bacterial pathogens. *Frontiers in Microbiology*, **7**.
- CHEN, Y.-C., XIE, X.-B., SHI, Q.-S., OUYANG, Y.-S. & CHEN, Y.-B. 2010. Species identification of industry spoilage microorganism and the resistance analysis. *Microbiology/Weishengwuxue Tongbao*, **37**, 1558-1565.
- CHENG, J., THANASSI, J. A., THOMA, C. L., BRADBURY, B. J., DESHPANDE, M. & PUCCI, M. J. 2007. Dual targeting of DNA gyrase and topoisomerase IV: target interactions of heteroaryl isothiazolones in *Staphylococcus aureus*. *Antimicrobial Agents and Chemotherapy*, **51**, 2445-2453.
- CHEUNG, V. W. N., XUE, B., HERNANDEZ-VALLADARES, M., GO, M. K., TUNG, A., AGUDA, A. H., ROBINSON, R. C. & YEW, W. S. 2014. Identification of polyketide inhibitors targeting 3-dehydroquinate dehydratase in the shikimate pathway of *Enterococcus faecalis*. *PloS One*, **9**, e103598.
- CHIN, C.-S., ALEXANDER, D. H., MARKS, P., KLAMMER, A. A., DRAKE, J., HEINER, C., CLUM, A., COPELAND, A., HUDDLESTON, J. & EICHLER, E. E. 2013. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nature Methods*, **10**, 563-569.
- CHOMICZEWSKA-SKÓRA, D., KRĘCISZ, B. & KIEĆ-ŚWIERCZYŃSKA, M. 2014. Isothiazolinones as causal factors of contact allergy epidemics in the 20th and 21st centuries. *Medycyna Pracy*, **65**, 543-554.
- CHOU, H. T., KWON, D.-H., HEGAZY, M. & LU, C.-D. 2008. Transcriptome analysis of agmatine and putrescine catabolism in *Pseudomonas aeruginosa* PAO1. *Journal of Bacteriology*, **190**, 1966-1975.
- CHUA, S. L., LIU, Y., YAM, J. K. H., CHEN, Y., VEJBORG, R. M., TAN, B. G. C., KJELLEBERG, S., TOLKER-NIELSEN, T., GIVSKOV, M. & YANG, L. 2014. Dispersed cells represent a distinct stage in the transition from bacterial biofilm to planktonic lifestyles. *Nature Communications*, **5**, 4462.
- CHUANCHUEN, R., BEINLICH, K., HOANG, T. T., BECHER, A., KARKHOFF-SCHWEIZER, R. R. & SCHWEIZER, H. P. 2001. Cross-resistance between triclosan and antibiotics in *Pseudomonas aeruginosa* is mediated by multidrug efflux pumps: exposure of a susceptible mutant strain to triclosan selects *nfxB* mutants overexpressing MexCD-OprJ. *Antimicrobial Agents and Chemotherapy*, **45**, 428-432.
- CHUANCHUEN, R., MURATA, T., GOTOH, N. & SCHWEIZER, H. P. 2005. Substrate-dependent utilization of OprM or OpmH by the *Pseudomonas aeruginosa* MexJK efflux pump. *Antimicrobial Agents and Chemotherapy*, **49**, 2133-2136.
- CHUANCHUEN, R., NARASAKI, C. T. & SCHWEIZER, H. P. 2002. The MexJK efflux pump of *Pseudomonas aeruginosa* requires OprM for antibiotic efflux but not for efflux of triclosan. *Journal of Bacteriology*, **184**, 5036-5044.
- CHUGANI, S., KIM, B. S., PHATTARASUKOL, S., BRITTNACHER, M. J., CHOI, S. H., HARWOOD, C. S. & GREENBERG, E. P. 2012. Strain-dependent diversity in the *Pseudomonas aeruginosa* quorum-sensing regulon. *Proceedings of the National Academy of Sciences*, **109**, E2823-E2831.
- CLOETE, T. E. 2003. Resistance mechanisms of bacteria to antimicrobial compounds. *International Biodeterioration & Biodegradation*, **51**, 277-282.
- COCHRAN, W., MCFETERS, G. & STEWART, P. 2000. Reduced susceptibility of thin *Pseudomonas aeruginosa* biofilms to hydrogen peroxide and monochloramine. *Journal of Applied Microbiology*, **88**, 22-30.
- COGGAN, K. A. & WOLFGANG, M. C. 2012. Global regulatory pathways and cross-talk control *Pseudomonas aeruginosa* environmental lifestyle and virulence phenotype. *Current Issues in Molecular Biology*, **14**, 47.
- COLIJN, C., BRANDES, A., ZUCKER, J., LUN, D. S., WEINER, B., FARHAT, M. R., CHENG, T.-Y., MOODY, D. B., MURRAY, M. & GALAGAN, J. E. 2009. Interpreting expression data with metabolic flux models: predicting *Mycobacterium tuberculosis* mycolic acid production. *PloS Computational Biology*, **5**, e1000489.
- COLLIER, P. J., AUSTIN, P. & GILBERT, P. 1991. Isothiazolone biocides: enzyme-inhibiting pro-drugs. *International Journal of Pharmaceutics*, **74**, 195-201.
- COLLIER, P., RAMSEY, A., AUSTIN, P. & GILBERT, P. 1990a. Growth inhibitory and biocidal activity of some isothiazolone biocides. *Journal of Applied Bacteriology*, **69**, 569-577.
- COLLIER, P., RAMSEY, A., WAIGH, R., DOUGLAS, K., AUSTIN, P. & GILBERT, P. 1990b. Chemical reactivity of some isothiazolone biocides. *Journal of Applied Microbiology*, **69**, 578-584.
- CONESA, A., MADRIGAL, P., TARAZONA, S., GOMEZ-CABRERO, D., CERVERA, A., MCPHERSON, A., SZCZEŚNIAK, M. W., GAFFNEY, D. J., ELO, L. L. & ZHANG, X. 2016. A survey of best practices for RNA-seq data analysis. *Genome Biology*, **17**, 13.

- CONNOR, T. R., LOMAN, N. J., THOMPSON, S., SMITH, A., SOUTHGATE, J., POPLAWSKI, R., BULL, M. J., RICHARDSON, E., ISMAIL, M. & ELWOOD-THOMPSON, S. 2016. CLIMB (the Cloud Infrastructure for Microbial Bioinformatics): an online resource for the medical microbiology community. *Microbial Genomics*, **2**.
- CONSORTIUM, U. 2017. UniProt: the universal protein knowledgebase. *Nucleic Acids Research*, **45**, D158-D169.
- COSTAGLIOLI, P., BARTHE, C., CLAVEROL, S., BRÖZEL, V. S., PERROT, M., CROUZET, M., BONNEU, M., GARBAY, B. & VILAIN, S. 2012. Evidence for the involvement of the anthranilate degradation pathway in *Pseudomonas aeruginosa* biofilm formation. *Microbiologyopen*, **1**, 326-339.
- COSTERTON, J. & LASHEN, E. 1984. Influence of biofilm on efficacy of biocides on corrosion-causing bacteria. *Materials Performance*, **23**, 13-17.
- CRICK, D. C., MAHAPATRA, S. & BRENNAN, P. J. 2001. Biosynthesis of the arabinogalactan-peptidoglycan complex of *Mycobacterium tuberculosis*. *Glycobiology*, **11**, 107R-118R.
- CROUCHER, N. J. & THOMSON, N. R. 2010. Studying bacterial transcriptomes using RNA-seq. *Current Opinion in Microbiology*, **13**, 619-624.
- CZOLKOSS, S., FRITZ, C., HÖLZL, G. & AKTAS, M. 2016. Two distinct cardiolipin synthases operate in *Agrobacterium tumefaciens*. *PLoS One*, **11**, e0160373.
- DADASHI, L. & DEHGHANZADEH, R. 2016. Investigating incidence of bacterial and fungal contamination in shared cosmetic kits available in the women beauty salons. *Health Promotion Perspectives*, **6**, 159.
- DAILEY, H. A., GERDES, S., DAILEY, T. A., BURCH, J. S. & PHILLIPS, J. D. 2015. Noncanonical coproporphyrin-dependent bacterial heme biosynthesis pathway that does not use protoporphyrin. *Proceedings of the National Academy of Sciences*, **112**, 2210-2215.
- DANTAS, G., SOMMER, M. O., OLUWASEGUN, R. D. & CHURCH, G. M. 2008. Bacteria subsisting on antibiotics. *Science*, **320**, 100-103.
- D'ARGENIO, D. A., CALFEE, M. W., RAINEY, P. B. & PESCI, E. C. 2002. Autolysis and autoaggregation in *Pseudomonas aeruginosa* colony morphology mutants. *Journal of Bacteriology*, **184**, 6481-6489.
- DAS, N. & CHANDRAN, P. 2010. Microbial degradation of petroleum hydrocarbon contaminants: an overview. *Biotechnology Research International*, **2011**.
- DAUGHTON, C. G. & TERNES, T. A. 1999. Pharmaceuticals and personal care products in the environment: agents of subtle change? *Environmental Health Perspectives*, **107**, 907.
- DE LORENZO, V., HERRERO, M., JAKUBZIK, U. & TIMMIS, K. N. 1990. Mini-Tn5 transposon derivatives for insertion mutagenesis, promoter probing, and chromosomal insertion of cloned DNA in gram-negative *eubacteria*. *Journal of Bacteriology*, **172**, 6568-6572.
- DEHIO, C. & MEYER, M. 1997. Maintenance of broad-host-range incompatibility group P and group Q plasmids and transposition of Tn5 in *Bartonella henselae* following conjugal plasmid transfer from *Escherichia coli*. *Journal of Bacteriology*, **179**, 538-540.
- DEKKER, F. J., GHIZZONI, M., VAN DER MEER, N., WISASTRA, R. & HAISMA, H. J. 2009. Inhibition of the PCAF histone acetyl transferase and cell proliferation by isothiazolones. *Bioorganic & medicinal chemistry*, **17**, 460-466.
- DELCOUR, A. H. 2009. Outer membrane permeability and antibiotic resistance. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, **1794**, 808-816.
- DEMBEK, M., BARQUIST, L., BOINETT, C. J., CAIN, A. K., MAYHO, M., LAWLEY, T. D., FAIRWEATHER, N. F. & FAGAN, R. P. 2015. High-throughput analysis of gene essentiality and sporulation in *Clostridium difficile*. *MBio*, **6**, e02383-14.
- DENYER, S. P. 1995. Mechanisms of action of antibacterial biocides. *International Biodeterioration & Biodegradation*, **36**, 227-245.
- DÉRASPE, M., ALEXANDER, D. C., XIONG, J., MA, J. H., LOW, D. E., JAMIESON, F. B. & ROY, P. H. 2014. Genomic analysis of *Pseudomonas aeruginosa* PA96, the host of carbapenem resistance plasmid pOZ176. *FEMS Microbiology Letters*, **356**, 212-216.
- DEVER, L. A. & DERMODY, T. S. 1991. Mechanisms of bacterial resistance to antibiotics. *Archives of Internal Medicine*, **151**, 886-895.
- DEVOID, S., OVERBEEK, R., DEJONGH, M., VONSTEIN, V., BEST, A. A. & HENRY, C. 2013. Automated genome annotation and metabolic model reconstruction in the SEED and Model SEED. *Systems Metabolic Engineering: Methods and Protocols*, 17-45.
- DÉZIEL, E., LÉPINE, F., MILOT, S., HE, J., MINDRINOS, M. N., TOMPKINS, R. G. & RAHME, L. G. 2004. Analysis of *Pseudomonas aeruginosa* 4-hydroxy-2-alkylquinolines (HAQs) reveals a role for 4-hydroxy-2-heptylquinoline in cell-to-cell communication. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 1339-1344.
- DI CESARE, A., LUNA, G. M., VIGNAROLI, C., PASQUAROLI, S., TOTA, S., PARONCINI, P. & BIAVASCO, F. 2013. Aquaculture can promote the presence and spread of antibiotic-resistant *Enterococci* in marine sediments. *PLoS One*, **8**, e62838.
- DIAZ-PEREZ, A., ZAVALA-HERNANDEZ, A., CERVANTES, C. & CAMPOS-GARCIA, J. 2004. The *gnyRDBHAL* cluster is involved in acyclic isoprenoid degradation in *Pseudomonas aeruginosa*. *Applied and Environmental Microbiology*, **70**, 5102-5110.
- DIGGLE, S. P., MATTHIJS, S., WRIGHT, V. J., FLETCHER, M. P., CHHABRA, S. R., LAMONT, I. L., KONG, X., HIDER, R. C., CORNELIS, P. & CÁMARA, M. 2007. The *Pseudomonas aeruginosa* 4-quinolone signal molecules HHQ and PQS play multifunctional roles in quorum sensing and iron entrapment. *Chemistry & Biology*, **14**, 87-96.
- DONLAN, R. M. & COSTERTON, J. W. 2002. Biofilms: survival mechanisms of clinically relevant microorganisms. *Clinical Microbiology Reviews*, **15**, 167-193.

- DONLAN, R. M. 2002. Biofilms: microbial life on surfaces. *Emerging Infectious Diseases*, **8**, 881.
- DURVE, A., NAPHADE, S., BHOT, M., VARGHESE, J. & CHANDRA, N. 2013. Plasmid curing and protein profiling of heavy metal tolerating bacterial isolates. *Archives of Applied Science Research*, **5**, 46-54.
- DZIVA, F., HAUSER, H., CONNOR, T. R., VAN DIEMEN, P. M., PRESCOTT, G., LANGRIDGE, G. C., ECKERT, S., CHAUDHURI, R. R., EWERS, C. & MELLATA, M. 2013. Sequencing and functional annotation of avian pathogenic *Escherichia coli* serogroup O78 strains reveal the evolution of *E. coli* lineages pathogenic for poultry via distinct mechanisms. *Infection and Immunity*, **81**, 838-849.
- EBRAHIM, A., BRUNK, E., TAN, J., O'BRIEN, E. J., KIM, D., SZUBIN, R., LERMAN, J. A., LECHNER, A., SASTRY, A. & BORDBAR, A. 2016. Multi-omic data integration enables discovery of hidden biological regularities. *Nature Communications*, **7**.
- ECKMANNS, T., OPPERT, M., MARTIN, M., AMOROSA, R., ZUSCHNEID, I., FREI, U., RÜDEN, H. & WEIST, K. 2008. An outbreak of hospital-acquired *Pseudomonas aeruginosa* infection caused by contaminated bottled water in intensive care units. *Clinical Microbiology and Infection*, **14**, 454-458.
- ECKWEILER, D., BUNK, B., SPRÖER, C., OVERMANN, J. & HÄUSSLER, S. 2014. Complete genome sequence of highly adherent *Pseudomonas aeruginosa* small-colony variant SCV20265. *Genome Announcements*, **2**, e01232-13.
- EDWARD, W. Y., AIRES, J. R. & NIKAIDO, H. 2003. AcrB multidrug efflux pump of *Escherichia coli*: composite substrate-binding cavity of exceptional flexibility generates its extremely wide substrate specificity. *Journal of Bacteriology*, **185**, 5657-5664.
- EL KHOURY, M., SWAIN, J., SAUTREY, G., ZIMMERMANN, L., VAN DER SMISSEN, P., DÉCOUT, J. & MINGEOT-LECLERCQ, M. 2017. Targeting Bacterial Cardiolipin Enriched Microdomains: An Antimicrobial Strategy Used by Amphiphilic Aminoglycoside Antibiotics. *Scientific Reports*, **7**, 10697.
- EL-HALFAWY, O. M. & VALVANO, M. A. 2013. Chemical communication of antibiotic resistance by a highly resistant subpopulation of bacterial cells. *PLoS One*, **8**, e68874.
- EL-HALFAWY, O. M. & VALVANO, M. A. 2014. Putrescine reduces antibiotic-induced oxidative stress as a mechanism of modulation of antibiotic resistance in *Burkholderia cenocepacia*. *Antimicrobial Agents and Chemotherapy*, **58**, 4162-4171.
- EL-HOUSSEINY, R., ABOULWAFI, M. M., ELKHATIB, W. & HASSOUNA, N. 2013. Recovery and Detection of Microbial Contaminants in Some Non-Sterile Pharmaceutical Products. *Archives of Clinical Microbiology*, **4**.
- ELROD, R. & BRAUN, A. C. 1942. *Pseudomonas aeruginosa*: its role as a plant pathogen. *Journal of Bacteriology*, **44**, 633.
- ENGLISH, A. C., RICHARDS, S., HAN, Y., WANG, M., VEE, V., QU, J., QIN, X., MUZNY, D. M., REID, J. G. & WORLEY, K. C. 2012. Mind the gap: upgrading genomes with Pacific Biosciences RS long-read sequencing technology. *PLoS One*, **7**, e47768.
- ESSAR, D., EBERLY, L., HADERO, A. & CRAWFORD, I. 1990. Identification and characterization of genes for a second anthranilate synthase in *Pseudomonas aeruginosa*: interchangeability of the two anthranilate synthases and evolutionary implications. *Journal of Bacteriology*, **172**, 884-900.
- EZRATY, B., GENNARIS, A., BARRAS, F. & COLLET, J.-F. 2017. Oxidative stress, protein damage and repair in bacteria. *Nature Reviews Microbiology*.
- FABREGAT, A., SIDIROPOULOS, K., GARAPATI, P., GILLESPIE, M., HAUSMANN, K., HAW, R., JASSAL, B., JUPE, S., KORNINGER, F. & MCKAY, S. 2015. The reactome pathway knowledgebase. *Nucleic Acids Research*, **44**, D481-D487.
- FARRINGTON, J., MARTZ, E., WELLS, S., ENNIS, C., HOLDER, J., LEVCHUK, J., AVIS, K., HOFFMAN, P., HITCHINS, A. & MADDEN, J. 1994. Ability of laboratory methods to predict in-use efficacy of antimicrobial preservatives in an experimental cosmetic. *Applied and Environmental Microbiology*, **60**, 4553-4558.
- FAVERO, M., CARSON, L., BOND, W. & PETERSEN, N. 1971. *Pseudomonas aeruginosa*: growth in distilled water from hospitals. *Science*, **173**, 836-838.
- FDA. 2017a. Drug Safety Communication: FDA warns about rare but serious allergic reactions with the skin antiseptic chlorhexidine gluconate [Online]. Available: <https://www.fda.gov/Drugs/DrugSafety/ucm530975.htm> [Accessed 10 July 2017].
- FDA. 2017b. Microbiological Safety and Cosmetics [Online]. Available: <https://www.fda.gov/cosmetics/productsingredients/potentialcontaminants/ucm433748.htm> [Accessed 10 July 2017].
- FENG, Y., JONKER, M. J., MOUSTAKAS, I., BRUL, S. & TER KUILE, B. H. 2016. Dynamics of mutations during development of resistance by *Pseudomonas aeruginosa* against five antibiotics. *Antimicrobial Agents and Chemotherapy*, **60**, 4229-4236.
- FIEGE, H., VOGES, H. W., HAMAMOTO, T., UMEMURA, S., IWATA, T., MIKI, H., FUJITA, Y., BUYSCH, H. J., GARBE, D. & PAULUS, W. 2000. Phenol derivatives. *Ullmann's Encyclopedia of Industrial Chemistry*. Wiley: VCH.
- FITZGERALD, K., DAVIES, A. & RUSSELL, A. 1992. Mechanism of action of chlorhexidine diacetate and phenoxyethanol singly and in combination against gram-negative bacteria. *Microbios*, **70**, 215-230.
- FLEMMING, H.-C. & WINGENDER, J. 2010. The biofilm matrix. *Nature Reviews Microbiology*, **8**, 623.
- FLEMMING, H.-C., NEU, T. R. & WOZNIAK, D. J. 2007. The EPS matrix: the "house of biofilm cells". *Journal of Bacteriology*, **189**, 7945-7947.
- FLORES, M., MORILLO, M. & CRESPO, M. 1997. Deterioration of raw materials and cosmetic products by preservative resistant microorganisms. *International Biodeterioration & Biodegradation*, **40**, 157-160.

- FLYNN, J. M., PHAN, C. & HUNTER, R. C. 2017. Genome-wide survey of *Pseudomonas aeruginosa* PA14 reveals a role for the glyoxylate pathway and extracellular proteases in the utilization of mucin. *Infection and Immunity*, IAI. 00182-17.
- FÖRSTER-FROMME, K., HÖSCHLE, B., MACK, C., BOTT, M., ARMBRUSTER, W. & JENDROSSEK, D. 2006. Identification of genes and proteins necessary for catabolism of acyclic terpenes and leucine/isovalerate in *Pseudomonas aeruginosa*. *Applied and Environmental Microbiology*, **72**, 4819-4828.
- FORWARD, J. A., BEHRENDT, M. C., WYBORN, N. R., CROSS, R. & KELLY, D. J. 1997. TRAP transporters: a new family of periplasmic solute transport systems encoded by the dctPQM genes of *Rhodobacter capsulatus* and by homologs in diverse gram-negative bacteria. *Journal of Bacteriology*, **179**, 5482-5493.
- FRANKLIN, M. J. & OHMAN, D. E. 1993. Identification of algF in the alginate biosynthetic gene cluster of *Pseudomonas aeruginosa* which is required for alginate acetylation. *Journal of Bacteriology*, **175**, 5057-5065.
- FRAUD, S., CAMPIGOTTO, A. J., CHEN, Z. & POOLE, K. 2008. MexCD-OprJ multidrug efflux system of *Pseudomonas aeruginosa*: involvement in chlorhexidine resistance and induction by membrane-damaging agents dependent upon the AlgU stress response sigma factor. *Antimicrobial Agents and Chemotherapy*, **52**, 4478-4482.
- FREDERICKSON, M., PARKER, E. J., HAWKINS, A. R., COGGINS, J. R. & ABELL, C. 1999. Selective inhibition of type II dehydroquinases. *The Journal of Organic Chemistry*, **64**, 2612-2613.
- FREILICH, S., KREIMER, A., BORENSTEIN, E., YOSEF, N., SHARAN, R., GOPHNA, U. & RUPPIN, E. 2009. Metabolic-network-driven analysis of bacterial ecological strategies. *Genome Biology*, **10**, R61.
- FRESCHI, L., JEUKENS, J., KUKAVICA-IBRULJ, I., BOYLE, B., DUPONT, M.-J., LAROCHE, J., LAROSE, S., MAAROUFI, H., FOTHERGILL, J. L. & MOORE, M. 2015. Clinical utilization of genomics data produced by the international *Pseudomonas aeruginosa* consortium. *Frontiers in Microbiology*, **6**.
- FRIDKIN, S. K. & GAYNES, R. P. 1999. Antimicrobial resistance in intensive care units. *Clinics in Chest Medicine*, **20**, 303-316.
- FUKUOKA, T., OHYA, S., NARITA, T., KATSUTA, M., IJIMA, M., MASUDA, N., YASUDA, H., TRIAS, J. & NIKAIIDO, H. 1993. Activity of the carbapenem panipenem and role of the OprD (D2) protein in its diffusion through the *Pseudomonas aeruginosa* outer membrane. *Antimicrobial Agents and Chemotherapy*, **37**, 322-327.
- FURDAS, S. D., SHEKFEH, S., BISSINGER, E.-M., WAGNER, J. M., SCHLIMME, S., VALKOV, V., HENDZEL, M., JUNG, M. & SIPPL, W. 2011. Synthesis and biological testing of novel pyridoisothiazolones as histone acetyltransferase inhibitors. *Bioorganic & Medicinal Chemistry*, **19**, 3678-3689.
- GALÁN-VÁSQUEZ, E., LUNA, B. & MARTÍNEZ-ANTONIO, A. 2011. The regulatory network of *Pseudomonas aeruginosa*. *Microbial Informatics and Experimentation*, **1**, 3.
- GALLAGHER, L. A. & MANOIL, C. 2001. *Pseudomonas aeruginosa* PAO1 kills *Caenorhabditis elegans* by cyanide poisoning. *Journal of Bacteriology*, **183**, 6207-6214.
- GALPERIN, M. Y., MAKAROVA, K. S., WOLF, Y. I. & KOONIN, E. V. 2014. Expanded microbial genome coverage and improved protein family annotation in the COG database. *Nucleic Acids Research*, **43**, D261-D269.
- GAO, C., HU, C., MA, C., SU, F., YU, H., JIANG, T., DOU, P., WANG, Y., QIN, T. & LV, M. 2012. Genome sequence of the lactate-utilizing *Pseudomonas aeruginosa* strain XMG. *Journal of Bacteriology*, **194**, 4751-4752.
- GEDI, V., MOON, J.-Y., LIM, W.-M., LEE, M.-Y., LEE, S.-C., KOO, B.-S., GOVINDWAR, S. & YOON, M.-Y. 2011. Identification and characterization of inhibitors of *Haemophilus influenzae* acetohydroxyacid synthase. *Enzyme and Microbial Technology*, **49**, 1-5.
- GERDES, K., RASMUSSEN, P. B. & MOLIN, S. 1986. Unique type of plasmid maintenance function: postsegregational killing of plasmid-free cells. *Proceedings of the National Academy of Sciences*, **83**, 3116-3120.
- GIBSON, B. W., MUNSON, R. S. & POST, D. M. 2006. Sialic acid ABC transporters in prokaryotes therapeutic targets. *Google Patents*. U.S. Patent Application 11/916,975.
- GILBERT, P. & MCBAIN, A. J. 2003. Potential impact of increased use of biocides in consumer products on prevalence of antibiotic resistance. *Clinical Microbiology Reviews*, **16**, 189-208.
- GILBERT, P., BEVERIDGE, E. & CRONE, P. 1976a. Inhibition of some respiration and dehydrogenase enzyme systems in *Escherichia coli* NCTC 5933 by phenoxyethanol. *Microbios*, **20**, 29-37.
- GILBERT, P., BEVERIDGE, E. & CRONE, P. 1977. Effect of phenoxyethanol on the permeability of *Escherichia coli* NCTC 5933 to inorganic ions. *Microbios*, **19**, 17-26.
- GILBERT, P., BEVERIDGE, E. & CRONE, P. 1979. Effect of 2-phenoxyethanol upon RNA, DNA and protein biosynthesis in *Escherichia coli* NCTC 5933. *Microbios*, **28**, 7-17.
- GILBERT, P., BEVERIDGE, E. & CRONE, P. 1976b. The action of phenoxyethanol upon respiration and dehydrogenase enzyme systems in *Escherichia coli* [proceedings]. *The Journal of Pharmacy and Pharmacology*, **28**, 51P-51P.
- GOLDOVÁ, J., ULRYCH, A., HERCÍK, K. & BRANNY, P. 2011. A eukaryotic-type signalling system of *Pseudomonas aeruginosa* contributes to oxidative stress resistance, intracellular survival and virulence. *BMC Genomics*, **12**, 437.

- GÓMEZ, M. I. & PRINCE, A. 2007. Opportunistic infections in lung disease: *Pseudomonas* infections in cystic fibrosis. *Current Opinion in Pharmacology*, **7**, 244-251.
- GOODERHAM, W. J. & HANCOCK, R. E. 2009. Regulation of virulence and antibiotic resistance by two-component regulatory systems in *Pseudomonas aeruginosa*. *FEMS Microbiology Reviews*, **33**, 279-294.
- GOODMAN, A. L., KULASEKARA, B., RIETSCH, A., BOYD, D., SMITH, R. S. & LORY, S. 2004. A signaling network reciprocally regulates genes associated with acute infection and chronic persistence in *Pseudomonas aeruginosa*. *Developmental Cell*, **7**, 745-754.
- GOVAN, J. R. & DERETIC, V. 1996. Microbial pathogenesis in cystic fibrosis: mucoid *Pseudomonas aeruginosa* and *Burkholderia cepacia*. *Microbiological Reviews*, **60**, 539-574.
- GREEN, S. K., SCHROTH, M. N., CHO, J. J., KOMINOS, S. D. & VITANZA-JACK, V. B. 1974. Agricultural plants and soil as a reservoir for *Pseudomonas aeruginosa*. *Applied Microbiology*, **28**, 987-991.
- GREINER, L., WATANABE, H., PHILLIPS, N., SHAO, J., MORGAN, A., ZALESKI, A., GIBSON, B. & APICELLA, M. 2004. Nontypeable *Haemophilus influenzae* strain 2019 produces a biofilm containing N-acetylneuraminic acid that may mimic sialylated O-linked glycans. *Infection and Immunity*, **72**, 4249-4260.
- GUO-LIANG, Z., YUE-TING, W., XIN-PING, Q. & QIN, M. 2005. Biodegradation of crude oil by *Pseudomonas aeruginosa* in the presence of rhamnolipids. *Journal of Zhejiang University-Science B*, **6**, 725-730.
- GUPTA, S. K., PADMANABHAN, B. R., DIENE, S. M., LOPEZ-ROJAS, R., KEMPF, M., LANDRAUD, L. & ROLAIN, J.-M. 2014. ARG-ANNOT, a new bioinformatic tool to discover antibiotic resistance genes in bacterial genomes. *Antimicrobial Agents and Chemotherapy*, **58**, 212-220.
- GUREVICH, A., SAVELIEV, V., VYAHHI, N. & TESLER, G. 2013. QUAST: quality assessment tool for genome assemblies. *Bioinformatics*, **29**, 1072-1075.
- GUSAROV, I., SHATALIN, K., STARODUBTSEVA, M. & NUDLER, E. 2009. Endogenous nitric oxide protects bacteria against a wide spectrum of antibiotics. *Science*, **325**, 1380-1384.
- GUTIERREZ, B., DOUTHWAITE, S. & GONZALEZ-ZORN, B. 2013. Indigenous and acquired modifications in the aminoglycoside binding sites of *Pseudomonas aeruginosa* rRNAs. *RNA biology*, **10**, 1324-1332.
- HAENNI, M., HOCQUET, D., PONSIN, C., CHOLLEY, P., GUYEUX, C., MADEC, J.-Y. & BERTRAND, X. 2015. Population structure and antimicrobial susceptibility of *Pseudomonas aeruginosa* from animal infections in France. *BMC Veterinary Research*, **11**, 9.
- HAGINS, J. M., SCOFFIELD, J., SUH, S.-J. & SILO-SUH, L. 2011. Malate synthase expression is deregulated in the *Pseudomonas aeruginosa* cystic fibrosis isolate FRD1. *Canadian Journal of Microbiology*, **57**, 186-195.
- HALLIN, P. F. & USSERY, D. W. 2004. CBS Genome Atlas Database: a dynamic storage for bioinformatic results and sequence data. *Bioinformatics*, **20**, 3682-3686.
- HALL-STOODLEY, L., COSTERTON, J. W. & STOODLEY, P. 2004. Bacterial biofilms: from the natural environment to infectious diseases. *Nature reviews. Microbiology*, **2**, 95.
- HANCOCK, R. 1998. Resistance Mechanisms in *Pseudomonas aeruginosa* and other Non-fermentative Gram-Negative Bacteria. *Clinical Infectious Diseases*, **27**, 93-99.
- HANCOCK, R. E. & SPEERT, D. P. 2000. Antibiotic resistance in *Pseudomonas aeruginosa*: mechanisms and impact on treatment. *Drug Resistance Updates*, **3**, 247-255.
- HARRISON, J. J., WADE, W. D., AKIERMAN, S., VACCHI-SUZZI, C., STREMIK, C. A., TURNER, R. J. & CERI, H. 2009. The chromosomal toxin gene *yafQ* is a determinant of multidrug tolerance for *Escherichia coli* growing in a biofilm. *Antimicrobial Agents and Chemotherapy*, **53**, 2253-2258.
- HARVEY, P. W. & EVERETT, D. J. 2004. Significance of the detection of esters of p-hydroxybenzoic acid (parabens) in human breast tumours. *Journal of Applied Toxicology*, **24**, 1-4.
- HÄUSSLER, S. & BECKER, T. 2008. The *Pseudomonas* Quinolone Signal (PQS) balances life and death in *Pseudomonas aeruginosa* populations. *PLoS Pathogens*, **4**, e1000166.
- HAZAN, R., QUE, Y. A., MAURA, D., STROBEL, B., MAJCHERCZYK, P. A., HOPPER, L. R., WILBUR, D. J., HREHA, T. N., BARQUERA, B. & RAHME, L. G. 2016. Auto poisoning of the respiratory chain by a quorum-sensing-regulated molecule favors biofilm formation and antibiotic tolerance. *Current Biology*, **26**, 195-206.
- HAZLETT, L. D. 2004. Corneal response to *Pseudomonas aeruginosa* infection. *Progress in Retinal and Eye Research*, **23**, 1-30.
- HE, J., BALDINI, R. L., DÉZIEL, E., SAUCIER, M., ZHANG, Q., LIBERATI, N. T., LEE, D., URBACH, J., GOODMAN, H. M. & RAHME, L. G. 2004. The broad host range pathogen *Pseudomonas aeruginosa* strain PA14 carries two pathogenicity islands harboring plant and animal virulence genes. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 2530-2535.
- HENRY, C. S., DEJONGH, M., BEST, A. A., FRYBARGER, P. M., LINSAY, B. & STEVENS, R. L. 2010. High-throughput generation, optimization and analysis of genome-scale metabolic models. *Nature Biotechnology*, **28**, 977-982.
- HERRERA, A. G. 2004. Microbiological analysis of cosmetics. *Public Health Microbiology: Methods and Protocols*, 293-295.

- HEURLIER, K., DÉNERVAUD, V., HAENNI, M., GUY, L., KRISHNAPILLAI, V. & HAAS, D. 2005. Quorum-sensing-negative (*lasR*) mutants of *Pseudomonas aeruginosa* avoid cell lysis and death. *Journal of Bacteriology*, **187**, 4875-4883.
- HIGGINS, C. F. 2007. Multiple molecular mechanisms for multidrug resistance transporters. *Nature*, **446**, 749-757.
- HINZ, A., LEE, S., JACOBY, K. & MANOIL, C. 2011. Membrane proteases and aminoglycoside antibiotic resistance. *Journal of Bacteriology*, **193**, 4790-4797.
- HOCQUET, D., VOGNE, C., EL GARCH, F., VEJUX, A., GOTOH, N., LEE, A., LOMOVSKAYA, O. & PLÉSIAT, P. 2003. MexXY-OprM efflux pump is necessary for adaptive resistance of *Pseudomonas aeruginosa* to aminoglycosides. *Antimicrobial Agents and Chemotherapy*, **47**, 1371-1375.
- HOVE-JENSEN, B., ROSENKRANTZ, T. J., HALDIMANN, A. & WANNER, B. L. 2003. *Escherichia coli* *phnN*, encoding ribose 1, 5-bisphosphokinase activity (phosphoribosyl diphosphate forming): dual role in phosphonate degradation and NAD biosynthesis pathways. *Journal of Bacteriology*, **185**, 2793-2801.
- HOYLE, B. D. & COSTERTON, J. W. 1991. Bacterial resistance to antibiotics: the role of biofilms. *Progress in Drug Research/Fortschritte der Arzneimittelforschung/Progrès des recherches pharmaceutiques*. Birkhäuser Basel: Springer.
- HUERTA-CEPAS, J., FORSLUND, K., SZKLARCZYK, D., JENSEN, L. J., VON MERING, C. & BORK, P. 2016a. Fast genome-wide functional annotation through orthology assignment by EggNOG-mapper. *bioRxiv*, 076331.
- HUERTA-CEPAS, J., SZKLARCZYK, D., FORSLUND, K., COOK, H., HELLER, D., WALTER, M. C., RATTEI, T., MENDE, D. R., SUNAGAWA, S. & KUHN, M. 2016b. EggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucleic Acids Research*, **44**, D286-D293.
- HUMMERJOHANN, J., KÜTTEL, E., QUADRONI, M., RAGALLER, J., LEISINGER, T. & KERTES, M. A. 1998. Regulation of the sulfate starvation response in *Pseudomonas aeruginosa*: role of cysteine biosynthetic intermediates. *Microbiology*, **144**, 1375-1386.
- HUNT, M., DE SILVA, N., OTTO, T. D., PARKHILL, J., KEANE, J. A. & HARRIS, S. R. 2015. Circlator: automated circularization of genome assemblies using long sequencing reads. *Genome Biology*, **16**, 294.
- HYDUKE, D. R., LEWIS, N. E. & PALSSON, B. Ø. 2013. Analysis of omics data with genome-scale models of metabolism. *Molecular BioSystems*, **9**, 167-174.
- HYTTIÄINEN, H., MONTESANO, M. & PALVA, E. T. 2001. Global regulators ExpA (GacA) and KdgR modulate extracellular enzyme gene expression through the RsmA-rsmB system in *Erwinia carotovora* *subsp. carotovora*. *Molecular Plant-Microbe Interactions*, **14**, 931-938.
- ICHISE, Y.-K., KOSUGE, T., UWATE, M., NAKAE, T. & MASEDA, H. 2015. Complete genome sequence of *Pseudomonas aeruginosa* strain 8380, isolated from the human gut. *Genome Announcements*, **3**, e00520-15.
- IKARASHI, Y., UCHINO, T. & NISHIMURA, T. 2009. Analysis of preservatives used in cosmetic products: salicylic acid, sodium benzoate, sodium dehydroacetate, potassium sorbate, phenoxyethanol, and parabens. *Kokuritsu Iyakuhiin Shokuhin Eisei Kenkyujo Hokoku= Bulletin of National Institute of Health Sciences*, **128**, 85-90.
- ITO, K. & AKIYAMA, Y. 2005. Cellular functions, mechanism of action, and regulation of FtsH protease. *Annual Review of Microbiology*, **59**, 211-231.
- JACOBS, M. A., ALWOOD, A., THAIPISUTTIKUL, I., SPENCER, D., HAUGEN, E., ERNST, S., WILL, O., KAUL, R., RAYMOND, C. & LEVY, R. 2003. Comprehensive transposon mutant library of *Pseudomonas aeruginosa*. *Proceedings of the National Academy of Sciences*, **100**, 14339-14344.
- JÁCOME, P. R. L. D. A., ALVES, L. R., CABRAL, A. B., LOPES, A. C. S. & MACIEL, M. A. V. 2012. Phenotypic and molecular characterization of antimicrobial resistance and virulence factors in *Pseudomonas aeruginosa* clinical isolates from Recife, State of Pernambuco, Brazil. *Revista da Sociedade Brasileira de Medicina Tropical*, **45**, 707-712.
- JENSEN, P. A. & PAPIN, J. A. 2011. Functional integration of a metabolic network model and expression data without arbitrary thresholding. *Bioinformatics*, **27**, 541-547.
- JENSEN, P. A., LUTZ, K. A. & PAPIN, J. A. 2011. TIGER: Toolbox for integrating genome-scale metabolic models, expression data, and transcriptional regulatory networks. *BMC Systems Biology*, **5**, 147.
- JEUKENS, J., BOYLE, B., BIANCONI, I., KUKAVICA-IBRULJ, I., TÜMMLER, B., BRAGONZI, A. & LEVESQUE, R. C. 2013. Complete genome sequence of persistent cystic fibrosis isolate *Pseudomonas aeruginosa* strain RP73. *Genome Announcements*, **1**, e00568-13.
- JEUKENS, J., BOYLE, B., KUKAVICA-IBRULJ, I., OUELLET, M. M., AARON, S. D., CHARETTE, S. J., FOTHERGILL, J. L., TUCKER, N. P., WINSTANLEY, C. & LEVESQUE, R. C. 2014. Comparative genomics of isolates of a *Pseudomonas aeruginosa* epidemic strain associated with chronic lung infections of cystic fibrosis patients. *PLoS One*, **9**, e87611.
- JEUKENS, J., FRESCHI, L., KUKAVICA-IBRULJ, I., EMOND-RHEAULT, J. G., TUCKER, N. P. & LEVESQUE, R. C. 2017. Genomics of antibiotic-resistance prediction in *Pseudomonas aeruginosa*. *Annals of the New York Academy of Sciences*.
- JIMENEZ, L. 2001. Molecular diagnosis of microbial contamination in cosmetic and pharmaceutical products: a review. *Journal of AOAC International*, **84**, 671-675.
- JIMENEZ, L. 2007. Microbial diversity in pharmaceutical product recalls and environments. *PDA Journal of Pharmaceutical Science and Technology*, **61**, 383-399.

- JOHANSSON, S., LIND, M. L., BOMAN, A. & LIDÉN, C. 2011. Preservatives and fragrances in selected consumer-available cosmetics and detergents. *Contact Dermatitis*, **64**, 265-272.
- JONES, C. J., NEWSOM, D., KELLY, B., IRIE, Y., JENNINGS, L. K., XU, B., LIMOLI, D. H., HARRISON, J. J., PARSEK, M. R. & WHITE, P. 2014. ChIP-Seq and RNA-Seq reveal an AmrZ-mediated mechanism for cyclic di-GMP synthesis and biofilm development by *Pseudomonas aeruginosa*. *PLoS Pathogens*, **10**, e1003984.
- JUHAS, M., EBERL, L. & CHURCH, G. M. 2012. Essential genes as antimicrobial targets and cornerstones of synthetic biology. *Trends in Biotechnology*, **30**, 601-607.
- JUHAS, M., EBERL, L. & TÜMMLER, B. 2005. Quorum sensing: the power of cooperation in the world of *Pseudomonas*. *Environmental Microbiology*, **7**, 459-471.
- JUHAS, M., VAN DER MEER, J. R., GAILLARD, M., HARDING, R. M., HOOD, D. W. & CROOK, D. W. 2009. Genomic islands: tools of bacterial horizontal gene transfer and evolution. *FEMS Microbiology Reviews*, **33**, 376-393.
- JUN, S.-R., WASSENAAR, T. M., NOOKAEW, I., HAUSER, L., WANCHAI, V., LAND, M., TIMM, C., LU, T.-Y. S., SCHADT, C. W. & DOKTYCZ, M. J. 2015. Comparative genome analysis of *Pseudomonas* genomes including Populus-associated isolates. *Applied and Environmental Microbiology*, 02612-02615.
- JURCISEK, J., GREINER, L., WATANABE, H., ZALESKI, A., APICELLA, M. A. & BAKALETZ, L. O. 2005. Role of sialic acid and complex carbohydrate biosynthesis in biofilm formation by nontypeable *Haemophilus influenzae* in the chinchilla middle ear. *Infection and Immunity*, **73**, 3210-3218.
- KADNER, R. J. & SHATTUCK-EIDENS, D. M. 1983. Genetic control of the hexose phosphate transport system of *Escherichia coli*: mapping of deletion and insertion mutations in the uhp region. *Journal of Bacteriology*, **155**, 1052-1061.
- KALELI, I., CEVAHIR, N., DEMIR, M., YILDIRIM, U. & SAHIN, R. 2007. Anticandidal activity of *Pseudomonas aeruginosa* strains isolated from clinical specimens. *Mycoses*, **50**, 74-78.
- KAMATH, S., CHEN, M. & CHAKRABARTY, A. 2000. Secretion of nucleoside diphosphate kinase by mucoid *Pseudomonas aeruginosa* 8821: involvement of a carboxy-terminal motif in secretion. *Journal of Bacteriology*, **182**, 3826-3831.
- KANEHISA, M., FURUMICHI, M., TANABE, M., SATO, Y. & MORISHIMA, K. 2017. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Research*, **45**, D353-D361.
- KANG, Y., DURFEE, T., GLASNER, J. D., QIU, Y., FRISCH, D., WINTERBERG, K. M. & BLATTNER, F. R. 2004. Systematic mutagenesis of the *Escherichia coli* genome. *Journal of Bacteriology*, **186**, 4921-4930.
- KARATAN, E. & WATNICK, P. 2009. Signals, regulatory networks, and materials that build and break bacterial biofilms. *Microbiology and Molecular Biology Reviews*, **73**, 310-347.
- KARNA, S. R., CHEN, T., CHEN, P., PEACOCK, T. J., ABERCROMBIE, J. J. & LEUNG, K. P. 2016. Genome Sequence of a Virulent *Pseudomonas aeruginosa* Strain, 12-4-4 (59), Isolated from the Blood Culture of a Burn Patient. *Genome Announcements*, **4**, e00079-16.
- KASHAF, S. S., ANGIONE, C. & LIÓ, P. 2017. Making life difficult for *Clostridium difficile*: augmenting the pathogen's metabolic model with transcriptomic and codon usage data for better therapeutic target characterization. *BMC Systems Biology*, **11**, 25.
- KELLY, D. J. & THOMAS, G. H. 2001. The tripartite ATP-independent periplasmic (TRAP) transporters of bacteria and archaea. *FEMS Microbiology Reviews*, **25**, 405-424.
- KELLY, M. & CLARKE, P. H. 1962. An inducible amidase produced by a strain of *Pseudomonas aeruginosa*. *Microbiology*, **27**, 305-316.
- KERR, K. & SNELLING, A. 2009. *Pseudomonas aeruginosa*: a formidable and ever-present adversary. *Journal of Hospital Infection*, **73**, 338-344.
- KHALEDI, A., SCHNIEDERJANS, M., POHL, S., RAINER, R., BODENHOFER, U., XIA, B., KLAWONN, F., BRUCHMANN, S., PREUSSE, M. & ECKWEILER, D. 2016. Transcriptome profiling of antimicrobial resistance in *Pseudomonas aeruginosa*. *Antimicrobial Agents and Chemotherapy*, **60**, 4722-4733.
- KHAN, N. H., ISHII, Y., KIMATA-KINO, N., ESAKI, H., NISHINO, T., NISHIMURA, M. & KOGURE, K. 2007. Isolation of *Pseudomonas aeruginosa* from Open Ocean and comparison with freshwater, clinical, and animal isolates. *Microbial Ecology*, **53**, 173-186.
- KHAN, R., KONG, H. G., JUNG, Y.-H., CHOI, J., BAEK, K.-Y., HWANG, E. C. & LEE, S.-W. 2016. Triclosan resistome from metagenome reveals diverse enoyl acyl carrier protein reductases and selective enrichment of triclosan resistance genes. *Scientific Reports*, **6**.
- KHATUA, B., GHOSHAL, A., BHATTACHARYA, K., MANDAL, C., SAHA, B., CROCKER, P. R. & MANDAL, C. 2010. Sialic acids acquired by *Pseudomonas aeruginosa* are involved in reduced complement deposition and siglec mediated host-cell recognition. *FEBS Letters*, **584**, 555-561.
- KING, A., LODOLA, A., CARMI, C., FU, J., MOR, M. & PIOMELLI, D. 2009. A critical cysteine residue in monoacylglycerol lipase is targeted by a new class of isothiazolinone-based enzyme inhibitors. *British Journal of Pharmacology*, **157**, 974-983.
- KING, J. D., KOCÍNCOVÁ, D., WESTMAN, E. L. & LAM, J. S. 2009. Lipopolysaccharide biosynthesis in *Pseudomonas aeruginosa*. *Innate Immunity*, **15**, 261-312.
- KLEMM, E. & DOUGAN, G. 2016. Advances in understanding bacterial pathogenesis gained from whole-genome sequencing and phylogenetics. *Cell Host & Microbe*, **19**, 599-610.

- KLOCKGETHER, J., CRAMER, N., WIEHLMANN, L., DAVENPORT, C. F. & TÜMMLER, B. 2011. *Pseudomonas aeruginosa* genomic structure and diversity. *Frontiers in Microbiology*, **2**.
- KLOCKGETHER, J., MUNDER, A., NEUGEBAUER, J., DAVENPORT, C. F., STANKE, F., LARBIG, K. D., HEEB, S., SCHÖCK, U., POHL, T. M. & WIEHLMANN, L. 2010. Genome diversity of *Pseudomonas aeruginosa* PAO1 laboratory strains. *Journal of Bacteriology*, **192**, 1113-1121.
- KOEBNIK, R., LOCHER, K. P. & VAN GELDER, P. 2000. Structure and function of bacterial outer membrane proteins: barrels in a nutshell. *Molecular Microbiology*, **37**, 239-253.
- KOHANSKI, M. A., DWYER, D. J. & COLLINS, J. J. 2010. How antibiotics kill bacteria: from targets to networks. *Nature Reviews Microbiology*, **8**, 423-35.
- KOHANSKI, M. A., DWYER, D. J., WIERZBOWSKI, J., COTTAREL, G. & COLLINS, J. J. 2008. Mistranslation of membrane proteins and two-component system activation trigger antibiotic-mediated cell death. *Cell*, **135**, 679-690.
- KÖHLER, T., EPP, S. F., CURTY, L. K. & PECHÈRE, J.-C. 1999. Characterization of MexT, the regulator of the MexE-MexF-OprN multidrug efflux system of *Pseudomonas aeruginosa*. *Journal of Bacteriology*, **181**, 6300-6305.
- KOMINOS, S. D., COPELAND, C. E., GROSIK, B. & POSTIC, B. 1972. Introduction of *Pseudomonas aeruginosa* into a hospital via vegetables. *Applied Microbiology*, **24**, 567-570.
- KONSTANTINIDIS, K. T. & TIEDJE, J. M. 2004. Trends between gene content and genome size in prokaryotic species with larger genomes. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 3160-3165.
- KOREN, S., WALENZ, B. P., BERLIN, K., MILLER, J. R., BERGMAN, N. H. & PHILLIPPY, A. M. 2017. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *bioRxiv*, 071282.
- KRAAK, W. & VAN DER ZEE, A. 2015. *Pseudomonas aeruginosa* strain S04 90 plasmid, complete sequence [Online]. Available: <https://www.ncbi.nlm.nih.gov/nuccore/CP011370.1> [Accessed 03 March 2017].
- KROWKA, J., LORETZ, L., BRZUSKA, K., ALMEIDA, J., DIEHL, M., GONSIOR, S., JOHNSON, A., SELAM, S., BADE, S. & CHAMP, S. 2014. Phenoxyethanol as a Safe and Important Preservative in Personal Care. *Cosm & Toil*, **129**, 24-7.
- KULAKOV, L. A., MCALISTER, M. B., OGDEN, K. L., LARKIN, M. J. & O'HANLON, J. F. 2002. Analysis of bacteria contaminating ultrapure water in industrial systems. *Applied and Environmental Microbiology*, **68**, 1548-1555.
- KUMAR, A., CHUA, K.-L. & SCHWEIZER, H. P. 2006. Method for regulated expression of single-copy efflux pump genes in a surrogate *Pseudomonas aeruginosa* strain: identification of the BpeEF-OprC chloramphenicol and trimethoprim efflux pump of *Burkholderia pseudomallei* 1026b. *Antimicrobial Agents and Chemotherapy*, **50**, 3460-3463.
- KUNG, V. L., OZER, E. A. & HAUSER, A. R. 2010. The accessory genome of *Pseudomonas aeruginosa*. *Microbiology and Molecular Biology Reviews*, **74**, 621-641.
- KUTTY, P. K., MOODY, B., GULLION, J. S., ZERVOS, M., AJLUNI, M., WASHBURN, R., SANDERSON, R., KAINER, M. A., POWELL, T. A. & CLARKE, C. F. 2007. Multistate outbreak of *Burkholderia cenocepacia* colonization and infection associated with the use of intrinsically contaminated alcohol-free mouthwash. *CHEST Journal*, **132**, 1825-1831.
- LABAUVE, A. E. & WARGO, M. J. 2012. Growth and laboratory maintenance of *Pseudomonas aeruginosa*. *Current Protocols in Microbiology*, **6E**, 1-8.
- LAM, J. S., TAYLOR, V. L., ISLAM, S. T., HAO, Y. & KOCÍNCOVÁ, D. 2011. Genetic and functional diversity of *Pseudomonas aeruginosa* lipopolysaccharide. *Frontiers in Microbiology*, **2**.
- LAMA, A., ALLEN, S. C., ATTENBURROW, G. E., BATES, M. P., COVINGTON, A. D. & ANTUNES, A. P. M. 2012. Survival and growth of *Bacillus cereus*, *Pseudomonas aeruginosa* and *Staphylococcus aureus* in extreme environments present during the early stages of a conventional leather making process. *Journal of American Leather Chemists Association*, **107**, 186-195.
- LAMBERT, P. A. 2002. Mechanisms of antibiotic resistance in *Pseudomonas aeruginosa*. *Journal of the Royal Society*, **95**, 22-26.
- LAMBERT, P. A. 2005. Bacterial resistance to antibiotics: modified target sites. *Advanced Drug Delivery Reviews*, **57**, 1471-1485.
- LAMERS, R. P., CAVALLARI, J. F. & BURROWS, L. L. 2013. The efflux inhibitor phenylalanine-arginine beta-naphthylamide (PAβN) permeabilizes the outer membrane of gram-negative bacteria. *PLoS One*, **8**, e60666.
- LANGAN, K. M., KOTSIMBOS, T. & PELEG, A. Y. 2015. Managing *Pseudomonas aeruginosa* respiratory infections in cystic fibrosis. *Current Opinion in Infectious Diseases*, **28**, 547-556.
- LANGRIDGE, G. C., PHAN, M.-D., TURNER, D. J., PERKINS, T. T., PARTS, L., HAASE, J., CHARLES, I., MASKELL, D. J., PETERS, S. E. & DOUGAN, G. 2009. Simultaneous assay of every *Salmonella Typhi* gene using one million transposon mutants. *Genome Research*, **19**, 2308-2316.
- LANGSRUD, S., STEINHAEUER, K., LÜTHJE, S., WEBER, K., GORONCY-BERMES, P. & HOLCK, A. L. 2016. Ethylhexylglycerin Impairs Membrane Integrity and Enhances the Lethal Effect of Phenoxyethanol. *PLoS One*, **11**, e0165228.
- LAOPAIBOON, L., SMITH, R. & HALL, S. 2001. A study of the effect of isothiazolones on the performance and characteristics of a laboratory-scale rotating biological contactor. *Journal of Applied Microbiology*, **91**, 93-103.

- LATIFI, A., WINSON, M. K., FOGILINO, M., BYCROFT, B. W., STEWART, G. S., LAZDUNSKI, A. & WILLIAMS, P. 1995. Multiple homologues of LuxR and LuxI control expression of virulence determinants and secondary metabolites through quorum sensing in *Pseudomonas aeruginosa* PAO1. *Molecular Microbiology*, **17**, 333-343.
- LEE, D. G., URBACH, J. M., WU, G., LIBERATI, N. T., FEINBAUM, R. L., MIYATA, S., DIGGINS, L. T., HE, J., SAUCIER, M. & DÉZIEL, E. 2006a. Genomic analysis reveals that *Pseudomonas aeruginosa* virulence is combinatorial. *Genome Biology*, **7**, R90.
- LEE, J. & ZHANG, L. 2015. The hierarchy quorum sensing network in *Pseudomonas aeruginosa*. *Protein & Cell*, **6**, 26-41.
- LEE, J. M., GIANCHANDANI, E. P. & PAPIN, J. A. 2006b. Flux balance analysis in the era of metabolomics. *Briefings in Bioinformatics*, **7**, 140-150.
- LETUNIC, I. & BORK, P. 2016. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Research*, **44**, W242-W245.
- LEVIN, J. Z., YASSOUR, M., ADICONIS, X., NUSBAUM, C., THOMPSON, D. A., FRIEDMAN, N., GNIRKE, A. & REGEV, A. 2010. Comprehensive comparative analysis of strand-specific RNA sequencing methods. *Nature Methods*, **7**, 709-715.
- LEVY, S. E. & MYERS, R. M. 2016. Advancements in next-generation sequencing. *Annual Review of Genomics and Human Genetics*, **17**, 95-115.
- LEWENZA, S., FALSAFI, R. K., WINSOR, G., GOODERHAM, W. J., MCPHEE, J. B., BRINKMAN, F. S. & HANCOCK, R. E. 2005. Construction of a mini-Tn5-*luxCDABE* mutant library in *Pseudomonas aeruginosa* PAO1: a tool for identifying differentially regulated genes. *Genome Research*, **15**, 583-589.
- LEWENZA, S., GARDY, J. L., BRINKMAN, F. S. & HANCOCK, R. E. 2005. Genome-wide identification of *Pseudomonas aeruginosa* exported proteins using a consensus computational strategy combined with a laboratory-based PhoA fusion screen. *Genome Research*, **15**, 321-329.
- LEWIS, N. E., NAGARAJAN, H. & PALSSON, B. O. 2012. Constraining the metabolic genotype-phenotype relationship using a phylogeny of *in silico* methods. *Nature reviews Microbiology*, **10**, 291.
- LI, G., LU, S., SHEN, M., LE, S., TAN, Y., LI, M., ZHAO, X., WANG, J., SHEN, W. & GUO, K. 2016. Complete genome sequence of *Pseudomonas aeruginosa* phage-resistant variant PA1RG. *Genome Announcements*, **4**, e01761-15.
- LI, G., SHEN, M., LE, S., TAN, Y., LI, M., ZHAO, X., SHEN, W., YANG, Y., WANG, J. & ZHU, H. 2016. Genomic analyses of multidrug resistant *Pseudomonas aeruginosa* PA1 resequenced by single-molecule real-time sequencing. *Bioscience Reports*, **36**, e00418.
- LI, H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv preprint arXiv:1303.3997*.
- LI, H., HANDSAKER, B., WYSOKER, A., FENNELLS, T., RUAN, J., HOMER, N., MARTH, G., ABECASIS, G. & DURBIN, R. 2009. The sequence alignment/map format and SAMtools. *Bioinformatics*, **25**, 2078-2079.
- LI, Y., MIMA, T., KOMORI, Y., MORITA, Y., KURODA, T., MIZUSHIMA, T. & TSUCHIYA, T. 2003. A new member of the tripartite multidrug efflux pumps, MexVW-OprM, in *Pseudomonas aeruginosa*. *Journal of Antimicrobial Chemotherapy*, **52**, 572-575.
- LIBERATI, N. T., URBACH, J. M., MIYATA, S., LEE, D. G., DRENKARD, E., WU, G., VILLANUEVA, J., WEI, T. & AUSUBEL, F. M. 2006. An ordered, nonredundant library of *Pseudomonas aeruginosa* strain PA14 transposon insertion mutants. *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 2833-2838.
- LIPPA, B., MORRIS, J., CORBETT, M., KWAN, T. A., NOE, M. C., SNOW, S. L., GANT, T. G., MANGIARACINA, M., COFFEY, H. A. & FOSTER, B. 2006. Discovery of novel isothiazole inhibitors of the TrkA kinase: structure-activity relationship, computer modeling, optimization, and identification of highly potent antagonists. *Bioorganic & Medicinal Chemistry Letters*, **16**, 3444-3448.
- LISTER, P., WOLTER, D. & HANSON, N. 2009. Antibacterial-Resistant *Pseudomonas aeruginosa*: Clinical Impact and Complex Regulation of Chromosomally Encoded Resistance Mechanisms. *Clinical Microbiology Reviews*, **22**, 582-610.
- LIU, H., LIANG, R., TAO, F., MA, C., LIU, Y., LIU, X. & LIU, J. 2012. Genome sequence of *Pseudomonas aeruginosa* strain SJTD-1, a bacterium capable of degrading long-chain alkanes and crude oil. *Journal of Bacteriology*, **194**, 4783-4784.
- LIU, L., AGREN, R., BORDEL, S. & NIELSEN, J. 2010. Use of genome-scale metabolic models for understanding microbial physiology. *FEBS Letters*, **584**, 2556-2564.
- LIU, L., CUI, Y., ZHENG, B., JIANG, S., YU, W., SHEN, P., JI, J., LI, L., QIN, N. & XIAO, Y. 2016. Analysis of tigecycline resistance development in clinical *Acinetobacter baumannii* isolates through a combined genomic and transcriptomic approach. *Scientific Reports*, **6**, 26930.
- LIU, X., YIN, Y., WU, J. & LIU, Z. 2014. Structure and mechanism of an intramembrane liponucleotide synthetase central for phospholipid biosynthesis. *Nature Communications*, **5**.
- LIU, Y., YANG, L. & MOLIN, S. 2010. Synergistic activities of an efflux pump inhibitor and iron chelators against *Pseudomonas aeruginosa* growth and biofilm formation. *Antimicrobial Agents and Chemotherapy*, **54**, 3960-3963.
- LIVERMORE, D. M. 2002. Multiple mechanisms of antimicrobial resistance in *Pseudomonas aeruginosa*: our worst nightmare? *Clinical Infectious Diseases*, **34**, 634-640.
- LLANES, C., KÖHLER, T., PATRY, I., DEHECQ, B., VAN DELDEN, C. & PLÉSIAT, P. 2011. Role of the efflux system MexEF-OprN in low level resistance of *Pseudomonas aeruginosa* to ciprofloxacin. *Antimicrobial Agents and Chemotherapy*, AAC. 00101-11.

- LOMOVSKAYA, O., WARREN, M. S., LEE, A., GALAZZO, J., FRONKO, R., LEE, M., BLAIS, J., CHO, D., CHAMBERLAND, S. & RENAULT, T. 2001. Identification and characterization of inhibitors of multidrug resistance efflux pumps in *Pseudomonas aeruginosa*: novel agents for combination therapy. *Antimicrobial Agents and Chemotherapy*, **45**, 105-116.
- LOVE, M. I., HUBER, W. & ANDERS, S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, **15**, 550.
- LOW, A. S., MACKENZIE, F. M., GOULD, I. M. & BOOTH, I. R. 2001. Protected environments allow parallel evolution of a bacterial pathogen in a patient subjected to long-term antibiotic therapy. *Molecular Microbiology*, **42**, 619-630.
- LÖYTYNOJA, A. 2014. Phylogeny-aware alignment with PRANK. *Multiple Sequence Alignment Methods*, 155-170.
- LU, S., LE, S., LI, G., SHEN, M., TAN, Y., ZHAO, X., WANG, J., SHEN, W., GUO, K. & YANG, Y. 2015. Complete genome sequence of *Pseudomonas aeruginosa* PA1, isolated from a patient with a respiratory tract infection. *Genome Announcements*, **3**, e01453-15.
- LUNDGREN, B. R., THORNTON, W., DORNAN, M. H., VILLEGAS-PENARANDA, L. R., BODDY, C. N. & NOMURA, C. T. 2013. Gene PA2449 is essential for glycine metabolism and pyocyanin biosynthesis in *Pseudomonas aeruginosa* PAO1. *Journal of Bacteriology*, **195**, 2087-2100.
- LUNDOV, M. D. & ZACHARIAE, C. 2008. Recalls of microbiologically contaminated cosmetics in EU from 2005 to May 2008. *International Journal of Cosmetic Science*, **30**, 471-474.
- LUNDOV, M. D., KOLARIK, B., BOSSI, R., GUNNARSEN, L. & JOHANSEN, J. D. 2014. Emission of isothiazolinones from water-based paints. *Environmental Science & Technology*, **48**, 6989-6994.
- LUNDOV, M. D., MOESBY, L., ZACHARIAE, C. & JOHANSEN, J. D. 2009. Contamination versus preservation of cosmetics: a review on legislation, usage, infections, and contact allergy. *Contact Dermatitis*, **60**, 70-78.
- LUNDOV, M., JOHANSEN, J., ZACHARIAE, C. & MOESBY, L. 2011. Low-level efficacy of cosmetic preservatives. *International Journal of Cosmetic Science*, **33**, 190-196.
- LYCZAK, J. B., CANNON, C. L. & PIER, G. B. 2000. Establishment of *Pseudomonas aeruginosa* infection: lessons from a versatile opportunist. *Microbes and Infection*, **2**, 1051-1060.
- MA, L., CONOVER, M., LU, H., PARSEK, M. R., BAYLES, K. & WOZNIK, D. J. 2009. Assembly and development of the *Pseudomonas aeruginosa* biofilm matrix. *PLoS Pathogens*, **5**, e1000354.
- MA, Q., ZHAI, Y., SCHNEIDER, J. C., RAMSEIER, T. M. & SAIER, M. H. 2003. Protein secretion systems of *Pseudomonas aeruginosa* and *P. fluorescens*. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, **1611**, 223-233.
- MACFARLANE, E. L., KWASNICKA, A., OCHS, M. M. & HANCOCK, R. E. 1999. PhoP-PhoQ homologues in *Pseudomonas aeruginosa* regulate expression of the outer-membrane protein OprH and polymyxin B resistance. *Molecular Microbiology*, **34**, 305-316.
- MACLEOD, D. L., NELSON, L. E., SHAWAR, R. M., LIN, B. B., LOCKWOOD, L. G., DIRKS, J. E., MILLER, G. H., BURNS, J. L. & GARBER, R. L. 2000. Aminoglycoside-resistance mechanisms for cystic fibrosis *Pseudomonas aeruginosa* isolates are unchanged by long-term, intermittent, inhaled tobramycin treatment. *The Journal of Infectious Diseases*, **181**, 1180-1184.
- MÄDER, U., NICOLAS, P., RICHARD, H., BESSIÈRES, P. & AYMERICH, S. 2011. Comprehensive identification and quantification of microbial transcriptomes by genome-wide unbiased methods. *Current Opinion in Biotechnology*, **22**, 32-41.
- MAH, T.-F., PITTS, B., PELLOCK, B., WALKER, G. C., STEWART, P. S. & O'TOOLE, G. A. 2003. A genetic basis for *Pseudomonas aeruginosa* biofilm antibiotic resistance. *Nature*, **426**, 306-310.
- MAHENTHIRALINGAM, E., CAMPBELL, M. E., FOSTER, J., LAM, J. S. & SPEERT, D. P. 1996. Random amplified polymorphic DNA typing of *Pseudomonas aeruginosa* isolates recovered from patients with cystic fibrosis. *Journal of Clinical Microbiology*, **34**, 1129-1135.
- MAILLARD, J. Y. 2002. Bacterial target sites for biocide action. *Journal of Applied Microbiology*, **92**, 16S-27S.
- MANTIONE, K. J., KREAM, R. M., KUZELOVA, H., PTACEK, R., RABOCH, J., SAMUEL, J. M. & STEFANO, G. B. 2014. Comparing bioinformatic gene expression profiling methods: microarray and RNA-Seq. *Medical Science Monitor Basic Research*, **20**, 138-141.
- MANUJ, K., GUNDERSON, C., TROUPE, J. & HUBER, M. E. 2006. Efficacy of contact lens disinfecting solutions against *Staphylococcus aureus* and *Pseudomonas aeruginosa*. *Eye & Contact Lens*, **32**, 216-218.
- MARQUES, A. M., CONGREGADO, F. & SIMONPUJOL, D. M. 1979. Antibiotic and heavy metal resistance of *Pseudomonas aeruginosa* isolated from soils. *Journal of Applied Microbiology*, **47**, 347-350.
- MARSHALL, B., STINTZI, A., GILMOUR, C., MEYER, J.-M. & POOLE, K. 2009. Citrate-mediated iron uptake in *Pseudomonas aeruginosa*: involvement of the citrate-inducible FecA receptor and the FeoB ferrous iron transporter. *Microbiology*, **155**, 305-315.
- MASUDA, N., SAKAGAWA, E., OHYA, S., GOTOH, N., TSUJIMOTO, H. & NISHINO, T. 2000. Substrate specificities of MexAB-OprM, MexCD-OprJ, and MexXY-oprM efflux pumps in *Pseudomonas aeruginosa*. *Antimicrobial Agents and Chemotherapy*, **44**, 3322-3327.
- MATHEE, K., NARASIMHAN, G., VALDES, C., QIU, X., MATEWISH, J. M., KOEHRSEN, M., ROKAS, A., YANDAVA, C. N., ENGELS, R. & ZENG, E. 2008. Dynamics of *Pseudomonas aeruginosa* genome evolution. *Proceedings of the National Academy of Sciences*, **105**, 3100-3105.
- MATTICK, J. S. 2002. Type IV pili and twitching motility. *Annual Reviews in Microbiology*, **56**, 289-314.

- MAURA, D., HAZAN, R., KITAO, T., BALLOK, A. E. & RAHME, L. G. 2016. Evidence for direct control of virulence and defense gene circuits by the *Pseudomonas aeruginosa* quorum sensing regulator, MvR. *Scientific Reports*, **6**.
- MC CAY, P. H., OCAMPO-SOSA, A. A. & FLEMING, G. T. 2010. Effect of subinhibitory concentrations of benzalkonium chloride on the competitiveness of *Pseudomonas aeruginosa* grown in continuous culture. *Microbiology*, **156**, 30-38.
- MCCARTHUR, A. G., WAGLECHNER, N., NIZAM, F., YAN, A., AZAD, M. A., BAYLAY, A. J., BHULLAR, K., CANOVA, M. J., DE PASCALE, G. & EJIM, L. 2013. The comprehensive antibiotic resistance database. *Antimicrobial Agents and Chemotherapy*, **57**, 3348-3357.
- MCBAIN, A., RICKARD, A. & GILBERT, P. 2002. Possible implications of biocide accumulation in the environment on the prevalence of bacterial antibiotic resistance. *Journal of Industrial Microbiology and Biotechnology*, **29**, 326-330.
- MCCARTHY, A. 2010. Third generation DNA sequencing: pacific biosciences' single molecule real time technology. *Chemistry & Biology*, **17**, 675-676.
- MCCOLLISTER, B. D., HOFFMAN, M., HUSAIN, M. & VÁZQUEZ-TORRES, A. 2011. Nitric oxide protects bacteria from aminoglycosides by blocking the energy-dependent phases of drug uptake. *Antimicrobial Agents and Chemotherapy*, **55**, 2189-2196.
- MCDONNELL, G. & RUSSELL, A. D. 1999. Antiseptics and disinfectants: activity, action, and resistance. *Clinical Microbiology Reviews*, **12**, 147-179.
- MCNEIL, M. 1999. Arabinogalactan in mycobacteria: structure, biosynthesis, and genetics. *Genetics of Bacterial Polysaccharides*. Boca Raton, Florida: CRC Press, 207-223.
- MCPHEE, J. B., LEWENZA, S. & HANCOCK, R. E. 2003. Cationic antimicrobial peptides activate a two-component regulatory system, PmrA-PmrB, that regulates resistance to polymyxin B and cationic antimicrobial peptides in *Pseudomonas aeruginosa*. *Molecular Microbiology*, **50**, 205-217.
- MEIGHEN, E. A. 1994. Genetics of bacterial bioluminescence. *Annual Review of Genetics*, **28**, 117-139.
- MENDUM, T. A., NEWCOMBE, J., MANNAN, A. A., KIERZEK, A. M. & MCFADDEN, J. 2011. Interrogation of global mutagenesis data with a genome scale model of *Neisseria meningitidis* to assess gene fitness in vitro and in sera. *Genome Biology*, **12**, R127.
- MESARICH, C. H., REES-GEORGE, J., GARDNER, P. P., GHOMI, F. A., GERTH, M. L., ANDERSEN, M. T., RIKKERINK, E. H., FINERAN, P. C. & TEMPLETON, M. D. 2017. Transposon insertion libraries for the characterization of mutants from the kiwifruit pathogen *Pseudomonas syringae* pv. *actinidiae*. *PLoS One*, **12**, e0172790.
- MESAROS, N., NORDMANN, P., PLÉSIAT, P., ROUSSEL-DELVALLEZ, M., VAN ELDERE, J., GLUPCZYNSKI, Y., VAN LAETHEM, Y., JACOBS, F., LEBECQUE, P. & MALFROOT, A. 2007. *Pseudomonas aeruginosa*: resistance and therapeutic options at the turn of the new millennium. *Clinical Microbiology and Infection*, **13**, 560-578.
- MIKAMI, E., GOTO, T., OHNO, T., MATSUMOTO, H. & NISHIDA, M. 2002. Simultaneous analysis of dehydroacetic acid, benzoic acid, sorbic acid and salicylic acid in cosmetic products by solid-phase extraction and high-performance liquid chromatography. *Journal of Pharmaceutical and Biomedical Analysis*, **28**, 261-267.
- MIKKELSEN, H., MCMULLAN, R. & FILLOUX, A. 2011a. The *Pseudomonas aeruginosa* reference strain PA14 displays increased virulence due to a mutation in *ladS*. *PLoS One*, **6**, e29113.
- MIKKELSEN, H., SIVANESON, M. & FILLOUX, A. 2011b. Key two-component regulatory systems that control biofilm formation in *Pseudomonas aeruginosa*. *Environmental Microbiology*, **13**, 1666-1681.
- MILES, A., MISRA, S. & IRWIN, J. 1938. The estimation of the bactericidal power of the blood. *Epidemiology & Infection*, **38**, 732-749.
- MILLER, C. L., CHEN, T., CHEN, P. & LEUNG, K. P. 2016. Genome sequence of highly virulent *Pseudomonas aeruginosa* strain VA-134, isolated from a burn patient. *Genome Announcements*, **4**, e01662-15.
- MIMA, T., JOSHI, S., GOMEZ-ESCALADA, M. & SCHWEIZER, H. P. 2007. Identification and characterization of TriABC-OpmH, a triclosan efflux pump of *Pseudomonas aeruginosa* requiring two membrane fusion proteins. *Journal of Bacteriology*, **189**, 7600-7609.
- MIMA, T., KOHIRA, N., LI, Y., SEKIYA, H., OGAWA, W., KURODA, T. & TSUCHIYA, T. 2009. Gene cloning and characteristics of the RND-type multidrug efflux pump MuxABC-OpmB possessing two RND components in *Pseudomonas aeruginosa*. *Microbiology*, **155**, 3509-3517.
- MIMA, T., SEKIYA, H., MIZUSHIMA, T., KURODA, T. & TSUCHIYA, T. 2005. Gene cloning and properties of the RND-type multidrug efflux pumps MexPQ-OpmE and MexMN-OprM from *Pseudomonas aeruginosa*. *Microbiology and Immunology*, **49**, 999-1002.
- MINE, T., MORITA, Y., KATAOKA, A., MIZUSHIMA, T. & TSUCHIYA, T. 1999. Expression in *Escherichia coli* of a new multidrug efflux pump, MexXY, from *Pseudomonas aeruginosa*. *Antimicrobial Agents and Chemotherapy*, **43**, 415-417.
- MIYAKOSHI, M., CHAO, Y. & VOGEL, J. 2015. Regulatory small RNAs from the 3' regions of bacterial mRNAs. *Current Opinion in Microbiology*, **24**, 132-139.
- MIYATA, S., CASEY, M., FRANK, D. W., AUSUBEL, F. M. & DRENKARD, E. 2003. Use of the *Galleria mellonella* caterpillar as a model host to study the role of the type III secretion system in *Pseudomonas aeruginosa* pathogenesis. *Infection and Immunity*, **71**, 2404-2413.
- MIYOSHI-AKIYAMA, T., KUWAHARA, T., TADA, T., KITAO, T. & KIRIKAE, T. 2011. Complete genome sequence of highly multidrug-resistant *Pseudomonas aeruginosa* NCGM2. S1, a representative strain of a cluster endemic to Japan. *Journal of Bacteriology*, **193**, 7010-7010.

- MOISI, M., LICHTENEGGER, S., TUTZ, S., SEPER, A., SCHILD, S. & REIDL, J. 2013. Characterizing the hexose-6-phosphate transport system of *Vibrio cholerae*, a utilization system for carbon and phosphate sources. *Journal of Bacteriology*, **195**, 1800-1808.
- MOLINA-HENARES, A. J., KRELL, T., EUGENIA GUAZZARONI, M., SEGURA, A. & RAMOS, J. L. 2006. Members of the IclR family of bacterial transcriptional regulators function as activators and/or repressors. *FEMS Microbiology Reviews*, **30**, 157-186.
- MOLLOY, M. P., GILLINGS, M., WILLCOX, M., WALSH, B. J. & WALES, N. S. 2000. Complementing genomics with proteomics: the membrane subproteome of *Pseudomonas aeruginosa* PAO1. *Electrophoresis*, **21**, 3797-3809.
- MONOGUE, M. L., KUTI, J. L. & NICOLAU, D. P. 2016. Optimizing antibiotic dosing strategies for the treatment of gram-negative infections in the era of resistance. *Expert Review of Clinical Pharmacology*, **9**, 459-476.
- MONTANARO, L., POGGI, A., VISAI, L., RAVAIOLI, S., CAMPOCCIA, D., SPEZIALE, P. & ARCIOLA, C. R. 2011. Extracellular DNA in biofilms. *The International Journal of Artificial Organs*, **34**, 824-831.
- MONTGOMERY, K., CHARLESWORTH, J. C., LEBARD, R., VISSCHER, P. T. & BURNS, B. P. 2013. Quorum sensing in extreme environments. *Life*, **3**, 131-148.
- MORADALI, M. F., GHODS, S. & REHM, B. H. 2017. *Pseudomonas aeruginosa* lifestyle: A paradigm for adaptation, survival, and persistence. *Frontiers in Cellular And Infection Microbiology*, **7**.
- MORALES, E. H., PINTO, C. A., LURASCHI, R., MUÑOZ-VILLAGRÁN, C. M., CORNEJO, F. A., SIMPKINS, S. W., NELSON, J., ARENAS, F. A., PIOTROWSKI, J. S. & MYERS, C. L. 2017. Accumulation of heme biosynthetic intermediates contributes to the antibacterial action of the metalloid tellurite. *Nature Communications*, **8**.
- MORITA, Y., MURATA, T., MIMA, T., SHIOTA, S., KURODA, T., MIZUSHIMA, T., GOTOH, N., NISHINO, T. & TSUCHIYA, T. 2003. Induction of *mexCD-oprJ* operon for a multidrug efflux pump by disinfectants in wild-type *Pseudomonas aeruginosa* PAO1. *Journal of Antimicrobial Chemotherapy*, **51**, 991-994.
- MORITA, Y., TOMIDA, J. & KAWAMURA, Y. 2012. Primary mechanisms mediating aminoglycoside resistance in the multidrug-resistant *Pseudomonas aeruginosa* clinical isolate PA7. *Microbiology*, **158**, 1071-1083.
- MORIYA, Y., ITOH, M., OKUDA, S., YOSHIZAWA, A. C. & KANEHISA, M. 2007. KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Research*, **35**, W182-W185.
- MORRISON, A. J. & WENZEL, R. P. 1984. Epidemiology of infections due to *Pseudomonas aeruginosa*. *Reviews of Infectious Diseases*, **6**, S627-S642.
- MOSQUERA-RENDÓN, J., RADA-BRAVO, A. M., CÁRDENAS-BRITO, S., CORREDOR, M., RESTREPO-PINEDA, E. & BENÍTEZ-PÁEZ, A. 2016. Pangenome-wide and molecular evolution analyses of the *Pseudomonas aeruginosa* species. *BMC Genomics*, **17**, 45.
- MOUNEIMNÉ, H., ROBERT, J., JARLIER, V. & CAMBAU, E. 1999. Type II topoisomerase mutations in ciprofloxacin-resistant strains of *Pseudomonas aeruginosa*. *Antimicrobial Agents and Chemotherapy*, **43**, 62-66.
- MULCAHY, H., CHARRON-MAZENOD, L. & LEWENZA, S. 2008. Extracellular DNA chelates cations and induces antibiotic resistance in *Pseudomonas aeruginosa* biofilms. *PLoS Pathogens*, **4**, e1000213.
- MULCAHY, L. R., ISABELLA, V. M. & LEWIS, K. 2014. *Pseudomonas aeruginosa* biofilms in disease. *Microbial Ecology*, **68**, 1-12.
- MULLIGAN, C., FISCHER, M. & THOMAS, G. H. 2011. Tripartite ATP-independent periplasmic (TRAP) transporters in bacteria and archaea. *FEMS Microbiology Reviews*, **35**, 68-86.
- MURUGAN, N., MALATHI, J., UMASHANKAR, V. & MADHAVAN, H. 2016. Unraveling genomic and phenotypic nature of multidrug-resistant (MDR) *Pseudomonas aeruginosa* VRFP04 isolated from keratitis patient. *Microbiological Research*, **193**, 140-149.
- MUTO, Y. & GOTO, S. 1986. Transformation by extracellular DNA produced by *Pseudomonas aeruginosa*. *Microbiology and immunology*, **30**, 621-628.
- NAIR, B. 2000. Final report on the safety assessment of Benzyl Alcohol, Benzoic Acid, and Sodium Benzoate. *International Journal of Toxicology*, **20**, 23-50.
- NAKANO, K., SHIROMA, A., SHIMOJI, M., TAMOTSU, H., ASHIMINE, N., OHKI, S., SHINZATO, M., MINAMI, M., NAKANISHI, T. & TERUYA, K. 2017. Advantages of genome sequencing by long-read sequencer using SMRT technology in medical area. *Human Cell*, 1-13.
- NAKANO, K., TERABAYASHI, Y., SHIROMA, A., SHIMOJI, M., TAMOTSU, H., ASHIMINE, N., OHKI, S., SHINZATO, M., TERUYA, K. & SATOU, K. 2015. First complete genome sequence of *Pseudomonas aeruginosa* (Schroeter 1872) Migula 1900 (DSM 50071T), determined using PacBio single-molecule real-time technology. *Genome Announcements*, **3**, e00932-15.
- NAUGHTON, S., PARKER, D., SEEMANN, T., THOMAS, T., TURNBULL, L., ROSE, B., BYE, P., CORDWELL, S., WHITCHURCH, C. & MANOS, J. 2011. *Pseudomonas aeruginosa* AES-1 exhibits increased virulence gene expression during chronic infection of cystic fibrosis lung. *PloS One*, **6**, e24526.
- NEHME, D. & POOLE, K. 2007. Assembly of the MexAB-OprM multidrug pump of *Pseudomonas aeruginosa*: component interactions defined by the study of pump mutant suppressors. *Journal of Bacteriology*, **189**, 6118-6127.
- NEZA, E. & CENTINI, M. 2016. Microbiologically contaminated and over-preserved cosmetic products according Rapex 2008–2014. *Cosmetics*, **3**, 3.

- NIKAIDO, H. & SAIER, M. 1992. Transport proteins in bacteria: common themes in their design. *Science-New York Then Washington*, **258**, 936-936.
- NIKAIDO, H. & TAKATSUKA, Y. 2009. Mechanisms of RND multidrug efflux pumps. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, **1794**, 769-781.
- NISHIJYO, T., HAAS, D. & ITOH, Y. 2001. The CbrA–CbrB two-component regulatory system controls the utilization of multiple carbon and nitrogen sources in *Pseudomonas aeruginosa*. *Molecular Microbiology*, **40**, 917-931.
- OBERHARDT, M. A., PUCHAŁKA, J., DOS SANTOS, V. A. M. & PAPIN, J. A. 2011. Reconciliation of genome-scale metabolic reconstructions for comparative systems analysis. *PLoS Computational Biology*, **7**, e1001116.
- OBERHARDT, M. A., PUCHAŁKA, J., FRYER, K. E., DOS SANTOS, V. A. M. & PAPIN, J. A. 2008. Genome-scale metabolic network analysis of the opportunistic pathogen *Pseudomonas aeruginosa* PAO1. *Journal of Bacteriology*, **190**, 2790-2803.
- OGLESBY-SHERROUSE, A. G., DJAPGNE, L., NGUYEN, A. T., VASIL, A. I. & VASIL, M. L. 2014. The complex interplay of iron, biofilm formation, and mucoidy affecting antimicrobial resistance of *Pseudomonas aeruginosa*. *Pathogens and Disease*, **70**, 307-320.
- OHTSUBO, Y., SATO, T., KISHIDA, K., TABATA, M., OGURA, Y., HAYASHI, T., TSUDA, M. & NAGATA, Y. 2014. Complete genome sequence of *Pseudomonas aeruginosa* MTB-1, isolated from a microbial community enriched by the technical formulation of hexachlorocyclohexane. *Genome Announcements*, **2**, e01130-13.
- OLIVER, A., CANTÓN, R., CAMPO, P., BAQUERO, F. & BLÁZQUEZ, J. 2000. High frequency of hypermutable *Pseudomonas aeruginosa* in cystic fibrosis lung infection. *Science*, **288**, 1251-1253.
- OLIVER, A., LEVIN, B. R., JUAN, C., BAQUERO, F. & BLÁZQUEZ, J. 2004. Hypermutation and the preexistence of antibiotic-resistant *Pseudomonas aeruginosa* mutants: implications for susceptibility testing and treatment of chronic infections. *Antimicrobial Agents and Chemotherapy*, **48**, 4226-4233.
- OPDAM, S., RICHELLE, A., KELLMAN, B., LI, S., ZIELINSKI, D. C. & LEWIS, N. E. 2017. A Systematic Evaluation of Methods for Tailoring Genome-Scale Metabolic Models. *Cell Systems*, **4**, 318-329. e6.
- ORGAD, O., OREN, Y., WALKER, S. L. & HERZBERG, M. 2011. The role of alginate in *Pseudomonas aeruginosa* EPS adherence, viscoelastic properties and cell attachment. *Biofouling*, **27**, 787-798.
- ORTH, D. S., KABARA, J. J., DENYER, S. P. & TAN, S. 2006. Cosmetic and drug microbiology, *Informa Health Care*.
- ORTH, J. D., THIELE, I. & PALSSON, B. Ø. 2010. What is flux balance analysis? *Nature Biotechnology*, **28**, 245-248.
- ORUS, P. & LERANOZ, S. 2005. Current trends in cosmetic microbiology. *International Microbiology*, **8**, 77-79.
- OVERHAGE, J., LEWENZA, S., MARR, A. K. & HANCOCK, R. E. 2007. Identification of genes involved in swarming motility using a *Pseudomonas aeruginosa* PAO1 mini-Tn5-lux mutant library. *Journal of Bacteriology*, **189**, 2164-2169.
- OZSOLAK, F. & MIŁOS, P. M. 2011. RNA sequencing: advances, challenges and opportunities. *Nature Reviews. Genetics*, **12**, 87.
- PAGE, A. J., CUMMINS, C. A., HUNT, M., WONG, V. K., REUTER, S., HOLDEN, M. T., FOOKES, M., FALUSH, D., KEANE, J. A. & PARKHILL, J. 2015. Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics*, **31**, 3691-3693.
- PAGÈS, J.-M., JAMES, C. E. & WINTERHALTER, M. 2008. The porin and the permeating antibiotic: a selective diffusion barrier in Gram-negative bacteria. *Nature reviews. Microbiology*, **6**, 893.
- PAL, C., BENGTSOON-PALME, J., KRISTIANSSON, E. & LARSSON, D. J. 2015. Co-occurrence of resistance genes to antibiotics, biocides and metals reveals novel insights into their co-selection potential. *BMC Genomics*, **16**, 964.
- PAL, C., BENGTSOON-PALME, J., RENSING, C., KRISTIANSSON, E. & LARSSON, D. J. 2013. BacMet: antibacterial biocide and metal resistance genes database. *Nucleic Acids Research*, **42**, D737-D743.
- PAMP, S. J., GJERMENSEN, M., JOHANSEN, H. K. & TOLKER-NIELSEN, T. 2008. Tolerance to the antimicrobial peptide colistin in *Pseudomonas aeruginosa* biofilms is linked to metabolically active cells, and depends on the *pmr* and *mexAB-oprM* genes. *Molecular Microbiology*, **68**, 223-240.
- PAPAGEORGIOU, S., VARVARESOU, A., TSIRIVAS, E. & DEMETZOS, C. 2010. New alternatives to cosmetics preservation. *Journal of Cosmetic Science*, **61**, 107.
- PASUMARTHI, R., CHANDRASEKARAN, S. & MUTNURI, S. 2013. Biodegradation of crude oil by *Pseudomonas aeruginosa* and *Escherichia fergusonii* isolated from the Goan coast. *Marine Pollution Bulletin*, **76**, 276-282.
- PATERSON, I. K., HOYLE, A., OCHOA, G., BAKER-AUSTIN, C. & TAYLOR, N. G. 2016. Optimising antibiotic usage to treat bacterial infections. *Scientific Reports*, **6**.
- PAULSEN, I. T. 2003. Multidrug efflux pumps and resistance: regulation and evolution. *Current Opinion in Microbiology*, **6**, 446-451.
- PEETERS, C., COOPER, V. S., HATCHER, P. J., VERHEYDE, B., CARLIER, A. & VANDAMME, P. 2017. Comparative genomics of *Burkholderia multivorans*, a ubiquitous pathogen with a highly conserved genomic structure. *PloS One*, **12**, e0176191.
- PELEG, A. Y., MIYAKIS, S., WARD, D. V., EARL, A. M., RUBIO, A., CAMERON, D. R., PILLAI, S., MOELLERING JR, R. C. & ELIOPOULOS, G. M. 2012. Whole genome characterization of the mechanisms of daptomycin resistance in clinical and laboratory derived isolates of *Staphylococcus aureus*. *PloS One*, **7**, e28316.

REFERENCES

- PESSI, G., WILLIAMS, F., HINDLE, Z., HEURLIER, K., HOLDEN, M. T., CÁMARA, M., HAAS, D. & WILLIAMS, P. 2001. The Global Posttranscriptional Regulator RsmA Modulates Production of Virulence Determinants and N-Acylhomoserine Lactones in *Pseudomonas aeruginosa*. *Journal of Bacteriology*, **183**, 6676-6683.
- PIDDOCK, L. J. 2006. Multidrug-resistance efflux pumps? not just for resistance. *Nature Reviews Microbiology*, **4**, 629-636.
- PIKUTA, E. V., HOOVER, R. B. & TANG, J. 2007. Microbial extremophiles at the limits of life. *Critical Reviews in Microbiology*, **33**, 183-209.
- PODNECKY, N. L., WUTHIEKANUN, V., PEACOCK, S. J. & SCHWEIZER, H. P. 2013. The BpeEF-OprC efflux pump is responsible for widespread trimethoprim resistance in clinical and environmental *Burkholderia pseudomallei* isolates. *Antimicrobial Agents and Chemotherapy*, **57**, 4381-4386.
- POOLE, K. 2001. Multidrug efflux pumps and antimicrobial resistance in *Pseudomonas aeruginosa* and related organisms. *Journal of Molecular Microbiology and Biotechnology*, **3**, 255-264.
- POOLE, K. 2002. Mechanisms of bacterial biocide and antibiotic resistance. *Journal of Applied Microbiology*, **92**, 55S-64S.
- POOLE, K. 2004. Efflux-mediated multiresistance in Gram-negative bacteria. *Clinical Microbiology and Infection*, **10**, 12-26.
- POOLE, K. 2005a. Aminoglycoside resistance in *Pseudomonas aeruginosa*. *Antimicrobial Agents and Chemotherapy*, **49**, 479-487.
- POOLE, K. 2005b. Efflux-mediated antimicrobial resistance. *Journal of Antimicrobial Chemotherapy*, **56**, 20-51.
- POOLE, K. 2011. *Pseudomonas aeruginosa*: resistance to the max. *Frontiers in Microbiology*, **2**, 1-13.
- POUDRIER, J. 1990. Final report on the safety assessment of phenoxyethanol. *Journal of the American College of Toxicology*, **9**, 259-277.
- PRESTA, L., BOSI, E., MANSOURI, L., DIJKSHOORN, L., FANI, R. & FONDI, M. 2017. Constraint-based modeling identifies new putative targets to fight colistin-resistant *A. baumannii* infections. *bioRxiv*, 111765.
- PREVENTION, C. F. D. C. A. 1998. Nosocomial *Burkholderia cepacia* infection and colonization associated with intrinsically contaminated mouthwash--Arizona, 1998. *MMWR. Morbidity and Mortality Weekly Report*, **47**, 926.
- PRICE, M. N., DEHAL, P. S. & ARKIN, A. P. 2010. FastTree 2--approximately maximum-likelihood trees for large alignments. *PLoS One*, **5**, e9490.
- PUCCI, M. J., PODOS, S. D., THANASSI, J. A., LEGGIO, M. J., BRADBURY, B. J. & DESHPANDE, M. 2011. *In vitro* and *in vivo* profiles of ACH-702, an isothiazoloquinolone, against bacterial pathogens. *Antimicrobial Agents and Chemotherapy*, **55**, 2860-2871.
- QUAIL, M. A., SMITH, M., COUPLAND, P., OTTO, T. D., HARRIS, S. R., CONNOR, T. R., BERTONI, A., SWERDLOW, H. P. & GU, Y. 2012. A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC Genomics*, **13**, 341.
- RAJA, C. E. & SELVAM, G. 2009. Plasmid profile and curing analysis of *Pseudomonas aeruginosa* as metal resistant. *International Journal of Environmental Science & Technology*, **6**, 259-266.
- RAJYAGURU, N., VELEZ, A., SANDIN, R. L., YACOUB, A. T. & GREENE, J. N. 2014. Pyorubin Producing *Pseudomonas* Scalp Infection. *Infectious Diseases in Clinical Practice*, **22**, e91-e92.
- RAMAN, K. & CHANDRA, N. 2009. Flux balance analysis of biological systems: applications and challenges. *Briefings in Bioinformatics*, **10**, 435-449.
- RAMIREZ, M. S. & TOLMASKY, M. E. 2010. Aminoglycoside modifying enzymes. *Drug Resistance Updates*, **13**, 151-171.
- RASAMIRAVAKA, T., LABTANI, Q., DUEZ, P. & EL JAZIRI, M. 2015. The formation of biofilms by *Pseudomonas aeruginosa*: a review of the natural and synthetic compounds interfering with control mechanisms. *BioMed Research International*, 2015.
- RAU, M. H., MARVIG, R. L., EHRlich, G. D., MOLIN, S. & JELSBak, L. 2012. Deletion and acquisition of genomic content during early stage adaptation of *Pseudomonas aeruginosa* to a human host environment. *Environmental Microbiology*, **14**, 2200-2211.
- REDDY, B. M., TANNEERU, K., MEETEI, P. A. & GURUPRASAD, L. 2012. 3D-QSAR and Molecular Docking Studies on Substituted Isothiazole Analogs as Inhibitors Against MEK-1 Kinase. *Chemical Biology & Drug Design*, **79**, 84-91.
- REES, D. C., JOHNSON, E. & LEWINSON, O. 2009. ABC transporters: the power to change. *Nature Reviews Molecular Cell Biology*, **10**, 218.
- REID, F. R. & WOOD, T. O. 1979. *Pseudomonas* corneal ulcer: the causative role of contaminated eye cosmetics. *Archives of Ophthalmology*, **97**, 1640-1641.
- REINHARD, E., WAEBER, R., NIEDERER, M., MAURER, T., MALY, P. & SCHERER, S. 2001. Preservation of products with MCI/MI in Switzerland. *Contact Dermatitis*, **45**, 257-264.
- RHOADS, A. & AU, K. F. 2015. PacBio sequencing and its applications. *Genomics, Proteomics & Bioinformatics*, **13**, 278-289.
- RIOU, M., AVRAIN, L., EL GARCH, F., GLUPCZYNSKI, Y., PIRNAY, J., DE VOS, D., TULKENS, P. & VAN BAMBEKE, F. 2010. Influence of antibiotic treatments on gene expression of RND efflux pumps in successive isolates of *Pseudomonas aeruginosa* collected from patients with nosocomial pneumonia hospitalized in Intensive Care Units from Belgian Teaching Hospitals. *ECCMID*, 10-13 April 2010, Vienna, Austria. P780.

- ROBINSON, M. D., MCCARTHY, D. J. & SMYTH, G. K. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, **26**, 139-140.
- ROCHA, I., FÖRSTER, J. & NIELSEN, J. 2008. Design and application of genome-scale reconstructed metabolic models. *Microbial Gene Essentiality: Protocols and Bioinformatics*, 409-431.
- RODRIGUE, A., QUENTIN, Y., LAZDUNSKI, A., MÉJEAN, V. & FOGILINO, M. 2000. Cell signalling by oligosaccharides. Two-component systems in *Pseudomonas aeruginosa*: why so many? *Trends in Microbiology*, **8**, 498-504.
- RODRÍGUEZ-MARTÍNEZ, J.-M., POIREL, L. & NORDMANN, P. 2009. Extended-spectrum cephalosporinases in *Pseudomonas aeruginosa*. *Antimicrobial Agents and Chemotherapy*, **53**, 1766-1771.
- ROMPF, A., HUNGERER, C., HOFFMANN, T., LINDENMEYER, M., RÖMLING, U., GROß, U., DOSS, M. O., ARAI, H., IGARASHI, Y. & JAHN, D. 1998. Regulation of *Pseudomonas aeruginosa* *hemF* and *hemN* by the dual action of the redox response regulators Anr and Dnr. *Molecular Microbiology*, **29**, 985-997.
- ROMSANG, A., ATICHARTPONGKUL, S., TRINACHARTVANIT, W., VATTANAVIBOON, P. & MONGKOLSUK, S. 2013. Gene expression and physiological role of *Pseudomonas aeruginosa* methionine sulfoxide reductases during oxidative stress. *Journal of Bacteriology*, **195**, 3299-3308.
- ROSENTHAL, V. D., AL-ABDELY, H. M., EL-KHOLY, A. A., ALKHAWAJA, S. A. A., LEBLEBICIOGLU, H., MEHTA, Y., RAI, V., HUNG, N. V., KANJ, S. S. & SALAMA, M. F. 2016. International Nosocomial Infection Control Consortium report, data summary of 50 countries for 2010-2015: Device-associated module. *American Journal of Infection Control*, **44**, 1495-1504.
- ROSSOLINI, G. & MANTENGOLI, E. 2005. Treatment and control of severe infections caused by multi-resistant *Pseudomonas aeruginosa*. *Clinical Microbiology and Infection*, **11**, 17-32.
- ROY, A. S., BARUAH, R., GOGOI, D., BORAH, M., SINGH, A. K. & BORUAH, H. P. D. 2013. Draft genome sequence of *Pseudomonas aeruginosa* strain N002, isolated from crude oil-contaminated soil from Geleky, Assam, India. *Genome Announcements*, **1**, e00104-12.
- ROY, P. H., TETU, S. G., LAROUCHE, A., ELBOURNE, L., TREMBLAY, S., REN, Q., DODSON, R., HARKINS, D., SHAY, R. & WATKINS, K. 2010. Complete genome sequence of the multiresistant taxonomic outlier *Pseudomonas aeruginosa* PA7. *PLoS One*, **5**, e8842.
- RUMBAUGH, K. P. 2014. Genomic complexity and plasticity ensure *Pseudomonas* success. *FEMS Microbiology Letters*, **356**, 141-143.
- RUSHTON, L., SASS, A., BALDWIN, A., DOWSON, C. G., DONOGHUE, D. & MAHENTHIRALINGAM, E. 2013. Key role for efflux in the preservative susceptibility and adaptive resistance of *Burkholderia cepacia* complex bacteria. *Antimicrobial Agents and Chemotherapy*, **57**, 2972-2980.
- RUSSELL, A. 2000. Do biocides select for antibiotic resistance? *Journal of Pharmacy and Pharmacology*, **52**, 227-233.
- RUSSELL, A. 2003a. Biocide use and antibiotic resistance: the relevance of laboratory findings to clinical and environmental situations. *The Lancet Infectious Diseases*, **3**, 794-803.
- RUSSELL, A. 2003b. Challenge testing: principles and practice. *International Journal of Cosmetic Science*, **25**, 147-153.
- RUSSELL, A. D. 2003c. Similarities and differences in the responses of microorganisms to biocides. *Journal of Antimicrobial Chemotherapy*, **52**, 750-63.
- SADIKOT, R. T., BLACKWELL, T. S., CHRISTMAN, J. W. & PRINCE, A. S. 2005. Pathogen–host interactions in *Pseudomonas aeruginosa* pneumonia. *American Journal of Respiratory And Critical Care Medicine*, **171**, 1209-1223.
- SAHA, S., THAVASI, R. & JAYALAKSHMI, S. 2008. Phenazine pigments from *Pseudomonas aeruginosa* and their application as antibacterial agent and food colourants. *Research Journal of Microbiology*, **3**, 122-128.
- SALIBA, A.-E., SANTOS, S. C. & VOGEL, J. 2017. New RNA-Seq approaches for the study of bacterial pathogens. *Current Opinion in Microbiology*, **35**, 78-87.
- SAMBROOK, J., FRITSCH, E. F. & MANIATIS, T. 1989. Molecular cloning, New York: Cold Spring Harbor Laboratory Press.
- SANGARI, F. J., CAYÓN, A. M., SEOANE, A. & GARCÍA-LOBO, J. M. 2010. *Brucella abortus ure2* region contains an acid-activated urea transporter and a nickel transport system. *BMC Microbiology*, **10**, 107.
- SANJAR, F., KARNA, S. R., CHEN, T., CHEN, P., ABERCROMBIE, J. J. & LEUNG, K. P. 2016. Whole-genome sequence of multidrug-resistant *Pseudomonas aeruginosa* strain BAMCPA07-48, isolated from a combat injury wound. *Genome Announcements*, **4**, e00547-16.
- SANTAJIT, S. & INDRAWATTANA, N. 2016. Mechanisms of antimicrobial resistance in ESKAPE pathogens. *BioMed research international*, 2016.
- SASS, A., MARCHBANK, A., TULLIS, E., LIPUMA, J. J. & MAHENTHIRALINGAM, E. 2011. Spontaneous and evolutionary changes in the antibiotic resistance of *Burkholderia cenocepacia* observed by global gene expression analysis. *BMC Genomics*, **12**, 373.
- SASSEVILLE, D. 2004. Hypersensitivity to preservatives. *Dermatologic therapy*, **17**, 251-263.
- SAUER, K. & CAMPER, A. K. 2001. Characterization of Phenotypic Changes in *Pseudomonas putida* in response to Surface-Associated Growth. *Journal of Bacteriology*, **183**, 6579-6589.
- SAUER, K., CAMPER, A. K., EHRlich, G. D., COSTERTON, J. W. & DAVIES, D. G. 2002. *Pseudomonas aeruginosa* displays multiple phenotypes during development as a biofilm. *Journal of Bacteriology*, **184**, 1140-1154.

- SCHELLENBERGER, J., QUE, R., FLEMING, R. M., THIELE, I., ORTH, J. D., FEIST, A. M., ZIELINSKI, D. C., BORDBAR, A., LEWIS, N. E. & RAHMANIAN, S. 2011. Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0. *Nature Protocols*, **6**, 1290-1307.
- SCHMIDT, B. J., EBRAHIM, A., METZ, T. O., ADKINS, J. N., PALSSON, B. Ø. & HYDUKE, D. R. 2013. GIM3E: condition-specific models of cellular metabolism developed from metabolomics and expression data. *Bioinformatics*, **29**, 2900-2908.
- SCHMIDT, K. D., TÜMMLER, B. & RÖMLING, U. 1996. Comparative genome mapping of *Pseudomonas aeruginosa* PAO with *P. aeruginosa* C, which belongs to a major clone in cystic fibrosis patients and aquatic habitats. *Journal of Bacteriology*, **178**, 85-93.
- SCHREIBER, K., KRIEGER, R., BENKERT, B., ESCHBACH, M., ARAI, H., SCHOBERT, M. & JAHN, D. 2007. The anaerobic regulatory network required for *Pseudomonas aeruginosa* nitrate respiration. *Journal of Bacteriology*, **189**, 4310-4314.
- SCHURCH, N. J., SCHOFIELD, P., GIERLIŃSKI, M., COLE, C., SHERSTNEV, A., SINGH, V., WROBEL, N., GHARBI, K., SIMPSON, G. G. & OWEN-HUGHES, T. 2016. How many biological replicates are needed in an RNA-Seq experiment and which differential expression tool should you use? *RNA*, **22**, 839-851.
- SCHWAN, W. R., WARRENER, P., KEUNZ, E., STOVER, C. K. & FOLGER, K. R. 2005. Mutations in the *cueA* gene encoding a copper homeostasis P-type ATPase reduce the pathogenicity of *Pseudomonas aeruginosa* in mice. *International Journal of Medical Microbiology*, **295**, 237-242.
- SCHWEIZER, H. P. 1998. Intrinsic resistance to inhibitors of fatty acid biosynthesis in *Pseudomonas aeruginosa* is due to efflux: application of a novel technique for generation of unmarked chromosomal mutations for the study of efflux systems. *Antimicrobial Agents and Chemotherapy*, **42**, 394-398.
- SCHWEIZER, H. P. 2001. Triclosan: a widely used biocide and its link to antibiotics. *FEMS Microbiology Letters*, **202**, 1-7.
- SCHWEIZER, H. P. 2003. Efflux as a mechanism of resistance to antimicrobials in *Pseudomonas aeruginosa* and related bacteria: unanswered questions. *Genetics and Molecular Research*, **2**, 48-62.
- SEEMANN, T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*, **30**, 2068-2069.
- SEILER, C. & BERENDONK, T. U. 2012. Heavy metal driven co-selection of antibiotic resistance in soil and water bodies impacted by agriculture and aquaculture. *Frontiers in Microbiology*, **3**.
- SEYEDMOHAMMAD, S., BORN, D. & VENTER, H. 2014. Expression, purification and functional reconstitution of FeoB, the ferrous iron transporter from *Pseudomonas aeruginosa*. *Protein Expression and Purification*, **101**, 138-145.
- SEYEDNASROLLAH, F., LAIHO, A. & ELO, L. L. 2013. Comparison of software packages for detecting differential expression in RNA-Seq studies. *Briefings in Bioinformatics*, **16**, 59-70.
- SHAHID, M. & MALIK, A. 2003. Multidrug-resistant *Pseudomonas aeruginosa* strains harbouring R-plasmids and AmpC β -lactamases isolated from hospitalised burn patients in a tertiary care hospital of North India. *FEMS Microbiology Letters*, **228**, 181-186.
- SHARMA, V., NORIEGA, C. E. & ROWE, J. J. 2006. Involvement of NarK1 and NarK2 proteins in transport of nitrate and nitrite in the denitrifying bacterium *Pseudomonas aeruginosa* PAO1. *Applied and Environmental Microbiology*, **72**, 695-701.
- SHEN, K., SAYEED, S., ANTALIS, P., GLADITZ, J., AHMED, A., DICE, B., JANTO, B., DOPICO, R., KEEFE, R. & HAYES, J. 2006. Extensive genomic plasticity in *Pseudomonas aeruginosa* revealed by identification and distribution studies of novel genes among clinical isolates. *Infection and Immunity*, **74**, 5272-5283.
- SHIN, S. C., AHN, D. H., KIM, S. J., LEE, H., OH, T.-J., LEE, J. E. & PARK, H. 2013. Advantages of single-molecule real-time sequencing in high-GC content genomes. *PLoS One*, **8**, e68824.
- SHLOMI, T., CABILI, M. N., HERRGÅRD, M. J., PALSSON, B. Ø. & RUPPIN, E. 2008. Network-based prediction of human tissue-specific metabolism. *Nature Biotechnology*, **26**, 1003-1010.
- SIEGERT, W. 2012. Microbiological quality management for the production of cosmetics and detergents. *SOFW Journal-Seifen Ole Fette Wachse*, **138**, 30.
- SIEGERT, W., GÜCKEL, A. & CARSTENS, S. 2011. A New Approach to Prevent Waterborne Contamination in Detergent Production. *SOFW Journal-Seifen Ole Fette Wachse*, **137**, 35.
- SIMM, R., MORR, M., KADER, A., NIMTZ, M. & RÖMLING, U. 2004. GGDEF and EAL domains inversely regulate cyclic di-GMP levels and transition from sessility to motility. *Molecular Microbiology*, **53**, 1123-1134.
- SINGH, S., SINGH, S. K., CHOWDHURY, I. & SINGH, R. 2017. Understanding the Mechanism of Bacterial Biofilms Resistance to Antimicrobial Agents. *The Open Microbiology Journal*, **11**, 53.
- SMART, R. & SPOONER, D. 1972. Microbiological spoilage in pharmaceuticals and cosmetics. *Journal of Cosmetic Science*, **23**, 721-737.
- SMITH, H. E. & BLAIR, J. M. 2013. Redundancy in the periplasmic adaptor proteins AcrA and AcrE provides resilience and an ability to export substrates of multidrug efflux. *Journal of Antimicrobial Chemotherapy*, **69**, 982-987.
- SONG, J., WARE, A. & LIU, S.-L. 2003. Wavelet to predict bacterial ori and ter: a tendency towards a physical balance. *BMC Genomics*, **4**, 17.
- SONG, R., HUA, Z., LI, H. & CHEN, J. 2006. Biodegradation of petroleum hydrocarbons by two *Pseudomonas aeruginosa* strains with different uptake modes. *Journal of Environmental Science and Health Part A*, **41**, 733-748.

- SOTO, S. M. 2013. Role of efflux pumps in the antibiotic resistance of bacteria embedded in a biofilm. *Virulence*, **4**, 223-229.
- SPENCER, D. H., KAS, A., SMITH, E. E., RAYMOND, C. K., SIMS, E. H., HASTINGS, M., BURNS, J. L., KAUL, R. & OLSON, M. V. 2003. Whole-genome sequence variation among multiple isolates of *Pseudomonas aeruginosa*. *Journal of Bacteriology*, **185**, 1316-1325.
- SPITZER, M., WILDENHAIN, J., RAPPILBER, J. & TYERS, M. 2014. BoxPlotR: a web tool for generation of box plots. *Nature Methods*, **11**, 121-122.
- SPOERING, A. L. & LEWIS, K. 2001. Biofilms and planktonic cells of *Pseudomonas aeruginosa* have similar resistance to killing by antimicrobials. *Journal of Bacteriology*, **183**, 6746-6751.
- SPOONER, R. & YILMAZ, Ö. 2012. Nucleoside-diphosphate-kinase: a pleiotropic effector in microbial colonization under interdisciplinary characterization. *Microbes and Infection*, **14**, 228-237.
- SRIKUMAR, R., KON, T., GOTOH, N. & POOLE, K. 1998. Expression of *Pseudomonas aeruginosa* Multidrug Efflux Pumps MexA-MexB-OprM and MexC-MexD-OprJ in a Multidrug-Sensitive *Escherichia coli* Strain. *Antimicrobial Agents and Chemotherapy*, **42**, 65-71.
- STANIER, R. Y., PALLERONI, N. J. & DOUDOROFF, M. 1966. The aerobic pseudomonads a taxonomic study. *Microbiology*, **43**, 159-271.
- STAVRI, M., PIDDOCK, L. J. & GIBBONS, S. 2006. Bacterial efflux pump inhibitors from natural sources. *Journal of Antimicrobial Chemotherapy*, **59**, 1247-1260.
- STEINBERG, D. C. 2008. Voluntary registration of cosmetics and 2007 frequency of preservative use. *Cosmetics and Toiletries*, **123**.
- STEPHENSON, J., HEARD, S., RICHARDS, M. & TABAQCHALI, S. 1984. Outbreak of septicaemia due to contaminated mouthwash. *British Medical Journal (Clinical Research ed.)*, **289**, 1584.
- STEWART, L., FORD, A., SANGAL, V., JEUKENS, J., BOYLE, B., KUKAVICA-IBRULJ, I., CAIM, S., CROSSMAN, L., HOSKISSON, P. A. & LEVESQUE, R. 2014. Draft genomes of 12 host-adapted and environmental isolates of *Pseudomonas aeruginosa* and their positions in the core genome phylogeny. *Pathogens and Disease*, **71**, 20-25.
- STEWART, P. S. & WILLIAM COSTERTON, J. 2001. Antibiotic resistance of bacteria in biofilms. *The Lancet*, **358**, 135-138.
- STIMSON, L., ROWLANDS, M. G., NEWBATT, Y. M., SMITH, N. F., RAYNAUD, F. I., ROGERS, P., BAVETSIAS, V., GORSUCH, S., JARMAN, M. & BANNISTER, A. 2005. Isothiazolones as inhibitors of PCAF and p300 histone acetyltransferase activity. *Molecular Cancer Therapeutics*, **4**, 1521-1532.
- STOCK, A. M., ROBINSON, V. L. & GOUDREAU, P. N. 2000. Two-component signal transduction. *Annual Review of Biochemistry*, **69**, 183-215.
- STOTHARD, P. & WISHART, D. S. 2005. Circular genome visualization and exploration using CGView. *Bioinformatics*, **21**, 537-539.
- STOVER, C., PHAM, X., ERWIN, A., MIZOGUCHI, S., WARRENER, P., HICKEY, M., BRINKMAN, F., HUFNAGLE, W., KOWALIK, D. & LAGROU, M. 2000. Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic pathogen. *Nature*, **406**, 959-964.
- STRATEVA, T. & YORDANOV, D. 2009. *Pseudomonas aeruginosa*- A Phenomenon of Bacterial Resistance. *Journal of Medical Microbiology*, **58**, 1133-1148.
- SU, L., JIA, W., HOU, C. & LEI, Y. 2011. Microbial biosensors: a review. *Biosensors and Bioelectronics*, **26**, 1788-1799.
- SUGIMURA, M., MASEDA, H., HANAKI, H. & NAKAE, T. 2008. Macrolide antibiotic-mediated downregulation of MexAB-OprM efflux pump expression in *Pseudomonas aeruginosa*. *Antimicrobial Agents and Chemotherapy*, **52**, 4141-4144.
- SUTTON, S. & JIMENEZ, L. 2012. MICROBIOLOGY-A Review of Reported Recalls Involving Microbiological Control 2004-2011 with Emphasis on FDA Considerations of "Objectionable Organisms". *American Pharmaceutical Review*, **15**, 42.
- SZKLARCZYK, D., MORRIS, J. H., COOK, H., KUHN, M., WYDER, S., SIMONOVIC, M., SANTOS, A., DONCHEVA, N. T., ROTH, A. & BORK, P. 2016. The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Research*, **45**, D362-D368.
- TADA, T., MIYOSHI-AKIYAMA, T., SHIMADA, K., SHIROMA, A., NAKANO, K., TERUYA, K., SATOU, K., HIRANO, T., SHIMOJIMA, M. & KIRIKAE, T. 2016. A Carbapenem-Resistant *Pseudomonas aeruginosa* Isolate Harboring Two Copies of blaIMP-34 Encoding a Metallo-β-Lactamase. *PloS One*, **11**, e0149385.
- TAKAMI, H., TANIGUCHI, T., ARAI, W., TAKEMOTO, K., MORIYA, Y. & GOTO, S. 2016. An automated system for evaluation of the potential functionome: MAPLE version 2.1.0. *DNA Research*, **23**, 467-475.
- TAKANASHI, M., SHIRAKI, M. & SAITO, T. 2009. Characterization of a novel 3-hydroxybutyrate dehydrogenase from *Ralstonia pickettii* T1. *Antonie van Leeuwenhoek*, **95**, 249-262.
- TAM, R. & SAIER, M. 1993. Structural, functional, and evolutionary relationships among extracellular solute-binding receptors of bacteria. *Microbiological Reviews*, **57**, 320-346.
- TAN, A. S. B., TUYSUZ, M. & OTUK, G. 2013. Investigation of preservative efficacy and microbiological content of some cosmetics found on the market. *Pakistan Journal of Pharmaceutical Sciences*, **26**, 153-157.
- TANG, M., SUN, J., SHIMIZU, K. & KADOTA, K. 2015. Evaluation of methods for differential expression analysis on multi-group RNA-Seq count data. *BMC Bioinformatics*, **16**, 360.

- TATUSOV, R. L., GALPERIN, M. Y., NATALE, D. A. & KOONIN, E. V. 2000. The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Research*, **28**, 33-36.
- TAYLOR, P. K., YEUNG, A. T. & HANCOCK, R. E. 2014. Antibiotic resistance in *Pseudomonas aeruginosa* biofilms: towards the development of novel anti-biofilm therapies. *Journal of Biotechnology*, **191**, 121-130.
- TAYLOR, P. W., HAMILTON-MILLER, J. M. & STAPLETON, P. D. 2005. Antimicrobial properties of green tea catechins. *Food Science and Technology Bulletin*, **2**, 71.
- THADEN, J. T., LORY, S. & GARDNER, T. S. 2010. Quorum-sensing regulation of a copper toxicity system in *Pseudomonas aeruginosa*. *Journal of Bacteriology*, **192**, 2557-2568.
- THAKER, M., SPANOGIANNOPOULOS, P. & WRIGHT, G. D. 2010. The tetracycline resistome. *Cellular and Molecular Life Sciences*, **67**, 419-431.
- THIELE, I. & PALSSON, B. Ø. 2010. A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nature Protocols*, **5**, 93-121.
- THOMAS, L. 2011. The Molecular basis for preservative resistance in *Burkholderia cepacia* complex bacteria. *PhD Thesis*, Cardiff University.
- TIELEN, P., STRATHMANN, M., JAEGER, K.-E., FLEMMING, H.-C. & WINGENDER, J. 2005. Alginate acetylation influences initial surface colonization by mucoid *Pseudomonas aeruginosa*. *Microbiological Research*, **160**, 165-176.
- TOLKER-NIELSEN, T., BRINCH, U. C., RAGAS, P. C., ANDERSEN, J. B., JACOBSEN, C. S. & MOLIN, S. 2000. Development and dynamics of *Pseudomonas* sp. biofilms. *Journal of Bacteriology*, **182**, 6482-6489.
- TOP, E. M. & SPRINGAEL, D. 2003. The role of mobile genetic elements in bacterial adaptation to xenobiotic organic compounds. *Current Opinion in Biotechnology*, **14**, 262-269.
- TRIPATHY, S., KUMAR, N., MOHANTY, S., SAMANTA, M., MANDAL, R. & MAITI, N. 2007. Characterisation of *Pseudomonas aeruginosa* isolated from freshwater culture systems. *Microbiological Research*, **162**, 391-396.
- TUMAH, H. 2009. Bacterial biocide resistance. *Journal of Chemotherapy*, **21**, 5-15.
- VALLET-GELY, I. & BOCCARD, F. 2013. Chromosomal organization and segregation in *Pseudomonas aeruginosa*. *PLoS Genetics*, **9**, e1003492.
- VALOT, B., ROHMER, L., JACOBS, M. A., MILLER, S. I., BERTRAND, X. & HOCQUET, D. 2014. Comparative genomic analysis of two multidrug-resistant clinical isolates of ST395 epidemic strain of *Pseudomonas aeruginosa* obtained 12 years apart. *Genome Announcements*, **2**, e00515-14.
- VAN ALST, N. E., PICARDO, K. F., IGLEWSKI, B. H. & HAIDARIS, C. G. 2007. Nitrate sensing and metabolism modulate motility, biofilm formation, and virulence in *Pseudomonas aeruginosa*. *Infection and Immunity*, **75**, 3780-3790.
- VAN ALST, N. E., SHERRILL, L. A., IGLEWSKI, B. H. & HAIDARIS, C. G. 2009. Compensatory periplasmic nitrate reductase activity supports anaerobic growth of *Pseudomonas aeruginosa* PAO1 in the absence of membrane nitrate reductase. *Canadian Journal of Microbiology*, **55**, 1133-1144.
- VAN DER MEER, J. R., DE VOS, W., HARAYAMA, S. & ZEHNDER, A. 1992. Molecular mechanisms of genetic adaptation to xenobiotic compounds. *Microbiological Reviews*, **56**, 677-694.
- VAN NIMWEGEN, E. 2006. Scaling laws in the functional content of genomes. *Power Laws, Scale-Free Networks and Genome Biology*. USA: Springer.
- VENTER, H., SHILLING, R. A., VELAMAKANNI, S., BALAKRISHNAN, L. & VAN VEEN, H. W. 2003. An ABC transporter with a secondary-active multidrug translocator domain. *Nature*, **426**, 866.
- VIDAL, F., MENSA, J., MARTINEZ, J., ALMELA, M., MARCO, F., GATELL, J., RICHART, C., SORIANO, E. & DE ANTA, M. J. 1999. *Pseudomonas aeruginosa* bacteremia in patients infected with human immunodeficiency virus type 1. *European Journal of Clinical Microbiology and Infectious Diseases*, **18**, 473-477.
- VIKRAM, A., BOMBERGER, J. M. & BIBBY, K. J. 2015. Efflux as a glutaraldehyde resistance mechanism in *Pseudomonas fluorescens* and *Pseudomonas aeruginosa* biofilms. *Antimicrobial Agents and Chemotherapy*, **59**, 3433-3440.
- VILLERET, V., BEEUMEN, J. V., CHESSA, J. P. & GERDAY, C. 1997. Preliminary crystal structure determination of the alkaline protease from the Antarctic psychrophile *Pseudomonas aeruginosa*. *Protein Science*, **6**, 2462-2464.
- VIMR, E. R. 2013. Unified theory of bacterial sialometabolism: how and why bacteria metabolize host sialic acids. *ISRN Microbiology*, 2013.
- VIMR, E. R., KALIVODA, K. A., DESZO, E. L. & STEENBERGEN, S. M. 2004. Diversity of microbial sialic acid metabolism. *Microbiology and Molecular Biology Reviews*, **68**, 132-153.
- VITAL-LOPEZ, F. G., REIFMAN, J. & WALLQVIST, A. 2015. Biofilm formation mechanisms of *Pseudomonas aeruginosa* predicted via genome-scale kinetic models of bacterial metabolism. *PLoS Computational Biology*, **11**, e1004452.
- VIVEK-ANANTH, R. & SAMAL, A. 2016. Advances in the integration of transcriptional regulatory information into genome-scale metabolic models. *Biosystems*, **147**, 1-10.

- WADE, D. S., CALFEE, M. W., ROCHA, E. R., LING, E. A., ENGSTROM, E., COLEMAN, J. P. & PESCI, E. C. 2005. Regulation of *Pseudomonas* quinolone signal synthesis in *Pseudomonas aeruginosa*. *Journal of Bacteriology*, **187**, 4372-4380.
- WAGENLEHNER, F. & NABER, K. 2006. Current challenges in the treatment of complicated urinary tract infections and prostatitis. *Clinical Microbiology and Infection*, **12**, 67-80.
- WAITE, R. D., PACCANARO, A., PAPAKONSTANTINOPOULOU, A., HURST, J. M., SAQI, M., LITTLER, E. & CURTIS, M. A. 2006. Clustering of *Pseudomonas aeruginosa* transcriptomes from planktonic cultures, developing and mature biofilms reveals distinct expression profiles. *BMC Genomics*, **7**, 162.
- WALTERS, M. C., ROE, F., BUGNICOURT, A., FRANKLIN, M. J. & STEWART, P. S. 2003. Contributions of antibiotic penetration, oxygen limitation, and low metabolic activity to tolerance of *Pseudomonas aeruginosa* biofilms to ciprofloxacin and tobramycin. *Antimicrobial Agents and Chemotherapy*, **47**, 317-323.
- WANG, C. L., MICHELS, P. C., DAWSON, S. C., KITASAKKUL, S., BAROSS, J. A., KEASLING, J. & CLARK, D. S. 1997. Cadmium removal by a new strain of *Pseudomonas aeruginosa* in aerobic culture. *Applied and Environmental Microbiology*, **63**, 4075-4078.
- WANG, W. & SHAO, Z. 2013. Enzymes and genes involved in aerobic alkane degradation. *Frontiers in Microbiology*, **4**.
- WANG, X. & QUINN, P. J. 2010. Lipopolysaccharide: Biosynthetic pathway and structure modification. *Progress in Lipid Research*, **49**, 97-107.
- WANG, Z., GERSTEIN, M. & SNYDER, M. 2009. RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics*, **10**, 57-63.
- WASI, S., TABREZ, S. & AHMAD, M. 2013. Use of *Pseudomonas* spp. for the bioremediation of environmental pollutants: a review. *Environmental Monitoring and Assessment*, **185**, 8147-8155.
- WATTAM, A. R., ABRAHAM, D., DALAY, O., DISZ, T. L., DRISCOLL, T., GABBARD, J. L., GILLESPIE, J. J., GOUGH, R., HIX, D. & KENYON, R. 2014. PATRIC, the bacterial bioinformatics database and analysis resource. *Nucleic Acids Research*, **42**, D581-D591.
- WATTAM, A. R., DAVIS, J. J., ASSAF, R., BOISVERT, S., BRETTIN, T., BUN, C., CONRAD, N., DIETRICH, E. M., DISZ, T. & GABBARD, J. L. 2017. Improvements to PATRIC, the all-bacterial Bioinformatics Database and Analysis Resource Center. *Nucleic Acids Research*, **45**, D535-D542.
- WEBB, J. S., THOMPSON, L. S., JAMES, S., CHARLTON, T., TOLKER-NIELSEN, T., KOCH, B., GIVSKOV, M. & KJELLEBERG, S. 2003. Cell death in *Pseudomonas aeruginosa* biofilm development. *Journal of Bacteriology*, **185**, 4585-4592.
- WEERDENBURG, E. M., ABDALLAH, A. M., RANGKUTI, F., EL GHANY, M. A., OTTO, T. D., ADROUB, S. A., MOLENAAR, D., UMMELS, R., TER VEEN, K. & VAN STEMPVOORT, G. 2015. Genome-wide transposon mutagenesis indicates that *Mycobacterium marinum* customizes its virulence mechanisms for survival and replication in different hosts. *Infection and Immunity*, **83**, 1778-1788.
- WEI, Q. & MA, L. Z. 2013. Biofilm matrix and its regulation in *Pseudomonas aeruginosa*. *International Journal of Molecular Sciences*, **14**, 20983-21005.
- WEISER, R. 2015. The resistance of *Pseudomonas aeruginosa* to preservatives used in industrial formulations. *PhD Thesis*, Cardiff University.
- WEITZ, H. J., RITCHIE, J. M., BAILEY, D. A., HORSBURGH, A. M., KILLHAM, K. & GLOVER, L. A. 2001. Construction of a modified mini-Tn5 *luxCDABE* transposon for the development of bacterial biosensors for ecotoxicity testing. *FEMS Microbiology Letters*, **197**, 159-165.
- WHITBY, J. & RAMPLING, A. 1972. *Pseudomonas aeruginosa* contamination in domestic and hospital environments. *The Lancet*, **299**, 15-17.
- WHITCHURCH, C. B., TOLKER-NIELSEN, T., RAGAS, P. C. & MATTICK, J. S. 2002. Extracellular DNA required for bacterial biofilm formation. *Science*, **295**, 1487-1487.
- WHITELEY, M., BANGERA, M. G., BUMGARNER, R. E., PARSEK, M. R., TEITZEL, G. M., LORY, S. & GREENBERG, E. 2001. Gene expression in *Pseudomonas aeruginosa* biofilms. *Nature*, **413**, 860-864.
- WHITELEY, M., LEE, K. M. & GREENBERG, E. 1999. Identification of genes controlled by quorum sensing in *Pseudomonas aeruginosa*. *Proceedings of the National Academy of Sciences*, **96**, 13904-13909.
- WHO. 2017. World Health Organisation (WHO) priority pathogens list for R&D of new antibiotics [Online]. Available: <http://www.who.int/mediacentre/news/releases/2017/bacteria-antibiotics-needed/en/> [Accessed 04 March 2017].
- WILLIAMS, T. M. 2007. The mechanism of action of isothiazolone biocides. *PowerPlant Chemistry*, **9**, 14.
- WILSON, S. A., WILLIAMS, R. J., PEARL, L. H. & DREW, R. E. 1995. Identification of Two New Genes in the *Pseudomonas aeruginosa* Amidase Operon, Encoding an ATPase (AmiB) and a Putative Integral Membrane Protein (AmiS). *Journal of Biological Chemistry*, **270**, 18818-18824.
- WILTON, M., CHARRON-MAZENOD, L., MOORE, R. & LEWENZA, S. 2016. Extracellular DNA acidifies biofilms and induces aminoglycoside resistance in *Pseudomonas aeruginosa*. *Antimicrobial Agents and Chemotherapy*, **60**, 544-553.
- WINDER, C., AL-ADHAM, I., ABDEL MALEK, S., BUULTJENS, T., HORROCKS, A. & COLLIER, P. 2000. Outer membrane protein shifts in biocide-resistant *Pseudomonas aeruginosa* PAO1. *Journal of Applied Microbiology*, **89**, 289-295.
- WINSON, M. K., SWIFT, S., HILL, P. J., SIMS, C. M., GRIESMAYR, G., BYCROFT, B. W., WILLIAMS, P. & STEWART, G. S. 1998. Engineering the *luxCDABE* genes from *Photobacterium luminescens* to provide a bioluminescent reporter for constitutive and promoter probe plasmids and mini-Tn5 constructs. *FEMS Microbiology Letters*, **163**, 193-202.

- WINSOR, G. L., GRIFFITHS, E. J., LO, R., DHILLON, B. K., SHAY, J. A. & BRINKMAN, F. S. 2016. Enhanced annotations and features for comparing thousands of *Pseudomonas* genomes in the *Pseudomonas* genome database. *Nucleic Acids Research*, **44**, D646-D653.
- WINSTANLEY, C., LANGILLE, M. G., FOTHERGILL, J. L., KUKAVICA-IBRULJ, I., PARADIS-BLEAU, C., SANSCHAGRIN, F., THOMSON, N. R., WINSOR, G. L., QUAIL, M. A. & LENNARD, N. 2009. Newly introduced genomic prophage islands are critical determinants of in vivo competitiveness in the Liverpool Epidemic Strain of *Pseudomonas aeruginosa*. *Genome Research*, **19**, 12-23.
- WITTE, W. 1998. Medical consequences of antibiotic use in agriculture. *Science*, **279**, 996-997.
- WONG, S., STREET, D. & DELGADO, S. I. 2000. Recalls of foods and cosmetics due to microbial contamination reported to the US Food and Drug Administration. *Journal of Food Protection*, **63**, 1113-1116.
- WRIGHT, G. D. 2005. Bacterial resistance to antibiotics: enzymatic degradation and modification. *Advanced Drug Delivery Reviews*, **57**, 1451-1470.
- WU, D.-Q., YE, J., OU, H.-Y., WEI, X., HUANG, X., HE, Y.-W. & XU, Y. 2011. Genomic analysis and temperature-dependent transcriptome profiles of the rhizosphere originating strain *Pseudomonas aeruginosa* M18. *BMC Genomics*, **12**, 438.
- XIONG, J., ALEXANDER, D. C., MA, J. H., DÉRASPE, M., LOW, D. E., JAMIESON, F. B. & ROY, P. H. 2013. Complete sequence of pOZ176, a 500-kilobase IncP-2 plasmid encoding IMP-9-mediated carbapenem resistance, from outbreak isolate *Pseudomonas aeruginosa* 96. *Antimicrobial Agents and Chemotherapy*, **57**, 3775-3782.
- XU, F., LIN, Q. & HOU, B. 2009. Synthesis and bioactivity of novel benzisothiazolone derivatives as potential microbicides. *Journal of Heterocyclic Chemistry*, **46**, 320-323.
- YAMAMOTO, K., HIRAO, K., OSHIMA, T., AIBA, H., UTSUMI, R. & ISHIHAMA, A. 2005. Functional characterization in vitro of all two-component signal transduction systems from *Escherichia coli*. *Journal of Biological Chemistry*, **280**, 1448-1456.
- YAN, S., APPLEBY, T., GUNIC, E., SHIM, J. H., TASU, T., KIM, H., RONG, F., CHEN, H., HAMATAKE, R. & WU, J. Z. 2007. Isothiazoles as active-site inhibitors of HCV NS5B polymerase. *Bioorganic & Medicinal Chemistry Letters*, **17**, 28-33.
- YAZDANKHAH, S., RUDI, K. & BERNHOFT, A. 2014. Zinc and copper in animal feed—development of resistance and co-resistance to antimicrobial agents in bacteria of animal origin. *Microbial Ecology in Health and Disease*, **25**, 25862.
- ZANKARI, E., HASMAN, H., COSENTINO, S., VESTERGAARD, M., RASMUSSEN, S., LUND, O., AARESTRUP, F. M. & LARSEN, M. V. 2012. Identification of acquired antimicrobial resistance genes. *Journal of Antimicrobial Chemotherapy*, **67**, 2640-2644.
- ZAPKA, C. A., CAMPBELL, E. J., MAXWELL, S. L., GERBA, C. P., DOLAN, M. J., ARBOGAST, J. W. & MACINGA, D. R. 2011. Bacterial hand contamination and transfer after use of contaminated bulk-soap-refillable dispensers. *Applied and Environmental Microbiology*, **77**, 2898-2904.
- ZEMKE, A. C. & BOMBERGER, J. M. 2016. Microbiology: social suicide for a good cause. *Current Biology*, **26**, R80-R82.
- ZGURSKAYA, H. I., LÓPEZ, C. A. & GNANAKARAN, S. 2015. Permeability barrier of Gram-negative cell envelopes and approaches to bypass it. *ACS Infectious Diseases*, **1**, 512-522.
- ZHANG, L. & MAH, T.-F. 2008. Involvement of a novel efflux system in biofilm-specific resistance to antibiotics. *Journal of Bacteriology*, **190**, 4447-4452.
- ZHANG, R. & LIN, Y. 2008. DEG 5.0, a database of essential genes in both prokaryotes and eukaryotes. *Nucleic Acids Research*, **37**, D455-D458.
- ZHANG, T., DING, Y., LI, T., WAN, Y., LI, W., CHEN, H. & ZHOU, R. 2012. A Fur-like protein PerR regulates two oxidative stress response related operons *dpr* and *metQIN* in *Streptococcus suis*. *BMC Microbiology*, **12**, 85.
- ZHANG, T., MURAIH, J. K., TISHBI, N., HERSKOWITZ, J., VICTOR, R. L., SILVERMAN, J., UWUMARENOGIE, S., TAYLOR, S. D., PALMER, M. & MINTZER, E. 2014. Cardiolipin prevents membrane translocation and permeabilization by daptomycin. *Journal of Biological Chemistry*, **289**, 11584-11591.
- ZHANG, Y.-M. & ROCK, C. O. 2004. Evaluation of epigallocatechin gallate and related plant polyphenols as inhibitors of the FabG and FabI reductases of bacterial type II fatty-acid synthase. *Journal of Biological Chemistry*, **279**, 30994-31001.
- ZHENG, Y., TSUJI, G., OPOKU-TEMENG, C. & SINTIM, H. O. 2016. Inhibition of *P. aeruginosa* c-di-GMP phosphodiesterase RocR and swarming motility by a benzoisothiazolinone derivative. *Chemical Science*, **7**, 6238-6244.
- ZHOU, G., SHI, Q.-S., HUANG, X.-M. & XIE, X.-B. 2015a. The three bacterial lines of defense against antimicrobial agents. *International Journal of Molecular Sciences*, **16**, 21711-21733.
- ZHOU, G., SHI, Q.-S., HUANG, X.-M. & XIE, X.-B. 2016. Comparison of transcriptomes of wild-type and isothiazolone-resistant *Pseudomonas aeruginosa* by using RNA-Seq. *Molecular Biology Reports*, **43**, 527-540.
- ZHOU, G., SHI, Q.-S., HUANG, X.-M., XIE, X.-B. & CHEN, Y.-B. 2015b. Insights into *Pseudomonas aeruginosa* ATCC9027 resistance to isothiazolones through proteomics. *Microbial Drug Resistance*, **21**, 140-148.
- ZHOU, Y., LIANG, Y., LYNCH, K. H., DENNIS, J. J. & WISHART, D. S. 2011. PHAST: a fast phage search tool. *Nucleic Acids Research*, **gkr485**.

REFERENCES

ZHULENKOV, D., RUDEVICA, Z., JAUDZEMS, K., TURKS, M. & LEONCHIKS, A. 2014. Discovery and structure–activity relationship studies of irreversible benzisothiazolinone-based inhibitors against *Staphylococcus aureus* sortase A transpeptidase. *Bioorganic & Medicinal Chemistry*, **22**, 5988-6003.