



COPYRIGHT AND USE OF THIS THESIS

This thesis must be used in accordance with the provisions of the Copyright Act 1968.

Reproduction of material protected by copyright may be an infringement of copyright and copyright owners may be entitled to take legal action against persons who infringe their copyright.

Section 51 (2) of the Copyright Act permits an authorized officer of a university library or archives to provide a copy (by communication or otherwise) of an unpublished thesis kept in the library or archives, to a person who satisfies the authorized officer that he or she requires the reproduction for the purposes of research or study.

The Copyright Act grants the creator of a work a number of moral rights, specifically the right of attribution, the right against false attribution and the right of integrity.

You may infringe the author's moral rights if you:

- fail to acknowledge the author of this thesis if you quote sections from the work
- attribute this thesis to another author
- subject this thesis to derogatory treatment which may prejudice the author's reputation

For further information contact the University's Copyright Service.

sydney.edu.au/copyright



INTERACTIONS OF INSERTION SEQUENCES TARGETING INTEGRON ASSOCIATED RECOMBINATION SITES

Mia Žerić

BSc. Molecular Biotechnology (Hons)

Faculty of Science

The University of Sydney, Australia.

A thesis submitted for fulfilment of the requirements of the degree of Doctor of
Philosophy

August 2015

STATEMENT OF CANDIDATE

This work is original and has not been submitted for a higher degree to any other university of institution. The work of others, when drawn upon, is referenced fully.

Signed

Mia Žerić

ACKNOWLEDGMENTS

I would like to thank my supervisor, Associate Professor Andrew J. Holmes, for his support throughout my candidature. Thank you for your time, your patients and your willingness to push me. You sir, are a vault of endless information, a human encyclopaedia! I walk away from my candidature knowing that I have learnt a lot from you about everything, and anything. Thanks Captain!

My appreciation also goes out to my associate supervisor, Dr. Nicholas V. Coleman. Your door was always open, you were always ready to give me tips, tricks and words of encouragement. I wish you the very best of luck with all your academic endeavours and hope to call you Professor Nick one day real soon. You deserve it Nick!

To past and present members of the Holmes and Coleman lab. Thanks to you ladies (and the occasional gentleman), I successfully put on a few kilograms during my candidature. You continued to feed me cake and tea, and it is fair to say that without it, my brain would have been deprived of glucose. I thank you all for your support and the endless cake supplies.

I would also like to thank Dr Katherine Phan for all the pushes and inspirational chats, for your time and willingness to help me dissolve any mental and writing blocks I was experiencing. Thanks Kitty! I owe you cake!

I would also like to thank our collaborators, Mitchell Brown, Professor Jon Iredell, Dr. Sally Partridge, Professor Hatch Stokes at CIID Westmead and the University of Technology Sydney for providing me with the clinical isolates. I also need to extend my thanks to Dr. Elena Martinez for helping me troubleshoot fosmid constructions.

To Associate Professor Jacquie Mathews for providing me with the necessary lab space to do my protein work. In particular I would like to thank Dr. Lorna Wilkinsons-White for being an amazing mentor. You have taught me all that I know about EMSA gels.

To Professor Richard Christopherson, for allowing me to work in his lab and to Dr. Trisha Almazri for teaching me how to perform MALDI-TOF Analysis. Thank you Trish for letting me fly the voyager!

To my family, my loving parents Mirsad and Edina Žerić for everything. You kept me alive, you never failed to provide, and you showed me love. To my sister, Aida Žerić, know that I love you and that I am immensely proud of you. Now go and conquer the world!

Finally I would like to thank Danya B. Rose. You have seen the ugly side of my candidate. You have seen the anxiety, you witnesses the ugly panic attacks. You never judged me for it, or called me weak. Your support, understanding and kind heart have helped me more than I can ever say.

ABBREVIATIONS

°C	Degrees Celsius
aa	amino acid
<i>aadB</i>	aminoglycoside adenytransferase B gene
Amp	Ampicillin
<i>attC</i>	attachment site cassette
<i>attI</i>	attachment site integron
ATCC	American Type Culture Collection
<i>bla</i>	Beta-lactamase gene
BGC	Bacterial Gene Cassette
BLAST	Basic Local Alignment Search Tool
bp	base pairs
β-	Beta-
β-Mecap	Beta-mecaptophenol
CI	Chromosomal Integron
Cm	Chloramphenicol
CTn	Compound Transposon
CTAB	Hexadecyltrimethylammonium bromide
DIG-6-dUTP	digoxigenin-6-dUTP
DNA	Deoxyribonucleic acid
dsDNA	double stranded deoxyribonucleic acid
DTT	Dithiothreitol
<i>E. coli</i>	<i>Escherichia coli</i>
EB	Elution Buffer
EDTA	Ethylenediaminetetraacetic acid
EMSA	Electrophoretic Mobility Shift Assay
EtOH	Ethanol
EtBr	Ethidium Bromide
Gc	Gene cassette
GI	Genomic Island
Gv	Genomovar
HGT	Horizontal gene transfer
ICE	Integrating Conjugative Element
Inc	Incompatibility
IntI	Integron integrase (protein)
<i>intI</i>	integron integrase (gene)
<i>intII</i>	Class 1 integron integrase gene
IR	Inverted Repeats
IS	Insertion Sequence
kb	kilobase
Km	Kanamycin
LB broth	Luria-Bertani broth
LGT	Lateral gene transfer
M	Molar
mM	millimolar
μM	micromolar
nM	nanomolar
pMol	picomolar
MBP	Maltose Binding protein
<i>MBP</i> Tpase	Maltose-Transposase fusion protein
MDR	Multidrug Resistance
MGE	Mobile Genetic Elements
MI	Mobile Integron

min	minute
MQ	MilliQ water
MRI	Multidrug Resistance Integron
NCBI	National Centre for Biotechnology Information
OD	Optical Density
orf	Open reading frame
Pc	Integron associated cassette promoter
PCI	<i>Pseudomonas</i> Chromosomal Integron
PCR	Polymerase chain reaction
ProK	Proteinase K
Ps.	<i>Pseudomonas</i>
Psmendo	<i>Pseudomonas mendocina</i>
PstQ	<i>Pseudomonas stutzeri</i> st <i>Q</i>
Pst587	<i>Pseudomonas stutzeri</i> ATCC17587
Pst641	<i>Pseudomonas stutzeri</i> ATCC17641
qacE	Quaternary ammonium resistance gene
res	Site-specific resolution site
RM	Relative mobility
SSBs	Single Stranded Binding Proteins
SSC	Standard Saline Citrate
ssDNA	single stranded deoxyribonucleic acid
SDS	Sodium dodecyl sulphate
st.	Strain
sulI	Sulfonamide resistance gene
TAE buffer	Tris Acetate EDTA buffer
TBE buffer	Tris Borate EDTA buffer
TE buffer	Tris EDTA buffer
TE(s)	Transposable element(s)
tni	Transposition gene
tnpR	Resolvase gene
tra	Conjugal transfer gene
Tn(s)	Transposon(s)
tpase	Transposase open reading frame
Tpase	Transposase protein
UV	Ultraviolet
VCR	<i>Vibrio cholerae</i> repeats
V	Volts
w/v	Weight/volume
v/v	Volume/volume
Δ	Deletion of
μg	microgram
μL	microliter
μm	micrometer

SYNOPSIS

In nature, microbes are constantly exposed to variable and stressful environments. Bacteria from the genus *Pseudomonas* are known for their ability to colonise multiple habitats and to rapidly adapt to new environments. Indeed, some members have crossed ecological barriers and established themselves in clinical settings. *Pseudomonas aeruginosa*, a problematic pathogen, is such an example. It bears intrinsic resistance to many drug classes; has the ability to acquire resistance through mobile genetic elements; has high resistance rates worldwide; and is frequently implicated in severe infections.

The ability of *Pseudomonas* to withstand antibiotic selective pressure in the clinical environment is primarily due to the presence of multidrug resistance integrons (MRIs). These genetic platforms carry genes encoding antibiotic resistance and can cross ecological and phylogenetic barriers through associations with transposons and conjugative plasmids. This interaction had resulted in an increased incidence of clinical class 1 integrons in *Pseudomonas aeruginosa*. This can ultimately result in the formation of complex mosaic elements and complex phenotypes.

However, studying the interaction between different mobile genetic elements and making predictions about the rate and dissemination of resistance genes is challenging. In the majority of cases, the source of these resistance genes is unknown. Furthermore, the biological processes governing the rate of resistance gene flow are unable to be quantified *in vivo*. Lastly, it is difficult to predict the success of resistance gene maintenance within a bacterial population. Thus, studying resistance gene flow over ecological and phylogenetic separations is very challenging.

Based on previous studies, it is known that gene flow occurs from bacteria in natural environments such as soil and water to clinical bacterial populations. Understanding the various intragenomic rearrangements and intergenomic transfers involved in this is especially relevant to the evolution of multidrug resistance. This study explores the IS1111-*attC* family of insertion sequences as a model for this process in Pseudomonads. IS elements, integrons, and gene cassettes were surveyed in *Pseudomonas* species from clinical, domestic and natural environments. The role of target site preferences and transposition mechanisms of the IS1111-*attC* elements in distribution patterns was investigated in models based on purified

transposase and synthetic nucleic acid molecules (*in vitro*) and using IS elements or target sites carried on various plasmid vectors (*in vivo*).

The IS1111-*attC* subgroup of insertion sequences was found to be overrepresented within non-clinical isolates of the genus *Pseudomonas* from either natural or domestic environments relative to other genera or to clinical *Pseudomonas* isolates. Within this group they were specifically found in strains that had chromosomal integrons. The pattern of distribution suggests that all instances of IS1111-*attC* elements in class 1 integrons represent recent invasion of *attC* sites that has occurred when a class 1 integron was present in the same cell as a chromosomal integron.

The IS1111-*attC* elements were shown to specifically recognize the *attC* recombination sites of integrons in binding assays and to specifically target the *attC* in mobility assays. Conditions affecting the rate of IS1111-*attC* movement between environmental and clinical *Pseudomonads* were also examined. The mechanism for IS1111-*attC* translocation was demonstrated to be site-specific recombination. Furthermore, the mode of IS1111-*attC* translocation was observed to echo the integron integrase in that the target site is a single-stranded form of the *attC* site. Significantly, the IS1111-*attC* transposase was found to preferentially bind the single strand forms of the top strand of the *attC* site, rather than the bottom strand *attC* site which has been shown to be the target of the integron integrase.

This is the first evidence for IS1111-*attC* mobility in *Pseudomonas* cells occurring via a similar mechanism to integron gene mobilization and illustrates the potential for these elements to both move between environmental (chromosomal) and clinical (plasmid borne) integrons as well as facilitate interactions between them.

TABLE OF CONTENTS

STATEMENT OF CANDIDATURE.....	i
ACKNOWLEDGEMENTS.....	ii
ABBREVIATIONS.....	iii
SYNOPSIS.....	v
TABLE OF CONTENTS.....	vii
LIST OF FIGURES.....	xi
LIST OF TABLES.....	xiv
Chapter 1 Introduction	1
1.1 Co-evolutionary arms race and antibiotic resistance: Red Queen hypothesis	1
1.2 Phylogenetic and ecological separation: old gene, new context	2
1.3 Genes and genome evolution	3
1.4 Resistance mechanisms	8
1.5 Overcoming ecological and phylogenetic separation: Horizontal gene transfer	9
1.6 Exploring the diversity of MGEs and complexity of the floating genome.....	13
1.6.1 Replicons-Plasmids	13
1.6.2 Transferrers-ICE.....	14
1.6.3 Translocators and accumulators-transposons, compound transposons, bacteriophages, insertion sequences, integrons	15
1.7 Site-specific recombination: a key process in multidrug resistance assembly and spread....	19
1.7.1 IntI Integrase, an unorthodox site-specific recombinase	22
1.8 A model system for the study of gene flow across ecological and phylogenetic barriers.....	23
1.8.1 Integron, the model system for gene flow between mobile and chromosomal contexts.....	24
1.8.2 CI-to-MRI flow: Analysis of the gene cassette metagenome	27
1.8.3 <i>attC</i> sites: handles and markers for gene flow.....	29
1.8.4 <i>Pseudomonas</i> : Model organism for gene flow	33
1.8.5 <i>IS1111-attC</i> elements target integron associated recombination sites	37
1.9 Two convergent enzymes, one recombination site.....	41
1.10 Aims	42
Chapter 2 Materials and Methods	43
2.1 General Material.....	43
2.1.1 Media.....	43
2.1.3 Solutions.....	44
2.1.4 Antibiotics	44
2.1.5 Equipment	49
2.2 Bacterial strains	49
2.2.1 Clinical Isolates	50
2.2.2 Isolation of <i>Pseudomonads</i> from domestic environments.....	50
2.3 High molecular weight genomic DNA extraction.....	53
2.3.1 Ammonium acetate purification of low molecular weight PCR amplicons	53
2.3.2 Crude DNA extraction via cell boil-lysing.....	53
2.4 Polymerase Chain Reaction.....	54
2.5 Agarose gel electrophoresis.....	54
2.6 Chemical transformation of bacterial cells.....	57
2.6.1 <i>E. coli</i> JM109 cells	57
2.6.2 Making <i>E. coli</i> Rosetta2 culture stocks for transformation experiments.....	58
2.6.3 Transformation of <i>E. coli</i> Rosetta2 with the pMAL and/or p _{MBP} ase plasmid.....	58
2.7 Electrical transformation of bacterial cells.....	58

2.7.1	Making <i>Ps. stutzeri</i> st. Q (<i>PstQ</i>) and <i>Ps. stutzeri</i> ATCC17641 (<i>Pst641</i>) culture stocks for electroporation experiments.....	58
2.7.2	Electro-transformation of <i>PstQ</i> and <i>Pst641</i>	58
2.8	Plasmid miniprep-Alkaline lysis	59
2.9	Isolation of genomic DNA for Fosmid library construction	61
2.9.1	Phenol: chloroform: isoamyl extraction of genomic DNA	61
2.9.2	Fosmid library construction.....	62
2.9.3	Fosmid screening via Southern hybridization	63
2.9.4	Fosmid purification	63
2.10	Fosmid sequencing.....	63
2.11	Bioinformatics analysis.....	64

Chapter 3 Distributions of IS1111-attC, mobile and chromosomal integrons within the bacterial genus *Pseudomonas*65

3.1	Introduction	65
3.1.1	Aims	67
3.2	Methods.....	68
3.2.1	Strain Collection.....	68
3.2.2	Genotyping domestic environmental <i>Pseudomonas</i> strains	68
3.2.3	Southern Hybridisation.....	69
3.2.4	Southern hybridization thresholds of detection	71
3.2.5	Polymerase chain reaction of gene fragments for sequencing and sequence analysis ..	72
3.3	Results	73
3.3.1	The non-clinical, environmental isolate set is composed of diverse members of the genus <i>Pseudomonas</i>	73
3.3.2	Distribution of IS1111-attC elements, MRIs and PCIs across three phylogenetic groups.....	74
3.3.3	IS1111-attCs are overrepresented in MRIs in clinical, and in PCIs in non-clinical <i>Pseudomonads</i> respectively.....	76
3.3.4	PCR detection and sequence analysis of IS1111-attC elements	77
3.3.5	Recovery of class 1 <i>intI1-attI1</i> junctions from all three sample sets.....	80
3.3.6	Phylogenetic analysis of the recovered <i>intI1-attI1</i> sequences reveals different MRI populations	84
3.3.7	Distinct IS1111-attC groups are associated with different <i>intI1</i> subgroups.....	85
3.3.8	Fine mapping of IS1111-attC genetic contexts	87
3.3.9	Identifying individual fosmid clones that are positive for IS1111-attC elements.....	88
3.3.10	Annotations of assembled contigs	92
3.3.11	Genome mapping of contigs confirms the absence of IS elements in fosmid clones....	92
3.4	Discussion	94
3.4.1	IS1111-attC elements are over-represented in <i>Pseudomonadaceae</i>	94
3.4.2	Identifying IS1111-attC source-sink relationships.....	96
3.4.3	Class 1 integron subgroups have different exposure to PCI gene cassette reservoirs ...	98
3.4.4	A proposed model for IS1111-attC spread across barriers.....	99
3.4.5	Further work.....	101

Chapter 4 Examining the *in vivo* translocation activity of IS1111-attC elements and mechanisms of recombination.....102

4.1	Introduction	102
4.2	Aims	105
4.3	Methods.....	106
4.3.1	Construction of “attC-traps”.....	106
4.3.2	Natural transformation of <i>Pst587</i>	109
4.3.3	Screening natural transformants for IS1111-attC translocation	110
4.3.4	PCI invasion by plasmid-borne ISPst6 in <i>PstQ</i> and <i>Pst641</i> via electroporation.....	111
4.4	Results	112

4.4.1	<i>Ps. stutzeri</i> ATCC 17587 is naturally transformable with double stranded plasmid containing an <i>attC</i> site.....	112
4.4.2	Restriction mapping detects mixed plasmid populations	115
4.4.3	Sequencing results confirm that <i>ISPst6</i> inserted into <i>Pseudomonas</i> -type <i>attC</i> sites... ..	117
4.4.4	<i>IS1111-attC</i> invasion of chromosomal <i>attC</i> array.....	119
4.4.5	Introduced IS elements are able to undergo initial stages of mobilization to form minicircles.....	120
4.4.6	Insertion into PCI <i>attC</i> sites was not detectable.....	122
4.5	Discussion	124
4.5.1	<i>IS1111-attC</i> elements translocate with regional-specificity into <i>attC</i> sites.....	124
4.5.2	Orientation-specific translocations is a feature of <i>ISPst6</i> elements	125
4.5.3	To jump or not to jump? Potential routes for single stranded <i>attC</i> site generation	126
4.5.4	Does <i>IS1111-attC</i> undergo a non- replicative (cut-and-paste) or a replicative (copy-and-paste) translocation mechanism?.....	128
4.5.5	Further work.....	129
Chapter 5 Characterization of recombination site-preferences of <i>Ps. stutzeri</i> <i>IS1111-attC</i> transposase.....		130
5.1	Introduction	130
5.2	Aims	133
5.3	Methods.....	134
5.3.1	PCR, restriction digestion and cloning of <i>IS1111-attC</i> tase into pMAL-c2X.....	134
5.3.2	Overexpression of recombinant proteins.....	134
5.3.3	Purification and identification of _{MBP} Tpase fusion protein.....	137
5.3.4	Electrophoretic Mobility Shift Assays (EMSA).....	138
5.4	Results	143
5.4.1	<i>in silico</i> analysis of <i>IS1111-attC</i> Transposase.....	143
5.4.2	Expression and purification of _{MBP} Tpase and MBP control	144
5.4.3	Mass Peptide fingerprint of MBP-Fusion protein	145
5.4.4	DNA Binding assays	149
5.4.5	EMSA competition assays.....	151
5.4.6	Testing the <i>IS1111-attC</i> minicircle junction as a binding partner for Tpase	157
5.5	Discussion	162
5.5.1	<i>IS1111-attC</i> transposase does not bind double stranded <i>attC</i> sites at a detectable level.....	162
5.5.2	<i>IS1111-attC</i> transposase binds strand specifically to the top strand <i>Pseudomonas attC</i> sites.....	163
5.5.3	Experimental evidence for <i>IS1111-attC</i> transposase interactions with classical <i>attC</i> sites is inconclusive.....	165
5.5.4	Factors impacting on EMSA sensitivity.....	166
5.5.5	Secondary binding partner for <i>IS1111-attC</i> translocation: the <i>IS1111-attC</i> minicircle junction.....	167
5.5.6	Further Work.....	171
Chapter 6 Conclusions		172
6.1	Overview	172
6.2	<i>IS1111-attC</i> : an epidemiological marker element for ecological barrier crossing in <i>Pseudomonadaceae</i>	173
6.3	Are there molecular-genetic barriers to <i>IS1111-attC</i> translocation between PCI and MRI contexts?.....	176
6.3.1	Uptake method impacts on <i>IS1111-attC</i> acquisition: natural transformation vs electroporation.....	177
6.3.2	The context of the <i>attC</i> site impacts on <i>IS1111-attC</i> translocation: topology specificity.....	178
6.3.3	<i>IS1111-attC</i> and micro-variations: a consequence of polymerase activity?.....	179

6.3.4	<i>IS1111-attC</i> may preferentially recognise <i>Pseudomonas</i> -type <i>attC</i> over <i>attC_{aadB}</i> sites.....	180
6.3.5	The importance of DNA structures and presentation context in the HGT landscape..	181
6.5	Convergent evolution, selective advantages of a “gene handle”.....	183
References.....		184
APPENDICES		209
APPENDIX A1: Characterization of isolates used in this study.....		220
APPENDIX A2: Exploring <i>IS1111-attC</i> translocation potential.....		280
APPENDIX A3: MBP Control.....		288

LIST OF FIGURES

CHAPTER 1	1
Figure 1.1 An adapted schematic representation of comparisons of strains from different genera and the same species	4
Figure 1.2 Size of shared extended genome and total genome is dependent on the number of isolates examined during comparative genome analysis	6
Figure 1.3 Intrinsic and acquired resistance mechanisms in <i>Pseudomonas aeruginosa</i>	9
Figure 1.4 Modes of horizontal gene transfer and their relationships to spatial proximity	11
Figure 1.5 Mobile genetic elements are diverse yet share similar features that may include transposase genes, inverted repeats (IR), resistance gene and <i>attC</i> recombination sites	13
Figure 1.6 General structure of integron-gene cassette system	17
Figure 1.7 XerC/D, Cre- <i>loxP</i> and IntI1 mediated recombination all require a precise cross-over of target DNA sites	21
Figure 1.8 Schematic illustration of the structure of <i>attI</i> and <i>attC_{aadB}</i> sites	23
Figure 1.9 Resistance markers vary between integrons and in most cases are flanked by the 5'- and 3'- conserved region (CS).....	25
Figure 1.10 Phylogenetic tree of known integron integrases (IntI)	28
Figure 1.11 Schematic representation of secondary structures of single stranded <i>attC</i> sites	31
Figure 1.12 Chromosomal integrons- reservoirs for resistance cassettes	32
Figure 1.13 Distribution of chromosomal integrons across the <i>Pseudomonas</i> lineage	36
Figure 1.14 Sharing ecological niches is postulated to influence HGT between Pseudomonads and enteric organisms	37
Figure 1.15 Gene flow between related organisms in comparison to that between non-related organisms	38
Figure 1.16 Schematic representation of ISP <i>st6</i> , a member of the IS1111- <i>attC</i> subgroup	39
Figure 1.17 Proposed model for IS1111- <i>attC</i> flow from an environmental to a clinical Pseudomonad.....	40
CHAPTER 2	43
Figure 2.1 Workflow for the isolation of Pseudomonads from environmental samples	52
Figure 2.2 Flowchart of gDNA extraction protocol from isolates for Fosmid constructions	62
Figure 2.3 Screening fosmids libraries for IS1111- <i>attC</i> elements.....	64
CHAPTER 3	65
Figure 3.1 Conceptual framework for the study using IS1111- <i>attC</i> as a model for gene flow across biological distances.....	66
Figure 3.2 Probes and their respective target sites used in southern hybridization experiments	70
Figure 3.3 IS1111- <i>attC</i> PCR screen across a subset of environmental and clinical <i>Pseudomonas</i> isolates	78
Figure 3.4 Dendrogram showing the relationship of the IS1111- <i>attC</i> group of elements to other members of the IS1111 family.....	81
Figure 3.5 <i>intI1-attI1</i> and <i>intI1</i> internal screen in a subset of 12 clinical Pseudomonads	82
Figure 3.6 <i>intI1-attI1</i> PCR screen of domestic and hospital environmental <i>Pseudomonas</i> isolates	83
Figure 3.7 Sub-tree of IS1111- <i>attC</i> tree from figure 3.6 with <i>intI1-attI1</i> groups overlaid.....	87
Figure 3.8 Identifying individual fosmid clones containing an IS1111- <i>attC</i> across five fosmid libraries	90
Figure 3.9 SNP tree of class 1 integrons.....	99
Figure 3.10 Postulated flow of IS1111- <i>attC</i> elements between environmental and clinical stains	100

CHAPTER 4	102
Figure 4.1 Diversity of <i>attC</i> sites in which IS1111- <i>attC</i> have been observed.....	103
Figure 4.2 Conceptual models copy-&-paste and cut-&-paste for IS1111- <i>attC</i> translocation between <i>attC</i> sites	105
Figure 4.3A Construction of <i>Pseudomonas-attC</i> traps	107
Figure 4.3B Construction of classical <i>attC_{aadB}</i> traps.....	108
Figure 4.4 Natural transformation of <i>Pst587</i> with <i>pattC</i> traps.....	109
Figure 4.5 Theoretical experimental design and recovery of IS1111- <i>attC</i> element translocation a using synthetic trap	110
Figure 4.6 Introduction of pPsattC1_IS into Pseudomonads via electroporation.....	111
Figure 4.7 Definitive assay screen for IS1111- <i>attC</i> translocation into <i>attC</i> -traps.....	114
Figure 4.8 <i>PvuII</i> predicted digests	115
Figure 4.9 1% agarose and 0.75% agarose topological analysis	116
Figure 4.10 Sequencing pPsattC1-IS confirmed the presence of ISPst6 within the <i>Pseudomonas</i> -type <i>attC</i> site	118
Figure 4.11 Translated amino acid sequence of the predicted and recovered <i>tpase</i> gene	117
Figure 4.12 <i>PstQ</i> and <i>Pst641</i> transformants contain ISPst6 minicircles	121
Figure 4.13 Minicircles from <i>PstQ</i> transformant 2 and <i>Pst641</i> transformant 9 recovered minicircles align with <i>attC</i> embedded ISPst6 from pPsattC1-IS.....	122
Figure 4.14 ISPst6 does not appear to have invaded the PstQ chromosomal arrays	123
Figure 4.15 Linear map of pBBR1MCS2 and orientations of cloned recombination sites	125
CHAPTER 5	130
Figure 5.1 IntI1-mediated recombination of <i>dsattI</i> and <i>ssattC_{bs}</i> compared to proposed IS1111- <i>attC</i> transposase mediated recombination of dsIS1111- <i>attC</i> minicircle junction and <i>ssattC_{ts}</i> or <i>ssattC_{bs}</i>	131
Figure 5.2 Comparison between <i>attC_{aadB}</i> and <i>Pseudomonas</i> -type (BGC101) <i>attC</i>	132
Figure 5.3 Expression construct used to produce a Maltose-binding transposase fusion protein.....	135
Figure 5.4 Flow diagram of the protein preparation strategy	136
Figure 5.5 Schematic representation of an EMSA competition assay.....	141
Figure 5.6 SDS-PAGE analysis reveals poor overexpression of _{MBP} Tpase fusion protein in Rosetta cells induced with 0.4 mM and 1 mM at 37°C and 25°C.....	144
Figure 5.7 SDS-PAGE of MBP control and _{MBP} Tpase fusion proteins, expressed at 16°C for 24 hours.....	145
Figure 5.8 SDS-PAGE analysis of purified _{MBP} Tpase fusion protein.....	146
Figure 5.9 Amino acid sequence of the _{MBP} Tpase fusion protein showing the major expected peptides and their relative mass: charge ratio.....	147
Figure 5.10 MALDI-TOF spectra of excised bands	148
Figure 5.11 EMSA of _{MBP} Tpase titrated against MBP- only control.....	150
Figure 5.12 EMSA of _{MBP} Tpase titrated against MBP control using ssDNA as binding substrates	152
Figure 5.13 Cold competition assay of single stranded top and bottom <i>Pseudomonas</i> -type <i>attC</i> site.....	154
Figure 5.14 Cold competition assay of single stranded top and bottom classical <i>attC_{aadB}</i> site.....	156
Figure 5.15 Schematic representation of the IS1111- <i>attC</i> minicircle and oligonucleotides used for _{MBP} Tpase DNA binding assays	157
Figure 5.16 Cold competition assay of double stranded IS minicircle junction DNA using 40 μM and 100 μM _{MBP} Tpase and MBP control	159
Figure 5.17 Cold competition assay of single stranded IS minicircle junction DNA using 40μM of _{MBP} Tpase and MBP-only control	161
Figure 5.18 A proposed model for IS1111- <i>attC</i> transposase mediated recombination with <i>ssattC</i> sites and double stranded IS junctions	170

CHAPTER 6172

Figure 6.1 Ecological and phylogenetic barriers impact on IS1111-attC distribution across
Proteobacteria populations.....174

Figure 6.2 Overcoming molecular-genetic barriers in a *Pseudomonas stutzeri* host
leads to successful IS1111-attC translocation into attC sites177

LIST OF TABLES

CHAPTER 1	1
Table 1.1 A summary of the mechanistic capacities of MGEs	19
Table 1.2 MGE and their independent evolution-enzymes and mechanisms	20
Table 1.3 Source and sink relationships of resistance genes	34
 CHAPTER 2	 43
Table 2.1 Media composition	43
Table 2.2 Solutions	44
Table 2.3 Antibiotics	45
Table 2.4 Equipment list and respective supplier	49
Table 2.5 Strains as reference controls for integron and insertion sequence assays.....	51
Table 2.6 PCR reagents and concentration	54
Table 2.7 Primers	55
Table 2.8 Details of PCR thermocycling protocols	56
Table 2.9 <i>E. coli</i> expression systems	57
Table 2.10 Plasmids used/generated	60
 CHAPTER 3	 65
Table 3.1 Strain collection	68
Table 3.2 Control strains.....	71
Table 3.3 Environmental isolates recovered on PIA from mop, sponge and swab samples.....	73
Table 3.4 Distribution of MRIs, PCI, and IS1111-attC across Pseudomonads based on Southern hybridization and PCR findings.....	75
Table 3.5 Observable SNPs in recovered <i>intI1</i> sequences based on pR388 class 1 integron In3 as the type sequence.....	84
Table 3.6 Summary of fosmid libraries	91
Table 3.7 <i>de novo</i> assembly summary	92
Table 3.8 Synteny of fosmid contigs to best reference genomes.....	93
Table 3.9 Examples of IS1111-attC elements in enteric bacteria found from <i>in silico</i> search...	97
 CHAPTER 4	 102
Table 4.1 Summary of the correlations between patterns of IS1111-attC occurrence and aspects of target sites, context and cell host.....	102
Table 4.2 Transformation efficiency of <i>Pst587</i> cells with synthetic <i>attC</i> -trap.....	112
Table 4.3 Recovery efficiency of IS1111-attC translocation into <i>pattC</i> traps in <i>Pst587</i>	113
Table 4.4 PCI characteristics of <i>Pst405</i> , <i>Pst641</i> and <i>PstQ</i>	119
 CHAPTER 5	 130
Table 5.1 Tested protein concentrations for corresponding oligonucleotide.....	139
Table 5.2 DNA structures tested in EMSA experiments	140
Table 5.3 Non-specific competitor EGG-m2F/R.....	142
Table 5.4 <i>in silico</i> predictions of IS1111-attC transposase and IntI1 using ProtPram ExpASy	143
Table 5.5 Summary of EMSA experiments.....	164

Chapter 1 Introduction

1.1 Co-evolutionary arms race and antibiotic resistance: Red Queen hypothesis

Environments are dynamic systems which change physically, chemically and biologically. For organisms to survive in ever changing environments requires them to adapt or die! Over multigenerational periods, evolutionary change in genomes is a major factor for survival. The evolution of antibiotic resistance represents a major challenge for the 21st century and can be thought of as an arms race between bacteria and humans. Such arms races have been recognised in evolutionary biology in the Red Queen hypothesis.

VanValen first formulated the Red Queen hypothesis, which proposed that continuing adaptation is necessary in order for a species to maintain its relative fitness within a co-evolving system (VanValen, 1973). In microbial ecosystems such as soil, bacteria and fungi produce antibiotics and other metabolites, as a means of competing for space and nutrients (Kümmerer, 2004). Microbes in such ecosystems are in an arms race to survive by developing competing mechanisms, which confer baseline resistance to their own antibiotics, and those released by their co-inhabitants. The nature of bacterial exposure to antibiotics has been changed by human activity through the extensive use and discharge of antibiotics into varying environments. While antibiotic resistance in natural ecosystems is a part of bacterial biology, antibiotic use and release by humans has exacerbated the resistance problem (Andersson & Hughes, 2014).

Since the discovery and industrialization of antibiotics, their exploitation and intensive use in clinics and agriculture has altered the nature of the arms race between antibiotic production and resistance. Antibiotics, with the aid of humans, are now moving between habitats and at higher concentrations than ever occurred naturally. This has dramatically changed the nature of the selective pressures for evolution of resistance; bacteria must adapt rapidly and co-evolve resistance to a range of antibiotic classes, leading to the emergence of multi-drug resistance (MDR) strains. The movement of genes between habitats by virtue of a series of specialized elements is now recognised as an important contributor to the emergence of MDR. For many organisms, increased incidence of MDR strains can largely be attributed to horizontal gene transfer (HGT) of resistance genes and subsequent formation of complex mosaic elements. This is particularly true of organisms that contribute to opportunistic

infections. Thus, excess amounts of antibiotic use has altered the bacterial response strategy by favouring HGT and the formation of MDR phenotypes (Andersson & Hughes, 2014).

Thus, when considering MDR one cannot just focus on a single clinical pathogen in the clinical system alone, but instead needs to consider numerous factors such as i) the ecological barriers the pathogen must overcome in order to establish in a clinical setting; ii) the intrinsic resistance mechanisms of the pathogen; iii) the origin of resistance gene; iv) the rate and success of resistance transfer across phylogenetic barriers; v) the mode of resistance transfer; vi) the maintenance of resistance phenotype in a population, as well as; vii) MDR organism spread from clinical setting into world-wide ecosystem (González-Zorn & Escudero, 2012). These observations highlight that MDR is a multifactorial problem involving many different species, genetic elements and habitats. It cannot be understood, or managed, by focussing on a few problem pathogens in the clinical context. A deeper appreciation and understanding of some of the rules governing the biological process of antibiotic resistance assembly and formation is required.

1.2 Phylogenetic and ecological separation: old gene, new context

Biological distance can be thought of as having two broad dimensions, phylogenetic (genetic) and ecological, both of which play key roles in MDR formation and dissemination. Phylogenetic separation can be defined as the separation of organisms based on evolutionary distance. This distance encompasses biochemical and molecular features within a bacterial species that have evolved over time. More importantly, bacteria with similar/identical genetic makeup are more likely to occupy the same ecological niches (Whiteley *et al.*, 2001), and are more likely to exchange genetic material (Andam & Gogarten, 2011). Ecologically disparate species are species inhabiting distinct niches that are spatially separated and have differing physiochemical conditions, such as pH, nutrients, oxygen availability etc. The ability of a bacterial family such as *Pseudomonadaceae*, to cross ecological barriers and establish in varying ecological niches can be related back to its pangenome complexity (Silby *et al.*, 2011).

Many antibiotic resistance genes are believed to have evolved within communities in non-human environments such as soil and freshwater. A diverse collection of genes encoding resistance to antibiotics such as β -lactams, tetracyclines and glycopeptides, predate the

industrialization of antibiotics (D'Costa *et al.*, 2011). The basic enzymology for such resistance exists in organisms that produce antibiotics and in the pathways for synthesis of antibiotic targets. For example, peptidoglycan modifying DD-peptidases (a penicillin-binding protein) are believed to be the ancestor protein to serine β -lactamases (Fisher *et al.*, 2005; Massova & Mobashery, 1998). Therefore, antibiotic selective pressures are inferred to have impacted on ancestral enzymes resulting in optimization of enzymatic functions such that they degrade antimicrobial agents. Additionally, antibiotic selective pressures have impacted on the transfer of these enzymes across biological distance and new context (Fisher *et al.*, 2005; Massova & Mobashery, 1998). In this model, natural selection has evolved a resistance function before the resistance gene was able to cross into the clinical setting.

When examining resistance elements and plasmids in clinical pathogens, it appears that specific, highly advantageous genes have been selected that result in beneficial resistance phenotypes. Given the specialized nature of resistance genes in pathogens, and the short evolutionary time scale since the introduction of antimicrobials for therapeutic use, it is widely postulated that antibiotic producing bacteria might be reservoirs of efficient resistance genes (Davies & Davies, 2010). Irrespective of the model for origins of resistance genes there are a series of key points to note. Modern resistance genes are almost certainly derived from multiple phylogenetically disparate organisms (Fournier *et al.*, 2006). These organisms did not, and do not, typically occupy the animal-human context, or cause clinical disease. Therefore, for a resistance gene to establish in a clinical context, it has to overcome issues of ecological and phylogenetic separations.

1.3 Genes and genome evolution

Bacterial reproduction, typically performed by binary fission, is a process wherein the cell grows to twice its size, and then splits in two. Whilst binary fission appears as a simple process, it must occur at the correct location in the cell in order to ensure that each progeny daughter cell receives a complete complement of genes with high fidelity (Angert, 2005). Given that bacterial reproduction is asexual, the relationship between the species concept and genome structure is often difficult to define. Genome evolution, via mutation, rearrangement or acquisition of new genetic material (Hacker & Carniel, 2001), is thus key to bacterial survival in an ever changing habitat. Similarities and/or differences between genomes are a reflection of the organisms' life history. Genome evolution can thus be used for species

classification, where members of the same bacterial species show genome relatedness. This relationship between genome composition of bacterial species and ecology provides the conceptual framework for understanding the evolution of multiple antibiotic resistance.

Mass genome sequencing has altered the way we classify and compare bacterial species and has highlighted the fact that bacterial genomes do not evolve, or change, at a constant rate over the whole chromosome (Mira *et al.*, 2010). Comparative genomics has repeatedly shown that where a bacterial species is defined by a set of phenotypic and ecological (habitat restrictions and preferences) features, then members of that species also share a significant proportion of their genomes. The concept of pangenomes can arbitrarily be defined as the total gene repertoire of a bacterial species (Medini *et al.*, 2005). Genome comparisons of different species has shown that genome similarity of members of the same “pheno-species” is higher than if members of different species are compared (Figure 1.1).

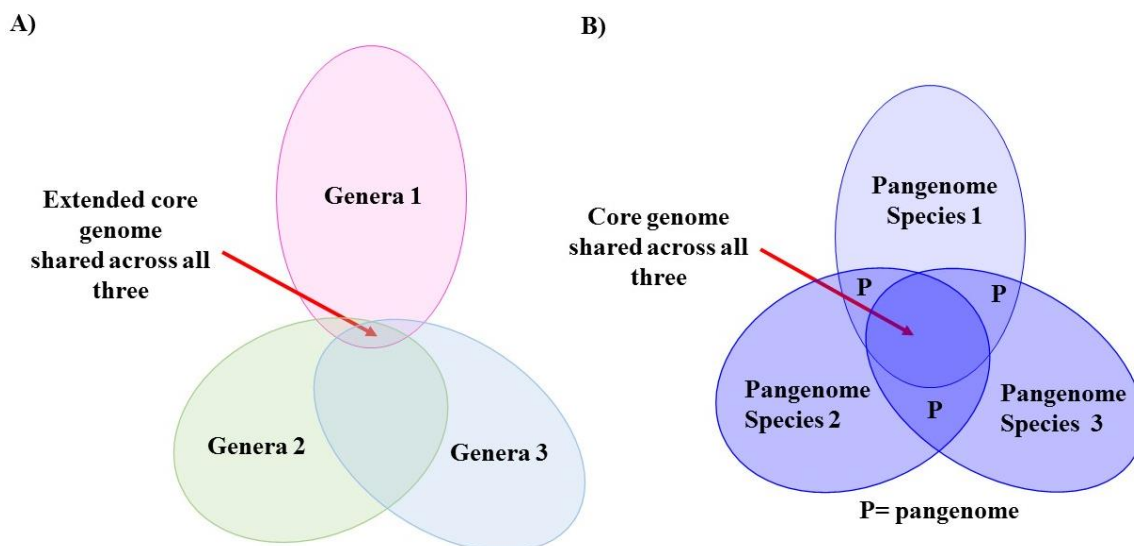


Figure 1.1: An adapted schematic representation of comparisons of strains from A) different genera and B) the same species.

Venn diagrams reveal a core set of genetic information that is shared by all members between and within a species. *P* represents pangenome overlap between two species.

Genes can form stable, long-term associations with chromosomes. This association is an ongoing process whereby organisms have assimilated stable gene sets with adaptive advantages for their given environments. The following key concepts arise from comparative genome analysis (Figure 1.1):

- 1) **The extended (ancient) core:** Some genes are always present. Chromosomal regions shared amongst all bacterial genera are defined as the extended core genome. When comparing more distantly related organisms, a smaller set of genetic information

“core genes” is still shared. It is estimated that the extended core genome contains in the order of 250 gene families, present in all bacterial genomes (Lapierre & Gogarten, 2009).

- 2) ***The core genome:*** Some genes define the unifying features of a species (core genome), others represent the total potential that may occur within a species (pangenome). The core genome is vertically inherited, and contains all the necessary genomic information for metabolic activity, housekeeping genes, cell envelope formation, binding proteins and phenotypic traits (Mira *et al.*, 2010; Wolfgang *et al.*, 2003). The pangenome includes all genes, both vertically and horizontally acquired. Both aspects of genome complexity provides insight into how tightly defined the ecological niche of a bacterial genus/species is and of how “old” the species is. Accurate description of the core and pangenome within a species is influenced by the number of strains examined. Comparative genome analysis of 26 Streptococcal strains has revealed the core genome plateaus at *ca* 600 genes (Lefebure & Stanhope, 2007). This is interesting given that each strain contains between 1697-2376 coding genes and that the *Streptococcus* pangenome was postulated to contain over 6000 genes (Lefebure & Stanhope, 2007). Analysis of these 26 strains revealed that the Streptococcal pangenome is open, that is, genome complexity increases with additional strains being considered and an asymptote is not reached (Figure 1.2). Comparative genome analysis revealed that open pangenomes are characteristic of environmentally sourced species, capable of inhabiting multiple ecological niches (Medini *et al.*, 2005). In comparison, the genome complexity within closed pangenomes does not alter, irrespective of the number of clinical isolates considered. In general terms clinically important organisms are clonal populations of organisms with tightly described species definitions and have closed pangenomes. (Figure 1.2).

- 3) ***The character “recent” genes:*** Some genes define the differences between two species (or between two populations of the same species such as genomovars or pathovars). These are present in a species core but not in an extended core and clonal variations in the character genomes are observed. When comparing sets of closely related species, the character genes are effectively the difference between these cores. 8000 functional gene families are estimated to make up the character genome (Lapierre & Gogarten, 2009). These gene families contain niche-specific genes which

are essential for survival and adaptation in particular ecological system. An example is the production and release of varying siderophores by different *Pseudomonad* species (Sullivant & Gara, 1992).

- 4) **The “other” genes:** These genes have arisen within a population through the presence of strong mutator alleles and/or floating genomes in order to increase mutation rates and persist in challenging environments, such as clinics. Broadly, the floating genome consists of mobile genetic elements, such as plasmids, transposons, insertion sequences, resistance gene cassettes *etc.*, all of which can act co-operatively and become readily dispersed across ecological and phylogenetic contexts (Liebert *et al.*, 1999). Generally, these other genes have resulted from rapid specialization in a single environment (Mena *et al.*, 2008).

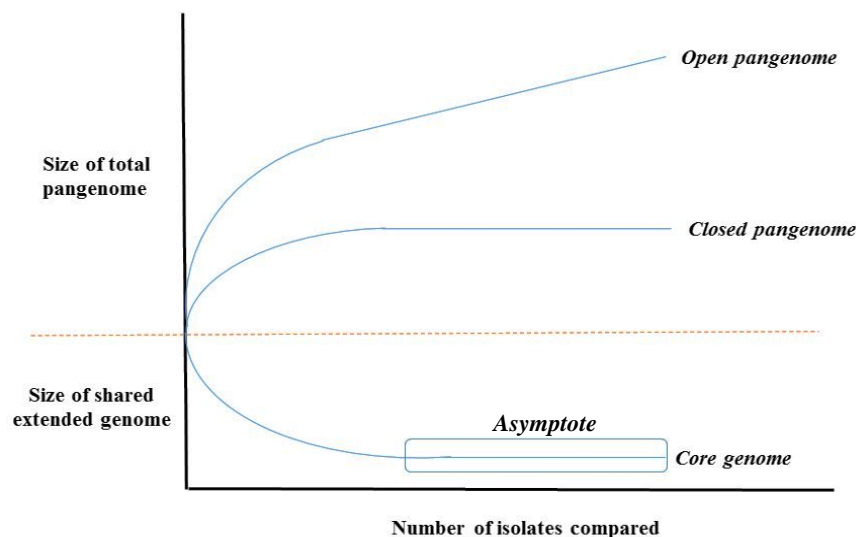


Figure 1.2: Size of shared extended genome and total genome is dependent on the number of isolates examined during comparative genome analysis.

The relationship between genome composition and species concept is thus a product of the bacteria and its ecological niche. Organisms that live in the same environment and do the same things share a high proportion of the same genes. The emergence of antibiotic resistance has involved the recruitment of phylogenetically disparate genes from both the core and acquired (floating) pools. Comparative genome analysis of pathogenic clones revealed niche specific characteristics (Hilker *et al.*, 2014). Acquisition of the genes allows clinical clones to encode adaptive capabilities in order to persist in an ever changing, hostile

environment. The recruitment of these floating genes is however poorly understood. How do we then recognise when a gene has come in and established itself as part of a species genome? Comparisons between genomes and identification of acquired genes can be used to make predictions about the organism's ability to cross ecological barriers, adapt and persist in a specific environment (Read & Ussery, 2006). Conceptually, this is influenced by the origin of this gene as well as its movement between genetic contexts, i.e. core to chromosomal framework. When comparing all genes within the core genome, it is difficult to determine gene origin or to predict gene flow. However, when comparing core genomes in a contextually ordered way (chromosomal framework), all genes are placed within an evolution context that allows us to compare and contrast gene flow and novelty. We can then define common ancestry, mode of gene acquisition, and make informative predictions of the adaptive response to "selective pressure" of these genes within/between species.

Acquisition of foreign DNA can be distinguished by virtue of its distinctive phylogeny and inconsistent presence within a bacterial species. Genetic differences between bacterial species are often observed in acquired genes and variation in genomes size between the species can often be explained by the presence of these foreign genes (Lukjancenko *et al.*, 2010). The cooperative nature of genetic elements comprising the acquired genome leads to the production of novel biochemical pathways, multi-antibiotic resistance phenotypes, and increased virulence (Kung *et al.*, 2010; Ozer *et al.*, 2014). More importantly, these elements are dispersed by horizontal gene transfer events; as such they are able to cross both ecological and molecular-genetic barriers, and establish in a range of prokaryotes (Gillings, 2013).

The integron-gene cassette system is distinctive in that it straddles the divide between core and floating genome. Briefly, chromosomal integrons (CIs) fit the conceptual definition of core genes (albeit recently acquired) (Gillings *et al.*, 2005; Holmes *et al.*, 2003; Mazel *et al.*, 1998; Vaisvila *et al.*, 2001). In contrast, gene cassettes associated with CIs almost never fit the concept of core genes and are part of the floating genome (Michael *et al.*, 2004). A plasmid borne multi-resistance integron (MRI) is floating and varies from strain to strain – i.e it is not core. However where the MRI has become embedded in chromosome and that becomes a feature of the strain (e.g. *Acinetobacter* global clone) then it fits the core concept (Diancourt *et al.*, 2010; Nemeč *et al.*, 2004). MRI platforms can be thought of as shuttles transferring genetic content between the floating and core genome. The primary interest is then in using the integron-gene cassette system in order to understand the forces involved in

the integration and spread of DNA across these genome components and bacterial species. Interactions between varying genome features need to be considered and their overall impact on speciation is vital to our understanding of adaptability, survival and resistance mechanisms in a selective environment.

1.4 Resistance mechanisms

Bacteria have evolved numerous mechanisms such as porins, efflux pumps, and cell permeability as a means of regulating cellular homeostasis. For instance, the AcrAB multidrug efflux pump in *Escherichia coli*, plays a physiological role in transporting acidic substrates out of the cell, thereby preventing toxicity (Du *et al.*, 2014). In comparison, as a means of bypassing low outer membrane permeability, *Pseudomonas aeruginosa* relies on porins to facilitate nutrient uptake under nutrient-limited conditions (Tamber *et al.*, 2006). The normal physiological roles of these mechanisms have adapted in such a way as to confer antibiotic resistance phenotypes. Efflux pumps, porins, and cell wall permeability are able to transport/channel/act on multiple substrates, including antibiotics (Hancock & Speert, 2000; Strateva & Yordanov, 2009). Thus bacteria have coping and resistance mechanisms against these bioactive molecules as a consequence of their natural physiology. Bacterial core genomes naturally encode resistance mechanisms to limit antibiotic susceptibility through an active process. These active resistance mechanisms includes antibiotic degradation or modification, and efflux (Figure 1.3) and are referred to as intrinsic resistance. Intrinsic resistance can be loosely defined as the natural ability of the microbe to resist the antibiotic by; i) lack of a drug target, ii) prevent antibiotic entry into the cell or iii) actively disable the antibiotic (D'Costa *et al.*, 2006; Lister *et al.*, 2009).

Whilst bacteria are able to utilise their native cellular machinery as defence mechanisms against antibiotics, in some instance these mechanisms are acquired from genetically distant sources. In contrast to intrinsic resistance, acquired resistance is typically more specialized and targeted to a specific role in protection against a given antibiotic rather than conferring generalized stress protection or by having a different primary function. Acquired resistance is also typically a result of strong selective pressures and consequently associated with the accessory genome (Breidenstein *et al.*, 2011). An example of this is the widespread occurrence of metallo-beta-lactamase resistance in *Ps. aeruginosa*, driven by plasmid-mediated transfer (Lagatolla *et al.*, 2004).

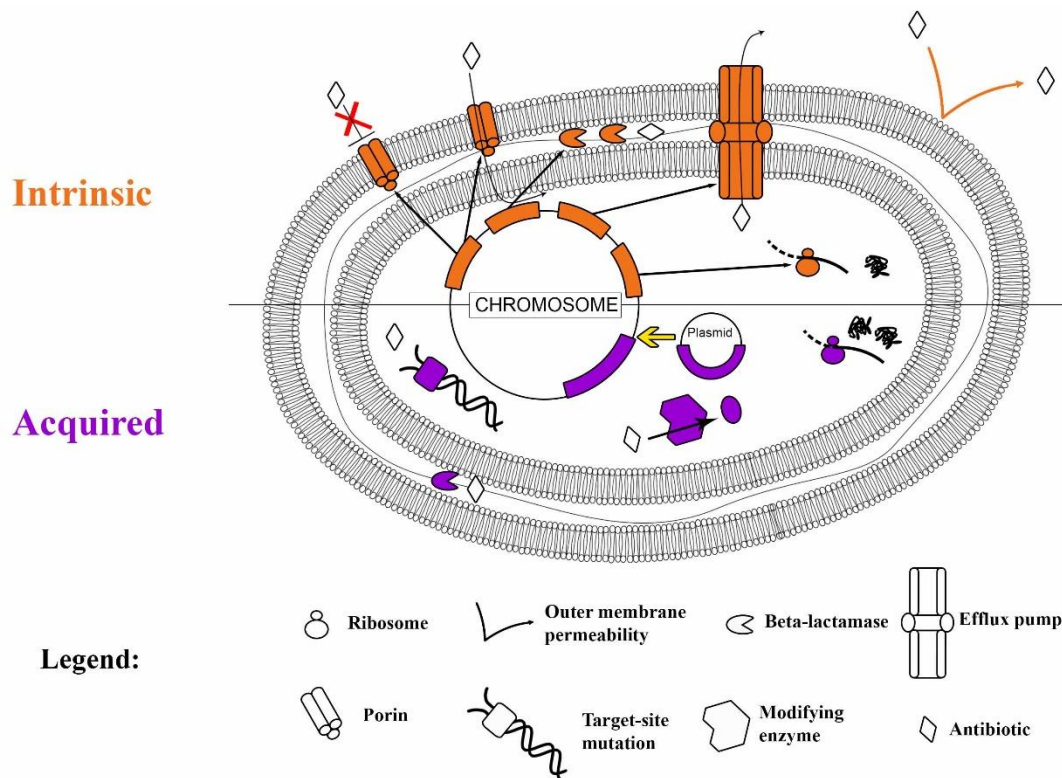


Figure 1.3: Intrinsic and acquired resistance mechanisms in *Pseudomonas aeruginosa*.

Intrinsic Panel: Resistance by non-specific means such as low outer membrane permeability, porin activity and chromosomally encoded genes against aminoglycosides, quinolones and beta-lactams. *Acquired Panel:* Resistance is drug specific and mediated by mobile genetic elements such as plasmids. Acquired resistance includes resistance to β -lactams; the presence of antibiotic-modifying enzymes; target site mutations; ribosomal mutations or modifications; and metabolic bypass mechanisms.

Plasmid mediated transfer of antibiotic resistance genes has been a major driving force in the successful dissemination of resistance in pathogenic clinical strains (Nikokar *et al.*, 2013) and has greatly driven the adaptation and genomic fitness of opportunistic pathogens. However, what determines which resistance genes become plasmid mediated? What are some of the biological forces that drive this selection, and add to the complexity of acquired resistance?

1.5 Overcoming ecological and phylogenetic separation: Horizontal gene transfer

Horizontal gene transfer (HGT) is a key driver of genomic innovation by promoting the formation of co-adapted combinations of genes in bacterial communities. However, both spatial separation as well as differences in the molecular-genetic makeup of bacteria will have an effect on HGT success. Two fundamental constraints for HGT are the susceptibility

of the DNA to attack/degradation and the acceptability of the DNA to the recipient cell. The three major mechanisms involved in HGT, transduction, conjugation and transformation, display different strategies to address the two constraints.

i) Susceptibility of DNA to attack and degradation: Strong habitat effect, some phylogenetic

DNA is susceptible to chemical and enzymatic degradation. Consequently ‘naked’, cell-free DNA degrades over time. Therefore the proximity of the recipient cell is an important factor in how readily DNA is translocated within genomes and/or between species (populations). During conjugation, depolymerisation of pilin subunits precedes DNA transfer through a conjugative channel (Firth *et al.*, 1996), thereby restricting the distance of separation between donor and recipient cells. The trade-off here is the requirement for physical proximity (μm distance), where close proximity drives the success of conjugative transfer (Coscollá *et al.*, 2011) (Figure 1.4). Conjugative transfer of DNA between a donor and recipient must occur within a defined ecological niche, and is also strongly dependent on a compatible transfer machinery between the two.

During transformation, extracellular DNA is taken up from the surrounding environment. Natural transformation occurs commonly in highly populated niches such as soil and biofilms (Demaneche *et al.*, 2001). In such environments, the DNA is not protected by a phage capsid or the host cell. Thus, DNA uptake has to occur before DNA “quality” is compromised. The success of natural transformation is however dependent on cell competency (Thomas & Nielsen, 2005), DNA uptake efficiency and cellular recombination machinery (Chen & Dubnau, 2004). For successful DNA transfer into a recipient cell to occur, the recipient must have the necessary outer membrane machinery for DNA uptake into the cytoplasm. Thus the molecular-genetic factors of the recipient play a vital role in natural transformation.

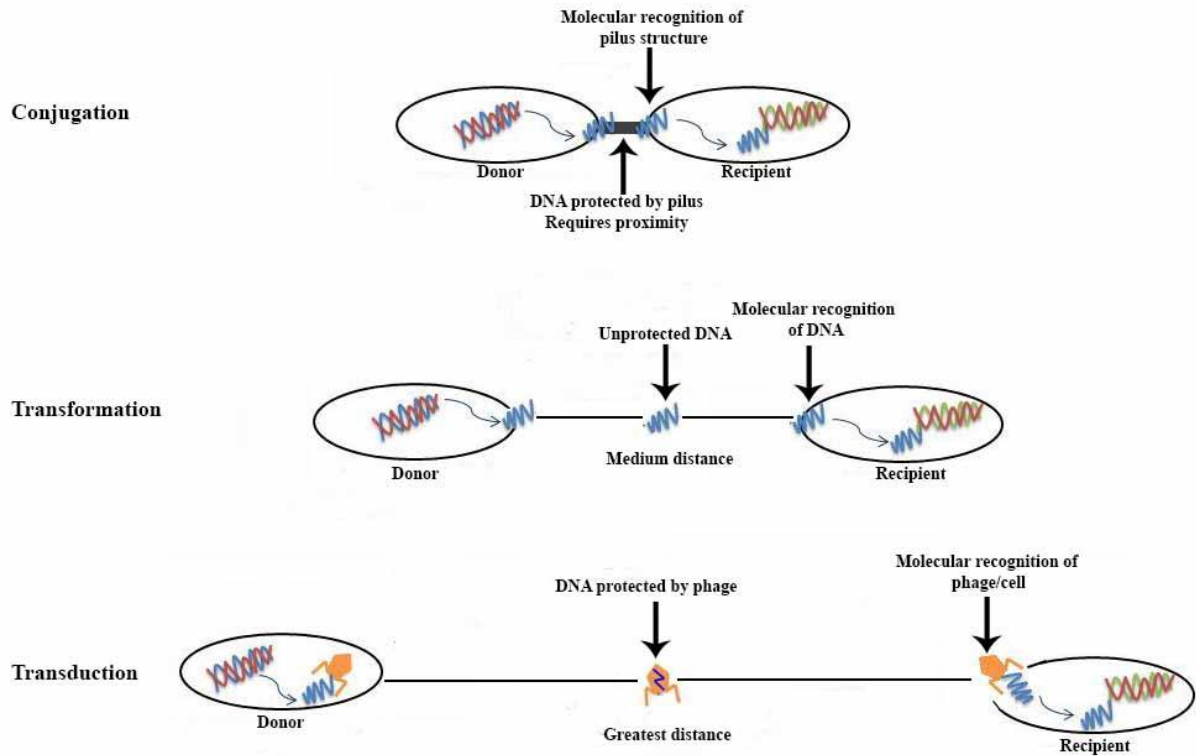


Figure 1.4: Modes of horizontal gene transfer and their relationships to spatial proximity.

Finally, during transduction, DNA is packaged in a phage particle that serves as protection in the extracellular environment (Brüssow *et al.*, 2004; Frost *et al.*, 2005). Transduction is thus not restricted to short distances as the phage capsid allows transmission across habitat separations, hence ecological separation is less of a constraint in this case. However, the trade-off is the constraint on DNA acceptability as the phage particle is dependent on specific molecular recognition machinery (Frost *et al.*, 2005). When the appropriate recognition machinery is present, phage mediated transfer of antibiotic resistance genes and pathogenicity islands has been noted to occur between distantly related bacterial species (Chen & Novick, 2009; Koskella & Meaden, 2013; Mazaheri Nezhad Fard *et al.*, 2011).

ii) Acceptability of DNA: Genetic and biochemical factors greatly influence stable maintenance of newly acquired DNA

A gene's fitness is constrained by its context. For a newly acquired gene to persist, it must confer an advantage to its host related to the host's environmental context. Integration of the gene into the chromosome must result in gene expression in order to continue this advantage. Gene maintenance is also related to integration with respect to cell division and replication machinery and finally, gene insertion is also impacted by surrounding genetic content

(adjacent genes within the chromosomal framework). Long-term maintenance and replication may occur through self-replication in the case of episomes, or through insertion into existing elements or chromosomal contexts (Eisen, 2000). Sharing common transcriptional and translational features, promoters and regulatory sequences, enhance the success of a HGT event by driving successful integration and expression of the newly acquired genes into the recipient cell (Andam & Gogarten, 2011; Iyer *et al.*, 2004; Olendzenski *et al.*, 2002). Therefore, the likelihood of a successful HGT event and DNA maintenance within a bacterial species is strongly correlated to the phylogenetic proximity of the partners (Coscollá *et al.*, 2011).

In some instances, the foreign DNA cannot be maintained within the new recipient. Host factors, restriction-modification systems as well as genomic sequence divergence all limit the success of incorporation of the newly acquired DNA into the genomic, chromosomal context (Matic *et al.*, 1996). Mismatch repair systems in some species have been noted to prevent homologous recombination based on insertions of distantly related genes. In the case of *Escherichia coli* and *Salmonella typhimurium*, the frequency of conjugational and transformational genetic exchange was found to be increased a thousand fold when the MutHLSU mismatch repair system was inactivated (Matic *et al.*, 1995; Rayssiguieret *et al.*, 1989).

Phylogenetic relationships can thus be used to predict both genetic and ecological relationships between donor and recipient. Both ecological and genetic conditions must be met, and newly acquired DNA must be stably maintained, before a selective advantage/phenotype such as antibiotic resistance emerges. The dispersal of resistance genes in the modern era indicates that rather striking barriers of ecological and genetic distance have been repeatedly overcome to result in a major problem for humanity. Yet there is very little understanding about how antibiotic resistance genes overcome these ecological and genetic barriers.

1.6 Exploring the diversity of MGEs and complexity of the floating genome

Mobile genetic elements (MGEs) facilitate HGT by overcoming some of the habitat and genetic barriers (Frost *et al.*, 2005). Despite the variability in MGEs (Figure 1.5), some recurrent structural and enzymatic features are frequently seen. Based on these features, MGEs can be broadly classified as replicons; transferrers; translocators and accumulators.

MGEs classified as replicons lower the molecular-genetic constraints related to stable maintenance. Replicons carry replication/partition genes that may be actively expressed in distantly related organisms (Tamminen *et al.*, 2012). Their presence allows for self-replication in a wide range of host cells and thus facilitate dissemination within and across habitats. Transferrers carry translocation genes with them and thereby possess autonomous mobility. Transferrers are also able to reduce the constraint of molecular-genetic distance through “sampling” available cell hosts. Similarly, translocators have the capacity to replicate in multiple and divergent hosts, found in varying ecological niches. Accumulators generate novelty through gene rearrangements, acquisition and altered expression and have the potential to increase microbial fitness (Nagy & Chandler, 2004, Schneider & Lenski, 2004).

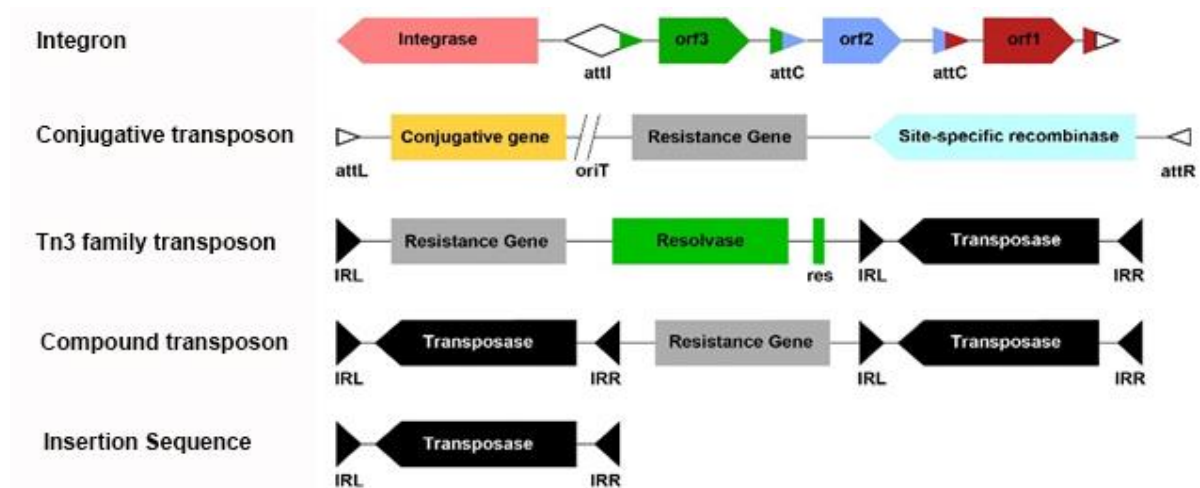


Figure 1.5: Mobile genetic elements are diverse yet share similar features that may include transposase genes, inverted repeats (IR), resistance genes and recombination sites (*attC*). These are all examples of MGEs that have genetic “handles” and thus possess the ability to translocate and recognise other MGEs that have abilities such as transfer and/or replication.

1.6.1 Replicons-Plasmids

Categorisation of plasmids is based on their autonomous replication mechanisms. Plasmids are extrachromosomal DNA elements, which are capable of replicating independently of the

genome. Plasmids categorized into the incompatibility groups IncP, IncQ and IncW, promote horizontal gene transfer of resistance genes among unrelated bacteria (Carattoli, 2013). The role of plasmids in the widespread dissemination of resistance to the major classes of antimicrobials, such as β -lactams and aminoglycosides, is well understood and documented (Carattoli, 2009). Broad-host-range plasmids in the IncP-1 family have been linked to the spread of resistance to a broad spectrum of antibiotics, heavy metals, and quaternary ammonium compounds used as disinfectants (Popowska & Krawczyk-Balska, 2013). Furthermore, the IncP-1 family (of plasmids) is able to replicate in numerous species of *Alpha*, *Beta*- and *Gamma*-*proteobacteria* and are thus able to persist in a diverse range of ecological and genetic niches. Whilst these plasmids can replicate in a diverse range of bacteria, their overall stability is varied and eventual loss of the plasmid can be observed (De Gelder *et al.*, 2007). Note also that not all broad-host range plasmids have the capacity to become as widespread as the IncP-1 family. Host range thus highlights constraints on overall distribution across ecological and genetic barriers.

Host factors and plasmid-related interactions such as conjugative transfer, plasmid partitioning as well as accumulations of mutations are linked to plasmid loss and acquisition (Koraimann, 2004; Williams & Thomas, 1992). Success of conjugative plasmid transfer is dependent on close cell-to-cell proximity, presence of pilus and molecular recognition of pilus by recipient cell. Plasmid maintenance is a trade-off between the cost of plasmid mobility and the advantage they provide the host (Svara & Rankin, 2011). Acquisition of MGEs such as transposons, integrons and insertion sequences can provide a means of ensuring that these plasmids are maintained within a population by conferring a selective and adaptive advantage to its host. Therefore the replicon “cargo DNA”, highlights constraints on the selective advantage it may provide to its host bacterium.

What remains largely unknown is how these MGE are acquired and assembled into replicons in the first place. If a plasmid does not provide an advantage, is likely to become a cell burden and may ultimately be lost from a population.

1.6.2 Transferrers-ICE

Integrative and Conjugative Elements (ICEs) are a group of self-transmissible genetic elements that must integrate into an existing replicon in order to achieve replication (Burrus

et al., 2002). ICEs are less dependent on Inc mechanisms for stable maintenance, thus are better able to co-exist with other replicons. ICEs act as vectors for gene transfer, ultimately resulting in the spread of antibiotic resistance genes (Larbig *et al.*, 2002), metabolic genes (Klockgether *et al.*, 2004), virulence and pathogenicity genes (Battle *et al.*, 2008), and genes for toxic compound transportation (Larbig *et al.*, 2002). ICEs with low target site specificity are termed conjugative transposons and are tens of kb in size. In comparison, ICEs with higher target sequence specificity and which are larger in size are termed genomic islands (GIs) (Wozniak & Waldor, 2010).

Their complex genetic makeup arises by assembly of a diverse range of MGEs, such as ISs, transposons (Tns) and antibiotic resistance genes (Juhas *et al.*, 2009; Pembroke *et al.*, 2002; Harada *et al.*, 2010). By means of autonomous mobility, transposons remove constraints via their ability to alter their genetic context (chromosome ↔ plasmid) (Douard *et al.*, 2010). However, autonomous mobility does not guarantee long term maintenance of ICEs in a bacterial population. Additionally, surface exclusion may prevent ICE entry into a bacterial cell and thereby reducing the likelihood of long-term establishment of the ICE within the species (Frost *et al.*, 1994). ICE mobility can result in the formation of novel genetic structures and the spread of virulence and resistance genes, but it still leaves us wondering how different MGE assemble into their ICE platform in the first place.

1.6.3 Translocators and accumulators-transposons, compound transposons, bacteriophages, insertion sequences, integrons

Translocators, such as Tns, compound transposons (CTNs), and Insertion sequences (ISs), have the innate ability to translocate from one genetic context (e.g. chromosomal) into a new location (e.g. plasmid). Often their translocation process results in the mobilization of additional genetic cargo. This acquisition requires a degree of stable maintenance of the translocators as well as its DNA “cargo”. Once presented with a new genetic context, recombination of the translocators and its cargo must however ensure that recipient integrity is maintained in order for the translocator to persist within a population. A classic example is the mobilization of resistance genes in the *class1* integron array by transposon Tn402 (Gillings *et al.*, 2009; Labbate *et al.*, 2008; Sajjad *et al.*, 2011). Translocation of such an integron array into a plasmid, impacts on dissemination of resistance within a bacterial population. Translocators alter both the genetic context of the resistance gene and its

ecological distribution to overcome both ecological and phylogenetic constraints of HGT. How exactly does the cargo DNA get into the translocator? What determines where and how fast the cargo is delivered?

i) Transposons and Compound Transposons

Transposons have undoubtedly been a source of genetic diversity, and Tn mobility is often a balance between positive and negative outcomes on the host. Tns can alter their fitness through association with other MGE such as plasmids. Tns mediate their own transfer within a genome, and integrate into unrelated DNA. Transposons can vary greatly in size and genetic makeup as they can harbour various accessory genes, such as metal resistance, antibiotic resistance and metabolic processing genes. Translocation of the Tn3 family is dependent on three key structural features; a *tnpA* gene for transposition, a *tnpR* gene for resolution, and short (15 to 40-bp) terminal inverted repeats (IR) (Bennett, 1999, 2004; Kleckner, 1981) (Figure 1.5). Translocation of antibiotic resistance genes in hospital, environment and commensal associated bacteria is mostly due to large broad-host-range plasmids and the transposons they carry (Minakhina *et al.*, 1999).

CTNs consist of a central region flanked by two identical or near identical insertion sequences (ISs) (Figure 1.5). This central region contains accessory genes for antibiotic resistance and metabolic pathways (Chalmers & Kleckner, 1996; Meer & Zehnder, 1991). The presence of the catabolic and antibiotic resistance genes on CTNs, suggest a mechanism by which gene clusters can be generated to form novel catabolic and resistance pathways and become widespread.

ii) Bacteriophages

Bacteriophages can also be thought of as translocators. Many GIs are thought to have arisen by bacteriophage phi (ϕ) integrase machinery (Boyd *et al.*, 2009). Recent evidence has emerged suggesting that bacteriophages in wastewater treatment plants are vehicles for the transfer of resistance genes that encode beta-lactamases (*blaTEM*, *blaCTX-M*) among Gram-negative pathogens, and penicillin-binding protein 2a (PBP2a), associated with methicillin resistance (*mecA* gene) in staphylococci (Colomer-Lluch *et al.*, 2011). This demonstrates that bacteriophage mediated HGT has the capacity to overcome both ecological and phylogenetic barriers. A constraint however is the bacteriophage capsid size, which limits the amount of cargo DNA that can be efficiently packaged and translocated.

iii) **Insertion Sequences (ISs): simplest, self-mobilisable transposable elements**

IS translocation can impact on bacterial genomes in numerous ways (Bartlett & Silverman, 1989; Mahillon & Chandler, 1998; Mills *et al.*, 1992; Sokol *et al.*, 1994). The primary function of ISs is to maintain themselves through random or site-specific translocation. IS are able to aggregate genes into mosaics, mobilize them, and thereby provide a positive net-effect to the host (Hacker & Carniel, 2001b; Klockgether *et al.*, 2004a; Venkatesan *et al.*, 2001). Despite being so widespread, IS translocation can be limited by phylogenetic separation (Mahillon & Chandler, 1998). Appropriate host factors may be necessary to drive the initial IS translocation into a higher-order translocation (Tn/CTn) or replicator.

iv) **Integrans and gene cassettes: accumulators of resistance genes**

In 1989, Stokes and Hall defined integrans as a genetic platform which served to accumulate genes. Most importantly, integrans harness the power of DNA translocation to overcome both ecological and phylogenetic boundaries. The minimal requirement for an active integron platform were identified as a core integron structure, *intI* and *attI* sites. Integrans are categorized by the amino acid sequence of the encoded DNA integrase. Site-specific recombination mediated by the integron integrase achieves directed insertion of gene cassettes into the integron platform, through the recognition of two distinct recombination sites, *attI* and *attC* (Collis *et al.*, 1993) (Figure 1.6).

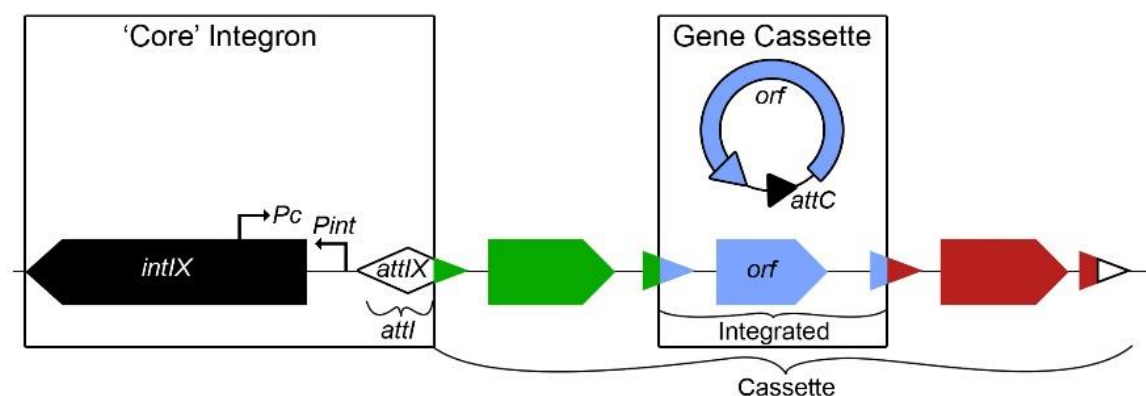


Figure 1.6: General structure of the integron-gene cassette system.

The *intI* gene encodes a tyrosine recombinase (integrase) that catalyses site-specific recombination between the *attI* and *attC* sites, resulting in integration or excision of gene cassettes.

Gene cassettes themselves are non-replicating, transcriptionally silent, circularised units containing a recombination site, *attC*, and one or more open reading frames (*orfs*) (Recchia &

Hall, 1995). The independence of gene cassettes is what makes integron accumulators. Conceptually a translocator goes to multiple different fixed points in the chromosome, but an accumulator brings multiple separate sites to one location. Thus in order to be expressed, gene cassettes must be integrated into the integron platform via site-specific recombination between *attI-attC* or *attC-attC* sites. Upon capture by the integron, gene cassettes are transformed from free circular molecules into linear integrated molecules. They are then replicated with the chromosome or plasmid hosting the integron (Collis & Hall, 1992). Once integrated, gene cassettes are expressed by the integron *P_c* promoter. The location and strength of this promoter varies between the different integron classes (Collis & Hall, 1995; Levesque *et al.*, 1994). The incorporation of many different gene cassettes, by multiple recombination events, results in an array of gene cassettes. Such arrays may range from 0 to 200 cassettes in size (Boucher *et al.*, 2007; Mazel *et al.*, 1998). The accumulative nature of integrons, when coupled with antibiotic selection in clinical systems, can facilitate selection for integron-mediated multidrug resistant pathogens (Krauland *et al.*, 2009; Leverstein-van Hall *et al.*, 2002; Livermore & Woodford, 2006). While an integron is able to recruit gene cassettes, it requires a translocator and a replicator to disseminate within/between bacterial populations (Frost *et al.*, 2005). An example of accumulators, translocators and replicators working co-operatively is found in plasmid pAX22. This plasmid contains integron In70 carrying resistance genes to metallo- β -lactamases. Translocation of this integron is believed to be mediated by Tn7017. Mobilization of this integron-transposon into the plasmid is believed to have occurred by an additional translocator, *ISPa17* (Di Pilato *et al.*, 2014).

The environmental and genetic factors that regulate successful conjugative transfer of antibiotic resistance genes and MGEs in bacterial populations are highly complex (Beaber *et al.*, 2004) and is often challenged with varying constraints. These constraints are largely dependent on interactions between multiple types of MGEs and differences in their mechanistic capacity (Table 1.1).

Table 1.1: A summary of the mechanistic capacities of MGEs

MGE	Genes/Enzymes	Function
Replicons	<i>rep/par</i> genes	Replication and partitioning of plasmid in cell
	Toxin/ antitoxin genes	Ensure plasmid selection and maintenance in population
Transferrers	Transfer/ conjugation genes	Vector for genetic information
Translocators	Transposase	Self-mediated translocation without replication
Accumulators	Integrase, <i>attI</i> , <i>attC</i> , gene cassettes	Recombination of genetic information into a meaningful context

1.7 Site-specific recombination: a key process in multidrug resistance assembly and spread

Despite structural, enzymatic and differences in target site preference, site-specific recombination is a recurrent theme for many MGE elements mentioned. Note that site-specific recombination is especially important in multi-drug resistance generation and spread (Hall & Collis, 1995; Sundstrom *et al.*, 1988). Site-specific recombination enables efficient interactions between different MGEs and determines which MGEs interact. Site-specific recombination is a conservative process during which DNA strand exchange occurs between two DNA sequences with a limited degree of sequence homology (Grindley *et al.*, 2006). Well known site-specific recombinases fall into one of two families: tyrosine recombinases (λ integrase family), and serine recombinases (resolvase family). Serine recombinases will not be examined in depth in this thesis. Instead it should simply be noted that serine recombinases require a serine within their active site to produce double strand breaks at the cross over sites before any exchange of DNA can be initiated (Grindley *et al.*, 2006).

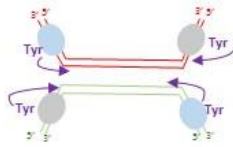
Tyrosine recombinases use an active tyrosine amino acid to cleave one strand of the duplex chain at a time (Star *et al.*, 1992). Members of this family include XerC and XerD, Cre of bacteriophage P1, FLP of yeasts, bacteriophage λ Int recombinases (Lee & Sadowski, 2005) as well as the integron integrase IntI. Tyrosine-recombinases exhibit strand selectivity during recombination reactions. One strand on each participating DNA is cleaved by a reactive tyrosine residue to form a 3'-phosphotyrosyl intermediate and a free 5'-hydroxyl group. Strand exchange between the two sites follows, generating a Holliday junction (HJ)

intermediate (Figure 1.7). The opposite strands then exchange to form recombinant products (Blakely, 2000; Grindley *et al.*, 2006).

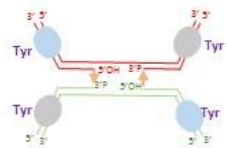
MGEs have evolved independently of one another (Table 1.2). Their fate and impact on HGT is strongly linked to their ability to target their insertion sites with some MGEs displaying striking site-specific targeting within conserved genes (Peters & Craig., 2001; Petrovski & Stanisich, 2010; Stanisich *et al.*, 1989) or site-structures (Tetu & Holmes, 2008).

Table 1.2: MGE and their independent evolution-enzymes and mechanisms

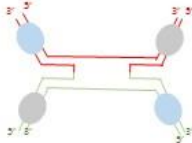
Recombinase Family	Examples of elements	Outcome	Reference
Serine	Tn3 (TnA or type II transposons) and $\gamma\delta$	One copy of the TE remains at the original site while the other copy is inserted at the new recombination <i>res</i> site	Grinsted <i>et al.</i> (1979)
Tyrosine	XerC/XerD	Crossover at the specific <i>dif</i> site present in the <i>E. coli</i> K-12 DNA replication terminus region	Blakely, (2000)
	Cre- <i>loxP</i>	Site-specific recombination between 34-bp <i>loxP</i> sites	Sternberg <i>et al.</i> (1981)
	Tn4555	conjugal transposition is hypothesized to occur via a circular intermediate	Tribble <i>et al.</i> (1997)
	Phage lambda integrase	Integration and excision of DNA at <i>attL</i> and <i>attR</i> attachment sites	Landy & Ross, (1977)
	Integron Integrase (IntI)	Integration and excision of gene cassettes into the integron array by <i>attIxattC</i> and <i>attCxattC</i> recombination	Demarre <i>et al.</i> (2007)
DDE	Tn7, Tn10, Tn5 IS3 family	Tn is excised from the original site and inserted at the new target site. Preference for the sequence 5'-GCTNAGC-3' as the target site	Bender & Kleckner, (1986), 1992; Kleckner <i>et al.</i> (1996)
DEDD	IS1111- <i>attC</i> family	Predicted to target <i>attC</i> structure via a Piv/MooV transposase homologue	Tetu & Holmes, (2008)

A) XerC/D Recombinationi) Synapsis & Attack 

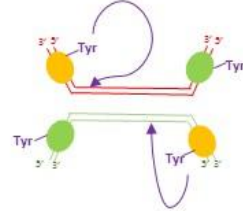
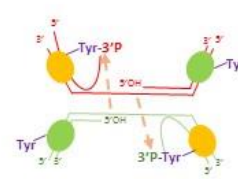
ii) Reactive Radical



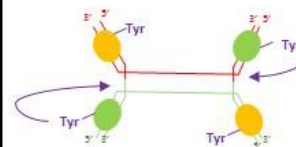
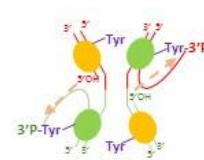
iii) Holliday Junction

**B) Cre-loxP Recombination**

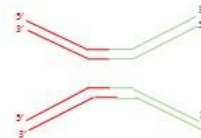
i) Synapsis

ii) 1st Attack

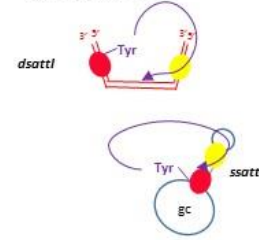
iii) Holliday Junction

iv) 2nd Attack

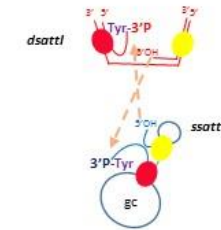
v) Recombinant

**C) Int1 attI x attC Recombination**

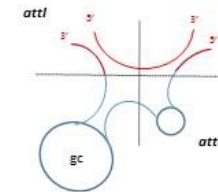
i) Synapsis



ii) Attack



iii) Holliday Junction



iv) Recombinant & Replication

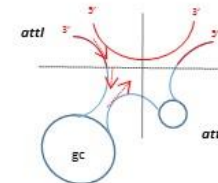


Figure 1.7: a) XerC/D, b) Cre-loxP, and c) Int1 mediated recombination all require a precise cross-over of target DNA sites. XerC/D and Cre-loxP require two double stranded DNA strands for site-specific recombination, whereas Int1 requires one double stranded attI and one single stranded attC hairpin. XerC/D is essential for chromosome segregation in *E. coli* and Cre-loxP is essential for bacteriophage P1 integration into the chromosome. i) Synapsis and attack involves the attachment of 4 enzymatic units and attack by the active tyrosine to form a 3'-P and 5'-OH radical. ii) Reactive Radical/Attack involves strand exchange driven by the reactive 5'-OH radical. iii) A Holliday junction forms iv) and v) is resolved by recombination and resolution.

1.7.1 *IntI Integrase, an unorthodox site-specific recombinase*

IntI recombination can be thought of as non-classical site-specific recombination. In contrast to XerC/XerD (Figure 1.7A) and Cre-*loxP* (Figure 1.7B) recombinases which require two double stranded DNAs, IntI1 mediated site-specific recombination requires one double stranded *attI* site and one single stranded, *attC* hairpin (Figure 1.7C). Francia *et al.*, (1999) demonstrated that IntI1 binds double stranded *attI* molecules, but only single stranded *attC* substrates, specifically the bottom strand, were bound by IntI1. This single stranded *attC_{bs}* adopts a double stranded DNA-like structure by the annealing the complementary IR, forming a structure similar to the canonical recombination site: two IR separated by a short spacer (Johansson *et al.*, 2004, Stokes *et al.*, 1997, Francia *et al.*, 1997). Resolution of the *Vibrio cholera* IntI (VchIntIA)-*attC* crystal structure further supported the hypothesis that insertion of gene cassettes requires a single stranded, folded back *attC* (MacDonald *et al.*, 2006).

The diverse range of *attC* sites in multidrug resistance integrons (MRIs) arrays highlights the ability of IntI1 to overcome phylogenetic boundaries. MRIs are functional integron platforms physically associated with mobile DNA elements (Tns), which can be transferred by conjugative plasmids. The lack of conservation between MRI *attC* sites, suggests that they originate from biologically distinct sources. In comparison, the ability of class 2 integrase (Int2) to overcome phylogenetic separations is believed to be limited given the high conservation, and relative simplicity of In2 arrays (Ramírez *et al.*, 2010).

Whilst IntI1 is able to recognize diverse *attC* sites, its action on *attI* sites is more specific (it only acts on *attI*). *attI* sites are highly conserved within an integron class (e.g. *attI* occurs in class 1 integrons) (Partridge *et al.*, 2000), and highly divergent between classes (Elsaied *et al.*, 2011). Note that *attI* and *attC* sites are canonical DNA targets, and that IntI-*attC* binding is dependent on the topological conformation of the *attC*, whereas IntI1-*attI* binding is dependent on *attI* sequence recognition. Four IntI1 binding sites, (weak, strong and two regions within the simple site) are involved in IntI1-*attI* recombination activity (Partridge *et al.*, 2000; Stokes *et al.*, 1997) (Figure 1.8). Binding sites 1 and 2 contain the core recombination site (GTTRRRY). Site 3 is the strong binding site while site 4 is the weak binding site and are direct repeats (DRs), which are adjacent to one another and are not conserved across all other *attI* sites (Stokes *et al.*, 2001).

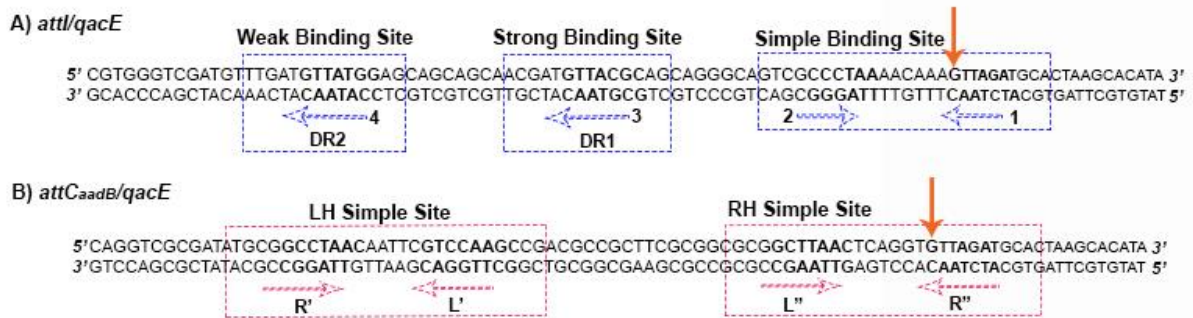


Figure 1.8: Schematic illustration of the structure of A) *attI* and B) *attCaadB* sites.

A) Experimentally determined strong and weak *IntI*-binding sites (Collis *et al.*, 1998) and a pair of direct repeats, DR1 and DR2, are indicated by blue dashed boxes and arrows.

B) The location of simple sites and their inverted repeats are shown by dashed red boxes and arrows. R'' contains the recombination crossover point GTTRRRY (marked with vertical orange arrows). R' contains the inverse core site with the consensus sequence RYYAAC. The *qacE* part of the sequences is represented by smaller font size (Bouvier *et al.*, 2009; Partridge *et al.*, 2000).

IntI specifically recognises cognate *attI*, which contrasts with its ability to recognise and bind variable, unrelated *attC* sites. The ability of *IntI* to overcome sequence and conformation constrains, as well as the formation of single stranded structures, are evolutionary driving forces and must be considered when studying the formation and dissemination of multi-drug resistance. In summary, the roles of site-specific recombination are in either precisely joining or precisely separating DNA targets. Their specificity for one or both target sites is therefore an indicator of their biological roles.

1.8 A model system for the study of gene flow across ecological and phylogenetic barriers

Studying the interaction between different mobile genetic elements and making predictions about the rate of dissemination of resistance genes is challenging. In the majority of cases the source of these resistance genes is unknown. *In vitro* quantification of antibiotic resistance and sensitivity are routinely performed on pure cultures in clinical laboratories (Ahmed *et al.*, 2013; Chaudhary *et al.*, 1998; Drago *et al.*, 2005). However, this a crude and often insensitive measure of resistance and thus cannot be regarded as a complete representation of resistance. *In vivo* quantification, whilst difficult to achieve and establish as a routine diagnostic approach, provides more information on resistance than does *in vitro* analysis (Nilsson *et al.*, 2003; Riou *et al.*, 2010). Understanding resistance at an *in vivo* level would result in more accurate drug administration, a localized form of treatment, and overall lower

resistance levels. Lastly, it is difficult to predict the success of resistance gene maintenance within a bacterial population. Understanding these factors is clearly important if we wish to make informative decisions about antibiotic administration and long-term efficacy in clinics. Given the multi-factorial nature of antibiotic resistance, studying resistance gene flow over ecological and phylogenetic separations requires the implementation of a few different model systems.

1.8.1 Integron, the model system for gene flow between mobile and chromosomal contexts

Based on the sequence of the encoded integrase (40%–58% identity), five different classes of mobile integrons (MIs) have been defined (Cambray *et al.*, 2010). MIs can broadly be defined as integron platforms located on mobile elements such as plasmids or Tns. Class 1, 2 and 3 integrons, exhibit a number of features not typical of the more diverse chromosomal integron classes. Class 2 integrons are typically associated with Tn7-like transposons (Ramírez *et al.*, 2010), while class 3 integrons are typically associated with Tn5053-like transposons (Collis *et al.*, 2002). Hall *et al.* (1999) first suggested that different classes of integrons reflect different origins from different evolutionary groups. However, despite representing different evolutionary origins, integron classes can share resistance gene cassettes. For instance, the *aadB* antibiotic resistance gene can often be found in a variety of mobile integrons. This highlights the relative importance of HGT and interactions between different integrons in driving the spread of antibiotic resistance (Partridge *et al.*, 2009). The majority of integrons encountered in the clinical context are class 1 integrons. This is due to their ability to capture and express a diverse range of resistance genes. They are also widespread in pathogenic and commensal populations (Goldstein *et al.*, 2001).

Class 1 integrons are observed in various ecological habitats (Stalder *et al.*, 2012), predominantly the clinical setting (Essen-Zandbergen, 2007), as well as different phylogenetic contexts, replicons, transmitters and translocators. The widespread dissemination of class 1 integrons (Int1) coincides with the widespread use of antibiotics and the selective pressure which they exert. As a response to these selective pressures and natural selection (Gillings *et al.*, 2008), class 1 integron arrays are constantly evolving at a phenomenal rate.

Class 1 integrons most commonly contain between 0 to 9 cassettes, almost all of which encode antibiotic resistance determinants. This diversity often reflects the selective pressures that the bacteria needed to overcome, as the arrays often match the prevailing antibiotics used in the environment from which the strain was recovered (White *et al.*, 2001; Xu *et al.*, 2013). Despite variations in gene cassette arrays and the different genetic contexts that class 1 integrons are found in, the 5' and 3' region of class 1 integrons are typically well conserved (Figure 1.9). They are associated with functional and non-functional transposons derived from Tn402, which can be further embedded in larger transposons such as Tn21.

Stokes *et al.* (2006) and Gillings *et al.* (2008) performed PCR-based surveys for class 1 integron integrase genes using non-clinical environmental isolates. These non-clinical class 1 integrons, differed from their clinical “relatives” in two ways. Firstly the gene cassette content differs between these integrons and secondly, they did not contain the 3' CS and Tn402-related flanking sequences. This work illustrated the importance of the Tn flanking gene in the mobilisation of the integron platform to the clinical environment where these resistance gene cassettes are selected.

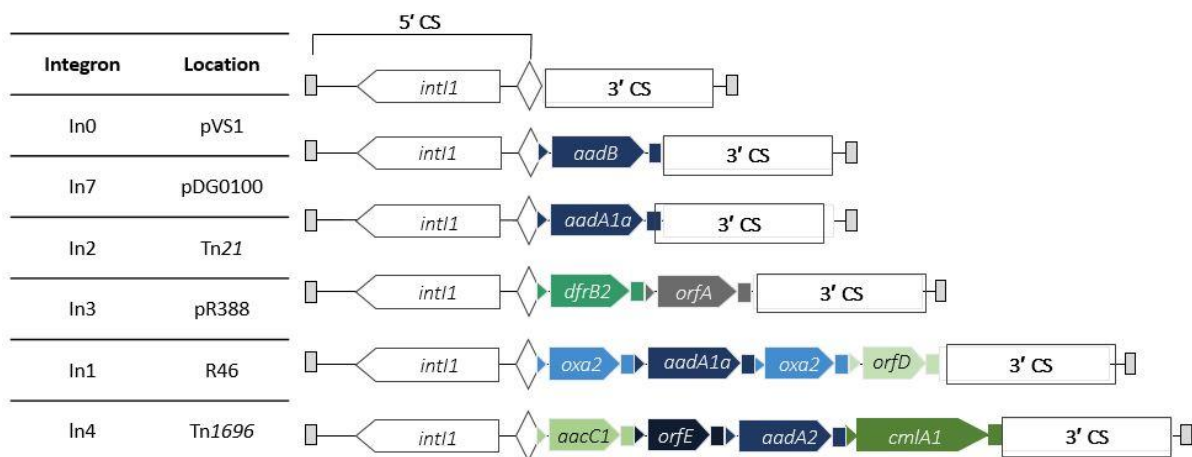


Figure 1.9: Resistance markers vary between integrons and in most cases are flanked by the 5'- and 3'- conserved region (CS).

Examples of class 1 integrons depicted here contain different resistance markers (*aadB*, *aadA1a*, *dfrB2*, *oxa2*, *aacC1*, *aadA2*, *cmlA1*) and their respective cross-over *attC* sites (rectangles and small triangles). These integrons are localized on different transposons and plasmid backbones. The 5'CS region contains the *intI1* and *attI1* region. 3'CS region generally encompasses the *sul1* resistance gene (sulphonamide), *qacE* resistance gene (quaternary ammonium), and *orf5* (unknown function).

Two fundamental questions have arisen from this research. Firstly, what was the genetic context/origin of the class 1 integron ancestor? Chromosomal class 1 integrons have commonly been observed in a wide range of nonpathogenic β -Proteobacteria, including

members of the genera *Hydrogenophaga*, *Aquabacterium*, *Acidovorax*, *Imtechium*, *Azoarcus*, and *Thauera* (Gillings, 2014). There is mounting evidence that suggests that a chromosomal class 1 integron within β -Proteobacteria was captured by a Tn402 transposon by site-specific recombination. This hybrid also contained the *qacE* cassette, which is able to confer resistance to disinfectants. Note that the use of disinfectants predates the use of antibiotics. Subsequent acquisition of the *sull* gene encoding resistance to sulfonamides, and *orf5*, with an unknown function, resulted in a partial deletion of the *qacE* cassette, thereby generating the 3'-conserved region. Further steps involved deletions, insertions of *tni*, and acquisitions of additional antibiotic resistance cassettes. Subsequent step involved the mobilization of the Tn402-integron hybrid into a replicon and/or other transposons such as Tn21 family. These events lead to the mobilization and widespread dissemination of the class 1 integron across ecological and phylogenetic contexts (Gillings *et al.*, 2008).

Secondly, what is the origin of resistance genes observed in clinical class 1 integrons? The answer to this question may also lie in the chromosomal framework of non-pathogenic bacteria! IntI homologs are present in 17% of available bacterial genomes and are prevalent in members of the γ -Proteobacteria (Cambray *et al.*, 2010) (Figure 1.10). These clade specific integron homologs show some features of acquired genes (gene cassette arrays), yet unlike their mobilized counterparts, these show signs of incorporation into chromosomal framework of its host (Heidelberg *et al.*, 2000). As such they are broadly referred to a Chromosomal integrons (CIs).

CIs generally exhibit patterns of stable vertical inheritance in bacterial lineages (Boucher *et al.*, 2007; Gillings *et al.*, 2005; Holmes *et al.*, 2003; Mazel *et al.*, 1998; Mazel, 2006). Their “patchy” distribution across different bacterial genera, suggests that collectively CIs have experienced multiple losses and undergone multiple gene transfer events. Yet their chromosomal context and evolutionary trajectory also shows their co-evolution within a species. CIs within a species are often observed in the same locus. For instance, the PA3672 locus is one of the insertion points for *Pseudomonas* chromosomal integron (PCI) in *Ps. stutzeri* strain Q, *Ps. stutzeri* BAM17 and other *Pseudomonads* (Holmes *et al.*, 2003). The co-evolutionary trajectories of 16S rRNA DNA, housekeeping gene *rpoA* and *intI*, further emphasise that chromosomal integrons are an ancient feature of bacterial genomes (Nemergut *et al.*, 2008). Their “fixed” genomic nature increases their long-term stability to selective pressures, and in turn provides increased opportunity for the evolution of multi-gene

phenotypes (Koenig *et al.*, 2009; Rosewarne *et al.*, 2010; Wright *et al.*, 2008). CIs are very closely tied to the habitats within the ecological niche of their “host” organisms. In contrast, mobilized MRIs are widely distributed between species and are maintained as a consequence of the selective advantage they provide in an antibiotic rich environment (Boucher *et al.*, 2007). Consequently MRIs are found in a diversity of habitats. Given that MRIs carry antibiotic resistance genes, and CIs contain cassettes of no known resistance function, what is the origin of resistance cassettes?

1.8.2 CI-to-MRI flow: Analysis of the gene cassette metagenome

The functions encoded by gene cassettes in the clinical context and found in MRIs is well understood. Comparatively, the importance of gene cassettes recovered from environmental samples and genetic contexts, remains poorly understood. Environmental gene cassettes are numerous, diverse and generally do not confer resistance phenotypes. The vast pool of environmental gene cassettes is thought to represent a reservoir where:

- i) All gene cassettes are potentially able to be captured by a MRI integron integrase. This is supported by the diversity of gene cassettes/*attC* demonstrated to be recombination substrates to IntI (Holmes *et al.*, 2003; Rowe-Magnus *et al.*, 2002; Stokes *et al.*, 2001).
- ii) Any gene can be captured within a cassette and become mobilized. The evidence to support this can be found in the sequence and structural diversity of ORFs: bioinformatic analysis suggests no limits to what is contained (Koenig *et al.*, 2008).
- iii) Under the right circumstance may provide its host with an evolutionary advantage. For example, by the expression of novel enzymes (Neild *et al.*, 2004; Robinson *et al.*, 2005).

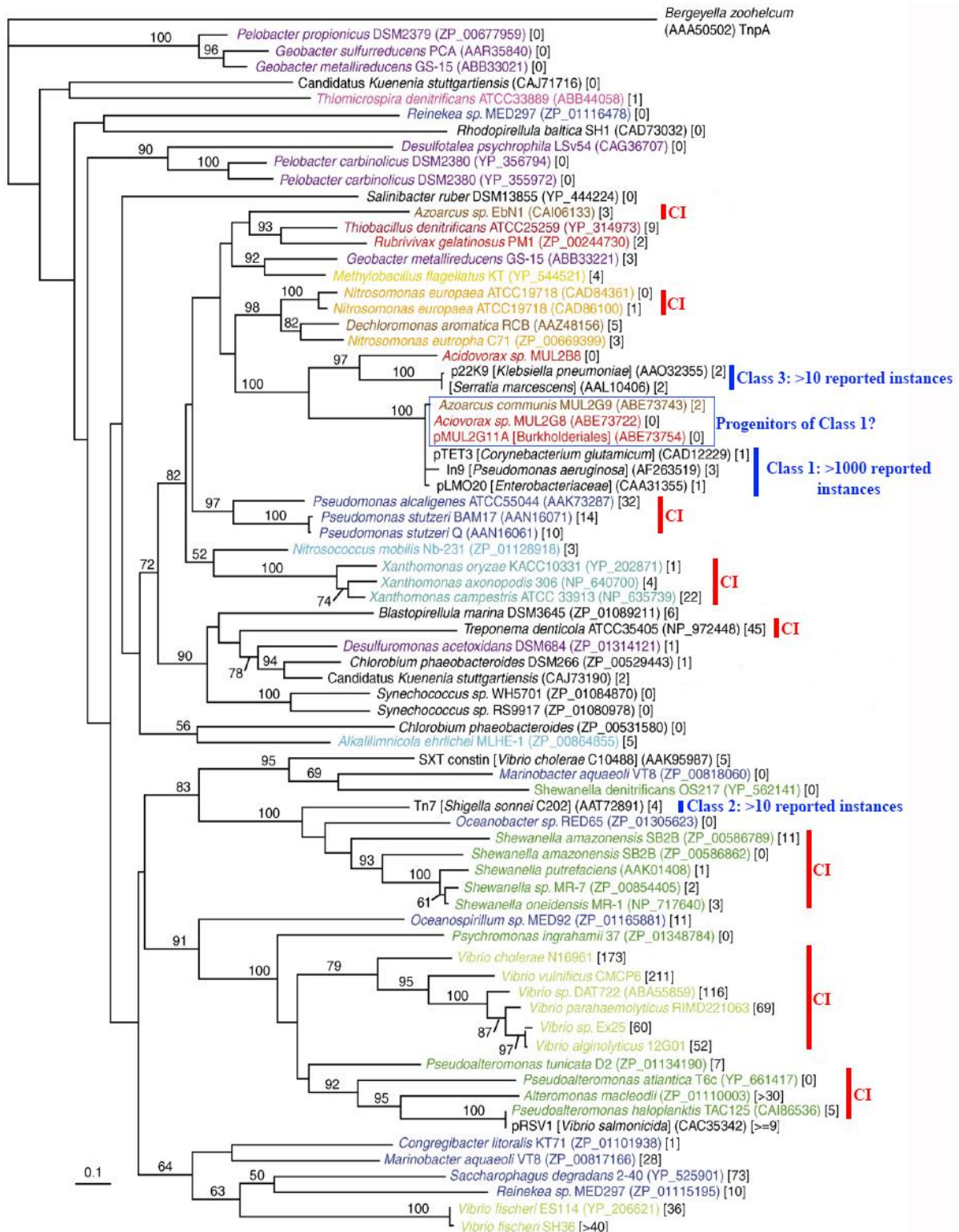


Figure 1.10: Phylogenetic tree of known integron integrases (IntI).

The highlighted group are progenitors of the class 1 integron scaffold. All other integrons are potential reservoirs of the gene cassettes that have been accumulated within this.

A single integron integrase is included for each bacterial species. The number of gene cassettes associated with the organism is next to the accession number. Red vertical lines represent clades where the integrase is strongly linked to a chromosomal location within that genus. Blue vertical lines represent class 1, 2 and 3 integron branches. Adapted from Boucher et al., 2007.

Studies on the cassette metagenome have demonstrated that the high levels of gene cassette diversity within a habitat may exceed the total number of genes present in the chromosome/s of any one member of that community. In a study conducted by Michael *et al.* (2004), all gene cassettes were found to be distinct from chromosomal genes and only 17% of these matched to protein families with known functions. The mechanistic evidence of the horizontal gene transfer event/s that led to the flow of resistance genes between environmental and clinical strains often remains unknown (Forsberg *et al.*, 2012). More studies are necessary to establish a strong environment-clinic connection (Davies & Davies, 2010). However, rather than mass screening the cassette metagenome, an alternative and more narrow screening approach may prove to be more informative. For active gene transfer to occur, MRIs and CIs have to co-exist in the same cell. Yet evidence for the co-occurrence of MRIs and CIs, and a gene exchange between them in a single cell is rare.

1.8.3 attC sites: handles and markers for gene flow

MRIs contain cassette arrays with highly divergent *attC* sites (Partridge *et al.*, 2009). Structural recognition of *attC* simple sites by MRI is an important feature of integrin mediated site-specific recombination. *attC* sites vary between 57-bp to 141-bp and display sequence variability primarily in the region between the highly conserved LH and RH simple sites (Recchia & Hall, 1995). These two sites are part of two potential antiparallel recombination core sites, R'-L' and L''-R'', previously termed 1L-2L (LH simple site) and 2R-1R (RH simple site) respectively (Bouvier *et al.*, 2009; Recchia & Sherratt, 2002; Stokes *et al.*, 1997). Despite *attC* sequence variability, the structural conservation of these simple sites leads to strong conservation of *attC* secondary structure (Figure 1.11). Also, this structural conservation ensures that extrahelical bases (G and T) necessary for IntI1 binding are always placed into the correct plane for covalent binding between IntI1-*attC* (Bouvier *et al.*, 2009; MacDonald *et al.*, 2006).

Integrins, being able to recognise a wide diversity of *attC* sites, are therefore accumulators of the diverse range of genes associated with *attC* sites. Critical to their success is preservation of contextual information of the gene after capture. This includes the integrity of the ORF (so that nothing is deleted or frame-shifted) and its correct orientation so as to allow read-through from Pc. Thus structural recognition and strand polarity of *attC* sites is an important feature of this site-specific recombination system.

The diversity of *attC* sequences within MRIs suggests that each integrated gene cassette and its respective *attC* site may be traced back to a unique, environmental origin (Figure 1.12). CIs are hypothesized to be the reservoir for resistance genes. There is high homology of *attC* sites within the cassette arrays of clade specific CI. This homology means that *attC* sites can be used as lineage specific markers. For instance *Pseudomonas stutzeri* sites range from 73-bp to 89-bp (Holmes *et al.*, 2003; Vaisvila *et al.*, 2001), while *Vibrio attC* sites, also known as *Vibrio cholerae* repeats (VCRs), are characteristically just over 100-bp long (Rowe-Magnus *et al.*, 2002).

Despite the size and diversity of the cassette metagenome, there are only a few cases where a clinically relevant resistance gene has been traced back to its original environment source (Mazel, 2004) (Table 1.3). These resistance genes were part of the core genome in the source organism and are now seen in a new context in a sink organism. However even then, the resistance gene sequence identity and chromosomal context is not 100% identical. Note that integrons typically accumulate gene cassettes by *IntI* mediated site-specific recombination, however other processes can also lead to them acquiring genes (Verdet *et al.*, 2000). Instead of utilizing resistance genes as a marker to study gene flow, *attC* sites and their structural conservation are hypothesized to be a more reliable marker to study gene flow between the different integron classes.

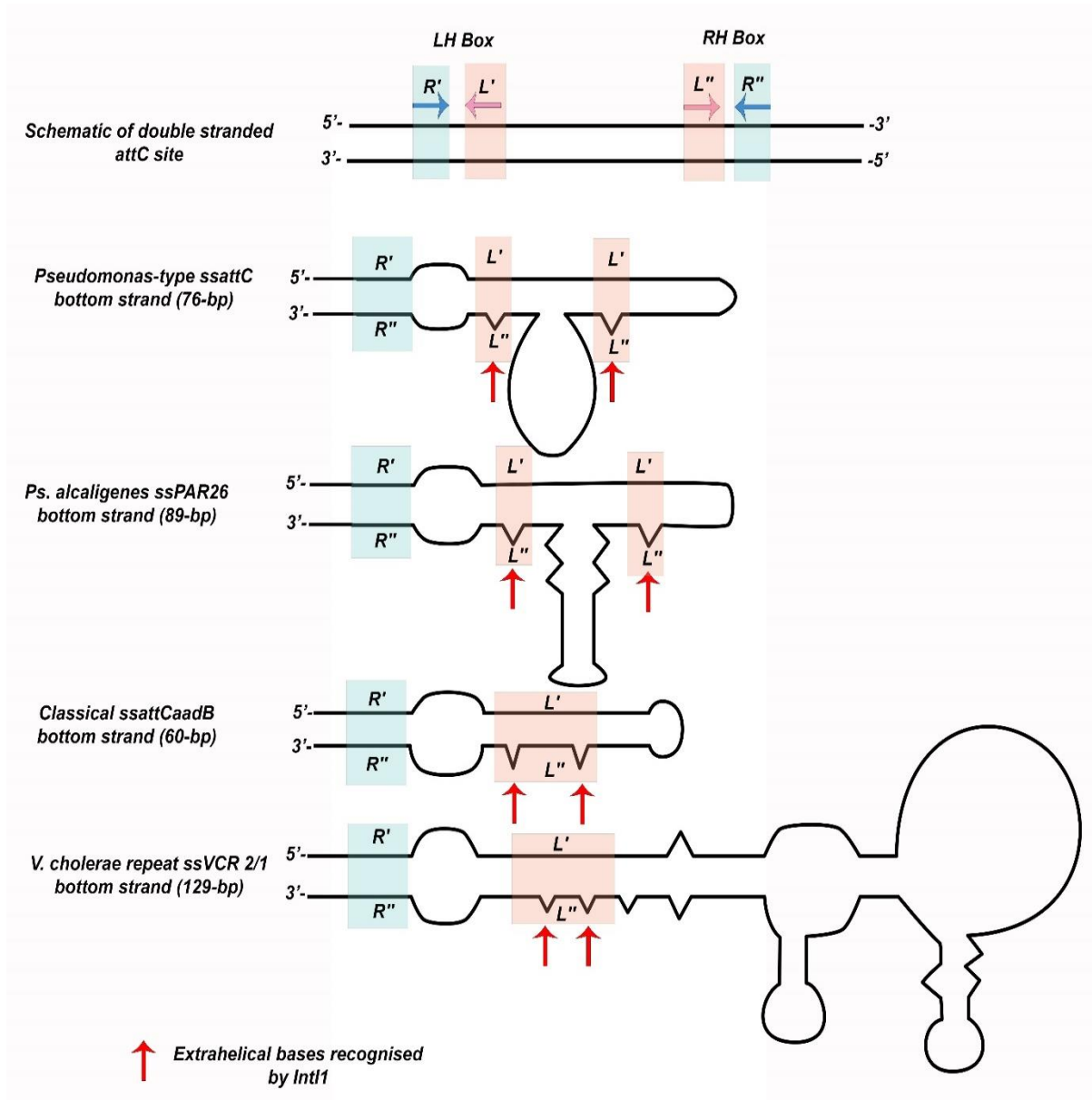


Figure 1.11: Schematic representation of secondary structures of single stranded *attC* sites. In all cases, the bottom strand is depicted. Despite the sequence variations and variable lengths between *Pseudomonas*-type *attC*, *Pseudomonas* PAR, classical and VCR sites, the RH and LH boxes are conserved throughout. Inverted repeats R' and R'' are indicated by blue boxes. Inverted repeats L' and L'' are indicated by pink boxes. Extrahelical bases recognized by *IntI1* are pointed out by red arrows. Adapted from Bouvier et al., 2009.

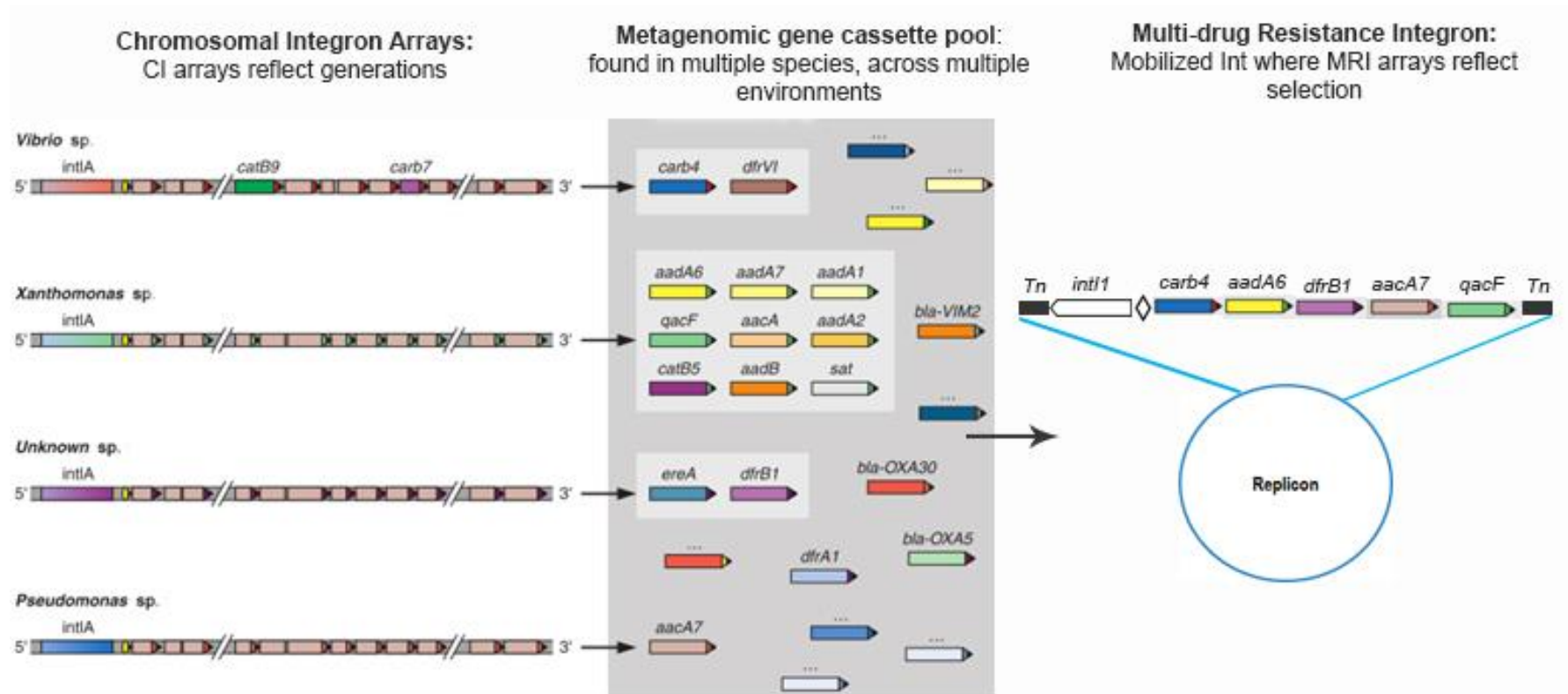


Figure 1.12: Chromosomal integrons- reservoirs for resistance cassettes.

Diverse *attC* sites are represented as triangles of different colour. Cassettes commonly observed in MRIs have *attC* sites that can be traced to homologous *attCs* within CI arrays. Adapted from Mazel, 2004.

1.8.4 *Pseudomonas*: Model organism for gene flow

Integrations are a paradigm for the movement of resistance genes between habitats and species. To model integron movement, exchange of gene cassettes across ecological and phylogenetic separation, and use *attC* sites as a marker for gene flow, there is a need for a bacterial genus that meets the following criteria: i) it must inhabit both environment and clinical environments, ii) it must contain both integron types, ie. MRIs and CI, iii) it must have both the capacity and opportunity to undergo HGT with opportunistic pathogens and iv) it must be found in habitats that intersect different environments with different levels of antibiotic selective pressures.

A feature of the bacterial genus *Pseudomonas* is the ecological flexibility seen across its member species. This flexibility is demonstrated by wide distribution in varying habitats (Spiers *et al.*, 2000), diverse carbon source range (Conrad *et al.*, 1979; Wang & Nomura, 2010), use of alternative electron acceptors (Lalucat *et al.*, 2006), intrinsic resistance mechanisms (Lister *et al.*, 2009) and a large and flexible genome and accessory genome ranging between 5.1 to 7-Mb (Mikkelsen *et al.*, 2011; Wiehlmann *et al.*, 2007). Members of the genus *Pseudomonas*, in particular *Ps. stutzeri*, contain multiple chromosomal integron arrays. *Pseudomonas* chromosomal integrons (PCIs) display a limited co-evolution within the *Pseudomonas* host lineage arrays (Holmes *et al.*, 2003; Vaisvila *et al.*, 2001) (Figure 1.13).

Table 1.3: Source and sink relationships of resistance genes. Adapted from Partridge, 2011.

Gene	Source	Sink	Mobilization by	Reference
<i>bla</i> _{SHV}	<i>Klebsiella pneumoniae</i>	<i>Enterobacteriaceae</i>	IS26	Ford & Avison (2004)
<i>bla</i> _{OXA-48}	<i>Shewanella spp.</i>	<i>Enterobacteriaceae</i>	IS1999	Aubert <i>et al.</i> (2006)
<i>bla</i> _{DHA-1}	<i>Morganella morganii</i>	<i>Salmonella enterica serovar enteritidis</i>	In6 and In7	Verdet <i>et al.</i> (2000)
<i>bla</i> _{CMY-2-like}	<i>Citrobacter freundii</i>	<i>Enterobacteriaceae</i>	?	Barlow & Hall (2002)
<i>bla</i> _{ACC-1}	<i>Hafnia alvei</i>	<i>Klebsiella pneumoniae</i> / <i>Enterobacteriaceae</i>	ISEcp1	Nadjar <i>et al.</i> (2000); Partridge (2007)
<i>bla</i> _{ACT-1}	<i>Enterobacter asburiae</i>	<i>Enterobacter spp.</i>	?	Rottman <i>et al.</i> (2002)
<i>bla</i> _{Fox}	<i>Aeromonas caviae</i>	<i>Aeromonas spp.</i> , <i>Pseudomonas spp.</i> & <i>Enterobacteriaceae</i>	?	Fosse <i>et al.</i> (2003); Maravić <i>et al.</i> (2013); Voolaid <i>et al.</i> (2013)
<i>bla</i> _{CTX-M-2}	<i>Kluyvera ascorbata</i>	<i>Enterobacteriaceae</i>	ISEcp1B	Lartigue <i>et al.</i> (2006)
<i>bla</i> _{CTX-M-3}	<i>Kluyvera ascorbata</i>	<i>Enterobacteriaceae</i>	Possibly ISEcp1	Rodríguez <i>et al.</i> (2004)
<i>bla</i> _{CTX-M-5}	<i>Kluyvera ascorbata</i>	<i>Enterobacteriaceae</i>	ISEcp1	Humeniuk <i>et al.</i> (2002)
<i>bla</i> _{CTX-M-8 Group}	<i>Kluyvera georgiana</i>	<i>Enterobacteriaceae</i>	ISEcp1-like & IS10-like	Poirel <i>et al.</i> (2002)
<i>bla</i> _{CTX-M-9 Group}	<i>Kluyvera georgiana</i>	<i>Kluyvera spp.</i>	IS3000, ISEcp1, IS903D, In60	Olson <i>et al.</i> (2005)
<i>bla</i> _{CTX-M-25}	<i>Kluyvera georgiana</i>	<i>Enterobacteriaceae</i>	ISEcp1 & ISEcp1B	Rodríguez <i>et al.</i> (2010)
<i>bla</i> _{OXA-119}	<i>Shewanella xiamenensis</i>	<i>Enterobacteriaceae</i>	ISShe2	Zong (2012)
<i>bla</i> _{TEM}	?	<i>Enterobacteriaceae</i>	Tn3, IS26	Bailey <i>et al.</i> (2011); Partridge & Hall (2005)
<i>qnrA</i>	<i>Shewanella algae</i>	<i>Enterobacteriaceae</i>	?	Poirel <i>et al.</i> (2005)
<i>qnr5</i>	<i>Vibrio splendidus</i>	<i>Enterobacteriaceae</i>	?	Cattoir <i>et al.</i> (2007)
<i>qnrB19</i>	<i>Vibrio splendidus</i> ?	<i>Enterobacteriaceae</i> & <i>Aeromonas spp.</i>	ISEcp1	Cattoir <i>et al.</i> (2008)

bla= β -lactamase resistance, *qnr*= quinolone resistance, ? = origin or mechanism of plasmid mobilization unknown.

However it should be noted that within the *Pseudomonas* bacterial lineage PCI distribution does not strictly follow the pattern of vertical inheritance. PCIs are overrepresented within the *Ps. stutzeri* species complex but less so in other Pseudomonads (Wilson, 2007). This bacterial clade contains a broad family of CIs where, closely related PCIs are observed within the same locus (*Ps. stutzeri* species complex). In comparison, unrelated PCIs are located within a different loci. This patchy PCI distribution among the various *Pseudomonas* species suggests that CIs are the result of multiple instances of acquisition by HGT from other bacterial genera, following the pattern of chromosomal drift.

Ps. aeruginosa and *Ps. putida* lack PCIs but frequently contain mobile class 1 integrons (Lee *et al.*, 2002; Davies & Davies, 2010). Chromosomal insertions of class 1 integron/transposon structures have also been noted in four distinct chromosomal loci including genomic islands in clinical *Ps. aeruginosa* isolates (Martinez *et al.*, 2012). Chromosomal acquisition of these MGE only adds to the complexity of the accessory genome of *Ps. aeruginosa* and contributes to the range of antibiotic resistance phenotypes and the overall spread of resistance between *Ps. aeruginosa* strains (Kung *et al.*, 2010).

Given that *Ps. aeruginosa* shares a common habitat, such as soil, with other non-pathogenic Pseudomonads, we can predict that HGT events between these would occur (Figure 1.14). To understand how genetic material crosses ecological and phylogenetic boundaries, the widespread and clinically relevant bacterial genus *Pseudomonas* is used as a model system for gene flow. Furthermore, it is important to understand HGT between clinical *Ps. aeruginosa* and the prevalent *Enterobacteriaceae* co-inhabitant.

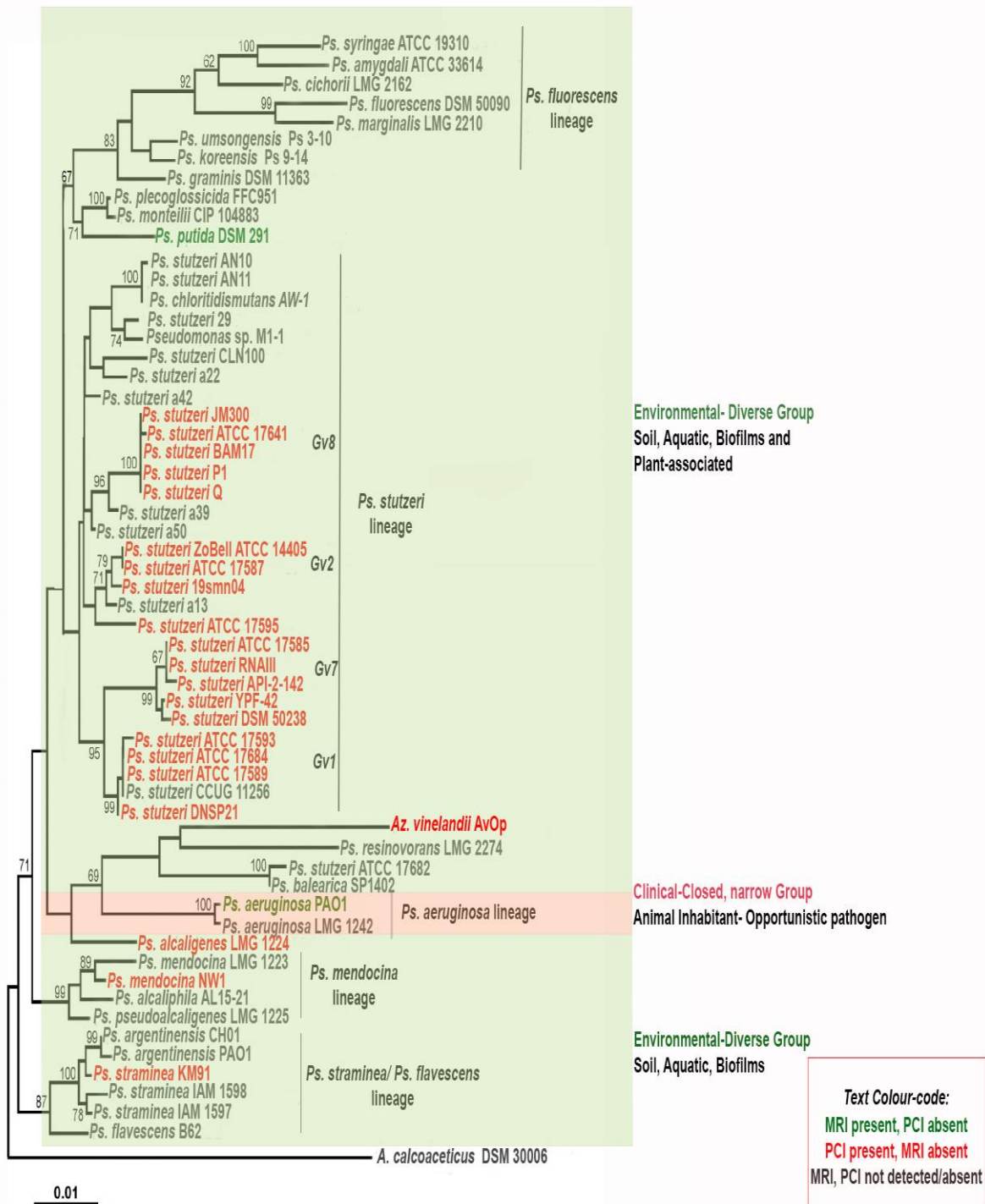


Figure 1.13: Distribution of chromosomal integrons across the *Pseudomonas* lineage. The tree was constructed from an alignment of 16S rDNA sequences spanning approximately 1300-bp, using maximum-likelihood and neighbour-joining analysis. All *Pseudomonas* spp. in which integrons are known to be either present or absent are indicated (see key). *Ps. aeruginosa* contains mobile integrons and lacks a PCI. Note – not all of the strains indicated above have been screened for the presence of integrons and are included here as representatives only. Adapted from Wilson, 2007.

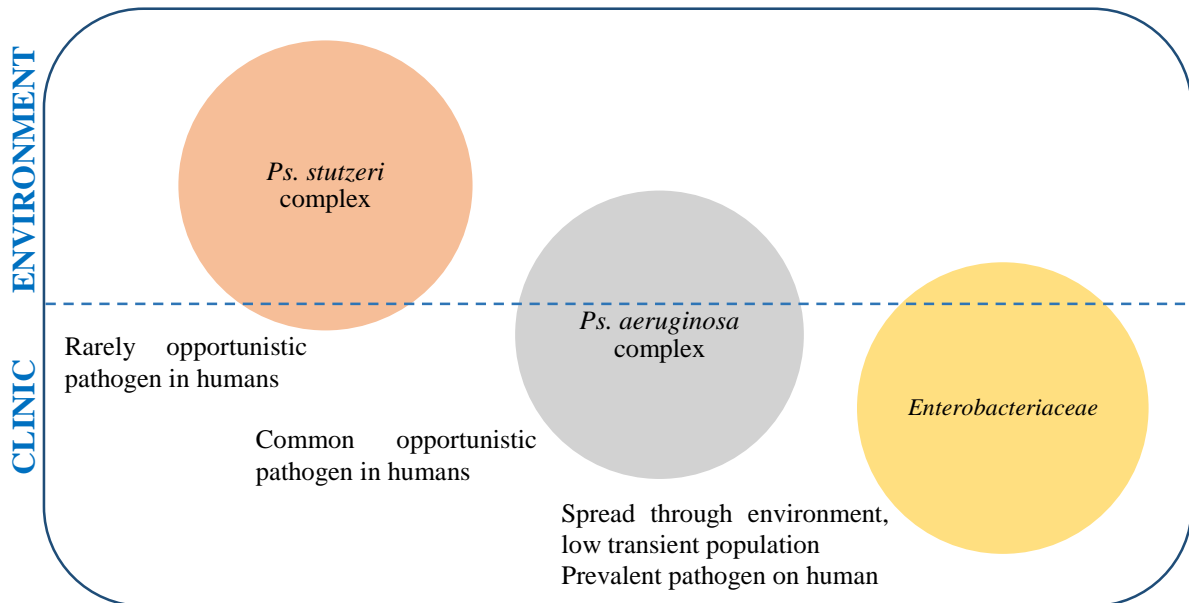


Figure 1.14: Sharing ecological niches is postulated to influence HGT between *Pseudomonads* and enteric organisms.

To illustrate this concept, consider a clinical environment where gene flow occurs between the primary inhabitants, two distinct *Pseudomonas* populations and an enteric population (Figure 1.15). Genetic exchange is more likely to occur between the *Pseudomonads* than between a *Pseudomonad* and an *Enterobacteriaceae*. This is because the *Pseudomonads* are more likely to occupy a common ecological habitat, such as a biofilm, and have commonalities at the molecular level in terms of codon usage and control elements for gene expression, as a result promoting gene flow between the *Pseudomonad* populations. Comparatively, while *Enterobacteriaceae* can inhabit the same ecosystem, they are phylogenetically distinct. These various features promote gene flow and are predicted to occur at a lower rate due to "incompatible" biological machineries, and may involve a subset of *Pseudomonas* species and a mode of horizontal gene transfer.

1.8.5 *IS1111-attC* elements target integron associated recombination sites

In characterizing CIs associated with *Pseudomonads*, Tetu & Holmes (2008) discovered a subfamily of insertion sequences, *IS1111-attC*, which appeared to be preferentially inserted within *Pseudomonas*-type *attC* sites. This site-specific targeting of *attC* sites is expected to significantly impact integrons, but more importantly provides additional support to the idea of using *attC* sites as markers for gene flow.

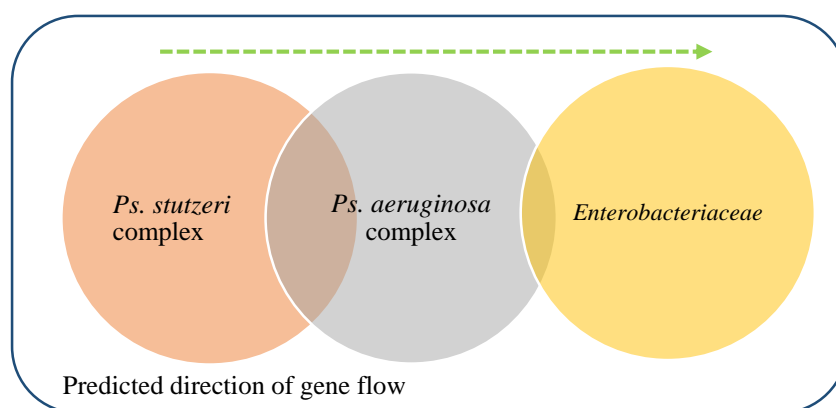


Figure 1.15: Gene flow between related organisms in comparison to that between non-related organisms.

Ecological and phylogenetic separation between *Pseudomonad* populations is minimal, thus gene flow between these two populations is predicted to readily occur. The enteric population shares the same ecological, antibiotic rich habitat as the clinical *Pseudomonas* population. Gene flow between these two populations is predicted to occur but at a low rate.

IS1111-*attC* elements were categorized as members of the IS110/IS492 family. This family of insertion sequences displays some key differences to other IS families. All members of the IS110/IS492 family either lack or have very small (<7-bp) terminal inverted repeats (IRs). Translocation of these ISs is site specific and does not generate a target site duplication, with the exception of IS492 which has been shown to generate 5-bp repeats (Mahillon & Chandler, 1998). Furthermore, IS110/IS492 recombinases (transposases) are not related to the site-specific recombinases of the λ integrase or resolvase/invertase families. Instead the transposase resembles a phosphoryltransferase similar to the DDE Transposase (Tpase), a site-specific invertase/transposase (Mahillon & Chandler, 1998; Tetu & Holmes, 2008). *ISPst6*, a member of the IS1111-*attC* subgroup (Tetu & Holmes, 2008), was first discovered in a PCI in *Ps. stutzeri* ATCC14405. This 1371-bp insertion sequence has a simple organization: a single open reading frame, encoding a putative transposase (347 amino acids), flanked by relatively long noncoding sequences and bounded by conserved sub-terminal inverted repeats. The noncoding regions are 53 and 240-bp at the right and left ends, respectively (Figure 1.16).

Importantly, *ISPst6* has apparent high target site specificity towards PCI *attC* sites. This target specificity allows these ISs to come into close contact with cassette encoded genes. Tetu and Holmes found *ISPst6* elements to be greatly over-represented in *Ps. stutzeri* genomovar 2 strains, but absent in *Ps. stutzeri* genomovar 8 strains.

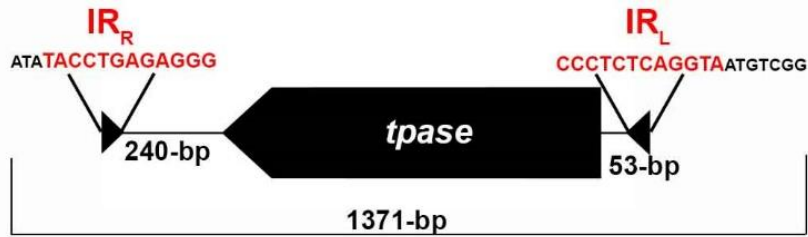


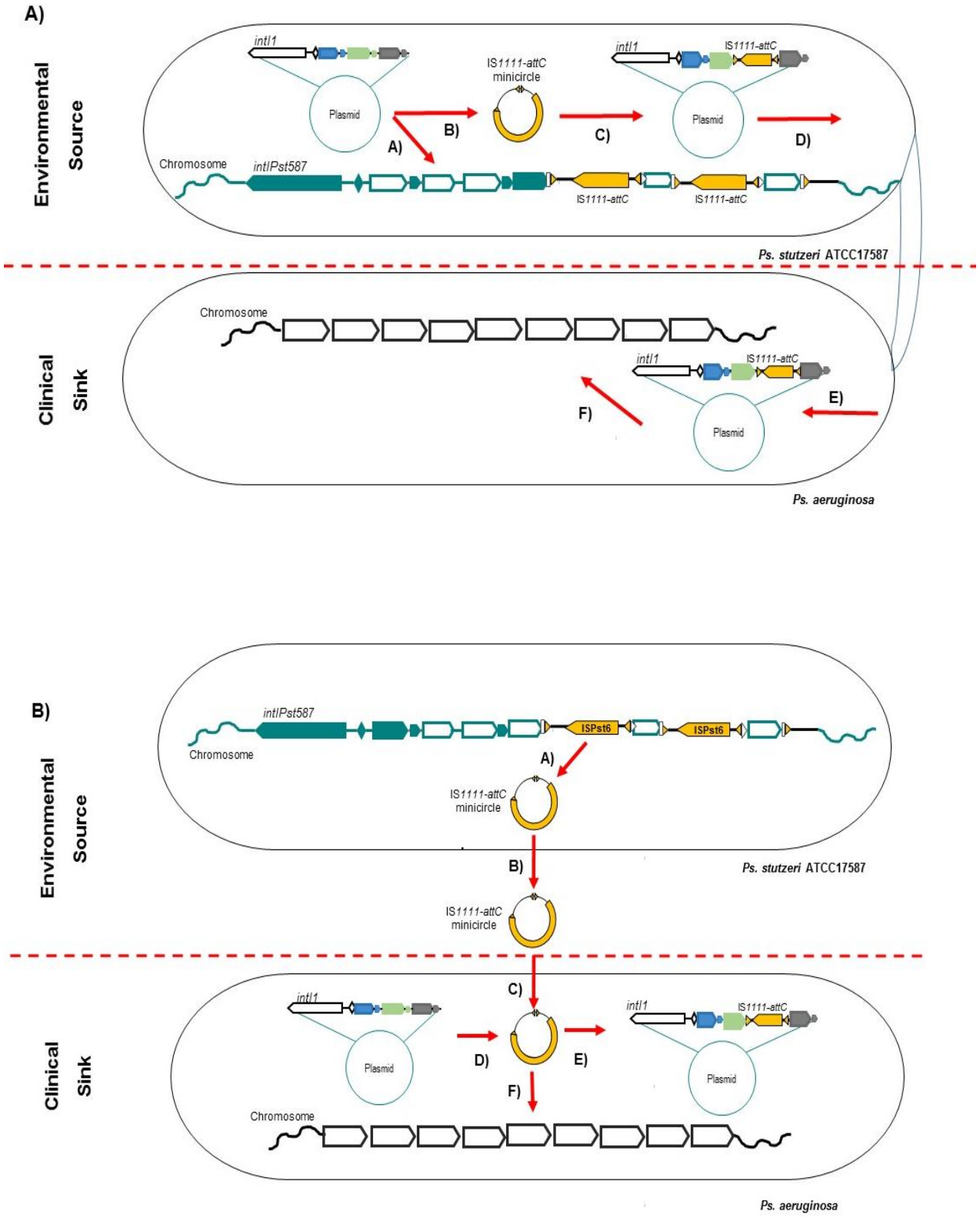
Figure 1.16: Schematic representation of *ISPst6*, a member of the *IS1111-attC* subgroup. The inverted repeat boundaries are indicated by red lettering.

An *ISPst6* footprint was observed in *Ps. stutzeri* genomovar 1 and 7 strains at the terminus of the array. Based on the distribution of *IS1111-attC* elements and the evolutionary radiation of the *Ps. stutzeri* species complex, Tetu and Holmes proposed that *IS1111-attC* elements undergo an infection-expansion-extinction cycle within the *Ps. stutzeri* complex.

At the time of the Tetu and Holmes study, *IS1111-attC* elements had been observed in nine instances of MRI (class 1 integrons). Of these, all but one, found in *Klebsiella pneumoniae*, were present within a *Pseudomonad*. In all *Ps. aeruginosa*-MRI cases, the *IS1111-attC* elements were embedded in resistance gene associated *attC*-sites found on class 1 integrons (Poirel *et al.*, 2005; Post & Hall, 2009). Thus, I postulate *attC*-targeted translocation of *IS1111-attC* elements can be exploited as a means of observing gene flow between different integron classes (*attC* sites) and *Pseudomonads* in different ecological habitats (Figure 1.17A & B).

Figure 1.17: Proposed models for *IS1111-attC* flow from an environmental to a clinical *Pseudomonad*.

- A) An MRI located on broad host range plasmid is acquired by an environmental *Pseudomonas* containing a *PCI+IS1111-attC*. The MRI acquires a) gene cassette and/or b) an *IS1111-attC* element. This c) newly acquired, *IS1111-attC* may be passed onto d) a clinical *Pseudomonas aeruginosa* or another *Pseudomonad* through conjugation. The plasmid + *IS1111-attC* may be e) stably maintained under antibiotic selective pressure or f) inserted into the chromosomal context.
- B) Under certain conditions, a chromosomally embedded *IS1111-attC* might excise as a) a minicircle and b) become transported out of the environmental *Pseudomonas*. c) Natural transformation favours the uptake of the IS minicircle into a new host such as *Pseudomonas aeruginosa*. If this new host already contains an d) MRI, the *IS1111-attC* element may insert into an available e) MI *attC* site. Alternatively, the IS element may insert into an f) unknown chromosomal loci.



1.9 Two convergent enzymes, one recombination site

The ability both of IntI and the IS1111-*attC* transposase to recognize *attC* target sites and integron platforms is an interesting example of convergent evolution. IS1111-*attC* site-specific insertion into CI and MRI *attC* sites, suggests that *attC* specificity is an ancestral feature of both the IS1111-*attC* transposase and IntI, heterologous proteins with different catalytic mechanisms. Based on amino acid sequence homology, IS110/IS482 transposases are related to the Piv/MooV family of DNA recombinases (Mahillon & Chandler, 1998). Briefly, Piv mediates site-specific inversion of a 2.1-kb chromosomal segment that encodes type 4 pilin in *Moraxella lacunata* and *Moraxella bovis* by targeting two identical 32-bp inverted repeats, *invL* and *invR*. Buchner *et al.* (2005) demonstrated the importance of the highly conserved amino acids D₉-E₅₉-D₁₀₁-D₁₀₄ for Piv catalytic activity. Notably, IS1111-*attC* transposase contains the conserved D-E-D-D motif (D₇-E₅₁-D₈₈-D₉₁), which is predicted to be essential for binding of IS1111-*attC* transposase to *attC* sites.

Recognition of the *attC* extrahelical bases G and T was linked to IntI recombination efficiency (MacDonald *et al.*, 2006). Thus, I hypothesise that like IntI integrase, IS1111-*attC* transposase also recognises the structure produced by single-stranded fold backs with extrahelical bases. It is possible that IS1111-*attC* transposase, like IntI, recognises and preferentially performs top-strand binding, having an equivalent site of attack. Alternatively, it could have a different site of attack and perform bottom-strand binding. Therefore studying transposase binding to an *attC* site may provide a paradigm for how single-stranded genetic material is recognized by convergent enzymes.

1.10 Aims

Understanding intragenomic rearrangements mediated by insertion sequences in bacterial genomes is especially relevant to the evolution of multidrug resistance. The complexity arises as mobile genetic elements do not ‘belong’ to a particular cell or lineage and have independent evolutionary trajectories when compared to phylogenetic trees. As a means of studying HGT between environmental and clinical pathogens, *Pseudomonas* is used in this study as a model organism. This study also explores the *IS1111-attC* subgroup as a model for this process in clinical, and environmental, non-host associated *Pseudomonads*.

Integrans and *IS1111-attC* elements together are postulated to comprise a model system for interaction between genetic elements. *IS1111-attC* elements have similar characteristics to gene cassettes; that is, they are predicted to form minicircle structures, and, upon insertion into an integron array, take up a linear form. *IS1111-attC* elements “obey” the same rules as gene cassettes by inserting site-specifically into the integron platform. This study explores the *IS1111-attC* family of insertion sequences as a model gene cassette for HGT processes. The broad aims of the study were:

- 1) To survey the distribution of mobile integrans, *Pseudomonas* chromosomal integrans and *IS1111-attC* elements across three distinct populations (clinical *Pseudomonadaceae*, non-pathogenic environmental *Pseudomonadaceae*, clinical enteric outgroup) and test for an association between IS presence/absence and bacterial genera.
- 2) To determine the mechanistic properties of *IS1111-attC* elements *in vivo* with altering genetic contexts.
- 3) To test *IS1111-attC* transposase *in vivo* binding to different target substrates and *attC* conformations (single vs double stranded DNA: *Pseudomonas*-type *attC*, *attC_{aadB}* and *IS1111-attC* minicircle junction).

Chapter 2 Materials and Methods

2.1 General Material

2.1.1 Media

All media (Table 2.1) were sterilised by autoclaving at 121 kPa for 15 min. Chemicals were purchased from Sigma-Aldrich or Ajax, and were of analytical grade. Plates were made by adding 1.5% agar. All media were made in RO water.

Table 2.1: Media composition

Media	Ingredients	Properties	Use
Luria-Bertani (LB) Broth	1% (w/v) casein peptone (tryptone) (Amyl Media, Australia), 0.5% (w/v) yeast extract (Amyl Media, Australia), 0.5% (w/v) NaCl. Adjust to 1 L with distilled water	pH 7.4	General purpose medium used to culture most strains in this study
Luria-Bertani (LBa) Agar	1% (w/v) casein peptone (tryptone) (Amyl Media, Australia), 0.5% (w/v) yeast extract (Amyl Media, Australia), 0.5% (w/v) NaCl. Adjust to 1 L with distilled water. Aliquot into 400 mL volumes and add 1.5% agar.	pH 7.4	General purpose medium used to culture most strains in this study.
LBa/Amp¹⁰⁰/IPTG/Xgal¹⁶⁰	Add 400 μ L of Amp ¹⁰⁰ stock, 400 μ L of 1 M IPTG stock, 200 μ L of 80 mg/mL X-gal stock to molten, cooled 400 mL LBa.	pH 7.4	Chemical transformation Blue-white selection
Pseudomonas Isolation Agar (PIA)	Bacteriological peptone 20.0 g (Amyl Media, Australia), MgCl ₂ 1.4 g, KSO ₄ 10.0 g, glycerol 10.0 g, Agar 15.0 g (Langdon & Co., Australia). Adjust to 1 L with distilled water	pH 7.0 \pm 0.2	Selective medium used for isolation of environmental and clinical Pseudomonads
Lysogeny Broth /0.2% Glucose (LB/Glu)	1% (w/v) casein peptone (tryptone), 0.5% (w/v) yeast extract, 0.5% (w/v) NaCl, 0.2% Glucose. Adjust to 1 L with distilled water	pH 7.4 sterilised using short glucose cycle, 115°C, 10 min, 101.5 kPa	Used to grow <i>E. coli</i> Rosetta2 cells transformed with pMAL vector

2.1.3 Solutions

All solutions were sterilised by autoclaving at 121°C, 101.5 kPa for 15 min unless specified otherwise. In cases where the solution could not be autoclaved, filtration was used for sterilisation, using a 0.20 µm filter (Millex, Germany). All chemicals were purchased from Sigma-Aldrich or Ajax, and were of analytical grade. All solutions were made in reverse osmosis (RO) water unless indicated otherwise. For recipes see Table 2.2.

2.1.4 Antibiotics

The antibiotics used in this work were purchased from Sigma-Aldrich. All antibiotics were filter sterilized using 0.22 µm filter (Millex, Germany) and stored at -20°C, except chemicals in non-aqueous solvents, these were not sterilised. Antibiotics were added to media after autoclaving from stock solutions.

Table 2.3: Antibiotics

Antibiotic	Abbreviation	Stock Concentration	Dissolvent
Ampicillin	Amp ¹⁰⁰	100 mg/mL	In MQ water
Carbenicillin	Carb ¹⁰⁰	100 mg/mL	In MQ water
Chloramphenicol	Cm ²⁵	25 mg/mL	In 100% Ethanol
Kanamycin	Km ⁵⁰	50 mg/mL	In MQ water

Table 2.2: Solutions

Name	Ingredients	Sterilization
1M Isopropyl-β-D-thiogalactopyranoside (IPTG) stock	Dissolve 2.38 g of IPTG in 8 mL of MQ water. Bring volumes to 10 mL. Aliquot into 1 mL volumes by filtration and store at -20°C.	Filtration
80 mg/mL 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-GAL) stock	Dissolve 80 mg of X-GAL in 1 mL DMSO. Filter sterilize and aliquot into 1 mL volume. Store at -20°C.	Filtration
10 M Ammonium acetate	Dissolve 385 g of ammonium acetate in 400 mL of RO. Adjust the volume to 500 mL with RO.	Filtration
1 M Tris- HCl pH 7.5 stock	Dissolve 121.14 g of Tris[hydroxymethyl]aminomethane hydrochloride in 700 ml of RO. Adjust pH to 7.5 with concentrated HCl. Bring volume to 1 L with RO.	Autoclave
1 M Tris-HCl pH 9.5	Dissolve 121.14 g of Tris[hydroxymethyl]aminomethane hydrochloride in 700 mL of RO. Adjust pH to 9.5 with concentrated 10 N NaOH. Bring volume to 1 L with RO.	Autoclave
1 M Tris-OH pH 11	Dissolve 121.14 g of tris[hydroxymethyl]aminomethane in 700 mL of RO. Check that pH is at 11. Bring volume to 1 L with RO.	Autoclave
100 mM Magnesium chloride stock (MgCl₂)	Dissolve 2.03 g of magnesium chloride in 80 mL of H ₂ O. Once dissolved, makeup up to 100 mL with H ₂ O.	Autoclave
0.5 M EDTA stock pH 8.0	Dissolve 186.12 g EDTA.Na ₂ .2H ₂ O in 800 mL RO. Adjust pH to 8.0 with 10 N NaOH. Adjust final volume to 1000 mL with RO.	Autoclave
1 M Boric acid	Dissolve 6.183 g boric acid in 80 mL RO. Adjust volume to 100 mL with RO.	
Elution buffer (EB)	5 mM Tris-HCl made from 1 M Tris-HCl stock. pH adjusted to pH 8.0 using 1 M Tris-base stock with pH 11.	Autoclave
TE buffer	10 mM Tris-HCl from 1 M Tris-HCl pH 7.5 stock, 1 mM EDTA from 0.5 M EDTA stock. Adjust to pH 8.0.	
10xTBE buffer	0.89 M Tris-HCl from 1 M Tris-HCl, 0.89 M Boric acid, 7 mM EDTA from 100 mM EDTA stock. Adjust to pH 8.0.	Autoclave
6X loading dye	0.25% (w/v) Bromophenol blue, 30% (w/v) Glycerol in MQ H ₂ O dissolved at 37°C with shaking at 200rpm until bromophenol blue is dissolved.	
Phosphate buffered saline (PBS) (pH 7.2-7.4)	Dissolve 8.0 g NaCl, 200 mg KCl, 1.44 g Na ₂ HPO ₄ O, 200 mg KH ₂ PO ₄ O in 900 mL RO. Adjust pH to 7.2-7.4 with concentrate HCl. Bring volume up to 1L with RO.	Autoclave
Ethidium bromide staining solution (0.5 μg/mL)	Add 50 μ L of 10 mg/mL Ethidium bromide solution (Amresco, USA) to 1-L of RO water. Mix gently.	

SOUTHERN HYBRIDIZATION REAGENTS		
Maleic acid buffer pH 7.5	Dissolve 10.73 g Maleic acid, 7.01 g NaCl in 600 mL MQ H ₂ O. Adjust pH to 7.5 with NaOH pellets. Finally, adjust the volume to 800 mL with MQ water.	Autoclave
Blocking solution	Dilute powdered (100x) Blocking Reagent (Roche) to 1x using Maleic acid buffer	Not sterilised, make fresh prior use
Saline sodium citrate (SSC) 20x pH 7.0 stock	Dissolve 175.3 g of NaCl and 88.2 g of tri-sodium citrate in 800 mL of MQ water. Adjust pH to 7.0 with 10 N solution of NaOH. Bring final volume to 1 L with MQ water.	
0.5x SSC (HIGH)	0.5x SSC from 20x SSC stock, 0.10% (w/v) SDS. Adjust the volume to 800 mL with sterilized MQ water.	
1x SSC (MEDIUM)	1x SSC from 20x SSC stock, 0.10% (w/v) SDS. Adjust the volume to 800 mL with sterilized MQ water.	
2x SSC (LOW)	2x SSC from 20x SSC stock, 0.10% SDS. Adjust the volume to 800 mL with sterilized MQ water.	
Washing buffer pH 7.5	Dissolve 10.73 g Maleic Acid, 7.01 g NaCl, Tween20 0.30% (w/v) in 600 mL MQ water. Adjust to pH 7.5 with NaOH pellets. Finally adjust the volume to 800 mL with sterilized MQ water.	Autoclave
Antibody solution	Anti-DIG-AP 1:1000 in blocking solution.	Not sterilized, make fresh prior use
Stripping solution	Dissolve 6.39 g NaOH and 0.8 g SDS in 800 mL MQ water.	
Detection buffer	Dissolve 2.34 g NaCl in 40 mL of 1 M Tris-HCl pH 9.5 stock. Bring volume to 400 mL in MQ water.	
PLASMID/ DNA PURIFICATION		
CTAB lysis buffer	50 mM Tris-HCl (pH 8.0), 10 mM EDTA, 1 M NaCl, 1% CTAB	
Cell lysis solution (plasmid preparation)	0.2 M NaOH, 1% SDS (prepared fresh from autoclaved 2 M NaOH and filter sterilised 10% SDS).	Not sterilised,
Precipitation solution	3 M potassium, 5 M acetate, (pH 4.8).	
1xPBS	Dissolve 8 g NaCl, 0.2 g KCl, 1.44 g Na ₂ PO ₄ , 0.24 g KH ₂ PO ₄ . Adjusted to 1 L distilled water. Adjust to pH 7.0.	Autoclave
Sodium acetate	3 M NaOAc, adjust to pH 4.8 with concentrated glacial acetic acid	
24:1 Chloroform/Isoamyl-alcohol	Mix 24 vol Chloroform to 1 vol Isopropanol and leave to stand overnight at room temperature.	Do not autoclave
Phenol/chloroform/isoamyl-alcohol	Solution comes pre-equilibrated with TE buffer at pH 8.0 (Ratio 25:24:1).	Ready-to-use
ROSETTA COMPETANCY		
KCM buffer	100 mM KCl, 30 mM CaCl ₂ , 50 mM MgCl ₂ . Store at -20°C.	Filter sterilized

Rosetta resuspension buffer	LB broth at pH 6.1 (autoclaved) supplemented with 10% PEG (3350 or 8000), 5% DMSO, 10 mM MgCl ₂ from 100 mM MgCl ₂ stock, 10 mM MgSO ₄ from 100 mM MgSO ₄ stock, 10% (w/v) glycerol. Store at 4°C.	Filter sterilized
AFFINITY CHROMATOGRAPHY (MBP COLUMN)		
100 mM Phenylmethylsulfonyl fluoride (PMSF)	Dissolve 17.42 mg of PMSF in DMSO to get a 100 mM solution. Filter sterilise and use immediately.	Filter sterilise
MBP column buffer	Combine 20 mM Tris-HCl from 1 M stock at pH 7.5, 200 mM NaCl from 5 M stock, 100 µL PMSF from 100 mM stock, and 1 mM DTT in 500 mL of sterile MQ water and add 100 µL of complete protease inhibitor cocktail (Sigma-Aldrich). Adjust pH to 7.4 and bring solution to 1 L.	Not sterilised.
20% Maltose solution	Dissolve 20 g of maltose in 100 mL RO. Sterilise and store at -20°C.	Filter sterilised.
MBP elution buffer	MBP column buffer supplemented with Maltose solution to get a final concentration of 50 mM maltose. Make up fresh just prior use.	Not sterilised.
SDS-PAGE		
SDS-PAGE 1x Tris glycine electrophoresis buffer	0.25 M Tris/HCl, 0.192 M glycine, 0.1% (w/v) SDS, adjust pH to 8.3. Store at 4°C.	Not sterilised
4x Resolving buffer (1.5 M)	Dissolve 181.7 g of Tris in 800 mL of RO, adjust pH to 8.8 with concentrated HCl acid. Bring solution to 1 L with RO.	Not sterilised
4x Stacking buffer (0.5 M)	Dissolve 60.6 g of Tris in 800 mL of RO, adjust pH to 6.8 with concentrated HCl. Bring solution to 1 L with RO.	Not sterilised.
10% SDS solution	Dissolve 10 g of SDS in 80 mL RO. Adjust pH to 7.2 with concentrated HCl and bring solution to 100 mL with RO. Store solution at RT.	Filter sterilise.
20% Ammonium persulfate (APS)	Dissolve 2 g ammonium persulfate in 10 mL of MQ water. Filter sterilise and aliquot into 1 mL volumes. Store at 4°C.	Filter sterilise.
SDS-PAGE 12% acrylamide gel (makes 2 resolving gel)	Combine 4 mL RO water, 3.3 mL 30% acrylamide/ bisacrylamide (19:1) (Sigma-Aldrich), 2.5 mL 4x Resolving buffer, 100 µL 10% SDS, 60 µL 20% APS, 4 µL TEMED.	Not sterilised, used fresh
SDS-PAGE 5% acrylamide gel (makes 2 stacking gel)	Combine 2.03 mL RO water, 0.44 mL 30% acrylamide/ bisacrylamide (19:1) (Sigma-Aldrich), 0.83 mL 4x stacking buffer, 33 µL 10% SDS, 20 µL 20% APS, 4 µL TEMED.	Not sterilised, used fresh
Colloidal coomassie blue staining solution	Combine 40% (v/v) methanol, 7% (v/v) glacial acetic acid, with 0.025% (w/v) Coomassie R250 Brilliant Blue in RO.	Not sterilised.
Coomassie high destain solution	Combine 40% (v/v) methanol, 7% (v/v) glacial acetic acid, in RO.	Not sterilised.
Coomassie low destain solution	Combine 10% (v/v) methanol, 7% (v/v) glacial acetic acid, in RO.	Not sterilised.
5x SDS gel loading buffer	0.125 mM Tris-Cl (pH 6.8), 100 mM DTT, 4% (w/v) SDS, 0.02% (w/v) Bromophenol blue,	Not sterilised.

	20% (v/v) glycerol. Aliquot in 1 mL volumes and store at -20°C.	
GEL SHIFT ASSAYS		
10x TBE (pH 8.3)	Dissolve 60.6 g Tris, 55.6 g Boric acid, 9.3 g EDTA in 800 mL RO. Adjust pH to 8.3 with concentrated HCl. Bring volume to 1 L.	Autoclave.
1xTBE (pH 8.3)	Make 1 L volume from 10x TBE pH 8.3 stock.	Autoclave.
10% Ammonium persulfate (APS)	Dissolve 1 g of ammonium persulfate in 10 mL MQ water. Filter sterilise, aliquot into 1 mL volumes and store at -20°C.	
EMSA gel (non-denaturing polyacrylamide gel, makes 1 gel)	Combine 9 mL of 40% (w/v) acrylamide/ bisacrylamide (19:1) (Sigma-Aldrich), 50.6 mL 1x TBE (pH 8.3), 400 µL 10% APS, 40 µL TEMED.	Not sterilised.
EMSA loading dye 6x	30% Ficoll, 0.1% Bromophenol blue, 0.25% Xylene cyanole in 50 mL MQ water.	Not sterilised.
EMSA running buffer (0.5x TBE, pH 8.3)	Make 1 L volume from 10x TBE pH 8.3 stock.	Not sterilised.
1 M Dithiothreitol (DTT)	Dissolve 1.55 g of DTT powder in 10 mL of deionized water. Filter sterilise and aliquot into 0.5 mL volumes. Store at -20°C.	Filter sterilised.
10x Reaction buffer (need 50 mL)	100 mM Hepes (pH 7.9), 100 mM NaCl and 50 mM MgCl ₂ in RO.	Autoclave.

2.1.5 Equipment

Details of equipment used in this study are shown in Table 2.4.

Table 2.4: Equipment list and respective supplier

Equipment Name	Manufacturer
Eppendorf BioPhotometer	Eppendorf (Germany)
Eppendorf Centrifuge 5424	Eppendorf (Germany)
Eppendorf Centrifuge 5810R	Eppendorf (Germany)
Gradient S Thermocycler	Eppendorf (Germany)
GS-800 Calibrated Densitometer	Bio-Rad (USA)
Junior Orbital Shaker	Lab-Line (USA)
Nanodrop 1000 Spectrophotometer	Thermo Scientific (USA)
Orbital Mixer Shaker	Ratek (Australia)
PowerPac HC	Bio-Rad (USA)
FLA 9000 (phosphoimager) Typhoon	GE Healthcare Life Sciences (USA)
Stratalinker® UV Crosslinker 1800	Stratagene (USA)
Voyager-DE™ STR Workstation	Applied Biosystems (USA)

2.2 Bacterial strains

All strains were grown in Luria- Bertani broth (LBb) or on LB agar. Clinical *Pseudomonads* and enteric isolates were grown at 37°C, and all other *Pseudomonads* including controls were grown at 30°C. A collection of 125 clinical *Ps. aeruginosa* strains (representing multiple patients, across three hospitals in Sydney) were isolated, identified and provided by collaborators around Sydney, Australia (Appendix Table A1). Also a group of 52 enteric isolates (outgroup) were provided by collaborators (Appendix Table A3). All control strains used in this study are outlined in Table 2.5. In summary, *Ps. stutzeri* strains (representing genomovars 6, 2, & 8) were used as control isolates for the presence of PCI and IS elements. *Ps. aeruginosa* isolates PAO1, MB 225158 Adelaide, and 216.2B were controls for IS1111-*attC*-like elements and mobile class 1 integrons. A conjugative plasmid carrying a class 1 integron (pR388) and a *Ps. fluorescens* (negative for class 1 integron, IS1111-*attC* and PCI) were also included in most analyses. Genomic DNA from *Thaura* species, containing a class

2 integron on transposon Tn7 was used as a class 2 control in preliminary Southern hybridization analysis. All strains were grown in LB broth or on LB agar, and were stored at -80°C in LB medium containing 15% (v/v) glycerol for long-term strain storage. The clinically relevant *Ps. aeruginosa* and gram negative unknowns were grown at 37°C , and all other Pseudomonads including controls were grown at 30°C

2.2.1 Clinical Isolates

All clinical *Ps. aeruginosa* strains were isolated by collaborators from the following hospitals: Sydney Adventist Hospital (sample set TS-) and Sydney Westmead Hospital (sample sets JIP- and SSP-). See Table A2 environmental (non-host associated) Pseudomonads were kindly provided by Mitch Brown from Westmead, Sydney (Table A2).

2.2.2 Isolation of Pseudomonads from domestic environments

An additional 68 environmental Pseudomonads (Table A2) were isolated from mop and sponge samples from 6 geographically distant domestic environments around Sydney, Australia (Figure 2.1). 1 g of mop thread or 1g of sponge was weighed out and resuspended by vigorous vortexing in 9 mL of 1x PBS. This suspension was serially diluted to range between 10^{-1} - 10^{-9} and 100 μL was spread plated per dilution onto PIA. Plates were first incubated overnight at 30°C , and then at room temperature for another 48 hrs. Colonies with different morphologies were selected, streaked onto fresh PIA plates and grown overnight at 30°C . Final purification was done on LB agar at 30°C , overnight.

Table 2.5 Strains as reference controls for integron and insertion sequence assays

Control strain	Shorthand nomenclature	Element			Reference
		MI	IS1111-attC	PCI	
<i>Ps. stutzeri</i> ATCC 14405	<i>Pst405</i>	-	+	+	Zobell & Upham, (1944)
<i>Ps. stutzeri</i> ATCC 17587	<i>Pst587</i>	-	+	+	Stanier <i>et al.</i> (1966)
<i>Ps. stutzeri</i> ATCC 17595	<i>Pst595</i>	-	+	+	Stanier <i>et al.</i> (1966)
<i>Ps. stutzeri</i> ATCC 17641	<i>Pst641</i>	-	-	+	Stanier <i>et al.</i> (1966)
<i>Ps. mendocina</i> NW1	<i>PsNW1</i>	-	+	+	Wilson (2007)
<i>Ps. stutzeri</i> RNAIII	<i>PstRNAIII</i>	-	+	+	Sikorski, pers. comm.
<i>Ps. straminea</i> KM91	<i>PsKM91</i>	-	+	+	McNicol (2002)
<i>Ps. stutzeri</i> DNSP21	<i>PstDNSP21</i>	-	-	+	Rius <i>et al.</i> (2001) (array and <i>attI</i> only, no <i>intIX</i>)
<i>Ps. stutzeri</i> strain Q	<i>PstQ</i>	-	-	+	Holmes <i>et al.</i> (2003)
<i>Ps. fluorescens</i> NCTC 7244	<i>Pfluro</i>	-	-	-	University of Sydney strain collection
<i>Ps. aeruginosa</i> NCTC 3756	<i>Paeru</i>	-	-	-	University of Sydney strain collection
<i>Ps. aeruginosa</i> PAO1	<i>PAO1</i>	-	Putative <i>tpase</i>	-	Stover <i>et al.</i> (2000); Winsor <i>et al.</i> (2011)
<i>Ps. aeruginosa</i> 216.2B	<i>216.2B</i>	+	-	-	Westmead Hospital strain collection
<i>E. coli</i> strain UB1637 pR388	<i>R388</i>	+	-	-	Datta & Hedges (1972)
<i>Thaura</i> sp. Tn7	<i>ColE Tn7</i>	<i>intI2</i>	-	-	L. Woolfended as per. comms.

Note: MIs, IS1111-attC elements and PCI in these strains were previously identified by Southern hybridisation screening, PCR screening and genome sequencing (Tetu, 2007; Wilson 2007; references as listed table).

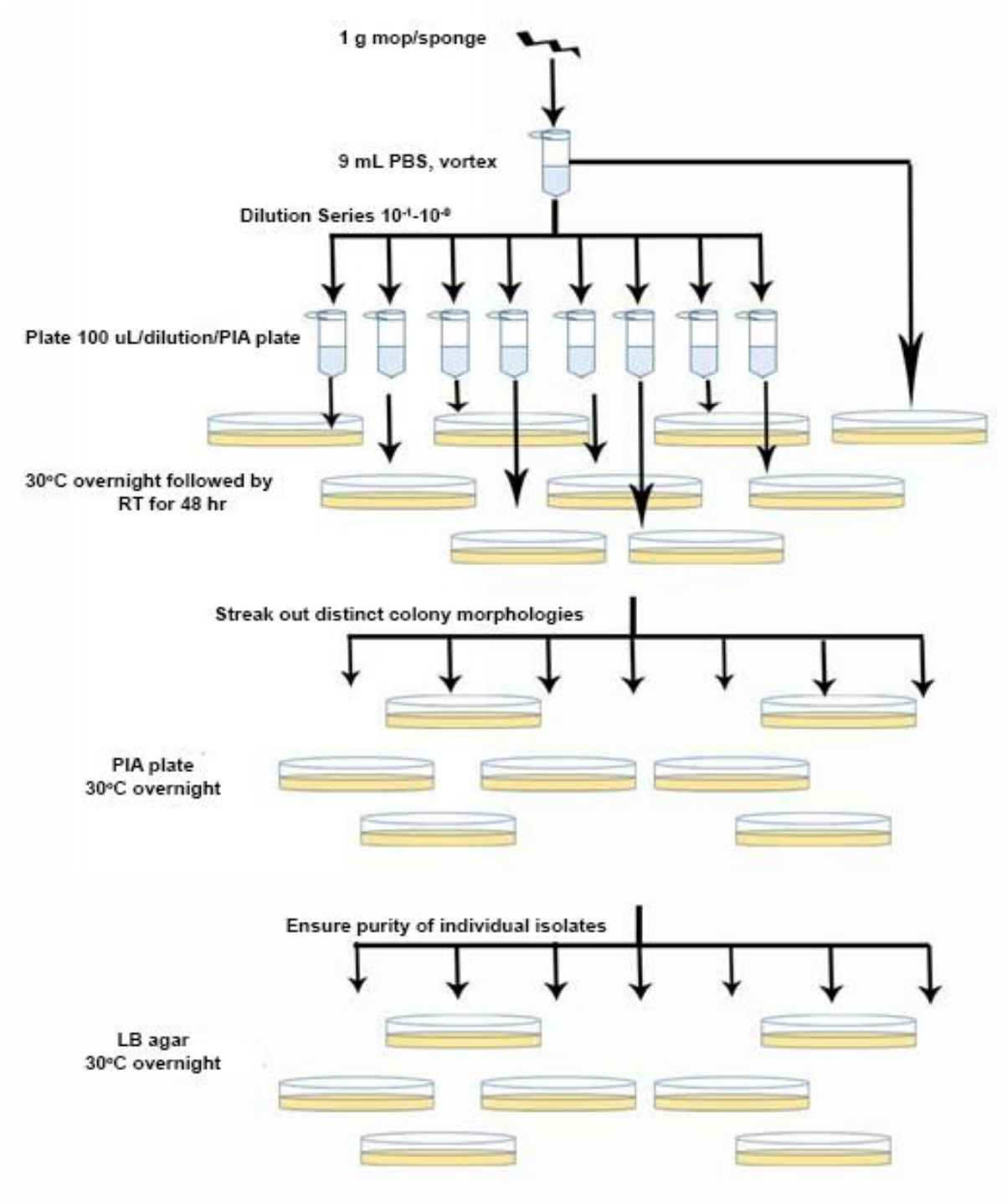


Figure 2.1: Workflow for the isolation of *Pseudomonads* from environmental samples. Weigh out 1 g of sample (mop thread or sponge) and suspend bacterial cell by vortexing vigorously in 1 mL of 1x PBS. Dilute the resuspended cell (10^{-1} - 10^{-9}) and spread plate 100 μ L of each dilution onto fresh PIA. For the first round of plating, incubate overnight at 30°C and then two days at room temperature (RT). For the second round, pick distinct colony morphologies and streak individually onto fresh PIA. To ensure purity for the last round streak plate onto LB agar and incubate at 30°C overnight. These plates were also used as master plates for subsequent experiments.

2.3 High molecular weight genomic DNA extraction

High molecular weight DNA was extracted from all isolates and control strains using a CTAB/phenol/chloroform method, modified from Sambrook and Russell (2001). Cells from 5 mL overnight LB cultures were pelleted using centrifugation, washed in several volumes of TE, and resuspended in 2 mL CTAB lysis buffer (50 mM Tris-HCL pH 8.0, 10 mM EDTA, 1 M NaCl, 1% CTAB). Lysozyme (1 mg/mL) and Proteinase K (ProK) (100 mg/mL) were added to the cell suspension, followed by an incubation at 37°C for 1-2 hrs, and extraction with one volume of chloroform:iso-amyl alcohol (25:1). After recovery of the aqueous phase, high molecular weight DNA was purified by a standard phenol/chloroform extraction and ethanol precipitation protocol as described previously (Sambrook & Russell, 2001). Recovered DNA was dissolved in 100 µL TE and stored at -20°C. DNA quantity and quality were determined using a Nanodrop spectrophotometer.

2.3.1 Ammonium acetate purification of low molecular weight PCR amplicons

PCR reactions were pooled to give a total of 400 µL ~1 µg of DNA. Equal volumes of phenol: chloroform: isoamyl alcohol (25:24:1) were added, inverted and then centrifuged for 3 min at 14200 rpm at room temperature. The top phase was extracted with an equal volume of chloroform. The centrifugation step was repeated and 0.2 volume of 10 M ammonium acetate and 2.5 volume of pure chilled ethanol were added to the top phase. This solution was inverted and chilled at 4°C for 30 min. The precipitate was recovered at 14200 rpm for 5 min at 4°C, and the pellet was washed with 70% ethanol. The centrifugation step was repeated again and the pellet was dried at 50°C for 5 min, before being dissolved in 1x TE (pH 8.0). The Protocol was amended from Sambrook and Russell, 2001.

2.3.2 Crude DNA extraction via cell boil-lysing

A sterile pipette tip was used to pick an individual colony, which was then resuspended in 100 µL of 100x DMSO stock. The resuspension was boiled at 99°C for 15 min to lyse bacterial cells and centrifuged at 14200 rpm for 5 min. The supernatant was removed and 1xTE was added. The DNA was stored at -20°C till subsequent use.

2.4 Polymerase Chain Reaction

All PCR amplification mixtures contained the reagents listed in Table 2.6, unless otherwise indicated. PCR reactions were performed in a total volume of 25 μ L unless otherwise indicated.

The PCR cycling conditions used varied depending on the primers (Table 2.7) used and are indicated for each assay (Table 2.8). The thermocycler used was a Gradient S Thermocycler (Eppendorf). Amplification products were separated using agarose gel electrophoresis.

Table 2.6: PCR reagents and concentration

Reagent	Concentration	Manufacturer
Primer	50 pmol per primer	IDT, USA
dNTP mix	200 nM per dNTP	New England Biolabs
MgCl ₂	2 mM	New England Biolabs
1x Thermopol® Buffer	20 mM Tris-HCl, 10 mM (NH ₄) ₂ SO ₄ , 10 mM KCl, 2 mM MgSO ₄ , 0.1% Triton® X-100, pH 8.8	New England Biolabs
Taq DNA polymerase	1 U	New England Biolabs
DNA	10-100 ng/ μ l	N/A

2.5 Agarose gel electrophoresis

Agarose gel electrophoresis was used to determine the size and concentration of all DNA recovered by direct extraction or amplification by PCR. Molecular grade, DNase and RNase free agarose (Bioline, USA) was used at concentrations between 1%-2% in 0.5xTBE buffer. All samples were run in a horizontal electrophoresis unit at 1-5 V/cm. 0.5 μ l of GelRed™ (10000x, Biotium, Ca) was used per 50 mL 0.5x TBE gel volume. Alternatively gels were post stained for 15 min in ethidium bromide staining solution. Fractioned DNA was then visualised under UV-light and an image captured using the Gel Doc™ XR+ Bio-Rad system (Bio-Rad, USA).

Table 2.7: Primers

Primer	Target	Sequence (5'-3')	Reference
CHAPTER 3			
BOX1AR	BOX repeats	CTACGGCAAGGCGACGCTGACG	Louws <i>et al.</i> (1994)
f27	16S rRNA DNA	AGAGTTTGATCMTGGCTCAG	Weisburg <i>et al.</i> (1991)
r1492	16S rRNA DNA	TACGGYTACCTTGTTACGACTT	Weisburg <i>et al.</i> (1991)
LW18	<i>intI</i> Q PstQ (AY129392)	TATCGATCCAAAGGACTCAGC	L. Woolfended as per com.
LW19	<i>intI</i> Q PstQ (AY129392)	CATGTTGCCTCGCAGTCTGG	L. Woolfended as per com.
LW14	<i>intI</i> 1 <i>E. coli</i> pR388 In3 (U12441.2)	CTTTGTTTTAGGGCGACTGC	L. Woolfended as per com.
LW15	<i>intI</i> 1 <i>E. coli</i> pR388 In3 (U12441.2)	ATGCCCGTTCCATACAGAAGC	L. Woolfended as per com.
ST87	<i>Pst405/Pst587 ISPst6</i>	ATATGGACTCTCCCCACAAG	Tetu & Holmes (2008)
ST117	<i>Pst405/Pst587 ISPst6</i>	TAATGGACTCTCCCCGCACC	Tetu & Holmes (2008)
ST8	<i>PstDNSP21 attC</i>	TGGCAGAAAAGTGTGATCTC	Tetu (2007)
ST9	<i>PstDNSP21 attC</i>	GAGTTATCTGCGTAAATATC	Tetu (2007)
ST10	<i>PsNW1 attC</i> (EF648216)	AAATGCACTCTAACAAGTCG	Tetu (2007)
ST11	<i>PsNW1 attC</i> (EF648216)	AGGGTCATGAGCATTACCTC	Tetu (2007)
ST12	<i>PstKM91 attC</i>	TTCAGGTTGAGTCGCTTACC	Tetu (2007)
ST13	<i>PstKM91 attC</i>	GATATTTTTTCATAAAGGTTTCTTA	Tetu (2007)
HS915	<i>intI</i> 1 gene	CGTGCCGTGATCGAAATCCAG	Martinez <i>et al.</i> (2013)
HS916	<i>intI</i> 1 gene	TTCGTGCCTTCATCCGTTTCC	Martinez <i>et al.</i> (2013)
CHAPTER 4			
ST96	BGC093 flanking DNA in <i>Pst587</i>	TTTCACGAGTATTCCCCGAGG	Tetu & Holmes (2008)
ST139	BGC093 flanking DNA in <i>Pst587</i>	AATATGACACCAAGCACTCC	Tetu & Holmes (2008)
MZ09	<i>attC_{aadB}</i> flanking DNA in pUS21	GCTTTCAGGTCGCGATATGC	This study
MZ10	<i>attC_{aadB}</i> flanking DNA in pUS21	TTCTGCAGAAGCTTGCGC	This study
AJH17	<i>PstQ attC</i> sites and array	CCCAGYGARCGARGYGAGCG	Holmes <i>et al.</i> (2003)
AJH27	<i>PstQ attC</i> sites and array	GGCTGAAGCCVGGCCCTTARC	Holmes <i>et al.</i> (2003)
NVC15	pBBR1MCS2vector backbone	GCTTCCGGCTCGTATGTTGTGTGG	N. Coleman as per com.
NVC16	pBBR1MCS2vector backbone	GCAAGGCGATTAAGTTGGGTAACG	N. Coleman as per com.
ST89	<i>ISPst6</i> minicircle junction	GAACGCTTGGCCCCYGTGC	Tetu & Holmes (2008)
ST123	<i>ISPst6</i> minicircle junction	CGGCAACCTGGTAAACGGAC	Tetu & Holmes (2008)
CHAPTER 5			
MZ03	start of <i>tpase</i> in <i>Pst405</i>	GCGGGATCCATGAAACGCATAGCGATTG (<i>Bam</i> HI site)	This study
MZ04	end of <i>tpase Pst405</i>	CCGCTGCAGTCAGGCGGGCTTTGTGC (<i>Pst</i> I site)	This study
ELF21	pMAL-c2X backbone	GTCAGACTGTTCGATGAAGCCC	E. Liew as per com.
ELF22	pMAL-c2X backbone	GCAACTGTTGGGAAGGGCGAT	E. Liew as per com.

Table 2.8: Details of PCR thermocycling protocols

Primer combinations	Target (expected size)	Thermocycling protocol
CHAPTER 3		
BOXA1R	<i>rep</i> regions (Variable)	94°C x 5 min, [94°C x 30 sec, 53°C x 30 sec, 72°C x 8 min] x 35 cycles, 72°C x 5 min x 1 cycle.
Southern hybridization: probe generation		
LW18/LW19	<i>PstQ intIPstQ</i> (500-bp)	94°C for 5 min, [94°C 30 sec, 56°C 30 sec, 72°C 1 min] x 30 cycles, 72°C for 5 min x 1 cycle
LW14/LW15	<i>E. coli</i> pR388 In3 <i>intII-attII</i> (551-bp)	94°C for 5 min, [94°C 30 sec, 58°C 30 sec, 72°C 1 min] x 30 cycles, 72°C for 5 min x 1 cycle
ST87/ST117	Intact <i>ISPst6</i> (1340-bp)	94°C for 5 min, [94°C 30 sec, 57°C 30 sec, 72°C 2.5 min] x 30 cycles, 72°C for 5 min x 1 cycle
ST8/9 probe	<i>PstDNSP21</i> (100-bp)	94°C for 5 min, [94°C 30 sec, 52°C 30 sec, 72°C 20 sec] x 30 cycles, 72°C for 5 min x 1 cycle
ST10/ST11	<i>PsNW1 attC</i> (100-bp)	94°C for 5 min, [94°C 30 sec, 60°C 30 sec, 72°C 20 sec] x 30 cycles, 72°C for 5 min x 1 cycle
ST12/ST13	<i>PstKM91 attC</i> (100-bp)	94°C for 5 min, [94°C 30 sec, 55°C 30 sec, 72°C 20 sec] x 30 cycles, 72°C for 5 min x 1 cycle
PCR screening for MGE		
LW14/LW15	<i>intII-attII</i> (551-bp)	94°C for 5 min, [94°C for 30 sec, 57°C for 30 sec, 72°C for 45 sec] x 30 cycles, 72°C for 5 min x 1 cycle
HS915/HS916	<i>intII</i> (371-bp)	94°C for 5 min, [94°C for 30 sec, 62°C for 30 sec, 72°C for 30 sec] x 30 cycles, 72°C for 5 min x 1 cycle
ST87/ST117	<i>ISPst6</i> (1346-bp)	94°C for 5 min, [94°C for 30 sec, 60°C for 30 sec, 72°C for 2.5 min] x 30 cycles, 72°C for 5 min x 1 cycle
ST96/ST139	BGC093 <i>attC</i> with flank (219-bp)	94°C 5 min, [94°C 30 sec, 60°C 30 sec, 72°C 20 sec] x 30 cycles, 72°C for 5 min x 1 cycle
CHAPTER 4		
MZ09/MZ10	<i>attC_{aadB}</i> with flanking DNA in pUS21	94°C 5 min, [94°C 30 sec, 57°C 30 sec, 72°C 10 sec] x 30 cycles, 72°C for 5 min x 1 cycle
NVC15/NVC16	Transformation screen	94°C for 5 min, [94°C 30 sec, 60°C 30 sec, 72°C 2 min] x 30 cycles, 72°C for 5 min x 1 cycle
AJH17/AJH27	<i>PstQ attC</i> sites and array (variable)	94°C for 5 min, [94°C 30 sec, 65°C 30 sec, 72°C 2 min] x 30 cycles, 72°C for 5 min x 1 cycle
NVC15/NVC16	pBBR1MCS2 vector backbone (variable)	94°C for 5 min, [94°C 30 sec, 55°C 30 sec, 72°C 2 min] x 30 cycles, 72°C for 5 min x 1 cycle
ST89/ST123	<i>ISPst6</i> minicircle junction (258-bp)	94°C for 5 min, [94°C 30 sec, 55°C 30 sec, 72°C 2 min] x 30 cycles, 72°C for 5 min x 1 cycle
CHAPTER 5		
MZ03/MZ04*	<i>Ps. stutzeri</i> ATCC 14405 <i>tpase</i>	98°C 2 min, [98°C 10 sec, 65°C 30 sec, 72°C for 30 sec] x 35 cycles, final 72°C for 5 min x 1 cycle
MZ03/ELF22	pMAL- <i>tpase</i> junction	94°C for 5 min, [94°C 30 sec, 55°C 30 sec, 72°C 2 min] x 30 cycles, 72°C for 5 min x 1 cycle
MZ04/ELF21	pMAL- <i>tpase</i> junction	94°C for 5 min, [94°C 30 sec, 55°C 30 sec, 72°C 2 min] x 30 cycles, 72°C for 5 min x 1 cycle

*Phusion™ High fidelity polymerase was used.

2.6 Chemical transformation of bacterial cells

See Table 2.9 below for information about the *E. coli* expression systems used in this study.

Table 2.9: *E. coli* expression systems

Strain	Shorthand	Use	Description	Reference
<i>E. coli</i> JM109	<i>JM109</i>	Used as a cloning expression system for pGEM-T Easy vector system	<i>recA1 supE44 endA1 hsdR17 gyrA96 relA1 thi Δ(lac-proAB)F9[traD36 proAB+ lacIq lacZ ΔM15];Gm^S, GFP⁻, Lac⁺</i>	Yanisch-Perron <i>et al.</i> (1985)
<i>E. coli</i> Rosetta2 (DE3)	<i>Rosetta</i>	Used for MBP-fusion protein expression	<i>F⁻ ompT hsdS_B(r_B⁻ m_B⁻) gal dcm (DE3) pLysSRARE2 (Cam^R)</i>	Novagen, Madison, WI
<i>E. coli</i> Epi300TM	<i>Epi300</i>	Used for fosmid library construction	<i>F⁻ mcrA Δ(mrr-hsdRMS-mcrBC) Φ80dlacZΔM15 ΔlacX74 recA1 endA1 araD139 Δ(ara, leu)7697 galU galK λ rpsL (Str^R) nupG trfA dhfr</i>	Epicentre, Madison, WI

2.6.1 *E. coli* JM109 cells

Fresh LB/Amp¹⁰⁰/IPTG/X-Gal plates were prepared and equilibrated to 37°C. 2 μL of ligation reaction was added to 50 μL of thawed *E. coli* JM109 cells on ice. The tube was gently flicked to mix and placed on ice for 20 min. The cells were heat-shocked for 45 sec in a water bath at 42°C. Tubes were immediately returned to ice for 2 min. 200 μL room-temperature LB medium was added to cells and incubated for 1.5 hrs at 37°C with shaking (~200 rpm). The transformation mix was then plated onto duplicate LB/Amp¹⁰⁰/IPTG/X-Gal plates and incubated overnight (16–24 hrs) at 37°C. Ligation and transformation efficiencies were assessed with reference to uncut vector, cut/re-ligated vector as appropriate. Individual colonies were picked based on blue-white selection and screened by boil-lysing DNA and using universal primer pair M13F (5'-GTAAAACGACGGCCAGT-3') and M13R (5'-AACAGCTATGACCATG-3'). Screening PCR conditions were as following; 1 cycle at 94°C 5 min, then 30 cycles of [94°C for 1 min, 55°C for 30 sec, 72°C for x seconds], and one final round of extension at 72°C for 5 min. Note x sec is determined by expected product size, i.e. 1 min/ 1-kb.

2.6.2 Making *E. coli* Rosetta2 culture stocks for transformation experiments

E. coli Rosetta2 cells were grown overnight in 10mL LB at 37°C with shaking at 200 rpm. Overnight culture was diluted 1 in 100 into fresh LB and grown to $\sim OD_{600nm} = 0.4$ at 37°C, 200 rpm. Cells were pelleted at 4000 rpm at 4°C and resuspended on ice in 1/20 volume with pre-chilled Rosetta resuspension buffer. Resuspended cells were aliquoted into 50 or 100 μ L volumes, snap frozen in liquid N₂ and stored at -80°C.

2.6.3 Transformation of *E. coli* Rosetta2 with the pMAL and/or pMBPTpase plasmid

50 μ L aliquots of chemically competent *E. coli* Rosetta2 cells were thawed on ice and 50 μ L KCM solution was added. 2 μ L of plasmid (pMAL or/and pMBPTpase) was added and mixed by flicking tube. The transformation reaction was incubated on ice for 20 min and heat shocked at 42°C for 90 sec. 200 μ L room temperature LB broth was added immediately thereafter and cells were recovered at 37°C for 2 hrs prior to plating onto LB agar containing carbenicillin (100 μ g/mL) and 2% glucose.

2.7 Electrical transformation of bacterial cells

2.7.1 Making *Ps. stutzeri* st. *Q* (*PstQ*) and *Ps. stutzeri* ATCC17641 (*Pst641*) culture stocks for electroporation experiments

PstQ and *Pst641* were streaked onto LB agar (no antibiotic) and grown at 30°C for 2 days. 50 mL LB broth was inoculated with single colonies of *PstQ/Pst641* and grown overnight at 30°C, 200 rpm. 1 L LB broth was inoculated with 1:100 overnight culture and cells were grown at 30°C, 200 rpm to an OD ~ 1.0 . The culture was pelleted in pre-cooled centrifuge bottles by spinning at 7500 rpm, 4°C, for 15 min and kept cool from this point onwards. The supernatant was discarded and the cell pellet washed in pre-chilled 10% Sucrose/1 mM MgSO₄, 1 mM HEPES solution. The centrifuge cycle and wash step were repeated two more times. Finally the cells were suspended in 2-3 mL of the solution above and aliquoted into 110 μ L stocks prior to storage at -80°C.

2.7.2 Electro-transformation of *PstQ* and *Pst641*

Electrocompetent *PstQ* and *Pst641* cells (50 μ L) were thawed on ice and 1 μ g of plasmid was added to the cells. The cells were mixed before transfer to pre-chilled electroporation cuvettes. The cells were electroporated at 200 Ohms, 2.5 Volts and a capacity of 25 μ FD.

Immediately after, 750 μ L LB broth was added to the cuvette, cells were mixed gently by pipetting and the mixture was transferred to a 1.5 mL microcentrifuge tube. The cells were incubated for 2 hrs at 30⁰C, 200 rpm. Cells were then plated onto LB agar supplemented with kanamycin at 50 μ g/ml (LB/Km⁵⁰) and incubated at 30⁰C for 4 days.

2.8 Plasmid miniprep-Alkaline lysis

E. coli JM109 cells containing appropriate plasmid constructs (Table 2.10) were grown in 50 mL LB/Km⁵⁰ overnight at 37⁰C. Cells were pelleted and washed in 2 mL TE and lysed in 2 mL of cell lysis solution at room temperature for 5 min. 3 mL of ice-cold precipitation solution was thoroughly mixed in and the total incubated on ice for 15 min. The supernatant was recovered after 15 min centrifugation at 4000 rpm and an equal volume of isopropanol was added and incubated on ice for 15 min. The sample was spun at 4000 rpm for 15 min. 5mL of 70% cold ethanol was added to the recovered pellet, and left to stand for 5 min at room temperature to ensure a maximum plasmid yield recovery.

Table 2.10 Plasmids used/generated

Plasmid	Backbone	Use	Size	Properties	Reference
pGEM-T Easy	pUC	General purpose cloning	3.0-kb	pBR322 <i>ori</i> , Amp ^R ; contains multiple cloning site within <i>lacZ</i> fragment allowing for blue/white selection	Promega Corporation
pUS21	pUC19	<i>aadB-attC</i> template	2.8-kb	pBR322 <i>ori</i> , Amp ^R ; amplified <i>aadB</i> 59-be (105-bp) at <i>XbaI/PstI</i> site in pUC19	Coleman & Holmes (2005)
pMAL-c2	pUC	MBP T _{pase} expression	6.6-kb	pBR322 <i>ori</i> , Amp ^R , <i>malE</i> gene with multiple cloning site, <i>Factor Xa</i> cleavage site	New England Biolabs
pPsattC1/ pPsattC2	pBBR1MCS2 Kovach <i>et al.</i> (1995)	<i>Pseudomonas</i> -type <i>attC</i> trap	5.35-kb	<i>mod</i> , <i>rep</i> , Km ^R , <i>Ps.</i> - like <i>attC</i> site (219-bp) cloned into <i>SmaI</i> site. In pPsattC1 the <i>attC</i> site is found in the opposite orientation to the <i>ori/rep</i> gene. In pPsattC2, the <i>attC</i> site is in the same orientation as the <i>ori/rep</i>	This study
pPsattC1-IS		ISP _{st6} containing trap	6.64-kb	<i>mod</i> , <i>rep</i> , Km ^R , <i>Ps.</i> - like <i>attC</i> site (219-bp) cloned into <i>SmaI</i> site with a translocated ISP _{st6} element	This study
paadB- attC1/ paadB-attC2		Classical <i>attC_{aadB}</i> trap	5.24-kb	<i>mod</i> , <i>rep</i> , Km ^R , classical <i>attC</i> site (100-bp) cloned into <i>SmaI</i> site. In paadB-attC1, the <i>attC</i> site is found in the appositive orientation to the <i>ori/rep</i> gene. In paadB-attC2, the <i>attC</i> site is in the same orientation as the <i>ori/rep</i>	This study

2.9 Isolation of genomic DNA for Fosmid library construction

The ISOLATE Genomic DNA Mini Kit from Bioline was used to extract high quality genomic DNA with amendments to the manufacturers protocol. Changes were as per figure 2.2 and the following; the culture was grown overnight at 30°C for environmental and 37°C for clinical isolates in 4 x 10 mL volumes (40 mL total). Cells were pelleted at 4000 rpm for 20 min at room temperature. Cell pellets were washed in 0.9% NaCl and the spin cycle was repeated. 800 µL Lysis Buffer D (kit supplied) and 300 µL 20% SDS was added per cell pellet. The tube was inverted several times and 50 µL ProK (kit supplied) was added, before vortexing and incubated for 1 hr at 37°C. ProK was inactivated at 50°C for 30 min. 400 µL Binding buffer (kit supplied) was added and mixed by inversion for 2 min. The content was transferred to the spin column provided which was then spun at 14000 rpm for 3 min. Eluate was discarded and 700 µL Wash buffer (kit supplied) was added and spun again as above. Wash was repeated once more. 100 µL of Elution buffer (kit supplied) was added per column and the column was incubated at 50°C for 10 min. Final elution step involved centrifugation at 14 000rpm for 3 min. The 4 x 100 µL elutions were pooled into one 400 µL total volume and purified further by Phenol: Chloroform: Isoamyl extraction (see 2.9.1).

2.9.1 *Phenol: chloroform: isoamyl extraction of genomic DNA*

To ensure that pure DNA was obtained, and the high levels of polysaccharide produced by Pseudomonad strains was removed prior to fosmid library construction, an additional cleaning step was carried out. 400 µL of phenol: chloroform: isoamyl alcohol (25:24:1) was added to the purified DNA from section 2.9, and mixed by gently inverting 10 times. The sample was centrifuged at 13000 rpm for 10 min at room temperature and 370 µL of the upper aqueous phase was collected gently. Equal volume of chloroform: isoamyl alcohol (24:1) was added, inverted and then centrifuged for 10 min at 13000 rpm. 350 µL of upper phase was collected again and equal volumes of chloroform: isoamyl alcohol (24:1) was added. The tube was inverted and spun at 13000 rpm for 10 min again. 320 µL of the upper aqueous phase was collected and 1/10th volume of 3 M Sodium acetate (pH 5.2) was added. The tube was mixed by gentle inversion. 2.5 volumes of -20°C 100% ethanol was added and mixed again by gently inverting. The tube was chilled at -20°C for 30 min and DNA was pelleted by centrifugation at 14000 rpm for 10 min at 4°C. The pellet was washed with 70% ethanol kept at -20°C, then the pellet was dried at 50°C for 5-10 min and finally resuspended in 25-30 µL of MQ water before storage at -20°C.

High molecular weight DNA extracted from individual isolates was analysed by gel electrophoresis. 0.8% Agarose in 0.5X TBE gels were prepared. 2 μ L of gDNA was loaded per isolate. An initial run at 100V for 5 min was done so as to allow all of the DNA to exit the well and enter the agarose medium. The gel volume was 50 mL set as a 5 cm x 10 cm gel and was then run at 18 V overnight. Once the run was complete, the gel was stained in ethidium bromide staining solution for 30 min and was visualized by UV-light using the Gel Doc® XR+ system (Bio-Rad, USA). Total DNA yields were estimated by comparison to DNA standards and control DNA.

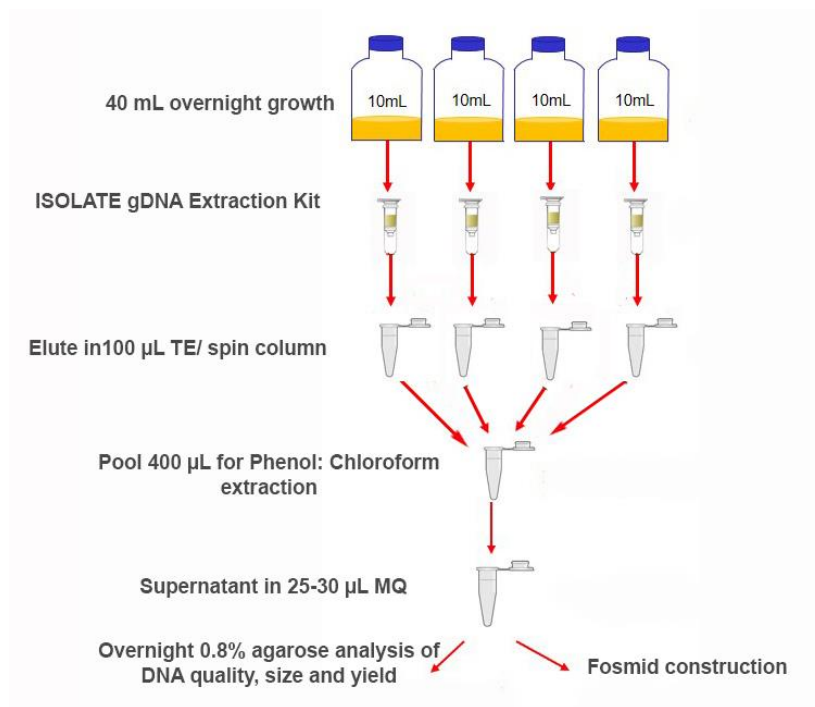


Figure 2.2: Flowchart of gDNA extraction protocol from isolates used in Fosmid construction. 40 mL of culture was grown per isolate overnight. Four separate Bioline columns were used to purify the gDNA. Elutions were pooled and an additional phenol: chloroform: Isoamyl alcohol step was added to ensure that the gDNA was pure. DNA quality and quantity was confirmed by gel electrophoresis prior to fosmid construction.

2.9.2 Fosmid library construction

To recover *IS1111-attC*, the genes flanking them, as well as potential integrons, Fosmid libraries were constructed using the CopyControl™ Fosmid Library Production Kit (Epicentre) according to the instructions supplied with the kit. For each Fosmid library constructed, 384 clones were selected for further analysis. All libraries were stored at -80°C in 96-well plates in 200 μ L of LB broth supplemented with 12.5 μ g/mL chloramphenicol, 10% L-arabinose and 15% glycerol. *E. coli* Epi300 (Epicentre) cells were transformed with

phage packaged fosmid and were grown overnight at 37°C as per manufacturer's instructions. Cells were diluted from 10^0 - 10^{-3} and plated onto LB agar supplemented with 12.5 µg/mL chloramphenicol (Cm) and grown overnight at 37°C. Individual colonies were picked and resuspended into 200 µL LB/12.5 µg Cm/10% L-arabinose wells of 96-well plates.

2.9.3 Fosmid screening via Southern hybridization

Screening for *IS1111-attC* elements in Fosmid clones was accomplished using a pooled 96-well plate method (Figure 2.3). Each Fosmid library (ie. per isolate) had 4x 96 well plates. Clones were grown overnight at 37°C, 200 rpm. For ease of screening, wells in individual columns were pooled. 5 µL per well was added to a microcentrifuge tube. In total 12 tubes were obtained per 96-well plate. All tubes were heated at 99°C for 10 min, and cellular debris was spun down at 14000 rpm for 3 min. The supernatant was reheated at 99°C for 10 min and chilled on ice for 1 min prior to spotting 3 µL onto a nylon membrane (Hybond N⁺, Amersham). The membrane was hybridized against a 1347-bp *IS1111-attC* probe as per section 3.2. A positive southern hybridization signal corresponds to a positive 96-well plate column. The 96-well plates were revisited and 10 µL was removed from the individual wells in the corresponding positive column, to give 8 tubes per positive column. The cells in the individual tubes were boil-lysed and hybridized as mentioned above. Fosmid was extracted as per manufacturer's instructions.

2.9.4 Fosmid purification

When purified fosmid was required for analysis such as restriction digestion and PCR, induction of the fosmid to high copy number was performed according to the Epicentre instructions. The fosmid was then extracted using an alkaline lysis plasmid purification protocol (see section 2.8). Purified fosmid DNA was stored in TE buffer at -20°C.

2.10 Fosmid sequencing

Glycerol stocks were sent to the commercial sequencing service Macrogen Inc. (Korea). This sequencing service extracted the fosmids from glycerol stocks, generated a shot gun library and used a 6X sequencing platform prior to assembling contigs.

2.11 Bioinformatics analysis

All sequences were analysed for the presence *IS1111-attC* elements and *intI* homologs to determine the nature of genes flanking the arrays using the various search options available on the NCBI website (Altschul et al., 1997). Geneious 6.1.5 ORF predict function was used to locate all open reading frames greater than 100 codons within the sequence data with an *atg*, *gtg* and *tgg* start codons. SnapGene® Viewer 2.3 DNA Sequence (GSL Biotech LLC, USA) was used to draw and annotate fosmid maps.

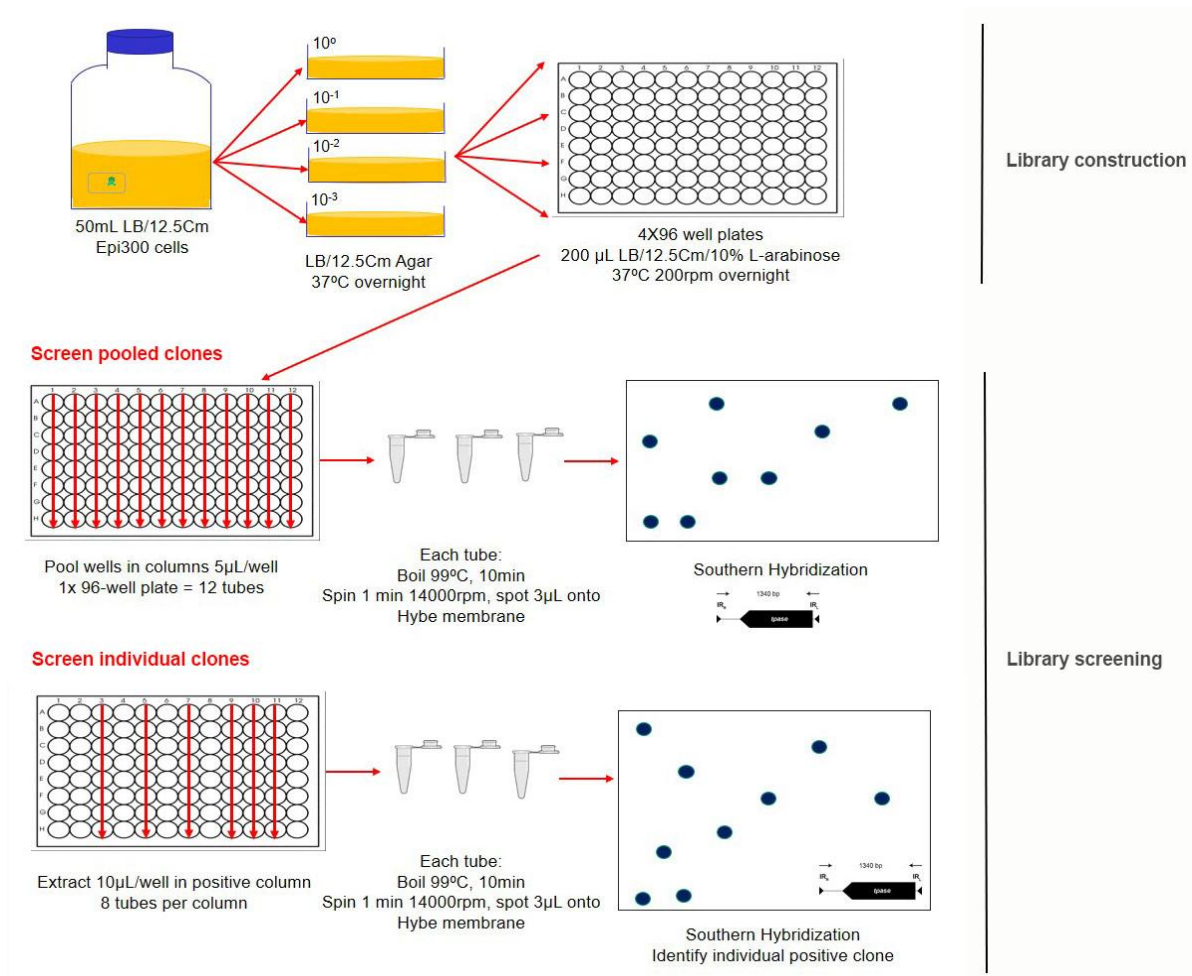


Figure 2.3: Screening fosmid libraries for *IS1111-attC* elements.

Putative clones, in 96-well plates, were pooled into 12 individual tubes, boil-lysed and spotted onto a N^+ -membrane for southern hybridization. Once a positive clone was identified, individual wells in a “positive” column are screened to identify the *E. coli* Epi300 clone that contained the *IS1111-attC* element.

Chapter 3 Distributions of IS1111-attC, mobile and chromosomal integrons within the bacterial genus *Pseudomonas*

3.1 Introduction

Acquired antibiotic resistance in many Proteobacteria is strongly associated with resistance genes on plasmid borne MRIs (Livermore, 2002; Odumosu *et al.*, 2013). It is evident these genes have been acquired from foreign sources across wide biological distances. However, understanding how resistance determinants are able to cross ecological barriers and come together in a given bacterial cell is poorly understood. Host related molecular and genetic factors are postulated to influence multiresistant phenotypes, yet these are also poorly understood.

Despite environmental organisms constituting a reservoir of resistance genes (Allen *et al.*, 2010; Cantón, 2009; Mazel, 2004), there are limited examples of resistance genes that are shared between non-pathogenic environmental and human pathogens (see Table 1.3). In the few cases that are characterized we can see two features: i) the source-sink organisms are genetically related, emphasizing that HGT is more prevalent in related organisms sharing ecological and molecular-genetic traits; and ii) additional “specialized” mobile elements have also been involved. For example, the *bla*_{CTX-M} resistance gene was translocated by *ISEcp1* from the source *Kluyvera ascorbata* to an *Enterobacteriaceae* sink. I postulate that the lack of observations of source-sink relationships for more phylogenetically and ecologically separated organisms is simply harder to observe. Transfer of a chromosomal framework (core) gene is easy to detect because that gene is always present in the source population. The integron-gene cassette system is both the major source for multi-drug resistance in Proteobacteria and a special challenge in the exploration of source-sink relationships. The gene cassettes in source populations (CIs) are not core genes. Thus they are not present in all cells and their movement from source to sink is much harder to observe.

Gene cassettes were hypothesized to have a discrete source population in environmental Pseudomonads (*Pseudomonas* 1) (Figure 3.1). For gene cassettes to cross into a related sink population (clinical *Pseudomonas* 2), they would conceivably need to overcome medium ecological distance where these two populations have overlapping, but distinct habitat preferences. Available data suggests a key ecological difference is the extent to which they are animal associated (Figure 1.13). Both groups are readily isolated from planktonic and

biofilm habitats in freshwater systems, but *Ps. aeruginosa* is by far more frequently encountered as an opportunistic pathogen, a low abundance member of the normal microbiota (Markou & Apidianakis, 2014) in the gut and impaired mucociliary clearance [notably in cystic fibrosis patients (Folkesson *et al.*, 2012) and ventilators (Planquette *et al.*, 2013)]. In the context of this study, a key genetic distinction between *Pseudomonas* populations is the presence of chromosomal integrons. These are typically, but variably, present in the *Ps. stutzeri* group but absent in the *Ps. aeruginosa* group. IS1111-attC are exploited as the “core” gene cassette. They are so frequently observed in PCIs that their movement is postulated to be readily observable. I used three groups of γ -Proteobacteria that represent three distinct categories of pair-wise biological separations. To study the flow of genetic material between ecological and molecular-genetic barriers, chromosomal integrons, MRIs and IS1111-attC elements were used as a model system for HGT and resistance acquisition.

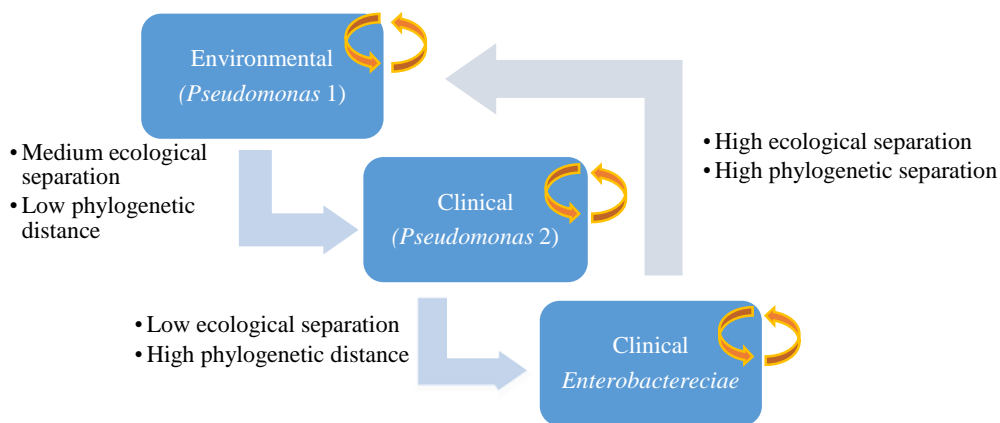


Figure 3.1 Conceptual framework for the study using IS1111-attC elements as a model for gene flow across biological distances.

IS flow between *Pseudomonas* populations is postulated to have to overcome medium ecological separations, but low genetic distances. IS flow between a clinical *Pseudomonas* population and enteric population is predicted to be limited by the high genetic differences despite reduced ecological distance. Ecological and phylogenetic distances between Enterobacteriaceae and population 1 is thus predicted to have a vast impact on IS1111-attC translocation. Blue arrows represent IS1111-attC flow between groups, orange arrows represent IS1111-attC flow between different strains/species within a group.

IS1111-attC and gene-cassettes are both genetic entities that are dependent on integrons for residence in a cell. The availability of IS residence sites (*attC*) is predicted to be limited in Pseudomonads with no PCIs and/or there is potential molecular incompatibility. As such, IS1111-attC spread between *Pseudomonas* populations 1 and 2 may need to overcome low levels of molecular-genetic separations. Furthermore, IS spread within a clinical system to a genetically unrelated group (*Enterobacteriaceae*) is predicted to be governed by low

ecological distance but high molecular-genetic distance (Figure 3.1). Thus by tracking its distribution across the three populations in Figure 3.1, we can assess the relative importance of ecological barriers in MDR formation. Secondly, by exploring the details of the genetic insertion points we can examine what molecular-genetic interactions were part of the *IS1111-attC* translocation process.

3.1.1 Aims

IS1111-attC was used as a model gene cassette to test the relative contributions of genetic and ecological aspects to gene flow and MDR formation in integrons. Characteristically, *IS1111-attC* members mimic gene cassettes as they site-specifically target integron associate recombination sites within PCIs and MRIs (Tetu & Holmes, 2008). The primary aim in this chapter was to identify potential source and sink elements (MRIs, PCIs and *IS1111-attC*) in three biologically distinct Proteobacteria populations. The fundamental question is: Is gene flow of *IS1111-attC* elements greater between phylogenetically related organisms (Pseudomonad to Pseudomonad) or between non-related organisms (Pseudomonad to non-Pseudomonad)?

The overall study design was based on the following predictions:

- 1) *IS1111-attC* elements are dependent on the presence of MRIs in clinically relevant Pseudomonads
- 2) *IS1111-attC* elements are dependent on the presence of PCIs in environmental Pseudomonads isolated from low selective pressure environments (domestic)
- 3) *IS1111-attC* elements are overrepresented in the genus *Pseudomonas* relative to their distribution in an enteric outgroup

3.2 Methods

For general media and solution recipes, as well as PCR conditions, refer back to Chapter 2.

3.2.1 Strain Collection

The strain collection primarily consisted of clinical isolates obtained from various patients and environmental isolates from inert objects such as sponge and mop samples (Table 3.1). All 125 clinical human isolates and 9 non-host associated clinical environmental isolates were kindly donated by collaborators. Sixty-three Pseudomonads were isolated as per section 2.2 from domestic sponge and mop samples. An additional six non-host associated isolates were recovered from hospital mop and swab samples. For more information on all isolates refer to the Appendix, Tables A1-A3.

Table 3.1 Strain collection

	<i>Clinical</i> ⁺	<i>Non-clinical</i>		
		<i>Hospital</i>	<i>Environmental Ps.</i> <i>(water, soil)</i>	<i>Domestic</i>
Pseudomonad	125	6	9	63
Enteric outgroup	52	0	0	0

+ From clinical specimens (blood, sputum, wounds etc.)

3.2.2 Genotyping domestic environmental Pseudomonas strains

Fingerprinting Pseudomonad isolates obtained from sponge and mop samples was done using a repetitive sequence based PCR using the BOXA1R sequence (BOX-PCR) (Versalovic *et al.*, 1994). This approach generates fingerprints that can discriminate bacteria at very fine resolution and can therefore distinguish different strains of the same species (Louws *et al.*, 1994; Gillings & Holley, 1997). The technique uses just one oligonucleotide BOXA1R, which acts as both a forward and reverse primer. Band-classification was done by manual correction of automated analysis with Quantity-One software. Cluster analysis of BOX-PCR banding patterns was done manually using Quantity-One software at a 1.5% tolerance level.

3.2.3 Southern Hybridisation

High molecular weight genomic DNA was extracted for all isolates as per section 2.3. DNA was quantified and its purity was checked by NanoDrop® spectrophotometry analysis and gel electrophoresis. Approximately 200-400 ng of high molecular weight genomic DNA of each isolate was denatured at 99°C for 15 min in MQ water and kept on ice during spotting onto a nylon membrane (Hybond N⁺, Amersham). The membrane was air-dried and the DNA was UV autocross linked to the membrane using the Stratalinker® UV Crosslinked (Stratagene, USA). The DNA was cross-linked with 1200 microjoules (x100) for a period of 45 sec using a 254-nm light source.

Southern Hybridisation was performed as described by Southern (1975) with modifications by Sambrook and Russel (2001). Probes for southern hybridisation were labelled with digoxigenin-6-dUTP (Roche) directly by PCR (Figure 3.2). The efficiency and yield of the labelling reaction was assessed using standard gel electrophoresis of amplified probes. Band intensity after staining with ethidium bromide was used to gauge the yield of the labelling reaction, and as the addition of DIG-dUTP to a PCR product affects electrophoresis, retardation of the labelled PCR product relative to an unlabelled control was used to assess labelling efficiency.

All hybridisations were performed in pre-dissolved and pre-heated *DIG Easy-Hyb* Granules (Roche) as per manufacturer's instructions. 25 ng/mL denatured DIG-labelled probe was added to dissolve *DIG Easy-Hyb* Granules in a total volume of 20 mL/100 cm² membrane size. This was incubated overnight at the hybridisation temperature of ~55°C. Hybridization, immunological detection and visualisation steps were performed using the DIG Nucleic Acid Detection Kit (Roche Applied Science) according to the manufacturer's instructions. Stringency washes, to remove unbound probe, were carried out as follows (20 mL/100 cm²): once for 5 min with 2x SSC, 0.1% SDS at 20°C, followed by two 15 min washes with 0.5x SSC, 0.1% SDS at 60°C. To visualise bound probe, the membrane was evenly soaked in CDP-Star solution (Roche Applied Science) and allowed to incubate for 10 min prior to exposure to X-ray film (Kodak X-OMAT, Henry Schein Regional). Multiple exposures, employing different incubation times, were performed following each hybridization, to optimise the signal strength.

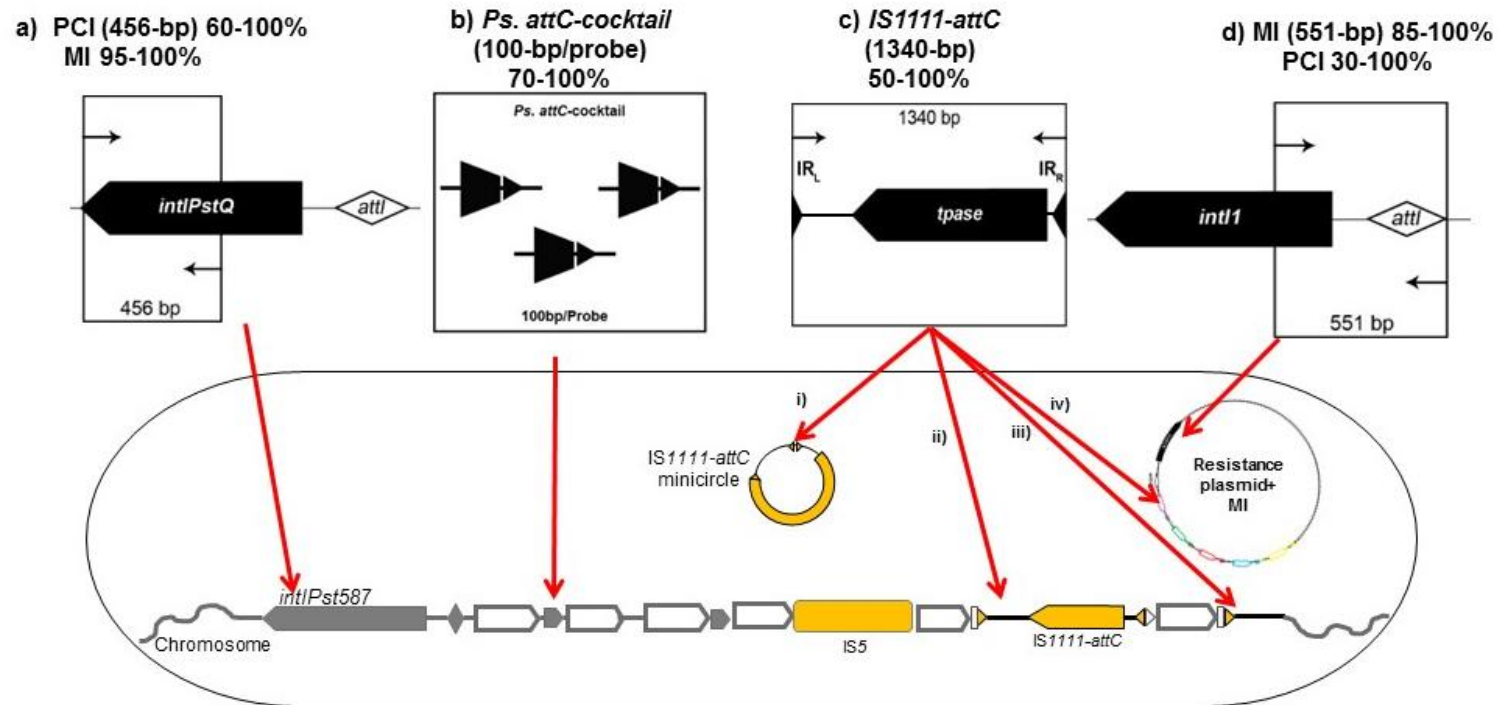


Figure 3.2 Probes and their respective target sites used in southern hybridization experiments.

% indicates the expected sequence variation in these targets for each probe.

A) PCI probe: based on a 456-bp region corresponding to the chromosomal integron integrase

B) attC cocktail probe (~100-bp per probe): three different *Ps.*-type attC sites from *Ps. mendocina* NW1, *Ps. straminea* KM91 and *Ps. stutzeri* DNSP21.

C) *IS1111-attC* probe (1.3-kb): detect different levels of divergences of *IS1111-attC* elements. An *IS1111-attC* may reside as i) non-integrated minicircle or embedded in chromosomal attC; ii) intact IS; iii) *IS1111-attC* footprint iv) located on MRI

d) MRI class 1 integron probe (550-bp): pR388 integron integrase region and attI recombination site.

3.2.4 Southern hybridization thresholds of detection

Southern hybridization signals will be influenced by the copy number, sequence divergence and sequence extent of the targeted genes. An extensive array of controls were used to interpret the hybridization screens (Table 3.2). A positive signal was determined to be true if the isolate had a signal stronger than the control strains, and was consistent with predicted sequence divergence. *Ps. stutzeri* ATCC 14405 (*Pst405*) and *Ps. stutzeri* ATCC 17587 (*Pst587*) contain at least three full intact copies of IS1111-attC elements. *Ps. stutzeri* ATCC 17595 (*Pst595*) and *Ps. stutzeri* RNAIII (*PstRNAIII*) contain an IS1111-attC footprint. These strains serve as positive controls for sequences that have greater than 50% identity over regions of 200 nucleotides and an IS1111-attC copy number of 1-3. In comparison *Ps. stutzeri* strain Q (*PstQ*) contains no IS1111-attC elements and *Ps. aeruginosa* PAO1 contains six transposases that are distantly related to IS1111-attC transposase. These two strains serve as negative controls for non-specific hybridization to sequences with less than 30% sequence homology. Similar detection thresholds are applied for the MRI, PCI and *Ps. attC* probes.

Table 3.2 Control strains

Control strain	ID	Element		
		MRI	IS1111-attC	PCI
<i>Ps. stutzeri</i> ATCC 14405	<i>Pst405</i>	-	+	+
<i>Ps. stutzeri</i> ATCC 17587	<i>Pst587</i>	-	+	+
<i>Ps. stutzeri</i> ATCC 17595	<i>Pst595</i>	-	+	+
<i>Ps. mendocina</i> NW1	<i>PsNW1</i>	-	+	+
<i>Ps. stutzeri</i> RNAIII	<i>PstRNAIII</i>	-	+	+
<i>Ps. straminea</i> KM91	<i>PsKM91</i>	-	+	+
<i>Ps. stutzeri</i> strain Q	<i>PstQ</i>	-	-	+
<i>Ps. fluorescens</i> NCTC 7244	<i>Pfluro</i>	-	-	-
<i>Ps. aeruginosa</i> PAO1	<i>PAO1</i>	-	-/+*	-
<i>Ps. aeruginosa</i> 216.2B	<i>216.2B</i>	+	-	-
<i>Ps. aeruginosa</i> MB	<i>PstMB</i>	+	+	-
<i>E. coli</i> R388	<i>R388</i>	+	-	-

**Ps. aeruginosa* PAO1 has 6 putative transposases which display approximately 60% sequence similarity and 38% identity to IS1111-attC transposase.

3.2.5 Polymerase chain reaction of gene fragments for sequencing and sequence analysis

Amplification of IS1111-attC and integron sequences was achieved by PCR using 10-100 ng of template DNA. Purification of the PCR product was carried out with Qiaquick PCR purification kit (Qiagen) prior to cloning into a pGEM-T® Easy vector system (Promega) (as per manufacturer's instructions). For difficult to sequence PCR products, or where multiple PCR products were obtained, the PCR product was cloned into pGEM-T Easy vector also. Automated sequencing was carried out by the Australian Genome Research Facility (Sydney, Australia) on an ABI 3730xl 95 capillary automated DNA sequencer. Geneious 6.1.5 software was used for all alignments and annotations of sequence reads, and the National Centre for Biotechnology Information (NCBI) BLAST website (www.ncbi.nlm.nih.gov/80/BLAST) utilizing both BLASTN and BLASTX programs were used to search the recovered nucleotide sequences for homologs.

3.3 Results

3.3.1 *The non-clinical, environmental isolate set is composed of diverse members of the genus Pseudomonas*

The primary aim was to compare and contrast the distribution of MRIs, PCIs and IS1111-attC elements between Pseudomonads occupying distinctly different habitats. All 125 clinical and 9 non-clinical isolates provided by collaborators had previously been identified as belonging to the genus *Pseudomonas*. An additional sample set of 63 non-clinical (domestic) isolates were recovered from mop and sponge samples from six suburban homes in Sydney, Australia. An additional 6 hospital environmental Pseudomonads were isolated in this study from a hospital panroom swab and bench swab sample. For more comprehensive information about these environmental isolates see Appendix Table A2. See Table 3.3 for a brief overview of the environmental isolate series isolated in this study using the selective medium, *Pseudomonas* Isolation agar (PIA).

Table 3.3 Environmental isolates recovered on PIA from mop, sponge and swab samples.

Isolate Series*	Habitat	Number of strains
AP	Hospital Panroom Swab	3
AS	Hospital Bench swab	3
ES	Domestic Sponge	5
FM	Domestic Mop	11
JM	Domestic Mop	11
JS	Domestic Sponge	7
LS	Domestic Sponge	15
MS	Domestic Sponge	6
NS	Domestic Sponge	8
TOTAL		69

*Abbreviation for person supplying the sample and sample type, e.g. Mia Sponge= MS

To exclude sibling isolates and confirm independence of the 69 non-clinical isolates, all were typed by BOX fingerprinting. Grouping of the PCR banding patterns was done manually using Quantity-One software and bands were analysed in comparison to each other and the 100-bp molecular weight markers. The results of the BOXA1R PCR indicated that the vast majority of environmental Pseudomonads are genetically distinct and independent isolates (53/69 different BOX patterns). See appendix Figure A1 for all BOXA1R profiles.

In summary, eight out of sixty-nine isolates (8/69), FM3, FM8, JM3, JM4, LS11, LS14, MS5 and PO6, consistently failed to give a BOX-PCR product despite producing a strong 16S rRNA DNA PCR amplicon. It is likely that these isolates do not contain repetitive regions to which the BOXA1R primer could anneal. Also, no observable difference in BOX-PCR banding patterns was seen for isolates AP1 and AP3. 16S rRNA DNA sequencing identified both isolates as *Ps. putida* W619 (iso. AP1) and *Ps. putida* ND6 (iso. AP3). Also both isolates hybridized positively only to the MI probe and no other probe. Similarly, AS1 and AS3 had identical BOX repetitive patterns and 16S rRNA DNA sequencing identified both isolates as *Pseudomonas* sp. AS1 and *Pseudomonas* sp. AS3. However, the southern hybridization “profiles” varied in that AS1 hybridized positively to both the MI and *IS1111-attC* probe and AS3 was negative for all probes tested.

Finally, LS5, LS12, LS15, NS1, NS2 and NS5 had identical patterns. Further differentiation of these isolates was made possible using southern hybridization (See section 3.3.2). For instance LS5 was identified as positive for *IS1111-attC* and *Pseudomonas*-type *attC* site. In comparison isolates LS12 and LS15 are negative for all mobile elements of interest. Isolate NS2, despite having an identical BOX rep pattern, was isolated from a different domestic source. Similarly, NS1 and NS5 display the same BOX rep pattern. However these two strains were of very little interest as they were found to lack PCI, MI, *Pseudomonas*-type *attC* and *IS1111-attC*. Finally, JS3 and JS6 had identical BOX patterns but JS3 hybridized positive for PCI and *Pseudomonas*-type *attCs*. JS6 was negative for all elements.

Use of PIA proved successful in recovering independent isolates. Note, where an isolate was of interest based on the southern hybridization data, the 16S rRNA DNA was sequenced to further validate that it indeed was a *Pseudomonad*. In conclusion, the environmental sample set is comprised of different members of the *Pseudomonadaceae* family.

3.3.2 Distribution of IS1111-attC elements, MRIs and PCIs across three phylogenetic groups

All isolates were examined by Southern hybridization and PCR for the presence of MRIs, *IS1111-attC* elements, PCIs and *Pseudomonas*-type *attC* sites (Table 3.4). An array of controls was used for every southern hybridization and that a positive signal was called such when the hybridization intensity was 50-100% of that of the controls (Figure 3.3).

Table 3.4 Distribution of MRIs, PCI, and IS1111-attC across Pseudomonads based on Southern hybridization and PCR findings.

Geographic group		No. isolates	Distribution (%) ^a					No. of recovered sequences ^b		
			IS1111-attC	MI	PCI	Ps. attC	MI+ IS1111-attC	PCI+ IS1111-attC	IS1111-attC	MI
Pseudomonadaceae	Clinical ^c	125	34 (27%)	24 (19%)	4 (3%)	2 (2%)	19 (15%)	2 (2%)	14	9
	Total	78	17 (22%)	11 (14%)	13 (17%)	9 (12%)	6 (8%)	9 (12%)	9	8
	Hospital	6	1 (17%)	6 (100%)	0 (0%)	0 (0%)	1 (17%)	0 (0%)	1	4
	Environment	9	3 (33%)	1 (11%)	2 (22%)	2 (22%)	0 (0%)	2 (22%)	3	0
	Domestic	63	18 (29%)	4 (6%)	11 (18%)	7 (11%)	4 (6%)	8 (13%)	5	3
Enterobacteriaceae	Outgroup	52	0 (0%)	11 (21%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0	2

a) Data are presented as a total percentage from southern hybridization and PCR screens, where bold denotes the data for each population.

b) Number of recovered sequences by PCR and sequencing.

c) All clinical isolates were derived from independent sample sets, annotated as TS, JIP, SSI and P (Table A1).

d) Non-clinical Pseudomonads are further subdivided into the following categories: hospital (swabs, panroom swab, sponge, mop); environmental (soil, water); and domestic (swab and sponge.)

24 clinical strains (19%, 24/125) gave a signal of comparable intensity to the positive MRI control and significantly different from all negative controls (Figure A2 Part A). In comparison, three percent of the clinical isolates (3%, 4/125) hybridized with a PCI probe and 2% (2/125) also probed positive for *Pseudomonas*-type attC sites (Figure A2 Parts C & D). Two *Pseudomonas aeruginosa* isolates SSI2.84 and JIP117 gave a weak signal with the PCI but not the *Pseudomonas attC* probe. These two isolates were found to carry an MI, hence the PCI signal is believed to be a cross-reaction with an MRI probe due to the low stringency conditions used during hybridization washes rather than the presence of a true PCI. MIs are often carried on high copy number plasmids, hence the weak signal detected under low stringency conditions. Isolate P15 (*Ps. stutzeri* ATCC 17588) and isolate P17

(*Pseudomonas stutzeri*), hybridized for PCI and *Pseudomonas*-type *attC*, but not MRI. 27% (34/125) of the clinical *Pseudomonads* emitted a positive hybridization signal for *IS1111-attC* elements (Figure A2 Part B).

In comparison, 14% (11/78) of non-clinical *Pseudomonads* positively hybridized to the class1 integron probe (Figure A3 Part A). 17% (13/78) produced a hybridization signal to PCI probe (Figure A3 Part C) and 12% (9/78) hybridized to the *Pseudomonas*-type *attC* cocktail probe (Figure A3 Part D). Finally, 22% (17/78) of the non-clinical isolates were positive for *IS1111-attC* elements (Figure A3 Part B).

The clinical enteric outgroup used in this study is a collection of 52 isolates obtained as part of an antibiotic cycling study at Westmead Hospital, Sydney (Ginn *et al.*, 2012). 21% (11/52) of this group was MI positive but were all negative for *IS1111-attC*, PCIs and *Pseudomonas*-type *attC* sites (Figure A4). Two isolates within this sample set gave a rather strong signal for an MI. 16S rRNA DNA sequencing identified these two isolates as *Enterobacter* sp. JUNL-1 and *Enterobacter cloacae* subspecies *cloacae* ATCC 13047.

Given these observations I conclude that chromosomal integrons, *IS1111-attC* and *Pseudomonas*-type *attCs* are all more prevalent within *Pseudomonads* than the *Enterobacteriaceae* outgroup.

3.3.3 *IS1111-attCs* are overrepresented in MRIs in clinical, and in PCIs in non-clinical *Pseudomonads* respectively

To determine if an *IS1111-attC* element can co-exist in cells with multiple types of integrons, mobile class1 vs. chromosomal integrons, southern hybridization data was examined for any isolates that hybridized to two or more probe combinations (Table 3.4).

15% (19/125) of clinical *Pseudomonads* hybridized to both an MI and *IS1111-attC* probe. 2% (2/125) of the clinical isolates were observed to hybridize to both a PCI and *IS1111-attC* element. Given the small percentage containing a PCI and IS, suggests that these are acquired at lower rates within the clinical *Pseudomonads*. The inverse was observed for the non-clinical *Pseudomonads*. 12% (9/78) of these isolates produced a hybridization signal for both the PCI and *IS1111-attC* element, whilst only 8% (6/78) produced positive signals for

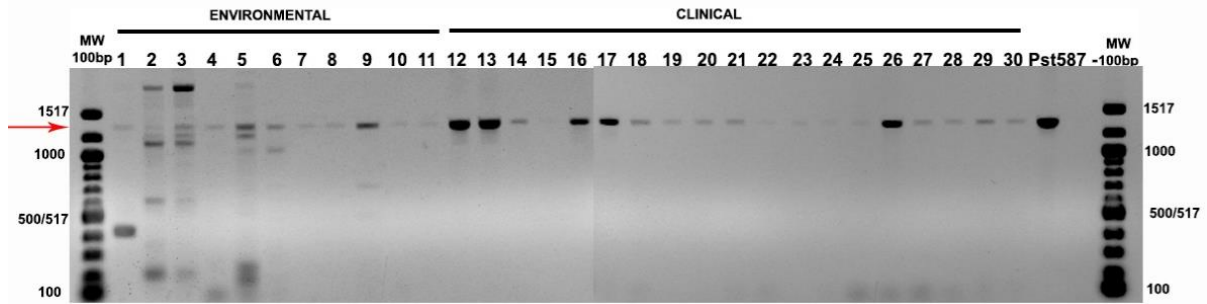
both MI and *IS1111-attC* elements. In the case of the enteric outgroup, no such association was observed as no *IS1111-attC* elements were recovered from these 52 isolates.

These data support previous findings of Tetu and Holmes (2008), where *IS1111-attC* were found to have a strong association with PCIs. This current study demonstrates that *IS1111-attC* are associated with MIs within clinical *Pseudomonads* and that generally *IS1111-attC* are over-represented within the *Pseudomonadaceae*.

3.3.4 PCR detection and sequence analysis of *IS1111-attC* elements

Screening and sequencing efforts targeting *IS1111-attC* hybridising fragments across all hybridisation positive clinical and environmental strains recovered a total of 23 strains. However two sequences from domestic *Pseudomonas* sp. JM7 and *Pseudomonas aeruginosa* SSI 2.37, were disregarded from further analysis as these were found to be non-specific products. Thus 21 IS elements spanning different levels of divergence, are considered to be a representation of *IS1111-attC* related sequences as observed by hybridisation intensity, and sequence divergence within this strain collection.

PCR recovery using primers targeting intact *IS1111-attC* from the non-clinical set generally resulted in multiple banding patterns (Figure 3.4, Lanes 1-6 & Figure A2). Tetu and Holmes (2008) found that multiple banding patterns are readily obtained when more than one *IS1111-attC* element is located within a PCI array. In comparison, recovery of *IS1111-attC* elements from host associated *Pseudomonas* isolates, resulted in a single band of the expected size (Figure 3.3, Lanes 12-30).



Lane	Isolate ID	Number of bands & corresponding size (bp)	Additional information where 1300-bp product was sequenced
1	JM1	2 bands (400-bp & 1300-bp)	
2	JM2	6 bands (200bp, 600-bp, 1000-bp, 1100-bp, 1300-bp, 1500+bp)	
3	JM6	6 bands (200bp, 600-bp, 1000-bp, 1100-bp, 1300-bp, 1500+bp)	
4	JM7	1 band (1300-bp)	Sequenced- non-specific product
5	JM10	7 bands (100-bp, 200bp, 600-bp, 1000-bp, 1100-bp, 1300-bp, 1500+bp)	
6	AS1	2 bands (1000-bp & 1300-bp)	Sequenced- <i>ISUnc1</i> -like element
7	LS5	1 band (1300-bp)	
8	LS6	1 band (1300-bp)	
9	LS8	1 band (1300-bp)	Sequenced-confirmed <i>ISUnc1</i> -like element
10	LS13	1 band (1300-bp)	
11	NS6	1 band (1300-bp)	Sequenced- non-specific product
12	TS464	1 band (1300-bp)	
13	TS491	1 band (1300-bp)	Sequenced- confirmed <i>ISPa21</i> -like
14	TS504	1 band (1300-bp)	
15	TS558	1 band (1300-bp)	Sequenced-confirmed <i>ISUnc1</i> -like
16	TS684	1 band (1300-bp)	
17	TS589	1 band (1300-bp)	
18	TS597	1 band (1300-bp)	
19	TS598	1 band (1300-bp)	
20	TS599	1 band (1300-bp)	
21	TS600	1 band (1300-bp)	
22	TS616	1 band (1300-bp)	
23	TS620	1 band (1300-bp)	
24	TS629	1 band (1300-bp)	
25	TS630	1 band (1300-bp)	
26	TS631	1 band (1300-bp)	
27	TS634	1 band (1300-bp)	
28	TS653	1 band (1300-bp)	
29	TS654	1 band (1300-bp)	
30	TS664	1 band (1300-bp)	
31	<i>Pst587</i>	1 band (1300-bp)	Positive control 1.3-kb <i>IS1111-attC</i>
32	-ve	0	Negative control MQ water only

Figure 3.3: *IS1111-attC* PCR screen across a subset of environmental and clinical *Pseudomonas* isolates.

All isolates were previously identified via southern hybridization to contain *IS1111-attC* elements.

Previously, Tetu and Holmes (2008) recovered a total of 9 IS elements and two IS-like fragments from three *Pseudomonas stutzeri* genomovar 2 strains. Using a criterion of 98% sequence identity, 21 independent *IS1111-attC* occurrences (14 host associated and 9 non-host associated) were identified in this study. All 21 were attributed to one of the previously described isoforms, *ISUnc1*-like, *ISPa21*-like and *ISPst6*-like. Sequences corresponding to the right hand non-coding region (Figure A7) were used to generate Figure 3.4 as this region is typically retained in degenerate *IS1111-attC* elements, so it is suitable for global analysis of IS sequences (Tetu & Holmes, 2008).

Eight clinical isolates (JIP004, JIP009, JIP041, JIP044, JIP045, JIP047, JIP073, TS558), one hospital environmental *Pseudomonas* sp. AS1, one environmental *Ch. meningosepticum* iso. PO6, and two domestic environmental isolates (*Delftia* sp. LS8 and *Pseudomonas* sp. FM4) were found to have *ISUnc1*-like elements. These recovered sequences represent divergent *ISUnc1*-like elements where sequences from *Ps. aeruginosa* isolate JIP004 and JIP041 are 97-98% identical to the reference *ISUnc1*-like sequences (*ISUnc1 Pa* and *ISUnc1 Psp7*). *ISUnc1*-like sequence from six remaining host associated isolates (JIP009, JIP044, JIP045, JIP047, JIP073, TS558) are 99% identical to the *ISUnc1*-like sequence recovered from domestic isolate, *Pseudomonas* sp. FM4. This isolate also hybridized to the PCI probe but not an MRI probe. Note that 16S rRNA DNA sequence analysis for isolate FM4 revealed 98% (1387/1409) identity to *Ps. putida* ATCC12633 type strain 16S rRNA DNA. A PCI association with a lineage within *Ps. putida* is thus the likely source for *ISUnc1*-like elements.

Two clinical *Ps. aeruginosa* isolates, JIP052 and TS491, have *ISPa21*-like elements. These IS elements have previously been observed in clinical *Ps. aeruginosa* isolates and in each case were in the *attC* site of a resistance cassette in a class 1 integron (*attC_{aacA7}*, *attC_{bla_{GES-9}}*, *attC_{aadA7}*) (Poirel *et al.*, 2005; Tetu & Holmes, 2008).

Finally, two domestic isolates (Unknown *Pseudomonas* sp. FM1 and FM3), two environmental isolates (*Chryseomonas luteola* iso. PO5, *Ps. mendocina* ATCC25411 iso. P13), and three clinical isolates (*Ps. stutzeri* ATCC17588 iso. P15, *Ps. aeruginosa* JIP117 and SSI2.84) have *IS1111-attC* elements with more than a 96% identity to *ISPst6*. Acquisition of an *ISPst6*-like element by a clinical strain reflects recent acquisition of the IS from an environmental *Ps. stutzeri* strain. Note however that the *ISPst6*-like sequence from

Ps. aeruginosa iso. SSI2.84 was not included in Figure 3.5 because the non-coding side was not recovered (see Appendix Figure A8). Nevertheless, the recovered IS sequence displayed 98% (825/834) sequence identity to *Pst587 ISPst6*.

Attempts to sequence the 16S rRNA DNA from isolates FM1 and FM3 were unsuccessful. Culture based work and microscopy of the culture stocks showed that they were contaminated by *Bacillus* species. This media is based on the production of pyocyanin, as such contamination by *Bacillus* sp. is believed to have occurred when FM1 and FM3 were incubated overnight in LB broth, before long-term storage in glycerol solution. The initial southern hybridization and sequencing efforts for FM1 and FM3 were not reproducible when genomic DNA was prepared freshly from the *Bacillus* contaminated glycerol stocks.

3.3.5 Recovery of class 1 *intI1-attI1* junctions from all three sample sets

To determine if the various IS*Unc1*-like and IS*Pa21*-like elements are in independent MRIs, I looked at fine scale patterns of relations to integrons. MRIs are the most comprehensively studied integron due to their high impact on the spread of antibiotic resistance (Stalder *et al.*, 2012). PCRs for the *intI1-attI1* junction were conducted as a means to validating the southern hybridization findings across all sample sets. However, the correlation between the hybridization data and PCR was inconsistent due to the *intI1* gene being prone to degeneration, deletion and insertional disruption. An additional *intI1* screening approach was thus conducted using *intI1* internal primers (HS915 and HS916) (Chowdhury *et al.*, 2009).

In figure 3.5A, 12 clinical isolates were screened using the LW primer pair (*intI1-attI1*). Six out of twelve gave an expected product size of ~550-bp. Using the HS primer pair (Figure 3.5B) additional samples (lanes 6, 7 and 9) gave a PCR product. Note that sample in lane 7 failed to give a product using the LW primer pair, however a weakly amplified ~370-bp HS915-HS916 product was observed.

Similar results were obtained in that when the clinical TS sample set was screened using the HS primer pair only two isolates (TS464 and TS491) produced a weakly amplified product of ~370-bp. However, when screened with the LW primers, 21% (4/19) of clinical TS isolates (TS464, TS491, TS589 and TS654) produced a PCR product of the correct size (Figure A6a).

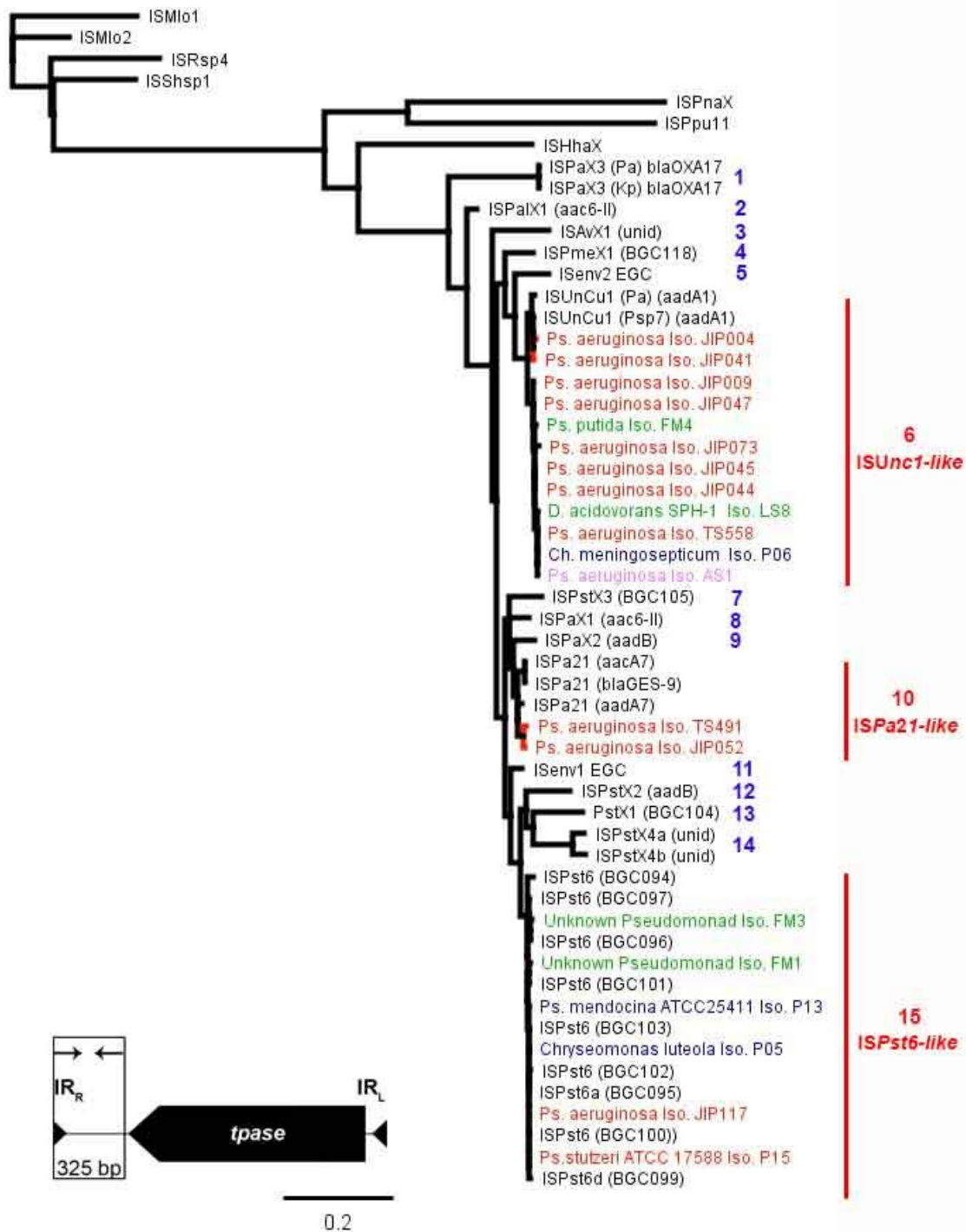
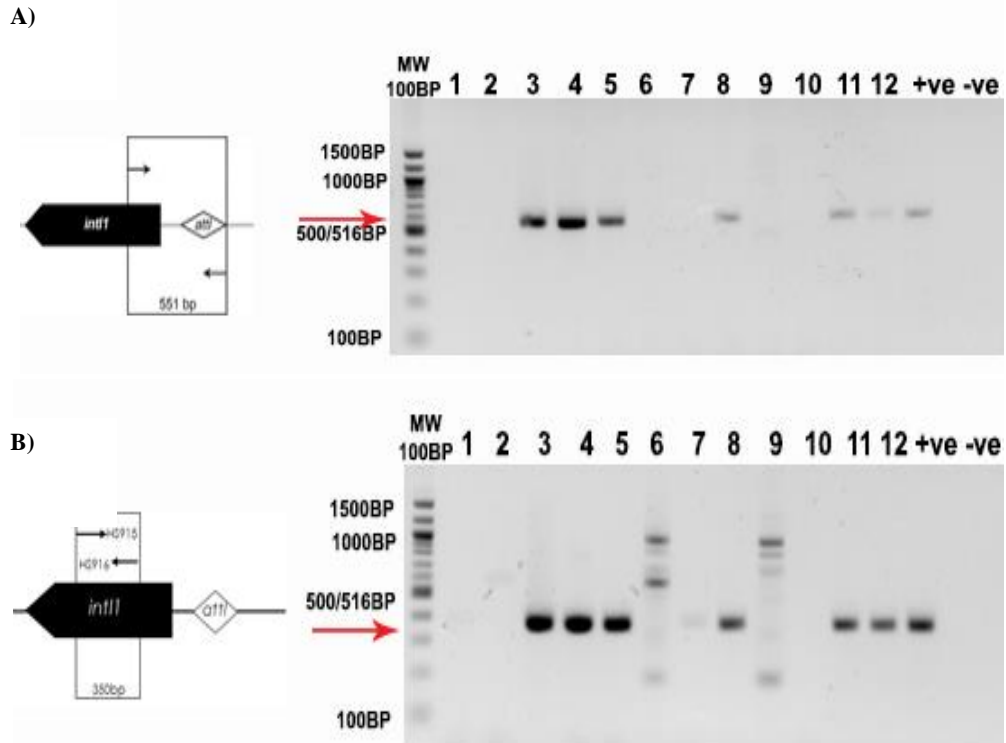


Figure 3.4: Dendrogram showing the relationship of the IS1111-attC group of elements to other members of the IS1111 family.

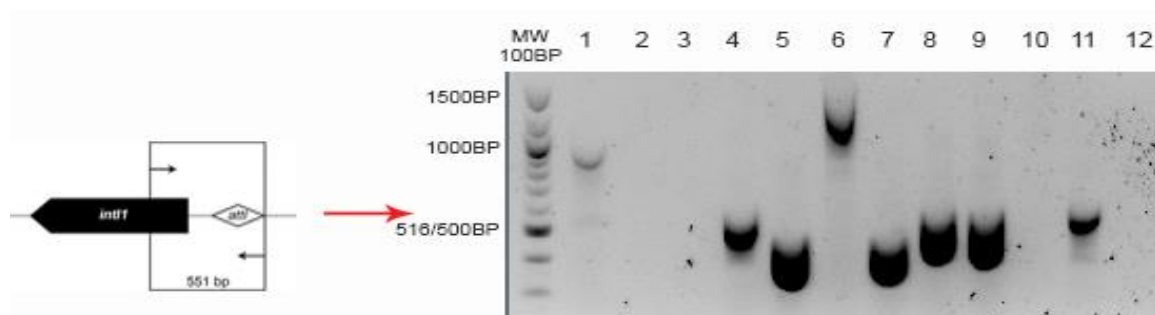
Phylogeny based on the alignment of 325 nucleotide positions. A gap alignment penalty of 10 was first applied but alignment was manually annotated. Tree was rooted using ISMol1. Most clinical IS-elements comprise a clade with ISUnc1-like elements (red). Two clinical isolates branch with ISPa21-like sequences. Domestic (green) and environmental (blue) *Ps.* isolates cluster with the known associated ISPst6-like. Figure A4 shows alignment used to generate above tree. Blue numbers represent different IS1111-attC subgroups. All recovered sequences fall into subgroups 6, 10 and 15.



Lane	Isolate ID	<i>intI1-attI1</i> (LW14/LW15) ~551bp	Internal <i>intI1</i> (HS915/HS916) ~350bp	Additional information
1	JIP031	-	-	
2	JIP038	-	-	
3	JIP041	+	+	Sequenced-confirmed <i>intI1-attI1</i>
4	JIP073	+	+	Sequenced- confirmed <i>intI1-attI1</i>
5	JIP080	+	+	Sequenced- confirmed <i>intI1-attI1</i>
6	JIP106	-	+	Array of products possibly due to insertions or deletions in the <i>intI1</i>
7	JIP117	-	+	Weakly amplified <i>intI1</i> using HS primers
8	SSI2.37	+	+	Sequenced- confirmed <i>intI1-attI1</i>
9	SSI2.84	-	+	Array of products possibly due to insertions or deletions in the <i>intI1</i>
10	SSI3.44	-	-	
11	SSI3.34	+	+	Sequenced- confirmed <i>intI1-attI1</i>
12	SSI5.34	+	+	Sequenced- confirmed <i>intI1-attI1</i>
13	R388 <i>intI1</i>	+	+	Positive control
14	No template control	-	-	No template control, no DNA only MQ water
Total positive		6	9	

Figure 3.5: A) *intI1-attI1* and B) *intI1* internal screen in a subset of 12 clinical *Pseudomonads*. Table outlines the PCR findings and sequencing results.

Note that using HS915/916 primer pair was found to be successful when screening clinical isolates for MRIs. However, these primers produced a range of non-specific PCR products ranging between 500-1500-bp, when used to screen the environmental population (Figure A6b). Similarly, screening environmental *Pseudomonads* with LW (*intI1-attI1*) primers resulted in variably sized PCR products. In Figure 3.6, two isolates produced a product size of 1-kb and 1.5-kb in size (lanes 1 and 6 respectively). Three domestic environmental *Pseudomonads* (JM2, JM6 and LS13) failed to give a PCR product (lanes 2, 3 and 10). Three isolates (lanes 4, 8 and 9) gave a PCR product of the expected size and two (lanes 5 and 7) resulted in PCR product approximately 100-bp smaller than expected.



Lane	Isolate ID	<i>intI1-attI1</i> (LW14/LW15) ~551bp	Additional information
1	JM1	+	Product is 2x expected size
2	JM2	-	
3	JM6	-	
4	JM10	+	Sequenced- confirmed <i>intI1-attI1</i>
5	AP1	+	Size is ~400-bp
6	AP2	+	Product is 1.5x in size- sequenced, contains a IS insertion in <i>intI1</i> gene
7	AP3	+	Sequenced- confirmed <i>intI1-attI1</i> with 105-bp deletion
8	AS1	+	Sequenced- confirmed <i>intI1-attI1</i>
9	AS2	+	Sequenced- confirmed <i>intI1-attI1</i>
10	LS13	-	DNA was diluted 10 fold and <i>intI1-attI1</i> was amplified and confirmed by sequencing at a later stage
11	R388 <i>intI1</i>	+	Positive control ~551-bp
12	NTC	-	No template control, no DNA only MQ water

Figure 3.6: *intI1-attI1* PCR screen of domestic and hospital environmental *Pseudomonas* isolates. Lanes 1-10: 10 environmental *Pseudomonads*. Lane 11 is positive control pR388 and lane 12 is NTC. Table Note gel is a Gel-red gel. The top of each band corresponds to the correct product size.

Different PCRs targeting different regions of the *intI1* gene were observed to give separate results and highlighted possible deletions and insertions. These deletions and insertions in

intI1 may also explain the observed inconsistencies with southern hybridization screening for MRIs.

3.3.6 Phylogenetic analysis of the recovered *intI1-attI1* sequences reveals different MRI populations

Nineteen *intI1-attI1* sequences were recovered across all sample sets (Table 3.5), and were analysed for single nucleotide polymorphisms (SNPs) in reference to the *E. coli* pR388 (In3) *intI1-attI1* sequence. In this study the pR388 *intI1-attI1* sequence (Accession U12441) is also referred to as the wild-type integron (WT). SNP positions are in reference to the starting point of alignment, nucleotide A of the wild-type integron (Figure A9 & A10). Given that the amplicon sequences were generated using Sanger sequencing, the SNPs called were also confirmed against the chromatogram files. Broadly, the *intI1-attI1* sequences can be divided into two groups based on a SNP at position 263.

Table 3.5: Observable SNPs in recovered *intI1* sequences based on pR388 class 1 integron In3 as the type sequence

Group	SNPs between <i>intI1</i> sequences	Clinical	Non-Clinical		Clinical Enteric	Total (n=19)
			Hospital	Domestic		
1	SNP263t	5	0	2	2	9
1a	SNP263t	4	-	1	1	6
1b	SNP263t + SNP268c	-	-	-	1	1
1c	SNP263t+ SNP288a	1	-	1	-	3
2	SNP263c	4	6	0	0	10
2a	SNP263c	4	3	-	-	7
2b	SNP263c+ Deletion	-	1	-	-	1
2c	SNP263c+Insertion (IS4)	-	1	-	-	1
2d	SNP263c+SNP113a	-	1	-	-	1

Group 1 is defined by thymine (T) at position 263 (SNP_{263t}). Five out of nine *intI1-attI1* sequences from clinical Pseudomonads, one domestic environmental Pseudomonad, and one enteric *intI1-attI1* were observed to have SNP_{263t}. Two subtypes (1b and 1c) of SNP_{263t} were also observed. Subtype 1b) has an additional SNP at position 268 where the wild-type nucleotide guanine is replaced by a cytosine. This 1b SNP subtype was only observed in the *intI1-attI1* sequence of the second outgroup isolate, *Enterobacter* sp. 11045. Subtype 1c) has an additional SNP at position 288 where an adenine is located. One clinical (*Ps. aeruginosa*

JIP080) and one domestic (*Brevundimonas* sp. LS13) isolate belong to the SNP_{263t}+ SNP_{288a} subtype.

Group 2 has a different SNP at position 263, where the thymine of the type sequence is replaced by cytosine (SNP_{263c}). Four clinical, four hospital environmental, and two environmental *intI1-attI1* sequences can be categorized into the 2a) SNP_{263c} group. However, further diversification can be seen across the six non-clinical sequences. One sequence was placed into subtype 2b). *Pseudomonas* sp. AP3 was found to have a 105 base pair deletion in the *intI1* gene between positions 329-434. Subtype 2c) was observed in the clinical environmental *Stenotrophomonas* sp. AP2 and involves an insertion of 956 base pairs between positions 329-434. This insertion has an 86% identity to IS4 transposase gene found on a plasmid pSWIT02 in *Sphingomonas wittichii* RW1. The final subtype is 2d) was observed only once in *Chryseobacterium meningosepticum* isolate P06. To date this may be the only example of an *intI1-attI1* sequence recovered from a *Chryseobacterium meningosepticum*.

Whilst sequencing based errors may introduce SNPs, the probability of this occurring across all 19 recovered *intI1-attI1* sequences is small. In all cases, PCR products were directly sequenced throughout different periods within this study. Thus, the observed SNPs are believed to be genuine.

3.3.7 Distinct IS1111-attC groups are associated with different intI1 subgroups

For 7/21 of the recovered IS1111-attC elements, MRI sequence data was available. Their association was analysed and mapped against a simplified version of the IS1111-attC tree in Figure 3.4 to generate Figure 3.7. This observation of IS-IntI co-association in independent strains means multiple acquisitions of the IS followed by spread of the IS-Int mosaic. Note that during the course of this study, collaborators at the University of Technology Sydney published the *intI1* sequence of the clinical isolate *Ps. aeruginosa* iso. TS491 (Martinez *et al.*, 2013).

There is an emerging pattern where the same IS phylotype is always seen in the same Int SNP type. Group 2 *intI1* sequences with SNP_{263c} were present in 5 of the 7 strains harbouring an IS of the IS*UncI*-like clade. The IS*UnCuI* group is found in at least three distinct integrons.

Two of these are the MRI subgroups (SNP_c and SNP_t) and one is a PCI in *Pseudomonas* sp. FM4. Isolate FM4 did not hybridize to an MRI probe, and PCR based recovery of the MRI was not successful. It did however hybridise to a PCI probe, thus the IS*Unc1*-like element in isolate FM4 is believed to be located on a PCI. Note that IS*Unc1* (*Psp7*) (AY139602) and IS*Unc1* (*Pa*) (DQ302723) are both located on a class 1 integron for which the integron integrase and *attI* sequence are not available; as such, the subgroup preference for these two reference sequences remains unknown. However based on their clustering with the recovered IS*Unc1*-like sequences in this study, they may show preference for integron group 1 (SNP_t) (Figure 3.7).

Group 1 *intI1* sequences with SNP_{263t} were present in clinical *Ps. aeruginosa* isolates TS491 and JIP052. These isolates were confirmed to have IS*Pa21*-like elements. *Ps. aeruginosa* isolate JIP041 possessed an IS*Unc1*-like element, and also was classified into the IntI group 2. This IS*Unc1*-like element branched separately from all other recovered IS*Unc1*-like sequences that target group 2 integrons. In comparison the IS*Pa21* group occur in at least two distinct integrons. The reference IS*Pa21* (AY920928) sequences embedded in *bla*_{GES-9} and *aacA7* sites are located on an MRI integron In109. When compared to the pR388 MRI sequence, In109 has the following SNP profile: SNP_{263c}+ SNP_{268g} + SNP_{288a} (Figure A10). The integron integrase In163 sequence for reference IS*Pa21* (*aadA7*; AY660529) is not available. Thus its SNP subgroup could not be determined. IS*Pa21*-like elements are associated with Int SNP_t and SNP_c subgroups (Figure 3.7).

Finally, IS*Pst6* group is overrepresented in PCI within the *Ps. stutzeri* complex. The recovered IS*Pst6*-like sequences were from the domestic and hospital environmental isolates FM3, FM1, P05, and P13. IS*Pst6*-like elements in clinical *Ps. aeruginosa* isolates JIP117 and SSI2.84 are predicted to reside in IntIs distinct from the two SNP sets seen in this study. *Ps. stutzeri* is thus believed to be a source for IS*Pst6*-like elements (Figure 3.7).

2) *Where both MRI and PCI exist in the same cell, do IS1111-attC elements jump between these different integrons?*

Two domestic environmental Pseudomonads, namely *Ps. alcaligenes* str. S3 (iso. JM2) and *Ps. alcaligenes* str. Y34 (iso. JM6), were chosen for fosmid construction. Southern hybridization data indicated that these two strains contain both MRI and PCIs, however PCR and sequencing efforts failed to recover these sequences. Attempts to PCR amplify *IS1111-attC* resulted in a banding pattern of various sized products suggesting that there may be more than one *IS1111-attC* element present. *IS1111-attC* PCR banding patterns differed in these isolates, suggesting that these two isolates have differences in their arrays and were hence chosen as candidates for fosmid construction.

3) *Was a single class 1 integron borne on a mobile genetic element responsible for the spread of ISPst6-like elements into human clinical isolates JIP117 and SSI2.84?*

Two epidemiologically unrelated, clinical *Ps. aeruginosa* isolates, SSI 2.84 and JIP117, were of particular interest. Partial *IS1111-attC* sequences were recovered from both clinical isolates suggesting a recent acquisition of the IS into the clinical system. NCBI BLAST analysis showed that the *IS1111-attC* sequences from isolate SSI 2.84 (Figure A8) and JIP117 were 99% (841/845 bases) and 99% identical (849/851 bases) identical to *Ps. stutzeri* strain ATCC 14405 *ISPst6* respectively. However, the MGEs (integron and plasmid) responsible in this acquisition is unknown.

3.3.9 Identifying individual fosmid clones that are positive for *IS1111-attC* elements

Each column of eight distinct fosmid clones over a number of 96-well plates per library was pooled for southern hybridization screening. This pooled column screening approach, with two rounds of hybridization, was used to efficiently identify individual fosmid clones positive for *IS1111-attC* elements (Figure A12 & Figure 3.8).

JM2 fosmid library, column 44 (C44) gave a positive signal with the *IS1111-attC* probe. All 8 clones comprising this pooled column were spotted out individually and clone G was identified as the only clone that produced an *IS1111-attC* hybridization signal. This fosmid clone was annotated as JM2C44G. Similarly for the JM6 library, column 39 (C39) was positive for an *IS1111-attC* element. When clones in C39 were screened individually, clone

D (JM6C39D) and H (JM2C39H) were positive for an IS element. In particular, clone JM6C39D hybridized strongly, matching the relative intensity of the positive controls *Pst587*, JM6 gDNA, *Ps. aeruginosa* MB216.2B as well as *Ps. aeruginosa* PAO1.

LS5 library, column 2 (C2) clones were screened individually, and clones E and F produced a faint hybridization signal when probed against an *IS1111-attC* probe. These were annotated as LS5C2E and LS5C2F. Note that hybridization negative controls, *E. coli* pR388 and *Ps. strameia* KM91 did not produce a hybridization signal as expected.

For clinical libraries SSI2.84 and JIP117, individual fosmid clones containing *IS1111-attC* elements were also identified by southern hybridization screening. SSI2.84C9 clones G and in particular H were *IS1111-attC* positive. Interestingly, for the JIP117 clone library a pooled column screen approach resulted in 38 out of 46 columns testing positive for an *IS1111-attC* element. Four columns, C9, C22, C34 and C47 were selected at random and all clones were spotted out individually and re-screened for the presence of *IS1111-attC* elements. C9 contained two positive clones C9F and C9H, C22 has three positive clones, C22A, C22C and C22E. C34 had five clones C34A, C34C, C34D, C34E, C34F (Figure 3.8B). Column 47 was interesting because initially the pooled clones gave an IS signal, but screened individually, no clone gave an *IS1111-attC* signal when hybridized with this IS probe. Therefore, the primary C47 result is believed to be due to non-specific binding of the probe to the DNA template.

Based on the hybridization data, five fosmid clones JM2C44G, JM6C39D, LS5C2E, JIP117C34A and SSI2.84C9H were selected for subsequent analysis involving restriction digestion (Figure A14) and screening for MRIs (Figure A13) and *IS1111-attC* elements (Figure A15). The five clones were also sequenced by Macrogen (Korea). See Table 3.6 for summary information regarding *IS1111-attC* screening results, genome fold coverage, contig number and total kb of sequence obtained, for these fosmids.

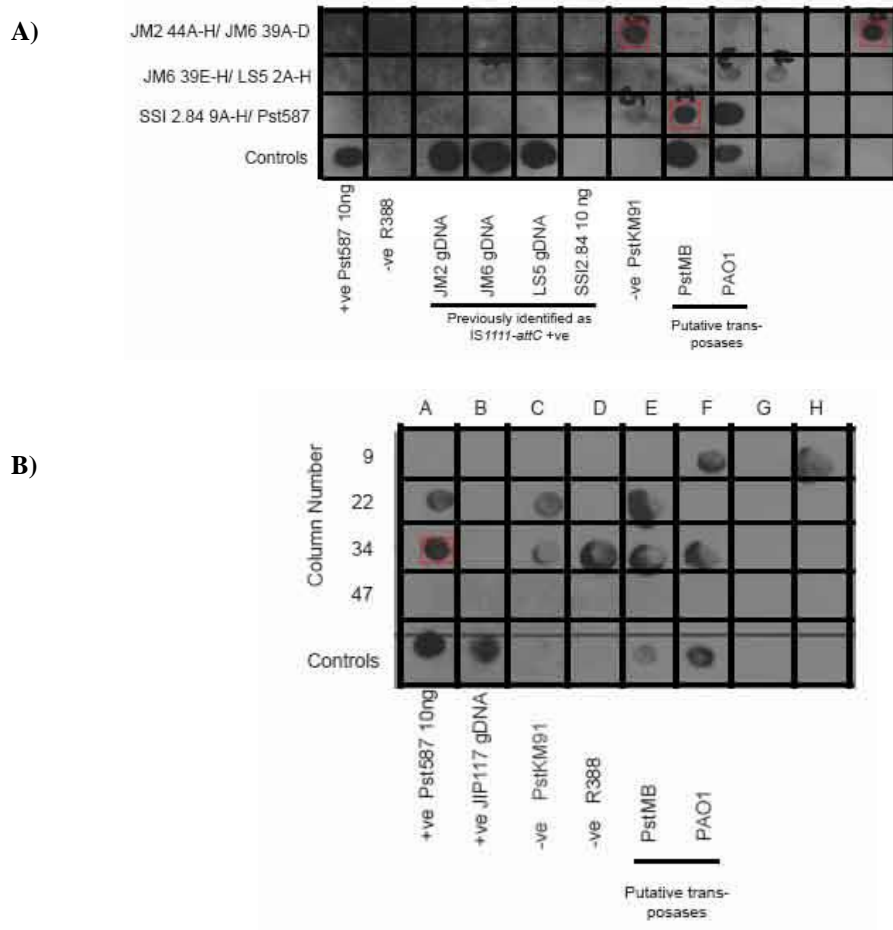


Figure 3.8: Identifying individual fosmid clones containing an IS1111-attC across five fosmid libraries.

A) JM2 C44, JM6 C39, LS5 C2 and SSI2.84 C9

B) JIP117 C9, C34, C47. The probe used was an intact ~1.3-kb ISPst6 generated from Pst587 gDNA. Red boxes denote fosmid clones of interest.

Table 3.6: Summary of fosmid libraries

	Strain ID	Total fosmid clones	Fold-genome coverage	Clone of interest	IS1111-attC presence		# of raw sequence reads	# of sequences reads used in assembly	Contigs in assembly	Fold coverage of fosmid	Total kb assembled
					Hybe	PCR					
Clinical	JIP117	384	1.92 ^a	JIP117C34A	+	+	384	273	6	6.1	32.769
	SSI2.84	480	2.40 ^a	SSI284C9H	+	+	384	255	2	7.4	35.134
Non-Clinical	JM2	384	2.40 ^b	JM2C44G	+	+	384	325	3	6.1	50.569
	JM6	480	3.00 ^b	JM6C39D	+	+	384	322	3	6.6	39.755
	LS5	192	1.20 ^c	LS5C2E	+	+	384	293	4	7.2	32.483

The fold-genome coverage was calculated based on an average insert size of 30-kb (see Appendix A6.2) and assuming a genome size of a) 6 Mb for *Ps. aeruginosa* (PAO1 reference), b) 4.8 Mb *Pseudomonas alcaligenes* (NBRC 14159 reference) c) 4.8 Mb *Stenotrophomonas maltophilia* (K279a reference). Fold coverage of fosmids = number of reads used in assembly x average length of reads = number of bases of sequence. Fold coverage between 6-8x means confidence in sequence coverage for fosmid.

3.3.10 Annotations of assembled contigs

A total of ~190-kb of sequence information was obtained from the five clone libraries. The provided assemblies did not contain the anticipated IS1111-attC sequences (Figure A16). To validate the contig assemblies, all were *de novo* re-assembled from raw reads. See appendix figure A17 for schematic figures of assembled *de novo* contigs. The assembly parameters were set to maximum sensitivity as per Appendix Table A5. For all of the five fosmid *de novo* assembly produced more than one contig (Table 3.7).

Table 3.7: *de novo* assembly summary

Pseudomonad group	Isolate IS	Fosmid clone ID	# of <i>de novo</i> contigs	<i>de novo</i> contig size (kb)
Clinical	JIP117	C34A	2	17.8 & 11.6
	SSI2.84	C9H	0*	0*
Environmental	JM2	C44G	0*	0*
	JM6	C39D	1	38.602
	LS5	C2E	2	10.8 & 21.66

*JM2C44G and SSI2.84C9H contigs could not be *de novo* assembled to produce a continuous read. Failure to assembly into contigs occurs mostly due to the lack of sequence homology between the individual contig ends (Baker, 2012).

For all individual contigs as well as continuous contigs, open reading frames (*orfs*) were identified using Glimmer plug-in in Geneious with default setting of bacterial “*atg*” start codon. A total of 178 *orfs* were analysed and identified using BLASTn NCBI search (Table A6). The assemblies were mapped to completed reference genomes (Section 3.3.11) to see if they were ‘missing’ IS1111-attC reads and to validate that isolates SSI2.84, JIP117, JM2 and JM6 were Pseudomonads and LS5 was *Stentrophomonas sp.*

3.3.11 Genome mapping of contigs confirms the absence of IS elements in fosmid clones

All contigs were mapped against completed reference genomes from *Ps. aeruginosa* PAO1 (NC_002516), *Ps. mendocina ymp* (NC_009439), *Ps. resinovorans* (NC_021499) and *Stenotrophomonas maltophilia* K279a (NC_010943). This mapping was undertaken to 1) validate the construction of the contigs and predicted *orfs*, 2) look for areas of sequence variation between contigs and genomes and 3) place the contigs in respect to the genome, i.e. GI vs chromosomal framework. Contigs had extensive synteny to published genomes (Figure A18) and no hot spots for integron or IS1111-attC insertions were identified. In all cases no

IS1111-attC and *intI1* sequence homologs were identified (Table 3.8). Three possibilities would account for the lack of *IS1111-attC* sequences in these fosmid:

- 1) Sent the wrong data set: All contigs were validated to be of *Pseudomonad* and *Stenotrophomonas* origin. This is consistent with these isolates being identified as such by 16S rRNA DNA sequencing. Thus this does not explain the lack of *IS1111-attC* sequence data in the sequenced fosmids.
- 2) Fosmid sequence is systematically biased against IS regions: The precedent for IS regions being difficult to assemble and sequence is based on the doctorate works by Tetu (2007) and Wilson (2007). In this current study, all raw reads were mapped against re-assembled scaffolds and the relative contigs sizes were accounted for.
- 3) Wrong fosmid clones were sequenced: whilst the contigs had no regions with integron signatures they did have foreign DNA signatures.

To recover fosmids containing *IS1111-attC* elements and *intI1*, perhaps more than 384 fosmids would also need to be screened per fosmid library, thereby increasing the overall fold coverage for large genomes such as those of the *Pseudomonadaceae* family.

Table 3.8: Synteny of fosmid contigs to best reference genomes

	Fosmid ID	Size recovery	Mean % identity*	Extent of synteny to best reference genome	Relevant features
Clinical	JIP117C34A	~30-kb	95-99%	2 loci separated by 3 Mb in PAO1 reference	No <i>IS1111-attC</i> or <i>intI1</i> homologue
	SSI284C9H	~35-kb	95-99%	1 locus with 4-kb deletion in PAO1 reference	No <i>IS1111-attC</i> or <i>intI1</i> homologue
Environmental	LS5C2E	~35-kb	95-99%	1 locus with 1.5-kb deletion in K279a reference	No <i>IS1111-attC</i> or <i>intI1</i> homologue
	JM2C44G	~49-kb	~60-65%	1 locus in NBRC106553/ymp extensive region of synteny identified	No <i>IS1111-attC</i> or <i>intI1</i> homologue
	JM6C39D	~39-kb	~60%	1 locus in NBRC 106553/ymp: no extensive regions of synteny identified	No <i>IS1111-attC</i> or <i>intI1</i> homologue

*Mean % identity to best reference genome. Synteny was based on genome mapping of all contigs against reference genomes. JIP and SSI contigs were mapped against 6 Mb for *Ps. aeruginosa* (PAO1) reference genome. LS5 contigs were mapped against 4.8 Mb *Stenotrophomonas maltophilia* (K279a) reference. JM2 and JM6 contigs were mapped to both 5 Mb *Ps. mendocina* ymp (NC_009439) and 6.3Mb *Ps. resinovorans* NBRC 106553 (NC_021499) genomes.

3.4 Discussion

3.4.1 IS1111-attC elements are over-represented in Pseudomonadaceae

Insertion sequences are commonly used as epidemiological and phylogenetic markers for the study of internal genome dynamics (da Silva Rabello *et al.*, 2010; Cerveau *et al.*, 2011; Kenna *et al.*, 2006; Vandelannoote *et al.*, 2014; Vorská *et al.*, 2001). In this study, IS1111-attC elements were used as an epidemiological and phylogenetic marker for between-population dynamics in opportunistically pathogenic γ -Proteobacteria, of the families *Pseudomonadaceae* and *Enterobacteriaceae*. I examined three strain collections that were defined by distinct aspects of their evolutionary ecology. Two *Pseudomonad* strains collections, from distinct ecological and phylogenetic backgrounds (clinical & environmental) and a clinical *Enterobacteriaceae* collection were screened for IS1111-attC elements. I examined two aspects of IS1111-attC distribution across these strain collections: i) Frequency - the number of IS1111-attC observations per independent bacterial strains, and; ii) Diversity - the number of independent IS1111-attC elements per strain collection. Collectively this approach provided insight into the nature of IS reservoirs in different bacterial communities; the extent and manner of movement between communities; and the extent to which they spread within communities.

IS1111-attC frequency shows that strains of the family *Pseudomonadaceae* are more likely to harbour the IS than *Enterobacteriaceae* (Table 3.4). To determine the statistical significance of this finding, a larger strain collection is required, however a trend is apparent. Comparable IS1111-attC frequencies were seen in both environmental and clinical *Pseudomonad* groups. Data were primarily based on hybridisation analyses, with a subset confirmed by sequence recovery. Both sequencing and hybridization data indicated that IS1111-attC elements were found in different contexts in the two *Pseudomonad* collections. The environmental group IS1111-attCs were observed in a chromosomal framework, whereas in the clinical *Pseudomonad* group they were in an accessory (mobile) context.

IS1111-attC diversity shows that there are differences in the distribution of different IS1111-attC elements within and between the three strain collections. Within the *Pseudomonadaceae* that reflect both ecological and phylogenetic dissimilarities. The clinical *Pseudomonas* set (*Ps. aeruginosa* lineage) was shown to have low IS1111-attC diversity, whereas the

environmental *Pseudomonas* set (*Ps. stutzeri* lineage) had high IS1111-attC diversity (Figure 3.3 & 3.4). The low diversity, combined with high frequency, implies a small number of independent IS elements circulating within the clinical population. This interpretation is consistent with the observation that IS1111-attC elements are MRI-borne within the clinical *Pseudomonas* group. In comparison, high diversity of IS1111-attC implies a large number of independent IS1111-attC elements in the environmental pool and implies environmental *Pseudomonads* are IS1111-attC reservoirs. Given that no repeat IS1111-attC observations were made in the environmental *Pseudomonas* group, this suggests very low level of IS spread within this group (Tetu & Holmes, 2008; Figure 3.4).

One interesting aspect of IS1111-attC diversity in the environmental *Pseudomonadaceae* is that it shows a punctuated distribution. ISs from the same genomovar are more closely related than those of distinct genomovars. For instance, the *ISPst6* subgroup is overrepresented in PCIs in *Ps. stutzeri* species complex genomovar 2 (Tetu & Holmes, 2008). The PCIs also show a similar punctuated distribution within the environmental population based on integron integrase sequence relationships (Wilson, 2007). Using the terminology of comparative species genomics, the distribution of IS1111-attC diversity in PCI/genomovars is consistent with their being nascent character genes (Figure 1.13). Although, there are no examples of two strains with the same IS/gene cassette association as would be predicted for ‘true’ character genes, the pattern is indicative of lineage-specific associations arising through repeated independent formation of IS1111-attC/ gene cassette associations within PCIs, rather than movement of the hybrid associations between the PCI of strains in a genomovar. Taken together, this indicates that in broad terms the environmental *Pseudomonads* are the reservoir for IS1111-attC elements, but at genomovar level we can also predict the specific *Pseudomonas* lineage that is the source for each clade of IS1111-attC elements. This capacity may provide insights to gene cassette flow between integrons.

In *Pseudomonas aeruginosa* the presence of IS1111-attC elements is predicted to be dependent on MRIs, since they lack PCIs. I observed that MRIs were similarly overrepresented in the clinical *Pseudomonads* and *Enterobacteriaceae* strain collections relative to the environmental *Pseudomonads*. However, IS1111-attCs were not seen at the same frequency in these two clinical collections. Two types of IS1111-attC elements, *ISPa21* and *ISUnc1* were seen in association with MRIs in my clinical *Ps. aeruginosa* strains and

none in the clinical *Enterobacteriaceae* collection. I postulate that the ability of an IS1111-attC to specifically move between gene cassettes within a cell means they essentially act as gene cassette tags and allow study of gene cassette flow between integrons and across bacterial populations in different habitats. In the clinical context, IS1111-attC elements can thus be used as evolutionary and epidemiological tools for the study of integron-mediated gene transfer in pathogenic Bacteria. Where an IS is observed in an MRI, its original source is predicted to be a Pseudomonad population with PCI. This source population for the IS can be inferred from its sequence relationship to IS1111-attC groups with known associations to PCIs, thus indicating historical movement of that integron.

3.4.2 Identifying IS1111-attC source-sink relationships

In addition to the two examples described above, *in silico* database searches indicate further examples of IS1111-attC elements in MRIs (Table 3.9). Importantly these include strains from outside the genus *Pseudomonas* and in most such cases the IS-gene cassette combination observed in a non-Pseudomonad has also been observed in an integron in a Pseudomonad. For example the *aadA1*-IS*Unc1* association has been seen in *E. coli*, *A. baumannii* and *Ps. aeruginosa*. The higher frequency of IS1111-attC in Pseudomonads, combined with the co-occurrence of identical IS/gene cassette associations in the clinical isolates and various *Pseudomonas* strains suggests that IS1111-attC elements on MRIs have spread from *Pseudomonas* in clinical environments (primarily *Ps. aeruginosa*). My data show that an environmental strain identified by 16S sequence as *Ps. putida* has an IS1111-attC in a PCI (Figure 3.4 & Table A2) with high sequence relationship to the IS*Unc1* element that has spread through one group of *Ps. aeruginosa* on class 1 integrons and onto enterics.

MRI carrying pathogenic *Pseudomonas putida* isolates have been documented (Lee *et al.*, 2002; Lombardi *et al.*, 2002; Ramos-gonzález *et al.*, 2006). Database analysis revealed that a clinical *Ps. putida*, HB3267 (CP003739.1), contains an MRI annotated as having an IS116/IS110/IS902 element. Closer examination revealed that this IS is embedded in a quaternary ammonium resistance gene (*qacE*) attC and is identical (1376/1376 bp) to IS*Knp4* (Post & Hall, 2009). Furthermore, the insertion point in attC_{qacE} is identical to that of IS*Knp4*. This IS should be annotated as an IS1111-attC member.

Table 3.9: Examples of IS1111-attC elements in enteric bacteria found from in silico search

Group ^{a)}	Database ID	Organism	IS	attC context	Genetic context	Size (bp)	% identity to reference	Reference
ISUnc1-like	AY139602	UnCul plasmid Psp7	ISUnc1	<i>aadA1</i>	InXX intI1	1381	100%	Tennstedt <i>et al.</i> (2003)
	DQ302723	<i>Ps. aeruginosa</i>	ISUnc1	<i>aadA1</i>	InXX intI1	1382	98.9%	Unpublished
	FJ267652.1	<i>Ps. aeruginosa</i>	ISPa21	<i>aac(6')</i>	<i>intI</i>	1382	98.8%	Unpublished
	AM932676.1	<i>E. coli</i>	ISUnc1	<i>aadA1</i>	<i>MRI intI1</i>	1379	99.7%	Kadlec & Schwarz (2008)
	KF921554.1	<i>E. coli</i>	ISUnc1-partial	<i>aadA1</i>	<i>MRI intI1</i>	1322	99.8%	Unpublished
	EU089665.1	<i>E. coli</i>	ISUnc1-partial	<i>aadA1</i>	<i>MRI intI1</i>	1071	99.6%	Moura <i>et al.</i> (2007)
	JN253504.1	<i>Ac. baumannii</i>	ISAb3	<i>aadA1</i>	<i>MRI intI</i>	1375	99.9%	Karah <i>et al.</i> (2011)
	KC675185.1 ^{d)}	<i>En. cloacae</i>	ISPa21	<i>aadA1</i>	<i>MRI inti1</i>	1381	99%	Unpublished
	KC675185.1 ^{c)}	<i>En. cloacae</i>	ISPa21	<i>bla_{OXA10}</i>	<i>MRI intI1</i>	1381	99%	Unpublished
ISPa21-like	AY920928	<i>Ps. aeruginosa</i> DEJ	ISPa21	<i>bla_{GES-9}</i>	In109 intI1	1374	100%	Poirel <i>et al.</i> (2005)
	AY920928	<i>Ps. aeruginosa</i> DEJ	ISPa21	<i>aacA7</i>	In109 intI1	1374	100%	Poirel <i>et al.</i> (2005)
	AY660529	<i>Ps. aeruginosa</i>	ISPa21	<i>aadA7</i>	In162 intI1	1375	99%	Unpublished
ISPaX1-like	AM296017	<i>P. aeruginosa</i>	ISPaX1	<i>aac(6)-II</i>	InXX intI1	1376	100%	Unpublished
ISPaX2-like	DQ522236	<i>Ps. aeruginosa</i>	ISPaX2	<i>aadB</i>	InXX intI1	1377	100%	Unpublished
	AM749812.1	<i>Ac. baumannii</i>	ISPa21	<i>aadB</i>	<i>MRI intI</i>	1377	100%	Toleman & Walsh (2007)
ISKnp4-like ^{b)}	EF408254	<i>K. pneumoniae</i>	ISKnp4	<i>qacH</i>	MRI intI1	1376	100%	Post & Hall (2009)
	JX275775.1	<i>En. cloacae</i>	ISPa21 ^{f)}	<i>Δqac1</i>	<i>MRI intI1</i>	1364	93.7%	Sonnevend <i>et al.</i> (2012)
	AJ704863.3 ^{e)}	<i>K. pneumoniae</i>	ISPa21 ^{f)}	<i>smr</i>	<i>MRI intI1</i>	1340	93.7%	Colinon <i>et al.</i> (2007)
	KF894700.1 ^{e)}	<i>K. pneumoniae</i>	ISPa21-partial ^{f)}	<i>smr</i>	<i>MRI intI1</i>	763	95.0%	Unpublished

Grey boxes highlight the reference sequence/s in each subgroup.

a) IS1111-attC subgroups are defined when the IS DNA >95% identical and protein >98% identical to the isoform founder

b) ISKnp4 is 97.7% identical to ISPaX2.

c) Match IS in *aac(6')* in FJ267652.1 *Ps. aeruginosa*

d) Match IS in *aadA1* in FJ267652.1 *Ps. aeruginosa*

e) Matches 100% to JX275775.1 in *En. cloacae*

f) Requires reclassification as ISKnp4-like elements.

To validate the hypothesis that *Ps. putida* is a source of IS*Unc1*, and to determine its PCI context, there is need for a comprehensive *Ps. putida* collection as seen with the *Ps. stutzeri* set (Tetu & Holmes, 2008).

The acquisition of IS1111-attC elements by MRIs does not appear to be restricted to any phylogroup and the intact is stably maintained. For instance, IS*Unc1*-like elements were recovered from *Chryseobacterium meningosepticum* PO6 and *Delftia* sp. LS8 (Figure 3.4). IS acquisition by *Ch. meningosepticum* is postulated to have occurred by a mobile class 1 integron (Figure A4) via a previous interaction with a *Ps. putida* reservoir. This is particularly interesting given that *Ch. meningosepticum* has been proposed as the primary source for metallo-beta-lactamase resistance genes in clinical *Ps. aeruginosa* (Bellais *et al.*, 2000; Poirel *et al.*, 2000). Here I provide additional support to the idea of active HGT between *Ch. meningosepticum* and Pseudomonads. Given that no MRI class 1 was detected in *Delftia* sp. LS8, IS*Unc1*-like acquisition here is postulated to have occurred via a different route. *Delftia* spp. are known to carry class 3 integrons (Xu *et al.*, 2007), thus the IS*Unc1*-like element may be residing in a class 3 mobile integron array. However, the distribution of class 3 integrons was not examined in this study as they are not commonly found in Pseudomonads. The low level of diversity in IS-gene cassette fusions suggests a low rate of loss or gain of IS1111-attC by MRI and suggests long-term stability of the IS-gene cassette within a clinical population.

3.4.3 Class 1 integron subgroups have different exposure to PCI gene cassette reservoirs

Data presented in this study suggests that IS1111-attC subgroup elements are associated specifically with integrons (Figure 3.9). I conclude that different Class 1 integron subgroups appear to be circulating differently, sampling from different ecological niches and thus have differing probabilities for IS1111-attC acquisition. IS1111-attC spread thus appears to be a dependent on the plasmid, transposon and the MRI in which it is located. Given that MRIs are seen on diverse replicons, we can expect some constraints on their movement, due to plasmid incompatibility mechanisms and host range. This is expected to impose some constraints on the exchange and distributions of IS1111-attC and of gene cassettes. Also selective pressure may result in the re-arrangement of gene cassettes on the plasmid and may produce different IS-gene cassette fusions.

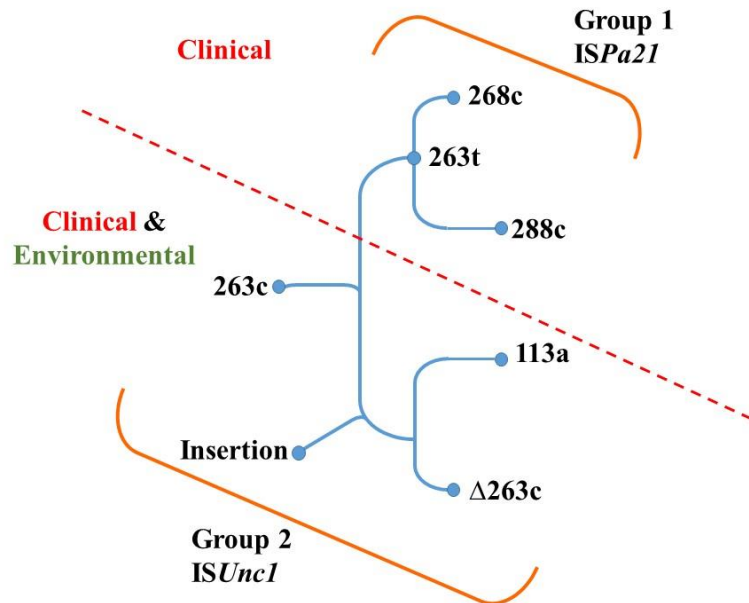


Figure 3.9: SNP tree of class I integrons.

Each tip represents a distinct variant of *intI1*, defined by accumulated mutations. IS elements have a dichotomous distribution that suggests the class I integrons have diverged into two clonal pools that are circulating differently.

3.4.4 A proposed model for IS1111-attC spread across barriers

Divergent IS1111-attC copies have been detected in varying integron arrays and suggests a very recent, and possibly ongoing, expansion. IS1111-attC distribution indicates that horizontal transfer occurs frequently and possibly simultaneously in different genomes. The success of IS1111-attC spread between different genera is postulated to be dependent on the fitness of the MRI carrier (Figure 3.10). Maintenance of the tight association between IS lineages and PCIs in *Pseudomonas* genomovars is believed to reflect vertical inheritance (Figure 3.10A) and PCIs coinciding with 16S rRNA DNA evolution (Figure 1.13). If the environmental *Pseudomonas* acquires an MRI containing replicon (Figure 3.10B), then the MRI can potentially acquire an IS1111-attC (Figure 3.10C). HGT of the newly formed IS/MRI association is then able to drive the dissemination of the newly embedded IS1111-attC into new populations (clinical *Pseudomonas*) (Figure 3.10D). Due to antibiotic selective pressures, the replicon is then shared between different clinical populations (Figure 3.10E), such as *Ps. aeruginosa* and *Enterobacteriaceae*. Thus, clinical pathogens become a sink for MRI mediated spread of IS1111-attC elements. This process could equally apply to gene cassette movement across ecological and phylogenetic contexts.

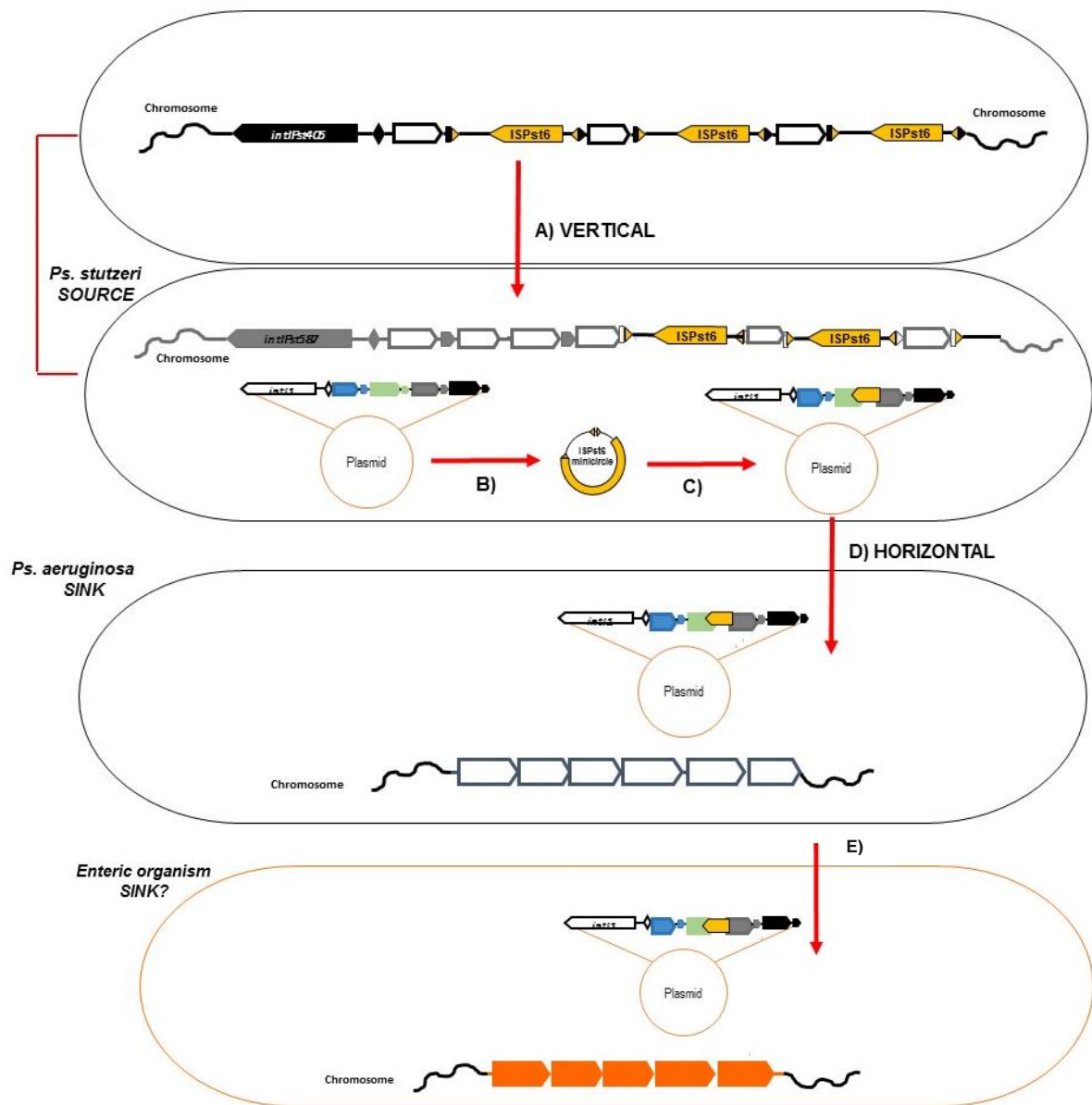


Figure 3.10: Postulated flow of IS1111-attC elements between environmental and clinical strains.

- A) IS1111-attC spread within the *Pseudomonas stutzeri* complex is postulated to occur by vertical inheritance.
- B) B-E) IS1111-attC spread across ecological and molecular-genetic barriers is postulated to occur by MRIs.

3.4.5 Further work

Why are multiple imperfect IS1111-attC replicates seen in the PCI-containing strains and only single perfect IS1111-attC replicates seen across the MRI-containing strains? The varying distribution and copy number of IS1111-attC between PCIs and MRIs, suggests that IS1111-attC translocation activity is not constant over time. This reflects the situation with gene cassettes generally where chromosomal integrons contain multiple copies of imperfectly replicated attC sites (with very diverse ORFs) but MRIs tend to contain single copies of near-perfect replicated attC sites with the same (or near perfect) gene. To explore this further, we need to consider the mode of IS translocation and host factors associated with translocation success. Based on the detection of IS*Pst6* minicircles (Tetu & Holmes, 2008) and work conducted on other members of the IS110 family (Partridge & Hall, 2003; Prosseda *et al.*, 2006), I predict that the IS1111-attC transposase in PCIs has a baseline activity that drives translocation of IS1111-attC elements through the formation of an IS mini-circle intermediate (Chapter 4). Insertion into various attC sites, suggests that the target mechanism does not depend on sequence conservation at or around the insertion point. Instead, I postulate that the recognition of a conserved attC structure drives the success of IS translocation, as seen with integron integrase preference for attC hairpins (Johansson *et al.*, 2004) (Chapters 4 and 5).

Chapter 4 Examining the *in vivo* translocation activity of IS1111-attC elements and mechanisms of recombination.

4.1 Introduction

My data, and all published observations, show that IS1111-attC elements are exclusively observed in attC sites (Poirel *et al.*, 2005; Post & Hall, 2009; Tetu & Holmes, 2008). Data presented in Chapter 3 demonstrated that IS1111-attCs are over-represented in environmental Pseudomonad chromosomal arrays, and also affirmed an interesting pattern that had previously been reported (Tetu & Holmes, 2008). Within the *Pseudomonas* 1 sample group the ISs are typically found in multiple near-identical copies whereas within either the *Pseudomonas* 2 sample group or the *Enterobacteriaceae* sample group the ISs are typically found in single copy. This pattern is postulated to reflect repeated insertion of the IS elements having occurred in some contexts and not in the others. This suggests that there are molecular-genetic constraints to IS activity that differ between these populations, potentially acting as further barriers to the net rate of gene movement between these populations. This prompted me to explore differences in the insertion context in each group and what clues this might give as to the mechanisms of translocation.

Table 4.1: Summary of the correlations between patterns of IS1111-attC occurrence and aspects of target sites, context and cell host (data from Tetu & Holmes, 2008)

Pattern of IS element	Biological context of the IS element					
	attC context		Replicon context		Cell context	
	Ps.-type	Other attC ^a	Chromosome	Episome	<i>Pseudomonadaceae</i>	<i>Enterobacteriaceae</i>
Multiple copies in cell	√√√	0 ^b	√√√	0 ^b	√√√	0
Single copy in cell	0 ^c	√√	0 ^c	√√√	√√ ^c	√√

√√√= 95-100% of instances of that pattern (e.g. multiple copies) are seen in the indicated context; √√=60% cases; √= 5-10% cases

a: Non-*Pseudomonas*-type attC sites in which IS1111-attC have been observed

b: *Ps. aeruginosa* DEJ (AY920928) contains two ISPa21 copies in other attC sites (*bla*_{GES-9} & *aacA7*) on a plasmid-borne *int1*

c: *Pseudomonas* strains (gv.1 & 7) with remnant footprints are an exception. They have no intact IS1111-attC copies and a single pseudocopy in a chromosome.

IS1111-attC elements have been seen in a total of 39 target sites contexts (Poirel *et al.*, 2005; Post & Hall, 2009; Tetu & Holmes, 2008, Table 3.9). Over 70% of these are in *Pseudomonas*-type attC sites, the remainder are in at least three distinct structural subclasses of attC sites (Figure 4.1). However, it is only in PCIs and *Pseudomonas*-type attC sites where

they have been observed in multiple adjacent copies (Table 4.1). This suggests that IS activity is different in different contexts. This phenomenon is unlikely to reflect absolute specificity to the *Pseudomonas*-type attC since ISs are also observed in other attC sites, but it might reflect much greater levels of activity (higher affinity). However the *Pseudomonas*-type attC site distribution co-varies with other factors including the replicon (a chromosome) and the host cell (*Pseudomonas* 1 group).

IS1111-attCs have also been seen on multiple replicons, including chromosomes and episomes. The majority are in the chromosome in *Pseudomonas*-type attC sites (Tetu & Holmes, 2008). The remainder are on class 1 integrons, in at least 19 different MRI arrays, carried on a range of translocators and replicons. Finally, IS1111-attCs are seen only in Proteobacteria to date. Within *Pseudomonadaceae*, IS1111-attCs are observed in MRIs in clinical *Pseudomonas*, whereas in environmental Pseudomonads, they are primarily observed in a PCI context. Outside of the *Pseudomonadaceae*, IS1111-attC are found to be associated with MRIs rather than chromosomal integrons/arrays. Here I aimed to explore which aspects of the molecular-genetic system, such as attC site preference, replicon preference and/or host cell preference, are able to influence IS1111-attC acquisition and insertion at specific sites.



Figure 4.1: Diversity of attC sites in which IS1111-attC have been observed. Alignment of *Pseudomonas*-type (black bar) and “other attC” (red bar) housing IS1111-attC elements. At least three distinct attC structures are seen in the “other attC” group (red numbers). Insertion site of IS1111-attC occurs between positions h and i (yellow arrow). Adapted from Tetu and Holmes (2008).

Tetu and Holmes (2008) observed a partial translocation event of an IS1111-attC element into a *Pseudomonas*-type attC, located on a fosmid, in an *E. coli* host. They defined partial translocation as being when; i) only ISPst6 IR_R junction in a plasmid trap in *E. coli* is observed by means of PCR screening but not the IR_L junction, and ii) no attC-trap plasmid, consistent with containing a single intact copy of ISPst6 in *E. coli*, is observed (Tetu &

Holmes, 2008). Their study highlighted the likelihood that *Pseudomonas* host machinery and host factors, such as expression levels of the native IS1111-attC T_pase, constrain IS1111-attC movement. In order to address which *Pseudomonas* host factors might be important in IS1111-attC translocation, we need to consider how ISs translocate and what factors are known to influence this. The observed spread of IS1111-attCs across PCI arrays may be explained by an opportunistic event, wherein a host cell was infiltrated by one IS1111-attC element at a time until all, or almost all, available PCI attC sites were occupied. Given that translocation rates of ISs are estimated to be rather low (10^{-5} - 10^{-8} per element per generation) (Souise *et al.*, 2013), the probability of acquiring an identical IS multiple times, and inserting it in the same location, diminishes as the number of insertion events increases. Hence a series of opportunistic events alone is an unlikely explanation for IS1111-attC distribution across PCI arrays. In a plasmid context, this may very well have occurred because only a single IS1111-attC is observed at a time. Instead I postulate that IS1111-attC distribution is more likely a feature of its translocation mechanism.

The relative homogeneity of IS1111-attC elements across PCI arrays (Tetu & Holmes, 2008) suggests that they translocate by a copy-&-paste mechanisms. During copy-&-paste, a single IS1111-attC molecule can enter a cell and expand within the genome (Figure 4.2A). However, multiple IS1111-attCs could reflect multiple independent acquisitions of a single molecule, or a single acquisition on a multi-copy plasmid. In contrast, IS1111-attC heterogeneity across a range of varying MRI attCs can be explained by cut-&-paste translocation. Here multiple IS1111-attC molecules must enter the cell and/or be formed in it during cut-&-paste translocation (Figure 4.2B). Note, however, that insertion sequences, such as IS2, translocate by both copy-&-paste and cut-&-paste mechanisms (Biel & Berg, 1984; Mahillon & Chandler, 1998). Similarly, IS1111-attC translocation may occur by both means; however, this remains to be determined.

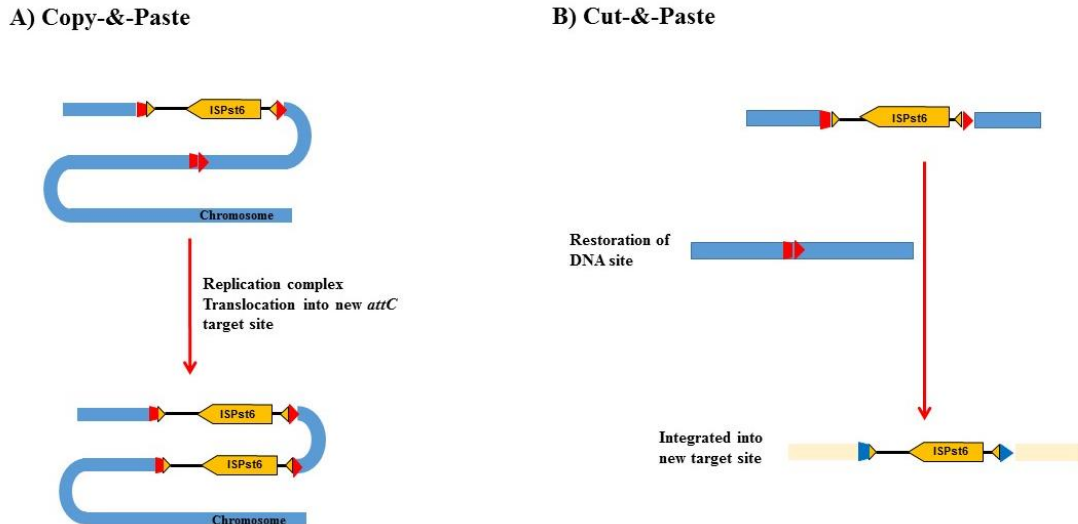


Figure 4.2: Conceptual models a) copy-&-paste and b) cut-&-paste for IS1111-attC translocation between attC sites.

- A) Topological flexibility of chromosomal DNA and host factors (Loot *et al.*, 2012) are postulated to play a vital role in bringing two PCI attCs within proximity of one another.
- B) Reduced number of target sites in MRIs, as well as accessory factors necessary for plasmid replication (Biel & Berg, 1984) are predicted to affect IS1111-attC translocation efficiency during cut-&-paste.

4.2 Aims

There are a number of precedents for the genomic context of IS elements influencing their success. These include DNA secondary structure, target site context and target site orientation (Partridge & Hall, 2003), mode of DNA replication, and accessory factors that help resolve the synaptic complex formed between the IS and its target DNA (Mahillon & Chandler, 1998). The relative importance of these influencing factors for any one IS is reflected by its translocation mechanisms.

The work presented in this chapter tests the following hypotheses:

- 1) IS1111-attC translocation requires accessory factors in *Pseudomonas*.
- 2) IS1111-attC elements preferentially insert into *Pseudomonas*-type attC sites.
- 3) The likelihood of repeat IS1111-attC translocation events is dependent on the copy number and context of the IS.

4.3 Methods

For general media and solution recipes, as well as PCR conditions, refer back to Chapter 2.

4.3.1 Construction of “*attC*-traps”

A *Pseudomonas attC*-site was amplified from *Ps. stutzeri* strain ATCC 17587 (BGC093) using primers ST96 and ST139 (Tetu & Holmes, 2008). The 219-bp PCR product consists of a 76-bp *Pseudomonas attC* site flanked by upstream (22-bp) and downstream (121-bp) sequences. This PCR product was cleaned using Qiagen PCR Purification Kit (as per manufacturer’s instructions). Similarly, an *attC_{aadB}* site was amplified from pUS21 using primers MZ09 and MZ10. Given the small product size of 100-bp, this PCR product was cleaned using ammonium acetate precipitation (see section 2.3.1).

To generate the synthetic *attC*-traps, 1 µg of purified PCR product was end-repaired using NEB Quick Blunt Mix (as per NEB instructions) and ligated overnight at 4°C with *SmaI*-cut pBBR1MCS2 using an insert to vector ratio of 5:1. Prior to ligation, digested pBBR1MCS2 vector was dephosphorylated using Antarctic phosphatase as per NEB instructions. 5 µL of the ligation reaction was transformed into chemically competent *E. coli* TOP10 cells by heat-shock at 42°C for 30 sec. Cells were recovered at 37°C for 1 hr in 200 µL LB and plated onto LB/Km⁵⁰ plates overnight at 37°C. PCR primers targeting the plasmid regions flanking the cloning site (NVC15 and NVC16) were used to screen for positive clones. Additionally, primer combinations using NVC15/NVC16 primers and *attC* primers ST96/ST139, MZ09/MZ10, were used to determine the orientation of the cloned *Pseudomonas*-type and classical *attC_{aadB}* sites respectively.

In total, two *Pseudomonas*-type *attC* and two classical *attC_{aadB}* traps were generated. The same site was cloned into pBBR1MCS2 in one or other orientation relative to the plasmids *ori* (*rep*) region (Figure 4.3A & B). The resulting vectors were termed pPsattC1 and pPsattC2, paadB_*attC*1 and paadB_*attC*2. Purified plasmid preparations were extracted from 50 mL LB overnight culture of *E. coli* JM109 cells (See section 2.7).

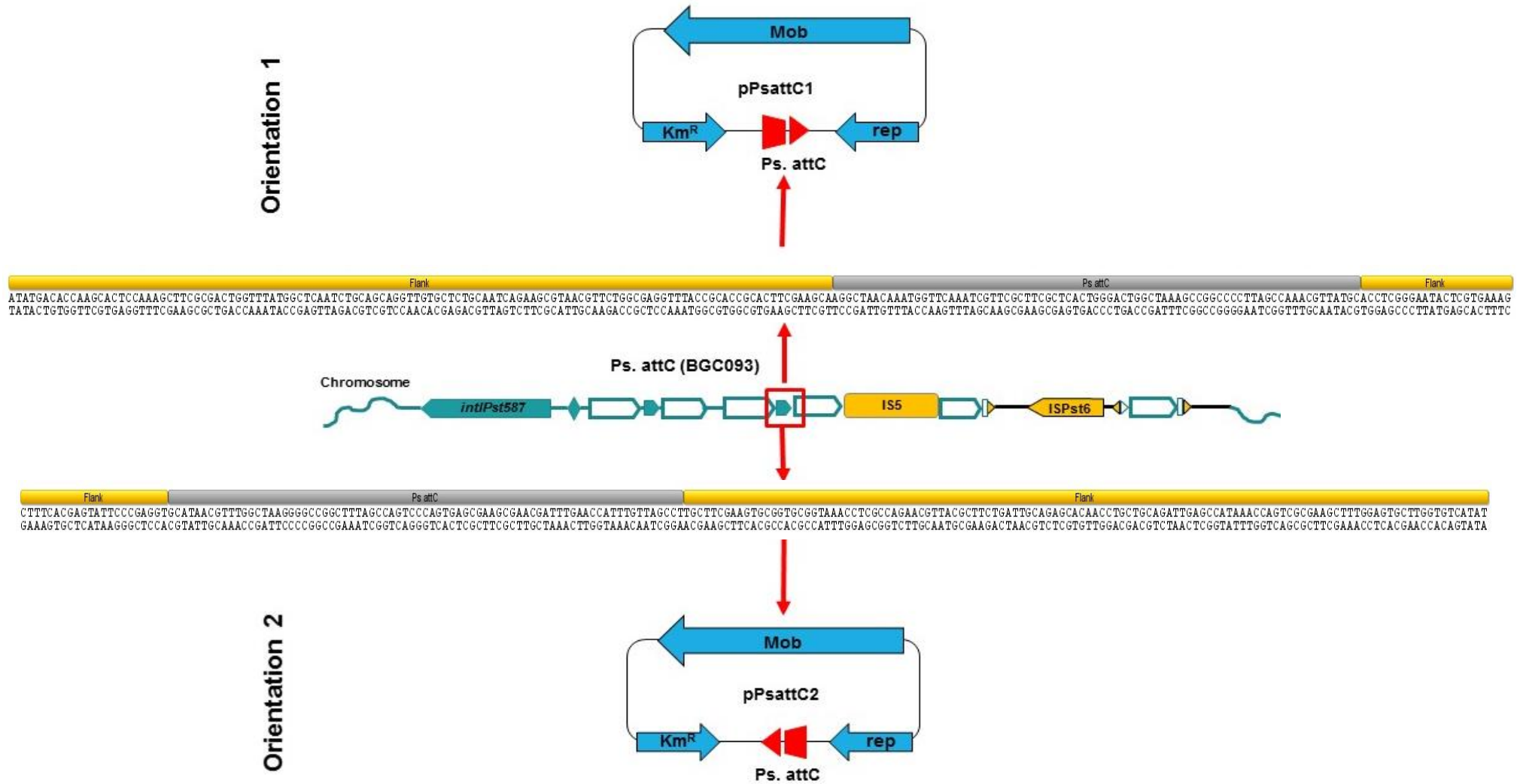


Figure 4.3A: Construction of *Pseudomonas*-attC traps.

Pseudomonas stutzeri ATCC17587 genomic DNA was used as PCR template with BGC093 specific primer ST96 and ST139 to give a 219-bp product consisting of the *Pseudomonas*-type attC site (76-bp) and flanking DNA (Red boxed attC site). Post PCR purification reaction, the product was cloned in one or other orientation in pBBR1MCS2. Each construct, pPsattC1 and pPsattC2 was used in subsequent natural transformation assays to detect IS1111-attC translocation.

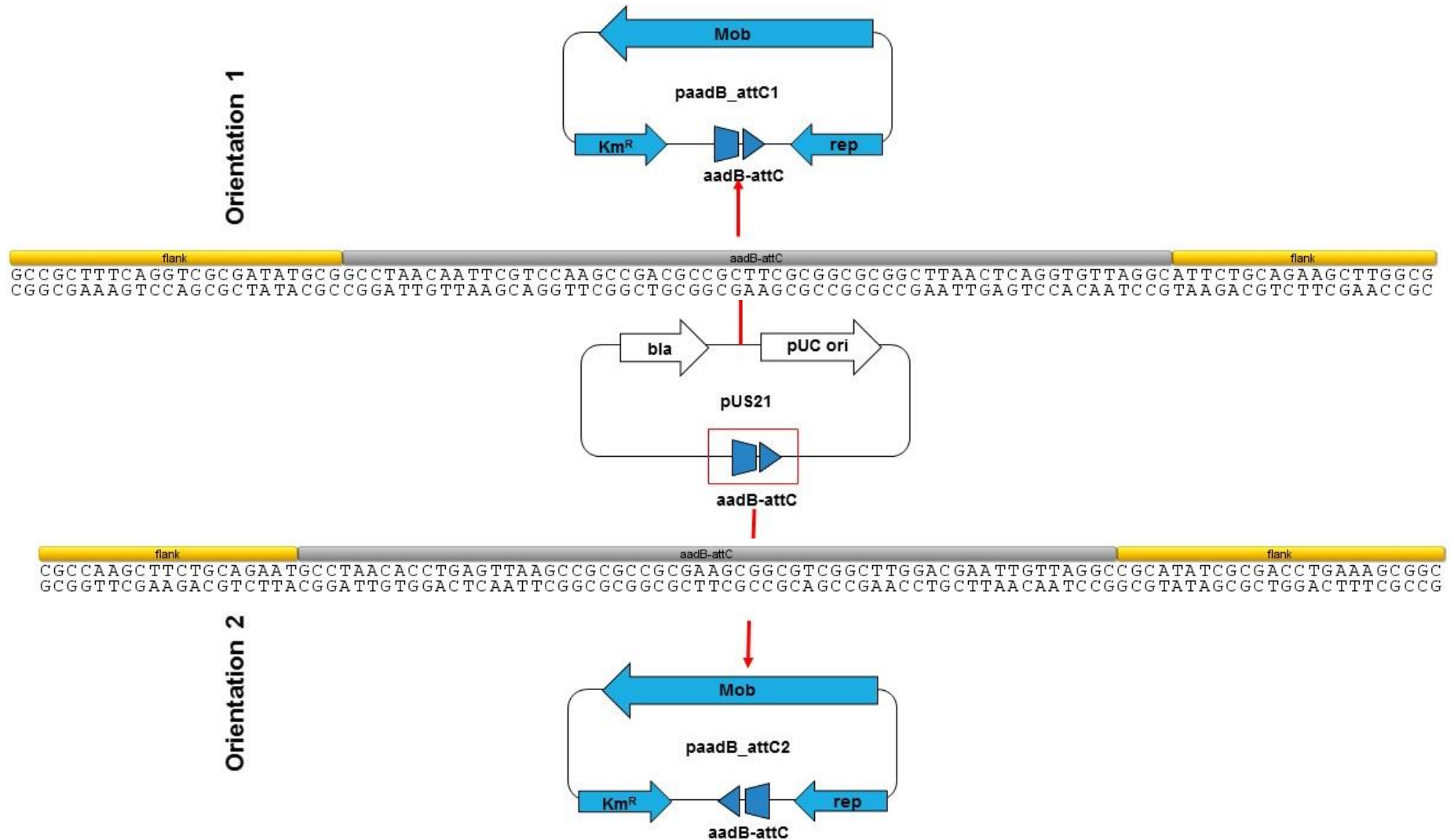


Figure 4.3B: Construction of classical *attC_{aadB}* traps.

Purified pUS21 plasmid DNA was used as PCR template with primers MZ09 and MZ10 to give a 100-bp product consisting of the classical *attC_{aadB}* site (60-bp) and flanking DNA (Red boxed *attC* site). Post PCR purification reaction, the product was cloned in one or other orientation in pBBR1MCS2. Each construct, *paadB_attC1* and *paadB_attC2* was used in subsequent natural transformation assays to detect IS1111-*attC* translocation.

4.3.2 Natural transformation of *Pst587*

Each synthetic trap was naturally transformed into *Pst587*. This strain is known to have a PCI, carrying one confirmed intact IS*Pst6* element and a truncated IS*Pst6* footprint. *Pst587* also contains at least 2, empty *attC* sites, BGC092 and BGC093. A single colony of *Pst587* was patched into a standard size of 1 cm diameter onto a LB plate and grown at 30°C for 24 hrs. *pattC* trap DNA (pPsattC or paadB_attC) was applied straight onto the colony as a solution of 1- μ g in 5 μ L of EB and allowed to absorb. The plates were incubated for 24 hrs (without inverting) at 30°C. The bacterial growth was harvested and resuspended in 300 μ L PBS and incubated on LB/Km50 plate (100 μ L/plate) (Figure 4.4). Plates were incubated at 30°C for 2 days, then examined for the appearance of transformants.

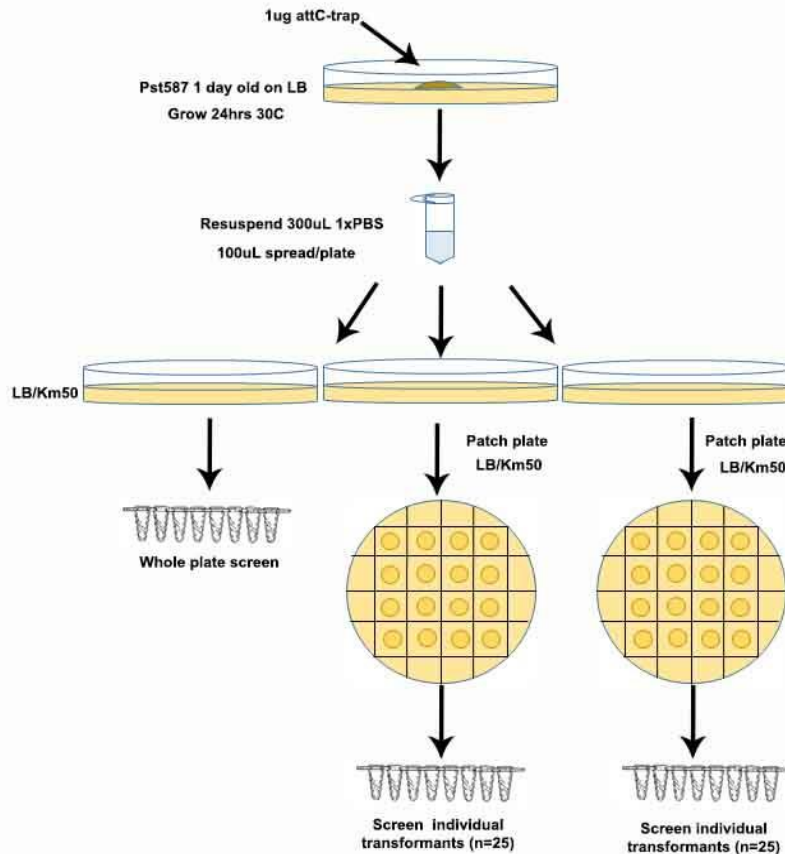


Figure 4.4: Natural transformation of *Pst587* with *pattC* traps.

pattC traps are naturally transformed into *Pst587* overnight at 30°C. The cell pellet is resuspended in PBS and equal volumes of resuspended cell mass are spread plated onto LB/Km⁵⁰. *pattC*-trap encode Km^R. After 2 days at 30°C, transformants are visible and are screened for a transposition event (see section 4.2.3). One plate was used for whole plate screens where the total cell mass was removed and used in PCR with plasmid flanking primers. Individual transformants were screened from the other two plates. In parallel, *Pst587* cells are transformed with 1xTE as a negative control and/or pBBR1MCS2 vector only to test the transformation efficiency of *Pst587*.

4.3.3 Screening natural transformants for IS1111-attC translocation

Crude DNA extractions were prepared for 50 colonies of transformant per experiment as outlined in section 2.3.2. 1 μ L of crude DNA preparation was used as PCR template with plasmid flanking primers (NVC15 and NVC16) to screen for successful translocation of ISPst6 into pattC- traps (Figure 4.5). Positive PCR product sizes were expected to be 1.8-kb and 1.7-kb for pPsattC-IS and paadB_attC-IS respectively. Where a PCR positive transformant was observed, its plasmid was recovered as per section 2.8. To generate a pure template for sequencing, purified plasmid DNA was used firstly with plasmid flanking primers NVC15 and NVC16. A second PCR was done by a nested PCR approach using primers ST96 and ST139 and the final PCR product was ligated into pGEM-T Easy vector as per manufacturer's instructions and cloned into *E.coli* JM109 as per section 2.6.1. Plasmid from a clone positive for the PCR product was purified and sequenced using M13F and M13R primers at AGRF Sydney Node, Australia.

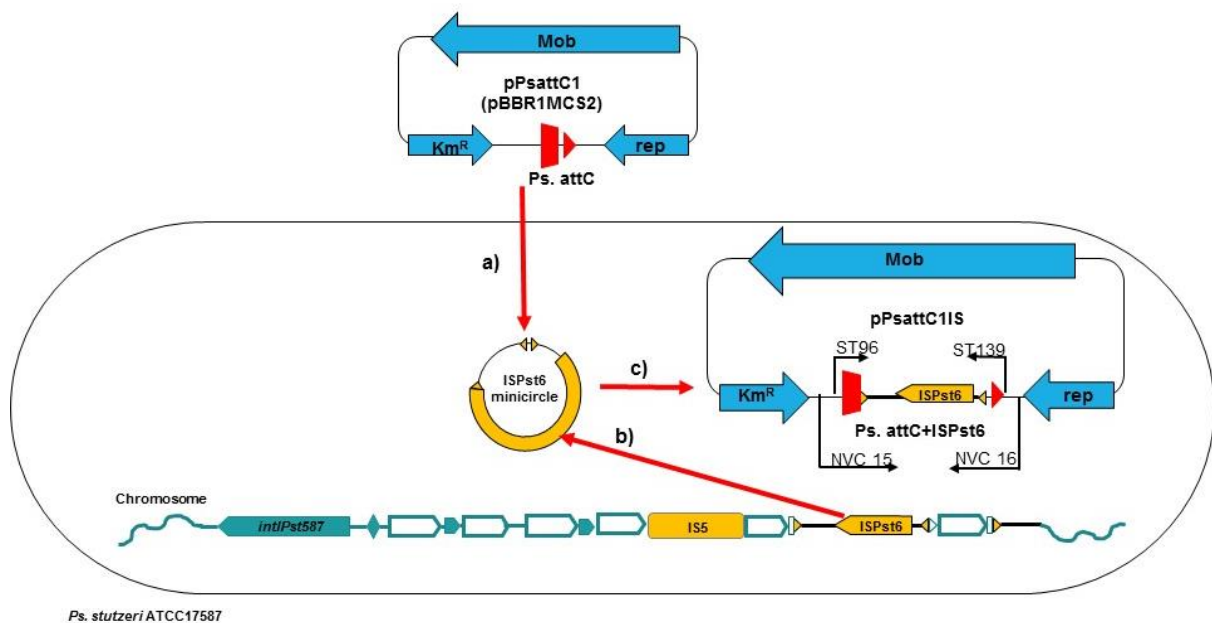


Figure 4.5: Theoretical experimental design and recovery of IS1111-attC element translocation using a synthetic trap.

a) Natural transformation of pattC traps into Pst587 cell b) ISPst6 is predicted to exist as minicircles within the cell c) ISPst6 site-specifically inserts into the pattC trap. Plasmid backbone primers, NVC15 and NVC16 are used to screen transformants for a translocation event into the pattC traps. Primers ST96 and ST139 are used in a nested PCR approach to give product of 1.5-kb corresponding to the attC and embedded ISPst6 element. The final recovered construct is termed pPsattC1S (or p_{pattCaadB} IS).

4.3.4 PCI invasion by plasmid-borne ISPst6 in *PstQ* and *Pst641* via electroporation.

To test ISPst6 translocation potential and ability to invade a *Pseudomonas* species (Figure 4.6), pPsattC1-IS was electroporated into *PstQ* and *Pst641* as per section 2.7.2. Previous profiling of *PstQ* and *Pst641* chromosomal integron arrays showed complete absence of IS1111-attC elements (Holmes *et al.*, 2003; Tetu & Holmes, 2008). Electroporation was used as the transformation mode as *PstQ* and *Pst641* do not show natural competency under the experimental conditions applied here. Electrocompetent cells of *PstQ* and *Pst641* (50 μ L), prepared as per section 2.6.1, were electroporated with 1 μ g of vector pPsattC1-IS as per section 2.6.2. As done previously for *Pst587* natural transformants, individual *PstQ* and *Pst641* transformants were patch-plated and boil-lysed to extract crude DNA. 1 μ L of crude DNA preparation was used as PCR template for the following PCR screens to test i) for ISPst6 disruption of the wild type *PstQ/Pst641* array using AJH17/AJH27 primer pair, ii) if ISPst6 translocation from pPsattC1-IS trap has restored the original synthetic attC site using NVC15/NVC16 primer pair and iii) if plasmid ISPst6 had generated minicircles within *PstQ* and *Pst641* using ST89/ST123 primer pair.

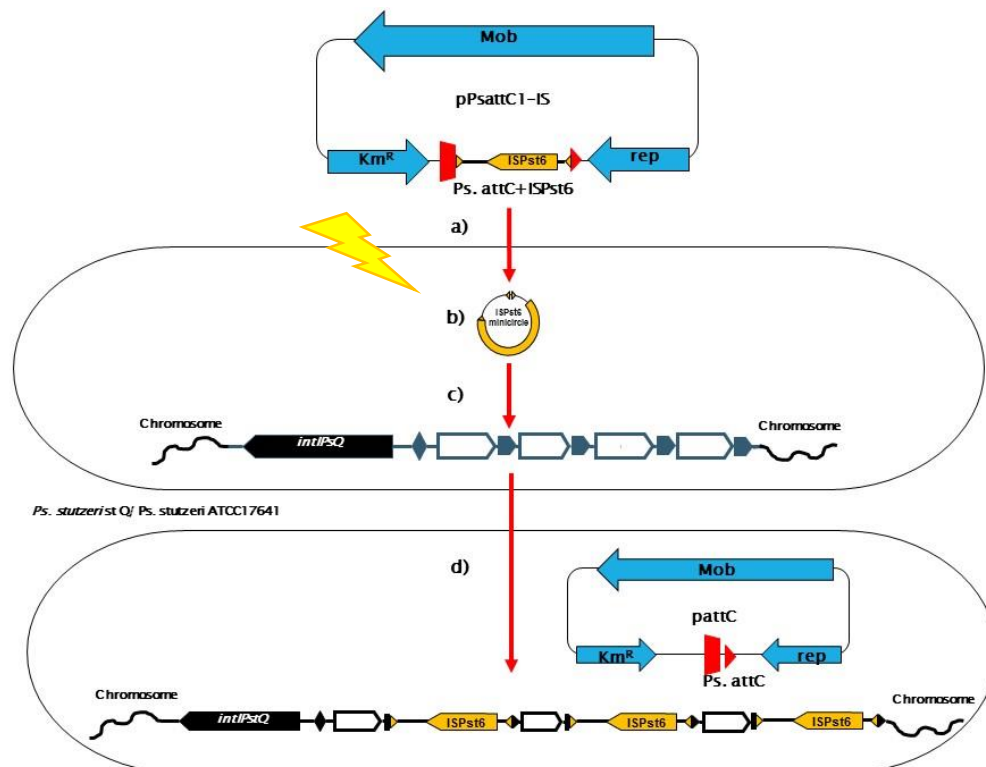


Figure 4.6: Introduction of pPsattC1_IS into *Pseudomonads* via electroporation.

a) Transformation of pPsattC1IS into *PstQ* and *Pst641* cells b) ISPst6 is hypothesised to translocate out of the plasmid c) ISPst6 is predicted to target the chromosomally located attC site. D) One or more attC sites of the array will be invaded by ISPst6. PCRs target IS minicircle, PCI array and pPsattC-IS.

4.4 Results

4.4.1 *Ps. stutzeri* ATCC 17587 is naturally transformable with double stranded plasmid containing an attC site

To test the translocation potential and target specificity of IS*Pst6* within its natural host *Pseudomonas*, purified plasmids, termed pattC traps, were added directly to naturally competent *Pst587* cells. On average, 106 and 70 *Pst587* transformant colonies were seen when transformed with 1 µg of pPsattC1 and pPsattC2 DNA respectively (Table 4.2). In contrast when *Pst587* cells were transformed with paadB_attC traps, the colony counts were 93 and 90 for paadB_attC1 and paadB_attC2 respectively. The difference in colony counts with the pPsattC traps may be significant but is more likely due to a slight variation in the trap concentration added to the cells or in the harvesting method. Transformation of *Pst587* with synthetic attC-traps was comparable to the 101-106 cfu/µg DNA transformation efficiency observed for *Pst587* with foreign gene cassettes (Gestal *et al.*, 2011). Where *Pst587* cells were transformed with TE buffer only, no transformants were obtained on the LB/Km⁵⁰ plates as wild-type *Pst587* is not intrinsically resistant to kanamycin.

Table 4.2: Transformation efficiency of *Pst587* cells with synthetic attC-trap

attC-type	Synthetic trap ^{a)}	Experiment 1 colony count ^{b)}	Experiment 2 Colony count ^{b)}	Transformation yield/µg DNA (Average colony count)
<i>Pseudomonas</i> -type	pPsattC1 (→)	97	115	106
	pPsattC2 (←)	69	71	70
attC _{aadB}	paadB_attC1 (→)	92	94	93
	paadB_attC2 (←)	81	99	90
Negative control	1xTE buffer only	0	0	0

a) (←/→) orientation of attC site in relation to rep site on pBBR1MCS2.

b) In each experiment, the colony count is an average of two plate counts of *Pst587* transformants.

A preliminary screen of the whole transformation plate did not reveal translocation of IS into attC-traps. That is, using plasmid flanking primers NVC15/NVC16, only PCR products corresponding to “empty” *Pseudomonas*-type (~550-bp) and attC_{aadB} (~420-bp) traps were observed (Figure A19). In order to increase the sensitivity of detection for an IS*Pst6* translocation event, individual transformants (50 per experiment) were screened instead of the total plate mass (see 4.3.2).

Translocation of IS*Pst6* was detected, at a rate of 4% (4/100) for *Pst587* cells transformed with pPsattC1. In comparison, detection of translocation into all other attC-traps was <1%.

ISPst6 translocation into pPsattC1 produced PCR products of both 550-bp and ~1.8-kb (Figures A20 & A21). This 1.3-kb increase corresponds to the size of an IS1111-attC element. Screening individual transformants did not overcome the PCR bias towards the smaller PCR product.

Primer combinations involving one flanking primer (NVC15/NVC16) and one primer targeting the IS transposase gene (ST123), was used to re-screen a subset of 50 *Pst587* transformants for IS translocation. This approach increased detection sensitivity so that 4% (2/50) of *Pst587* cells transformed with paadB_attC1 were positive for translocation (Table 4.3, Figure 4.7), where previously the detection rate was below 1% using flanking primers only. Similarly, sensitivity was improved for *Pst587* transformed with pPsattC1. Detection levels for pPsattC2 and paadB_attC2 remained below 1% (Figure A22).

Table 4.3: Recovery efficiency of IS1111-attC translocation into pattC traps in *Pst587*

attC type	Trap	Experiment 1^{a)} Flanking	Experiment 2^{a)} Flanking	Total^{b)}	Experiment 2^{c)} Flanking/IS
<i>Ps.-</i> <i>type</i> (73bp)	pPsattC1 (→)	6% (3/50)*	2% (1/50)	4% (4/100)	4% (2/50)
	pPsattC2 (←)	<1% (0/50)	<1% (0/50)	<1% (0/100)	<1% (0/50)
<i>attC_{aadB}</i> (60bp)	paadB_attC1 (→)	<1% (0/50)	<1% (0/50)	<1% (0/100)	4% (2/50)
	paadB_attC2 (←)	<1% (0/50)	<1% (0/50)	<1% (0/100)	<1% (0/50)

a) All positive transformants identified were from one plate in each experiment.

b) Based on findings in experiment 1 and 2.

c) Transformants from experiment 2 were rescreened using one primer in the plasmid backbone (NVC15/NVC16) and one primer in the ISPst6 (ST123).

* See appendix A2.1.

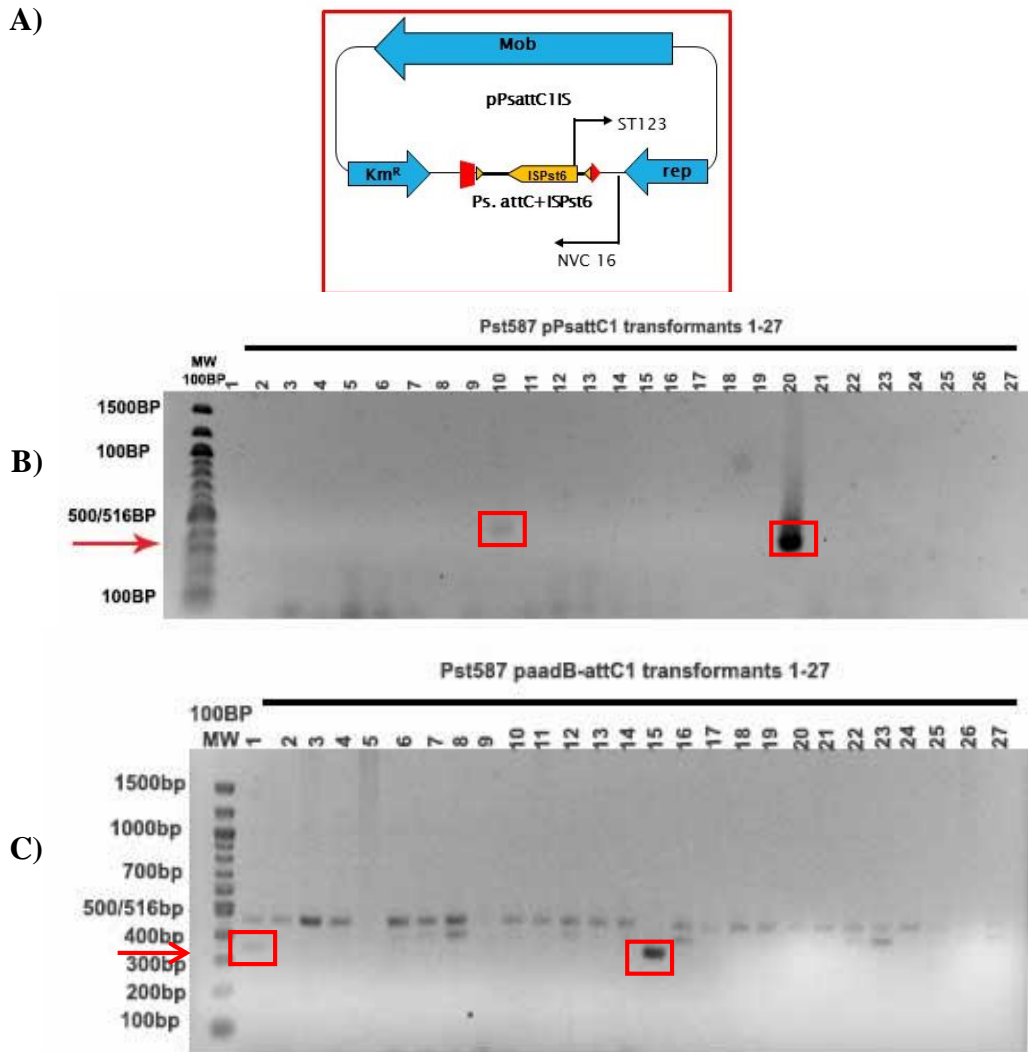


Figure 4.7: Definitive assay screen for IS1111-attC translocation into attC-traps.

A) Schematic representation of primer combination used to detect translocation. Primers NVC16 and ST123 were used on boil-lysed *Pst587* cells. Red arrow in parts B and C corresponds to the correct product size.

B) *Pst587* transformant 10 (NT10) and transformant 20 (NT20) are positive for pPsattC1-IS translocation. *Pst587* transformants 1-27 are shown here. C) *Pst587* transformant 1 (NT1) and transformant 15 (NT15) are positive for paadB_attC1-IS translocation. *Pst587* transformants 1-27 are shown here. Note that there was no available positive control for this PCR, however *in silico* analysis predicted products of 353-bp (for pPsattC1) and 298-bp (for paadB_attC1). Additional bands (400-500-bp) are non-specific products, and were observed in the no template control DNA also.

Recovered pPsattC1-IS plasmid traps from transformants NT10 and NT20 were further characterized by restriction digestion mapping (Figure 4.8). The *Pst587* NT1-recovered plasmid preparation paadB_attC1-IS was further analyzed by PCR only (Figure A23) and was found to contain a mixed plasmid population consisting of IS invaded and non-invaded attC_{aadB} sites. The *Pst587* NT15 plasmid population could not be recovered as the colony was not viable due to prolonged storage on agar plates.

4.4.2 Restriction mapping detects mixed plasmid populations

Figure 4.8 shows that *Pst*587 transformants NT10 and NT20 both had a mixed plasmid population composed of pPsattC1 and the IS invaded, pPsattC1-IS. The banding patterns of *Pvu*II digests of NT10 and NT20 plasmids (Figure 4.8D) matched *in silico* predictions for those of pPsattC1-IS (Figure 4.8A) and pPsattC (Figure 4.8B). In comparison, pPsattC1 only contained the predicted five bands ranging from 360-bp to 2071-bp.

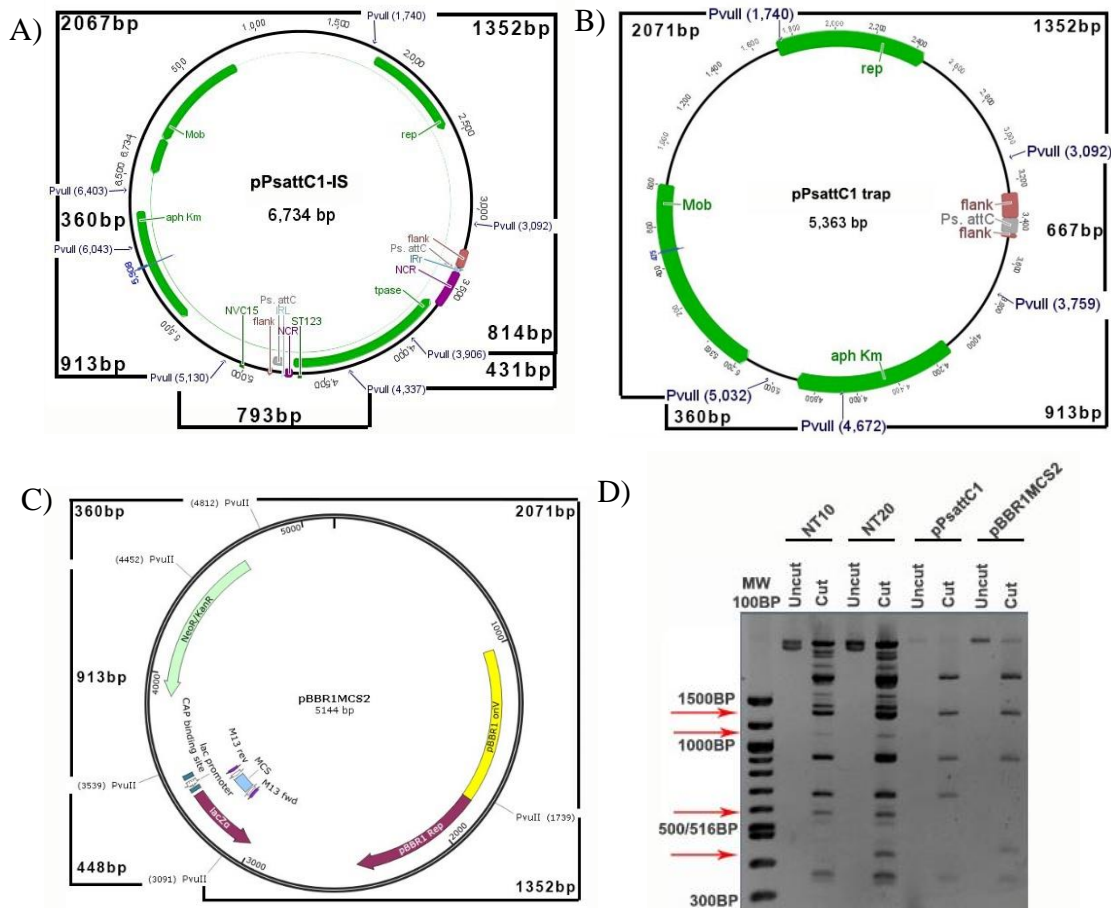


Figure 4.8: *Pvu*II predicted digests for plasmids a) pPsattC1-IS b) pPsattC c) pBBR1MCS2. D) *Pvu*II-digested plasmid from NT10 and NT20.

The profile confirms presence of a mixed plasmid population, consisting of pPsattC and pPsattC-IS. Arrows indicate additional bands in NT10 and NT20 plasmid preparation that correspond to presence of pPsattC1-IS.

The mixed plasmid populations from NT10 and NT20 were separately transformed into *E. coli* JM109 to recover the pPsattC1-IS component. Characterization of only NT20 pPsattC1-IS will be discussed from this point onwards. The plasmid topology of pPsattC1-IS was analysed by gel electrophoresis (Figure 4.9).

To demonstrate that an intact IS had invaded the pPsattC-1 trap, and that a 1.3-kb increase in plasmid size was a result of trap invasion, purified plasmid preparations were run out on low percentage agarose gels (Figure 4.9). During agarose gel electrophoresis, purified pPsattC1-IS ran at a molecular weight of greater than 10-kb despite its predicted size being 6-7-kb, while the pPsattC1 and pBBR1MCS2 controls produced faint bands of the correct sizes (5.4-kb and 5.1-kb respectively) (Figure 4.9A). pPsattC1-IS also ran at a far greater molecular weight than predicted under lower percentage agarose, while the pPsattC1 and pBBR1MCS2 controls produced bands of the expected size (Figure 4.9B). In both gels, the purified, uncut pPsattC1-IS showed multiple electrophoretic forms than ran significantly slower than linear 10 kb DNA markers. This is consistent with the insertion of the IS into the trap producing a double-stranded molecule with multiple topological forms. To confirm that a single IS had invaded the trap, restriction digestion using *Bam*HI was performed on the pPsattC1-IS vector. Cut pPsattC1-IS vector produced a band of the expected size (6.7-kb). However, a higher molecular weight band (>10-kb) corresponding to uncut pPsattC1-IS was also seen (Figure 4.9C). This incomplete digestion could be due to the newly inserted IS interfering with digestion, as the *attC* is situated 9-bp downstream of the *Bam*HI site, or excess pPsattC1-IS in the digest. *Bam*HI-cut pBBR1MCS2 also produced a band of the expected size (5.1-kb).

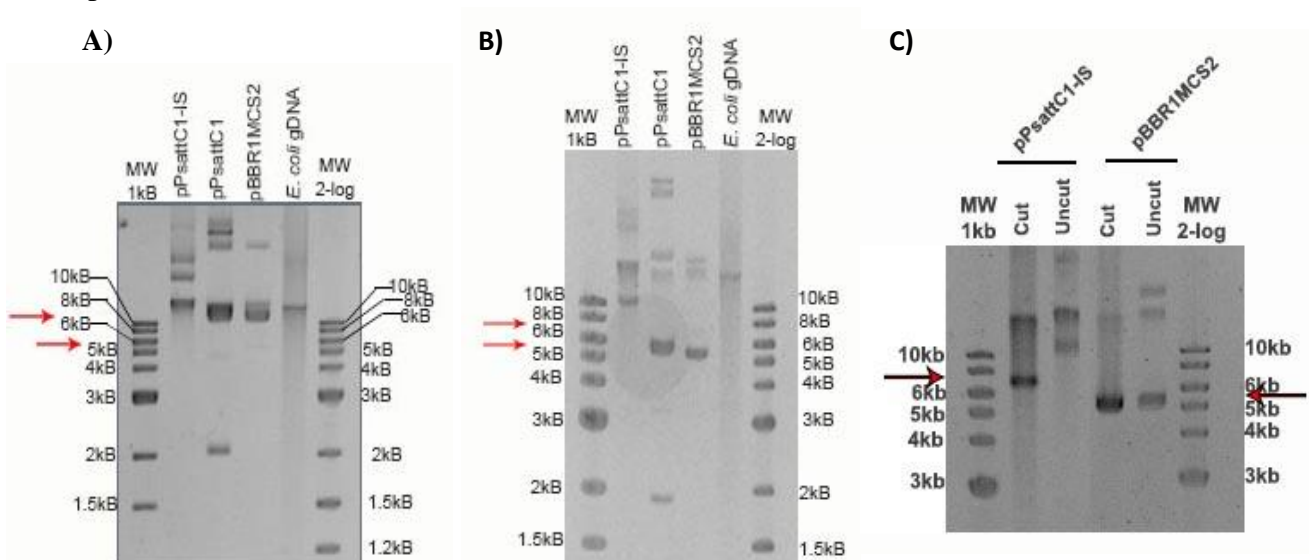


Figure 4.9: a) 1% agarose b) 0.75% agarose topological analysis of pPsattC1-IS, pPsattC1, pBBR1MCS2.

pPsattC1-IS was predicted to travel at 6.7-kb, pPsattC1 5.3-kb and pBBR1MCS2 5.1-kb. c) 1% agarose gel of *Bam*HI digest of pPsattC1-IS and pBBR1MCS2 plasmids.

*Bam*HI cut pPsattC1-IS (predicted size 6734-bp) *Bam*HI cut pBBR1MCS2 vector (size 5144-bp). 1.5-kb increase in pPsattC1-IS corresponds to the cloned in 219-bp *attC* with flanking region and inserted 1.3-kb IS*Pst6* element. *E. coli* gDNA was loaded as control.

The increase in relative size of the pPsattC1-IS plasmid in comparison to the pPsattC1 plasmid, as seen by *Bam*HI profiling, confirmed that an IS1111-attC element had indeed invaded the synthetic *Pseudomonas*-type attC trap.

4.4.3 Sequencing results confirm that ISPst6 inserted into Pseudomonas-type attC sites

To determine the precise nature of insertion of ISPst6 into the synthetic pPsattC1 trap, plasmid flanking primers were used to amplify the 1.8-kb region containing the “trapped” ISPst6 element. However, these primers are prone to mispriming, hence a nested PCR approach was also used to generate a PCR product of 1.5-kb. This sequenced product corresponded to a complete ISPst6 insertion into the *Pseudomonas*-type attC site. All relevant features such as the flanking region, *Pseudomonas*-attC site, IS non-coding regions, inverted repeats and transposase gene (*tpase*) were identified.

The recovered sequence from *Pst587* transformant NT20 was analysed by BLASTn was 99% (1367/1371 bases) identical to a chromosomally embedded ISPst6 (BGC094) from *Ps. stutzeri* ATCC 17587 (EF648209.2) (Figure 4.10). Five ISs were previously known from *Pst587* (Tetu & Holmes, 2008). This recovered sequence was not identical to the published IS from *Pst587*, thus suggesting the capture of a previously unsequenced IS1111-attC copy. One base difference was observed in the *tpase* ORF, where a thymine was replaced by an adenine at position 446. Six frame translation of the recovered *tpase* with the adenine substitution does not result in frame shift or stop codon formation, however, it does alter the amino acid from a negatively charged glutamic acid to a hydrophobic valine (Figure 4.11).

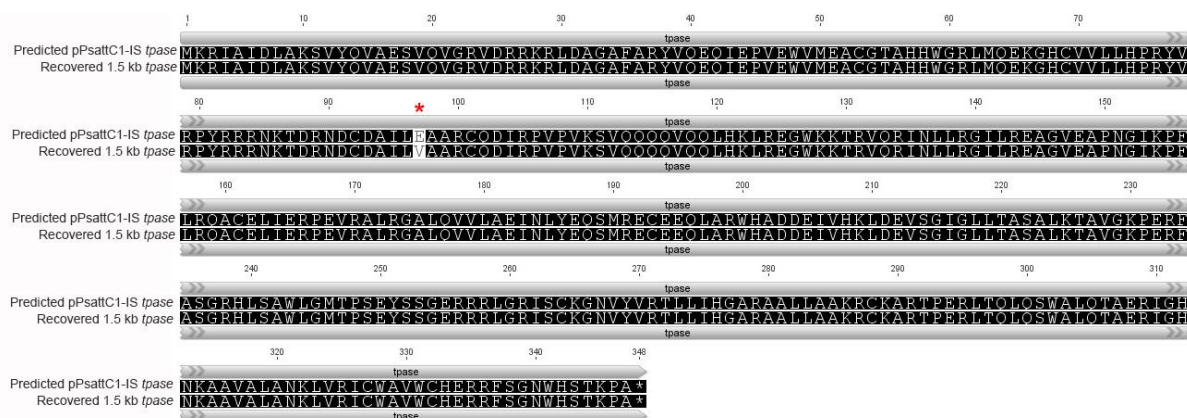


Figure 4.11: Translated amino acid sequence of the predicted and recovered *tpase* gene.

Note the single amino acid variation at position 98, where a valine was recovered instead of the predicted glutamic acid.

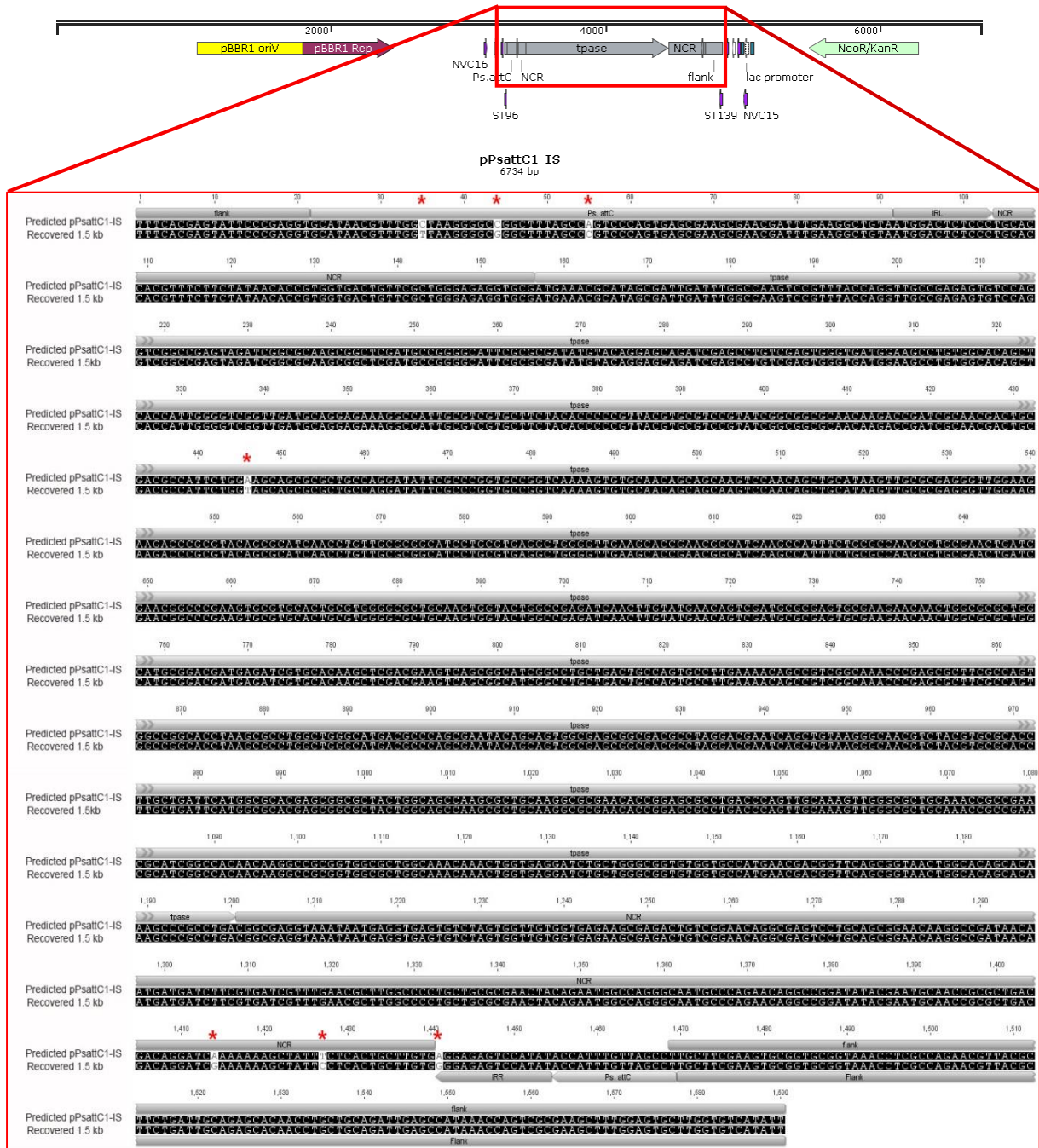


Figure 4.10: Sequencing pPsattC1-IS confirmed the presence of ISPst6 within the *Pseudomonas*-type attC site.

A linear plasmid map of the predicted pPsattC1-IS is shown. The boxed region corresponds to the 1.5-kb sequenced region. Sequence alignment of the recovered 1.5-kb product is shown relative to the predicted pPsattC1-IS sequence. Nucleotide differences between the two are annotated with an asterisk. The alignment was generated using Geneious.

Two base differences are also observed in the non-coding region adjacent to IR_L where guanine and cytosine replace an adenine at position 1414 and thymine at position 1427 respectively. Also in IR_L, an adenine is replaced by a guanine in the recovered sequence at position 1441. Interestingly, this identical three base difference is observed in IS*Pst6* elements in *Ps. stutzeri* ATCC 14405. The extent of variations observed here is comparable to that seen within other PCI assays. Finally, three base differences were noted in the recovered *Pseudomonas*-type attC site at positions 35, 44 and 55. Differences in the *Pseudomonas*-attC site may have occurred as a result of a replication error during translocation. Sequencing or PCR errors are also not uncommon when dealing with MGE. These results demonstrate that IS*Pst6* is able to insert into both *Pseudomonas*-type attC and classical attC_{aadB} sites, however, there appears to be a high efficiency for the *Pseudomonas*-type attC. It is noteworthy that insertion appears to be dependent on the orientation of the target site in the plasmid.

4.4.4 IS1111-attC invasion of chromosomal attC array

To test the translocation potential of IS1111-attC elements and their ability to invade PCI arrays, the IS1111-attC element was delivered into three different *Ps. stutzeri* strains by electroporation of the plasmid pPsattC1-IS (Table 4.4).

Table 4.4: PCI characteristics of *Pst405*, *Pst641* and *PstQ*

Strain ^{a)}	Gv.	IntI	IS1111-attC Presence ^{b)}	IS1111-attC minicircles	attC sites available ^{c)}	Competency ^{d)}
<i>Pst405</i>	2	Not active	4 intact copies	Present	>2	Electrocompetent
<i>Pst641</i>	8	Active	None	Absent	>2	Natural* Electrocompetent
<i>PstQ</i>	8	Active	None	Absent	>2	Electrocompetent

a) *Pseudomonas stutzeri* ATCC14405, *Pseudomonas stutzeri* ATTC17641 (Tetu & Holmes, 2008), *Pseudomonas stutzeri* st. *Q* (Holmes *et al.*, 2003).

b) Represents the presence (and copy number) or absence of IS1111-attC elements within PCIs in these wild-type strains.

c) Availability is simply the number of empty (thus available) PCI attC sites into which the introduced IS1111-attC can insert.

d) *Pst641* is naturally competent for small linear DNA, but not circular DNA under the conditions used in this experiment

4.4.4.1 IS1111-attC is stably maintained in the plasmid

An initial experiment was conducted in which pPsattC1-IS was electroporated into *Pst405*. This is a strain in which the IS naturally exists and is presumed to be fully active with respect to IS mobility. Fifty *Pst405* transformants were screened for a loss of IS from pPsattC1-IS.

All fifty transformants maintained intact pPsattC1-IS constructs (Figure A24), and no transformant containing an empty *pattC* trap was detected. Given that wild-type *Pst405* contains IS minicircles as well as multiple IS*Pst6* elements within its PCI array, it was not possible to test for other signs of activity such as the formation of minicircles by the plasmid-introduced IS.

4.4.5 Introduced IS elements are able to undergo initial stages of mobilization to form minicircles.

To observe IS*Pst6* proliferation across PCIs, pPsattC1-IS was electroporated into *PstQ* and *Pst641*. Both of these strains have at least 2 available *attC* sites for insertion, and lack IS1111-*attC* elements. As with *Pst405*, no loss of IS*Pst6* from pPsattC1-IS was observed after screening fifty *PstQ* transformants with plasmid-flanking primers (Figure A24). *PstQ* and *Pst641* transformants were also screened for the presence of IS*Pst6* minicircles and variations within their PCI arrays. 8% (4/50) of *PstQ* and 25% (3/12) *Pst641* transformants gave the predicted PCR product of ~258-bp when IS*Pst6* minicircle primers were used (ST89 and ST123) (Figure A25). The *PstQ* positive transformants were annotated as *PstQET2*, *PstQET8*, *PstQET49* and *PstQET50*. *Pst641* positive transformants were annotated as *Pst641ET5*, *Pst641ET9* and *Pst641ET11*. High molecular weight DNA was extracted from these six of the seven positive transformants and the minicircle screen was repeated (Figure 4.12). The *Pst641ET5* plasmid population could not be recovered as the colony was not viable due to prolonged storage on agar.

A PCR product corresponding to an IS*Pst6* minicircle junction region was present in the positive control strain *Pst587* and absent in the wild type *PstQ* strain. pPsattC1-IS purified plasmid was added as control and gave two false positive products (~1500-bp and 900-bp). *In silico* analysis indicated that these products are the result of non-specific binding of the primers to the *mob* and *rep* genes on pBBR1MCS2. Transformants *PstQET2*, *PstQET8*, *PstQET49*, *PstQET50* have composite banding patterns. That is, all four transformants display the 2 plasmid bands (1500-bp & 900-bp), as well as the 350⁺-bp and ~500⁺-bp band present in *PstQ* WT and finally the ~258-bp minicircle band. *Pst641ET9* and *Pst641ET11* all produce a band of 258-bp, corresponding to the minicircle. The minicircle PCR bands from *PstQET2* and *Pst641ET9* were cloned and sequenced for identification

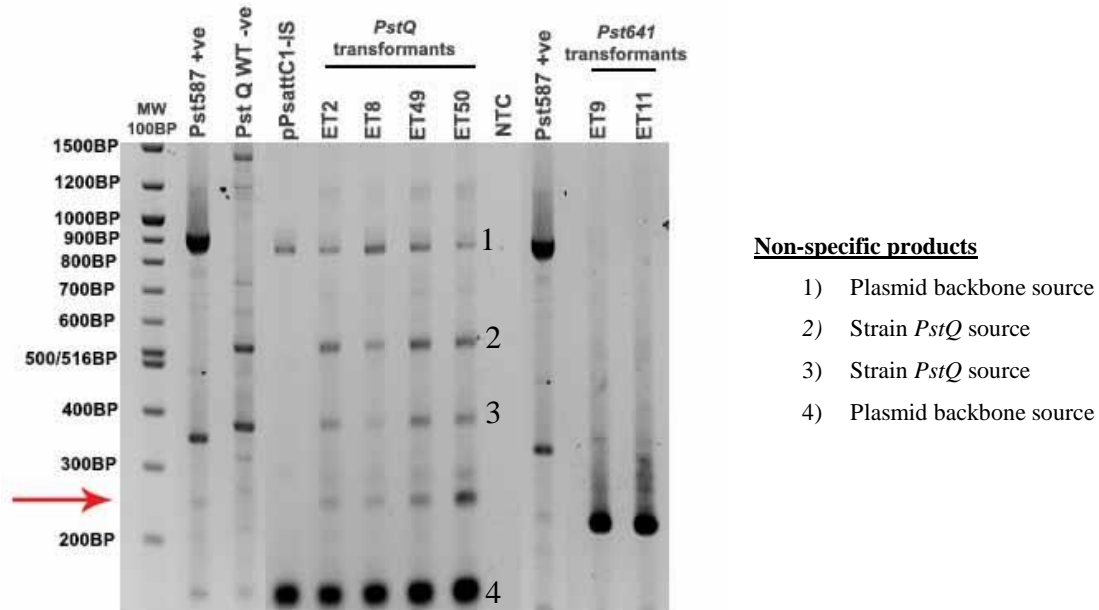


Figure 4.12: *PstQ* and *Pst641* transformants contain ISPst6 minicircles. Expect product size is 258-bp (red arrow). *Pst587* positive controls, *PstQ* wild-type was used as a negative control as well as pPsattC1-IS purified plasmid construct. NTC: no template control, red arrow points to cloned and sequenced bands.

4.4.5.1 Sequencing confirms the presence of ISPst6 minicircles in transformants *PstQ*ET2 and *Pst641*ET9.

The 258-bp PCR product from *PstQ*ET2 and *Pst641*ET9 (Figure 4.12) was cloned into pGEM-T Easy and sequenced. In both cases the product contained all the expected features of an ISPst6 minicircle - these are IRs, non-coding regions (NCRs) and *tpase* gene (Figure A26). Splitting the 258-bp sequence at the 10 base pair junction resulted in two fragments annotated as *PstQ*ET2 1-137-bp and *PstQ*ET2 138-258-bp, and *Pst641*ET9 1-137-bp and *Pst641*ET9 138-258. The respective sequences completely aligned to the intact ISPst6 located on pPsattC1-IS (Figure 4.13). I conclude that the 258-bp minicircle present in the *PstQ* and *Pst641* transformants originated from activity of the IS element introduced to the cells on the plasmid.

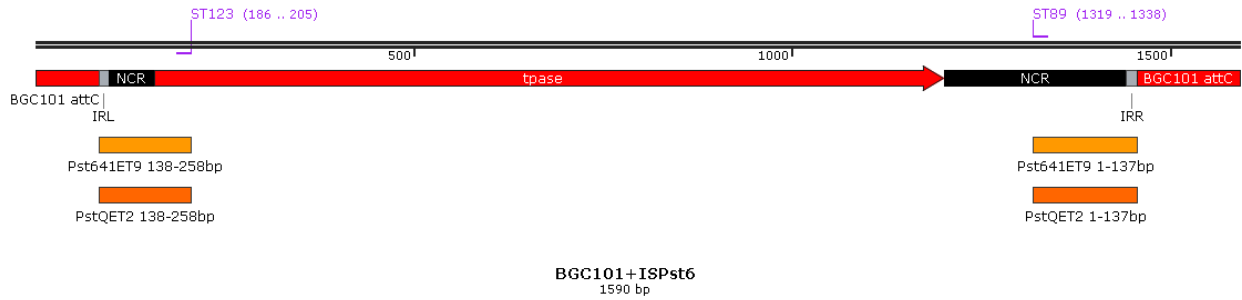


Figure 4.13: Minicircles from *PstQ* transformant 2 and *Pst641* transformant 9 recovered minicircles align with attC embedded ISPst6 from pPsattC1-IS.

4.4.6 Insertion into PCI attC sites was not detectable.

A “global” assay for IS invasion was performed in which the PCIs were screened using previously published primers, AJH17 and AJH27, to target *PstQ* integron arrays. Four transformants, *PstQET2*, *PstQET8*, *PstQET49* and *PstQET50*, and all twelve *Pst641ET* were examined for differences in their integron array by comparing their PCR banding pattern to that of the wild-type *PstQ* and *Pst641*. No differences in the PCR banding patterns were observed (Figure 4.14). Thus IS1111-attC proliferation across the *PstQ* and *Pst641* PCIs was not detected.

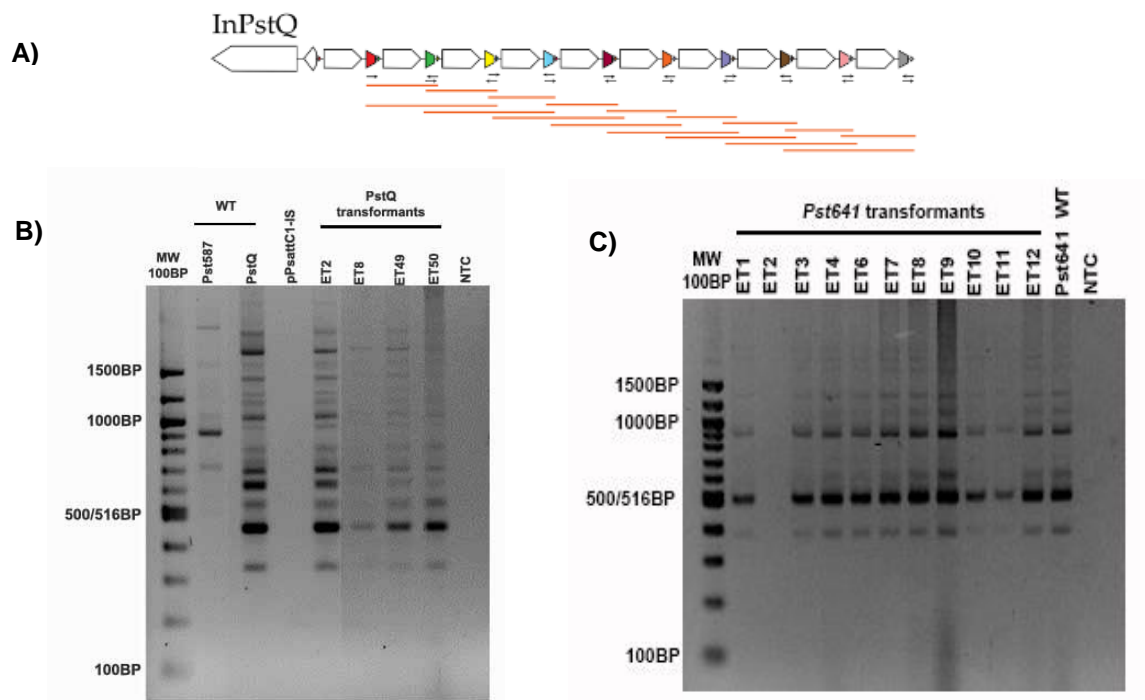


Figure 4.14: ISPst6 does not appear to have invaded the PstQ chromosomal arrays.

a) Schematic representation of InPstQ (Holmes et al., 2003). This array contains 10 gene cassettes. Primers AJH17 and AJH27 were used to amplify the array. These primers bind Ps.-type attC-sites. Resulting PCR product is a PCR array of varying sizes.

b) InPstQ array PCR results show no variation between wild type PstQ strain and transformants PstQET2, PstQET8, PstQET49, PstQET50 arrays. Pst587 and PstQ positive controls for PCI arrays. pPsattC1-IS plasmid DNA negative control. NTC is the no template control.

c) InPst641 array PCR results show no variation between wild-type Pst641 strain and transformants. Pst641 was the positive control for PCI arrays. Note Pst641ET5 is not included as the culture was not viable. NTC is the no template control.

4.5 Discussion

4.5.1 IS1111-attC elements translocate with regional-specificity into attC sites

Translocation of insertion sequences plays a major role in the horizontal spread of antibiotic resistance genes amongst prokaryotes (Soki, 2013). In chapter 3, I established a correlation between PCI-IS1111-attC in environmental and MRI-IS1111-attC co-occurrence in clinical *Pseudomonads*. Previously, Tetu & Holmes (2008) demonstrated partial translocation of IS1111-attC when using an *E. coli* host and a single copy fosmid containing a PCI array. Their results implied that all conditions for complete translocation were not met.

In my study, intact translocation of ISPst6 was detected in the *Pseudomonas*-type attC trap when a low copy plasmid pBBR1 (~20-30 copies per cell) (Kovach *et al.*, 1995; Lefebvre & Valvano, 2002) was used as the trap backbone and introduced into a *Pseudomonas* host. More importantly, in my study, IS1111-attC translocation was confirmed as site-specific in the one instance studied, emphasising the importance of site-specific recombination in the overall dissemination of these IS elements. Furthermore, no target site-duplications of the *Pseudomonas*-type attC site were identified upon ISPst6 insertion. This is in line with the general characteristic of previously studied IS110/IS1111 insertion sequences (Bruton & Chater, 1987; Henderson *et al.*, 1989; Mahillon & Chandler, 1998; Partridge & Hall, 2003). IS1111-attC was translocated to two distinct structural types of attC sites: *Pseudomonas*-type and classical attC_{aadB}. Translocation into the *Pseudomonas*-type was confirmed by sequencing, and by PCR mapping for the classical site. IS1111-attC translocation activity into and out of the attC-trap was observed across four different *Ps. stutzeri* strains containing different PCI arrays. If IS1111-attC translocation is “dependant” on *Pseudomonas* host factors, it may explain why IS1111-attC translocation into non-*Pseudomonads* (enteric) is not as readily observed. Spread of IS1111-attC elements into non-*Pseudomonads* is thus believed to be the results of MRI driven movement, rather than IS1111-attC only.

4.5.2 Orientation-specific translocations is a feature of ISPst6 elements

Four synthetic *attC* traps were transformed into a naturally competent *Ps. stutzeri* ATCC 17587 containing PCI-embedded ISPst6 elements. These traps differed in their respective *attC* sites, *Pseudomonas*-type and classical *attC_{aadB}*, and in the respective orientations in which these sites were cloned into the pBB1MCS2 plasmid backbone (Figure 4.15).

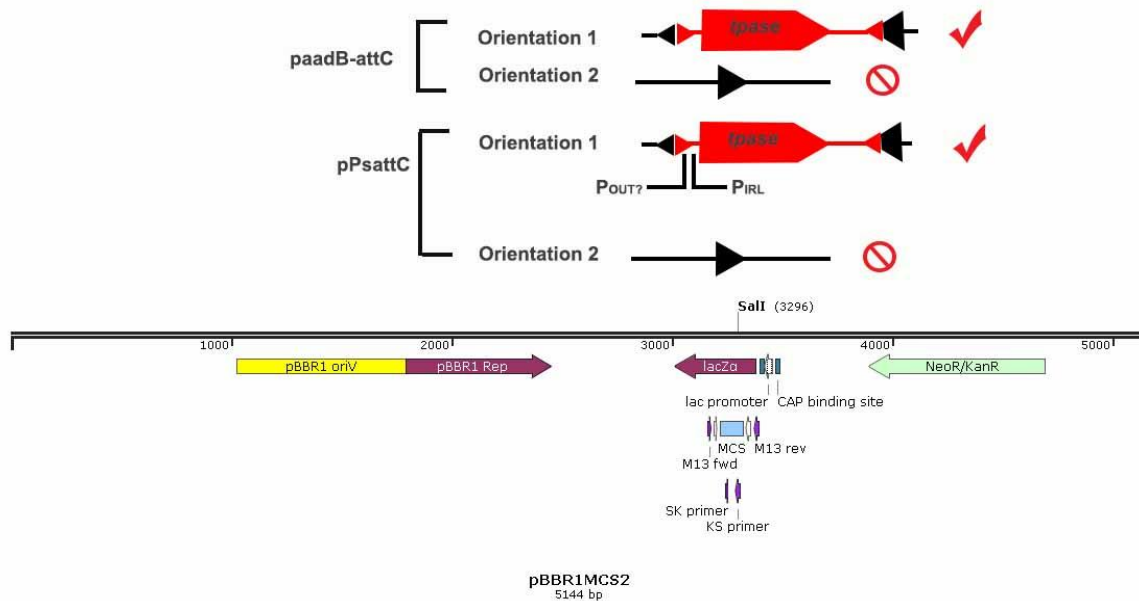


Figure 4.15: Linear map of pBBR1MCS2 and orientations of cloned recombination sites (*attC_{aadB}* and *Pseudomonas*-type *attC*).

IS1111-*attC* translocation was observed in *attC* sites cloned in orientation 1 but not in orientation 2.

ISPst6 translocation was observed in both types of *attC* sites when they were cloned in orientation 1, implying that translocation is not strictly dependent on the *attC* sequence (Figure 4.15). *attC* sequence variation may have an effect on the IS1111-*attC* T_{pase} ability in preferentially recognizing the *Pseudomonas*-type over *attC_{aadB}* sites. It was harder to detect IS translocation into *attC_{aadB}* by PCR assays using flanking primers, suggesting that there was a lower plasmid population with *attC_{aadB}*-IS translocation. Translocation was not observed when the sites were cloned in orientation 2, implying that some other molecular features were necessary for successful translocation. This orientation-specific translocation suggests that IS1111-*attC* elements are able to recognise single stranded *attC* targets. I postulate that the formation and recognition of a single stranded *attC* structure and its accessibility to the IS transposases drives IS translocation into orientation 1. Given what is known about IntI strand preference towards single stranded *attC* sites (Francia *et al.*, 1999), it may also be that IS1111-*attC* transposase preferentially binds single stranded *attCs*.

4.5.3 *To jump or not to jump? Potential routes for single stranded attC site generation*

In order to account for the observed orientation-specificity by ISPst6 towards orientation 1 attC sites, scenarios that may result in the production of single stranded attC sites, and thus lead to ISPst6 translocation, are discussed here:

Case 1: Natural transformation and ssDNA uptake promote ISPst6 transposition

Although, *Ps. stutzeri* is one of the better known models in population genetics (Lalucat *et al.*, 2006), the mechanisms of transformation have not been studied. Based on other model systems and the presence of genome homologs, predictions can be made as to the mechanisms involved in the active uptake of foreign DNA into the bacterial cell. Initially, double stranded DNA is bound to the cell surface, and is subsequently pulled through the secretin pore by retraction of a pilus, such as the type 4 pilus (T4P) (Chen & Dubnau, 2004). One strand of DNA is trafficked in a single stranded state into the cell, while the other strand is degraded. If natural transformation is the mechanism by which single stranded attC site is generated as the plasmid is taken up, then there is predicted to be a strand bias.

There is precedent for this in *Neisseria gonorrhoeae*, which has been shown to take up DNA through the recognition of a non-palindromic 10 or 12 nucleotide DNA uptake sequence (DUS10 or DUS12). More interestingly, the ssDUS12 Crick strand was observed to enhance transformation efficiency in comparison to the ssDUS12 Watson strand (Duffin & Seifert, 2012). Assuming that plasmid uptake in *Ps. stutzeri* follows an analogous mechanism, then one would predict single stranded attC-traps enter the *Pseudomonas* cytoplasm and yield a transient attC hairpin. This hairpin is predicted to be necessary for transposase recognition and site-specific recombination of IS1111-attC into the attC sites. However, DNA uptake does not explain why *Pst587* transformants with mixed plasmid populations, containing empty and invaded attC sites, were observed.

Case 2: Plasmid replication exposes the attC site for ISPst6 attack

To ensure that a fixed concentration of a low copy number plasmid is maintained in the bacterial population, it has to be replicated once per cycle and effectively partitioned to the daughter cells (Pinto *et al.*, 2012). For instance, ColE1 plasmid replication in *E. coli* is initiated by antisense RNAII and a repertoire of proteins that monitor plasmid number and ensure that the plasmid population is maintained (Del Solar *et al.*, 1998). This is achieved by RNAII binding to the target strand and acting as an attachment point for DNA polymerase I

(Cesareni *et al.*, 1991). However, antisense RNAI binding to RNAII, alters the structural conformation of RNAII, which can no longer act as a binding partner for DNA polymerase I (Cesareni *et al.*, 1991), thus resulting in inhibition of leading strand synthesis, and halting plasmid replication. This arrest results in the temporary formation of single stranded plasmid. However, this RNAI-RNAII complex formation is reversible (Xu *et al.*, 1993).

While the precise mechanism of pBBR1 plasmid replication is still unknown, there is some evidence to suggest that it does not occur via the rolling circle method (Antoine & Locht, 1992a; Antoine & Locht, 1992b; Szpirer *et al.*, 2001). I postulate that as the replication machinery attaches to pBBR1, the double stranded plasmid is disrupted to generate single strands. It may be maintained in a single stranded state due to the *attC* site acting as an antisense molecule. There is a precedent for replication leading to differential propensity for a single stranded *attC* to be exposed and able to form hair-pins (Bikard *et al.*, 2010). If pBBR1 replication results in single stranded *attC* formation and subsequent IS invasion, then this could explain why transformants containing mixed plasmid populations were recovered.

Partridge and Hall (2003) observed that the transposition efficiency of IS4321, an IS1111 subgroup member, increased when the 38-bp TIR target site was cloned into pACYC18:Tn21 in an orientation facing away from the *ori* site. They too postulated that the overall orientation of the single stranded site impacts on promoter placement and plasmid replication. Similarly, I predict that as pBBR1 is disrupted, single stranded *attC* sites form bulged hairpin structures, which may, depending on their strand location, pose problems by destabilizing replication machinery or interfering with the antisense RNA (Gerhart *et al.*, 1994). Alternatively, if *attC* site orientation had no effect on plasmid replication, it may be that inverse translocation of IS*Pst6* had a negative effect on replication by acting as an inhibitor or replication terminator. An inhibitory effect on plasmid replication would result in a lower detection threshold of IS translocation due to a reduction in plasmid copy numbers, whereas termination of plasmid replication would result in eventual loss of the plasmid from the *Pseudomonas* population (del Solar *et al.*, 1998), thus no detection of IS1111-*attC* translocation.

Case 3: Formation of single stranded attC during transcription by RNA polymerase

A transcription terminator is a section of nucleic acid sequence that initiates the release of newly synthesised mRNA from the transcription complex, freeing RNA polymerase and related transcriptional machinery (Henkin, 1996). Rho-dependent and Rho-independent

transcriptional terminators are widely distributed sequences across prokaryotic genomes (Henkin, 1996). This discussion will not go into the mechanistic differences between Rho-dependent and Rho-independent transcription terminators, but will merely highlight that in Rho-independent termination, there is a precedent for the formation of stem-loops on one strand during transcription (Henkin, 1996; Von Hippel, 1998).

I propose that the secondary structure formed by single stranded *attC* sites has the capacity to resemble transcriptional terminators, thereby promoting dissociation of RNA polymerase through allosteric effects and thus affecting pBBR1 replication. This may explain why IS1111-*attC* elements were only observed in *attC* sites cloned in orientation 1. Hairpins of *attC* sites cloned in orientation 2 were perhaps more likely to result in replication fork stalling, and RNA polymerase destabilization, thereby reducing the overall stability of the mRNA-DNA-RNA polymerase complex. Hence no IS1111-*attC* translocation was observed in orientation 2.

4.5.4 Does IS1111-attC undergo a non-replicative (cut-and-paste) or a replicative (copy-and-paste) translocation mechanism?

A postulated feature of IS1111 elements is the ability to move via circular intermediates, where the terminal inverted repeats of a single IS are brought together and are then separated by a characteristic number of base pairs of flanking sequence (Prosseda *et al.*, 2006; Tetu & Holmes, 2008). In this study, IS1111-*attC* minicircles were detected in *Ps. stutzeri* st. Q and *Ps. stutzeri* ATCC 17641 transformants after the introduction of the IS via a pPsattC1-IS trap. In contrast, IS minicircles were absent in the wild-type strains.

ISPst6 minicircles are predicted to be non-replicating, so integration of the IS into a chromosomal framework or plasmid is necessary for its replication and spread. However, in this study, integration of IS1111-*attC* into PCIs arrays was not observed. To overcome extinction within a cell/population, IS1111-*attC* integration into the PCI was predicted to be essential for maintenance and transposase expression (Thomas & Nielsen, 2005). What molecular-genetic factor controls IS1111-*attC* integration into a PCI? IS1111-*attC* may only invade single-stranded *Pseudomonas-attC* sites during genomic DNA replication, repair, or transcription. Realistically, there is only a small window of opportunity for successful invasion of IS1111-*attC* to occur!

SI hypothesised that if *ISPst6* translocated through a cut-&-paste mechanism, then the original *attC* trap would be restored. However, restoration was not observed and only the intact plasmid-embedded IS-*attC* was detected. The presence of IS minicircles within the transformants implied that *ISPst6* transposase was actively expressed in at least the early stages of movement. As the plasmid replicated, a newly synthesised IS copy would have been generated on both the leading and lagging strand. It is quite possible that one of these newly copied ISs was able to excise as a minicircle once the replication fork had passed. This would ensure that one IS copy was maintained within one strand, whilst allowing evacuation from the other strand. This echoes the cut-&-paste nature of transposon *Tn10*. *Tn10* translocates only after the replication fork passes, leaving a hemi-methylated copy behind in each sister chromosome (Bender & Kleckner, 1986). The co-occurrence of intact copies of plasmid embedded *ISPst6* and minicircles, may well be explained by a non-replicative mechanisms such as that of *Tn10*. However, at this stage, this hypothesis requires further validation.

4.5.5 Further work

I have detected *ISPst6* translocation by demonstrating the *in vivo* capture of the IS by a plasmid trap in its native *Pseudomonas* host. Furthermore, I have also detected the formation of *ISPst6* minicircles, in previously “un-infected” *Ps. stutzeri* strain *Q* and strain ATCC17641, following the introduction of a ptrap-IS plasmid. I have confirmed that the final outcome of *ISPst6* transposition is site-specific recombination into *attC*, and I predict that IS minicircle are an intermediate step in transposition, and that they can form in multiple cell types. However, the sequential steps leading up to IS transposition into *attC* sites are still unknown and a few key questions remain unanswered:

- 1) What precisely are the steps involved in *ISPst6* translocation?
- 2) Is IS1111-*attC* transposase target recognition purely sequence-specific or is it structural?
- 3) Is IS1111-*attC* transposase recognition of different *attC* sites as flexible as the Int11 target range?

Chapter 5 Characterization of recombination site-preferences of *Ps. stutzeri* IS1111-attC transposase

5.1 Introduction

Thus far I have demonstrated activity of IS1111-attC towards both *Pseudomonas*-type and attC_{aadB} sites in laboratory cultures. I have also shown that the frequency of IS1111-attC insertion is subject to attC orientation where translocation was observed in attC sites cloned in one orientation in the plasmid multiple cloning site (Figure 4.13), but not the other. DNA orientation (also referred to as polarity in this study) is biologically significant, thus the underlying causes were explored. This dependency on polarity could be a selective consequence of effects on surrounding genes/functions, such as IS insertion, potentially affecting plasmid replication (Collis *et al.*, 2001). Alternatively, it could be a mechanistic feature of IS1111-attC recombination mechanism; for example, only one strand may be involved in the reaction, and its presentation is affected by orientation specific processes.

There is precedence for a mechanistic origin of polarity with integron integrase mediated recombination between a double stranded (ds) attI and a single stranded (ss) attC (Cambray *et al.*, 2010; Johansson *et al.*, 2004). IntI-mediated recombination requires only three DNA strands, where one of these single strands contains a palindrome able to form an energetically stable duplex (hairpin), thereby producing a structure that is equivalent to a double stranded DNA molecule. The single-stranded DNA is the bottom strand of attC (Francia *et al.*, 1999) and its preferential targeting results in a biologically important outcome where the orientation of the incoming cassette ORF is preserved with respect to Pc (Partridge *et al.*, 2009).

Unlike IntI, which is a tyrosine recombinase, IS1111-attC encodes a transposase with a conserved D-E-D-D motif that is essential for activity of the Piv/MooV family of DNA recombinases (Buchner *et al.*, 2005; Tobiasson *et al.*, 2001). In this study, two mechanisms for the observed polarity were considered. Firstly, biased attC strand formation was considered where the hypothesis is that single-stranded attC sites are also an IS1111-attC target and the formation of this strand is subject to orientation effects. Based on the IntI model, I predict that the IS1111-attC transposase will only bind to one strand of the attC site as well as to a double stranded IS1111-attC minicircle (Figure 5.1). Secondly, specific recognition of and preference for some attC subclasses over others was hypothesised to play a role in the distribution patterns of IS1111-attC across PCI and MRIs. Work presented in chapter 4

revealed that IS1111-attC transposition was not attC sequence dependent; however, the findings alluded to the importance of attC structural recognition.

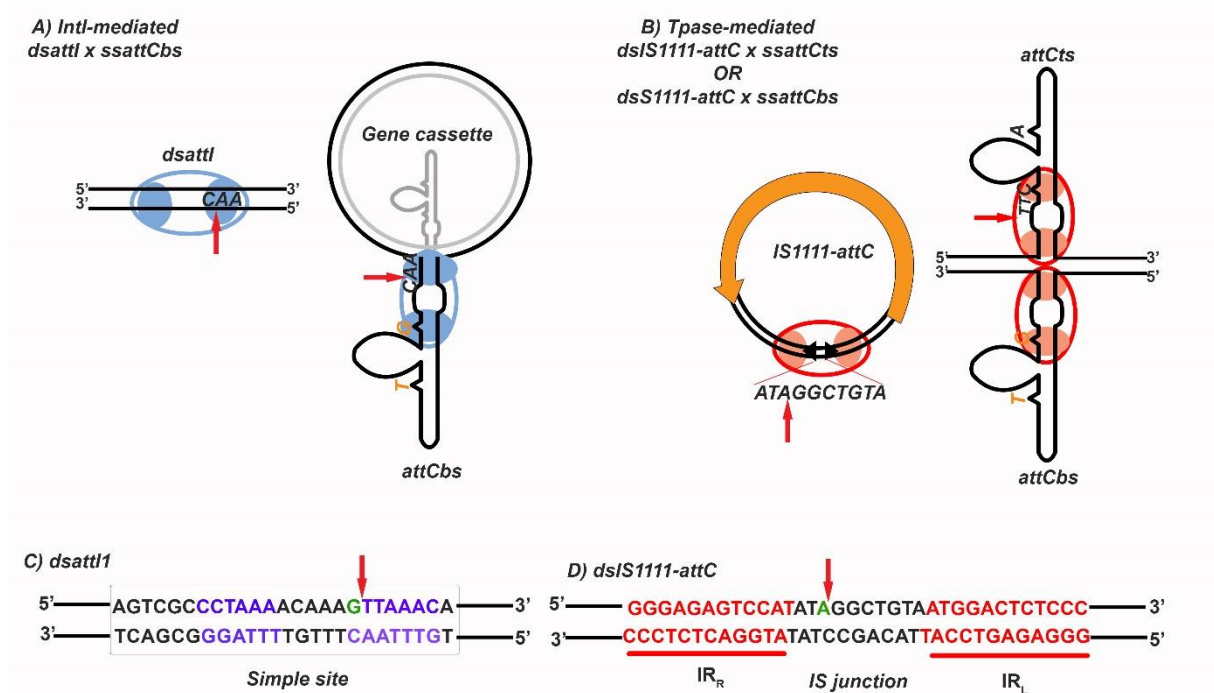


Figure 5.1: a) IntI-mediated recombination of dsattI and ssattC_{bs} compared to proposed B) IS1111-attC transposase mediated recombination of dsIS1111-attC minicircle junction and ssattC_{ts} or ssattC_{bs}.

In a) only one (black) attC strand participates. In b) both the attC top strand (ts) and bottom strand (bs) can participate. Arrows point to the nucleotide pair involved in recombination cross-over where nucleotides highlighted in green in panel c) and d) are key nucleotides.

Given that both top and bottom strands of attC are possibilities for IS1111-attC Tpnase recognition, I aimed to explore whether or not a single strand was a recombination partner, and, if so, which strand that is. IS1111-attC inserts into the precise attC palindromic position (positions *h* and *i*) despite variation in attC sequences (Tetu & Holmes, 2008). The presence of extrahelical bases is important for specificity and polarity in the IntI system (Francia *et al.*, 1999; MacDonald *et al.*, 2006) and is predicted to be essential for IS1111-attC transposase binding also. If Tpnase and IntI both target the same attC bottom strand, and recognize the same extrahelical bases (G and T) (Figure 5.2 A), this is predicted to result in a competitive environment between these two heterologous enzymes. If IS1111-attC Tpnase preferentially recognises the attC top strand, it would have a spatially separated binding site as illustrated in Figure 5.1, thereby reducing any competitiveness with IntI. attC structure and strand availability is thus postulated to play a key role in Tpnase binding and overcoming molecular-

genetic constraints. For instance, the role of the variable terminal structure (VTS) in influencing single-stranded template formation and affinity for IntI attachment has been demonstrated (Bouvier *et al.*, 2009; Loot *et al.*, 2010). Thus it is unlikely that the polarity comes from the IS1111-attC minicircle. Tpmase is postulated to be able to recognise a tag between the two 12-base repeats, thereby reducing overall polarity. High tolerance to sequence variation, specificity of insertion point and orientation bias appear to be shared features of IntI and IS1111-attC Tpmase.

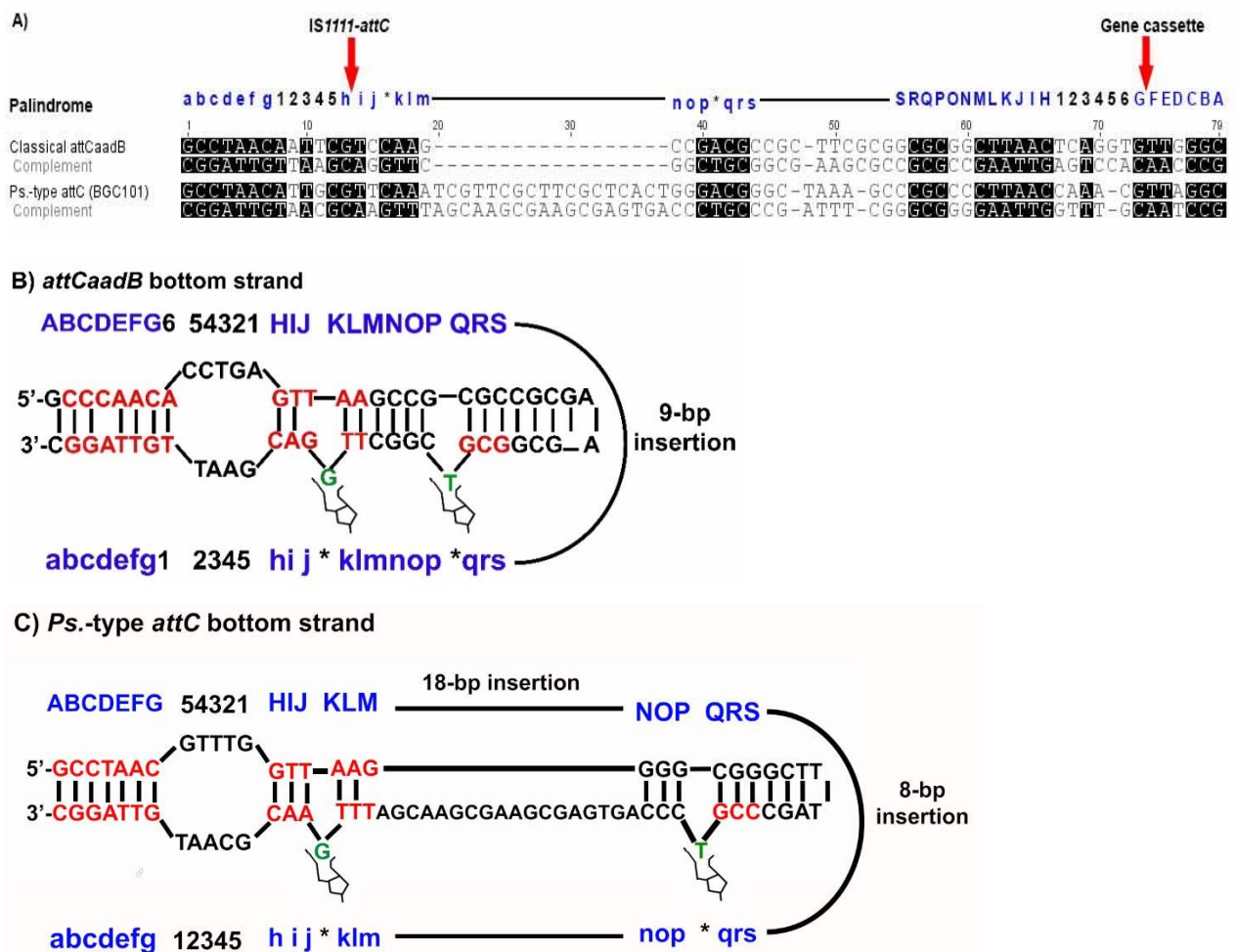


Figure 5.2: Comparison between *attCaadB* and *Pseudomonas*-type (BGC101) *attC*
a) classical *attCaadB* and *Pseudomonas*-type *attC* aligned according to secondary structure position prediction (blue). Black shading notes sequence identity between the two *attC* sites. Arrows point to IS1111-attC and Gene cassette insertion points
b) classical *attCaadB*. *attCaadB* has a 9-bp insertion between positions *s* and *S*.
c) *Pseudomonas*-type *attC* has an 18-bp insertion at *m-n* and an 8-bp insertion at *s-S*.
Structural similarities between these two *attC* sites and the flipped out extrahelical G and T nucleotides necessary for IntI docking (green). Bases in red depict protein-phosphate contacts of IntI. Adapted from Bouvier *et al.*, 2009.

5.2 Aims

Given the biological importance of integrons in bacterial evolution and adaptation, understanding how IS1111-attC elements and in particular the transposase interact with integron systems will provide molecular insights into this complex system. Firstly, an *in vitro* system for IS1111-attC transposase expression would need to be developed based on IntI research (Gravel *et al.*, 1998). The aims were:

- 1) To determine IS1111-attC Tase sequence specificity or preference to potential recombination substrates.
- 2) To study the binding of IS1111-attC Tase to double stranded recombination substrates.
- 3) To study the binding of IS1111-attC Tase to single stranded recombination substrates.

5.3 Methods

For general media and solution recipes, as well as PCR conditions, refer back to Chapter 2.

5.3.1 PCR, restriction digestion and cloning of IS1111-attC *tpase* into pMAL-c2X

A 1044-bp DNA fragment encoding IS1111-attC transposase was amplified by PCR from *Pst405* (BGC101) genomic DNA by using the primers MZ03 and MZ04. The PCR product was purified using Qiagen PCR Purification kit, and a double digest using *Bam*HI and *Pst*I (New England BioLabs, Canada) was performed using 500 ng of purified PCR product at 37°C, overnight. Enzyme inactivation was performed at 65°C for 20 min prior to an additional round of purification. 500 ng of pMAL-c2X (NEB) was also digested overnight using *Bam*HI/*Pst*I at 37°C. Digested vector was subsequently dephosphorylated for 1 hr at 37°C using Antarctic phosphatase (New England BioLabs, Canada) as per manufacturer's instructions.

Ligation of insert and vector was done at a ratio of 3:1 at 4°C overnight using T4 DNA ligase, 10x T4 DNA Ligation buffer, 10x ATP. The ligation mixture was transformed into chemically competent *E. coli* JM100 cells by heat shock at 42°C (Section 2.6.1). Plating was done on LB/Amp¹⁰⁰ at 37°C overnight. Positive ligation clones were identified using a PCR screen where one primer targets the *tpase* gene and other is a plasmid backbone primer. Primer combinations were MZ03/ELF22, and MZ04/ELF21. Positive clone was grown overnight in 50 mL LB/Amp¹⁰⁰ and the plasmid was purified as described in section 2.7. Purified constructs (Figure 5.3) were stored at -20°C and were used to transform *E. coli* Rosetta cells for protein expression.

5.3.2 Overexpression of recombinant proteins

5.3.2.1 Cell expression system- *E. coli* Rosetta 2 cells

The *E. coli* Rosetta 2 strain was used as the cell expression system. This strain is a BL21 derivative designed to enhance the expression of eukaryotic proteins that contain codons rarely used in *E. coli*. It supplies tRNAs for 7 rare codons (AGA, AGG, AUA, CUA, GGA, CCC, CGG), which were all strongly represented in the *tpase* gene. 50 µL of Rosetta cells were thawed on ice and 50 µL of KCM buffer was added to the cells with 1 µL pMAL-Tpase construct. As a control, pMAL-c2X was added to a separate tube of Rosetta cells. The cell suspension was mixed and incubated on ice for 20min. Cells were then heat shocked at 42°C

for 90 sec, 200 μ L of LB was immediately added and cells incubated at 37°C for 1.5 hrs before being plated onto LB/Carb¹⁰⁰/Cm^{12.5}/Glu^{2%}. Transformants were recovered overnight at 37°C.

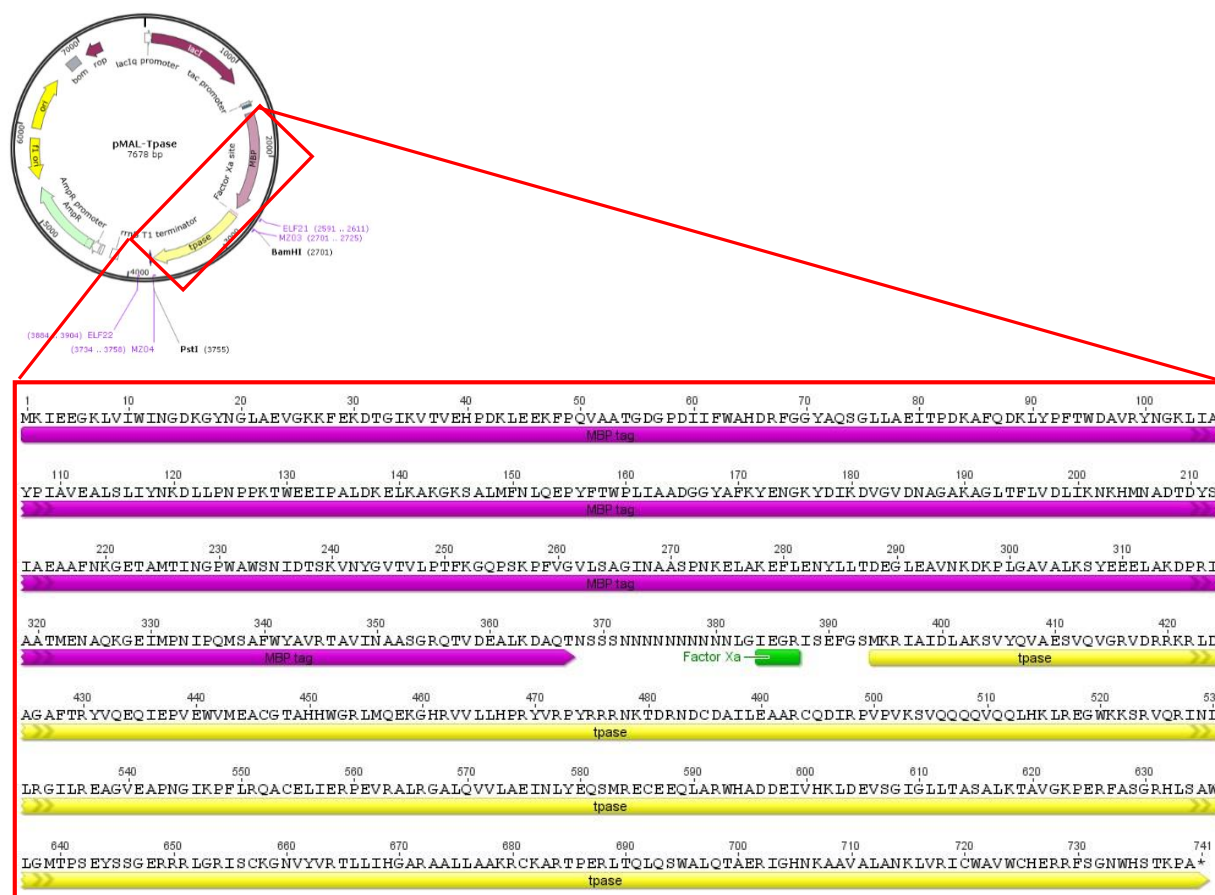


Figure 5.3: Expression construct used to produce a Maltose-binding transposase fusion protein. The BamHI and PstI sites were used to clone the tpase gene into the pMAL-c2X vector backbone. Blown out panel is the corresponding amino acid sequence of the ISPst6 Pst405 tpase. MBP-tag amino acid sequence is underlined in purple, Factor Xa site in green and Tpase in yellow.

5.3.2.2 Expression of fusion protein

Figure 5.4 outlines the expression and purification steps of MBPtpase from Rosetta cells. Single colonies of transformed *E. coli* Rosetta 2 were used to inoculate LB containing carbenicillin (100 μ g/mL), chloramphenicol (12.5 mg/mL) and 2% glucose and shaken overnight at 180 rpm and 37°C. Six litres of LB media containing carbenicillin (100 μ g/mL), chloramphenicol (12.5 mg/mL) and glucose (0.2%) was inoculated with the overnight culture to a starting OD_{600nm} of 0.05. Cultures were shaken at 180 rpm at 37°C until an OD_{600nm} of 0.3 was reached. The temperature was then lowered to 16°C and cells were grown to an OD_{600nm} of 0.5 still shaking at 180 rpm for approximately 2 hrs. Overexpression of recombinant

protein was induced by addition of isopropyl β -D-1-thiogalactopyranoside (IPTG) to 0.3 mM, the culture was incubated overnight (~20 hrs). Cells were harvested by centrifugation (4000 g, 20 min, 4°C), resuspended in 30 mL MBP column buffer and stored at -20°C. Overexpression was confirmed by SDS-Page analysis on 10% polyacrylamide gel run at 110 Volts 40 min, stained for 30 min in Coomassie blue and destained in low destain solution overnight.

5.3.2.3 Cell lysis

Cell pellets were thawed in ice water baths and DNase I (1 mg/mL) (Sigma-Aldrich, USA) was added to a final concentration of 10 μ g/mL. Cells were incubated on ice for 10 min before sonication in 8x15 sec bursts with a 30 sec resting period in between the bursts. Soluble and insoluble fractions were separated by centrifugation at 4000 g, 20 min at 4°C.

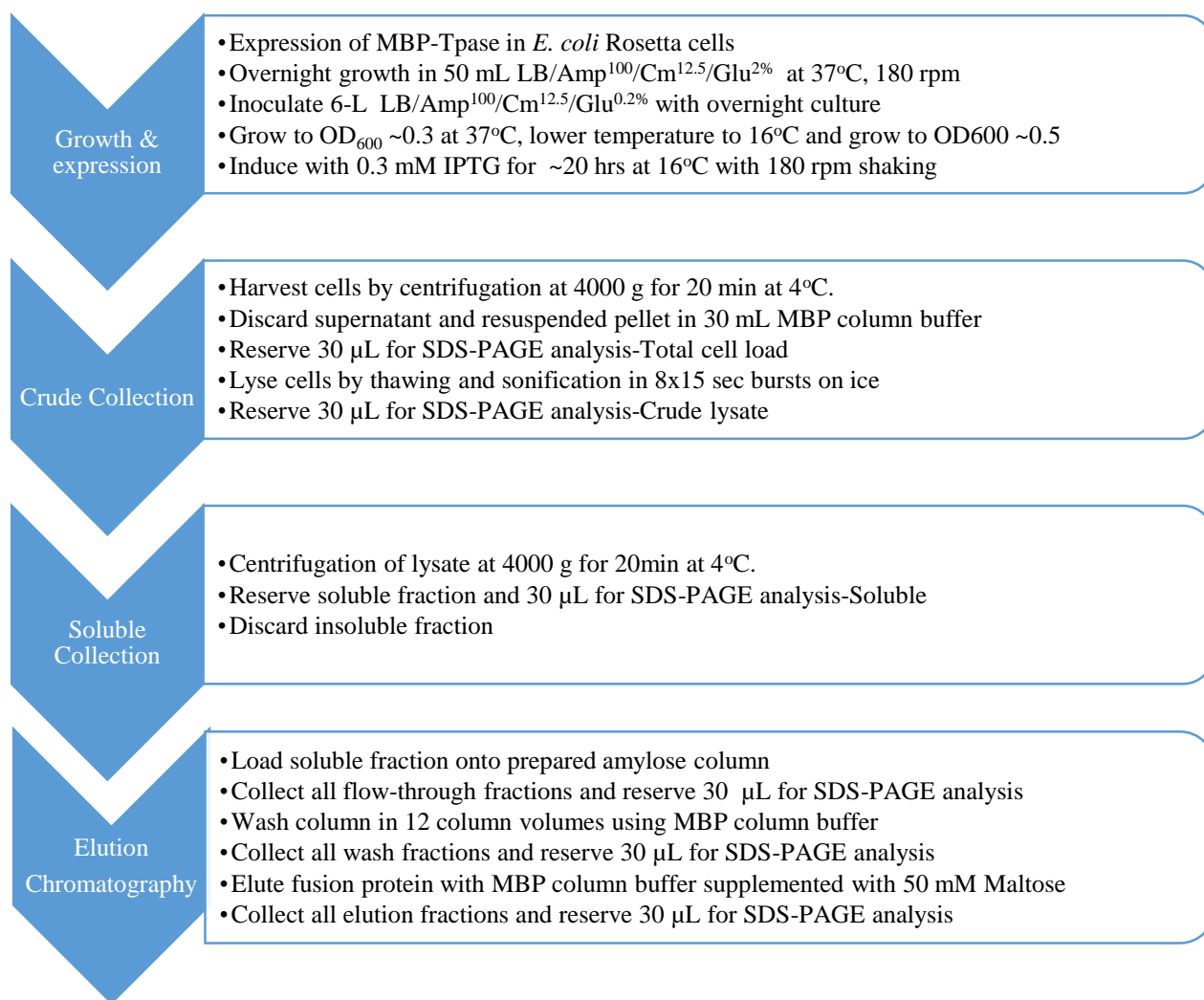


Figure 5.4: Flow diagram of the protein preparation strategy. At each purification step, 30 μ L of liquid is reserved for subsequent SDS-PAGE analysis.

5.3.3 Purification and identification of MBPTpase fusion protein

5.3.3.1 Purification using amylose affinity chromatography

The amylose resin was poured into a 2.5 x 10 cm column. The column was prepared by washing with 5 column volumes of MBP column buffer, 3 volumes of reverse osmosis water, followed by 3 volumes of MBP column buffer. The soluble protein fraction was then loaded onto the column and all of the flow-through was collected. The column was then washed in 12 column volumes of MBP column buffer. Finally, the fusion protein was eluted with MBP column buffer supplemented with 50 mM maltose. 10 fractions of 1 mL each were collected and purification was confirmed by SDS-PAGE analysis.

5.3.3.2 In-gel trypsin digestion

Protein bands of interest were excised from the gel with a sharp scalpel using a clean scalpel per protein band of interest. Protein strips were cut into 1-2 mm³ cubes and put into 1.5 mL microcentrifuge tubes. Gel cubes were destained using 500 μ L of 60% (v/v) 50 mM ammonium bicarbonate (Sigma-Aldrich) and 100 μ L 40% acetonitrile for 1 hr at room temperature. If needed, this step was repeated 2-4 times until all Coomassie blue was washed out of the gel cubes. Gel cubes were washed by adding 100 μ L 100% (v/v) acetonitrile and vortexing to mix. The above washing step was repeated until the gel plugs were solid white in colour. Gel plugs were completely dried using vacuum centrifugation (30min, 40°C, Concentrator 5301, Eppendorf). Dehydrated cubes were rehydrated in 8 μ L of 50 mM ammonium bicarbonate containing 12 ng/ μ L modified sequencing grade trypsin (New England BioLabs, Canada) for 60 min at 4°C. This will bring trypsin to its active pH 7-9. Excess trypsin solution was removed by centrifugation and the rehydrated gel cubes were incubated overnight at 37°C in 20 μ L of 50 mM ammonium bicarbonate.

5.3.3.3 Peptide Mass Fingerprinting

For MALDI-TOF MS analysis, peptides were desalted and concentrated using C₁₈ micro-columns (PerfectPure, Eppendorf) according to the manufacturer's protocol. Peptides were eluted onto a MALDI plate with 2 μ L of matrix solution (10 mg/mL α -cyano-4-hydroxycinnamic acid in 0.1% TFA and 50% acetonitrile). Mass spectra were acquired in the mass: charge range of 800-3500 m/z on a Q-STAR XL Hybrid mass spectrometer equipped with a MALDI ion source (Applied Biosystems Inc., Foster City, USA). The instrument was

calibrated using [Glu¹] Fibrinopeptide B peaks 175.1200 and 1570.6774 (Sigma-Aldrich, St. Louis, MO, USA). The resulting protein mass spectra were manually inspected and analysed with Analyst QS 1.1 software (Applied Biosystems, Bardon, Australia). The generated monoisotopic peak masses were compared to the predicted *in silico* digests of MBPase fusion protein.

5.3.4 Electrophoretic Mobility Shift Assays (EMSA)

5.3.4.1 Radiolabelling of oligonucleotides

All oligonucleotides were manufactured by IDT (Coralville, Iowa) (Table 5.2). Oligonucleotides were resuspended in sterile MQ water to produce ~100 μ M stocks. Double stranded (ds) DNA fragments for binding assays were generated by heating equimolar concentrations of top and bottom strands oligonucleotides in a total volume of 50 μ L with 1 μ L of 100 mM NaCl added. The DNA mixture was heated for 20 min at 95°C, and cooled gradually in the heat block overnight to room temperature to generate dsDNA. Single stranded and double stranded oligonucleotides were 5'-end labelled with a ~20 pmol [γ -³²P] ATP (300 μ Ci/mmol) in T4 polynucleotide kinase (PNK) buffer (70mM Tris-HCl, 10mM MgCl₂, 5mM DTT, pH6.7). T4 PNK was added to 500 mU/ μ L and incubated at 37°C for 1 hr, then inactivated by heating at 65°C for 20 min. Unincorporated [γ -³²P] ATP was removed by centrifugation in a mini Quick Spin oligo column (Perkin-Elmer) for 4 min at 1000 g. Radiolabelled oligonucleotides were stored at ~20°C.

5.3.4.2 Protein-DNA sample preparations

To test MBPase binding to labelled oligonucleotides, binding reactions were set up using a constant concentration of ³²P-radiolabelled oligonucleotide (150 pmol) and increasing concentrations of protein (Table 5.1). Fusion protein and oligonucleotide were added to a reaction mixture (10 mM HEPES, 50 mM KCl, 2.5 mM MgCl₂, 100 mM DTT, 67 μ g/mL BSA, pH 7.7) in a final volume of 25 μ L. Samples were briefly spun down and EMSA 6x loading dye was added just prior to loading onto a Native EMSA gel.

Table 5.1: Tested protein concentrations for corresponding oligonucleotide

<i>attC</i>	Strand	Protein Concentrations tested											
<i>Pseudomonas</i> -type	<i>ds</i>	20 nM	50 nM	100 nM	200 nM	400 nM	600 nM	800 nM	1 μ M	2 μ M	3 μ M	4 μ M	5 μ M
	<i>ssTop</i>	200 nM	400 nM	600 nM	800 nM	1 μ M	2 μ M	3 μ M	4 μ M	5 μ M	7.5 μ M	10 μ M	20 μ M
	<i>ssBottom</i>	200 nM	400 nM	600 nM	800 nM	1 μ M	2 μ M	3 μ M	4 μ M	5 μ M	7.5 μ M	10 μ M	20 μ M
Classical <i>attC_{andB}</i>	<i>ds</i>	100 nM	200 nM	400 nM	600 nM	800 nM	1 μ M	2 μ M	4 μ M	5 μ M	10 μ M	20 μ M	40 μ M
	<i>ssTop</i>	1 μ M	2 μ M	3 μ M	4 μ M	5 μ M	7.5 μ M	10 μ M	20 μ M	25 μ M	30 μ M	35 μ M	40 μ M
	<i>ssBottom</i>	1 μ M	2 μ M	3 μ M	4 μ M	5 μ M	7.5 μ M	10 μ M	20 μ M	25 μ M	30 μ M	35 μ M	40 μ M

5.3.4.3 EMSA Native PAGE

A 180 mm x 160 mm x 1.5 mm, 8% (w/v) non-denaturing polyacrylamide (acrylamide/bisacrylamide 19:1) gel made in 1x TBE buffer (5 mM Tris, 9 mM boric acid, 0.25 EDTA pH7.8) was pre-run in 0.5x TBE at 110 volts for 30 min. Samples were subjected to electrophoresis at 24 mA for 2.5 hrs at room temperature. The gel was transferred to Whatman filter paper and exposed on a phosphor-imaging screen for 16 hrs before imaging with a Typhoon Phosphorimager (GE Healthcare).

5.3.4.4 EMSA Native PAGE - Competition assays

To control for specificity of binding of transposase, EMSA competition assays were performed (Figure 5.5). An equal concentration (40 μ M) of MBP_{TPase} and MBP control was used across all binding reactions. All labelling and binding reactions were conducted as per section 5.3.4.1 and 5.3.4.2, however increasing concentrations (10x-50x) of unlabelled oligonucleotide (cold target) were also added. In the final reaction, 100x non-specific competitor, *EGG-m2F* or *EGG-m2R* or *EGG-m2* (Table 5.3) was added. If the transposase bound specifically to the target site, this interaction will be unaffected by the presence of a non-specific competitor target.

Table 5.2: DNA structures tested in EMSA experiments

	Sites	Strand	Sequence (5'-3')	Additional comments
<i>Ps.-type attC</i> (76-bp)	<i>BGC101T</i>	<i>Top</i>	5' GCCTAACATTGCGTTCAAATCGTTCGCTTCGCTCACTGGGACGGGCTAAAGCCCGCCCCTTAACCAAACGTTAGGC 3'	ISPst6 already resides within this attC-site in <i>Pst405</i> . Transposase used in this study was cloned from this strain.
	<i>BGC101B</i>	<i>bottom</i>	5' <u>GCCTAACGTTT</u> GGTTAAGGGGCGGGCTTAGCCCGACCCAGTGAGCGAAGCGAACGATTGAACGCAATGTTAGGC 3'	
	<i>BGC101</i>	<i>dsDNA</i>	5' GCCTAACATTGCGTTCAAATCGTTCGCTTCGCTCACTGGGACGGGCTAAAGCCCGCCCCTTAACCAAACGTTAGGC 3' 3' CGGATTGTAACGCAAGTTTAGCAAGCGAAGCGAGTGACCCTGCCCGATTTCGGGCGGGGAATTGGTTGC <u>AATCCG</u> 5'	
Classical attC _{aadB} (60-bp)	<i>aadBT</i>	<i>Top</i>	5' GCCTAACAAATTCGTCCAAGCCGACCCGCTTCGCGGCGCGGCTTAACTCAGGTGTTGGGC 3'	This site naturally harbours an IS1111-attC (ISPax2). ISPax2 transposase is 79% identical to ISPst6 transposase purified in this study.
	<i>aadBB</i>	<i>bottom</i>	5' <u>GCCCAACACCT</u> GAGTTAAGCCGCGCCGGAAGCGGCGTCGGCTTGGACGAATTGTTAGGC 3'	
	<i>aadB</i>	<i>dsDNA</i>	5' GCCTAACAAATTCGTCCAAGCCGACCCGCTTCGCGGCGCGGCTTAACTCAGGTGTTGGGC 3' 3' CGGATTGTTAAGCAGGTTGCGGCTGGGCGAAGCGCCGCGCCGAATTGAGTCCAC <u>AACCCG</u> 5'	
ISJunction (50-bp)	<i>ISJuncT</i>	<i>Top</i>	5' TGCTTGTTGGGGAGAGTCCATATAGGCTGTAATGGACTCTCCCTGCACCAC 3'	This site is formed in the minicircle of IS1111-attC elements and is strongly conserved. It is the cognate one for the cloned protein.
	<i>ISJuncB</i>	<i>bottom</i>	5' GTGGTGCAGGGAGAGTCCATTACAGCCTATATGGACTCTCCCAACAAGCA 3'	
	<i>ISJunc</i>	<i>dsDNA</i>	5' TGCTTGTTGGGGAGAGTCCATATAGGCTGTAATGGACTCTCCCTGCACCAC 3' 3' ACGAACACCCCTCTCAGGTATATCCGACATTACCTGAGAGGGACGTGGT 5'	

The sequences are all derived from integrated attC sites within PCIs and MRI arrays. This is the form in which they are predicted to appear as substrates for the IS1111-attC Tpsase. The underlined region is seen by IntI during integrative reactions between attC x attI. Also note that T4 Polynucleotide kinase (PNK) efficiently catalyses the transfer of the gamma-phosphate from ATP to the 5'-OH group on both single and double stranded DNA (Berkner & Folk, 1977).

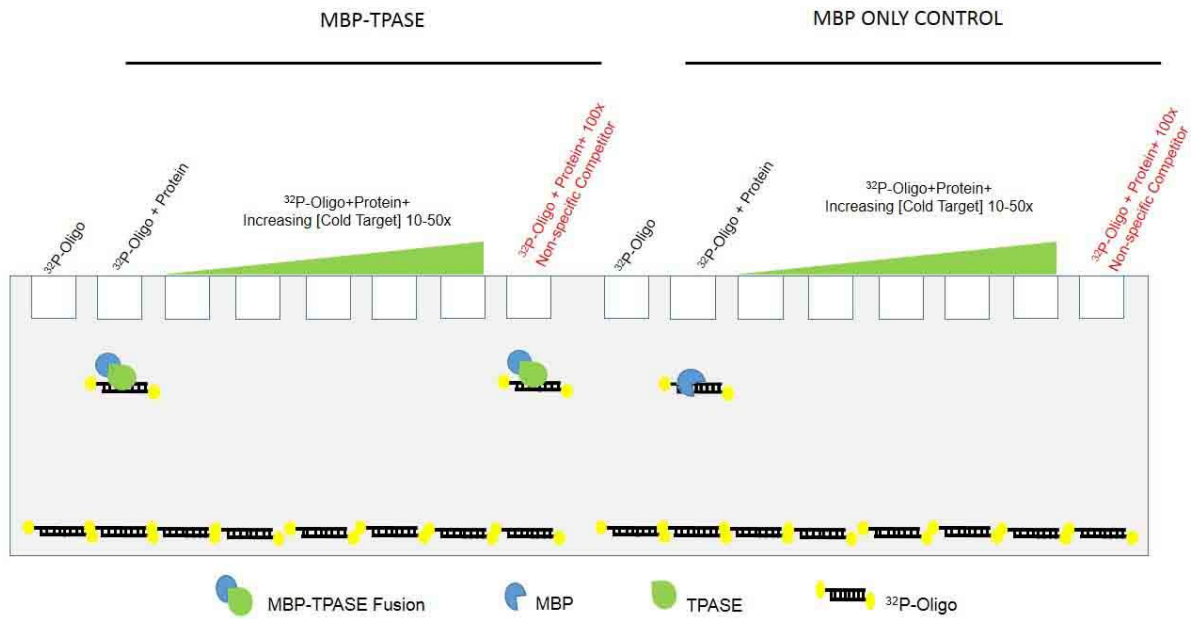
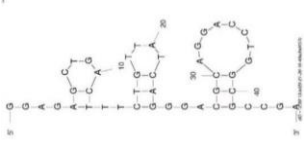
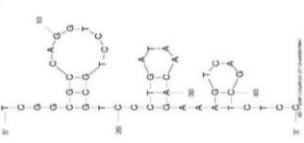


Figure 5.5: Schematic representation of an EMSA competition assay.

If $_{MBP}Tpase$ binds a hot target, the protein-DNA complex is detected on the native EMSA. This binding will disassociate with the addition of increasing concentrations (10x-50x) of cold target due to the specific binding of the cold target to the T $_{pase}$, hence binding will not be detected. If the $_{MBP}Tpase-attC$ interaction is unaffected by the presence of a 100x excess non-specific competitor, the protein-DNA complex will be detected on the native EMSA.

In the case of the MBP only control, the addition of all competing agents is expected to interfere with detection, thus no shift in mobility is expected to occur.

Table 5.3: Non-specific competitor EGG-m2F/R

Name (45bp)	Strand	Nucleotide sequence	M-Fold Structural Prediction	Additional Comments
EGG-m2F	top	5'GGAGAGCTGACTTTCTGTTATCAGGGACGCAGGAGGAGGCGCCGA3'		The 'sense' strand in the original source*
EGG-m2R	bottom	5'TCGGCGCCTCCTCCTGCGTCCCTGATAACAGAAAGTCAGCTCTCC3'		The 'anti-sense: strand in the original source*
EGG-m2	dsDNA	5'GGAGAGCTGACTTTCTGTTATCAGGGACGCAGGAGGAGGCGCCGA3' 3'CCTCTCGACTGAAAGACAATAGTCCCTGCGTCCCTCCTCCGCGGCT3'		The dsDNA molecule was formed by annealing equimolar amounts of the two oligos under conditions described in 5.2.4.4.

*This nucleotide sequence includes a zinc finger binding motif. This zinc finger binding motif is unrelated to attC and does not have any palindromic properties. Also, the topology between the EGG-m2 competitor differs to that of the IS binding partners in Table 5.2. The secondary structure needed so as to present the two extra-helical residues to Tpmase do not form in the EGGm2 sequence.

5.4 Results

5.4.1 *in silico* analysis of IS1111-attC Transposase

DNA recombinases have been frequently found to be challenging to purify due to their hydrophobic nature, large size and/or need for accessory protein binding partners (Table 5.4). Based on *in silico* analysis using <http://web.expasy.org/protparam/>, the ISPst6 transposase was predicted to be ~38 kDa in size and have a weak hydrophilic property (-.478 GRAVY). Similarly, IntI1 is also predicted to be ~39 kDa in size and have a weak hydrophilic property (-.313 GRAVY). Weak hydrophobicity in both recombinases is indicative of them having a hydrophobic core region (Messier & Roy, 2001). Furthermore, both proteins have similar isoelectric values greater than 10. Previous workers had found that inclusion of soluble domains such as FLAG or MBP improved IntI solubility and purification (Collis *et al.*, 1998; Gravel *et al.*, 1998). Given this, a Maltose Binding Protein (MBP) tag was selected as an appropriate protein tag to increase the overall solubility of IS T_{pase}, and a protease site, Factor Xa, was included to enable recovery if necessary.

Table 5.4: *in silico* predictions of IS1111-attC transposase and IntI1 using ProtPram ExpASy

	T_{pase}	IntI1*
Number of amino acids (aa)	347	337
Molecular weight (Da)	39679.8	38381.1
Theoretical pI	10.14	10.25
Total number of negatively charged residues (Asp + Glu)	36	32
Total number of positively charged residues (Arg + Lys)	61	48
Ext. coefficient	60430 M ⁻¹ cm ⁻¹ , at 280 nm	57410 M ⁻¹ cm ⁻¹ , at 280 nm
Instability index (II)	55.02 unstable	46.85 unstable
Aliphatic index	90.26	90.00
Grand average of hydropathicity (GRAVY)	-0.478 Hydrophilic	-0.313 Hydrophilic

*IntI1, class I integron integrase from *Ps. aeruginosa* plasmid pVS1 (Accession number AAC44315.1) (Bissonnette & Roy, 1992).

17.5% (61/347 aa) and 14.2% (48/337 aa) of the predicted T_{pase} and IntI1 sequence respectively are positively charged amino acids. To avoid potential for non-specific interaction with DNA during protein purification, DNase I was included. Furthermore, given the G+C richness of *tpase* gene, *E. coli* Rosetta 2 cells were used as the expression host strain.

5.4.2 Expression and purification of *MBP*Tpase and *MBP* control

Initial attempts to express transposase at 37°C resulted in barely detectable levels of *MBP*Tpase fusion protein in the crude cell lysate by SDS-PAGE analysis (Figure 5.6A). Altering the concentration of IPTG from 0.4 mM to 1 mM also failed to enhance expression. Induction at high temperature is believed to result in packaging of *MBP*Tpase into inclusion bodies. Attempts to purify the *MBP*Tpase from inclusion bodies using 8M Urea were successful (data not shown) however, the final yields of folded-soluble *MBP*Tpase were low ($\leq 1 \mu\text{M}$). Additionally the use of urea may pose a problem as it can denature the transposase, thereby possibly making it inactive for DNA binding. To avoid the formation of inclusion bodies by “leaky expression”, glucose was added to the overnight culture medium, to suppress the *lac*-promoter, and the induction temperature was lowered to 25°C. This resulted in higher *MBP*Tpase yields (Figure 5.6B).

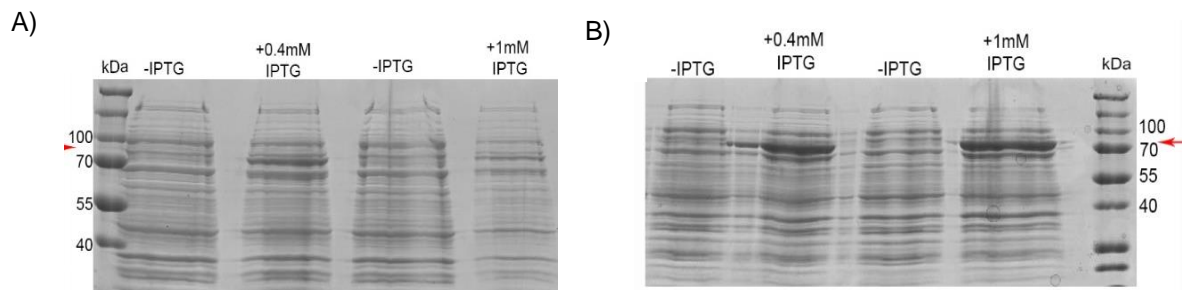


Figure 5.6: SDS-PAGE analysis reveals poor overexpression of *MBP*Tpase fusion protein in Rosetta cells induced with 0.4 mM and 1 mM at A) 37°C and B) 25°C. Samples were taken before (-IPTG) and after (+IPTG) induction. Arrow indicates expected product of ~81 kDa. 30 μL of cell mass was boil-lysed in 30 μL SDS loading dye at 99°C for 15 min. 10 μL of crude extract was loaded onto the SDS-PAGE gel.

Initial attempts to recover the fusion protein by affinity chromatography gave very poor yields. Most of the fusion protein was lost in the insoluble fraction. To improve final *MBP*Tpase yield, the culture conditions for induction and expression were lowered to 16°C and the growth period extended to a maximum of 24 hours. Additionally the culture volume was increased from 50 mL to 6-L (1x 1-L culture flasks). This resulted in significantly improved yield of *MBP*Tpase fusion protein in whole cell lysates (Figure 5.7).

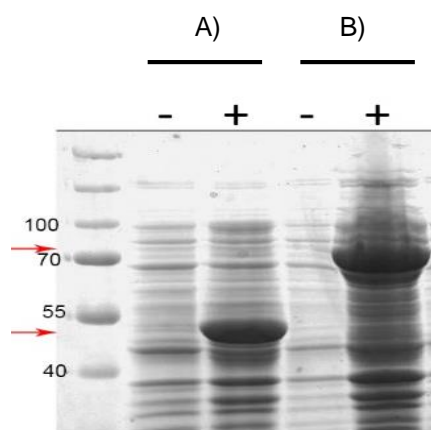


Figure 5.7: SDS-PAGE of A) MBP control and B) $_{MBP}Tpase$ fusion proteins, expressed at 16°C for 24 hours.

Samples were taken before (-IPTG) and after (+0.2 mM IPTG) induction. Expressed bands of the expected size (A) ~42kDa and (B) ~81kDa are observed in post induction lanes. 30 μ L of cell mass was boil-lysed in 30 μ L SDS, 10 μ L of crude extract was loaded onto the SDS-PAGE gel.

An initial low speed centrifugation step was performed to remove insoluble protein from the cell lysate prior to affinity chromatography. Loading the protein onto the column and subsequently washing the column did result in some loss of the $_{MBP}Tpase$ fusion (Figure 5.8A, lane 4). However, a loss was to be expected given that not all of the fusion protein was predicted to be folded correctly. To elute the maximum concentration of the fusion protein from the amylose column, the concentration of maltose in the elution buffer used was increased from 10 mM to 50 mM (Figure 5.8A, lanes 5-9). Upon elution with 50 mM maltose buffer, the fusion protein ran at the expected ~81 kDa size, with the greatest concentration of protein found consistently in fractions 2 and 3 (lanes 6 and 7). However, additional bands with molecular weights of less than ~81 kDa were unexpectedly observed (Figure 5.8B). In parallel, the MBP control was eluted as a single band of ~42 kDa by the addition of 50 mM maltose buffer (Figure A27).

5.4.3 Mass Peptide fingerprint of MBP-Fusion protein

A single band of ~81 kDa corresponding to $_{MBP}Tpase$ was expected, however as per SDS-PAGE analysis multiple bands were observed. To verify that the ~81 kDa band was the expected $_{MBP}Tpase$ fusion protein, MALDI-TOF analysis was conducted. Also the smaller bands corresponding to estimated weights of 60 kDa, 39 kDa, and 30 kDa (Figure 5.8B) were randomly chosen and also analysed.

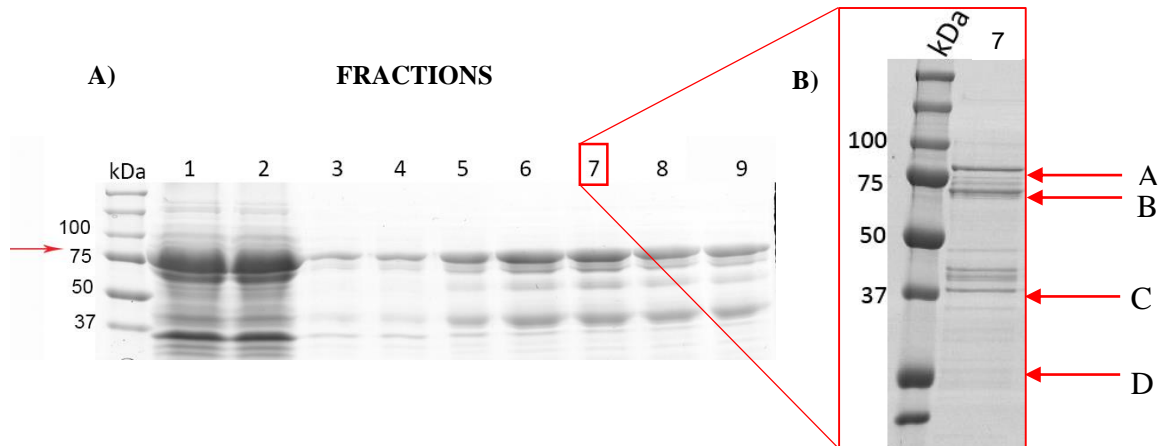


Figure 5.8: SDS-PAGE analysis of purified *MBPtpase* fusion protein.

- A) *MBPtpase* was expressed in *E. coli* Rosetta2. Fraction 1: Cell lysate (soluble and insoluble fraction). Fraction 2: soluble fraction only. Fraction 3: flow through amylose resin. Fraction 4: Column wash. Fractions 5-9: *MBPtpase* was eluted with 50 mM maltose. Expected protein size is ~81 kDa.
- B) SDS-PAGE of *MBPtpase* elution fraction 7. Arrows point to bands that were used for MALDI-TOF analysis. A= ~81 kDa; B= 60 kDa; C= 39 kDa; D= 30 kDa

Comparing the MALDI-TOF spectra of the four analysed bands suggested that the three smaller protein products are derived from *MBPtpase* fusion protein (Figures 5.9 & 5.10). Major peaks observed in the 81 kDa spectra, such as 1644.848 (m/z) and 1302.534 (m/z), were absent in the other spectra. Matching the MALDI-TOF spectra to the predicted *in silico* digest of *MBPtpase* amino acid sequences suggests that the protein was being truncated from the C'-terminus (Figure 5.9). Despite the addition of protease inhibitor cocktails during affinity purification of the fusion protein, this "truncated" banding pattern was continuously observed. Although multiple bands were obtained, the intact MBP-Fusion protein was estimated to range between 50%-70% of total protein yield based on SDS-PAGE analysis. Only elution fractions containing this estimated range and the same total protein concentration, as determined by spectrophotometry, were used for subsequent DNA binding assays. To determine the protein concentration the following formula was used: [(reading at 260:280)/predicted *MBPtpase* extinction coefficient $M^{-1} cm^{-1}$] $\times 10^6 = [mg/mL]$.

Truncated proteins are often the result of *in vivo* endoproteolytic cleavage. Despite the addition of protease cocktail inhibitors, truncated *MBPtpase* forms were present. This implies that these smaller forms are products of protein cleavage, or products of incomplete translation since fusion to the N-terminus can reduce the efficiency of protein translation (Terpe, 2003). The formation of truncated versions of the *MBPtpase* may have been reduced if

the tag was fused to the C-terminus instead. The use of a vector which tags both the C and N-terminal ends of the T_pase may protect the protein from degradation (Hamilton *et al.*, 2002; Podmore & Reynolds, 2002). A Polyhistidine-tag (His-tag) was considered at this stage, however the associated purification procedure requires the use of denaturants, high levels of salt and imidazole, all of which may interfere with the DNA binding assay. Imidazole in particular can lead to protein aggregates (Hefti *et al.*, 2001). Thus, EMSA were carried out using a MBP bound T_pase, despite the presence of multiple MBP T_pase protein variations.

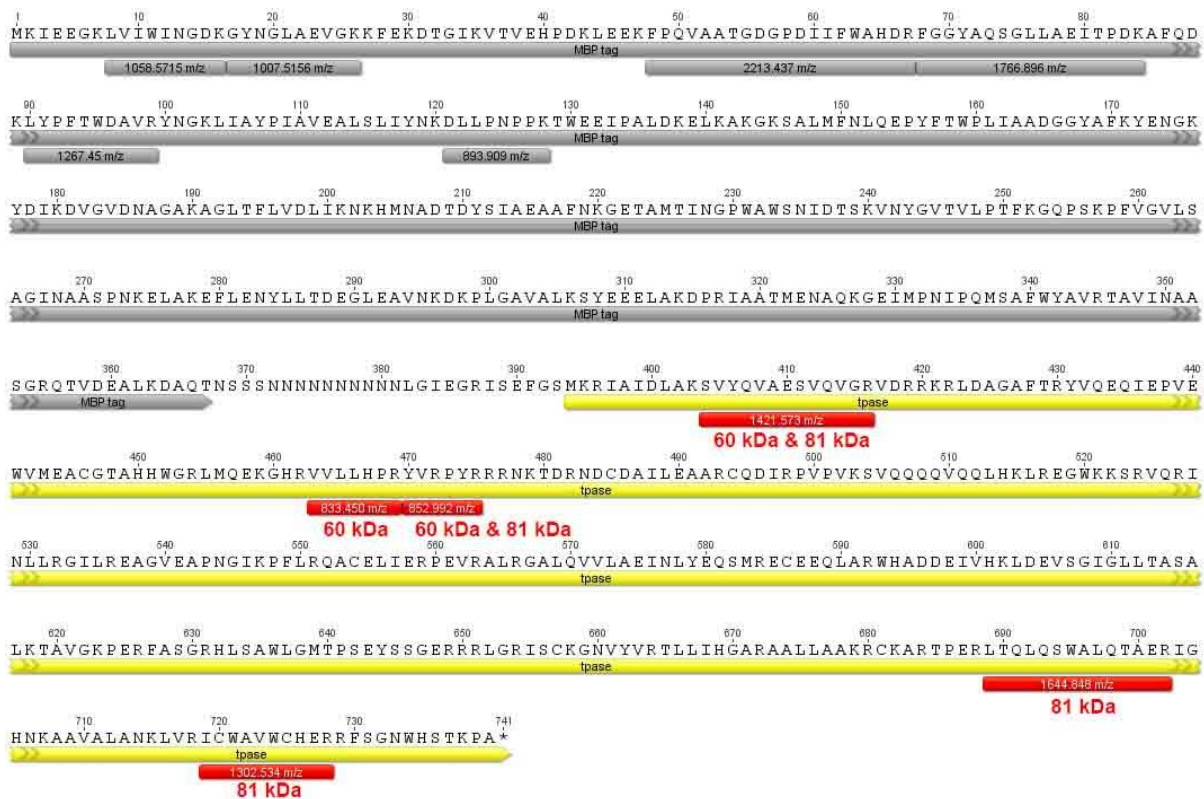


Figure 5.9: Amino acid sequence of the MBP T_pase fusion protein showing the major expected peptides and their relative mass: charge ratio.

Peptide prediction from *in silico* digest of MBP T_pase amino acid sequence using PeptideMass (SwissProt). Yellow denotes the transposase amino acid sequence. Grey peptides are present in all excised bands and correspond to the MBP tag. Red peptides are observed in the 60kDa and 81 kDa bands (see figure 5.10 for MALDI-TOP spectra).

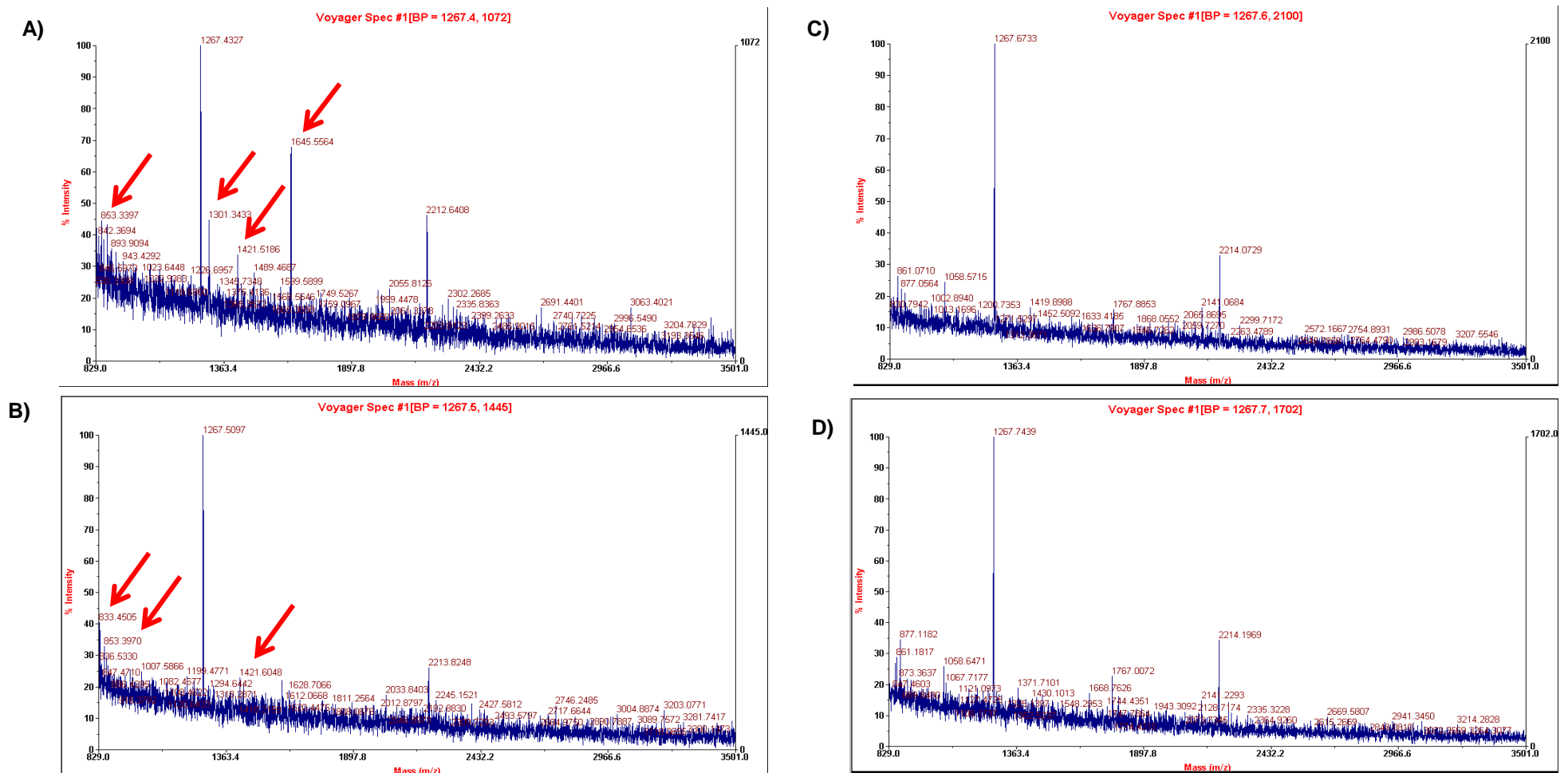


Figure 5.10: MALDI-TOF spectra of excised bands a) 81kDa, B) 60kDa, C) 39kDa, D) 30kDa. Peaks 1645.5504 (m/z) and 1301.3433 (m/z) found in 81kDa band are missing absent in all other spectra. MBP_Tpase peptides denoted by arrows.

5.4.4 DNA Binding assays

To test if $_{MBP}Tpase$ displays an *attC* site and strand preference, DNA binding assays (EMSAs) were conducted. The fusion protein was purified either the day prior, or on the day of the EMSA to ensure a maximum yield of intact $_{MBP}Tpase$, as approximated by SDS-PAGE and spectrophotometry. The first assays were performed using total protein titrated (20 nM – 40 μ M) against a constant concentration of dsDNA labelled 5' with 32 -Phosphorous. Titrated MBP was used as a control (200 nM – 40 μ M). Double stranded DNA substrates of *Pseudomonas*-type *attC* and classical *attC_{aadB}* sites were tested first.

5.4.4.1 Interaction of $_{MBP}Tpase$ with dsDNA *attC* sites

Purified $_{MBP}Tpase$ was titrated from 20 nM-5 μ M to test its binding specificity to double stranded *Pseudomonas*-type *attC* sites (Figure 5.11A). All 32 -P labelled dsDNA was found in a single band migrating far down the native PAGE gel and no additional bands were observed, thus both strands were labelled with equal efficiency and/or were not denatured. When the purified $_{MBP}Tpase$ was added to the double stranded *Pseudomonas*-type *attC* fragment, no retardation was observed. The MBP control titrated at a range of concentrations between 0.02 to 2 μ M displayed no other bands but the labelled dsDNA.

Similarly, $_{MBP}Tpase$ binding specificity to double stranded classical *attC* sites was also tested. $_{MBP}Tpase$ was titrated at a higher concentration range, from 100 nM-40 μ M but despite this, all 32 -P labelled *attC_{aadB}* dsDNA was found in a single band migrating far down the native page gel and no additional bands were observed (Figure 5.11B).

The MBP control, however was found to bind the dsDNA. The MBP protein was titrated between 2 μ M to 40 μ M. The more protein added, the more obvious the apparent shift was, and the less unbound DNA target was visible on the native gel. For example, where 2 μ M MBP was added, a shift is observed as well as residual unbound target. At 40 μ M MBP, all of the *dsattC_{aadB}* appears to be bound. The ability of MBP to bind *dsattC_{aadB}* was reproducible (data not shown). Note these findings suggest that the presence of IS1111-*attC* transposase impairs the binding observed with the MBP control.

Based on these EMSAs, no mobility shift was observed in the presence of the transposase for either the *Pseudomonas*-type or the classical *attC_{aadB}* sites. A shift was observed when the

MBP control was added to *dsattC_{aadB}* but this shift is inferred to be non-specific. _{MBP}Tpase protein does not interact with the two tested double stranded DNA molecules. The MBP protein does show interaction at high concentrations with the double stranded *attC_{aadB}* site (*Pseudomonas*-type *attC* not tested at high concentrations greater than 5 μ M). The presence of the Tpase in the fusion appears to block the MBP interaction.

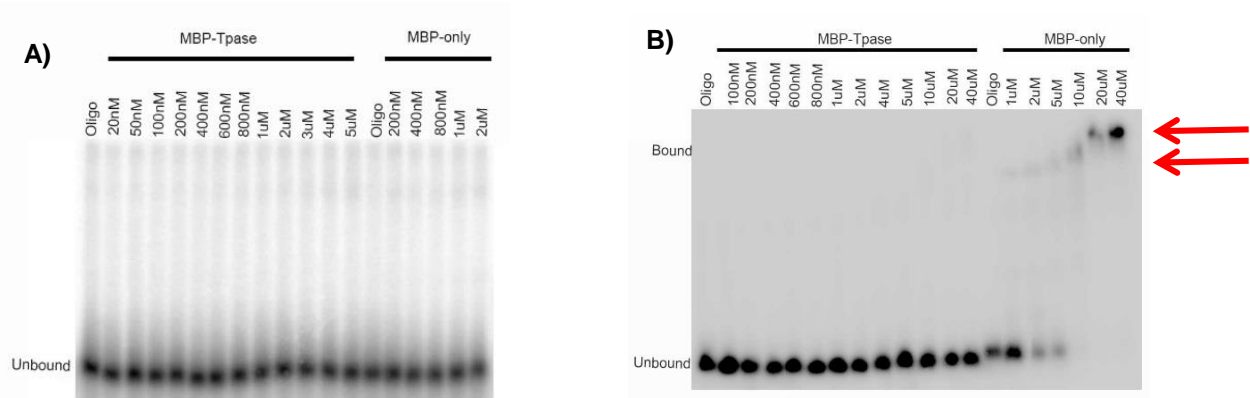


Figure 5.11: EMSA of _{MBP}Tpase titrated against MBP- only control.

A) *ds Pseudomonas*-type *attC* site. Constant concentration of 100 μ mol of labelled DNA was used across all lanes. B) *ds classical attC_{aadB}* site. Constant concentration of 100 μ mol of labelled DNA was used across all lanes. MBP control binds *ds attC_{aadB}* as indicated by the arrows. Control Oligo lanes have no added protein.

To determine the effect of secondary *attC*-structures and DNA binding specificity of the purified transposase, single-stranded DNA molecules of *Pseudomonas*-type *attC* and classical *attC_{aadB}* sites were tested next. These sequences contain imperfect inverted repeats with the capacity to form hairpin structures under the conditions of this assay.

5.4.4.2 Interaction of _{MBP}Tpase with ssDNA *attC* sites

In Figure 5.12A binding of _{MBP}Tpase to the *Pseudomonas*-type *attC* top strand occurs in concentration range 7.5 μ M to 20 μ M as evidenced by a shift in DNA mobility. In Figure 5.12B, binding of _{MBP}Tpase to the *Pseudomonas*-type *attC* bottom strand occurs in the 10 μ M and 20 μ M test lanes. However, in both instances, there was a lot of unbound DNA migrating as a single band. Furthermore, in both experiments, the MBP control was observed to bind both top and bottom strands of *Pseudomonas*-type *attC* sites non-specifically between tested concentrations of 5 μ M to 20 μ M. Thus it is difficult to exclude the possibility that the observed binding of the _{MBP}Tpase to labelled DNA target may also be non-specific.

For EMSA's using the classical *attC_{aadB}* site, the _{MBP}Tpase concentrations ranged from 1 μ M to 40 μ M. Where concentrations of _{MBP}Tpase ranged between 20-40 μ M, some protein and DNA target remained in the well. In Figure 5.12C, a shift was observed when _{MBP}Tpase, ranging between 2 μ M to 40 μ M, was bound to the classical *attC* top strand. However, note that the MBP control was binding far greater amounts of the labelled single strand as indicated by the pronounced shifting of labelled target. Similarly, for the bottom single strand in Figure 5.12D, a shift is observed for _{MBP}Tpase from 2 μ M to 40 μ M. As seen previously, this shift also occurs for the MBP control.

To summarise, the _{MBP}Tpase protein influences the mobility of single-stranded *attC* sites but not double-stranded. However, it is possible that this effect is due to the MBP moiety since the control MBP also influences mobility. To test for the specificity of transposase binding to single-stranded DNA, and exclude non-specific interaction with the MBP domain, competition experiments were carried out with an excess of cold (non-labelled) target DNA (i.e. *Pseudomonas*-type and classical *attC_{aadB}* site) and non-specific competitor, *EGG-m2F/R*.

5.4.5 EMSA competition assays

To test that the interactions in Figure 5.12 included *attC*-specific effects that were dependent on the transposase moiety of the fusion protein, cold competition assays were done. These competition assays compare the ability of cold (unlabelled) oligonucleotides that are specific or non-specific to interfere with the _{MBP}Tpase + *attC* interaction. If the interaction is specific, addition of unlabelled *attC* target (cold target) will displace the binding and no migrating band will be observed. If an unlabelled non-specific competitor is added, it will not displace the binding and thus the bound complex will be retained. In all competition assays the concentration of fusion protein _{MBP}Tpase and MBP-only control was standardized to 40 μ M.

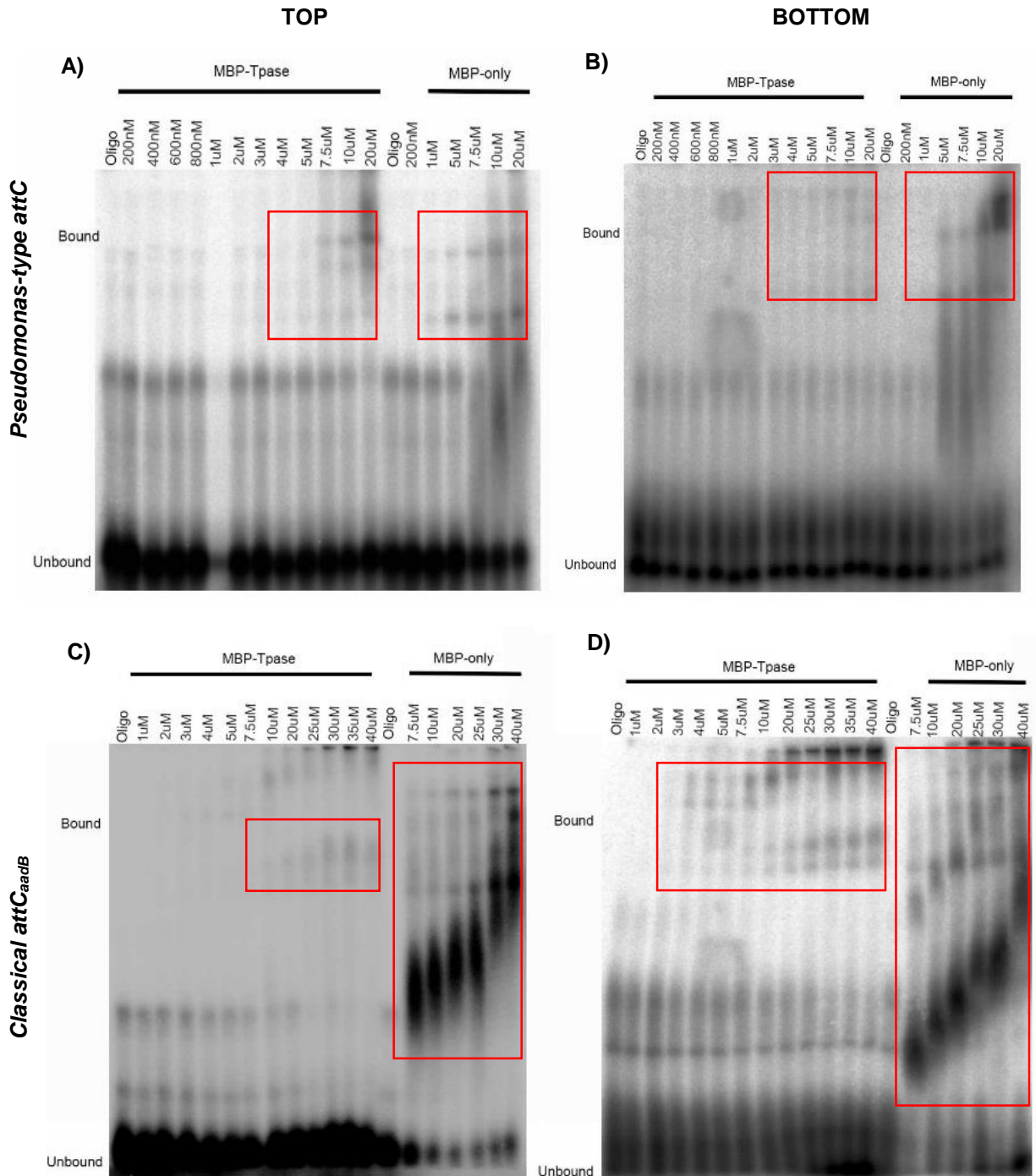


Figure 5.12: EMSA of $_{MBP}Tpase$ titrated against MBP control using ssDNA as binding substrates.

A) ssDNA *Ps. specific attC* top strand, B) ssDNA *Ps. specific attC* bottom strand, C) ssDNA *attCaadB* top strand, D) ssDNA *attCaadB* bottom strand. Red boxes note area of shift.

5.4.5.1 $_{MBP}Tpase$ binds reproducibly to the top strand of *Pseudomonas-type attC* sites.

In Figure 5.13A, lane 1 shows that the labelled target (*ss Ps. attC* top strand) runs as a single band. Lane 2 shows a band shift when $_{MBP}Tpase$ or MBP is added. Lanes 3, 4, 5, 6, 7 all show

loss of band shift, indicating competition between the cold target (unlabelled ss *Ps. attC* top strand) and the labelled target for protein-binding. Lane 8 includes an unlabelled oligonucleotide (*EGG-m2F*) that is unrelated to *attC* sites. Addition of non-specific competitor was reproducibly demonstrated to have minimal/no impact on $_{MBP}Tpase + ss Ps. attC$ top strand binding whilst interfering with $MBP + ss Ps. attC$ top strand (see insert). This represents the key experimental test and notably the non-specific competitor does not interfere with $_{MBP}Tpase$.

These findings suggest that the MBP interaction is non-specific but the $Tpase$ one is specific. Even at half the concentration of $_{MBP}Tpase$ (20 μM), binding to the top strand was recoverable when challenged with 100-fold non-specific competitor (see insert). The interaction between $_{MBP}Tpase$ and labelled target is blocked by cold target (10-fold excess) but not by non-target. In comparison, the interaction between MBP only is blocked equally by both cold competitors.

5.4.5.2 $_{MBP}Tpase$ does not bind to the bottom strand of *Pseudomonas*-type *attC* sites.

In contrast, binding between $_{MBP}Tpase$ and the bottom strand was outcompeted by non-specific competitor *EGG-m2R* (Figure 5.13B). Lane 1 shows the labelled test target (ss *Ps. attC* bottom strand) runs as a single band. Lane 2 shows the presence of the fusion protein causes a band shift. As seen previously, $_{MBP}Tpase$ shifts the labelled target with greater efficiency than in MBP-only control. Lanes 3, 4, 5, 6, 7 all show no band shift indicating competition between the cold target (unlabelled *Ps. attC* bottom strand) and the labelled target for protein-binding. This is comparable in both panels. Addition of 100-fold excess non-specific competitor, *EGG-m2R*, in lane 8 outcompetes the labelled target for binding for $_{MBP}Tpase$ fusion and MBP. The interaction between $_{MBP}Tpase$ and labelled target, and MBP and labelled target is blocked by cold target (10-fold excess) as well as non-specific competitor. I conclude that the $_{MBP}Tpase$ fusion is interacting specifically with *Pseudomonas*-type *attC* and this interaction is dependent on the $Tpase$ domain. Furthermore, this interaction also displays a strand preference where binding of $_{MBP}Tpase$ to single stranded top strand was recovered

	Hot Target							
	Cold Target							Cold NS _C
Lanes	1	2	3	4	5	6	7	8
NS _C	-	-	-	-	-	-	-	100
O _H	+	+	+	+	+	+	+	+
P	-	+	+	+	+	+	+	+
O _C	-	-	10	20	30	40	50	-

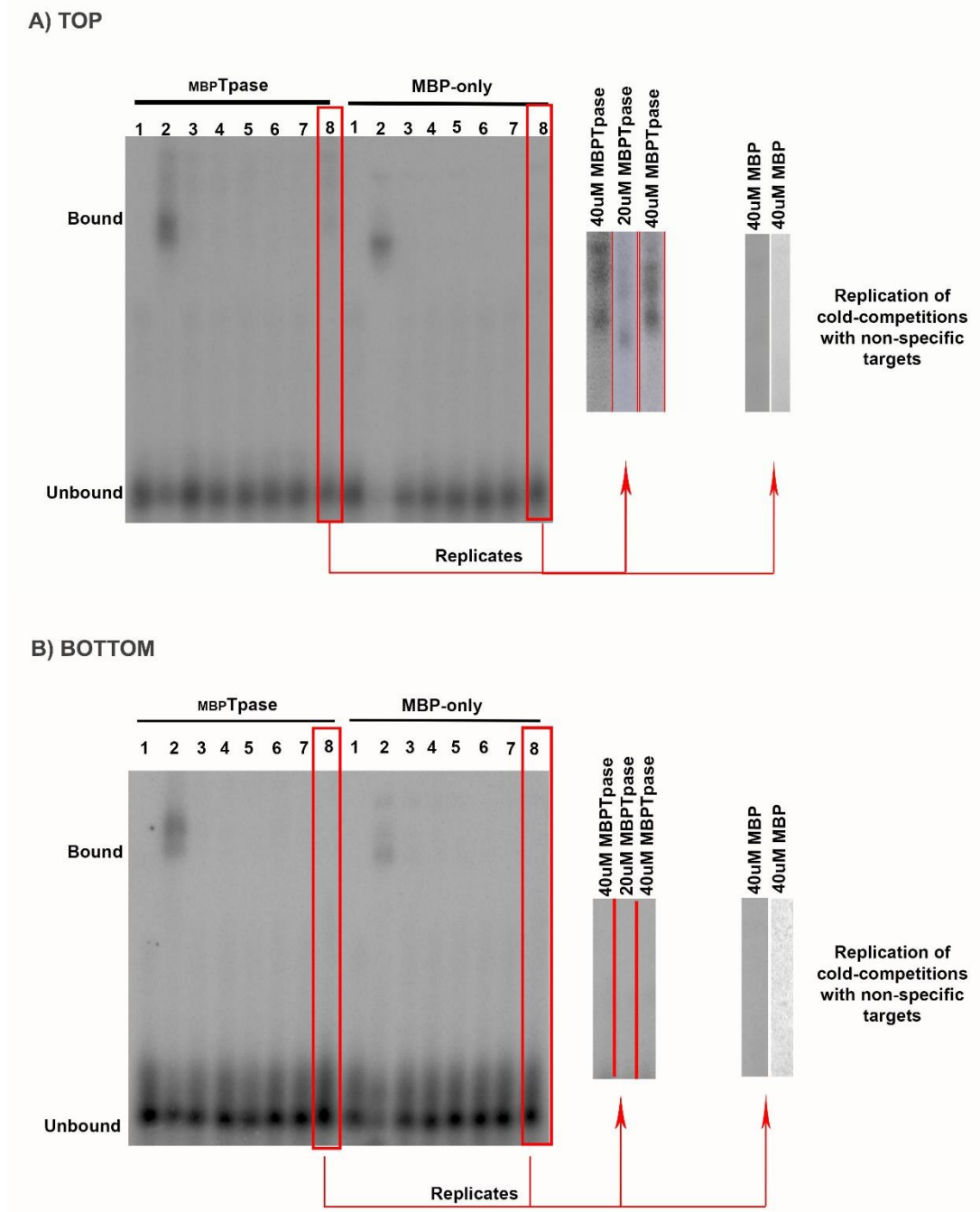


Figure 5.13: Cold competition assay of single stranded A) top and B) bottom *Pseudomonas*-type attC site. MBPase binding was demonstrated reproducibly to the top strand and not the bottom. This binding at half concentration of MBPase (20 μ M). For binding efficiency in each gel compare test lanes MBPase to control MBP lanes.

Key: NS_C= Non-specific competitor, O_H= Oligo Hot (labelled), P= Protein, O_C=Oligo Cold target (10-50x). Note that binding data for 20 μ M MBP is unavailable.

5.4.5.3 *MBP*Tpase binding to single stranded classical *attC* site is inconclusive

In Fig 5.14 a similar gel shift is seen for the *attC_{aadB}* test with both *MBP*Tpase and MBP only. However in this case I was unable to show a differential response between target and non-specific competitors. This could reflect either reduced sensitivity of detection or lower affinity of the Tpase to the target site.

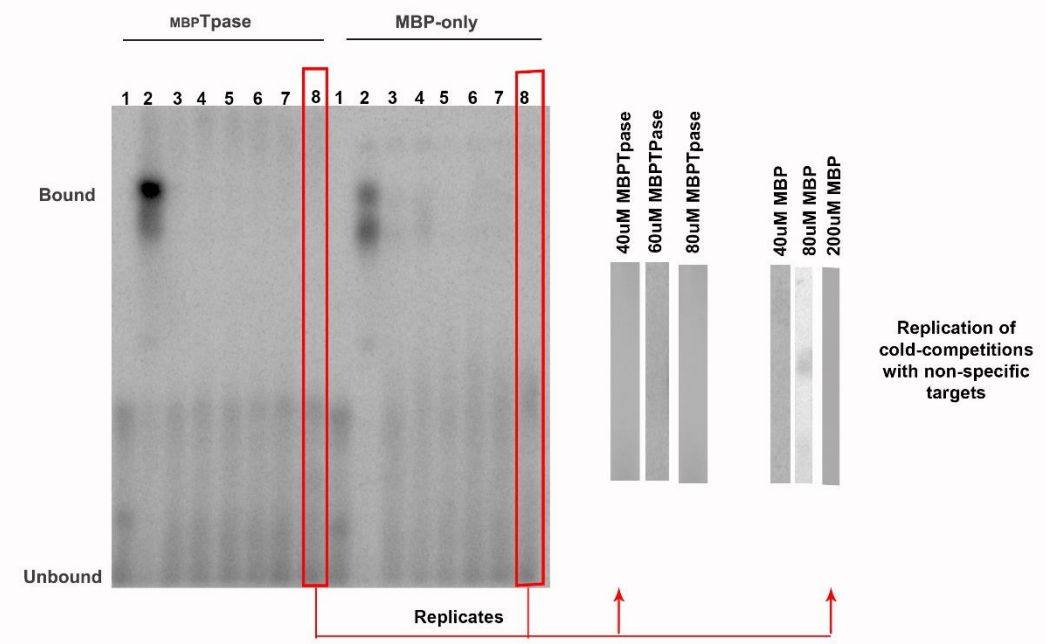
In Figure 5.14A, lane 1, the single stranded labelled *attC_{aadBts}* test target can be seen migrating as two weakly labelled bands, rather than one strongly labelled band as was seen for the *Ps. attC* test. This banding pattern implies that two distinct forms of the labelled target are present. These could be variations in length (degradation or poor synthesis) or differing conformations. If only one of these forms were a target for the Tpase, then this would effectively reduce the sensitivity of detection for binding. Lane 2, shows that the presence of the fusion protein causes a band shift and that the intensity of the shifted band is greater than in MBP-only panel. In both cold competition assays, 10-fold excess of cold target was enough to displace the binding to the labelled *attC_{aadB}* top and bottom strands (lanes 3-7). In Figure 5.14A at 20-fold excess cold competitor, MBP control binding to *attC_{aadB}* the top strand was displaced (lane 4). In comparison 30-fold addition of the cold competitor displaced MBP binding to the bottom strand (Figure 5.14B). However, a 100-fold excess addition of the non-specific competitor was enough to displace the binding of *MBP*Tpase or MBP to both either *attC_{aadB}* strands (lane 8).

The fusion protein *MBP*Tpase caused a gel shift with both top and bottom strand. The interaction was stronger for the top strand, which is consistent with preferential binding, but I was unable to demonstrate specificity of the interaction. I cannot exclude the possibility that both types of *attC* site are targeted by the Tpase.

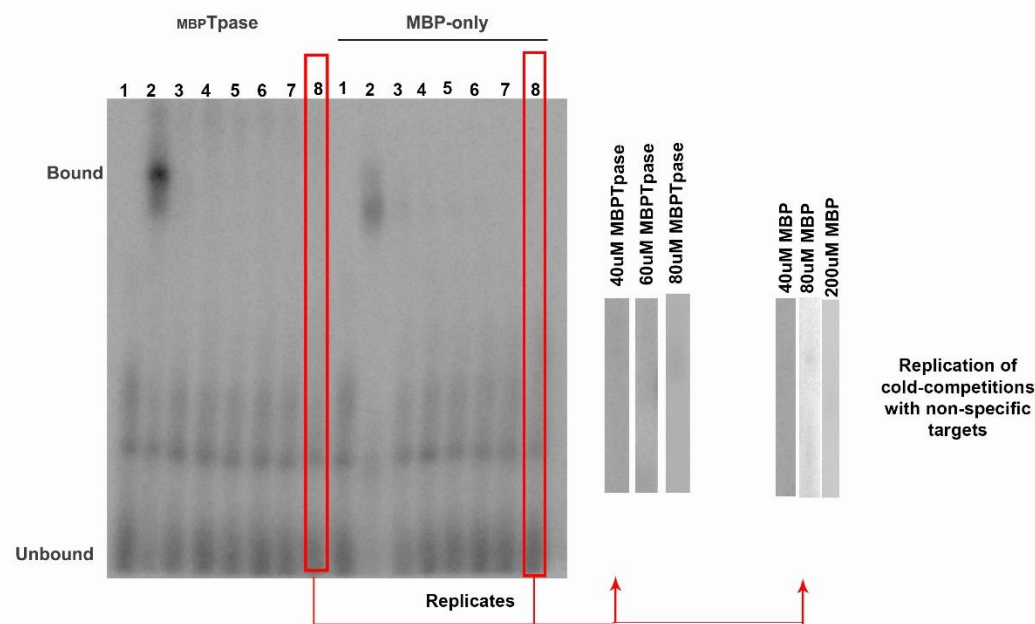
*Figure 5.14: Cold competition assay of single stranded A) top and B) bottom classical attC_{aadB} site. MBP*Tpase binding to either strand was not detected upon the addition of non-specific competitor. Increasing concentrations (60 μ M and 80 μ M) of *MBP*Tpase was also outcompeted for strand binding by the non-specific competitor Key: NS_C= Non-specific competitor, O_H= Oligo Hot (labelled), P= Protein, O_C=Oligo Cold target (10-50x). Across all cold competition assays 40 μ M *MBP*Tpase and 40 μ M MBP were tested.

	Hot Target							
	Cold Target							Cold NSc
Lanes	1	2	3	4	5	6	7	8
NSc	-	-	-	-	-	-	-	100
O _H	+	+	+	+	+	+	+	+
P	-	+	+	+	+	+	+	+
O _c	-	-	10	20	30	40	50	-

A) TOP



B) BOTTOM



5.4.6 Testing the IS1111-attC minicircle junction as a binding partner for T_pase

For successful recombination of an IS1111-attC element and an attC site to occur, a second partner site in the IS element is obviously required (Figure 5.1). Various lines of evidence (Partridge & Hall, 2003; Tetu & Holmes, 2008) and work presented in chapter 4, implicate a circular intermediate as the reaction partner for recombination. Here I tested the sequence formed by IS excision as a target for the MBP_Tpase (Figure 5.15).

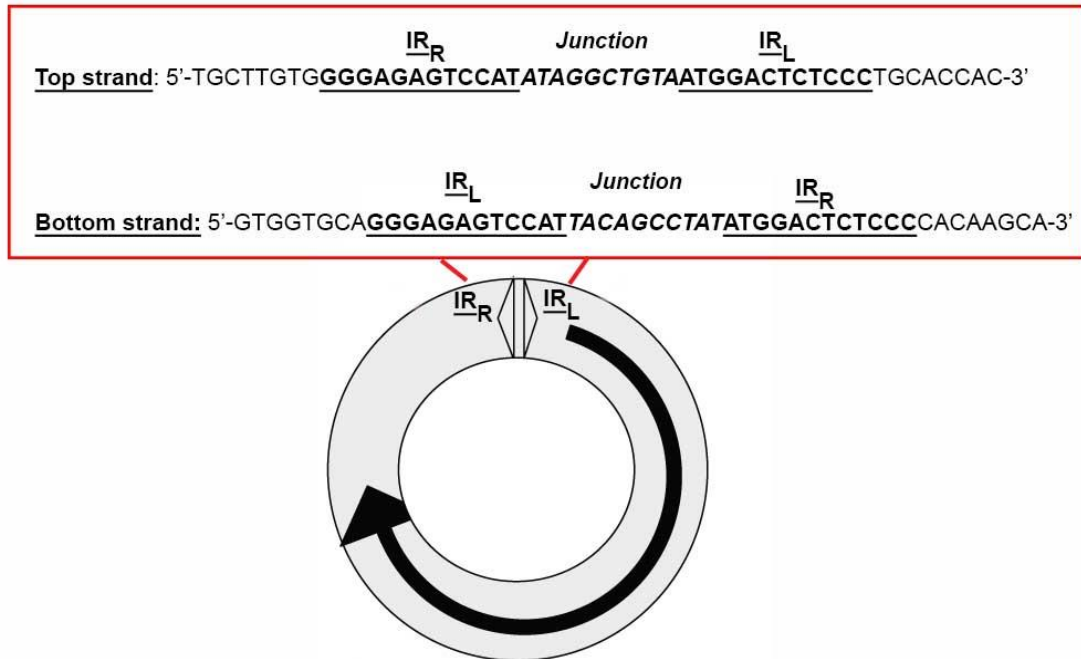


Figure 5.15: Schematic representation of the IS1111-attC minicircle and oligonucleotides used for MBP_Tpase DNA binding assays.

Sequences encompassing the junction region of the top and bottom strands are shown in the red box. The inverted repeat regions are underlined and the junction formed upon circularisation is italicized.

Binding assays in which the MBP_Tpase was titrated across a range of concentrations were not conducted for the IS Junction DNA. Previous work mentioned above, demonstrated the non-specific binding of the MBP tag towards the DNA substrates. Consequently, only cold competition assays were carried out to test for interactions between the IS junction and transposase.

5.4.6.1 *MBP*Tpase binding to ds IS junction cannot be outcompeted using cold target or a non-specific competitor.

dsIS cold target outcompeted *MBP*Tpase binding when tested at 40 μ M and 100 μ M *MBP*Tpase (Figure 5.16). In Figure 5.16A, lane 1 shows the labelled ds IS junction running as a single band. Lane 2 shows a band shift and that the shift intensity appears greater than in the *MBP*-only panel. Note the presence of a fast migrating (low retardation) band in lane 2. This low retardation band is not present in the *MBP*-only experiment. Both fast and slow migrating bands were eliminated upon the addition of cold targets (lanes 3-7). Similarly, a 100-fold excess of ds stranded non-specific competitor displaced binding of *MBP*Tpase, but it did not eliminate the fast migrating (low retardation) band. This pattern is consistent with specific binding of *MBP*Tpase, but the mobility is much faster than expected for a fusion protein of its size (~81 kDa). However a Tpase fragment that retains binding could potentially explain this (Figure 5.8). A labelled protein lane (with no oligo) was not able to be run here but would help resolve this. In comparison, *MBP*-only binding was consistently observed across all test lanes. The addition of cold competitors (specific and non-specific) made little difference to the ability of *MBP* to bind the hot DNA target, hence the observed smearing in lanes 2-8 in the *MBP* test set.

In order to increase the sensitivity of EMSA detection and determine if the fast migrating band is reproducible, the same experiment was repeated using 100 μ M *MBP*Tpase (Figure 5.16B). As seen previously, band retardation for the fusion protein was at a constant size, implying that a stoichiometric oligo-protein complex was formed. The fast migrating (low retardation) band was observed when binding was challenged with 100-fold non-specific competitor, suggesting that this binding represents a metastable association between the fusion and labelled double stranded IS junction DNA. *MBP*-only binding to target DNA was as seen previously. However, note that at 100 μ M *MBP*, addition of cold competitors did not displace non-specific binding (lanes 3-8, control panel).

	Hot Target							
	Cold Target							Cold NSc
Lanes	1	2	3	4	5	6	7	8
NSc	-	-	-	-	-	-	-	100
O _H	+	+	+	+	+	+	+	+
P	-	+	+	+	+	+	+	+
O _C	-	-	10	20	30	40	50	-

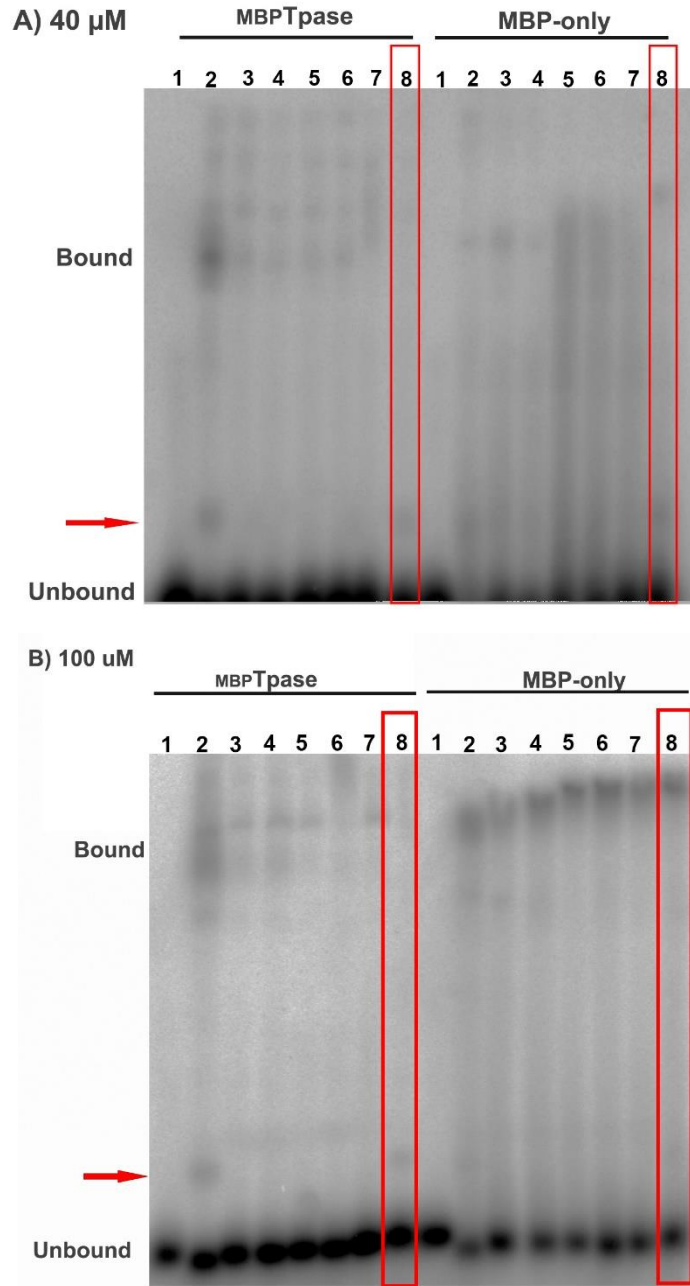


Figure 5.16: Cold competition assay of double stranded IS minicircle junction DNA using A) 40 μM and B) 100 μM *MBPase* and MBP control.

Note the fast migrating band (red arrow) was only observed in lanes 2 and 8 of *MBPase* binding to dsIS junction DNA. Key: NSc= Non-specific competitor, O_H= Oligo Hot (labelled), P= Protein, O_C=Oligo Cold target (10-50x). In B) increasing concentrations of cold target for MBP only, results in slower migration (lanes 2-8, hence the apparent increase in size).

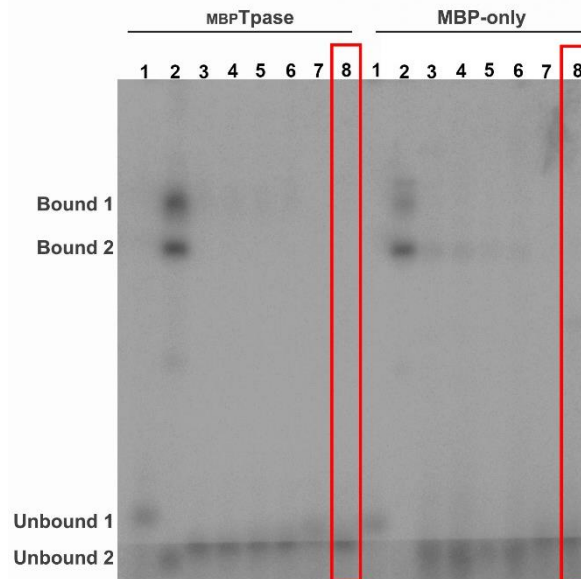
5.4.6.2 *MBP*Tpase does not bind IS Junction DNA when presented in a single stranded conformation

The *MBP*Tpase binding assay with single stranded IS junction template revealed no evidence for specific binding to either the top or bottom strand of the DNA region formed at the IS minicircle junction. In Figure 5.17A, lane 1 in both test sets shows labelled top single strand oligonucleotide running as a single band (Unbound 1), but also running slower in comparison to labelled oligo in all other test lanes (Unbound 2). This may be due to uneven polymerization at the bottom of the gel, hence also the visible horizontal line running across the bottom of the gel, and/or salt concentration differences between the binding reactions and the oligonucleotide stock. Lane 2 in *MBP*Tpase set, shows that the presence of the fusion protein causes two band shifts (Bound 1 & Bound 2) and an unbound oligonucleotide band at a faster migration state. In lane 2 of the MBP test, all of the hot oligonucleotide has been bound (unbound bands absent) by MBP resulting in two bound complexes. The addition of cold target or single stranded non-specific competitor resulted in displacement of hot oligonucleotide from both bound complexes in the *MBP*Tpase test set (lanes 3-8). However, 10-40 fold excess of cold targets failed to displace bound 2 complex binding for MBP-only (lanes 3-6). 50-fold cold targets and 100-fold addition of single stranded non-specific competitor *EGG-m2F/R* displaced binding of MBP bound complex 2 (lane 8).

In Figures 5.17B, lane 1 shows labelled bottom single strand oligonucleotide running as a single band (Unbound 1), but running marginally slower in comparison to labelled oligo in all other test lanes (Unbound 2). Lane 2 shows binding of *MBP*Tpase protein to the bottom strand also causes a two band shift (Bound 1 & Bound 2) and an unbound oligonucleotide band at a faster migration state (Unbound 2). In lane 2 of MBP-only test, all of the hot oligonucleotide has been bound (unbound bands absent). Addition of cold target and non-specific competitor displaced bound 1 & 2 complexes for *MBP*Tpase (lanes 3-8). In comparison, addition of cold target to MBP only displaced bound 1 complex (lanes 3-7) but not bound 2. Bound 2 complex was only displaced at 100-fold addition of single stranded non-specific competitor *EGG-m2F/R* (lane 8).

	Hot Target							
			Cold Target					Cold NSc
Lanes	1	2	3	4	5	6	7	8
NSc	-	-	-	-	-	-	-	100
O _H	+	+	+	+	+	+	+	+
P	-	+	+	+	+	+	+	+
O _C	-	-	10	20	30	40	50	-

A) TOP



B) BOTTOM

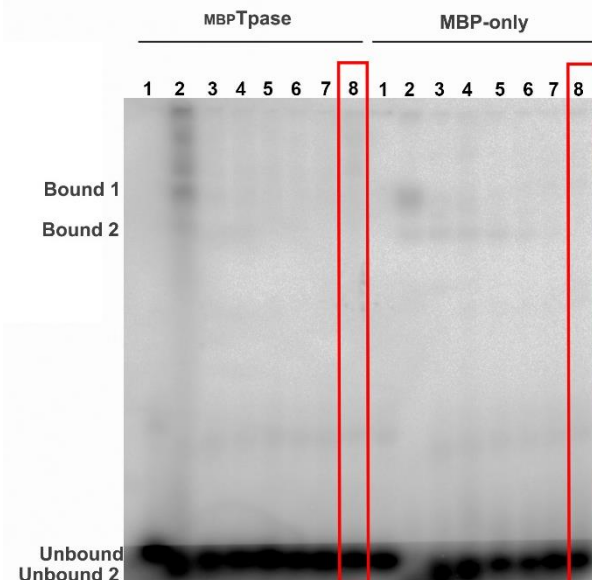


Figure 5.17: Cold competition assay of single stranded IS minicircle junction DNA using 40µM of MBPase and MBP-only control.

Key: NSc= Non-specific competitor, O_H= Oligo Hot (labelled), P= Protein, O_C=Oligo Cold target (10-50x). Across all cold competition assays 40 µM MBPase and 40 µM MBP were tested.

5.5 Discussion

In chapter 4, I demonstrated the potential of IS1111-attC elements to translocate from a native attC site in a *Pseudomonas* chromosomal array to an artificially introduced site on a plasmid. Transposases of the IS110/IS492 family possess the conserved D-E-D-D motif (D₇-E₅₁-D₈₈-D₉₁) that has been demonstrated to be essential for activity of the Piv/MooV family of DNA recombinases (Buchner *et al.*, 2005; Mahillon & Chandler, 1998; Tobiason *et al.*, 2001). The Piv/MooV recombinase does not share the conserved amino acid motifs of site-specific tyrosine (Y) or serine (S) recombinase family. Tpsases in general are predicted to be poorly expressed. The low-level expression of Tpsase may be sufficient for translocation but not for ease of protein isolation. Thus Tpsase purification from the *Pseudomonas* wild-type strain is predicted to be an ineffective method. Instead an MBP-expression system was developed and used successfully to purify the *Ps. stutzeri* IS1111-attC transposase as a maltose binding fusion protein (MBPTpsase).

To determine the DNA binding specificity of purified IS1111-attC transposase, its behaviour was explored via electrophoretic mobility shift assays (EMSAs) with nine different DNA targets. Binding between MBPTpsase and single stranded attC sites was observed to differ in regards to an MBP control. This is also the first study where transposase was shown to bind preferentially to the top strand over the bottom strand of *Pseudomonas*-type attC sites. Also, transposase binding to the double stranded IS minicircle junction was observed but not to the single stranded forms.

5.5.1 IS1111-attC transposase does not bind double stranded attC sites at a detectable level.

Francia *et al.* (1999) first demonstrated the binding of the integron integrase to attI and attC sites. They found that purified IntI1 was able to bind the double stranded DNA fragment containing an attI, but was unable to bind double stranded attC_{aadA1}. In this present study, no retardation was observed when purified MBPTpsase was added to double stranded *Pseudomonas*-type and attC_{aadB} sites. Thus as seen with IntI1, double stranded attC sites are not a binding partner for IS1111-attC transposase.

5.5.2 *IS1111-attC* transposase binds strand specifically to the top strand *Pseudomonas attC* sites

By means of competition assays, Francia *et al.* (1999) also demonstrated preferential binding of IntI1 to single stranded *attC* sites. In my study, competition assays were successfully used to demonstrate the specificity of transposase binding to single stranded top strand of *Pseudomonas*-type *attC*. Relative mobility, calculated as the ratio of the distances travelled on EMSAs by the bound and unbound oligonucleotides, was calculated for recovered binding and compared to that of the unchallenged $_{MBP}Tpase$. Thus the observed shift was a fixed shift, where the relative mobility was the same regardless of the dose of competing agent (Table 5.5). Note in calculating relative mobility, the midpoint of the diffuse band was used.

Binding was reproducibly observed and was also seen at half the $_{MBP}Tpase$ protein concentration. The recovered binding would always appear as a diffuse rather than a defined band. I interpret this diffuse band to represent specific interactions between the *Tpase* moiety of the fusion protein and the oligonucleotide. In cases where the target template, such as RNA (Bendak *et al.*, 2012) or *attC* sites (Bikard *et al.*, 2010), form hair-pin structures, diffuse EMSA bands are likely to occur. Diffuse bands may also result where the protein conformation is not stable, or the stoichiometry is variable (Hellman & Fried, 2007b). In comparison, defined bands of consistent relative mobility are expected where a binding of 1:1 stoichiometry with a stable conformation exists. Binding of the transposase to the bottom strand of *Pseudomonas*-type *attC* was not detected by competition assays. When the $_{MBP}$ -*Pseudomonas*-type *attC* interaction was challenged with the cold competitors, the amount of alteration to mobility varied with the competitor dose. This interaction is thus believed to be dose-dependent (Table 5.5).

In summary, the findings presented here are most interesting as the integron integrase preferentially binds to the bottom strand of *attC* sites (Francia *et al.*, 1999) whereas in comparison, *Tpase* does not, it binds to the top strand.

Table 5.5: Summary of EMSA experiments

DNA Target	<i>Pseudomonas-type attC</i>			Classical <i>attC_{aadB}</i>			IS minicircle junction		
Strand conformation	<i>ds</i>	<i>SS_{Top}</i>	<i>SS_{Bottom}</i>	<i>ds</i>	<i>SS_{Top}</i>	<i>SS_{Bottom}</i>	<i>ds</i>	<i>SS_{Top}</i>	<i>SS_{Bottom}</i>
MBP Tpase	Nothing	Fixed shift	Dose dependent shift	Nothing	Fixed shift	Dose-dependent shift	Fixed shift + oligo shift	Dose-dependent shift	Dose-dependent shift
		Specific Out-competed	Specific Out-competed		Partial-competition	Out-competed	Out-competed shift but no Oligo shift	Out-competed	Out-competed
		Non-specific NOT competed	Non-specific Out-competed						
MBP	Nothing	Dose dependent shift	Dose dependent shift	Dose dependent shift at high concentrations	Dose dependent shift	Weakly Dose dependent shift	Dose dependent shift at high concentrations + no oligo shift	Dose dependent shift	Dose dependent shift
		Out-competed	Out-competed		Out-competed	Out-competed	Not-outcompeted	Out-competed	Out-competed
Figure	5.11A	5.13A	5.13B	5.11B	5.14A	5.14B	5.16	5.17A	5.17B

Grey boxes highlight differences and observations between *MBP*Tpase and *MBP* control. Orange box highlights key finding.

Key definitions:

Fixed shift = relative migration remains the same regardless of dose (Relative mobility (*R_m*) constant), where Relative mobility was calculated as the distance migrated by bound hot oligonucleotide band divided by the distance migrated by the unbound hot oligonucleotide.

Dose dependent shift = amount of alteration to migration varies with dose (Relative mobility (*R_m*) varies), where Relative mobility was calculated as the distance migrated by bound hot oligonucleotide band divided by the distance migrated by the unbound hot oligonucleotide.

Out-competed = a heterologous molecule blocks binding to hot target

Partial-competition = a heterologous molecule partially blocks binding to hot target, weak interaction between protein and target

Not-competed = a heterologous molecule does not compete and block binding to hot target

5.5.3 *Experimental evidence for IS1111-attC transposase interactions with classical attC sites is inconclusive*

The occurrence of IS1111-*attC* elements is increasingly being noted in clinically relevant contexts (Martinez *et al.*, 2012; Post & Hall, 2009; Santos *et al.*, 2010; Tetu & Holmes, 2008; Chapter 3). Here I explored the ability of T_pase to bind to different *attC_{aadB}* structural targets.

Initial binding assays using double stranded *attC_{aadB}* sites showed that, like the integrase, T_pase does not bind. However, the MBP control was found to bind non-specifically, and such binding was dependent on dose at high concentrations of protein. MBP also bound single stranded *attC_{aadB}* conformations rather strongly and also in a dose-dependent manner. I hypothesised that this is primarily due to MBP containing ~11% positively charged amino acids, Arginine and Lysine (42/387) which have the capacity to bind rather strongly to negatively charged DNA. When challenged with a cold, non-specific competitor, the binding was dissociated, adding more evidence to MBP binding acting non-specifically towards the tested DNA targets.

EMSAs were used to test the binding specificity of T_pase to single-stranded *attC_{aadB}* sites. Cold targets displaced the observed binding of T_pase to *attC_{aadB}* sites as expected, however the cold, non-specific competitor unexpectedly displaced binding too. This is surprising given that in nature *attC_{aadB}* sites are found to be occupied by IS1111-*attC* members. In principal, binding of transposase to *attC_{aadB}* sites should occur if it interacts with the target nucleic acid with greater affinity than its competitor and secondary binding does not discriminate between the sequences (Alves & Cunha, 2012). The absence of a gel shift in the presence of the non-specific competitor may reflect the low detection sensitivity of EMSAs, combined with a lower affinity for *attC_{aadB}* and/or the formation of a stable secondary structure by *attC_{aadB}*.

5.5.4 Factors impacting on EMSA sensitivity

attC_{aadB} sites are known to be a target for IS1111-attCs (Tetu & Holmes, 2008), so why was no convincing binding observed in my study?

a) Lower labelling efficiency and/or oligonucleotide concentration: EMSA analysis revealed two different conformations of single stranded *attC_{aadB}* sites for both the top and bottom strands. Oligonucleotide labelling was found to be at half-intensity when compared to *Ps-type attC* labelling. Binding to *single-stranded attC_{aadB}* was therefore expected to have reduced intensity, with less oligonucleotide being bound. Specific binding between IS1111-attC transposase and *attC_{aadB}* site may still be present but was below the detection limits of an EMSA when challenged with a cold competitor.

b) Greater instability of T_pase-attC_{aadB} bound complex results in a diffuse band and/or an absent band: DNA-protein samples which are not at chemical equilibrium during the electrophoresis step are more likely to dissociate (Hellman & Fried, 2007a). Complete dissociation would result in no detection whereas partial, or slow dissociation, can result in the underestimation of the binding density (Hellman & Fried, 2007b). In chapter 4, I presented circumstantial evidence for a lower level of interaction of IS1111-attC with *attC_{aadB}* traps in my transformation experiments. Dissociation of DNA-protein complexes may also be playing a role in *in vivo* interactions. To prevent dissociation small neutral additives such as glycerol or sucrose can be added *in vitro* to stabilize labile proteins and thereby enhance protein-DNA stability (Garner & Rau, 1995; Vossen *et al.*, 1997).

c) Multiple single stranded attC_{aadB} conformers with differing band positions and band intensities were observed: It is not uncommon for secondary structure misfolding to occur (Sinan *et al.*, 2011). Such structures may exist in solution as a mixture of differentially base-paired states of equilibrium (Gell *et al.*, 2008). Once misfolding of *attC_{aadB}* occurs, the site may no longer be available for T_pase binding, compromising the formation and detection of a metastable T_pase- *attC_{aadB}* complex. For misfolded *attC_{aadB}* to become available to IS1111-attC T_pase, *attC_{aadB}* may need to unfold so as to allow T_pase to attach. Heat and/or chemical application (Sinan *et al.*, 2011), as well as DNA-binding proteins can be used to refold the single-stranded hairpin, thereby making it more favourable for transposase recognition and binding.

5.5.5 Secondary binding partner for IS1111-attC translocation: the IS1111-attC minicircle junction

Recombination reactions performed by integron integrase, revealed that it mediates recombination between non-canonical substrates (Loot *et al.*, 2012). These non-canonical substrates involve a palindromic single-stranded *attC* site and a double stranded *attI* site. I postulated that the IS1111-attC element itself is the second binding substrate for recombination with a single stranded top strand of the *Pseudomonas*-type *attC*. The circularization of an IS1111-attC generates a region of 10 bases flanked by inverted repeats. Such regional characteristics are typical of targets for tyrosine recombinase (Grindley *et al.*, 2006). I hypothesised that this region is essential for recombination. IS minicircles are universally believed to be an intermediate phase/structure of the IS prior to linearization and integration of the IS into its site-specific target (Partridge & Hall, 2003).

Evidence suggests that ISs undergo a structural change from a double stranded conformation to a single stranded form (Lewis & Grindley, 1997; Lewis *et al.*, 2004; Lewis *et al.*, 2012). I have demonstrated that IS1111-attC T_pase does not bind single stranded IS junction DNA when challenged with a cold competitor. This raised the possibility that a double stranded IS junction is required during recombination. Cold target competitors were unable to displace all slow migrating T_pase-double stranded IS junction binding. Whilst the non-specific competitor was successful in displacing slow migrating binding, it was unable to displace the fast migrating T_pase-double stranded IS junction complex. Given the preservation of the fast migrating band suggests that there is weak binding of the transposase protein to double stranded IS junction DNA, and that this DNA is a template for T_pase mediated recombination with single stranded *attC* sites. This band results from the formation of a synaptic complex between T_pase and the double stranded target, and possibly represents the formation of a covalently closed linear molecule as seen with IntI1 binding (Bouvier *et al.*, 2009; Johansson *et al.*, 2004).

The translocation/activity assays and EMSA findings highlight similarities between T_pase-*attC* and the IntI-*attC* system. ISP_{st6} transposase possess the four acidic residues D-E-D-D (Tetu & Holmes, 2008) that are essential for Piv catalysis of inversion and RuvC-like Holliday junction resolvases (Buchner *et al.*, 2005). Based on what is known about Piv, RuvC and integron-mediated recombination models, I propose a model for IS1111-attC T_pase mediated recombination with *attC* sites (Figure 5.18), where I predict that IS1111-attC

recombination involves a Holliday junction intermediate, although this remains to be validated experimentally. During IntI mediated recombination, two IntI moieties bind, with two contact points at *attI* and two at *attC*, and initiate classical site-specific recombination steps as catalysed by other tyrosine recombinases (Collis *et al.*, 2002; Loot *et al.*, 2012; MacDonald *et al.*, 2006). I predict that two T_{pase} moieties are also needed at each *attC* and IS junction site (Figure 5.18i). Initial cleavage and strand transfer would most likely require T_{pase} binding to both the double stranded IS-minicircle as well as single-strand *attC_{ts}*. Subsequent steps are predicted to involve single-strand cleavage of *attC_{ts}* and IS junction, thereby resulting in a 5'-OH radical group (Figure 5.18ii). This reactive hydroxyl group would be involved in strand transfer via nucleophilic attack of the 3'-Phosphate group (Figure 5.18iii). The proposed model involves the formation of a Holliday junction (Figure 5.18iv). A second round of strand cleavage and exchange is predicted to result in IS1111-attC excision out of the *attC* site. Holliday junctions mediated by IntI are resolved by a replicative step, thus ensuring complete integration of the gene cassette into the integron platform (Cambray *et al.*, 2010; Loot *et al.*, 2012; MacDonald *et al.*, 2006; Mazel, 2006). Therefore, replicative resolution of the IS1111-attC Holliday junction is similarly predicted to occur.

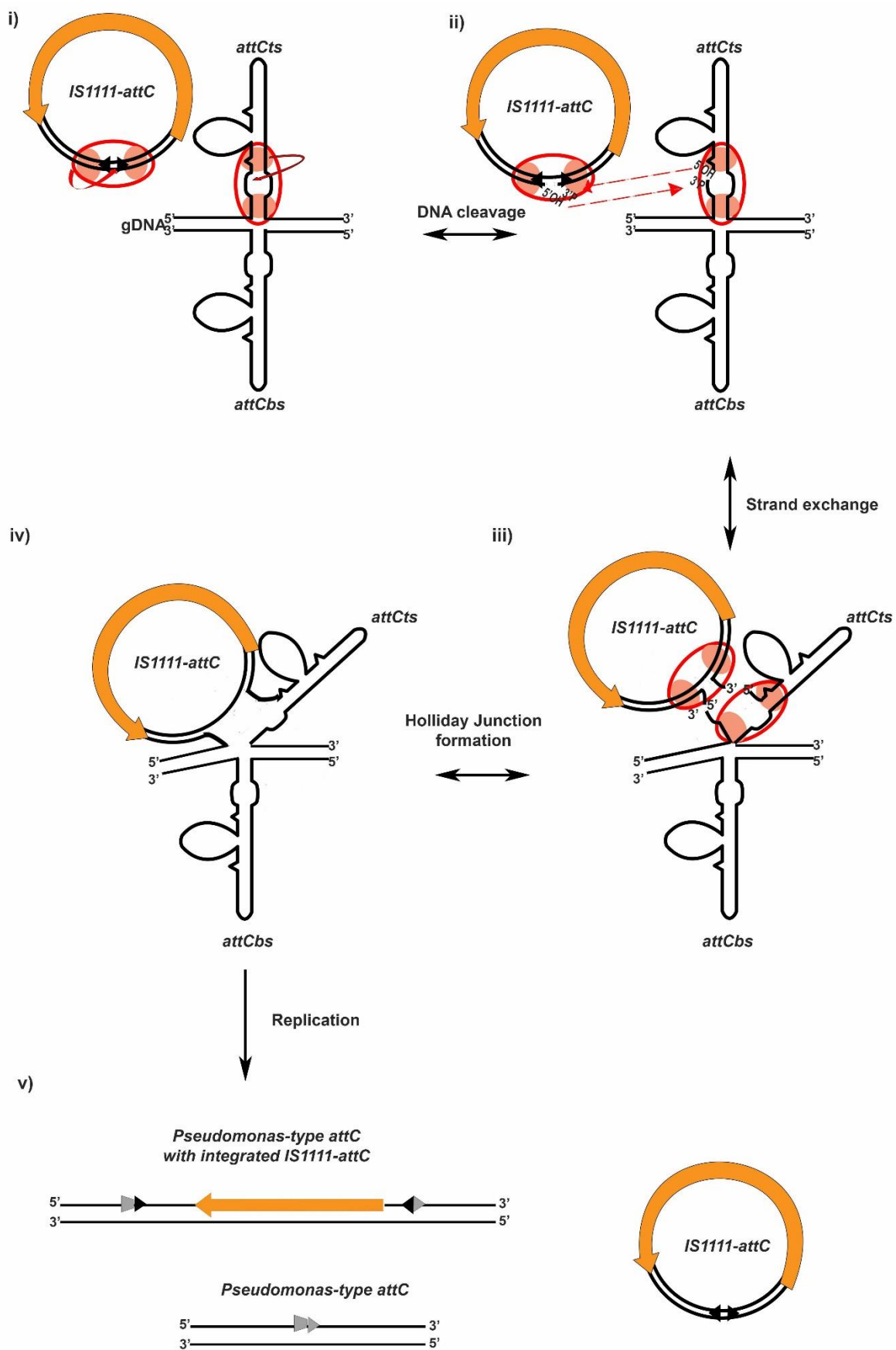
The extrahelical G20" in VCRs was demonstrated to play a critical role in IntI binding to *attC* bottom strand (Bouvier *et al.*, 2009). It ensures proper geometrical assembly of VChIntIA to *Vibrio attC* through insertion into the hydrophobic pocket of VChIntIA, thus resulting in an interaction with the integrase protein (MacDonald *et al.*, 2006). Given that IS1111-attC elements insert site-specifically into the same location within *attC* sites (between positions *h* and *i* in the palindrome) (Tetu & Holmes, 2008; Chapter 4), I predict that this event is dependent on the extrahelical bases of the *attC_{ts}*. I postulate that the conserved extrahelical bases (C and A) in the top strand play a role in docking the *attC* site into the transposase. It is plausible that T_{pase}-*attC* binding to the *Pseudomonas*-type *attC* top strand variable region is mainly due to non-specific interactions. To determine the types of T_{pase}-*attC* interactions that occur, DNA target mutagenesis and transposase binding to these modified sites would have to be analysed.

The mechanistic details and host factors involved in IS1111-attC translocation are at present unknown. Accessory proteins are postulated to play a vital role in IS1111-attC translocation success and replication fork resolution. In *E. coli* and *Ps. stutzeri* strain Q and ATCC 17641, partial IS translocation was observed. Partial translocation is a recombination that has not

proceeded to complete resolution of the partners. It is seen by formation of at least one covalent join between strand partners, but not formation of a completely resolved co-integrate. Partial translocation is experimentally observed when; i) a PCR recovery of only ISPst6 IR_R junction in a plasmid trap is observed but not the IR_L junction in *E.coli* (Tetu & Holmes, 2008), ii) no attC-trap plasmid, consistent with containing a single intact copy of ISPst6 in *E.coli*, is observed (Tetu & Holmes, 2008), or iii) ISPst6 minicircles are observed, but no subsequent translocation into PCI arrays in *PstQ* and *Pst641* (Chapter 4). Partial translocation of ISPst6 in these strains suggests that they lack enzymes involved in replication fork resolution, thereby resulting in replication arrest and partial/no IS1111-attC translocation (Loot *et al.*, 2012). Despite variations between the different mechanisms of site-specific recombination, accessory proteins are known to assist in the assembly of the initial complex and to drive the overall recombination reaction in the desired direction (Grindley *et al.*, 2006).

Figure 5.18: A proposed model for IS1111-attC transposase mediated recombination with ssattC sites and double stranded IS junctions. This model is based on the working model for IntII-mediated site-specific recombination by Loot et al., 2012.

- i) IS1111-attC transposase dimers (red ovals) bind top attC strand and the IS1111-attC minicircle molecule. Each Tpase has two contact points per DNA target (pink ovals).*
- ii) Transposase mediated-DNA hydrolysis of one strand at both sites leads to the generation of reactive 5'-hydroxyl groups and strand transfer between attC top strand and IS1111-attC.*
- iii) Formation of a Holliday junction structure.*
- iv) Host DNA ligase activity repairs any nicks.*
- v) Subsequent replication and recombination of IS into additional site may occur.*



5.5.6 Further Work

Here I have demonstrated that the requirements for IS1111-*attC* transposase mediated recombination echo those of IntI1. IS1111-*attC* Tase recombination requires a single stranded, *Pseudomonas*-type *attC* site (top strand) hairpin and a double stranded IS1111-*attC* minicircle. The ability of IntI and Tase to preferentially target opposite *attC* strands, is predicted to reduce competition between them and drives strand orientation preference. Whilst I was able to demonstrate that IS1111-*attC* transposase acts in a similar fashion to IntI protein and targets *attC* sites strand specifically, questions remain to be answered:

1) *Is attC site recognition by IS1111 transposase specific?*

Work presented here suggests that transposase binding is strand and structure specific. DNA fingerprinting assays would determine the exact binding region of Tase to the top strand of *attC* sites.

2) *Are the extrahelical bases in the top Pseudomonas attC site responsible for transposase docking?*

Work conducted using purified integron integrase has highlighted the importance of the conserved nucleotides G and T (Mazel, 2006) for protein binding and docking to single stranded *attC* sites. The exact point of transposase protein-*attC* DNA interaction to date remains unknown. To study this interaction, site alterations can be made to the *Pseudomonas*-type *attC* site and tested individually using EMSAs.

3) *What role do the conserved amino acids D-E-D-D play in transposase activity on attC sites?*

These four acidic residues were shown to be crucial for Piv mediated inversion, however their role in IS translocation remains unknown. A mutagenesis library of transposase protein would shed insight into the role of these amino acids in *attC* site interaction.

4) *How is the site-specific recombination of IS1111-attC and attC site resolved?*

The current model for integron mediated recombination of *attI* and *attC* sites, as well as Piv mediated catalysis of inversion of *invL* and *invR* sites, involve a Holliday junction as an intermediate state. To investigate if IS1111-*attC* elements pass through an intermediate Holliday junction state *in vitro*, perhaps it may be possible, although challenging, to crystallise the transposase-HJ intermediate and determine its X-ray structure. Comparable studies have achieved this for the Cre-*loxP* site specific recombination system (Gopaul *et al.*, 1998).

Chapter 6 Conclusions

6.1 Overview

In this study I aimed to explore the role of horizontal gene transfer in the evolution of multi-drug resistance. Bacterial multi-drug resistance is a world-wide problem, often associated with severe and untreatable infections. Treatment is very challenging due to the limited susceptibility of bacteria to antimicrobial agents and the high frequency of antibiotic resistance during therapy (Aloush *et al.*, 2006). Bacterial population response to administered drug therapy; likelihood of MDR selection and rate of occurrence as a consequence of treatment; and origin of resistance genes, are just a few of the confounding factors of MDR. Many aspects of MDR remain poorly understood. In particular, factors that may facilitate/hinder movement of resistance genes, such as biological distances between different organisms and the rate of gene recruitment, require further investigating.

In this study, biological distance was defined as the sum of i) molecular and genetic compatibility and ii) evolutionary and ecological relatedness between microorganisms. Biological distance encompasses spatial separations wherein phylogenetically related organisms may share similar genes and transcriptional machinery because they occupy distinct ecological niches (Friedrich *et al.*, 2001) but never meet in high numbers. Also considered were the evolutionary mechanics of organisms wherein ecologically similar organisms that are distantly related, such as Cyanobacteria and *Pseudomonadaceae*, may frequently co-exist (Gallucci & Paerl, 1983) but differences in basic metabolism and cell processes impact on their ability to exchange DNA. Both intracellular and ecological contexts were predicted to have been overcome with respect to assembly of multi-resistance elements. The general assumption was that the vast majority of clinically relevant MDR is encoded by genes that originated in the environment, yet there are very few examples of direct gene exchange between these two ecosystems (Perry & Wright, 2013). The few examples available, are linked to various replicons and integrons (Stalder *et al.*, 2012). Thus the fundamental question examined in this study was how genetic elements have evolved special features to facilitate movement across greater biological distances.

Movement of genes between biologically distinct groups of γ -Proteobacteria implicated in multi-drug resistance was explored. The main bacterial groups were environmental and clinical *Pseudomonas* populations as well as a phylogenetically distant *Enterobacteriaceae*

clinical group. Integron-associated resistance across these groups was used as a model system to explore MDR formation. I specifically focussed on integrons, gene cassettes and the *IS1111-attC* elements as markers for gene cassette movement between groups of differing biological distance. The first series of experiments was based on the premise that PCI, *Pseudomonas*-type *attC* sites and *IS1111-attC* elements had a distinct source in environmental *Pseudomonads*. I aimed to test the hypothesis that *IS1111-attC* elements would move from the environmental pool into *Ps. aeruginosa* with higher frequency when compared to *Enterobacteriaceae*. A secondary aspect investigated was whether such movement would be dependent on mobilized class 1 integrons.

6.2 *IS1111-attC*: an epidemiological marker element for ecological barrier crossing in *Pseudomonadaceae*

To model our understanding of MDR formation and dissemination, the rate and movement of resistance genes from one ecological context to another must be addressed. However, as the environmental source for clinically relevant antibiotic resistance genes remains largely unknown, we are unable to examine gene flow in “real time” across ecological boundaries. Initial observations of *IS1111-attC* distribution in the *Ps. stutzeri* complex (Tetu & Holmes, 2008) and MRIs led to the hypothesis that these ISs could be used as marker elements for gene cassette flow between *Pseudomonads*. I hypothesised that *IS1111-attC* mobility would occur without the need for selection/selective pressures allowing for its use in this study to explore gene flow across ecological barriers.

I have shown *IS1111-attC* distribution includes *Pseudomonads* outside of the *stutzeri* species complex and is strongly associated with PCIs. The *Pseudomonads* studied were isolated from a variety of sample sites, ranging from soil, domestic sponges, mops and various tissues samples. *IS1111-attC* distributions were observed to vary across distantly related *Pseudomonads*, thus demonstrating that *IS1111-attC* can cross genetic barriers. The data from my study indicated *IS1111-attC* associated with PCIs in environmentally sourced isolates, but for clinical *Pseudomonas aeruginosa* isolates association was reliant on the presence of MRI.

My data showed that *IS1111-attC* exists across the three conceptual groups in our model, but at different frequencies and in different contexts (Figure 6.1). From this I conclude that

IS1111-attC elements are not constrained by ecological separations and are able to establish in a variety of Pseudomonad populations. Thus, the observed frequency and distribution of IS1111-attC highlights the importance of ecological niche and compatibility in genomic machinery in facilitating IS dissemination. A larger, more explicitly targeted sample collection would be needed to formally test the significance of the observed IS frequency and diversity differences across the sample populations. This would also allow us to use IS1111-attC elements to measure rates of gene flow across biological differences.

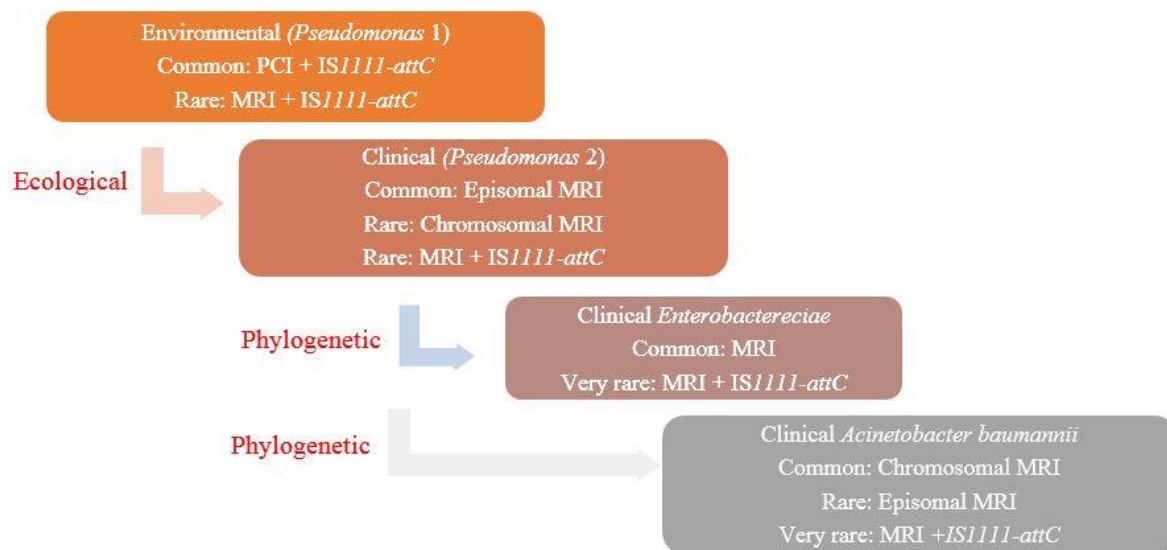


Figure 6.1: Ecological and phylogenetic barriers impact on IS1111-attC distribution across Proteobacteria populations.

Common, rare, and very rare frequencies were arbitrarily defined by the observed occurrence of IS1111-attCs, PCIs and MRIs in this study, published studies and NCBI database searches.

In this study, non-identical IS1111-attC elements were detected, indicating multiple independent acquisitions. However, these elements clustered to previously identified IS1111-attC subgroups. Their distribution was consistent with the model of a diverse group of organisms (clinical Pseudomonads) acquiring IS1111-attC elements via a specific route (class 1 MRI) from a narrow source pool (environmental Pseudomonads group) multiple times. The observation that diverse IS1111-attC elements were located on specific mobile complexes (*intI1* subtypes) indicated that mobile elements were required to facilitate their transfer across ecological boundaries. Environmental *Ps. stutzeri* were hence considered an IS1111-attC reservoir from which *Ps. aeruginosa* has preferential access relative to *Enterobacteriaceae*. In this study, recovery of an environmental source for ISPst6-like elements in MDR *Ps.*

aeruginosa clinical isolates supported the idea that clinical *Pseudomonas* has access to the vast MGE pool from environmental sources.

A subpopulation of *Enterobacteriaceae* were also screened for IS1111-*attC* elements, however none were recovered. Recent database analysis performed in this study revealed two cases where clinical *Acinetobacter baumannii* strains had acquired an IS1111-*attC*. To date no IS1111-*attC* elements have been detected in *Acinetobacter baumannii* resistance islands (Hamidian & Hall, 2011; Krizova & Nemeč, 2010; Nigro & Hall, 2012). The two mentioned ISs were embedded within *attC_{aadA1}* (IS*Aba3* Accession number: JN253504.1) and *attC_{aadB}* (IS*Pa21* Accession number: AM74982.12) sites in different MRI arrays. Both *intI1* arrays were previously seen in episomal contexts in other clinical γ -Proteobacteria, such as *Ps. aeruginosa* and *E. coli*. As such, I conclude that while these ISs were not embedded in a new target site context and were highly unlikely to be in a new replicon context, they are definitely in a new cell context.

Antibiotic and bactericide contamination of the environment is predicted to be a driver for the increased recruitment of antibiotic resistance phenotypes (Davies & Davies, 2010). Selective pressure from human activity may influence the composition of the environmental floating genome and thus the balance of the resulting gene transfer events (da Costa *et al.*, 2013; Perry & Wright, 2013). Clinical settings are highly selective and hostile environments to organisms not already adapted to their pressures. An organism removed from a clinical habitat into an environment such as soil, may encounter different selective pressures such as metal, antibiotic contamination and competition from the native population (Hibbing *et al.*, 2010; Krishna *et al.*, 2012). The presence of MGEs in clinical isolates provides these organisms with an advantage in sequestering novel elements from the floating genome. We can thus expect clinical organisms to persist in the environment (da Costa *et al.*, 2013) due to their molecular-genetic makeup. Subsequent transportation back into the clinical/animal system is expected to occur with the organism having acquired new “molecular armour” through HGT with the native population.

In comparison, environmental organisms are only expected to encounter excessive levels of antibiotic pressure when they move from their ecological niche to a clinical setting. For them to persist in this new niche, they may need to undergo mutations and/or acquire the necessary genetic makeup in order to survive. As they do not have all of the necessary components of

the floating genome in a meaningful context, I postulate that the potential for the long-term establishment of these environmental organisms is greatly reduced. My results support the hypothesis that organisms travel from clinical into environmental habitats, where they may acquire new genetic material at a greater rate in comparison to environmental organisms crossing into the clinic/animal system.

6.3 Are there molecular-genetic barriers to IS1111-attC translocation between PCI and MRI contexts?

IS1111-attC subgroups have the capacity to recognise and become associated with different attCs. This highlights a parallel in MRI integrase recognition of different attC sites (Biskri *et al.*, 2005). Yet IS1111-attC distribution between PCI and MRI attCs is highly variable (Tetu & Holmes, 2008). Also, there is no evidence for dynamic sequence changes involving the IS in MRI arrays, despite it appearing to be common in PCI arrays. This raises the question as to whether there are barriers to IS1111-attC activity in *Ps. aeruginosa*, enterics, or class 1 integrons that restrict further movement. If IS1111-attC elements move independently of MRIs under rate limiting conditions, then they should be present in multiple sites in *Pseudomonas* MRIs. If however, the MRI capture of IS1111-attC from a PCI source is not rate-limiting, we should see IS1111-attC elements in *Enterobacteriaceae* at similar frequencies.

Firstly, I examined whether or not molecular-genetic barriers are a rate-limiting factor to IS acquisition and post-acquisition activity. Acquisition of IS1111-attC elements by the synthetic trap was confirmed to occur with mild or no dependence on attC site structure. Translocation efficiency varied when same/similar IS1111-attC and attC-target site partners were tested in two different species, *Ps. stutzeri* and *E. coli*. The difference in IS1111-attC translocation activity is thus believed to be influenced by host cell factors (Figure 6.2).

Secondly, I examined the ability of IS1111-attC to recognise target site context and/or secondary structure. IS1111-attC translocation was found to be a result of both genetic context and topology of DNA. Interestingly, whilst translocation was observed from a chromosomal *Pseudomonas*-type attC into an episomal attC context, the reverse was not observed. Once integrated into an episomal context, IS1111-attC post-translocation activity appeared to be altered. The target sites were the same in both bacterial species tested, but PCI

invasion by *IS1111-attC* was not detected. Differences in translocation were seen when natural transformation and electroporation were used, which implies that the formation of a single-stranded *attC* template is critical in *IS1111-attC* translocation success.

6.3.1 Uptake method impacts on *IS1111-attC* acquisition: natural transformation vs electroporation

In the Tetu & Holmes (2008) study only partial translocation of the IS was observed. This implied that while the IS was potentially functional, not all conditions for complete translocation had been met. To test the hypothesis that host cell differences affect *IS1111-attC* translocation success, a natural host species *Ps. stutzeri* ATCC 17587 was used. In my study, complete translocation of *IS1111-attC* elements was found to be stably maintained. Complete translocation is said to have occurred when a single intact *ISPst6* element, including both IR_L and IR_R junctions, are detected in the *attC* plasmid trap. Recovery of the intact IS element suggests that the molecular-genetic barrier is not a limiting factor when considering *IS1111-attC* translocation from a chromosomal to an episomal context in a native *Pseudomonas* host. Instead, I propose that host associated factors are a major driving force in *IS1111-attC* translocation success (Figure 6.2).

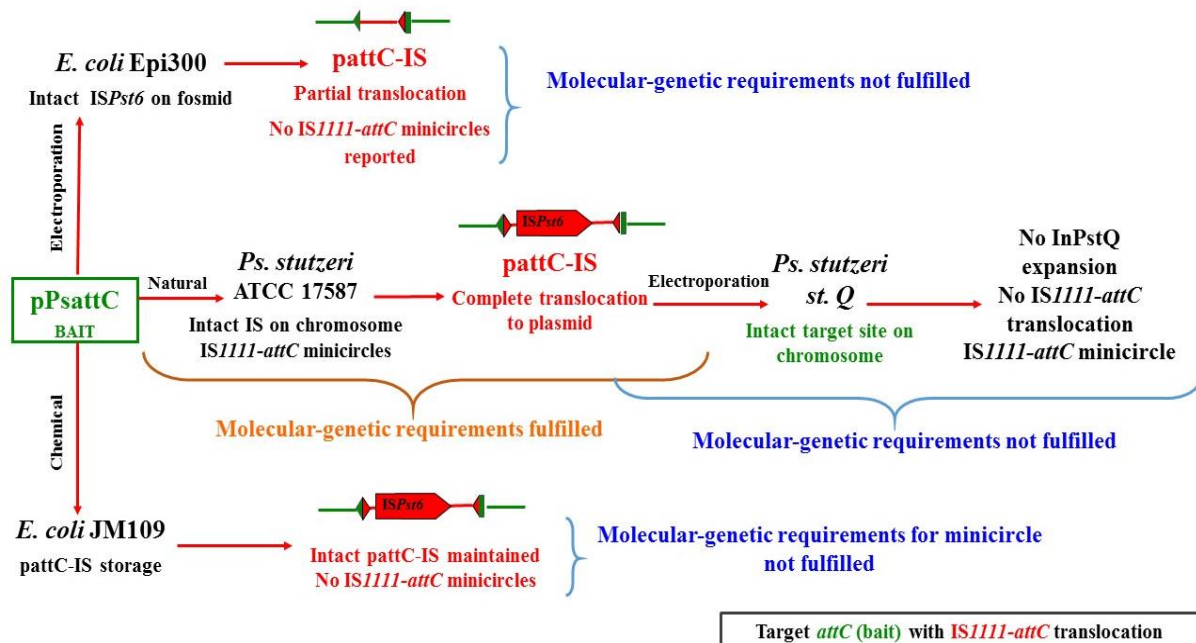


Figure 6.2: Overcoming molecular-genetic barriers in a *Pseudomonas stutzeri* host leads to successful *IS1111-attC* translocation into *attC* sites.

E. coli JM109 was used as a host strain for long-term storage of *pattC-IS*. *IS1111-attC* minicircles were only observed in a *Pseudomonas* host, suggesting that host associated factors drive *Tpase* expression. Note electroporation of *pattC* trap into *E. coli Epi300* was performed by Tetu & Holmes (2008).

As a further test, I transformed the plasmid-borne IS-*attC* into *Pseudomonas stutzeri* species. Transformation was only possible via electroporation, which raised the question as to the importance of presentation context of the IS-*attC*, rather than host cell properties. A precedent for this exists in IntI mediated recombination (Collis *et al.*, 2001). Whilst no IS invasion of PCI arrays was detected, when the trap was electroporated into a *Pseudomonas* host (Figure 6.2), IS1111-*attC* minicircles were detected. This implies that the PCI array is “protected” from IS invasion and/or another *Pseudomonas* factor is required. In the case of host factors, the availability and recognition of a single stranded PCI-*attC* hairpin by IS1111-*attC* is postulated to affect IS translocation success. Additionally, if *Pseudomonas* integrase IntIPstX and IS transposase target the same *attC* strand, steric interference may occur. Thus competition between these two recombinases may affect IS1111-*attC* translocation.

In summary, IS1111-*attC* translocation into synthetic, plasmid-borne *attC* sites was observed to occur in a site-specific fashion. These findings also support the notion that the occurrence of a single IS1111-*attC* within an MRI in a Pseudomonad reflects the acquisition of a single IS from a PCI, as demonstrated here, or the acquisition of a pre-existing IS-gene cassette fusion (Tetu & Holmes, 2008).

6.3.2 The context of the attC site impacts on IS1111-attC translocation: topology specificity

The orientation in which *attC* was presented to IS1111-*attC* was noted to affect translocation efficiency (Chapter 4). IS1111-*attC* translocation preferentially occurred when the synthetic *attC* sites were cloned into one orientation. Similar results have been noted for IS4321 translocation (Partridge & Hall, 2003) and IntI1 mediated recombination at synthetic *attC* sites (Collis *et al.*, 2001). If recombination by IS1111-*attC* occurs in a strand-specific manner, then strand topology of the *attC* hairpin will differ between the two cloned orientations. Orientation preference is thus believed to be a feature of plasmid topology, where one orientation exposes the *attC* cruciform in a way which “presents” the *attC* top strand for transposase binding and recombination. Another possible explanation is that IS1111-*attC* translocation into one *attC* orientation ensures that the IS1111-*attC* P_{IRL} and putative P_{OUT} promoter do not interfere with the *rep* region of the pBBR1 vector and thereby negatively affect plasmid fitness (Collis *et al.*, 2001).

Furthermore, these *in vivo* findings add support to bioinformatic analysis of IS1111-*attC* orientation specific translocation into MRI and CI *attC* sites (Post & Hall, 2009; Tetu & Holmes, 2008). Orientation bias may reflect consequences of the mechanism by which T_{psa} mediates site-specific recombination. There are two main consequences predicted to arise due to preferential targeting of single stranded *attC* sites. Firstly, the fitness of IS1111-*attC* is dependent on biological processes that produce a single stranded *attC*. Secondly, orientation of IS1111-*attC* may cause polar effects on cellular processes, through the disruption of expression of any adjacent genes (Aubert *et al.*, 2003; Kallastu *et al.*, 1998). IS1111-*attC* orientation is a product of which *attC* strand is joined, as well as to which IS1111-*attC* the strand is joined. I predict that the top (non-coding) IS strand is the target for *attC* recombination (Chapter 5, Figure 5.18). Also, IS1111-*attC* direct repeats may act as targets for RNA polymerase, thereby pausing RNA polymerase activity and consequently leaving the IS top strand free to act as the *attC* recombination partner strand. Orientation preference is therefore an important feature of IS1111-*attC* translocation.

6.3.3 IS1111-*attC* and micro-variations: a consequence of polymerase activity?

Resolution of *attC-attI* junctions in IntI-mediated recombination was demonstrated to occur by replication, involving potentially multiple enzymes (Loot *et al.*, 2012). A feature of IntI recombination substrates is that they show micro-variation. The *attC* sites and also the cassette ORFs each show levels of heterogeneity that are higher than other genetic features in the same replicons. In a chromosomal integron, ‘families’ of *attC* sites are typically observed (XCR, VCR, PCR, *etc.*) (Gillings *et al.*, 2005; Rowe-Magnus *et al.*, 2002; Vaisvila *et al.*, 2001). Within one genome, the sequence variation across the *attC* sites is greater than the sequence variation between the *intI* genes of different strains (*Ps. stutzeri*- all observed genomovars; *Vibrio cholerae* and *Xanthomonas campestris* -all observed pathovars). In Class 1 integrons, for those gene cassettes that have been widely disseminated (such as *aadB* and *bla_{OXA}*) multiple variants have been observed in the different integron occurrences (Figure 1.12), but the *intI* sequence and the *attI* sequence show higher conservation.

The same pattern is seen for IS1111-*attC*, whereby micro-variation is seen between cases where translocation is inferred to have occurred (i.e. between independent locations). Significantly the ‘captured’ element observed in my study was also non-identical to known

instances in *Ps. stutzeri* ATCC 17587 (Figure 4.10). *In vivo* activity of error prone DNA polymerases during DNA replication in *Ps. stutzeri* may result in minor IS1111-*attC* and *Pseudomonas*-type *attC* sequence differences across PCI arrays. Whilst there is no direct evidence of error prone replication in *Ps. stutzeri*, it is plausible that such polymerase activity can be used to explain the cases of imperfectly copied gene cassettes and *attC* sites across PCIs. Adaptation of a *Pseudomonas* population under environmental and selective pressures may result in stress-induced activation of error-prone DNA polymerases (Foster, 2000; Rosenberg, 2001; Tegova *et al.*, 2004; Tenaillon *et al.*, 2001) and result in IS-gene cassette genetic difference in population members. Interestingly, near perfect IS copies are observed in MRIs. IS variation is only introduced during translocation so if there is no ongoing movement there is no change. I propose that high fidelity polymerases involved in plasmid replication (Allen *et al.*, 2011) are less likely to introduce errors/sequence differences in IS1111-*attC* elements, in comparison to their chromosomal counterparts.

I propose that the occurrence of IS1111-*attC*-gene cassette micro-variants can be explained if the mechanism of co-integrate resolution in both IntI-mediated and Transposase-mediated recombination is inherently error prone. If the ability to resolve the co-integrate is limited to a specific subset of DNA polymerases then this could conceivably restrict the range of cells in which active recombination involving IntI or IS1111-*attC* can take place.

6.3.4 IS1111-*attC* may preferentially recognise *Pseudomonas*-type *attC* over *attC*_{*aadB*} sites

One of the most biologically significant features of IntI is its ability to recognise a wide variety of *attC* sites (Jacquier *et al.*, 2009; Larouche & Roy, 2011). *In vivo* assays in this study demonstrated IS1111-*attC* ability to translocate into *Pseudomonas* and classical *attC* sites with equal propensity. *In vitro* EMSAs were used to test the binding specificity of T_pase to double stranded and single-stranded *Pseudomonas* and classical *attC* site. Whilst the results for IS1111-*attC* transposase binding to *Pseudomonas*-type *attC* were consistent with my predictions, the results from T_pase binding to *attC*_{*aadB*} were somewhat inconclusive. The latter was surprising, given that *attC*_{*aadB*} sites were found to be occupied by IS1111-*attC* members (Tetu & Holmes, 2008). In principal, binding of transposase to *attC*_{*aadB*} sites should be possible if transposase interacts with the *attC* target with greater affinity than the competitor, and when the secondary binding does not discriminate between the sequences (Alves & Cunha, 2012). It is possible that additional molecular-genetic factors are necessary

for T_pase-*attC*_{*aadB*} binding. In summary, *in vivo* data demonstrated that IS1111-*attC* transposase has a flexible *attC* substrate range. *In vitro* data together with *in vivo* data indicate that a structural conformation in *Pseudomonas attC* top strands is essential for T_pase binding. The precise single stranded conformations required may be more difficult to form for *attC*_{*aadB*} sites, thus creating a degree of site preference. However, further experimentation is required before any conclusions can be made.

6.3.5 *The importance of DNA structures and presentation context in the HGT landscape*

Although the concept of the gene is still evolving, one feature that has been true since the introduction of Mendelian genetics is that the gene is a unit. As such it must have defined sequence boundaries. Primary sequence boundaries do not change topologically (at least not on timescales relevant to an individual cell's physiology), while secondary sequence boundaries (secondary structures) do. Secondary structures formed by intrastrand base pairing are now known to often be critical for biological processes such as replication, transcription, CTX-phage integration, IS608 transposition, as well as integron mediated recombination of *attC* sites (Bouvier *et al.*, 2009; Guynet *et al.*, 2008; Val *et al.*, 2005; Bikard *et al.*, 2010). The use of secondary structures, that are transiently present, creates a scenario whereby temporal processes become important to cell outcomes and have led to optimisation of transfer events with other cell processes.

Various processes can result in the transient production of single stranded *attC* hairpins (Bikard *et al.*, 2010; Loot *et al.*, 2010). DNA supercoiling and lagging strand synthesis can also favour the extrusion of an *attC* cruciform (Loot *et al.*, 2013). The ability of IntI to recognise and recombine *attC* and *attI* sites, is a function of DNA topologies of both target sites. Topological properties of *attC* sites can be influenced by a few factors:

- i) *Stability of single-stranded structure:* Formation of an unstable *attC* hairpin is more likely to yield an unfavourable substrate for integron integrase (Loot *et al.*, 2010). Multiple single-stranded *attC*_{*aadB*} conformations were observed in EMSAs, some of which are hypothesised to be unfavourable for T_pase binding potential.
- ii) *Resistance of double stranded structure:* Length of the variable terminal structure (VTS) affects IntI binding and *attI-attC* recombination. The larger the VTS, the less efficient the DNA melting and the lower the recombination frequency (Loot *et al.*,

2010). Inadequate melting of double stranded *attC* can lead to formation of unstable *attC* structures, which is unfavourable for T_pase binding.

- iii) *Unwinding or stabilizing factors that promote either i) or ii):* Ability to “attract” DNA binding proteins that destabilize the hairpin structure, such as RecA and Single Stranded Binding proteins (SSBs) (Cambray *et al.*, 2011; Dickey *et al.*, 2013; Marceau, 2012; Loot *et al.*, 2010; Loot *et al.*, 2013; Nowak *et al.*, 2014). SSBs prevent DNA hairpin formation by stabilizing the single stranded DNA, and directing RecA binding to the hairpin (Kowalczykowski & Krupp, 1987; Meyer & Laine, 1990). SSBs bind *attC* sites by structural recognition, rather than through sequence specificity (Cambray *et al.*, 2011). This highlights that spatial presentation of contact points for binding is more important than the neighbouring bases.
- iv) *Ability of IntI to displace SSBs/ from attC:* Purified IntI has the capacity to capture and stabilize the *attC* hairpin, thus counterbalancing the effect of SSBs (Loot *et al.*, 2013). I postulate that *Pseudomonas* SSBs have an impact on the success of IS1111-*attC* binding to single stranded *attC* sites thereby inducing an equilibrium state shift of the hairpin and destabilizing it in an effort to make it a linear DNA molecule. This binding is antagonistic, given that transposase binding to single stranded *attC* is believed to be dependent on hairpin formation. Hence transposase was unable to mediate the site-specific recombination of IS1111-*attC* into PCI-*attC*s. However, if IntI and T_pase act co-operatively, the ability of IntI to displace SSBs from *attC* sites would promote the formation of hairpins, thus allowing T_pase to bind and drive IS1111-*attC* translocation into *attC* strands.

In conclusion, evolutionary adaptation of IS1111-*attC* transposase has allowed the enzyme to recognise DNA structures (cruciform) as gene handles for successful translocation. IS1111-*attC* fitness and translocation is structure dependent in a manner that mimics IntI activity.

6.5 Convergent evolution, selective advantages of a “gene handle”.

The significance of secondary structure formation in single stranded *attC* sites, and their role in integron mediated recombination and excision of gene cassettes, is well understood (Bouvier *et al.*, 2009; Johansson *et al.*, 2004; Larouche & Roy, 2011; Loot *et al.*, 2013). As *attC* structure was hypothesized to be key a factor in IS1111-*attC* transposase recognition and binding, secondary structure formation was found to be particularly interesting in this study also. The ability of IntI and T_{pase} to target *attC* sites is expected to promote formation of mosaic genetic elements that ultimately result in complex phenotypes such as MDR. In this study, I have demonstrated that *attC* sites are a shared ancestral feature recognized by both the IntI and T_{pase} recombinases. Preferential binding of IS1111-*attC* T_{pase} to the *attC_{ts}* strand was shown repeatedly in mobility shift assays. This is in contrast to IntI binding, where the preference was for the *attC* bottom strand (Francia *et al.*, 1999; Johansson *et al.*, 2004).

As in the IntI binding assay, double stranded *attC* sites were not recognised by T_{pase} under *in vitro* conditions. Furthermore, a fast migrating mobility-shifted complex was observed when the IS1111-*attC* minicircle junction was tested as a T_{pase} binding partner. This was evidence that electrophoretically distinct subpopulations of T_{pase}-IS junction interactions exist, as shown by shifting oligonucleotide equilibrium. As a result, T_{pase} mediated site-specific recombination between the IS and *attC* site is believed to involve a double stranded IS minicircle and a folded-back *attC_{ts}*. In conclusion, this study presents the first evidence of the molecular requirements for IS1111-*attC* site-specific recombination with integron associated recombination sites. T_{pase} *attC* binding ability has shed significant insight into the nature of specific DNA-protein interactions.

The results presented throughout this thesis highlight that there is a dynamic and ongoing process of horizontal transfer of genetic material amongst environmental and clinical Pseudomonads. Furthermore, this study also highlights the detection of movement with which MGEs move between genetic contexts. Understanding the driving forces behind IS1111-*attC* distribution, abundance, and the molecular details involved in transposase interaction with the integron platform has shed light on the mechanisms that govern gene cassette flow between ecologically and phylogenetically distant Pseudomonads.

References

- Ahmed, Z., Khan, S. S., & Khan, M.** (2013). *In vitro* trials of some antimicrobial combinations against *Staphylococcus aureus* and *Pseudomonas aeruginosa*. *Saudi Journal of Biological Sciences*, **20**(1), 79–83.
- Allen, H. K., Donato, J., Wang, H. H., Cloud-Hansen, K. A., Davies, J., & Handelsman, J.** (2010). Call of the wild: antibiotic resistance genes in natural environments. *Nature Reviews. Microbiology*, **8**(4), 251–259.
- Allen, J.M., Simcha, D. M., Ericson, N. G., Alexander, D. L., Marquette, J. T., Van Biber, B. P., Troll, C. J., Karchin, R., Bielas, J. H., Loeb, L. A., & Camps, M.** (2011). Roles of DNA polymerase I in leading and lagging-strand replication defined by a high-resolution mutation footprint of ColE1 plasmid replication. *Nucleic Acid Research*, **39**(16), 7020-7033.
- Aloush, V., Navon-Venezia, S., Seigman-Igra, Y., Cabili, S., & Carmeli, Y.** (2006). Multidrug-resistant *Pseudomonas aeruginosa*: risk factors and clinical impact. *Antimicrob Agents Chemother*, **50**(1), 43–48.
- Alves, C., and Cunha, C.** (2012). Electrophoretic Mobility Shift Assay: Analyzing Protein - Nucleic Acid Interactions, Gel Electrophoresis - Advanced Techniques, Dr. S. Magdeldin (Ed.), ISBN: 978-953-51-0457-5, InTech, Available from: <http://www.intechopen.com/books/gel-electrophoresis-advancedtechniques/electrophoretic-mobility-shift-assay-analyzing-protein-nucleic-acid-interactions>
- Andam, C. P., & Gogarten, J. P.** (2011). Biased gene transfer in microbial evolution. *Nature Reviews. Microbiology*, **9**(7), 543–555.
- Andersson, D. I., & Hughes, D.** (2014). Microbiological effects of sublethal levels of antibiotics. *Nature Reviews. Microbiology*, **12**(7), 465–478.
- Angert, E. R.** (2005). Alternatives to binary fission in bacteria. *Nature Reviews. Microbiology*, **3**(3), 214–224.
- Antoine, R., & Locht, C.** (1992). Isolation and molecular characterization of a novel broad-host-range plasmid from *Bordetella bronchiseptica* with sequence similarities to plasmids from gram-positive organisms. *Molecular Microbiology*, **6**(13), 1785–1799.
- Aubert, D., Naas, T., & Nordmann, P.** (2003). IS1999 increases expression of the extended-spectrum beta-lactamase VEB-1 in *Pseudomonas aeruginosa*. *Journal of Bacteriology*. **185**(17), 5314–5319.
- Aubert, D., Naas, T., Héritier, C., Poirel, L., & Nordmann, P.** (2006). Functional characterization of IS1999, an IS4 family element involved in mobilization and expression of beta-lactam resistance genes. *Journal of Bacteriology*, **188**(18), 6506–6614.

- Bailey, J. K., Pinyon, J. L., Anantham, S., & Hall, R. M.** (2011). Distribution of the *bla*TEM gene and *bla*TEM-containing transposons in commensal *Escherichia coli*. *The Journal of Antimicrobial Chemotherapy*, **66**(4), 745–751.
- Baker, M.** (2012). *De novo* genome assembly: what every biologist should know. *Nature Methods*, **9**(4), 333–337.
- Barlow, M., & Hall, B. G.** (2002). Origin and evolution of the AmpC β -lactamases of *Citrobacter freundii*. *Antimicrobial Agents and Chemotherapy*, **46**(5), 1190–1198.
- Bartlett, D. H., & Silverman, M.** (1989). Nucleotide sequence of IS492, a novel insertion sequence causing variation in extracellular polysaccharide production in the marine bacterium *Pseudomonas atlantica*. *Journal of Bacteriology*, **171**(3), 1763–1766.
- Battle, S. E., Meyer, F., Rello, J., Kung, V. L., & Hauser, A. R.** (2008). Hybrid pathogenicity island PAGI-5 contributes to the highly virulent phenotype of a *Pseudomonas aeruginosa* isolate in mammals. *Journal of Bacteriology*, **190**, 7130–7140.
- Beaber, J. W., Hochhut, B., & Waldor, M. K.** (2004). SOS response promotes horizontal dissemination of antibiotic resistance genes. *Nature*, **427**(6969), 72–74.
- Bellais, S., Aubert, D., Naas, T., & Nordmann, P.** (2000). Molecular and biochemical heterogeneity of class B carbapenem-hydrolyzing β -lactamases in *Chryseobacterium meningosepticum*. *Antimicrobial Agents and Chemotherapy*, **44**(7), 1878–1886.
- Bendak, K., Loughlin, F. E., Cheung, V., O’Connell, M. R., Crossley, M., & Mackay, J. P.** (2012). A rapid method for assessing the RNA-binding potential of a protein. *Nucleic Acids Research*, **40**(14), e105.
- Bender, J., & Kleckner, N.** (1986). Genetic evidence that Tn10 transposes by a nonreplicative mechanism. *Cell*, **45**(6), 801–815.
- Bender, J., & Kleckner, N.** (1992). Tn10 insertion specificity is strongly dependent upon sequences immediately adjacent to the target-site consensus sequence. *Proceedings of the National Academy of Sciences of the United States of America*, **89**(17), 7996–8000.
- Bennett, P. M.** (1999). Integrons and gene cassettes: a genetic construction kit for bacteria. *Journal of Antimicrobial Chemotherapy*, **43**, 1–4.
- Bennett, P. M.** (2004). Genome plasticity: insertion sequence elements, transposons and integrons, and DNA rearrangement. *Methods in Molecular Biology*, **266**, 71–113.
- Berkner, K. L., & Folk, R.** (1977). Polynucleotide kinase exchange reaction: quantitative assay for restriction endonuclease-generated 5'-phosphoryl termini in DNA. *The Journal of Biological Chemistry*, **252**, 3176–3184.
- Biel, S. W., & Berg, D. E.** (1984). Mechanism of IS1 transposition in *E. coli*: choice between simple insertion and cointegration. *Genetics*, **108**, 319–330.

- Bikard, D., Loot, C., Baharoglu, Z., & Mazel, D.** (2010). Folded DNA in action: hairpin formation and biological functions in prokaryotes. *Microbiology and Molecular Biology Reviews : MMBR*, **74**(4), 570–588.
- Biskri, L., Bouvier, M., Guérout, A., Boisnard, S., & Mazel, D.** (2005). Comparative study of class 1 integron and *Vibrio cholerae* superintegron integrase activities. *Journal of Bacteriology*, **187**(5), 1740–1750.
- Bissonnette, L., & Roy, P. H.** (1992). Characterization of In0 of *Pseudomonas aeruginosa* plasmid pVS1, an ancestor of integrons of multiresistance plasmids and transposons of gram-negative bacteria. *Journal of Bacteriology*, **174**(4), 1248–1257.
- Blakely, G. W.** (2000). Sequential strand exchange by XerC and XerD during site-specific recombination at *dif*. *Journal of Biological Chemistry*, **275**(14), 9930–9936.
- Boucher, Y., Labbate, M., Koenig, J. E., & Stokes, H. W.** (2007). Integrons: mobilizable platforms that promote genetic diversity in bacteria. *Trends in Microbiology*, **15**(7), 301–309.
- Bouvier, M., Ducos-Galand, M., Loot, C., Bikard, D., & Mazel, D.** (2009). Structural features of single-stranded integron cassette *attC* sites and their role in strand selection. *PLoS Genetics*, **5**(9), e1000632.
- Boyd, E. F., Almagro-Moreno, S., & Parent, M. A.** (2009). Genomic islands are dynamic, ancient integrative elements in bacterial evolution. *Trends in Microbiology*, **17**(2), 47–53.
- Breidenstein, E. B. M., de la Fuente-Núñez, C., & Hancock, R. E. W.** (2011). *Pseudomonas aeruginosa*: all roads lead to resistance. *Trends in Microbiology*, **19**(8), 419–426.
- Brüssow, H., Canchaya, C., Hardt, W., & Bru, H.** (2004). Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. *Microbiology and Molecular Biology Reviews*, **68**(3), 560–602.
- Bruton, C. J., & Chater, K. F.** (1987). Nucleotide sequence of IS110, an insertion sequence of *Streptomyces coelicolor* A3(2). *Nucleic Acids Research*, **15**, 7053–7065.
- Buchner, J. M., Robertson, A. E., Poynter, D. J., Denniston, S. S., & Karls, A. C.** (2005). Piv site-specific invertase requires a DEDD motif analogous to the catalytic center of the RuvC Holliday junction resolvases, **187**(10), 3431–3437.
- Burrus, V., Pavlovic, G., Decaris, B., & Guédon, G.** (2002). Conjugative transposons: the tip of the iceberg. *Molecular Microbiology*, **46**(3), 601–610.
- Cambray, G., Guerout, A.-M., & Mazel, D.** (2010). Integrons. *Annual Review of Genetics*, **44**, 141–166.
- Cambray, G., Sanchez-Alberola, N., Campoy, S., Guerin, E., Da Re, S., González-Zorn, B., Ploy, M. C., Barbée, J., Mazel, D., & Erill, I.** (2011). Prevalence of SOS-mediated

- control of integron integrase expression as an adaptive trait of chromosomal and mobile integrons. *Mobile DNA*, **2**(1), 6.
- Cantón, R.** (2009). Antibiotic resistance genes from the environment: a perspective through newly identified antibiotic resistance mechanisms in the clinical setting. *Clinical Microbiology and Infection*, **15**, 20–25.
- Carattoli, A.** (2009). Resistance plasmid families in *Enterobacteriaceae*. *Antimicrobial Agents and Chemotherapy*, **53**, 2227–2238.
- Carattoli, A.** (2013). Plasmids and the spread of resistance. *International Journal of Medical Microbiology*, **303**(6-7), 298–304.
- Cattoir, V., Nordmann, P., Silva-Sanchez, J., Espinal, P., & Poirel, L.** (2008). ISEcp1-mediated transposition of *qnrB*-like gene in *Escherichia coli*. *Antimicrobial Agents and Chemotherapy*, **52**(8), 2929–2932.
- Cattoir, V., Poirel, L., Mazel, D., Soussy, C.-J., & Nordmann, P.** (2007). *Vibrio splendidus* as the source of plasmid-mediated *QnrS*-like quinolone resistance determinants. *Antimicrobial Agents and Chemotherapy*, **51**(7), 2650–2651.
- Cerveau, N., Leclercq, S., Leroy, E., Bouchon, D., & Cordaux, R.** (2011). Short- and long-term evolutionary dynamics of bacterial insertion sequences: insights from *Wolbachia* endosymbionts. *Genome Biology and Evolution*, **3**, 1175–1186.
- Cesareni, G., Helmer-Citterich, M., & Castagnoli, L.** (1991). Control of ColE1 plasmid replication by antisense RNA. *Trends in Genetics*, **7**(7), 230–235.
- Chalmers, R. M., & Kleckner, N.** (1996). IS10/Tn10 transposition efficiently accommodates diverse transposon end configurations. *The EMBO Journal*, **15**, 5112–5122.
- Chaudhary, N. A., Flynn, H. W., Murray, T. G., Tabandeh, H., Mello, M. O., & Miller, D.** (1998). Emerging ciprofloxacin-resistant *Pseudomonas aeruginosa*. *American Journal of Ophthalmology*, **128**(4), 509–510.
- Chen, I., & Dubnau, D.** (2004). DNA uptake during bacterial transformation. *Nature Reviews. Microbiology*, **2**(3), 241–9.
- Chen, J., & Novick, R. P.** (2009). Phage-mediated intergeneric transfer of toxin genes. *Science*, **323**, 139–141.
- Coleman, N. V., & Holmes, A. J.** (2005). The native *Pseudomonas stutzeri* strain Q chromosomal integron can capture and express cassette-associated genes. *Microbiology*, **151**(6), 1853–1864.
- Colinon, C., Miriagou, V., Carattoli, A., Luzzaro, F., & Rossolini, G. M.** (2007). Characterization of the IncA/C plasmid pCC416 encoding VIM-4 and CMY-4 beta-lactamases. *The Journal of Antimicrobial Chemotherapy*, **60**(2), 258–62.

- Collis, C. M., Grammaticopoulos, G., Briton, J., Stokes, H. W., & Hall, R. M.** (1993). Site-specific insertion of gene cassettes into integrons. *Molecular Microbiology*, **9**(1), 41–52.
- Collis, C. M., & Hall, R. M.** (1992). Gene cassettes from the insert region of integrons are excised as covalently closed circles. *Molecular Microbiology*, **6**(19), 2875–2885.
- Collis, C. M., & Hall, R. M.** (1995). Expression of antibiotic resistance genes in the integrated cassettes of integrons. *Antimicrobial Agents and Chemotherapy*, **39**, 155–162.
- Collis, C. M., Kim, M. J., Stokes, H. W., & Hall, R. M.** (1998). Binding of the purified integron DNA integrase Int11 to integron- and cassette-associated recombination sites. *Molecular Microbiology*, **29**(2), 477–490.
- Collis, C. M., Kim, M.-J., Partridge, S. R., Stokes, H. W., & Hall, R. M.** (2002). Characterization of the class 3 integron and the site-specific recombination system it determines. *Journal of Bacteriology*, **184**(11), 3017–3026.
- Collis, C. M., Recchia, G. D., Kim, M., Hall, R. M., & Stokes, H. W.** (2001). Efficiency of recombination reactions catalyzed by class 1 integron integrase. *Journal of Bacteriology*, **183**(8), 2535–2542.
- Colomer-Lluch, M., Imamovic, L., Jofre, J., & Muniesa, M.** (2011). Bacteriophages carrying antibiotic resistance genes in fecal waste from cattle, pigs, and poultry. *Antimicrobial Agents and Chemotherapy*, **55**(10), 4908–4911.
- Conrad, R. S., Wulf, R. G., & Clay, D. L.** (1979). Effects of carbon sources on antibiotic resistance in *Pseudomonas aeruginosa*. *Antimicrobial Agents and Chemotherapy*, **15**(1), 59–66.
- Coscollá, M., Comas, I., & González-Candelas, F.** (2011). Quantifying nonvertical inheritance in the evolution of *Legionella pneumophila*. *Molecular Biology and Evolution*, **28**(2), 985–1001.
- D'Costa, V. M., King, C. E., Kalan, L., Morar, M., Sung, W. W., Schwarz, C., Forese, D., Zazula, G., Calmels, F., Debruyne, R., Golding, G. B., Poinar, H. N., & Wright, G. D.** (2011). Antibiotic resistance is ancient. *Nature*, **477**, 457–461.
- D'Costa, V. M., McGrann, K. M., Hughes, D. W., & Wright, G. D.** (2006). Sampling the antibiotic resistome. *Science*, **311**(5759), 374–377.
- da Costa, P. M., Loureiro, L., & Matos, A. J. F.** (2013). Transfer of multidrug-resistant bacteria between intermingled ecological niches: The interface between humans, animals and environment. *International Journal of Environmental Research and Public Health*, **10**(1), 278–294.
- da Silva Rabello, M. C., de Oliveira, R. S., Silva, R. M., & Leao, S. C.** (2010). Natural occurrence of horizontal transfer of *Mycobacterium avium*- specific insertion sequence IS1245 to *Mycobacterium kansasii*. *Journal of Clinical Microbiology*, **48**(6), 2257–2259.

- Datta, N., & Hedges, R. W.** (1972). Trimethoprim resistance conferred by W plasmids in *Enterobacteriaceae*. *Journal of General Microbiology*, **72**, 349–355.
- Davies, J., & Davies, D.** (2010). Origins and evolution of antibiotic resistance. *Microbiology and Molecular Biology Reviews*, **74**(3), 417–433.
- De Gelder, L., Ponciano, J. M., Joyce, P., & Top, E.** (2007). Stability of a promiscuous plasmid in different hosts: no guarantee for a long-term relationship. *Microbiology*, **153**(2), 452–463.
- Del Solar, G., Giraldo, R., Ruiz-echevarría, M. J., Espinosa, M., & Díaz-orejas, R.** (1998). Replication and control of circular bacterial plasmids. *Microbiology and Molecular Biology Reviews*, **62**(2), 434–464.
- Demaneche, S., Kay, E., Gourbiere, F., & Simonet, P.** (2001). Natural transformation of *Pseudomonas fluorescens* and *Agrobacterium tumefaciens* in Soil. *Applied and Environmental Microbiology*, **67**(6), 2617–2621.
- Demarre, G., Frumerie, C., Gopaul, D. N., & Mazel, D.** (2007). Identification of key structural determinants of the IntI1 integron integrase that influence *attC* x *attII* recombination efficiency. *Nucleic Acids Research*, **35**(19), 6475–6489.
- Diancourt, L., Passet, V., Nemeč, A., Dijkshoorn, L., & Brisse, S.** (2010). The population structure of *Acinetobacter baumannii*: expanding multiresistant clones from an ancestral susceptible genetic pool. *PloS one*, **5**(4), e10034.
- Di Pilato, V., Pollini, S., & Rossolini, G. M.** (2014). Characterization of plasmid pAX22, encoding VIM-1 metallo- β -lactamase, reveals a new putative mechanism of In70 integron mobilization. *Journal of Antimicrobial Chemotherapy*, **69**(1), 67–71.
- Dickey, T. H., Altschuler, S. E., & Wuttke, D. S.** (2013). Single-stranded DNA-binding proteins: multiple domains for multiple functions. *Structure*, **21**(7), 1074–1084.
- Didier, M.** (2006). Integrons: agents of bacterial evolution. *Nature Reviews Microbiology*, **4**(8), 608–620.
- Dobrindt, U., Hochhut, B., Hentschel, U., & Hacker, J.** (2004). Genomic islands in pathogenetic and environmental microorganisms. *Nature Reviews Microbiology*, **2**(5), 414–424.
- Douard, G., Praud, K., Cloeckert, A., & Doublet, B.** (2010). The *Salmonella* genomic island 1 is specifically mobilized in trans by the IncA/C multidrug resistance plasmid family. *PloS One*, **5**(12), e15302.
- Drago, L., De Vecchi, E., Nicola, L., Tocalli, L., & Gismondo, M. R.** (2005). *In vitro* selection of resistance in *Pseudomonas aeruginosa* and *Acinetobacter* spp. by levofloxacin and ciprofloxacin alone and in combination with beta-lactams and amikacin. *The Journal of Antimicrobial Chemotherapy*, **56**(2), 353–359.

- Du, D., Wang, Z., James, N. R., Voss, J. E., Klimont, E., Ohene-Agyei, T., Venter, H., Chiu, W., & Luisi, B. F.** (2014). Structure of the AcrAB–TolC multidrug efflux pump. *Nature*, **509**, 512–515.
- Duffin, P. M., & Seifert, H. S.** (2012). Genetic transformation of *Neisseria gonorrhoeae* shows a strand preference. *FEMS Microbiol Letter*, **334**(1), 44–48.
- Eisen, J.** (2000). Horizontal gene transfer among microbial genomes: new insights from complete genome analysis. *Current Opinion in Genetics & Development*, **10**(6), 606–611.
- Elsaied, H., Stokes, H. W., Kitamura, K., Kurusu, Y., Kamagata, Y., & Maruyama, A.** (2011). Marine integrons containing novel integrase genes, attachment sites, *attI*, and associated gene cassettes in polluted sediments from Suez and Tokyo Bays. *The ISME Journal*, **5**(7), 1162–1177.
- Essen-Zandbergen, A., Smith, H., Veldman, K., & Mevius, D.** (2007). Occurrence and characteristics of class 1, 2 and 3 integrons in *Escherichia coli*, *Salmonella* and *Campylobacter* spp. in the Netherlands. *Journal of Antimicrobial Chemotherapy*, **59**, 746–750.
- Firth, N., Ippen-ihler, K., & Skurray, R. A.** (1996). Structure and function of the F factor and mechanism of conjugation. In F. C. Neidhardt, R. Curtiss, J. L. Ingraham, E. C. C. Lin, K. B. Low, M. Boris, ... H. E. Umbarger (Eds.), *Escherichia coli and Salmonella Cellular and Molecular Biology* (2nd ed., pp. 2377–2401). Washington, D.C.: ASM Press.
- Fisher, J. F., Meroueh, S. O., & Mobashery, S.** (2005). Bacterial resistance to β -lactam antibiotics: compelling opportunism, compelling opportunity. *Chemical Reviews*, **105**, 395–424.
- Folkesson, A., Jelsbak, L., Yang, L., Johansen, H. K., Ciofu, O., Høiby, N., & Molin, S.** (2012). Adaptation of *Pseudomonas aeruginosa* to the cystic fibrosis airway: an evolutionary perspective. *Nature Reviews. Microbiology*, **10**(12), 841–851.
- Ford, P. J., & Avison, M. B.** (2004). Evolutionary mapping of the SHV beta-lactamase and evidence for two separate IS26-dependent *bla*_{SHV} mobilization events from the *Klebsiella pneumoniae* chromosome. *Journal of Antimicrobial Chemotherapy*, **54**(1), 69–75.
- Forsberg, K. J., Reyes, A., Wang, B., Selleck, E. M., Sommer, M. O. A., & Dantas, G.** (2012). The shared antibiotic resistome of soil bacteria and human pathogens. *Science*, **337**(6098), 1107–1111.
- Fosse, T., Giraud-Morin, C., Madinier, I., & Labia, R.** (2003). Sequence analysis and biochemical characterization of chromosomal CAV-1 (*Aeromonas caviae*), the parental cephalosporinase of plasmid-mediated AmpC “FOX” cluster. *FEMS Microbiology Letters*, **222**(1), 93–98.

- Foster, P. L.** (2000). Adaptive mutation in *Escherichia coli*. *Cold Spring Harbor Symposia on Quantitative Biology*, **65**, 21-29.
- Fournier, P.-E., Vallenet, D., Barbe, V., Audic, S., Ogata, H., Poirel, L., Richet, H., Robert, C., Mangenot, S., Abergel, C., Nordmann, P., Weissenbach, J., Raoult, D., & Claverie, J.-M.** (2006). Comparative genomics of multidrug resistance in *Acinetobacter baumannii*. *PLoS Genetics*, **2**(1), e7.
- Francia, M. V., Zabala, J. C., de la Cruz, F., & García Lobo, J. M.** (1999). The IntI1 integron integrase preferentially binds single-stranded DNA of the *attC* site. *Journal of Bacteriology*, **181**(21), 6844–6849.
- Friedrich, A., Hartsch, T., & Averhoff, B.** (2001). Natural transformation in mesophilic and thermophilic bacteria: identification and characterization of novel, closely related competence genes in *Acinetobacter* sp. strain BD413 and *Thermus thermophilus* HB27. *Applied and Environmental Microbiology*, **67**(7), 3140–3148.
- Frost, L. S., Leplae, R., Summers, A. O., & Toussaint, A.** (2005). Mobile genetic elements: the agents of open source evolution. *Nature Reviews Microbiology*, **3**, 722–732.
- Gallucci, K. K., & Paerl, H. W.** (1983). *Pseudomonas aeruginosa* chemotaxis associated with blooms of N₂-fixing blue-green algae (Cyanobacteria). *Applied Environmental Microbiology*, **45**(2), 557–562.
- Garner, M. M., & Rau, D. C.** (1995). Water release associated with specific binding of *gal* repressor. *The EMBO Journal*, **14**(6), 1257–1263.
- Gerhart, E., Wagner, H., & Simons, R. W.** (1994). Antisense RNA control in bacteria, phages, and plasmids. *Annual Reviews Microbiology*, **48**, 713–742.
- Gestal, A. M., Liew, E. F., & Coleman, N. V.** (2011). Natural transformation with synthetic gene cassettes: new tools for integron research and biotechnology. *Microbiology*, **157**(12), 3349–3360.
- Gibert, I., Barbé, J., & Casadesús, J.** (1990). Distribution of insertion sequence IS200 in *Salmonella* and *Shigella*. *Journal of General Microbiology*, **136**(12), 2555–2560.
- Gillings, M., Boucher, Y., Labbate, M., Holmes, A., Krishnan, S., Holley, M., & Stokes, H. W.** (2008). The evolution of class 1 integrons and the rise of antibiotic resistance. *Journal of Bacteriology*, **190**(14), 5095–5100.
- Gillings, M. R.** (2013). Evolutionary consequences of antibiotic use for the resistome, mobilome and microbial pangenome. *Frontiers in Microbiology*, **4**, 4.
- Gillings, M. R.** (2014). Integrons: past, present, and future. *Microbiology and Molecular Biology Reviews*, **78**(2), 257–277.

- Gillings, M. R., Xuejun, D., Hardwick, S. A., Holley, M. P., & Stokes, H. W.** (2009). Gene cassettes encoding resistance to quaternary ammonium compounds: a role in the origin of clinical class 1 integrons? *The ISME Journal*, **3**(2), 209–215.
- Ginn, A. N., Wiklendt, A. M., Gidding, H. F., George, N., O’Driscoll, J. S., Partridge, S. R., O’Toole, B. I., Perri, R. A., Faoagali, J., Gallagher, J. E., Lipman, J., & Iredell, J. R.** (2012). The ecology of antibiotic use in the ICU: homogeneous prescribing of cefepime but not tazocin selects for antibiotic resistant infection. *PLoS One*, **7**(6), e38719.
- González-Zorn, B., & Escudero, J. A.** (2012). Ecology of antimicrobial resistance: humans, animals, food and environment. *International Microbiology*, **15**, 101–109.
- Gopaul, D. N., Guo, F., & Van Duyne, G. D.** (1998). Structure of the Holliday junction intermediate in Cre-loxP site-specific recombination. *The EMBO Journal*, **17**(14), 4175–4187.
- Gravel, A., Fournier, B., & Roy, P. H.** (1998). DNA complexes obtained with the integron integrase IntI1 at the *attI1* site. *Nucleic Acids Research*, **26**(19), 4347–4355.
- Grindley, N. D. F., Whiteson, K. L., & Rice, P. A.** (2006). Mechanisms of site-specific recombination. *Annual Review of Biochemistry*, **75**, 567–605.
- Hacker, J., & Carniel, E.** (2001). Ecological fitness, genomic islands and bacterial pathogenicity. A Darwinian view of the evolution of microbes. *EMBO Reports*, **2**(5), 376–381.
- Hall, R. M., & Collis, C. M.** (1995). Mobile gene cassettes and integrons: capture and spread of genes by site-specific recombination. *Molecular Microbiology*, **15**(4), 593–600.
- Hamidian, M., & Hall, R. M.** (2011). AbaR4 replaces AbaR3 in a carbapenem-resistant *Acinetobacter baumannii* isolate belonging to global clone 1 from an Australian hospital. *Journal of Antimicrobial Chemotherapy*, **66**(11), 2484–2491.
- Hamilton, S. R., O’Donnell, J. B. J., Hammet, A., Stapleton, D., Habinowski, S. A., Means, A. R., Kemp, B. E. & Witters, L. A.** (2002). AMP-activated protein kinase: detection with recombinant AMPK alpha1 subunit. *Biochemical and Biophysical Research Communications*, **293**(3), 892–898.
- Hancock, R. E. W., & Speert, D. P.** (2000). Antibiotic resistance in *Pseudomonas aeruginosa*: mechanisms and impact on treatment. *Drug Resistance Updates*, **3**(4), 247–255.
- Harada, S., Ishii, Y., Saga, T., Tateda, K., & Yamaguchi, K.** (2010). Chromosomally encoded *bla*CMY-2 located on a novel SXT/R391-related integrating conjugative element in a *Proteus mirabilis* clinical isolate. *Antimicrobial Agents and Chemotherapy*, **54**(9), 3545–3550.

- Hedges, R. W., & Datta, N.** (1972). R factors conferring resistance to trimethoprim but not sulphonamides. *Journal of General Microbiology*, **73**, 573-575.
- Hefti, M. H., Caroline, J. G., der Toorn, V. V., Dixon, R., & Vervoort, J.** (2001). A novel purification method for histidine-tagged proteins containing a thrombin cleavage site. *Analytical Biochemistry*, **295**(2), 180-185.
- Heidelberg, J. F., Eisen, J. A., Nelson, W. C., Clayton, R. A., Gwinn, M. L., & Al., E.** (2000). DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*. *Nature*, **406**, 477-483.
- Hellman, L. M., & Fried, M. G.** (2007). Electrophoretic mobility shift assay (EMSA) for detecting protein-nucleic acid interactions. *Nature Protocols*, **2**, 1849 - 1861.
- Henderson, D. J., Lydiate, D. J., & Hopwood, D. A.** (1989). Structural and functional analysis of the minicircle, a transposable element of *Streptomyces coelicolor* A3(2). *Molecular Microbiology*, **3**, 1307-1318.
- Henkin, T. M.** (1996). Control of transcription termination in prokaryotes. *Annual Review of Biochemistry*, **30**, 35-57.
- Hibbing, M. E., Fuqua, C., Parsek, M. R., & Peterson, S. B.** (2010). Bacterial competition: surviving and thriving in the microbial jungle. *Nature Reviews Microbiology*, **8**(1), 15-25.
- Hilker, R., Munder, A., Klockgether, J., Losada, P. M., Chouvarine, P., Cramer, N., Davenport, C. F., Dethlefsen, S., Fischer, S., Peng, H., Schönfelder, T., Türk, O., Wiehlmann, L., Wölbeling, F., Gulbins, E., Goesmann, A., & Tümmler, B.** (2015). Interclonal gradient of virulence in the *Pseudomonas aeruginosa* pangenome from disease and environment. *Environmental Microbiology*, **17**(1), 29-46.
- Holmes, A. J., Gillings, M. R., Nield, B. S., Mabbutt, B. C., Helena, K. M., & Stokes, H. W.** (2003). The gene cassette metagenome is a basic resource for bacterial genome evolution. *Environmental Microbiology*, **5**(5), 383-394.
- Holmes, A. J., Holley, M. P., Mahon, A., Nield, B., Gillings, M., & Stokes, H. W.** (2003). Recombination activity of a distinctive integron-gene cassette system associated with *Pseudomonas stutzeri* populations in soil. *Journal of Bacteriology*, **185**(3), 918-928.
- Humeniuk, C., Arlet, G., Gautier, V., Labia, R., Philippon, A., & Grimont, P.** (2002). Beta-lactamases of *Kluyvera ascorbata*, probable progenitors of some plasmid-encoded CTX-M types. *Antimicrobial Agents and Chemotherapy*, **49**(9), 3045-3049.
- Iyer, L. M., Koonin, E. V., & Aravind, L.** (2004). Evolution of bacterial RNA polymerase: implications for large-scale bacterial phylogeny, domain accretion, and horizontal gene transfer. *Gene*, **335**, 73-88.
- Jacquier, H., Zaoui, C., Sanson-le Pors, M.-J., Mazel, D., & Berçot, B.** (2009). Translation regulation of integrons gene cassette expression by the *attC* sites. *Molecular Microbiology*, **72**(6), 1475-1486.

- Johansson, C., Kamali-Moghaddam, M., & Sundström, L.** (2004). Integron integrase binds to bulged hairpin DNA. *Nucleic Acids Research*, **32**(13), 4033–4043.
- 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**(2), 376–393.
- Kadlec, K., & Schwarz, S.** (2008). Analysis and distribution of class 1 and class 2 integrons and associated gene cassettes among *Escherichia coli* isolates from swine, horses, cats and dogs collected in the BfT-GermVet monitoring study. *The Journal of Antimicrobial Chemotherapy*, **62**(3), 469–473.
- Kallastu, A., Horak, R., & Kivisaar, M.** (1998). Identification and characterization of IS1411, a new insertion sequence which causes transcriptional activation of the phenol degradation genes in *Pseudomonas putida*. *Journal of Bacteriology*, **180**(20), 5306–5312.
- Kenna, D. T., Yesilkaya, H., Forbes, K. J., Barcus, V. A., Vandamme, P., & Govan, J. R. W.** (2006). Distribution and genomic location of active insertion sequences in the *Burkholderia cepacia* complex. *Journal of Medical Microbiology*, **55**, 1–10.
- Kleckner, N.** (1981). Transposable elements in prokaryotes. *Annual Review of Genetics*, **15**, 341–404.
- Klockgether, J., Reva, O., Larbig, K., & Tummeler, B.** (2004). Sequence analysis of the mobile genome island pKLC102 of *Pseudomonas aeruginosa* C. *Journal of Bacteriology*, **186**, 518–534.
- Koenig, J. E., Boucher, Y., Charlebois, R. L., Nesbo, C., Zhaxybayeva, O., Baptiste, E., Spencer, M., Joss, M. J., Stokes, H. W., & Doolittle, W. F.** (2008). Integron-associated gene cassettes in Halifax harbour: assessment of a mobile gene pool in marine sediments. *Environmental Microbiology*, **10**, 1024–1038.
- Koenig, J. E., Sharp, C., Dlutek, M., Curtis, B., Joss, M., Boucher, Y., & Doolittle, W. F.** (2009). Integron gene cassettes and degradation of compounds associated with industrial waste: the case of the Sydney tar ponds. *PLoS One*, **4**(4), e5276.
- Koskella, B., & Meaden, S.** (2013). Understanding bacteriophage specificity in natural microbial communities. *Viruses*, **5**(3), 806–823.
- Kovach, M. E., Elzer, P. H., Hill, D. S., Robertson, G. T., Farris, M. A., Roop, R. M., & Peterson, K. M.** (1995). Four new derivatives of the broad-host-range cloning vector pBBRMCS, carrying different antibiotic-resistance cassettes. *Gene*, **166**, 175–176.
- Krishna, M. P., Varghese, R., & Hatha, A. A. M.** (2012). Heavy metal tolerance and multiple drug resistance of heterotrophic bacterial isolates from metal contaminated soil. *The South Pacific Journal of Natural and Applied Sciences*, **30**(1), 58–64.

- Krauland, M. G., Marsh, J. W., Paterson, D. L., & Harrison, L. H.** (2009). Integron-mediated multidrug resistance in a global collection of nontyphoidal *Salmonella enterica* isolates. *Emerging Infectious Diseases*, **15**(3), 388–396.
- Krizova, L., & Nemeč, A.** (2010). A 63 kb genomic resistance island found in a multidrug-resistant *Acinetobacter baumannii* isolate of European clone I from 1977. *The Journal of Antimicrobial Chemotherapy*, **65**(9), 1915–1918.
- Kümmerer, K.** (2004). Resistance in the environment. *The Journal of Antimicrobial Chemotherapy*, **54**(2), 311–320.
- Kung, V. L., Ozer, E. A., & Hauser, A. R.** (2010). The accessory genome of *Pseudomonas aeruginosa*. *Microbiology and Molecular Biology Reviews*, **74**(4), 621–641.
- Labbate, M., Roy Chowdhury, P., & Stokes, H. W.** (2008). A class 1 integron present in a human commensal has a hybrid transposition module compared to Tn402: evidence of interaction with mobile DNA from natural environments. *Journal of Bacteriology*, **190**(15), 5318–5327.
- Lagatolla, C., Tonin, E. A., Monti-Bragadin, C., Dolzani, L., Gombac, F., Bearzi, C., Edalucci, E., Gionechetti, F., & Rossolini, G. M.** (2004). Endemic carbapenem-resistant *Pseudomonas aeruginosa* with acquired metallo-beta-lactamase determinants in European hospitals. *Emerging Infectious Diseases*, **10**(3), 3–6.
- Lalucat, J., Bennasar, A., Bosch, R., García-Valdés, E., & Palleroni, N. J.** (2006). Biology of *Pseudomonas stutzeri*. *Microbiology and Molecular Biology Reviews*, **70**(2), 510–547.
- Landy, A., & Ross, W.** (1977). Viral integration and excision: structure of the lambda *att* sites. *Science*, **197**, 1147–1160.
- Lapierre, P., & Gogarten, J. P.** (2009). Estimating the size of the bacterial pangenome. *Trends in Genetics*, **25**, 107–110.
- Larbig, K. D., Christmann, A., Johann, A., Klockgether, J., Hartsch, T., Merkl, R., Wehlmann, L., Fritz, H. J., & Tummler, B.** (2002). Gene islands integrated into tRNA(Gly) genes confer genome diversity on a *Pseudomonas aeruginosa* clone. *Journal of Bacteriology*, **184**, 6665–6680.
- Larouche, A., & Roy, P. H.** (2011). Effect of *attC* structure on cassette excision by integron integrases. *Mobile DNA*, **2**(1), 3.
- Lartigue, M., Poirel, L., Aubert, D., & Nordmann, P.** (2006). *In vitro* analysis of ISE $cplB$ -mediated mobilization of naturally occurring β -lactamase gene *bla*_{CTX-M} of *Kluyvera ascorbata*. *Antimicrobial Agents and Chemotherapy*, **50**(4), 1282–1286.
- Lee, K., Lim, J. B., Yum, J. H., Yong, D., Chong, Y., Kim, J. M., & Livermore, D. M.** (2002). *bla*_{VIM-2} cassette-containing novel integrons in metallo- β -lactamase-producing *Pseudomonas aeruginosa* and *Pseudomonas putida* isolates disseminated in a Korean hospital. *Antimicrobial Agents and Chemotherapy*, **46**(4), 1053–1058.

- Lee, L., & Sadowski, P. D.** (2005). Strand selection by the tyrosine recombinases. *Proceedings of the National Academy of Science of the United States of America*, **80**, 1–42.
- Lefebvre, M. D., & Valvano, M. A.** (2002). Construction and evaluation of plasmid vectors optimized for constitutive and regulated gene expression in *Burkholderia cepacia* complex isolates. *Applied and Environmental Microbiology*, **68**(12), 5956–5964.
- Lefebvre, T., & Stanhope, M. J.** (2007). Evolution of the core and pan-genome of *Streptococcus*: positive selection, recombination, and genome composition. *Genome Biology*, **8**, R71.
- Leverstein-van Hall, M. A., Box, A. T. A., Blok, H. E. M., Paauw, A., Fluit, A. C., & Verhoef, J.** (2002). Evidence of extensive interspecies transfer of integron-mediated antimicrobial resistance genes among multidrug-resistant *Enterobacteriaceae* in a clinical setting. *The Journal of Infectious Diseases*, **186**(1), 49–56.
- Levesque, C., Brassard, S., Lapointe, J., & Roy, P. H.** (1994). Diversity and relative strength of tandem promoters for the antibiotic-resistance genes of several integrons. *Gene*, **142**, 49–54.
- Lewis, L. A., & Grindley, N. D.** (1997). Two abundant intramolecular transposition products, resulting from reactions initiated at a single end, suggest that IS2 transposes by an unconventional pathway. *Molecular Microbiology*, **25**(3), 517–529.
- Lewis, L. A., Astatke, M., Umekubo, P. T., Alvi, S., Saby, R., Afrose, J., Oliveria, P. H., Monteiro, G. A., & Prazeres, D. M. F.** (2012). Protein-DNA interactions define the mechanistic aspects of circle formation and insertion reactions in IS2 transposition. *Mobile DNA*, **3**(1), 1.
- Lewis, L. A., Cylin, E., Lee, H. K., Saby, R., Wong, W., & Grindley, N. D. F.** (2004). The left end of IS2: a compromise between transpositional activity and an essential promoter function that regulates the transposition pathway. *Journal of Bacteriology*, **186**(3), 858–865.
- Liebert, C. A., Hall, R. M., & Summer, A. O.** (1999). Transposon Tn21, flagship of the floating genome. *Microbiology and Molecular Biology Reviews*, **63**(3), 507–522.
- Lister, P. D., Wolter, D. J., & Hanson, N. D.** (2009). Antibacterial-resistant *Pseudomonas aeruginosa*: clinical impact and complex regulation of chromosomally encoded resistance mechanisms. *Clinical Microbiology Reviews*, **22**(4), 582–610.
- Livermore, D. M.** (2002). Multiple mechanisms of antimicrobial resistance in *Pseudomonas aeruginosa*: our worst nightmare? *Clinical Infectious Diseases*, **34**(5), 634–640.
- Livermore, D. M., & Woodford, N.** (2006). The beta-lactamase threat in *Enterobacteriaceae*, *Pseudomonas* and *Acinetobacter*. *Trends in Microbiology*, **14**(9), 413–420.

- Lombardi, G., Luzzaro, F., Riccio, M. L., Perilli, M., Coli, A., Amicosante, G., Rossolini, G. M., & Toniolo, A.** (2002). Nosocomial infections caused by multidrug-resistant isolates of *Pseudomonas putida* producing VIM-2 and VIM-4 metallo- β -lactamase. *The Journal of Antimicrobial Chemotherapy*, **61**(3), 749-751.
- Loot, C., Bikard, D., Rachlin, A., & Mazel, D.** (2010). Cellular pathways controlling integron cassette site folding. *The EMBO Journal*, **29**(15), 2623–2634.
- Loot, C., Ducos-Galand, M., Escudero, J. A., Bouvier, M., & Mazel, D.** (2012). Replicative resolution of integron cassette insertion. *Nucleic Acids Research*, **40**(17), 8361–8370.
- Loot, C., Parissi, V., Escudero, J. A., Amarir-Bouhram, J., Bikard, D., & Mazel, D.** (2013). The integron integrase efficiently prevents the melting effect of *Escherichia coli* SSB protein on folded *attC* sites. *Journal of Bacteriology*, **196**(4), 762-771.
- Louws, F. J., Fulbright, D. W., Stephens, C. T., & de Bruijn, F. J.** (1994). Specific genomic fingerprints of phytopathogenic *Xanthomonas* and *Pseudomonas* pathovars and strains generated with repetitive sequences and PCR. *Applied and Environmental Microbiology*, **60**(7), 2286-2295.
- Lukjancenko, O., Wassenaar, T. M., & Ussery, D. W.** (2010). Comparison of 61 sequenced *Escherichia coli* genomes. *Microbial Ecology*, **60**(4), 708–720.
- MacDonald, D., Demarre, G., Bouvier, M., Mazel, D., & Gopaul, D. N.** (2006). Structural basis for broad DNA-specificity in integron recombination. *Nature*, **440**(7088), 1157–1162.
- Mahillon, J., & Chandler, M.** (1998). Insertion Sequences. *Microbiology and Molecular Biology Reviews*, **62**(3), 725–744.
- Maravić, A., Skočibušić, M., Samanić, I., Fredotović, Z., Cvjetan, S., Jutronić, M., & Puizina, J.** (2013). *Aeromonas* spp. simultaneously harbouring *bla*(CTX-M-15), *bla*(SHV-12), *bla*(PER-1) and *bla*(FOX-2), in wild-growing Mediterranean mussel (*Mytilus galloprovincialis*) from Adriatic Sea, Croatia. *International Journal of Food Microbiology*, **166**(2), 301–308.
- Marceau, A. H.** (2012). Functions of single-stranded DNA-binding proteins in DNA replication, recombination, and repair. In J. M. Walker (Ed.), *Single-Stranded DNA Binding Proteins. Methods and Protocols* (1–21). Springer New York.
- Markou, P., & Apidianakis, Y.** (2014). Pathogenesis of intestinal *Pseudomonas aeruginosa* infection in patients with cancer. *Frontiers in Cellular and Infection Microbiology*, **3**, 115.
- Martinez, E., Marquez, C., Ingold, A., Merlino, J., Djordjevic, S. P., Stokes, H. W., & Chowdhury, P. R.** (2012). Diverse mobilized class 1 integrons are common in the chromosomes of pathogenic *Pseudomonas aeruginosa* clinical isolates. *Antimicrobial Agents and Chemotherapy*, **56**(4), 2169–2172.

- Martinez, E., Pérez, J. E., Márquez, C., Vilacoba, E., Centrón, D., Leal, A. L., Saavedra, C., Tovar, C., Vanegas, N., & Stokes, H. W.** (2013). Emerging and existing mechanisms co-operate in generating diverse β -lactam resistance phenotypes in geographically dispersed and genetically disparate *Pseudomonas aeruginosa* strains. *Journal of Global Antimicrobial Resistance*, **1**(3), 135–142.
- Massova, I., & Mobashery, S.** (1998). Kinship and diversification of bacterial penicillin-binding proteins and β -lactamases. *Antimicrobial Agents and Chemotherapy*, **42**, 1–17.
- Matic, I., Rayssiguier, C., & Radman, M.** (1995). Interspecies gene exchange in bacteria: the role of SOS and mismatch repair systems in evolution of species. *Cell*, **80**(3), 507–515.
- Matic, I., Taddei, F., & Radman, M.** (1996). Genetic barriers among bacteria. *Trends in Microbiology*, **4**(2), 69–72.
- Mazaheri Nezhad Fard, R., Barton, M., & Heuzenroeder, M.** (2011). Bacteriophage-mediated transduction of antibiotic resistance in enterococci. *Letters in Applied Microbiology*, **52**, 559–564.
- Mazel, D.** (2004). Integrons and the origin of antibiotic resistance gene cassettes. *ASM News*, **70**(11), 520–525.
- Mazel, D., Dychinco, B., Webb, V. A., & Davies, J.** (1998). A distinctive class of integron in the *Vibrio cholerae* genome. *Science*, **280**, 605–608.
- McNicol, K.** (2002) In *Department of Biological Sciences*. Macquarie University, Sydney, pp.36
- Medini, D., Donati, C., Tettelin, H., Masignani, V., & Rappuoli, R.** (2005). The microbial pan-genome. *Current Opinion in Genetics & Development*, **15**(6), 589–594.
- van der Meer, J. R., Zehnder, A. J., & de Vos, W. M.** (1991). Identification of a novel composite transposable element, Tn5280, carrying chlorobenzene dioxygenase genes of *Pseudomonas* sp. strain P51. *Journal of Bacteriology*, **173**(22), 7077–7083.
- Mena, A., Smith, E. E., Burns, J. L., Speert, D. P., Moskowitz, S. M., Perez, J. L., & Oliver, A.** (2008). Genetic adaptation of *Pseudomonas aeruginosa* to the airways of cystic fibrosis patients is catalyzed by hypermutation. *Journal of Bacteriology*, **190**(24), 7910–7917.
- Messier, N., & Roy, P. H.** (2001). Integron integrases possess a unique additional domain necessary for activity. *Journal of Bacteriology*, **183**(22), 6699–6706.
- Michael, C. A., Gillings, M. R., Holmes, A. J., Hughes, L., Andrew, N. R., Holley, M. P., & Stokes, H. W.** (2004). Mobile gene cassettes: a fundamental resource for bacterial evolution. *American Naturalist*, **164**, 1–12.

- Mikkelsen, H., McMullan, R., & Filloux, A.** (2011). The *Pseudomonas aeruginosa* reference strain PA14 displays increased virulence due to a mutation in *ladS*. *PloS One*, **6**(12), e29113.
- Mills, J. A., Venkatesan, M. M., Baron, L. S., & Buysse, J. M.** (1992). Spontaneous insertion of an *IS1*-like element into the *virF* gene is responsible for avirulence in opaque colonial variants of *Shigella flexneri* 2a. *Infection and Immunity*, **60**, 175–182.
- Minakhina, S., Kholodii, G., Mindlin, S., Yurieva, O., & Nikiforov, V. T.** (1999). Tn5053 family transposons are *res* site hunters sensing plasmidal *res* sites occupied by cognate resolvases. *Molecular Microbiology*, **33**, 1059–1068.
- Mira, A., Martín-Cuadrado, A. B., Auria, G. D., & Rodríguez-Valera, F.** (2010). The bacterial pan-genome: a new paradigm in microbiology. *International Microbiology*, **13**, 45–57.
- Moura, A., Henriques, I., Ribeiro, R., & Correia, A.** (2007). Prevalence and characterization of integrons from bacteria isolated from a slaughterhouse wastewater treatment plant. *The Journal of Antimicrobial Chemotherapy*, **60**(6), 1243–1250.
- Nadjar, D., Rouveau, M., Verdet, C., Donay, L., Herrmann, J., Lagrange, P. H., Philippon, A., & Arlet, G.** (2000). Outbreak of *Klebsiella pneumoniae* producing transferable AmpC-type beta-lactamase (ACC-1) originating from *Hafnia alvei*. *FEMS Microbiology Letters*, **187**(1), 35–40.
- Nemec, A., Dolzani, L., Brisse, S., van den Broek, P., & Dijkshoorn, L.** (2004). Diversity of aminoglycoside-resistance genes and their association with class 1 integrons among strains of pan-European *Acinetobacter baumannii* clones. *Journal of Medical Microbiology*, **53**(12), 1233–1240.
- Nemergut, D. R., Robeson, M. S., Kysela, R. F., Martin, A. P., Schmidt, S. K., & Knight, R.** (2008). Insights and inferences about integron evolution from genomic data. *BMC Genomics*, **9**, 261.
- Nield, B. S., Willows, R. D., Torda, A. E., Gillings, M. R., Holmes, A. J., Nevalainen, K. M. H., Stokes, H. W., & Mabbutt, B. C.** (2004). New enzymes from environmental cassette arrays: functional attributes of a phosphotransferase and an RNA-methyltransferase. *Protein Science*, **13**, 1651–1659.
- Nigro, S. J., & Hall, R. M.** (2012). Tn6167, an antibiotic resistance island in an Australian carbapenem-resistant *Acinetobacter baumannii* GC2, ST92 isolate. *The Journal of Antimicrobial Chemotherapy*, **67**(6), 1342–1346.
- Nikokar, I., Tishayar, A., Flakiyan, Z., Alijani, K., Rehana-Banisaeed, S., Hosseinpour, M., Amir-Alvaei, S., & Araghian, A.** (2013). Antibiotic resistance and frequency of class 1 integrons among *Pseudomonas aeruginosa*, isolated from burn patients in Guilan, Iran. *Iranian Journal of Microbiology*, **5**(1), 36–41.

- Nilsson, A. I., Berg, O. G., Aspevall, O., Kahlmeter, G., & Andersson, D. I.** (2003). Biological costs and mechanisms of fosfomycin resistance in *Escherichia coli*. *Antimicrobial Agents and Chemotherapy*, **47**(9), 2850–2858.
- Nowak, M., Olszewski, M., Śpibida, M., & Kur, J.** (2014). Characterization of single-stranded DNA-binding proteins from the psychrophilic bacteria *Desulfotalea psychrophila*, *Flavobacterium psychrophilum*, *Psychrobacter arcticus*, *Psychrobacter cryohalolentis*, *Psychromonas ingrahamii*, *Psychroflexus torquis*, and Pho. *BMC Microbiology*, **14**, 91.
- Odumosu, B. T., Adeniyi, B. A., & Chandra, R.** (2013). Analysis of integrons and associated gene cassettes in clinical isolates of multidrug resistant *Pseudomonas aeruginosa* from Southwest Nigeria. *Annals of Clinical Microbiology and Antimicrobials*, **12**(1), 29.
- Olendzenski, L., Zhaxybayeva, O., & Gogarten, J. P.** (2002). Horizontal gene transfer. In M. Syvanen & C. I. Kado (Eds.), (pp. 427–435). New York: Academic.
- Olson, A. B., Silverman, M., Boyd, D. A., McGeer, A., Willey, B. M., Daneman, N., & Mulvey, M. R.** (2005). Identification of a progenitor of the CTX-M-9 group of extended-spectrum β -lactamases from *Kluyvera georgiana* isolated in Guyana. *Antimicrobial Agents and Chemotherapy*, **49**(5), 2112–2115.
- Ozer, E. A., Allen, J. P., & Hauser, A. R.** (2014). Characterization of the core and accessory genomes of *Pseudomonas aeruginosa* using bioinformatic tools Spine and AGEnt. *BMC Genomics*, **15**(1), 737.
- Partridge, S. R.** (2007). Genetic environment of *ISEcp1* and *bla_{ACC-1}*. *Antimicrobial Agents and Chemotherapy*, **51**(7), 2658–2659.
- Partridge, S. R.** (2011). Analysis of antibiotic resistance regions in gram-negative bacteria. *FEMS Microbiology Reviews*, **35**(5), 820–855.
- Partridge, S. R., & Hall, R. M.** (2003). The *IS1111* family members *IS4321* and *IS5075* have subterminal inverted repeats and target the terminal inverted repeats of *Tn21* family transposons. *Journal of Bacteriology*, **185**(21), 6371–6384.
- Partridge, S. R., & Hall, R. M.** (2005). Evolution of transposons containing *bla_{TEM}* genes. *Antimicrobial Agents and Chemotherapy*, **49**(3), 1267–1269.
- Partridge, S. R., Recchia, G. D., Scaramuzzi, C., Collis, C. M., Stokes, H. W., & Hall, R. M.** (2000). Definition of the *attI1* site of class 1 integrons. *Microbiology*, **146**(1), 2855–2864.
- Partridge, S. R., Tsafnat, G., Coiera, E., & Iredell, J. R.** (2009). Gene cassettes and cassette arrays in mobile resistance integrons. *FEMS Microbiology Reviews*, **33**(4), 757–784.

- Pembroke, J. T., MacMahon, C., & McGrath, B.** (2002). The role of conjugative transposons in the *Enterobacteriaceae*. *Cellular and Molecular Life Sciences*, **59**(12), 2055–2064.
- Perry, J. A., & Wright, G. D.** (2013). The antibiotic resistance “mobilome”: searching for the link between environment and clinic. *Frontiers in Microbiology*, **4**, 138.
- Peters, J. E., & Craig, N. L.** (2001). Tn7: smarter than we thought. *Nature Reviews Molecular Cell Biology*, **2**, 806–814.
- Petrovski, S., & Stanisich, V. A.** (2010). Tn502 and Tn512 are *res* site hunters that provide evidence of resolvase-independent transposition to random sites. *Journal of Bacteriology*, **192**, 1865–1874.
- Pinto, U. M., Pappas, K. M., & Winans, S. C.** (2012). The ABCs of plasmid replication and segregation. *Nature Reviews Microbiology*, **10**(11), 755–765.
- Planquette, B., Timsit, J.-F., Misset, B. Y., Schwebel, C., Azoulay, E., Adrie, C., Vesin, A., Jamali, S., Zahar, J. R., Allaouchiche, B., Souweine, B., Darmon, M, Dumenil, A. S., Goldgran-Toledano, D., Mourvillier, B. H., Bédos, J. P., & OUTCOMEREA Study Group** (2013). *Pseudomonas aeruginosa* ventilator-associated pneumonia. Predictive factors of treatment failure. *American Journal of Respiratory and Critical Care Medicine*, **188**(1), 69–76.
- Podmore, A. H., & Reynolds, P. E.** (2002). Purification and characterization of VanXYC, a D,D-dipeptidase/D,D-carboxypeptidase in vancomycin-resistant *Enterococcus gallinarum* BM4175. *European Journal of Biochemistry*, **269**(11), 2740–2746.
- Poirel, L., Brinas, L., Fortineau, N., & Nordmann, P.** (2005). Integron-encoded GES-type extended-spectrum β -lactamase with increased activity toward Aztreonam in *Pseudomonas aeruginosa*. *Antimicrobial Agents and Chemotherapy*, **49**(8), 3593–3597.
- Poirel, L., Brinas, L., Verlinde, A., Ide, L., & Nordmann, P.** (2005). BEL-1 , a novel clavulanic acid-inhibited extended-spectrum β -lactamase, and the class 1 integron In120 in *Pseudomonas aeruginosa*. *Antimicrobial Agents and Chemotherapy*, **49**(9), 3743–3748.
- Poirel, L., Kämpfer, P., & Nordmann, P.** (2002). Chromosome-encoded ambler class A β -lactamase of *Kluyvera georgiana*, a probable progenitor of a subgroup of CTX-M extended-spectrum β -lactamases. *Antimicrobial Agents and Chemotherapy*, **46**(12), 4038–4040.
- Poirel, L., Naas, T., Nicolas, D., Collet, L., Bellais, S., Cavallo, J., & Nordmann, P.** (2000). Characterization of VIM-2, a carbapenem-hydrolyzing metallo-beta-lactamase and its plasmid- and integron-borne gene from a *Pseudomonas aeruginosa* clinical isolate in France. *Antimicrobial Agents and Chemotherapy*, **44**(4), 891–897.
- Poirel, L., Todriguez-Martinez, J.-M., Mammeri, H., Liard, A., & Nordmann, P.** (2005). Origin of plasmid-mediated quinolone resistance determinant QnrA. *Antimicrobial Agents and Chemotherapy*, **49**(8), 3523–3525.

- Popowska, M., & Krawczyk-Balska, A.** (2013). Broad-host-range IncP-1 plasmids and their resistance potential. *Frontiers in Microbiology*, **4**, 44.
- Post, V., & Hall, R. M.** (2009). Insertion sequences in the IS1111 family that target the *attC* recombination sites of integron-associated gene cassettes. *FEMS Microbiology Letters*, **290**(2), 182–187.
- Prosseda, G., Latella, M. C., Casalino, M., Nicoletti, M., Michienzi, S., & Colonna, B.** (2006). Plasticity of the P *junc* promoter of ISEc11, a new insertion sequence of the IS1111 family. *Journal of Bacteriology*, **188**(13), 4681–4689.
- Ramírez, M. S., Piñeiro, S., & Centrón, D.** (2010). Novel insights about class 2 integrons from experimental and genomic epidemiology. *Antimicrobial Agents and Chemotherapy*, **54**(2), 699–706.
- Ramos-gonzález, M. I., Campos, M. J., Ramos, L., Espinosa-urgel, M., & Ramos, J. L.** (2006). Characterization of the *Pseudomonas putida* mobile genetic element IS*Ppu10*: an occupant of repetitive extragenic palindromic sequences. *Journal of Bacteriology*, **188**(1), 37-44.
- Rayssiguier, C., Thaler, D. S., & Radman, M.** (1989). The barrier to recombination between *Escherichia coli* and *Salmonella typhimurium* is disrupted in mismatch-repair mutants. *Nature*, **342**, 396–342.
- Read, T. D., & Ussery, D. W.** (2006). Opening the pan-genomics box. *Current Opinion in Microbiology*, **9**, 496–498.
- Recchia, G. D., & Hall, R. M.** (1995). Gene cassettes: a new class of mobile element. *Microbiology*, **141**(12), 3015–3027.
- Recchia, G. D., & Sherratt, D. J.** (2002). Chapter 9- Gene acquisition in bacteria by integron-mediated site-specific recombination. In: Craig, N. L., Craigie, R., Gellert, M., Lambowitz, A. M., eds. *Mobile DNA II*. Washington, DC: ASM Press. 162-176.
- Riou, M., Carbonnelle, S., Avrain, L., Mesaros, N., Pirnay, J.-P., Bilocq, F., De Vos, D., Simon, A., Pierard, D., Jacobs, F., Dediste, A., Tulkens, P. M., Van Bambeke, F., & Glupczynski, Y.** (2010). *In vivo* development of antimicrobial resistance in *Pseudomonas aeruginosa* strains isolated from the lower respiratory tract of Intensive Care Unit patients with nosocomial pneumonia and receiving antipseudomonal therapy. *International Journal of Antimicrobial Agents*, **36**(6), 513–522.
- Rius, N., Fuste, M. C., Guasp, C., Lalucat, J., & Loren, J. G.** (2001). Clonal population structure of *Pseudomonas stutzeri*, a species with exceptiona genetic diversity. *Journal of Bacteriology*, **183**, 736-744.
- Robinson, A., Wu, P. S.-C., Harrop, S. J., Schaeffer, P. M., Dosztányi, Z., Gillings, M. R., Holmes, A. J., Nevalainen, K. M. H., Stokes, H. W., Gottfried, O., Dixon, N. E., Curmi, P. M. G., & Mabbutt, B. C.** (2005). Integron-associated mobile gene cassettes code for folded proteins: the structure of Bal32a, a new member of the adaptable a+ β barrel family. *Journal of Molecular Biology*, **346**(5), 1229-1241.

- Rodríguez, M. M., Power, P., Radice, M., Vay, C., Famiglietti, A., Galleni, M., Ayala, J. A., & Gutkind, G.** (2004). Chromosome-encoded CTX-M-3 from *Kluyvera ascorbata*: a possible origin of plasmid-borne CTX-M-1-Derived Cefotaximases. *Antimicrobial Agents and Chemotherapy*, **48**(12), 4895–4897.
- Rodríguez, M. M., Power, P., Sader, H., Galleni, M., & Gutkind, G.** (2010). Novel chromosome-encoded CTX-M-78 beta-lactamase from a *Kluyvera georgiana* clinical isolate as a putative origin of CTX-M-25 subgroup. *Antimicrobial Agents and Chemotherapy*, **54**(7), 3070–3071.
- Rosenberg, S. M.** (2001). Evolving responsively: adaptive mutation. *Nature Reviews Genetics*, **2**, 504-515.
- Rosewarne, C. P., Pettigrove, V., Stokes, H. W., & Parsons, Y. M.** (2010). Class 1 integrons in benthic bacterial communities: abundance, association with Tn402-like transposition modules and evidence for coselection with heavy-metal resistance. *FEMS Microbiology Ecology*, **72**(1), 35–46.
- Rottman, M., Benzerara, Y., Hanau-Berçot, B., Bizet, C., Philippon, A., & Arlet, G.** (2002). Chromosomal *ampC* genes in *Enterobacter* species other than *Enterobacter cloacae*, and ancestral association of the ACT-1 plasmid-encoded cephalosporinase to *Enterobacter asburiae*. *FEMS Microbiol Letter*, **210**(1), 87–92.
- Rowe-Magnus, D. A., Guerout, A.-M., & Mazel, D.** (2002). Bacterial resistance evolution by recruitment of superintegron gene cassettes. *Molecular Microbiology*, **43**(6), 1657–1669.
- Sajjad, A., Holley, M. P., Labbate, M., Stokes, H. W., & Gillings, M. R.** (2011). Preclinical class 1 integron with a complete Tn402-like transposition module. *Applied and Environmental Microbiology*, **77**(1), 335–337.
- Sambrook, J., & Russell, D. W.** (2001). Molecular cloning: a laboratory manual. *Cold Spring Harbor Laboratory Press*, New York. 3rd ed.
- Sambrook, J., & Russell, D. W.** (2006). Purification of PCR Products in preparation for cloning. *Cold Spring Harbor Protocols*, New York **1**, Ch8.25.
- Santos, C., Caetano, T., Ferreira, S., & Mendo, S.** (2010). Tn5090-like class 1 integron carrying *bla*(VIM-2) in a *Pseudomonas putida* strain from Portugal. *Clinical Microbiology and Infection*, **16**(10), 1558–1561.
- Silby, M. W., Winstanley, C., Godfrey, S. A. C., Levy, S. B., & Jackson, R. W.** (2011). *Pseudomonas* genomes: diverse and adaptable. *FEMS Microbiology Reviews*, **35**(4), 652–680.
- Soki, J.** (2013). Extended role of insertion sequence elements in the antibiotic resistance of *Bacteriodes*. *World Journal of Clinical Infectious Diseases*, **3**(1), 1–12.

- Sokol, P. A., Luan, M. Z., Storey, D. G., & Thirukkumaran, P.** (1994). Genetic rearrangement associated with in vivo mucoid conversion of *Pseudomonas aeruginosa* PAO is due to insertion elements. *Journal of Bacteriology*, **176**(3), 553–562.
- Sonnevend, A., Ghazawi, A., Yahfoufi, N., Al-Baloushi, A., Hashmey, R., Mathew, M., Tariq, W. Z., & Pal, T.** (2012). VIM-4 carbapenemase-producing *Enterobacter cloacae* in the United Arab Emirates. *Clinical Infectious Diseases*, **18**(12), 494–496.
- Sousa, A., Bourgard, C., Wahl, L. M., & Gordo, I.** (2013). Rates of transposition in *Escherichia coli*. *Biology Letter*, **9**, 1-5.
- Southern, E. M.** (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *Journal of Molecular Biology*, **98**, 503-517.
- Spiers, A. J., Buckling, A., & Rainey, P. B.** (2000). The causes of *Pseudomonas* diversity. *Microbiology*, **146**, 2345–2350.
- Spilker, T., Coenye, T., Vandamme, P., Lipuma, J., & Lipuma, J. J.** (2004). PCR-based assay for differentiation of *Pseudomonas aeruginosa* from other *Pseudomonas* species recovered from cystic fibrosis patients. *Journal of Clinical Microbiology*, **42**(5), 2074-2079.
- Stalder, T., Barraud, O., Casellas, M., Dagot, C., & Ploy, M.-C.** (2012). Integron involvement in environmental spread of antibiotic resistance. *Frontiers in Microbiology*, **3**(April), 119.
- Stanier, R. Y., Palleroni, N. J., & Doudoroff, M.** (1966). The aerobic Pseudomonads: a taxonomic study. *Journal of General Microbiology*, **43**, 159-271.
- Stanisich, V. A., Arwas, R., Bennett, P. M., & de la Cruz, F.** (1989). Characterization of *Pseudomonas* mercury-resistance transposon Tn502, which has a preferred insertion site in RP1. *Journal of General Microbiology*, **135**, 2909–2915.
- Stark, W. M., Boocock, M. R., & Sherratt, D. J.** (1992). Catalysis by site-specific recombinases. *Trends in Genetics*, **8**, 432–439.
- Sternberg, N., Hamilton, D., Austin, S., Yarmolinsky, M., & Hoess, R.** (1981). Site-specific recombination and its role in the life cycle of bacteriophage P1. *Cold Spring Harbor Symposia on Quantitative Biology*, **1**, 297-309.
- Stokes, H. W., Holmes, A. J., Nield, B. S., Marita, P., Nevalainen, K. M. H., Mabbutt, B. C., & Gillings, M. R.** (2001). Gene cassette PCR : sequence-independent recovery of entire genes from environmental DNA. *Applied and Environmental Microbiology*, **67**(11), 5240-5246.
- Stokes, H. W., O’Gorman, D. B., Recchia, G. D., Parsekhian, M., & Hall, R. M.** (1997). Structure and function of 59-base element recombination sites associated with mobile gene cassettes. *Molecular Microbiology*, **26**(4), 731–45.

- Stover, C. K., Pham, X. Q., Erwin, A. L., Mizoguchi, S. D., Warrenner, P., Hickey, M. J., Brinkman, F. S., Hufnagle, W. O., Kowalik, D. J., Lagrou, M., Garber, R. L., Tolenino, E., Westbrock-Wadman, S., Yuan, Y., Brody, L. L., Coulter, S. N., Folger, K. R., Kas, K., Lim, R., Smith, K., Spencer, D., Wong, G. K., Wu, Z., Paulsen, I. T., Reizer, J., Saier, M. H., Hancock, R. E., Lory, S., & Olson, M. V. (2000). Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic pathogen. *Nature*, **406**(6799), 959–964.
- Strateva, T., & Yordanov, D. (2009). *Pseudomonas aeruginosa* - a phenomenon of bacterial resistance. *Journal of Medical Microbiology*, **58**(9), 1133–1148.
- Sullivant, D. J. O., & Gara, F. O. (1992). Traits of fluorescent *Pseudomonas* spp. involved in suppression of plant root pathogens. *Microbiological Reviews*, **56**(4), 662–676.
- Sundström, L., Rådström, P., Swedberg, G., & Sköld, O. (1988). Site-specific recombination promotes linkage between trimethoprim- and sulfonamide resistance genes. Sequence characterization of *dhfrV* and *sulI* and a recombination active locus of Tn21. *Molecular & General Genetics*, **322**(2-3), 191–201.
- Svara, F., & Rankin, D. J. (2011). The evolution of plasmid-carried antibiotic resistance. *BMC Evolutionary Biology*, **11**(1), 130.
- Szpirer, C. Y., Faelen, M., & Couturier, M. (2001). Mobilization function of the pBHR1 plasmid, a derivative of the broad-host-range plasmid pBBR1. *Journal of Bacteriology*, **183**(6), 2101–2110.
- Tamber, S., Ochs, M. M., & Robert, E. W. (2006). Role of the novel OprD family of porins in nutrient uptake in *Pseudomonas aeruginosa*. *Journal of Bacteriology*, **188**(1), 45–54.
- Tamminen, M., Virta, M., Fani, R., & Fondi, M. (2012). Large-scale analysis of plasmid relationships through gene-sharing networks. *Molecular Biology and Evolution*, **29**(4), 1225–1240.
- Tegova, R., Tover, A., Tarassova, K., Tark, M., & Kivisaar, M. (2004). Involvement of error-prone DNA polymerase IV in stationary-phase mutagenesis in *Pseudomonas putida*. *Journal of Bacteriology*, **186**(9), 2735–2744.
- Tenaillon, O., Taddei, F., Radman, M., & Matic, I. (2001). Second-order selection in bacterial evolution: selection acting on mutation and recombination rates in the course of adaptation. *Research in Microbiology*, **152**, 11–16.
- Tennstedt, T., Szczepanowski, R., Braun, S., Pühler, A., & Schlüter, A. (2003). Occurrence of integron-associated resistance gene cassette located on antibiotic resistance plasmids isolated from a wastewater treatment plant. *FEMS Microbiology Ecology*, **45**, 239–252.
- Terpe, K. (2003). Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. *Applied Microbiology and Biotechnology*, **60**(5), 523–533.

- Tetu, S. G.** (2007). Gene cassettes as an evolutionary resource for *Pseudomonas* species. PhD Thesis, School of Molecular and Microbial Bioscience, The University of Sydney.
- Tetu, S. G., & Holmes, A. J.** (2008). A family of insertion sequences that impacts integrons by specific targeting of gene cassette recombination sites, the IS1111-attC group. *Journal of Bacteriology*, **190**(14), 4959–4970.
- Thomas, C. M., & Nielsen, K. M.** (2005). Mechanisms of, and barriers to, horizontal gene transfer between bacteria. *Nature Reviews Microbiology*, **3**(9), 711–721.
- Tobiason, D. M., Buchner, J. M., Thiel, W. H., Gernert, K. M., & Karls, A. C. G.** (2001). Conserved amino acid motifs from the novel Piv/MooV family of transposases and site-specific recombinases are required for catalysis of DNA inversion by Piv. *Molecular Microbiology*, **39**(3), 641–651.
- Toleman, M. A., Vinodh, H., Sekar, U., Kamat, V., & Walsh, T. R.** (2007). *bla*_{VIM-2}-harboring integrons isolated in India, Russia, and the United State arise from an ancestral class 1 integron predating the formation of the 3' Conserved Sequence. *Antimicrobial Agents and Chemotherapy*, **51**(7), 2636–2638.
- Tribble, G. D., Parker, A. C., & Smith, C. J.** (1997). The Bacteroides mobilizable transposon Tn4555 integrates by a site-specific recombination mechanism similar to that of the gram-positive bacterial element Tn916. *Journal of Bacteriology*, **179**(8), 2731–2739.
- Vaisvila, R., Morgan, R. D., Posfai, J., & Raleigh, E. A.** (2001). Discovery and distribution of super-integrons among Pseudomonads. *Molecular Microbiology*, **42**(3), 587–601.
- Vandelannoote, K., Jordaens, K., Bomans, P., Leirs, H., Durnez, L., Affolabi, D., Sopoh, D., Aquiar, J., Phanzu, D. M., Kibadi, K., Eyangoh, S., Manou, L. B., Phillips, R. O., Adjei, O., Ablordey, A., Rigouts, L., Portaels, F., Eddyani, M., & de Jong B. C.** (2014). Insertion sequence element single nucleotide polymorphism typing provides insights into the population structure and evolution of *Mycobacterium ulcerans* across Africa. *Applied and Environmental Microbiology*, **80**(3), 1197–1209.
- VanValen, L.** (1973). A New Evolutionary Law. *Evolutionary Theory*, **1**, 1–30.
- Venkatesan, M. M., Goldberg, M. B., Rose, D. J., Grotbeck, E. J., Burland, V., & Blattner, F. R.** (2001). Complete DNA sequence and analysis of the large virulence plasmid of *Shigella flexneri*. *Infection & Immunity*, **69**(5), 3271–3285.
- Verdet, C., Arlet, G., Barnaud, G., Lagrange, P. H., & Philippon, A.** (2000). A novel integron in *Salmonella enterica* serovar Enteritidis, carrying the *bla*_{DHA-1} gene and its regulator gene *ampR*, originated from *Morganella morganii*. *Antimicrobial Agents and Chemotherapy*, **44**(1), 222–225.
- Versalovic, J., Schneider, M., De Bruijn, F. J., & Lupski, J. R.** (1994). Genomic fingerprinting of bacteria using repetitive sequence-based polymerase chain reaction. *Methods in Molecular and Cellular Biology*, **5**, 25–40.

- Von Hippel, P. H.** (1998). An integrated model of the transcription complex in elongation, termination, and editing. *Science*, **281**, 660–665.
- Voolaid, V., Tenson, T., & Kisand, V.** (2013). *Aeromonas* and *Pseudomonas* species carriers of *ampC* FOX genes in aquatic environments. *Applied and Environmental Microbiology*, **79**(3), 1055–1057.
- Vorská, L. D., Artoš, M. B., Artin, G. M., Rler, W. E., & Avlík, I. P.** (2001). Strategies for differentiation, identification and typing of medically important species of mycobacteria by molecular methods, **46**(11-12), 309–328.
- Vossen, K., Wolz, R., Daugherty, M., & Fried, M.** (1997). Role of macromolecular hydration in the binding of the *Escherichia coli* cyclic AMP receptor to DNA. *Biochemistry*, **36**(39), 11640–11647.
- Wang, Q., & Nomura, C. T.** (2010). Monitoring differences in gene expression levels and polyhydroxyalkanoate (PHA) production in *Pseudomonas putida* KT2440 grown on different carbon sources. *Journal of Bioscience and Bioengineering*, **110**(6), 653–659.
- Weisburg, W. G., Barns, S. M., Pelletier, D. A., & Lane, D. J.** (1991). 16S ribosomal DNA amplification for phylogenetic study. *Journal of Bacteriology*, **173**(2), 697–703.
- White, P. A., McIver, C. J., & Rawlinson, W. D.** (2001). Integrons and Gene Cassettes in the *Enterobacteriaceae*. *Antimicrobial Agents and Chemotherapy*, **45**(9), 2658–2661.
- Whiteley, A. S., Wiles, S., Lilley, A. K., Philp, J., & Bailey, M. J.** (2001). Ecological and physiological analyses of *Pseudomonad* species within a phenol remediation system. *Journal of Microbiological Methods*, **44**(1), 79–88.
- Wiehmann, L., Wagner, G., Cramer, N., Siebert, B., Gudowius, P., Morales, G., Köhler, T., van Delden, C., Weinel, C., Slickers, P., & Tümmler, B.** (2007). Population structure of *Pseudomonas aeruginosa*. *Proceedings of the National Academy of Sciences of the United States of America*, **104**(19), 8101–8106.
- Wilson, N.** (2007). Integrons in *Pseudomonads* are associated with hotspots of genomic diversity. PhD thesis, *School of Molecular & Microbial Bioscience, The University of Sydney*.
- Winsor, G. L., Lam, D. K., Fleming, L., Lo, R., Whiteside, M. D., Yu, N. Y., Hancock, R. E., & Brinkman, F. S.** (2011). *Pseudomonas* Genome Database: improved comparative analysis and population genomics capability for *Pseudomonas* genomes. *Nucleic Acids Research*, **39**(Database issues), D596–600.
- Wolfgang, M. C., Kulasekara, B. R., Liang, X., Boyd, D., Wu, K., Yang, Q., Miyada, C. G., & Lory, S.** (2003). Conservation of genome content and virulence determinants among clinical and environmental isolates of *Pseudomonas aeruginosa*. *Proceedings of the National Academy Science of the United States of America*, **100**(14), 8484–8489.
- Wright, M. S., Baker-Austin, C., Lindell, A. H., Stepanauskas, R., Stokes, H. W., & McArthur, J. V.** (2008). Influence of industrial contamination on mobile genetic

- elements: class 1 integron abundance and gene cassette structure in aquatic bacterial communities. *The ISME Journal*, **2**(4), 417–428.
- Xu, F., Lin-Chao, S., & Cohen, S. N.** (1993). The *Escherichia coli* *pcnB* gene promotes adenylation of antisense RNAI of ColE1-type plasmids *in vivo* and degradation of RNAI decay intermediates. *Proceedings of the National Academy of Sciences of the United States of America*, **90**(14), 6756–6760.
- Xu, H., Davies, J., & Miao, V.** (2007). Molecular characterization of class 3 integrons from *Delftia* spp. *Journal of Bacteriology*, **189**(17), 6276–6283.
- Xu, Y., Luo, Q., & Zhou, M.** (2013). Identification and characterization of integron-mediated antibiotic resistance in the phytopathogen *Xanthomonas oryzae* pv. *oryzae*. *PloS One*, **8**(2), e55962.
- Yanisch-Perron, C., Vieira, J., & Messing, J.** (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene*, **33**, 103–119.
- Zobell, C. E., & Upham, H. C.** (1944). A list of marine bacteria including descriptions of sixty new species. *Bulletin of the Scripps Institute of Oceanography*, **5**, 239–292.
- Zong, Z.** (2012). Discovery of *bla*(OXA-199), a chromosome-based *bla*(OXA-48)-like variant, in *Shewanella xiamenensis*. *PloS One*, **7**(10), e48280.

APPENDICES

All of the isolates used in this study were screened for class 1 integrons (*intI1*), *Pseudomonas* chromosomal integrons (PCIs), *Pseudomonas attC* sites (*Ps. attC-attC*), and IS1111-*attC* elements by southern hybridisation and PCR. Where a clinical isolate tested positive for given element by with Southern hybridization or PCR, it is annotated as +.

Table A1: Clinical, host associated *Pseudomonads* used in this study

Ref.	Organism	Date	Clinical Site	Hospital	<i>intI1</i>	PCI	<i>Ps. attC-attC</i>	IS1111- <i>attC</i>
TS464	<i>Ps. aeruginosa</i>	29/01/2010	Catheter specimen urine	San Adventist	+	-	-	+
TS491	<i>Ps. aeruginosa</i>	17/03/2010	Swab wound	San Adventist	+	-	-	+
TS504	<i>Ps. aeruginosa</i>	7/02/2010	Swab left heel	San Adventist	+	-	-	+
TS558	<i>Ps. aeruginosa</i>	21/03/2010	Midstream urine	San Adventist	+	-	-	+
TS589	<i>Ps. aeruginosa</i>	23/04/2010	Midstream urine	San Adventist	+	-	-	+
TS597	<i>Ps. aeruginosa</i>	3/05/2010	Fluid pleural fluid	San Adventist	+	-	-	+
TS599	<i>Ps. aeruginosa</i>	3/05/2010	Swab left lateral	San Adventist	-	-	-	+
TS600	<i>Ps. aeruginosa</i>	3/05/2010	Midstream urine	San Adventist	-	-	-	+
TS616	<i>Ps. aeruginosa</i>	3/05/2010	Unknown source	San Adventist	-	-	-	+
TS620	<i>Ps. aeruginosa</i>	17/05/2010	Swab vaginal	San Adventist	-	-	-	+
TS629	<i>Ps. aeruginosa</i>	21/05/2010	Fluid left drain	San Adventist	+	-	-	+
TS630	<i>Ps. aeruginosa</i>	24/05/2010	Swab left antrum	San Adventist	+	-	-	+
TS631	<i>Ps. aeruginosa</i>	25/05/2010	Midstream urine	San Adventist	-	-	-	+
TS634	<i>Ps. aeruginosa</i>	25/05/2010	Swab small ulcer	San Adventist	+	-	-	+
TS653	<i>Ps. aeruginosa</i>	16/06/2010	Swab left leg ulcer	San Adventist	-	-	-	+
TS654	<i>Ps. aeruginosa</i>	15/06/2010	Urine unknown meth	San Adventist	-	-	-	+
TS664	<i>Ps. aeruginosa</i>	22/06/2010	Swab right ear	San Adventist	-	-	-	+
TS684	<i>Ps. aeruginosa</i>	4/07/2010	Midstream urine	San Adventist	-	-	-	+
TS741	<i>Ps. aeruginosa</i>	18/08/2010	Swab right foot	San Adventist	-	-	-	+
JIP004	<i>Ps. aeruginosa</i>	8/03/2005	Urine	Westmead	-	-	-	-
JIP006	<i>Ps. aeruginosa</i>	9/03/2005	Wound	Westmead	-	-	-	+
JIP009	<i>Ps. aeruginosa</i>	11/03/2005	NBAL	Westmead	-	-	-	+

JIP012	<i>Ps. aeruginosa</i>	25/03/2005	Blood	Westmead	-	-	-	-
JIP013	<i>Ps. aeruginosa</i>	22/03/2005	Urine	Westmead	-	-	-	-
JIP014	<i>Ps. aeruginosa</i>	22/03/2005	Right heel tissue	Westmead	-	-	-	-
JIP016	<i>Ps. aeruginosa</i>	22/03/2005	Toe tissue	Westmead	-	-	-	-
JIP019	<i>Ps. aeruginosa</i>	6/04/2005	Foot tissue	Westmead	-	-	-	-
JIP027	<i>Ps. aeruginosa</i>	8/07/2005	Blood	Westmead	-	-	-	-
JIP028	<i>Ps. aeruginosa</i>	13/08/2005	SP	Westmead	-	-	-	-
JIP029	<i>Ps. aeruginosa</i>	11/08/2005	Mandible wound	Westmead	-	-	-	-
JIP031	<i>Ps. aeruginosa</i>	15/08/2005	SP	Westmead	+	-	-	+
JIP032	<i>Ps. aeruginosa</i>	12/07/2005	Wound	Westmead	-	-	-	-
JIP033	<i>Ps. aeruginosa</i>	20/08/2005	Urine	Westmead	-	-	-	-
JIP037	<i>Ps. aeruginosa</i>	20/08/2005	SP	Westmead	-	-	-	-
JIP038	<i>Ps. aeruginosa</i>	19/07/2005	Wound	Westmead	+	-	-	-
JIP040	<i>Ps. aeruginosa</i>	25/09/2005	Tip	Westmead	-	-	-	-
JIP041	<i>Ps. aeruginosa</i>	27/10/2005	Tissue	Westmead	+	-	-	+
JIP044	<i>Ps. aeruginosa</i>	22/01/2006	Fluid bag	Westmead	+	-	-	+
JIP045	<i>Ps. aeruginosa</i>	5/02/2006	Wound	Westmead	+	+	-	+
JIP047	<i>Ps. aeruginosa</i>	16/02/2006	Wound	Westmead	-	-	-	+
JIP051	<i>Ps. aeruginosa</i>	28/02/2006	Wound	Westmead	-	-	-	-
JIP052	<i>Ps. aeruginosa</i>	9/03/2006	Wound	Westmead	+	+	-	+
JIP055	<i>Ps. aeruginosa</i>	1/08/2006	Wound	Westmead	-	-	-	-
JIP056	<i>Ps. aeruginosa</i>	19/08/2006	SP	Westmead	-	-	-	-
JIP067	<i>Ps. aeruginosa</i>	3/11/2006	Urine	Westmead	-	-	-	-
JIP068	<i>Ps. aeruginosa</i>	9/11/2006	SP	Westmead	-	-	-	-
JIP070	<i>Ps. aeruginosa</i>	29/11/2006	Blood	Westmead	-	-	-	-
JIP071	<i>Ps. aeruginosa</i>	6/06/2006	Systemic-pulmonary collateral flow	Westmead	-	-	-	-
JIP072	<i>Ps. aeruginosa</i>	18/12/2006	Urine	Westmead	-	-	-	-
JIP073	<i>Ps. aeruginosa</i>	8/01/2007	Foot tissue	Westmead	+	-	-	+
JIP079	<i>Ps. aeruginosa</i>	22/01/2007	SP	Westmead	-	-	-	-
JIP080	<i>Ps. aeruginosa</i>	23/01/2007	SP	Westmead	+	-	-	+
JIP104	<i>Ps. aeruginosa</i>	20/03/2007	Blood	Westmead	-	-	-	-
JIP106	<i>Ps. aeruginosa</i>	30/03/2007	Urine	Westmead	+	-	-	-
JIP108	<i>Ps. aeruginosa</i>	14/03/2007	Blood	Westmead	-	-	-	-

JIP111	<i>Ps. aeruginosa</i>	3/04/2007	Sputum	Westmead	-	-	-	-
JIP112	<i>Ps. aeruginosa</i>	8/04/2007	Sputum	Westmead	-	-	-	-
JIP113	<i>Ps. aeruginosa</i>	16/04/2007	Sputum	Westmead	-	-	-	-
JIP116	<i>Ps. aeruginosa</i>	17/04/2007	Wound	Westmead	-	-	-	-
JIP117	<i>Ps. aeruginosa</i>	19/04/2007	Abdominal swab	Westmead	+	-	-	+
JIP118	<i>Ps. aeruginosa</i>	24/04/2007	Wound	Westmead	-	-	-	-
JIP119	<i>Ps. aeruginosa</i>	25/04/2007	Wound	Westmead	-	-	-	-
JIP120	<i>Ps. aeruginosa</i>	2/05/2007	Swab	Westmead	-	-	-	-
JIP122	<i>Ps. aeruginosa</i>	11/05/2007	Urine	Westmead	-	-	-	-
JIP124	<i>Ps. aeruginosa</i>	11/05/2007	Groin	Westmead	-	-	-	-
JIP132	<i>Ps. aeruginosa</i>	5/06/2007	NBAL	Westmead	-	-	-	-
JIP134	<i>Ps. aeruginosa</i>	26/06/2007	Sputum	Westmead	-	-	-	-
JIP136	<i>Ps. aeruginosa</i>	13/07/2007	Wound	Westmead	-	-	-	-
JIP137	<i>Ps. aeruginosa</i>	18/07/2007	Wound	Westmead	-	-	-	-
JIP138	<i>Ps. aeruginosa</i>	12/09/2007	Swab	Westmead	-	-	-	-
JIP139	<i>Ps. aeruginosa</i>	13/09/2007	Leg	Westmead	-	-	-	-
JIP141	<i>Ps. aeruginosa</i>	15/09/2007	Urine	Westmead	-	-	-	-
SSI 1.25	<i>Ps. aeruginosa</i>	25-Apr-04	Respiratory C Sp	Westmead	-	-	-	-
SSI 1.28	<i>Ps. aeruginosa</i>	27-Apr-04	Urine	Westmead	-	-	-	-
SSI 1.48	<i>Ps. aeruginosa</i>	24-May-04	Respiratory C Sp	Westmead	-	-	-	-
SSI 1.58	<i>Ps. aeruginosa</i>	08-Jun-04	Neck collection	Westmead	-	-	-	-
SSI 1.63	<i>Ps. aeruginosa</i>	13-Jun-04	Respiratory C Sp	Westmead	-	-	-	-
SSI 1.65	<i>Ps. aeruginosa</i>	18-Jun-04	Respiratory C Sp	Westmead	-	-	-	-
SSI 1.69	<i>Ps. aeruginosa</i>	20-Jun-04	Respiratory C Sp	Westmead	-	-	-	-
SSI 1.9	<i>Ps. aeruginosa</i>	03-Apr-04	Respiratory C Sp	Westmead	-	-	-	-
SSI 1.95	<i>Ps. aeruginosa</i>	05-Jul-04	Respiratory C Sp	Westmead	-	-	-	-
SSI 2.1	<i>Ps. aeruginosa</i>	10-Jul-04	Respiratory C NBAL	Westmead	-	-	-	-
SSI 2.100	<i>Ps. aeruginosa</i>	13-Oct-04	Respiratory C NBAL	Westmead	-	-	-	-
SSI 2.17	<i>Ps. aeruginosa</i>	29-Jul-04	Respiratory C NBAL	Westmead	-	-	-	-
SSI 2.32	<i>Ps. aeruginosa</i>	03-Aug-04	Respiratory C NBAL	Westmead	-	-	-	-
SSI 2.37	<i>Ps. aeruginosa</i>	06-Aug-04	Respiratory C NBAL	Westmead	+	-	-	+
SSI 2.51	<i>Ps. aeruginosa</i>	16-Aug-04	Respiratory C NBAL	Westmead	-	-	-	-
SSI 2.84	<i>Ps. aeruginosa</i>	16-Sep-04	Blood	Westmead	+	+	-	+
SSI 2.87	<i>Ps. aeruginosa</i>	10-Oct-04	Respiratory C Sp	Westmead	-	-	-	-

SSI 2.96	<i>Ps. aeruginosa</i>	11-Oct-04	Other C WS leg	Westmead	-	-	-	-
SSI 3.14	<i>Ps. aeruginosa</i>	13-Nov-04	Other C WS knee	Westmead	-	-	-	-
SSI 3.35	<i>Ps. aeruginosa</i>	08-Dec-04	Blood dialysis fluid	Westmead	+	-	-	-
SSI 3.37	<i>Ps. aeruginosa</i>	07-Dec-04	Respiratory C NBAL	Westmead	-	-	-	-
SSI 3.38	<i>Ps. aeruginosa</i>	07-Dec-04	Catheter Urine	Westmead	-	-	-	-
SSI 3.40	<i>Ps. aeruginosa</i>	13-Dec-04	Respiratory C NBAL	Westmead	-	-	-	-
SSI 3.44	<i>Ps. aeruginosa</i>	25-Dec-04	Respiratory C SP	Westmead	+	-	-	-
SSI 3.53	<i>Ps. aeruginosa</i>	07-Jan-05	Respiratory C NBAL	Westmead	-	-	-	-
SSI 3.62	<i>Ps. aeruginosa</i>	16-Jan-05	Other C Tip	Westmead	-	-	-	-
SSI 3.71	<i>Ps. aeruginosa</i>	22-Jan-05	Respiratory C Sp	Westmead	-	-	-	-
SSI 3.72	<i>Ps. aeruginosa</i>	21-Jan-05	Blood	Westmead	-	-	-	-
SSI 3.77	<i>Ps. aeruginosa</i>	14-Feb-05	Foot wound	Westmead	-	-	-	-
SSI 3.86	<i>Ps. aeruginosa</i>	21-Feb-05	Other C WS wound Mid	Westmead	-	-	-	-
SSI 3.9	<i>Ps. aeruginosa</i>	30-Oct-04	Other C WS wound	Westmead	-	-	-	-
SSI 3.91	<i>Ps. aeruginosa</i>	28-Feb-05	Respiratory C SP	Westmead	-	-	-	-
SSI 4.100	<i>Ps. aeruginosa</i>	06-Jun-05	Other C WS wound	Westmead	-	-	-	-
SSI 4.32	<i>Ps. aeruginosa</i>	29-Mar-05	Respiratory C SP	Westmead	-	-	-	-
SSI 4.37	<i>Ps. aeruginosa</i>	06-Apr-05	Tissue left stump	Westmead	-	-	-	-
SSI 4.41	<i>Ps. aeruginosa</i>	11-Apr-05	Respiratory sputum	Westmead	-	-	-	-
SSI 4.46	<i>Ps. aeruginosa</i>	20-Apr-05	Trachey wounds	Westmead	-	-	-	-
SSI 4.58	<i>Ps. aeruginosa</i>	28-Apr-05	Plerural fluid	Westmead	-	-	-	-
SSI 4.62	<i>Ps. aeruginosa</i>	05-May-05	Other C WS wound	Westmead	-	-	-	-
SSI 4.63	<i>Ps. aeruginosa</i>	03-May-05	Unknown	Westmead	-	-	-	-
SSI 4.69	<i>Ps. aeruginosa</i>	10-May-05	Urine	Westmead	-	+	-	-
SSI 4.77	<i>Ps. aeruginosa</i>	17-May-05	Respiratory C Sp	Westmead	-	-	-	-
SSI 5.25	<i>Ps. aeruginosa</i>	08-Jul-05	Respiratory C NBAL	Westmead	-	-	-	-
SSI 5.32	<i>Ps. aeruginosa</i>	07-Jul-05	Respiratory C Sp	Westmead	-	-	-	-
SSI 5.34	<i>Ps. aeruginosa</i>	12-Jul-05	Sternal Wound	Westmead	+	-	-	-
SSI 5.39	<i>Ps. aeruginosa</i>	19-Jul-05	Respiratory C NBAL	Westmead	-	+	-	-
SSI 5.41	<i>Ps. aeruginosa</i>	24-Jul-05	Respiratory C Sp	Westmead	-	-	-	-
P1	<i>Chryseobacterium indologenes</i>	28/04/2004	Bronchial washing	Westmead	-	-	-	-
P2	<i>Ps. mendocina</i>	28/03/2004	Blood culture	Westmead	-	-	-	-
P4	<i>Ps. fluorescens</i>	28/11/2003	Bronchoalveolar lavag	Westmead	-	-	-	-

P15	<i>Ps. stutzeri</i> ATCC 17588	15/01/1999	Human spinal fluid	Westmead	-	+	+	+
P16	<i>Ps. aeruginosa</i>	16/02/2006	Clinical Sputum	Westmead	-	-	-	-
P17	<i>Ps. stutzeri</i>	21/02/2007	Eye swab	Westmead	-	+	+	+

Note: All of the clinical *Pseudomonas* isolates were kindly provided by our collaborators. These clinical isolates were previously assigned a reference (Ref.) identity. TS, JIP, SSI refer to different antibiotic cycling studies carried out the respective Hospitals, from which these *Pseudomonas* were isolated. P was abbreviated form of *Pseudomonas*. Date refers to the date the isolate was recovered from clinical site at the given hospital. Medical abbreviations: NBAL: Non-bronchoscopic bronchoalveolar lavage; C: catheter; Sp: Spontaneous pneumothorax; WS: WoundStat.

Table A2: Non-clinical, environmental isolates used in this study

Isolate	BLAST HIT Organism	Accession number	BLASTn % identity	Date	Sample Source	<i>intI1</i>	PCI	<i>Ps attC-attC</i>	IS1111- <i>attC</i>
<i>Pseudomonas</i> sp. JM1	<i>Ps. alcaliphila</i>	KM248339.1	99 (1066/1068bp)	May-2011	Jake's Mop	-	+	-	-
<i>Pseudomonas</i> sp. JM2	<i>Ps. alcaligenes</i> str. S3	DQ115541.1	99 (826/837bp)	May-2011	Jake's Mop	+	+	+	+
<i>Pseudomonas</i> sp. JM3	Unknown	N/A	N/A	May-2011	Jake's Mop	-	-	-	-
<i>Pseudomonas</i> sp. JM4	Unknown	N/A	N/A	May-2011	Jake's Mop	-	-	-	-
<i>Pseudomonas</i> sp. JM5	Unknown	N/A	N/A	May-2011	Jake's Mop	-	-	-	-
<i>Pseudomonas</i> sp. JM6	<i>Ps. alcaligenes</i> str. Y34	FJ830845.1	99 (899/901bp)	May-2011	Jake's Mop	+	+	+	+
<i>Pseudomonas</i> sp. JM7	Unknown	N/A	N/A	May-2011	Jake's Mop	-	-	-	+
<i>Pseudomonas</i> sp. JM8	Unknown	N/A	N/A	May-2011	Jake's Mop	-	-	-	-
<i>Pseudomonas</i> sp. JM9	Unknown	N/A	N/A	May-2011	Jake's Mop	-	-	-	-
<i>Pseudomonas</i> sp. JM10	<i>Ps. sp. N3(2012b)</i>	JN820155.1	99 (892/904bp)	May-2011	Jake's Mop	+	-	+	+
<i>Pseudomonas</i> sp. JM11	Unknown	N/A	N/A	May-2011	Jake's Mop	-	-	-	-

<i>Pseudomonas</i> sp. JS1	<i>Unknown</i>	N/A	N/A	May-2011	Jake's Sponge	-	-	-	-
<i>Pseudomonas</i> sp. JS2	<i>Unknown</i>	N/A	N/A	May-2011	Jake's Sponge	-	+	+	-
<i>Pseudomonas</i> sp. JS3	<i>Unknown</i>	N/A	N/A	May-2011	Jake's Sponge	-	+	+	-
<i>Pseudomonas</i> sp. JS4	<i>Unknown</i>	N/A	N/A	May-2011	Jake's Sponge	-	-	-	-
<i>Pseudomonas</i> sp. JS5	<i>Unknown</i>	N/A	N/A	May-2011	Jake's Sponge	-	-	-	-
<i>Pseudomonas</i> sp. JS6	<i>Unknown</i>	N/A	N/A	May-2011	Jake's Sponge	-	-	-	-
<i>Pseudomonas</i> sp. JS7	<i>Unknown</i>	N/A	N/A	May-2011	Jake's Sponge	-	-	-	-
<i>Pseudomonas</i> sp. AP1	<i>Uncultured Ps. sp. clone AV_5N-G03</i>	EU341207.1	100 (1400/1400bp)	May-2011	RNSH Panroom	+	-	-	-
<i>Stenotrophomonas</i> sp. AP2	<i>Stenotrophomonas maltophilia JV3</i>	NC015947.1	100 (1400/1413bp)	May-2011	RNSH Panroom	+	-	-	-
<i>Pseudomonas</i> sp. AP3	<i>Ps. putida WZX_19</i>	EF440613.1	99 (1391/1397bp)	May-2011	RNSH Panroom	+	-	-	-
<i>Pseudomonas</i> sp. AS1	<i>Ps. aeruginosa ACZ01</i>	KM434772.1	100 (612/612bp)	May-2011	RNSH Sink Panroom	+	-	-	+
<i>Pseudomonas</i> sp. AS2	<i>Ps. aeruginosa PA94</i>	KM013815.1	99 (1389/1390bp)	May-2011	RNSH Sink Panroom	+	-	-	-
<i>Pseudomonas</i> sp. AS3	<i>Ps. aeruginosa HE21</i>	LN624810.2	100 (1076/1076)	May-2011	RNSH Sink Panroom	-	-	-	-
<i>Pseudomonas</i> sp. ES1	<i>Uncultured bacterium gene for 16S rRNA, partial sequence, clone: HSM-SS-011</i>	AB238774.1	99 (1079/1081bp)	May-2011	Yi Vee's Sponge	-	-	-	-
<i>Pseudomonas</i> sp. ES2	<i>Unknown</i>	N/A	N/A	May-2011	Yi Vee's Sponge	-	-	-	-
<i>Pseudomonas</i> sp. ES3	<i>Unknown</i>	N/A	N/A	May-2011	Yi Vee's Sponge	-	-	-	-

<i>Pseudomonas</i> sp. ES4	<i>Unknown</i>	N/A	N/A	May-2011	Yi Vee's Sponge	-	-	-	-
<i>Pseudomonas</i> sp. ES5	<i>Unknown</i>	N/A	N/A	May-2011	Yi Vee's Sponge	-	-	-	-
<i>Pseudomonas</i> sp. FM1	<i>Unknown</i>	N/A	N/A	Jan-2012	Feyza's Mop	-	+	-	+
<i>Pseudomonas</i> sp. FM2	<i>Unknown</i>	N/A	N/A	Jan-2012	Feyza's Mop	-	+	-	-
<i>Pseudomonas</i> sp. FM3	<i>Unknown</i>	N/A	N/A	Jan-2012	Feyza's Mop	-	+	-	+
<i>Pseudomonas</i> sp. FM4	<i>Ps. putida ND6</i>	CP003588.1	99 (1399/1403)	Jan-2012	Feyza's Mop	-	+	-	+
<i>Pseudomonas</i> sp. FM5	<i>Unknown</i>	N/A	N/A	Jan-2012	Feyza's Mop	-	-	-	-
<i>Pseudomonas</i> sp. FM6	<i>Unknown</i>	N/A	N/A	Jan-2012	Feyza's Mop	-	-	-	+
<i>Pseudomonas</i> sp. FM7	<i>Unknown</i>	N/A	N/A	Jan-2012	Feyza's Mop	-	-	-	-
<i>Pseudomonas</i> sp. FM8	<i>Unknown</i>	N/A	N/A	Jan-2012	Feyza's Mop	-	-	-	-
<i>Pseudomonas</i> sp. FM9	<i>Unknown</i>	N/A	N/A	Jan-2012	Feyza's Mop	-	-	-	-
<i>Pseudomonas</i> sp. FM10	<i>Unknown</i>	N/A	N/A	Jan-2012	Feyza's Mop	-	-	-	-
<i>Pseudomonas</i> sp. FM11	<i>Unknown</i>	N/A	N/A	Jan-2012	Feyza's Mop	-	-	-	-
<i>Pseudomonas</i> sp. MS1	<i>Unknown</i>	N/A	N/A	Jan-2011	Mia's Sponge	-	+	-	-
<i>Pseudomonas</i> sp. MS2	<i>Unknown</i>	N/A	N/A	Jan-2011	Mia's Sponge	-	-	-	-
<i>Pseudomonas</i> sp. MS3	<i>Uncultured gamma proteobacterium from sea squirt Microcosmus sp. clone MspS123</i>	AY770721.1	99 (899/905p)	Jan-2011	Mia's Sponge	-	-	-	-

<i>Pseudomonas</i> sp. MS4	<i>Unknown</i>	N/A	N/A	Jan-2011	Mia's Sponge	-	-	-	-
<i>Pseudomonas</i> sp. MS5	<i>Unknown</i>	N/A	N/A	Jan-2011	Mia's Sponge	-	-	-	-
<i>Pseudomonas</i> sp. MS6	<i>Unknown</i>	N/A	N/A	Jan-2011	Mia's Sponge	-	-	-	-
<i>Pseudomonas</i> sp. NS1	<i>Unknown</i>	N/A	N/A	May-2011	Nick's Sponge	-	-	-	-
<i>Pseudomonas</i> sp. NS2	<i>Unknown</i>	N/A	N/A	May-2011	Nick's Sponge	-	-	-	-
<i>Pseudomonas</i> sp. NS3	<i>Unknown</i>	N/A	N/A	May-2011	Nick's Sponge	-	-	-	-
<i>Pseudomonas</i> sp. NS4	<i>Unknown</i>	N/A	N/A	May-2011	Nick's Sponge	-	-	-	-
<i>Pseudomonas</i> sp. NS5	<i>Unknown</i>	N/A	N/A	May-2011	Nick's Sponge	-	-	-	-
<i>Agrobacterium</i> sp. NS6	<i>Agrobacterium</i> sp. <i>DB14</i>	JQ4375441.1	99(1258/1260bp)	May-2011	Nick's Sponge	-	+	-	+
<i>Pseudomonas</i> sp. NS7	<i>Unknown</i>	N/A	N/A	May-2011	Nick's Sponge	-	-	-	-
<i>Pseudomonas</i> sp. NS8	<i>Unknown</i>	N/A	N/A	May-2011	Nick's Sponge	-	-	-	-
<i>Pseudomonas</i> sp. LS1	<i>Unknown</i>	N/A	N/A	June-2011	Laura's Sponge	-	-	-	-
<i>Pseudomonas</i> sp. LS2	<i>Unknown</i>	N/A	N/A	June-2011	Laura's Sponge	-	-	-	-
<i>Pseudomonas</i> sp. LS3	<i>Unknown</i>	N/A	N/A	June-2011	Laura's Sponge	-	-	-	-
<i>Pseudomonas</i> sp. LS4	<i>Unknown</i>	N/A	N/A	June-2011	Laura's Sponge	-	-	+	-
<i>Stenotrophomonas</i> sp. LS5	<i>Stenotrophomonas</i> <i>maltophilia K279a</i>	NC010943.1	99 (1077/1081bp)	June-2011	Laura's Sponge	-	-	+	+
<i>Stenotrophomonas</i> sp. LS6	<i>Stenotrophomonas</i> sp. <i>U1369-101122-</i>	JQ082152.1	99 (1074/1081bp)	June-2011	Laura's Sponge	-	-	-	+

	SW177								
<i>Pseudomonas</i> sp. LS7	<i>Unknown</i>	N/A	N/A	June-2011	Laura's Sponge	-	-	-	-
<i>Delftia</i> sp. LS8	<i>Delftia acidovorans</i> SPH-1	NR074691.1	99 (1387/1391bp)	June-2011	Laura's Sponge	-	-	-	+
<i>Pseudomonas</i> sp. LS9	<i>Unknown</i>	N/A	N/A	June-2011	Laura's Sponge	-	-	-	-
<i>Pseudomonas</i> sp. LS10	<i>Unknown</i>	N/A	N/A	June-2011	Laura's Sponge	-	-	-	-
<i>Pseudomonas</i> sp. LS11	<i>Unknown</i>	N/A	N/A	June-2011	Laura's Sponge	-	-	-	-
<i>Pseudomonas</i> sp. LS12	<i>Unknown</i>	N/A	N/A	June-2011	Laura's Sponge	-	-	-	-
<i>Brevundimonas</i> sp. LS13	<i>Brevundimonas vesicularis</i> G1-1-80	KC494327.1	99 (1026/1030bp)	June-2011	Laura's Sponge	+	-	-	+
<i>Pseudomonas</i> sp. LS14	<i>Unknown</i>	N/A	N/A	June-2011	Laura's Sponge	-	-	-	-
<i>Pseudomonas</i> sp. LS15	<i>Unknown</i>	N/A	N/A	June-2011	Laura's Sponge	-	-	-	-
P05	<i>Chryseomonas luteola</i>	N/A	N/A	5/04/2005	Enviro.	-	-	-	+
P06	<i>Chryseobacterium meningosepticum</i>	N/A	N/A	1/07/1993	Water	+	+	+	+
P07	<i>Ps. alcaligenes</i>	N/A	N/A	Unknown	Enviro.	+	-	-	-
P08	<i>Ps. oryzihabitans</i>	N/A	N/A	Unknown	Enviro.	-	-	-	-
P09	<i>Ps. putida</i>	N/A	N/A	1/07/1993	Enviro.	-	-	-	-
P10	<i>Brevundimonas diminuta</i>	N/A	N/A	Unknown	Enviro.	-	-	-	-
P11	<i>Ps. aureofacium</i> ATCC13985	N/A	N/A	20/01/1999	Sandstone rock	-	-	-	-
P13	<i>Ps. mendocina</i> ATCC 25411	N/A	N/A	15/01/1999	Soil enrichment	-	+	+	+
P14	<i>Ps. putida</i> ATCC 12683	N/A	N/A	20/01/1999	Activated sludge	-	-	-	-

N/A: not available; Enviro: Environmental; RNS: Royal North Shore Hospitals. All *Pseudomonad* isolated annotated as “P#” were kindly provided by the Brown Lab.

Table A3: Enteric (outgroup) isolates used in this study

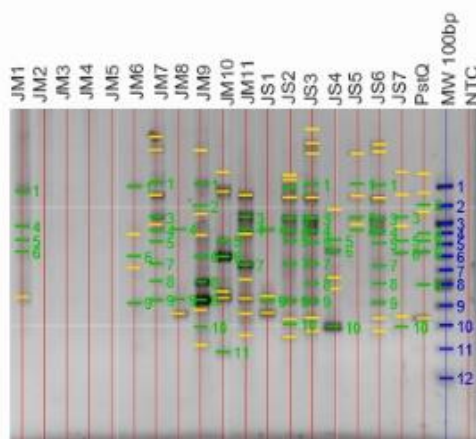
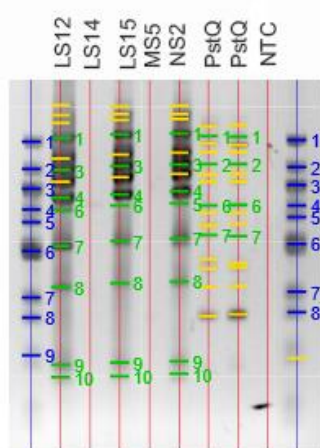
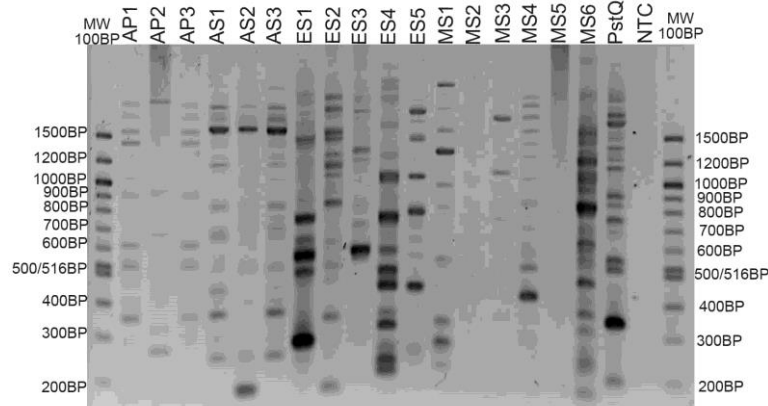
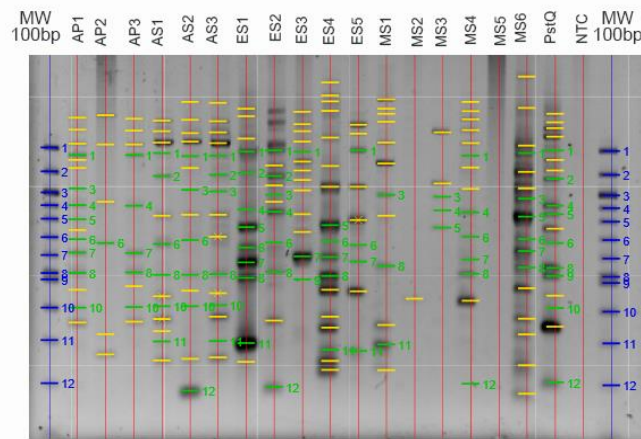
Ref.	Organisms	Date	Source	Lab	<i>intI1</i>	<i>PCI</i>	<i>Ps. attC-attC</i>	<i>IS1111-attC</i>
2045	Unknown Gram Negative	16/6/09	Unknown	Iredell Lab Westmead Hospital	-	-	-	-
2064	Unknown Gram Negative	16/6/09	Unknown	Iredell Lab Westmead Hospital	-	-	-	-
4060	Unknown Gram Negative	16/6/09	Unknown	Iredell Lab Westmead Hospital	-	-	-	-
6066	Unknown Gram Negative	16/6/09	Unknown	Iredell Lab Westmead Hospital	-	-	-	-
9023	Unknown Gram Negative	16/6/09	Unknown	Iredell Lab Westmead Hospital	-	-	-	-
11045	Unknown Gram Negative	16/6/09	Unknown	Iredell Lab Westmead Hospital	+	-	-	-
11047	Unknown Gram Negative	16/6/09	Unknown	Iredell Lab Westmead Hospital	+	-	-	-
13047	Unknown Gram Negative	16/6/09	Unknown	Iredell Lab Westmead Hospital	-	-	-	-
13090	Unknown Gram Negative	16/6/09	Unknown	Iredell Lab Westmead Hospital	-	-	-	-
14025	Unknown Gram Negative	16/6/09	Unknown	Iredell Lab Westmead Hospital	-	-	-	-
14050i	Unknown Gram Negative	16/6/09	Unknown	Iredell Lab Westmead Hospital	-	-	-	-
14050ii	Unknown Gram Negative	16/6/09	Unknown	Iredell Lab Westmead Hospital	-	-	-	-
15025	Unknown Gram Negative	16/6/09	Unknown	Iredell Lab Westmead Hospital	-	-	-	-
15053	Unknown Gram Negative	16/6/09	Unknown	Iredell Lab Westmead Hospital	-	-	-	-
15065	Unknown Gram Negative	16/6/09	Unknown	Iredell Lab Westmead Hospital	-	-	-	-
15066	Unknown Gram Negative	16/6/09	Unknown	Iredell Lab Westmead Hospital	-	-	-	-
15084	Unknown Gram Negative	16/6/09	Unknown	Iredell Lab Westmead Hospital	-	-	-	-
16075	Unknown Gram Negative	16/6/09	Unknown	Iredell Lab Westmead Hospital	-	-	-	-
18063	Unknown Gram Negative	16/6/09	Unknown	Iredell Lab Westmead Hospital	-	-	-	-
18064	Unknown Gram Negative	16/6/09	Unknown	Iredell Lab Westmead Hospital	-	-	-	-
18084	Unknown Gram Negative	16/6/09	Unknown	Iredell Lab Westmead Hospital	-	-	-	-
19072	Unknown Gram Negative	16/6/09	Unknown	Iredell Lab Westmead Hospital	-	-	-	-
19073	Unknown Gram Negative	16/6/09	Unknown	Iredell Lab Westmead Hospital	-	-	-	-
20019	Unknown Gram Negative	16/6/09	Unknown	Iredell Lab Westmead Hospital	-	-	-	-
20023	Unknown Gram Negative	16/6/09	Unknown	Iredell Lab Westmead Hospital	-	-	-	-
20060	Unknown Gram Negative	16/6/09	Unknown	Iredell Lab Westmead Hospital	-	-	-	-
20061	Unknown Gram Negative	16/6/09	Unknown	Iredell Lab Westmead Hospital	-	-	-	-

22041	Unknown Gram Negative	16/6/09	Unknown	Iredell Lab Westmead Hospital	-	-	-	-
22042	Unknown Gram Negative	16/6/09	Unknown	Iredell Lab Westmead Hospital	-	-	-	-
23069	Unknown Gram Negative	16/6/09	Unknown	Iredell Lab Westmead Hospital	+	-	-	-
24025	Unknown Gram Negative	16/6/09	Unknown	Iredell Lab Westmead Hospital	-	-	-	-
25077	Unknown Gram Negative	16/6/09	Unknown	Iredell Lab Westmead Hospital	+	-	-	-
25078	Unknown Gram Negative	16/6/09	Unknown	Iredell Lab Westmead Hospital	-	-	-	-
26073	Unknown Gram Negative	16/6/09	Unknown	Iredell Lab Westmead Hospital	-	-	-	-
27010	Unknown Gram Negative	16/6/09	Unknown	Iredell Lab Westmead Hospital	-	-	-	-
27023	Unknown Gram Negative	16/6/09	Unknown	Iredell Lab Westmead Hospital	-	-	-	-
27092	Unknown Gram Negative	16/6/09	Unknown	Iredell Lab Westmead Hospital	+	-	-	-
28076	Unknown Gram Negative	16/6/09	Unknown	Iredell Lab Westmead Hospital	+	-	-	-
31044	Unknown Gram Negative	16/6/09	Unknown	Iredell Lab Westmead Hospital	-	-	-	-
33023	Unknown Gram Negative	16/6/09	Unknown	Iredell Lab Westmead Hospital	-	-	-	-
33024	Unknown Gram Negative	16/6/09	Unknown	Iredell Lab Westmead Hospital	+	-	-	-
33094	Unknown Gram Negative	16/6/09	Unknown	Iredell Lab Westmead Hospital	+	-	-	-
33100	Unknown Gram Negative	16/6/09	Unknown	Iredell Lab Westmead Hospital	-	-	-	-
35070w	Unknown Gram Negative	16/6/09	Unknown	Iredell Lab Westmead Hospital	-	-	-	-
35070y	Unknown Gram Negative	16/6/09	Unknown	Iredell Lab Westmead Hospital	+	-	-	-
38085	Unknown Gram Negative	17/6/09	Unknown	Iredell Lab Westmead Hospital	+	-	-	-
38763	Unknown Gram Negative	18/6/09	Unknown	Iredell Lab Westmead Hospital	-	-	-	-
39010	Unknown Gram Negative	19/6/09	Unknown	Iredell Lab Westmead Hospital	-	-	-	-
52098	Unknown Gram Negative	20/6/09	Unknown	Iredell Lab Westmead Hospital	-	-	-	-
53049	Unknown Gram Negative	21/6/09	Unknown	Iredell Lab Westmead Hospital	+	-	-	-
54026	Unknown Gram Negative	22/6/09	Unknown	Iredell Lab Westmead Hospital	-	-	-	-
54062	Unknown Gram Negative	23/6/09	Unknown	Iredell Lab Westmead Hospital	-	-	-	-

APPENDIX A1: Characterization of isolates used in this study

APPENDIX A1.1 BOX-REP analysis of environmental *Pseudomonads*

Genomic fingerprints were generated for all environmental *Pseudomonas* strains in the collection using PCR. Fingerprints were generated using BOX-PCR, and involved the single primer BOXA1R. The number and size of all bands generated using BOX-PCR were determined using the Quantity-1 software package (BioRad). Individual bands were detected using the automatic detection function within Quantity-1 (green) and any undetected bands were manually flagged (yellow). Bands in each isolate were matched to all other bands within the individual gels.



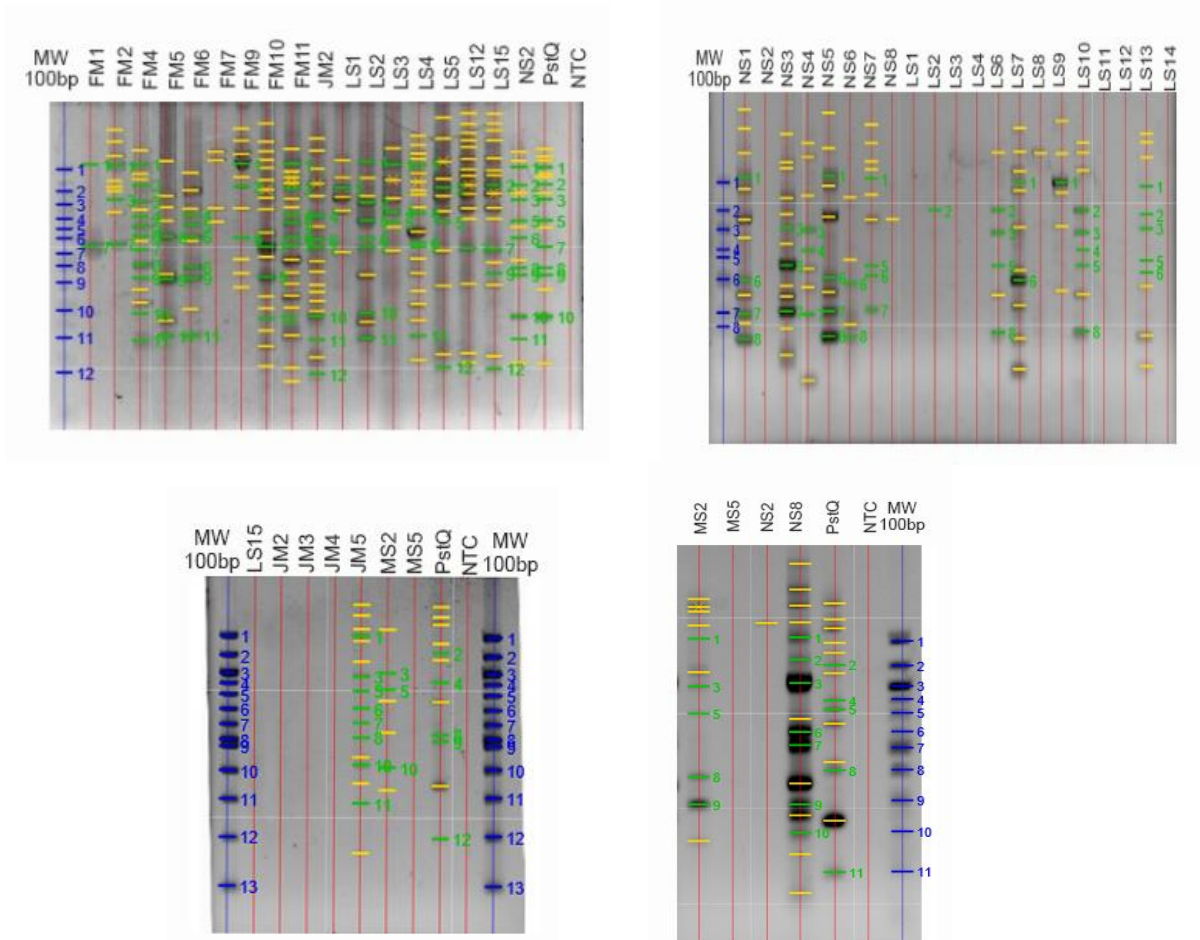


Figure A1: BOXA1R fingerprinting of environmental *Pseudomonads*. The positive control for this experiment was purified *PstQ* gDNA and the negative control contained MQ water as the template (No template control NTC). Isolate names are given at the top of the gel.

APPENDIX A1.2 Southern hybridization detection and recovery of MIs, IS1111-attC, PCIs and Ps.-type attC sites

Southern hybridization was used to screen both the clinical and non-clinical *Pseudomonadaceae* groups as well as the *Enterobacteriaceae* outgroup. All membranes were screen with the following probes:

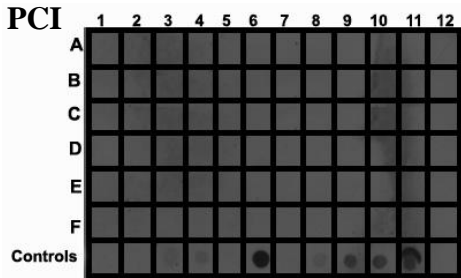
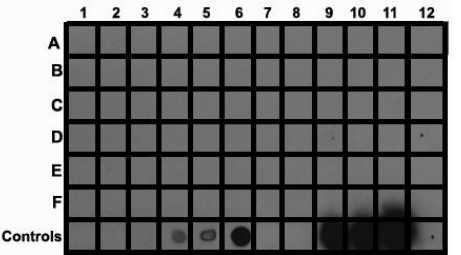
- *MI probe based on pR388 551-bp*
- *IS1111-attC probe based on either Pst587 or Pst405 gDNA 1340-bp in size*
- *PCI probe based on PstQ gDNA 456-bp*
- *Pseudomonas-type attC cocktail: composed of three different Pseudomonas attC sites from Ps. streamia, Ps. mendocina and Ps. stutzeri 100-bp each site.*

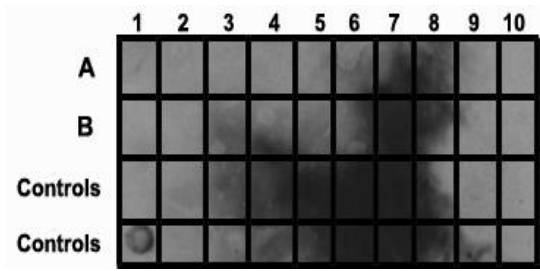
The following controls were used:

- *Pst405*
- *Pst587*
- *Pst595*
- *PstRNAIII*
- *Pst641*
- *PstDNSP21*
- *PsKM91*
- *PsNW1*
- *PA01*
- *PsMB*
- *Ps216.2B*
- *PstQ*
- *R388 (intI1)*
- *Psfluro*
- *Thaura (Tn7 intI2)*
- *Probe: To test probe labelling efficiency 200 ng of denatured probe (IS1111-attC/intI1/PCI/Ps.- attC) was spotted onto the membrane only where noted.*

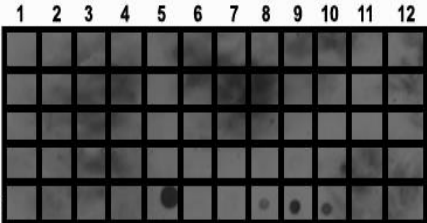
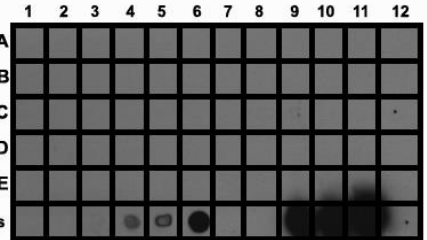
Equal concentrations of 400 ng gDNA was spotted per isolate and membranes were hybridized against respective probes.

Clinical TS (n=19) and JIP (n=54) sample set: PCI and *Ps. -attC* screen

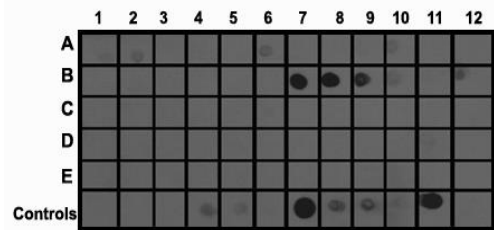
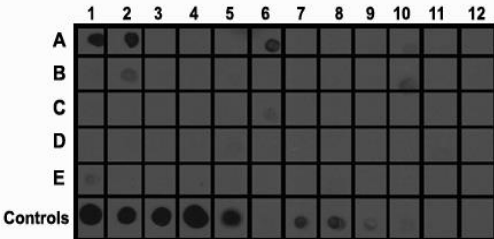
		1	2	3	4	5	6	7	8	9	10	11	12
 <p>PCI</p>	A	TS464	TS491	TS504	TS558	TS589	TS597	TS598	TS599	TS600	TS616	TS620	TS629
	B	TS630	TS631	TS634	TS653	TS654	TS664	TS684	JIP 004	JIP 006	JIP 009	JIP 012	JIP 013
	C	JIP 014	JIP 016	JIP 019	JIP 027	JIP 028	JIP 029	JIP 031	JIP 032	JIP 033	JIP 037	JIP 038	JIP 040
	D	JIP 041	JIP 044	JIP 045	JIP 047	JIP 051	JIP 052	JIP 055	JIP 056	JIP 067	JIP 068	JIP 070	JIP 071
	E	JIP 072	JIP 073	JIP 079	JIP 080	JIP 104	JIP 106	JIP 108	JIP 111	JIP 112	JIP 113	JIP 116	JIP 117
	F	JIP 118	JIP 119	JIP 120	JIP 122	JIP 124	JIP 132	JIP 134	JIP 136	JIP 137	JIP 138	JIP 139	JIP 141
	Controls	JIP 147	<i>PstQ</i>	<i>Pst595</i>	<i>PstDN SP21</i>	<i>PsNW1</i>	<i>PsKM 91</i>	<i>PAO1</i>	<i>216.2B</i>	<i>PstDNS P21 probe</i>	<i>PsNW1 probe</i>	<i>PsKM 91 probe</i>	
 <p><i>Ps. attC</i></p>													

Clinical TS (n=19) sample set: MI screen (See Figure 3.4 for IS1111-attC PCR Screen)											
		1	2	3	4	5	6	7	8	9	10
	A	TS464	TS491	TS504	TS558	TS589	TS597	TS598	TS599	TS600	TS616
	B	TS620	TS629	TS630	TS631	TS634	TS653	TS654	TS664	TS684	
	Controls	<i>Pst405</i>	<i>Pst587</i>	<i>Pst595</i>	<i>Pst641</i>	<i>PsNW1</i>	<i>PAO1</i>	<i>Ps216.2B</i>	<i>PsMB</i>	<i>PstQ</i>	<i>Pfluro</i>
	Controls	<i>R388</i>	Tn7								

Clinical (SSI n=42) sample set: MI and IS1111-attC screen									
		1	2	3	4	5	6	7	8
	A		SSI 1.25	SSI 1.48	SSI 1.55	SSI 1.58	SSI 1.69	SSI 1.90	
	B	SSI 2.01	SSI 2.17	SSI 2.32	SSI 2.37	SSI 2.51	SSI 2.84A	SSI 2.87	SSI 2.96
	C	SSI 2.100	SSI 3.10	SSI 3.14	SSI 3.35	SSI 3.37	SSI 3.38	SSI 3.40	SSI 3.44
	D	SSI 3.53	SSI 3.62	SSI 3.71	SSI 3.72	SSI 3.77	SSI 3.86	SSI 3.91	SSI 4.32
	E	SSI 4.37	SSI 4.46	SSI 4.58	SSI 4.62	SSI 4.63	SSI 4.69	SSI 4.77	SSI 1.95
	Controls	SSI 4.100	SSI 5.32	SSI 5.34	SSI 5.39	<i>Pst405</i>	<i>Pst587</i>	<i>Pst595</i>	<i>R388</i>
	Controls				<i>PsMB</i>	<i>Pfluro</i>			

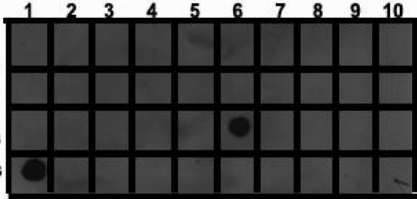
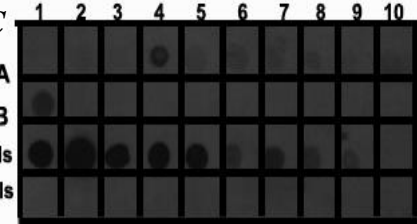
Clinical SSI (n=47) sample set: PCI and <i>Ps. attC</i> screen													
		1	2	3	4	5	6	7	8	9	10	11	12
PCI 	A	TS464	TS491	TS504	TS558	TS589	TS597	TS598	TS599	TS600	TS616	TS620	TS629
	B	TS630	TS631	TS634	TS653	TS654	TS664	TS684	JIP 004	JIP 006	JIP 009	JIP 012	JIP 013
	C	JIP 014	JIP 016	JIP 019	JIP 027	JIP 028	JIP 029	JIP 031	JIP 032	JIP 033	JIP 037	JIP 038	JIP 040
	D	JIP 041	JIP 044	JIP 045	JIP 047	JIP 051	JIP0 052	JIP 055	JIP 056	JIP 067	JIP 068	JIP 070	JIP 071
	Controls	JIP 141	<i>PstQ</i>	<i>Pst405</i>	<i>PstDN SP21</i>	<i>PsNW I</i>	<i>PsKM9 I</i>	<i>PAO1</i>	<i>Ps216. 2B</i>	<i>PstDN SP21 Probe</i>	<i>PsNW I Probe</i>	<i>PsKM 91 Probe</i>	
<i>Ps. attC</i> 	E	JIP 072	JIP 073	JIP 079	JIP 080	JIP 104	JIP 106	JIP 108	JIP 111	JIP 112	JIP 113	JIP 116	JIP 117
	F	JIP 118	JIP 119	JIP 120	JIP 122	JIP 124	SSP 1.28	JIP 132	JIP 134	JIP 136	JIP 137	JIP 138	JIP 139
	Controls	JIP 141	<i>PstQ</i>	<i>Pst405</i>	<i>PstDN SP21</i>	<i>PsNW I</i>	<i>PsKM9 I</i>	<i>PAO1</i>	<i>Ps216. 2B</i>	<i>PstDN SP21 Probe</i>	<i>PsNW I Probe</i>	<i>PsKM 91 Probe</i>	

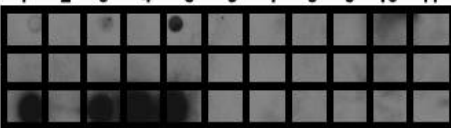
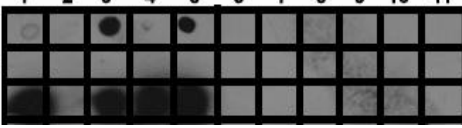
Clinical JIP (n=53) sample set: MI and IS1111-attC screen													
		1	2	3	4	5	6	7	8	9	10	11	12
MI	A												
	B												
	C												
	D												
	E												
Controls													
		1	2	3	4	5	6	7	8	9	10	11	12
	A	JIP 004	JIP 006	JIP 009	JIP 012	JIP 013	JIP 014	JIP 016	JIP 019	JIP 027	JIP 028	JIP 029	JIP 031
	B	JIP 032	JIP 033	JIP 037	JIP 038	JIP 040	JIP 041	JIP 044	JIP 045	JIP 047	JIP 051	JIP 052	JIP 055
	C	JIP 056	JIP 067	JIP 068	JIP 070	JIP 071	JIP 072	JIP 073	JIP 079	JIP 080	JIP 104	JIP 106	JIP 108
	D	JIP 111	JIP 112	JIP 113	JIP 116	JIP 117	JIP 118	JIP 119	JIP 120	JIP 122	JIP 124	SSI 1.28	JIP 132
	E	JIP 134	JIP 136	JIP 137	JIP 138	JIP 139	JIP 141	JIP 147					
Controls		<i>Pst405</i>	<i>Pst587</i>	<i>Pst595</i>	<i>Pst641</i>	<i>PsNW1</i>	<i>PAO1</i>	<i>Ps216.2B</i>	<i>PsMB</i>	<i>PstQ</i>	<i>Pfluro</i>	<i>R388</i>	Tn7
IS1111-attC	A												
	B												
	C												
	D												
	E												
Controls													

Domestic (n=58) sample set: MI and IS1111-attC screen													
MI 		1	2	3	4	5	6	7	8	9	10	11	12
	A	JM1	JM2	JM3	JM4	JM5	JM6	JM7	JM8	JM9	JM10	JM11	JS1
	B	JS2	JS3	JS4	JS5	JS6	JS7	AP1	AP2	AP3	AS1	AS2	AS3
	C	ES1	ES2	ES3	ES4	ES5	MS1	MS2	MS3	MS4	MS5	MS6	N1
	D	NS2	NS3	NS4	NS5	NS6	NS7	NS8	LS1	LS2	LS3	LS4	LS5
	E	LS6	LS7	LS8	LS9	LS10	LS11	LS12	LS13	LS14	LS15		
Controls													
IS1111-attC 	Controls	<i>Pst405</i>	<i>Pst587</i>	<i>Pst595</i>	<i>Pst641</i>	<i>PsNW1</i>	<i>PAO1</i>	<i>Ps216.2B</i>	<i>PsMB</i>	<i>PstQ</i>	<i>Pfluro</i>	R388	Tn7

Domestic (n=58) sample set: PCI screen													
		1	2	3	4	5	6	7	8	9	10	11	12
	A	JM1	JM2	JM3	JM4	JM5	JM6	JM7	JM8	JM9	JM10	JM11	JS1
	B	JS2	JS3	JS4	JS5	JS6	JS7	AP1	AP2	AP3	AS1	AS2	AS3
	C	ES1	ES2	ES3	ES4	ES5	MS1	MS2	MS3	MS4	MS5	MS6	N1
	D	NS2	NS3	NS4	NS5	NS6	NS7	NS8	LS1	LS2	LS3	LS4	LS5
	E	LS6	LS7	LS8	LS9	LS10	LS11	LS12	LS13	LS14	LS15		
	Controls	<i>Pst405</i>	<i>Pst587</i>	<i>Pst595</i>	<i>Pst641</i>	<i>PsNW1</i>	<i>PAO1</i>	<i>PsMB</i>	<i>Ps21.2B</i>	<i>Psfluro</i>	<i>PstQ</i>	<i>R388</i>	

Domestic (n=58) same set: <i>Ps. attC</i> screen													
		1	2	3	4	5	6	7	8	9	10	11	12
	A	JM1	JM2	JM3	JM4	JM5	JM6	JM7	JM8	JM9	JM10	JM11	JS1
	B	JS2	JS3	JS4	JS5	JS6	JS7	AP1	AP2	AP3	AS1	AS2	AS3
	C	ES1	ES2	ES3	ES4	ES5	MS1	MS2	MS3	MS4	MS5	MS6	N1
	D	NS2	NS3	NS4	NS5	NS6	NS7	NS8	LS1	LS2	LS3	LS4	LS5
	E	LS6	LS7	LS8	LS9	LS10	LS11	LS12	LS13	LS14	LS15		
	Controls	<i>Pst405</i>	<i>Pst587</i>	<i>Pst641</i>	<i>PstDN SP21</i>	<i>PsKM9 1</i>	<i>PsNW 1</i>	<i>PAO1</i>	<i>PsMB</i>	<i>Ps216 .2B</i>	<i>PsNW 1</i>	<i>PstQ</i>	<i>R388</i>

Domestic FM (n=11) sample set: MI and IS1111-attC screen											
MI 		1	2	3	4	5	6	7	8	9	10
	A	FM1	FM2	FM3	FM4	FM5	FM6	FM7	FM8	FM9	FM10
	B	FM11									
	Controls	<i>Pst405</i>	<i>Pst587</i>	<i>Pst595</i>	<i>Pst641</i>	<i>PsNW1</i>	<i>PAO1</i>	<i>Ps216.2B</i>	<i>PsMB</i>	<i>PstQ</i>	<i>Pfluro</i>
IS1111-attC 											
	A										
	B										
	Controls	<i>R388</i>	Tn7								

Domestic FM (n=11) sample set: PCI and <i>Ps. attC</i> screen												
PCI 		1	2	3	4	5	6	7	8	9	10	11
	A	FM1	FM2	FM3	FM4	FM5	FM6	FM7	FM8	FM9	FM10	FM11
	Controls	<i>PstQ</i>	<i>Pst405</i>	<i>PstDN SP21</i>	<i>PsNW1</i>	<i>PsKM 91</i>	<i>R388</i>	<i>PAO1</i>	<i>PsMB</i>	<i>Ps216.2B</i>		
<i>Ps. attC</i> 												
	A											
	Control Probes	<i>intPst Q</i>		<i>PstDN SP21</i>	<i>PsNW1</i>	<i>PsKM 91</i>						

Clinical P (n=6) & Clinical Environmental P (n=9) samples set: MI, IS1111-attC and <i>Ps. attC</i> screen											
MI 1 2 3 4 5 6 7 8 9 10 A B Controls Controls											
		1	2	3	4	5	6	7	8	9	10
IS1111-attC 1 2 3 4 5 6 7 8 9 10 A B Controls Controls	A	P01	P02	P04	P05	P06	P07	P08	P09	P10	P11
	B	P13	P14	P15	P17						
<i>Ps. attC</i> 1 2 3 4 5 6 7 8 9 10 A B Controls Controls	Controls	<i>Pst</i> 405	<i>Pst</i> 587	<i>Pst</i> 595	<i>Pst</i> 641	<i>PsNW1</i>	<i>PAO1</i>	<i>Ps216.</i> 2B	<i>PsMB</i>	<i>PstQ</i>	<i>Pfluro</i>
	Controls	R388									

Clinical P (n=6) & Clinical Environmental P (n=9) sample set: PCI screen									
		1	2	3	4	5	6	7	8
A									
	A	P01	P02	P04	P05	P06	P07	P08	P09
B									
	B	P10	P11	P13	P14	P15	P17		
Controls									
Controls		<i>Pst405</i>	<i>Pst587</i>	<i>Pst595</i>	<i>Pst641</i>	<i>PsNW1</i>	<i>PAO1</i>	<i>Ps216.2B</i>	<i>PsMB</i>
Controls									
Controls		<i>PstQ</i>	<i>Pfluro</i>	R388					

Outgroup Enterobacteriaceae (n = 55) sample set: IS1111-attC screen													
		1	2	3	4	5	6	7	8	9	10	11	12
A													
	A	2045	2064	4060	6066	9023	11045	11047	13047	13090	14025	14050 i	14050 ii
	B	15025	15023	15065	15066	15084	16075	18063	18064	18084	19072	19073	20019
	C	20023	20060	20061	22041	22042	23069	24025	25077	25078	26073	27010	27023
	D	27092	28076	31044	33023	33024	33094	33100	35070 W	35070 Y	38085	38763	39010
E													
E		52098	53049	54026	54062								
Controls													
Controls		<i>Pst405</i>	<i>Pst587</i>	<i>Pst595</i>	RNAII I	<i>PsNW1</i>	<i>PsKM 9I</i>	<i>Ps216.2B</i>	<i>PsMB</i>	<i>PAO1</i>	<i>PstQ</i>	<i>Ps fluro</i>	R388

Note: W: white colony morphology, Y: yellow colony morphology

Outgroup Enterobacteriaceae (n = 55) sample set: MI, PCI and *Ps.-attC* screen

MI												
	1	2	3	4	5	6	7	8	9	10	11	12
A												
B												
C												
D												
E												
Controls												

PCI												
	1	2	3	4	5	6	7	8	9	10	11	12
A												
B												
C												
D												
E												
Controls												

<i>Ps.-attC</i>												
	1	2	3	4	5	6	7	8	9	10	11	12
A												
B												
C												
D												
E												
Controls												

	1	2	3	4	5	6	7	8	9	10	11	12
A	2045	2064	4060	6066	9023	11045	11047	13047	13090	14025	14050 i	14050 ii
B	15025	15023	15065	15066	15084	16075	18063	18064	18084	19072	19073	20019
C	20023	20060	20061	22041	22042	23069	24025	25077	25078	26073	27010	27023
D	27092	28076	31044	33023	33024	33094	33100	35070w hite	35070y ellow	38085	38763	39010
E	52098	53049	54026	54062								
Controls	<i>Pst405</i>	<i>Pst595</i>	<i>Pst641</i>	<i>RNAIII</i>	<i>PAO1</i>	<i>Ps216.2B</i>	<i>PsMB</i>	<i>PstDNS P21</i>	<i>PsNW I</i>	<i>PsKM 91</i>	<i>R388</i>	<i>PstQ</i>

APPENDIX A1.3 PCR detection and recovery of MIs (*intI1-attI1* boundaries) and *IS1111-attC* elements

PCR was used to recover *IS1111-attC* as well as MI sequences from positively hybridized clinical and environmental *Pseudomonas*.

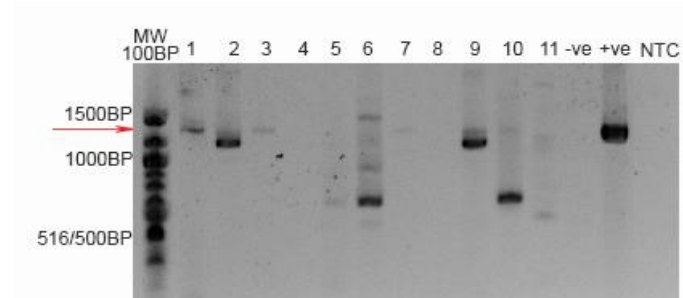


Figure A5: *IS1111-attC* PCR screen of FM domestic sample set. Expected product size is 1340bp (red arrow). Primers ST87 and ST117 were used. Lanes 1-11: FM1-FM11 isolates. -ve is the *PstQ* negative control, +ve is the *Pst587* gDNA positive control, NTC is the no template control where MQ was used as template.

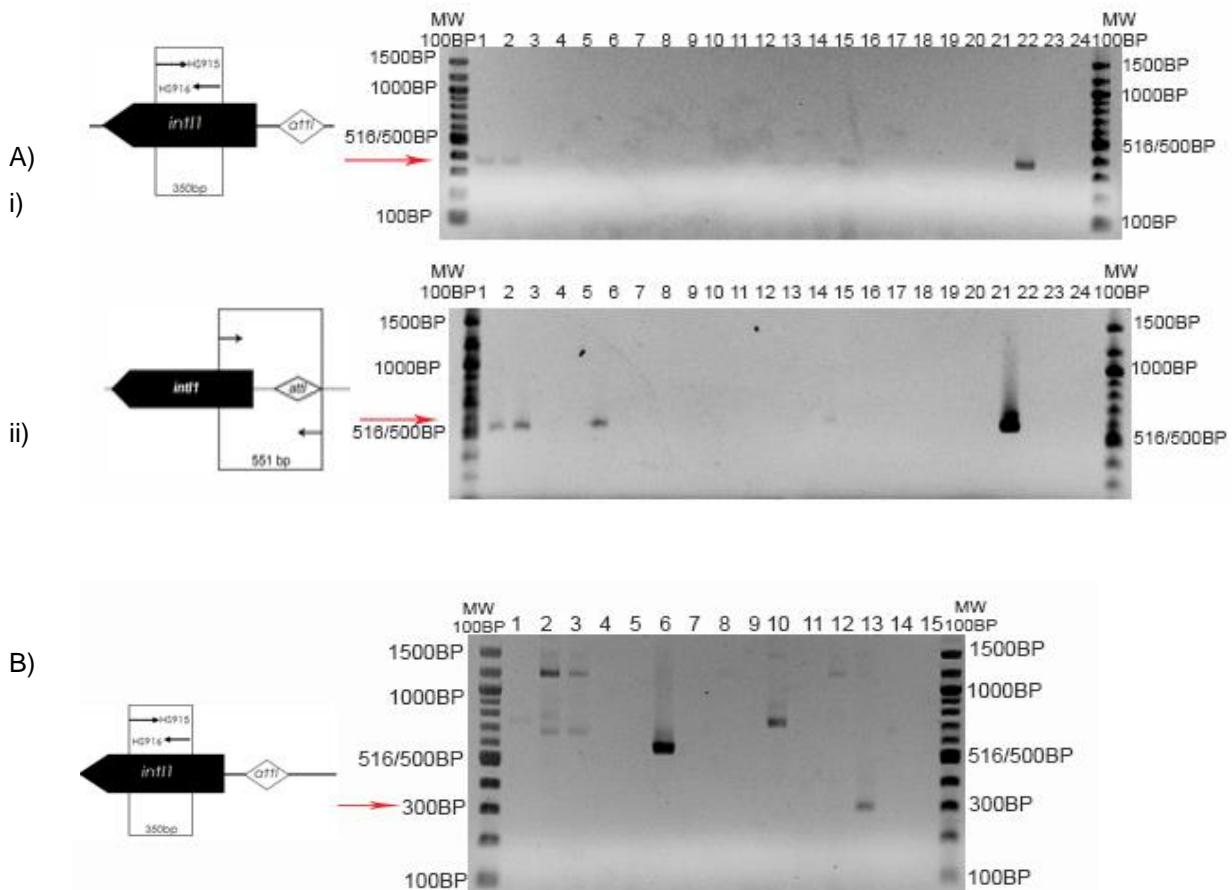


Figure A6: MRI screen using *intI1* primer pairs LW14/15 and HS914 /915. Positive control used was purified genomic DNA from *E. coli* pR388 (lane 22 in A, lane 13 in B). Expected product size is indicated by red arrows.

A) Clinical *Pseudomonas aeruginosa* TS sample set screened with i) HS915/916 & ii) LW14/15 primer pairs.

B) Environmental sample FM sample set screened with HS915/916 primers.

APPENDIX A1.4 Sequence alignments

The nucleotide sequence of the putative IS and MIs were used to search the National Centre for Biotechnology Information (NCBI) BLAST website utilizing BLASTN program.

https://blast.ncbi.nlm.nih.gov/Blast.cgi?CMD=Web&PAGE_TYPE=BlastHome

Multiple sequence alignment was performed using the default settings of the Geneious Alignment module of the Geneious 6.1.5 package, with a gap open penalty of 6 and a gap extension penalty of 3. The *IS1111-attC* alignment was generated using 65% sequence similarity setting. Alignments were manually optimized in the alignment editor. Phylogenetic analyses were performed using Geneious Tree Builder and the Jukes-Cantor genetic distance mode. The tree building method is Neighbor-Joining with an appropriate outgroup as the root. The tree can be seen in Chapter 3 results. Alignments are shown in Figures A7-A10.

	1	10	20	30	40	48	56	66	76
Ps. stutzeri ATCC14405 ISPst6	GGCTGTAATGGACTCTCCCTSCACCACGTTTCTTCTATA----ACTCCTGCGTSACTGTTTCGTGGGAGAGGTGCGATGAAA								
REV Ps. aeruginosa Iso. JIP004	GGGAGACTCGGTATGAGA								
FWD Ps. aeruginosa Iso. JIP009									
REV Ps. aeruginosa Iso. JIP041									
REV Ps. aeruginosa Iso. JIP044									
REV Ps. aeruginosa Iso. JIP045									
REV Ps. aeruginosa Iso. JIP047	GGGAGACTCGGTATGAAA								
REV Ps. aeruginosa Iso. JIP052									
REV Ps. aeruginosa Iso. JIP073									
REV Ps. aeruginosa Iso. JIP117									
FWD Ps. aeruginosa Iso. SSI2.84	TAAATGGACTCTCCCGCACCACGTTTCTTCTATA----ACTCCTGCGTSACTGTTTCGTGGGAGAGGTGCGATGAAA								
FWD Ps. aeruginosa Iso. TS491	GTCGTTTTCGCGGACAGTCAAATGACA								
FWD Ps. aeruginosa Iso. TS558	TAAATGGACTCTCCCGCACCAC---CTTCTATACAAGACATCTGCGT-AGCTGAAATGGGAGACTCGGTATGAAA								
FWD Ps. aeruginosa Iso. AS1	TAAATGGACTCTCCCGCACCAC---CTTCTATACAAGACATCTGCGT-AGCTGAAATGGGAGACTCGGTATGAAA								
FWD Unknown Pseudomonad Iso. FM1	CGAATTAATGGACTCTCCCGCACCACGTTTCTTCTATA----ACTCCTGCGTSACTGTTTCGTGGGAGAGGTGCGATGAAA								
FWD Unknown Pseudomonad Iso. FM3	TAAATGGACTCTCCCGCACCACGTTTCTTCTATA----ACACCTGCGTSACTGTTTCGTGGGAGAGGTGCGATGAAA								
FWD Ps. putida Iso. FM4	TAAATGGACTCTCCCGCACCAC---CTTCTATAAGCCAGACATCTGCGT-AGCTGAAATGGGAGACTCGGTATGAAA								
FWD D. acidovorans SPH-1 Iso. LS8	TAAATGGACTCTCCCGCACCAC---CTTCTATACAAGACATCTGCGT-AGCTGAAATGGGAGACTCGGTATGAAA								
REV Ch. luteola Iso. P05									
REV Ch. meningosepticum Iso.P06									
FWD Ps. mendocina ATCC25411 Iso. P13									
REV Ps. stutzeri ATCC17588 Iso. P15									
	90	100	110	120	130	140	150	160	
	88	96	108	118	128	138	148	158	
Ps. stutzeri ATCC14405 ISPst6	CGCAATCCGANTTCANTTGGCCAAATCCGTTTACCAAGTTGCCGAGAGTCTCCAGGTGGGCGAGSTTCAATCGCCCAAGCGGG								
REV Ps. aeruginosa Iso. JIP004	CGCAATCCAGTTCANTTGGCCAAAATCCGTTTACCAAGTTGCCGAGAGTCTCCAGGTGGGCGAGSTTGTGCAAGCCCAAGCGGG								
FWD Ps. aeruginosa Iso. JIP009	CGAATCCAGTTCANTTGGCAAAAATCCGTTTACCAAGTTGCCGAGAGTCTCCAGGTGGGCGAGSTTGTGCAAGCCCAAGCGGG								
REV Ps. aeruginosa Iso. JIP041									
REV Ps. aeruginosa Iso. JIP044									
REV Ps. aeruginosa Iso. JIP045									
REV Ps. aeruginosa Iso. JIP047	CGCAATCCAGTTCANTTGGCCAAAATCCGTTTACCAAGTTGCCGAGAGTCTCCAGGTGGGCGAGSTTGTGCAAGCCCAAGCGGG								
REV Ps. aeruginosa Iso. JIP052									
REV Ps. aeruginosa Iso. JIP073									
REV Ps. aeruginosa Iso. JIP117									
FWD Ps. aeruginosa Iso. SSI2.84	CGCAATCCGANTTCANTTGGCCAAATCCGTTTACCAAGTTGCCGAGAGTCTCCAGGTGGGCGAGSTTCAATCGCCCAAGCGGG								
FWD Ps. aeruginosa Iso. TS491	CGCAATCCAAATTCANTTGGCCAAATCCGTTTACCAAGTTGCCGAGAGTCTCCAGGTGGGCGAGSTTGTGCAAGCCCAAGCGGG								
FWD Ps. aeruginosa Iso. TS558	CGCAATCCAGTTCANTTGGCCAAAATCCGTTTACCAAGTTGCCGAGAGTCTCCAGGTGGGCGAGSTTGTGCAAGCCCAAGCGGG								
FWD Ps. aeruginosa Iso. AS1	CGCAATCCAGTTCANTTGGCCAAAATCCGTTTACCAAGTTGCCGAGAGTCTCCAGGTGGGCGAGSTTGTGCAAGCCCAAGCGGG								
FWD Unknown Pseudomonad Iso. FM1	CGCAATCCGANTTCANTTGGCCAAATCCGTTTACCAAGTTGCCGAGAGTCTCCAGGTGGGCGAGSTTCAATCGCCCAAGCGGG								
FWD Unknown Pseudomonad Iso. FM3	CGCAATCCGANTTCANTTGGCCAAATCCGTTTACCAAGTTGCCGAGAGTCTCCAGGTGGGCGAGSTTGTGCAAGCCCAAGCGGG								
FWD Ps. putida Iso. FM4	CGCAATCCAGTTCANTTGGCCAAAATCCGTTTACCAAGTTGCCGAGAGTCTCCAGGTGGGCGAGSTTGTGCAAGCCCAAGCGGG								
FWD D. acidovorans SPH-1 Iso. LS8	CGCAATCCAGTTCANTTGGCCAAAATCCGTTTACCAAGTTGCCGAGAGTCTCCAGGTGGGCGAGSTTGTGCAAGCCCAAGCGGG								
REV Ch. luteola Iso. P05									
REV Ch. meningosepticum Iso.P06									
FWD Ps. mendocina ATCC25411 Iso. P13									
REV Ps. stutzeri ATCC17588 Iso. P15									

	168	176	186	196	206	216	226	236	
Ps. stutzeri ATCC14405 ISPst6	TCGATGCCGGGGCA	TTTACGGCATATG	TACAGGAGCAGA	TCCAACC	GTCCAGTGGGTGATGGA	AGCCCTGTGGCACGGCTCA			
REV Ps. aeruginosa Iso. JIP004	TGAACCGAGCAGGGCTTCCGGGCATATA	TACAGGAGCAG	CCGAACC	GTCCAGTGGGTGATGGA	AGCCCTGTGGCACGGCCA				
FWD Ps. aeruginosa Iso. JIP009	TGAACCGAGCAGGGCTTCCGGGCATATA	TACAGGAGCAG	CCGAACC	GTCCAGTGGGTGATGGA	AGCCCTGTGGCACGGCCA				
REV Ps. aeruginosa Iso. JIP041									
REV Ps. aeruginosa Iso. JIP044									
REV Ps. aeruginosa Iso. JIP045									
REV Ps. aeruginosa Iso. JIP047	TGAACCGAGCAGGGCTTCCGGGCATATA	TACAGGAGCAG	CCGAACC	GTCCAGTGGGTGATGGA	AGCCCTGTGGCACGGCCA				
REV Ps. aeruginosa Iso. JIP052									
REV Ps. aeruginosa Iso. JIP073									
REV Ps. aeruginosa Iso. JIP117									
FWD Ps. aeruginosa Iso. SSI2.84	TCGATGCCGGGGCA	TTTACGGCATATG	TACAGGAGCAGA	TCCAACC	GTCCAGTGGGTGATGGA	AGCCCTGTGGCACGGCTCA			
FWD Ps. aeruginosa Iso. TS491	TCGATGCCGGGGCA	TTTACGGCATATG	TACAGGAGCAG	CCGCACC	GTCCAGTGGGTGATGGA	AGCCCTGTGGCACGGCCA			
FWD Ps. aeruginosa Iso. TS558	TGAACCGAGCAGGGCTTCCGGGCATATA	TACAGGAGCAG	CCGAACC	GTCCAGTGGGTGATGGA	AGCCCTGTGGCACGGCCA				
FWD Ps. aeruginosa Iso. AS1	TGAACCGAGCAGGGCTTCCGGGCATATA	TACAGGAGCAG	CCGAACC	GTCCAGTGGGTGATGGA	AGCCCTGTGGCACGGCCA				
FWD Unknown Pseudomonad Iso. FM1	TCGATGCCGGGGCA	TTTACGGCATATG	TACAGGAGCAGA	TCCAACC	GTCCAGTGGGTGATGGA	AGCCCTGTGGCACGGCTCA			
FWD Unknown Pseudomonad Iso. FM3	TCGATGCCGGGGCA	TTTACGGCATATG	TACAGGAGCAGA	TCCAACC	GTCCAGTGGGTGATGGA	AGCCCTGTGGCACAGCTCA			
FWD Ps. putida Iso. FM4	TGAACCGAGCAGGGCTTCCGGGCATATA	TACAGGAGCAG	CCGAACC	GTCCAGTGGGTGATGGA	AGCCCTGTGGCACGGCCA				
FWD D. acidovorans SPH-1 Iso. LS8	TGAACCGAGCAGGGCTTCCGGGCATATA	TACAGGAGCAG	CCGAACC	GTCCAGTGGGTGATGGA	AGCCCTGTGGCACGGCCA				
REV Ch. luteola Iso. P05									
REV Ch. meningosepticum Iso.P06									
FWD Ps. mendocina ATCC25411 Iso. P13									
REV Ps. stutzeri ATCC17588 Iso. P15									
	250	260	270	280	290	300	310	320	
	246	256	266	276	286	296	306	316	
Ps. stutzeri ATCC14405 ISPst6	CCATAGGGCCCGGTGATSCAGS	AAAAAAGCCCATG	GCCTCGTCCCTTA	CACTCC	TGATACCTGCG	TATCCG	CCGG	TATCCG	CCGGCC
REV Ps. aeruginosa Iso. JIP004	CTACAGGGCCCGGTGATSCAGS	AAAAAAGCCCATG	GCCTCGTCCCTTA	CACTCC	TGATACCTGCG	TATCCG	CCGG	TATCCG	CCGGCC
FWD Ps. aeruginosa Iso. JIP009	CTACAGGGCCCGGTGATSCAGS	AAAAAAGCCCATG	GCCTCGTCCCTTA	CACTCC	TGATACCTGCG	TATCCG	CCGG	TATCCG	CCGGCC
REV Ps. aeruginosa Iso. JIP041									
REV Ps. aeruginosa Iso. JIP044									
REV Ps. aeruginosa Iso. JIP045									
REV Ps. aeruginosa Iso. JIP047	CTACAGGGCCCGGTGATSCAGS	AAAAAAGCCCATG	GCCTCGTCCCTTA	CACTCC	TGATACCTGCG	TATCCG	CCGG	TATCCG	CCGGCC
REV Ps. aeruginosa Iso. JIP052									
REV Ps. aeruginosa Iso. JIP073									
REV Ps. aeruginosa Iso. JIP117									
FWD Ps. aeruginosa Iso. SSI2.84	CCATAGGGCCCGGTGATSCAGS	AAAAAAGCCCATG	GCCTCGTCCCTTA	CACTCC	TGATACCTGCG	TATCCG	CCGG	TATCCG	CCGGCC
FWD Ps. aeruginosa Iso. TS491	CTACAGGGCCCGGTGATSCAGS	AAAAAAGCCCATG	GCCTCGTCCCTTA	CACTCC	TGATACCTGCG	TATCCG	CCGG	TATCCG	CCGGCC
FWD Ps. aeruginosa Iso. TS558	CTACAGGGCCCGGTGATSCAGS	AAAAAAGCCCATG	GCCTCGTCCCTTA	CACTCC	TGATACCTGCG	TATCCG	CCGG	TATCCG	CCGGCC
FWD Ps. aeruginosa Iso. AS1	CTACAGGGCCCGGTGATSCAGS	AAAAAAGCCCATG	GCCTCGTCCCTTA	CACTCC	TGATACCTGCG	TATCCG	CCGG	TATCCG	CCGGCC
FWD Unknown Pseudomonad Iso. FM1	CCATAGGGCCCGGTGATSCAGS	AAAAAAGCCCATG	GCCTCGTCCCTTA	CACTCC	TGATACCTGCG	TATCCG	CCGG	TATCCG	CCGGCC
FWD Unknown Pseudomonad Iso. FM3	CCATAGGGCCCGGTGATSCAGS	AAAAAAGCCCATG	GCCTCGTCCCTTA	CACTCC	TGATACCTGCG	TATCCG	CCGG	TATCCG	CCGGCC
FWD Ps. putida Iso. FM4	CTACAGGGCCCGGTGATSCAGS	AAAAAAGCCCATG	GCCTCGTCCCTTA	CACTCC	TGATACCTGCG	TATCCG	CCGG	TATCCG	CCGGCC
FWD D. acidovorans SPH-1 Iso. LS8	CTACAGGGCCCGGTGATSCAGS	AAAAAAGCCCATG	GCCTCGTCCCTTA	CACTCC	TGATACCTGCG	TATCCG	CCGG	TATCCG	CCGGCC
REV Ch. luteola Iso. P05									
REV Ch. meningosepticum Iso.P06									
FWD Ps. mendocina ATCC25411 Iso. P13									
REV Ps. stutzeri ATCC17588 Iso. P15									

	320	330	340	350	360	370	380	390	400
Ps. stutzeri ATCC14405 ISPst6	AACAAGACCGATCGCAAC	SACTGCGAGGCCAT	CTCGAAGCAGCAGGTGG	ACGATATTCATCC	GTGCCAGTCAAAGTG				
REV Ps. aeruginosa Iso. JIP004	AACAAGACCGACCGCAAT	SACTGCGAGGCCAT	CTCGAAGCGCAGGTGG	ACGATATTCATCC	GTGCCAGTCAAAGCC				
FWD Ps. aeruginosa Iso. JIP009	AACAAGACCGACCGCAAT	SACTGCGAGGCCAT	CTCGAAGCGCAGGTGG	ACGATATTCATCC	GTGCCAGTCAAAGCC				
REV Ps. aeruginosa Iso. JIP041									
REV Ps. aeruginosa Iso. JIP044									
REV Ps. aeruginosa Iso. JIP045									
REV Ps. aeruginosa Iso. JIP047	AACAAGACCGACCGCAAT	SACTGCGAGGCCAT	CTCGAAGCGCAGGTGG	ACGATATTCATCC	GTGCCAGTCAAAGCC				
REV Ps. aeruginosa Iso. JIP052									
REV Ps. aeruginosa Iso. JIP073									
REV Ps. aeruginosa Iso. JIP117									
FWD Ps. aeruginosa Iso. SS12.84	AACAAGACCGATCGCAAC	SACTGCGAGGCCAT	CTCGAAGCAGCAGGTGG	ACGATATTCATCC	GTGCCAGTCAAAGTG				
FWD Ps. aeruginosa Iso. TS491	AACAAGACCGACCGCAAT	SACTGCGAGGCCAT	CTCGAAGCGCAGGTGG	ACGATATTCATCC	GTGCCAGTCAAAGCC				
FWD Ps. aeruginosa Iso. TS558	AACAAGACCGACCGCAAT	SACTGCGAGGCCAT	CTCGAAGCGCAGGTGG	ACGATATTCATCC	GTGCCAGTCAAAGCC				
FWD Ps. aeruginosa Iso. AS1	AACAAGACCGACCGCAAT	SACTGCGAGGCCAT	CTCGAAGCGCAGGTGG	ACGATATTCATCC	GTGCCAGTCAAAGCC				
FWD Unknown Pseudomonad Iso. FM1	AACAAGACCGATCGCAAC	SACTGCGAGGCCAT	CTCGAAGCAGCAGGTGG	ACGATATTCATCC	GTGCCAGTCAAAGTG				
FWD Unknown Pseudomonad Iso. FM3	AACAAGACCGATCGCAAC	SACTGCGAGGCCAT	CTCGAAGCAGCAGGTGG	ACGATATTCATCC	GTGCCAGTCAAAGTG				
FWD Ps. putida Iso. FM4	AACAAGACCGACCGCAAT	SACTGCGAGGCCAT	CTCGAAGCGCAGGTGG	ACGATATTCATCC	GTGCCAGTCAAAGCC				
FWD D. acidovorans SPH-1 Iso. LS8	AACAAGACCGACCGCAAT	SACTGCGAGGCCAT	CTCGAAGCGCAGGTGG	ACGATATTCATCC	GTGCCAGTCAAAGCC				
REV Ch. luteola Iso. P05									
REV Ch. meningosepticum Iso.P06									
FWD Ps. mendocina ATCC25411 Iso. P13									
REV Ps. stutzeri ATCC17588 Iso. P15									

	420	430	440	450	460	470	480	490
Ps. stutzeri ATCC14405 ISPst6	TGCAACAGGAGCAA	GTCCAACAGCTGCCATAA	GTTCGGCGAGGGT	TGGAGGAA	GTCGCGGTG	CAACG	ATCAACCTCTGCC	
REV Ps. aeruginosa Iso. JIP004	ACAGCAGCAACAACTGCCAACAGTGGCACGGGC	TCCGCGAAACCTGGAGAAAG	TCGCACTCAGCG	ATCAACCTCTGCC				
FWD Ps. aeruginosa Iso. JIP009	ACAGCAGCAACAACTGCCAACAGTGGCACGGGC	TCCGCGAAACCTGGAGAAAG	TCGCACTCAGCG	ATCAACCTCTGCC				
REV Ps. aeruginosa Iso. JIP041								
REV Ps. aeruginosa Iso. JIP044								
REV Ps. aeruginosa Iso. JIP045								
REV Ps. aeruginosa Iso. JIP047	ACAGCAGCAACAACTGCCAACAGTGGCACGGGC	TCCGCGAAACCTGGAGAAAG	TCGCACTCAGCG	ATCAACCTCTGCC				
REV Ps. aeruginosa Iso. JIP052								
REV Ps. aeruginosa Iso. JIP073								
REV Ps. aeruginosa Iso. JIP117								
FWD Ps. aeruginosa Iso. SS12.84	TGCAACAGGAGCAA	GTCCAACAGCTGCCATAA	GTTCGGCGAGGGT	TGGAGGAA	GTCGCGGTG	CAACG	ATCAACCTCTGCC	
FWD Ps. aeruginosa Iso. TS491	TGCAACAGGAGCAA	GTCCAACAGCTGCCATAA	GTTCGGCGAGGGT	TGGAGGAA	GTCGCGGTG	CAACG	ATCAACCTCTGCC	
FWD Ps. aeruginosa Iso. TS558	ATAGCAGCAACAACTGCCAACAGTGGCACGGGC	TCCGCGAAACCTGGAGAAAG	TCGCACTCAGCG	ATCAACCTCTGCC				
FWD Ps. aeruginosa Iso. AS1	ATAGCAGCAACAACTGCCAACAGTGGCACGGGC	TCCGCGAAACCTGGAGAAAG	TCGCACTCAGCG	ATCAACCTCTGCC				
FWD Unknown Pseudomonad Iso. FM1	TGCAACAGGAGCAA	GTCCAACAGCTGCCATAA	GTTCGGCGAGGGT	TGGAGGAA	GTCGCGGTG	CAACG	ATCAACCTCTGCC	
FWD Unknown Pseudomonad Iso. FM3	TGCAACAGGAGCAA	GTCCAACAGCTGCCATAA	GTTCGGCGAGGGT	TGGAGGAA	GTCGCGGTG	CAACG	ATCAACCTCTGCC	
FWD Ps. putida Iso. FM4	ACAGCAGCAACAACTGCCAACAGTGGCACGGGC	TCCGCGAAACCTGGAGAAAG	TCGCACTCAGCG	ATCAACCTCTGCC				
FWD D. acidovorans SPH-1 Iso. LS8	ATAGCAGCAACAACTGCCAACAGTGGCACGGGC	TCCGCGAAACCTGGAGAAAG	TCGCACTCAGCG	ATCAACCTCTGCC				
REV Ch. luteola Iso. P05								
REV Ch. meningosepticum Iso.P06								
FWD Ps. mendocina ATCC25411 Iso. P13								
REV Ps. stutzeri ATCC17588 Iso. P15								

	498	508	518	525	535	544	554	564	
Ps. stutzeri ATCC14405 ISPst6	CGCCATCC	TCCGTCAGGGTGGGGTTGAA	SCACC	GAA	CGCATCAAGGCATT	TTCCGGCAAA	CGT	SCGAATCAATCGAAC	
REV Ps. aeruginosa Iso. JIP004	CGCTC	NATTCGGTCAAGCCGGTATC	GAGSCCC	CGC	TCCACCGCCGCCTTTA	ATCCGGCC	--	SCA	
FWD Ps. aeruginosa Iso. JIP009	CGCTC	NATTCGGTCAAGCCGGTATC	GAGSCCC	CGC	TCCACCGCCGCCTTTA	ATCCGGCC	--	SCA	
REV Ps. aeruginosa Iso. JIP041									
REV Ps. aeruginosa Iso. JIP044									
REV Ps. aeruginosa Iso. JIP045									
REV Ps. aeruginosa Iso. JIP047	CGCTC	NATTCGGTCAAGCCGGTATC	GAGSCCC	CGC	TCCACCGCCGCCTTTA	ATCCGGCC	--	SCA	
REV Ps. aeruginosa Iso. JIP052									
REV Ps. aeruginosa Iso. JIP073									
REV Ps. aeruginosa Iso. JIP117									
FWD Ps. aeruginosa Iso. SS12.84	CGCCATCC	TCCGTCAGGGTGGGGTTGAA	SCACC	GAA	CGCATCAAGGCATT	TTCCGGCAAA	CGT	SCGAATCAATCGAAC	
FWD Ps. aeruginosa Iso. TS491	CGCCATCC	TCCGTCAGGGTGGGGTTGAA	SCACC	GAA	CGCATCAAGGCATT	TTCCGGCAAA	CGT	SCGAATCAATCGAAC	
FWD Ps. aeruginosa Iso. TS558	CGCTC	NATTCGGTCAAGCCGGTATC	GAGSCCC	CGC	TCCACCGCCGCCTTTA	ATCCGGCC	--	SCA	
FWD Ps. aeruginosa Iso. AS1	CGCTC	NATTCGGTCAAGCCGGTATC	GAGSCCC	CGC	TCCACCGCCGCCTTTA	ATCCGGCC	--	SCA	
FWD Unknown Pseudomonad Iso. FM1	CGCCATCC	TCCGTCAGGGTGGGGTTGAA	SCACC	GAA	CGCATCAAGGCATT	TTCCGGCAAA	CGT	SCGAATCAATCGAAC	
FWD Unknown Pseudomonad Iso. FM3	CGCCATCC	TCCGTCAGGGTGGGGTTGAA	SCACC	GAA	CGCATCAAGGCATT	TTCCGGCAAA	CGT	SCGAATCAATCGAAC	
FWD Ps. putida Iso. FM4	CGCTC	NATTCGGTCAAGCCGGTATC	GAGSCCC	CGC	TCCACCGCCGCCTTTA	ATCCGGCC	--	SCA	
FWD D. acidovorans SPH-1 Iso. LS8	CGCTC	NATTCGGTCAAGCCGGTATC	GAGSCCC	CGC	TCCACCGCCGCCTTTA	ATCCGGCC	--	SCA	
REV Ch. luteola Iso. P05									
REV Ch. meningosepticum Iso.P06	CGCTC	NATTCGGTCAAGCCGGTATC	GAGSCCC	CGC	TCCACCGCCGCCTTTA	ATCCGGCC	--	SCA	
FWD Ps. mendocina ATCC25411 Iso. P13									
REV Ps. stutzeri ATCC17588 Iso. P15									
	580	590	600	610	620	630	640	650	
Ps. stutzeri ATCC14405 ISPst6	GCCCCGAA	GATGC	GTGCAGTGC	GTGGGC	SCGCAAG	TGGTA	CTGGCCGA	GATCAAG	TCTATGAAACAGTGC
REV Ps. aeruginosa Iso. JIP004	A	CCCGA	GCTAG	CACTCC	CCGAGCC	STCT	SCGCAATAT	CTGC	CTGGCCGA
FWD Ps. aeruginosa Iso. JIP009	A	CCCGA	GCTAG	CACTCC	CCGAGCC	STCT	SCGCAATAT	CTGC	CTGGCCGA
REV Ps. aeruginosa Iso. JIP041									
REV Ps. aeruginosa Iso. JIP044									
REV Ps. aeruginosa Iso. JIP045									
REV Ps. aeruginosa Iso. JIP047	A	CCCGA	GCTAG	CACTCC	CCGAGCC	STCT	SCGCAATAT	CTGC	CTGGCCGA
REV Ps. aeruginosa Iso. JIP052									
REV Ps. aeruginosa Iso. JIP073									
REV Ps. aeruginosa Iso. JIP117	G	CCCCGAA	GATGC	GTGCAGTGC	GTGGGC	SCGCAAG	TGGTA	CTGGCCGA	GATCAAG
FWD Ps. aeruginosa Iso. SS12.84	G	CCCCGAA	GATGC	GTGCAGTGC	GTGGGC	SCGCAAG	TGGTA	CTGGCCGA	GATCAAG
FWD Ps. aeruginosa Iso. TS491	G	CCCCGAA	GATGC	GTGCAGTGC	GTGGGC	SCGCAAG	TGGTA	CTGGCCGA	GATCAAG
FWD Ps. aeruginosa Iso. TS558	A	CCCGA	GCTAG	CACTCC	CCGAGCC	STCT	SCGCAATAT	CTGC	CTGGCCGA
FWD Ps. aeruginosa Iso. AS1	A	CCCGA	GCTAG	CACTCC	CCGAGCC	STCT	SCGCAATAT	CTGC	CTGGCCGA
FWD Unknown Pseudomonad Iso. FM1	G	CCCCGAA	GATGC	GTGCAGTGC	GTGGGC	SCGCAAG	TGGTA	CTGGCCGA	GATCAAG
FWD Unknown Pseudomonad Iso. FM3	G	CCCCGAA	GATGC	GTGCAGTGC	GTGGGC	SCGCAAG	TGGTA	CTGGCCGA	GATCAAG
FWD Ps. putida Iso. FM4	A	CCCGA	GCTAG	CACTCC	CCGAGCC	STCT	SCGCAATAT	CTGC	CTGGCCGA
FWD D. acidovorans SPH-1 Iso. LS8	A	CCCGA	GCTAG	CACTCC	CCGAGCC	STCT	SCGCAATAT	CTGC	CTGGCCGA
REV Ch. luteola Iso. P05									
REV Ch. meningosepticum Iso.P06	A	CCCGA	GCTAG	CACTCC	CCGAGCC	STCT	SCGCAATAT	CTGC	CTGGCCGA
FWD Ps. mendocina ATCC25411 Iso. P13	G	CCCCGAA	GATGC	GTGCAGTGC	GTGGGC	SCGCAAG	TGGTA	CTGGCCGA	GATCAAG
REV Ps. stutzeri ATCC17588 Iso. P15	G	CCCCGAA	GATGC	GTGCAGTGC	GTGGGC	SCGCAAG	TGGTA	CTGGCCGA	GATCAAG
	574	584	594	604	614	624	634	644	

	854	864	874	884	894	704	714	724					
Ps. stutzeri ATCC14405 ISPst6	CGAA	GAACA	CTGGCG	CGGTGGCATGC	GACGATGA	ATCGTA	CGCAGGCT	GACFA	STCAGGGCATT	GGCTGCTGACT			
REV Ps. aeruginosa Iso. JIP004	CGAGCA	GACAG	CTCAAAA	CGGTGGCATGC	GACGATGA	ATCGTA	CGCAGGCT	GATFA	STCAGGGCATT	GGCTGCTGACT			
FWD Ps. aeruginosa Iso. JIP009	CGAGCA	GACAG	CTCAAAA	CGGTGGCATGC	GACGATGA	ATCGTA	CGCAGGCT	GATFA	STCAGGGCATT	GGCTGCTGACT			
REV Ps. aeruginosa Iso. JIP041	CGAGCA	GACAG	CTCAAAA	CGGTGGCATGC	GACGATGA	ATCGTA	CGCAGGCT	GATFA	STCAGGGCATT	GGCTGCTGACT			
REV Ps. aeruginosa Iso. JIP044	CGAGCA	GACAG	CTCAAAA	CGGTGGCATGC	GACGATGA	ATCGTA	CGCAGGCT	GATFA	STCAGGGCATT	GGCTGCTGACT			
REV Ps. aeruginosa Iso. JIP045	CGAGCA	GACAG	CTCAAAA	CGGTGGCATGC	GACGATGA	ATCGTA	CGCAGGCT	GATFA	STCAGGGCATT	GGCTGCTGACT			
REV Ps. aeruginosa Iso. JIP047	CGAGCA	GACAG	CTCAAAA	CGGTGGCATGC	GACGATGA	ATCGTA	CGCAGGCT	GATFA	STCAGGGCATT	GGCTGCTGACT			
REV Ps. aeruginosa Iso. JIP052	CGAGCA	GACAG	CTCAAAA	CGGTGGCATGC	GACGATGA	ATCGTA	CGCAGGCT	GATFA	STCAGGGCATT	GGCTGCTGACT			
REV Ps. aeruginosa Iso. JIP073				SATGC	FA	GATGA	ATCGTA	CGCAGGCT	GACFA	STCAGGGCATT	GGCTGCTGACT		
REV Ps. aeruginosa Iso. JIP117	CGAA	GAACA	CTGGCG	CGGTGGCATGC	GACGATGA	ATCGTA	CGCAGGCT	GACFA	STCAGGGCATT	GGCTGCTGACT			
FWD Ps. aeruginosa Iso. SSI2.84	CGAA	GAACA	CTGGCG	CGGTGGCATGC	GACGATGA	ATCGTA	CGCAGGCT	GACFA	STCAGGGCATT	GGCTGCTGACT			
FWD Ps. aeruginosa Iso. TS491	CGAGCA	GACAG	CTCAAAA	CGGTGGCATGC	GACGATGA	ATCGTA	CGCAGGCT	GACFA	STCAGGGCATT	GGCTGCTGACT			
FWD Ps. aeruginosa Iso. TS558	CGAGCA	GACAG	CTCAAAA	CGGTGGCATGC	GACGATGA	ATCGTA	CGCAGGCT	GACFA	STCAGGGCATT	GGCTGCTGACT			
FWD Ps. aeruginosa Iso. AS1	CGAGCA	GACAG	CTCAAAA	CGGTGGCATGC	GACGATGA	ATCGTA	CGCAGGCT	GACFA	STCAGGGCATT	GGCTGCTGACT			
FWD Unknown Pseudomonad Iso. FM1	CGAA	GAACA	CTGGCG	CGGTGGCATGC	GACGATGA	ATCGTA	CGCAGGCT	GACFA	STCAGGGCATT	GGCTGCTGACT			
FWD Unknown Pseudomonad Iso. FM3	CGAA	GAACA	CTGGCG	CGGTGGCATGC	GACGATGA	ATCGTA	CGCAGGCT	GACFA	STCAGGGCATT	GGCTGCTGACT			
FWD Ps. putida Iso. FM4	CGAGCA	GACAG	CTCAAAA	CGGTGGCATGC	GACGATGA	ATCGTA	CGCAGGCT	GATFA	STCAGGGCATT	GGCTGCTGACT			
FWD D. acidovorans SPH-1 Iso. LS8	CGAGCA	GACAG	CTCAAAA	CGGTGGCATGC	GACGATGA	ATCGTA	CGCAGGCT	GATFA	STCAGGGCATT	GGCTGCTGACT			
REV Ch. luteola Iso. P05	CGAA	GAACA	CTGGCG	CGGTGGCATGC	GACGATGA	ATCGTA	CGCAGGCT	GACFA	STCAGGGCATT	GGCTGCTGACT			
REV Ch. meningosepticum Iso.P06	CGAGCA	GACAG	CTCAAAA	CGGTGGCATGC	GACGATGA	ATCGTA	CGCAGGCT	GATFA	STCAGGGCATT	GGCTGCTGACT			
FWD Ps. mendocina ATCC25411 Iso. P13	CGAA	GAACA	CTGGCG	CGGTGGCATGC	GACGATGA	ATCGTA	CGCAGGCT	GACFA	STCAGGGCATT	GGCTGCTGACT			
REV Ps. stutzeri ATCC17588 Iso. P15	CGAA	GAACA	CTGGCG	CGGTGGCATGC	GACGATGA	ATCGTA	CGCAGGCT	GACFA	STCAGGGCATT	GGCTGCTGACT			
	740	750	760	770	780	790	800	810	820				
	733	743	753	763	773	783	792	802	811				
Ps. stutzeri ATCC14405 ISPst6	-GCCAG	TCCC	TGAAAACA	GGCGT	GGCAAA	CCCCAGCCGNT	CCCAGGG	CGG	FA	-CTTA	AGC	CCCAGGGGATGGG	-CATGA
REV Ps. aeruginosa Iso. JIP004	-GCCAG	TCCC	TGAAAACA	GGCGT	GGCAAA	CCCCAGCCGNT	CCCAGGG	CGG	FA	-CTTA	AGC	CCCAGGGGATGGG	-CATGA
FWD Ps. aeruginosa Iso. JIP009	-GCCAG	TCCC	TGAAAACA	GGCGT	GGCAAA	CCCCAGCCGNT	CCCAGGG	CGG	FA	-CTTA	AGC	CCCAGGGGATGGG	-CATGA
REV Ps. aeruginosa Iso. JIP041	-GCCAG	TCCC	TGAAAACA	GGCGT	GGCAAA	CCCCAGCCGNT	CCCAGGG	CGG	FA	-CTTA	AGC	CCCAGGGGATGGG	-CATGA
REV Ps. aeruginosa Iso. JIP044	-GCCAG	TCCC	TGAAAACA	GGCGT	GGCAAA	CCCCAGCCGNT	CCCAGGG	CGG	FA	-CTTA	AGC	CCCAGGGGATGGG	-CATGA
REV Ps. aeruginosa Iso. JIP045	-GCCAG	TCCC	TGAAAACA	GGCGT	GGCAAA	CCCCAGCCGNT	CCCAGGG	CGG	FA	-CTTA	AGC	CCCAGGGGATGGG	-CATGA
REV Ps. aeruginosa Iso. JIP047	-GCCAG	TCCC	TGAAAACA	GGCGT	GGCAAA	CCCCAGCCGNT	CCCAGGG	CGG	FA	-CTTA	AGC	CCCAGGGGATGGG	-CATGA
REV Ps. aeruginosa Iso. JIP052	-GCCAG	TCCC	TGAAAACA	GGCGT	GGCAAA	CCCCAGCCGNT	CCCAGGG	CGG	FA	-CTTA	AGC	CCCAGGGGATGGG	-CATGA
REV Ps. aeruginosa Iso. JIP073	-GCCAG	TCCC	TGAAAACA	GGCGT	GGCAAA	CCCCAGCCGNT	CCCAGGG	CGG	FA	-CTTA	AGC	CCCAGGGGATGGG	-CATGA
REV Ps. aeruginosa Iso. JIP117	-GCCAG	TCCC	TGAAAACA	GGCGT	GGCAAA	CCCCAGCCGNT	CCCAGGG	CGG	FA	-CTTA	AGC	CCCAGGGGATGGG	-CATGA
FWD Ps. aeruginosa Iso. SSI2.84	-GCCAG	TCCC	TGAAAACA	GGCGT	GGCAAA	CCCCAGCCGNT	CCCAGGG	CGG	FA	-CTTA	AGC	CCCAGGGGATGGG	-CATGA
FWD Ps. aeruginosa Iso. TS491	-GCCAG	TCCC	TGAAAACA	GGCGT	GGCAAA	CCCCAGCCGNT	CCCAGGG	CGG	FA	-CTTA	AGC	CCCAGGGGATGGG	-CATGA
FWD Ps. aeruginosa Iso. TS558	-GCCAG	TCCC	TGAAAACA	GGCGT	GGCAAA	CCCCAGCCGNT	CCCAGGG	CGG	FA	-CTTA	AGC	CCCAGGGGATGGG	-CATGA
FWD Ps. aeruginosa Iso. AS1	-GCCAG	TCCC	TGAAAACA	GGCGT	GGCAAA	CCCCAGCCGNT	CCCAGGG	CGG	FA	-CTTA	AGC	CCCAGGGGATGGG	-CATGA
FWD Unknown Pseudomonad Iso. FM1	-GCCAG	TCCC	TGAAAACA	GGCGT	GGCAAA	CCCCAGCCGNT	CCCAGGG	CGG	FA	-CTTA	AGC	CCCAGGGGATGGG	-CATGA
FWD Unknown Pseudomonad Iso. FM3	-GCCAG	TCCC	TGAAAACA	GGCGT	GGCAAA	CCCCAGCCGNT	CCCAGGG	CGG	FA	-CTTA	AGC	CCCAGGGGATGGG	-CATGA
FWD Ps. putida Iso. FM4	-GCCAG	TCCC	TGAAAACA	GGCGT	GGCAAA	CCCCAGCCGNT	CCCAGGG	CGG	FA	-CTTA	AGC	CCCAGGGGATGGG	-CATGA
FWD D. acidovorans SPH-1 Iso. LS8	-GCCAG	TCCC	TGAAAACA	GGCGT	GGCAAA	CCCCAGCCGNT	CCCAGGG	CGG	FA	-CTTA	AGC	CCCAGGGGATGGG	-CATGA
REV Ch. luteola Iso. P05	-GCCAG	TCCC	TGAAAACA	GGCGT	GGCAAA	CCCCAGCCGNT	CCCAGGG	CGG	FA	-CTTA	AGC	CCCAGGGGATGGG	-CATGA
REV Ch. meningosepticum Iso.P06	-GCCAG	TCCC	TGAAAACA	GGCGT	GGCAAA	CCCCAGCCGNT	CCCAGGG	CGG	FA	-CTTA	AGC	CCCAGGGGATGGG	-CATGA
FWD Ps. mendocina ATCC25411 Iso. P13	-GCCAG	TCCC	TGAAAACA	GGCGT	GGCAAA	CCCCAGCCGNT	CCCAGGG	CGG	FA	-CTTA	AGC	CCCAGGGGATGGG	-CATGA
REV Ps. stutzeri ATCC17588 Iso. P15	-GCCAG	TCCC	TGAAAACA	GGCGT	GGCAAA	CCCCAGCCGNT	CCCAGGG	CGG	FA	-CTTA	AGC	CCCAGGGGATGGG	-CATGA

Ps. stutzeri ATCC14405 ISPst6	CGCCCAEGGAATACAGGAGTGGCGGCGGACGCCTAGGAGGAATCAGCTTAAAGGGCAAGCTCTAETGCGCCACCTTGCT
REV Ps. aeruginosa Iso. JIP004	CCCGGCAGGAATTCAGGAGTGGCGGCGGCAAGCTCCGGCAGATCAGCTTAAAGGGCAAGCTCTAETGCGCCACCTTGCT
FWD Ps. aeruginosa Iso. JIP009	CCCGGCAGGAATTCAGGAGTGGCGGCGGCAAGCTCCGGCAGATCAGCTTAAAGGGCAAGCTCTAETGCGCCACCTTGCT
REV Ps. aeruginosa Iso. JIP041	CCCGGCAGGAATTCAGGAGTGGCGGCGGCAAGCTCCGGCAGATCAGCTTAAAGGGCAAGCTCTAETGCGCCACCTTGCT
REV Ps. aeruginosa Iso. JIP044	CCCGGCAGGAATTCAGGAGTGGCGGCGGCAAGCTCCGGCAGATCAGCTTAAAGGGCAAGCTCTAETGCGCCACCTTGCT
REV Ps. aeruginosa Iso. JIP045	CCCGGCAGGAATTCAGGAGTGGCGGCGGCAAGCTCCGGCAGATCAGCTTAAAGGGCAAGCTCTAETGCGCCACCTTGCT
REV Ps. aeruginosa Iso. JIP047	CCCGGCAGGAATTCAGGAGTGGCGGCGGCAAGCTCCGGCAGATCAGCTTAAAGGGCAAGCTCTAETGCGCCACCTTGCT
REV Ps. aeruginosa Iso. JIP052	CGCCCAEGGAATACAGGAGTGGCGGCGGCTGGCGCTTAAAGGGCAAGCTCTAETGCGCCACCTTGCT
REV Ps. aeruginosa Iso. JIP073	CGCCCAEGGAATACAGGAGTGGCGGCGGCTGGCGCTTAAAGGGCAAGCTCTAETGCGCCACCTTGCT
REV Ps. aeruginosa Iso. JIP117	CGCCCAEGGAATACAGGAGTGGCGGCGGCTGGCGCTTAAAGGGCAAGCTCTAETGCGCCACCTTGCT
FWD Ps. aeruginosa Iso. SSI2.84	CGCCCAEGGAATACAGGAGTGGCGGCGGCTGGCGCTTAAAGGGCAAGCTCTAETGCGCCACCTTGCT
FWD Ps. aeruginosa Iso. TS491	CGCCCAEGGAATACAGGAGTGGCGGCGGCTGGCGCTTAAAGGGCAAGCTCTAETGCGCCACCTTGCT
FWD Ps. aeruginosa Iso. TS558	CCCGGCAGGAATTCAGGAGTGGCGGCGTSCAAGCTCCGGCAGATCAGCTTAAAGGGCAAGCTCTAETGCGCCACCTTGCT
FWD Ps. aeruginosa Iso. AS1	CCCGGCAGGAATTCAGGAGTGGCGGCGTSCAAGCTCCGGCAGATCAGCTTAAAGGGCAAGCTCTAETGCGCCACCTTGCT
FWD Unknown Pseudomonad Iso. FM1	CGCCCAEGGAATACAGGAGTGGCGGCGGCTGGCGCTTAAAGGGCAAGCTCTAETGCGCCACCTTGCT
FWD Unknown Pseudomonad Iso. FM3	CGCCCAEGGAATACAGGAGTGGCGGCGGCTGGCGCTTAAAGGGCAAGCTCTAETGCGCCACCTTGCT
FWD Ps. putida Iso. FM4	CCCGGCAGGAATTCAGGAGTGGCGGCGGCAAGCTCCGGCAGATCAGCTTAAAGGGCAAGCTCTAETGCGCCACCTTGCT
FWD D. acidovorans SPH-1 Iso. LS8	CCCGGCAGGAATTCAGGAGTGGCGGCGGCTSCAAGCTCCGGCAGATCAGCTTAAAGGGCAAGCTCTAETGCGCCACCTTGCT
REV Ch. luteola Iso. P05	CCCGGCAGGAATTCAGGAGTGGCGGCGTSCAAGCTCCGGCAGATCAGCTTAAAGGGCAAGCTCTAETGCGCCACCTTGCT
REV Ch. meningosepticum Iso.P06	CCCGGCAGGAATTCAGGAGTGGCGGCGTSCAAGCTCCGGCAGATCAGCTTAAAGGGCAAGCTCTAETGCGCCACCTTGCT
FWD Ps. mendocina ATCC25411 Iso. P13	CGCCCAEGGAATACAGGAGTGGCGGCGGCTGGCGCTTAAAGGGCAAGCTCTAETGCGCCACCTTGCT
REV Ps. stutzeri ATCC17588 Iso. P15	CGCCCAEGGAATACAGGAGTGGCGGCGGCTGGCGCTTAAAGGGCAAGCTCTAETGCGCCACCTTGCT
Ps. stutzeri ATCC14405 ISPst6	GATTCAGGGGACCGAAGCGCGCNACTGGGACCCAGCCGTCAGGGGCGAAGACCGGAGCGGCCTGACCACTTCGAAAGT
REV Ps. aeruginosa Iso. JIP004	ATTCAGGGGTCAGGGGCAAGCTTTCCTGGCTCCACAAAGCCGTCAGGGGCGTACCCGGGAAAAGCTGACCGAGCTTCGAAAGC
FWD Ps. aeruginosa Iso. JIP009	ATTCAGGGGTCAGGGGCAAGCTTTCCTGGCTCCACAAAGCCGTCAGGGGCGTACCCGGGAAAAGCTGACCGAGCTTCGAAAGC
REV Ps. aeruginosa Iso. JIP041	ATTCAGGGGTCAGGGGCAAGCTTTCCTGGCTCCACAAAGCCGTCAGGGGCGTACCCGGGAAAAGCTGACCGAGCTTCGAAAGC
REV Ps. aeruginosa Iso. JIP044	ATTCAGGGGTCAGGGGCAAGCTTTCCTGGCTCCACAAAGCCGTCAGGGGCGTACCCGGGAAAAGCTGACCGAGCTTCGAAAGC
REV Ps. aeruginosa Iso. JIP045	ATTCAGGGGTCAGGGGCAAGCTTTCCTGGCTCCACAAAGCCGTCAGGGGCGTACCCGGGAAAAGCTGACCGAGCTTCGAAAGC
REV Ps. aeruginosa Iso. JIP047	ATTCAGGGGTCAGGGGCAAGCTTTCCTGGCTCCACAAAGCCGTCAGGGGCGTACCCGGGAAAAGCTGACCGAGCTTCGAAAGC
REV Ps. aeruginosa Iso. JIP052	GATTCAGGGGACCGAAGCGCGCCTGCTGGCTCCACAAAGCCGTCAGGGGCGTACCCGGGAAAAGCTGACCGAGCTTCGAAAGC
REV Ps. aeruginosa Iso. JIP073	GATTCAGGGGACCGAAGCGCGCCTGCTGGCTCCACAAAGCCGTCAGGGGCGTACCCGGGAAAAGCTGACCGAGCTTCGAAAGC
REV Ps. aeruginosa Iso. JIP117	GATTCAGGGGACCGAAGCGCGCCTGCTGGCTCCACAAAGCCGTCAGGGGCGTACCCGGGAAAAGCTGACCGAGCTTCGAAAGC
FWD Ps. aeruginosa Iso. SSI2.84	GATTCAGGGGACCGAAGCGCGCCTGCTGGCTCCACAAAGCCGTCAGGGGCGTACCCGGGAAAAGCTGACCGAGCTTCGAAAGC
FWD Ps. aeruginosa Iso. TS491	ATTCAGGGGTCAGGGGCAAGCTTTCCTGGCTCCACAAAGCCGTCAGGGGCGTACCCGGGAAAAGCTGACCGAGCTTCGAAAGC
FWD Ps. aeruginosa Iso. TS558	ATTCAGGGGTCAGGGGCAAGCTTTCCTGGCTCCACAAAGCCGTCAGGGGCGTACCCGGGAAAAGCTGACCGAGCTTCGAAAGC
FWD Ps. aeruginosa Iso. AS1	ATTCAGGGGTCAGGGGCAAGCTTTCCTGGCTCCACAAAGCCGTCAGGGGCGTACCCGGGAAAAGCTGACCGAGCTTCGAAAGC
FWD Unknown Pseudomonad Iso. FM1	GGATTCAGGGGACCGAAGCGCGCCTGCTGGCTCCACAAAGCCGTCAGGGGCGTACCCGGGAAAAGCTGACCGAGCTTCGAAAGC
FWD Unknown Pseudomonad Iso. FM3	GGATTCAGGGGACCGAAGCGCGCCTGCTGGCTCCACAAAGCCGTCAGGGGCGTACCCGGGAAAAGCTGACCGAGCTTCGAAAGC
FWD Ps. putida Iso. FM4	ATTCAGGGGTCAGGGGCAAGCTTTCCTGGCTCCACAAAGCCGTCAGGGGCGTACCCGGGAAAAGCTGACCGAGCTTCGAAAGC
FWD D. acidovorans SPH-1 Iso. LS8	ATTCAGGGGTCAGGGGCAAGCTTTCCTGGCTCCACAAAGCCGTCAGGGGCGTACCCGGGAAAAGCTGACCGAGCTTCGAAAGC
REV Ch. luteola Iso. P05	GGATTCAGGGGACCGAAGCGCGCCTGCTGGCTCCACAAAGCCGTCAGGGGCGTACCCGGGAAAAGCTGACCGAGCTTCGAAAGC
REV Ch. meningosepticum Iso.P06	GGATTCAGGGGACCGAAGCGCGCCTGCTGGCTCCACAAAGCCGTCAGGGGCGTACCCGGGAAAAGCTGACCGAGCTTCGAAAGC
FWD Ps. mendocina ATCC25411 Iso. P13	GGATTCAGGGGACCGAAGCGCGCCTGCTGGCTCCACAAAGCCGTCAGGGGCGTACCCGGGAAAAGCTGACCGAGCTTCGAAAGC
REV Ps. stutzeri ATCC17588 Iso. P15	GGATTCAGGGGACCGAAGCGCGCCTGCTGGCTCCACAAAGCCGTCAGGGGCGTACCCGGGAAAAGCTGACCGAGCTTCGAAAGC

	981	990	1,000	1,010	1,020	1,030	1,040	1,050
Ps. stutzeri ATCC14405 ISPst6	TGGGC	GCACAAACCGCC	GAAAGGCATGGGGCACAACAAGGC	CGCGTGGCC	GGCTGGCC	AGCAAATGGGTG	GGATCTGTGG	
REV Ps. aeruginosa Iso. JIP004	TGGGC	CGACGCAACGGCG	GCGGGCATGGGGCACAACAAGGC	TGCGGTA	GCCGTGGC	AGCAAATGGGTG	GGATCTGTGG	
FWD Ps. aeruginosa Iso. JIP009	TGGGC	GTGACCACGGCG	GCGGGCATGGGGCACAACAAGGC	TGCGGTA	GCCGTGGC	AGCAAATGGGTG	GGATCTGTGG	
REV Ps. aeruginosa Iso. JIP041	TGGGC	CGACGCAACGGCG	GCGGGCATGGGGCACAACAAGGC	TGCGGTA	GCCGTGGC	AGCAAATGGGTG	GGATCTGTGG	
REV Ps. aeruginosa Iso. JIP044	TGGGC	CGACGCAACGGCG	GCGGGCATGGGGCACAACAAGGC	TGCGGTA	GCCGTGGC	AGCAAATGGGTG	GGATCTGTGG	
REV Ps. aeruginosa Iso. JIP045	TGGGC	CGACGCAACGGCG	GCGGGCATGGGGCACAACAAGGC	TGCGGTA	GCCGTGGC	AGCAAATGGGTG	GGATCTGTGG	
REV Ps. aeruginosa Iso. JIP047	TGGGC	CGACGCAACGGCG	GCGGGCATGGGGCACAACAAGGC	TGCGGTA	GCCGTGGC	AGCAAATGGGTG	GGATCTGTGG	
REV Ps. aeruginosa Iso. JIP052	TGGGC	CGACGCAACGGCG	GCGGGCATGGGGCACAACAAGGC	TGCGGTA	GCCGTGGC	AGCAAATGGGTG	GGATCTGTGG	
REV Ps. aeruginosa Iso. JIP073	TGGGC	CGACGCAACGGCG	GAAAGGCATGGGGCACAACAAGGC	TGCGGTA	GCCGTGGC	AGCAAATGGGTG	GGATCTGTGG	
REV Ps. aeruginosa Iso. JIP117	TGGGC	GCACAAACCGCC	GAAAGGCATGGGGCACAACAAGGC	CGCGTGGCC	GGCTGGCC	AGCAAATGGGTG	GGATCTGTGG	
FWD Ps. aeruginosa Iso. SS12.84	TGGGC	CGACGCAACGGCG	GAAAGGCATGGGGCACAACAAGGC	TGCGGTA	GCCGTGGC	AGCAAATGGGTG	GGATCTGTGG	
FWD Ps. aeruginosa Iso. TS491	TGGGC	CGACGCAACGGCG	GCGGGCATGGGGCACAACAAGGC	TGCGGTA	GCCGTGGC	AGCAAATGGGTG	GGATCTGTGG	
FWD Ps. aeruginosa Iso. TS558	TGGGC	CGACGCAACGGCG	GCGGGCATGGGGCACAACAAGGC	TGCGGTA	GCCGTGGC	AGCAAATGGGTG	GGATCTGTGG	
FWD Ps. aeruginosa Iso. AS1	TGGGC	CGACGCAACGGCG	GCGGGCATGGGGCACAACAAGGC	TGCGGTA	GCCGTGGC	AGCAAATGGGTG	GGATCTGTGG	
FWD Unknown Pseudomonad Iso. FM1	TGGGC	GCACAAACCGCC	GAAAGGCATGGGGCACAACAAGGC	CGCGTGGCC	GGCTGGCC	AGCAAATGGGTG	GGATCTGTGG	
FWD Unknown Pseudomonad Iso. FM3	TGGGC	ACACGCAACGGCG	GAAAGGCATGGGGCACAACAAGGC	CGCGTGGCC	GGCTGGCC	AGCAAATGGGTG	GGATCTGTGG	
FWD Ps. putida Iso. FM4	TGGGC	CGACGCAACGGCG	GCGGGCATGGGGCACAACAAGGC	TGCGGTA	GCCGTGGC	AGCAAATGGGTG	GGATCTGTGG	
FWD D. acidovorans SPH-1 Iso. LS8	TGGGC	CGACGCAACGGCG	GCGGGCATGGGGCACAACAAGGC	TGCGGTA	GCCGTGGC	AGCAAATGGGTG	GGATCTGTGG	
REV Ch. luteola Iso. P05	TGGGC	GCACAAACCGCC	GAAAGGCATGGGGCACAACAAGGC	CGCGTGGCC	GGCTGGCC	AGCAAATGGGTG	GGATCTGTGG	
REV Ch. meningosepticum Iso.P06	TGGGC	CGACGCAACGGCG	GCGGGCATGGGGCACAACAAGGC	TGCGGTA	GCCGTGGC	AGCAAATGGGTG	GGATCTGTGG	
FWD Ps. mendocina ATCC25411 Iso. P13	TGGGC	GCACAAACCGCC	GAAAGGCATGGGGCACAACAAGGC	CGCGTGGCC	GGCTGGCC	AGCAAATGGGTG	GGATCTGTGG	
REV Ps. stutzeri ATCC17588 Iso. P15	TGGGC	GCACAAACCGCC	GAAAGGCATGGGGCACAACAAGGC	CGCGTGGCC	GGCTGGCC	AGCAAATGGGTG	GGATCTGTGG	
	1,070	1,080	1,090	1,100	1,110	1,120	1,130	1,140
Ps. stutzeri ATCC14405 ISPst6	CCGCTGTGGTGGCCATGCA	CGGATTCAG	CGGTTAACTGGCA	AGCAGCAAAGCCGGCCGCA	GGCGA	GGTA	-----	AAT
REV Ps. aeruginosa Iso. JIP004	CCGCTGTGGTGGCCATGCA	CGGATTCAG	CGGTTAACTGGCA	AGCAGCAAAGCCGGCCGCA	GGCGA	GGTA	-----	AAT
FWD Ps. aeruginosa Iso. JIP009	CCGCTGTGGTGGCCATGCA	CGGATTCAG	CGGTTAACTGGCA	AGCAGCAAAGCCGGCCGCA	GGCGA	GGTA	-----	AAT
REV Ps. aeruginosa Iso. JIP041	CCGCTGTGGTGGCCATGCA	CGGATTCAG	CGGTTAACTGGCA	AGCAGCAAAGCCGGCCGCA	GGCGA	GGTA	-----	AAT
REV Ps. aeruginosa Iso. JIP044	CCGCTGTGGTGGCCATGCA	CGGATTCAG	CGGTTAACTGGCA	AGCAGCAAAGCCGGCCGCA	GGCGA	GGTA	-----	AAT
REV Ps. aeruginosa Iso. JIP045	CCGCTGTGGTGGCCATGCA	CGGATTCAG	CGGTTAACTGGCA	AGCAGCAAAGCCGGCCGCA	GGCGA	GGTA	-----	AAT
REV Ps. aeruginosa Iso. JIP047	CCGCTGTGGTGGCCATGCA	CGGATTCAG	CGGTTAACTGGCA	AGCAGCAAAGCCGGCCGCA	GGCGA	GGTA	-----	AAT
REV Ps. aeruginosa Iso. JIP052	CCGCTGTGGTGGCCATGCA	CGGATTCAG	CGGTTAACTGGCA	AGCAGCAAAGCCGGCCGCA	GGCGA	GGTA	-----	AAT
REV Ps. aeruginosa Iso. JIP073	CCGCTGTGGTGGCCATGCA	CGGATTCAG	CGGTTAACTGGCA	AGCAGCAAAGCCGGCCGCA	GGCGA	GGTA	-----	AAT
REV Ps. aeruginosa Iso. JIP117	CCGCTGTGGTGGCCATGCA	CGGATTCAG	CGGTTAACTGGCA	AGCAGCAAAGCCGGCCGCA	GGCGA	GGTA	-----	AAT
FWD Ps. aeruginosa Iso. SS12.84	CCGCTGTGGTGGCCATGCA	CGGATTCAG	CGGTTAACTGGCA	AGCAGCAAAGCCGGCCGCA	GGCGA	GGTA	-----	AAT
FWD Ps. aeruginosa Iso. TS491	CCGCTGTGGTGGCCATGCA	CGGATTCAG	CGGTTAACTGGCA	AGCAGCAAAGCCGGCCGCA	GGCGA	GGTA	-----	AAT
FWD Ps. aeruginosa Iso. TS558	CCGCTGTGGTGGCCATGCA	CGGATTCAG	CGGTTAACTGGCA	AGCAGCAAAGCCGGCCGCA	GGCGA	GGTA	-----	AAT
FWD Ps. aeruginosa Iso. AS1	CCGCTGTGGTGGCCATGCA	CGGATTCAG	CGGTTAACTGGCA	AGCAGCAAAGCCGGCCGCA	GGCGA	GGTA	-----	AAT
FWD Unknown Pseudomonad Iso. FM1	CCGCTGTGGTGGCCATGCA	CGGATTCAG	CGGTTAACTGGCA	AGCAGCAAAGCCGGCCGCA	GGCGA	GGTA	-----	AAT
FWD Unknown Pseudomonad Iso. FM3	CCGCTGTGGTGGCCATGCA	CGGATTCAG	CGGTTAACTGGCA	AGCAGCAAAGCCGGCCGCA	GGCGA	GGTA	-----	AAT
FWD Ps. putida Iso. FM4	CCGCTGTGGTGGCCATGCA	CGGATTCAG	CGGTTAACTGGCA	AGCAGCAAAGCCGGCCGCA	GGCGA	GGTA	-----	AAT
FWD D. acidovorans SPH-1 Iso. LS8	CCGCTGTGGTGGCCATGCA	CGGATTCAG	CGGTTAACTGGCA	AGCAGCAAAGCCGGCCGCA	GGCGA	GGTA	-----	AAT
REV Ch. luteola Iso. P05	CCGCTGTGGTGGCCATGCA	CGGATTCAG	CGGTTAACTGGCA	AGCAGCAAAGCCGGCCGCA	GGCGA	GGTA	-----	AAT
REV Ch. meningosepticum Iso.P06	CCGCTGTGGTGGCCATGCA	CGGATTCAG	CGGTTAACTGGCA	AGCAGCAAAGCCGGCCGCA	GGCGA	GGTA	-----	AAT
FWD Ps. mendocina ATCC25411 Iso. P13	CCGCTGTGGTGGCCATGCA	CGGATTCAG	CGGTTAACTGGCA	AGCAGCAAAGCCGGCCGCA	GGCGA	GGTA	-----	AAT
REV Ps. stutzeri ATCC17588 Iso. P15	CCGCTGTGGTGGCCATGCA	CGGATTCAG	CGGTTAACTGGCA	AGCAGCAAAGCCGGCCGCA	GGCGA	GGTA	-----	AAT

	1,131	1,141	1,148	1,158	1,168	1,178	1,188	1,198	1,208
Ps. stutzeri ATCC14405 ISPst6	AA TGAGG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG
REV Ps. aeruginosa Iso. JIP004	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG
FWD Ps. aeruginosa Iso. JIP009	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG
REV Ps. aeruginosa Iso. JIP041	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG
REV Ps. aeruginosa Iso. JIP044	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG
REV Ps. aeruginosa Iso. JIP045	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG
REV Ps. aeruginosa Iso. JIP047	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG
REV Ps. aeruginosa Iso. JIP052	AGGAAGAC	AGGAAGAC	AGGAAGAC	AGGAAGAC	AGGAAGAC	AGGAAGAC	AGGAAGAC	AGGAAGAC	AGGAAGAC
REV Ps. aeruginosa Iso. JIP073	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG
REV Ps. aeruginosa Iso. JIP117	AA TGAGG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG
FWD Ps. aeruginosa Iso. SS12.84	AGGAAGAC	AGGAAGAC	AGGAAGAC	AGGAAGAC	AGGAAGAC	AGGAAGAC	AGGAAGAC	AGGAAGAC	AGGAAGAC
FWD Ps. aeruginosa Iso. TS491	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG
FWD Ps. aeruginosa Iso. TS558	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG
FWD Ps. aeruginosa Iso. AS1	AA TGAGG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG
FWD Unknown Pseudomonad Iso. FM1	AA TGAGG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG
FWD Unknown Pseudomonad Iso. FM3	AA TGAGG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG
FWD Ps. putida Iso. FM4	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG
FWD D. acidovorans SPH-1 Iso. LS8	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG
REV Ch. luteola Iso. P05	AA TGAGG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG
REV Ch. meningosepticum Iso.P06	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG
FWD Ps. mendocina ATCC25411 Iso. P13	AA TGAGG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG
REV Ps. stutzeri ATCC17588 Iso. P15	AA TGAGG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG	GGCAAGAG
	1,240	1,240	1,250	1,260	1,270	1,280	1,290	1,300	1,310
Ps. stutzeri ATCC14405 ISPst6	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT
REV Ps. aeruginosa Iso. JIP004	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT
FWD Ps. aeruginosa Iso. JIP009	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT
REV Ps. aeruginosa Iso. JIP041	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT
REV Ps. aeruginosa Iso. JIP044	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT
REV Ps. aeruginosa Iso. JIP045	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT
REV Ps. aeruginosa Iso. JIP047	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT
REV Ps. aeruginosa Iso. JIP052	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT
REV Ps. aeruginosa Iso. JIP073	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT
REV Ps. aeruginosa Iso. JIP117	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT
FWD Ps. aeruginosa Iso. SS12.84	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT
FWD Ps. aeruginosa Iso. TS491	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT
FWD Ps. aeruginosa Iso. TS558	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT
FWD Ps. aeruginosa Iso. AS1	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT
FWD Unknown Pseudomonad Iso. FM1	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT
FWD Unknown Pseudomonad Iso. FM3	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT
FWD Ps. putida Iso. FM4	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT
FWD D. acidovorans SPH-1 Iso. LS8	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT
REV Ch. luteola Iso. P05	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT
REV Ch. meningosepticum Iso.P06	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT
FWD Ps. mendocina ATCC25411 Iso. P13	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT
REV Ps. stutzeri ATCC17588 Iso. P15	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT	AACAAT
	1,218	1,228	1,238	1,248	1,258	1,268	1,278	1,285	

	1,294	1,304	1,314	1,324	1,334	1,344	1,350	1,360
Ps. stutzeri ATCC14405 ISPst6	ACA	GGCCGGATATACGAATGCAACCGCGCTGAG	GCAGGATCG	GAA	AAAAGCTTTTC	---	STCACT	CTTGTGGGGAGAGT
REV Ps. aeruginosa Iso. JIP004	ACA	GGCCGGATATACGAATGCAACCGG						
FWD Ps. aeruginosa Iso. JIP009	ACA	GGCCGGATATACGAATGCAACCGCGCT						
REV Ps. aeruginosa Iso. JIP041	ACA	GGCCGGATATACGAATGCAACCGCGCTGAG	GCAGGATCG	---	AAAAGCAATTC	CTTTG	STCACT	CTTGTGGGGAGAGT
REV Ps. aeruginosa Iso. JIP044	ACA	GGCCGGATATACGAATGCAACCGCGCTGAG	GCAGGATCG	---	AAAAGCAATTC	CTTTG	STCACT	CTTGTGGGGAGAGT
REV Ps. aeruginosa Iso. JIP045	ACA	GGCCGGATATACGAATGCAACCGCGCTGAG	GCAGGATCG	---	AAAAGCAATTC	CTTTG	STCACT	CTTGTGGGGAGAGT
REV Ps. aeruginosa Iso. JIP047	ACA	GGCCGGATATACGAATGCAACCGCGCT						
REV Ps. aeruginosa Iso. JIP052	ACA	GGCCGGATATACGAATGCAACCGCGCTGAG	GCAGGATCG	GAA	AAAAGCTTTTC	---	STCACT	CTTGTGGGGAGAGT
REV Ps. aeruginosa Iso. JIP073	ACA	GGCCGGATATACGAATGCAACCGCGCTGAG	GCAGGATCG	---	AAAAGCAATTC	CTTTG	STCACT	CTTGTGGGGAGAGT
REV Ps. aeruginosa Iso. JIP117	ACA	GGCCGGATATACGAATGCAACCGCGCTGAG	GCAGGATCG	GAA	AAAAGCTTTTC	---	STCACT	CTTGTGGGGAGAGT
FWD Ps. aeruginosa Iso. SSI2.84	ACA	GGCCGGATATACGAATGCAACCGCGCTGAG	GCAGGATCG					
FWD Ps. aeruginosa Iso. TS491	ACA	GGCCGGATATACGAATGCAACCGCGCTGAG	GCAGGATCG	---	AAAAGCAATTC	CTTTG	STCACT	CTTGTGGGGAGAGT
FWD Ps. aeruginosa Iso. TS558	ACA	GGCCGGATATACGAATGCAACCGCGCTGAG	GCAGGATCG	---	AAAAGCAATTC	CTTTG	STCACT	CTTGTGGGGAGAGT
FWD Ps. aeruginosa Iso. AS1	ACA	GGCCGGATATACGAATGCAACCGCGCTGAG	GCAGGATCG	---	AAAAGCAATTC	CTTTG	STCACT	CTTGTGGGGAGAGT
FWD Unknown Pseudomonad Iso. FM1	ACA	GGCCGGATATACGAATGCAACCGCGCTGAG	GCAGGATCG	GAA	AAAAGCTTTTC	---	STCACT	CTTGTGGGGAGAGT
FWD Unknown Pseudomonad Iso. FM3	ACA	GGCCGGATATACGAATGCAACCGCGCTGAG	GCAGGATCG	---	AAAAGCAATTC	CTTTG	STCACT	CTTGTGGGGAGAGT
FWD Ps. putida Iso. FM4	ACA	GGCCGGATATACGAATGCAACCGCGCTGAG	GCAGGATCG	---	AAAAGCAATTC	CTTTG	STCACT	CTTGTGGGGAGAGT
FWD D. acidovorans SPH-1 Iso. LS8	ACA	GGCCGGATATACGAATGCAACCGCGCTGAG	GCAGGATCG	---	AAAAGCAATTC	CTTTG	STCACT	CTTGTGGGGAGAGT
REV Ch. luteola Iso. P05	ACA	GGCCGGATATACGAATGCAACCGCGCTGAG	GCAGGATCG	GAA	AAAAGCTTTTC	---	STCACT	CTTGTGGGGAGAGT
REV Ch. meningosepticum Iso.P06	ACA	GGCCGGATATACGAATGCAACCGCGCTGAG	GCAGGATCG	---	AAAAGCAATTC	CTTTG	STCACT	CTTGTGGGGAGAGT
FWD Ps. mendocina ATCC25411 Iso. P13	ACA	GGCCGGATATACGAATGCAACCGCGCTGAG	GCAGGATCG	GAA	AAAAGCTTTTC	---	STCACT	CTTGTGGGGAGAGT
REV Ps. stutzeri ATCC17588 Iso. P15	ACA	GGCCGGATATACGAATGCAACCGCGCTGAG	GCAGGATCG	GAA	AAAAGCTTTTC	---	STCACT	CTTGTGGGGAGAGT
		1,401						
		1,371						
Ps. stutzeri ATCC14405 ISPst6	CCATATA							
REV Ps. aeruginosa Iso. JIP004	CCATATA							
FWD Ps. aeruginosa Iso. JIP009	CCATATA							
REV Ps. aeruginosa Iso. JIP041	CCATATA							
REV Ps. aeruginosa Iso. JIP044	CCATATA							
REV Ps. aeruginosa Iso. JIP045	CCATATA							
REV Ps. aeruginosa Iso. JIP047	CCATATA							
REV Ps. aeruginosa Iso. JIP052	CCATATA							
REV Ps. aeruginosa Iso. JIP073	CCATATA							
REV Ps. aeruginosa Iso. JIP117	CCATATA							
FWD Ps. aeruginosa Iso. SSI2.84	CCATATA							
FWD Ps. aeruginosa Iso. TS491	CCATATA							
FWD Ps. aeruginosa Iso. TS558	CCATATA							
FWD Ps. aeruginosa Iso. AS1	CCATATA							
FWD Unknown Pseudomonad Iso. FM1	CCATATA							
FWD Unknown Pseudomonad Iso. FM3	CCATATA							
FWD Ps. putida Iso. FM4	CCATATA							
FWD D. acidovorans SPH-1 Iso. LS8	CCATATA							
REV Ch. luteola Iso. P05	CCATATA							
REV Ch. meningosepticum Iso.P06	CCATATA							
FWD Ps. mendocina ATCC25411 Iso. P13	CCATATA							
REV Ps. stutzeri ATCC17588 Iso. P15	CCATATA							

Figure A7: Sequence alignment of IS1111-attC elements from clinical and environmental isolates. The alignment was generated using Geneious software. All sequences were mapped against Ps. stutzeri ATCC14405 ISPst6 sequence. Ps. aeruginosa isolates JIP, SSI, TS are all clinically relevant. Isolates AS, P05, P06, P13 and P15 are clinical environmental isolates. Finally isolates FM and LS are domestic environmental isolates.

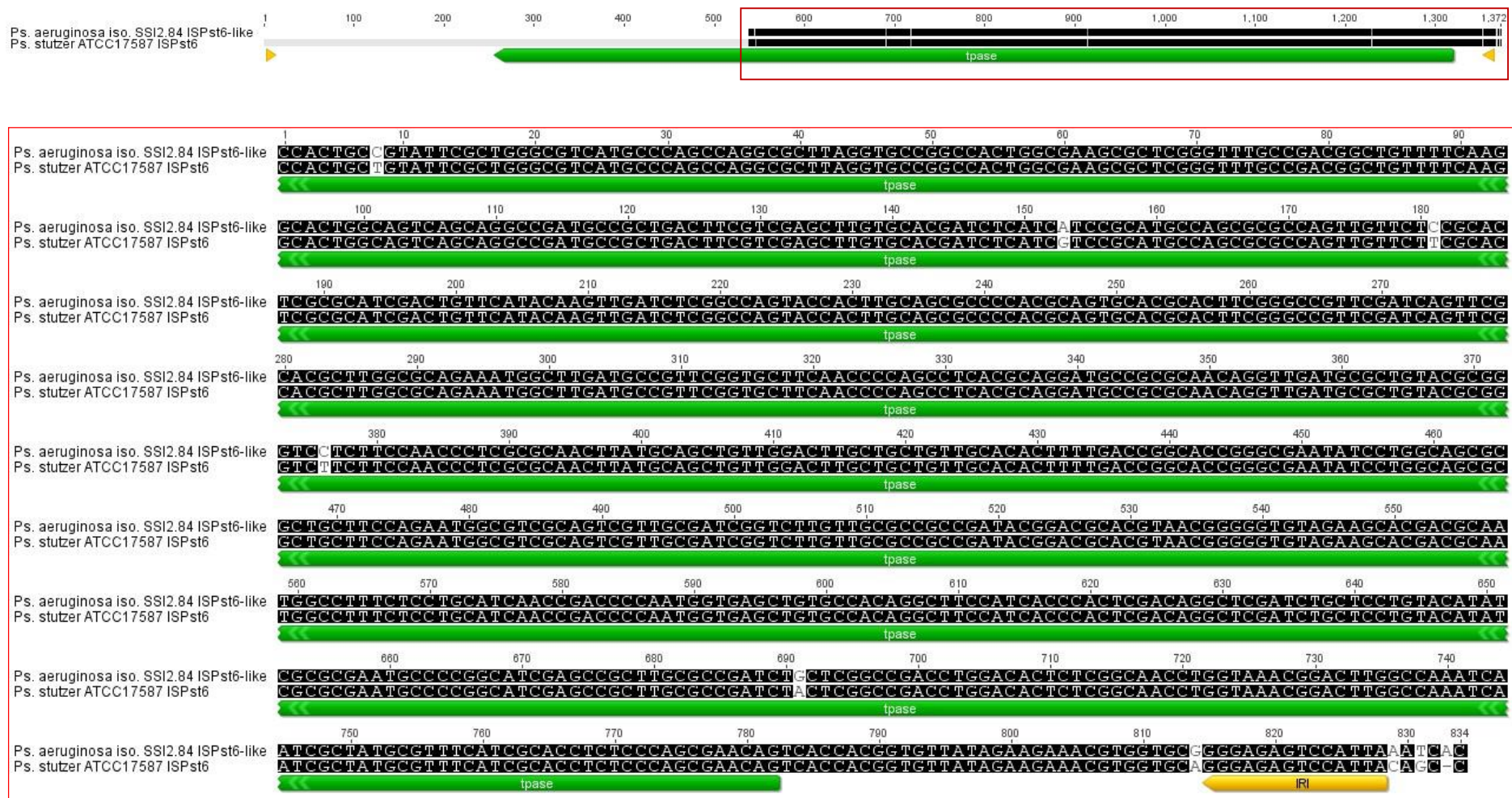


Figure A8: a) schematic and b) alignment of *Pseudomonas stutzeri* ATCC17587 ISPst6 sequence and recovered ISPst6-like from human clinical isolate *Pseudomonas aeruginosa* SSI2.84. Sequence alignment is based on 834 nucleotide bases and displays 98% (825/834) sequence identity.

		200	210	220	230	240	250	260	*	*	270	280	*	290			
		CACCGC	TCGCGT	TCGGTCA	AGTTCT	GTGAAC	CAGTTG	CCGTGAG	CGGATA	CGCTAC	TTGCATT	TACAGY	AACCGA	ACAGGG	CTTATG	TCAC	TGGGTT
E. coli pR388 Int1 and attL		CACCGC	TCGCGT	TCGGTCA	AGTTCT	GTGAAC	CAGTTG	CCGTGAG	CGGATA	CGCTAC	TTGCATT	TACAG	AACCGA	ACAGGG	CTTATG	TCAC	TGGGTT
Ps. aeruginosa Iso. JIP041		CACCGC	TCGCGT	TCGGTCA	AGTTCT	GTGAAC	CAGTTG	CCGTGAG	CGGATA	CGCTAC	TTGCATT	TACAG	AACCGA	ACAGGG	CTTATG	TCAC	TGGGTT
Ps. aeruginosa Iso. JIP044		CACCGC	TCGCGT	TCGGTCA	AGTTCT	GTGAAC	CAGTTG	CCGTGAG	CGGATA	CGCTAC	TTGCATT	TACAG	AACCGA	ACAGGG	CTTATG	TCAC	TGGGTT
Ps. aeruginosa Iso. JIP045		CACCGC	TCGCGT	TCGGTCA	AGTTCT	GTGAAC	CAGTTG	CCGTGAG	CGGATA	CGCTAC	TTGCATT	TACAG	AACCGA	ACAGGG	CTTATG	TCAC	TGGGTT
Ps. aeruginosa Iso. JIP052		CACCGC	TCGCGT	TCGGTCA	AGTTCT	GTGAAC	CAGTTG	CCGTGAG	CGGATA	CGCTAC	TTGCATT	TACAG	AACCGA	ACAGGG	CTTATG	TCAC	TGGGTT
Ps. aeruginosa Iso. JIP073		CACCGC	TCGCGT	TCGGTCA	AGTTCT	GTGAAC	CAGTTG	CCGTGAG	CGGATA	CGCTAC	TTGCATT	TACAG	AACCGA	ACAGGG	CTTATG	TCAC	TGGGTT
Ps. aeruginosa Iso. JIP080		CACCGC	TCGCGT	TCGGTCA	AGTTCT	GTGAAC	CAGTTG	CCGTGAG	CGGATA	CGCTAC	TTGCATT	TACAG	AACCGA	ACAGGG	CTTATG	TCAC	TGGGTT
Ps. aeruginosa Iso. SSI 2.37		CACCGC	TCGCGT	TCGGTCA	AGTTCT	GTGAAC	CAGTTG	CCGTGAG	CGGATA	CGCTAC	TTGCATT	TACAG	AACCGA	ACAGGG	CTTATG	TCAC	TGGGTT
Ps. aeruginosa Iso. SSI 3.34		CACCGC	TCGCGT	TCGGTCA	AGTTCT	GTGAAC	CAGTTG	CCGTGAG	CGGATA	CGCTAC	TTGCATT	TACAG	AACCGA	ACAGGG	CTTATG	TCAC	TGGGTT
Ps. aeruginosa Iso. SSI 5.34		CACCGC	TCGCGT	TCGGTCA	AGTTCT	GTGAAC	CAGTTG	CCGTGAG	CGGATA	CGCTAC	TTGCATT	TACAG	AACCGA	ACAGGG	CTTATG	TCAC	TGGGTT
St. maltophilia Iso. AP2		CACCGC	TCGCGT	TCGGTCA	AGTTCT	GTGAAC	CAGTTG	CCGTGAG	CGGATA	CGCTAC	TTGCATT	TACAG	AACCGA	ACAGGG	CTTATG	TCAC	TGGGTT
Ps. putida ND6 Iso. AP3		CACCGC	TCGCGT	TCGGTCA	AGTTCT	GTGAAC	CAGTTG	CCGTGAG	CGGATA	CGCTAC	TTGCATT	TACAG	AACCGA	ACAGGG	CTTATG	TCAC	TGGGTT
Ps. aeruginosa Iso. AS1		CACCGC	TCGCGT	TCGGTCA	AGTTCT	GTGAAC	CAGTTG	CCGTGAG	CGGATA	CGCTAC	TTGCATT	TACAG	AACCGA	ACAGGG	CTTATG	TCAC	TGGGTT
Ps. aeruginosa Iso. AS2		CACCGC	TCGCGT	TCGGTCA	AGTTCT	GTGAAC	CAGTTG	CCGTGAG	CGGATA	CGCTAC	TTGCATT	TACAG	AACCGA	ACAGGG	CTTATG	TCAC	TGGGTT
Ps. sp. N3 Iso. JM10		CACCGC	TCGCGT	TCGGTCA	AGTTCT	GTGAAC	CAGTTG	CCGTGAG	CGGATA	CGCTAC	TTGCATT	TACAG	AACCGA	ACAGGG	CTTATG	TCAC	TGGGTT
Br. vesicularis Iso. LS13		CACCGC	TCGCGT	TCGGTCA	AGTTCT	GTGAAC	CAGTTG	CCGTGAG	CGGATA	CGCTAC	TTGCATT	TACAG	AACCGA	ACAGGG	CTTATG	TCAC	TGGGTT
Ch. meningi Iso. PO6		CACCGC	TCGCGT	TCGGTCA	AGTTCT	GTGAAC	CAGTTG	CCGTGAG	CGGATA	CGCTAC	TTGCATT	TACAG	AACCGA	ACAGGG	CTTATG	TCAC	TGGGTT
Ps. alcaligenes Iso. P07		CACCGC	TCGCGT	TCGGTCA	AGTTCT	GTGAAC	CAGTTG	CCGTGAG	CGGATA	CGCTAC	TTGCATT	TACAG	AACCGA	ACAGGG	CTTATG	TCAC	TGGGTT
Enterobacter sp. Iso. 11045		CACCGC	TCGCGT	TCGGTCA	AGTTCT	GTGAAC	CAGTTG	CCGTGAG	CGGATA	CGCTAC	TTGCATT	TACAG	AACCGA	ACAGGG	CTTATG	TCAC	TGGGTT
En. cloacae Iso. 11047		CACCGC	TCGCGT	TCGGTCA	AGTTCT	GTGAAC	CAGTTG	CCGTGAG	CGGATA	CGCTAC	TTGCATT	TACAG	AACCGA	ACAGGG	CTTATG	TCAC	TGGGTT
En. cloacae pUL3AT Int3 and attI3	TC	CGTCC	CCATCC	ATCC	ATCC	ATCC	ATCC	ATCC	ATCC	ATCC	ATCC	ATCC	ATCC	ATCC	ATCC	ATCC	ATCC
K. pneumoniae Int3 and attI3	TC	CGTCC	CCATCC	ATCC	ATCC	ATCC	ATCC	ATCC	ATCC	ATCC	ATCC	ATCC	ATCC	ATCC	ATCC	ATCC	ATCC
		300	310	320	330	340	350	360			370	380		390			
		GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG		
E. coli pR388 Int1 and attL		GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG		
Ps. aeruginosa Iso. JIP041		GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG		
Ps. aeruginosa Iso. JIP044		GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG		
Ps. aeruginosa Iso. JIP045		GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG		
Ps. aeruginosa Iso. JIP052		GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG		
Ps. aeruginosa Iso. JIP073		GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG		
Ps. aeruginosa Iso. JIP080		GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG		
Ps. aeruginosa Iso. SSI 2.37		GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG		
Ps. aeruginosa Iso. SSI 3.34		GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG		
Ps. aeruginosa Iso. SSI 5.34		GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG		
St. maltophilia Iso. AP2		GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG		
Ps. putida ND6 Iso. AP3		GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG		
Ps. aeruginosa Iso. AS1		GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG		
Ps. aeruginosa Iso. AS2		GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG		
Ps. sp. N3 Iso. JM10		GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG		
Br. vesicularis Iso. LS13		GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG		
Ch. meningi Iso. PO6		GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG		
Ps. alcaligenes Iso. P07		GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG		
Enterobacter sp. Iso. 11045		GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG		
En. cloacae Iso. 11047		GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG	GTGGCT	CAACCG		
En. cloacae pUL3AT Int3 and attI3	AG	CCATT	GTTGG	TGGAC	GGCCCG	CAGC	ATGGG	GGGTAT	TCCAG	ATCCG	CGGAAA	TGGG	CAAG	CTGA	AGTC	GAGG	GT
K. pneumoniae Int3 and attI3	AG	CCATT	GTTGG	TGGAC	GGCCCG	CAGC	ATGGG	GGGTAT	TCCAG	ATCCG	CGGAAA	TGGG	CAAG	CTGA	AGTC	GAGG	GT

	400	410	420	430	440	450	460	470	480	490
	ACGAGCGCAAAGGTTTCGGTCTCCACGCATCGTCAG									
E. coli pR388 IntI1 and attI	ACGAGCGCAAAGGTTTCGGTCTCCACGCATCGTCAG									
Ps. aeruginosa Iso. JIP041	ACGAGCGCAAAGGTTTCGGTCTCCACGCATCGTCAG									
Ps. aeruginosa Iso. JIP044	ACGAGCGCAAAGGTTTCGGTCTCCACGCATCGTCAG									
Ps. aeruginosa Iso. JIP045	ACGAGCGCAAAGGTTTCGGTCTCCACGCATCGTCAG									
Ps. aeruginosa Iso. JIP052	ACGAGCGCAAAGGTTTCGGTCTCCACGCATCGTCAG									
Ps. aeruginosa Iso. JIP073	ACGAGCGCAAAGGTTTCGGTCTCCACGCATCGTCAG									
Ps. aeruginosa Iso. JIP080	ACGAGCGCAAAGGTTTCGGTCTCCACGCATCGTCAG									
Ps. aeruginosa Iso. SSI 2.37	ACGAGCGCAAAGGTTTCGGTCTCCACGCATCGTCAG									
Ps. aeruginosa Iso. SSI 3.34	ACGAGCGCAAAGGTTTCGGTCTCCACGCATCGTCAG									
Ps. aeruginosa Iso. SSI 5.34	ACGAGCGCAAAGGTTTCGGTCTCCACGCATCGTCAG									
St. maltophilia Iso. AP2	ACGAGCGCAAAGGTTTCGGTCTCCACGCATCGTCAGAGCCTGATCCGAAATTAAGTTGAGTGGTATCAGCGGGTTAGCTTGACGCCAGCCAGGTCATTG									
Ps. putida ND6 Iso. AP3	ACGAGCGCAAAGGTTTCGGTCTCCACGCATCGTCAG									
Ps. aeruginosa Iso. AS1	ACGAGCGCAAAGGTTTCGGTCTCCACGCATCGTCAG									
Ps. aeruginosa Iso. AS2	ACGAGCGCAAAGGTTTCGGTCTCCACGCATCGTCAG									
Ps. sp. N3 Iso. JM10	ACGAGCGCAAAGGTTTCGGTCTCCACGCATCGTCAG									
Br. vesicularis Iso. LS13	ACGAGCGCAAAGGTTTCGGTCTCCACGCATCGTCAG									
Ch. meningi Iso. PO6	ACGAGCGCAAAGGTTTCGGTCTCCACGCATCGTCAG									
Ps. alcaligenes Iso. P07	ACGAGCGCAAAGGTTTCGGTCTCCACGCATCGTCAG									
Enterobacter sp. Iso. 11045	ACGAGCGCAAAGGTTTCGGTCTCCACGCATCGTCAG									
En. cloacae Iso. 11047	ACGAGCGCAAAGGTTTCGGTCTCCACGCATCGTCAG									
En. cloacae pUL3AT IntI3 and attI3	CCGAGAAGCAAGTGGCCCGGCGCACCSACCGCSAG									
K. pneumoniae IntI3 and attI3	CCGAGAAGCAAGTGGCCCGGCGCACCSACCGCSAG									
	500	510	520	530	540	550	560	570	580	590

E. coli pR388 IntI1 and attI	-----									
Ps. aeruginosa Iso. JIP041	-----									
Ps. aeruginosa Iso. JIP044	-----									
Ps. aeruginosa Iso. JIP045	-----									
Ps. aeruginosa Iso. JIP052	-----									
Ps. aeruginosa Iso. JIP073	-----									
Ps. aeruginosa Iso. JIP080	-----									
Ps. aeruginosa Iso. SSI 2.37	-----									
Ps. aeruginosa Iso. SSI 3.34	-----									
Ps. aeruginosa Iso. SSI 5.34	-----									
St. maltophilia Iso. AP2	ATTTCAGGGGTTTTTTCGAAAACTTCGGGAGATCAACGATGTGGACCGATACCACTCGGGCGCAGTATGCCCGTGGGAAC TGGCTTTGCCAAGCGATTT									
Ps. putida ND6 Iso. AP3	-----									
Ps. aeruginosa Iso. AS1	-----									
Ps. aeruginosa Iso. AS2	-----									
Ps. sp. N3 Iso. JM10	-----									
Br. vesicularis Iso. LS13	-----									
Ch. meningi Iso. PO6	-----									
Ps. alcaligenes Iso. P07	-----									
Enterobacter sp. Iso. 11045	-----									
En. cloacae Iso. 11047	-----									
En. cloacae pUL3AT IntI3 and attI3	-----									
K. pneumoniae IntI3 and attI3	-----									

	600	610	620	630	640	650	660	670	680	690
E. coli pR388 IntI1 and attI	-----									
Ps. aeruginosa Iso. JIP041	-----									
Ps. aeruginosa Iso. JIP044	-----									
Ps. aeruginosa Iso. JIP045	-----									
Ps. aeruginosa Iso. JIP052	-----									
Ps. aeruginosa Iso. JIP073	-----									
Ps. aeruginosa Iso. JIP080	-----									
Ps. aeruginosa Iso. SSI 2.37	-----									
Ps. aeruginosa Iso. SSI 3.34	-----									
Ps. aeruginosa Iso. SSI 5.34	-----									
St. maltophilia Iso. AP 2	GACCGATGCGCGAATGGGCCACTGCTGGAGCCGTTCTTTCCGCCAGCATGCGCATGTGGGCCGCCC									
Ps. putida ND6 Iso. AP3	CGCGCAAGTGGCCACTCAGGCGGATTGTCGAGGGCGAT									
Ps. aeruginosa Iso. AS1	-----									
Ps. aeruginosa Iso. AS2	-----									
Ps. sp. N3 Iso. JM10	-----									
Br. vesicularis Iso. LS13	-----									
Ch. meningi Iso. P06	-----									
Ps. alcaligenes Iso. P07	-----									
Enterobacter sp. Iso. 11045	-----									
En. cloacae Iso. 11047	-----									
En. cloacae pUL3AT IntI3 and attI3	-----									
K. pneumoniae IntI3 and attI3	-----									
	700	710	720	730	740	750	760	770	780	790
E. coli pR388 IntI1 and attI	-----									
Ps. aeruginosa Iso. JIP041	-----									
Ps. aeruginosa Iso. JIP044	-----									
Ps. aeruginosa Iso. JIP045	-----									
Ps. aeruginosa Iso. JIP052	-----									
Ps. aeruginosa Iso. JIP073	-----									
Ps. aeruginosa Iso. JIP080	-----									
Ps. aeruginosa Iso. SSI 2.37	-----									
Ps. aeruginosa Iso. SSI 3.34	-----									
Ps. aeruginosa Iso. SSI 5.34	-----									
St. maltophilia Iso. AP 2	CCTGTAATCTGCTGCGTGGCGGCCTGCGCGGATGCTGCGCGCTGCTTTCCGCCGGTCTCGACGGTGC									
Ps. putida ND6 Iso. AP3	GCGGCCTGGTTCTACCTGTGGCGGGACAA									
Ps. aeruginosa Iso. AS1	-----									
Ps. aeruginosa Iso. AS2	-----									
Ps. sp. N3 Iso. JM10	-----									
Br. vesicularis Iso. LS13	-----									
Ch. meningi Iso. P06	-----									
Ps. alcaligenes Iso. P07	-----									
Enterobacter sp. Iso. 11045	-----									
En. cloacae Iso. 11047	-----									
En. cloacae pUL3AT IntI3 and attI3	-----									
K. pneumoniae IntI3 and attI3	-----									

	800	810	820	830	840	850	860	870	880	890
<i>E. coli</i> pR388 IntI1 and attI										
<i>Ps. aeruginosa</i> Iso. JIP041										
<i>Ps. aeruginosa</i> Iso. JIP044										
<i>Ps. aeruginosa</i> Iso. JIP045										
<i>Ps. aeruginosa</i> Iso. JIP052										
<i>Ps. aeruginosa</i> Iso. JIP073										
<i>Ps. aeruginosa</i> Iso. JIP080										
<i>Ps. aeruginosa</i> Iso. SSI 2.37										
<i>Ps. aeruginosa</i> Iso. SSI 3.34										
<i>Ps. aeruginosa</i> Iso. SSI 5.34										
<i>St. maltophilia</i> Iso. AP2	T	A	G	G	C	T	G	T	G	G
<i>Ps. putida</i> ND6 Iso. AP3	T	G	C	A	T	T	G	A	A	C
<i>Ps. aeruginosa</i> Iso. AS1	T	G	C	A	T	G	C	T	G	T
<i>Ps. aeruginosa</i> Iso. AS2	T	G	C	A	T	G	C	T	G	T
<i>Ps. sp. N3</i> Iso. JM10	T	G	C	A	T	G	C	T	G	T
<i>Br. vesicularis</i> Iso. LS13	T	G	C	A	T	G	C	T	G	T
<i>Ch. meningi</i> Iso. PO6	T	G	C	A	T	G	C	T	G	T
<i>Ps. alcaligenes</i> Iso. P07	T	G	C	A	T	G	C	T	G	T
<i>Enterobacter</i> sp. Iso. 11045	T	G	C	A	T	G	C	T	G	T
<i>En. cloacae</i> Iso. 11047	T	G	C	A	T	G	C	T	G	T
<i>En. cloacae</i> pUL3AT IntI3 and attI3	T	G	C	A	T	G	C	T	G	T
<i>K. pneumoniae</i> IntI3 and attI3	T	G	C	A	T	G	C	T	G	T
	900	910	920	930	940	950	960	970	980	990
<i>E. coli</i> pR388 IntI1 and attI										
<i>Ps. aeruginosa</i> Iso. JIP041										
<i>Ps. aeruginosa</i> Iso. JIP044										
<i>Ps. aeruginosa</i> Iso. JIP045										
<i>Ps. aeruginosa</i> Iso. JIP052										
<i>Ps. aeruginosa</i> Iso. JIP073										
<i>Ps. aeruginosa</i> Iso. JIP080										
<i>Ps. aeruginosa</i> Iso. SSI 2.37										
<i>Ps. aeruginosa</i> Iso. SSI 3.34										
<i>Ps. aeruginosa</i> Iso. SSI 5.34										
<i>St. maltophilia</i> Iso. AP2	G	T	C	A	A	A	C	C	A	C
<i>Ps. putida</i> ND6 Iso. AP3	G	T	C	A	A	A	C	C	A	C
<i>Ps. aeruginosa</i> Iso. AS1	G	T	C	A	A	A	C	C	A	C
<i>Ps. aeruginosa</i> Iso. AS2	G	T	C	A	A	A	C	C	A	C
<i>Ps. sp. N3</i> Iso. JM10	G	T	C	A	A	A	C	C	A	C
<i>Br. vesicularis</i> Iso. LS13	G	T	C	A	A	A	C	C	A	C
<i>Ch. meningi</i> Iso. PO6	G	T	C	A	A	A	C	C	A	C
<i>Ps. alcaligenes</i> Iso. P07	G	T	C	A	A	A	C	C	A	C
<i>Enterobacter</i> sp. Iso. 11045	G	T	C	A	A	A	C	C	A	C
<i>En. cloacae</i> Iso. 11047	G	T	C	A	A	A	C	C	A	C
<i>En. cloacae</i> pUL3AT IntI3 and attI3	G	T	C	A	A	A	C	C	A	C
<i>K. pneumoniae</i> IntI3 and attI3	G	T	C	A	A	A	C	C	A	C

	1,000	1,010	1,020	1,030	1,040	1,050	1,060	1,070	1,080	
E. coli pR388 IntI1 and attI										
Ps. aeruginosa Iso. JIP041										
Ps. aeruginosa Iso. JIP044										
Ps. aeruginosa Iso. JIP045										
Ps. aeruginosa Iso. JIP052										
Ps. aeruginosa Iso. JIP073										
Ps. aeruginosa Iso. JIP080										
Ps. aeruginosa Iso. SSI 2.37										
Ps. aeruginosa Iso. SSI 3.34										
Ps. aeruginosa Iso. SSI 5.34										
St. maltophilia Iso. AP2	GTCCATGCGGGTAGTCCACACCGCCGACATCCAGGACCGTGACGGTGC	CGCCGCTGGTGC	TGGCCGAGATCATCAAGCGCTTCCCGTGGCTGCGCCATGTC							
Ps. putida ND6 Iso. AP3										
Ps. aeruginosa Iso. AS1										
Ps. aeruginosa Iso. AS2										
Ps. sp. N3 Iso. JM10										
Br. vesicularis Iso. LS13										
Ch. meningo Iso. PO6										
Ps. alcaligenes Iso. P07										
Enterobacter sp. Iso. 11045										
En. cloacae Iso. 11047										
En. cloacae pUL3AT IntI3 and attI3										
K. pneumoniae IntI3 and attI3										
	1,090	1,100	1,110	1,120	1,130	1,140	1,150	1,160	1,170	1,180
E. coli pR388 IntI1 and attI										
Ps. aeruginosa Iso. JIP041										
Ps. aeruginosa Iso. JIP044										
Ps. aeruginosa Iso. JIP045										
Ps. aeruginosa Iso. JIP052										
Ps. aeruginosa Iso. JIP073										
Ps. aeruginosa Iso. JIP080										
Ps. aeruginosa Iso. SSI 2.37										
Ps. aeruginosa Iso. SSI 3.34										
Ps. aeruginosa Iso. SSI 5.34										
St. maltophilia Iso. AP2	TTCCCGATGGTGGATATGCTGGCGACAAGCTCAGGGACGCCCTCCGGCGGATCGGCAAA	TGGACCGT	CGAGATCGTCAAACGGTCAGATGCCGCAAGA							
Ps. putida ND6 Iso. AP3										
Ps. aeruginosa Iso. AS1										
Ps. aeruginosa Iso. AS2										
Ps. sp. N3 Iso. JM10										
Br. vesicularis Iso. LS13										
Ch. meningo Iso. PO6										
Ps. alcaligenes Iso. P07										
Enterobacter sp. Iso. 11045										
En. cloacae Iso. 11047										
En. cloacae pUL3AT IntI3 and attI3										
K. pneumoniae IntI3 and attI3										

K. pneumoniae Intl3 and attI3

1,190 1,200 1,210 1,220 1,230 1,240 1,250 1,260 1,270 1,280

- E. coli pR388 Intl1 and attI
- Ps. aeruginosa Iso. JIP041
- Ps. aeruginosa Iso. JIP044
- Ps. aeruginosa Iso. JIP045
- Ps. aeruginosa Iso. JIP052
- Ps. aeruginosa Iso. JIP073
- Ps. aeruginosa Iso. JIP080
- Ps. aeruginosa Iso. SSI 2.37
- Ps. aeruginosa Iso. SSI 3.34
- Ps. aeruginosa Iso. SSI 5.34
- St. maltophilia Iso. AP2
- Ps. putida ND6 Iso. AP3
- Ps. aeruginosa Iso. AS1
- Ps. aeruginosa Iso. AS2
- Ps. sp. N3 Iso. JM10
- Br. vesicularis Iso. LS13
- Ch. meningi Iso. PO6
- Ps. alcaligenes Iso. P07
- Enterobacter sp. Iso. 11045
- En. cloacae Iso. 11047
- En. cloacae pUL3AT Intl3 and attI3
- K. pneumoniae Intl3 and attI3

GGGTTTGTGTCCTCCCGCGCCGTTGGGTTGTCGAACGCACACTGGCTTGGCTCAATCGGAACCGCCGCTCGCCAAGGACTTCGAGCAGACCATCGCA

1,290 1,300 1,310 1,320 1,330 1,340 1,350 1,360 1,370 1,380

- E. coli pR388 Intl1 and attI
- Ps. aeruginosa Iso. JIP041
- Ps. aeruginosa Iso. JIP044
- Ps. aeruginosa Iso. JIP045
- Ps. aeruginosa Iso. JIP052
- Ps. aeruginosa Iso. JIP073
- Ps. aeruginosa Iso. JIP080
- Ps. aeruginosa Iso. SSI 2.37
- Ps. aeruginosa Iso. SSI 3.34
- Ps. aeruginosa Iso. SSI 5.34
- St. maltophilia Iso. AP2
- Ps. putida ND6 Iso. AP3
- Ps. aeruginosa Iso. AS1
- Ps. aeruginosa Iso. AS2
- Ps. sp. N3 Iso. JM10
- Br. vesicularis Iso. LS13
- Ch. meningi Iso. PO6
- Ps. alcaligenes Iso. P07
- Enterobacter sp. Iso. 11045
- En. cloacae Iso. 11047
- En. cloacae pUL3AT Intl3 and attI3
- K. pneumoniae Intl3 and attI3

TCGGCTACCGCATGGCTCTTCATCGCCTCGATCCAGCTCTTCGCACGCCGCATCGCAAGGCCATGAAATCACGCCGGATAATTATGAATCAGACTCTCA

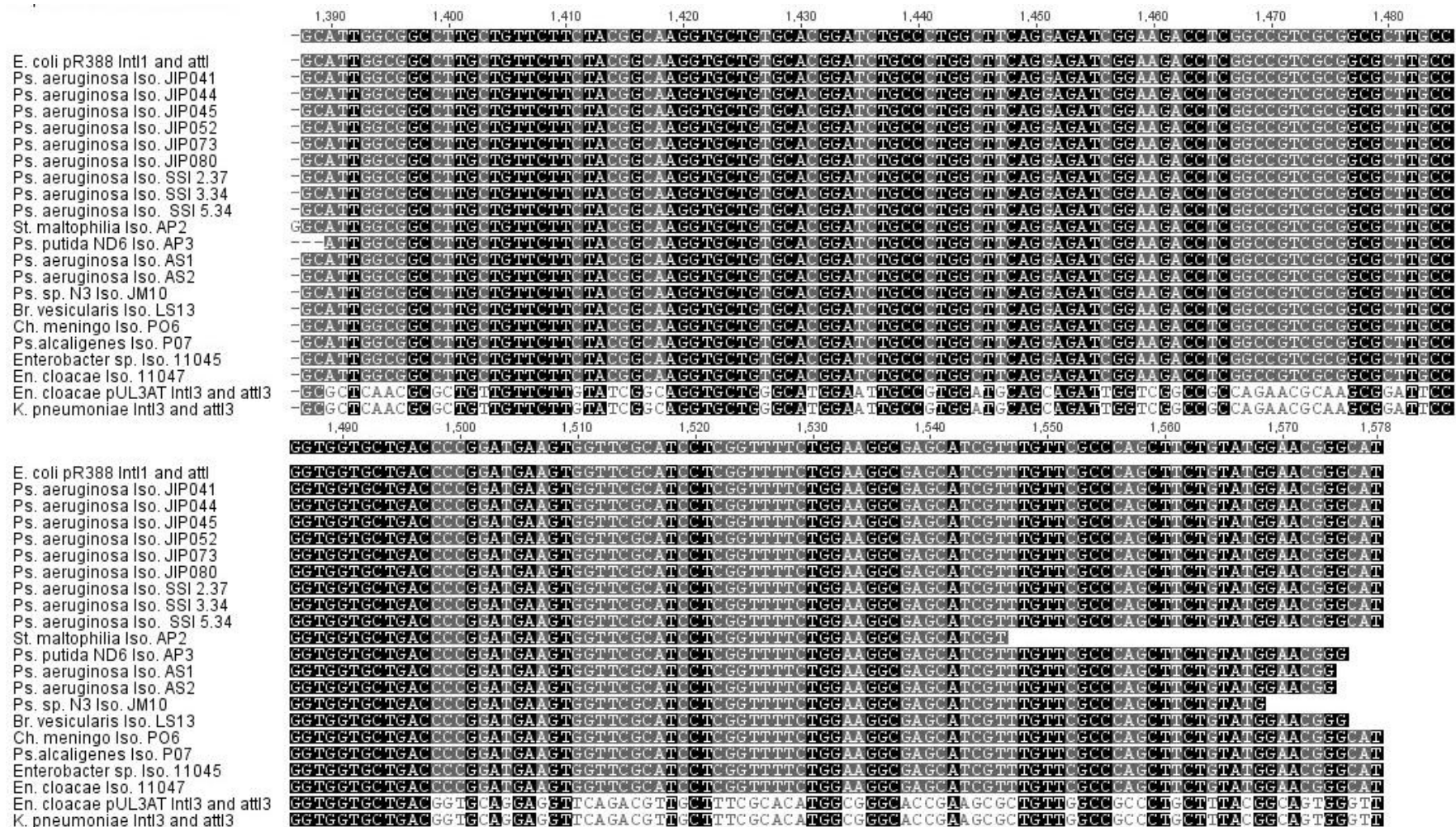


Figure A9: Sequence alignment of intII-attII recovered sequences from clinical and environmental isolates. Alignment was generated using Geneious software. All sequences were mapped against E. coli pR388 intII-attII sequence. Ps. aeruginosa isolates JIP and SSI are clinically relevant. Isolates AP, AS, P06 and P07 are clinical environmental isolates. Isolates JM and LS are domestic environmental isolates and 11045 and 11047 are two enteric isolates from outgroup. To examine the divergence of the recovered sequences, class 3 integron sequences were also added from En. cloacae and K. pneumoniae (intI3-attI3). Note IS4-like insertion in isolates St. maltophilia iso. AP2. SNPs at positions 113, 263, 268 and 288 are marked by red asterisks.

	SNP Positions: 263	268	10	16	20	288	30																									
Ps. aeruginosa Iso. JIP052	T	T	A	C	G	A	A	C	C	G	A	A	C	A	G	G	C	T	T	A	T	G	T	C	C	A	C	T	G	G	G	T
Pseudomonas aeruginosa TS491	T	T	A	C	G	A	A	C	C	G	A	A	C	A	G	G	C	T	T	A	T	G	T	C	A	A	C	T	G	G	G	T
In109 integrase ISPa21	C	T	T	A	C	C	A	A	C	C	G	A	A	C	A	G	G	C	T	T	A	T	G	T	C	C	A	C	T	G	G	T
R388 intl1 and attI	T	T	A	C	G	A	A	C	C	G	A	A	C	A	G	G	C	T	T	A	T	G	T	C	A	A	C	T	G	G	T	

Figure A10: SNP comparison between clinical isolates JIP052 and TS491, class 1 integron In109 and class 1 integron In3 type sequence.

Appendix A1.5 Fosmid library construction and screening

To determine the genetic context of *IS1111-attC* elements in two clinical (SSI2.84 and JIP117) and three environmental *Pseudomonas* (JM2, JM6 and LS5) isolates, large-insert clone/fosmid libraries were constructed. See below for library construction methodology.

A1.5.1 Obtaining high molecular weight DNA for fosmid by ISOLATE genomic DNA kit by Bioline.

The ISOLATE Genomic DNA Mini Kit from Bioline was used to extract high molecular DNA for the construction of fosmids. Adequate yields ranging from 100 ng/ μ L to 300 ng/ μ L of gDNA were obtained (as estimated by gel electrophoresis). More importantly, the extracted gDNA was found to be of the correct size range of 20-40-kb (Figure A11). As a result there was no need for additional DNA shearing or separation by PFGE. The obtained gDNA was subsequently used for fosmid construction.

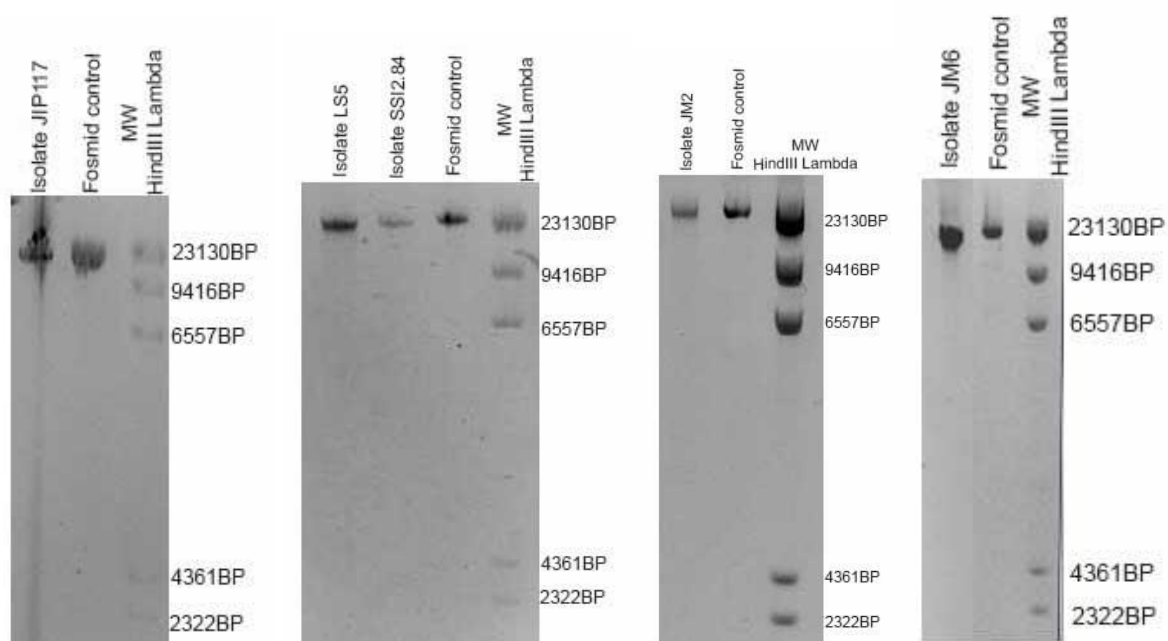


Figure A11: Gel electrophoresis analysis confirms presence of high molecular weight genomic DNA for clinical isolates *Pseudomonas aeruginosa* JIP117 and SSI2.84 as well as environmental isolates LS5, JM2 and JM6. Fosmid control DNA is 42-kb in size and supplied as 100 ng/ μ L (supplied by Epicenter).

A1.5.2 Whole plate screening using *ISPst6* as a probe

1920 fosmid clones across the five fosmid libraries were screened for *IS1111-attC* using a pooled column approach as represented in methods section 2.10 (Figure A12). Each positive spot on a southern hybridization corresponded to one column in the 96-well plate. Each column consisted of 8 different fosmid clones. Once a positive column is identified, the southern hybridization screen was repeated, wherein all of the individual fosmid clones were spotted to identify the clone carrying the IS. Southern hybridization using the *Pseudomonas ISPst6* probe indicated the presence of ISs in all 5 fosmid libraries.

Genomic DNA derived from *Pseudomonas stutzeri* ATCC17587 (*Pst587*), was used as the positive control for the *IS1111-attC* and hybridized intensely at all times. *Pseudomonas aeruginosa* PAO1 also produced a hybridization signal as it has six transposase genes as part of its genome, *Pseudomonas stutzeri* DNSP21 was also found to be positive. *E. coli* plasmid R388 DNA, *Pseudomonas stutzeri* st. Q (*PstQ*), and *Pseudomonas mendocina* KM91 are *IS1111-attC* element negative and did not produce a hybridization signal as expected. For the JIP117 fosmid screen, genomic DNA from the original *Pseudomonas* isolates was spotted as an additional control.

For the JIP117 library, southern hybridization was performed under very stringent conditions (60°C hybridization, 65°C wash temperature in 1x SSC buffer). Consequently, only two of the positive controls (*Pst587* and MB216.2B) gave a positive signal but the expected genomic DNA did not. Despite the stringent conditions, 36 of the 48 screened columns hybridized strongly to the *IS1111-attC* probe.

Additionally, the pooled columns in all five fosmid libraries were hybridized against a mobile class 1 integron probe as per conditions applied in Chapter 3 (Figure A13). Individual columns across all fosmid libraries were identified as being positive for a class 1 integron probe but were not pursued further, as the primary goal was to identify fosmid clones with *IS1111-attC* elements.

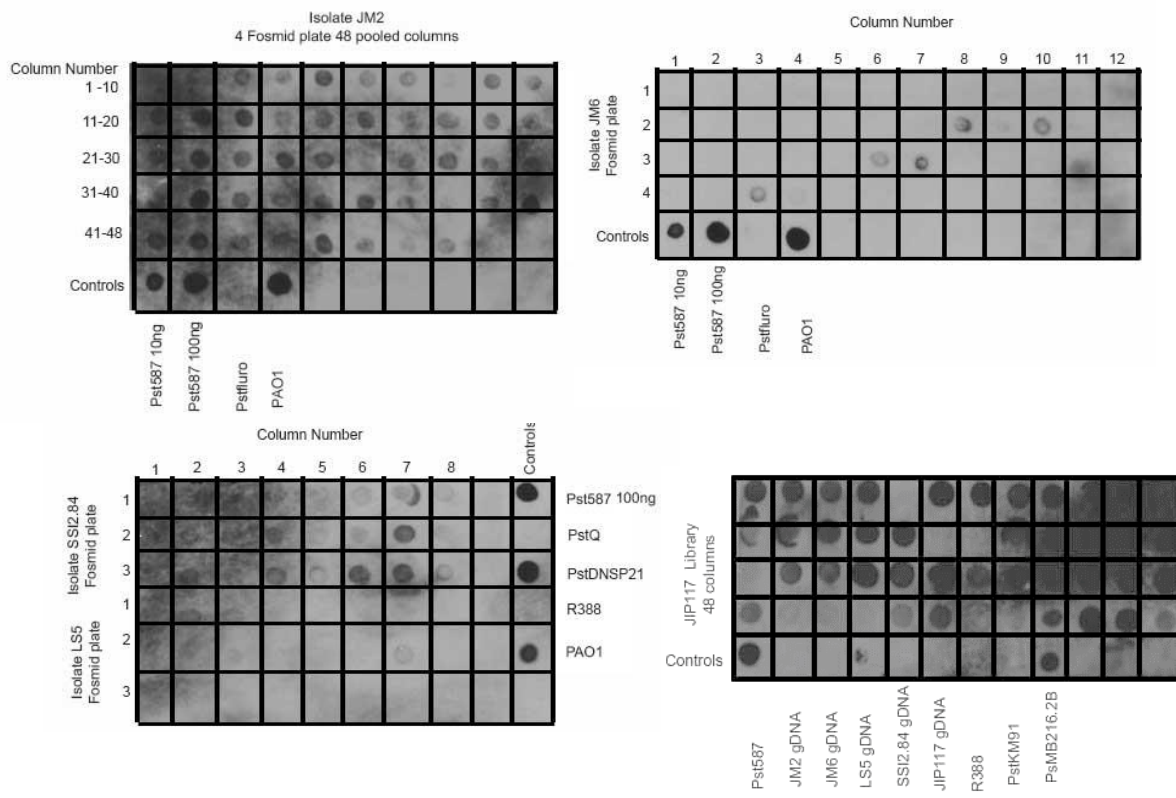


Figure A12: Screening pooled fosmid columns with an IS1111-attC probe across five fosmid libraries. The probe used was an intact ~1.3-kb ISPst6 generated from Pst587 gDNA. Top left: JM2 fosmid library, top right: JM6 fosmid library; bottom left: SSI2.84 and LS5 libraries and bottom right: JIP117 fosmid library screen.

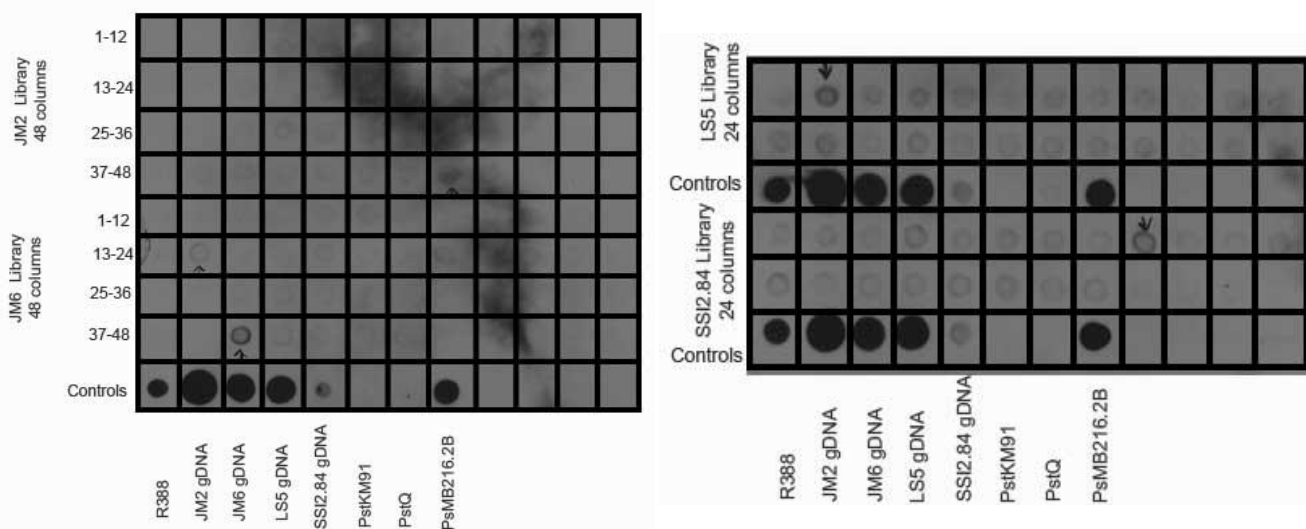


Figure A13: Pooled column screen for MRIs. Based on a 516-bp probe using LW14 and LW15 primers and *E. coli* pR388 intI1 as template DNA. All controls other than PstKM91 and PstQ should be hybridizing and producing a signal for the MI probe. JM2C44, JM6C14 and JM6C39 were of interest. LS5C2 and SSI2.84C9 were of interest. Individual clones were screened for presence of the MI.

A1.5.3 Estimation of genomic coverage based on fosmid insert size.

For each library, five or more clones were selected at random in order to estimate the average insert size and genomic coverage of each library. This was achieved by purifying the selected fosmid constructs and performing a *Bam*HI restriction digest (Figure A14). The size of each clone was estimated by adding the sizes of all bands observed after electrophoresis in a 1.2% agarose gel. The genomic coverage of each library was estimated by calculating the total amount of DNA (in megabases) present in each library and dividing this total by the estimated total genome size of the relevant strain.

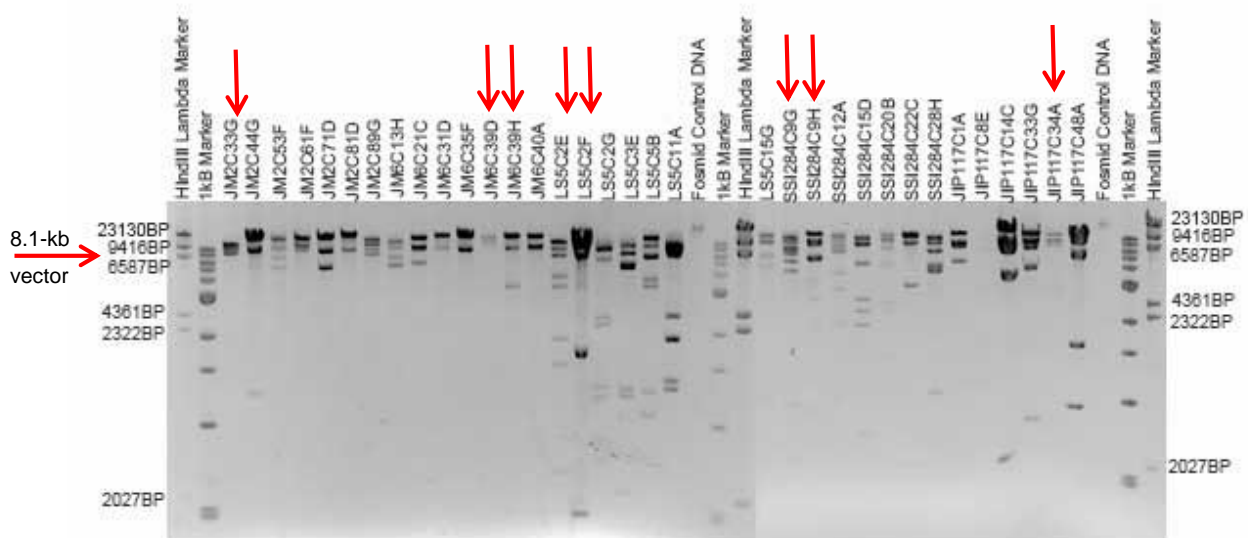


Figure A14: *Bam*HI digest of randomly selected as well as IS1111-attC identified fosmid clones for each strain for which fosmid libraries were constructed.

A band of ~8.1-kb is present in every fosmid digest and represents the entire fosmid vector pCC1FOS™. The remaining bands in each lane represent the cloned DNA. The sum of these bands gives the size for each clone. An average insert size of 30-kb was assumed for all fosmid clones. Arrows point to fosmid clones chosen for further analysis.

A1.5.4 IS1111-attC and mobile integron screen in fosmid clones of interest

Preliminary attempts to screen the 1920 fosmid clones via IS1111-attC and MI PCR resulted in non-specific PCR products being generated. Hence southern hybridization was used as a primary means of identifying IS positive fosmid clones. Once identified, IS1111-attC primers (ST87 and ST117) were used to screen for the presence of an IS from purified fosmid DNA. Interestingly this PCR was successful for the fosmid clones screened, with no instances of misprimed products such as those seen when genomic DNA was used as the PCR template. This observation is hypothesized to be due to the reduced amount of potential mispriming targets available in the small clones (~30-kb) compared to entire genomes (~4-Mb).

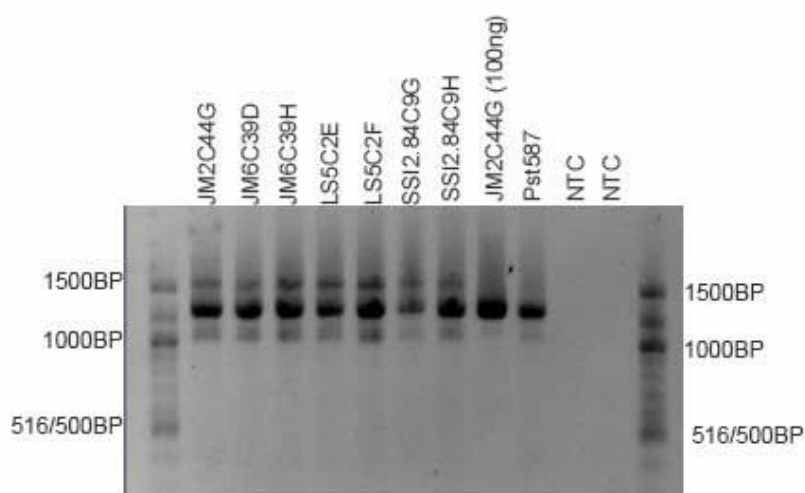
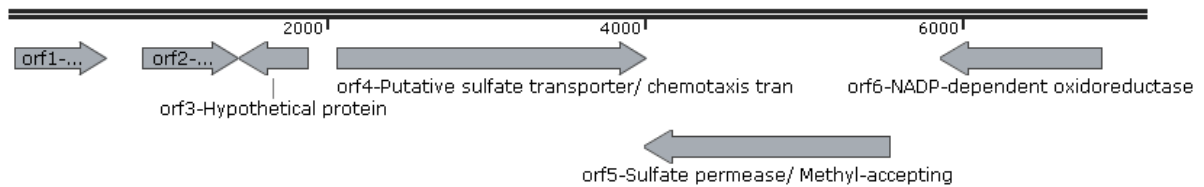


Figure A15: IS1111-attC PCR screen in purified fosmid clones.

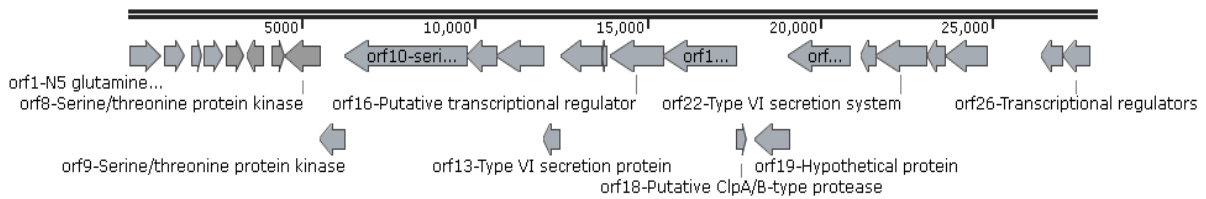
Expected product size is 1.3-kb. Pst587 was used as IS1111-attC positive control. NTC is the no template control where MQ was used as template. Expected product size was 1.3-kb. Note JM2C44G has 10x more gDNA than JM2C44G (100ng). When less gDNA is used as PCR template, the bands at 1.5 and 1.0-kb are not observed.

A1.5.5 Scale diagrams of individual fosmid contigs

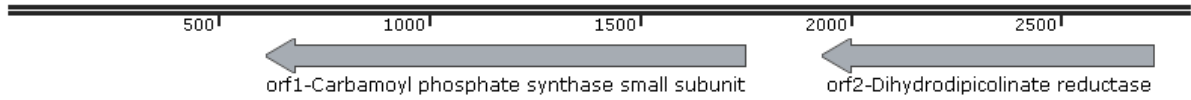
In each array diagram all features are shown to scale (arrays are not to scale relative to each other). Presence of ORFs was predicted using Glimmer plug-in in Geneious. ORFs were NCBI BLAST searched. The names of original isolate, size of contig as well as contig number are shown below the annotated array.



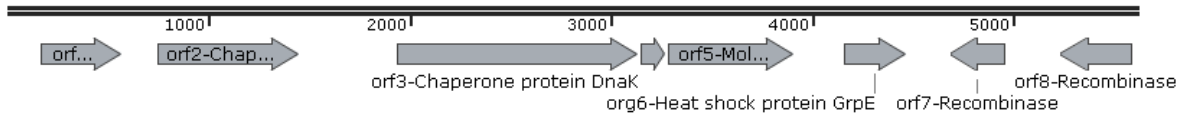
Ps. aeruginosa Iso. SSI2.84 Fosmid Contig 7
7150 bp



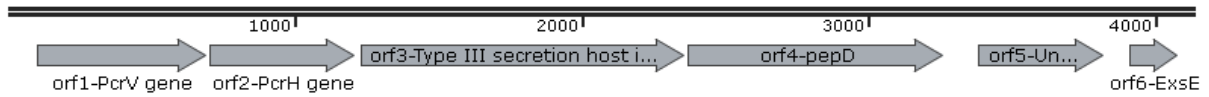
Ps. aeruginosa Iso. SSI2.84 Fosmid Contig B
27,984 bp



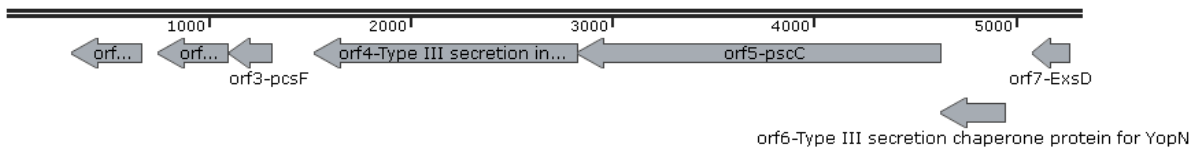
Ps. aeruginosa Iso. JIP117 Contig12
2802 bp



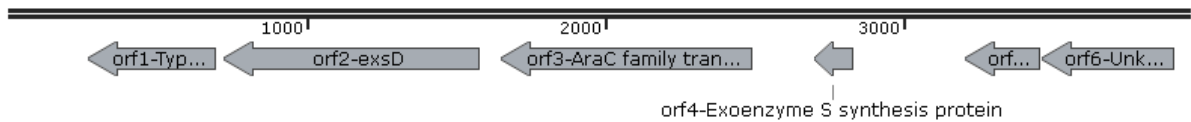
Ps. aeruginosa Iso. JIP117 Contig 13
5614 bp



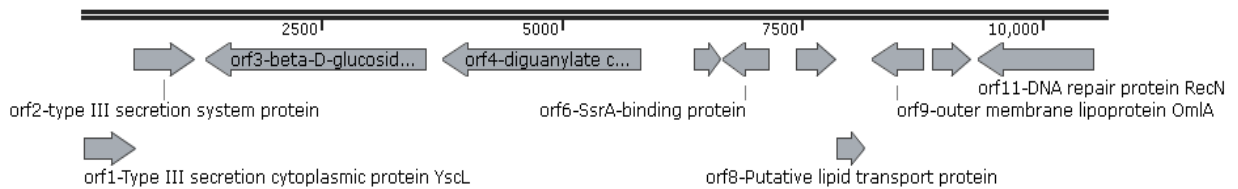
Ps. aeruginosa Iso. JIP117 Contig 14
4132 bp



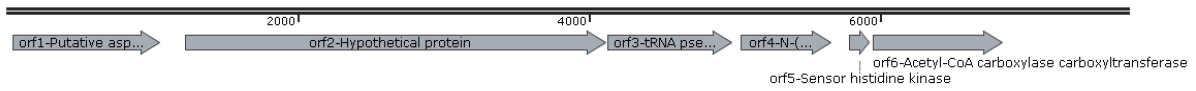
Ps. aeruginosa Iso. JIP117 Contig 15
5321 bp



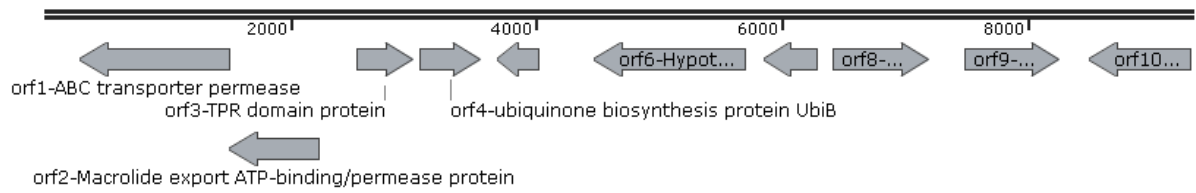
Ps. aeruginosa Iso. JIP117 Contig 16
3953 bp



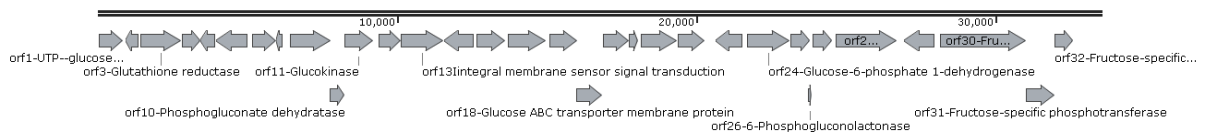
Ps. aeruginosa Iso. JIP117 Contig 17
10,667 bp



Ps. alcaligenes strain S3 Iso. JM2 Contig 18
7704 bp



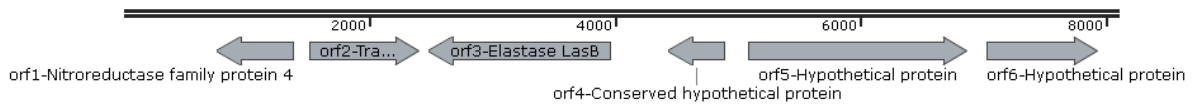
Ps. alcaligenes strain S3 Iso. JM2 Contig 19
9337 bp



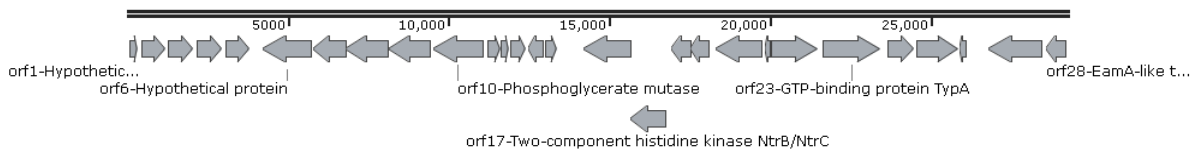
Ps. alcaligenes strain S3 Iso. JM2 Contig 20
33,528 bp



Ps. alcaligenes Iso. JM6 Contig2
2384 bp



Ps. alcaligenes Iso. JM6 Contig 3
8092 bp



Ps. alcaligenes Iso. JM6 Contig4
29,279 bp

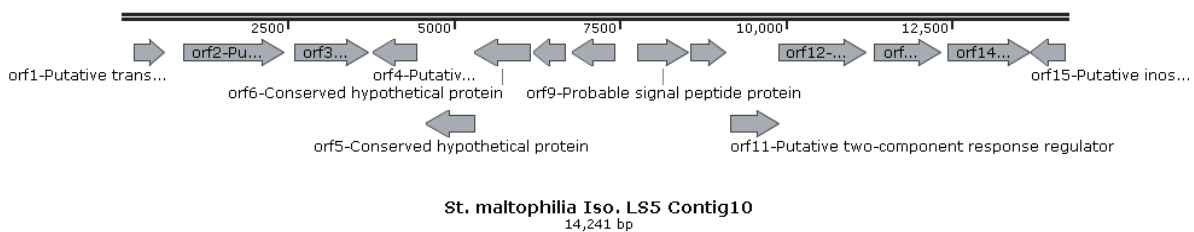
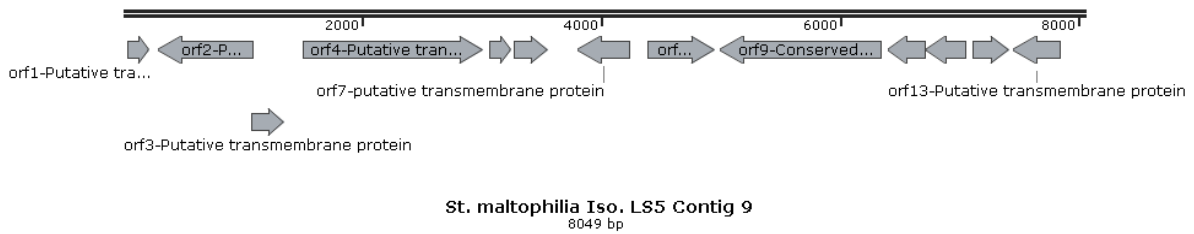
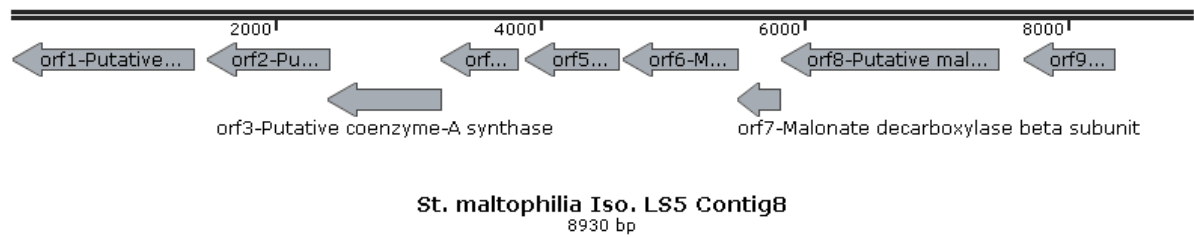
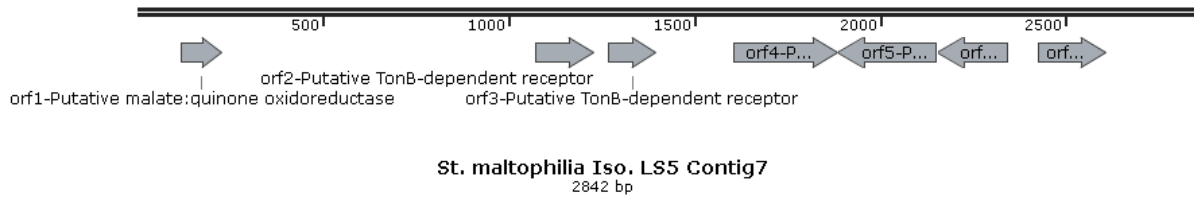


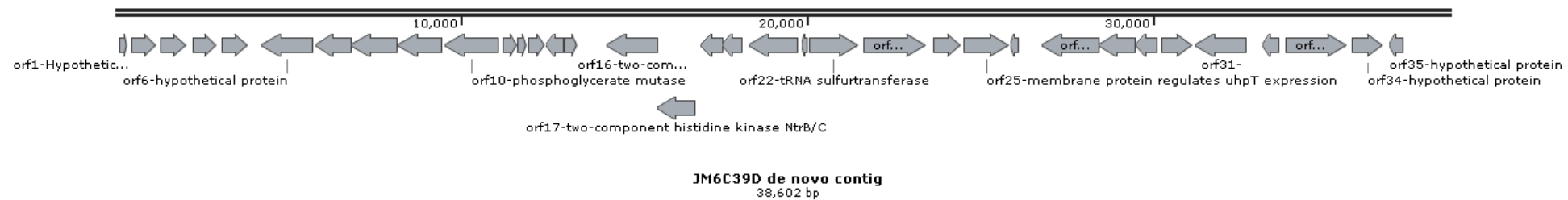
Figure A16: Individual contig reads as assembled by Macrogen and validated by raw read de novo assembly. All orfs were predicted using Glimmer plug-in in Geneious. Contig maps were generated using SnapGene.

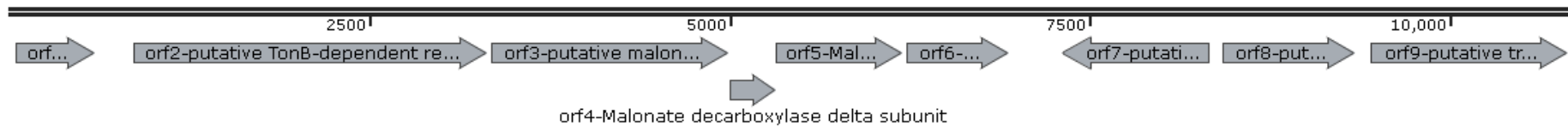
A1.5.6 Scale diagrams of de novo assembled contigs

Each array shown below is the product of *de novo* assembly of the individual contigs shown earlier in section A1.5.5. All contigs were assembled using *de novo* assembly function in Geneious using stringent assembly conditions outlined in Table A5. In each array diagram all features are shown to scale however arrays are not to scale relative to each other. Presence of ORFs were predicted using the Glimmer plug-in in Geneious. ORFs were NCBI BLAST searched. The names of original isolate, size of contig as well as contig number are shown below the annotated array. For fosmids JM2 and SSI2.84, a continuous contig could not be generated. For LS5 and JIP117, two continuous contigs were generated.

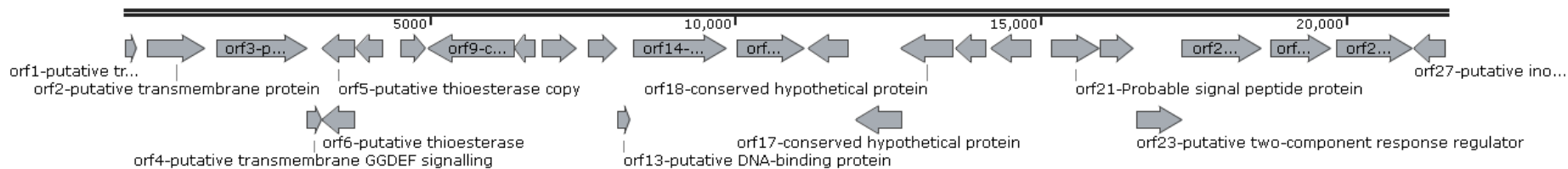
Table A5: Stringent conditions for *de novo* assembly using Geneious

Condition	Setting
Maximum gaps per read	20%
Maximum overlap	100bp
Word length	10
Ignore words repeated more than x times	1000
Maximum mismatch per read	50%
Maximum gap size	5
Minimum overlap identity	80%
Index word length	10
Reanalyse threshold	0
Maximum ambiguity	16
Use paired read distances to improve assembly	Activated
Do not merge contigs where there is a variant with coverage over x	~6
Merge homopolymer variants	Activated





LS5C2E de novo contig 1
10,824 bp



LS5C2E de novo contig 2
21,659 bp

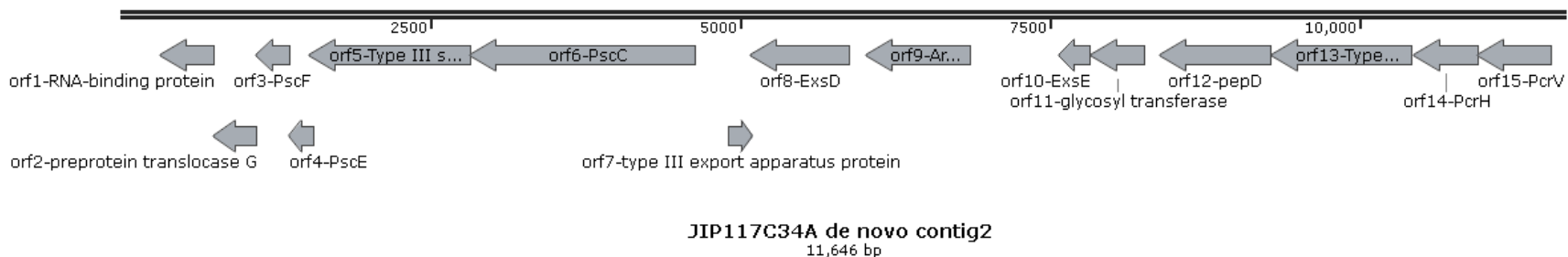
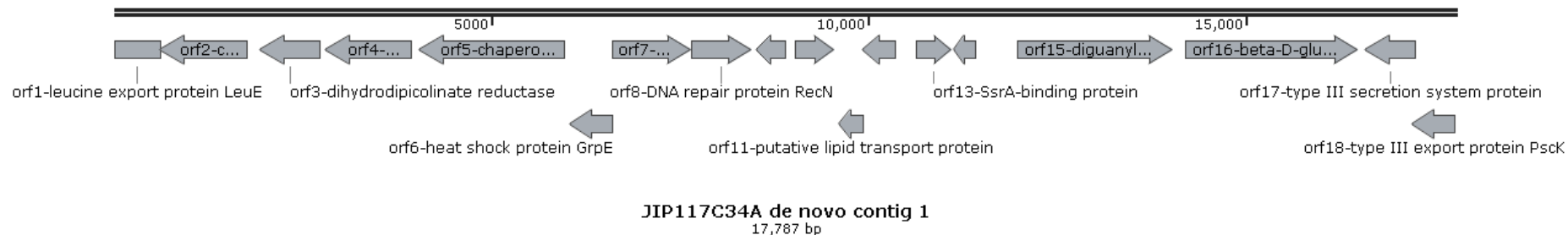


Figure A17: *de novo* assembled contigs based on individual contigs sequenced by Macrogen, Korea. Orfs were predicted using Glimmer plug-in in Geneious. Maps were constructed using SnapGene.

A1.5.7 Assigning *orfs* for all contig reads: see table A6

All *orfs* from the SSI284 and JIP117 libraries had high sequence identity to previously characterised *orfs* from various *Ps. aeruginosa* strains. The majority of the *orfs* from these two fosmid clones showed homology to transporter and secretion proteins, heat shock proteins, chaperone, kinases as well as proteases to name a few. SSI2.84C9H *orfs* contain nine hypothetical proteins whereas JIP117C34A had none. These hypothetical proteins are not homologous to an IS1111-*attC* transposase gene or protein. Note, JIP117C34A contig 1 *orf7* has a 99% (1042/1043bp) homology to a recombinase from *Ps. aeruginosa* YL84.

All thirty-six *orfs* from LS5C2E contigs showed sequence homology either to *St. maltophilia* K27a or *St. maltophilia* D457. None of the identified *orfs* matched an IS1111-*attC* transposase or any other recombinase enzyme. *orfs* showed high homology to transmembrane signaling regulatory proteins, malonate decarboxylase subunits, coenzymeA synthases as well as five conserved hypothetical proteins to name a few.

Fosmid clones JM2C44G and JM6C39D were generated using *Ps. alcaligenes* genomic DNA. Surprisingly, none of the 42 *orfs* from the JM2C44G or the 35 *orfs* from JM6C39D showed sequence homology to any *Ps. alcaligenes* sequence information in NCBI. Instead, these *orfs* displayed high homology to various *Ps. resinovorans* genes. This is perhaps to be expected given that *Ps. resinovorans* and *Ps. alcaligenes* cluster together based on their 16S rRNA DNA (Spilker *et al.*, 2004). *Orfs* were also identified as matching other *Pseudomonas* species. JM2C44G contig 19 *orf8* had a 74% (574/779) identity to a Mo regulatory and gene from *Azotobacter vinlandii* CA. Contig 20 *orf8* was 75% (50/67) identical to a hypothetical protein from *Thermobispora bispora* DSM43833 and *orf10* was 82% (406/496) identical to *Azotobacter vinelandii* phosphogluconate dehydratase. Finally, *orf15* had a 73% (180/246) sequence homology to a transcriptional regulator from the AraC family in *Burkholderia* sp.383.

Similarly, JM6C39D contained *orf24* that had a 77% (141/182) identity to a putative histidinol-phosphate phosphatase from *Sphingomonas wittichii* RW1. *Orf25* was 74% (895/1215) identical to a membrane protein in *Serratia marcescens*. *Orf28* was 74% (145/195) identical to a hypothetical protein in *Kitasatospora setae* KM-6054, *orfs* 29 and 30 had an 80% (498/625) and 78% (679/874) homology to an *Achromobacter xylosoxidans* A8 nitroreductase protein and a regulatory helix-turn-helix protein respectively. As with the previous three fosmid clones, none of the *orfs* in either JM2C44G or JM6C39H fosmid clones showed sequence homology to any IS elements.

Table A6: BLASTn Hits for predicted orfs across the five libraries

ID	BLASTN hit	Organism	Accession number	% identity	E-value
<i>Pseudomonas aeruginosa</i> isolate SSI2.84 fosmid SSI2.84C9H contig 7 (size 7150-bp)					
<i>orf1</i>	tRNA 2-selenouridine synthase/ glutathione-dependent formaldehyde-activating protein	<i>Ps. aeruginosa</i> PA3818	HG530068.1	582/585(99%)	0
<i>orf2</i>	Hypothetical protein	<i>Ps. aeruginosa</i> M18	CP002496.1	609/609(100%)	0
<i>orf3</i>	Hypothetical Protein	<i>Ps. aeruginosa</i> LES431	CP006937.1	426/426(100%)	0
<i>orf4</i>	putative sulfate transporter/ chemotaxis transducer	<i>Ps. aeruginosa</i> DK2	CP003149.1	1952/1959(99%)	0
<i>orf5</i>	Sulfate permease/ Methyl-accepting chemotaxis protein I (serine chemoreceptor)	<i>Ps. aeruginosa</i> SCV20265	CP006931.1	1536/1542(99%)	0
<i>orf6</i>	NADP-dependent oxidoreductase	<i>Ps. aeruginosa</i> YL84	CP007147.1	1001/1005(99%)	0
<i>Pseudomonas aeruginosa</i> isolate SSI2.84 fosmid SSI2.84C9H contig 8 (size 27984-bp)					
<i>orf1</i>	N5-glutamine-S-adenosyl-L-methionine-dependent methyltransferase	<i>Ps. aeruginosa</i> RP73	CP006245.1	936/936(100%)	0
<i>orf2</i>	Isochorismatase	<i>Ps. aeruginosa</i> SCV20265	CP006931.1	596/597(99%)	0
<i>orf3</i>	hypothetical protein	<i>Ps. aeruginosa</i> YL84	CP007147.1	321/321(100%)	3E-162
<i>orf4</i>	DNA mismatch repair protein MutS	<i>Ps. aeruginosa</i> YL84	CP007147.1	558/558(100%)	0
<i>orf5</i>	GTP cyclohydrolase I	<i>Ps. aeruginosa</i> PA38182	HG530068.1	546/546(100%)	0
<i>orf6</i>	hypothetical protein	<i>Ps. eruginosa</i> LES431	CP006937.1	462/462(100%)	0
<i>orf7</i>	hypothetical protein	<i>Ps. aeruginosa</i> RP73	CP006245.1	381/381(100%)	0
<i>orf8</i>	serine/threonine protein kinase	<i>Ps. aeruginosa</i> YL84	CP007147.1	989/990(99%)	0
<i>orf9</i>	IcmF-related protein Serine/threonine protein phosphatase	<i>Ps. aeruginosa</i> PA1R	CP004055.1	724/729(99%)	0
<i>orf10</i>	serine/threonine phosphoprotein phosphatase Stp1	<i>Ps. aeruginosa</i> M18	CP002496.1	3506/3528(99%)	0
<i>orf11</i>	putative outer membrane protein	<i>Ps. aeruginosa</i> PA38182	HG530068.1	869/870(99%)	0
<i>orf12</i>	protein ImpJ	<i>Ps. aeruginosa</i> PA38182	HG530068.1	1330/1332(99%)	0
<i>orf13</i>	type VI secretion protein	<i>Ps. aeruginosa</i> LES431	CP006937.1	468/468(100%)	0
<i>orf14</i>	FHA domain protein	<i>Ps. aeruginosa</i> PAO1-VE13	CP006832.1	1191/1194(99%)	0
<i>orf15</i>	hypothetical protein	<i>Ps. aeruginosa</i> PA38182	HG530068.1	117/117(100%)	6E-52
<i>orf16</i>	putative transcriptional regulator	<i>Ps. aeruginosa</i> PA38182	HG530068.1	1508/1512(99%)	0
<i>orf17</i>	putative ClpA/B-type protease	<i>Ps. aeruginosa</i> M18	CP002496.1	2130/2133(99%)	0
<i>orf18</i>	putative ClpA/B-type protease	<i>Ps. aeruginosa</i> PA38182	HG530068.1	314/315(99%)	3E-156

<i>orf19</i>	hypothetical protein	<i>Ps. aeruginosa</i> PAO1-VE13	CP006832.1	1006/1008(99%)	0
<i>orf20</i>	hypothetical protein	<i>Ps. aeruginosa</i> DK2	CP003149.1	1785/1788(99%)	0
<i>orf21</i>	Uncharacterized protein	<i>Ps. aeruginosa</i> SCV20265	CP006931.1	408/408(100%)	0
<i>orf22</i>	type VI secretion system	<i>Ps. aeruginosa</i> LES431	CP006937.1	1475/1476(99%)	0
<i>orf23</i>	hypothetical protein	<i>Ps. aeruginosa</i> YL84	CP007147.1	507/507(100%)	0
<i>orf24</i>	type VI secretion system ImpA	<i>Ps. aeruginosa</i> LES431	CP006937.1	1178/1179(99%)	0
<i>orf25</i>	glutathione S-transferase	<i>Ps. aeruginosa</i> YL84	CP007147.1	603/603(100%)	0
<i>orf26</i>	Transcriptional regulators containing a DNA-binding HTH domain and an aminotransferase domain (MocR family)	<i>Ps. aeruginosa</i> PA1R	CP004055.1	764/765(99%)	0
<i>Pseudomonas aeruginosa</i> isolate JIP117 fosmid JIP117C34A contig 1 (size 17787-bp)					
<i>orf1</i>	Leucine export protein LeuE	<i>Ps. aeruginosa</i> LESlike4	CP006985.1	611/612(99%)	0
<i>orf2</i>	carbamoyl phosphate synthase small subunit	<i>Ps. aeruginosa</i> YL84	CP007147.1	1132/1137(99%)	0
<i>orf3</i>	dihydrodipicolinate reductase	<i>Ps. aeruginosa</i> PAO1-VE13	CP006832.1	786/789(99%)	0
<i>orf4</i>	molecular chaperone DnaJ	<i>Ps. aeruginosa</i> RP73	CP006245.1	1133/1134(99%)	0
<i>orf5</i>	chaperone protein DnaK	<i>Ps. aeruginosa</i> c7447m	CP006728.1	1877/1878(99%)	0
<i>orf6</i>	heat shock protein GrpE	<i>Ps. aeruginosa</i> MTB-1	CP006853.1	560/561(99%)	0
<i>orf7</i>	Recombinase	<i>Ps. aeruginosa</i> YL84	CP007147.1	1042/1043(99%)	0
<i>orf8</i>	DNA repair protein RecN	<i>Ps. aeruginosa</i> PA1R	CP004055.1	801/804(99%)	0
<i>orf9</i>	Fur family transcriptional regulator	<i>Ps. aeruginosa</i> YL84	CP007147.1	366/366(100%)	0
<i>orf10</i>	outer membrane lipoprotein OmlA	<i>Ps. aeruginosa</i> PA38182	HG530068.1	531/531(100%)	0
<i>orf11</i>	putative lipid transport protein	<i>Ps. aeruginosa</i> PA38182	HG530068.1	306/306(100%)	5E-158
<i>orf12</i>	ribosome association toxin RatA	<i>Ps. aeruginosa</i> LESlike4	CP006985.1	435/435(100%)	0
<i>orf13</i>	SsrA-binding protein	<i>Ps. aeruginosa</i> YL84	CP007147.1	480/480(100%)	0
<i>orf14</i>	Lactate-responsive regulator LldR in Enterobacteria	<i>Ps. aeruginosa</i> PA1R	CP004055.1	276/276(100%)	2E-141
<i>orf15</i>	diguanylate cyclase domain protein	<i>Ps. aeruginosa</i> PAO1-VE13	CP006832.1	1875/1877(99%)	0
<i>orf16</i>	beta-D-glucoside glucohydrolase	<i>Ps. aeruginosa</i> DK2	CP003149.1	1877/1877(100%)	0
<i>orf17</i>	type III secretion system protein	<i>Ps. aeruginosa</i> LESlike4	CP006985.1	645/645(100%)	0
<i>orf18</i>	type III export protein PscK	<i>Ps. aeruginosa</i> LESlike4	CP006985.1	548/549(99%)	0
<i>Pseudomonas aeruginosa</i> isolate JIP117 fosmid JIP117C34A contig 2 (size 11646-bp)					
<i>orf1</i>	RNA-binding protein	<i>Ps. aeruginosa</i> LESlike4	CP006985.1	432/432(100%)	0
<i>orf2</i>	preprotein translocase G	<i>Ps. aeruginosa</i> LESlike4	CP006985.1	348/348(100%)	0
<i>orf3</i>	exoenzyme S secretion (psc) locus pscF	<i>Ps. aeruginosa</i>	PAU56077	258/258(100%)	2E-131

<i>orf4</i>	exoenzyme S secretion (psc) locus pscE	<i>Ps. aeruginosa</i> PA38182	HG530068.1	201/201(100%)	8E-100
<i>orf5</i>	Type III secretion inner membrane protein (YscD, flagellar export components)	<i>Ps. aeruginosa</i> PA1R	CP004055.1	1299/1299(100%)	0
<i>orf6</i>	exoenzyme S secretion (psc) locus pscC	<i>Ps. aeruginosa</i>	PAU56077	1803/1803(100%)	0
<i>orf7</i>	type III export apparatus protein	<i>Ps. aeruginosa</i> LESlike4	CP006985.1	201/204(99%)	2E-96
<i>orf8</i>	exoenzyme S secretion (psc) locus exsD	<i>Ps. aeruginosa</i>	PAU56077	792/792(100%)	0
<i>orf9</i>	AraC family transcriptional regulator	<i>Ps. aeruginosa</i> MTB-1	CP006853.1	837/837(100%)	0
<i>orf10</i>	exoenzyme S secretion (psc) locus exsE	<i>Ps. aeruginosa</i> YL84	CP007147.1	246/246(100%)	1E-124
<i>orf11</i>	glycosyl transferase	<i>Ps. aeruginosa</i> LESlike4	CP006985.1	438/438(100%)	0
<i>orf12</i>	pepD gene	<i>Ps. aeruginosa</i> c7447m	CP006728.1	888/888(100%)	0
<i>orf13</i>	Type III secretion host injection protein (YopB)	<i>Ps. aeruginosa</i> PA1R	CP004055.1	1131/1131(100%)	0
<i>orf14</i>	pcrH genes	<i>Ps. aeruginosa</i>	AF010149.1	507/507(100%)	0
<i>orf15</i>	pcrV gene	<i>Ps. aeruginosa</i>	AF010149.1	591/591(100%)	0
<i>Pseudomonas alcaligenes</i> isolate JM2 fosmid JM2C44G contig 18 (size 7704-bp)					
<i>orf1</i>	putative aspartate-semialdehyde dehydrogenase	<i>Ps. resinovorans</i> NBRC 106553 DNA	AP013068.1	767/1024(75%)	0
<i>orf2</i>	Hypothetical protein	<i>Ps. resinovorans</i> NBRC 106553 DNA	AP013068.1	2181/2991(73%)	0
<i>orf3</i>	tRNA pseudouridine synthase A	<i>Ps. fulva</i> 12-X	CP002727.1	679/820(83%)	0
<i>orf4</i>	N-(5'-phosphoribosyl)anthranilate isomerase	<i>Ps. aeruginosa</i> RP73	CP006245.1	466/611(76%)	5E-119
<i>orf5</i>	sensor histidine kinase	<i>Ps. poae</i> RE*1-1-14	CP004045.1	33/41(80%)	9.8
<i>orf6</i>	acetyl-CoA carboxylase carboxyltransferase beta subunit folylpolyglutamate synthetase	<i>Ps. resinovorans</i> NBRC 106553 DNA	AP013068.1	777/887(88%)	0
<i>Pseudomonas alcaligenes</i> isolate JM2 fosmid JM2C44G contig 19 (size 9337-bp)					
<i>orf1</i>	ABC transporter permease	<i>Ps. aeruginosa</i> YL84	CP007147.1	840/1175(71%)	4E-150
<i>orf2</i>	macrolide export ATP-binding/permease protein MacB	<i>Ps. protegens</i> CHA0	CP003190.1	529/674(78%)	1E-153
<i>orf3</i>	TPR domain protein	<i>Ps. protegens</i> CHA0	CP003190.1	45/56(80%)	0.021
<i>orf4</i>	ubiquinone biosynthesis protein UbiB	<i>Ps. sp.</i> TKP	CP006852.1	171/246(70%)	4E-18
<i>orf5</i>	Hypothetical protein	<i>Ps. resinovorans</i> NBRC 106553	AP013068.1	263/316(83%)	1E-85
<i>orf6</i>	Hypothetical protein	<i>Ps. sp.</i> TKP	CP006852.1	216/326(66%)	3E-11
<i>orf7</i>	molybdenum-pterin-binding protein MopA	<i>Ps. protegens</i> CHA0	CP003190.1	332/431(77%)	2E-84
<i>orf8</i>	Mo regulation, Mo processing homeostasis	<i>Azotobacter vinelandii</i> CA	CP005094.1	574/779(74%)	9E-124

<i>orf9</i>	Hypothetical protein	<i>Ps. protegens</i> CHA0	CP003190.1	526/738(71%)	5E-95
<i>orf10</i>	Shikimate dehydrogenase substrate binding domain protein	<i>Ps. fulva</i> 12-X	CP002727.1	592/776(76%)	1E-154
<i>Pseudomonas alcaligenes</i> isolate JM2 fosmid JM2C44G contig 20 (size 33528-bp)					
<i>orf1</i>	UTP--glucose-1-phosphate uridylyltransferase	<i>Ps. resinovorans</i> NBRC 106553	AP013068.1	720/774(93%)	0
<i>orf2</i>	Hypothetical protein	<i>Ps. resinovorans</i> NBRC 106553	AP013068.1	354/419(84%)	9E-127
<i>orf3</i>	glutathione reductase	<i>Ps. resinovorans</i> NBRC 106553	AP013068.1	1170/1352(87%)	0
<i>orf4</i>	hypothetical protein	<i>Ps. resinovorans</i> NBRC 106553	AP013068.1	519/624(83%)	0
<i>orf5</i>	putative hydroxymethylpyrimidine transporter CytX	<i>Ps. putida</i> DOT-T1E	CP003734.1	48/60(80%)	5E-4
<i>orf6</i>	hypothetical protein; putative signal peptide	<i>Ps. entomophila</i> str. L48	CT573326.1	222/305(73%)	4E-34
<i>orf7</i>	putative hydrolase	<i>Ps. resinovorans</i> NBRC 106553	AP013068.1	658/800(82%)	0
<i>orf8</i>	hypothetical protein	<i>Thermobispora bispora</i> DSM 43833	CP001874.1	50/67(75%)	0.27
<i>orf9</i>	6-phosphogluconate dehydratase	<i>Ps. resinovorans</i> NBRC 106553	AP013068.1	1164/1360(86%)	0
<i>orf10</i>	phosphogluconate dehydratase	<i>Azotobacter vinelandii</i>	CP005095.1	406/496(82%)	8E-135
<i>orf11</i>	Glucokinase	<i>Ps. stutzeri</i> RCH2	CP003071.1	716/954(75%)	2E-172
<i>orf12</i>	Two-component response regulator GltR	<i>Ps. stutzeri</i> ATCC 17588 = LMG 11199	CP002881.1	610/725(84%)	0
<i>orf13</i>	integral membrane sensor signal transduction histidine kinase	<i>Ps. mendocina</i> ymp	CP000680.1	1188/1406(84%)	0
<i>orf14</i>	ABC transporter ATP-binding protein	<i>Ps. sp.</i> TKP	CP006852.1	33/37(89%)	0.044
<i>orf15</i>	transcriptional regulator, AraC family	<i>Burkholderia</i> sp. 383	CP000152.1	180/246(73%)	7E-25
<i>orf16</i>	glucose-binding protein	<i>Ps. mendocina</i> ymp	CP000680.1	1033/1254(82%)	0
<i>orf17</i>	glucose ABC transporter membrane protein	<i>Ps. mendocina</i> ymp	CP000680.1	813/907(90%)	0
<i>orf18</i>	glucose ABC transporter membrane protein	<i>Ps. mendocina</i> ymp	CP000680.1	738/819(90%)	0
<i>orf19</i>	sn-glycerol-3-phosphate import ATP-binding protein UgpC	<i>Ps. protegens</i> CHA0	CP003190.1	723/813(89%)	0
<i>orf20</i>	ABC transporter, ATP-binding component	<i>Ps. brassicacearum</i> subsp. <i>brassicacearum</i> NFM421	CP002585.1	242/322(75%)	7E-51
<i>orf21</i>	porin, LamB type Aldose 1-epimerase	<i>Ps. mendocina</i> ymp	CP000680.1	972/1149(85%)	0
<i>orf22</i>	Aldose 1-epimerase	<i>Ps. mendocina</i> ymp	CP000680.1	572/783(73%)	2E-113
<i>orf23</i>	RpiR family transcriptional regulator HexR	<i>Ps. resinovorans</i> NBRC 106553	AP013068.1	791/870(91%)	0
<i>orf24</i>	glucose-6-phosphate 1-dehydrogenas	<i>Ps. resinovorans</i> NBRC 106553	AP013068.1	1224/1436(85%)	0

<i>orf25</i>	glucose-6-phosphate 1-dehydrogenase 6-phosphogluconolactonase	<i>Ps. resinovorans</i> NBRC 106553	AP013068.1	516/651(79%)	7E-156
<i>orf26</i>	DNA 6-phosphogluconolactonase	<i>Ps. resinovorans</i> NBRC 106553	AP013068.1	94/116(81%)	1E-21
<i>orf27</i>	2-dehydro-3-deoxy-phosphogluconate/4-hydroxy- 2-oxoglutarate aldolase	<i>Ps. resinovorans</i> NBRC 106553	AP013068.1	538/654(82%)	9E-180
<i>orf28</i>	glyceraldehyde-3-phosphate dehydrogenase	<i>Ps. denitrificans</i> ATCC 13867	CP004143.1	1386/1600(87%)	0
<i>orf29</i>	fructose transport system repressor FruR	<i>Ps. resinovorans</i> NBRC 106553	AP013068.1	753/963(78%)	0
<i>orf30</i>	fructose-specific phosphotransferase system enzyme IIA/HPr/I component	<i>Ps. resinovorans</i> NBRC 106553	AP013068.1	1412/1886(75%)	0
<i>orf31</i>	fructose-specific phosphotransferase system enzyme IIA/HP component 1- phosphofructokinase	<i>Ps. resinovorans</i> NBRC 106553	AP013068.1	730/960(76%)	0
<i>orf32</i>	fructose-specific phosphotransferase system enzyme IIBC component	<i>Ps. resinovorans</i> NBRC 106553	AP013068.1	495/635(78%)	1E-139
<i>Pseudomonas alcaligenes</i> isolate JM6 fosmid JM6C39D contig 1 (size 38602-bp)					
<i>orf1</i>	Hypothetical protein	<i>Ps. resinovorans</i> NBRC 106553	AP013068.1	230/254(91%)	1E-88
<i>orf2</i>	1-(5-phosphoribosyl)-5-[(5-phosphoribosylamino) methylideneamino] imidazole-4-carboxamide isomerase	<i>Ps. resinovorans</i> NBRC 106553	AP013068.1	684/719(95%)	0
<i>orf3</i>	imidazole glycerol phosphate synthase subunit HisF	<i>Ps. resinovorans</i> NBRC 106553	AP013068.1	698/771(91%)	0
<i>orf4</i>	putative ABC transporter substrate-binding protein	<i>Ps. resinovorans</i> NBRC 106553	AP013068.1	592/721(82%)	8E-170
<i>orf5</i>	putative ABC transporter substrate-binding protein	<i>Ps. resinovorans</i> NBRC 106553	AP013068.1	670/755(89%)	0
<i>orf6</i>	hypothetical protein	<i>Ps. aeruginosa</i> RP73	CP006245.1	786/1122(70%)	7E-29
<i>orf7</i>	hypothetical protein	<i>Ps. resinovorans</i> NBRC 106553	AP013068.1	572/681(84%)	0
<i>orf8</i>	putative S41 family peptidase	<i>Ps. resinovorans</i> NBRC 106553	AP013068.1	1162/1335(87%)	0
<i>orf9</i>	peptidase M23 family protein	<i>Ps. resinovorans</i> NBRC 106553	AP013068.1	1038/1273(82%)	0
<i>orf10</i>	2,3-bisphosphoglycerate-independent- phosphoglycerate mutase	<i>Ps. resinovorans</i> NBRC 106553	AP013068.1	1403/1539(91%)	0
<i>orf11</i>	hypothetical protein	<i>Ps. resinovorans</i> NBRC 106553	AP013068.1	341/415(82%)	2E-94
<i>orf12</i>	glutaredoxin 3	<i>Ps. fulva</i> 12-X	CP002727.1	210/249(84%)	1E-59
<i>orf13</i>	protein-export protein SecB	<i>Ps. resinovorans</i> NBRC 106553	AP013068.1	398/439(91%)	2E-163

<i>orf14</i>	tRNA (cytidine(34)-2'-O)-methyltransferase	<i>Ps. resinovorans</i> NBRC 106553	AP013068.1	408/463(88%)	8E-153
<i>orf15</i>	conserved hypothetical protein	<i>Ps. putida</i> GB-1	CP000926.1	219/313(70%)	1E-28
<i>orf16</i>	two-component histidine kinase NtrB two-component response regulator NtrC	<i>Ps. resinovorans</i> NBRC 106553	AP013068.1	1316/1437(92%)	0
<i>orf17</i>	two-component histidine kinase NtrB two-component response regulator NtrC	<i>Ps. resinovorans</i> NBRC 106553	AP013068.1	982/1086(90%)	0
<i>orf18</i>	hypothetical protein	<i>Ps. resinovorans</i> NBRC 106553	AP013068.1	467/580(81%)	4E-146
<i>orf19</i>	hypothetical protein	<i>Ps. resinovorans</i> NBRC 106553	AP013068.1	400/545(73%)	3E-77
<i>orf20</i>	glutamine synthetase	<i>Ps. denitrificans</i> ATCC 13867	CP004143.1	1312/1410(93%)	0
<i>orf21</i>	Glutamine synthase	<i>Ps. denitrificans</i> ATCC 13867	CP004143.1	1312/1410(93%)	0
<i>orf22</i>	tRNA sulfurtransferase	<i>Ps. resinovorans</i> NBRC 106553	AP013068.1	1292/1461(88%)	0
<i>orf23</i>	GTP-binding protein TypA	<i>Ps. resinovorans</i> NBRC 106553	AP013068.1	1735/1818(95%)	0
<i>orf24</i>	histidinol-phosphate phosphatase, putative	<i>Sphingomonas wittichii</i> RW1	CP000699.1	141/182(77%)	8E-16
<i>orf25</i>	membrane protein regulates uhpT expression	<i>Serratia marcescens</i> subsp. <i>marcescens</i> Db11	HG326223.1	895/1215(74%)	4E-115
<i>orf26</i>	Uncharacterized protein family UPF0126	<i>Ps. fulva</i> 12-X	CP002727.1	143/197(73%)	0
<i>orf27</i>	methyl-accepting chemotaxis protein	<i>Ps. mendocina</i> NK-01	CP002620.1	1377/1643(84%)	0
<i>orf28</i>	hypothetical protein	<i>Kitasatospora setae</i> KM-6054	AP010968.1	145/195(74%)	1E-9
<i>orf29</i>	nitroreductase family protein 4	<i>Achromobacter xylosoxidans</i> A8	CP002287.1	498/625(80%)	3E-119
<i>orf30</i>	Bacterial regulatory helix-turn-helix protein, LysR family	<i>Achromobacter xylosoxidans</i> A8	CP002287.1	679/874(78%)	9E-141
<i>orf31</i>	Elastase LasB	<i>Ps. aeruginosa</i> PA7	CP000744.1	769/971(79%)	0
<i>orf32</i>	Cell division protein	<i>Ps. brassicacearum</i> strain DF41	CP007410.1	317/436(73%)	1E-56
<i>orf33</i>	hypothetical protein	<i>Ps. resinovorans</i> NBRC 106553	AP013068.1	1389/1695(82%)	0
<i>orf34</i>	hypothetical protein	<i>Ps. resinovorans</i> NBRC 106553	AP013068.1	722/876(82%)	0
<i>orf35</i>	hypothetical protein	<i>Ps. resinovorans</i> NBRC 106553	AP013068.1	336/389(86%)	1E-115
<i>Stenotrophomonas maltophilia</i> isolate LS5 fosmid LS5C2E contig 1(size 14241-bp)					
<i>orf1</i>	putative malate:quinone oxidoreductase	<i>St. maltophilia</i> K279a	AM743169.1	548/549(99%)	0
<i>orf2</i>	putative TonB-dependent receptor for Fe(III)-coprogen, Fe(III)-ferrioxamine B and Fe(III)-rhodotruclic acid	<i>St. maltophilia</i> K279a	AM743169.1	1869/1879(99%)	0
<i>orf3</i>	putative malonate decarboxylase alpha-subunit	<i>St. maltophilia</i> K279a	AM743169.1	1624/1644(99%)	0
<i>orf4</i>	Malonate decarboxylase delta subunit	<i>St. maltophilia</i> D457	HE798556.1	318/321(99%)	3E-161
<i>orf5</i>	Malonate decarboxylase beta subunit	<i>St. maltophilia</i> D457	HE798556.1	857/873(98%)	0

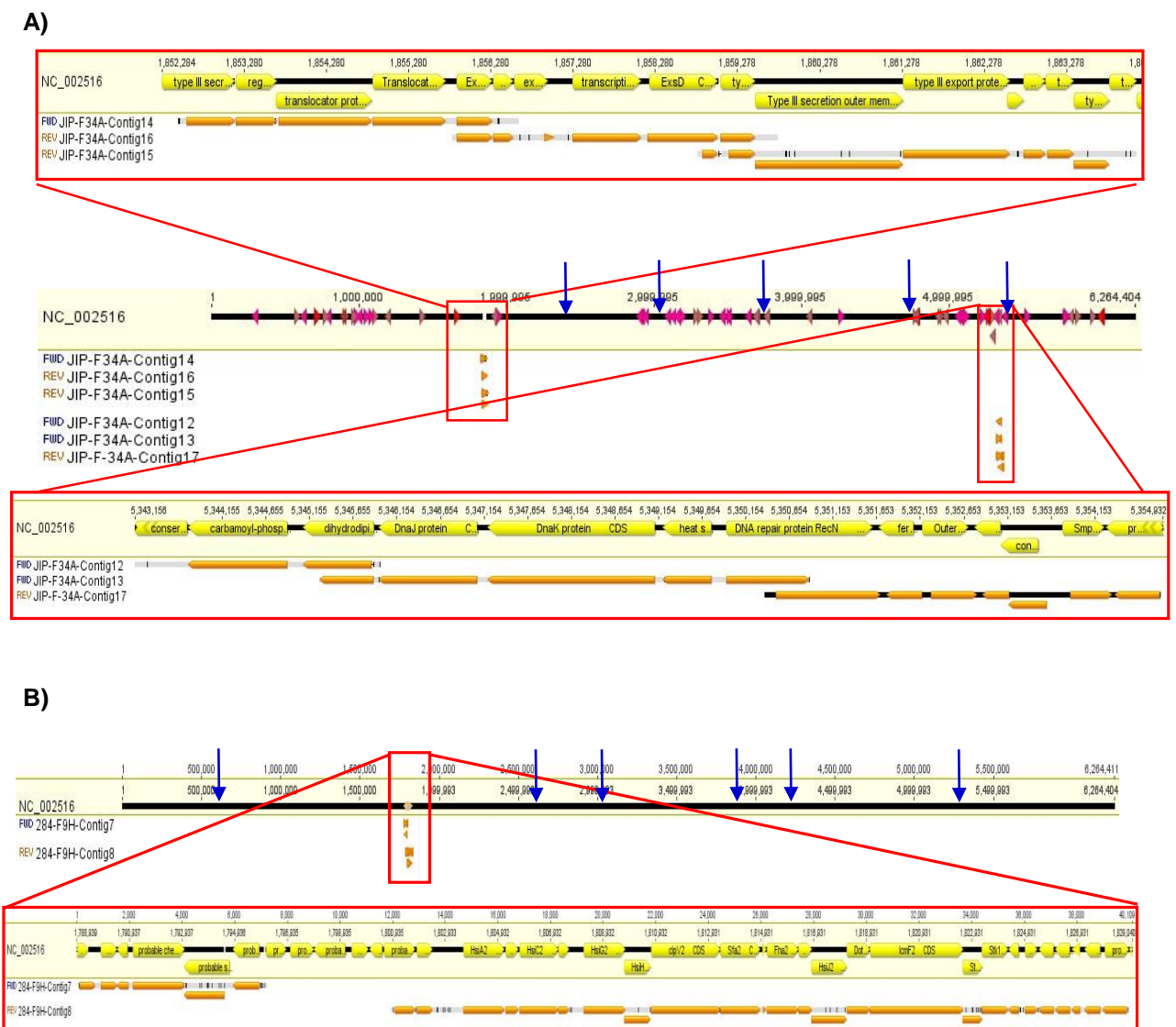
<i>orf6</i>	putative malonate decarboxylase gamma-subunit	<i>St. maltophilia</i> K279a	AM743169.1	711/711(100%)	0
<i>orf7</i>	putative coenzyme-A synthase	<i>St. maltophilia</i> K279a	AM743169.1	988/1014(97%)	0
<i>orf8</i>	putative malonyl CoA-acyl carrier protein transacylase	<i>St. maltophilia</i> K279a	AM743169.1	881/922(96%)	0
<i>orf9</i>	putative transmembrane dicarboxylate carrier protein	<i>St. maltophilia</i> K279a	AM743169.1	1351/1368(99%)	0
<i>Stenotrophomonas maltophilia</i> isolate LS5 fosmid LS5C2E contig 1(size 14241-bp)					
<i>orf1</i>	putative transmembrane GGDEF domain signalling protein	<i>St. maltophilia</i> K279a	AM743169.1	190/191(99%)	1E-92
<i>orf2</i>	putative transmembrane GGDEF domain signalling protein	<i>St. maltophilia</i> K279a	AM743169.1	124/126(98%)	5E-55
<i>orf3</i>	putative transmembrane GGDEF domain signalling protein	<i>St. maltophilia</i> K279a	AM743169.1	960/966(99%)	0
<i>orf4</i>	putative transmembrane GGDEF signalling regulatory protein	<i>St. maltophilia</i> K279a	AM743169.1	1504/1506(99%)	0
<i>orf5</i>	putative transmembrane GGDEF signalling regulatory protein	<i>St. maltophilia</i> K279a	AM743169.1	258/258(100%)	2E-131
<i>orf6</i>	putative thioesterase	<i>St. maltophilia</i> K279a	AM743169.1	507/507(100%)	0
<i>orf7</i>	putative transmembrane protein	<i>St. maltophilia</i> K279a	AM743169.1	435/435(100%)	0
<i>orf8</i>	conserved hypothetical protein	<i>St. maltophilia</i> K279a	AM743169.1	432/432(100%)	0
<i>orf9</i>	conserved hypothetical protein	<i>St. maltophilia</i> K279a	AM743169.1	1381/1395(99%)	0
<i>orf10</i>	transcriptional regulator, ArsR family	<i>St. maltophilia</i> D457	HE798556.1	307/310(99%)	3E-155
<i>orf11</i>	conserved hypothetical exported protein	<i>St. maltophilia</i> K279a	AM743169.1	576/582(99%)	0
<i>orf12</i>	putative transmembrane protein	<i>St. maltophilia</i> K279a	AM743169.1	466/468(99%)	0
<i>orf13</i>	putative DNA-binding protein	<i>St. maltophilia</i> K279a	AM743169.1	224/225(99%)	2E-111
<i>orf14</i>	putative transmembrane transporter domain protein	<i>St. maltophilia</i> K279a	AM743169.1	1520/1539(99%)	0
<i>orf15</i>	putative transmembrane phosphoesterase	<i>St. maltophilia</i> K279a	AM743169.1	1108/1128(98%)	0
<i>orf16</i>	putative transmembrane protein	<i>St. maltophilia</i> K279a	AM743169.1	635/639(99%)	0
<i>orf17</i>	conserved hypothetical protein	<i>St. maltophilia</i> K279a	AM743169.1	705/736(96%)	0
<i>orf18</i>	conserved hypothetical protein	<i>St. maltophilia</i> K279a	AM743169.1	827/846(98%)	0
<i>orf19</i>	conserved hypothetical protein	<i>St. maltophilia</i> K279a	AM743169.1	471/477(99%)	0
<i>orf20</i>	putative transmembrane DoxX family protein	<i>St. maltophilia</i> K279a	AM743169.1	623/646(96%)	0
<i>orf21</i>	Probable signal peptide protein	<i>St. maltophilia</i> D457	HE798556.1	775/792(98%)	0

<i>orf22</i>	putative exported thioredoxin	<i>St. maltophilia</i> K279a	AM743169.1	520/546(95%)	0
<i>orf23</i>	putative two-component response regulator transcriptional regulatory protein	<i>St. maltophilia</i> K279a	AM743169.1	739/741(99%)	0
<i>orf24</i>	putative transmembrane sensor histidine kinase transcriptional regulatory protein	<i>St. maltophilia</i> K279a	AM743169.1	1298/1314(99%)	0
<i>orf25</i>	putative exported protein	<i>St. maltophilia</i> K279a	AM743169.1	978/1014(96%)	0
<i>orf26</i>	putative Major Facilitator Superfamily transporter	<i>St. maltophilia</i> K279a	AM743169.1	1230/1260(98%)	0
<i>orf27</i>	putative inosine-uridine preferring nucleoside hydrolase	<i>St. maltophilia</i> K279a	AM743169.1	508/513(99%)	0

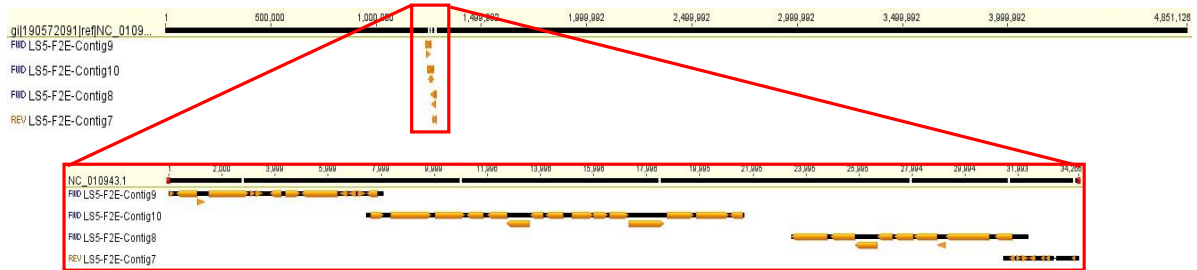
A1.5.8 Assembled contigs aligned against *Pseudomonas* genomes

To validate contig assembly and predicted *orf* identities for all five fosmid clones, contigs were aligned against genome sequences of *Ps. aeruginosa* PAO1 (NC_002516), *Ps. resinovorans* (NC_021499), and *Ps. mendocina ymp* (NC_009439). Alignments were generated using the Geneious Software package and up to 5 iterations were performed per alignment.

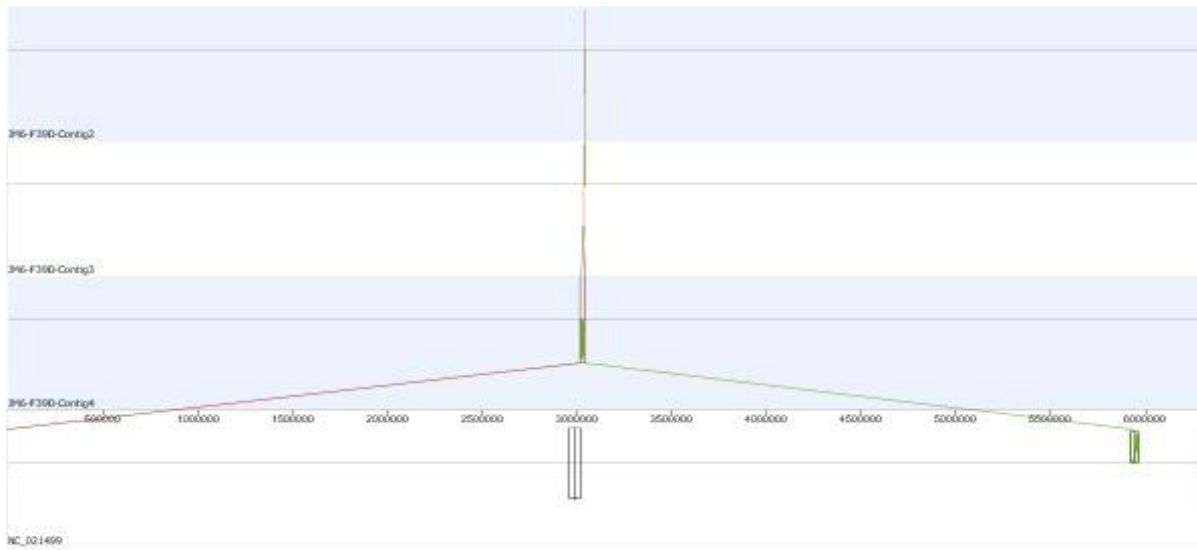
For fosmid clone JIP117F34A, contigs 12-17 were aligned against the completed *Ps. aeruginosa* PAO1 genome. The contigs align with extensive synteny at 2 loci, separated by 3-Mb to the reference genome. Similarly, SSI284C9H contigs align with greater than 90% mean identity to one loci within PAO1 genome with a 4-kb deletion between the contigs within this loci in the original *Ps. aeruginosa* SSI2.84 isolate.



C)



D)



E)

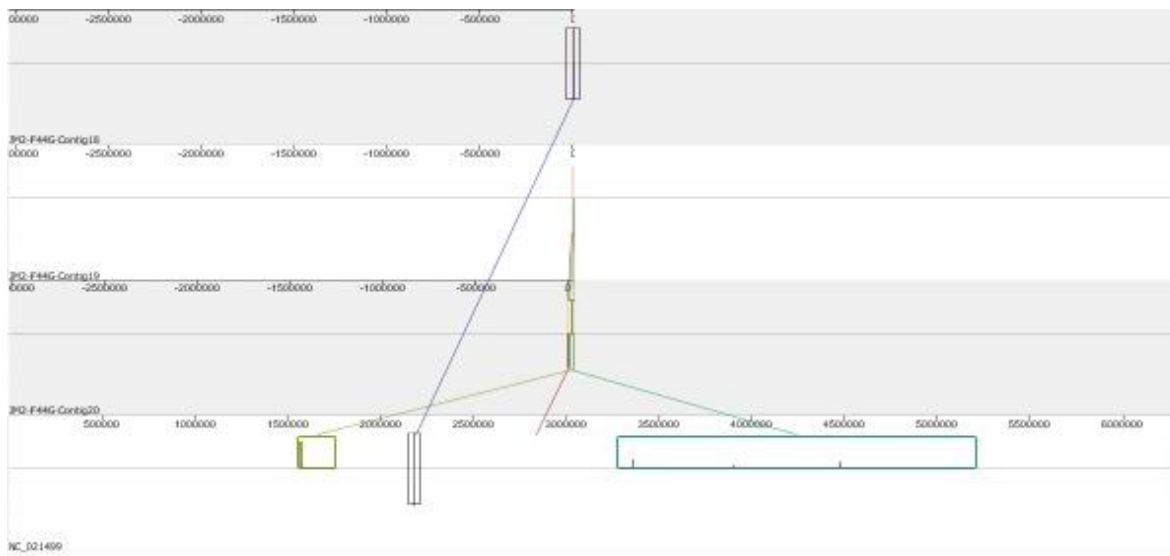


Figure A18: Genome mapping of JM2 and JM6 contigs against Ps. resinovorans NBC (NC_021499)

- A) *JIP117 fosmid contigs 12-17 aligned against Ps. aeruginosa PAO1 genome (NC_002516). Contigs 14, 15 and 16 aligned between the 1 852 294-1 865 423bp region. Contigs 12, 13 and 17 align between 5 343 156-5 361 217bp. Orange stretches indicated predicted orfs in contigs. Blue arrows indicate position of putative transposase in PAO1 genome.*
- B) *SSI284 fosmid contigs 7 and 8 aligned against Ps. aeruginosa PAO1 (NC_002516). A 4-kb deletion separates the contigs*
- C) *LS5C2E contigs align with greater than 90% mean identity to one loci within St. malthophilia K279a reference genome. All four contigs align between 1 248 699- 1 282 959bp with a 1.5-kb deletion between contigs 10 and 8.*
- D) *JM6 fosmid contigs 2, 3 and 4 aligned against reference genome. Contig 2 aligns with 1 loci at moderate synteny of 63%.*
- E) *JM2 fosmid contigs 18, 19 and 20 aligned against reference genome. Contig 18 is has aligned to 1 loci with moderate synteny of 63%.*

APPENDIX A2: Exploring *IS1111-attC* translocation potential

Appendix A2.1 Screening *Pst587* natural transformants for *IS1111-attC* translocation using plasmid flanking primers, NVC15 and NVC16

Initial attempts to detect *IS1111-attC* translocation involved the use of crude total plate DNA as the PCR template and the flanking NVC15 and NVC16 primer pair (Figure A19). However, this method was not sensitive enough to detect IS translocation. Instead, individual transformants were screened for translocation (Figures A20 & A21). Screening individual transformants by these means identified *IS1111-attC* positive translocation into pPsattC1 trap only. Note that no positive controls were available for any of these PCR screens. Expected sizes were 1.8 and 1.7 kb for pPsattC and paadB-attC translocation respectively.

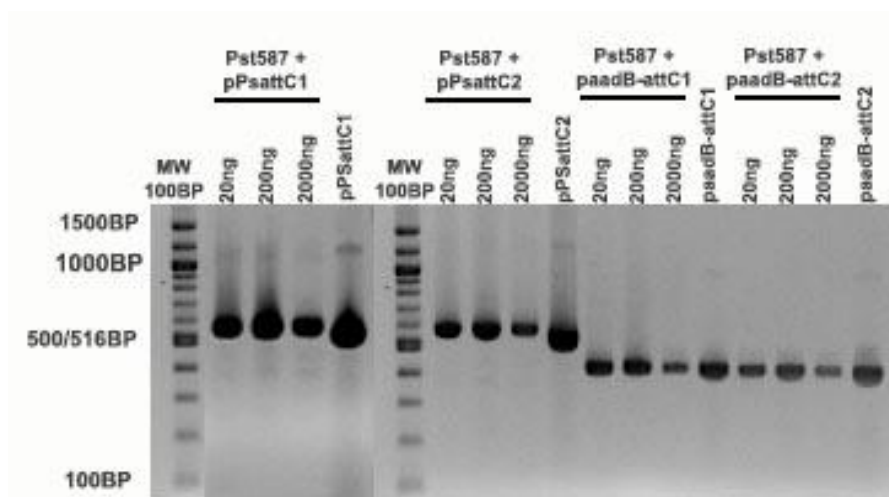


Figure A19: Crude total plate harvest and plasmid flanking primers were not sensitive enough for the detection of an *attC*-trap invasion.

Empty *pPsattC* traps and *paadB-attC* traps generate PCR products of 550-bp and 420-bp respectively. PCR product was generated using NVC15 and NVC 16 primers and total plate boil lysed DNA. In all screens, 20 ng, 200 ng and 2000 ng of boil lysed DNA was used as PCR template. Purified *pattC* traps (*pPsattC1*, *pPsattC2*, *paadb-attC1* and *paadB-atttc2*) were used as a control.

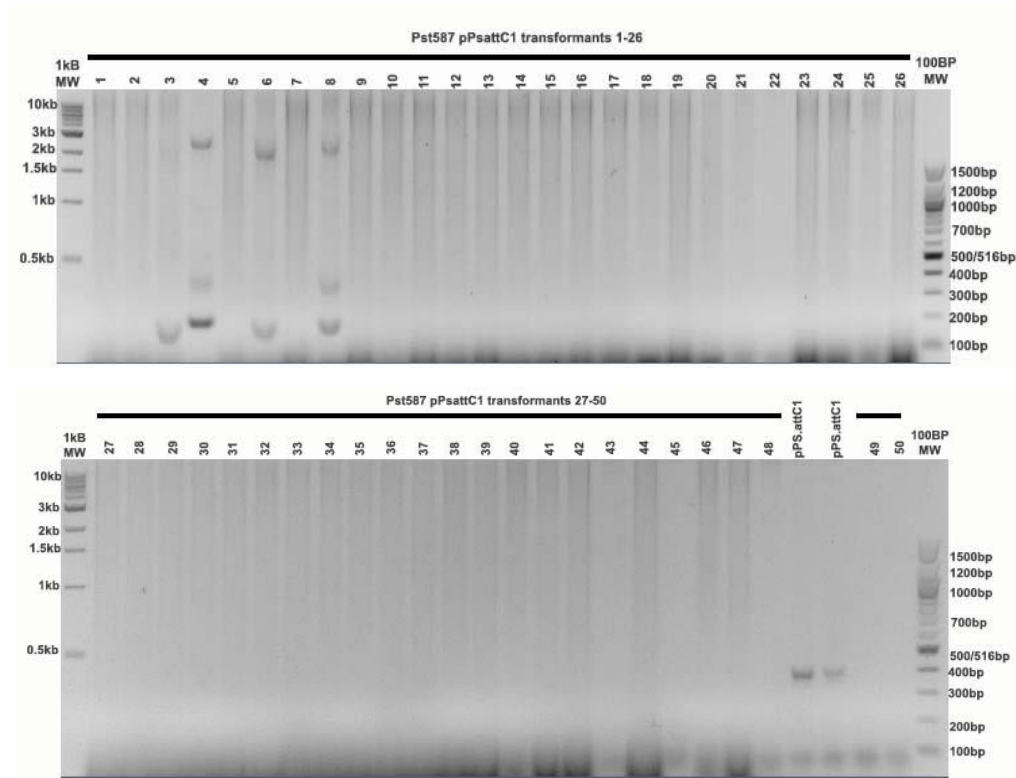
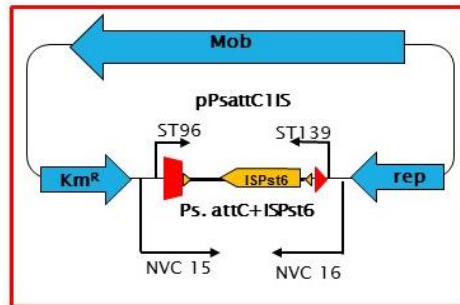


Figure A20: Identifying IS1111-attC translocation positive *Pst587* transformants by use of *pattC* flanking primers. Lane 51 and 52: pPs-attC1 constructs only. Expected product size is ~1.8-kb. Boil lysed DNA was used as PCR template with primers NVC15 and NVC16. Transformants 4, 6 and 8 are positive for intact translocation. Lanes 1-50: correspond to 50 *Pst581* transformants screened.

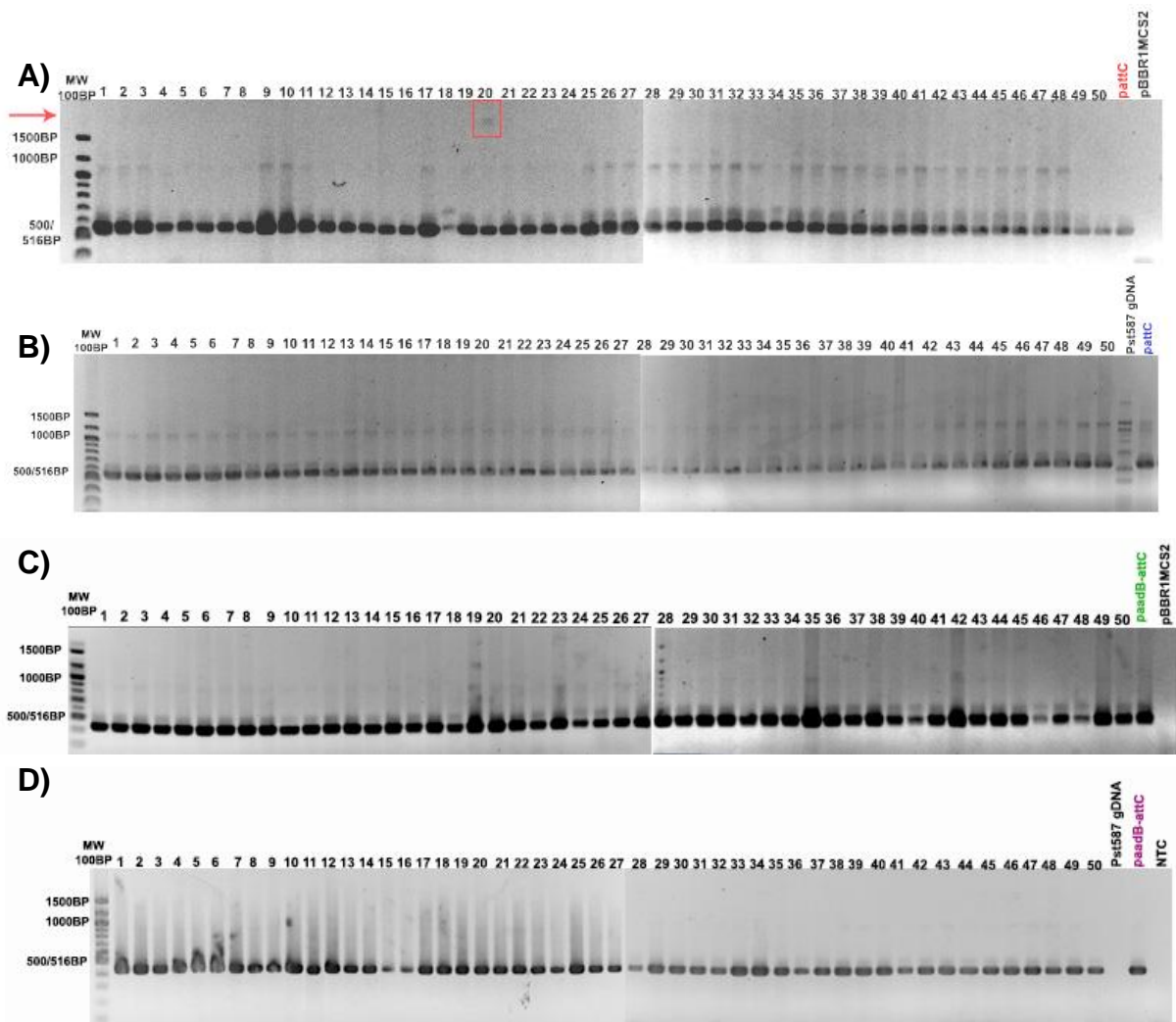


Figure A21: Screening a) *pPsattC1* b) *pPSattC2* c) *paadB-attC1* d) *paadB-attC2* transformants for intact *ISPst6* translocation. All showed a PCR signal for empty *attC* site.

Only lane 20 of panel (a) showed a product of 1.8-kB, indicating insertion of an *ISPst6* element into the *Pseudomonas attC* site. This transformant was annotated as Natural Transformant 20 (NT20) *pPsattC1-IS* and has a mixed plasmid population composed of *pPsattC1* and *pPsattC1-IS*.

A2.2 Trap-IS junction PCR did not detect *IS1111-attC* translocation positive *Pst587* transformants (pPsattC2 and paadB_attC2 traps).

An alternative screening method, other than using *pattC* flanking primers NVC 15 and NVC16, is the use of one flanking primer with an *ISPsst6* specific primer (ST123). NVC15 and ST123, were used to detect *IS1111-attC* translocation positive *Pst587* transformed with pPsattC2 (←) and paadB_attC2 (←). In each case, 50 transformants were screened with this primer combination. However, no positive transformants were detected. Note, there is no available positive control for this PCR. *In silico* prediction of the expected product sizes are 384-bp and 332-bp for pPsattC2 and paadB_attC2 respectively.

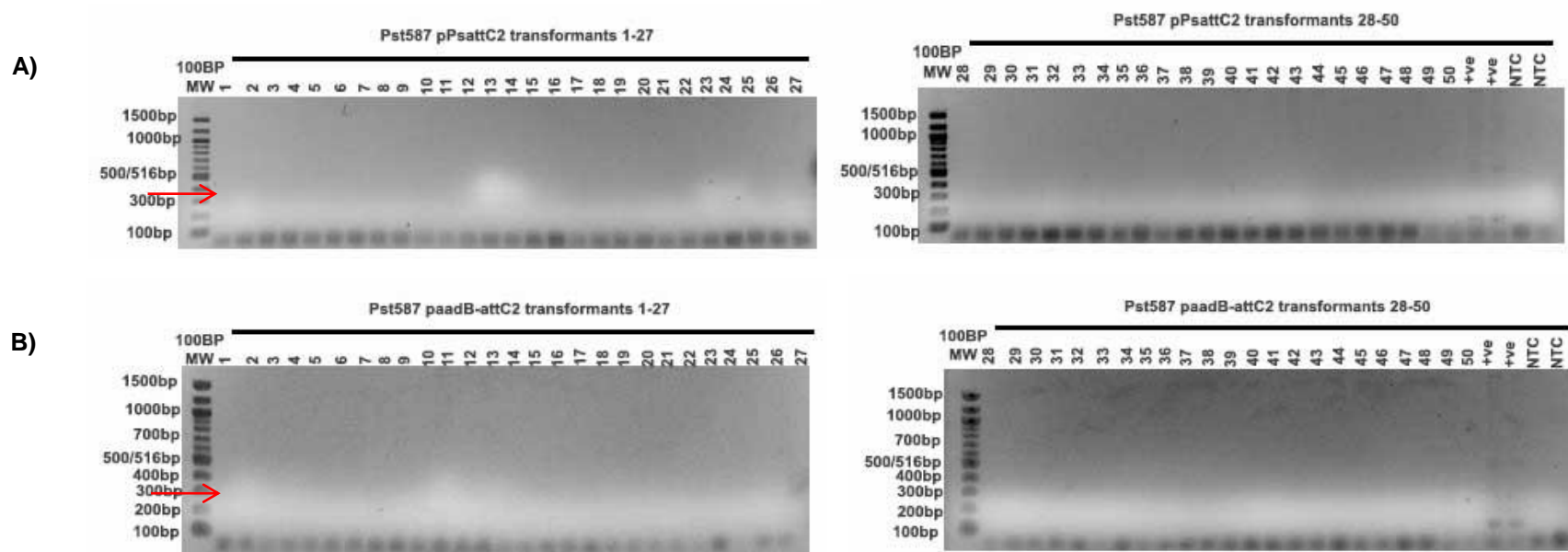


Figure A22: Trap-IS junction PCR did not detect *IS1111-attC* translocation positive *Pst587* transformed with a) pPsattC2 and b) paadB_attC2 traps. Lanes 1-50: correspond to 50 *Pst587* transformants. Lanes 51 and 52: *pattC* trap plasmid DNA. Lanes 53 and 54: NTC not template control. Boil lysed DNA was used as PCR template with primers NVC15 and NVC16.

A2.3 Mixed plasmid population, paadB_attC1 and paadB_attC1-IS is present in *Pst587* NT1

Junction screening *Pst587* transformants with paad_attC1 was expected to produce a single PCR product of ~280-bp. Instead, the product was over 300-bp. An additional band of ~500-bp was observed, however this band was also present in the negative control. The plasmid was recovered from *Pst587* NT1 and was rescreened for both the plasmid-IS junction and flanking product. The expected product size of ~298-bp corresponding to the plasmid-IS junction was observed when purified *Pst587* NT1 plasmid preparation was used as template DNA. In comparison the positive control pPsattC1-IS produced a product whose size was greater by ~100-bp than expected. Furthermore, screening the NT1 plasmid preparation using plasmid flanking primer (NVC15/NVC16) gave a PCR product of ~400-bp, corresponding to empty *attC_{aadB}* flanking size.

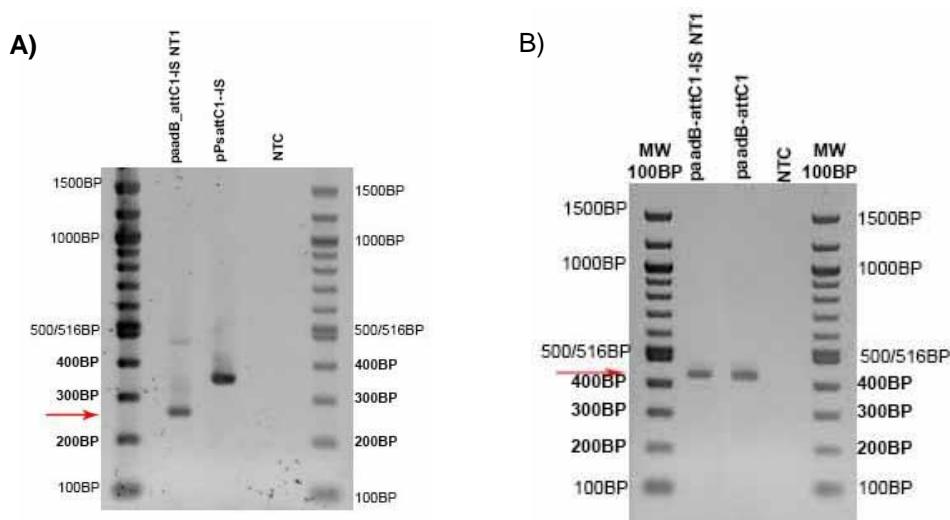
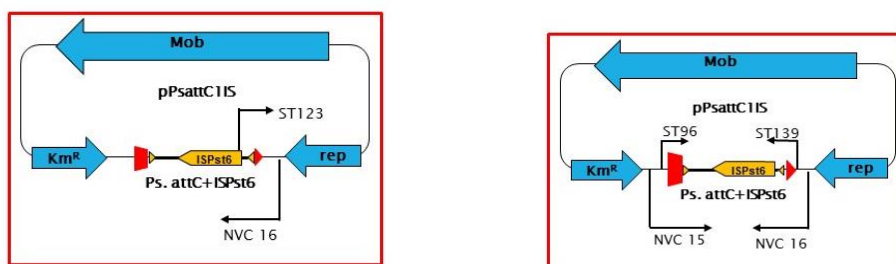


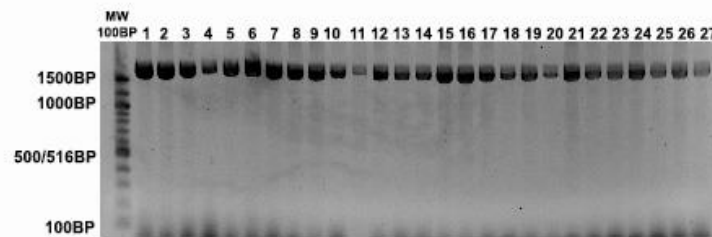
Figure A23: Junction and flanking PCR NT1 plasmid DNA suggests the presence of a mixed plasmid population.

A) Junction PCR for *ISPst6* element within *paadB-attC1* trap. Primers *NVC16* and *ST123*. Lane 1: *paadB_attC1-IS* plasmid DNA from NT1, lane 2: *pPsattC1-IS* control DNA Lane 3: NTC (no template DNA). Note lanes 1 is expected to be ~100-bp smaller to *pPsattC1-IS*. B) Flanking PCR using *NVC15* and *NVC16* presence of empty *aadB-attC1* site. Lane 1: *paadB_attC1-IS* plasmid DNA from NT1, lane 2: *paadB_attC1* control, Lane 3: NTC (no template DNA). Empty *attC* site is ~400-bp.

A2.4 *Pst405* and *PstQ* transformants maintain intact pPsattC1-IS constructs.

Flanking primers (NVC15/NVC16) were used to screen for a loss of *IS1111-attC* elements from the introduced construct. If *IS1111-attC* translocation was by a cut-&-paste mechanism, then restoration of an empty *Pseudomonas*-type *attC* site was expected. The PCR product size would have dropped from ~1.8-kb to ~400-bp. In all transformants screened, restoration was not observed.

A)



B)

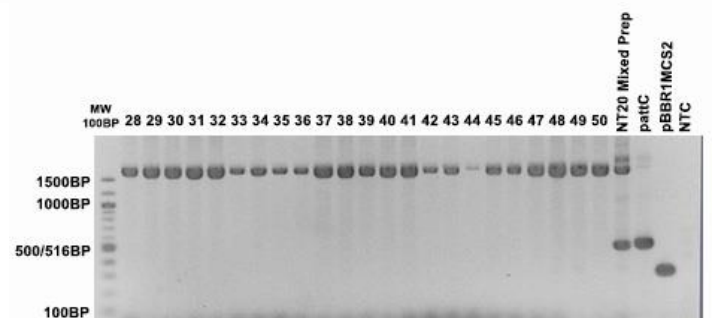
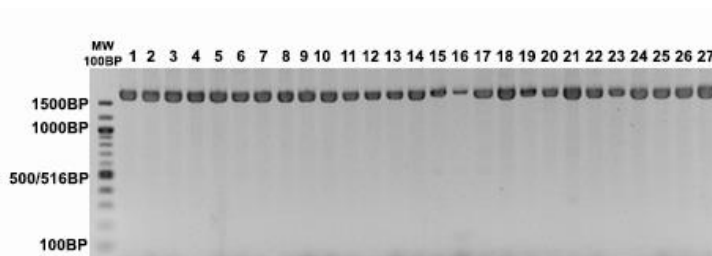
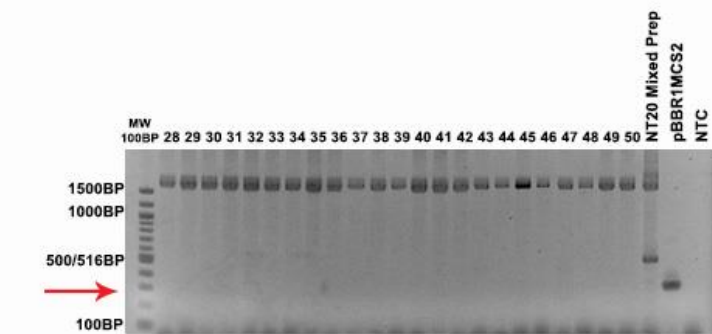


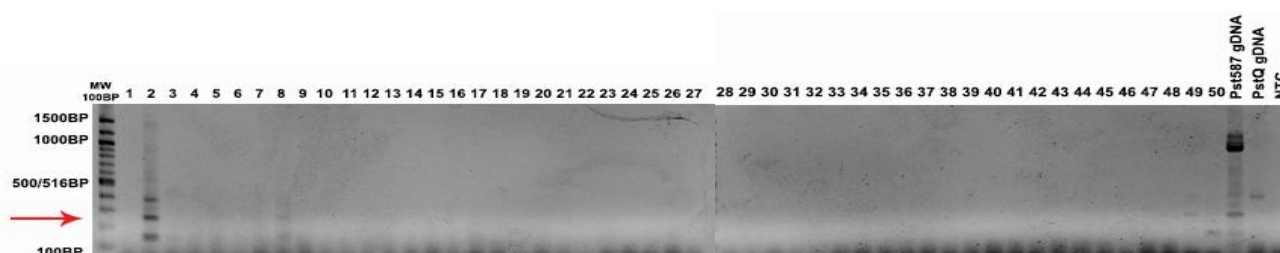
Figure A24: A) *Pst405* transformants and B) *PstQ* transformants maintain intact pPsattC1-IS constructs.

Lanes 1-50: correspond to 50 transformants screened. Boil lysed DNA was used as PCR template. Lane 51: NT20 mixed plasmid preparation used as a control where a transformant may contain a mixed plasmid preparation of pPsattC1-IS and pPsattC1. Lane 52: *pattC* constructs only. Lane 53 pBBR1MCS2 vector and Lane 54: no template control.

A2.5 Detection of IS*Pst6* minicircles in *PstQ* and *Pst641* transformants

To screen for the formation of IS*Pst6* minicircles in *PstQ* and *Pst641* electrocompetent transformants, individual transformants were picked, DNA was prepared by boil lysis prior to PCR. IS*Pst6* specific minicircle primers ST89 and ST123 were used. These outwards facing (reverse) primers anneal at either end of the IS element. A product of ~280-bp is only possible if the IS has formed a minicircle.

A) *PstQ*



B) *Pst641*

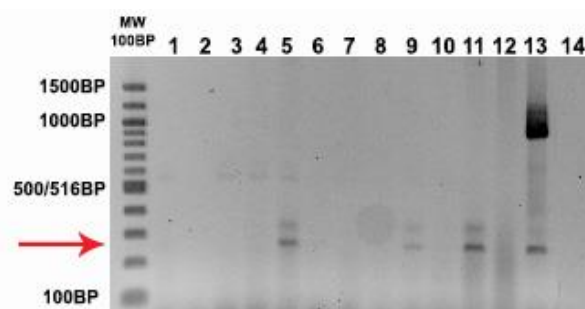


Figure A25:

- a) *PstQ* transformants 2, 8, 49 and 50 potentially have IS*Pst6* minicircles.
Lanes 1-50: correspond to 50 *PstQ* transformants screened. Lane 51: *Pst587* gDNA positive for IS*Pst6* minicircles at 258-bp. Lane 52: *PstQ* gDNA negative for IS*Pst6* minicircle. Lane 53: no template control.
- b) *Pst641* transformants 5, 9 and 11 potentially have IS*Pst6* minicircles.
Lanes 1-12: correspond to 12 *Pst641* transformants screened. Lane 13: *Pst587* gDNA positive control for IS*Pst6* minicircles. Lane 14: no template control.

A2.5.1 Recovered *ISPst6* minicircle junction sequences from *PstQ* and *Pst641* transformants

ISPst6 minicircle junction DNA was recovered from PstQ ET2 and Pst641 ET9. In both sequences, three nucleotide base differences were present when compared to the wild-type *Pst587* sequence (Figure A26). At the alignment position 95 a G replaced an A, at position 109 a C replaced a T and at position 123 a G replaced an A.

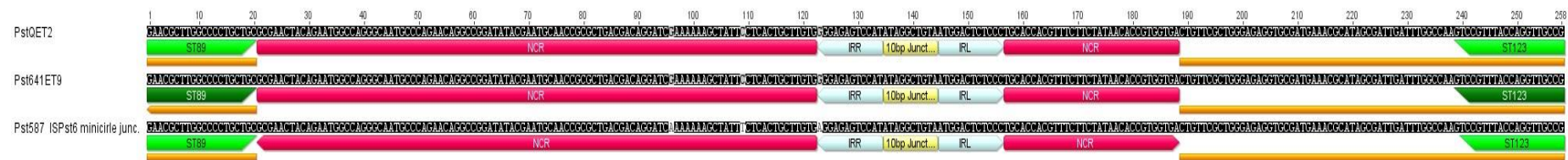


Figure A26: Alignment of *PstQET2*, *Pst641ET9* and *Pst587* *ISPst6* minicircle junctions. Primer sequences *ST89* and *ST123* are annotated in green. Non-coding regions (*NCR*) are annotated in hot pink, the inverted repeats (*IRR* and *IRL*) are in light blue and the 10-bp junction that forms upon circularization is shown in yellow.

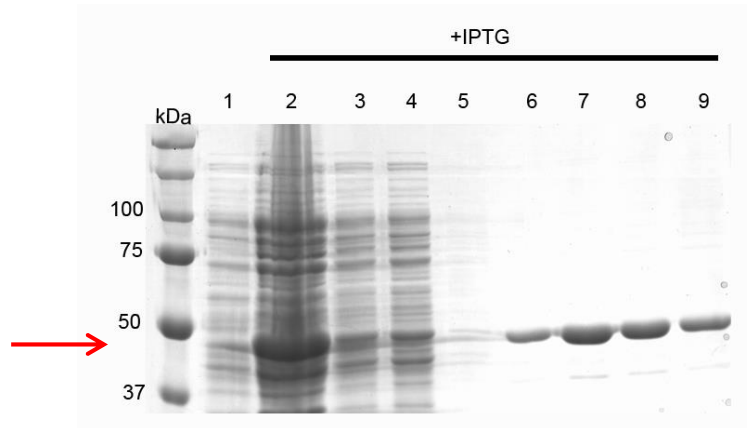
APPENDIX A3: MBP Control**Appendix A3.1 Expression of MBP as a control**

Figure A27: SDS-PAGE analysis of purified MBP protein. MBP was expressed in E.coli Rosetta2. Lane 1: Uninduced Rosetta2 cells. Lane 2: Cell lysate was separated by centrifugation to give soluble fraction. Lanes 3-4: Flow through amylose resin, which was washed in excess column buffer Lane 5. Lanes 6-9: MBP was eluted with 50mM maltose. Expected protein size is ~42 kDa.