

**Antibiotic resistance in global clone 2 *Acinetobacter*
baumannii from Singapore**

Grace Annie Bell Blackwell

A thesis submitted in the fulfilment of the requirements for the degree of

Doctor of Philosophy

Faculty of Science

School of Life and Environmental Science

The University of Sydney

NSW, Australia

2018

Declaration

This is to certify that to the best of my knowledge; the content of this thesis is my own work.

This thesis has not been submitted for any degree or other purposes.

I certify that the intellectual content of this thesis is the product of my own work and contains no material previously published or written by another person, unless an appropriate reference is stated in the text.

Grace Blackwell

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my supervisor Professor Ruth Hall. The mentorship she has provided for me over the past four years has been invaluable, particularly the time she has spent helping me interpret data, solve problems and critique my writing. Without her, this thesis would not have been possible and I feel privileged to have been one of her students. I wish to acknowledge the past and present members of the Hall laboratory for their assistance and support, and for making the lab a fun environment. I would particularly like to thank Steven and Mohammad for always being available to answer any questions I had, and Robert who was a constant companion through my PhD. I would also like to thank Chris, Carol, Stephanie, Jade, Isabella and Rebecca for making the time spent at the lab enjoyable. Furthermore, I would like to thank the staff and fellow students in G.08 who contributed to this building having such a wonderful environment.

I would like to thank my parents, Frances and George. Their support and encouragement through my entire education, not just my PhD, was invaluable and it would not have been possible to get through it without them. I would also like to mention my siblings, Thomas, Clare, Emily, Patrick and Isabel, who without your impact in my life, I would not be who I am today. Finally, I would like to thank my partner Michael for all the love, support, encouragement, proof-reading and patience he has provided, particularly during the writing of this thesis.

PUBLICATIONS

Publications generated during the course of this study

Papers:

1. **Blackwell, G.A.**, Nigro, S.J., Hall, R.M. Evolution of AbGRI2-0, the progenitor of the AbGRI2 resistance island in global clone 2 of *Acinetobacter baumannii*. *Antimicrobial Agents and Chemotherapy*, 2015; 60(3), 1421-1429. PMID: 26666934
2. **Blackwell, G.A.**, Hamidian, M., Hall, R.M. IncM plasmid R1215 is the source of chromosomally located regions containing multiple antibiotic resistance genes in the globally disseminated *Acinetobacter baumannii* GC1 and GC2 clones. *mSphere*, 2016; 1(3), e00117-16. PMID: 27303751
3. **Blackwell, G.A.**, Holt, K.E., Bentley, S.D., Hsu, L.Y., Hall, R.M. Variants of AbGRI3 carrying the *armA* gene in extensively antibiotic resistant *Acinetobacter baumannii* from Singapore. *Journal of Antimicrobial Chemotherapy*, 2017, 72(4), 1031-1039. PMID: 28073968
4. **Blackwell, G.A.**, Hall, R.M. The *tet39* determinant and the *msrE-mphE* genes in *Acinetobacter* plasmids are each part of discrete modules flanked by inversely-oriented *pdif* (XerC-XerD) sites. *Antimicrobial Agents and Chemotherapy*, 2017 doi:10.1128/AAC.00780-17. PMID: 28533235

GenBank database submissions:

1. **Blackwell G.A.**, Holt K.E., Pickard D. and Hall R.M. *Acinetobacter baumannii* strain RUH134 KL9 capsule biosynthesis locus, genomic resistance island AbGRI1-1 (Tn6166), genomic resistance island AbGRI2-0a, genomic resistance island AbGRI2-0b, and OCL1 outer-core biosynthesis locus. Update to GenBank accession number: JN247441; bases 48,558 to 60,783; 60,884 to 88,292; 39,633 bp. Submitted 30/06/2015.
2. **Blackwell G.A.** and Hall, R.M. *Serratia marcescens* strain NCTC 50331 plasmid R1215, complete sequence. GenBank accession number: KU315015; 95,855 bp. Submitted 17/12/2015.
3. **Blackwell G.A.** and Hall R.M. *Acinetobacter baumannii* strain SGH0701 genomic resistance island AbGRI3, complete sequence. GenBank accession number: KX011025; bases 30477 to 58194; 27,717 bp. Submitted 31/03/2016.
4. **Blackwell G.A.** and Hall R.M. *Acinetobacter baumannii* strain SGH0908 genomic resistance island AbGRI3, complete sequence. GenBank accession number: KX011026; 13,418 bp. Submitted 31/03/2016.
5. **Blackwell G.A.** and Hall R.M. *Acinetobacter baumannii* strain SGH1111 genomic resistance island AbGRI3, complete sequence. GenBank accession number: KX011027; 21,973 bp. Submitted 31/03/2016.
6. **Blackwell G.A.** and Hall R.M. *Acinetobacter baumannii* strain SGH0823 plasmid pS30-1, complete sequence. GenBank accession number: KY617771; 18,234 bp. Submitted 16/02/2017.

AUTHORSHIP ATTRIBUTION STATEMENT

Chapter 6 of this thesis is published as:

Blackwell, G.A., Holt, K.E., Bentley, S.D., Hsu, L.Y., Hall, R.M. Variants of AbGRI3 carrying the *armA* gene in extensively antibiotic resistant *Acinetobacter baumannii* from Singapore. *Journal of Antimicrobial Chemotherapy*, 2017, 72(4), 1031-1039.

I performed the experimental work, analysed the data, prepared figures and tables and wrote the drafts of the MS.

Chapter 10 of this thesis is published as:

Blackwell, G.A., Hall, R.M. The *tet39* determinant and the *msrE-mphE* genes in *Acinetobacter* plasmids are each part of discrete modules flanked by inversely-oriented *pdif* (XerC-XerD) sites. *Antimicrobial Agents and Chemotherapy*, 2017 doi:10.1128/AAC.00780-17.

I performed the experimental work, analysed the data, prepared figures and tables and wrote the drafts of the MS.

Grace Blackwell, 27/10/17

As supervisor for the candidature upon which this thesis is based, I can confirm that the authorship attribution statements above are correct.

Prof. Ruth Hall, 27/10/17

ABSTRACT

Extensively and pan antibiotic resistant *Acinetobacter baumannii* belonging to the global clone GC2 present a significant challenge in the treatment of infections worldwide. This study aimed to analyse the causes of antibiotic resistance in a selection of 20 extensively antibiotic resistant GC2 isolates from Singapore. One isolate was from 1996 and others were from 2004-2011. All carried multiple antibiotic resistance genes including the carbapenem resistance gene *oxa23*.

To determine the relationships between these isolates the structures of the AbGRI1, AbGRI2 and AbGRI3 resistance islands were examined, along with ISAbal locations and MLST profiles (Oxford scheme). Plasmid diversity was also examined. Most isolates were ST208 or a single locus variant and one was a triple locus variant. All had ISAbal upstream of *ampC* (resistance to third generation cephalosporins) and a version of AbGRI1 and AbGRI2. However, AbGRI3 including the *armA* gene (resistance to all clinically relevant aminoglycosides) was present in 15 isolates.

Shared features allowed evolutionary lineages to be inferred. The 1996 ST208 isolate shared many features with most Australian isolates including AbGRI2-1, and 11 of 12 conserved ISAbal positions, and hence is closely related to their progenitor. Four other isolates had AbGRI2-1 but ISAbal in only 8 conserved positions. The remaining 15 isolates contain an AbGRI3 variant and an IS26-derived deletion variant of AbGRI2-1. They share 3 of the conserved ISAbals and 2 more ISAbal positions specific to the AbGRI3 lineage. Six AbGRI3 variants were detected. One was associated with 6 extra ISAbal positions.

RepAci6 plasmids of 2 known backbone types and 2 novel types were found in the collection; most were transfer proficient. Of the 14 plasmids, 2 were cryptic, 4 carried *aphA6* (amikacin

resistance) in *TnaphA6*, 10 had *oxa23* in *Tn2006* in AbaR4 and one had *oxa23* in *Tn2008*. Two novel transposons were also identified. All AbGRI3 isolates had a small RepAci1 plasmid with no mobilisation genes and one carrying *Tn2006* was mobilised by a RepAci6 plasmid enabling a shared putative *oriT* to be found.

The *tet39* tetracycline resistance determinant and the macrolide resistance gene pair *msrE-mphE* were found in a novel plasmid in one isolate. These resistance genes are in *dif* modules, bounded by inversely-oriented XerC-XerD binding sites, termed *pdif* sites. Numerous *dif* modules containing many different orfs, including various toxin/antitoxin systems, were identified in *Acinetobacter* plasmids. Two types of *dif* modules, C and D modules, that usually alternate were defined.

ABBREVIATIONS

Units

bp	Base pair
kb	Kilo base
Mb	Mega base
M	Molar
Mol	Moles
U	Units of activity (of enzyme)
v/v	Volume per volume
w/v	Weight per volume

Terms

aa	Amino acid
AbaR	<i>A. baumannii</i> resistance island
AbGRI	<i>A. baumannii</i> genomic resistance island
CDS	Calibrated Dichotomous Sensitivity
CLSI	Clinical and Laboratory Standards Institute
CR	Common region
5'-CS	5'-conserved segment
3'-CS	3'-conserved segment
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphate
DR	Direct repeat
EC	European clone
GC	Global clone
EDTA	Ethylenediaminetetraacetic acid
EtBr	Ethidium bromide
HGT	Horizontal gene transfer
ICU	Intensive care unit
In	Integron
IS	Insertion sequence
IR	Inverted repeat
LB	Luria Bertani
MAR	Multiply antibiotic resistant
MARR	Multiple antibiotic resistance region
MIC	Minimum inhibitory concentration
MLST	Multi-locus sequence typing
MQ	Milli-Q water
NCBI	National Centre for Biotechnology information
nt	Nucleotide
orf	Open reading frame
PCR	Polymerase chain reaction
PFGE	Pulse-field gel electrophoresis
RI	Resistance island
SDS	Sodium dodecyl sulfate
SNP	Single nucleotide polymorphism
sp.	Species
ST	Sequence type
TBE	Tris-borate EDTA
TE	Tris-EDTA
Tn	Transposon
Tris	Tris(hydroxymethyl)aminomethane
UV	Ultra-violet
V	Volts

Antibiotics

IPM	Imipenem
MER	Meropenem
Dor	Doripenem
Ap	Ampicillin
SAM	Ampicillin/sulbactam
TIM	Ticarcillin-clavulanic acid
TZP	Piperacillin-tazobactam
CTX	Cefotaxime
CAZ	Ceftazidime
CRO	Ceftriaxone
FEP	Cefepime
Sm	Streptomycin
Sp	Spectinomycin
Su	Compound sulphonamides
Tc	Tetracycline
Mc	Minocycline
Dc	Doxycycline
Tp	Trimethoprim
Cl	Chloramphenicol
Fl	Florfenicol
Km	Kanamycin
Nm	Neomycin
Gm	Gentamicin
Ne	Netilmicin
Ak	Amikacin
Tm	Tobramycin
Nx	Nalidixic acid
CIP	Ciprofloxacin
LVX	Levofloxacin
Rif	Rifampicin
CLAV	Clavulanic acid

TABLE OF CONTENTS

Chapter 1	General Introduction	1
1.1	The genus <i>Acinetobacter</i>	2
1.2	<i>A. baumannii</i>	3
1.3	Antibiotic resistance in <i>A. baumannii</i>	5
1.3.1	Intrinsic resistance	6
1.3.2	Resistance to cephalosporins	6
1.3.3	Carbapenem resistance	7
1.3.4	Aminoglycoside resistance	10
1.3.5	Quinolones and fluoroquinolones	13
1.3.6	Tetracycline resistance	13
1.3.7	Sulphonamide and trimethoprim resistance	13
1.3.8	Colistin resistance	14
1.3.9	Rifampicin resistance	14
1.3.10	Chloramphenicol resistance	14
1.3.11	Macrolide resistance	15
1.3.12	Additional efflux pumps involved in antibiotic resistance	15
1.4	<i>A. baumannii</i> clones	15
1.5	<i>A. baumannii</i> genomics	17
1.6	Chromosomal resistance islands	20
1.6.1	GC1 AbaR islands	20
1.6.2	AbaR4, Tn6022 and Tn6021	22
1.6.3	Chromosomal resistance islands in GC2	23
1.6.3.1	Resistance islands in <i>comM</i>	23
1.6.3.2	AbGRI2 resistance island	26
1.6.3.3	The origin of AbGRI2	29
1.6.3.4	A third chromosomal resistance island	30
1.7	Tn6019 and Tn6021 and Tn6022 family	31
1.8	IS elements	32
1.8.1	IS26	32
1.8.2	ISAbal	33
1.9	Plasmids	34
1.9.1	<i>Acinetobacter</i> plasmids	35
1.9.1.1	Bertini PCR-based plasmid replicon typing scheme	36
1.9.1.2	RepAci1 and RepAci2 plasmids	36
1.9.1.3	RepAci6 plasmids	36
1.9.2	Conjugative plasmids not included in the Bertini PCR-based replicon typing scheme	37
1.10	Epidemiology of GC2 <i>A. baumannii</i> from Australia and South-east Asia	37
1.10.1	GC2 <i>A. baumannii</i> in Australia	37
1.10.2	GC2 <i>A. baumannii</i> in South-east Asia	38
1.11	Sequencing technologies	39
1.11.1	Second generation sequencing technology	39
1.11.2	Assembly of second generation sequencing	40
1.11.3	Third generation sequencing technology	41
1.12	Aims and scope	42
Chapter 2	Materials and Methods	44
2.1	Materials	45
2.1.1	Chemicals and reagents	45
2.1.2	Buffers and other solutions	45
2.1.3	Growth media	45
2.1.4	Enzymes	46
2.2	Bacterial isolates	47
2.3	Bacterial characterisation and manipulation	49
2.3.1	Bacterial growth	49
2.3.2	Antibiotic resistance phenotype determination using disk diffusion	50

2.3.3	Antibiotic resistance phenotype determination using Etest	50
2.3.4	Conjugation	51
2.3.5	Transformation	51
2.3.6	Plasmid stability assay	52
2.4	DNA preparation and analysis	53
2.4.1	DNA extraction	53
2.4.1.1	Genomic DNA for PCR and sequencing	53
2.4.1.2	Boil preparation	54
2.4.1.3	DNA purification technique of PCR products	54
2.4.1.4	Plasmid DNA	54
2.4.1.4.1	Plasmid DNA isolated from <i>E. coli</i>	54
2.4.1.4.2	Plasmid DNA isolated from <i>A. baumannii</i>	55
2.4.2	Polymerase Chain reaction	55
2.4.2.1	Primers	55
2.4.2.2	Standard PCR	56
2.4.2.3	Long-range PCR	56
2.4.3	Restriction endonuclease digestion	57
2.4.4	Agarose gel electrophoresis	57
2.4.4.1	PCR products and enzyme digestions	57
2.4.4.2	Plasmid DNA	58
2.5	DNA sequencing and sequence assembly	58
2.5.1	Sanger sequencing	58
2.5.2	Whole genome sequencing and sequence assembly	59
2.5.2.1	Illumina MiSeq	59
2.5.2.2	SPAdes assembly	59
2.6	Bioinformatic analysis	59
Chapter 3	Phenotypic and genotypic characterisation	61
3.1	Introduction	62
3.2	Phenotypic resistance profile	63
3.2.1	Antibiotic resistance phenotype determination using disk diffusion	63
3.2.2	Colistin resistance	65
3.3	Genotypic resistance profile	65
3.3.1	IS upstream of <i>ampC</i>	65
3.3.2	Resistance gene screening	66
3.3.2.1	Aminoglycoside resistance	67
3.3.3	Tn2006 screening	69
3.3.4	TnaphA6 screening	70
3.4	Detection of resistance gene content using the draft genome sequences	71
3.4.1	Resistance genes detected using ResFinder	71
3.4.2	Mutations leading to antibiotic resistance	73
3.4.2.1	Quinolone and fluoroquinolone resistance	73
3.4.2.2	Rifampicin resistance in S35	73
3.5	MLST	73
3.5.1	Institut Pasteur	74
3.5.2	Oxford	74
3.6	Discussion	75
Chapter 4	AbGRI1 and other chromosomal locations of <i>oxa23</i>	79
4.1	Introduction	80
4.2	Screening for Tn6022 and AbaR4	83
4.2.1	Detection of Tn6022 and/or Tn6022Δ1	83
4.2.2	Detection of AbaR4 or AbaR4Δ	84
4.2.3	Is an intact <i>sup</i> gene present?	85
4.3	AbGRI1 assembly	86
4.3.1	AbGRI1-3 in S21 and AbGRI1-4 in S27 and S34	87
4.3.2	AbGRI1-5 in S24 and S29	89
4.3.3	AbGRI1-1 in S38 and S39	91
4.3.4	Prediction of AbGRI1 variant forms in isolates not sequenced by Illumina HiSeq	93
4.3.4.1	Isolates without the <i>sul2</i> gene	94
4.3.4.2	Isolates containing the <i>sul2</i> gene	94
4.3.4.2.1	AbGRI1-7	96

4.4	Examination of the draft genomes from Illumina MiSeq sequencing	97
4.4.1	AbGRI1-8 in S28	98
4.4.2	AbGRI1 variants assembled	99
4.5	Additional chromosomal locations of <i>oxa23</i> containing transposons	101
4.5.1	AbaR4 in S22 and S23	101
4.5.2	AbaR4 in S31, S38 and S39	102
4.5.3	Tn2006 in S34	103
4.6	Discussion	104
Chapter 5	AbGRI2	107
5.1	Introduction	108
5.2	Assembly of AbGRI2 regions	110
5.2.1	AbGRI2-1 in S21, S24 and S29	110
5.2.2	AbGRI2-12 in S34	112
5.2.3	AbGRI2-13 in S38 and S39	114
5.2.4	No AbGRI2 variant in S27	116
5.3	PCR mapping of AbGRI2 forms in isolates not sequenced by Illumina HiSeq	117
5.3.1	AbGRI2-12 with a different adjacent deletion	119
5.3.2	Structure of AbGRI2-14, precursor of AbGRI2-13	120
5.4	Structures from the Illumina MiSeq draft genomes	121
5.4.1	No AbGRI2 variant in S35, S36 and S37	121
5.5	Discussion	122
Chapter 6	Characterisation of AbGRI3	126
6.1	Introduction	127
6.2	Transposons containing <i>armA</i> in GC2 <i>A. baumannii</i>	129
6.2.1	AbGRI3 in TYTH-1: AbGRI3-1	129
6.2.2	AbGRI3 in UH9907: AbGRI3-2	131
6.2.3	pJAB07104, p2JAB0868 and pMDR_ZJ06	132
6.3	AbGRI3 in the isolates from Singapore	133
6.3.1	AbGRI3-2i in S27	134
6.3.2	AbGRI3-3i in S38 and S39	138
6.3.3	AbGRI3-4 in S34	140
6.3.4	Predictions of AbGRI3 in isolates not sequenced by HiSeq	141
6.3.5	Final assembly using the draft genomes	144
6.3.5.1	AbGRI3-5 in S31	143
6.3.5.2	AbGRI3-6 in S25	144
6.3.6	Assembled AbGRI3 variants	146
6.4	AbGRI3 in completed genomes in Genbank	148
6.4.1	Examination of the chromosome in MDR_ZJ06, BJAB0868 and BJAB07104 at the AbGRI3 insertion site	151
6.4.2	A separate acquisition of a Tn6180 variant in a GC2 isolate	151
6.5	Discussion	153
Chapter 7	ISAbal mapping	158
7.1	Introduction	159
7.2	ISAbal mapping of the Singapore genomes	160
7.3	ISAbal positions in the AbGRI3 lineage	165
7.3.1	ISAbal positions 58 and 62 are indicative of AbGRI3 lineage	165
7.3.2	Six additional positions are indicative of AbGRI3-4	166
7.4	Relationship between GC2 isolates from Singapore and Australia	167
7.4.1	S21 contains 11 of the 12 ISAbal positions core to the east coast of Australian isolates	167
7.4.2	S24, S26 and S29 share positions with isolates from Melbourne and Adelaide	168
7.5	Discussion	168
Chapter 8	RepAci6 plasmids	173
8.1	Introduction	174
8.2	Screening for RepAci6 plasmids	176
8.3	Type 1 plasmids	180
8.3.1	Assembly	181
8.3.2	9.3 kb transposon in pS26-1 and pS29-1	184
8.3.3	Plasmid comparison and conjugation	187
8.4	Type 2 plasmids	189

8.4.1	Assembly	189
8.4.2	Backbone comparison and conjugation	191
8.5	The plasmid pS21-2	195
8.5.1	pS21-2 conjugation	198
8.5.2	pS21-2 is a new backbone type	198
8.5.3	Tn6177	200
8.6	pS30-2	203
8.6.1	New backbone type	204
8.6.2	Conjugation	206
8.7	Discussion	206
8.7.1	Type 1 plasmids	207
8.7.2	Type 2 plasmids	210
8.7.3	Other backbone types	210
8.7.4	Conserved orientation of AbaR4 transposon	213
Chapter 9	RepAci1 and RepAci2 plasmids	214
9.1	Introduction	215
9.2	Detection of RepAci1 and RepAci2 plasmids	217
9.3	Plasmid assembly and analysis	218
9.3.1	pAb-G7-1 type (RepAci1)	218
9.3.1.1	pS32-1 includes Tn2006	219
9.3.2	pD72-1 type (RepAci2)	220
9.3.2.1	pS21-1 contains TnaphA6	220
9.4	Plasmid compatibility	222
9.4.1	pS21-1 (RepAci1) and pS32-1 (RepAci2) are compatible	223
9.5	Plasmid mobilisation	225
9.5.1	Comparison of mobilisation genes	225
9.5.2	pS21-1 is not mobilised by a RepAci6 plasmid	227
9.5.3	Mobilisation of pS32-1 by RepAci6 plasmids	228
9.6	Putative <i>oriT</i> region of pS32-1	232
9.7	Discussion	234
Chapter 10	pS30-1	237
10.1	Introduction	238
10.2	pS30-1	240
10.3	pS30-1 transformation	244
10.4	RepAci6 plasmid does not mobilise pS30-1	246
10.5	XerC-XerD binding sites in pS30-1	247
10.5.1	<i>pdif</i> sites bound <i>tet39</i>	248
10.5.2	<i>pdif</i> sites bound the <i>msrE-mphE</i> gene pair	251
10.5.3	Additional <i>dif</i> modules in pS30-1	254
10.6	Two types of <i>dif</i> modules	258
10.6.1	C and D modules alternate	258
10.7	XerC-XerD experiments	262
10.7.1	pRCH51-1 and pBAL_204-1	262
10.7.2	pRCH51-1 and pBAL_204-1 in AB307_0294	263
10.8	ISAjo2 and ISAb32 target <i>dif</i> sites	265
10.8.1	ISAjo2 and relatives	265
10.8.2	ISAb32 and close relatives	268
10.9	Discussion	269
Chapter 11	General Discussion	273
11.1	Overview	274
11.2	Acquisition and evolution of antibiotic resistance islands in GC2 <i>A. baumannii</i>	278
11.3	Plasmids carrying antibiotic resistance genes	280
11.4	Plasmid evolution and movement	281
11.5	<i>dif</i> modules	282
11.6	Globalisation	282
11.7	Links to Australian isolates	283
11.8	Future directions	283
References		285
Appendix I		304
Appendix II		317

A2.1	AbGRI3-7 in D86	318
A2.2	RepAci6 plasmid assembly	319
A2.2.1	Type 1 plasmid assembly of pS29-1 and pS26-1	319
A2.2.2	Assembly of Type 2 plasmids	322
A2.2.2.1	Assembly of pS27-2 and pS34-2 from HiSeq sequencing	322
A2.2.2.2	Assembly of pA85-3 plasmids using MiSeq sequencing data	325
Appendix III		327

LIST OF FIGURES

Chapter 1		
1.1	The five structures harbouring <i>oxa23</i> in <i>A. baumannii</i>	8
1.2	Typical cassette array found in class 1 integrons in <i>A. baumannii</i>	11
1.3	Structure of <i>TnaphA6</i>	12
1.4	Schematic of the <i>A. baumannii</i> chromosome showing the location of genes used in the Pasteur and the Oxford MLST schemes	17
1.5	Completed <i>A. baumannii</i> genomes in GenBank	18
1.6	Structure of AbaR	21
1.7	Structure of the <i>Tn6018</i> flanked MARR of AbaR0	22
1.8	Structure of AbGRI1-2	24
1.9	AbGRI1 variants	25
1.10	AbGRI2	27
1.11	AbGRI2-0 is ancestral to AbGRI2-1	28
1.12	Relationship of the resistance island of R1215 to AbGRI2-0*	29
1.13	IS26 intramolecular transposition	33
Chapter 3		
3.1	ISAbal upstream of <i>ampC</i>	66
3.2	<i>Tn2006</i> screening	70
3.3	Structure of <i>TnaphA6</i>	71
Chapter 4		
4.1	AbaR4 structure	80
4.2	Variants of AbGRI1	81
4.3	Screening for <i>Tn6022</i> and/or <i>Tn6022Δ1</i>	84
4.4	Screening for <i>Tn2006</i> in AbaR4 and AbaR4Δ1	85
4.5	AbGRI1 mapping for S21, S27 and S34	88
4.6	Assembly of AbGRI1-5 in S24 and S29	90
4.7	Assembly of AbGRI1-1 in S38 and S39	93
4.8	Screening PCRs for the right end of AbGRI1	93
4.9	Structure of AbGRI1-7	97
4.10	Structure of AbGRI1-8 in S28	98
4.11	AbGRI1 variants	100
4.12	Location of AbaR4 in the chromosome of S22 and S23	102
4.13	Location of AbaR4 in the chromosome of S31, S38 and S39	103
4.14	Location of <i>Tn2006</i> in the chromosome of S34	104
Chapter 5		
5.1	Structure of AbGRI2-0 and AbGRI2-1	108
5.2	Expected contigs for the AbGRI2-1 structure	110
5.3	AbGRI2-1 contigs in S21, S29 and S24	111
5.4	Assembly of AbGRI2-12 in S34	113
5.5	Assembly of AbGRI2-13 in S38 and S39	115
5.6	Assembly of the region around AbGRI2 in S27	117
5.7	Amplification of the right end of AbGRI2-12 using primers RH881 and RH2008	119
5.8	Relationship between AbGRI2-13 and AbGRI2-14	120
5.9	AbGRI2 region of S35, S36 and S37	122
5.10	Structures in the AbGRI2 region	123
Chapter 6		
6.1	Structure of <i>Tn1548</i>	128
6.2	AbGRI3 and its chromosomal location	130
6.3	Comparison of <i>Tn1548</i> and <i>Tn6180</i>	131
6.4	AbGRI3-2 and <i>Tn6179</i>	132
6.5	Is the <i>atr</i> gene intact	134
6.6	Draft assembly of AbGRI3 contigs in S27	135
6.7	Assembly of AbGRI3 in S27	136
6.8	Comparison of AbGRI3-2 and 2i	137
6.9	Assembly of AbGRI3-3i in S38 and S39	139
6.1	Assembly of AbGRI3-4	141
6.11	Structure of AbGRI3-1i	142
6.12	Assembly of AbGRI3-5i	143
6.13	AbGRI3-5i in S31	145

6.14	AbGRI3-6 in S25	146
6.15	Six AbGRI3 variants	148
6.16	Variant of Tn6180 found in XH859	152
6.17	pMUR050, a potential source for the transposons in AbGRI3	155
Chapter 7		
7.1	Shared ISAbal positions	162
7.2	Location of ISAbal positions 1-5 and 12	163
7.3	Location of ISAbal position 58 and 62	165
7.4	Location of ISAbal positions 1-14	168
7.5	ML tree, island structure and ISAbal positions for 27 GC2 <i>A. baumannii</i>	170
Chapter 8		
8.1	RepAci6 plasmid backbone and repeat sequences	174
8.2	Comparison of the two RepAci6 plasmid backbone types	175
8.3	3 Screening PCRs over repeat 1, 2 and 3	177
8.4	Screening pAb-G7-2 location of TnaphA6	180
8.5	Three fragments of the RepAci6 backbone produced by assembly of Illumina HiSeq reads	181
8.6	Assembly of pS24-1	182
8.7	Linear comparison of pS24-1, pS26-1 and pS29-1	183
8.8	Features of the 9.3 kb novel transposon	184
8.9	Homology of Orf1 and 2 to the transposases of the IS21 family	186
8.10	Linear comparison of Type 2 plasmids	190
8.11	Assembly of pS21-2 contigs	196
8.12	Linear schematic of pS21-2	197
8.13	pS21-2 represents a new and third RepAci6 backbone type	299
8.14	A fourth repeat and novel sequence	200
8.15	Identification of the inverted repeat of the 35.6 kb transposon	201
8.16	Structure of the Tn6177 copper resistance transposon	201
8.17	Linear schematic of pS30-2	204
8.18	pS30-2 represents a new and fourth RepAci6 backbone type	205
8.19	Location of insertions in the type 1 backbone	208
Chapter 9		
9.1	pAb-G7-1 and pD72-1 comparison	216
9.2	Detection of pAb-G7-1 and pD72-1	218
9.3	pS32-1 contains Tn2006	220
9.4	pS21-1 contains TnaphA6	221
9.5	RepAci1 and RepAci2 iteron sequences	223
9.6	AB307_0294 containing pS21-1 and pS32-1	224
9.7	Plasmid visualisation in transconjugant	230
9.8	Restriction digest of transconjugant	230
9.9	Putative <i>oriT</i> region of pS32-1 and pS32-2	233
Chapter 10		
10.1	Sequences containing the <i>oxa24</i> and <i>oxa58</i> carbapenemase genes	239
10.2	Assembly and structure of pS30-1	241
10.3	Erythromycin resistance	246
10.4	pS30-1 <i>pdif</i> sites	248
10.5	The <i>tet39 dif</i> module	250
10.6	<i>msrE-mphe dif</i> module	252
10.7	Comparison of the region containing <i>msrE-mphe</i> in pS30-1 and Tn1548/Tn6180	253
10.8	Comparison of the sequence at the right end of the <i>ser</i> module	256
10.9	Comparison of pM131-5 and pS30-1	257
10.1	Alternating C and D modules in pS30-1	258
10.11	The <i>oxa58</i> gene is located in a <i>dif</i> module in pMAD	259
10.12	Runs of C and D modules	261
10.13	Structure of pRCH51-1 and pBAL_204-1	263

10.14	Alignment of ISAjo2 insertion sites in pS30-1 and EF102240	265
10.15	IR of ISAjo2-1	266
10.16	IR sequence of ISAb32 and its relationship to that of ISAjo2-1	268
10.17	Variation in the <i>oxa58</i> module	270
Chapter 11		
11.1	Timeline of when antibiotics suitable for Gram-negative pathogens were first used clinically	278
Appendix I		
A1.1	Unique positions of ISAb1	313
Appendix II		
A2.1	AbGRI3 in D86 and chromosomal inversion	319
A2.2	Assembly of pS29-1 contigs	321
A2.3	Assembly of pS27-2 contigs	323
A2.4	Assembly of pS34-2 contigs	325
A2.5	Contig 23 and 44 of pS32-2 and contig 24 and 43 of pS36-2	326

LIST OF TABLES

Chapter 1		
1.1	Antibiotic categories and antibiotics used to define resistance in <i>Acinetobacter</i> sp.	5
1.2	Aminoglycoside resistance genes, their typical phenotypes and context	11
Chapter 2		
2.1	Buffers and solutions	45
2.2	Growth media	46
2.3	Antibiotics used in this study and the final concentrations	46
2.4	GC2 isolates from Singapore used for experimental work in this study	47
2.5	Features of recipient strains	49
Chapter 3		
3.1	GC2 isolates from Singapore General Hospital	62
3.2	Antibiotics tested by disk diffusion	64
3.3	Aminoglycoside resistance profiles	64
3.4	Colistin MIC	65
3.5	Resistance genes detected in addition to <i>oxa23</i> , <i>tetA(B)</i> and <i>strA-strB</i>	67
3.6	Phenotypes conferred by aminoglycoside resistance genes	68
3.7	Two groups based on the <i>armA</i> and <i>aphA6</i> resistance genes	68
3.8	Resistance genes detected using ResFinder	72
3.9	MLST profiles using the Institut Pasteur scheme	74
3.10	MLST profiles using the Oxford scheme	75
3.11	Connection between the <i>gpi</i> allele and capsule locus replacement	75
3.12	Characteristics of group 1 isolates containing <i>aphA6</i>	77
3.13	Characteristics of group 2 containing <i>armA</i>	78
Chapter 4		
4.1	Groups based on screening results	86
4.2	Contigs of S21, S27 and S34 matching RI _{MDR-TJ}	87
4.3	Contigs of S24 and S29 matching AbGRI1-3	89
4.4	Contigs of S38 and S39 matching AbGRI1-1	92
4.5	Mapping AbGRI1 for the isolates without <i>sul2</i>	94
4.6	Mapping AbGRI1 for isolates with <i>sul2</i>	95
4.7	AbGRI1 variants detected	99
4.8	Location of <i>oxa23</i> containing transposons	105
Chapter 5		
5.1	Contigs of AbGRI2 in S24, S21 and S29	111
5.2	Contigs of AbGRI2 in S34	112
5.3	Contigs of AbGRI2 in S38/S39	114
5.4	AbGRI2-1 screening PCRs	117
5.5	Results of screening PCRs used to assemble additional AbGRI2 variants	118
5.6	Two groups within the collection	124
Chapter 6		
6.1	S27 contigs with IS26 fragments	134
6.2	S38 and S39 contigs with IS26 fragments	138
6.3	Mapping PCRs for isolates without <i>aphA1b</i>	142
6.4	PCR screening of AbGRI3 in isolates containing <i>armA</i> and <i>aphA1b</i>	143
6.5	AbGRI3 contigs of S25	145
6.6	Completed GC2 genomes containing AbGRI3 variants as of July 2017	149
6.7	Features isolates containing AbGRI3	157
Chapter 7		
7.1	ISAbal positions in the chromosome of 20 isolates	161
Chapter 8		
8.1	Properties of the isolates containing <i>repAci6</i> plasmids	178
8.2	Contigs matching RepAci6 plasmids in S24, S26, S29 and S30	181
8.3	AbaR4 location in the type 1 plasmids	184
8.4	Open reading frames and predicted products encoded by the 9.3 kb transposon	185
8.5	Sequences containing the 9.3 kb transposon	187
8.6	Type 1 plasmid backbone comparisons	188
8.7	Type 1 plasmid conjugations	189
8.8	Type 2 RepAci6 plasmids	191
8.9	Type 2 AbaR4 locations	191

8.10	Type 2 backbone identity	192
8.11	Conjugation of type 2 plasmids	194
8.12	S21 contigs matching RepAci6 plasmid	195
8.13	pS21-2 insertions	197
8.14	Plasmids with repeat 4 and the 2.2kb of sequence	200
8.15	Open reading frames and predicted products encoded by Tn6177	202
8.16	Relationship of the transposition genes of Tn6177 to Tn6019 and Tn7	203
8.17	RepAci6 plasmid contig of S30	203
8.18	pS30-2 conjugation	206
8.19	Type 1 plasmids in GenBank as of July 2017	209
8.20	Type 2 plasmids in GenBank as of July 2017	210
8.21	Other RepAci6 plasmids in GenBank as of July 2017	211
Chapter 9		
9.1	RepAci1 and RepAci2 plasmids	215
9.2	Contigs matching pAb-G7-1.	219
9.3	Transformation of AB307_0294 Rf ^R with pS32-1 and pS21-1	223
9.4	<i>In silico</i> SmaI digest	224
9.5	Cell counts over 5 serial passages (110 generations)	225
9.6	Mobilisation gene and protein lengths	226
9.7	Mob proteins related to those of pS21-1	227
9.8	Construction of A25 (pS21-1 and pS27-2)	227
9.9	A25 (pS21-1, pS27-2) and AB307_2094 (Nx ^R) mating	228
9.10	Transconjugant screening	229
9.11	Construction of 17978 ^{Rif} -A (pS32-1, pAb-G7-2)	231
9.12	Frequency of pS32-1 mobilisation	232
Chapter 10		
10.1	Contigs of pS30-1	240
10.2	Open reading frames and predicted products encoded by pS30-1	242
10.3	Comparison of OrfQ of pS30-1 to the product of the orf following the <i>rep</i> gene in selected plasmids	243
10.4	Transformation of pS30-1	245
10.5	pS30-1 mobilisation experiments	247
10.6	Sequences containing the <i>tet39 dif</i> module	249
10.7	Comparison of <i>pdif</i> sites surrounding the <i>tet39</i> determinant	251
10.8	Plasmids containing the <i>msrE</i> and <i>mphE dif</i> module	251
10.9	<i>pdif</i> sites surrounding the <i>msrE-mphE</i> determinant	253
10.10	The <i>higB</i> and <i>higA</i> toxin/antitoxin <i>dif</i> module	254
10.11	The <i>ser dif</i> module	255
10.12	<i>pdif</i> sites 6 and 7 bound a <i>dif</i> module	257
10.13	Relatives of ISAjo2 in GenBank	267
10.14	Relatives of ISAb32 in GenBank.	267
10.15	<i>dif</i> modules in <i>Acinetobacter</i> plasmids	269
Chapter 11		
11.1	Features of the 20 GC2 isolates from Singapore General Hospital	275
11.2	Location of <i>oxa23</i> containing transposons	277
Appendix I		
A1.1	Completed <i>A. baumannii</i> genomes as of July 2017	305
A1.2	Results of CDS disk diffusion for GC2 isolates from Singapore General Hospital	308
A1.3	Isolates sequenced for this study by HiSeq	312
A1.4	Isolates sequences for this study by MiSeq	312
A1.5	Positions of ISAb1 in XH386	314
Appendix II		
A2.1	Contigs matching RepAci6 plasmids in S26 and S29	321
A2.2	Contigs matching RepAci6 plasmids in S27 and S34	325
A2.3	Contig/s matching RepAci6 backbone and assembly PCRs	327

Appendix III

A3.1	Primer pairs used for the detection of resistance genes	328
A3.2	Primer pairs used to detect Tn6022, AbaR4 and AbaR4Δ	329
A3.3	Primer pairs used for mapping AbGRI1-1	330
A3.4	Primer pairs used for mapping AbGRI1-3	330
A3.5	Primer pairs used for mapping AbGRI1-4	331
A3.6	Primer pairs used for mapping AbGRI1-5	331
A3.7	Primer pairs used for mapping AbGRI1-6	332
A3.8	Primer pairs used for mapping AbGRI1-7	332
A3.9	Primer pairs used for mapping AbGRI1-8	333
A3.10	Primer pairs used for mapping AbGRI2-1	334
A3.11	Primer pairs used for mapping AbGRI2-12	334
A3.12	Primer pairs used for mapping AbGRI2-13	335
A3.13	Primer pairs used for mapping AbGRI2-14	335
A3.14	Primer pairs used for mapping those isolates with only IS26 in the AbGRI2 region	335
A3.15	Primer pairs used for AbGRI3 assembly	336
A3.16	Primer pairs used assembly of RepAci6 plasmids	338
A3.17	Small plasmid assembly	340

CHAPTER

1

General Introduction

1.1 The genus *Acinetobacter*

The genus *Acinetobacter* and its species have had a complicated history of taxonomic change. The first description of bacteria from this genus was from 1911 by a Dutch microbiologist Beijerinck, who isolated an organism from soil using an enriched calcium-acetate medium (15). He called the organism *Micrococcus calcoaceticus* (15). Over the next 40 or so years similar bacteria have been assigned many different names and the sub-committee on the taxonomy of Moraxella and Allied Bacteria officially acknowledged the *Acinetobacter* genus in 1971 (19, 122).

Acinetobacter is located within the family of *Moraxellaceae* in the order of *Gammaproteobacteria* (19). The genus of *Acinetobacter* comprises gram negative, aerobic, catalase positive, oxidase negative, non-fermenting, non-fastidious bacteria that have a GC content of 39-47% (19, 122, 207). They are typically rod-shaped during exponential phase but are more coccoid in shape at stationary phase (122). On solid media they usually form smooth, sometimes mucoid, gray-white colonies (207). Species of *Acinetobacter* have been recovered from virtually all soil and water samples, can be members of the normal flora and some have roles as human pathogens (9, 13, 258, 271).

Today there are 49 validly named species of *Acinetobacter* (43). Phenotypic methods of speciation were unreliable and even genotypic methods used in the 1980s such as DNA-DNA hybridisation separated 12 genospecies of *Acinetobacter* that were themselves composed of several species (30). One group, called the *A. calcoaceticus* – *A. baumannii* complex (Acb) (77, 78), included *A. baumannii*, *A. calcoaceticus*, *A. pittii* (genomic species 3) and *A. nosocomialis* (genomic species 13 TU) (185). The grouping of clinically relevant *Acinetobacter* species (*A. baumannii*, *A. pittii* and *A. nosocomialis*) with the environmental

species *A. calcoaceticus* is undesirable. In addition, accurate identification of the three clinically relevant Acb species is important for epidemiology and for effective infection control. Many methods have been used in clinical laboratories to distinguish these species including amplification of species-specific regions including the *gyrB* gene (114, 117), or for the specific *A. baumannii oxaAb* carbapenemase gene (263). More recently, MALDI-TOF MS (matrix assisted laser desorption ionisation – time of flight mass spectrometry) instruments have been used to effectively distinguish these species (257).

1.2 *A. baumannii*

A. baumannii causes the majority of clinical infections caused by *Acinetobacter* species and is an important nosocomial pathogen that causes opportunistic infections. These infections commonly occur in patients who are immunocompromised or in intensive care units (ICUs). Other predisposing factors that increase the likelihood of *A. baumannii* infection include old age, premature birth, the heavy use of antibiotics following surgery, indwelling catheters, wounds and length of hospital and/or ICU stay (75). *A. baumannii* is an important cause of hospital acquired pneumonia particularly in patients requiring mechanical ventilation (207). This bacterium is also commonly associated with skin infections, particularly burn wounds or surgical wounds and serious wound infections can lead to bacteraemia (9). *A. baumannii* can also be associated with secondary meningitis and is occasionally the cause of urinary tract infections. Rarely, community acquired *A. baumannii* infections have been reported and excessive alcoholism is a risk factor (8).

Unlike other species of *Acinetobacter*, *A. baumannii* is rarely isolated from non-clinical environmental sources and the natural reservoir of this bacterium is not known (258, 271). During an outbreak, *A. baumannii* can be recovered from hospital surfaces and the skin of

patients and hospital personnel (14). During an outbreak, *A. baumannii* can be recovered from hospital surfaces and the skin of patients and hospital personnel (14). Once these bacteria are present in the hospital environment they are very hard to eradicate due to their high level of resistance to desiccation and disinfection (258).

The success of *A. baumannii* as a nosocomial pathogen is likely due to its ability to persist in the hospital environment and its high level of antibiotic resistance, though it was this second characteristic that lead to specific studies of *A. baumannii* in the 1990s, particularly those that were carbapenem resistant (258). Current isolates frequently only remain susceptible to colistin (9) and ominously, isolates resistant to all antibiotics have been detected (84). The severity of antibiotic resistance in *A. baumannii*, has lead its inclusion in the ESKAPE group of bacteria (224). This group includes *Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa* and *Enterobacter* species, which are able to effectively ‘escape’ the effects of antibiotics (224). Increasingly, the therapeutic options for these pathogens are so limited that older previously discarded drugs that are associated with significant toxicity, such as colistin are being used (29). A 2017 report from the World Health Organisation, classified the threat of carbapenem-resistant *A. baumannii* as critical with the urgent need for the developments of new treatments.

The development of standardised terminology to classify and describe antibiotic resistance in *A. baumannii* in 2011 enabled accurate comparison of surveillance data world-wide (166). This classification system was based on the resistance to 9 useful categories of antibiotic for treatment of *Acinetobacter* infections listed in Table 1.1 (166). Multiple antibiotic resistance (MAR) was defined as being resistant to one or more antibiotics in at least 3 categories, extensively antibiotic resistance (XAR) means the bacteria were resistant to an antibiotic in 7

out of the 9 categories, and resistance to all antibiotics means that the bacteria are pan antibiotic resistant (PAR) (166).

Table 1.1 Antibiotic categories and antibiotics used to define resistance in *Acinetobacter* sp.

Antibiotic category	Antibiotics
Aminoglycosides	Gentamicin Tobramycin Amikacin Netilmicin
Carbapenems	Imipenem Meropenem Doripenem
Fluoroquinolones	Ciprofloxacin Levofloxacin
Antipseudomonal penicillins + β -lactamase inhibitors	Piperacillin-azobactam Ticarcillin-clavulanic acid
Extended-spectrum cephalosporins	Cefotaxime Ceftriaxon Ceftaxidime Cefepime
Folate pathway inhibitors	Trimethoprim-sulphamethoxazole
Penicillins + β -lactamase inhibitors	Ampicillin-sulbactam
Polymyxins	Colistin Polymyxin B
Tetracyclines	Tetracycline Doxycycline Minocycline

1.3 Antibiotic resistance in *A. baumannii*

Antibiotic resistance in *A. baumannii* can be due to both intrinsic and acquired resistance. Bacteria can acquire resistance to an antibiotic by enzymatically altering or degrading the antibiotic, modifying the target site, or via decreased antibiotic accumulation either through decreased uptake or increased efflux, overproduction of the target enzyme or acquiring an alternate pathway to those inhibited by an antimicrobial agent (170, 246, 276). Antibiotic

resistance can be acquired through chromosomal DNA mutations, which alter existing bacterial proteins or through the acquisition of new DNA carrying resistance genes from other bacteria (170). This transfer of genetic material between bacteria, known as horizontal gene transfer (HGT), increases the genetic flexibility of these organisms.

1.3.1 Intrinsic resistance

Based on early studies of *Acinetobacter* isolates (237, 266), it has been reported that *A. baumannii* are intrinsically resistant to aminopenicillin, 1st and 2nd generation cephalosporins and chloramphenicol (52). From these studies, it is difficult to determine if the genes conferring this resistance are intrinsic or acquired but it is likely that they are due to the chromosomally encoded AmpC cephalosporinase and CraA efflux pump (described below). A recent study has identified a novel aminoglycoside nucleotidyltransferase, intrinsic to most species of *Acinetobacter*, an *ant(3'')-II* or *aadA* gene, which confers resistance to spectinomycin and streptomycin (283).

1.3.2 Resistance to cephalosporins

A. baumannii usually contain a chromosomally-encoded cephalosporinase, called AmpC (27, 211). The *ampC* gene encoding this protein is constitutively expressed at low levels and does not confer resistance to 3rd generation cephalosporins including ceftazidime and cefotaxime (27, 165). However, increased expression of this gene, usually by the insertion of an ISAbal1 upstream providing a promoter stronger than the intrinsic one, results in high level resistance to these antibiotic (89, 165, 211). Despite being located upstream of many different *ampC* alleles, ISAbal1 is always found in the same position, 9 bp upstream of the start codon of the *ampC* gene (89). Insertion of ISAbal25 upstream of *ampC* has also been shown to increase expression and contribute resistance to 3rd generation cephalosporins (95, 165).

However, increasing the expression of their own native *ampC* gene is not the only way isolates of *A. baumannii* have become resistant to 3rd generation cephalosporins. There is evidence that some isolates have acquired an activated *ampC* region (with ISAba1 or ISAba125 already upstream) from another isolate by homologous recombination (91, 95). A composite transposon, named Tn6168, containing the *ampC* gene flanked by two directly orientated copies of ISAba1, has also been identified (92).

Some *A. baumannii* have also acquired extended-spectrum β -lactamases (ESBLs) including *bla*_{PER}, *bla*_{VEB}, *bla*_{CTX-M} and rarely *bla*_{SHV} or *bla*_{TEM} type genes (211).

1.3.3 Carbapenem resistance

Carbapenem antibiotics have historically been the first line of treatment of multiply antibiotic resistant infections. Resistance to carbapenems can be conferred by several β -lactamases but the most common are Class D β -lactamases (oxacillinases) (215). In general, these enzymes hydrolyse imipenem faster than meropenem (215).

A. baumannii isolates encode an intrinsic class D β -lactamase, often referred to as OXA-51 (31, 61) that is specific to the species of *A. baumannii*. This gene has been used to identify the organism (263) although there are many variants of this gene. The gene encoding all the different alleles of OXA-51 will be referred to as *oxaAb*. Like the intrinsic *ampC* gene, the *oxaAb* gene does not confer clinical levels of resistance to carbapenems, however upregulation of this gene either by ISAba125 or ISAba1, can result in reduced susceptibility to carbapenems (194, 211).

The most commonly acquired and hence most well studied of class the D β -lactamases in *A. baumannii* include the OXA-23, OXA-24 (and the identical OXA-40) and the OXA-58 β -lactamases (61, 215). Recently, two other Class D beta-lactamases have been identified in *A. baumannii*, OXA-143 (116) and OXA-235 (115).

The *oxa23* gene, and its numerous variants, have been detected worldwide and has been found as part of various transposon structures located both in the chromosome and in plasmids (61, 194, 215). OXA-23 was first identified in Scotland, 1985, in an imipenem resistant *A. baumannii*, though it was originally called ARI 1 (205). The *oxa23* gene has been detected in Tn2006, Tn2007, Tn2008, Tn2008B, Tn2009 (194, 215) (Fig 1.1). All of these transposons, except Tn2007 which has ISAb4 upstream, have ISAb1 upstream of *oxa23* (194). The *oxa23* gene has been mobilised from the chromosome of *Acinetobacter radioresistens*, in which it is intrinsic (212). In the *A. radioresistens* sequences to date, there is no IS upstream of the *oxa23* gene and these isolates are susceptible to carbapenems, indicating that an overexpression of OXA-23 is required to confer resistance (194).

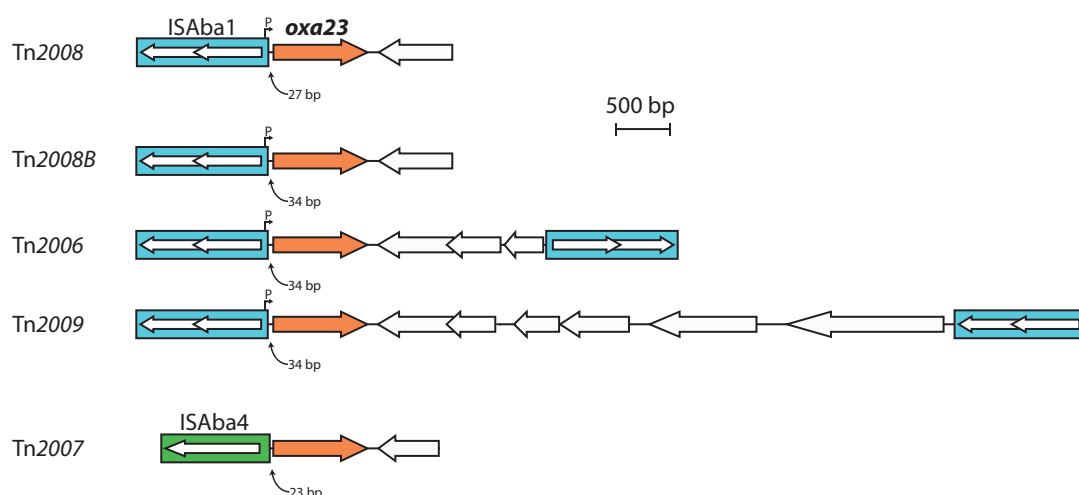


Figure 1.1 The five structures harbouring *oxa23* in *A. baumannii*. Arrows indicate orientations and extents of genes and open reading frames and the *oxa23* gene is in orange. The blue box is ISAb1 and the green box is ISAb4 and the internal arrows indicate the orientation of the transposase. The curved arrow below indicates the distance between the IS and the start codon of the *oxa23* gene. The arrow designated 'P' indicates the position and orientation of the strong promoter in ISAb1. Figure adapted from Nigro *et al.* (194).

Only Tn2006 and Tn2009 are true class I or composite transposons, with two copies of ISAbal flanking the internal segment containing *oxa23* (Fig 1.1). Tn2006 has been experimentally proven to transpose, and like ISAbal, produces a 9 bp target site duplication (TSD) (181). A characteristic TSD surrounding Tn2008, Tn2008B and Tn2009 in different locations indicates that these transposons are also mobile (180, 194). Tn2006, is often located within a larger transposon Tn6022 and this entire structure is named AbaR4 (3). AbaR4 itself is mobile and has been seen in both the chromosome and on plasmids of *A. baumannii* (3, 88, 99, 140, 232, 238).

The OXA-24 group was the second group of class D β -lactamases in *A. baumannii*, identified in 1997 in an outbreak in Spain (28). Most isolates containing OXA-24 are resistant to carbapenems although this is likely due a number of factors including efflux pumps (215). Genes encoding OXA-24 or related enzymes (OXA-40 and OXA-72) have been detected in plasmids from various species of *Acinetobacter* (46, 85, 172) and unlike *oxa23*, it is not usually associated with any IS. Instead, an internal segment encoding OXA-24 is found flanked by conserved inverted repeats of 28 bp that were found to be homologous to XerC-XerD binding sites (46). These sites contain the 11 bp binding sites for the XerC and XerD recombinase separated by a 6 bp spacer (46). The segment containing the *oxa24* gene and surrounded by XerC-XerD binding sites has also been detected in two locations in the RepAci1 plasmid pAB120 (220), as well as in plasmids of novel Rep types (85, 172). It has been suggested that this module is mobile and the XerC-XerD binding sites may have a role in its mobilisation (46, 172).

The first OXA-58 enzyme was identified in France in 2003 (214) and similar to OXA-24 enzymes it only hydrolyses carbapenems at low levels, so clinical levels of resistance in isolates

encoding these enzymes is usually multifaceted (61, 215). Variants of OXA-58 have since been identified including OXA-96, OXA-97 and OXA-164 (61, 215). The genes encoding OXA-58 variants have been detected in many different contexts. Most are often associated with the ISAb₃, though this is often interrupted by other insertion sequences such as ISAb₂, IS1008 and IS18 (61, 216). Genes encoding the OXA-58 enzymes have mostly been identified in plasmids (20, 64, 168, 216, 281).

Class B β -lactamases (metallo- β -lactamases) are powerful carbapenemases. Four types of these enzymes have been reported in *A. baumannii*, IMP-like, VIM-like, SIM-1 and NDM-type enzymes (211). The genes encoding the first three are found in gene cassettes in class 1 integrons (211) and NDM is associated with ISAb₁₂₅, which provides the gene with a strong promoter (58). Rarely, class A carbapenemases including KPC (226) and GES-type enzymes (26) have been detected.

1.3.4 Aminoglycoside resistance

A variety of aminoglycoside resistance genes have been detected in *Acinetobacter* and are listed in Table 1.2. Most of these genes confer resistance to a subset of aminoglycoside antibiotics through enzymatic modification and encode phosphotransferases, acetyltransferases and adenylyltransferases (239). However, the *armA* gene is a 16S rRNA methyltransferase that provides ribosomal protection through methylation of 16S rRNA, which impedes the binding of aminoglycosides to the 30S subunit (54, 72). This gene confers resistance to all clinically relevant aminoglycosides (72).

Table 1.2 Aminoglycoside resistance genes, their typical phenotypes and context

Gene	Confers resistance to	Context
<i>aacC1</i>	Gentamicin	Gene cassette
<i>aacC2</i>	Gentamicin, netilmicin, tobramycin	
<i>aadA1</i>	Streptomycin, spectinomycin	Gene cassette
<i>aphA1b</i>	Kanamycin, neomycin	Tn6020, Tn4352
<i>aphA6</i>	Amikacin, kanamycin, neomycin	TnaphA6
<i>aacA4a</i> ¹	Amikacin, netilmicin, tobramycin	Gene cassette
<i>aacA4b</i> ¹	Gentamicin, netilmicin, tobramycin	Gene cassette
<i>aadB</i>	Gentamicin, kanamycin, tobramycin	Gene cassette in secondary site
<i>aac(6')-Im</i>	Amikacin, netilmicin, tobramycin	Gene cassette
<i>armA</i>	Amikacin, gentamicin, netilmicin, tobramycin, kanamycin	Tn1548-like

¹The substrate specificity of the *aacA4* changes based on a single base change that results in an amino acid change.

Many of the aminoglycoside resistance determinants listed in Table 1.2 are located in gene cassettes in class 1 integrons and are often located within specific cassette arrays (211). The *aacC1* and *aadA1* genes are often found together in the cassette array of *aacC1*-orfP-orfQ-*aadA1* or *aacC1*-orfP-orfQ-*aadA1* in a class 1 integron, with the sulphonamide resistance gene *sul1* in the 3'-conserved segment (3'-CS) (Fig 1.2) (3, 183, 188, 217, 219, 262, 265). Previously, the orfP and orfQ gene cassettes have been referred to as orfX and orfX' (217, 219, 265).

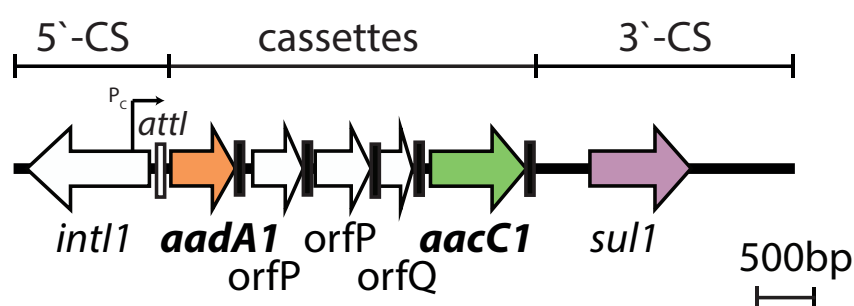


Figure 1.2 Typical cassette array found in class 1 integrons in *A. baumannii*. Arrows indicate extent and orientation of genes and open reading frames. The *aadA1* gene is orange, *aacC1* is green and the *sul1* gene is mauve. The sequence of the 5'-CS, gene cassettes and 3'-CS are indicated by the bars above. The *attI* site is shown by the open box and the black bars indicate the *attC* site of each gene cassette. The location and orientation of the promoter is indicated by *P_c*.

Isolates containing either of these two cassette arrays commonly also contain the *aphA1b* gene (127, 183, 265), which was later shown to be located in Tn6020 (188, 217), a 3.07 kb compound transposon flanked by IS26s.

Another frequently detected cassette array found in a class 1 integron includes the *aacA4-catB8-aadA1* array (232, 262), that in addition to conferring resistance aminoglycosides also contains the chloramphenicol resistance gene *catB8*. This array is often found associated with the *armA* gene within a Tn1548-like transposon that is flanked by IS26s (136, 285, 286).

Rarely, other cassette arrays containing *aacA4-orfO-bla_{OXA-20}* (129) or *aac(6')-Im-aadA1* genes (105) have been detected in other class 1 integrons. The *aadB* gene cassette is inserted in a secondary site in the small plasmid pRAY (236). The original sequence of pRAY contained a number of errors and a correct sequence pRAY* was released (100). A number of different variants eg. pRAY*-v1 and pRAY*-v2 have been detected (100).

The *aphA6* gene (147, 148) in *Acinetobacter* is most commonly located within TnaphA6, a composite transposon flanked by copies of ISAbal25 (Fig 1.3) (197). It is thought that the *aphA6* gene originated in *A. guillouiae*, an amikacin susceptible environmental species and has been mobilised by ISAbal25, resulting in over-expression of the *aphA6* gene and high level resistance (279). This transposon is frequently found on plasmids (90, 97, 148, 192, 277).

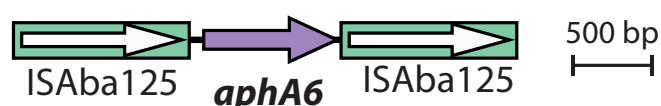


Figure 1.3 Structure of TnaphA6. The purple arrow indicates the extent and orientation of the *aphA6* gene. The green boxes represent ISAbal25 and the internal arrow indicates the orientation of the transposase

1.3.5 Quinolones and fluoroquinolones

Resistance to quinolones and fluoroquinolones can occur through mutations in the genes encoding the bacterial gyrase (*gyrA* and *gyrB*) and topoisomerase IV (*parA* and *parC*), that lower the binding affinity of these antibiotics to the enzyme-DNA complex (211). The most commonly reported mutations that have been experimentally shown to confer resistance to fluoroquinolones such as ciprofloxacin are those that result in a Ser-81-Leu substitution in GyrA and a Ser-84-Leu substitution in ParC (267, 268).

1.3.6 Tetracycline resistance

Tetracycline resistance in *A. baumannii* is commonly due to the presence of the TetA and TetB resistance determinants which are efflux pumps that specifically remove tetracycline from the cell (169, 211). The TetB determinant additionally confers resistance to minocycline (225). Another efflux pump, the Tet39 determinant, has been detected in *A. baumannii* isolates and like TetA, only confers resistance to tetracycline (5, 98). The TetM determinant has also been seen and confers resistance through ribosomal protection (223). Resistance to tigecycline, a modified tetracycline known as a glycycline, is often associated with overexpression of AdeABC and AdeIJK efflux pumps (described below) (209).

1.3.7 Sulphonamide and trimethoprim resistance

Sulphonamide resistance is frequently detected in *A. baumannii* (207, 211). Class 1 integrons which harbour the *sulI* resistance gene in the 3'-CS are frequently detected (36, 86, 120, 128, 193, 219, 252). Sulphonamide resistance in *Acinetobacter* can also be conferred by the *sul2* gene (191, 198), which is usually adjacent to a full or partial CR2, in a fragment derived from *GI**sul2* (189). Resistance to trimethoprim can be associated with *dhfr* genes located in gene

cassettes in integrons (86, 259) or overexpression of intrinsic efflux pumps (described below) (44).

1.3.8 Colistin resistance

The main mechanism of polymyxin resistance involves modification of lipid A (LPS) which reduces that net negative charge of the outer membrane (211). Point mutations and frameshift in the *pmrA* and *pmrB* genes, that form a two-component system involved in regulating the modification of LPS, have been shown to cause an increase in colistin MICs (4, 221). There is also evidence that resistance to colistin may also be attributed by complete loss of LPS by mutations in genes involved in lipid A biosynthesis (*lpxA*, *lpxC* and *lpxD*) (245, 270).

1.3.9 Rifampicin resistance

Resistance to rifampicin is most frequently due to mutations leading to amino acid changes in the active site of bacterial RNA polymerase and a number of different mutations have been associated with increase in MICs (80). Rarely, the acquired *arr-2* resistance gene which enzymatically inactivates rifampicin, has been detected in a cassette in class 1 integrons (121, 154).

1.3.10 Chloramphenicol resistance

An intrinsic efflux pump of the major facilitator superfamily, CraA (‘chloramphenicol resistance *Acinetobacter*’), was identified in *A. baumannii* (227). This efflux pump has high affinity to chloramphenicol and has been shown to confer high levels of resistance (227). Despite intrinsic resistance, chloramphenicol resistance genes encoding efflux pumps such as *cmlA* genes or those encoding a chloramphenicol-modifying enzyme such as *cat* genes, have been detected in *A. baumannii* (44, 217, 286).

1.3.11 Macrolide resistance

Although *A. baumannii* are intrinsically resistant to macrolides, acquired resistance genes conferring resistance to this class of antibiotics have been detected. These include the *msrE* and *mphE* resistance genes (285, 286).

1.3.12 Additional efflux pumps involved in antibiotic resistance

Efflux pumps from a number of different families including three efflux pumps from the resistance-nodulation-cell division (RND) family, AdeABC, AdeIJK and AdeFGH, as well as AbeM, from the multidrug and toxic compound extrusion family (MATE) and AbeS, a member of the major facilitator superfamily (MFS) have been detected in *A. baumannii* (44). Over expression of AdeABC efflux pump contributes to resistance to a broad-range of antibiotics including aminoglycosides, β -lactams, fluoroquinolones, erythromycin, chloramphenicol, tetracycline and tigecycline (278). Overexpression of AdeABC has been shown to provide a strong synergistic effect with aminoglycoside and carbapenem resistance genes (44). Upregulation of the AdeFGH pump also contributes to resistance to fluoroquinolones, chloramphenicol, trimethoprim and sulphonamides (44). The AdeIJK pump is constitutively expressed (47) and contributes to intrinsic resistance against β -lactams, fluoroquinolones, tetracyclines, tigecycline, erythromycin, chloramphenicol, trimethoprim, sulphonamides and rifampicin. AbeS has been shown to contribute to resistance against chloramphenicol, fluoroquinolones and erythromycin (247) and in addition the substrate range of AbeM also includes kanamycin and trimethoprim (250).

1.4 *A. baumannii* clones

Early analysis of different outbreaks of *A. baumannii* isolates were performed by a variety of molecular methods including ribotyping, pulsed-field gel electrophoresis (PFGE), randomly

amplified polymorphic DNA (RAPD) and amplified fragment length polymorphism (AFLP) (271). These older methods successfully differentiated isolates but they have limitations as they can be quite cumbersome, time consuming and expensive to perform. Studies employing these methods recognised that isolates from unrelated outbreaks shared similar characteristics (52). Typing of isolates from Northern Europe found that most of the outbreak and MAR isolates were members of two clones, I and II (182, 184). By the early 2000's, studies had detected isolates belonging to clone I and clone II across numerous countries across Europe including Spain, Italy, Greece, Portugal, Turkey and South Africa (182). Hence, these two clones became known as European clones (EC) I and II. Evidence of a third clone, EC III was also identified by ribotyping, though isolates belonging to this group were less common (182).

In 2007, a PCR typing scheme was developed that distinguished isolates into 3 groups which corresponded to the three clonal complexes (261). This scheme was based on the sequences of the *ompA*, *csuE* and *oxaAb* genes and sequence group (SG) 1 corresponded to EC II and SG2 corresponded to EC I (261). EC III corresponded to SG3. Other new and simple PCR and sequence based methods facilitated the analysis of larger, more diverse strain collections from around the world. Publications from 2009-2010 showed that EC I and II could be detected in hospitals on different continents (51, 70, 113, 180, 208, 217, 219), and henceforth these clones will be called global clones (GC) 1 and 2.

Two multi-locus sequence typing (MLST) schemes have been developed for *A. baumannii* (12, 51) and use the sequence of seven house-keeping genes, some of which are shared (indicated on Fig 1.4). In the Institut Pasteur (IP) scheme ST1 correlates with GC1 isolates and ST2 with GC2 isolates, with closely related isolates being a single or double locus variants (SLV, DLV) (51). Profiles of GC1 and GC2 isolates can be quite variable using the Oxford Scheme (104)

and this is due to the high levels of variation in the *gyrB* allele and in the *gpi* allele which is located in the capsule locus which has been shown to frequently switch (119, 137). The *A. baumannii* MLST database (<http://pubmlst.org/abaumannii>) currently contains 2,227 unique sequences (Aug, 2017).

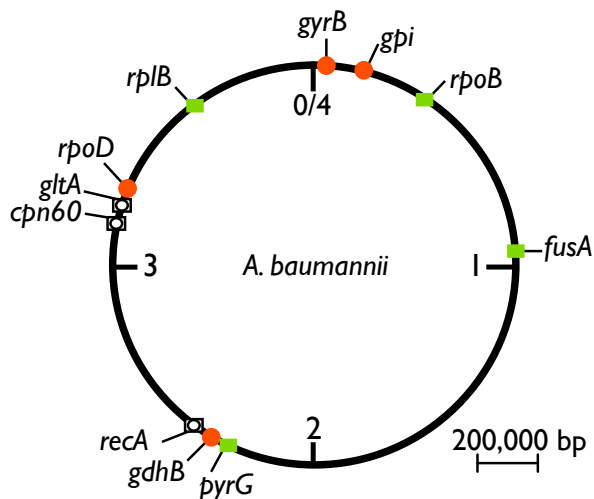


Fig 1.4 Schematic of the *A. baumannii* chromosome showing the location of genes used in the Pasteur and the Oxford MLST schemes. White circles in a black box show genes used in both schemes while the genes indicated by the green squares are used in the Pasteur Scheme and the red dots indicate genes used in the Oxford Scheme.

Most carbapenem resistant isolates are members of the global clones and the majority of these are GC2 (113, 131). The carbapenemase gene most frequently associated with resistance in GC2 isolates is the *oxa23* gene. Other problematic STs that are not members of the global clones have been reported including ST25 and ST78 (50), and ST85 (25).

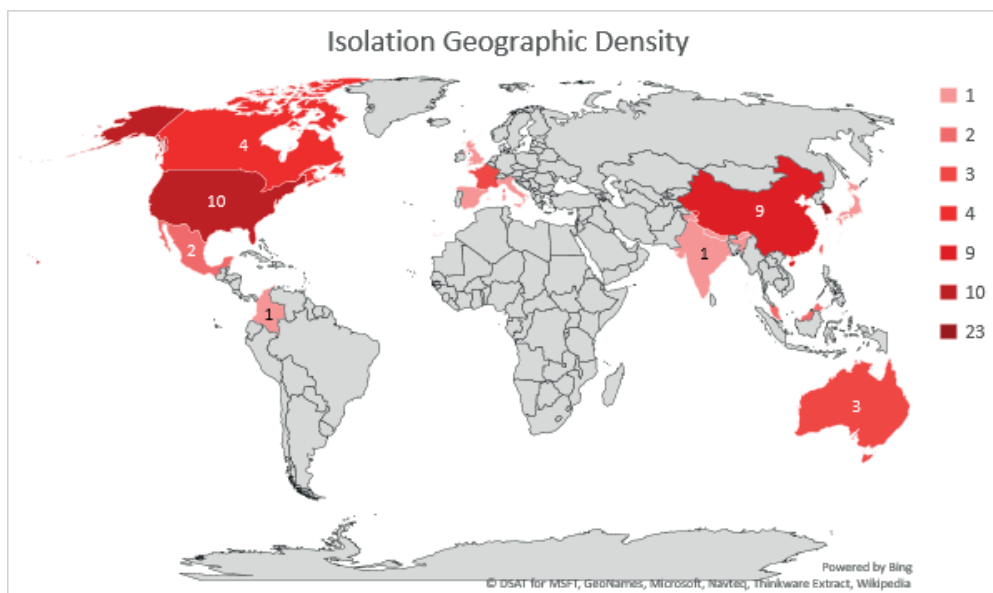
1.5 *A. baumannii* genomics

Whole genome sequence data has enabled in-depth comparison of bacterial isolates. The first *A. baumannii* genome was sequenced in 2006 (67), and at the advent of this project there were 17 completed genomes, 9 of which were GC2. However, as of July 2017 there are now 71

A.



B.



C.

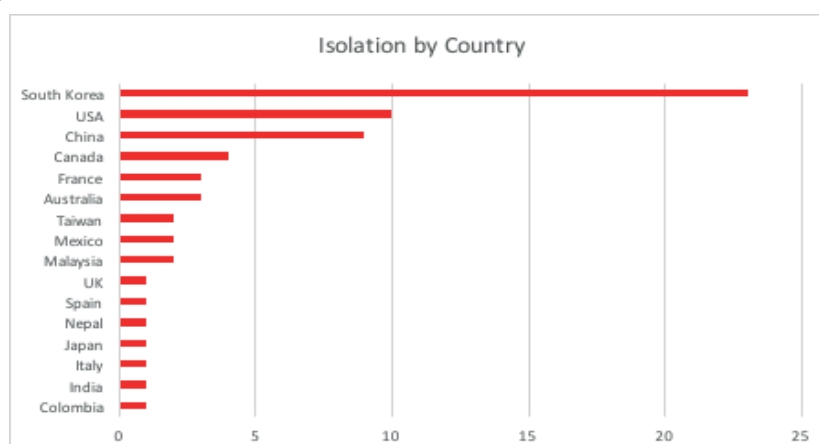


Figure 1.5 Completed *A. baumannii* genomes in GenBank. A. Year of submission of genome. B. Geographic density of isolation. C. Isolation by country.

completed genomes in GenBank (Fig 1.5A) and 39 of them are GC2 isolates (Appendix I; Table A1.1).

There are also over 2,000 draft *A. baumannii* genomes (<https://www.ncbi.nlm.nih.gov/genome/403>). At first these genomes were sequenced using less accurate and cumbersome sequencing technologies including 454 and Sanger whole shot-gun sequencing, but as sequencing technology improved, Illumina HiSeq and MiSeq platforms were used. In general, the HiSeq and MiSeq genomes were more accurate than the earlier genomes, but many may have suffered from assembly errors. From 2013 onwards, genomes sequenced by PacBio, a third-generation technology, began to be released and are the most accurate due to the long reads produced. A description of the differences between these sequencing technologies is provided in Section 1.11.

The Fig 1.5B and C show the location of isolation of the completed genomes and *A. baumannii* from 16 countries have been sequenced. Most countries have below 5 sequenced genomes (Fig 1.5C). China and USA have 9 and 10 genomes respectively, but South Korea currently has the highest count of 23 (Fig 1.5C), and all but one of them released prior to 2016 (Appendix I; Table A1.1).

The genome of *A. baumannii* isolates ranges from 3.7 Mb to 4.3 Mb. Whole genome phylogenies reveal separate clusters for GC1 and GC2 isolates and are consistent with MLST data (130, 162, 228, 254). Attempts to define the core genome of *A. baumannii* varied based on number and identity of strains analysed, but overall evidence suggests that it lies between 1,455 and 2,688 genes (3, 63, 130). This is around 60% of the average strain genome content and mostly encodes proteins involved in metabolic and cellular processes (9). However, the

pan genome, including both the core and accessory genome for *A. baumannii* is impressively large (over 8,800 genes) (9) and will likely continue to grow as more sequences become available. This accessory genome includes prophage regions, siderophore clusters, heavy metal resistance clusters and resistance islands and other mobile genetic elements (3, 130). Variation in gene content within chromosomal islands carrying antibiotic resistance genes has been a large focus of the analysis of the accessory genome of *A. baumannii*.

1.6 Chromosomal resistance islands

Both GC1 and GC2 isolates contain their own characteristic resistance islands located in the chromosome that contain clusters of antibiotic resistance genes. These regions are composed of fragments of transposons, IS, integrons and other mobile elements such as common region (CR) elements that either contain or are associated with specific antibiotic resistance genes.

1.6.1 GC1 AbaR islands

Most antibiotic resistant GC1 isolates contain a resistance island, named AbaR (*A. baumannii* resistance island) inserted in a chromosomal ATPase, later renamed *comM*, and surrounded by the 5 bp TSD of 5'-ACCGC-3' (Fig 1.6) (3, 67, 217). The first multi-antibiotic resistant GC1 isolate analysed was AYE, from France in 2001, and it was found to contain an 86 kb island in *comM*, designated AbaR1 (67). AbaR1 contains genes conferring resistance to various antibiotics and heavy metals and these genes were often in fragments of transposons (67). All of the antibiotic resistance genes in AbaR1 are located in a central region designated MARR (multiple antibiotic region) that is flanked by two copies of Tn6018 (previously named ISPu12-like elements) (217) and forms a transposon (Fig 1.6). This transposon is inserted in the *uspA* gene of Tn6019 and it is Tn6019 that is inserted in the *comM* gene and forms the backbone of this island (217). In addition, a region containing arsenate/arsenite resistance genes is located

between *tniE* and the *uspA* gene of Tn6019. The transposon Tn6019 has imperfect 26 bp IRs and the first 5 genes likely encode its transposition functions and have been called *tniA*, *tniC*, *tniB*, *tniD* and *tniE* (88). The gene named *sup* in Tn6019 (217) encodes a sulphate permease which had previously been incorrectly annotated as *sulI*.

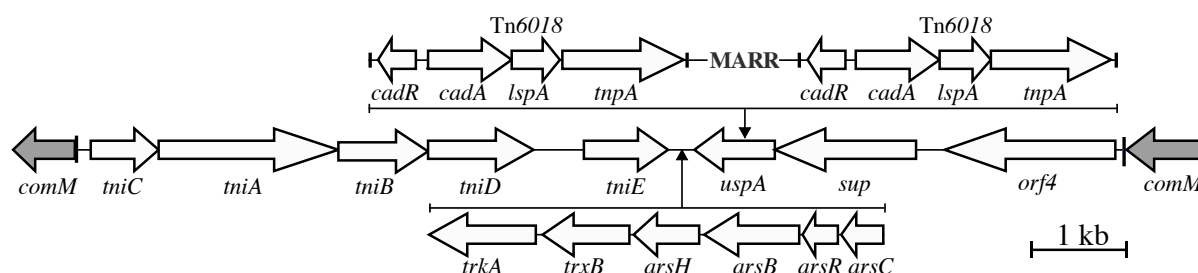


Figure 1.6 Structure of AbaR. The arrows indicate the extents and orientations of genes. The grey arrows are the fragments of the chromosomal *comM* gene. IRs are shown by vertical lines and the vertical arrow above indicates the insertion point of the Tn6018 flanked MARR, while the vertical arrow below so the position of genes involved in arsenic resistance.

Many variants of AbaR islands have since been discovered, but AbaR1 is still the largest (3, 145, 146, 198, 217, 219), and it appears that relative to other AbaRs, AbaR1 has acquired an additional region of antibiotic resistance genes. A comparison of AbaR1, AbaR3 (AB0057, isolated in America) (3) and AbaR5 (characterised in the Australian isolate 3208) (217) showed that a portion of the MARR region present in AbaR3 was missing in the other two islands. Hence, it was concluded that AbaR3 was ancestral (217). However, recent analysis has revealed that AbaR3 cannot be the ancestral to AbaR1 and AbaR5, as AbaR3 contains a 108 bp deletion in the 5'-CS of the integron, and this sequence was present in the other islands (103). The ancestor of AbaR3, AbaR1 and AbaR5 is AbaR0. This structure, characterised in an Australian isolate (WM98), does not contain the 108 bp deletion but in all other ways is identical to AbaR3 (103).

The MARR region of AbaR0, shown in Fig 1.7 is made up of fragments of known transposons which are indicated above. These include the mercury resistance module of Tn1696, the *tet(A)* resistance determinant in a fragment of Tn1721, the *catA1* gene in a partial Tn2670, *bla*_{TEM} within a partial Tn1 and a complete Tn6020 containing the *aphA1b* gene (Fig 1.7). A class 1 integron containing *aacC1*-orfP-orfP-orfQ-*aadA1* gene cassettes with *sul1* in the 3'-CS is also present (Fig 1.7) (67, 217). An additional fragment of the 3'-CS is present containing a second copy of *sul1*. Within the MARR of AbaR0, there are three copies of IS26 and much of the variation reported in other AbaR structures is due to IS26 mediated deletions.

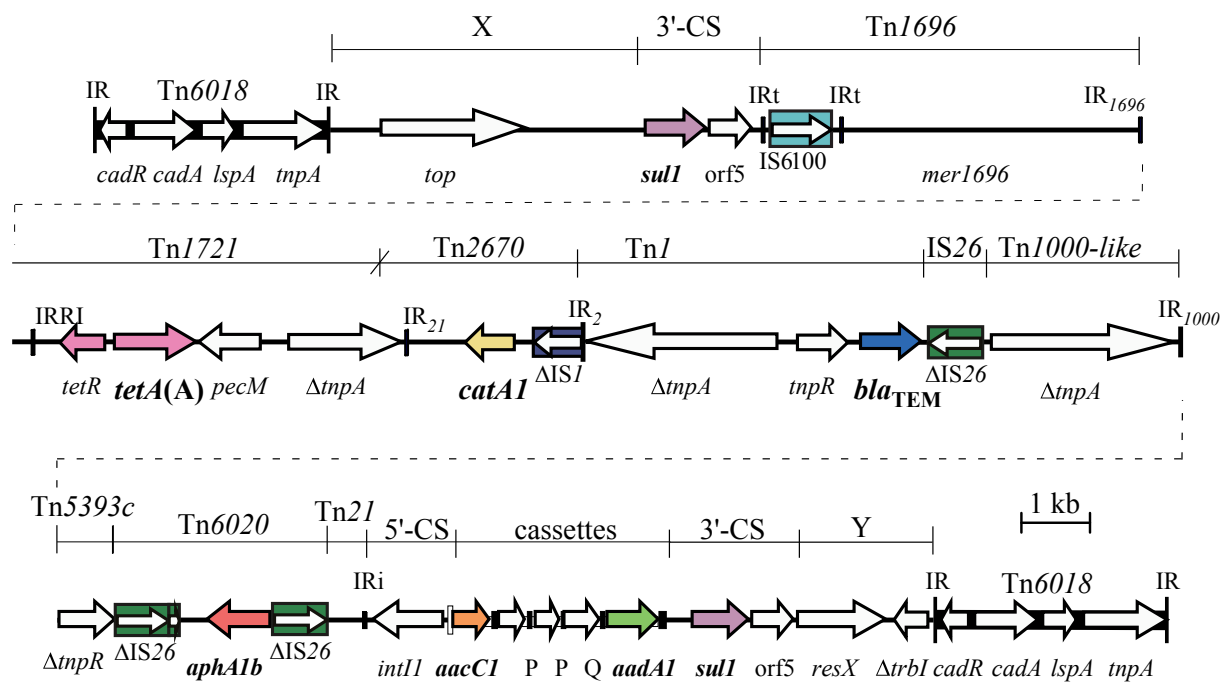


Figure 1.7 Structure of the Tn6018 flanked MARR of AbaR0. Arrows indicate the extents and orientations of genes and open reading frames. The *sul1* gene is purple, the *tet(A)* determinant is in pink, the *catA1* gene is yellow, the *bla*_{TEM} genes is blue, the *aphA1b* gene is red-orange and *aacC1* gene and *aadA1* gene are orange and green respectively. Coloured boxes indicate insertion sequences and the internal arrow indicates the transposase gene. Vertical bars indicate IRs and the open box indicates the *attI* site. The lines above represent the regions derived from known transposons with vertical lines indicate the boundaries between segments and X showing a recombination crossover.

1.6.2 AbaR4, Tn6022 and Tn6021

In a GC1 isolate another transposon has been detected in the *comM* gene and surrounded by the same 5 bp direct repeat as AbaR islands (88). This island is called AbaR4, consists of the

backbone transposon Tn6022 with Tn2006 containing the *oxa23* gene inserted in the *sup* gene (3, 88). AbaR4 was first identified in a secondary location in the GC1 isolate AB0057, which had AbaR3 in *comM* (3). Tn6022 is related to Tn6019 but is more closely related to another transposon, Tn6021 (88) that was identified in the *comM* gene of ATCC 17978 (244). Hence, Tn6019, Tn6021 and Tn6022 are part of the same family of transposons that target the *comM* gene.

1.6.3 Chromosomal resistance islands in GC2

1.6.3.1 Resistance islands in *comM*

GC2 isolates also contain a resistance island inserted in the *comM* gene and surrounded by the same 5 bp TSD as AbaR islands. The first completed GC2 genome was of an Italian isolate ACICU (129). This isolate had part of an AbaR island and that was called AbaR2. AbaR2 is lacking the left end of most AbaR islands and contains Tn6020 and a class 1 integron containing the *aacA4-orfO-bla_{OXA-20}* cassette array, followed by Tn6018 and the right end of Tn6019 (129). ACICU is the only sequenced GC2 that contains an AbaR type resistance island in *comM* and hence appears to be of a separate lineage to all of the other sequenced GC2 isolates.

The majority of GC2 isolates contain a different island in this position. Early analysis of the island in *comM* of GC2 isolates differentiated it from the AbaR islands of GC1 isolates as Tn6022 rather than Tn6019 could be detected at the left end and the *uspA* gene was intact as no copies of Tn6018 were present (217). Furthermore, the resistance content of the island differed from that found in AbaR islands. Most GC2 isolates contain the *tet(B)* determinant, the *strB-strA* genes (190, 238, 260, 285) and some also include the *sul2* gene (125, 191, 193).

Many of these structures have been assigned AbaR numbers or called AbaR4-like islands in the literature, though most do not resemble to AbaR islands.

Broadly, there appear to be two main types of these islands in GC2 isolates. The first type, named Tn6166 and later AbGRI1-2 (*A. baumannii* genomic island 1), has the structure shown in Fig 1.8 (191). This 37 kb island contains Tn6022Δ1 at the left end (has a 2.85 kb deletion in respect to Tn6022), followed by an orf region, and a fragment of a transposon 86% identical to that of Tn6022 (191) and recently this has been named Tn6173 (94). Another portion of this transposon, containing orf4b bounds the right end of this island (94). These two fragments of Tn6173 are separated by a region derived from GIsul2 with a fragment of Tn10 containing the *tetA*(B) determinant within it, adjacent to a segment of Tn5393 containing the *strB* and *strA* genes (Fig 1.8) (191). Interestingly, a large portion of this island has recently been discovered in the conjugative plasmids pD4, pA293-10, pAB3 and the resistance island in these plasmids is likely the origin of AbGRI1 (94). In addition, AbGRI1-2 also contains ISAbal7 inserted in the orf between the delta *tniE* and *uspA* genes in Tn6022Δ1 and Tn2006 containing the *oxa23* gene in the orf region (Fig 1.8) (191). So far, this arrangement appears to be unique to Australian isolates from the East Coast (193).

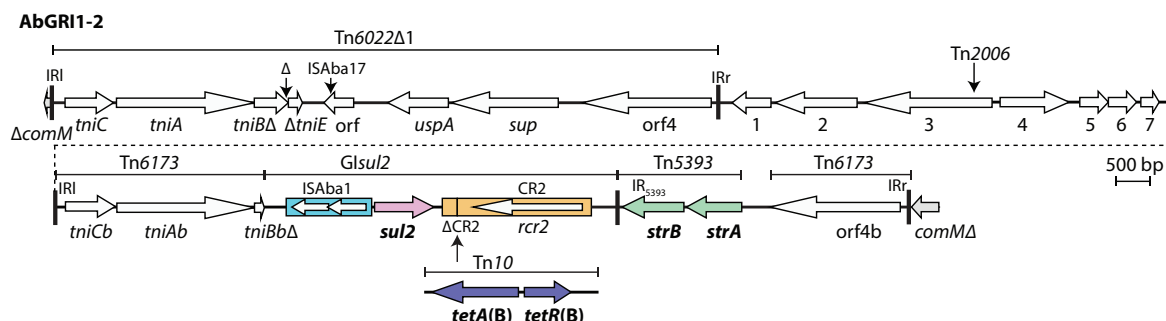


Figure 1.8 Structure of AbGRI1-2. Arrows indicate extents and orientations of genes and open reading frames. The *sul2* gene is pink, the *tetA*(B) and *tetR*(B) genes are dark blue and the *strB* and *strA* genes are light green. The light blue box represents ISAbal1, the yellow boxes indicate CR2 elements and the arrows inside represent the orientation of the transposase and rolling circle replicase, respectively. Vertical bars indicate IRs. Bars above indicate the origin of segments and vertical arrows indicate the deletion (Δ) and location of insertions. Figure made to scale from GenBank accession number JN96883.

The second group of these islands appear to have derived from a structure that has acquired a second copy of Tn6022, which is inserted in the last 6 bp of the *tetA(B)* gene. The 41.6 kb resistance island in *comM* of MDR-TJ, called RI_{MDR-TJ}, is an example of this (Fig 1.9A) (125). The stop codon for the *tetA(B)* gene now lies 18 bp within the IRr of Tn6022 and these additional bases do not appear to alter the function of the protein. A closely-related version of this island, called AbaR24, was assembled from isolates from Latvia and had AbaR4 instead of Tn6022 in this position (Fig 1.9A) (232).

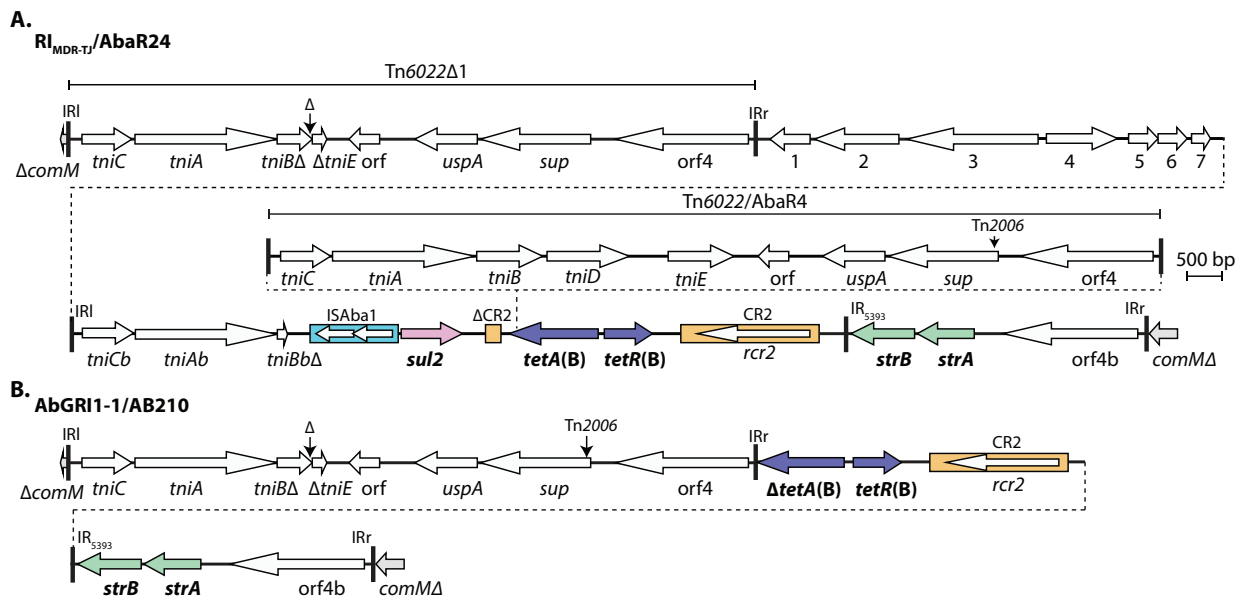


Figure 1.9 AbGRI1 variants. Arrows indicate extents and orientations of genes and open reading frames. The *sul2* gene is pink, the *tetA(B)* and *tetR(B)* genes are dark blue and the *strB* and *strA* genes are light green. The light blue box represents ISAbal, the yellow boxes indicate CR2 elements and the arrows inside represent the orientation of the transposase and rolling circle replicase, respectively. Vertical bars indicate IRs. Bars above indicate the origin of segments and vertical arrows indicate the deletion (Δ) and location of insertions. Figure made to scale from GenBank accession number CP003500 for RI_{MDR-TJ} and from GenBank accession number JN247441 for AbGRI1-1.

The variant detected in A320, the type strain for GC2 isolates and the oldest known isolate from 1982 was named AbGRI1-1 (previously referred to as Tn6166) (190) (Fig 1.9B). This variant was only 17.6 kb and has likely derived from an island related to RI_{MDR-TJ} as the *tetA(B)* gene in AbGRI1-1 is also missing 6 bp (190). It is likely that recombination between Tn6022Δ1 and Tn6022 of RI_{MDR-TJ} has removed the *orf* region, the fragment of Tn6173 and the *sul2* gene

to form AbGRI1-1. Variants of AbGRI1-1 containing Tn2006 in exactly same the position in the *sup* gene as in AbaR4, was detected in the isolate AB210 from the UK (260), as well as in isolates from Lithuania where it was called AbaR4b (238) (Fig1.9B). Another variant detected in another Lithuanian isolate is virtually identical to AbGRI1-1 but instead of Tn6022Δ1, it contains a full Tn6022 at the left end of the island. It is likely that a slightly different recombination event between Tn6022Δ1 and Tn6022 from RI_{MDR-TJ} has resulted in the evolution of this structure.

Another variant detected in 1656-2 (202), an isolate from South Korea, has evolved from RI_{MDR-TJ}, through the acquisition of an additional of a region containing the *bla_{PER}* gene between *strA* and *orf4b*, as well as the loss of the full Tn6022 and the *tetA(B)* determinant, likely through recombination between the partial and full CR2. The structure present in the Chinese isolates MDR-ZJ06 (285), named AbaR22, has a full copy of Tn6022 the left end of the island, followed by the *orf* region and the Tn6173 fragment. However, the *tniBb* gene in ΔTn6173 has recombined with the *tniB* gene of the second copy of Tn6022, forming a hybrid transposon, and the *sul2* region has been lost. The right end of the island is the same as RI_{MDR-TJ}. This implies that variants of this island with two full copies of Tn6022 may also be circulating.

1.6.3.2 AbGRI2 resistance island

This resistance island was only discovered and characterised in 2013 and is located in another region of the chromosome (Fig 1.10B) (188), however it was noted that a segment of the surrounding chromosome was missing. Unlike the islands in the *comM* gene, this resistance island was not located in a backbone transposon but was bounded by two copies of IS26 (Fig 1.10A).

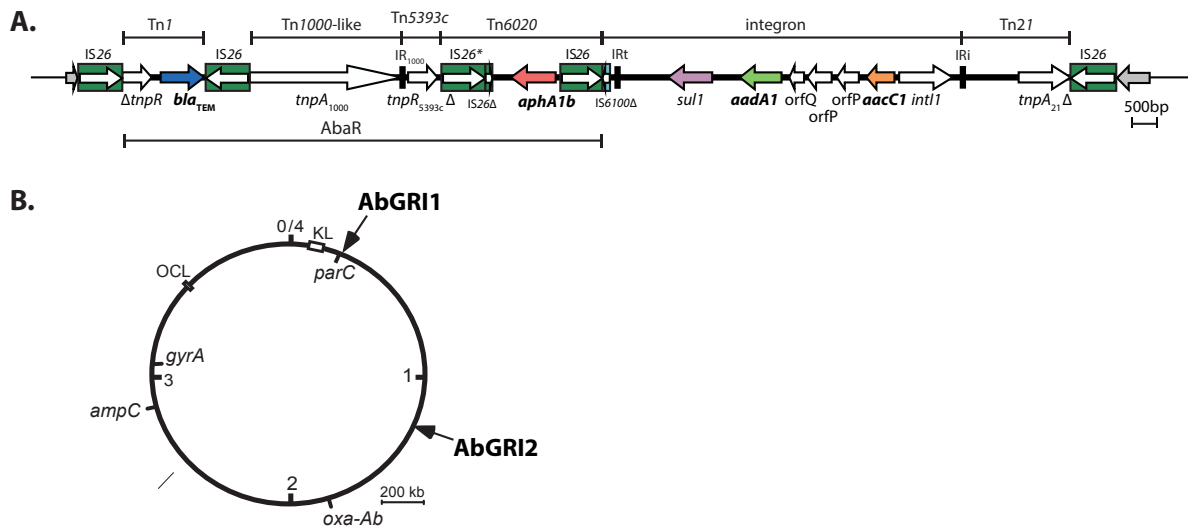


Figure 1.10 AbGRI2. A. Structure of AbGRI2-1. B. Schematic of the GC2 chromosome indicating the location of AbGRI1 and AbGRI2. In A. the arrows indicate extents and orientations of genes and open reading frames. The *bla_{TEM}* gene is blue, the *aphA1b* gene is red-orange, the *sul1* gene is purple and the *aadA1* and *aacC1* gene are green and orange, respectively. The IS are represented by coloured boxes and the internal arrow indicates the orientation of the transposase gene. *IS26*, *IS1* and *IS6100* are shown by the green, dark blue and light blue boxes respectively. Vertical bars indicate IRs and the bars above indicate the origin of segments. The dashed lines and the black triangle indicate sequence missing in AbGRI2-1. Figure made to scale from GenBank accession number JX869489 for AbGRI2-1. In B. the arrows indicate the location of AbGRI1 and AbGRI2.

The first characterised version of this resistance island was AbGRI2-1 and assembled from the draft genome of two Australian isolates, A91 (2004) and WM99c (1999) (188). AbGRI2-1 was 19.5 kb and was composed of fragments of known transposons containing resistance genes as well as three internal copies of *IS26* (Fig 1.10A). This island contains the *aphA1b* gene in Tn6020 (217), a fragment of Tn1 containing the *bla_{TEM}* gene and a class 1 integron containing the cassette array *aacC1*-*orfP*-*orfP*-*orfQ*-*aadA1* and the *sul1* gene in the 3'-CS (Fig 1.10A) (188). It was noted that a segment of AbGRI2-1 matched the MARR of AbaR islands. Two other variants of AbGRI2 were described in this work (188), AbGRI2-2 from the complete genome of MDR-TJ and AbGRI2-3 found in the complete genome of MDR-ZJ06. Both of these variants could have arisen from AbGRI2-1 via *IS26* mediated deletions in addition to cassette switching for AbGRI2-3.

Recently, the structure of the AbGRI2 region in the GC2 type strain A320, isolated in 1982, was published (24). The AbGRI2 region in A320 had been separated into two separate islands, AbGRI2-0a and AbGRI2-0b by an IS26 mediated inversion (24). Reversal of this inversion *in silico* revealed a 32.4 kb resistance island AbGRI2-0 (Fig 1.11) that was inserted in the gene corresponding the ABA1_01225 in A1 (GenBank accession number CP010781) and surrounded by an 8 bp TSD (5'-CGCCAACT-3') (24). This is the original insertion site of AbGRI2 and alludes to the important role of IS26 in its acquisition of this island into the GC2 chromosome.

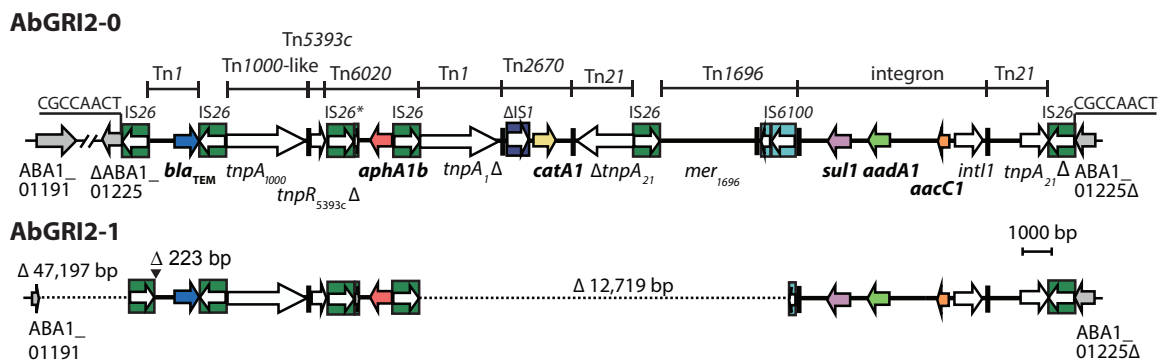


Figure 1.11 AbGRI2-0 is ancestral to AbGRI2-1. Arrows indicate extents and orientations of genes and open reading frames. The bla_{TEM} gene is blue, the aphA1b gene is red-orange, the catA1 gene is yellow, the sul1 gene is purple and the aadA1 and aacC1 gene are green and orange, respectively. The IS are represented by coloured boxes and the internal arrow indicates the orientation of transposase. IS26, IS1 and IS6100 are shown by the green, dark blue and light blue boxes, respectively. Vertical bars indicate IRs and the bars above indicate the origin of segments. The dashed lines and the black triangle indicate sequence missing in AbGRI2-1. Figure made adapted from Blackwell *et al.* (24).

In addition to the resistance genes found in AbGRI2-1, AbGRI2-0 also contains the catA1 gene located in a fragment from Tn2670 and a mercury resistance module from Tn1696 (Fig 1.11) (24). AbGRI2-0 is ancestral to AbGRI2-1, and AbGRI2-1 can be derived from AbGRI2-0 via a number of events. An IS26 mediated deletion has removed a 12.7 kb central fragment from the island, that includes the catA1 gene and the mer module. The IS26 at the left end of AbGRI2-1 has switched orientation and 223 bp of AbGRI2-0 next to it is missing. In addition,

compared to AbGRI2-0, AbGRI2-1 has 47.2 kb of chromosomal sequence to the left of the island missing.

The events detailed above that lead to the construction of AbGRI2-1 from AbGRI2-0 are quite specific and most of the reported structures from America and Asia have an island derived from AbGRI2-1 (24, 188, 193, 277). The island present in 1656-2 in Korea (202), has gained an addition IS26 bounded segment that includes the *aac6* gene cassette (24).

1.6.3.3 The origin of AbGRI2

Recently, a resistance island from the M1 plasmid R1215 was assembled (top line in Fig 1.12) (22). R1215 was isolated from *Serratia marcescens*, prior to 1980. The resistance region in R1215 is 37.8 kb in size and interrupts the *mucB* gene. This region is made up of the same fragments of transposons as both AbGRI2-0 and the centre of the MARR region of AbaR0 (22).

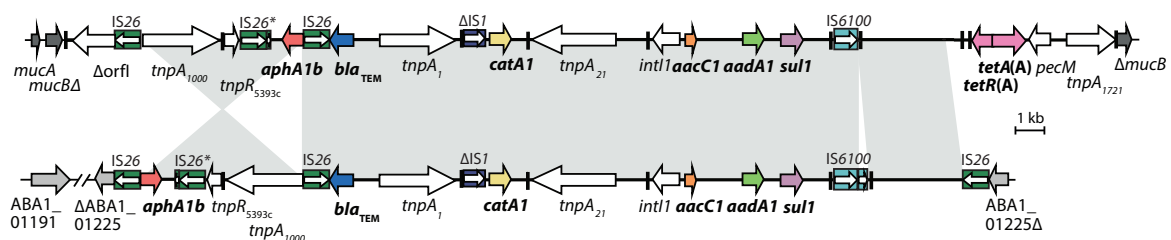


Figure 1.12 Relationship of the resistance island of R1215 to AbGRI2-0*. The R1215 resistance island is shown above and AbGRI2-0* is shown below. Arrows indicate extents and orientations of genes and open reading frames. The *bla*_{TEM} gene is blue, the *aphA1b* gene is red-orange, the *catA1* gene is yellow, the *sul1* gene is purple and the *aadA1* and *aacC1* gene are green and orange respectively. The IS are represented by coloured boxes and the internal arrow indicates the transposase gene. IS26, IS1 and IS6100 are shown by the green, dark blue and light blue boxes respectively. Vertical bars indicate IRs. Figure adapted from Blackwell *et al.* (22).

A comparison of AbGRI2-0*, which is a predicted precursor structure of AbGRI2-0 produced by the *in silico* reversal of two IS26 mediated inversions (24), displays the strong relationship between these two regions (Fig 1.12). The orientation of a segment including the *aphA1b* gene

is different in the two islands but as this segment is located between inversely-oriented copies of IS26, it could have inverted through homologous recombination (Fig 1.12). The R1215 resistance island is also lacking Δ IS6100 adjacent to IS6100 but this has been shown to be lost by recombination (32, 204). Hence, the resistance island in R1215 is closely related to the progenitor of AbGRI2-0* and this region was likely acquired into the GC2 chromosome through the action of IS26 (22). It appears that a single GC2 *A. baumannii* cell became multiply-antibiotic resistant in a single step, prior to 1982, and all modern isolates have evolved from this cell as they all contain a remnant of the AbGRI2 island (22, 194). Current evidence does not indicate whether 1 or 2 was first.

The centre of the MARR of AbaR0 can also be constructed from the resistance island in R1215, through its evolution is more complex. This involves inversions through homologous recombination, TU formation followed by TU integration by homologous recombination between a *sulI* gene already present in the AbaR precursor (22). Hence, GC1 isolates also became resistant in a single step to multiple antibiotics.

1.6.3.4 A third chromosomal resistance island

Recently, a third chromosomal island has been identified in some American isolates. This island was reported to be located between the reading frames corresponding to ACICU_02398 and ACICU_02399 in ACICU (GenBank accession number CP000863) (277). This resistance island contains the *armA* gene which encodes a 16S RNA methyltransferase which confers resistance to all clinically relevant aminoglycosides (72). The *armA* gene is located in an IS26 bounded transposon that is related to Tn1548 (73) and hence is a second IS26 bounded resistance island in these isolates. In addition to the *armA* gene, this region is also associated with macrolide resistance genes *msrE-mphE*, a class 1 integron containing the *aacA4*, *catB8*

and *aadA1* gene cassettes (conferring resistance to amikacin, tobramycin and netilmicin; chloramphenicol; streptomycin and spectinomycin) with the sulphonamide resistance gene *sul1* in the 5'-CS and sometimes also with *aphA1b* gene (resistance to kanamycin and neomycin) (277). The authors also noted that this resistance island was present in the completed genome of TYTH-1 from China, indicating that this region was not just specific to American GC2 isolates (277). The *armA* gene has been detected in other GC2 isolates, but the location of these gene was not determined (53, 136).

1.7 Tn6019, Tn6021 and Tn6022 family

The transposons Tn6019, Tn6021 and Tn6022 (and AbaR4) all appear to target the same position in the *comM* gene and are surrounded by a 5 bp TSD of 5' ACCGC 3' (88). These transposons are all bounded by approximately 25 bp inverted repeats (IR) that end with 5'-TGT-3' and a fragment of this IR is repeated multiple times at both ends of the transposon. They also share a conserved structure with 5 genes, *tniC*, *tniA*, *tniB*, *tniD* and *tniE*, that likely encode the transposition functions.

The size and outer sequence of the IR, the size of the TSD and targeted mode of transposition of these *A. baumannii* transposons is similar to Tn402, Tn5053 and Tn7 (45), and for ease these transposons will be called Class III transposons. In addition, TniA and TniB of the *A. baumannii* transposons show homology with those of Tn402 and Tn5053. Transposon Tn402 is both a transposon and a class 1 integron (176, 222). The transposition module of Tn402 encodes four proteins, TniA, TniB, TniQ and TniR. TniA of Tn402 is 25% identical to TnsB of Tn7, (222) and encodes a transposase. However, instead of encoding a restriction enzyme like TnsA, Tn402 instead contains a resolvase, TniR (222) and a *res* site.

The transposition module of Tn7 encodes five proteins called TnsA, TnsB, TnsC, TnsD and TnsE, which facilitate transfer of this element (45, 210). The essential TnsA and TnsB proteins resemble a restriction enzyme and a transposase, respectively (210). Transposition of Tn7 can occur in a targeted manner using TnsD, where Tn7 moves to its preferred single site in a bacterial chromosome (45), or in a non-specific manner, utilising TnsE, where the transposon preferentially moves to a random site in a conjugative plasmid (210).

1.8 IS elements

1.8.1 IS26

IS26 plays a key role in the evolution of resistance islands in *A. baumannii*. IS26 is a member of the IS6 family and is 820 bp in length with perfect 14 bp terminal IR, contains a single gene (*tnp26*) that encodes a 234 aa transposase and generates an 8 bp TSD upon insertion (177). IS26 has two modes of transposition, replicative and conservative mode. The replicative mode involves the formation of a cointegrate between the DNA molecule it resides in and the target molecule, duplicating an IS26 copy and generating a TSD (177). In the replicative mode, when the target of IS26 transposition is on the same DNA molecule as the donor IS, i.e. intramolecular transposition, either a deletion or inversion reaction occurs (10) depending on which strand is attacked by each end of the IS (Fig 1.13). Sequence deletion or inversion by IS26 has contributed to the evolution and rearrangement of many plasmids and chromosomes (22, 24, 109, 193, 218).

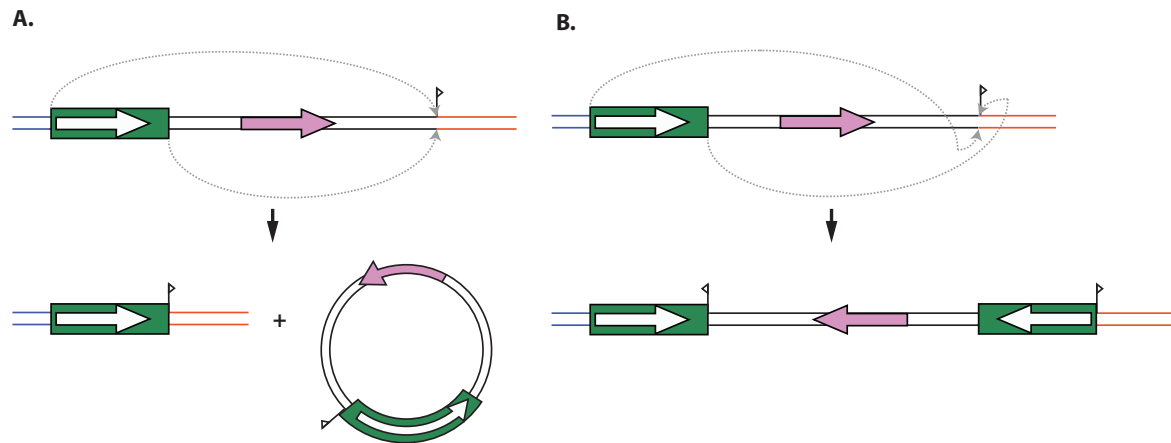


Figure 1.13 IS26 intramolecular transposition. Show the two outcomes of transposition of IS26 within the same DNA molecule, A. a deletion event, or B. an inversion event. These two outcomes are determined based on what strands are attacked. The pink arrow indicates a gene and different coloured lines are different segments of DNA sequence. The green box is IS26 and the internal arrow indicates the orientation of the transposase. The flags indicate the TSD.

The recent discovery of the conservative mode of IS26 transposition invokes a circular translocatable unit (TU) consisting of a single IS26 and a unique DNA segment (potentially carrying an antibiotic resistance gene), which then preferentially inserts itself adjacent to another IS26 in the target sequence so that both IS26s are in the same orientation, with no duplication of the IS26 or of target bases (107). The result of this transposition reaction has the same structure as a class I transposon (107) and accounts for the structures of tandem arrays of the same DNA fragment flanked by IS26s (20, 163, 171, 287) as well as ‘overlapping’ IS26 compound transposons with too few IS26s (32, 106, 217). Formation of TUs can occur via the replicative mode of IS26 transposition (deletion reaction) or through homologous recombination between two IS26s flanking a DNA sequence.

1.8.2 ISAbal

ISAbal was first identified in *A. baumannii* isolates (111) and is frequently detected in multiple copies in the chromosome. ISAbal is 1,180 bp in size and includes two overlapping open reading frames that encode the transposase. ISAbal is a member of the IS4 family (167) and

causes a 9 bp TSD at the site of transposition (111). This IS is able to affect the expression of neighbouring genes as it contains an outwardly facing promoter that includes both the -35 and -10 sequence (111). This IS is frequently detected upstream of the chromosomal *ampC* gene in clinical isolates, increasing its expression and contributing to third generation cephalosporin resistance (42, 89). ISAbal is frequently associated with the *oxa23* gene and is a component of the *oxa23*-containing transposons Tn2006, Tn2008 and Tn2009 (Fig 1.1) (194).

1.9 Plasmids

In general, plasmids are circular double-stranded DNA molecules that replicate independently of the host chromosome (16, 69). Most plasmids do not encode functions that are vital for the cell so can be lost and gained without any deleterious effect (69). Plasmids can exhibit broad host range, where they can transfer and survive in vastly different bacterial species, such as those found in the family of Enterobacteriaceae, or narrow host range where they are limited to a small number of similar bacterial species (16).

Some plasmids encode the functions required for transfer between cells and these plasmids are termed conjugative or self-transmissible. Conjugative plasmids have their own relaxase (and helper protein), origin of transfer (*oriT* site), type four coupling protein (T4CP) and type four secretions system (T4SS) conjugative apparatus so are generally required to be larger than 30 kb (69, 242). Some smaller plasmids are termed mobilisable (242), as they are able to be transferred by conjugation when a suitable helper conjugative plasmid is present in the same cell. Mobilisable plasmids can have their own *oriT* and relaxase but are recognised by the T4CP of a suitable conjugative plasmid, termed mobilisation *in cis* (242). However, in mobilisation *in trans*, the small plasmid has an *oriT* sequence that mimics the co-residing conjugative plasmid and relies on the conjugative plasmids relaxase, T4CP and T4SS (200).

1.9.1 *Acinetobacter* plasmids

The plasmids of *Acinetobacter* are specific to the genus and other plasmids found in Enterobacteriaceae are unable to replicate in *Acinetobacter* (21). *Acinetobacter* plasmids range from small cryptic plasmids, small plasmids containing one or two resistance genes, and large plasmids that may carry a number of resistance genes. Three different sets of conjugative transfer genes have been detected in *Acinetobacter* plasmids (87, 97, 126) and many of the small plasmids carry homologues of mobilisation genes (relaxase and helper proteins).

So far, *Acinetobacter* plasmids have been characterised based on their replication initiation protein. In 2010 Bertini et al. (21) analysed 18 *A. baumannii* plasmids available in GenBank and characterised 9 different replicase types which were designated RepAci1-9 as well as identifying a number of other Rep types that were not assigned Aci numbers (21). Some plasmids contained two different replication initiation proteins. A limitation of the replicon typing (21) was that, by definition, it did not include plasmids without a known replication initiation gene. The small plasmid pRAY (6 kb), containing the *aadB* gene (tobramycin, gentamicin and kanamycin resistance), contains no replication initiation protein but does encode mobilisation proteins.

Due to the decrease in the cost of genome sequencing over the last 7 years many more plasmid sequences are now available. As of July of 2017, there are now over 183 sequenced *A. baumannii* plasmids in GenBank and the analysis of Bertini *et al.* (21) likely under represents the diversity of the plasmids found in this species.

1.9.1.1 Bertini PCR-based plasmid replicon typing scheme

This scheme characterised 9 different replicase types but all of the Rep proteins belonged to one of two protein families, Pfam03090 or Pfam0105. The RepAci6 protein was the only representative of Pfam03090 and does not contain any iterons (21). The other replicase types detected were members of the Rep-3 superfamily, Pfam0105, and all of these plasmids included iteron sequences upstream of the *rep* gene (21). A number of these plasmids contain orthologues of the MobA or MobS-MobL mobilisation proteins indicating that they may be mobilised by a suitable, self-transmissible plasmid (21). The RepAci9 plasmid pMAC (9,540 bp), isolated from ATCC 19606, has been well characterised (56). The genes and regions performing replication functions, plasmid mobilisation as well as involved in organic peroxide resistance were shown experimentally (56).

1.9.1.2 RepAci1 and RepAci2 plasmids

The cryptic RepAci1 plasmid pAB0057 (3) is detected quite frequently in sequenced isolates and the pAb-G7-1 variant (97) is seen in GC1 isolates from Australia. Most Australian GC2 isolates contain a cryptic RepAci2 plasmid, represented by pD72-1 (192, 193). However, RepAci1 and RepAci2 plasmids carrying the *oxa24* resistance gene have also been reported (46, 179, 220).

1.9.1.3 RepAci6 plasmids

RepAci6 plasmids are common in sequenced isolates and *oxa23* containing transposons, Tn2008, Tn2006 and AbaR4, as well as *aphA6*, located in TnaphA6 have been detected in these plasmids (90, 97, 99, 192, 195, 277). These plasmids carry an MPF_F type transfer region and have been shown to conjugate (97). The mobilisation proteins encoded by RepAci6 plasmids are closest to those of IncW and hence have been named TrwB and TrwC (97). Within this

family of plasmids, two slightly different backbone types have been detected (192). The two backbones are distinguished based on SNPs throughout the backbone and differ by more than 1 % over the majority of the sequence but some small regions are up to 80 % different (192). The type 2 backbone also contains a small insertion. The type 1 backbone is represented by pAb-G7-2 (97) and the type 2 backbone by pA85-3 (99).

1.9.2 Conjugative plasmids not included in the Bertini PCR-based replicon typing scheme

A second group of conjugative plasmids were identified as they carried the NDM-1 carbapenemase resistance gene. These include pNDM-BJ01 (46 kb) (123) and pNDM-AB (47 kb) (284). These plasmids contain a novel Rep type and a VirB type 4 secretion system. Interestingly, both conjugated into and were maintained in *E. coli* (123, 284). This is the only type of *Acinetobacter* plasmid that has been shown to replicate outside the genus.

Recently, a group of related large plasmids, pAB3, pAB04-1 (272), pIOMTU433 (240), pD4-1 (93) and pA297-1 (87) were reported. Despite their large size (>130 kb), these plasmids do not contain any gene encoding a homologue of a protein known to be involved in replication initiation (87). However, they do contain a full set of genes encoding conjugative transfer functions and pAB3, pAB04-1 and pA297-1 have been shown to conjugate (87, 272).

1.10 Epidemiology of GC2 *A. baumannii* from Australia and South-east Asia

1.10.1 GC2 *A. baumannii* in Australia

The majority of the genomic analysis of GC2 *A. baumannii* in Australia has focused on isolates from the east coast of the country. A study analysing 61 isolates, from 6 different hospitals,

revealed that they form a homogeneous group based on sharing a number of unique characteristics (193). All but one was ST208 (*gltA-1*, *gyrB-3*, *gdhB-3*, *recA-2*, *cpn60-2*, *gpi-97* and *rpoD-3*) and one was ST425 (SLV of ST208 in the *gpi* allele). Every isolate contained the AbGRI1-2 structure including the characteristic ISAbal7 and Tn2006 location (Fig 1.8) (193). All of the isolates contained either AbGRI2-1 or an island structure that could be derived AbGRI2-1 through IS26 mediated deletions or inversions. None of these isolates contained the *armA* resistance gene. Based on the close homology between the isolates from the east coast of Australia it is likely that all of these isolates stem from a common ancestor (193). This suggests a single import event and that all of the current isolates have evolved from this ancestor creating the different sub-lineages detected.

1.10.2 GC2 *A. baumannii* in South-east Asia

A. baumannii is a major cause of hospital acquired infection in South-east Asia and was the dominant cause of hospital-acquired pneumonia or ventilator assisted pneumonia (VAP) in Thailand and Malaysia from a report in 2010 (38). A recent study indicates that carbapenem-resistant *A. baumannii* were the major cause of VAP in Vietnam, causing 46% of infections (187).

Isolates reported as CC92 or ST92 are predominant in South-East Asian countries including Malaysia, Singapore, Thailand and Vietnam (134). However, ST92 is actually ST208 and the discrepancy is due to the inclusion of the primer sequence in the *gpi* allele (101). ST195, a SLV of ST208, has been frequently detected in Malaysia and Thailand (139, 150). The *oxa23* gene, or one of its variants, was commonly associated with ST208 or one of its SLVs (134, 139, 143). The *armA* gene has been detected in isolates from Vietnam (253), and Malaysia (150).

Completed genomes of isolates from Malaysia have been released but little is currently known about the structures of resistance islands in isolates from Singapore, Thailand or Vietnam.

1.11 Sequencing technologies

Sequencing technologies have advanced rapidly over the last 50 years, increasing from the ability to sequence short oligonucleotides to millions of bases. Over this time there has been arguably 3 generations of sequencing technologies and the third generation has only become commercially available over the last 3 years.

Sanger sequencing was one of the first generation of sequencing technology. This method relies on synthesizing DNA on a single-stranded template while randomly incorporating chain terminators (230). The sequencing of overlapping fragments, which are then aligned into one contiguous sequence or contig in silico is referred to as shot-gun sequencing and was how the first bacterial genome, that of *Haemophilus influenza* was determined (65). This technology is still used today to sequence small DNA sequences such as a PCR amplicons but the maximum read length from Sanger sequencing is just under 1 kb (110).

1.11.1 Second generation sequencing technology

These technologies are grouped together based on the kind of template used for sequencing reactions (164). Platforms of the second generation depend on the production of libraries of clonally amplified templates and include 454 and Illumina (originally Solexa) platforms. Each of these technologies have the same work flow but may perform each step differently particularly in their chemistry of sequencing and the approach to used read the sequence (164). Firstly, libraries are produced from template DNA that has been fragmented, tagged, immobilized and amplified, followed by sequencing (48, 110, 164).

The Roche 454 system used the luminescent method for measuring pyrophosphate synthesis, also called pyrosequencing (199). High-end instruments of this technology produced long reads of 700-800 bp but with a relatively high error rate in terms of poly-bases longer than 6 bp (164). Illumina use Solexa chemistry, which includes reversible termination of sequencing products (17). HiSeq platforms produce large amounts of reads at an average length of 150 bp and the sequencing process can take up to a week (164). MiSeq sequencing on the other hand, produces fewer reads than HiSeq but with a larger read length of 300 bp (164). Despite the relatively short-read length of HiSeq and MiSeq, they have the advantage of producing paired-end (PE) data, in which the sequence at both ends of each DNA cluster is recorded (142, 164).

1.11.2 Assembly of second generation sequencing

Assembly of the short reads produced by second generation sequencing are often best assembled by programs involving De Bruijn graphs (40). These programs, such as Velvet and the newer SPAdes, work by splitting reads into even shorter units of a particular length k (k -mers) (40). The k -mer length that will best fit different read sets will vary (59). The programs builds a graph representing each k -mer prefixes and suffixes as nodes then draws edges that represent k -mers having a particular prefix and suffix (40). It then finds paths through the graph by alignment of overlapping k -mers and these paths form contigs (40, 282). Paired-end data is also incorporated correcting paths through the graph. Unlike Velvet, which uses a fixed k -mer value (282), SPAdes iterates over values of k -mer sizes and is able to incorporate paired-end data earlier in the assembly (11).

One of the major limitations of second generation sequencing is the length of the reads that are produced. No matter how good the assembly programs is, it is not able to resolve repeated

sequences larger than the read size, and as most repeats are longer, this results in breaks in the assembly where draft genomes of many contigs are produced. In addition, regions with repeated sequence, such as resistance islands, are often the focus of analysis, making their correct assembly critical for subsequent studies. Assembly of draft genomes is often laborious and expensive requiring manual confirmation of each junction via PCR and Sanger sequencing of the amplicons. Also, some repeat sequences are much larger than the lengths that can be resolved reliably by PCR.

1.11.3 Third generation sequencing technology

Pacific Biosciences (PacBio) single molecule real-time sequencing or SMRT sequencing is the most established of this generation. This method of sequencing has no amplification step and involves imaging the continuous incorporation of dye-labelled nucleotides during DNA synthesis (60). Zero-mode waveguide detectors (ZMW) obtain sequence information as phospholinked nucleotides are incorporated into the growing primer strand (48, 60). Improvements in the chemistry used in this technology now produces large reads (average of 10 kb) and although the individual read error is 10-15%, increased coverage allows for high cumulative accuracy rates (>99.99%) (152). Assembly of long read data is different to that of short read data. SMRT HGAP assembly program is the package released by Pacific Biosciences which first preassembles the longest reads as 'seed' reads, then corrects errors by tilling the small reads to these (37). These corrected 'seed' reads are then assembled together into a genome (37). Alternatively, the SMRT reads can be combined with shorter reads produced by second generation techniques in a hybrid assembly. The detection of DNA methylation data from SMRT sequencing has been well-established (66). Currently, the main limitation of SMRT is the cost relative to second generation technologies but this will likely continue to drop with time.

The idea for Nanopore sequencing was proposed 20 years ago but significant hurdles have had to be overcome. Useable devices, such as GridION or MiniION have only been produced recently and the technology is still evolving and improving (48). Nanopores use voltage variation produced when ssDNA is fed through the nanopore via an enzyme (233). The run time is short and as there is no amplification step the direct examination of DNA sequence modifications (such as methylation) can be examined, however this has only recently emerged as an application for Nanopore (48). The reads produced are super long but have a high error rate. However, they are useful to tile low error rate short-read data against and programs such as UniCycler (273) are designed for this.

1.12 Aims and scope

Extensively and pan antibiotic resistant *Acinetobacter baumannii* belonging to the global clone GC2 present a significant challenge in the treatment of infections worldwide. The GC2 populations from Australia and certain areas in USA have been studied in depth but the examination of populations from China are only beginning to emerge and little is known about those from Singapore, Korea or Taiwan. This study aimed to analyse the causes of antibiotic resistance in a selection of 20 extensively antibiotic resistant GC2 isolates from Singapore General Hospital to increase the understanding of *A. baumannii* from this region. These isolates were from a period of 15 years, from 1996 to 2011, and all were known to carry the carbapenem resistance gene *oxa23*. The in-depth understanding of antibiotic resistance in this group could then underpin a broader analysis of GC2 *A. baumannii* from Singapore and South-east Asia.

Specific aims were to:

- Determine the complete resistance profiles and complement of antibiotic resistance genes in a collection of 20 GC2 isolates
- Determine the contexts of all acquired antibiotic resistance mutations or genes

- Determine the structure(s) of the known chromosomal islands AbGRI1 and AbGRI2
- To determine if any isolates contain the more recently acquired third chromosomal resistance island AbGRI3 and characterise it and its variants in detail
- Examine the plasmids that contribute to antibiotic resistance and to define the structure of any regions containing antibiotic resistance genes
- Explore the relationships between the isolates using MLST and mapping of the positions of ISAbal.

CHAPTER

2

Materials and methods

2.1 Materials

2.1.1 Chemicals and reagents

The chemicals utilised throughout this project were of molecular biology grade and were purchased from Promega (Madison, WI, USA), Sigma-Aldrich (St Louis, MO, USA) and Bio-Rad (Hercules, CA, USA).

2.1.2 Buffers and other solutions

The buffers and solutions were made up with H₂O treated by reverse osmosis (Milli-Q H₂O), according to the protocols in Sambrook *et al.* (229) (Table 2.1).

Table 2.1 Buffers and solutions

Buffer/solution	Components
0.5X TBE buffer	45 mM Tris 45 mM boric acid 1mM EDTA
1X TAE buffer	40mM Tris 20mM acetic acid 1mM EDTA
TE buffer	10mM Tris-HCl, pH 7.5 1mM EDTA, pH 8
Loading dye	0.1% (w/v) bromophenol blue 0.2% (w/v) xylene 20 mg/mL ficoll
Saline	0.9% (w/v) NaCl
Elution Buffer (EB)	10 mM Tris-HCl, pH 8.5
Solution 1 (plasmid preparation)	50mM glucose 25mM Tris-HCl (pH 8.0) 10mM EDTA (pH 8.0)
Solution II (plasmid preparation)	0.2 N NaOH and 1% SDS
Solution III (plasmid preparation)	Potassium acetate (3M potassium 5M acetate)

2.1.3 Growth media

Media purchased from Becton Dickinson (Cockeysville, MD, USA) included Bacto™ Yeast Extract and Bacto™ Agar, while Tryptone and Sensitest Agar were acquired from Oxoid (Basingstocke, Hampshire, UK). Growth media was prepared according to

the instructions provided by the manufacturer and autoclaved at 121°C at 130 kPa for 20 minutes.

Table 2.2 Growth media

Media	Components
Luria-Bertani Broth (LB)	1% (w/v) tryptone 0.5% (w/v) yeast extract 1% NaCl
Luria-Bertani Agar (LA)	LB containing 1.5% (w/v) agar
Sensitest Agar	3.2% (w/v) Sensitest Agar in H ₂ O
Mueller Hinton Agar (MHA)	3.8% (w/v) Mueller Hinton Agar in H ₂ O
Mueller Hinton Broth (cation-adjusted)	2.2% (w/v) Mueller Hinton Broth (cation-adjusted) in H ₂ O

When necessary, antibiotics were added after the media had cooled to 60°C. The final concentration of antibiotics in the media are provided in Table 2.3. Most antibiotics were added to LB or LA, however when sulphamethoxazole or erythromycin were supplemented, MHA or MHB was used.

Table 2.3 Antibiotics used in this study and the final concentrations

Antibiotic or inhibitor	Abbreviation	Diluent	Stock conc. (mg/mL)	Final conc. (µg/mL)
Tetracycline ¹	Tc	50% ethanol	10	4
Kanamycin	Km	MQ H ₂ O	50	20
Ticarcillin	TIC	MQ H ₂ O	100	100 ²
Clavulanic acid	CLAV	MQ H ₂ O	25	25
Sulphamethoxazole	Su	MQ H ₂ O, NaOH	25	100
Rifampycin ¹	Rif	100% methanol	10	100
Nalidixic acid	Nx	0.5M NaOH	25	25
Erythromycin	Ery	100% ethanol	40	var ³

¹Tetracycline and rifampicin were stored in a lightproof container to prevent degradation.

²Was reduced to 40 µg/mL when selecting for pBAL_204-1.

³Concentration tested varied from 2 µg/mL to up to 128 µg/mL.

2.1.4 Enzymes

Taq polymerase and all restriction endonucleases with the appropriate buffers were obtained from New England Biolabs (Ipswich, USA). High fidelity Phusion™

polymerase was acquired from Finnzymer (Vantaa, Finland). Proteinase K was obtained from Promega (Madison, MI, USA). RNase A and lysozyme were purchased from Sigma-Aldrich (St Louis, MO, USA).

2.2 Bacterial isolates

Experimental work was conducted on twenty GC2 *Acinetobacter baumannii* isolates (listed in Table 2.4), isolated between 1996 and 2011 from Singapore General Hospital that were kindly supplied by Dr Li Yang Hsu (National University Health System, Singapore).

Table 2.4 GC2 isolates from Singapore used for experimental work in this study

Isolate ¹	Name ²	Source ³	Year
SGH9601 ¹	S21	Unknown	1996
SGH0402	S22	ETTA	2004
SGH0403	S23	ETTA	2004
SGH0601 ¹	S24	Trach	2006
SGH0602	S25	ETTA	2006
SGH0604	S26	ETTA	2006
SGH0701 ¹	S27	Blood	2007
SGH0702	S28	Blood	2007
SGH0703 ¹	S29	Blood	2007
SGH0823	S30	ETTA	2008
SGH0825	S31	Trach	2008
SGH0905	S32	Blood	2009
SGH0907	S33	Blood	2009
SGH0908 ¹	S34	Blood	2009
SGH1010	S35	Sputum	2010
SGH1011	S36	Blood	2010
SGH1012	S37	ETTA	2010
SGH1111 ¹	S38	Sputum	2011
SGH1112 ¹	S39	Sputum	2011
SGH1113	S40	Blood	2011

¹Draft genome sequence was available at the start of the project.

²For ease, this is the name by which the isolate will be referred to from here on.

³Abbreviations include tracheal aspiration (Trach) and Endotracheal aspiration (ETTA).

Seven isolates, S21, S24, S27, S29, S34, S38, S39, had draft sequences available at upon the commencement of this project. Genomic DNA for these isolates had been

sequenced using Illumina HiSeq at the Wellcome Trust Sanger Institute. Dr Kathryn Holt (University of Melbourne) assembled the paired-end reads of 100 bp into contigs using Velvet (282).

Additional isolates that all had draft genome sequence available were examined throughout the course of this study. This included D86, isolated from Royal North Shore Hospital in 2006 and was kindly provided by Peter Huntington. RCH51-1 (102) was isolated from Royal Children's Hospital (no isolation date) in Brisbane was kindly supplied by Dr Mohammad Katouli. The isolate BAL_204 (235), from Vietnam was isolated in 2010 and was kindly provided by Dr Steven Baker. Draft genome sequence were available for each of these isolates.

Features of strains used for conjugation and transformation recipients are listed in Table 2.5 and include ATCC 19606 (234) which was isolated prior to 1948, ATCC 17978 (3) which was isolated in 1951 and AB307_0294 (3) which was isolated in 1994 in Buffalo, USA. A25, which was isolated in 1987 in the United Kingdom, was kindly provided by Dr Kevin Towner, was also used in the project.

Table 2.5 Features of recipient strains

Strain	Resistance	Plasmids
ATCC 17978 ¹	Su, Ap,	pAB1 (CP000522), pAB2 (CP000523), pAB3 (Su ^R) (CP012005)
ATCC 17978 ^{Rif}	Su, Ap, Rif	pAB1, pAB2, pAB3 (Su ^R)
ATCC 17978-A	Ap	pAB1, pAB2
ATCC 17978 ^{Rif} -A	Ap, Rif	pAB1, pAB2
ATCC 19606 ^{Rif}	Ap, Sm, Su, Rif	pMAC (AY541809) ²
AB307_0294	Ap, Sp, CTX, CAZ, CRO, Nx	-
AB307_0294 ^{Rif}	Ap, Sp, CTX, CAZ, CRO, Nx, Rif	-
A25	Ap, Sm, Sp, Su, CTX, CAZ, Tc, Tp	-

Ampicillin, Ap; Ceftazidime, CAZ; Cefotaxime, CTX; Ceftriaxone, CRO; Nalidixic acid, Ns; Rifampicin, Rif; Spectinomycin, Sp; Specintomycin, Sm; Sulphamethoxazole, Su; Tetracycline, Tc; Trimethoprim, Tm.

¹GenBank accession for completed genome is CP012004

²Plasmid profile indicates that there may be another but no sequence has been confirmed.

Rifampicin resistant mutants of ATCC 17978 and ATCC 19606 were already in the Hall laboratory collection. In addition, a variant of ATCC 17978 that had spontaneously lost pAB3 (containing sulphonamide resistance) and was now sulphonamide susceptible is designated ATCC 17978-A (Table 2.5). A rifampicin resistance version of this strain, ATCC 17978^{Rif}-A had also been constructed previously.

2.3 Bacterial characterisation and manipulation

2.3.1 Bacterial growth

Isolates of *A. baumannii* were routinely cultured overnight in LB supplemented with appropriate antibiotics at the final concentrations listed in Table 2.3. To store isolates long-term, an equal volume of overnight culture (750 µL) was mixed with 50% glycerol in a Cryovial (Thermo Scientific, Rochester, USA) and stored at -80°C.

2.3.2 Antibiotic resistance phenotype determination using disk diffusion

The Calibrated Dichotomous Sensitivity (CDS) (<http://web.med.unsw.edu/cdstest>) disk diffusion assay was used to quantify the resistance of each isolate to a set of 30 antibiotics. Plates were made up with 20 mL Sensitest agar. Saline (4 mLs) was inoculated with single colonies from an isolate of *A. baumannii* and successively poured onto 5 plates. Once dry, each plate was stamped with 6 antibiotic disks, for a total of 30 different antibiotic disks tested. Following overnight incubation at 37°C, the annular radius of the zone clearing was measured from the edge of each disk. According to the CDS guidelines, resistance to an antibiotic is indicated by an annular radius of less than or equal to 6 mm. The antibiotics used were; ampicillin (25 µg), streptomycin (25 µg), spectinomycin (25 µg), trimethoprim (5 µg), tetracycline (30 µg), minocycline (30 µg), doxycycline (30 µg), compound sulphonamides (300 µg), chloramphenicol (30 µg), florfenicol (30 µg), kanamycin (30 µg), neomycin (30 µg) amikacin (30 µg), tobramycin (10 µg), netilmicin (30 µg), gentamicin (10 µg), ciprofloxacin (5 µg), nalidixic acid (30 µg), levofloxacin (5 µg), imipenem (10 µg), meropenem (10 µg), doripenem (10 µg), piperacillin-tazobactam (110 µg), Timentin (ticarcillin-clavulanic acid; 85 µg), ampicillin-sulbactam (20 µg), cefotaxime (30 µg), ceftazidime (30 µg), ceftriaxone (30 µg) , cefepime (30 µg) and rifampicin (30 µg) (Oxoid, Basingshire, UK).

2.3.3 Antibiotic resistance phenotype determination using Etest

Minimal inhibitory concentrations (MIC) were determined for colistin (BioMérieux). Single colonies from an isolate of *A. baumannii* were resuspended in 2 mL saline and then used to inoculate a plate made up with 20 mL cation-adjusted MHA. Once dry, a single Etest strip was placed on the surface of the agar and incubated overnight at 37°C.

The MIC was measured, according to the reading guidelines from bBioMérieux (www.biomerieux.com), from the point where the ellipse intersects the scale on the strip.

2.3.4 Conjugation

Equal volumes of donor and recipient cells were mixed on a LA plate and incubated overnight at 37°C. Cells were harvested in 0.9% saline and serially diluted to 10⁻⁸. 100 µL of each dilution was plated onto LA plates supplemented with an antibiotic to select for the recipient (most commonly rifampicin, 100 µg/mL) and an antibiotic, at the concentrations listed in Table 2.3, that the target plasmid confers resistance to in order to isolate transconjugants. Conjugation frequency was calculated as transconjugants per donor with experiments repeated in triplicate. To ensure the randomly selected transconjugants were not spontaneous rifampicin mutants of the donor, colonies were purified and screened for growth on LA supplemented with antibiotics that the donor was resistant to and the recipient was susceptible to. Further confirmation was performed via PCR screening of boil preparations (Section 2.4.1.2) and at least two transconjugants from each conjugation were frozen at -80°C (Section 2.3.1).

2.3.5 Transformation

An overnight culture (Section 2.3.1) was diluted 1:100 in fresh LB. Following growth at 37°C to exponential phase (OD₆₀₀: 0.5-0.7) the cells were pelleted by centrifugation at 4500 rpm for 10 minutes at 4°C. 10 mL of pre-chilled (0°C on ice) 10% glycerol (v/v) was used to resuspend the pellet, followed by centrifuged at 12,000 rpm for 10 minutes at 4°C. This washing step of resuspension in ice-cold 10% glycerol followed by centrifugation was repeated (3 times), concentrating the cells and the final

resuspension was in 300 μ L of pre-chilled 10% glycerol. This was separated into 60 μ L aliquots which were stored at -80°C or kept on ice for immediate use. This protocol follows a slight modification to the Manufacturer's Instructions (Biorad, California).

Electroporation was performed by adding 2 μ L of Genomic DNA (Section 2.4.1.1) or Plasmid DNA (Section 2.4.1.4.2) to an aliquot of electro-competent *A. baumannii* cells. The transformation mixture was transferred into a pre-chilled (0°C on ice) electroporation cuvette and electroporation was performed using a BioRad gene pulser unit with the following parameters, 0.2 cm cuvettes, 2.5 kV, 25 μ F, 200 Ω . One mL of LB was added to the transformation mix and was recovered for 1 hour at 37°C with shaking at 200 rpm. 100 μ L and 10 μ L of transformation reaction were plated on LA plates supplemented with the appropriate antibiotics (concentrations listed in Table 2.3) to select for transformants, subsequently screened by PCR with at least three transformants frozen at -80°C (Section 2.3.1).

2.3.6 Plasmid stability assay

Starter cultures of the strain containing both pS21-1 (kanamycin resistant) and pS32-1 (ticarcillin resistant) with the appropriate selections for both plasmids were grown overnight at 37°C. 100 μ L of the 10^{-5} dilution of starter culture (approximately 1,000 cells) was used to inoculate LB without antibiotic selection and grown overnight (approximately 22 generations). Cultures were diluted, and a fresh LB inoculated with approximately 1,000 cells, and this was performed daily for five cycles. The CFU/mL of cells resistant to kanamycin, ticarcillin and both kanamycin and ticarcillin, was calculated daily by plating serial dilutions on LA plates containing the three different combinations of antibiotics as well as LA alone to determine the total concentration.

Following the 5th subculture, 50 colonies grown on LA were patched onto LA plates containing kanamycin, ticarcillin and both kanamycin and ticarcillin to see if pS21-1 and pS32-1 had both been retained. PCR screening from boil preparations (Section 2.4.1.2) of the patches were used to confirm the presence of the two plasmids.

2.4 DNA preparation and analysis

2.4.1 DNA extraction

2.4.1.1 Genomic DNA for PCR and sequencing

The CTAB (Cetyltrimethyl ammonium bromide) genomic DNA purification method was used to extract chromosomal DNA. 1.5 mL of an overnight culture of an *A. baumannii* isolate (Section 2.3.1) was harvested by centrifugation (23,000 rpm, 30 seconds) and resuspended in 500 µL of TE. Lysozyme (20 µL of 10 mg/mL), RNase (20 µL of 10 mg/mL) and Proteinase K (5 µL of 10 mg/mL) was added to the cell suspension, followed by 100 µL of 10% SDS solution. The samples were incubated for 1 hour at 37°C followed by a second incubation at 50°C for another hour. 200 µL of 5M NaCl followed by 100 µL of 10% CTAB was added and incubated at 65°C for a further 10 minutes. The solution was extracted with an equal volume of chloroform:isomyl alcohol (24:1) solution, followed by a second extraction in a Heavy Phase Lock Gel (5prime, Gaithersburg, USA) tube using an equal volume of phenol:chloroform:isoamyl alcohol (25:24:1). The DNA was then precipitated using 0.6 volumes of isopropanol, followed by a wash with 70% ethanol. Once the ethanol was removed, the dried pellet was resuspended in TE buffer and stored at 4°C. Gel-electrophoresis and the Thermo Scientific Nanodrop spectrophotometer (Thermo Fisher Scientific, Wilmington, Delaware, USA) was used to verify the concentration and purity of the extracted DNA. An A_{260}/A_{280} of 1.8-1.9 was considered suitable for

DNA sequencing. For use in PCR, the DNA was diluted in sterile Milli-Q water to a concentration of 10-50 ng/ μ L (Section 2.4.2).

2.4.1.2 Boil preparation

To screen potential transformants, transconjugants as well colonies of specific AbGRI3 containing isolates, DNA was obtained by boiled preparation methods. A single colony or a single colony size of cell mass from patch plates were picked and placed into an eppendorf tube containing 200 μ L of sterile Milli-Q water with a pre-sterilised loop. Following a 10 minute incubation at 100°C, the contents of the tube were centrifuged at high speed and stored at 4°C.

2.4.1.3 DNA purification technique of PCR products

PCR products to be used for sequencing were cut from agarose gels using a clean scalpel and purified using the QIAquick Gel Extraction Kit (Qiagen, Hilden, Germany). The spin column protocol was followed according to the manufacturer's instructions (QIA quick spin Handbook, 2007). The Thermo Scientific Nanodrop spectrophotometer (Thermo Fisher Scientific, Wilmington, Delaware, USA) was used to determine the concentration of the purified DNA.

2.4.1.4 Plasmid DNA

2.4.1.4.1 Plasmid DNA isolated from *E. coli*

1.5 mL of overnight culture (Section 2.3.1) was harvested by centrifugation, and the pellet was resuspended in 100 μ L of solution I. 200 μ L of freshly made solution II was added to the suspension and left on ice for 10 mins. 150 μ L of solution III was added followed by a further 3-5 minutes incubation on ice. Following centrifugation the

supernatant was extracted with an equal volume of phenol:chloroform:isoamyl alcohol (25:24:1). The DNA was then precipitated using 2 volumes of 100% ethanol, followed by a wash with 70% ethanol. Once the ethanol was removed, the dried pellet was resuspended in TE buffer containing 20 µg/mL RNase. The plasmid DNA was then stored at 4°C. Whole plasmid DNA was visualised using gel electrophoresis (Section 2.4.4.2).

2.4.1.4.2 Plasmid DNA isolated from *A. baumannii*

A 20 mL volume of LB was inoculated with 1/100 volume of overnight culture (Section 2.3.1) and shaken at 37°C to an OD₆₀₀ of 0.4-0.5 was reached. 15mL of the culture was harvested by centrifugation and then the preparation was performed as described in Section 2.4.1.2.1 with the volumes scaled up ten-fold.

2.4.2 Polymerase Chain Reaction (PCR)

2.4.2.1 Primers

All primers were synthesised by Integrated DNA Technologies (Coralville, Iowa, USA). The lyophilised primers were resuspended in TE to a concentration of 500 µM and stored at -20°C. 50 µM working stocks were prepared by dilution with sterile Milli-Q water and were also stored at -20°C. The Primer 3 program (<http://bioinfo.ut.ee/primer3-0.4.0/primer3/>) was used to design primers made throughout this project. All primers were designed to have a standard annealing temperature of 60°C and be approximately 20-23 bp in length. All primers used in this project are listed in Appendix III.

2.4.2.2 Standard PCR

PCR amplification reactions were performed in a total volume of 12.5 μL of 1x ThermoPol PCR reaction buffer containing 200 μM dNTP mix, 1 μM of each primer, 1U Taq polymerase and template DNA. If the DNA template was derived from Section 2.1.1.1, 0.5 μL of diluted DNA was added. However, if the DNA template was derived from boil preparation (Section 2.4.1.2), 5 μL of template was used. Eppendorf Mastercycler, Eppendorf Mastercycler Nexus (Eppendorf, Hamburg, Germany) or a Biorad MyCyclerTM (BioRad, Hercules, USA) series machines were used to perform PCR reactions. The general cycling conditions are as follows; an initial denaturation time of 3 minutes at 94°C, followed by 30 cycles of denaturation at 94°C (30 seconds), primer annealing at 60 °C (30 seconds) and extension at 72°C (extension time varied based on expected amplicon size, where 1 minute was allowed for 1 kb of sequence) and a final extension time of 5 minutes for 72 °C. Gel electrophoresis (Section 2.4.4.1) was used to visualise amplicons and the products were identified by comparison to known standards. A DNA-free control was always included and when available a known control for the target region was used and. If the PCR amplicon was used for sequencing, the reaction was scaled up to 25 μL .

2.4.2.3 Long-range PCR

A high fidelity polymerase was used for PCR amplicons that were predicted to be 3 kb or larger. This reaction was made up to a total volume of 10 μL of 1x HF buffer containing 200 μM dNTP mix, 0.5 μM of each primer and 0.02 U of Phusion polymerase and template DNA. For template DNA derived from Section 2.4.1.1, 0.5 μL of diluted DNA was added, while if derived from Section 2.4.1.2, 5 μL of DNA template was used. Eppendorf Mastercycler, Eppendorf Mastercycler Nexus

(Eppendorf, Hamburg, Germany) or a Biorad MyCyclerTM (BioRad, Hercules, USA) series machines were used to perform PCR reactions. The general cycling conditions are as follows; an initial denaturation time of 3 minutes at 98°C, followed by 35 cycles of denaturation at 98°C (10 seconds), primer annealing at 60 °C (30 seconds) and extension at 72°C (extension time varied based on expected amplicon size, where 1 minute was allowed for 1 kb of sequence) and a final extension time of 5 minutes for 72 °C. Gel electrophoresis (Section 2.4.4.1) was used to visualise amplicons and the products were identified by comparison to known standards. A DNA-free control was always included and when available a known control for the target region was used and a DNA-free control was always included. If the PCR amplicon was used for sequencing, the reaction was scaled up to 20 µL.

2.4.3 Restriction endonuclease digestion

PCR amplicons and plasmid DNA were used as templates for digestion with 5-10 U of restriction endonuclease in 1x buffer and if recommended by the Manufacturer 1x BSA, in a total volume of 20-30 µL. The digestion mixture was incubated in a water bath for 1-4 hours at the optimal temperature for the enzyme used. Agarose gel electrophoresis was used to separate and visualise the digestion. However, the gel electrophoresis protocol varied depending upon predicted product size, between Section 2.4.4.1 (product smaller than 23 kb) and Section 2.4.4.2 (larger than 23 kb).

2.4.4 Agarose gel electrophoresis

2.4.4.1 PCR products and enzyme digestions

The gels used were 1% (w/v) molecular grade agarose (Bioline) dissolved by heating in 0.5x TBE. For visualisation of the DNA, ethidium bromide was added to the cooled

liquid gel at a final concentration of 0.1 µg/mL. DNA (5-25 µL) mixed with 2 µL of loading dye, was loaded into the wells of the set gel and allowed to run at 180 V for approximately 30-45 minutes. The bands on the gel were then visualised using a UV transilluminator and imaged with the Gel Doc XR program (BioRad, Hercules, USA). For comparison, each gel was loaded with 4 µL of 100 bp or 1 kb ladder (New England BioLabs, Ipswich, USA). λ DNA-HindIII Digest (heated at 60°C for 3 minutes before loading 2.5 µL) was also used as a standard when digests were examined.

2.4.4.2 Plasmid DNA

The gels used were 0.7% (w/v) agarose dissolved in 1x TAE for 1 minute. 15 µL of the sample mixed with 2 µL of loading dye was added to the wells of the set gel and allowed to run at 120 V for 1.5-3 hours (depending on the size of the gel). Gels were post-stained with an ethidium bromide solution (0.5 µg/mL) for 20 minutes, then de-stained in water for an hour before visualisation with an UV transilluminator and imaged using a Molecular Imager Gel Doc XR System. The sizing standard for these experiments was a plasmid preparation of the *Escherichia coli* strain 39R861+ (178) that contains 6 plasmids of known or estimated size (2, 7, 38, 66, 95, 155 kb, with the largest unable to enter the gel under the conditions used).

2.5 DNA sequencing and sequence assembly

2.5.1 Sanger sequencing

The sequencing reaction mixture was made up to a total volume of 12 µL and contained 12-18 ng (PCR product <500 bp) or 30-70 ng (PCR product >500 bp) of purified DNA (Section 2.4.1.3). Automated sequencing was performed on an Applied Biosystems

AB3730xl DNA Analyzer using the Big Dye system at the Sydney node of the Australian Genome Research Facility (AGRF).

2.5.2 Whole genome sequencing and sequence assembly

2.5.2.1 Illumina MiSeq

DNA was isolated as per the process outlined in Section 2.4.1.1. 100 µL aliquots, containing 2-5 µg of DNA, were then sent to the Sydney node of the AGRF and were sequenced using Illumina MiSeq, generating paired-end reads of 150 bp.

2.5.3.2 SPAdes assembly

Paired-end reads from MiSeq sequencing were assembled using SPAdes (v.3.5.0) (11) using k-mers of 21, 33, 55, 77, 99 and 127, and a coverage cut-off of 5. Assembly quality was assessed by QUAST (<http://quast.sourceforge.net/quast>).

2.6 Bioinformatics analysis

The contigs assembled from the Illumina HiSeq paired end sequencing and the contigs from Illumina MiSeq paired end sequencing were used to form a database of genomic data for used with the stand-alone BLAST command line application version 2.2.25+ (33). In addition, DNA and amino acid sequencers were analysed using the online BLAST facility (<http://blast.ncbi.nlm.nih.gov/BLAST.cgi>). Pairwise or multiple sequence alignments were performed using ClustalW2 (<http://www.ebi.ac.uk/Tools/msa/clustalw2/>) or using the Artemis Comparison Tool version 12.0.0 (34). The Multi-Locus Sequence Typing (MLST) of each sequence in both the Institut Pasteur and the Oxford scheme was determined online (<https://pubmlst.org>) or by employing a python script, mlstBLAST

(<http://sourceforge.net/projects/srst/files/mlstBLAST/>), which uses blastn to extract profiles from assembled contigs. Resistance genes were identified in assemblies online using ResFinder (<https://cge.cbs.dtu.dk/services/ResFinder/>) (280). ISMapper (108) was used to find the locations of insertion sequences in the paired end Illumina HiSeq short read sequences, while ISseeker (1) was used to locate insertion sequences in draft genomes assembled from Illumina MiSeq sequencing. DNA chromatograms from Sanger sequencing runs were viewed using Sequencher 5.1 (Gene Codes Corp, Ann Arbor, MI, USA) and then trimmed manually. Sequences were assembled using Sequencher and then visualised and annotated using Gene Construction Kit version 4.0.3 (Textco, West Lebanon, NH, USA). SnapGene viewer version 3.3.3 (GSL Biotech LLC, Chicago, USA) was used to annotate circular plasmid sequences.

CHAPTER

3

Phenotypic and genotypic characterisation

3.1 Introduction

Dr Li Yang Hsu kindly supplied a diverse collection of 20 GC2 *A. baumannii* from Singapore General Hospital (SGH) (Table 3.1). One isolate was from 1996 and the remaining 19 were isolated from 2004 to 2011. All isolates had been determined to be resistant to carbapenems and the draft genomes of 7 that had been sequenced at the Wellcome Trust Sanger Institute by Illumina HiSeq were available (bold in Table 3.1). Further information on these isolates are listed in Chapter 2; Table 2.2.

Table 3.1 GC2 isolates from Singapore General Hospital

Isolate	Source¹	Year	GC	<i>oxa23</i>
S21	Unknown	1996	2	+
S22	ETTA	2004	2	+
S23	ETTA	2004	2	+
S24	Trach	2006	2	+
S25	ETTA	2006	2	+
S26	ETTA	2006	2	+
S27	Blood	2007	2	+
S28	Blood	2007	2	+
S29	Blood	2007	2	+
S30	ETTA	2008	2	+
S31	Trach	2008	2	+
S32	Blood	2009	2	+
S33	Blood	2009	2	+
S34	Blood	2009	2	+
S35	Sputum	2010	2	+
S36	Blood	2010	2	+
S37	ETTA	2010	2	+
S38	Sputum	2011	2	+
S39	Sputum	2011	2	+
S40	Blood	2011	2	+

The draft genome of those in bold are available.

¹Tracheal aspiration (Trach) and Endotracheal aspiration (ETTA).

The initial screening of these isolates was performed by Matthew Wynn and showed that all isolates belonged to Turton group 1 (261) confirming that they were members of GC2. They were all found to contain the *oxa23* gene, which accounted for the detected carbapenem resistance and preliminary resistance phenotype determination using disk diffusion was performed (Appendix I; Table A1.2).

The work described in this chapter aimed to complete the phenotypic and genotypic analysis of antibiotic resistance in these isolates. This included completing the full spectrum of antibiotic resistance phenotypes to determine their level of resistance. Genes involved in antibiotic resistance were first detected by PCR screening and then by searching the draft genome sequences as they became available. The relationships between the isolates was examined through MLST.

3.2 Phenotypic resistance profiles

3.2.1 Antibiotic resistance phenotype determination using disk diffusion

Each isolate was screened for resistance 30 different antibiotics listed in Table 3.2 using calibrated Dichotomous Sensitivity (CDS) disk diffusion. The full results showing the annular radii of the zone of inhibition are listed in Appendix I; Table A1.2. Resistance to most antibiotics had been tested previously (by Matthew Wynn and shown in black in Appendix I; Table A1.2) but others were performed during this project (shown in red in Appendix I; Table A1.2). According to the CDS guidelines (<http://web.med.unsw.edu.au/cdstest>), resistance to an antibiotic is indicated by an annular radius of less than 6 mm.

The 20 isolates were all resistant to imipenem, meropenem and doripenem, consistent with the presence of the *oxa23* gene (Table 3.2). They were all also resistant to every extended spectrum cephalosporin, antipseudomonal penicillin with β -lactamase inhibitor, penicillin and β -lactamase inhibitor, quinolone, fluoroquinolone, tetracycline and folate pathway inhibitor tested (Table 3.2). As expected for *A. baumannii*, the isolates were all resistant to both phenicol antibiotics tested. A single isolate, S35, was resistant to rifampicin.

Table 3.2 Antibiotics tested by disk diffusion

Class	Antibiotics
Carbapenems	Imipenem, meropenem, doripenem
Extended-spectrum cephalosporins	Cefotaxime, ceftriaxone, ceftazidime, cefepime
Antipseudomonal penicillins + β -lactamase inhibitors	Piperacillin-tazobactam, ticarcillin-clavulanic acid
Penicillins and β -lactamase inhibitors	Ampicillin/sulbactam, ampicillin
Quinolones	Nalidixic acid
Fluoroquinolones	Ciprofloxacin, levofloxacin
Tetracyclines	Tetracycline, doxycycline, minocycline
Aminoglycosides	Gentamicin, tobramycin, amikacin, netilmicin, kanamycin, neomycin, streptomycin, spectinomycin
Folate pathway inhibitors	Trimethoprim, sulphamethoxazole
Phenicol	Chloramphenicol and florfenicol
Rifampin	Rifampicin

Those in red are a part of the antibiotic categories and antibiotics used to define resistance in *Acinetobacter* (Chapter 1; Table 1.1)

The aminoglycoside resistance profile of the 20 isolates varied and these profiles distinguished four groups in the collection (Table 3.3).

Table 3.3 Aminoglycoside resistance profiles

Isolates	Aminoglycoside resistance							
	Ak	Gm	Km	Net	Nm	Tm	Sm	Sp
S21, S24, S26, S29, S30	+	+	+	-	+	-	+	+
S22, S23, S27, S31-S33, S40	+	+	+	+	+	+	+	+
S25, S28, S35-S39	+	+	+	+	-	+	+	+
S34	+	+	+	+	+	+	+	-

Ak, amikacin; Gm, gentamicin; Net, netilmicin; Nm, neomycin; Sm, streptomycin; Sp, spectinomycin; Tm, tobramycin.

Eight of the nine classes of antibiotics used to define antibiotic resistance in *A. baumannii* are highlighted in red in Table 3.2. As each of the 20 isolates are resistant to all relevant antibiotics in the eight classes, they are at least classified as extensively antibiotic resistant (XAR) (166). If they are additionally found to be resistant to the antibiotics in the ninth class they would be considered pan resistant. The ninth class of applicable antibiotics is polymyxins, which includes colistin and polymyxin B, but due to poor agar diffusion, the determination of resistance to these antibiotics by disk diffusion is not reliable (71). Instead, susceptibility to colistin was measured using Etest.

3.2.3 Colistin resistance

The MIC of colistin was determined for each isolate using Etest strips (bioMerieux) and the CLSI breakpoints of $\leq 2 \mu\text{g/mL}$ as susceptible and $>2 \mu\text{g/mL}$ as resistant were applied (255). All but S35 had an MIC of $\leq 0.75 \mu\text{g/mL}$ of colistin, and were classed as susceptible (Table 3.4). Hence these 19 isolates are XAR. However, the MIC of S35 was $3 \mu\text{g/mL}$, indicating that this isolate was resistant to colistin (Table 3.4). To be classed as pan resistant S35 would have to be also resistant to polymixin B, but this was not available for testing, hence, S35 is still classified as XAR.

Table 3.4 Colistin MIC

Isolates	Colistin MIC ($\mu\text{g/mL}$)
S37	0.25
S21, S23, S28, S31-S34, S36, S38-S40	0.38
S22, S24, S30	0.5
S25-S27, S29	0.75
S35	3

3.3 Genotypic resistance profile

3.3.1 IS upstream of *ampC*

The isolates were screened for ISAbal upstream of the *ampC* gene (Fig 3.1), which is known to increase the expression of *ampC* and confer resistance to third generation cephalosporins (42, 89). The published primer pair RH581 and RH582 (92) produced an amplicon of approximately 2.5 kb in all isolates, consistent with the predicted amplicon size of 2,583 bp for ISAbal upstream of *ampC*. If ISAbal was not present, the amplicon would only be 1,403 bp. Resistance to third generation cephalosporins in these 20 isolates is likely due to the increased expression of *ampC* by ISAbal.

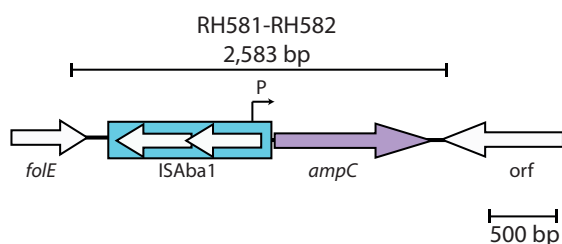


Figure 3.1 ISAbal upstream of *ampC*. Arrows indicate extent and orientation of genes and open reading frames. The *ampC* gene is shown in purple. The blue box is ISAbal and the internal arrows indicate the transposase orientation. The bent arrow indicates the orientation and location of the promoter. Primers used to detect if ISAbal is upstream of *ampC* are shown by the bars above along with the PCR amplicon size.

3.3.2 Resistance gene screening

In an attempt to account for the resistance profiles detected, the presence of nine resistance genes were screened for in each isolate using the gene specific primer combinations shown in Appendix III; Table A3.1. These included the sulphonamide resistance genes *sul1* and *sul2*, tetracycline resistance gene *tetA*(B), an ampicillin resistance gene *bla*_{TEM} and a selected group of aminoglycosides resistance genes including *aacC1*, *aadA1*, *aacA4*, *aphA1b*, *strA-strB*, *aphA6*, and *armA*. All of these genes have previously been detected in GC2 *A. baumannii* isolates.

The different combinations of resistance genes that were detected are shown in Table 3.5. In addition to the *oxa23* gene, all isolates contained the *tetA*(B) gene, accounting for resistance to tetracycline, doxycycline and minocycline, and the *strA-strB* resistance genes, causing resistance to streptomycin. Sulphonamide resistance was accounted for by *sul1* or *sul2* genes, though some isolates contained both. Some isolates included the *bla*_{TEM} gene, but this did not appear to affect their ampicillin resistance to profiles if absent. The aminoglycoside resistance gene content varied between isolates (Table 3.5) and is described in Section 3.3.2.1.

Table 3.5 Resistance genes detected in addition to *oxa23*, *tetA(B)* and *strA-strB*

Isolates	<i>sul1</i>	<i>sul2</i>	<i>bla</i> _{TEM}	Aminoglycoside resistance genes						
				<i>aacC1</i>	<i>aadA1</i>	<i>aphA6</i>	<i>armA</i>	<i>aacA4</i>	<i>strA-strB</i>	<i>aphA1b</i>
S21, S24, S26, S29	+	+	+	+	+	+	-	-	+	+
S30	+	-	+	+	+	+	-	-	+	+
S22, S23	+	+	+	+	+	-	+	+	+	+
S25, S28	+	+	+	+	+	-	+	+	+	-
S27	+	+	-	-	+	-	+	+	+	-
S31	+	-	-	+	+	-	+	+	+	-
S32-S34, S40	-	+	+	-	-	-	+	-	+	+
S35-S37	-	+	-	-	-	-	+	-	+	-
S38, S39	+	-	-	+	+	-	+	+	+	-

Acquired resistance genes do not fully account for resistance to trimethoprim and some of the β -lactam combinations (Table 3.5). They also do not account for resistance to quinolone and fluoroquinolone antibiotics or for rifampicin, however, in *Acinetobacter*, mutational resistance is the most common mechanism of resistance. Mutational resistance to these antibiotics is examined further in Section 3.4.2.1 and Section 3.4.2.2, after the genome sequence became available.

3.3.2.1 Aminoglycoside resistance

Aminoglycoside resistance determinants usually only confer resistance to a subset of aminoglycosides (shown in Table 3.6), so combinations of these genes would be required to confer the spectrum of aminoglycoside resistance detected in the 20 Singapore isolates (Table 3.2). Even the *armA* gene on its own, which confers resistance to all clinically relevant aminoglycoside antibiotics (Table 3.6), does not account for resistance to neomycin, streptomycin and spectinomycin.

Table 3.6 Phenotypes conferred by aminoglycoside resistance genes

Gene	Confers resistance to
<i>aacC1</i>	Gm
<i>aphA1b</i>	Km, Nm
<i>aphA6</i>	Ak, Km, Nm
<i>aacA4</i>	Ak or Gm ¹ , Net, Tm
<i>aadB</i>	Km, Gm, Tm
<i>armA</i>	Ak, Gm, Net, Tm, Km
<i>aadA1</i>	Sm, Sp
<i>strA-strB</i>	Sm

Ak, amikacin; Gm, gentamicin; Net, netilmicin; Nm, neomycin; Sm, streptomycin; Sp, spectinomycin; Tm, tobramycin.

¹The substrate specificity of *aacA4* changes based on a single base difference that results in an amino acid change.

The 20 isolates fall into two groups based on the presence or absence of *armA* and *aphA6* (Table 3.7). The isolates in the first group, S21, S24, S26, S29 and S30, lack *armA* but contain *aphA6* (orange in Table 3.7). The *aphA6* gene confers resistance to amikacin which is a clinically important antibiotic, as well as to kanamycin and neomycin. They all also contain the *aphA1b* gene, which also confers resistance to kanamycin and neomycin. Gentamicin resistance is conferred by the *aacC1* genes in these isolates and streptomycin and spectinomycin resistance can be accounted for by the *aadA1* and *strA* and *strB* genes. The absence of any genes conferring resistance to tobramycin or netilmicin in these isolates is consistent with their phenotype (Table 3.2).

Table 3.7 Two groups based on the *armA* and *aphA6* resistance genes

G ¹	Isolates	<i>oxa23</i> , <i>tetA(B)</i> , <i>strA-strB</i>	<i>aphA6</i>	<i>armA</i>	<i>aphA1b</i>	<i>aacC1</i>	<i>aadA1</i>	<i>aacA4</i>	<i>sul1</i>	<i>sul2</i>	<i>bla</i> _{TEM}
1	S21, S24, S26, S29	+	+	-	+	+	+	-	+	+	+
	S30	+	+	-	+	+	+	-	+	-	+
2	S22, S23	+	-	+	+	+	+	+	+	+	+
	S25, S28	+	-	+	-	+	+	+	+	+	+
	S27	+	-	+	+	-	+	+	+	+	-
	S31	+	-	+	+	+	+	+	+	-	-
	S32-S34, S40	+	-	+	-	-	-	-	-	+	+
	S35- S37	+	-	+	-	-	-	-	-	+	-
	S38, S39	+	-	+	-	+	+	+	+	-	-

Orange box indicates isolates with *aphA6* and are group 1, while the green box shows isolates with *armA* that are group 2.

¹Group.

The second group, with 15 isolates, all contain the *armA* gene (green in Table 3.7), which accounts for their resistance to amikacin, gentamicin, tobramycin and netilmicin, which are all clinically relevant antibiotics. In addition, some isolates also included the *aacA4* and/or the *aacC1* genes but as these confer resistance to a subset of the same antibiotics as *armA*, their presence cannot be inferred from phenotype alone. The subgroup that are neomycin resistant contain the *aphA1b* gene, as resistance to this antibiotic is not conferred by *armA*.

Most isolates, regardless of its group, include the *aadA1* gene, which confers resistance to streptomycin and spectinomycin, though they are not used clinically in humans (144). A subgroup within the 15 isolates consisting of S32, S33, S34 and S40 lack the *aadA1* gene (Table 3.7), but only S34 is susceptible to spectinomycin (Table 3.2).

3.3.3 Tn2006 screening

The *oxa23* gene is commonly located in Tn2006, a transposon bounded by inversely oriented copies of ISAbal (61, 194, 215). All isolates were screened for Tn2006 using published primer pairs (88), ISAbalB paired with *oxa23F* and *oxa23R* paired with ISAbalB (Fig 3.2A). This screening confirmed that each isolate had an ISAbal upstream of *oxa23* (Fig 3.2B) and all but one isolate produced an amplicon of the expected size for the ISAbal position downstream of *oxa23* in Tn2006 (Fig 3.2C). Thus, in 19 of the isolates the *oxa23* gene is located in Tn2006. The amplicon for the downstream ISAbal for S30 was not as crisp and as strong as the others (Fig 3.2C). As the left end of Tn2006 is conserved in S30, the *oxa23* gene in this isolate may be located within another *oxa23* containing transposon (194).

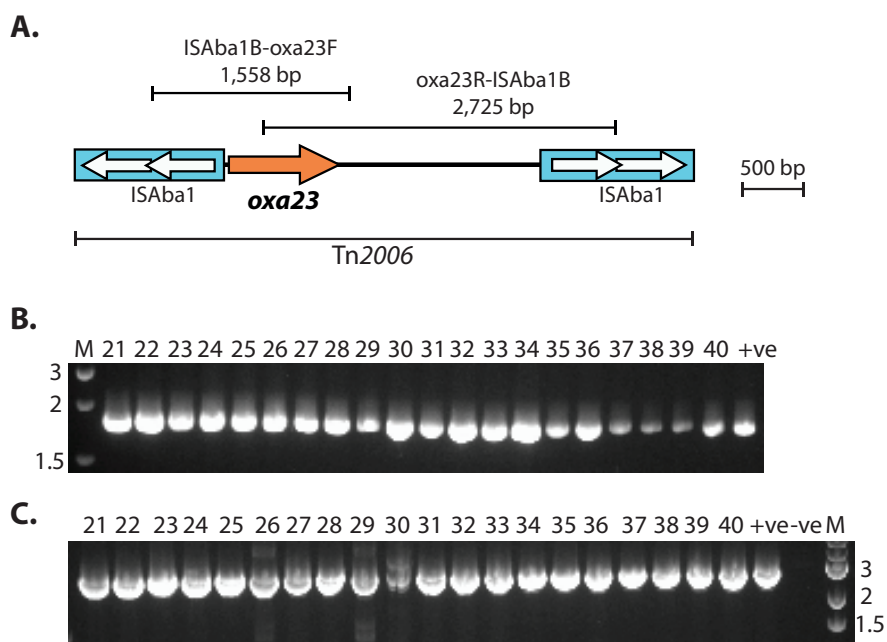


Figure 3.2 Tn2006 screening. A. Location of primers for detection of Tn2006. The orange arrow indicates the *oxa23* gene and the blue boxes represent ISAbal with the internal arrows showing the direction of the transposase. Primers used to detect Tn2006 are shown by the bars above along with the predicted amplicon size. B. Screening the left end of Tn2006 using primers ISAbalB and oxa23F to amplify a 1,558 bp amplicon. C. oxa23R and ISAbalB were paired to amplify the right end of Tn2006. The sizes of molecular markers on the left and right are in kb and isolate numbers are shown above their lanes.

3.3.4 TnaphA6 screening

The *aphA6* gene is most commonly located in TnaphA6, which is a class 1 transposon bounded at each end by ISAbal25 (Fig 3.3) (197). Two previously designed primer pairs, RH577 and aphA6R (703 bp amplicon), and aphA6F and RH576 (1,001 bp amplicon) (Fig 3.3) (197), were used screen for TnaphA6 in the five *aphA6* containing isolates. S21, S24, S26, S29 and S30 produced amplicons consistent with the expected sizes for both PCRs, indicating that the *aphA6* gene was located in TnaphA6 in these isolates.

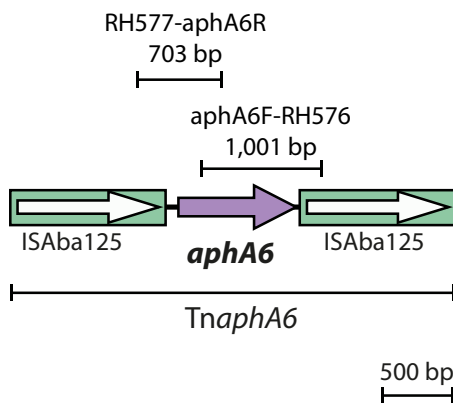


Figure 3.3 Structure of *TnaphA6*. The purple arrow indicates the extent and orientation of the *aphA6* gene. ISAbal25 are shown as green boxes with the internal arrows indicating the orientation of the transposase. Primer combinations used to detect *TnaphA6* are shown by the black bars above, along with the amplicon size.

3.4 Detection of resistance gene content using the draft genome sequences

The draft genomes of S21, S24, S27, S29, S34, S38 and S39 which were available at the start of this project had been sequenced using Illumina HiSeq and assembled using Velvet and summary of these assemblies and accession numbers are provided in Appendix I, Table A1.3. Draft genomes of the remaining 13 of these isolates were not available until the end of 2016. The Illumina MiSeq platform was used to sequence the latter isolates and their draft genomes were assembled using SPAdes (details of the assemblies can be found in Appendix I, Table A1.4). The analysis described in the following sections was first performed on the seven sequences, but was completed on the remaining 13 once their sequences were available.

3.4.1 Resistance genes detected using ResFinder

Analysis of the draft genomes of the 20 isolates confirmed the presence of the resistance genes detected by PCR screening, and revealed that 16 isolates had additional resistance genes (blue in Table 3.8). All of the *armA* containing isolates (those in group 2 in Table 3.7) also included the *msrE* and *mphE* macrolide resistance genes. When present, the *armA*, *msrE*, *mphE*, *sull*,

aadA1, *catB8* and *aacA4* genes were always on the same contig (underlined in Table 3.8). The isolate S30 from the *aphA6* group also had the *msrE* and *mphE* genes, but they were not associated with *armA*. Instead, it was seen on the same contig as the *tet39* tetracycline resistance determinant (underlined in Table 3.8). Examination of the draft genome sequence of S30 revealed that *oxa23* was in fact located in Tn2008. This transposon only contains the upstream ISAbal and no downstream ISAbal (194), which is consistent with earlier screening (Section 3.3.3).

Table 3.8 Resistance genes detected using ResFinder

Group	Isolates	Detected resistance genes
1	S21, S24, S26, S29	<i>sul2</i> , <i>strA-strB</i> , <i>tetA(B)</i> , <i>bla</i> _{TEM} , <i>aphA1b</i> , <i>aacC1</i> , <i>aadA1</i> , <i>sul1</i> , <i>aphA6</i> , <i>oxa23</i>
	S30	<i>strA-strB</i> , <i>tetA(B)</i> , <i>bla</i> _{TEM} , <i>aphA1b</i> , <i>aacC1</i> , <i>aadA1</i> , <i>sul1</i> , <i>aphA6</i> , <i>oxa23</i> , <i>msrE-mphE</i> , <i>tet39</i>
2	S22, S23	<i>sul2</i> , <i>strA-strB</i> , <i>tetA(B)</i> , <i>bla</i> _{TEM} , <i>aphA1b</i> , <i>aacC1</i> , <i>oxa23</i> <i>armA</i> , <i>msrE-mphE</i> , <i>sul1</i> , <i>aadA1</i> , <i>catB8</i> , <i>aacA4</i>
	S25, S28	<i>sul2</i> , <i>strA-strB</i> , <i>tetA(B)</i> , <i>bla</i> _{TEM} , <i>aacC1</i> , <i>oxa23</i> <i>armA</i> , <i>msrE-mphE</i> , <i>sul1</i> , <i>aadA1</i> , <i>catB8</i> , <i>aacA4</i>
	S27	<i>sul2</i> , <i>strA-strB</i> , <i>tetA(B)</i> , <i>aphA1b</i> , <i>oxa23</i> <i>armA</i> , <i>msrE-mphE</i> , <i>sul1</i> , <i>aadA1</i> , <i>catB8</i> , <i>aacA4</i>
	S31	<i>strA-strB</i> , <i>tetA(B)</i> , <i>aphA1b</i> , <i>aacC1</i> , <i>oxa23</i> <i>armA</i> , <i>msrE-mphE</i> , <i>sul1</i> , <i>aadA1</i> , <i>catB8</i> , <i>aacA4</i>
	S32, S33, S34, S40	<i>sul2</i> , <i>strA-strB</i> , <i>tetA(B)</i> , <i>bla</i> _{TEM} , <i>aphA1b</i> , <i>oxa23</i> , <i>armA</i> , <i>msrE-mphE</i>
	S35, S36, S37	<i>sul2</i> , <i>strA-strB</i> , <i>tetA(B)</i> , <i>oxa23</i> , <i>armA</i> , <i>msrE-mphE</i>
	S38, S39	<i>strA-strB</i> , <i>tetA(B)</i> , <i>aacC1</i> , <i>oxa23</i> <i>armA</i> , <i>msrE-mphE</i> , <i>sul1</i> , <i>aadA1</i> , <i>catB8</i> , <i>aacA4</i>

Blue genes indicate addition resistance genes identified by ResFinder and underlined genes are on the same contig.

Interestingly, the presence of *catB8* would not have affected the phenotypic characteristics of the isolates that contain them, as all isolates are intrinsically resistant to chloramphenicol. Macrolide antibiotics such as erythromycin are not used to treat Gram-negatives because they are slow to traverse the cell wall and are only effective against Gram-positives (170). The

ability of the *msrE* and *mphE* genes to confer resistance to erythromycin was examined further in Chapter 10; Section 10.3.

3.4.2 Mutations leading to antibiotic resistance

3.4.2.1 Quinolone and fluoroquinolone resistance

Mutations in *gyrA* and *parC* were detected using standalone BLAST searches with the sequence of the fluoroquinolone susceptible A320 as a query. All isolates had the single base mutation that leads to the Ser81Leu substitution in GyrA. All but S24, also had the single base change that leads to Ser84Leu in ParC. These mutations have been shown to be equivalent to those in *E. coli* (Ser83Leu in GyrA and Ser80Leu in ParC), which are known to cause high levels of resistance (267, 268). Hence, these mutations account for the quinolone and fluoroquinolone resistance in 19 of the 20 isolates.

3.4.2.2 Rifampicin resistance in S35

Mutations in the *rpoB* gene have been shown to confer resistance to rifampicin in *A. baumannii* (79). Sequence comparison to *rpoB* in the rifampicin susceptible isolate A320 using standalone BLAST revealed that all but S35 had an *rpoB* sequence identical to that of A320, consistent with their susceptible phenotype. However, the sequence of *rpoB* in S35 had two single base differences that cause the changes Leu520Ser and Asp525Gly in RpoB. The Asp535Gly mutation has previously been detected in rifampicin resistant *A. baumannii* (79) and likely accounts for S35 resistance to this antibiotic.

3.5 MLST

The pubMLST website (<https://pubmlst.org/Abaumannii>) was used to determine the ST of each isolate using both the Institut Pasteur and the Oxford schemes.

3.5.1 Institut Pasteur scheme

A total of 18 isolates were ST2 when the Institut Pasteur scheme was used (Table 3.9). However, the two remaining isolates S24 and S34, were each a single locus variant (SLV) of ST2 differing at the *fusA* and *cpn60* allele, respectively (Table 2.9). In both isolates, the different allele present is the ST1 (GC1) allele, indicating that recombination may have occurred in these isolates, incorporating DNA from a GC1 isolate. Despite the two differences at the single alleles, the 20 isolates are all GC2.

Table 3.9 MLST profiles using the Institut Pasteur scheme

Isolates	Year	ST (IP)	Alleles						
			<i>cpn60</i>	<i>fusA</i>	<i>gltA</i>	<i>pyrG</i>	<i>recA</i>	<i>rplB</i>	<i>rpoB</i>
18 isolates	1996-2011	2	2	2	2	2	2	2	2
S24	2006	641	2	1	2	2	2	2	2
S34	2009	98	1	2	2	2	2	2	2

3.5.2 Oxford scheme

Using the Oxford scheme there was much more variation within the collection, particularly in the *gpi* allele (Table 3.10). Seven different STs were detected including ST208 (five isolates), ST218 (four isolates), ST195 (six isolates) and a single example of ST369 and ST1128. In addition, two novel STs were detected, submitted to the pubMLST website, and assigned ST1165 (S26 and S29) and ST1166 (S34). All of the ST variants, except S34, are SLVs of ST208 with the variation confined to the *gpi* allele. However, ST1166 is a triple locus variant (TLV) of ST208 in the *cpn60*, *gpi* and *rpoD* alleles indicating that additional recombination has occurred in this isolate (Table 3.10).

Table 3.10 MLST profiles using the Oxford scheme

Isolate	Year	ST (IP)	ST (Ox)	Oxford alleles						
				<i>gltA</i>	<i>gyrB</i>	<i>gdhB</i>	<i>recA</i>	<i>cpn60</i>	<i>gpi</i>	<i>rpoD</i>
S21	1996	2	208	1	3	3	2	2	97	3
S22	2004	2	218	1	3	3	2	2	102	3
S23	2004	2	218	1	3	3	2	2	102	3
S24	2006	641	208	1	3	3	2	2	97	3
S25	2006	2	218	1	3	3	3	2	102	3
S26	2006	2	1165	1	3	3	2	2	164	3
S27	2007	2	218	1	3	3	2	2	102	3
S28	2007	2	1128	1	3	3	2	2	260	3
S29	2007	2	1165	1	3	3	2	2	164	3
S30	2008	2	369	1	3	3	2	2	106	3
S31	2008	2	208	1	3	3	2	2	97	3
S32	2009	2	195	1	3	3	2	2	96	3
S33	2009	2	195	1	3	3	2	2	96	3
S34	2009	98	1166	1	3	3	2	4	106	2
S35	2010	2	195	1	3	3	2	2	96	3
S36	2010	2	195	1	3	3	2	2	96	3
S37	2010	2	195	1	3	3	2	2	96	3
S38	2011	2	208	1	3	3	2	2	97	3
S39	2011	2	208	1	3	3	2	2	97	3
S40	2011	2	195	1	3	3	2	2	96	3

As the *gpi* allele lies in the capsule locus (KL), variation in this allele generally accompanies a replacement of the capsule gene cluster (3, 119, 137). The KL present in each isolate was kindly determined by Dr Johanna Kenyon (personal communication) and each change in the *gpi* allele is consistent with a different capsule being produced (Table 3.11).

Table 3.11 Connection between the *gpi* allele and capsule locus replacement

Isolates	ST (Oxford)	<i>gpi</i> allele	KL
<u>S21</u> , <u>S24</u> , S31, S38, S39	208	97	2
S22, S23, S25, S27	218	102	7
<u>S26</u> , <u>S29</u>	1165	164	73
S28	1128	260	32
<u>S30</u> , S34	369, 1166	106	49
S32, S33, S35, S36, S37, S40	195	96	3

The isolates in the *aphA6* group are underlined.

3.6 Discussion

The 20 GC2 *A. baumannii* isolates from Singapore General Hospital were all classed as XAR and all but one isolate remains susceptible to colistin. The majority of the antibiotic resistance

in these isolates can be accounted for by intrinsic resistance, acquired antibiotic resistance genes, specific mutations in chromosomal genes as well as increased expression of the chromosomal *ampC* gene.

However, resistance to some antibiotics was not able to be accounted for by the analysis in this chapter. This included the isolates resistance to trimethoprim and some of the β -lactam combinations. It is possible that a combination of resistance genes and efflux pumps such as AdeABC, AdeIJK, AdeFGH and AbeM (44) are contributing to these resistances. The resistance of S24 to fluoroquinolones could be accounted for by mutations in *gyrB* and *parE* (133, 141) or increased expression of efflux pumps (44, 247, 278) have been shown to contribute to resistance to these antibiotics. The cause of colistin resistance in S35 was also not examined during the course of this project but mutations in *pmrACB* locus have been shown to contribute resistance to this antibiotic (211).

The lack of resistance of S34 to spectinomycin is quite unusual, despite this isolates lack of an *aadA1* gene. Recently, an aminoglycoside nucleotidyltransferase (AadA) intrinsic to *A. baumannii*, ANT(3'')-II, has been identified and shown to contribute to resistance to streptomycin and spectinomycin (283). Expression of the gene encoding ANT(3'')-II in *E. coli* showed an increase in spectinomycin MIC from 32 to 2,048 $\mu\text{g/mL}$ (283). A gene matching *ant(3'')-II* was present in all 20 isolates and raises the question of why S34 is susceptible to spectinomycin. S32, S33 and S40 encode the same *ant(3'')-II* gene which is 98% identical (770/786) to that in S34. These SNPs cause 6 aa substitutions in the protein of S34 which may have a deleterious effect on the functioning of the protein and could potentially account for the spectinomycin susceptibility in this isolate.

Within the antibiotic resistance gene content of these isolates, there is also a high amount of redundancy. Many isolates contain both *sul1* and *sul2*, when one of these genes alone confers high levels of sulphonamide resistance. The isolate, S30, also contains two tetracycline resistance gene and many of the *armA* containing isolates contain additional aminoglycoside resistance genes which are already covered by the spectrum of *armA*. Furthermore, some isolates have also acquired the chloramphenicol resistance gene *catB8* which is superfluous as *A. baumannii* are intrinsically resistant to this antibiotic.

Two main populations can be distinguished in this collection based on aminoglycoside resistance genes. Isolates in the first group lack the *armA* gene and carry the *aphA6* gene located in *TnaphA6* (Table 3.12), while in the second group isolates contain *armA* but no *aphA6* gene (Table 3.12).

Table 3.12 Characteristics of group 1 isolates containing *aphA6*

Isolate	Year	ST (IP)	ST (Ox)	KL	Resistance genes and likely location
S21	1996	2	208	2	AbGRI1: <i>sul2</i> , <i>strA-strB</i> , <i>tetA(B)</i>
S24	2006	641	208	2	AbGRI2: <i>bla</i> _{TEM} , <i>aphA1b</i> , <i>aacC1</i> , <i>aadA1</i> , <i>sul1</i>
S26	2006	2	1165	73	Other: <i>aphA6</i> , <i>oxa23</i>
S29	2007	2	1165	73	
S30	2008	2	369	49	AbGRI1: <i>strA-strB</i> , <i>tetA(B)</i> AbGRI2: <i>bla</i> _{TEM} , <i>aphA1b</i> , <i>aacC1</i> , <i>aadA1</i> , <i>sul1</i> Other: <i>aphA6</i> , <i>oxa23</i> , <i>tet39</i> , <i>msrE-mphE</i>

The majority of the antibiotic resistance genes in group 1 isolates are those characteristically found in the AbGRI1 or AbGRI2 resistance islands, and will likely assemble into these regions (Table 3.12). In addition to these resistance islands, they all have *TnaphA6*, which has been detected on RepAci6 plasmids and the *oxa23* gene has been detected in many locations either in plasmids or in the chromosome. However, in this group there was lots of deviation when the sequence type in both the Institut Pasteur and the Oxford scheme were considered (Table 3.12). Only S29 and S26 appear to potentially be siblings.

Some of the resistance genes present in the *armA* positive isolates can be assigned to AbGRI1 and AbGRI2 island, though S35-S37 lack all of the characteristic AbGRI2 genes (Table 3.13).

Table 3.13 Characteristics of group 2 containing *armA*

Isolate	Year	ST (IP)	ST (Ox)	KL	Resistance genes and likely location
S22	2004	2	218	7	AbGRI1: <i>sul2</i> , <i>strA-strB</i> , <i>tetA</i> (B)
S23	2004	2	218	7	AbGRI2: <i>bla</i> _{TEM} , <i>aphA1b</i> , <i>aacC1</i> 3 rd island: <i>armA</i> , <i>aacA4</i> , <i>catB8</i> , <i>msrE-mphE</i> , <i>aadA1</i> , <i>sul1</i> Other: <i>oxa23</i>
S25	2006	2	218	7	AbGRI1: <i>sul2</i> , <i>strA-strB</i> , <i>tetA</i> (B)
S28	2007	2	1128	32	AbGRI2: <i>bla</i> _{TEM} , <i>aacC1</i> 3 rd island: <i>armA</i> , <i>aacA4</i> , <i>catB8</i> , <i>msrE-mphE</i> , <i>aadA1</i> , <i>sul1</i> Other: <i>oxa23</i>
S27	2007	2	218	7	AbGRI1: <i>sul2</i> , <i>strA-strB</i> , <i>tetA</i> (B) AbGRI2: <i>aphA1b</i> 3 rd island: <i>armA</i> , <i>aacA4</i> , <i>catB8</i> , <i>msrE-mphE</i> , <i>aadA1</i> , <i>sul1</i> Other: <i>oxa23</i>
S31	2008	2	208	2	AbGRI1: <i>strA-strB</i> , <i>tetA</i> (B) AbGRI2: <i>aphA1b</i> , <i>aacC1</i> 3 rd island: <i>armA</i> , <i>aacA4</i> , <i>catB8</i> , <i>msrE-mphE</i> , <i>aadA1</i> , <i>sul1</i> Other: <i>oxa23</i>
S32	2009	2	195	3	AbGRI1: <i>sul2</i> , <i>strA-strB</i> , <i>tetA</i> (B)
S33	2009	2	195	3	AbGRI2: <i>bla</i> _{TEM} , <i>aphA1b</i>
S34	2009	98	1166	49	3 rd island: <i>armA</i> , <i>msrE-mphE</i> Other: <i>oxa23</i>
S35	2010	2	195	3	AbGRI1: <i>sul2</i> , <i>strA-strB</i> , <i>tetA</i> (B)
S36	2010	2	195	3	AbGRI2: -
S37	2010	2	195	3	3 rd island: <i>armA</i> , <i>msrE-mphE</i> Other: <i>oxa23</i>
S38	2011	2	208	2	AbGRI1: <i>strA-strB</i> , <i>tetA</i> (B)
S39	2011	2	208	2	AbGRI2: <i>aacC1</i> 3 rd island: <i>armA</i> , <i>aacA4</i> , <i>catB8</i> , <i>msrE-mphE</i> , <i>aadA1</i> , <i>sul1</i> Other: <i>oxa23</i>

However, quite a large portion of these genes were contiguous with one another and had the same gene content as a third chromosomally located resistance island described recently (277). Hence, these isolates may also contain a version of this resistance island. Some isolates in this group contain the *aphA1b* (red in Table 3.13) which has been associated with both AbGRI2 (24, 188) and this new resistance island (277), and could potentially be in either region or both. Again, they all contain the *oxa23* gene in Tn2006 which could be potentially located in the chromosome or in RepAci6 plasmids. The diversity in the Oxford MLST profiles of these 15 isolates provided further evidence of the variation present in this collection.

CHAPTER

4

AbGRI1 and other chromosomal locations of *oxa23*

4.1 Introduction

The *oxa23* carbapenemase resistance gene in *A. baumannii* has been detected in a number of different transposons, including Tn2006, Tn2007, Tn2008, Tn2008B and Tn2009 (194, 215). All but Tn2007, which has ISAbA4 upstream, have ISAbA1 upstream of *oxa23*, and Tn2006 and Tn2009 also have a copy of ISAbA1 downstream of this gene (61, 194, 215). Tn2006 has been experimentally shown to transpose and like ISAbA1 it creates a 9 bp TSD (181). The *oxa23* gene is commonly detected in carbapenem resistant *A. baumannii*, both in the chromosome and in plasmids (194).

Tn2006 can also be found in a larger transposon Tn6022 and the entire structure was named AbaR4 (3). Tn6022 is an 11,998 bp transposon that encodes five genes likely involved in its transposition; *tniC*, *tniA*, *tniB*, *tniD* and *tniE*, as well as the *uspA* and *sup* genes (Fig 4.1) (88). In AbaR4, Tn2006 interrupts the *sup* gene and is surrounded by the 9 bp TSD 5'-CCCGCGAAT-3' (Fig 4.1). AbaR4 has been detected in the *comM* gene (88), but it was first identified in a different chromosomal location in AB0057 which already had AbaR3 in *comM* (AB57_0566, GenBank accession number CP001182) (3).

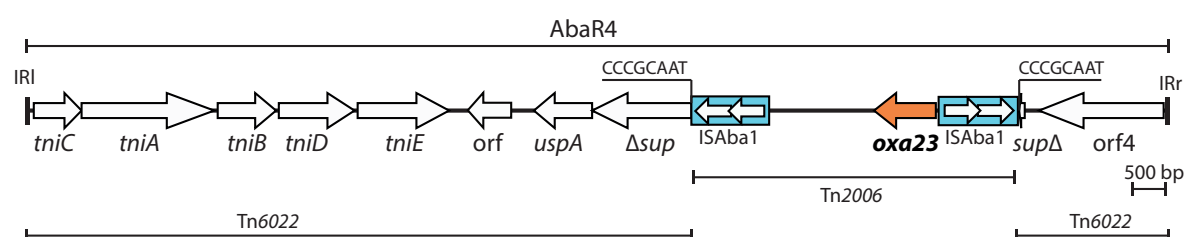


Figure 4.1 AbaR4 structure. Arrows indicate the extent and orientation of genes and open reading frames. The *oxa23* gene is orange. ISAbA1 are shown as light blue boxes and the internal arrows represent the transposase. Vertical bars indicate IRs. Bars above and below indicate extents of transposons. Figure made to scale from GenBank accession number JN96883.

Most GC2 isolates contain a resistance island in *comM*, and in all but one of the genomes completed so far, the resistance island contains Tn6022 or a specific deletion derivative Tn6022Δ1 at one end (190). Tn6022Δ1 (190) is missing 2.85 kb of the transposition module,

including part of *tniB* and *tniE* and all of *tniD* (260) (Fig 4.2A). Hence, to distinguish it from the AbaR islands present in *comM* GC1 isolates, the island variants containing Tn6022 or Tn6022Δ1 in the *comM* gene in GC2 isolates are called AbGRI1 and variants are numbered (191).

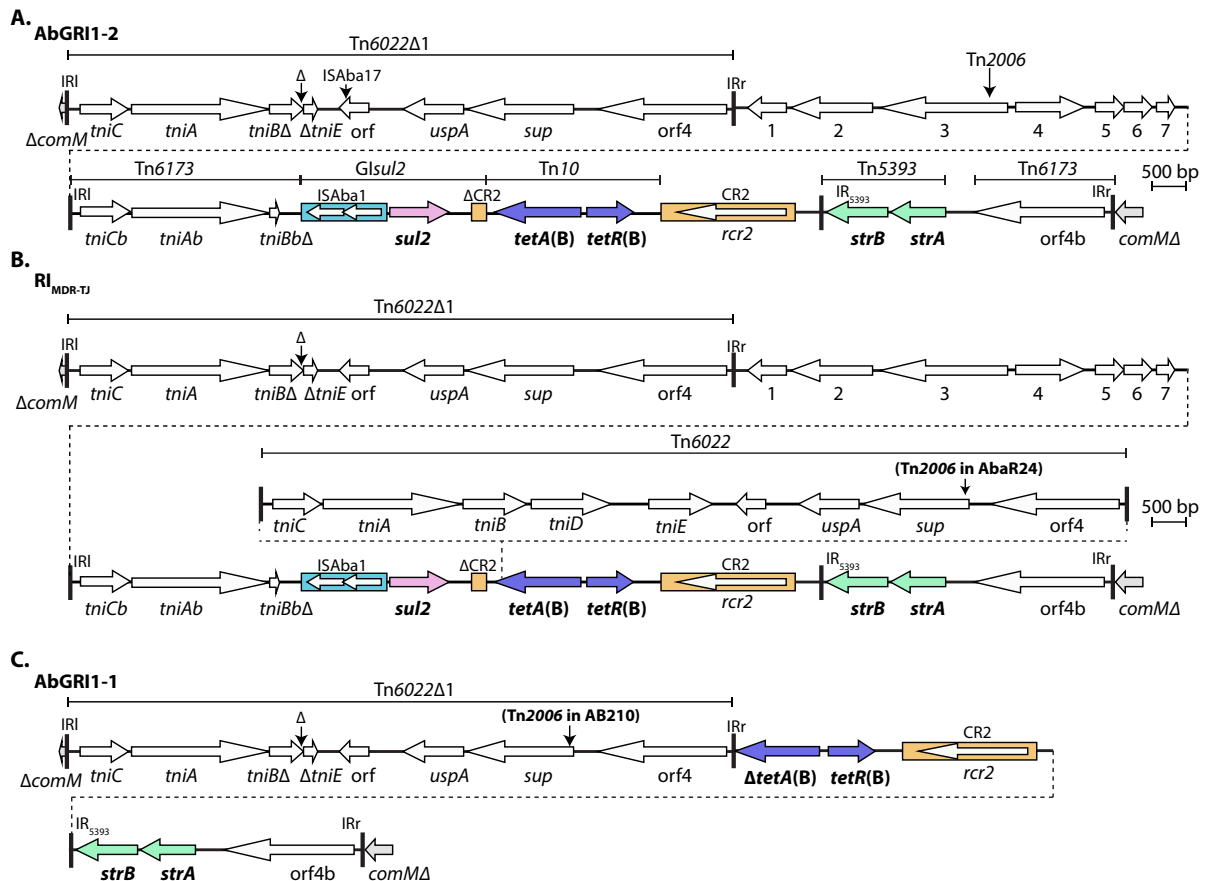


Figure 4.2. Variants of AbGRI1. Structure of A. AbGRI1-2, B. RI_{MDR-TJ} and C. AbGRI1. Arrows indicate extent and orientation of genes and open reading frames. The *sul2* gene is pink, the *tetA(B)* and *tetR(B)* genes are dark blue and the *strB* and *strA* genes are light green. Adjacent chromosomal genes are grey. The light blue box represents ISAbal, the yellow boxes indicate CR2 elements and the internal arrows indicate the transposase and rolling circle replicase, respectively. Vertical bars indicate IRs. Bars above indicate the origin of segments and vertical arrows indicate the deletion (Δ) and location of insertions. Figures made to scale from GenBank accession number JN968483 for AbGRI1-2, GenBank accession number CP003500 for RI_{MDR-TJ} and GenBank accession number JN247441 for AbGRI1-1.

Most Australian GC2 isolates contain AbGRI1-2 (37 kb in size and originally named Tn6167) (193), which is composed of Tn6022Δ1, an orf region, part of the transposition module of Tn6173 (94), a fragment originating from GIsul2, including the *sul2* resistance gene, the *tetB*

tetracycline resistance determinant from Tn10 and the *strA* and *strB* genes (conferring resistance to streptomycin) from Tn5393 (Fig 2A) (191). In addition, AbGRI1-2 has two unique insertions that are not seen in any other known AbGRI1 variant. The first is ISAbal7 downstream of $\Delta tniE$ and the second is Tn2006 inserted in orf 8 (vertical arrows Fig 4.2A) (191).

A larger island (41.6 kb), called RI_{MDR-TJ}, was found in the completed genome of MDR-TJ (125) and has the same basic structure as AbGRI1-2 but lacks ISAbal7 and Tn2006 (Fig 2B). In addition, RI_{MDR-TJ} has a complete copy of Tn6022 inserted in the end of the *tetA(B)* gene separating the last 6 bases. A replacement stop codon for the *tetA(B)* gene is found 16-18 bp into the IR of Tn6022 (Fig 4.2B). A closely-related version of this island, originally called AbaR24, was assembled from isolates from Latvia but had AbaR4 instead of Tn6022 in the same position (Fig 4.2B) (232).

The variant detected in A320, the oldest GC2 isolate from 1982, was AbGRI1-1 (originally named Tn6166) (190). AbGRI1-1 is only 17.6 kb and has likely derived from an island related to RI_{MDR-TJ} as the *tetA(B)* in AbGRI1-1 is also missing 6 bp (190) (Fig 4.2C). AbGRI1-1 could arise from RI_{MDR-TJ} through recombination between Tn6022 Δ 1 and the full Tn6022. Interestingly, AbGRI1-1 containing Tn2006 in exactly the same position in the *sup* gene as in AbaR4, had been detected earlier in the isolate AB210 from the UK (260) as well as in isolates from Lithuania, where it was called AbaR4b (238). For clarity, I will refer to Tn2006 inserted in Tn6022 Δ 1 as AbaR4 Δ 1.

The work in this chapter focuses on determining the location of the *oxa23* gene in the chromosome of the 20 GC2 isolates from Singapore. Previous screening (Chapter 3; Section

3.3.3) had shown that the *oxa23* gene was located in Tn2006 in all isolates except S30. The Tn2006 containing isolates were screened to determine if this transposon was located in AbaR4 or AbaR4Δ1. As both AbaR4 and AbaR4Δ1 have been detected in AbGRI1, this island was assembled in all isolates. Further locations of transposons containing *oxa23* in the chromosome were also determined.

4.2 Screening for Tn6022 and AbaR4

4.2.1 Detection of Tn6022 and/or Tn6022Δ1

Tn2006 has been seen in both Tn6022 as AbaR4 or in Tn6022Δ1 as AbaR4Δ1. The 20 isolates were screened for Tn6022 and Tn6022Δ1 using the published primers RH910 and RH587 (191). These primers bind to either side of the 2.85 kb deletion in Tn6022Δ1 and distinguish these two forms based on size (Fig 4.3A). For a full Tn6022, a product 3,410 bp is amplified, while for Tn6022Δ1 the amplicon is only 560 bp. The majority of the isolates contain both a full Tn6022 and Tn6022Δ1 (Fig 4.3B), but S24, S26 and S29 contain only Tn6022 and S30 contains only Tn6022Δ1.

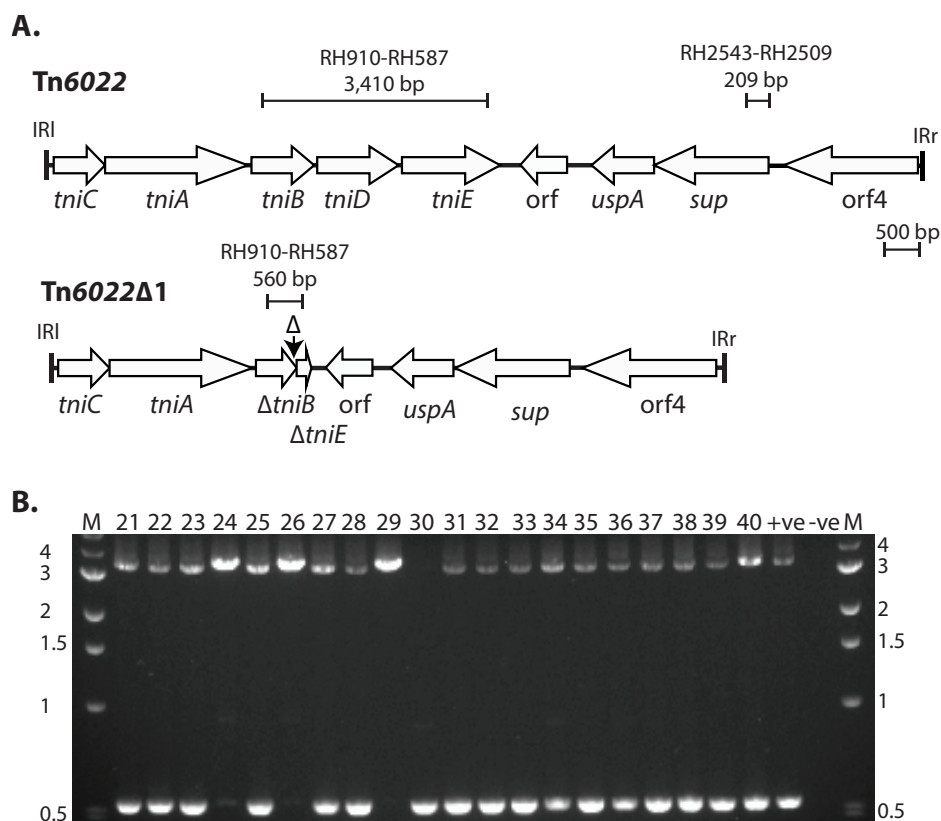


Figure 4.3. Screening for Tn6022 and/or Tn6022Δ1 A. Location of primers to detect Tn6022 and Tn6022Δ1. Arrows indicate extent and orientation of genes and open reading frames. Primers used to detect version of Tn6022 are shown by the bars above along with the PCR amplicon size. Vertical bars indicate IRs and the triangle indicates the position of the 2.85 kb deletion in Tn6022Δ1. B. The primers RH910 and RH587 were paired to amplify over a 3,410 bp region in Tn6022 and a 560 bp region in Tn6022Δ1. Isolate numbers are shown above the lanes and sizes of the molecular marker on the left and right are in kb.

4.2.2 Detection of AbaR4 or AbaR4Δ1

The isolates were screened for Tn2006 in AbaR4 or AbaR4Δ1 using a primer in Tn2006, RH743, with RH910 in the start of the *tniB* gene that is present in both Tn6022 and Tn6022Δ1 (Fig 4.4A). The isolates S21-S29 and S34, produced an amplicon between 9 to 10 kb, consistent with the 9,363 bp predicted amplicon size of Tn2006 in Tn6022 (Fig 4.4B, lanes 2 and 3). Hence, Tn2006 is located in AbaR4 in these isolates. However, in S31, S32, S33, S35, S36, S37 and S40, a 9-10 kb amplicon and an amplicon between 6 to 8 kb, were produced indicating that Tn2006 was present in both AbaR4 and AbaR4Δ1 were present (Fig 4.3B, lanes 4 and 5). An amplicon was not produced for S30, and this is consistent with the absence of Tn2006 in this isolate (Fig 4.4B, lane 1).

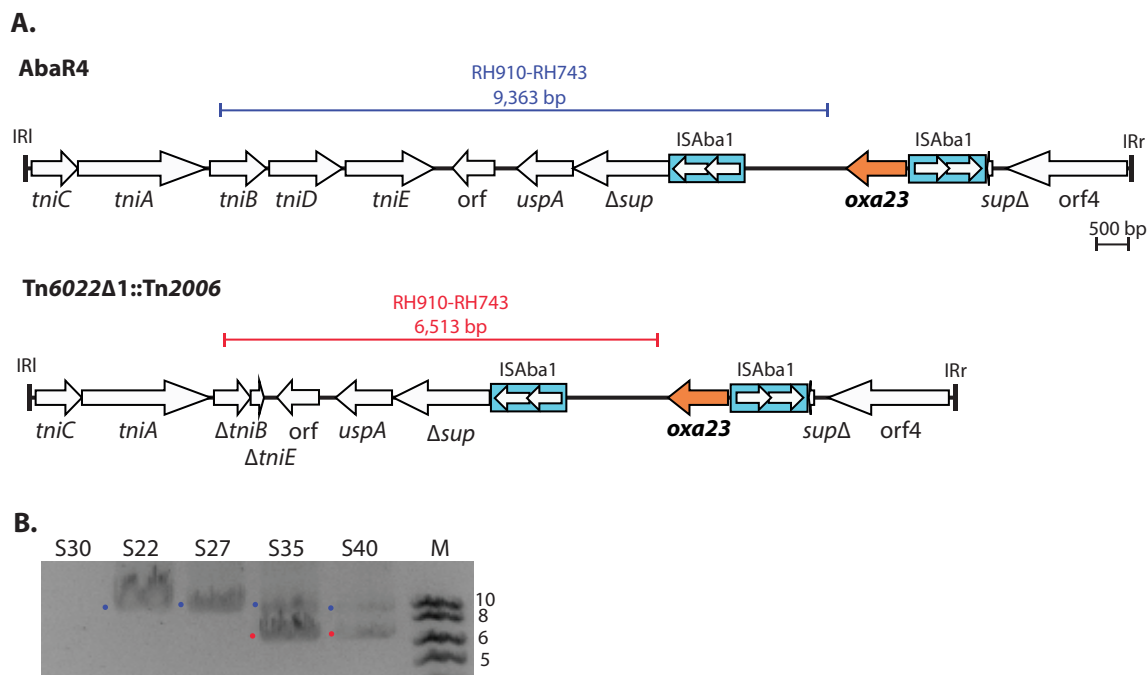


Figure 4.4 Screening for Tn2006 in AbaR4 and AbaR4 Δ 1. Arrows indicate extent and orientation of genes and open reading frames. The *oxa23* gene is orange. The light blue box represents ISAba1 and the internal arrows indicate the orientation of the transposase. Vertical bars indicate IRs. Primers used to detect Tn2006 in AbaR4 or AbaR4 Δ 1 are shown by the bars above along with the PCR amplicon size. B. The primers RH743 and RH910 were paired to amplify a 9,363 bp region in AbaR4 and a 6,513 bp region in AbaR4 Δ 1. Isolate numbers are shown above the lanes and sizes of the molecular marker on right are in kb. The blue and red dots indicate the amplicon for AbaR4 or AbaR4 Δ 1, respectively.

4.2.3 Is an intact *sup* gene present?

The 20 isolates were screened for an intact *sup* gene to determine if they contained a copy of Tn6022 or Tn6022 Δ 1 without Tn2006. A set of primers, RH2523 and RH2509 were designed in the *sup* gene, on either side of the insertion site of Tn2006. The isolates S21-S30, 34, S38 and S39 produced an amplicon of approximately 200 bp, consistent with the predicted product size of 209 bp for an intact *sup* gene (Table 4.1). Thus, these isolates contain at least one copy of Tn6022 or Tn6022 Δ 1 without Tn2006. However, S31-S33, S35-S37 and S40 failed to produce an amplicon, indicating that every copy of Tn6022 or Tn6022 Δ 1 in these isolates likely has Tn2006 inserted in the *sup* gene (Table 4.1).

The screening results are summarised in Table 4.1. Included in Table 4.1 is the presence of *sul2* gene in each isolate as this feature distinguishes AbGRI1-1 or RI_{MDR-TJ}. Six groups were

differentiated based on the pattern of their screening results, but only three of these groups included isolates that were sequenced by Illumina HiSeq that were available at the start of the project (isolates in bold, Table 4.1).

Table 4.1 Groups based on screening results

Group	Isolate	<i>sul2</i>	Tn2006	Tn6022	Tn6022Δ1	AbaR4	AbaR4Δ1	Intact <i>sup</i>
1	S21 , S22, S23, S25, S27 , S28, S34	+	+	+	+	+	-	+
2	S24 , S26, S29	+	+	+	-	+	-	+
3	S32, S33, S35, S36, S37, S40	+	+	+	+	+	+	-
4	S38 , S39	-	+	+	+	+	-	+
5	S31	-	+	+	+	+	+	-
6	S30	-	-	-	+	-	-	+

Isolates in bold has draft genome sequence available from the start of the project.

As both AbaR4 and AbaR4Δ1 had been detected in the AbGRI1 island of GC2 *A. baumannii*, this island was assembled in the 20 isolates either from HiSeq data or PCR mapping with later confirmation from MiSeq data.

4.3 AbGRI1 assembly

S30, S31, S38 and S39 do not contain the *sul2* gene and are likely to contain a resistance island resembling the shorter AbGRI1-1 (Fig 4.2C). The other isolates have the *sul2* gene and were more likely to resemble RI_{MDR-TJ}. The AbGRI1 regions in the seven isolates sequenced by HiSeq (those in bold in Table 4.1) were assembled using the appropriate scaffold of RI_{MDR-TJ} or AbGRI1-1. Through this process contigs indicating additional locations of AbaR4 were identified and their locations are further examined in Section 4.5.

4.3.1 AbGRI1-3 in S21 and AbGRI1-4 in S27 and S34

The draft genomes of the three group 1 isolates (Table 4.1), S21, S27 and S34, had 10 contigs that were retrieved using RI_{MDR-TJ} as query (Table 4.2). Eight of these contigs tiled to RI_{MDR-TJ}, spanning the entire region (Fig 4.5A). Two further contigs of all three isolates (listed in Table 4.2) had either the left or right IR of Tn6022 at one end, but the rest of the sequence matched the RepAci6 plasmid backbone. The increased coverage of the contig containing the *oxa23* gene in these isolates indicates that AbaR4 may be present in a RepAci6 plasmid and these contigs are further examined in Chapter 8.

Table 4.2 Contigs of S21, S27 and S34 matching RI_{MDR-TJ}

S21 contigs (bp)	S27 contigs (bp)	S34 contig (bp)	Content	Edges	
				L	R
17 (31,877)	27 (31,874)	46 (31,869)	LH flank	ND	IRI Tn6022
51 (3,281)	21 (3,282)	58 (3,282)	<i>tniC</i> , <i>tniA</i> , <i>tniB</i> Δ	IRI Tn6022	<i>tniB</i> Δ
34 (3,640)	12 (3,640)	45 (3,640)	Δ <i>tniE</i> , <i>uspA</i> , <i>sup</i> Δ	<i>tniE</i> Δ	<i>sup</i> Δ
11 (2,235)	10 (2,236)	33 (2,236)	Δ <i>sup</i> , <i>orf4</i>	<i>sup</i> Δ	IRr Tn6022
27 (9,581)	43 (9,569)	53 (9,578)	<i>orf</i> region, ΔTn6173	IRr Tn6022	ISAbal-IRr
105 (1,617)	88 (1,617)	104 (1,620)	<i>sul2</i> , CR2Δ	ISAbal-IRI	IRI Tn6022
66 (3,012)	17 (3,009)	59 (3,009)	Δ <i>tniB</i> , <i>tniD</i> , <i>tniE</i> Δ	Δ <i>tniB</i>	Δ <i>tniE</i>
52 (65,573)	46 (65,569)	73 (65,569)	<i>tetA</i> (B), <i>tetR</i> (B), CR2, <i>strB</i> , <i>strA</i> , <i>orf4b</i> , RH flank	IRr Tn6022	ND
10 (63,797)	50 (9,882)	22 (37,778)	RepAci6 plasmid backbone	ND	IRI Tn6022
38 (12,151)	19 (40,921)	8 (13,024)	RepAci6 plasmid backbone	IRr Tn6022	ND

The AbGRI1 region of S21, S27 and S34 was mapped using the linkage PCRs indicated by maroon bars in Fig 4.5A. The amplicons produced were of the expected size for the left end of RI_{MDR-TJ}. However, differences arose at the right end of the resistance island. When RH583 (located in *tniD*) was combined with a primer in *tetA*(B), *tetB*-F, only S21 produced an amplicon around 8 kb, consistent with the predicted size of 8,074 bp (orange bar in Fig 4.5A). Hence S21, had the same AbGRI1 island as RI_{MDR-TJ}. For consistency, this structure has been assigned an AbGRI1 number and is called AbGRI1-3.

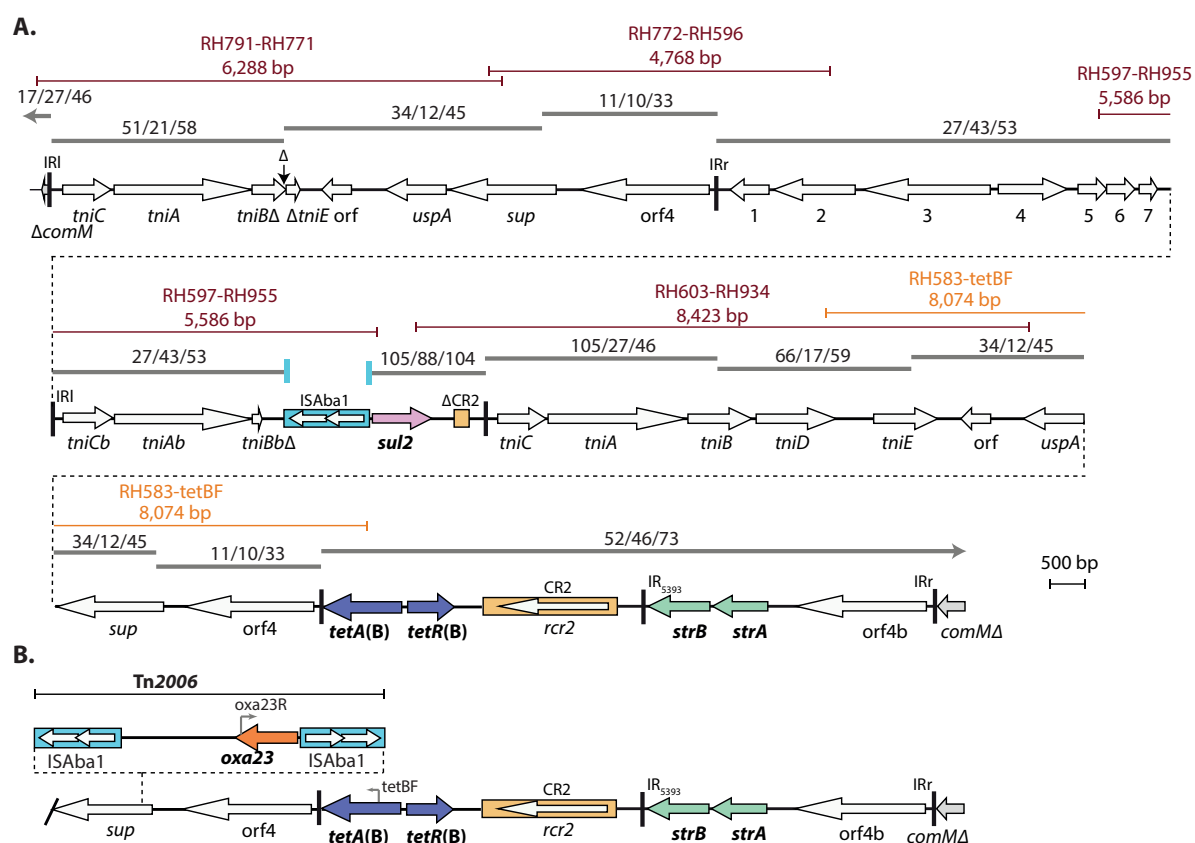


Figure 4.5 AbGRI1 mapping for S21, S27 and S34. A. AbGRI1 contigs in S21, S27 and S34 and PCRs used to assemble them. Arrows indicate extent and orientation of genes and open reading frames. The *sul2* gene is pink, the *tetA(B)* and *tetR(B)* genes are dark blue and the *strB* and *strA* genes are light green. Adjacent chromosomal genes are grey. The light blue box represents ISAbal1, the yellow boxes indicate CR2 elements and the internal arrows indicate the orientation of the transposase and rolling circle replicase, respectively. Vertical bars indicate IRs and the triangle indicates the position of the 2.85 kb deletion in Tn6022Δ1. Grey lines above indicate S21, S27 and S34 contigs and the blue bars indicate ISAbal1 fragments. The primers used to link sequences are shown above by the maroon bar and orange bars along with the predicted amplicon size. B. Tn2006 is inserted in the AbaR4 position in Tn6022 in S27 and S34. The *oxa23* gene in orange and grey arrows indicate the position of primers.

The isolates S27 and S34 failed to amplify a product when RH583 and tetBF were combined. However, they did produce an amplicon when RH743, an internal primer of Tn2006, was combined with tetBF (Fig 4.5B). Both isolates produced an amplicon between 5 and 6 kb, consistent with the predicted size of 5,205 bp. Sequence of the amplicon confirmed that Tn2006 was inserted directly adjacent to 5'-CCCGCAAT-3' which is the 9 bp TSD of AbaR4 (Fig 4.1). Hence, there is strong evidence that AbaR4 was present in S27 and S34 instead of the full copy of Tn6022 in the AbGRI1 of S21. However, primers linking Tn2006 to the *sul2* gene (RH743 and RH603) failed to amplify a product of the predicted size of 12,907 bp. This is not surprising

due to the predicted amplicon size, and sequencing S27 and S34 using long read technology such as PacBio or Oxford nanopore will be required to confirm the correct assembly of AbGRI1 in these isolates. The structure of AbGRI1 deduced from the genome data by PCR mapping of S27 and S34 is the same as that of one called AbaR24 (232). To avoid confusion with the AbaR-type islands in GC1 isolates, AbaR24 has been assigned AbGRI1-4.

4.3.2 AbGRI1-5 in S24 and S29

Nine and 10 contigs that had sequence matching AbGRI1-3 were retrieved from the draft genomes of S24 and S29, respectively (Table 4.3). These two isolates were in group 2 of Table 4.1. As for S21, S27 and S34, two of the contigs from S24 and S29 contained sequence matching RepAci6 plasmid backbone with IRI or IRr of Tn6022 at one end. Hence, it is likely that AbaR4 is located only in a RepAci6 plasmid in these isolates and the increased coverage of the *oxa23* containing contig supports this. These plasmids are examined further in Chapter 8.

Table 4.3 Contigs of S24 and S29 matching AbGRI1-3

S24 contigs (bp)	S29 contigs (bp)	Content	Edges	
			L	R
24 (31,874)	22 (31,862)	LH flank	ND	IRI Tn6022
138 (1,657)	91 (1,583)	<i>tniC</i> , <i>tniA</i> Δ	IRI Tn6022	<i>tniA</i> Δ
64 (1,677)	51 (1,665)	Δ <i>tniAb</i> , <i>tniBb</i> Δ	Δ <i>tniAb</i>	ISAbal-IRr
97 (1,618)	70 (1,604)	<i>sul2</i> , CR2Δ	ISAbal-IRI	IRI Tn6022
49 (8,114)	18 (8,188)	Δ <i>tniA</i> , <i>tniD</i> , <i>tniE</i> , <i>tniE</i> , <i>uspA</i> , <i>sup</i> Δ	Δ <i>tniA</i>	<i>sup</i> Δ
51(2,236)	92 (2,237)	Δ <i>sup</i> , orf4	Δ <i>sup</i>	IRr Tn6022
69 (65,570)	-	<i>tetA</i> (B), <i>tetR</i> (B), CR2, <i>strB</i> , <i>strA</i> , orf4b, RH flank	IRr Tn6022	ND
-	46 (65,569)	<i>tetA</i> (B), <i>tetR</i> (B), CR2, <i>strB</i> , <i>strA</i>	IRr Tn6022	ISAbal-IRI
-	6 (59,219)	orf4b, RH flank	ISAbal-IRr	ND
14 (9,864)	50 (9,882)	RepAci6 plasmid backbone	ND	IRI Tn6022
65 (1,752)	19 (40,921)	RepAci6 plasmid backbone	IRr Tn6022	ND

The remainder of the contigs of S24 and S29 in Table 4.3 tiled to AbGRI1-3 (Fig 4.6A). Consistent with the earlier screening (Table 4.1) there were no contigs for Tn6022Δ1. There were also no contigs matching the orf region or *tniCb* (Fig 4.6A). However, there was a smaller

contig 138/91, that at one edge had IRI of Tn6022 and $\Delta tniA$ at the other, and another contig 64/51, that contained sequence of $\Delta tniAb$, $tniBb\Delta$ and a fragment of ISAbal (Fig 4.6A).

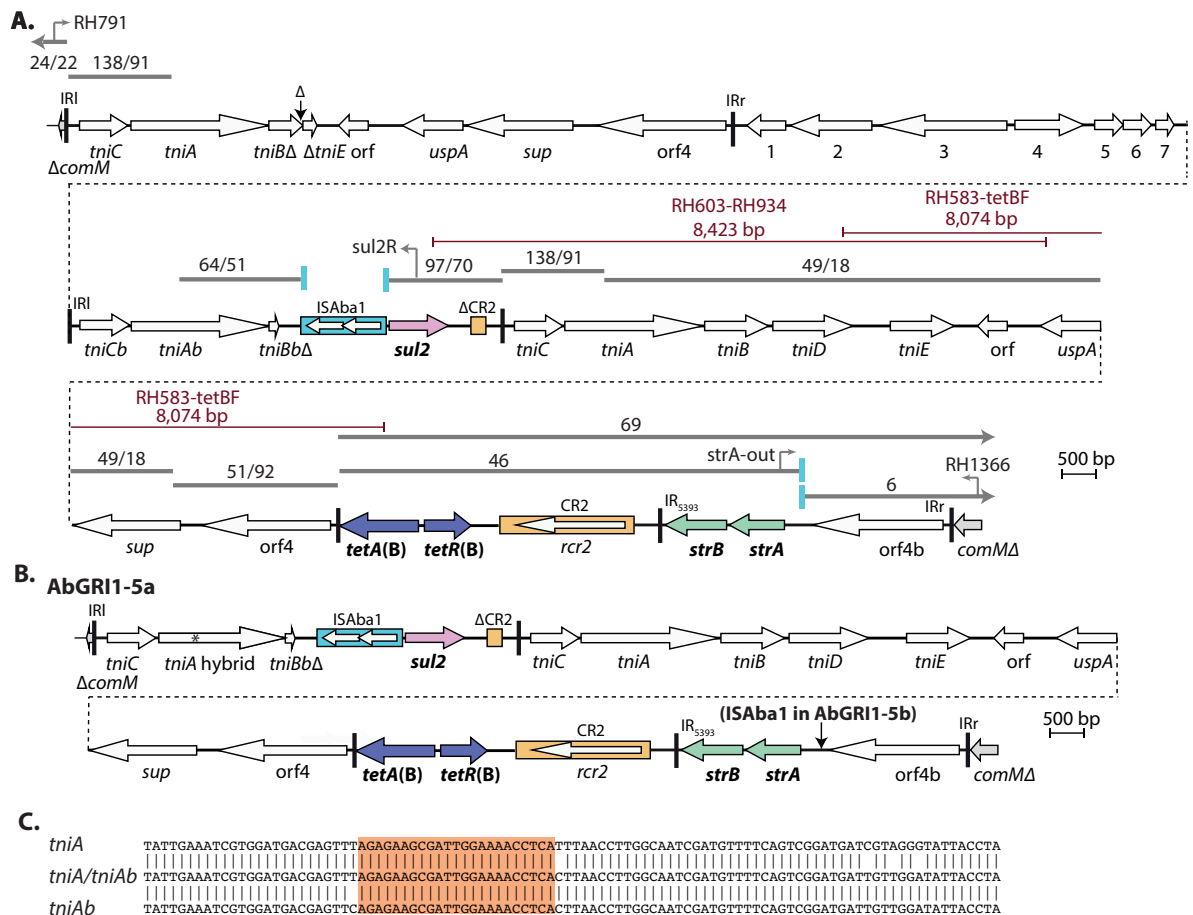


Figure 4.6 Assembly of AbGRI1-5 in S24 and S29. A. AbGRI1 contigs in S24 and S29. B. Structure of AbGRI1-5a. C. Site of recombination AbGRI1-5. In A. and B. arrows indicate extent and orientation of genes and open reading frames. The *sul2* gene is pink, the *tetA(B)* and *tetR(B)* genes are dark blue and the *strB* and *strA* genes are light green. Adjacent chromosomal genes are grey. The light blue box represents ISAbal, the yellow boxes indicate CR2 elements and the internal arrows indicate the orientation of the transposase and rolling circle replicase, respectively. Vertical bars indicate IRs and the triangle indicates the position of the 2.85 kb deletion in Tn6022Δ1. Grey lines above indicate S24 and S29 contigs and the blue bars indicate ISAbal fragments. The primers used to link sequences are shown above by the maroon bar and orange bars along with the predicted amplicon size. Some primer locations are indicated by grey arrows. The asterisk indicates the site of recombination. In C. the orange box indicates the 23 bp where recombination has occurred.

A primer in the sequence flanking AbGR1 on the left, RH791, combined with a primer in the *sul2* gene (Fig 4.6A), successfully amplified a 5 kb product, indicating that part of Tn6022Δ1, the orf region and a fragment of Δ Tn6173 had been lost at the left end of AbGRI1 in S24 and S29 (Fig 4.6B). As *tniA* (1,911 bp) and *tniAb* (1,911bp) share 94% nucleotide identity (1805/1912, 2 gaps), it is likely that recombination between these two genes has removed the

intervening sequence. Sequence of the amplicon revealed a 23 bp region where the recombination has occurred (orange box in Fig 4.6C) and created a hybrid *tniA* (1,911 bp) composed of at least the first 768 bp of *tniA* (Tn6022) and at least the last 1,120 bp of *tniAb* (Tn6179).

The mapping PCRs RH603 with RH934 (8,423 bp) and RH58 with tetBF (8,074 bp), confirmed that a full copy of Tn6022, without Tn2006, was located between the *sul2* gene and the *tetB* determinant in S24 and S29. These PCRs completed the assembly of the AbGRI1 region in S24 as the contig containing the *tet(B)* determinant (contig 69), also contained the remainder of the right end and junction of AbGRI1. This AbGRI1 structure has been named AbGRI1-5a (Fig 4.6B).

However, in S29, the *tetB* containing contig, contig 46, ends with a fragment of ISAbal just after the *strA* gene (Fig 4.6A). Contig 6 also has a fragment of ISAbal adjacent to the orf4b gene and includes sequence to the right of AbGRI1 (Fig 4.6A). The primer strA-out was combined with RH1366 in *comM* and amplified a product just above 4 kb, consistent with the predicted amplicon size of 4,122 bp if an ISAbal was present in this region in S29 (Fig 4.6B). The amplicon sequence confirmed the presence of an ISAbal inserted 132 bp upstream of the start codon of *strA* and was surrounded by the 9 bp TSD 5'-ATATTAATT-3'. The acquisition of this extra copy of ISAbal is the only difference between the AbGRI1 in S24 and S29, and thus the variant in S29 has been called AbGRI1-5b (Fig 4.6B).

4.3.3 AbGRI1-1 in S38 and S39

Eight contigs matching AbGRI1 were retrieved from both of the draft genomes of S38 and S39 (Table 4.4). None of these contigs contained the *sul2* gene or the orf region indicating that S38

and S39 contained AbGRI1-1 and five of the contigs tiled successfully to this island variant (Fig 4.7). However, two of the remaining contigs from these two isolates matched a segment of chromosome distant from AbGRI1 (Table 4.4) and had either IRL or IRr of Tn6022 at one end. The remaining additional contig, contig 48 in S38 and contig 16 in S39 contained the segment of Tn6022 missing in Tn6022Δ1. Each of these three contigs from S38 and S39 were examined further in Section 4.5.2.

Table 4.4 Contigs of S38 and S39 matching AbGRI1-1

S38 contigs (bp)	S39 contigs (bp)	Content	Edges	
			L	R
7 (31,878)	13 (31,874)	LH flank	ND	IRl Tn6022
57 (3,282)	37 (3,640)	<i>tniC</i> , <i>tniA</i> , <i>tniB</i> Δ	IRl Tn6022	<i>tniB</i> Δ
27 (3,640)	35 (3,282)	Δ <i>tniE</i> , <i>uspA</i> , <i>sup</i> Δ	Δ <i>tniE</i>	<i>sup</i> Δ
12 (2,236)	11 (2,236)	Δ <i>sup</i> , orf4	Δ <i>sup</i>	IRr Tn6022
48 (3,013)	16 (3,009)	Δ <i>tniB</i> , <i>tniD</i> , <i>tniE</i> Δ	Δ <i>tniB</i>	<i>tniE</i> Δ
11 (65,574)	8 (65,490)	<i>tetA</i> (B), <i>tetR</i> (B), CR2, <i>strB</i> , <i>strA</i> , orf4b, RH flank	IRr Tn6022	ND
20 (40,066)	18 (40,100)	Chromosome	ND	IRl Tn6022
31 (10,482)	27 (10,398)	Chromosome	IRr Tn6022	ND

The sequence to the left of AbGRI1 was linked to *sup* and produced an amplicon slightly larger than 6 kb, consistent with the predicted product size of 6,288 bp for a Tn6022Δ1 (Fig. 4.7). Hence, Tn6022Δ1 is at the left end of AbGRI1 in S38 and S39. Using the primers RH910 and tetBF (Fig 4.7), Tn6022Δ1 was linked to *tetA*(B), amplifying a product between 6 to 8 kb, consistent with the predicted size of 7,120 bp. Hence, S38 and S39 contained AbGRI1-1.

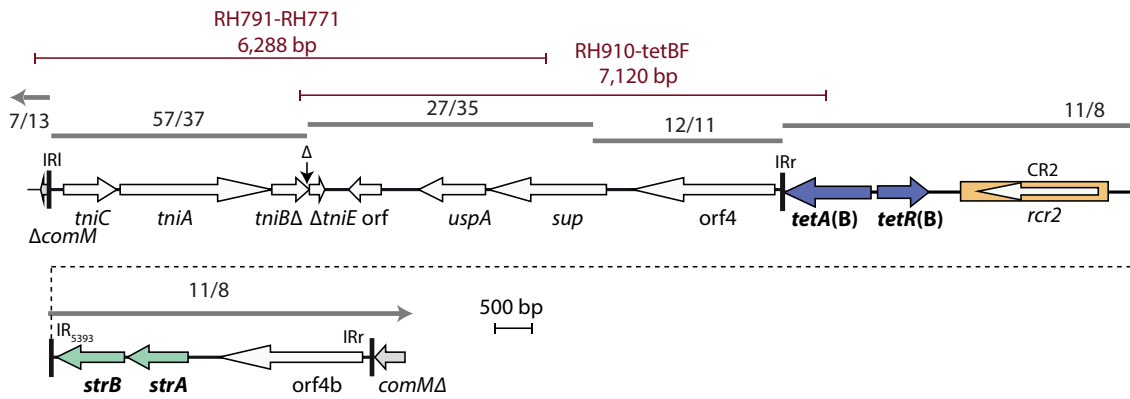


Figure 4.7 Assembly of AbGRII-1 in S38 and S39. Arrows indicate extent and orientation of genes and open reading frames. The *sul2* gene is pink, the *tetA(B)* and *tetR(B)* genes are dark blue and the *strB* and *strA* genes are light green. Adjacent chromosomal genes are grey. The light blue box represents ISAbal, the yellow boxes indicate CR2 elements and the internal arrows indicate the orientation the transposase and rolling circle replicase, respectively. Vertical bars indicate IRs and the triangle indicates the position of the 2.85 kb deletion in Tn6022Δ1. Grey lines above indicate S38 and S39 contigs. The primers used to link sequences are shown above by the maroon bars along with PCR amplicon size.

4.3.4 Prediction of AbGRII variant forms in isolates not sequenced by Illumina HiSeq

The mapping strategies used per Nigro *et al.* (191) and used to assemble the AbGRII variants in the isolates sequenced by Illumina HiSeq were applied to the 13 isolates without genome sequence data to predict their AbGRII structure. Two additional screening PCRs were used to confirm the structure at the right end of AbGRII (Fig 4.8). The primers tetBF and strBout (4,616 bp) and strAout and RH1366 (2,942 bp), were used to link the *tet(B)* determinant to *strB* and *strA* to the chromosome to the right of AbGRII, respectively (Fig 4.8) (191). Whether the *sul2* gene was present was also considered (Table 4.1).

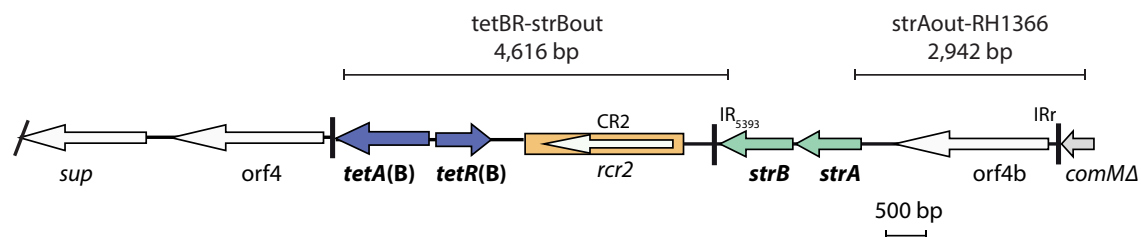


Figure 4.8 Screening PCRs for the right end of AbGRII. Arrows indicate extent and orientation of genes and open reading frames. The *tetA(B)* and *tetR(B)* genes are dark blue and the *strB* and *strA* genes are light green. Adjacent chromosomal genes are grey. CR2 is shown by the yellow box and the internal arrow indicates the rolling circle replicase. Vertical bars indicate IRs. The primers used to link sequences are shown above by black bars along with the predicted amplicon size.

4.3.4.1 Isolates without the *sul2* gene

Both S30 (group 6) and S31 (group 5) did not contain the *sul2* gene and so the mapping strategies for AbGRI1-1 were applied to these isolates (Table 4.5). The amplicons produced for S30 perfectly matched the predicted sizes for AbGRI1-1 and hence this isolate contained this version of AbGRI1. However, for S31, an amplicon matching the predicted size was produced for all primer pairs, except for RH910 and tetBF that link *tniB* to *tetA(B)*. From previous screening, S31 does not contain an intact *sup* gene, and has a copy of AbaR4Δ1. A 5 kb amplicon for oxa23R with tetBF, linking Tn2006 to *tetA(B)*, was produced and was consistent with the predicted size of 5,016 bp if AbaR4Δ1 is present. Hence, AbGRI1 in S31 differed to AbGRI1-1 by the presence of Tn2006 in AbaR4Δ1. The structure in S31 is identical to that of AB210 (Fig 4.2C) (260) and this version of AbGRI1 has been named AbGRI1-6.

Table 4.5 Mapping AbGRI1 for the isolates without *sul2*

Isolate	LH flank- <i>sup</i> (Tn6022Δ1)	<i>tniB</i> - <i>tetA(B)</i> (Tn6022Δ1)	Tn2006- <i>tetA(B)</i>	<i>tetA(B)</i> - <i>strB</i>	<i>strA</i> - RH flank	AbGRI1 version
	RH791- RH771 (6,288 bp)	RH910- tetBF (7,120 bp)	oxa23R- tetBF (5,016 bp)	tetBR- strBout (4,616 bp)	strAout- RH1366 (2,942 bp)	
S30	+	+	ND	+	+	1
S31	+	-	+	+	+	6

4.3.4.2 Isolates containing the *sul2* gene

The mapping strategies attempted on the isolates with *sul2* are shown in Table 4.6. Previous screening had revealed that S26 did not contain Tn6022Δ1 (Fig 4.3), so the failure of this isolate to produce an amplicon linking the left flank of AbGRI1 to the *sup* gene (RH791-RH771) was not surprising (Table 4.6). S26 belonged to the same group (group 2, Table 4.1) as S24 and S29 which contained AbGRI2-5 and so S26 was screened for this structure. S26 gave a product round 5 kb for RH791 (*comM* gene) and sul2R (*sul2* gene), consistent with the predicted amplicon size for recombination between *tniA* and *tniAb* present in AbGRI1-5. A copy of

Table 4.6 Mapping AbGRI1 for isolates with *sul2*

Isolate	LH flank- <i>sup</i> (Tn6022ΔI)	LH flank- <i>sul2</i>	<i>sup</i> -orf region	Tn2006- orf region	orf region- <i>tniCb</i>	ΔTn6173- <i>sul2</i>	<i>sul2-uspA</i> (Tn6022)	<i>tniD</i> - <i>tetA</i> (B)	<i>oxa23</i> - <i>tetA</i> (B)	<i>tetA</i> (B)- <i>strB</i>	<i>strA</i> -RH flank	AbGRI1 version
	RH791- RH771 (6,288 bp)	RH791- <i>sul2</i> R (5,099 bp)	RH772- RH596 (4,768 bp)	<i>oxa23</i> R- RH596 (7,720 bp)	RH792- RH597 (1,566 bp)	RH597- RH955 (5,596 bp)	RH603- RH934 (8,423 bp)	RH583- tetBF (8,074 bp)	<i>oxa23</i> R- tetBF (5,019 bp)	tetBR- strBout (4,616 bp)	strAout- RH1366 (2,942 bp)	
S26	-	+	ND	ND	ND	ND	+	+	ND	+	+ ¹	5
S22	+	ND	+	ND	+	+	+	+	ND	+	+	3
S23	+	ND	+	ND	+	+	+	+	ND	+	+	3
S25	+	ND	+	ND	+	+	+	+	ND	+	+	3
S28	+	ND	+	ND	+	+	+	-	-	+	-	?
S32	+	ND	-	+	+	+	+	-	+	+	+	7
S33	+	ND	-	+	+	+	+	-	+	+	+	7
S35	+	ND	-	+	+	+	+	-	+	+	+	7
S36	+	ND	-	+	+	+	+	-	+	+	+	7
S37	+	ND	-	+	+	+	+	-	+	+	+	7
S40	+	ND	-	+	+	+	+	-	+	+	+	7

ND denotes that this arrangement was not tested as conflicting arrangements had already been confirmed in this isolate.

¹Product was approximately 1 kb larger than the predicted size.

ISAba1 was present between *strA* and *orf4b*, indicating that S26 likely contained AbGRI1-5b (Table 4.6).

Three of the four remaining group 1 isolates (Table 4.1), S22, S23 and S25, produced amplicons consistent with AbGRI1-3 (Table 4.6). The fourth isolate, S28, failed to produce amplicons of the predicted sizes for two of the assembly PCRs for AbGRI1-3 (Table 4.6).

These were linking Tn6022 to *tetA*(B) (RH583 and *tetB*F) and linking *strA* to the right flanking sequence of AbGRI1 (*strA*out with RH1366). The structure present S28 was examined further in Section 4.4.1 after the draft genome became available.

4.3.4.2.1 AbGRI1-7

The six remaining isolates in Table 4.6 were all group 3 in Table 4.1 and were found to contain a novel AbGRI1 structure AbGRI1-7 (Fig 4.9). Like AbGRI1-4, AbaR4 instead of the full length Tn6022 was inserted in the end of the *tetA*(B) gene. However, AbGRI1-7 harboured AbaR4Δ1 instead of Tn6022Δ1 at the left end of the island. The presence of AbGRI1-7 in these 6 isolates is consistent with results of earlier screening including the presence of both AbaR4 and AbaR4Δ1, as well as not containing a full *sup* gene (Table 4.1).

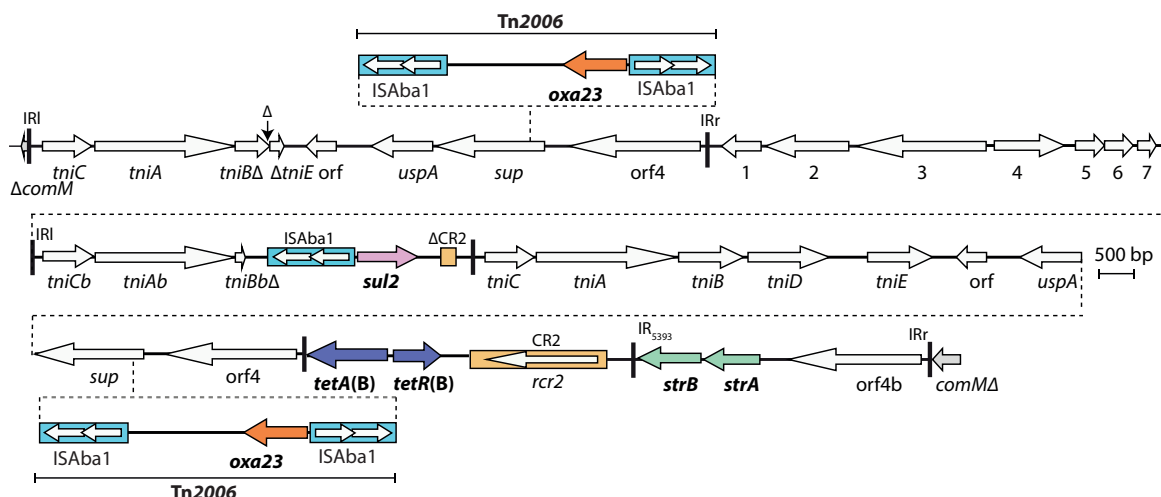


Figure 4.9 Structure of AbGRII-7. Arrows indicate extent and orientation of genes and open reading frames. The *oxa23* gene is orange, the *sul2* gene is pink, the *tetA(B)* and *tetR(B)* genes are dark blue and the *strB* and *strA* genes are light green. Adjacent chromosomal genes are grey. The light blue boxes represent ISAbal1, the yellow box indicates CR2 elements and the internal arrows represent the transposase and rolling circle replicase, respectively. Vertical bars indicate IRs and the triangle indicates the position of the 2.85 kb deletion in Tn6022Δ1.

4.4 Examination of the draft genomes from Illumina MiSeq sequencing

When the draft genome sequences of the isolates in Table 4.5 and Table 4.6 became available at the end of 2016, they were compared to the AbGRII regions assembled by PCR mapping. The draft genomes of S22, S23, S25, S26, S30, S31 S32, S33, S35, S36, S37 and S40 were consistent with predicted AbGRII structures. Additional information from the draft genome of S28 facilitated further assembly of the AbGRII region in this isolate (described in Section 4.4.1).

Analysis of the draft genomes of several isolates revealed additional copies of AbaR4. S31 contained two contigs matching chromosomal sequence, similar to those in S38 and S39 (Table 4.4), that had the IRs of Tn6022 at one end and these are examined in further detail in Section 4.5.2. The draft genomes of S22 and S23 had two additional chromosomal contigs that contained the IRL and IRR of Tn6022 at one end and these are examined in more detail in Section 4.5.1. The draft genomes of S22, S23, S25, S28, and S40 each included a pair contigs matching RepAci6 plasmids that had a copy of IRL or IRR of Tn6022 at one end, and these contigs are

further analysed in Chapter 8. The *oxa23* gene in Tn2008 was located on a contig containing sequence matching RepAci6 plasmid in the draft genome of S30 (Chapter 8).

4.4.1 AbGRI1-8 in S28

The contigs of the draft genome of S28 indicated that Tn2006 had inserted in the intergenic region between *strA* and *orf4b*, accounting for the amplification failure using the primers *strA* out and RH1366 (Table 4.6). Primers that anneal to the internal segment of Tn2006 were combined with *strA*-out and RH1366 to orientate the transposon (Fig 4.10). Amplicons of the expected size were amplified when *oxa23*R and *strA*-out (predicted amplicon size of 2,517 bp), and RH743 and RH1366 (5,023 bp predicted amplicon size) were paired. A 9 bp target site duplication of 5'-AAAAAAGG-3' surrounded Tn2006. As this is a new Tn2006 insertion position, the structure of AbGRI1 in S28 is assigned AbGRI1-8.

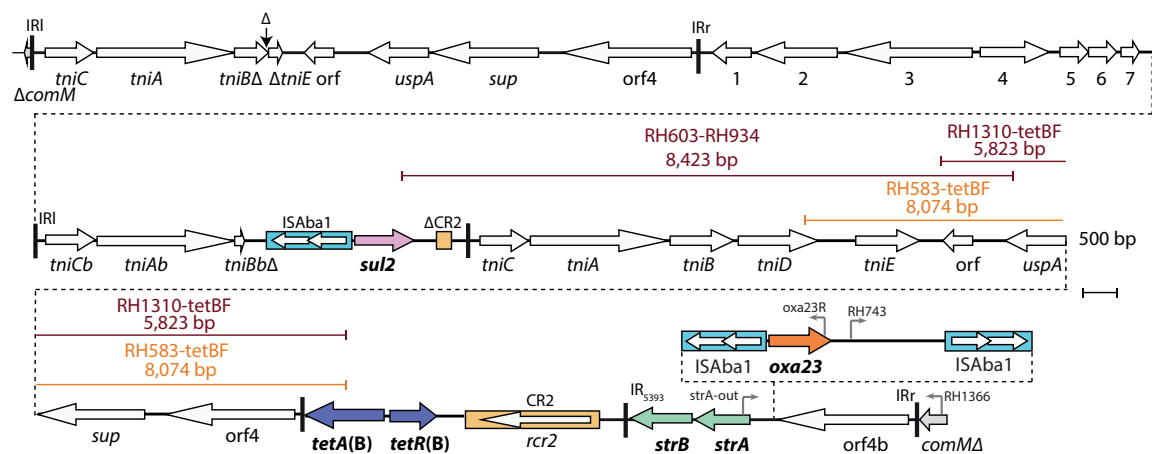


Figure 4.10 Structure of AbGRI1-8 in S28. Arrows indicate extent and orientation of genes and open reading frames. The *oxa23* gene is orange, the *sul2* gene is pink, the *tetA(B)* and *tetR(B)* genes are dark blue and the *strB* and *strA* genes are light green. Adjacent chromosomal genes are grey. The light blue box represents ISAbal, the yellow boxes indicate CR2 elements and the internal arrows indicate the orientation of the transposase and rolling circle replicase, respectively. Vertical bars indicate IRs and the triangle indicates the position of the 2.85 kb deletion in Tn6022Δ1. The primers used to link sequences are shown above by the maroon and orange bars along with the PCR amplicon size. The binding sites of some primers are indicated by grey arrows.

The draft genome of S28 did not provide any indication of why the primer combination of RH583 and tetBF failed to amplify a product (orange bars in Fig 4.10) (Table 4.6). However,

when RH1310 and tetBF were paired, a product of the predicted size was amplified (maroon bars in Fig 4.10). As the region spanned by this PCR, overlaps with the amplicon produced using RH603 and RH934 (maroon bars Fig 4.10), it was assumed that a full Tn6022 was inserted at the end of *tetA*(B) and the failure of RH583 to tetBF to amplify was not examined further. Sequencing S28 using Oxford nanopore or PacBio could confirm the structure of AbGRI1-8 in this isolate. AbGRI1-8 has the same structure as AbGRI1-3, but with a new insertion of Tn2006.

4.4.2 AbGRI1 variants

The AbGRI1 variants assembled in the 20 isolates are listed in Table 4.7 and compared in Fig 4.11. All of the variants detected can be derived from AbGRI1-3 either by addition of Tn2006 or through deletions of segments by recombination. In fact, the only difference between AbGRI1-3, AbGRI1-4, AbGRI1-7 and AbGRI1-8 is presence of Tn2006 in different locations (Fig 4.11, Table 4.7). Homologous recombination between *tniA* and *tniAb* has removed the intervening sequence in AbGRI1-5. AbGRI1-1 and AbGRI1-6 have lost a large internal segment of AbGRI1-3, likely through recombination between Tn6022Δ1 and Tn6022, but AbGRI1-6 also contains Tn2006 (Table 4.7).

Table 4.7 AbGRI1 variants detected

Isolates	AbGRI1 version	<i>oxa23</i>	<i>oxa23</i> location
S21, S22, S23, S25, S27, S34	3	-	
S32, S33, S35, S36, S37, S40	4	+	AbaR4
S28	7	+	AbaR4Δ1, AbaR4
S24	8	+	Tn2006
S26, S29	5a	-	
S30, S38, S39	5b	-	
S31	1	-	
	6	+	AbaR4Δ1

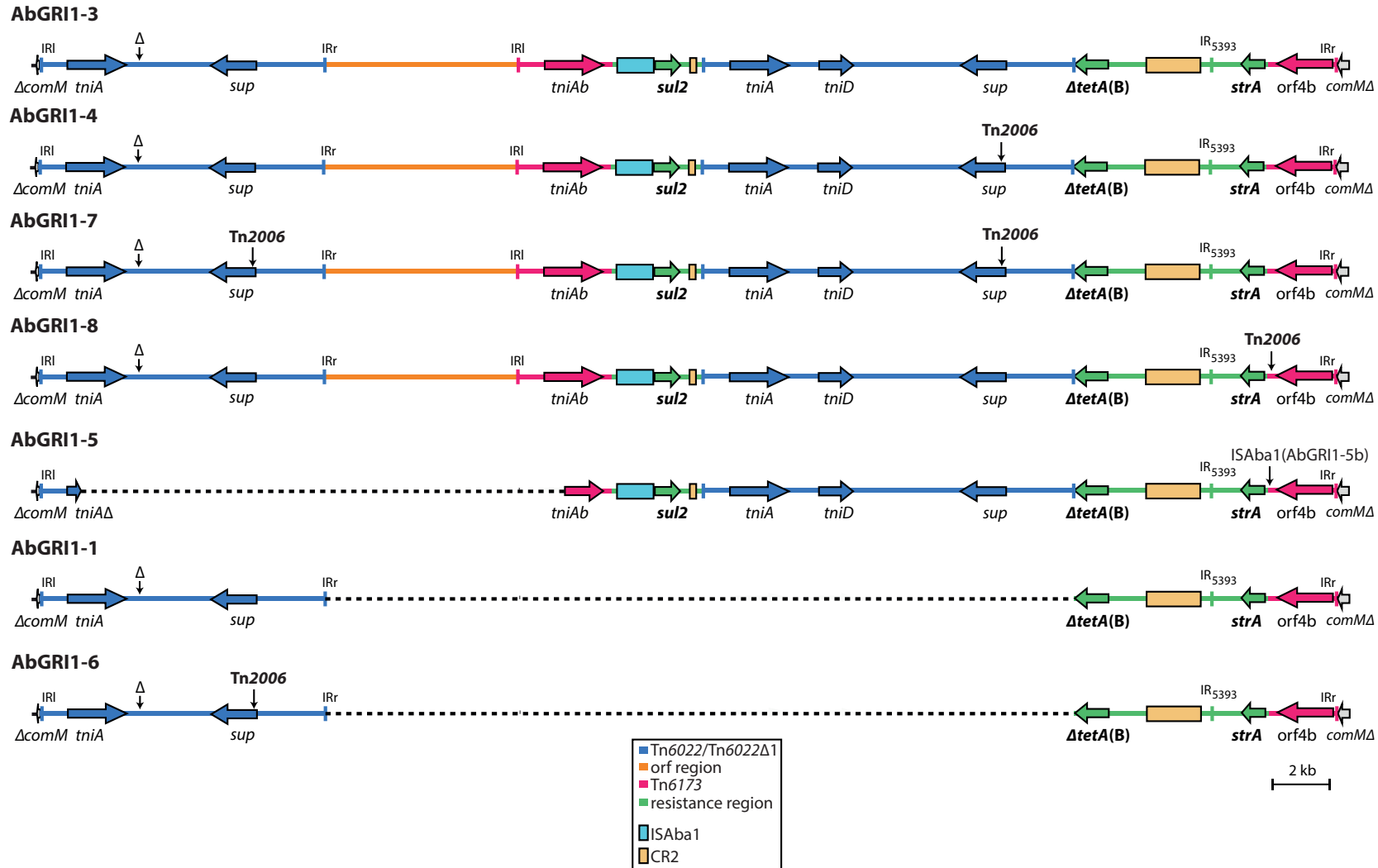


Figure 4.11 AbGRI1 variants. Coloured lines indicate different segments of AbGRI1 as shown the key. Labelled arrows above the line indicate insertion position of Tn2006 and ISAbal and the triangle indicates the position of the 2.85 kb deletion. Dashed lines indicate missing segments.

4.5 Additional chromosomal locations of *oxa23* containing transposons

The coverage of the contig containing *oxa23* in some isolates indicated that multiple copies of this gene may present and/or that they may be located in plasmids that have a higher sequence coverage. These draft genomes contained additional contigs that indicated that *oxa23* containing transposons may be located elsewhere in the genome. A subset of these contigs matched RepAci6 plasmids and these will be addressed in Chapter 8. The work in this section is focused on finding the location of *oxa23* containing transposons Tn2006 or AbaR4 in the chromosome at locations outside of AbGRI1.

4.5.1 AbaR4 in S22 and S23

Examination of the draft genomes of S22 and S23 revealed that both of these isolates contained an additional contig with the IRL of Tn6022 at one end (contig 39 in S22 and contig 44 in S23) and another contig with IRr of Tn6022 at one edge (contig 17 and contig 24 in S22 and S23, respectively). These contigs matched chromosomal sequence present in the completed genome of XH386 (GC2 isolate that is ST208 in the Oxford scheme and sequenced by PacBio, GenBank accession number CP010779) and hence were investigated as a chromosomal location of AbaR4. The presence of AbaR4 was confirmed by PCR mapping as indicated in Fig 4.12A. AbaR4 in S22 and S23 was not surrounded by a 5 bp TSD and compared to XH386, a 26.3 kb region was missing (Fig 4.12A). The approximate position of AbaR4 is indicated in a schematic of the chromosome of XH386 (Fig 4.12B).

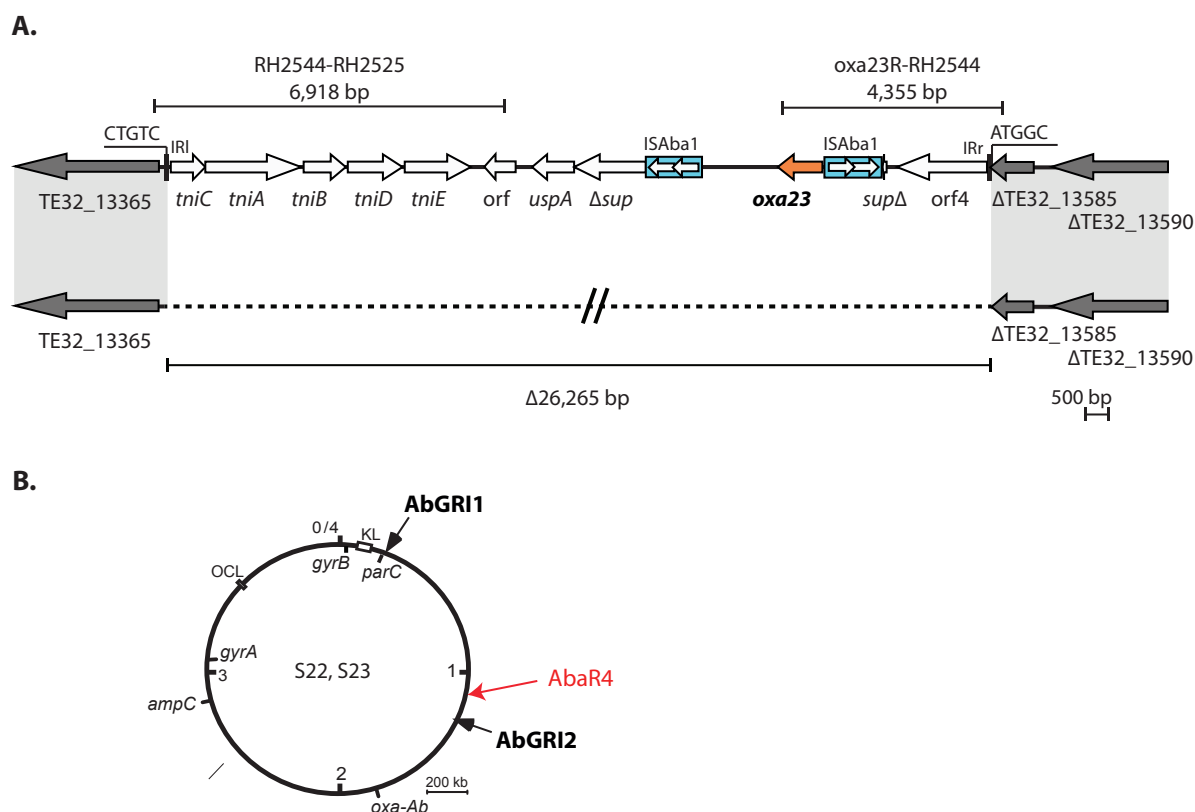


Figure 4.12 Location of AbaR4 in the chromosome of S22 and S23. Assembly and location of AbaR4 in S22 and S23. Arrows indicate extent and orientation of genes and open reading frames. The *xox23* gene is orange and adjacent chromosomal genes are grey. ISAbal are shown as light blue boxes and the internal arrows indicate the orientation of the transposase. Vertical bars indicate IRs. Locus tags of XH386 (GenBank accession number CP010779) are provided. B. Location of AbaR4 on a schematic of the chromosome. Arrows indicate the approximate location of resistance islands.

4.5.2 AbaR4 in S31, S38 and S39

The two additional contigs of S38 and S39 that contained IRl or IRr at one edge (contig 20 and 31 of S38 and 18 and 27 of S39) indicated that these isolates may also contain a copy of AbaR4 in the chromosome. The draft genome of S31 also included similar contigs. PCR mapping (indicated in Fig 4.13A) confirmed the presence of AbaR4, surrounded a 5 bp TSD 5'-CACAG-3' and inserted in the intergenic region between TE32_00030 and TE32_00035 in XH386 just prior to the *gyrB* gene in these three isolates (Fig 4.13B).

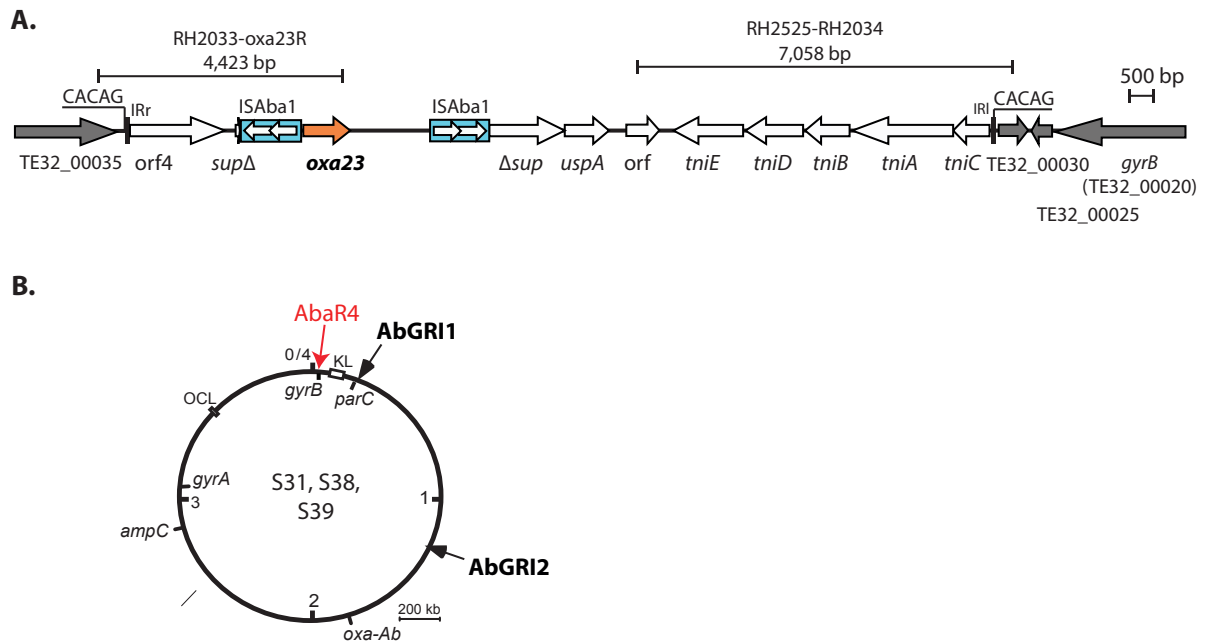


Figure 4.13 Location of AbaR4 in the chromosome of S31, S38 and S39. Assembly and location of AbaR4 in S31, S38 and S39. Arrows indicate extent and orientation of genes and open reading frames. The *oxa23* gene is orange and adjacent chromosomal genes are grey. ISAbal are shown as light blue boxes and the internal arrows indicate the orientation of the transposase. Vertical bars indicate IRs. Locus tags of XH386 (GenBank accession number CP010779) are provided. B. Location of AbaR4 on a schematic of the chromosome. Arrows indicate the approximate location of resistance islands.

4.5.3 Tn2006 in S34

Bioinformatic pairing of contigs containing fragments of ISAbal with matching 9 bp target site duplication from the draft genome of S34 revealed two contigs (contig 21 and contig 148) containing chromosomal sequence that had the same TSD (5'-CGCAAACAC-3') adjacent to a fragment of ISAbal, but both fragments of ISAbal were from the 5'-end of this IS. This is indicative of Tn2006 and PCR mapping confirmed that it was present at this position (Fig 4.14A). The sequence of the chromosome surrounding this transposon in S34 was found only in a subset of GC1 and GC2 isolates and was present in the complete genome of A85, a GC1 sequenced by PacBio (GenBank accession CP021782). Tn2006 is inserted between two open reading frames CBI29_00939 and CBI29_0090 of A85 and the approximate position of this transposon is indicated in Fig 4.14B.

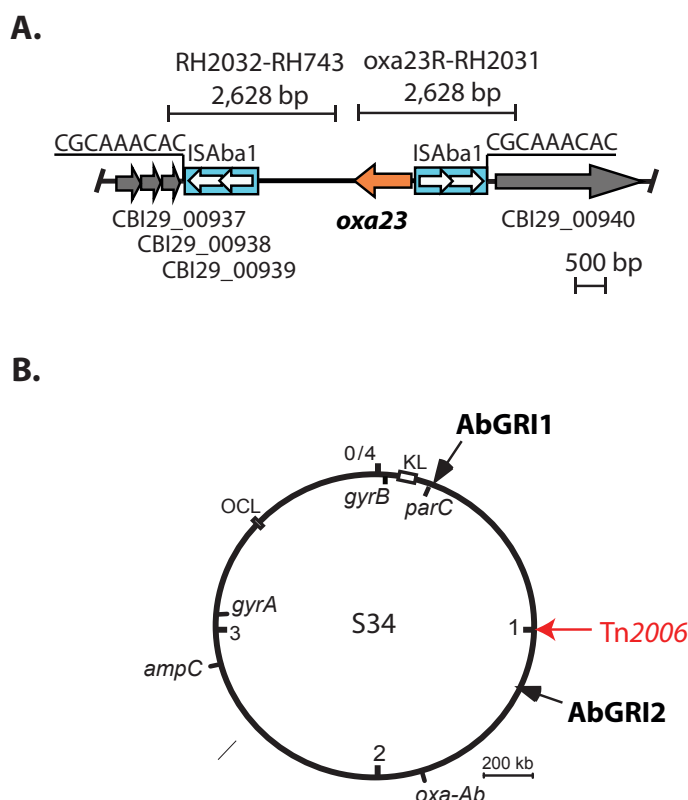


Figure 4.14 Location of Tn2006 in the chromosome of S34. Assembly and location of Tn2006 in S34. Arrows indicate extent and orientation of genes and open reading frames. The *oxa23* gene is orange and adjacent chromosomal genes are grey. ISAbal are shown as light blue boxes and the internal arrows represent the transposase. Vertical bars indicate IRs. Locus tags of A85 (GenBank accession number CP021782) are provided. B. Tn2006 location on a schematic of the chromosome. Figure made using genome sequence of A85 and the arrows indicate the approximate location the GC2 resistance islands, AbGRI1, AbGRI2 and AbGRI3, and Tn2006.

4.6 Discussion

A version of AbGRI1 was detected in each of the 20 isolates examined and 7 different variants of this island were characterised (Table 4.8). A number of these variants matched those that had been published previously AbGRI1-1, AbGRI1-3, AbGRI1-4 and AbGRI1-6, but the later three island structures had been called other names. These variants were renamed here for consistency and to distinguish them from the unrelated AbaR islands and Tn6019 backbone. However, the variants AbGRI1-5, AbGRI1-7 and AbGRI1-8 were first characterised in this project. AbGRI1 variants -1 and -3 were found in both isolates containing *armA* and those where *armA* is absent (bold in Table 4.8).

Table 4.8 Location of *oxa23* containing transposons

Isolate	AbGRI1 version	Context of <i>oxa23</i> in AbGRI1	Additional chromosomal locations	RepAci6 plasmid?
S21 , S25	3			AbaR4
S22, S23	3		AbaR4 (1)	AbaR4
S27	4	AbaR4		AbaR4
S34	4	AbaR4	Tn2006	AbaR4
S24, S26, S29	5			AbaR4
S28	8	Tn2006		AbaR4
S31	6	AbaR4Δ1	AbaR4 (2)	
S38, S39	1		AbaR4 (2)	
S32, S33, S35, S36, S37	7	AbaR4, AbaR4Δ1		
S40	7	AbaR4, AbaR4Δ1		AbaR4
S30	1			Tn2008

Those in bold are the five isolates without the *armA* gene.

The first three isolates, S21, S22 and S23 contain AbGRI1-3. The only difference between this variant and AbGRI1-4, AbGRI1-7 and AbGRI1-8 is the addition of the *oxa23* containing transposon Tn2006. AbGRI1-4 and AbGRI1-7 have acquired AbaR4 instead of the full length Tn6022 and AbGRI1-7 also contains AbaR4Δ1 instead of Tn6022Δ1. The switch between Tn6022 and AbaR4 has likely arisen by recombination as these transposons are located in the same position within the *tetA(B)* gene. As it is unlikely that Tn2006 has jumped into the same position in the *sup* gene of both AbaR4Δ1 and AbaR4, it is probable that recombination also accounts for the formation of AbaR4Δ1. However, AbGRI1-8 differs to AbGRI1-3 by a new acquisition of Tn2006. Tn2006 in AbGRI1-8 is in a different location to AbGRI1-2 (191) and represents a second transposition event of Tn2006 into AbGRI1.

Recombination in the AbGRI1-3 structure between *tniA* of Tn6022Δ1 and *tniAb* of Tn6173 accounts for the segment missing from AbGRI1-5. An AbGRI1 structure with the same recombination junction has been detected some Australian isolates from Melbourne and Adelaide that have different characteristics to those originating from hospitals from the East-coast of Australia (Dr Steven Nigro, personal communication). AbGRI1-1 is likely derived

from AbGRI1-3 by recombination between Tn6022Δ1 and Tn6022. AbGRI1-6 differs from AbGRI1-3 as AbaR4Δ1 instead of Tn6022Δ1 is present.

In addition, two other chromosomal locations for AbaR4 were detected, as was another insertion of Tn2006. Isolates that share the same chromosomal AbaR4 location are likely closely-related. There was also evidence from the draft genomes that 8 of the isolates may also harbour an RepAci6 plasmid carrying AbaR4 or Tn2008 (Table 4.8). The AbaR4 transposon has previously been detected on pA85-3 (97) and Tn2008 has been seen in other RepAci6 plasmids (277).

All isolates carry at least one copy of the *oxa23* gene, while most contain two or even three copies of this gene. A recent publication has shown that an increase in the copy number of *oxa23* does not appear to enhance carbapenem resistance (124).

CHAPTER

5

AbGRI2

5.1 Introduction

The AbGRI2 resistance island is bounded by copies of IS26 and contains variable resistance gene content. The first variant of this island, AbGRI2-1, was characterised in the Australian isolates WM99c (1999) and A91 (2004). AbGRI2-1 is 19.5 kb and composed of fragments from a number of transposons carrying resistance genes along with three internal copies of IS26 (second line in Fig 5.1). AbGRI2-1 contains the *aphA1b* gene (kanamycin resistance) in Tn6020 (217), *bla*_{TEM} (ampicillin resistance) and a class 1 integron with cassettes *aacC1*-orfP-orfP-orfQ-*aadA1* (conferring resistance to gentamicin, and spectinomycin and streptomycin) and *sulI* (sulphonamide resistance) in the 3'-CS (188). Two other variants of AbGRI2 were identified in this work, AbGRI2-2 from MDR-TJ and AbGRI2-3 found in MDR-ZJ06 (188). Both of these variants could have arisen from AbGRI2-1 via IS26 mediated deletions, in addition to cassette switching for AbGRI2-3.

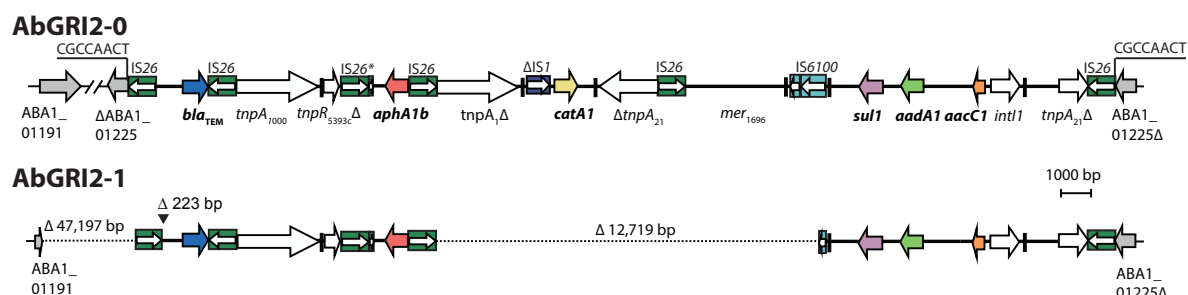


Figure 5.1 Structure of AbGRI2-0 and AbGRI2-1. Arrows indicate the extent and orientation of genes and open reading frames. The *bla*_{TEM} gene is blue, the *aphA1b* gene is red-orange, the *sulI* gene is mauve and the *aadA1* and *aacC1* genes are light green and orange, respectively. Adjacent chromosomal genes are grey. The dark green box represents IS26 and the light blue box is IS6100, with internal arrows indicating the transposase orientation. IS26* is 3 bp different to the standard sequence of IS26. Vertical bars indicate IRs and the TSD surrounding AbGRI2-0 is shown above. Dashed lines represent the extent of deletions relative to AbGRI2-0 and the size is indicated above. Figure made to scale from GenBank accession number JN247441 for AbGRI2-0 and GenBank accession number JX869489 for AbGRI2-1 and has been adapted from Blackwell *et al.* (24)

Recently, the AbGRI2 region of A320 (RUH134), the GC2 type strain from 1982, was assembled (24). The resistance island in A320 had been separated into two pieces, AbGRI2-0a and AbGRI2-0b, by an IS26 mediated inversion. Reversal of this inversion *in silico* revealed a

larger 32.4 kb resistance island, AbGRI2-0 (24) (top line Fig 5.1). As AbGRI2-0 was surrounded by an 8 bp TSD, the original location of this resistance island in the gene corresponding to ABA1_01225 in A1 (GenBank accession number CP010781) was revealed. The GC1 chromosome of A1 (isolated in 1982 and sequenced by PacBio) was used as a reference genome for this region as all GC2 genomes to date contain deletions of various sizes in this location.

In addition to the genes in AbGRI2-1, AbGRI2-0 contained the *catA1* chloramphenicol resistance gene and a mercury resistance region from Tn1696 (24). A comparison revealed that AbGRI2-0 is ancestral to AbGRI2-1 and AbGRI2-1 can be derived from AbGRI2-0 through a small number of events (Fig 5.1). An internal IS26 mediated deletion has removed 12.7 kb of AbGRI2-0, including the *catA1* gene and the mercury resistance module. The IS26 bounding the left end of AbGRI2-1 has switched orientation and 223 bp of AbGRI2-0 sequence next to it is missing. AbGRI2-1 retains the original right boundary of AbGRI2-0 but 47,197 bp is missing from the chromosome flanking the left of AbGRI2-1 (Fig 5.1). The events that have resulted in the configuration at the left end of AbGRI2-1 are quite specific, and most of the reported AbGRI2 structures found in isolates from Asia and America have an AbGRI2 island derived from AbGRI2-1 (24, 193, 277).

The work in this chapter focused on the assembly and analysis of the AbGRI2 region in the 20 GC2 isolates from Singapore. The relationship of these AbGRI2 versions to AbGRI2-1 was also examined.

5.2 Assembly of AbGRI2 regions

In draft genomes assembled from short read data, the sequence of AbGRI2 would be fragmented into several contigs due to the multiple copies of IS26. If AbGRI2-1 is present in any of these isolates it would likely be in four fragments, shown in Fig 5.2 (labels a-d). Published primer combinations (extents and product sizes indicated in Fig 5.2) had already been designed to map the AbGRI2-1 region (188). The available draft genomes of S21, S24, S26, S27, S29, S34, S38 and S39 were queried for IS26 sequence and the contigs retrieved were tiled to AbGRI2-1. Additional contigs containing IS26 fragments were also retrieved from the draft genomes S27, S34, S38 and S39 as these isolates contain the *armA* gene in another IS26 bounded resistance island, which is the focus of Chapter 6.

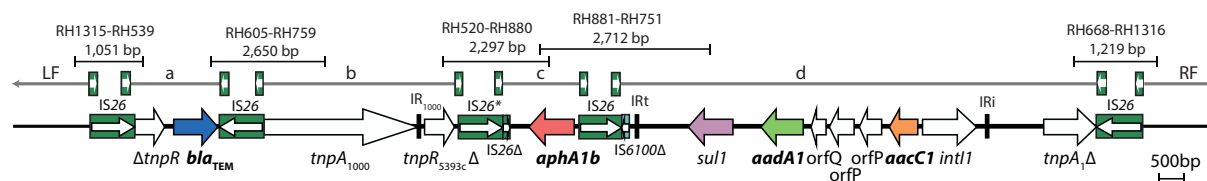


Figure 5.2 Expected contigs for the AbGRI2-1 structure. Arrows indicate the extent and orientation of genes and open reading frames. The *bla*_{TEM} gene is blue, the *aphA1b* gene is red-orange, the *sul1* gene is mauve and the *aadA1* and *aacC1* genes are light green and orange, respectively. The dark green box represents IS26 and the light blue box is part of IS6100, with internal arrows indicating the transposase orientation. IS26* is 3 bp different to the standard sequence of IS26. Vertical arrows indicate IRs. Grey lines above indicate the different fragments of AbGRI2-1. Primer combinations able to join the segments in this order are shown above by the black bars, along with the amplicon size.

5.2.1 AbGRI2-1 in S21, S24 and S29

The draft genome of S24 contained 6 contigs (Table 5.1) that tiled perfectly to AbGRI2-1 and its flanking chromosome (Fig 5.3A) while S21 and S29 each contained 7 contigs (Table 5.1) that matched all of AbGRI2-1 with the exception of the sequence covering segment d (over the integron) which had split at the *orfP* cassettes of the integron (Fig 5.3B). The presence of two copies of the *orfP* cassette was confirmed in the three isolates as they produced a 4.5 kb amplicon using primers RH754 and R1 (Fig 5.3B), consistent with the predicted amplicon size

of 4,549 bp, compared to the 4,011 bp amplicon produced for the case where only one orf P is present.

Table 5.1 Contigs of AbGRI2 in S24, S21 and S29

AbGRI2-1 fragment	Contig (size in bp)		
	S24	S21	S29
LF	197 (72,159)	253 (72,157)	134 (85,447)
a	63 (1,734)	74 (1,738)	133 (1,722)
b	38 (3,727)	37 (3,731)	110 (3,716)
c	77 (1,586)	71 (1,590)	95 (1,574)
d	95 (8,420)	-	-
d1	-	110 (3,918)	65 (3,902)
d2	-	24 (4,126)	15 (4,110)
RF	25 (25,215)	23 (24,119)	59 (25,203)

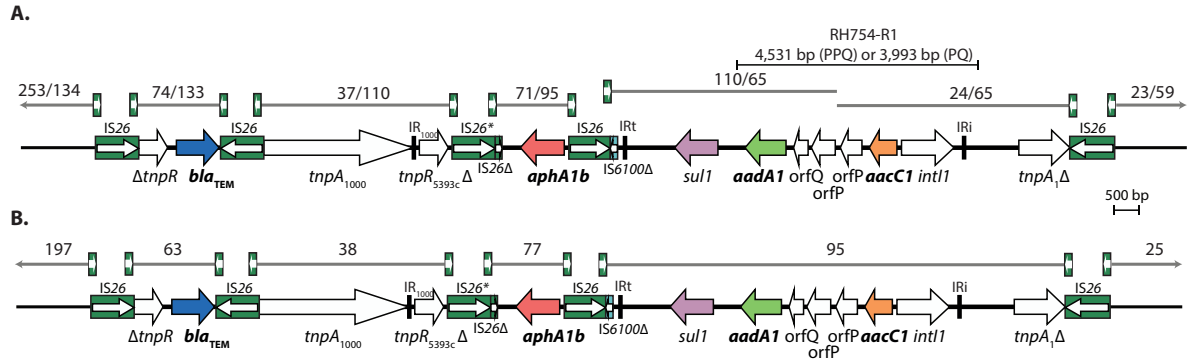


Figure 5.3 AbGRI2-1 contigs in S21, S29 and S24. A. Contigs of S21 and S29 tiled to AbGRI2-1. B. S24 contigs tiled to AbGRI2-1. Arrows indicate the extent and orientation of genes and open reading frames. The *bla*_{TEM} gene is blue, the *aphA1b* gene is red-orange, the *sul1* gene is mauve and the *aadA1* and *aacC1* genes are light green and orange, respectively. The dark green box represents IS26 and the light blue box is part of IS6100, with internal arrows indicating the transposase orientation. IS26* is 3 bp different to the standard sequence of IS26. Vertical bars indicate IRs. Grey lines above indicate the extents of different contigs in S21, S29 and S24. In A, the primer combination spanning the integron is shown above by the black bar, along with two predicted amplicon sizes.

Primers described previously (188) and indicated in Fig 5.2, were used to link segments. Each amplicon was consistent with the predicted product size for AbGRI2-1, where by sequencing these amplicons confirmed which IS26 isoform was present. Similar to AbGRI2-1, all but one copy of IS26 had the standard sequence, and the variant form IS26* (differs from IS26 at three nucleotide positions, G459A, G613A and G614A) was located between *tnpR*_{5393c} and *aphA1b*. Hence, these three isolates contain AbGRI2-1.

5.2.2 AbGRI2-12 in S34

The draft genome of S34 contained 7 contigs that end with fragments of IS26 (Table 5.2), but two of these did not tile to AbGRI2-1 and were later assembled into a third chromosomal resistance region (Chapter 6).

Table 5.2 Contigs of AbGRI2 in S34

AbGRI2-1 fragment	Contig (size in bp)
LF	136 (83,830) ¹
a	71 (1,734)
b	47 (3,648)
c	61 (1,412) ²
d	-
RF	6 (21,880) ³

¹30,672 bp of sequence is missing compared to the left flank of AbGRI2-1.

²Missing the 175 bp of Δ IS26.

³3,333 bp of sequence is missing compared to the right flank of AbGRI2-1.

The draft genome of S34 contained contigs matching fragment a, b and c, but no contig matched fragment d (containing the integron) (Table 5.2). Fragment a and b (contig 71 and 47) were linked using the published primer combination (indicated in Fig 5.4A). Published primers were also used to link fragment b with c (contig 47 and 61), but in S34 the amplicon was about approximately 200 bp smaller than the predicted size of 2,122 bp (Fig 5.4B). Sequencing of this smaller amplicon revealed that IS26* was directly adjacent to *aphA1b*, and that IS26 Δ (175 bp) was missing in S34. IS26 Δ could have been lost by homologous recombination.

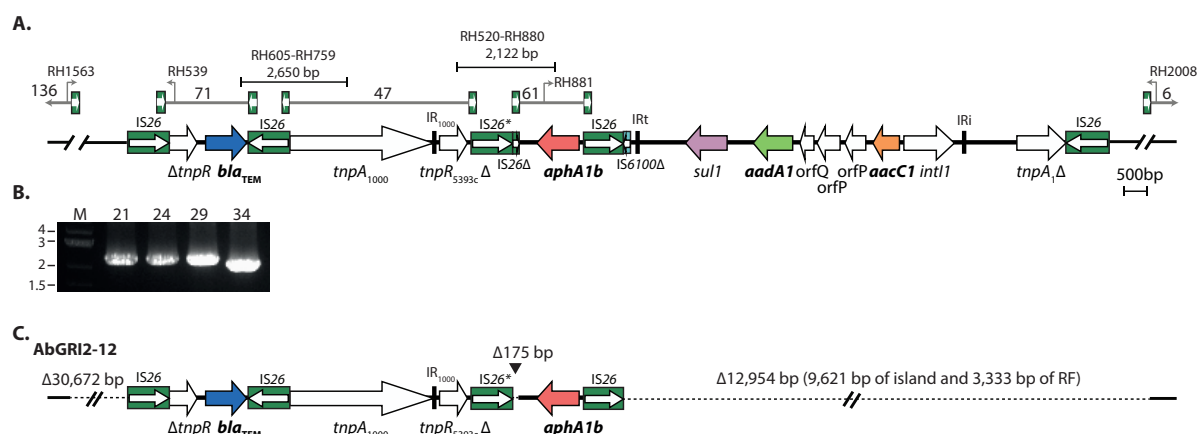


Figure 5.4 Assembly of AbGRI2-12 in S34. A. Contigs of S34 tiled to AbGRI2-1. Arrows indicate the extent and orientation of genes and open reading frames. The *bla_{TEM}* gene is blue, the *aphA1b* gene is red-orange, the *sul1* gene is mauve and the *aadA1* and *aacC1* genes are light green and orange, respectively. The dark green box represents IS26 and the light blue box is part of IS6100, with internal arrows indicating the transposase orientation. Vertical bars indicate IRs. Grey lines above indicate the extent of different contigs in S34 and grey arrows indicate the position and orientation of new primer pairs. Known primer combinations used to link sequences are shown above by the black bars, along with PCR amplicon size. B. Screening results of RH520 and RH880. Isolate numbers are shown above the lanes and sizes of the molecular marker on the left are in kb. C. Structure of AbGRI2-12. Dashed lines represent the extent of deletions relative to AbGRI2-1 and the deletion size is indicated above.

The contigs containing sequence matching the chromosome to the left and right of AbGRI2 (123 and 6 in Table 5.2), revealed that deletions in both flanks had occurred in S34 relative to AbGRI2-1. Hence, a new primer RH1563 was used with RH539 to join the contigs 136 and 71, allowing for detection of the left junction (Fig 5.4A). The amplicon was approximately 1 kb, consistent with the predicted product size of 1,046 bp. The right junction of AbGRI2 in S34 was confirmed by sequencing the 1.5 kb amplicon produced using RH881 and another new primer RH2008 (1,581 bp predicted product size) (Fig 5.4A).

The variant of AbGRI2 in S34 has been named AbGRI2-12 (Fig 5.4C). Compared to AbGRI2-1, AbGRI2-12 is missing IS26Δ and has a 9.6 kb deletion in the resistance region that has removed all of segment d and a further 3.3 kb of sequence to the right of the island. There is also an additional 30.7 kb deletion of sequence flanking on the left of AbGRI2-12. Most of this variation has likely been caused by IS26 deleting adjacent sequence.

5.2.3 AbGRI2-13 in S38 and S39

Using IS26 as a query, 7 contigs were retrieved from the draft genomes of S38 and S39. Only three of these contigs tiled on to AbGRI2-1 and adjacent chromosome (Table 5.3, Fig 5.5A), with none of them matching fragment a or c. The remaining four contigs were assembled into a different resistance island (Chapter 6).

Table 5.3 Contigs of AbGRI2 in S38/S39

AbGRI2-1 fragment	Contig (size in bp)	
	S38	S39
LF	41 (33,756) ¹	42 (33,750) ¹
a	-	-
b	41 (33,756) ¹	42 (33,750) ¹
c	-	-
d	56 (4,492)	65 (4,458)
RF	33 (25,218)	29 (25,214)

¹This appears to be a hybrid contig and sequence from a directly adjacent to sequence from c.

Contig 41 in S38 and contig 42 in S39 appeared to be a hybrid of the left flanking sequence of AbGRI2 and fragment b (Fig 5.5A). A large section of the sequence matched chromosomal sequence to 1,935 bp to the left of AbGRI2-1, then continued directly into sequence matching a part of a fragment b, containing a partial *tnpA*₁₀₀₀ and then *tnpR*_{5393c}Δ and a fragment of IS26 (Fig 5.5A).

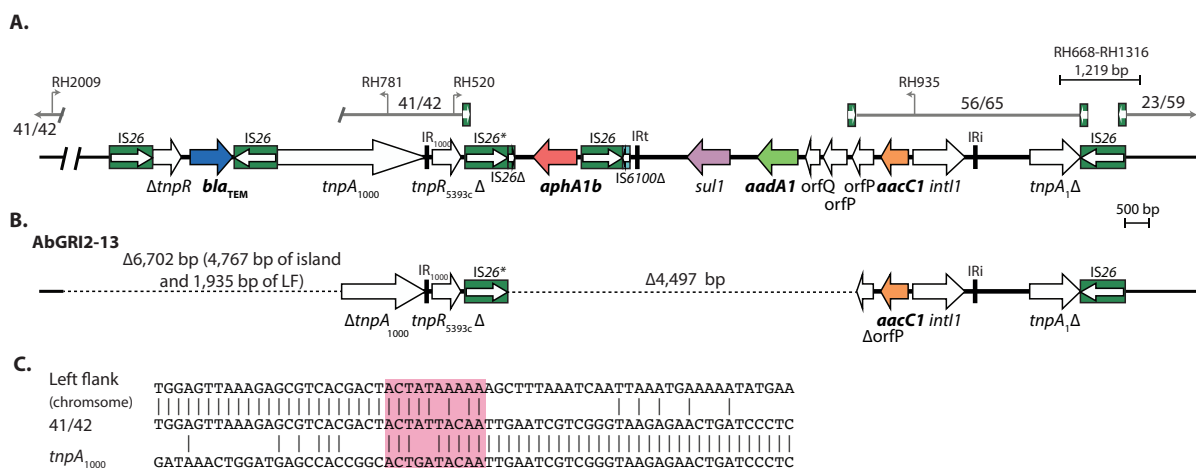


Figure 5.5 Assembly of AbGRI2-13 in S38 and S39. A. Location of AbGRI2 contigs in S38 and S39. Arrows indicate the extent and orientation of genes and open reading frames. The *bla*_{TEM} gene is blue, the *aphA1b* gene is red-orange, the *sul1* gene is mauve and the *aadA1* and *aacC1* genes are light green and orange, respectively. The dark green box represents IS26 and the light blue box is part of IS6100, with the internal arrows indicating the transposase orientation. Vertical bars indicate IRs. Grey lines above indicate the extent of different contigs in S38 and S39 and grey arrows indicate the binding sites and orientations of new primer pairs. Known primer combinations used to link sequences are shown above by the black bars, along with amplicon size. B. Structure of AbGRI2-13. Dashed lines represent the extent of deletions relative to AbGRI2-1 and the deletion size is indicated above. C. Putative site of recombination of AbGRI2-13. The pink box highlights the 10 bp region with 7 of 8 bp identical that could have facilitated the recombination.

To confirm the correct assembly of these contigs, the region was amplified using a new primer RH2009 with RH781 Fig 5.5A. This combination yields a product around 1.2 kb (consistent with the predicted product size of 1,223 bp) and the amplicon sequence confirmed the contig assembly. Comparison between the sequences of chromosome on the left of AbGRI2, contig 41/42 and *tnpA*₁₀₀₀, revealed a 10 bp region with 7 or 8 bp identical (pink box in Fig 5.5C). This short region of homology could have facilitated recombination, removing 4,767 bp of the left end of the resistance island and 1,935 bp of the flanking chromosomal sequence relative to AbGRI2-1, a total deletion of 6,702 bp.

S38 and S39 each had a contig (56 or 65 for S38 and S39, respectively) matching a part of fragment d (Table 5.3 and Fig 5.5A). The primer RH520 was combined with RH935 to link contig 41/42 to contig 56/65 and produced amplicons just below 2 kb, consistent with the predicted amplicon size of 1,881 bp. Sequence of the amplicon matched IS26*. The other end

of contig 56/65 was joined to contig 23/59 containing the right flank of AbGRI2-1 using the published primer combination (mapping PCR indicated above in Fig 5.5A) and the sequence of the amplicon was consistent with a standard IS26. A 4,497 bp segment including the *aphA1b*, *sulI* and *aadA1* genes is missing from the AbGRI2 region of S38 and S39, likely due to the IS26* causing an adjacent deletion (Fig 5.5B). The variant of AbGRI2 in S38 and S39 was named AbGRI2-13.

5.2.4 No AbGRI2 variant in S27

The draft genome of S27 contained six contigs ending with fragments of IS26. Contig 131 (125,012 bp) matched sequence found to the left of AbGRI2-1 and contig 28 (19,716 bp) matched the chromosome to the right, but is missing 5,498 bp compared to the region flanking AbGRI2-1 (Fig 5.6). Another contig contained *aphA1b* with fragments of IS26 at each edge. However, the *aphA1b* gene had one SNP compared to that of *aphA1b* in Tn6020, and the segment between the bounding copies of IS26 was of a different size to Tn6020 (described in more detail in Chapter 6; Section 6.2.2). Instead, this *aphA1b* contig and the three remaining contigs containing fragments of IS26 were assembled into the third resistance region (see Chapter 6). Hence, S27 did not contain any contigs matching the internal sequence of the AbGRI2-1 island. Contig 131 was joined directly to contig 28 using the primers RH1315 and RH2007 (Fig 5.6). The amplicon produced was just over 1 kb, consistent with the predicted size of 1,048 bp and its sequence revealed the presence of IS26. Hence, only a single IS26 is left in the vicinity of AbGRI2 in S27. The orientation of the IS26 and the conservation of the left boundary of AbGRI2-1 is consistent with it being the IS26 from the left end of AbGRI2 (Fig 5.6) as the right bounding IS26 is in the opposite orientation (Fig 5.2).

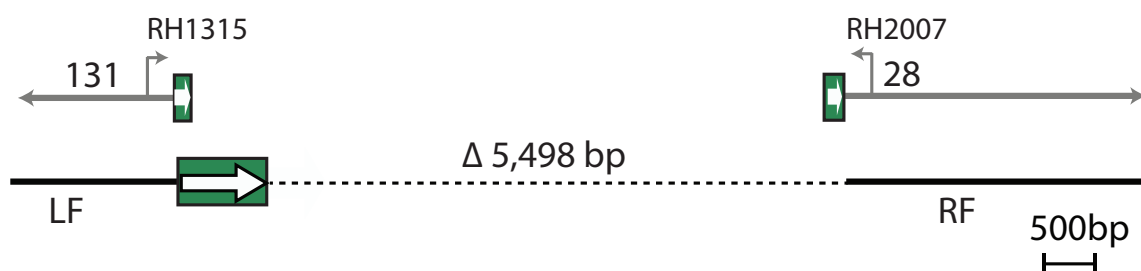


Figure 5.6 Assembly of the region around AbGRI2 in S27. The dark green box represents IS26 and the internal arrows indicate the transposase orientation. Grey lines above indicate the extent of different contigs in S27. Grey arrows indicate the binding sites and orientations of primers. Deletions of surrounding chromosomal sequence relative to AbGRI2-1 are indicated by dashed lines.

5.3 PCR mapping of AbGRI2 forms in isolates not sequenced by Illumina HiSeq

The 13 isolates without draft genome sequence available (by Illumina HiSeq) were screened using various primer combinations. First, the published primer combinations for AbGRI2-1 were used (Table 5.4) (188), followed by the primer combinations devised for the configurations described above (also summarised in Table 5.5), with the appropriate sequenced isolates used as controls.

Table 5.4 AbGRI2-1 screening PCRs

Isolate	LF-a	a-b	b-c	c-d	d-RF	AbGRI2 matches pattern
	RH1315- RH539	RH605- RH759	RH520- RH880	RH881- RH751	RH668- RH1316	
	1,051 bp	2,650 bp	2,297 bp	2,712 bp	1,219 bp	
S22	+	+	-	-	+	new
S23	+	+	-	-	+	new
S25	+	+	-	-	+	new
S26	+	+	+	+	+	1
S28	+	+	-	-	+	new
S30	+	+	+	+	+	1
S31	-	-	-	-	+	13
S32	-	+	+ ¹	-	-	12
S33	-	+	+ ¹	-	-	12
S35	-	-	-	-	-	-
S36	-	-	-	-	-	-
S37	-	-	-	-	-	-
S40	-	+	+ ¹	-	-	12

¹100-200 bp smaller than the predicted size.

The results for S26 and S30 were consistent with the presence of AbGRI2-1 structures (Table 5.4). In addition to S21, S24 and S29 there are now five isolates with this arrangement, all of which are a part of group 1 from Chapter 3; Table 3.7, which lack *armA* and contain *aphA6*.

Four different patterns from the AbGRI2-1 screening PCRs were detected in the remaining *armA* positive isolates (Table 5.4). Two of these patterns match those of AbGRI2-12 and AbGRI2-13, one was novel, and in the fourth no junctions were detected. Where relevant, additional linkages were attempted for these isolates using primer combinations utilised in the assembly of the AbGRI2 regions in S27, S34 (AbGRI2-12) and S38/S39 (AbGRI2-13) and for the detection of deletions in the surrounding chromosome (Table 5.5).

Table 5.5 Results of screening PCRs used to assemble additional AbGRI2 variants

Isolate	LF (AbGRI2-12)- a	AbGRI2-13 LF-hybrid	c- RF (AbGRI2-12)	b-Δd	LF- RF (S27)	Predicted AbGRI2 variant
	RH1563-RH539 (1,046 bp)	RH2009- RH781 (1,223 bp)	RH881-RH2008 (1,581 bp)	RH520- RH935 (1,881 bp)	RH1315- RH2007 (1,048 bp)	
S22	ND	ND	ND	+	ND	14
S23	ND	ND	ND	+	ND	14
S25	ND	ND	ND	+	ND	14
S28	ND	ND	ND	+	ND	14
S31	-	+	ND	+	ND	13
S32	+	ND	+ ¹	ND	ND	12b
S33	+	ND	+ ¹	ND	ND	12b
S35	-	-	-	-	-	?
S36	-	-	-	-	-	?
S37	-	-	-	-	-	?
S40	+	ND	+ ¹	ND	ND	12b

ND denotes that this arrangement was not tested as an arrangement had already been confirmed in this isolate.

¹Product was approximately 200 bp larger than the predicted size.

Three isolates S35, S36 and S37 did not produce any amplicons for the mapping arrangements attempted (Table 5.5). This is consistent with earlier gene screening (Chapter 3; Table 3.5) which showed that these isolates lacked *bla*_{TEM}, *aphA1b*, *aacCI*, *aadA1* and *sulI* and suggests that like S27, they only contain an IS26 at the chromosomal AbGRI2 position. However, as a linkage between the chromosome on the left and on the right of the IS26 in S27 (RH1315-

RH2007) failed to amplify in S35, S36 and S37 (Table 5.5), it was concluded that they contain a different deletion in the region.

5.3.1 AbGRI2-12 with a smaller adjacent deletion

The isolates S32, S33 and S40 produced amplicons similar to those of AbGRI2-12 (Table 5.4), including the smaller amplicon size when RH520 was paired with RH880, indicating that IS26 Δ of Tn6020 was missing. These three isolates also produced an amplicon of the expected size when RH1563 and RH539 were combined (Table 5.5), linking the left flanking chromosomal sequence to fragment a, indicating that they also had a 30.6 kb deletion in the flanking sequence to the left of the resistance island compared to AbGRI2-1. The *aphA1b* gene was connected directly to the flanking sequence on the right of the island using the primer combination RH881 and RH2008. However, the amplicon produced for S32, S33 and S40 was approximately 200 bp larger than for S34 (Fig 5.7). The sequence of the larger amplicon revealed that these three isolates had a smaller 3,142 bp deletion to the right of the island compared to AbGRI2-1, rather than the 3,333 bp deletion in S34. As the actual content of the resistance region is the same, the AbGRI2 variant in S34 was designated AbGRI2-12a and in S32, S33 and S40 it was named AbGRI2-12b.

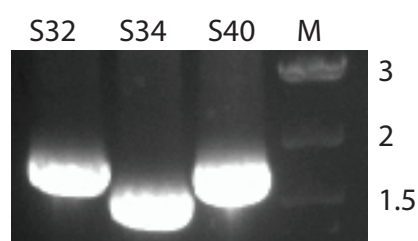


Figure 5.7 Amplification of the right end of AbGRI2-12 using primers RH881 and RH2008. The isolate numbers are shown above their lanes and the sizes of the molecular markers to the right are in kb.

5.3.2 Structure of AbGRI2-14, precursor of AbGRI2-13

The isolate S31 produced amplicons consistent with the predicted size for the two characteristic linkages of AbGRI2-13. This included the hybrid left flank containing *tnpA*₁₀₀₀Δ (RH1563-RH539) and the connection between *tnpR*_{5393c}Δ and the partial integron (RH520-RH935) (Table 5.5). Hence, S31 likely contains AbGRI2-13 (Fig 5.8A).

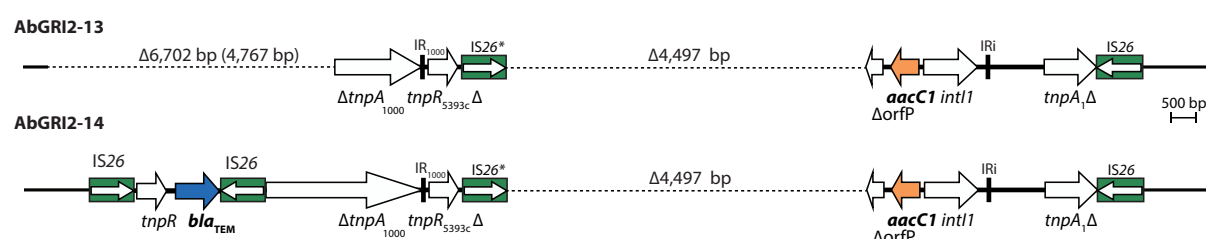


Figure 5.8 Relationship between AbGRI2-13 and AbGRI2-14. Arrows indicate the extent and orientation of genes and open reading frames. The *bla*_{TEM} gene is blue, the *aphA1b* gene is red-orange, the *sulI* gene is mauve and the *aadA1* and *aacC1* genes are light green and orange, respectively. The dark green box represents IS26 with internal arrows indicating the transposase orientation. Vertical bars indicate IRs. Dashed lines represent the extent of deletions relative to AbGRI2-1 with the deletion size indicated above.

The isolates S22, S23, S25 and S28, also produced an amplicon of the predicted size for the linkage of *tnpR*_{5393c}Δ to the partial integron (RH520-RH935, Table 5.5) seen in AbGRI2-13. Sequence of the amplicon revealed that the island in these isolates includes the same 4,497 bp internal deletion as AbGRI2-13, which had removed the *aphA1b*, *sulI* and *aadA1* resistance genes (Fig 5.8B). However, the left end of AbGRI2 was consistent with the arrangement of AbGRI2-1 (Table 5.4) and the variant in S22, S23, S25 and S28 was named AbGRI2-14 (Fig 5.8B). Hence, it is possible that AbGRI2-13 has arisen from AbGRI2-14 through the deletion at the left end of AbGRI2-13 likely caused by recombination between a short segment of homologous sequence (Fig 5.5C).

5.4 Structures from the Illumina MiSeq draft genomes

The AbGRI2 regions assembled by PCR mapping in the 10 isolates described above were compared to the draft genome sequences of these islands when they became available at the end of 2016. Assembly of the draft genome sequences of S22, S23, S25, S26, S28, S30-S33 and S40 confirmed they had AbGRI2 structures identical to those predicted using PCR. Additional information from the draft genomes of S35, S36 and S37 facilitated the assembly of the regions surrounding AbGRI2 in these isolates.

5.4.1 No AbGRI2 variant in S35, S36 and S37

Examination of the draft genomes of S35, S36 and S37 revealed that, like S27, only a single IS26 is present in the region of the chromosome usually occupied by AbGRI2 and that none of the internal sequence of AbGRI2-1 has been retained. The two chromosomal contigs that contained fragments of this IS26 were contig 26 (47,409 bp) and contig 32 (17,239 bp) in S35 (Fig 5.9A) and equivalent contigs were present in the draft genomes of S36 and S37. The primer RH1563 and the new primer RH2548 were combined to link these two segments and produced an amplicon just under 2 kb, consistent with the predicted amplicon size of 1,850 bp. The sequence of this amplicon was compared to AbGRI2-1, revealing 30,672 bp of the chromosome to the left is missing, an identical deletion as the one found adjacent to AbGRI2-12a (Fig 5.9B). In addition, 8,115 bp of chromosomal sequence to the right of the island has also been lost relative to AbGRI2-1 (Fig 5.9B). The arrangement present in S35, S36 and S37 could have been derived from AbGRI2-12 by the IS26 at the left of the island, which is in the correct orientation, causing a deletion that removed the remainder of the island as well as an extra 4.7 kb of chromosome on the right.

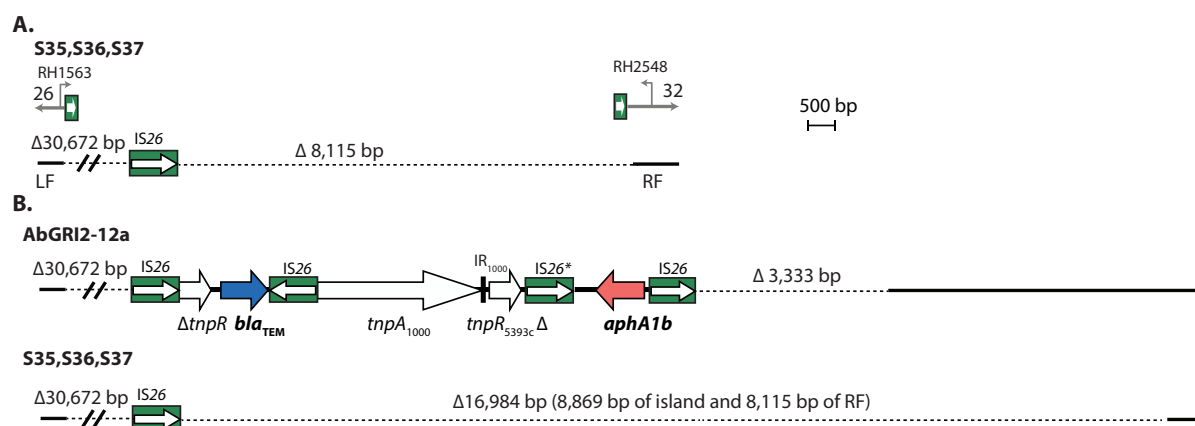


Figure 5.9 AbGRI2 region of S35, S36 and S37. A. Contigs matching the left and right flanking sequence of S35. The dark green box represents IS26 with the internal arrows indicating the transposase orientation. Grey lines above indicate the extent of different contigs in S35. Grey arrows show the binding sites and orientations of primers. Deletions of surrounding chromosomal sequence are represented by dashed lines and the size is relative to AbGRI2-1 is above. B. Relationship between AbGRI2-12 and the AbGRI2 region of S35, S36 and S37. Arrows indicate the extent and orientation of genes and open reading frames. The *bla*_{TEM} gene is blue and the *aphA1b* gene is red-orange. Vertical bars indicate IRs. Deletions are indicated by dashed lines the sizes of deletions of the chromosome are relative to AbGRI2-1, but the deletion size of the island is relative to AbGRI2-12.

5.5 Discussion

The 20 GC2 isolates from Singapore each contained a version of AbGRI2 or had an IS26 in the same region of the chromosome, indicating that an AbGRI2 island had been present (Fig 5.10). AbGRI2-1 was present in the five isolates without the *armA* gene and with *aphA6*. Within the isolates that contain *armA* there were three novel AbGRI2 variants (AbGRI2-12, AbGRI2-13 and AbGRI2-14) and two different arrangements with only a single IS26 in the region. A comparison of these variants is shown in Fig 5.10A. All of the variants can be derived from AbGRI2-1 via deletions that, in all but one case, could be caused by IS26. In the outlying case of AbGRI2-13, a deletion event likely facilitated by a short region of homology has occurred and this is the first case of non-IS26 mediated variation in AbGRI2.

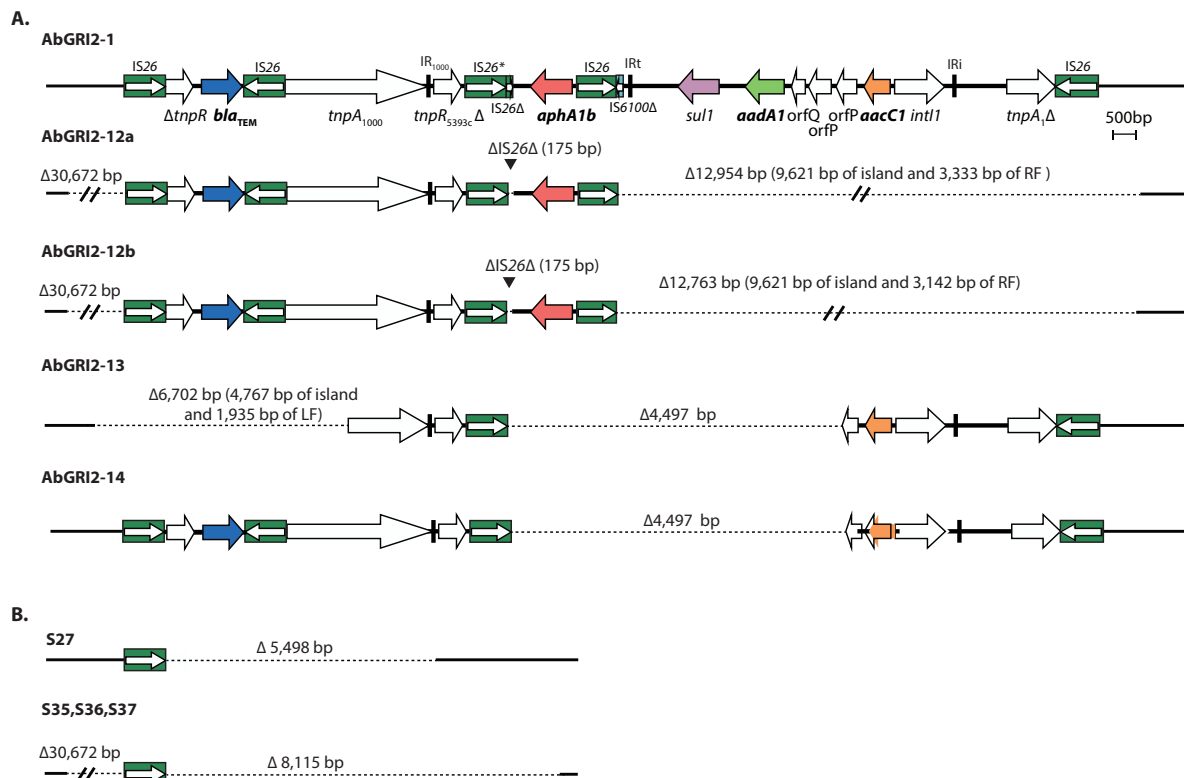


Figure 5.10 Structures in the AbGRI2 region. A. AbGRI2 variants detected. B. Isolates where only a single IS26 is present. Arrows indicate the extent and orientation of genes and open reading frames. The *bla*_{TEM} gene is blue, the *aphA1b* gene is red-orange, the *sul1* gene is mauve and the *aadA1* and *aacC1* genes are light green and orange, respectively. The dark green box represents IS26 and the light blue box is part of IS6100, with the internal arrows indicating the transposase orientation. Vertical bars indicate IRs. Dashed lines represent the extent of deletions relative to AbGRI2-1, with the deletion size indicated above.

In addition to the missing 47.2 kb segment of chromosome to the left of AbGRI2-1, relative to AbGRI2-0 (24), the AbGRI2-13 and AbGRI2-12 variants have lost additional 1.9 kb and 30.7 kb, respectively. Hence, AbGRI2-12 is missing 77.9 kb of sequence relative to AbGRI2-0. As the fitness of these isolates containing this variant does not appear to be affected, it can be concluded that this region of the chromosome is non-essential.

The nature of the variation of AbGRI2 provides additional evidence that the collection can be separated into two groups, which were initially defined by the presence or absence of *aphA6* and *armA* (Table 5.6). The five isolates that lack *armA*, and hence the resistance island that contains this gene, possess AbGRI2-1 and are the first report of this variant in isolates from hospitals outside of Australia. Furthermore, it was found in an isolate from 1996 (S21), which

is three years earlier than the oldest documented isolate with AbGRI2-1 seen in Australia (188, 193). This indicates a relationship between the common XAR GC2 isolates found across the eastern seaboard of Australia and some GC2 isolates from South-East Asia, particularly Singapore. Further work is required to examine the nature of the relationship between the GC2 isolates seen in Australia and South-East Asia.

Table 5.6 Two groups within the collection

Group	<i>armA</i>	Isolates	AbGRI2	AbGRI1
1	-	S21, S24, S26, S29, S30	1	3, 5, 1
2	+	S32, S33, S34, S40	12	7, 3,
	+	S31, S38, S39	13	6, 1
	+	S22, S23, S25, S28	14	3, 8
	+	S27	IS26 (1)	3
	+	S35, S36, S37	IS26 (2)	7

The 15 isolates with *armA* in the second group either contain an AbGRI2 variant smaller than AbGRI2-1 or have only a single IS26 remaining in the AbGRI2 region (Table 5.6; Fig 5.10A and B). All of the AbGRI2 variants in group 2 could be derived from a precursor structure corresponding to AbGRI2-14 (Fig 5.10), before Tn6020 was lost. However, MDR-TJ (125), which was also isolated in Asia and includes *armA* (Chapter 6; Section 6.2.1), has AbGRI2-2 which contains the entire integron fragment and Δ IS6100 (188). Hence, the only AbGRI2 structure that accounts for all the components found in the AbGRI2 regions of isolates with *armA* is AbGRI2-1. Consequently, the cell that originally acquired the *armA* containing island described in the next chapter must have had AbGRI2-1.

Tracking the AbGRI2 region over time indicates that it entered the GC2 chromosome prior to 1982 (24). By 1996, the AbGRI2-1 structure, characterised by the inversion and deletion at the left end, had arisen. The AbGRI2 structures in more recent isolates in Asia, America and Australia appear to have derived from AbGRI2-1 (193, 277), which suggests that either this

island variant was successful or the strain containing this variant was successful for other reasons, resulting in its subsequent and apparent globalisation. However, there are limited sequences of modern and historic GC2 isolates from Europe, so it is currently unclear if the GC2 population in Europe is also part of the AbGRI2-1 lineage or if other sub-lineages of this clone are present.

CHAPTER

6

Characterisation of AbGRI3

6.1 Introduction

There are many different antibiotics within the aminoglycoside class but most resistance genes confer resistance to only a subset of them. However, often these genes are found in combination with one another. In the AbGRI2-1 island, the *aacC1* (resistance to gentamicin), *aadA1* (streptomycin and spectinomycin resistance) and *aphA1b* (kanamycin and neomycin resistance) genes are co-located. However, the emergence of acquired 16S rRNA methyltransferases that confer resistance to most aminoglycoside antibiotics via one gene is most concerning. These genes provide ribosomal protection by hampering the binding of aminoglycosides to the 30S subunit (54).

ArmA is a 16S rRNA methyltransferase that was first characterised from a *Klebsiella pneumoniae* isolated from France in 2000 and shown to confer resistance to 4,6-disubstituted deoxystreptamines; gentamicin, kanamycin, amikacin, tobramycin, isepamicin, netilmicin and sisomicin (72). Subsequently, the context of the *armA* gene in an IncL/M plasmid, pCTX-M3 (GenBank accession number AF550415), from a 2002 *Citrobacter freundii* isolate from Poland (82) was characterised. The *armA* gene was located in a 16.6 kb composite transposon that was bounded by two directly-oriented copies of IS26 and called Tn1548 (Fig 6.1) (73). Transposon Tn1548 was subsequently reannotated (203) and in addition to *armA* also contains the macrolide resistance genes *msrE* and *mphE*, and a class 1 integron containing *sulI* (sulphonamide resistance), *aadA2* (resistance to streptomycin and spectinomycin) and *dfrA12* (resistance to trimethoprim) (Fig 6.1). Later, a putative *rep* gene, annotated as *repAciN* (55), was identified at the left end of the transposon (Fig 6.1). However, this reading frame includes 73 bp from the end of IS26 and only part of a *rep* gene. This reading frame is related to the *repAci2* gene (74% DNA identity), but it is missing the first 174 bp due to the insertion of IS26 and it had been shown previously that *repAciN* does not encode a functional replication

initiation protein (174). A closely related transposon containing *armA* but only carrying the *aadA2* gene cassette was also found in an IncN plasmid, pMUR050 in an *E. coli* from Spain in 2002 (83).

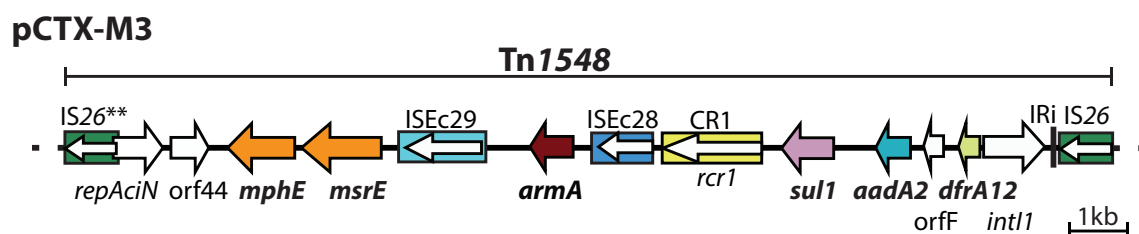


Figure 6.1 Structure of Tn1548. Arrows indicate the extent and orientation of genes and open reading frames. The *mphE* and *msrE* genes are orange, the *armA* gene is maroon, *sul1* gene is mauve and the *aadA2* and *dfrA12* genes are aqua and lime green, respectively. Boxes represent IS and CR elements and the internal arrows indicate the transposase or the rolling circle replicase. IS26 is dark green, ISEc29 and ISEc28 are light and dark blue and CR1 is yellow. IS26** is 2 bp different to the standard sequence of IS26. Vertical bars indicate inverted repeats and the extent of Tn1548 is shown above. Figure made to scale from pCTX-M3 (GenBank accession number AF550415).

The earliest *A. baumannii* isolates that have been shown to carry the *armA* gene are from 2003 and were isolated in South Korea (153). Since then, the *armA* gene has been reported in strains from North America (53, 277), Japan (253) China (125, 285, 286), and India (231). This gene was also found in isolates from Norway recovered from patients that had previously been hospitalised in Asia (136). Where MLST data using the Institut Pasteur (IP) scheme is available, most isolates containing the *armA* gene are ST2 and are consequently members of GC2. In the completed genomes of MDR-ZJ06 (285), BJAB0868 and BJAB07104 (286), the *armA* gene was reported to be on a plasmid. However, using isolate UH9907 as their exemplar (AY0H000000000), Wright *et al.* (277) recently described a third chromosomally-located island containing the *armA* gene in GC2 *A. baumannii* from the USA. This resistance island was described as located between ACICU_02398 and ACICU_02399 (using ACICU genome as a reference, GenBank accession number CP000863). These authors (277) also identified a previously unnoted, related island that was in the same location in the completed genome of

TYTH-1 (161), which was isolated in China. In both the chromosome and the sequences annotated as plasmids, the *armA* gene is in an IS26-bounded transposon that is described as similar to Tn1548, usually including the *aadA1*, *aacA4*, *sul1*, *msrE*, *mphE* and *catB8* genes (277). This can also be associated with the *aphA1b* gene.

The work presented in this chapter focuses on characterising the context of the *armA* gene in GC2 *A. baumannii*. This first involved analysing in detail the regions known to contain the *armA* gene, including those in TYTH-1, UH9907 and pJAB07104, and comparing these regions to Tn1548. This analysis then guided the assembly of the regions of the chromosome containing the *armA* gene, which was in 15 of the GC2 Singapore isolates (Chapter 3; Table 3.5). The extent of variation in this *armA* containing transposon and the region surrounding it in the chromosome was further examined through analysis of the sequences available in GenBank.

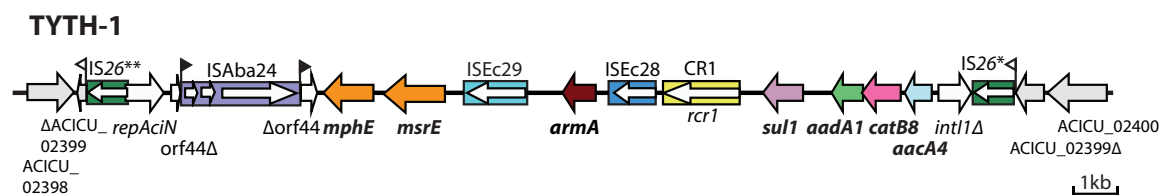
6.2 Transposons containing *armA* in GC2 *A. baumannii*

6.2.1 AbGRI3 in TYTH-1: AbGRI3-1

The region containing the *armA* gene in the completed genome TYTH-1 (GenBank accession number CP003856) was analysed and found to be located in a 19 kb IS26 bounded transposon (Fig 6.2A). The transposon in TYTH-1 was surrounded by an 8 bp TSD (5' CTCATCCT 3') (Fig 6.2A) and it was found to interrupt the open reading frame ACICU_02399. Thus, it is not situated between ACICU_02398 and ACICU_02399 as reported previously (277). The reading frame ACICU_02399 codes for a putative GNAT family acetyltransferase, which for simplicity will be referred to as *atr* hereafter. An almost identical transposon (containing 4 SNPs) was present in the completed genome of MDR-TJ (GenBank accession number CP003500), also interrupting the *atr* gene, and surrounded by the same 8 bp duplication as in TYTH-1. As the

armA transposon represents the third chromosomally-located resistance island in GC2 *A. baumannii*, for simplicity and clarity, I will call it AbGRI3 and the location for AbGRI3 with respect to the other two resistance islands is shown in Fig 6.2B. The form of the AbGRI3 island present in TYTH-1 and MDR-TJ is called AbGRI3-1.

A.



B.

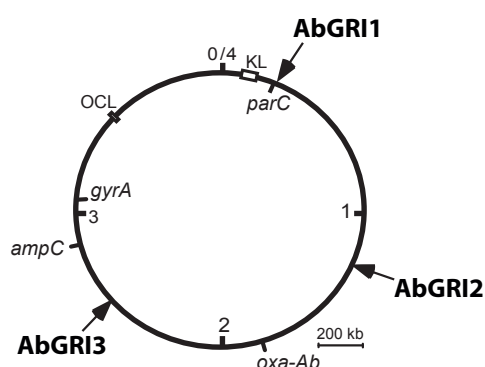


Figure 6.2 AbGRI3 and its chromosomal location. A. Structure of AbGRI3-1. Arrows indicate the extent and orientation of genes and open reading frames. The *mphE* and *msrE* genes are orange, the *armA* gene is maroon, *sul1* gene is purple and the *aadA1*, *catB8* and *aacA4* genes are green, pink and pale blue, respectively. Adjacent chromosomal genes are grey. Boxes represent IS and CR elements and the internal arrows indicate the transposase or the rolling circle replicase. IS26 is dark green, ISAb24 is purple, ISEc29 and ISEc28 are light and dark blue and CR1 is yellow. IS26** is 2 bp different to the standard sequence of IS26, while IS26* is 3 bp different. Black and white flags indicate matching TSD. B. Location of AbGRI3 on the GC2 chromosome. Arrows indicate the location of resistance island and figure is made to scale from TYTH-1 (GenBank accession number CP003500).

The sequence of the *armA* containing transposon of AbGRI3-1 was compared to that of Tn1548 (Fig 6.3). Several differences distinguish these two transposons and hence we have named the transposon in TYTH-1 Tn6180. Compared to Tn1548, the integron in Tn6180 carries a different set of cassettes *aadA1-catB8-aacA4*, and a deletion at the right end of the central segment has removed 216 bp of the *int11* gene (Fig 6.3). Another difference is the insertion of a novel IS, ISAb24, in orf44 in Tn6180 (Fig 6.3). ISAb24 is a member of the IS66 family and is 2,421 bp in length with almost perfect inverted repeats of 22 bp with 21 bp identical.

This IS is surrounded by an 8 bp TSD. The sequences of the IS26 elements bounding both Tn1548 and Tn6180 were also examined. The IS26 on the right of Tn1548 is the standard IS26 sequence, while the corresponding IS26 in Tn6180 differs at three nucleotide positions (G459A, G613A and G614A) and is referred to as IS26*. The IS26 on the left of both transposons harbours two of these differences (G613A and G614A) compared to the standard IS26 sequence and, here this isoform is called IS26**.

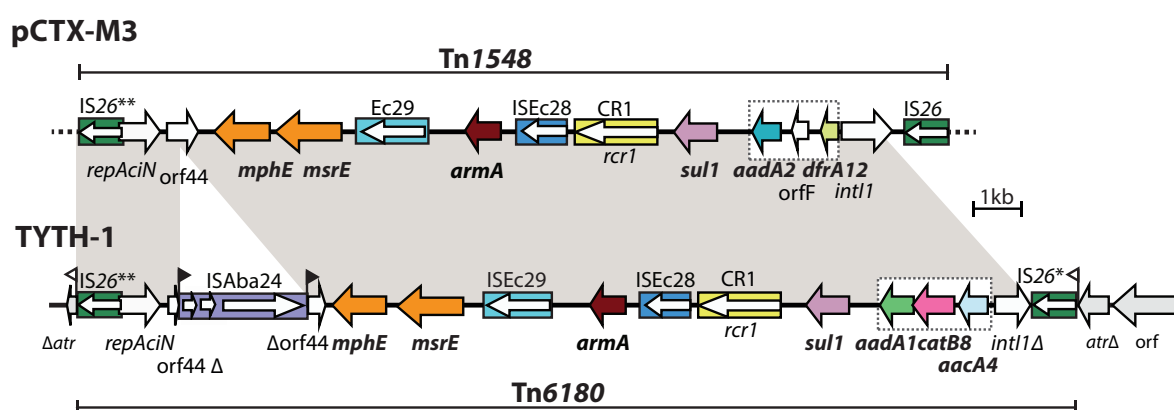


Figure 6.3 Comparison of Tn1548 and Tn6180. Arrows indicate the extent and orientation of genes and open reading frames. The *mphE* and *msrE* genes are orange, the *arma* gene is maroon, *sul1* gene is mauve, the *aadA2* and *dfrA12* genes are aqua and lime green, respectively, while the *aadA1*, *catB8* and *aacA4* genes are green, pink and pale blue, respectively. Adjacent chromosomal genes are shown in grey. Boxes represent IS and CR elements and the internal arrows indicate the transposase or the rolling circle replicase. IS26 is dark green, IS26** is 2 bp different to the standard sequence of IS26, while IS26* is 3 bp different. Black and white flags indicate matching TSD. Lines above and below indicate the extents of transposons and grey shading indicates matching sequence. Figure made to scale from pCTX-M3 (GenBank accession number AF550415) and TYTH-1 (GenBank accession number CP003500).

6.2.2 AbGRI3 in UH9907: AbGRI3-2

Examination of the *armA* containing transposon in UH9907 (277), revealed this isolate contained AbGRI3 in the same location as TYTH-1 and MDR-TJ, i.e. interrupting the *atr* gene (Fig 6.4A). However, in addition to containing Tn6180, UH9907 also includes an *aphA1b* containing transposon that shares one IS26 with the right IS of Tn6180, and the whole structure is surrounded by the same 8 bp TSD as TYTH-1 and MDR-TJ (Fig 6.4A). We named this 2,706 bp *aphA1b* containing transposon Tn6179 and Fig 6.4B shows a comparison of Tn6179 to Tn6020 (217), the transposon containing *aphA1b* that is found in AbGRI2 of GC2 *A*.

resembling a replication initiation gene is the inactive orf *repAciN*, which does not encode a functional replication initiation protein (174). Hence, the assignment of these sequences as plasmids is highly questionable and this is examined further in Section 6.4.1.

6.3 AbGRI3 in Singapore isolates

Previous screening revealed that 15 of the isolates, S22, S23, S25, S27, S28, S31-S40, contain the *armA* gene (Chapter 3; Table 3.5). Primers, RH2001 and RH2004, were designed in the chromosomal region surrounding the insertion location of AbGRI3 in the *atr* gene and were used to screen the 20 GC2 Singapore isolates. If the *atr* gene is intact, the amplicon is predicted to be 400 bp in size. The isolates S21, S24, S26, S29 and S30 produced an amplicon of this size which is consistent with the absence of *armA* and AbGRI3 in these isolates (Fig 6.5). The *atr* gene in the other 15 *armA* positive isolates is interrupted, indicating that they contain a version of AbGRI3 or have lost the priming sites (Fig 6.5).

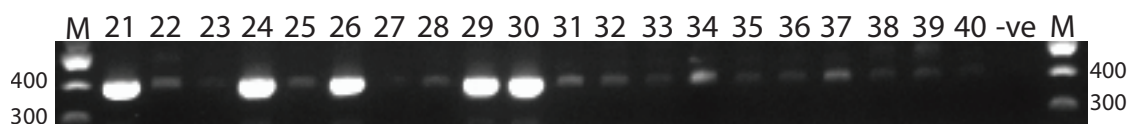


Figure 6.5 Is the *atr* gene intact? Screening the AbGRI3 insertion site using primers RH2001 and RH2004. Isolate numbers are shown above the lanes and sizes of the molecular marker to the left and right are in bp.

The assembly of AbGRI3 regions in the 15 *armA* positive isolates was first performed by assembling those of S27, S34, S38 and S39 whose draft genomes were available initially. The majority of the AbGRI3 regions in the remaining isolates were deduced by PCR screening while for two isolates, additional information was required that was gained once draft genomes were available.

6.3.1 AbGRI3-2i in S27

The draft genome of S27 contained five contigs that had a fragment of IS26 on at least one end (Table 6.1) but were not assembled as part of the AbGRI2 (Chapter 5; Section 5.2.4).

Table 6.1 S27 contigs with IS26 fragments

Contig	Size (kb)	Content	Edges	
			Left	Right
3	133.6	LH flank	ND	Δatr , IS26-IRr
30	17.6	Tn6180	IS26-IRl	IS26-IRr
80	1.1	Tn6179	IS26-IRl	IS26-IRr
37	3.3	RH flank	IS26-IRl, <i>atr</i> Δ	Δasr^l , IS26-IRl
5	142	RH chromosome	IS26-IRr, <i>asr</i> Δ	ND

[†]*asr* stands for aspartate racemase

Through a comparison of these contigs to AbGRI3-2 (Fig 6.6), it was clear that S27 contained both Tn6180 and Tn6179. The chromosomal contigs, 3 (133 kb) and 37 (3.27 kb), each contained part of the *atr* gene adjacent to an IS26 fragment and the boundaries between IS26 and the chromosome-derived sequence were the same as in AbGRI3-1 and -2. However, another fragment of IS26 was at the other end of the short contig 37 and the sequence next to the IS26 matched a part of an aspartate racemase gene (corresponding to ACICU_02402 in ACICU), here called *asr* for simplicity, which is found 3.1 kb from *atr* in ACICU. The remainder of the *asr* gene, along with a TSD, was found adjacent to an IS26 outer end on chromosomal contig 5 (142 kb). Therefore, an additional copy of IS26 interrupts the *asr* gene, in the chromosome a short distance to the right of the island in S27 (Fig 6.6).

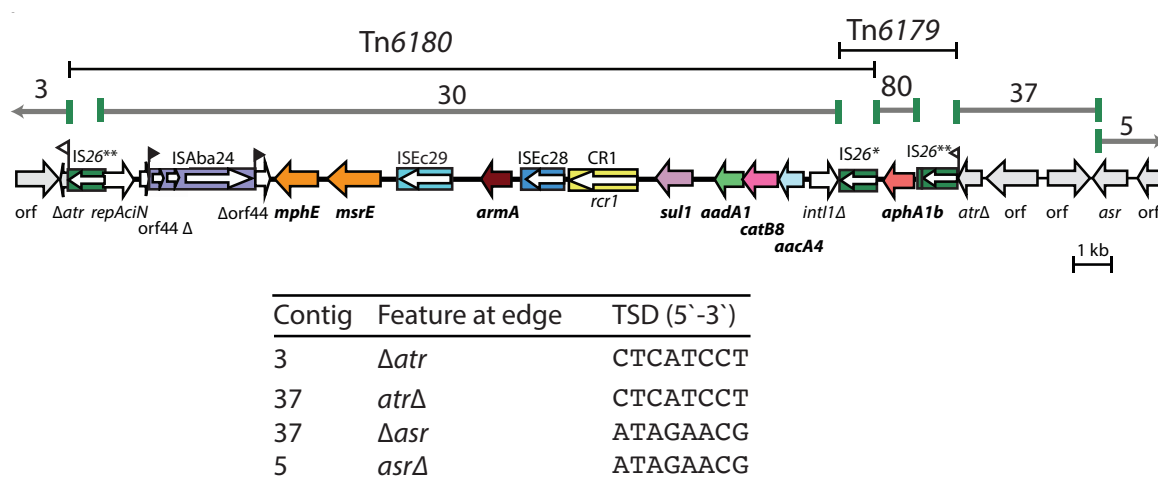
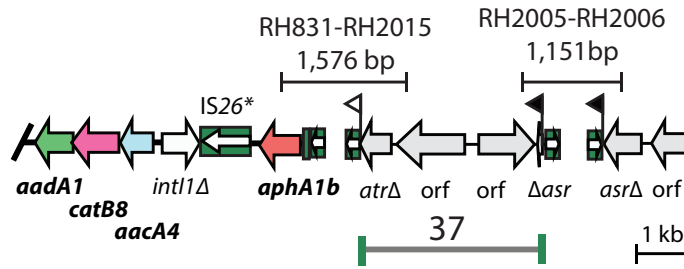


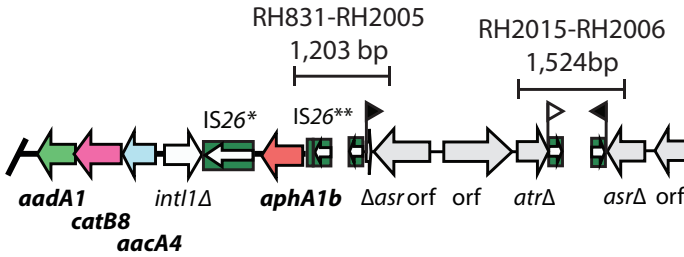
Figure 6.6 Draft assembly of AbGRI3 contigs in S27. Arrows indicate the extent and orientation of genes and open reading frames. The *mphE* and *msrE* genes are orange, the *armA* gene is maroon, *sul1* gene is mauve, the *aphA1b* gene is red-orange, and the *aadA1*, *catB8* and *aacA4* genes are green, pink and pale blue, respectively. Chromosomal genes are shown in grey. Boxes represent IS and CR elements and the internal arrows indicate the transposase or the rolling circle replicase. IS26 is dark green, ISAba24 is purple, ISEc29 and ISEc28 are light and dark blue and CR1 is yellow. IS26** is 2 bp different to the standard sequence of IS26, while IS26* is 3 bp different. Black and white flags indicate matching TSD. Grey lines above indicate the extent of different contigs in S27 and the green bars represent fragments of IS26. Black bars above show the extent of transposons.

The *repAciN* end of Tn6180 (contig 30) was linked to the *atr* end of contig 3 (Fig 6.6) using the newly designed primers RH2001 and RH2002. These primers were expected to amplify a 1,323 bp product and the approximately 1.3 kb amplicon produced was sequenced, revealing the presence of an IS26**, as in AbGRI3-1 and AbGRI3-2. The other end of Tn6180 was joined to Tn6179 (contig 80) using RH2003 and RH880, which produced an amplicon of 2 kb, consistent with the expected size of 1,943 bp. The IS26 isoform was IS26*. Linkages could be confirmed between Tn6179 and the remainder of *atr* (RH831-RH2015, 1,576 bp) and over the additional IS26, linking the fragments of the *asr* gene together (RH2005-RH2006, 1,151 bp) (Fig 6.7A and 6.7C). However, when the orientation of contig 37 was inverted as in Fig 6.7B, which is possible due to the opposite orientation of the two bounding IS26 sequences, amplicons of the expected size using RH831 and RH2005 (1,203 bp) and RH2015 and RH2006 (1,524 bp) were also recovered (Fig 6.7B and C).

A.



B.



C.

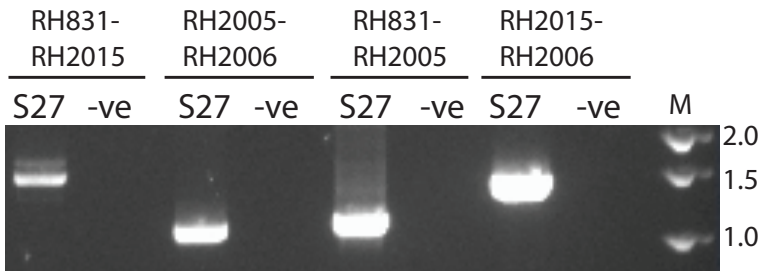


Figure 6.7 Assembly of AbGRI3 in S27. A. Assembly PCRs performed. B. Potential inversion in the sequence surrounding AbGRI3. C. Assembly AbGRI3 in S27. Arrows indicate the extent and orientation of genes and open reading frames. The *aphA1b* gene is red-orange and the *aadA1*, *catB8* and *aacA4* genes are green, pink and pale blue, respectively. Adjacent chromosomal genes are shown in grey. Boxes represent IS and CR elements and the internal arrows indicate the transposase or the rolling circle replicase. IS26 is dark green, ISAb24 is purple, ISEc29 and ISEc28 are light and dark blue and CR1 is yellow. IS26** is 2 bp different to the standard sequence of IS26, while IS26* is 3 bp different. Black and white flags indicate matching target site duplications. The primers used to link sequences are shown above by the black bars, along with PCR amplicon size. The extent of contig 37 is indicated by the grey line. In C, primer combinations are shown above the gel and the isolates are shown above their lanes. The primer combinations of RH831 and RH2015, and RH2005 and RH2006, screen for contig 37 in one orientation while RH831 and RH2005, and RH2015 and RH2006, screen for contig 37 in the other orientation. Isolate numbers are shown above the lanes and sizes of the molecular marker on the right are in kb.

The amplicons from the reactions of both orientations of contig 37 were sequenced and regardless of the orientation of contig 37, the right IS26 of Tn6179 is IS26** as expected, as is the additional IS26 downstream of the resistance island.

Due to the mechanisms of IS26 transposition, it is very difficult to introduce just a single IS26 into a sequence. It is likely that through the replicative transposition pathway of IS26, the IS26** to the right of Tn6179 has caused an inversion through intramolecular transposition to a site in the *asr* gene, duplicating the 8 bp target site and the IS26 involved (Refer to Chapter 1; Fig 1.13). The identical sequence of the instigating and duplicated IS26 supports this mechanism. Hence, it is most likely that the 3.1 kb of chromosomal sequence of contig 37 was originally in inverse orientation as shown in Fig 6.8 and as this is the only difference between AbGRI3 in S27 and AbGRI3-2, it is called AbGRI3-2i.

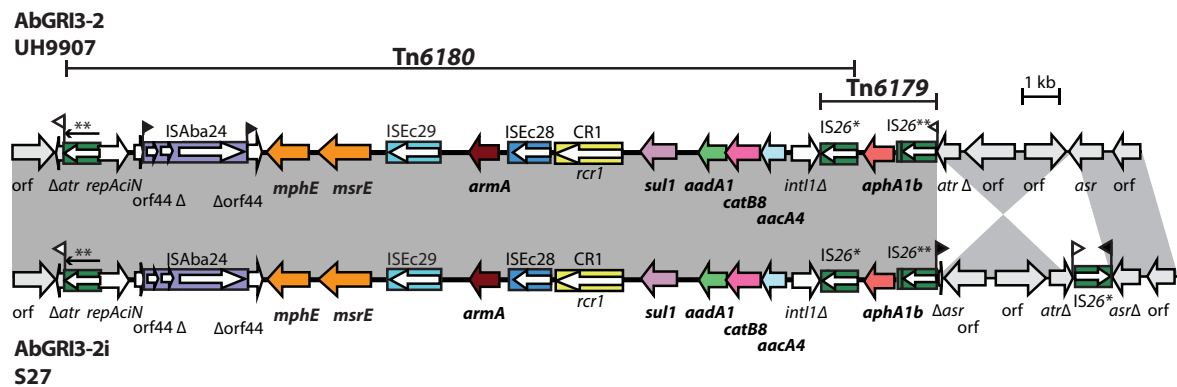


Figure 6.8 Comparison of AbGRI3-2 and 2i. Arrows indicate the extent and orientation of genes and open reading frames. The *repAciN* gene is blue, the *mphE* and *msrE* genes are orange, the *armA* gene is maroon, the *sul1* gene is mauve, the *aphA1b* gene is red-orange, and the *aadA1*, *catB8* and *aacA4* genes are green, pink and pale blue, respectively. Adjacent chromosomal genes are shown in grey. Boxes represent IS and CR elements and the internal arrows indicate the transposase or the rolling circle replicase. IS26 is dark green, ISAb24 is purple, ISEc29 and ISEc28 are light and dark blue and CR1 is yellow. IS26** is 2 bp different to the standard sequence of IS26, while IS26* is 3 bp different. Black and white flags indicate matching target site duplications. Grey boxes indicate matching sequence.

As the IS26 elements bounding Δasr and *atr* Δ segment are in opposite orientation, the internal sequence could be flipping backwards and forwards due to either homologous recombination or the action of IS26. To attempt to determine the major direction of this segment, boiled DNA preparations for 10 single colonies of S27 were checked for amplification of both orientations. Amplicons were detected for both orientations of this sequence in each colony. Though a possible explanation for the detection of amplicons in both orientations could be due to PCR

artefacts, this is unlikely as other segments that have been inverted did not produce amplicons of the re-inverted orientation (for example, see Section 6.3.5.2). Instead, a portion of the population could be in the re-orientated state. It is unlikely that that this is homologous recombination, as it does not occur at a high enough frequency, and the action of IS26 is possible. Cloning of this fragment or long-range sequencing such as Nanopore or PacBio could help unambiguously assign the orientation of this segment, if that is possible.

6.3.2 AbGRI3-3i in S38 and S39

S38 and S39 each had four identical contigs with fragments of IS26 that remained following AbGRI2 assembly (Table 6.2). One of these was the 3.3 kb *atr-asr* fragment found in S27.

Table 6.2 S38 and S39 contigs with IS26 fragments

Contig (size kb)		Content	Edges	
S38	S39		Left	Right
10 (133.6)	9 (133.6)	LH flank	ND	Δatr , IS26-IRr
17 (13.8)	15 (13.8)	$\Delta Tn6180$	IS26-IRl, $\Delta ISAb24$	<i>aacA4</i> , IS26-IRr
44 (3.3)	50 (3.3)	RH flank	IS26-IRl <i>atr</i> Δ ,	Δasr , IS26-IRl
16 (175)	19 (175)	RH chromosome	IS26-IRr, <i>asr</i> Δ^1	ND

¹126 bp closer to the start of the gene than in contig 5 of S27.

The contig carrying *armA* was shorter than in S27 and a comparison revealed that there were deletions at both ends of *Tn6180* in these isolates requiring new primers (RH2010 at the left end and at the right end RH2011) to be designed for assembly (Fig 6.9A). In addition, *Tn6179* is not present.

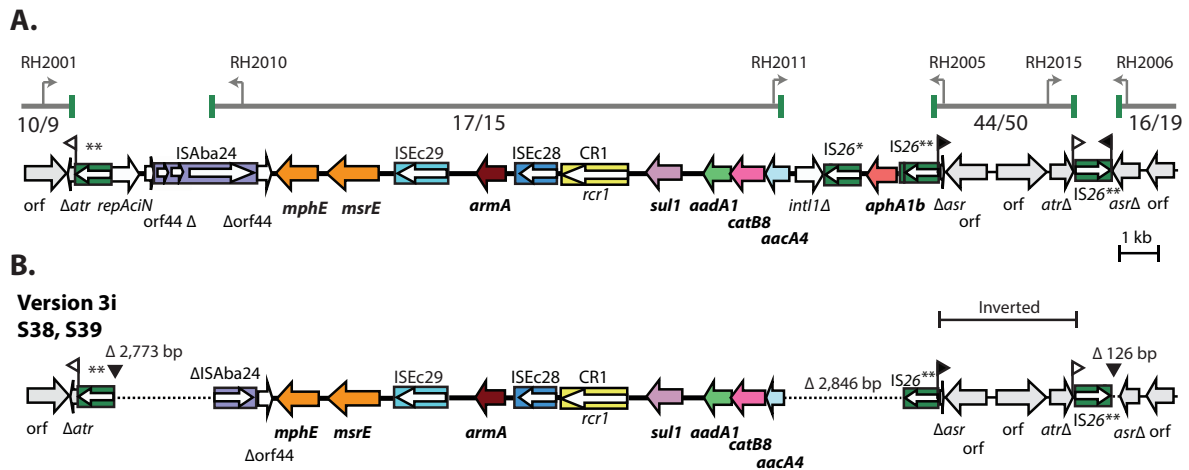


Figure 6.9 Assembly of AbGRI3-3i in S38 and S39. A. AbGRI3 contigs in S38 and S39. Arrows indicate the extent and orientation of genes and open reading frames. The *mphE* and *msrE* genes are orange, the *armA* gene is maroon, the *sul1* gene is mauve, the *aphA1b* gene is red-orange, and the *aadA1*, *catB8* and *aacA4* genes are green, pink and pale blue, respectively. Adjacent chromosomal genes are shown in grey. Boxes represent IS and CR elements and the internal arrows indicate the transposase or the rolling circle replicase. IS26 is dark green, ISAbA24 is purple, ISEc29 and ISEc28 are light and dark blue and CR1 is yellow. IS26** is 2 bp different to the standard sequence of IS26, while IS26* is 3 bp different. Black and white flags indicate matching TSD. Grey lines above indicate the extent of different contigs in S38 and S39, green bars indicate fragments of IS26 and grey arrows indicate the binding sites and orientations of primers. B. Structure of AbGRI3-3i. Dashed lines represent the extent of deletions relative to AbGRI3-2i with the deletion size indicated above.

The chromosome on the left of AbGRI3 (contig 10 in S38 and 9 in S39) was linked to the Δ ISAbA24 end of the truncated Tn6180 (contig 17/15) using RH2001 and RH2010, which produced an amplicon slightly larger than 1.5 kb, consistent with the predicted size of 1,605 bp. The other end of Tn6180, also contained a deletion and was linked to the Δ asr end of contig 44/50 using RH2011 and RH2005. This combination is predicted to produce a 1,007 bp product and the amplicon was approximately 1 kb in size. The Δ atr end of contig 44/50 was linked to contig 16/19 using RH2015 and RH2006 which should produce an amplicon of 1,398 bp, and the product just under 1.5 kb. This version of AbGRI3 has been called AbGRI3-3i (Fig 6.9B). As in the assembly of AbGRI3-2i, the 3.1 kb segment containing Δ asr and Δ atr could be detected in both orientations (i.e. amplicons of the expected size for RH2011-RH2015 and RH2005-RH2006 were also recovered). Amplicon sequence revealed the presence of IS26 isoforms indicated in Fig 6.9B.

AbGRI3-3i is easily derived from 2i by the action of IS26. The IS26 on the left of Tn6180 has caused a 2,773 bp deletion that extends into ISAb24 (776/2,421 bp remaining). The IS26 on the right of Tn6179 has removed *aphA1b*, the remainder of the *intI1* gene, the *attI1* site and 7 bp of the *aacA4* gene cassette, leaving the *aacA4* reading frame intact. It is likely that without the promoter in *intI1*, the cassettes are no longer expressed but this is unable to be distinguished in the phenotype of these isolates due to the presence of *armA*. The extra IS26** has also removed 126 bp of *asrΔ* (Fig 6.9B).

6.3.3 AbGRI3-4 in S34

The draft genome of S34 had only two contigs ending with an IS26 fragment unaccounted for, contig 17 (122 kb) containing the chromosomal sequence from the left flank of AbGRI3-2i and contig 24 (122 kb), which contained the *armA* gene and chromosomal sequence from the right-hand side (Fig 6.10A). They were linked using RH2001 and RH2002, as above. Examination of contig 24 revealed sequence matching Tn6180 up to 696 bp into ISEc28 (201 bp missing) directly adjacent to sequence matching chromosomal sequence located 3,112 bp away from the resistance island (Fig 6.10A). The correct assembly of this contig was confirmed by sequencing the amplicon (predicted size 1,934 bp) produced by the primer combination of RH2012 and RH2014. This version of AbGRI3, called AbGRI3-4, has lost 7,649 bp of Tn6180 as well as Tn6179 and 4,765 bp of the right flanking chromosome relative to AbGRI3-2i (Fig 6.10B). An alignment of the sequence of ISEc28, the region surrounding the recombination patch of contig 24 of S34 and the chromosomal sequence downstream of the island revealed that there was a 4 out of 5 bp identity (boxed in pink in Fig 6.10C) which could have facilitated the event that caused the deletion.

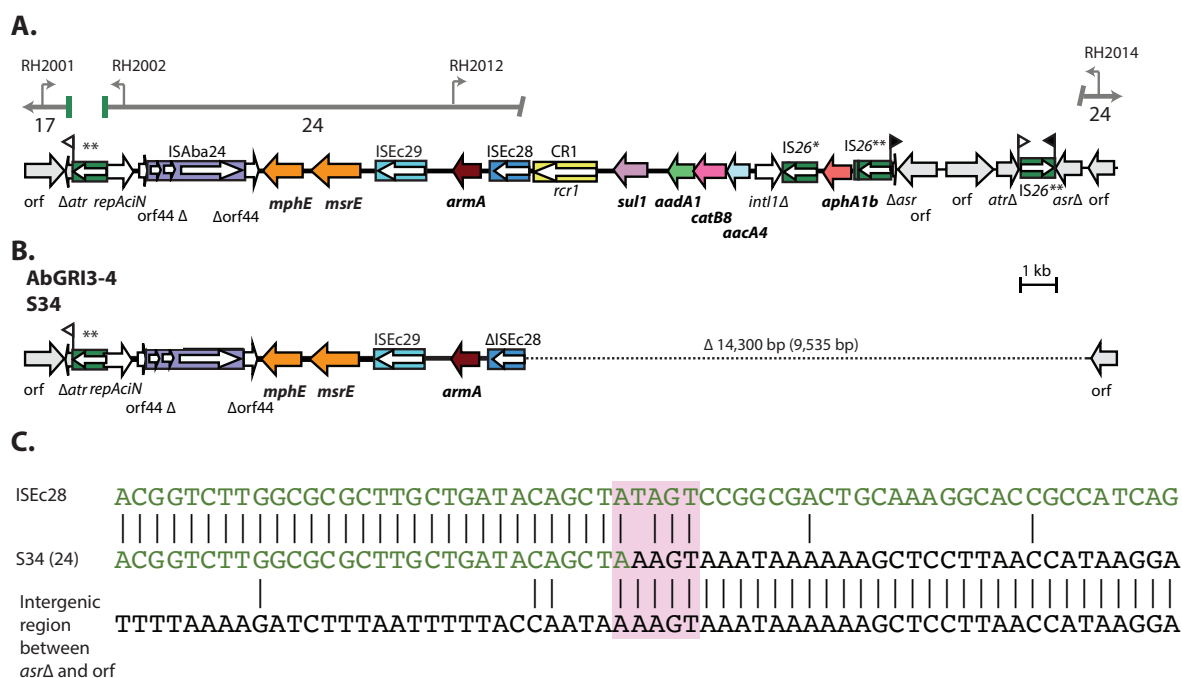


Figure 6.10 Assembly of AbGRI3-4. A. AbGRI3 contigs in S34. Arrows indicate the extent and orientation of genes and open reading frames. The *mphE* and *msrE* genes are orange, the *armA* gene is maroon, the *sul1* gene is mauve, the *aphA1b* gene is red-orange, and the *aadA1*, *catB8* and *aacA4* genes are green, pink and pale blue, respectively. Adjacent chromosomal genes are shown in grey. Boxes represent IS and CR elements and the internal arrows indicate the transposase or the rolling circle replicase. IS26 is dark green, ISAb24 is purple, ISec29 and ISec28 are light and dark blue and CR1 is yellow. IS26** is 2 bp different to the standard sequence of IS26*, while IS26 is 3 bp different. Black and white flags indicate matching TSD. Grey lines above indicate the extent of different contigs in 34, green bars indicate fragments of IS26 and grey arrows indicate the binding sites and orientations of primers. B. Structure of AbGRI3-4. Dashed lines represent the extent of deletions relative to AbGRI3-2i, with the total size of the deletion indicated above and the missing portion of AbGRI3-2 in brackets. C. Putative site of recombination AbGRI3-4. Green nucleotides indicate sequence of ISec28 and the pink box highlights the 4/5 bp that could have facilitated the deletion event.

6.3.4 Prediction of AbGRI3 variant forms in isolates not sequenced by Illumina HiSeq

Mapping strategies used to assemble AbGRI3 in the isolates sequenced by Illumina HiSeq were used to predict the structures of AbGRI3 in eight isolates prior to the availability of sequence. The mapping strategies differed slightly depending on whether the isolates also had an *aphA1b* gene that was not accounted for in their AbGRI2 variant (Table 6.3 or Table 6.4). The isolates S22, S23 and S31 have an *aphA1b* gene that is likely to be located in this island (Table 6.4).

Isolate	Δatr (LHC) to <i>repAciN</i> RH2001-RH2002 (1,323 bp)	AbGRI3-4 RJ RH2012-RH2014 (1,934 bp)	Tn6180 to Δasr RH2003-RH2005 (1,193 bp)	Δatr to Δasr RH2015-RH2006 (1,524 bp)	AbGRI3 version
S25	-	-	-	-	?
S28	+	-	+ ¹	+ ¹	1i
S32	+	+	ND	ND	4
S33	+	+	ND	ND	4
S35	+	+	ND	ND	4
S36	+	+	ND	ND	4
S37	+	+	ND	ND	4
S40	+	+	ND	ND	4

The isolate S25 did not produce an amplicon for any of the mapping PCRs and the assembly of its AbGRI3 variant in described is Section 6.3.5.2. However, six of the eight isolates without the *aphA1b* gene contained AbGRI3 variant 4 (Table 6.3). The isolate S28 harboured an AbGRI3 variant containing only Tn6180, like AbGRI3-1 found in TYTH-1 and MDR-TJ. However, this isolate also contains the additional IS26 interrupting the *asr* gene that has caused an inversion of 3.1 kb of chromosomal sequence. Therefore, the variant in S28 was called AbGRI3-1i (Fig 6.11). As the additional IS26 is IS26**, it is likely that previously Tn6179 was present in this isolate to provide the IS26** for the intramolecular inversion reaction. The shared IS26* of Tn6180 and Tn6179 could then have caused a deletion, removing the Tn6179 TU (*aphA1b* and one IS26).

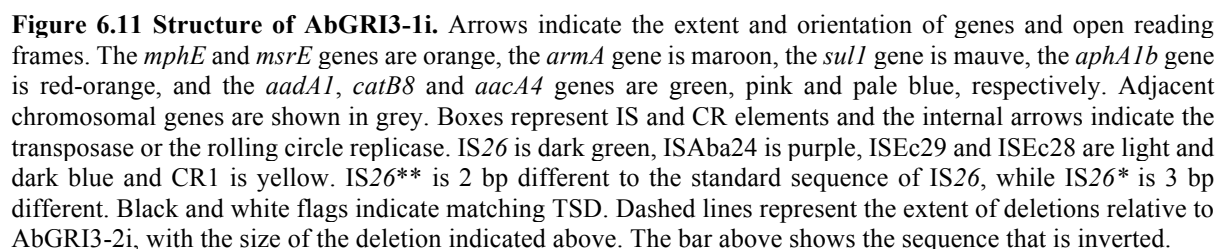


Table 6.4 PCR screening of AbGRI3 in isolates containing *armA* and *aphA1b*

Isolate	Δatr (LHC) to <i>repAciN</i> RH2001- RH2002 (1,323 bp)	Δatr (LHC) to $\Delta ISAb24$ RH2001- RH2010 (1,650 bp)	Tn6180 to Tn6179 RH2003- RH880 (1,934 bp)	Tn6179 to Δasr RH831- RH2005 (1,203 bp)	Δatr to <i>asr</i> Δ RH2015- RH2006 (1,524 bp)	AbGRI3 version
S22	+	ND	+	+ ¹	+ ¹	2i
S23	+	ND	+	+ ¹	+ ¹	2i
S31	-	+ ²	+	+ ³	-	?

¹ Amplicons of the expected size were also produced using RH831-RH2015 (1,576 bp) and RH2005-RH2006 (1,151 bp).

² Amplicon was just larger than 2 kb.

³ Amplicon on the expected size for the S38/S39 deletion in *asr* Δ was also produced using RH2005-RH2006.

AbGRI3-2i, with the adjacent inversion, was also detected in S22 and S23 (Table 6.4). The AbGRI3 of S31 was able to be partially assembled using mapping PCRs (Table 6.4) and contained a Δ Tn6180 and Tn6179. The boundary between the chromosome on the left and Tn6180 was amplified using the same primers as for AbGRI3-3i (RH2001-RH2010, see Fig 6.9), which has a deletion at the left end of Tn6180. However, the product for S31 was larger than for AbGRI3-3i. Sequencing of the amplicon revealed that S31 also had a deletion at the end of Tn6180 but had retained 541 bp more of ISAb24 (1,316 of the full 2,421 bp ISAb24) than AbGRI3-3i (Fig 6.12).

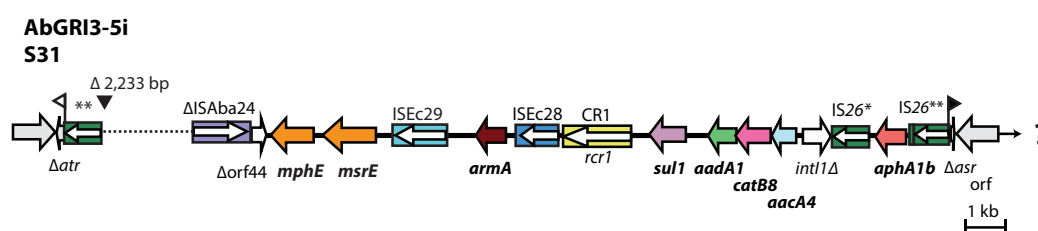


Figure 6.12 Assembly of AbGRI3-5i. Arrows indicate the extent and orientation of genes and open reading frames. The *mphE* and *msrE* genes are orange, the *armA* gene is maroon, the *sul1* gene is mauve, the *aphA1b* gene is red-orange, and the *aadA1*, *catB8* and *aacA4* genes are green, pink and pale blue, respectively. Adjacent chromosomal genes are shown in grey. Boxes represent IS and CR elements and the internal arrows indicate the transposase or the rolling circle replicase. IS26 is dark green, ISAb24 is purple, ISEc29 and ISEc28 are light and dark blue and CR1 is yellow. IS26** is 2 bp different to the standard sequence of IS26, while IS26* is 3 bp different. Black and white flags indicate matching TSD. Dashed lines represent the extent of deletions relative to AbGRI3-2i, with the total size of the deletion indicated above. The bar above shows the sequence that is inverted.

The other end of Tn6179 could be joined to Δasr (RH831-RH2005) but $asr\Delta$ could also be linked to Δasr (RH2005-RH2006) and the amplicon produced was the same size as the deletion AbGRI3-3i. However, no amplicon was produced when RH2015 and RH831 or RH2015 and RH2006 were paired, indicating that a deletion may be present in S31 which has removed the binding site of RH2015.

6.3.5 Final assembly using the draft genomes.

The AbGRI3 regions mapped by PCR were compared to the draft genomes of these isolates when the Illumina MiSeq data became available. The draft genomes were consistent with the AbGRI3 regions assembled in S22, S23, S28, S32, S33, S35, S36, S37 and S40. Additional information from the draft genomes of S31 and S25 facilitated assembly of AbGRI3 in these isolates.

6.3.5.1 AbGRI3-5i in S31

Examination of the draft genome of S31 revealed that there was a 785 bp deletion, likely caused by IS26, that had removed $atr\Delta$ and the binding site for RH2015. Amplification, using a new primer RH2016 in combination with RH2006 (Fig 6.13A), completed the assembly of a novel variant named AbGRI3-5i (Fig 6.13B), and revealed that 126 bp of $asr\Delta$ was also missing, likely caused by a deletion by the adjacent IS26**. The segment between the right of the island and the extra IS26** could also be detected in both orientations, despite containing a 785 bp deletion at one end (Fig 6.13B).

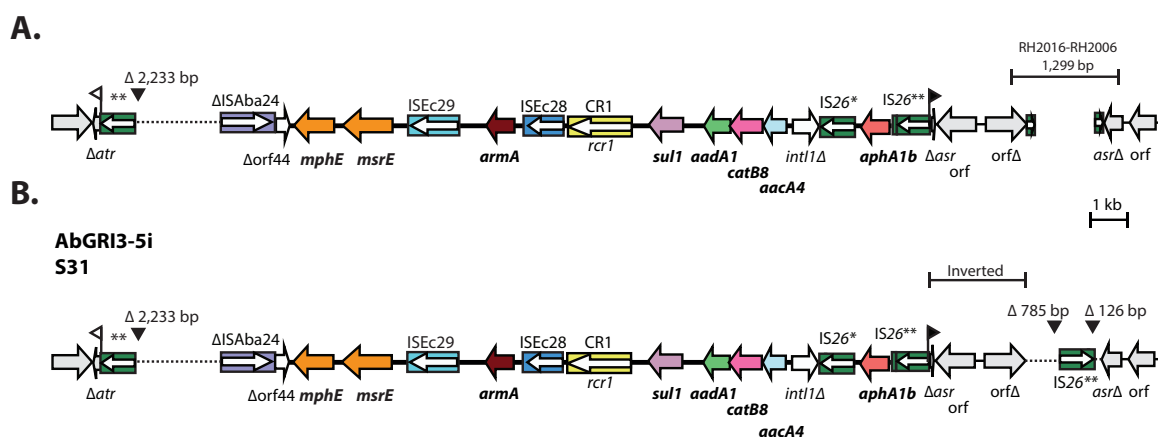


Figure 6.13 AbGRI3-5i in S31. A. Final link to assemble AbGRI3 in S31. B. Structure of AbGRI3-5i. Arrows indicate the extent and orientation of genes and open reading frames. The *mphE* and *msrE* genes are orange, the *armA* gene is maroon, the *sul1* gene is mauve, the *aphA1b* gene is red-orange, and the *aadA1*, *catB8* and *aacA4* genes are green, pink and pale blue, respectively. Adjacent chromosomal genes are shown in grey. Boxes represent IS and CR elements and the internal arrows indicate the transposase or the rolling circle replicase. IS26 is dark green, ISAba24 is purple, ISEc29 and ISEc28 are light and dark blue and CR1 is yellow. IS26** is 2 bp different to the standard sequence of IS26, while IS26* is 3 bp different. Black and white flags indicate matching TSD. Dashed lines represent the extent of deletions relative to AbGRI3-2i, with the total size of the deletion indicated above. In A, the bar above indicated the primers along with amplicon size while in B the bar above shows the sequence that is inverted.

6.3.5.2 AbGRI3-6 in S25

A search in the draft genome of S25 retrieved 4 contigs containing fragments of IS26 (Table 6.5) which were tiled to AbGRI3-2i and surrounds (Fig 6.14A). Tn6180 was separated into two contigs, 56 and 31, which had IS26 on both edges, indicating the presence of an additional IS26 within the transposon. However, no sequence matching Tn6179 or sequence containing *Δatr-asr* were present.

Table 6.5 AbGRI3 contigs of S25

Contig	Length	Features	Edges	
			Left	Right
14	133826	LH flank	NA	IS26-IRr
56	559	<i>ΔrepAciN</i>	IS26-IRl	IS26-IRl
31	16187	<i>ΔTn6180</i>	IS26-IRr	IS26-IRl
9	175053	RH chromosome	IS26-IRr	NA

Results of previous screening indicated that contig 56 and 14 with sequence of chromosome to the left of the island could not be linked using RH2001 and RH2002 (Table 6.3, Fig 6.14B).

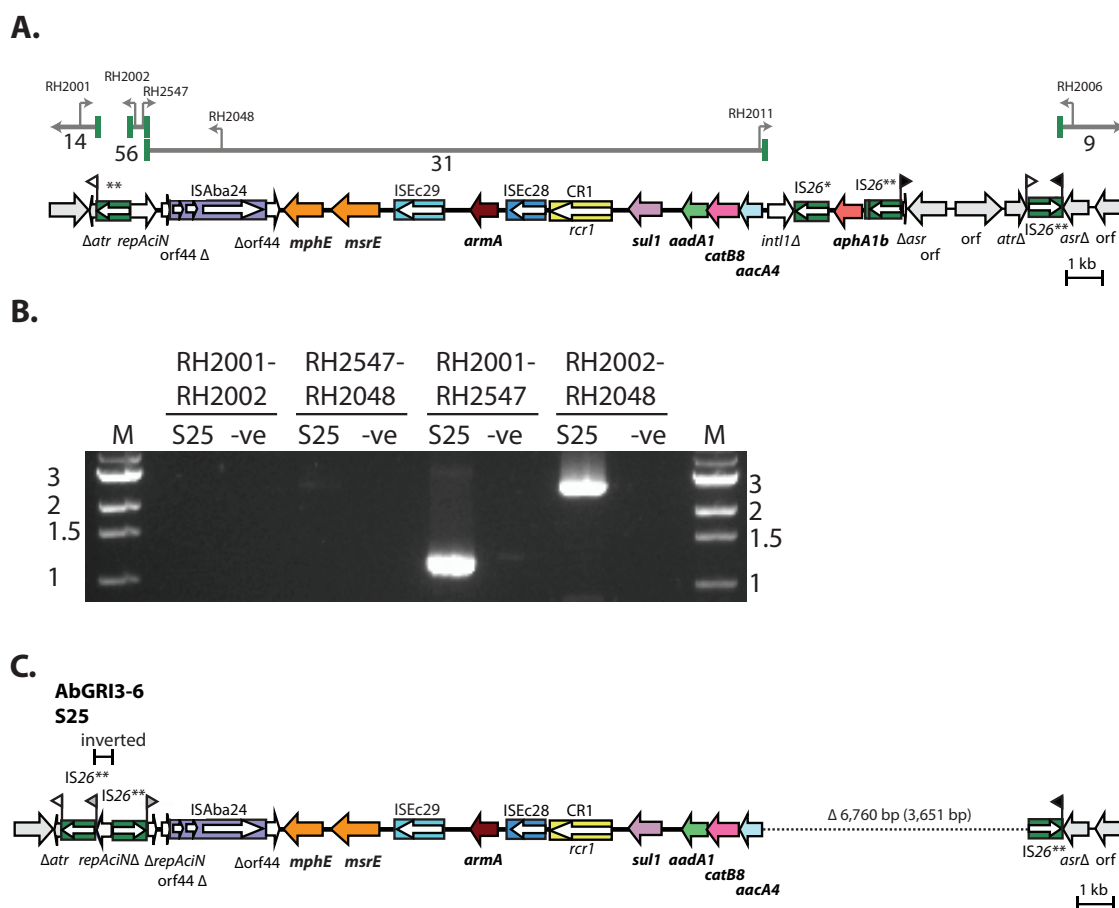


Figure 6.14 AbGRI3-6 in S25. A. AbGRI3 contigs in S25. B. Assembly of the left end of AbGRI3-6. C. Structure of AbGRI3-6. For A. and C. arrows indicate the extent and orientation of genes and open reading frames. The *repAciN* gene is blue, the *mphE* and *msrE* genes are orange, the *armA* gene is maroon, the *sul1* gene is mauve, the *aphA1b* gene is red-orange, and the *aadA1*, *catB8* and *aacA4* genes are green, pink and pale blue, respectively. Adjacent chromosomal genes are shown in grey. Boxes represent IS and CR elements and the internal arrows indicate the transposase or the rolling circle replicase. IS26 is dark green, ISAbA24 is purple, ISEc29 and ISEc28 are light and dark blue and CR1 is yellow. IS26** is 2 bp different to the standard sequence of IS26. Black and white flags indicate matching TSD. In A., grey lines above indicate the extent of different contigs in S25, green bars indicate fragments of IS26 and grey arrows indicate the binding sites and orientations of primers. B. depicts the screening PCRs to orient contig 56 in S25. The primer combinations are shown above the gel and the isolates are shown above the lane. The primer combination RH2001-RH2002 and RH2547-RH2048 will amplify contig 56 in one direction, while RH2001-RH2547 and RH2002-RH2048 will amplify it in the other. Sizes of the molecular marker to the left and right are in kb. In C. dashed lines represent the extent of deletions relative to AbGRI3-2i, with the total size of the deletion indicated above and the missing portion of AbGRI3-2 in brackets.

Instead, contig 14 was linked to contig 56 using RH2001 and RH2547, which detect the opposite orientation. This produced an amplicon approximately 1,200 bp in size. (Fig 6.14B). Hence, it is likely that through intramolecular transposition the IS26** at the left end of Tn6180 has targeted a site in *repAciN*, duplicating itself and an 8 bp target site and inverting the intervening sequence (Fig 6.14C).

The other end of 56 was then linked with the remainder of Tn6180 (contig 31) using RH2002 and RH2048, producing an amplicon of a size consistent with the expected 2,638 bp (Fig 6.14B). The *intI1*Δ and all but 8 bp of the *attI* site is missing at the integron end of Tn6180. RH2011 was used to link this end of contig 31 directly to contig 9 which had the right boundary of the additional IS26 in *asr*Δ. Hence, a deletion of 9.7 kb is present in this isolate removing the 3.1 kb of chromosomal sequence, Tn6179 and *intI1*Δ (Fig 6.14C). The AbGRI3-6 variant in S25 can be derived from AbGRI3-2i by the action of IS26.

6.3.6 Assembled AbGRI3 variants

Overall, 6 variants of AbGRI3 were detected in the 15 isolates that were *armA* positive (Fig 6.15). Some of these variants were only found in one isolate (AbGRI3-1i, AbGRI5i and AbGRI3-6) while the other variants were found in two or more isolates. The flanking chromosomal sequence on the left of each variant was always the same but differences have arisen both within the island and in the chromosomal sequence to the right of the resistance island. Most of this variation can be accounted for by the action of IS26 either causing deletions or inversions in adjacent sequences. However, the deletion in AbGRI3-4, which was present in 6 isolates, is due to a different type of event. All variants, except AbGRI3-4 have clearly arisen from AbGRI3-2i as they all contain the IS26** interrupting the *asr* gene.

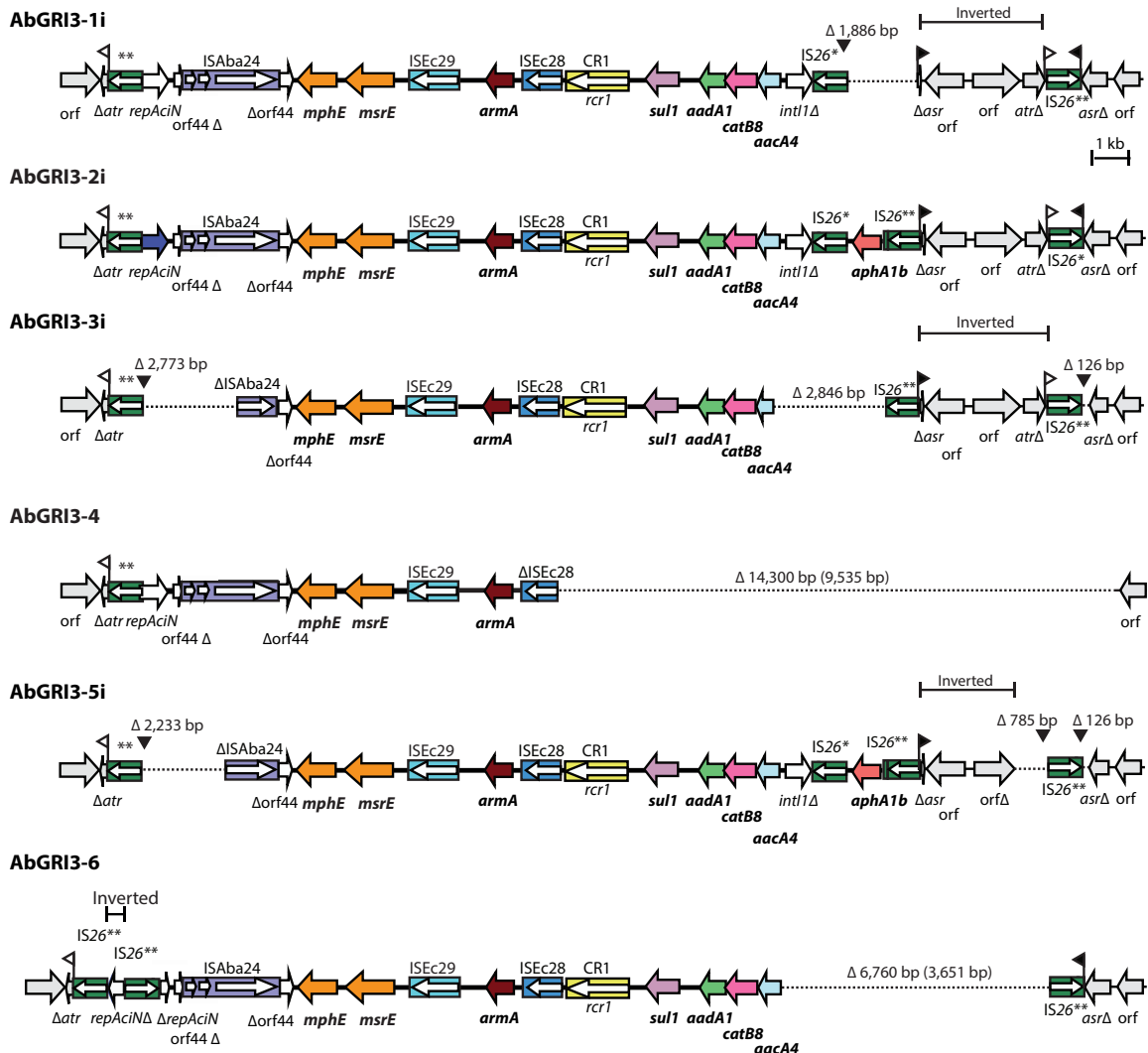


Figure 6.15 Six AbGRI3 variants. Arrows indicate the extent and orientation of genes and open reading frames. The *mphE* and *msrE* genes are orange, the *armA* gene is maroon, the *sul1* gene is mauve, the *aphA1b* gene is red-orange, and the *aadA1*, *catB8* and *aacA4* genes are green, pink and pale blue, respectively. Adjacent chromosomal genes are shown in grey. Boxes represent IS and CR elements and the internal arrows indicate the transposase or the rolling circle replicase. IS26 is dark green, ISAbA24 is purple, ISEc29 and ISEc28 are light and dark blue and CR1 is yellow. IS26** is 2 bp different to the standard sequence of IS26*, while IS26 is 3 bp different. Black and white flags indicate matching TSD. Dashed lines represent the extent of deletions relative to AbGRI3-2i, with the total size of the deletion indicated above and the missing portion of AbGRI3-2 in brackets. The black bar above indicates the extent of inverted sequence.

6.4 AbGRI3 in completed GC2 genomes in GenBank

A search of the non-redundant database of GenBank using the sequence of AbGRI3-2, revealed that in addition to TYTH-1 and MDR-TJ described in Section 6.2.1, a version of AbGRI3 could be discerned in 31 other completed GC2 genomes (Table 6.6). Only two of these have been described previously. The region of AbGRI3 in NCGM_237 containing Tn6180 with deletions

Table 6.6 Completed GC2 genomes containing AbGRI3 variants as of July 2017

Isolate	Year of Isolation	Location	AbGRI3			Accession	Reference
			Chromosome left	Content	Chromosome right		
MDR-ZJ06	2006	Zhejiang, China	+	IS26 ¹	Δ25,935 bp	CP001937	(285)
BJAB07104	2007	Beijing, China	+	IS26 ¹	+	CP003846	(286)
BJAB0868	2008	Beijing, China	+	IS26 ¹	+	CP003849	(286)
MDR-TJ	<2011	Tianjin, China	+	Tn6180	+	CP003500	(125)
TYTH-1	2008	Taiwan	+	Tn6180	+	CP003856	(160)
NCGM_237	2012	Japan	Δ2,814 bp	Tn6180	Δ36,971 bp	AP013357	(253)
XH586	2010	Guangdong, China	+	Tn6180, Tn6179	+	CP014541	
ORAB01	2012	Oregon, USA	+	Tn6180, Tn6179	+	CP015483	
KPN10P02143	2012	Daegu, South Korea	+ ²	Tn6180, Tn6179	+	CP013924	(157)
XH386	2014	Guangzhou, China	+	Tn6180, Tn6179	+	CP010779	(62)
DU202	<2014	Korea	+	Tn6180, Tn6179	+	CP017152	(155)
CMC-CR-MDR-Ab4	<2016	USA	+	Tn6180, Tn6179	+	CP016295	
XDR-BJ83	2007	Beijing, China	+	Tn6180, Tn6179	+	CP018421	
AF-673	2008	Arizona, USA	+	Tn6180, Tn6179	+	CP018256	
CMC-CR-MDR-Ab66	<2016	USA	+	Tn6180, Tn6179	Δ243 bp	CP016300	
XDR-BJ83	2007	Beijing, China	+	Tn6180, Tn6179	+	CP018421	
SSA6	2013	Seoul, South Korea	+	Tn6180 ³ , Tn6179	Δ35,691 bp	CP020591	
CMC-MDR-Ab59	<2016	USA	+	Tn6180, Tn6179x2	+	CP016298	
XH857 ⁴	2010	Guangdong, China	+	Tn6180, Tn6179	Δ18,304 bp	CP014540	
HRAB-85	2014	Beijing, China	+	ΔTn6180 ⁵ , Tn6179	+	CP018143	
YU-R612	2014	Seoul, Korea	Δ2,850 bp	2 copies of ΔTn6180 ⁶	+	CP014215	(156)
USA2	2013	Seoul, South Korea	+	ΔTn6180 ⁷	Δ1,155 bp	CP020592	
JBA13	2013	Jeonbuk, South Korea	Δ1,004 bp ⁸	ΔTn6180 ⁷	+	CP020584	
SSA12	2013	Seoul, South Korea	Δ1,004 bp ⁸	ΔTn6180 ⁷	+	CP020578	
KAB07	2015	Jeonbuk, South Korea	Δ1,004 bp ⁸	ΔTn6180 ⁷	+	CP017654	
15A34	2014	South Korea	+	Tn6180 ^{3,9}	Δ35,691 bp ¹⁰	CP020590	
SAA14	2013	Seoul, South Korea	+ ¹¹	ΔTn6180 ^{7,12}	+	CP020579	
KAB04	2015	Busan, South Korea	+ ¹¹	ΔTn6180 ¹³	+	CP017648	
KAB02	2015	Daegon, South Korea	Δ2,850 bp ¹⁴	ΔTn6180 ^{7,15}	Δ1,1391 bp	CP017644	
15A5	2013	South Korea	+ ¹⁶	ΔTn6180 ⁷	+	CP020574	

SSMA17	2013	Seoul, South Korea	$\Delta 3,286$ bp ¹⁷	Δ Tn6180 ⁷	+	CP020581
KAB06	2015	Daegu, South Korea	$\Delta 4,203$ bp	Δ Tn6180 ^{7, 18}	$\Delta 312$ bp	CP017652
KAB05	2015	Daegu, South Korea	$\Delta 4,203$ bp	Δ Tn6180 ^{7, 18}	$\Delta 312$ bp	CP017650
KAB08	2015	Gyeonggi, South Korea	+	Δ Tn6180 ¹⁹	+	CP017656
KAB03	2015	Daejeon, South Korea	$\Delta 785$ bp	version 4 ²⁰	$\Delta 3,937$ bp	CP017646
KAB01	2015	Daejeon, South Korea	$\Delta 785$ bp	version 4 ²¹	$\Delta 3,937$ bp	CP017642

¹These genomes include a Tn6180-Tn6179 TU form said to be a plasmid.

²2,858 bp has been inverted by intramolecular transposition of IS26.

³Missing the 19 duplication in *attII*.

⁴Is a triple locus variant of ST2 (Institute Pasteur scheme).

⁵Missing 2,013 bp at the left end.

⁶First copy of Δ Tn6180 is missing 4,729 bp from the left end and is missing 772 bp from the right end, and the second copy of Δ Tn6180 is also missing 4,729 bp from the left but instead is only missing 434 bp from the right end.

⁷Missing 434 bp at the right end.

⁸Inversion of 1,854 bp caused by intramolecular transposition of IS26.

⁹1,035 bp inverted at the left end due to intramolecular transposition of IS26.

¹⁰Located approximately 2.5 Mb away, due to an IS26 mediated inversion.

¹¹An IS26 surrounded by TSD is located 2,858 bp away.

¹²Missing 1,929 bp at the left end.

¹³Missing 2,013 bp at left end, followed by a 2,165 bp inversion and another 156 bp deletion.

¹⁴IS26 is in the opposite orientation.

¹⁵4,250 bp missing at the left.

¹⁶2,505 bp has been inverted caused by intramolecular transposition of IS26.

¹⁷Located approximately 1 Mb away, due to an IS26 mediated inversion.

¹⁸Missing 7,693 bp from left end.

¹⁹Missing 678 bp left end, 356 bp from the right and 37 bp of IS26 adjacent to the right IS26.

²⁰Missing 10 bp at the left end.

²¹Missing 6,436 bp at the left end.

in both flanking sequences (253) and YU-R612 (156) includes two partial copies of a Δ Tn6180 with a deletion in the flanking sequence to the left. The remainder of AbGRI3 variants in Table 6.6 were analysed in this study. Overall, 14 isolates contained a variant related to AbGRI3-2, i.e. had all or part of both Tn6180 and Tn6179, 16 isolates had an island related to AbGRI3-1, i.e. with only Tn6180, and KAB03 and KAB01 from South Korea carried AbGRI3-4. However, deletions and/or inversions caused by IS26, have contributed to the large variation both within Tn6180 and Tn6179, as well as in the surrounding chromosome (Table 6.6). However, none of the genomes listed in Table 6.6 contain an IS26 interrupting the *asr* gene. The location of this IS26 and the resulting difficulty in assigning the direction of the intervening sequence appears to be specific to the isolates from SGH.

6.4.1 Examination of the chromosome in MDR_ZJ06, BJAB0868 and BJAB07104 at the AbGRI3 insertion site

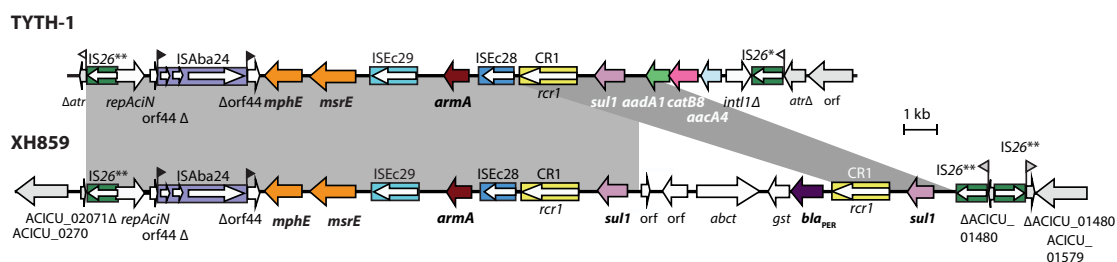
The region of the chromosome where AbGRI3 is located was examined in MDR_ZJ06, BJAB0868 and BJAB07104, which were claimed to contain plasmids composed of Tn6180 and Tn6179 (Section 6.2.3). In the genomes of BJAB0868 and BJAB07104, an IS26 is present in the location of AbGRI3 interrupting the *atr* gene and surrounded by the same 8 bp TSD. MDR-ZJ06 also has a suitably located IS26 but 25.9 kb of adjacent chromosomal sequence is missing. Thus, it is likely that the proposed plasmid forms actually belong in the chromosome, as AbGRI3, and this should be investigated.

6.4.2 An independent acquisition of a Tn6180 variant in a GC2 isolate

A transposon related to Tn6180 was also identified in the GC2 isolate XH859 from China (GenBank accession number CP014539). The transposon in XH859 had ISAba24 in the same position as in Tn6180, interrupting orf44. However, the *intI1* gene has been lost and there is an

additional segment containing the *bla*_{PER} resistance gene (resistance to extended spectrum cephalosporins) downstream of the *sulI* gene in XH859, followed by a second copy of CR1 and *sulI*, then up to 8 bp into the *aadA1* gene cassette before the right IS26** (bottom line, Fig 6.16A).

A.



B.

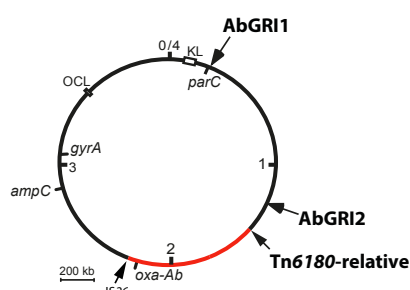


Figure 6.16 Variant of Tn6180 found in XH859. A. Comparison of Tn6180 and the *armA* containing transposon in XH859. Arrows indicate the extent and orientation of genes and open reading frames. The *mphE* and *msrE* genes are orange, the *armA* gene is maroon, the *sulI* gene is mauve, the *aphA1b* gene is red-orange, the *bla*_{PER} gene in dark purple and the *aadA1*, *catB8* and *aacA4* genes are green, pink and pale blue, respectively. Adjacent chromosomal genes are shown in grey. Boxes represent IS and CR elements and the internal arrows indicate the transposase or the rolling circle replicase. IS26 is dark green, ISAb24 is purple, ISEc29 and ISEc28 are light and dark blue and CR1 is yellow. IS26** is 2 bp different to the standard sequence of IS26. Black and white flags indicate matching TSD. Grey boxes indicate regions of identical sequence. B. Schematic of XH859 chromosome. The inverted section of the chromosome is in red and insertions are indicated by arrows. Figure made from XH859 (GenBank accession number CP014539).

The *atr* gene of XH859 is intact so this transposon is not in the same location as AbGRI3. Instead, the chromosomal sequence flanking the right end of the Tn6180 relative in XH859 is located 280 kb from the AbGRI2 insertion location (Fig 6.16B). The sequence that would be expected at the left of the island was adjacent to an IS26, 600 kb away. The IS26 at the left of the transposon in XH589 has likely caused an inversion of chromosomal sequence through intramolecular transposition. Unfortunately, the TSD that would have arisen from this event is not present due to more than one deletion events. Following the *in-silico* reversal of this inversion, the location for the original insertion of the Tn6180-relative in the GC2 chromosome

can be narrowed down to a 13 kb region. Nonetheless, it represents a second acquisition of transposon related to Tn6180 in GC2 isolates.

6.5 Discussion

AbGRI3 is the third and most recently acquired resistance island in GC2 *A. baumannii*. Evidence suggests that the AbGRI3 island, bringing with it the *armA* resistance gene, entered the GC2 chromosome in the early 2000s, likely in Asia. The isolates examined in this study, S22 and S23 from 2004, containing AbGRI3-2i that include both Tn6180 and Tn6179, are now the earliest sequenced examples of GC2 *A. baumannii* containing this third resistance island. The origin of these isolates in Singapore is consistent with the original acquisition of this island by a GC2 isolate in Asia. The *armA* gene has been detected by PCR in earlier isolates from 2003 in Korea. (153). However, as the location of the gene was not determined, it is not known if *armA* was located in AbGRI3 or within the related transposon in XH859 that is found in the second location in the GC2 chromosome. In addition, the *armA* gene in another transposon related to Tn1548, has been found in a set of large conjugative plasmids that have no known *rep* gene (135, 240, 272). This Tn1548-like transposon carries the *cmlA5-arr2* gene cassettes (chloramphenicol and rifampicin resistance) as well as the *bla_{PER}* gene (resistance to extended spectrum cephalosporins) (135, 240, 272) and is located within another transposon in these plasmids. Current whole genome sequence data indicates that the AbGRI3 island entered the GC2 chromosome prior to 2004 but additional information about isolates from the early 2000s should enable this date to be refined.

The presence of an 8 bp target site duplication in the surrounding sequence of the resistance island indicates that, like AbGRI2, AbGRI3 also entered the chromosome by the action of IS26. What is less clear, is if Tn6180 and Tn6179 entered together or individually. Both routes are

plausible using either the replicative or the conservative mechanism of transposition of IS26. A translocatable unit (TU) made up of a circular piece of DNA containing both Tn6180 and Tn6179 with only two IS26s could have inserted into the *atr* gene, duplicating the target site and one copy of IS26**. The plasmid pMUR050, an IncN plasmid isolated in Spain in 2002 (83), contains a transposon related to Tn1548 along with Tn6179. The transposons also overlap in pMUR050, but share the IS26 at the *repAciN* end of the Tn1548-like transposon rather than the *intI1* end (top line Fig 6.17). The *armA* containing transposon of pMUR050 is also missing 216 bp of *intI1*, the same truncation as found in Tn6180, but it only contains a single gene cassette *aadA2* in its integron. Because of this, it was investigated as a potential source for both transposons of AbGRI3-2.

Hypothetically, homologous recombination between the IS26s at the bounding ends of these transposons in pMUR050 could form a circular TU that was composed of the Tn1548-related transposon and Tn6179 but only two IS26s (Fig 6.17). If the central IS26 instigated the insertion of the TU into the *asr* gene (IS26-2 in Fig 6.17), the order of the transposons would be reversed and the configuration seen in AbGRI3 created. However, the addition of ISAb24 and the switching of the cassettes in the integron to be those in AbGRI3 would still be required, as would the change of IS26 to IS26*.

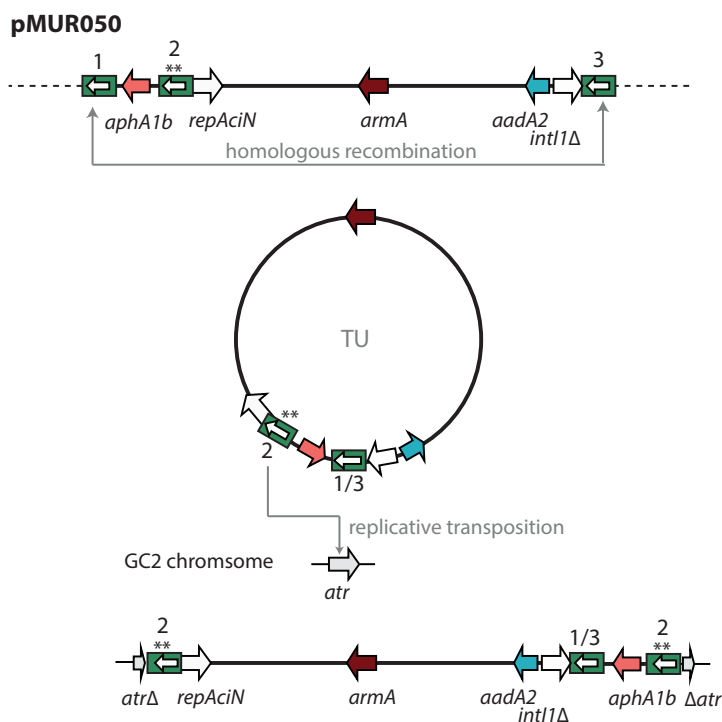


Figure 6.17 pMUR050, a potential source for the transposons in AbGRI3. Arrows indicate the extent and orientation of genes. The *armA* gene is maroon and the *aadA2* and the *aphA1b* genes are aqua and red-orange, respectively. The green boxes represent IS26 and the internal arrow indicates the orientation of the transposase. IS26** is 2 bp different to the standard sequence of IS26. Figure not to scale.

Most of the island variants assembled from the 15 *armA* positive Singapore isolates, also contain an additional copy of IS26**, that is located 3.1 kb away in from the resistance island in the chromosome and the intervening sequence has been inverted. This has likely occurred by intramolecular transposition of the IS26** at the right of Tn6179, which in one step, has duplicated itself and the target site and caused the inversion. A surprising finding was the chromosomal segment that has been inverted could be detected in both orientations using PCR, and further experiments such as cloning the fragment or sequencing using long-range platforms such as PacBio or Oxford Nanopore could be used to determine if this segment is actually inverting frequently. This additional IS26** was not present in any of the genomes examined in GenBank as of July 2017 and appears to be unique to SGH isolates. Hence, its presence and the chromosomal inversion associated with it, can be used to define a sub-lineage of AbGRI3-containing GC2 isolates. As this marker is present in the current earliest known isolates with

AbGRI3 (S22 and S23 from 2004) we can conclude it was acquired soon after the acquisition of AbGRI3. It would be interesting to determine if more recent isolates from SGH are members of this lineage or instead are a part of the original AbGRI3 lineage.

After most of the work in this chapter had been completed, a manuscript that had also analysed the AbGRI3 region of *A. baumannii* from Sweden was published (135). The authors identified a structure identical to AbGRI3-2 structure (including both Tn6180 and Tn6179) in the isolate A071 (GenBank accession number KT317075) and named the complete structure Tn6279 (135). Interestingly, the same study also reported that circular forms of the TU of AbGRI3-2 (Tn6180 and Tn6179 with only two copies of IS26) as well as just AbGRI3-1 (Tn6180 with one copy of IS26) could be detected by PCR in this isolate. These circular TUs are consistent with the “plasmids” described in Section 6.4.1 and could account for the annotation of these sequences as plasmids.

Since its acquisition, the AbGRI3 region has continued to evolve *in situ*. As for AbGRI2, this variation has mainly occurred through deletions occurring adjacent to copies of IS26. Other versions of AbGRI3 found in *armA* positive Singapore isolates are smaller than AbGRI3-2i. They have clearly evolved from AbGRI3-2i, by various deletions of sequence adjacent to copies of IS26, as versions 1i, 3i, 5i all contain the adjacent chromosomal inversion. AbGRI3-6 has also likely evolved from AbGRI3-2i, as the right boundary of this island is in the same position as the additional IS26**. Whether AbGRI3-4 is also derived from AbGRI3-2i cannot be discerned, as the characteristic deletion in AbGRI3-4 extends past the inverted chromosomal segment. This deletion has now been detected in isolates from Singapore, South Korea, India (231) and Sweden (135), suggesting that the sub-lineage containing AbGRI3-4 has dispersed quite successfully.

The 15 isolates that have AbGRI3 do not contain AbGRI2-1. Instead, they have different variants that can be derived AbGRI2-1 (Chapter 5; Fig 5.10) and contain different segments of AbGRI2-1. Hence, it is likely that the cell which first acquired AbGRI3 contained AbGRI2-1, or something very closely related to it, as the *sulI* and *aadA1* genes may already have been missing. Since the acquisition of AbGRI3, deletions in the AbGRI2 region have occurred and in four isolates, no resistance genes have been retained in this region (Table 6.7).

Table 6.7 Features isolates containing AbGRI3

Isolate	Year	ST(Ox)	AbGRI1	AbGRI2	AbGRI3	Additional <i>oxa23</i> locations in the chromosome
S22	2004	218	3	14	2i	AbaR4 (1)
S23	2004	218	3	14	2i	AbaR4 (1)
S25	2006	218	3	14	6	
S27	2007	218	4	IS26a ³	2i	
S28	2007	1128 ¹	8	14	1i	
S31	2008	208	6	13	5i	AbaR4 (2)
S32	2009	195 ¹	7	12b	4	
S33	2009	195 ¹	7	12b	4	
S34	2009	1166 ²	4	12a	4	Tn2006
S35	2010	195 ¹	7	IS26b ⁴	4	
S36	2010	195 ¹	7	IS26b ⁴	4	
S37	2010	195 ¹	7	IS26b ⁴	4	
S38	2011	208	1	13	3i	AbaR4 (2)
S39	2011	208	1	13	3i	AbaR4 (2)
S40	2011	195 ¹	7	12b	4	

¹SLV of ST208 in the *gpi* allele.

²TLV of ST208 in the *cpn60*, *gpi* and *rpoD* alleles.

³A single IS26 is present in the AbGRI2 region and 5,498 bp of the right flank of AbGRI2-1 is missing.

⁴A single IS26 is present in the AbGRI2 region and 30.6 kb of sequence of the left flank and 8.1 kb of the right flank of AbGRI2-1 is missing.

The range of antibiotic resistance phenotypes conferred by genes in the AbGRI3 region replace and extend those conferred by the genes in AbGRI2-1, so deletions in the AbGRI2 region do not contribute to a reduction in phenotypic resistance. Due to extensive variation in AbGRI1 in isolates containing AbGRI3 (Table 6.7), it is less clear what AbGRI1 structure was present in the cell that first acquired AbGRI3, though I postulate that this island likely resembled AbGRI1-3, containing both a copy of Tn6022 and Tn6022Δ, but not Tn2006.

CHAPTER

7

ISAbal mapping

7.1 Introduction

ISAbal is an IS often found associated with *A. baumannii*. ISAbal is 1,180 bp in size and creates a 9 bp TSD. Examination of the distribution of copies of ISAbal in genomes has been shown to be a useful method to discriminate *A. baumannii* isolates (2, 111). Original methods involved digestion of genomic DNA followed by hybridisation with an ISAbal probe and the different profile patterns were indicative of different ISAbal insertion sites (111). However, more recently ISAbal positions have been manually identified in whole genome sequences (2, 277). Two different bioinformatic programs have been developed to automatically identify the insertion positions in genome sequences, ISMapper which uses unassembled reads (108) and ISseeker which uses assembled contigs (1). These programs identify reads or contigs that contain sequence derived from the selected IS. Attempts are then made to pair these sequences to make a complete IS, i.e. with a left and right end, by looking for proximity of the position of the flanking sequences in the reference sequence (1, 108). IS positions present in the reference genome as well as novel positions are able to be identified by these programs and the output provides the IS locations in relation to the reference genome. Unless a copy of Tn2006 is present in the reference genome these programs miss Tn2006 due to the inverse orientation of the bounding copies of ISAbal and manual finishing is needed.

The ISAbal profiles of the 20 SGH isolates was determined using ISMapper for those sequenced by Illumina HiSeq and ISseeker for those sequenced by Illumina MiSeq. Profiles were then compared to elucidate relationships among these 20 isolates and to determine if they were consistent with the groupings from earlier chapters. In addition, these profiles were compared to representatives of XAR GC2 isolates from Australia to examine additional connections.

7.2 ISAbal mapping of the Singapore genomes

The genome of XH386 (GenBank accession number CP010779) was selected as the reference genome for both ISMapper and ISseeker as it was sequenced by PacBio and is ST208, and the majority of the Singapore isolates are ST208 or an SLV of this ST (Chapter 3, Section 3.5.2). ISAbal insertion locations were detected using the ISMapper program on the reads generated by Illumina HiSeq sequencing for S21, S24, S27, S29, S34, S38 and S39. The positions of ISAbal in the 13 isolates sequenced by Illumina MiSeq were located using ISseeker on the assembled contigs. Each position of ISAbal detected by the programs was manually checked to confirm the TSD, orientation and location in the reference genome XH386. The copies of ISAbal flanking *Tn2006* (which are inversely oriented and not easily identifiable by the programs) and ISAbal locations in plasmids are examined in more detail in Chapter 4 and Chapter 8, and were not included in the analysis presented in this chapter.

Overall, 74 different ISAbal insertion locations in the chromosome were detected in the 20 isolates and 31 of these positions were unique to one isolate (Table 7.1). All isolates contained at least 10 ISAbal insertions and the most found in one isolate was 19 (Table 7.1).

Isolates with the same coloured shading in Table 7.1 have ISAbal insertions in identical positions or share a high proportion of their ISAbal insertion positions. This indicates that these isolates are closely related and supports the groupings found in earlier chapters. For example, S22 and S23 contain the same AbGRI1, AbGRI2 and AbGRI3 structures as one another, as well as an AbaR4 in a specific position in the chromosome. S31, S38 and S39 share 10 ISAbal insertion positions (S31 contains a unique ISAbal insertion as well) and contain AbaR4 in another identical chromosomal location.

Table 7.1 ISAbal positions in the chromosome of 20 isolates

Isolate	Total positions	Unique positions
S21	13	4
S22	10	-
S23	10	-
S24	19	7
S25	10	-
S26	19	1
S27	13	4
S28	11	1
S29	18	-
S30	13	7
S31	10	1
S32	18	1
S33	17	- ¹
S34	16	2
S35	18	2
S36	16	-
S37	16	-
S38	10	-
S39	10	-
S40	15	1

Isolates with the same shading have identical ISAbal positions.

¹Has one side of position 305 (present in S30) and the other of position 308 (found in S32). S33 likely had both positions but has lost an ISAbal and the 722 bp between the sites.

The positions of ISAbal insertions unique to each isolate are shown in Appendix I; Fig A1.1 and Appendix I; Table A1.3 details the location of the each ISAbal insertion in the genome of XH386, the TSD and the locus tags of the surrounding open reading frames. The positions of ISAbal shared between two or more of the 20 isolates from SGH isolates are shown in Fig 7.1. Certain ISAbal positions have been colour coded to aid in the descriptions of the groups within this collection. A filled box indicates that the isolate contains a copy of ISAbal in that position, except in the case of the grey boxes of S33 (Fig 7.1). The two grey boxes indicate that an ISAbal in S33 is flanked by one side of position 305 and the other side of position 308. Hence, the most likely explanation for this arrangement is that S33 previously had an ISAbal at position 305 and 308 but they have recombined and the intervening sequence (approximately 200 bp) and one ISAbal has been lost.

All of the Singapore isolates have an ISAbal in position 1, which corresponds to the ISAbal upstream of *ampC*, as well as in positions 2, 3, 4 and 12 (Fig 7.1). XH386 and additional Australian isolates included for comparison (see below) also have ISAbal in these positions. The location of these five positions as well as position 5, which corresponds to the ISAbal adjacent to the *sul2* gene in AbGRI1 and is present in most isolates, are depicted on the schematic of the GC2 chromosome in Fig 7.2. The isolates lacking the ISAbal insertion at position 5 (S30, S31, S38 and S39) have either AbGRI1-1 or AbGRI1-6 that have lost the region containing *sul2* and this ISAbal. Therefore, the presence of ISAbal in position 5 is not as useful for epidemiology as it is located in a resistance island that is subject to variation. Unlike position 1, the impact of these other four common ISAbal insertions is unclear. Positions 2 and 3 do not interrupt any reading frames and are not appropriately positioned to increase the expression of any neighbouring genes. The ISAbal in position 4 truncates an open reading frame annotated as a transcriptional regulator and provides it with an in-frame stop codon that translates a protein 14 aa shorter than for the native stop codon. An open reading frame annotated as a translation elongation factor is interrupted by ISAbal in position 12.

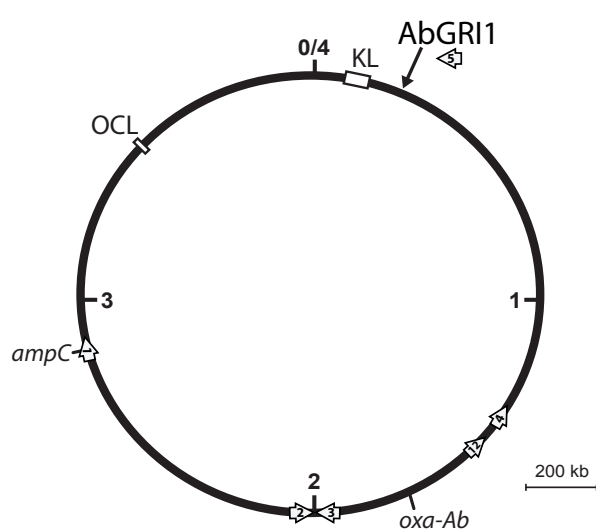


Figure 7.2 Location of ISAbal positions 1-5 and 12. Block arrows show the orientation and approximate insertion location of copies of ISAbal on a schematic of the GC2 chromosome. The location of AbGRI1 is indicated by the black arrow.

The positions of some of the additional ISAbal1 copies present in these isolates help distinguish groups in the Singapore collection. Overall, the general trend remained consistent with that in earlier chapters as four of the five *aphA6* containing isolates (shaded in pink in Fig 7.1) shared four additional ISAbal1 insertions with one another (indicated by the orange boxes in Fig 7.1). The presence of ISAbal1 in two other positions (green boxes in Fig 7.1) were common to the *armA* containing isolates (shaded in green in Fig 7.1). However, the one exception to this was S30. The ISAbal1 pattern of this isolate did not match the other *aphA6* containing isolates and only contained the ISAbal1 positions shared by both the *aphA6* and *armA* containing isolates. This suggests that S30 is from an earlier lineage that predates movement of ISAbal1 to the orange or green ISAbal1 positions.

Additional ISAbal1 positions shared between the isolates (shown in black to the right of Fig 7.1) indicate a closer relationship between them. For example, the 6 ISAbal1 positions shared between S26 and S29 (Fig 7.1) may indicate that these isolates are siblings.

The ISAbal1 profiles of a selected group of Australian isolates (kindly provided by Dr Steven Nigro) were included to compare to the ISAbal1 profiles of the 20 SGH isolates. This included WM99c, isolated in 1999 from Westmead Hospital in Sydney, in which AbGRI2-1 was characterised (188). The ISAbal1 profile of WM99c represents the profile of larger group of GC2 *A. baumannii* isolated from the east coast of Australia. The ISAbal1 profile of G2 from Alfred Hospital in Melbourne (2002) and K16 (2004) from the Royal Adelaide Hospital were also included as representatives of the GC2 populations in these hospitals. G2 and K16 have AbGRI1-5 (personal communication, Dr Steven Nigro) which is same AbGRI1 structure seen in S24, S26 and S29 indicating that they may be related to the Singapore isolates. D86, isolated from Royal North Shore Hospital in Sydney in 2010 was also included for comparison as the

armA gene was identified in the draft genome of this isolate using ResFinder. Assembly of the AbGRI3 region in D86 (described in Appendix II; Section A2.1), named AbGRI3-7, revealed that *armA* was located in an island related to and likely derived from AbGRI3-4. The right end of AbGRI3-7 had the same deletion as AbGRI3-4 but left end of the island was different.

7.3 ISAbal positions in the AbGRI3 lineage

7.3.1 ISAbal positions 58 and 62 are indicative of AbGRI3 lineage

The 15 Singapore isolates highlighted in green in Fig 7.1 contain the *armA* gene in an AbGRI3 resistance island and ISAbal in positions 58 and 62 (green boxes in Fig 7.1 and the insertion location is indicated in the chromosome in Fig 7.3) that were not in the *aphA6*-containing Singapore isolates. In addition, ISAbal was in these two positions in XH386 from China which carries AbGRI3-2 (Chapter 6, Table 6.6) (23). There were also present in the sporadic isolate from Sydney, D86, which was found to contain AbGRI3-7 (Appendix II; Section A2.1). It was concluded that ISAbal present in these two positions may be in all AbGRI3-containing isolates.

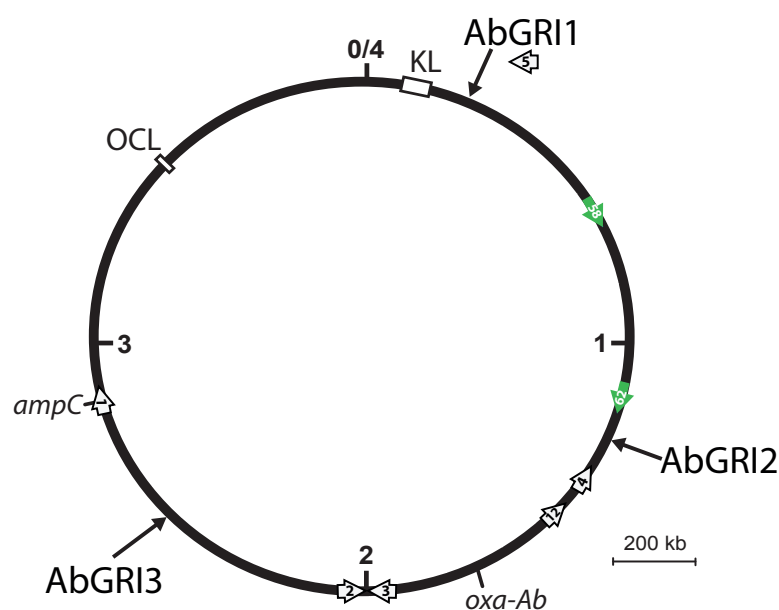


Figure 7.3 Location of ISAbal position 58 and 62. Coloured block arrows show the orientation and approximate insertion location of copies of ISAbal on a schematic of the GC2 chromosome. Black arrows indicate the location of AbGRI1, AbGRI2 and AbGRI3.

To test this hypothesis, the completed GC2 genomes that contain AbGRI3 (listed in Chapter 6; Table 6.6) were examined for ISAbal in positions 58 and 62. All genomes contained ISAbal at these two chromosomal locations.

In the course of this analysis, three additional completed genomes, CBA7 (CP020586), AB345299 (CP014291) and XH860 (CP014538), that did not carry *armA* but had ISAbal in these two positions were identified. Closer examination of the region of the chromosome where AbGRI3 is located revealed that these isolates contain a suitably located IS26, indicating that AbGRI3 had been present but its content was subsequently lost. Hence, as these two ISAbal positions are found only in isolates with AbGRI3 or a remnant of this island, they are a marker for the lineage of GC2 that contains AbGRI3. Hence, either the cell that first acquired AbGRI3 already had ISAbal in these positions or gained these ISAbal insertions soon after the acquisition of AbGRI3.

The GC2 isolate XH859 (GenBank accession number CP01459), which contains a relative of Tn6180 in a location in the chromosome different to that of AbGRI3 (Chapter 6; Section 6.4.2), was also examined to determine if it carried ISAbal in the two positions characteristic of AbGRI3 containing isolates. XH859 did not have ISAbal at these positions, providing additional evidence that the acquisition of the Tn6180 relative in this isolate was a separate event to that of AbGRI3.

7.3.2 Six additional ISAbal insertions are associated with AbGRI3-4

In addition to containing ISAbal at position 58 and 62 (green boxes in Fig 7.1), 8 of the 15 AbGRI3 containing SGH isolates share an ISAbal at position 206 (brown box in Fig 7.1). The other 7 isolates, S32-S37 and S40, share six other ISAbal insertions (yellow boxes in Fig 7.1).

All of the latter group contain AbGRI3-4 that is characterised by a specific deletion at the right end of the island. D86 also includes ISAbal at these six positions and has an AbGRI3 variant containing an identical deletion at the right end of the island (Appendix II; Section A2.1). Hence, these six insertions of ISAbal are markers for the sub-lineage of AbGRI3-4 containing isolates.

7.4 Relationship between GC2 isolates from Singapore and Australia

7.4.1 S21 contains 11 of the 12 ISAbal positions core to the east coast Australian isolates

In addition to ISAbal upstream of *ampC* (position 1) and next to *sul2* in AbGRI1 (position 5), the majority of Australian isolates have ISAbal in 12 other shared positions that are indicated by the white, orange, light and dark blue arrows in Fig 7.4. These include positions 2-4 and 7-14. The isolate S21 contains ISAbal in all of these positions, except position 11 (indicated in dark blue in Fig 7.4). In addition, S21 contains AbGRI2-1, which to date had only been seen in Australia and the ISAbal data further connects this isolate from 1996 to a precursor of the majority of isolates from the Australian eastern seaboard.

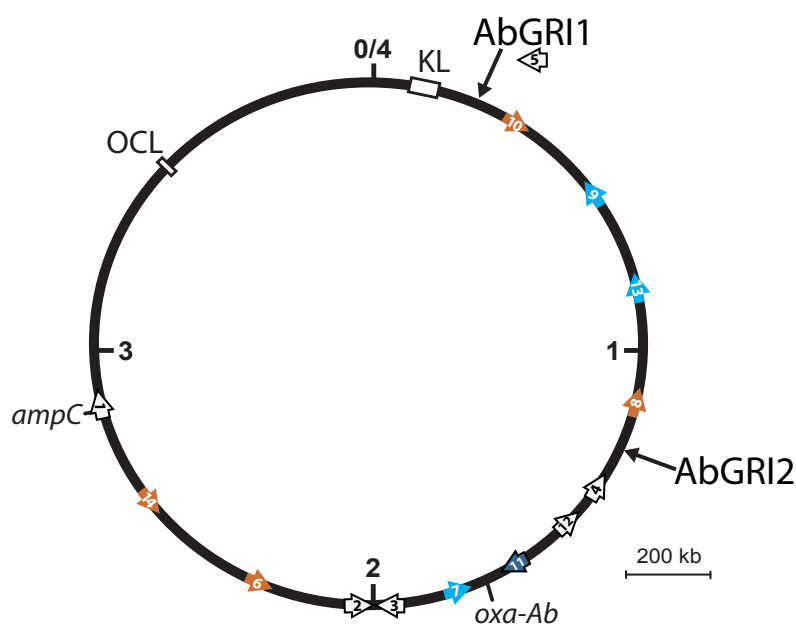


Figure 7.4 Location of ISAbal positions 1-14. Coloured block arrows show the orientation and approximate insertion location of copies of ISAbal on a schematic of the GC2 chromosome. The colour scheme is consistent with that used in Fig 7.1. Black arrows indicate the location of AbGRI1 and AbGRI2.

7.4.2 S24, S26 and S29 share positions with isolates from Melbourne and Adelaide

S24, S26 and S29 contain ISAbal in only 8 of the 12 positions core to the east coast Australian isolates (Fig 7.1). Instead, these 5 isolates have an ISAbal at positions 207, 208 and 209 (indicated by the purple boxes in Fig 7.1). The isolates from Adelaide (Hospital K) and Melbourne (Hospital G), represented by K16 and G2, respectively, also share these three ISAbal positions. The three isolates from SGH and K16 and G2 all contain AbGRI1-5, which arose from AbGRI1-3 by a recombination between *tniC* and *tniCb*, and ISAbal in these three positions may have been present in the cell in which this recombination event occurred. However, the reverse order could also be possible and without additional information there is no way to determine which was first.

7.5 Discussion

The mapping of the positions of insertions of ISAbal in the chromosome of the 20 isolates in the Singapore collection confirmed and strengthened the relationships between these isolates.

Combining the ISAbal profile with the structures of AbGRI1, AbGRI2 and, if present, AbGRI3, allowed key lineage specific insertions to be identified, such as the two ISAbal insertions characteristically present with AbGRI3. A further example of this is the six additional ISAbal insertions for the isolates with AbGRI3-4.

All of the isolates examined in this chapter, except A320, shared five (or six, if the copy of ISAbal adjacent to *sul2* in AbGRI1 was present) ISAbal positions. Hence, all of these isolates, which includes isolates originating from Singapore, China and Australia, have a common ancestor. Apart from the ISAbal upstream of *ampC*, the impact of these other shared ISAbal positions is unclear and needs to be investigated further. As A320, isolated in 1980 (190), contains the ancestral AbGRI2 structure, perhaps this isolate predates the acquisition of ISAbal at these positions.

The structure of the chromosomal resistance islands and ISAbal profiles of the 20 SGH isolates, as well as those of the selected GC2 isolates from Australia, was amalgamated with a genome wide phylogenetic tree and is depicted in Fig 7.5. The core genome phylogeny of these isolates (kindly provided by Dr Steven Nigro) was generated from the draft genome reads by SNP (single nucleotide polymorphism) calling against reference genome XH386 (GenBank accession number CP010779), followed by a maximum likelihood phylogenetic inference on the SNP alignments using RAxML to build a genome-wide phylogenetic tree.

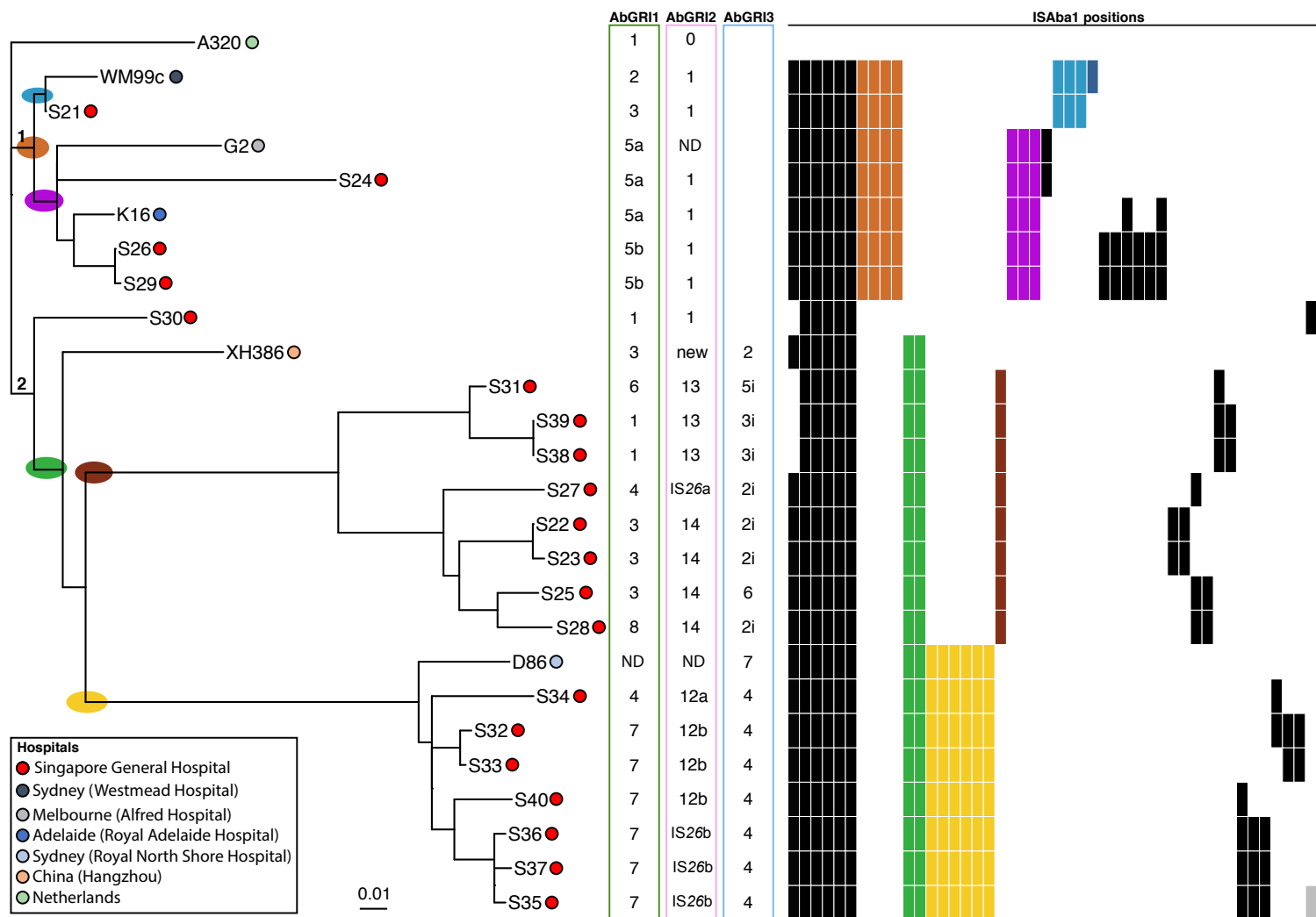


Figure 7.5 ML tree, island structure and ISAbal positions for 27 GC2 *A. baumannii*. Origin of isolates is indicated by the coloured circles adjacent to isolate names. ISAbal insertions are shown by coloured rectangles and ovals of matching colours on the tree indicate branching events characteristic of these positions. ML tree generated by Dr S. Nigro using Gubbins and RAxML.

The isolates from Singapore, Australia and XH386 from China cluster into two main clades. The first clade includes the isolates S21, S24, S26 and S29 (all have *aphA6* and lack *armA*) as well as the representative Australian isolates WM99c, G2 and K16 (Fig 7.5). These isolates all contained an ISAbal in the four positions indicated in orange in Fig 7.5 and these IS were likely present in their progenitor (branch point indicated by the orange oval in the tree in Fig 7.5). However, these isolates then break into two sub-clades, identifiable by the insertions of either three ISAbal indicated by the light blue boxes or by the three insertions indicated by the purple boxes (Fig 7.5).

The sub-clade characterised by the four ISAbal positions shown in blue in Fig 7.5, includes S21 and WM99c, the latter of which is a representative of the homogenous east coast Australian GC2 population. This provides additional evidence of the relationship between the common XAR GC2 isolates found in the eastern seaboard of Australia and S21 from Singapore, indicating that these isolates have a common source.

The second sub-clade, based on the presence of the insertions of ISAbal in purple (Fig 7.4), includes S24, S26 and S29 with G2 (2002) from Melbourne and K16 (2004) from Adelaide. These isolates also share the same AbGRI1 structure, AbGRI1-5, that contains a deletion that is unlikely to arise in precisely the same location twice. This suggests a strong connection between this three GC2 isolates from Singapore and those recovered at hospitals in Adelaide and Melbourne.

Using only the ISAbal data, S30 could be representative of the ancestral lineage of both clades. However, based on the core phylogeny, S30 lies within the second clade (Fig 7.5), but all isolates in the second clade, except S30, contain AbGRI3 and hence *armA*. This suggests that

S30 may be related to the cell that first acquired AbGRI3. This isolate also contains AbGRI2-1, which is the AbGRI2 structure that was predicted to be present in the ancestral AbGRI3 containing isolates. The AbGRI3-containing isolates separate away from S30 by the addition of the two ISAbal insertions indicated in green (Fig 7.5).

The Singapore isolates containing AbGRI3 split into two sub-clades based on the insertion of the ISAbal indicated by either the maroon box or the six copies of ISAbal at the positions indicated by the yellow boxes in Fig 7.5. The progenitor of the yellow clade likely contained AbGRI3-4, which is another island structure that arose via a unique deletion that is unlikely to occur twice. XH386, isolated in China, sits on its own branch (Fig 7.5), but if additional sequences from related isolates were included, such as those in Chapter 6; Table 6.6, they may form another sub-clade with their own characteristic positions of ISAbal.

Examination of the ISAbal profiles isolates from the Wright *et al.* (277) study also revealed clade specific ISAbal insertions. Unfortunately, as the authors used a different reference genome, TYTH-1 and as the specific coordinates of each position were not provided, it is difficult to determine if any of these isolates from America shared positions with the isolates examined in this chapter and this should be looked at in the future.

The congruence of the SNP tree, chromosomal resistance island structure and ISAbal profiles demonstrates that the recording of ISAbal positions is a potentially powerful epidemiological tool that can be used to define and track the evolution of GC2 lineages.

CHAPTER

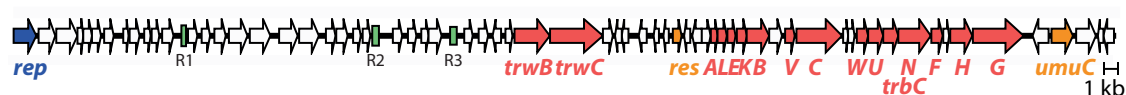
8

RepAci6 plasmids

8.1 Introduction

The plasmid originally used to represent the RepAci6 family of conjugative plasmids is pACICU2 (GenBank accession number CP000865) (21). However, the pACICU2 sequence is incomplete (90). Members of this plasmid family have been detected in *Acinetobacter* isolated from all over the world (21, 90, 129, 192, 277) and the first complete plasmid pAb-G7-2 (GenBank accession number KF669606) provides the backbone sequence (97). The RepAci6 plasmid backbone carries the *repAci6* gene, which encodes RepAci6 replication initiation protein matching Pfam0309 (Fig 8.1, blue arrow), and two separate regions with genes involved in conjugal transfer (red arrows in Fig 8.1) (97). The first region encodes the mobilisation genes *trwC* and *trwB* which most closely resemble those of IncW and nick the plasmid at the origin of transfer, while the second region includes the *tra* genes which are responsible for mating pair formation (MPF_F) of the type IV secretion system (97). The backbone is comprised of three segments separated by a repeated sequence (green boxes R1, R2 and R3 in Fig 8.1). Repeat 1 is only partial and 225 bp, while repeats 2 and 3 are two full copies of 422 or 423 bp (99). Primers have previously been designed (99, 192) to amplify these repeat regions as this facilitates assembly from short-read sequencing technologies.

A.



B.

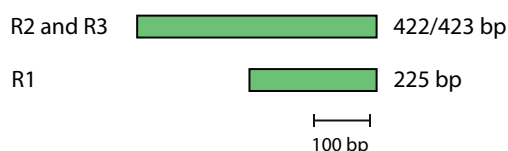


Figure 8.1 RepAci6 plasmid backbone and repeat sequences. A. Structure of the RepAci6 backbone and B. Repeat 1, 2 and 3. Arrows indicate the extent and orientation of genes and open reading frames. The *repAci6* gene is blue, genes involved in conjugal transfer are red and the *res* and *umuC* gene are orange. Green boxes indicate the three repeat sequences. Figure made to scale using pAb-G7-2 (GenBank accession number KF669606).

Although RepAci6 plasmids have a conserved structure, two related backbone types, type 1 represented by pAb-G7-2 (97) and type 2 exemplified by pA85-3 (99) have been found. These two types are distinguished based on SNPs throughout the backbone as well as short insertions (Fig 8.2). Differences in the sequence surrounding the third repeat (Fig 8.2) can produce amplicons of different sizes that easily distinguish between the type 1 and 2 backbones (99).

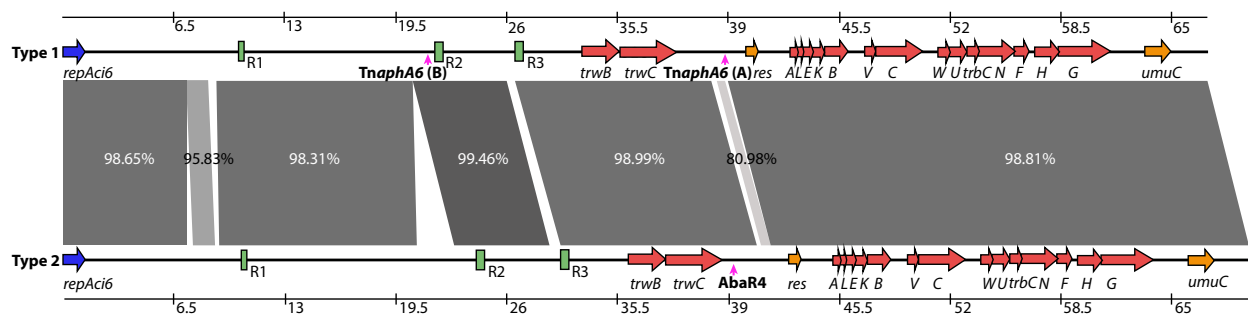


Figure 8.2 Comparison of the two RepAci6 plasmid backbone types. Arrows indicate the extent and orientation of genes. The *repAci6* gene is blue, genes involved in conjugal transfer are red and the *res* and *umuC* gene are orange. Green boxes indicate the three repeat sequences. Regions of homology are shown by grey boxes and the numbers display DNA identity. Vertical pink arrows show the insertion location *TnaphA6* and *AbaR4*. Figure is made to scale using pAb-G7-2 (GenBank accession number KF669606) for type 1 and from pA85-3 (GenBank accession number KJ493819) for type 2.

RepAci6 plasmids are important as some have been associated with the carriage of antibiotic resistance genes, including the *aphA6* gene (confers resistance to amikacin, kanamycin and neomycin) and the *oxa23* gene (confers resistance to carbapenem antibiotics).

The *aphA6* gene is located in *TnaphA6* (197), and so far, this transposon has been found in two different positions in the Type 1 backbone in RepAci6 plasmids. The first position (*TnaphA6* (A) in top line of Fig 8.2), was detected in a plasmid from an Australian isolate, pAb-G7-2 (99) and later in pABUH1 from the US (196, 277). pACICU2 was originally reported to be cryptic (129). However, it was later shown to carry *Tnaph6* in the same position as pAb-G7-2 (90) and the two plasmids are likely identical. The second position of *TnaphA6* (*TnaphA6* (B) in top

line of Fig 8.2) was found in pD72-2 (192) and pD46-3 (196), which are both type 1 plasmids from isolates recovered at different hospitals within in Australia.

The *oxa23* gene has been detected in various transposon contexts in RepAci6 plasmids. These include *oxa23* within Tn2008 (277), Tn2009 (125) and Tn2006 (196) or AbaR4 (Tn2006 inserted in Tn6022) (99). In some RepAci6 plasmids both the *oxa23* and the *aphA6* gene are both present (196, 277).

The aim of the work presented in this chapter was to examine if RepAci6 plasmids are found in the 20 GC2 isolates from SGH and, if they are present, to determine what backbone they have and whether any of them harbour TnaphA6 or *oxa23* containing transposons.

8.2 Screening for RepAci6 plasmids

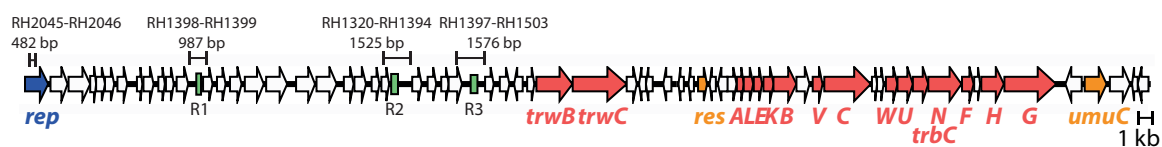
The primers RH2045 and RH2046 were designed to amplify a 482 bp region of the *repAci6* gene and were used to screen the 20 isolates (Fig 8.3A and B). S20-S30, S32, S34, S36 and S40 produced an amplicon around 500 bp indicating that they contain a *repAci6* plasmid. S31, S33, S35, S37-S39 did not generate any amplicon indicating that these isolates do not have a *repAci6* plasmid and this was later confirmed when the draft genomes of these isolates were examined. Repeat 1, 2 and 3 in the isolates that were positive for *repAci6* were then screened for (shown in Fig 8.3A and summarised in Table 8.1).

The primers spanning R3, RH1397 and RH1503, will produce a 1,575 bp amplicon if the plasmid is type 1, while if it is type 2 a 1,774 bp amplicon will be generated (Fig 8.3A). S24, S26, S29 and S30 produced an amplicon close to 1.5 kb in size indicating that they contained type 1 plasmids (Fig 8.3B). However, S21, S22, S25, S27, S28, S32, S34 and S40 produced an

amplicon approximately 1.8 kb in size matching the type 2 control (Fig 8.3B). For repeat 2, the primers RH1320 and RH1394 were used to amplify a region of 1,534 bp in type 2 and 1,525 bp in type 1 plasmids (Fig 8.3A). All isolates screened, except S30, produced an amplicon around 1.5 kb in size. The RepAci6 plasmid in S30 may contain additional sequence in this region or the primer binding sites may be absent from the backbone of S30.

A.

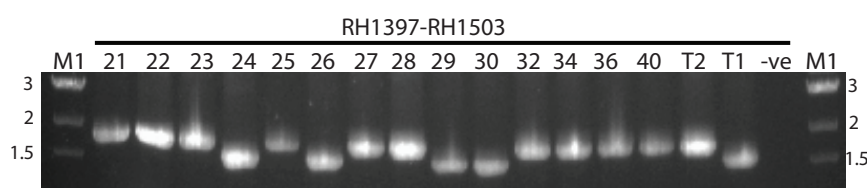
Type 1



Type 2



B.



C.

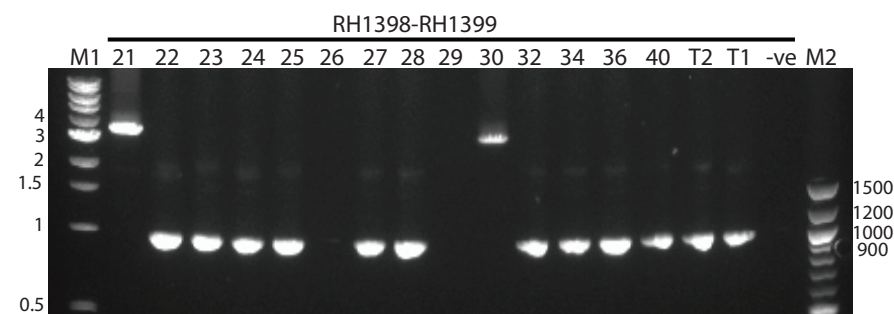


Figure 8.3 Screening PCRs over repeat 1, 2 and 3. A. Location of the primers on the plasmid backbone. Arrows indicate the extent and orientations of genes and open reading frames. The *repAci6* gene is blue, genes involved in conjugal transfer are red and the *res* and *umuC* gene are orange. Green boxes indicate the three repeat sequences. Primers used to span the repeat sequences as well as detect the *repAci6* gene are shown by the bars above along with the PCR amplicon size. B. Screening repeat 3 using primers RH1397 and RH1503. The amplicon is 1,576 bp for type 1 and 1,774 bp for type 2. C. RH1398 and RH1399 were paired to detect an amplicon of 987 or 991 bp depending if the type 1 or 2 backbone is present, respectively. The sizes of molecular markers on the left and right in B. are in kb, as are the sizes on the left of C. Those on the right in C. are in bp. Isolate numbers are shown above their lanes with T1 and T2 representing type 1 and type 2, respectively.

When the primers RH1398 and RH1399 were used to amplify repeat 1, which should yield a predicted product size of 978 bp or 991 bp depending if the plasmid is type 1 or type 2 respectively (Fig 8.3A), the results were variable (Fig 8.3C). The isolates S22-S25, S27, S28, S34, S36 and S40 amplified a product around 1 kb, matching the expected amplicon for either backbone type (Fig 8.3B). However, S21 and S30 generated a product approximately 3-4 kb, while S26 and S29 failed to produce an amplicon (Fig 8.3B). It appears that S21 and S30 have an additional 2-3 kb of sequence in the region of repeat 1 (Table 8.1). S26 and S29 may also contain additional sequence in this region that may be too large to be amplified under the conditions used here, or there may be variability in the region of the plasmid backbone where the primers should bind in the plasmids of these two isolates therefore preventing proper annealing.

Table 8.1 Properties of the isolates containing *repAci6* plasmids

Isolate	ST (ox)	AbGRI3	Repeat 1	Repeat 2	Repeat 3 ¹	TnaphA6 pAb-G7-2 position
S24	208	-	+	+	+(1)	+
S26	1165	-	-	+	+(1)	+
S29	1165	-	-	+	+(1)	+
S30	369	-	+ ²	-	+(1)	+
S21	208	-	+ ²	+	+(2)	-
S22	218	+	+	+	+(2)	-
S23	218	+	+	+	+(2)	-
S25	218	+	+	+	+(2)	-
S27	218	+	+	+	+(2)	-
S28	1128	+	+	+	+(2)	-
S32	195	+	+	+	+(2)	-
S34	1166	+	+	+	+(2)	-
S36	195	+	+	+	+(2)	-
S40	195	+	+	+	+(2)	-

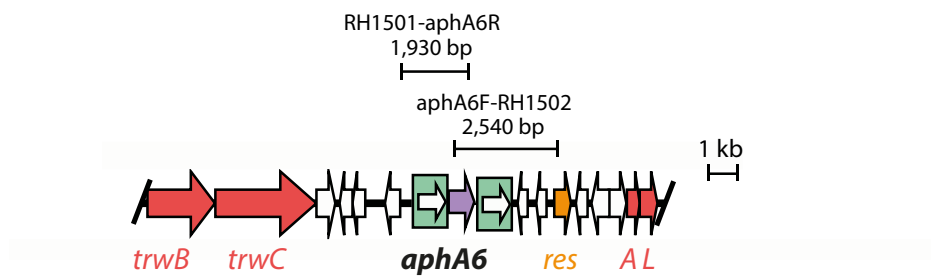
¹The backbone type is in the parenthesis.

²Amplicon is 2-3 kb larger than expected.

PCR screening results separated the RepAci6 plasmid of the 14 isolates into backbone types based on the size of amplicon over repeat 3 (Table 8.1). However, amplification of repeat 1 and 2 revealed additional differences (summarised in Table 8.1).

The type 1 plasmid pAb-G7-2 contains *TnaphA6* (location (A) indicated by the vertical arrow in the top line of Fig 8.2). From previous screening S21, S24, S26, S29 and S30 also contain the *aphA6* gene located in *TnaphA6* (Chapter 3; Section 3.3.4). These five isolates (four containing a type 1 plasmid and one containing a type 2 plasmid) were screened using published primers (RH1501 and *aphA6R*, and *aphA6F* and RH1502, which should amplify a 1,930 bp region and a 2,540 bp region (97), respectively), to determine if *TnaphA6* is located in the same place in the backbone of the *repAci6* plasmids (Fig 8.4). The four isolates containing a type 1 plasmid produced amplicons of the predicted sizes for both junctions of this transposon (Fig 8.4B) and hence harboured *TnaphA6* in the same position as pAb-G7-2 (Table 8.1). *TnaphA6* in the S21 may be located in a different position of the backbone of its type 2 RepAci6 plasmid or elsewhere in the genome of this isolate.

A.



B.

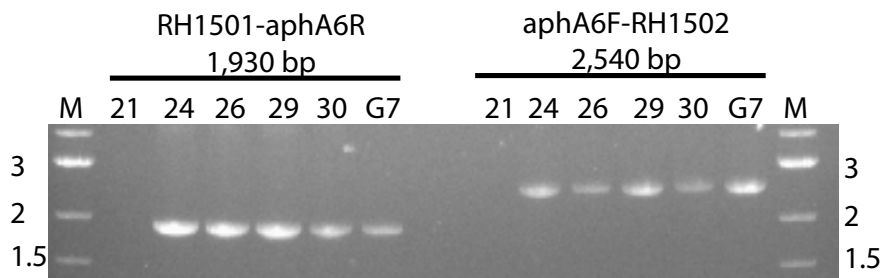


Figure 8.4 Screening pAb-G7-2 location of *TnaphA6*. A. Location of *TnaphA6* in pAb-G7-2. Arrows indicate extent and orientation of genes and open reading frames. Genes involved in conjugal transfer are red, the *res* gene is orange and the *aphA6* gene is in purple. ISAbal25 is represented by a green box and the internal arrow indicates the transposase orientation. Primers used to detect *TnaphA6* are shown above by the bars along with the PCR amplicon size. B. Amplicons produced using RH1501 and aphA6R and aphA6F and RH1502 if *TnaphA6* is in the same location as in G7. The sizes of the molecular marker are shown to the left and right in kb and isolate numbers are shown above their lanes.

8.3 Type 1 plasmids

Based on screening over repeat 3, the isolates S24, S26, S29 and S30 contained a type 1 RepAci6 plasmid carrying *TnaphA6*. The number of contigs matching *repAci6* plasmids in each isolate depended on what technology was used to sequence the isolates, either HiSeq or MiSeq sequencing platforms, and if additional insertions were present. If the draft genome of the isolate was sequenced by HiSeq, the RepAci6 plasmid was in at least three fragments (F1, F2 and F3 in Fig 8.5). Due to the longer read length of Illumina MiSeq, there were generally fewer contigs as the assembly did not always break at the repeat sequences or over some insertions.

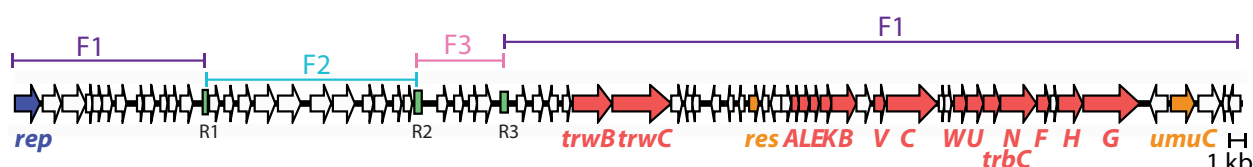


Figure 8.5 Three fragments of the RepAci6 backbone produced by assembly of Illumina HiSeq reads. Arrows indicate the extent and orientation of genes and open reading frames. The *repAci6* gene is blue, genes involved in conjugal transfer are red and the *res* and *umuC* genes are orange. Green boxes indicate the three repeat sequences. The lines above of different colours indicate fragment 1 (purple), fragment 2 (blue) and fragment 3 (pink).

8.3.1 Assembly

The type 1 backbone was used as a query against the draft genomes of S24, S26, S29 and S30.

The contigs were tiled to the type 1 backbone and the features at the edges of each contig were examined (Table 8.2). The draft assembly of the RepAci6 plasmids in S24 and S29, which were produced by Illumina HiSeq, were each in five contigs. Due to the longer read length of Illumina MiSeq the draft RepAci6 plasmid of S26 was in fewer contigs and in S30, it was only in one. The plasmid in S30 had a number of differences to the type 1 backbone and this plasmid is described in more detail in Section 8.6.

Table 8.2 Contigs matching RepAci6 plasmid backbone in S24, S26, S29 and S30

Isolate	Contig number	Contig length (bp)	Description	Contig edges ³	
				Left	Right
S24 ¹	14	9,864	F2	R1	IRl of Tn6022
	15	38,664	F1, <i>tra</i> region, <i>repAci6</i>	ISAbal25	R1
	36	4,476	F3	R2	R3
	43	12,056	F1, <i>trw</i> genes	R3	ISAbal25
	65	1,752	F2	IRr of Tn6022	R2
S29 ¹	55	978	<i>aphA6</i>	ISAbal25	ISAbal25
	12	4,464	F3	R2	R3
	23	11,441	F2	R1	R2
	25	11,970	F1, <i>trw</i> genes	R3	ISAbal25
	29	45,852	F1 (extra 10 kb), <i>tra</i> region, <i>repAci6</i>	IRr of Tn6022	R1
S26 ²	39	2,235	F1, <i>res</i>	ISAbal25	IRl of Tn6022
	63	972	<i>aphA6</i>	ISAbal25	ISAbal25
	20	62,163	F1 (<i>tra</i> region, <i>repAci6</i>), F2, F3	IRr of Tn6022	repeat 3
	38	17,312	F1 (<i>trw</i> genes, <i>TnaphA6</i>)	Repeat 3	IRl of Tn6022
S30 ²	87	424	R2 and R3 sequence		
	18	73,043	F1 (<i>TnaphA6</i>), F2 and F3	ISAbal1 (IRl)	ISAbal1 (IRr)

¹Sequenced by Illumina HiSeq.

²Sequenced by Illumina MiSeq.

³R1, R2 and R3 stand for repeat 1, 2 and 3, respectively.

The five RepAci6 contigs of S24 (Fig 8.6A) were almost completely assembled by sequencing the amplicons of the screening PCR over the three repeats and *TnaphA6*. However, there was an additional break just upstream of the second repeat (Fig 8.6A). Contig 14 and 65 had the IR sequence of Tn6022 at one end, and adjacent to this sequence was a 4 bp target site duplication (5'-GTCG-3'), indicating that AbaR4 was present in this location.

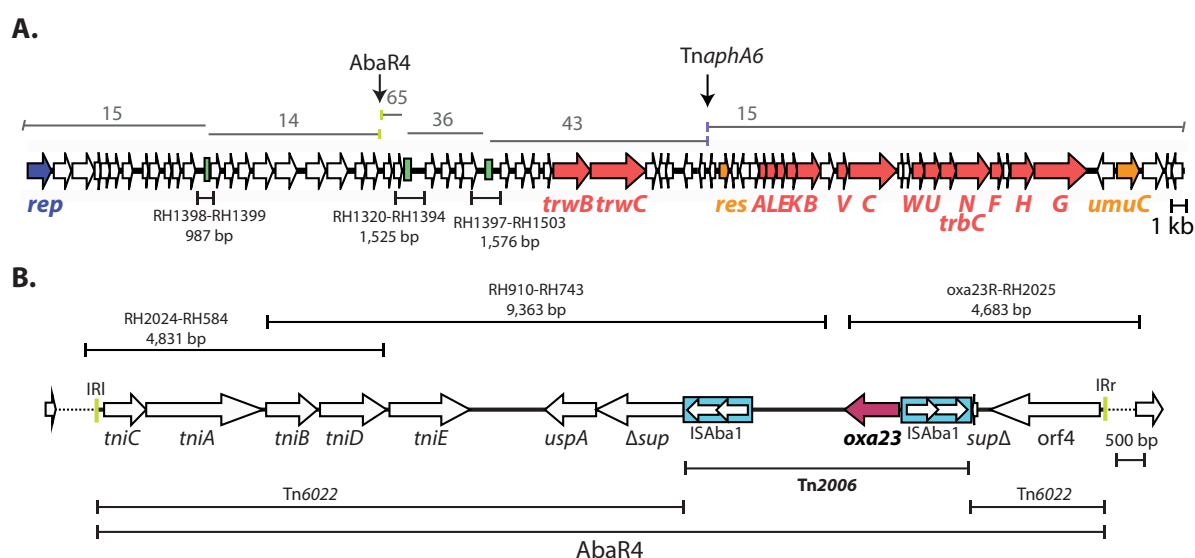


Figure 8.6 Assembly of pS24-1. A. pS24-1 contigs tiled against the type 1 backbone. Arrows indicate the extent and orientation of genes and open reading frames. The *repAci6* gene is blue, genes involved in conjugal transfer are red and the *res* and *umuC* genes are orange. Green boxes indicate the three repeat sequences. Grey lines above indicate the extent of different pS24-1 contigs. The black bars underneath show the primers used for assembly along with PCR amplicon size. Vertical arrows indicate the insertion location of AbaR4 and *TnaphA6*. B. Assembly of AbaR4 in pS24-1. The *oxa23* gene is magenta and the IR sequences of Tn6022 are shown in olive green. The blue box indicates ISAbal and the internal arrows indicate the transposase orientation. The primers used to span AbaR4 are shown above by the black bars, along with PCR amplicon size. The lines below indicate the transposons.

To map this AbaR4 location a new primer, RH2024, was designed at the end of contig 14 and combined with RH584 to produce an amplicon 5,034 bp in size (Fig 8.6B). Oxa23R and another new primer RH2025 designed at the end of contig 65 should produce a 4,683 bp amplicon if AbaR4 is present (Fig 8.6B). In both cases, products were of the expected size and were sequenced. The internal AbaR4 PCR, RH910 and RH743, which should produce an amplicon 9,363 bp in size was performed and an amplicon of the expected size was generated.

Therefore, the RepAci6 plasmid in S24, named pS24-1, is 86,916 bp in size and contains *TnaphA6* as well as AbaR4 (Fig 8.7, top line).

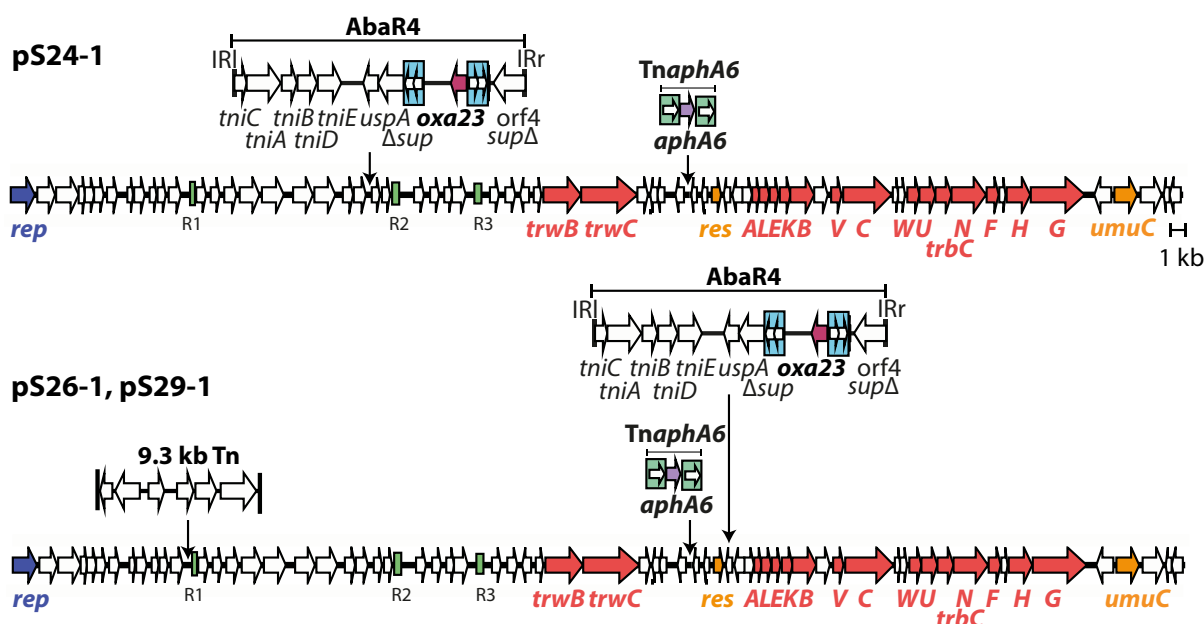


Figure 8.7 Linear comparison of pS24-1, pS26-1 and pS29-1. Arrows indicate the extent and orientation of genes and open reading frames. The *repAci6* gene is blue, genes involved in conjugal transfer are red, *res* and *umuC* genes are orange and the *oxa23* gene is magenta. Green boxes indicate the three repeat sequences. Vertical arrows indicate the insertion of a transposon and the identity of the transposon is shown above. ISAbal25 and ISAbal1 are represented by the green and light blue box respectively, and the internal arrows indicate the transposase orientation.

The RepAci6 plasmids in S26 and S29, called pS26-1 and pS29-1, were assembled using a similar process and details of each assembly can be found in Appendix III; A2.2.1. The plasmids pS26-1 (96,203 bp) and pS29-1 (96,201 bp) also contained AbaR4 in the same position and surrounded by a 5 bp target site duplication (5'-AATTG-3') but this was a different location to that in pS24-1 (Fig 8.7). AbaR4, in both positions in the type 1 plasmids, had interrupted an open reading frame, however the products of both have no known functions (Table 8.3). An additional 9.3 kb of sequence was located just upstream of repeat 1 in pS26-1 and pS29-1 (Fig 8.7), accounting for the lack of amplification of this repeat in these isolates (Table 8.1).

Table 8.3 AbaR4 location in the type 1 plasmids

Plasmid	Duplication (5'-3')	Location of AbaR4	Description of interrupted protein	Novel position?
pS24-1	GTCG	Interrupts a gene encoding a 209 aa protein.	No conserved domains – ORF	Y
pS26-1/ pS29-1	AATTG	Interrupts a gene encoding a 144 aa protein.	No conserved domains - ORF	Y

8.3.2 9.3 kb transposon in pS26-1 and pS29-1

The boundaries of the additional sequence prior to repeat 1 in pS26-1 and pS29-1 was defined by comparison to the type 1 backbone. The insert was found to be 9,279 bp in size and surrounded by a 7 bp duplication (5'-ATTCTAG-3'). An examination of the end of the insert revealed a 75 bp inverted repeat with 64 bp identical (bases highlighted in green in Fig 8.8A). Six open reading frames were identified in this 9.3 kb region (Fig 8.8B) and pBLAST was used to predict their putative function where possible (Table 8.4).

A.

IRI TGCAAAATACTGGATAAAACAACCACCCATTCTGCTCCTACCATACCAGTCATTCTTGTTTCGTGCTATACCCTATAATTAAAAATAAAAGGGAGA
 IRr TGCATATTCCTGCATAAACATACCACCCATTCTTATTCCTACTATACCAGCCATTCTGTTTCGTGCTATACCCTTGAGTTAAACTGATAAAGGCT

64/75 bp

B.

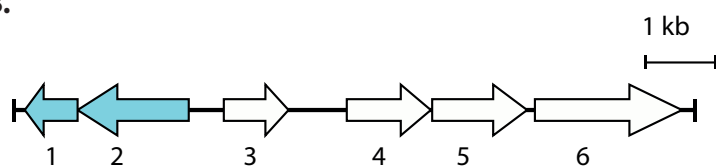


Figure 8.8 Features of the 9.3 kb novel transposon. A. Identification of the 75 bp inverted repeat with 64 bp identical. The bases highlighted in green are the IR. B. Schematic of the 9.3 kb transposon. Arrows indicate the extent and orientation of genes and open reading frames, those in blue are likely involved in transposition. Vertical lines indicate the IR sequences.

Table 8.4 Open reading frames and predicted products encoded by the 9.3 kb transposon

orf	AA	Superfamily	Multi-domains	Description
1	246	IstB_IS21_ATP	DnaA	transposase
2	522	COG4584	COG4584	
3	315	Arc superfamily		
4	382	Cupin_like		
5	422	Type_I_hlyD	EmrA	Efflux/secretion proteins
6	709	SunT	NHLM_micro_ABC1	Bacteriocin/antibiotic exporter

Orf 1 hit the IstB protein of IS21, which is the second protein encoded by IS21 family IS, while Orf 2 matched a transposase. IS21 family IS encode two proteins necessary for transposition, IstA, the transposase and IstB, a helper protein, and both are necessary for efficient and accurate transposition (18). The two reading frames encoding IstB and IstA are just overlapped, as are orf 1 and 2 in the 9.3 kb insert. The first two proteins from the insert were queried against the ISFinder database to examine if they showed any homology to those of the IS21 family. Orf 2 matched with 34% aa identity to IstA of ISPpu23 (Fig 8.9A) while Orf 1 matched with 37% aa identity to IstB of ISPpu23 (Fig 8.9B). ISPpu23 is a member of the IS21 family and these matches extend over most of the IstA and IstB proteins (Fig 8.9A and B). IS21 family members produce target site duplications ranging from 2-9 bp in size.

As Orf 1 and 2 show similarity to the transposase and helper protein of IS21 family IS, it is likely that they are responsible for the transposition functions of this 9.3 kb insert. However, as they are moving a much larger region including additional genes, this 9.3 kb insert is a novel transposon. The proteins encoded by orf 3 and 4 do not have any strong matches. The products of orf 5 and 6 contain domains matching efflux/secretion proteins and bacteriocin/antibiotic exporter, respectively, and could potentially be involved in the virulence of these isolates (Table 8.4).

A.

```

orf_1      MLKQSVKDKLERLKLKGFRLTAEASLDQMLAQGKSFEDFLDFCLDAELNDRNRRIERLF
ISPpu23_IstB MLPHPTLDKLTQLRLTGMKLALAEQLKTPDIDSLSFVERLGLLDRELTERDDKRLSSRL
          ** : . ***: *:*.***: .*. :. ** : *: : * *.** :*: . :

orf_1      SIAKFRYPtarLEHIDYAPERKIKKADISRYANCEWISDCRGLIISGATGVKSWLACAL
ISPpu23_IstB RQARLK-HNACLEIDYRSPRGLDKALILQLSSGQWLRDGLNLIINGPTGVGKTWLACAL
          **: . * *.*** * :.*** : :. :*: * .***. * *****:*****

orf_1      GAEACHLSYSAMYFNAISLFDEISSAITLGEIKDFKKKVCSTKLLIIDDGGLGGFNQSLA
ISPpu23_IstB AHQACREGYSVRYLRRLLEELGLAHGDRFAKLMSGYAKTDLLILDWGLAPFNAEQR
          . :*: .*. *. :. :*: . * *. :. . .*.***:***. ** .

orf_1      PSFLEIIDKQSQSGGLLITSQVPKRLWFEQFQDPTLADAVMDRILHRSYELEISGRSYRE
ISPpu23_IstB RDMLELLDDRYGQRSTIVTSQMPVDNWHELIGDPTLADAILDRLVHNAYRINLKGESMRK
          .:***:*. :. :****: * *: *****:***:*.***:*. * *:

orf_1      KIYAQ-----
ISPpu23_IstB QTKKLTTPGTSD
          :

```

B.

```

orf_2      MRIETDKIRQIVRL-LNLGHSQRSTAIQARANRSSVKTFIDKLT----LFPIDNSELNLF
ISPpu23_IstA ----MRKIREVRLKFEVGLSARQIAVSQIGRVTVGDYLNRFASGLAWPCSLSDS---
          ****:*** :*: * *. *. :. :. :*: :*: :*: :*: :*: :*:

orf_2      TNLELLDYFEISRTPTY-PTRKIYPDFEYINQELKKRDMTELLWQEYIAQYKEGLSYSR
ISPpu23_IstA ---ELEQ-QLFPPPPAVPSEQRPLPDWSVHAELERRPGVTLLWQEYRLSQPKGFQYSW
          ** : : * : : : ** :.*** :*: :.*** ***** . :*:**

orf_2      FCHVFNFHFNKKHASMRQTFKSGEVLLVDFCGRTMEITSPVDGSKSYAQVFVGLGASAY
ISPpu23_IstA FCEHYRAWQGLDVMRQEHVRGEKLFVDYAGQTVPVIDRHSGEIRQAQVFVAVLGASSY
          **. :. :*: . * *. * :. * *.***:***: :. .*. *****.***:*.

orf_2      TFAYAVPSQRVEHWLECFIKTFDHIGVGPATVITDNLKSAVLKNNKTGLIFQKDFEDFAA
ISPpu23_IstA TFAEATWSQQLPDWLGSHVRCFAFLGGVPEIVVPDNLRSVAVSKSHRYEPDINPSYRDIAE
          *** *. ***: .** .. :. :.*** * :.***:*** *. :. :.***:

orf_2      YYDFAIIPTRPRHPKDKGLAEVSVQIVQRAVLASLRNQKFFSIEELNQAIQEKMDVINRK
ISPpu23_IstA HYGVAVVPARARKPRDKAKAEVGVQVVERWILAALNRNQFFSLDELNSAIALLLERLNR
          :*.***:*** *:***. ***.***: * :***:***:***:***.*** : : **:

orf_2      TTRRFTISRYEQFLVLDKDISPLPLYPYELCSWKRHVRVSEFYRVEYEGNQYSVPYTYI
ISPpu23_IstA PFRKLPGSRHSAFEALDRPALSPLEQPYVYAEWKKA-RVHIDYHVEVDGHYYSVPYQLV
          *: **:. * .** :**** ** ..*: ** *:** *: ***** :

orf_2      HLHVLDLKVTSIAIHIFYEREKIAEHAISHGICLDICLDEHMSPAHRAQKGLSKDEIIHWA
ISPpu23_IstA KKQLEVRLTARTVECFHANQRVASHLRSPHKGRHSTQTEHMPKSHREHAEWTPQRLIRWA
          : :****: ** . * :.***: * . *** **: : :.***:

orf_2      NRAGPNTSHYVCHILNQKRDRLARNIKSLNKLKRWVDNQKSHCLEEACDFASQRQIYALD
ISPpu23_IstA EQTGPHTAGVISHILERRIHPTQGYRACLGLRLG-KTHGEVRLELACRRALSLGACSYK
          :*:***: :.***: :. :. : : :. :. * * * * . :.

orf_2      RLQRIANSAYQIQQQPL-SDALSHIPTSHHNIRGADYYLQNSGVAHA
ISPpu23_IstA SLESILRQG---LENLPLAQQLPLLPDDHANLRGPGYYH-----
          *: * :. :*: ** . * * . * *.** .**

```

Figure 8.9 Homology of Orf1 and 2 to the transposases of the IS27 family. A. Similarity of Orf 1 to IstB of ISPpu23 and B. Similarity of Orf 2 to IstA of ISPpu23.

A search of the nucleotide database of GenBank revealed that four other *A. baumannii* isolates from South Korea and China contain a copy of this 9.3 kb Tn (Table 8.5). This transposon was found in the same location in pUSA15-1, another RepAci6 plasmid that was isolated in Seoul in South Korea in 2013. The RepAci6 plasmids present in KAB04 and KAB08 also contain a

version of this transposon. However, in these two plasmids it is located 102 bp downstream of repeat 3, surrounded by an 8 bp TSD (Table 8.5). In addition, orf 4 of the 9.3 kb transposon in these plasmids is interrupted by ISAbal0. KAB04 and KAB08 were isolated from Busan and Gyeonoggi respectively, in South Korea in 2015. A copy of the 9.3 kb transposon was found in a third location in the genome of XH857 (GenBank accession number CP014540) which was isolated in Guangdong in China in 2010. In this isolate the transposon was surrounded by a 7 bp direct repeat of 5' AAAC TAG 3' (Table 8.5).

Table 8.5 Sequences containing the 9.3 kb transposon

Sequence	Tn size (kb)	Location of Tn	TSD (5'-3')	Geographical origin	GenBank accession number
pS26-1, pS29-1	9.3	43 bp prior to R1	ATTCTAG	Singapore	
pUSA15-1	9.3	43 bp prior to R1	ATTCTAG	Seoul, South Korea	CP020594
pKAB04	10.3 ¹	RepAci6 plasmid, 102 bp after R3	GGATAGGC	Busan, South Korea	CP017649
pKAB08	10.3 ¹	RepAci6 plasmid, 102 bp after R3	GGATAGGC	Gyeonoggi, South Korea	CP017657
XH857	9.3	Chromosome	AAAC TAG	Guangdong, China	CP014540

¹An ISAbal0 interrupts orf 4.

8.3.3 Plasmid comparison and conjugation

The backbone of pAb-G7-2 is 67,025 bp and when the inserts of pS24-1, pS26-1 and pS29-1 were removed the backbones of these three plasmids were the same size. These backbones were then compared to that of pAb-G7-2, but only minor differences were detected (Table 8.6). All three plasmids had a SNP in *traH* causing an amino acid substitution in the translated protein, while pS24-1 and pS29-1 had additional SNPs.

Table 8.6 Type 1 plasmid backbone comparisons

Plasmid	Backbone size (bp)	Identity to backbone of pAb-G7-2	Location of SNPs relative to pAb-G7-2 (base change; aa change)
pS24-1	67,025	67,022/67,025	- <i>trwC</i> (T to A; N→K) - <i>traB</i> (T to C; S→L) - <i>traH</i> (A to T; N→Y)
pS26-1	67,025	67,024/67,025	- <i>traH</i> (A to T; N→Y)
pS29-1	67,025	67,017/67,025	- 5 SNPs and 2 gaps in a 909 bp region over repeat 2 - <i>traH</i> (A to T; N→Y)

As pS24-1, pS26-1 and pS29-1 confer resistance to ticarcillin (*oxa23* in AbaR4) and kanamycin (*aphA6* in Tn*aphA6*) their conjugation proficiency was able to be examined. Each isolate was mated with ATCC 19606^{Rif} (rifampicin resistant) and transconjugants were selected for using plates supplemented with ticarcillin and rifampicin.

Ticarcillin and rifampicin resistant transconjugants were recovered from the S24 mating (Table 8.7). These colonies were additionally found resistant to kanamycin, indicating that the *aphA6* gene is also present. PCR screening confirmed the presence of the *oxa23* (*oxa23R-oxa23F*) and the *aphA6* (*aphA6F-aphA6R*) genes. Hence, pS24-1 is transfer proficient indicating that the SNP in *traH* is not detrimental for conjugation (Table 8.6). A conjugation frequency of pS24-1 of 2.09×10^{-5} transconjugants/donor was determined by selecting the ticarcillin and nalidixic acid transconjugants of three independent matings of ATCC 19606^{Rif} (pS24-1) with AB307_0294^{Nx} (nalidixic acid resistant) (Table 8.7).

Table 8.7 Type 1 plasmid conjugations

Plasmid	Resistance	Mating of each strain with ATCC 19606 ^{Rif}		Mating of ATCC 19606 ^{Rif} (RepAci6 plasmid) with AB307_0294	
		Transconjugants (TIC ^{Rif})	Conjugation frequency ¹	Conjugation frequency ¹	Mean Conjugation frequency ²
pS24-1	TIC, Km	+	4.55x10 ⁻⁶	1.18x10 ⁻⁵ 3.86x10 ⁻⁵ 6.00x10 ⁻⁶	2.09x10 ⁻⁵
pS26-1	TIC, Km	-	<1.84x10 ⁻¹⁰		
pS29-1	TIC, Km	-	<1.25x10 ⁻¹⁰		

TIC, ticarcillin; Km, kanamycin.

¹Conjugation frequency expressed as transconjugants per donor. Each frequency is from an independent determination.

²Mean frequency expressed as transconjugants per donor calculated from three independent determinations.

However, when either S26 or S29 was used as a donor and mated with ATCC 19606^{Rif}, following the removal of rifampicin-resistant mutants of the donor, no transconjugants were recovered. Hence, both pS26-1 and pS29-1 were unable to transfer at a detectable frequency (Table 8.7). The only SNP causing an amino acid substitution in pS26-1 and pS29-1 is in *traH* (Table 8.6) but the presence of this SNP does not abolish conjugation in pS24-1. An obvious difference between S24-1 and pS26-1/pS29-1 is the location of AbaR4. The reading frame AbaR4 interrupts in S26-1/pS29-1 plasmids is located downstream of the *res* gene and within the transfer region (Table 8.3). Although, it encodes a protein of unknown function this protein may be required for the self-transmission of the plasmid.

8.4 Type 2 plasmids

8.4.1 Assembly

The contigs matching RepAci6 plasmid in S21, S22, S23, S25, S27, S28, S32, S34, S36 and S40 were retrieved, tiled to the type 2 backbone. Edges of the contigs were analysed and assembled using a similar process to pS24-1. The specifics of the assembly of each plasmid can be found in Appendix II, A2.2. Linearized versions of 9 type 2 plasmids (pS22-2, pS23-2,

Table 8.8 Type 2 RepAci6 plasmids

Plasmid	Plasmid size (bp)	Inserts ¹	
		AbaR4	Other
pS32-2	70,833		ISAba19
pS36-2	70,833		ISAba19
pS22-2	86,340	AbaR4 (1)	
pS23-2	86,340	AbaR4 (1)	
pS40-2	86,340	AbaR4 (1)	
pS25-2	87,529	AbaR4 (1)	ISAba1
pS27-2	87,530	AbaR4 (1)	ISAba1
pS28-2	87,529	AbaR4 (1)	ISAba1
pS34-2	86,336	AbaR4 (2)	

¹The number inside the bracket indicates the position of AbaR4 insertion.

The two locations of AbaR4 in type 2 plasmids both interrupt genes encoding proteins of unknown function (Table 8.9).

Table 8.9 Type 2 AbaR4 locations

Plasmid	Duplication (5'-3')	Location of AbaR4	Description of interrupted protein	Novel position?
pS27-2, pS22-2, pS23-2, pS25-2, pS28-2, pS40-2	CCATT	Interrupts a gene encoding 108 aa protein	No conserved domain - ORF	N (Same as pA85-3)
pS34-2	TCCAC	Interrupts a gene encoding a 372 aa protein.	No conserved domains - ORF	Y

8.4.2 Backbone comparisons and conjugation

The backbone of pA85-3, once AbaR4 and the 5 bp direct repeat is removed, is 69,524 bp in size. By removing the inserts present in each of the nine type 2 plasmids, along with a copy of the direct repeat, the backbone of each plasmid was revealed (Table 8.10). These backbones were then queried against the pA85-3 backbone to detect any differences (Table 8.10). All of the plasmids were above 99.9% nucleotide identity to pA85-3.

Table 8.10 Type 2 backbone identity

Plasmid	Backbone size (bp)	Identity to backbone of pA85-3	Location of SNPs or gaps (base change, aa change)
pS22-2/ pS23-2/ pS25-2/ pS40-2	69,523	69,523/69,524 (1 gap)	- Missing T in repeat 2
pS27-2	69,524	69,523/69,525 (2 gaps)	- Missing T in repeat 2 - Extra A in an intergenic region
pS28-2	69,523	69,522/69,524 (1 gap)	- orf (A to T, M→L) - Missing T in repeat 2
pS32-2/ pS36-2 ¹	69,521	69,510/69,525 (10 gaps)	-Spread over backbone (intergenic regions, repeats or cause single aa substitutions in ORFs)
pS34-2	69,519	69484/69532 (21 gaps)	-truncates an ORF (392aa instead of 426aa) -Spread over backbone (intergenic regions, repeats or cause single aa substitutions in ORFs)

¹Are one SNP different from one another.

The nine plasmids in Table 8.10 are missing a base in repeat 2 compared to pA85-3. This missing base is the only difference between the backbone of pA85-3 and that of pS22-2, pS23-2, pS25-2 and pS40-2. The plasmid pS27-2 has gained an extra A in an intergenic region, while a SNP in an orf which causes a single aa substitution in a hypothetical located in the region between repeat 1 and repeat 2 (marked by an asterisk in Fig 8.10) is present in pS28-2 (Table 8.10).

The backbone of pS32-2 and pS36-2 are 69,521 bp in size and only have one difference between them. This SNP is located in an orf and is a silent mutation. Compared to the backbone of pA85-3 there are a greater number of differences in pS32-2 and pS36-2. The nucleotide identity to pA85-3 of pS32-2 and pS36-2 is 69,510/69,525 bp and 69,511/69,525 bp respectively, and both contain 5 single nucleotide gaps (Table 8.10). These 14 or 15 SNPs are spread over the backbone of these two plasmids and are located either in intergenic regions, repeats or cause single aa substitutions in orfs of unknown function.

The backbone of pS34-2 is 69,519 bp in size and its identity to pA85-3 is 69,484/69,524 bp, with 21 gaps (Table 8.10). A large portion of these SNPs, 36, were concentrated in a 50 bp region resulting in the truncation of an orf from 426 aa in pA85-3 to 392 aa in pS34-2 (indicated by the cross in Fig 8.10). Again, the protein it encodes has no known function. The other SNPs are spread throughout the backbone and, like those in pS32-2 and pS36-2, are located in either intergenic regions, repeats or cause single aa substitutions in orfs of unknown function.

The ability of these plasmids to conjugate was determined for the seven of these plasmids that contained a selectable marker, the *oxa23* gene in AbaR4, conferring resistance to ticarcillin. Each isolate was mated with ATCC 19606^{Rif} (rifampicin resistant) and transconjugants were selected on media supplemented with rifampicin and ticarcillin. Plasmids pS22-2, pS23-2, pS40-2, pS27-2, pS25-2 and pS40-2 transferred to the recipient (Table 8.11).

As a number of these plasmids are closely related to one another and have the same insertions in the same positions in the plasmid backbone, a representative of each group (bold in Table 8.11) was selected and conjugation frequencies utilising the same backgrounds were determined. ATCC 19606^{Rif} containing pS22-2, pS27-2 or pS34-2 was mated with AB307_0294^{Nx} (nalidixic acid resistant) and transconjugants were selected using L-agar supplemented with nalidixic acid and ticarcillin. The representative of type 2 plasmids conjugated at a higher frequency than the type 1 plasmids pS24-1 (Table 8.7 and Table 8.11). The plasmid pS40-2 transferred at the highest frequency of 2.02×10^{-3} transconjugants/donor while pS22-2 and pS27-2 transferred at a frequency of 8.05×10^{-4} and 2.02×10^{-4} transconjugants/donor, respectively (Table 8.11).

Table 8.11 Conjugation of type 2 plasmids

Plasmid	Inserts	Resistance	Mating of each strain with ATCC 19606 ^{Rif}		Mating of ATCC 19606 ^{Rif} (RepAci6 plasmid) with AB307 0294	
			Transconjugants (TIC ^{Rif})	Conjugation frequency ²	Conjugation frequency ²	Mean Conjugation frequency ³
pS32-2	ISAba19	-	+ ¹	ND		
pS36-2	ISAba19	-	ND			
pS22-2	AbaR4(1)	TIC	+	1.49x10 ⁻⁴	2.22x10 ⁻⁴ 8.04x10 ⁻⁴ 1.39x10 ⁻³	8.05x10 ⁻⁴
pS23-2	AbaR4(1)	TIC	+	3.20x10 ⁻⁵	ND	
pS40-2	AbaR4(1)	TIC	+	6.13x10 ⁻⁵	ND	
pS27-2	AbaR4(1) ISAba1	TIC	+	2.43x10 ⁻⁷	1.31x10 ⁻⁴ 7.26x10 ⁻⁴ 3.33x10 ⁻⁴	3.96x10 ⁻⁴
pS25-2	AbaR4(1) ISAba1	TIC	+	1.04x10 ⁻⁵	ND	
pS28-2	AbaR4(1) ISAba1	TIC	-	<1.80x10 ⁻¹⁰		
pS34-2	AbaR4(2)	TIC	+	2.95x10 ⁻⁵	4.38x10 ⁻³ 1.28x10 ⁻³ 3.87x10 ⁻⁴	2.02x10 ⁻³

TIC; ticarcillin.

¹pS32-2 was able to mobilise a small antibiotic resistant plasmid. Details in Chapter 9; Section 9.5.3.

²Conjugation frequency expressed as transconjugants per donor. Each frequency is from a single determination.

³Mean frequency expressed as transconjugants per donor calculated from three independent determinations.

However, the plasmid pS28-2 did not transfer above the limit of detection (Table 8.11). The inability of pS28-2 to transfer is not accounted for by the location of its inserts, as they are in the same position as pS27-2 and pS25-2 and both of these plasmids can conjugate. In addition, pS28-2 is only one SNP away from pS25-2 (Table 8.10) and this difference causes an amino acid substitution in a protein of unknown function (marked by an asterisk in Fig 8.10). However, whether this orf performs a vital role in conjugation needs closer examination. It is possible that additional effects such as fertility inhibition are playing a role.

Despite the absence of any antibiotic resistance genes, pS32-2 was determined to be self-transmissible through mobilisation of another co-residing plasmid, pS32-1, which contains the *oxa23* carbapenem resistance gene as a selectable marker. This is described in more detail in

Chapter 9; Section 9.5.3. As pS32-2 and pS36-2 are only one SNP different (Table 8.10), it is likely that pS36-2 is also conjugative.

8.5 The plasmid pS21-2

Despite producing the amplicon size expected for type 2 plasmids over repeat 3 (Section 8.2), the RepAci6 backbone contigs of S21 (Table 8.12) were a better match to the type 1 backbone, although regions of difference were still present (Fig 8.11A).

Table 8.12 S21 contigs matching RepAci6 plasmid

Contig number	Contig length (bp)	Kmer coverage	Description	Contig edges (L, R) ^{1,2}
10	63,879	31.8	F1 (extra 35.6 kb), <i>tra</i> region	ISAbal2-like, IRI of Tn6022
20	11,091	31.9	F3	R3, ISAbal2-like
38	12,233	31.8	F1, <i>repAci6</i>	IRr of Tn6022
41	365	59.6	Repeat sequence	
50	11,828	31.8	F2	R1, R3
55	4,672	30.9	F1, <i>trw</i> genes	R2, R3
120	2,272	29.8	Novel	R sequence at both ends

¹Contig edges from Left to Right.

²R1, R2 and R3 stand for repeat 1, 2 and 3 respectively

Sequencing the >3 kb product of the reaction between RH1398-RH1399 over Repeat 1 in S21 (Section 8.2), revealed that an additional contig (120), with repeat sequence on its edges, was present between contig 38 and contig 50. The RepAci6 plasmid in S21 contains 2.2 kb of sequence adjacent to repeat 1, followed by another partial repeat sequence which has been called repeat 4.

In addition to the breaks in the assembly at repeat 2 and 3 which had been screened previously, there were two additional breaks in pS21-2. Assembly of contig 20 and 10, using RH1391 and RH2021, revealed an ISAbal2-like element, surrounded by a 9 bp target site duplication (5'

GTTCATAGA 3') (Fig 8.11A) (Table 8.13). The sequence of this IS is 97.3% (1011/1039) identical to the sequence of ISAba12.

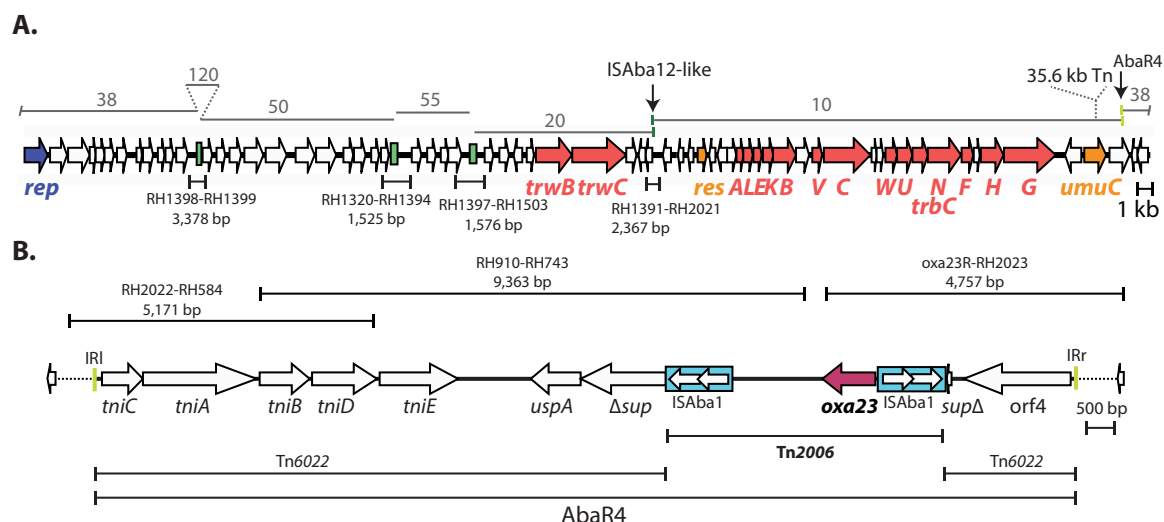


Figure 8.11 Assembly of pS21-2 contigs. A. pS21-2 contigs tiled to pAb-G7-2. Arrows indicate the extent and orientation of genes and open reading frames. The *repAci6* gene is blue, genes involved in conjugal transfer are red and the *res* and *umuC* genes are orange. Green boxes indicate the three repeat sequences. Grey lines above indicate the extent of different pS21-2 contigs. The black bars underneath show the primers used for assembly along with PCR amplicon size. Vertical arrows indicate the insertion location of *AbaR4* and *ISAba12*-like element. B. Assembly of *AbaR4* in pS21-2. The *oxa23* gene is magenta and the IR sequences of *Tn6022* are shown in olive green. The blue box indicates *ISAba1* and the internal arrows show the transposase orientation. The primers used to span *AbaR4* are shown above by the black bars, along with PCR amplicon size. The lines below indicate the transposons.

The break between contig 10 and 38 revealed another location for *AbaR4* (Table 8.13). The IR of *Tn6022* was found on the left and right edges of contig 38 and 10, respectively, as was a 5 bp TSD (5'-TCTAC-3') (Table 8.13). This region was then screened for the presence of *AbaR4* (Fig 8.11B) using a new primer on the end of contig 10, RH2022, with RH583 to generate a 5,171 bp amplicon. The right end of *AbaR4* was detected using *oxa23R* with a new primer on the end of contig 38 to amplify a 4,757 bp region (Fig 8.11B). In both cases, an amplicon of the expected size was produced and sequenced, confirming the presence of *AbaR4* in this location in pS21-2. The internal *AbaR4* PCR of RH910 and RH743, linking *Tn2006* to *tniA*, was also performed and produced an amplicon at the approximate size of 10 kb and the

predicted size was 9,363 bp. Again, AbaR4 in pS21-2 interrupts an orf, but the function of the product of this orf is unknown (Table 8.13).

Table 8.13 pS21-2 insertions

Insertion	Location	Target site duplication (5'-3')	Description of interrupted protein	Novel position?
ISAbal2-like	Intergenic region (between 2 orfs located 3 or 4 downstream of <i>trwC</i>)	GTTCATAGA	N/A	Y
35.6 kb Tn	<i>umuC</i>	GCTTA	Error-prone DNA polV subunit	Y
AbaR4	Interrupts a gene encoding a 372 aa protein. [†]	TCTAC	No conserved domains- ORF	Y

[†]AbaR4 in pS34-2 also interrupts this gene but in a different position.

An additional 35.6 kb of sequence (Fig 8.11A) was also identified in contig 10 and this transposon is described in detail in Section 8.5.3. The plasmid pS21-2 is 123,432 bp in size (Fig 8.12) and contains an insertion adjacent to repeat 1, followed by repeat 4, an ISAbal2-like IS, 35.6 kb of additional sequence and a new insertion location for AbaR4 (Table 8.13).

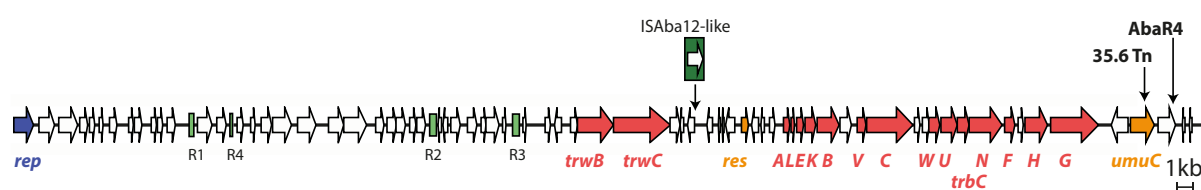


Figure 8.12 Linear schematic of pS21-2. Arrows indicate the extent and orientation of genes and open reading frames. The *repAci6* gene is blue, genes involved in conjugal transfer are red and the *res* and *umuC* genes are orange. Light green boxes indicate the three repeat sequences, while the dark green box represents ISAbal2-like and the internal arrow indicates the transposase orientation. Vertical arrows show the location of insertions with the identity of the insertion is shown above.

8.5.1 pS21-2 conjugation

The ability of ticarcillin resistance to be transferred by the conjugation of pS21-2 was tested by mating S21 with rifampicin resistant ATCC 19606^{Rif}. From three independent experiments, no ticarcillin and rifampicin resistant transconjugants were recovered, averaging a conjugation frequency for pS21-2 of less than 6.55×10^{-9} transconjugants/donor. Although the backbone of pS21-2 is different to type 1 and 2 plasmids, and this is addressed in more detail in the following section, all of the genes involved in transfer appear to be intact, with no premature stop codons. Hence, it may be the location of the insertions, particularly that of the 35.6 kb of extra sequence, that has diminished the self-transmissibility of pS21-2 (Fig 8.12).

8.5.2 pS21-2 is a new backbone type

The backbone of pS21-1 is 69,951 and is significantly different to both type 1 and type 2 plasmids (Fig 8.13). The region between repeat 3 and 1, encoding genes involved in transfer and the replication initiation protein in pS21-2 is within 1% difference to type 1 plasmids. The greatest region of variation appears to be between repeat 1 and 3 with approximately a 2-3.3% difference to type 1, and a 3-4% difference to type 2 (Fig 8.13), and the additional segment is within this region. Hence, pS21-2 represents a new backbone type, type 3.

Repeat 1 and 4 are partial compared to repeat 2 and 3. Repeat 4 is even smaller than repeat 1, but they are both fragments of the right end of the complete repeat sequence (Fig 8.14A). The additional 2.2 kb sequence between repeat 1 and repeat 4 in pS21-2 has two reading frames (Fig 8.14B). The first produces a 363 aa hypothetical protein which hits the CysH superfamily (COG0175) and is most commonly annotated as a phosphohydrolase. The second open reading frame is commonly annotated as hypothetical protein and translates to a 263 aa protein.

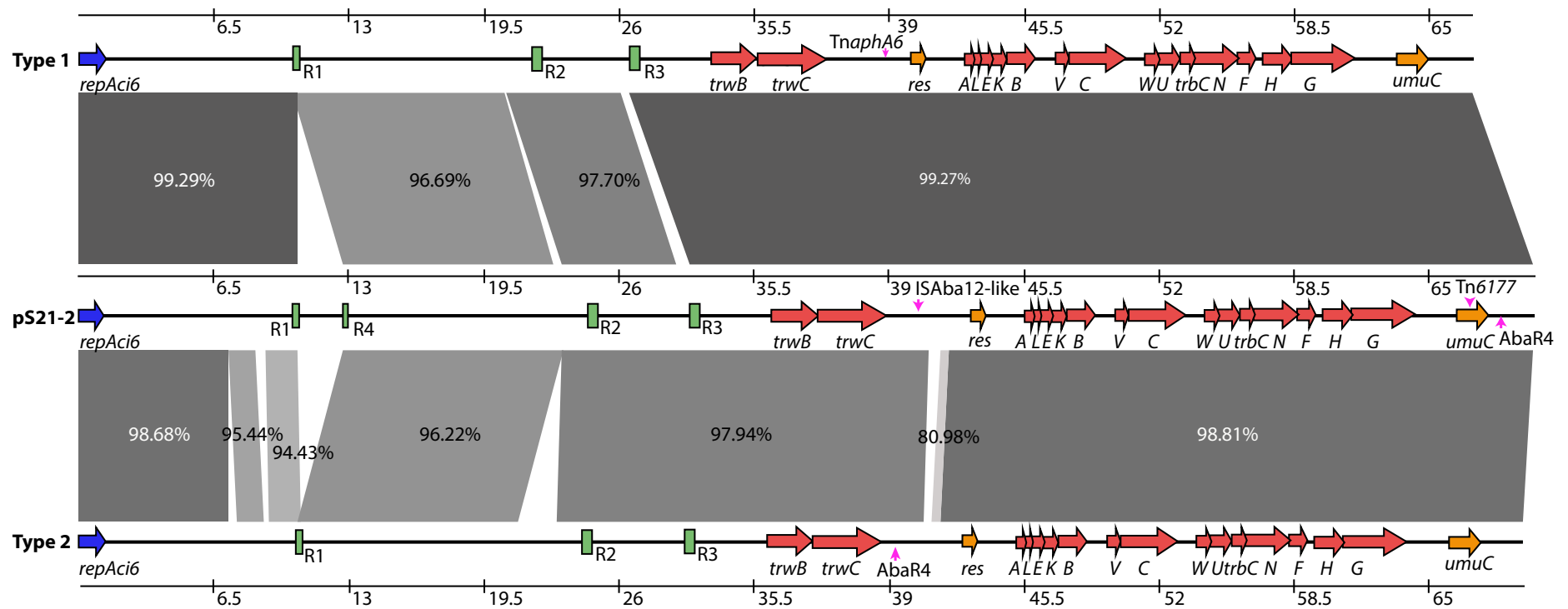


Figure 8.13 pS21-2 represents a new and third RepAci6 backbone type. Arrows indicate the extent and orientation of genes and open reading frames. The *repAci6* gene is blue, genes involved in conjugal transfer are red and the *res* and *umuC* genes are orange. Green boxes indicate the three repeat sequences. Pink vertical arrows show the location of insertions with the identity of the insertion shown above. Regions of homology are shown by grey boxes and the numbers display DNA identity. Figure is made to scale from GenBank accession number KF669606 (pAb-G7-2 for type 1) and KJ493819 (pA85-3 for type 2).

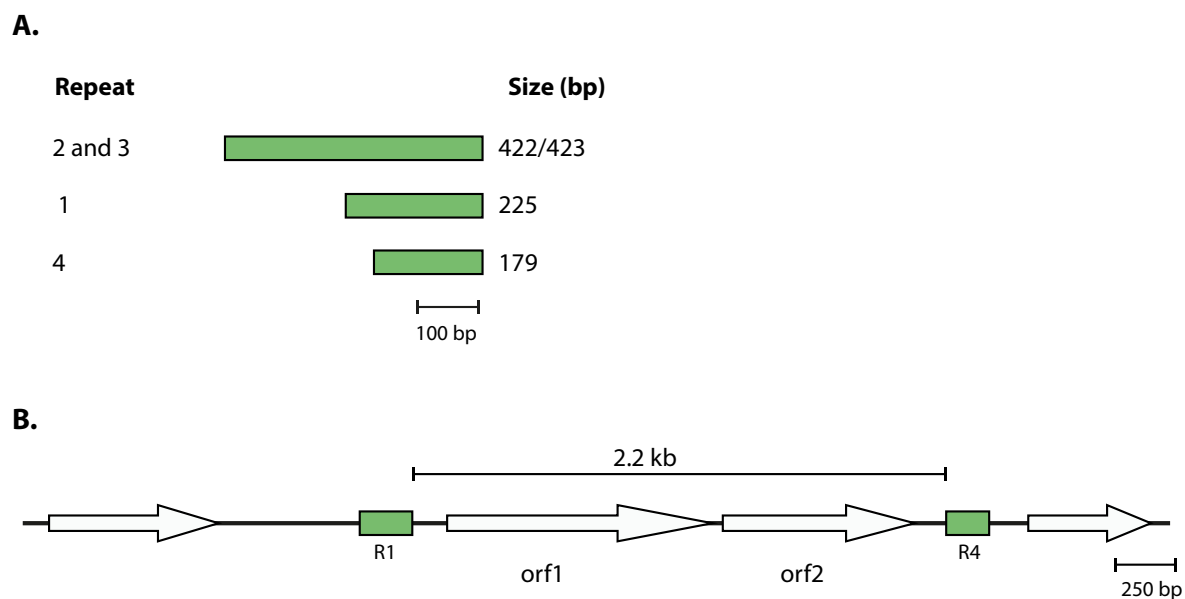


Figure 8.14 A fourth repeat and novel sequence. A. Comparison of repeat lengths. Green boxes indicate repeat sequences. B. Additional 2.2 kb between repeat 1 and repeat 4. Arrows indicate open reading frames.

This segment along with repeat 4 is also found in 11 plasmids from *A. baumannii* isolated in South Korea (Table 8.14).

Table 8.14 Plasmids with repeat 4 and the 2.2kb of sequence

Plasmid	Location	GenBank Accession Number	Reference
pUSA2_1	Seoul, South Korea	CP020593	
p15A5_1	South Korea	CP020573	
pYU-R612-1	Seoul, South Korea	CP014216	(156)
p15A34_1	South Korea	CP020589	
pSSA12_1	Seoul, South Korea	CP020577	
pKAB07	Jeonju, South Korea	CP017655	
pKAB02	Daejeon, South Korea	CP017645	
pABKp1	South Korea	CP001922	(202)
pKAB05	Daegu, South Korea	CP017651	
pKAB01	Daegu, South Korea	CP017643	
pKAB06	Daegu, South Korea	CP017653	

8.5.3 Tn6177

35.6 kb of additional sequence was present downstream of the *tra* genes in the backbone of pS21-1. Comparison to the type 1 backbone enabled the boundaries of this extra sequence to be defined as well as its precise location. The extra sequence interrupts the *umuC* gene and is

surrounded by a 5 bp target site duplication (5'-GCTTA-3') (Table 8.13). The outer ends of the 35.6 kb (Fig 8.15) were then examined to determine if inverted repeats (IRs) were present. An imperfect 26 bp IR (blue in Fig 8.15), with 24 bp identical, was identified starting with 5' TGT 3'. This IR is about the same size and starts with the same bases as Tn6019, the backbone transposon of AbaR4. Partial duplications of the IR sequence were also detected in the sequence internal to the IRs (Fig 8.15).



Figure 8.15 Identification of the inverted repeat of the 35.6 kb transposon. The bases highlighted in blue show the terminal IR sequences. The top line shows IRl while the bottom line is IRr sequence. Arrows indicate internal partial copies of the IR sequence and the bases in bold are identical to the sequence of the IR.

Due to the presence of an IR, target site duplication and the size of the insert, it is likely a transposon and has been named Tn6177. Tn6177 harbours 26 open reading frames and pBLAST was used to attempt to determine their putative function (Figure 8.16, Table 8.15).

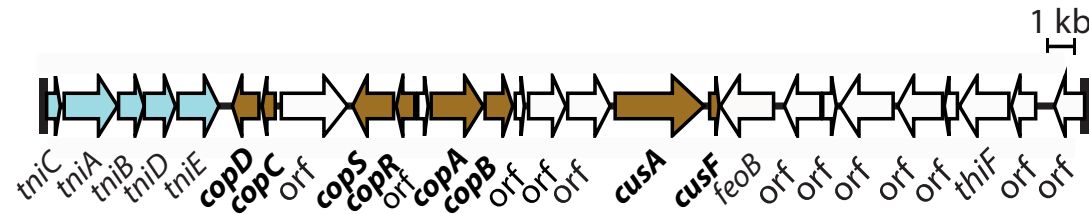


Figure 8.16 Structure of the Tn6177 copper resistance transposon. Arrows indicate the extent and orientation of genes and open reading frames. Transposition genes are in blue while known genes involved in copper resistance are in brown. Vertical lines indicate the IR sequences.

Table 8.15 Open reading frames and predicted products encoded by Tn6177

orf	AA	Superfamily	Description	Product
1	237	Tn7_Tnp_TnsA (pfam08722)		TniC
2	629	Rve (pfam00665)		TniA
3	283	TniB (pfam05621)		TniB
4	361	TniQ (pfam06527)		TniD
5	480	-		TniE
6	328	CopD (pfam05425)		CopD
7	126	CopC (pfam04234)		CopC
8	785	ATPase-1B1_Cu (TIGR01511)	copper-translocating P-type ATPase	Orf
9	475	czts_silS_copS (TICR01386)		CopS
10	227	cstR_silR_copR (TICGR01387)		CopR
11	127	PcoB (COG3667)	uncharacterized protein involved in copper resistance	Orf
12	627	copper_res_A (TIGR01480)		CopA
13	306	CopB (pfam05275)		CopB
14	153	-		Orf
15	434	TolC (COG1538)	TolC family protein	Orf
16	499	HlyD_D23 (pfam16576)	barrel-sandwich domain of CusB HlyD membrane-fusion	Orf
17	1048	CusA (COG3696)		CusA
18	120	CusF_Ec (pfam11604)		CusF
19	609	FeoB (COG0370)		FeoB
20	438	Cation_efflux (pfam01545)	Co/Zn/Cd efflux Cd(II)/Pb(II)-responsive	Orf
21	130	HTH_CadR-PbrR (TIGR02047)	transcriptional regulator	Orf
22	601	-		Orf
23	533	HNH_2 (pfam13391)	HNH endonuclease	Orf
24	163	Prok-JAB (pfam14464)		Orf
25	576	ThiF (pfam00899)		ThiF
26	346	NT_2-5OAS_ClassI-CCAase (cd05400)	nucleotidyltransferase	Orf
27	368	HTH_XRE (cd0093)		Orf

The first 5 open reading frames show homology to the transposition proteins produced by the backbone transposon of AbaR4, Tn6019 (Table 8.16), and two of them show shorter regions of similarity to the transposition proteins of Tn7. Similar to their role in Tn6019, these proteins likely form the transposition module of this transposon and facilitate its movement.

Table 8.16 Relationship of the predicted transposition proteins of Tn6177, Tn6019 and Tn7

35.6 kb Tni proteins	Similarity to respective proteins in Tn6019	Similarity to respective proteins in Tn7 ¹
TniC (2358aa)	37% 81/218	NS
TniA (630aa)	45% 245/548	23% 129/571
TniB (384aa)	40% 114/283	22% 56/252
TniD (362aa)	31% 112/360	NS
TniE (481aa)	65% 226/466	NS

¹*E. coli* strain 3.5-R3 transposon Tn7, GenBank accession number KX117211. TniA, TniB, TniD and TniE were matched to TnsA, TnsB, TnsC, TnsD and TnsE, respectively. Non-significant matches are noted by NS.

A number of the later open reading frames have homology to either Cop or Cus proteins which are known to be involved in copper resistance (241) (Table 8.15, Fig 8.16). However, copper resistance was not examined. This transposon also includes *feoB* gene involved in iron transport and a *thiF* gene which is involved in thiamine synthesis.

8.6 pS30-2

Only one contig matching the RepAci6 plasmid backbone was retrieved from the draft genome of S30 sequenced by Illumina MiSeq (Table 8.17). This contig fragments of ISAbal on both ends. Adjacent to the IRr end of the ISAbal was sequence of F1 just prior to repeat 2 while on next to the IRL end of ISAbal was the *oxa23* gene located in Tn2008 and then repeat 2.

Table 8.17 RepAci6 plasmid contig of S30

Isolate	Contig	Length (bp)	Description	Edges
S30	18	73,043	F1 (TnaphA6), F2 and F3	ISAbal (IRl), ISAbal (IRr)

To circularise the sequence, the *oxa23R* primer was combined with RH1320 and produced an approximately 2 kb amplicon that matched the predicted size and was sequenced. The *oxa23*

gene in pS30-2 was in Tn2008A, instead of AbaR4. Tn2008A in pS30-2 was surrounded by a 9 bp target site duplication (5'-AGATAATTA-3') and this is a novel insertion site for this transposon (194) that is located just prior to repeat 2 (Fig 8.17). The different context of the *oxa23* gene in pS30-2 is consistent with the findings from the earlier *oxa23* screening Chapter 3; Section 3.3.3. pS30-2 also includes the same 2.2 kb adjacent to repeat 1, followed by repeat 4, as in pS21-2 (Fig 8.17).

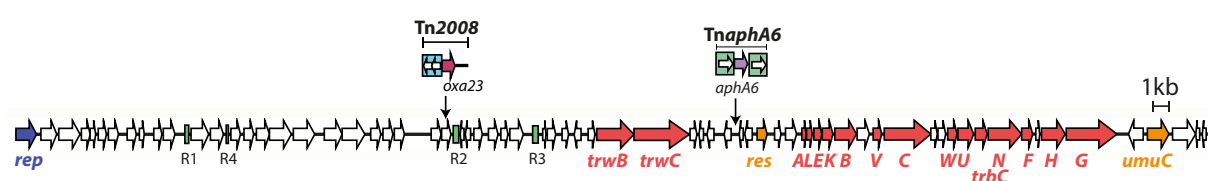


Figure 8.17 Linear schematic of pS30-2. Arrows indicate the extent and orientation of genes and open reading frames. The *repAci6* gene is blue, genes involved in conjugal transfer are red and the *res* and *umuC* genes are orange. The *oxa23* and *aphA6* gene are magenta and purple, respectively. Green boxes indicate the three repeat sequences. The light blue and green boxes represent ISAbal and ISAbal25, respectively, and the internal arrow indicates the transposase orientation. Vertical arrows show the location of insertions with the identity of the insertion shown above.

8.6.1 New backbone type 4

The plasmid pS30-2 amplified the type 1 amplicon size over R3 in Section 8.2, but its 71,056 bp backbone appears to have key features of the three different backbone types (Fig 8.18). A large section of the backbone, including the *tra* region (boxed in yellow in Fig 8.18), is within .3% to that of type 1 plasmids, and pS30-2 contains *TnaphA6* in the same location as pAb-G7-2 (indicated by the pink arrow in Fig 8.18). However, there are sections of greater difference and pS30-2 contains extra sequence compared to type 1 plasmids. Overall, pS30-2 is more distantly related to type 2 plasmids but contains an additional segment upstream of repeat 2 (boxed in blue in Fig 8.18) that is present in this plasmid type. Thirdly, the extra sequence following repeat 1, as well as the partial duplication of this repeat is closely related to that of pS21-2 (green in Fig 8.18). The backbone of pS30-2 appears to be a fourth and hybrid backbone type, with a large segment closely related to the type 1 backbone but also contains segments

from the type 2 and 3 backbones, demonstrating the recombination that can occur in these closely related plasmids.

8.6.2 Conjugation

Transconjugants that were rifampicin and ticarcillin resistant were recovered from the mating of S30 with rifampicin resistant ATCC 19606^{Rif}. The mean conjugation frequency of pS30-2 was 8.38×10^{-2} transconjugants per donor was averaged from three independent matings of ATCC 19606^{Rif} (pS30-2) with AB307_0294 (Table 8.18). This frequency is much higher than that determined for the other plasmids, indicating that conjugation may have somehow been de-repressed in pS30-2.

Table 8.18 pS30-2 conjugation

Plasmid	Inserts	Resistance	Mating of S30 with ATCC 19606 ^{Rif}		Mating of ATCC 19606-Rf ^R (pS30-2) with AB307_0294	
			Transconjugants (TIC ^{Rif})	Conjugation frequency ¹	Conjugation frequency ¹	Mean Conjugation frequency ²
pS30-2	Tn2008, TnaphA6	TIC Km	+	2.34×10^{-5}	1.18×10^{-1} 7.5×10^{-2} 5.75×10^{-2}	8.38×10^{-2}

TIC, ticarcillin; Km, kanamycin.

¹Conjugation frequency is expressed as transconjugants per donor. Each frequency is from a single determination.

²Mean frequency expressed as transconjugants per donor calculated from three independent determinations.

8.7 Discussion

Fourteen RepAci6 plasmids from the isolates examined in this project were assembled and characterised. The backbones of most resembled the two known backbone types, including three type 1 plasmids and 9 type 2 plasmids. However, single examples of two novel backbone types, pS21-2 (type 3) and pS30-2 (type 4), were also identified. Plasmids in this family clearly contribute to the spread of *oxa23* and *aphA6*. Twelve of the plasmids carry the *oxa23* gene, either in Tn2008 or AbaR4 and hence, can contribute to carbapenem resistance. The AbaR4 transposon has been repeatedly acquired by RepAci6 plasmids as it was found in two positions

in the type 1 backbone, two positions in the backbone of type 2 plasmids, as well as in another different location in pS21-2. In addition to carrying carbapenem resistance, the three type 1 plasmids and the type 4 plasmid pS30-2, also contain the amikacin resistance transposon *TnaphA6*. Two novel transposons, one containing copper resistance genes, were also detected in the RepAci6 plasmids examined. The majority of these plasmids were shown to be conjugative and hence RepAci6 plasmids are playing a significant role in the spread of genes conferring resistance to carbapenems and amikacin, and both of these antibiotics are important for the treatment of *A. baumannii* infections. The reasons for failure of pS26-1, pS29-1, pS28-2 and pS21-2 to conjugate will require further investigation.

8.7.1 Type 1 plasmids

Three of the *TnaphA6* containing plasmids, pS24-1, pS26-1 and pS29-1, have a type 1 backbone and are the first plasmids of this type to contain AbaR4. Twelve other RepAci6 plasmids in the non-redundant database of GenBank were found to also be members of this backbone type (Table 8.19). These plasmids were isolated from isolates of different STs indicating spread (Table 8.19). A number of different transposons and IS have inserted into type 1 plasmids (Table 8.19) and the location of these transposons are indicated in Fig 8.19. Two of the type 1 RepAci6 plasmids in GenBank did not contain *TnaphA6*, but when this transposon was present, it was commonly found in position A (originally identified in pAb-G7-2) (Table 8.19, Fig 8.19) (97). However, there were also two examples of *TnaphA6* in position B (the pD72-1 *TnaphA6* location) (192) and a single example of the position C insertion location in pA105-1(135), indicating that this transposon has been picked up by type 1 RepAci6 plasmids on at least three occasions (Table 8.19, Fig 8.19). A 3.3 kb transposon containing *aphA6* in pAba3207b (35) is bounded by directly-orientated copies of ISAba125, but the internal segment shares only 98% nucleotide identity to that of *TnaphA6* and is 310 bp

longer (Table 8.19). Hence, this *aphA6*-containing transposon in pAba3207b is the second time that ISAbal25 has mobilised the *aphA6* gene from *A. guillouiae* (279).

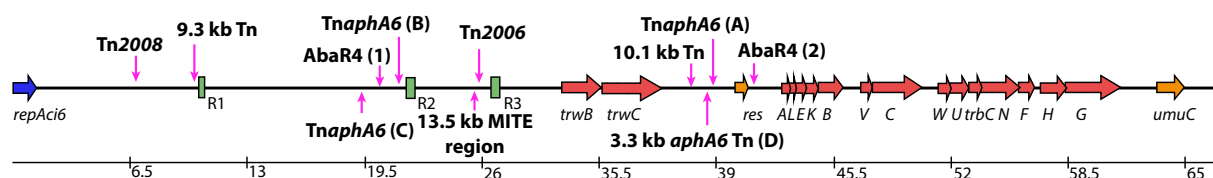


Figure 8.19 Location of insertions in the type 1 backbone. Arrows indicate the extent and orientation of genes. The *repAci6* gene is blue, genes involved in conjugal transfer are red and the *res* and *umuC* genes are orange. Green boxes indicate the three repeat sequences. Vertical pink arrows show the locations of insertions in Type 1 plasmids listed in Table 8.19. Figure is made to scale using pAb-G7-2 (GenBank accession number KF669606).

The plasmids pS24-1, pS26-1 and pS29-1, also contain the *oxa23* containing transposon AbaR4, in two novel locations. AbaR4 in the position found in pS26-1 and pS29-1 was also present in a plasmid from a South Korean isolate, pUSA15_1 (Table 8.19). In fact, pUSA15_1 also contains the 9.3 kb transposon and the only difference between pUSA15_1 and pS26-1/pS29-1 is the presence of 2 copies of ISAbal3 in pUSA15 (Table 8.19).

However, AbaR4 is not the only vehicle carrying the *oxa23* gene found in Type 1 RepAci6 plasmids. The plasmid pD46-3 was previously reported to contain Tn2006 (196). Tn2008 has been reported in the plasmid pABUH1-74 (GenBank accession number NZ_AYOH01000010.1) (277) and this Tn2008 location was also found in pCMCVTab2-Ab4 and pCMCVTab2-Ab66 (Table 8.19). Other resistance genes have been acquired by p1AB5075 and pAB8098 in a 13.5 kb region flanked by mite sequences, which contains the *sull*, *dfrA7*-*aacA4*-*bla*_{GES-11}, *aadA2*-*cmlA1*-*aadB* as well as the *strA*, *strB* resistance genes (74). The plasmid pC13-2 contains Tn6166, a transposon carrying copper resistance genes, which is quite different to Tn6167 identified in pS21-2.

Table 8.19 Type 1 plasmids in GenBank as of July 2017

Plasmid	ST of isolate (IP) ¹	Size (bp)	<i>TnaphA6</i> location ²	Additional insertions	Location	Year	GenBank Accession number	Reference
pAb-G7-2	1	70,100	A		Melbourne, Australia	2002	KF669606	(97)
pS24-1	2	86,916	A	AbaR4 (1)	Singapore	2006		This study
pS26-1/pS29-1	2	96,203	A	AbaR4 (2), 9.3 kb Tn	Singapore	2006/7		This study
pCMCVTab2-Ab4	2	74,090	A	Tn2008 ³ , ISAbal	Virginia, USA	<2017	CP016297	
pC13-2	2	103,871	A	Tn6176 ⁴	Sydney, Australia	2006	KU549175	(194)
p1AB5075	1	83,610	A	13.5 kb region bounded by mites ⁵	USA	2008	CP008707	(74)
pAB8098	499	82,667 ⁶	A	13.5 kb region bounded by mites ⁵	Tunisia	2010	KY022424	
pUSA15_1	1	98,301	A	AbaR4 (2) ⁷ , 9.3 kb Tn, 2xISAbal3	Seoul, South Korea	2013	CP020594	
pAB34299	2	84,967 ⁸	A	ISAbal	-	2015	CP014292	
pD72-2	2	70,102	B		Sydney, Australia	2010	KM051846	(192)
pD46-3	25	74,916	B	Tn2006 ⁹	Sydney, Australia	2013	KM977710	(196)
pA105-1	636	70,098	C		Sweden	2008	KR535992	(135)
pAba3207b	422	80,546	D ¹⁰	10.1 kb additional sequence ¹¹	Monterrey, Mexico	2010	CP015366	(35)
pCMCVTab2-Ab66	2	73,188	- ¹²	Tn2008 ³ , 2xISAbal25	Virginia, USA	<2017	CP016302	
pMAL-2	492				Serbia	<2016	KX230794	(57)

¹Using the Institut Pasteur MLST scheme, of the isolate containing the plasmid. ST492 is a single locus variant of ST2, while ST636 is a triple locus variant. ST25, ST499 and ST422 are not related to ST1 or ST2.

²pAb-G7-2 location is position A, position B is the location found in pD72-2 while position C and D are unique to pA105-1 and pAba3207b, respectively.

³Interrupts an orf 3,055 bp prior to R1.

⁴33.8 kb transposon that contains copper resistance genes, but is different to Tn6177.

⁵Inserted 770 bp prior to repeat 2 and the region contains the *sulI*, *dfrA7*-*aacA4*-*bla*_{GES-11}, *strA*, *strB*, *aadA2*-*cmlA1*-*aadB* resistance genes.

⁶Is missing the centre 942 bp of the ISAbal25 on the right in *TnaphA6*.

⁷AbaR4 is located in the same position as in pS26-2 and pS29-2 as is the 9.3 kb Tn.

⁸Duplication of 13.7 kb of plasmid backbone.

⁹Located 728 bp upstream of repeat 2 and interrupts an orf.

¹⁰The ISAbal25 bounded transposon is 3,382 bp in size and has 310 bp extra prior to the right ISAbal25 compared to *TnaphA6*.

¹¹Located 2,395 bp prior to the *TnaphA6*(A) insertion site and is surrounded by a 9 bp duplication.

¹²Has two copies of ISAbal25 in opposite orientation and the 4.7 kb of backbone sequence between them is inverted. If the sequence is re-inverted then the one of the copies of ISAbal25 is in the same position as *TnaphA6* in pAb-G7-2.

8.7.2 Type 2 plasmids

Nine of the RepAci6 plasmids characterised in this work had a type 2 plasmid and this backbone type was only present in the isolates that contain a version of AbGRI3. Six of these plasmids harboured AbaR4 in the same location as pA85-3 (location 1) (99), while a novel insertion site of AbaR4 was detected in pS34-2 (location 2). In addition to pA85-3, only three other type 2 plasmid sequences were found in GenBank (Table 8.20). Like pS32-2 and pS36-2, none of these plasmids contained resistance genes but other insertions were present (Table 8.20).

Table 8.20 Type 2 plasmids in GenBank as of July 2017

Plasmid	Size (bp)	Additional insertions	Location	Year	GenBank accession number	Reference
pA85-3	86,334	AbaR4 (1)	Sydney, Australia	2003	KJ493819	(99)
pS22-2, pS23-2, pS40-2	86,340	AbaR4 (1)	Singapore	2004-2011		This study
pS25-2, pS27-2, pS28-2	87,529	AbaR4 (1), ISAbal	Singapore	2006-2007		This study
pS34-2	86,336	AbaR4 (2)	Singapore	2009		This study
pS32-2, pS36-2	70,833	ISAbal9	Singapore	2009		
pJBA13_1	90,972	20.4 kb Tn ¹ , ISAbal10	Jeonbuk, South Korea	2013	CP020582	
pSSMA17_1	90,973	20.4 kb Tn ¹ , ISAbal10	Seoul, South Korea	2013	CP020580	
pKAB04	121,485 ²	9.3 kb Tn ³ , 20.4 kb Tn ¹ , 2xISAbal10	Busan, South Korea	2015	CP017649	

¹Class 3 family transposon that also includes a copy of ISAbal10.

²19,166 bp of the backbone is duplicated.

³Located downstream of repeat 3 and orf 4 of the transposon is interrupted by ISAbal10.

8.7.3 Further backbone types

Two novel backbone types, pS21-2 and pS30-2, were also discovered in the isolates examined in this chapter. The backbone of pS21-2 has large regions of sequence that are different (up to 4-6% nucleotide different) to both type 1 and type 2 plasmids, as well as 2.2 kb of extra sequence adjacent to repeat 1 and a copy of repeat 4. The backbone of pS30-2 appears to be

Table 8.21 Other RepAci6 plasmids in GenBank as of July 2017

Plasmid	Size (bp)	Backbone features				Backbone type	Additional insertions	Location	Year	GenBank accession number	Reference
		2.2 kb and R4	1.2 kb prior to R2	1 kb downstream of R3	<i>parB</i> downstream of <i>umuC</i>						
pS21-2	123,432	+				3	Tn6167 ¹ , AbaR4 ²	Singapore	1996		This study
pS30-2	76,933	+	+			4	TnaphA6(A) ³ , Tn2008 ⁴	Singapore	2008		This study
pKAB08	101,406		+	+	+	5	9.3 kb Tn ⁵ , 20.4 kb Tn ⁶ , ISAbal0	Gyeonggi, South Korea	2015	CP017657	
pUSA2_1	74,241	+	+		+	6		Seoul, South Korea	2013	CP020593	
p15A5_1	74,241	+	+		+	6		South Korea	2013	CP020573	
pYU-R612-1	74,241	+	+		+	6		Seoul, South Korea	2014	CP014216	(156)
pKAB05	70,884	+			+	7		Daegu, South Korea	2015	CP017651	
pKAB01	70,875	+			+	7		Daejeon, South Korea	2015	CP017643	
pKAB06	70,884	+			+	7		Daegu, South Korea	2015	CP017653	
pABKp1	74,451	+		+	+ ⁷	8	ISAbal	South Korea	<2010	CP001922	
pKAB07	72,073	+		+	+	8		Jeonju, South Korea	2015	CP017655	
p15A34_1	72,076	+		+	+	8		South Korea	2014	CP020589	
pSSA12_1	73,264	+		+	+ ⁷	8		Seoul, South Korea	2013	CP020577	
pKAB03	73,891				+	9	3xISAbal0	Daejeon, South Korea	2015	CP017647	
pAB_CC	65,890 ⁸				+	9		China	2013	KF889012	(161)
pAb04-2	87,569				+	10	AbaR4(1), ISAbal25 ⁹	Edmonton, Canada	2012	CP012008	(272)

¹Interrupts the *umuC* gene.

²Interrupts the orf downstream of *umuC*.

³Located in the same position as in pAb-G7-2.

⁴Interrupts the orf prior to repeat 2.

⁵Located 6,355 bp away from repeat 3 and orf 4 of the transposon is interrupted by ISAbal0.

⁶Class 3 family transposon that include a copy of ISAbal0.

⁷Adjacent to *parB* is ISAbal.

⁸Missing 5,977 bp of backbone including repeat 2.

⁹ISAbal25 is in the same position as *TnaphA6*(B) in pD72-2.

a hybrid of the three backbone types, indicating the recombination that can occur between these closely-related plasmids.

Examination of 14 additional RepAci6 plasmids in GenBank, revealed a number of further backbone types, based on combinations of specific backbone features and the high sequence similarity between them (Table 8.21). A subset of RepAci6 plasmids also contain a *parB* partitioning gene that is located downstream of *umuC* (Table 8.21). The plasmid pAb04-2 (type 10 in Table 8.21), appears to be another hybrid as it contains AbaR4 in the pA85-3 position (position 1), has *parB* but also contains an ISAba125 in the same location as *TnaphA6* in the type 1 plasmid pD72-1 (position B). The backbones of RepAci6 plasmids are diverse and further work will be required to characterise them fully.

8.7.4 Conserved orientation of AbaR4 transposon

Overall, AbaR4 has been found in a total of five positions in the different RepAci6 plasmid backbones. In each location, the AbaR4 transposon is always in the same orientation with respect to the *repAci6* gene, with IRL of Tn6022 on the left and IRr on the right. This orientation specific insertion into plasmid sequence has also been observed in Tn7 transposons (210). Interestingly, Tn6167, the class III transposon containing a *tnt* transposition module in pS21-2, is also in the same orientation. This suggests that the *Acinetobacter* class III transposons may also target the lagging strand.

CHAPTER

9

RepAci1 and RepAci2 plasmids

9.1 Introduction

A number of small plasmids that are cryptic or contain resistance genes have been identified in *Acinetobacter* and have been classified using the sequence of their replication initiation proteins (Rep proteins) and assigned RepAci numbers (21). Plasmids with a Rep protein matching RepAci1 such as pACICU1 (129) and pAB0057 (3) or a RepAci2 type such as pABVA01 (21), are fairly common in sequenced isolates (Table 9.1). The RepAci1 and RepAci2 proteins are 85% aa identical. The plasmid pABVA01 contains the carbapenem resistance gene *oxa24* located between two inversely-oriented XerC-XerD binding sites (46). XerC-XerD binding sites have also been detected in plasmids with RepAci1 (46, 220). In genomes of *A. baumannii* from Australia a cryptic RepAci1 plasmid pAb-G7-1, is prevalent in GC1 isolates (Dr M. Hamidian, personal communication) (Fig 9.1). The plasmid pAb-G7-1 is almost identical to pAB0057, except pAB0057 is missing two bases and contains a SNP. However, the missing bases were determined to be a sequencing error (Dr M. Hamidian, personal communication) and during the writing of this chapter a corrected sequence was released under GenBank accession number CP001183. In GC2 isolates from Australia a cryptic RepAci2 plasmid, named pD72-1, is most common (193) (Fig 9.1).

Table 9.1 RepAci1 and RepAci2 plasmids

Rep type	Plasmid	Size (bp)	Resistance genes	Mobilisation genes	GenBank Accession number
RepAci1	pACICU1 ¹	28,279	<i>oxa58</i>	-	CP000864
	pAB0057	8,729	-	-	CP001183
	pAB-G7-1	8,731	-	-	KJ586856
RepAci2	pABVA01	8,963	<i>oxa24</i>	-	FM210331
	pD72-1	10,967	-	<i>mobA, mobC</i>	KM051986

¹Contains two replication initiation genes and is likely a fusion of two plasmids.

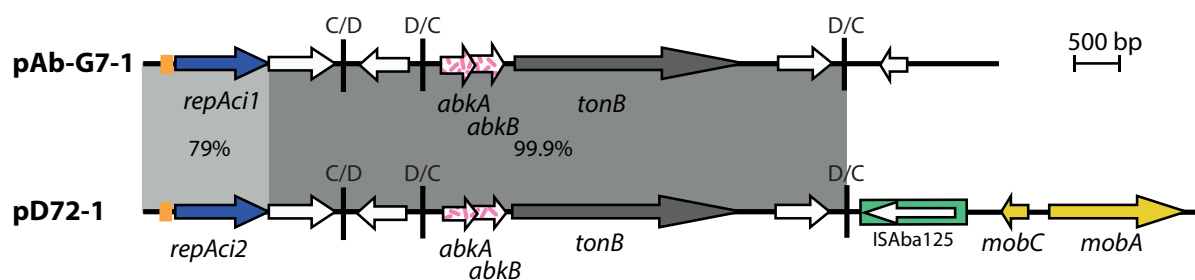


Figure 9.1 pAb-G7-1 and pD72-1 comparison. Arrows indicate the extent and orientation of genes and open reading frames. Genes encoding putative toxin/antitoxin modules are pink patterned and *rep* genes are shown in blue. The *tonB* gene is grey and genes showing homology to mobilisation proteins are yellow. The green box represents ISAbal25 and the internal arrow shows the transposase orientation. The orange box indicates the iteron sequence. Vertical lines indicate XerC-XerD binding sites and the orientation of these sites are shown above. The extents of regions with significant DNA identities are shown in grey.

Genes involved in plasmids mobilisation (*mob* genes) have been detected on several *Acinetobacter* plasmids, including pD72-1 (192) (Fig 9.1), the *repAci9* plasmid pMAC (56) and pRAY* (100) which does not contain a recognisable *rep* gene. Most of these have not been experimentally characterised, but the mobilisation region of pMAC has been shown to be functional (56). The plasmid pMAC encodes a *mobA/L* gene that produces a protein related to the MobA protein of RSF1010 and pSC101 (56). The region encoding this gene and an additional region thought to contain the *oriT* (origin of transfer) sequence was cloned into pUC18 and shown to be mobilised by pRK2073 (a derivative of RK2, an IncP plasmid) in a tri-parental mating at very low frequency (56). However, IncP plasmids are not normally found in *Acinetobacter* isolates.

The aim of the work presented in this chapter was to determine if any small RepAci1 and RepAci2 plasmids are present in the 20 GC2 isolates from Singapore General Hospital and if so to characterise their structure. Plasmid incompatibility and mobility were examined using RepAci1 and RepAci2 plasmids containing resistance genes.

9.2 Detection of RepAci1 and RepAci2 plasmids

To detect and distinguish between the RepAci1 and RepAci2 cryptic plasmids represented by pAb-G7-1 and pD72-1 respectively, primers were designed in a unique region of each plasmid backbone (Fig 9.2A). For detection of pAb-G7-1, the primers RH2035 and RH2036 were paired to amplify a region 328 bp in size and an amplicon slightly larger than 300 bp was amplified in S22, S23, S25, S27, S28 and S31-S40 (Fig 9.2B), indicating that they contained a plasmid of the pAbG7-1 type.

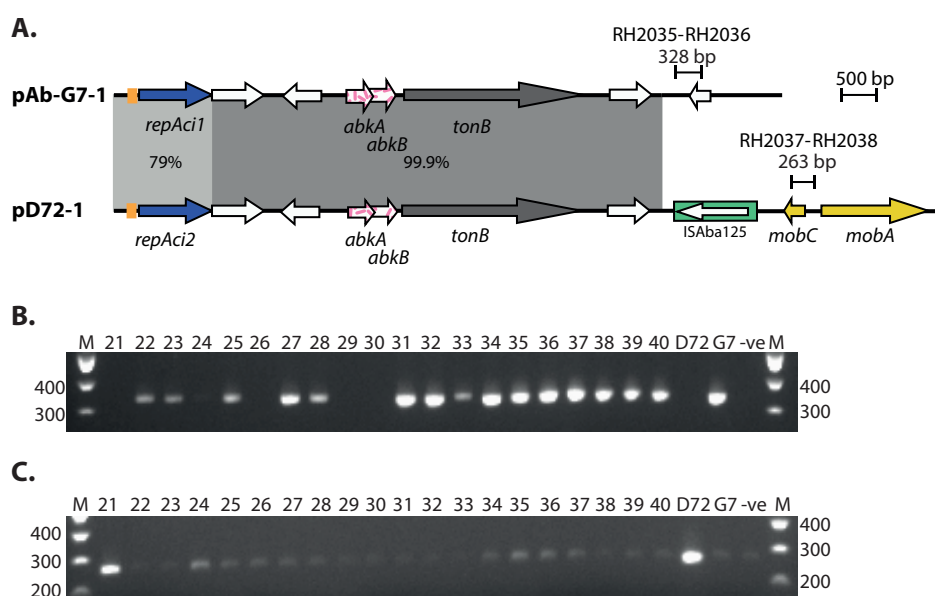


Figure 9.2 Detection of pAb-G7-1 and pD72-1. A. Comparison of pAb-G7-2 and pD72-1. Arrows indicate the extent and orientation of genes and open reading frames. Genes encoding putative toxin/antitoxin modules are pink patterned and *rep* genes are shown in blue. The *tonB* gene is grey and genes showing homology to mobilisation proteins are yellow. The green box represents ISAbal25 and the internal arrow is the transposase gene. The orange box indicates the iteron sequence. Vertical lines indicate XerC-XerD binding sites and the orientation of these sites are shown above. Primers used to differentiate the backbone types are shown by the bars above, along with PCR amplicon size. The extent of regions with significant DNA identities are shown in grey. B. Screening for a 328 bp region of the backbone of pAb-G7-2 using primers RH2035 and RH2036. C. Screening for backbone of pD72-1. RH2037 and RH2038 were paired to detect an amplicon 263 bp in size. In B. and C. the sizes of the molecular markers on the left and right are in bp and isolates are shown above their lanes.

The primers RH2037 and RH2038 were designed to amplify a 263 bp region of pD72-1, including part of the *mobC* gene. When these primers were used in combination, S21 was the only isolate to produce a strong amplicon that was around 250 bp in size, consistent with that

of pD72-1 (Fig 9.2C). From these screening results, S24, S26, S29 and S30 did not contain a plasmid containing a similar backbone to that of pAb-G7-2 or pD72-1. The absence of a RepAci1 or RepAci2 plasmid in these four isolates was later confirmed when the draft genome sequence became available.

9.3 Plasmid assembly and analysis

9.3.1 pAbG7-1 type (RepAci1)

The entire backbone of pAb-G7-1 was retrieved in a single contig from the assemblies of S27, S34, S38 and S39 (Table 9.2). As the assembly program is unable to complete a circular sequence, the assembly of each plasmid broke randomly and at a different point in the backbone for each plasmid. The ends of the contigs overlapped by approximately 80 bp in each isolate. The overlapping sequence was removed and the contigs were circularised *in silico* and reopened at the same point to facilitate comparisons. These four plasmids, named pS27-1, pS34-1, pS38-1 and pS39-1, were identical to each other and were 8,731 bp in size. They all had one SNP relative to pAb-G7-1 sequence (position 1,399, C→A, silent mutation) in the orf adjacent to *repAci1* (Fig 9.2A). This is the same SNP that is present in pAB0057.

Once the MiSeq sequences became available, contigs containing pAb-G7-1 sequence were also retrieved (Table 9.2) and all but the plasmid in S32 were assembled using the same method as above. The plasmids present in S22, S23, S25, S28, S31 and S33-S40 (named pS22-1 to pS40-1) were identical to those in the initial four isolates, each containing the pAB0057 SNP. The edges of contig 40 in S32 (highlighted in orange in Table 9.2) did not overlap with each other, instead ending with fragments of the same end of ISAbal.

Table 9.2 Contigs matching pAb-G7-1

Isolate ¹	Contig	Length	Kmer coverage	Final length (bp)
S22	32	8,858	1297.3200	8,731
S23	32	8,858	818.4330	8,731
S25	34	8,858	1057.7100	8,731
S27	1	8,811	139.4231	8,731
S28	37	8,858	418.8120	8,731
S31	33	8,858	510.0670	8,731
S32	40	9,080	541.0410	>8,731
S33	36	8,858	845.5050	8,731
S34	13	8,811	80.4217	8,731
S35	37	8,858	921.4150	8,731
S36	38	8,858	1180.2400	8,731
S37	37	8,858	1008.0900	8,731
S38	3	8,813	77.8566	8,731
S39	4	8,811	80.8631	8,731
S40	39	8,858	946.5960	8,731

¹Isolates in bold were sequenced by Illumina HiSeq and were assembled using Velvet. The others were sequenced using Illumina MiSeq and were assembled using SPAdes.

9.3.1.1 pS32-1 includes Tn2006

The fragments of ISAbal at the ends of contig 40 of the S32 genome were in inverse orientation to each other, but a 9 bp direct repeat (5' AATTCTAG 3') was adjacent to each fragment of ISAbal. This is consistent with the presence of the *oxa23* containing transposon Tn2006, which from earlier screening is present in S32 (Chapter 3; Section 3.3.3). If Tn2006 was located here, from the sequence of its direct repeat, it would be a novel insertion position for this transposon (194). Furthermore, the kmer coverage of the *oxa23* containing contig is 618.6430, which is approximately equivalent to that of the backbone of pS32-1 (Table 9.2). The insertion in the backbone of pS32-1 directly follows the stop codon of the *tonB* gene, leaving the reading frame intact. Two primers were designed facing outwards at each end of the contig (RH2505 and RH2506) and paired with internal Tn2006 primers (Figure 9.3A), to determine the orientation of this transposon in pS32-1.

(location of primers indicated in the top line in Fig 9.4C). When the two primers amplify over the single ISAbal25 in pD72-1 the product is 1,251 bp, however, the amplicon of pS21-1 was approximately 2 kb larger than this (Fig 9.4A), indicating that additional sequence was present.

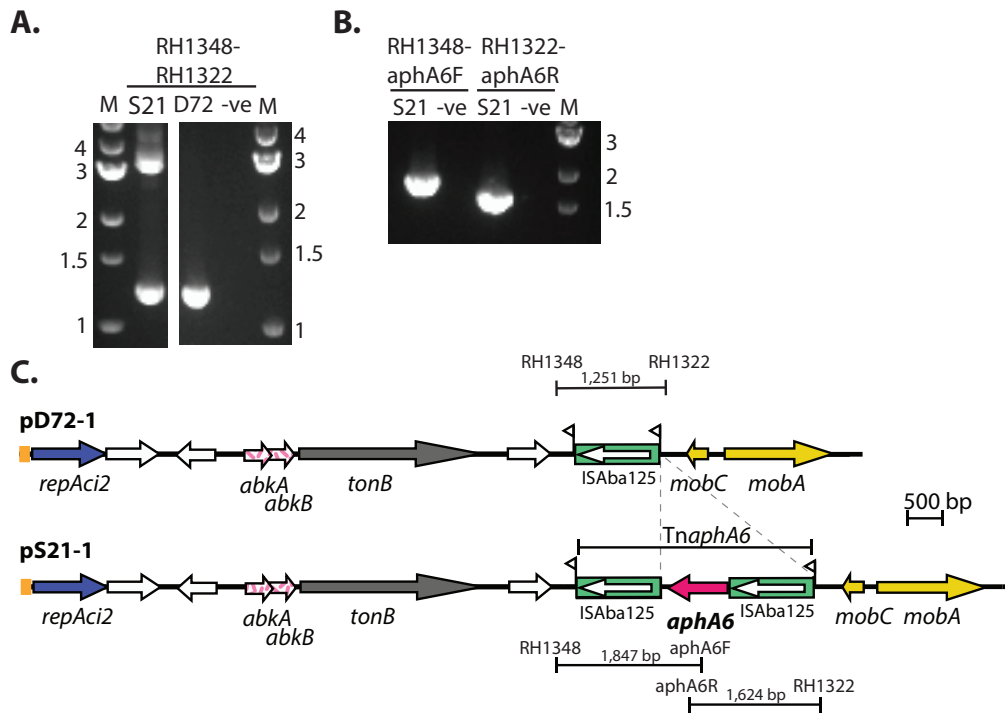


Figure 9.4 pS21-1 contains *TnaphA6*. A. pS21-1 has sequence additional to ISAbal25. The primer combination of RH1348 and RH1322 produce an amplicon size of 1,251 bp when only a single ISAbal25 is present. If *TnaphA6* is present the amplicon is 3,236 bp. B. Confirmation of *TnaphA6* in pS21-1. RH1348 and aphA6F and RH1322 and aphA6R were paired producing predicted amplicons of 1,847 bp and 1,624 bp respectively. In A. and B. the sizes in kb of the molecular marker are shown to the left and right. The isolates are shown above their lane. C. Comparison of pD72-1 and pS21-1. Arrows indicate the extent and orientation of genes and open reading frames. Genes encoding putative toxin/antitoxin modules are pink patterned and *rep* genes are shown in blue. The *tonB* and *aphA6* genes are coloured grey and magenta. The green boxes represent ISAbal25 and the internal arrow indicates the transposase orientation. An orange box indicates the iteron sequence and matching flags indicate TSD. Primers used to span ISAbal25 or *TnaphA6* are shown by the bars above along with predicted amplicon size. Vertical lines indicate XerC-XerD binding sites and the orientation of these sites are indicated above.

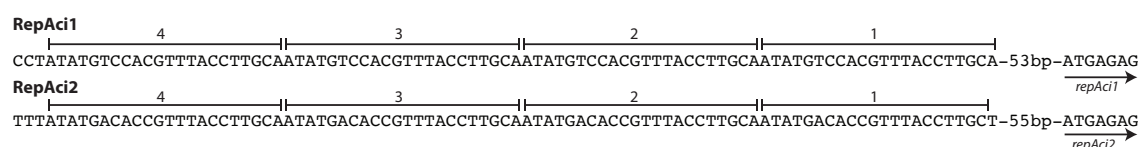
Earlier screening (Chapter 3; Section 3.3.4) showed that S21 contains the transposon *TnaphA6*, which is 3,072 bp in length and is bounded by directly orientated copies of ISAbal25. The kmer coverage of the contig containing *aphA6* in S21 is 90.3650, which is equivalent to that of the contig of the plasmid backbone, indicating it is likely to be present here. Each of the backbone primers was paired with one in *TnaphA6* (Fig 9.4B and C), yielding products around

2 and 1.5 kb, consistent with the predicted amplicon sizes of 1,824 bp and 1,624 bp respectively, indicating that *TnaphA6* was present. Hence, pS21-1 is 12,643 bp in size and the only difference between pS21-1 and pD72-1 is the presence of *TnaphA6* in the same location as ISAbal25 in pD72-1 (Fig 9.4C). Like pD72-1, pS21-2 also contains the potential mobilisation genes *mobA* and *mobC*. The MobA protein of pS21-1 shares 32% aa identity to that of the MobA of RSF1010 and its MobC protein is 40% identical to MobC of RSF1010.

9.4 Plasmid compatibility

The *repAci1* and *repAci2* genes are 80% identical (772/961) and the proteins they encode are 85% (270/316 aa) identical to one another. Approximately 50 bp upstream of the *repAci1* and *repAci2* genes are iteron sequences. Iteron sequences are the binding sites for the plasmid-encoded Rep proteins and have control properties (49). There are four identical repeats of 22 bp in RepAci1 plasmids (top line in Fig 9.5A), while RepAci2 plasmids have three identical repeats of the 22 bp iteron sequence and one of 21 bp, where the final A is missing (bottom line in Fig 9.5A). The almost identical iteron sequence (19 out of 22 bp) shared by the RepAci1 and RepAci2 plasmids (comparison shown in Fig 9.5B), and the high identity of their replication initiation proteins may result in the incompatibility of these two plasmids and to date these two Rep types have not been observed in the same isolate. The availability of antibiotic resistant plasmids pS21-1 and pS32-1 allowed the compatibility of RepAci1 and RepAci2 plasmids to be examined for the first time.

A.



B.

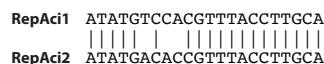


Figure 9.5 RepAci1 and RepAci2 iteron sequences. A. Structure and sequence of the iteron region in RepAci1 and RepAci2 plasmids. The bars above indicate the iteron sequences and the arrows below represent the start of the *repAci1* or *repAci2* gene. B. Comparison of iteron sequence.

9.4.1 pS21-1 (RepAci1) and pS32-1 (RepAci2) are compatible

pS32-1 was transformed into AB307_0294 and its presence was confirmed by PCR screening (Table 9.3). Then AB307_0294 containing pS32-1 was transformed with pS21-1. Transformants recovered were both ticarcillin and kanamycin resistant, and the boundaries of Tn2006 in pS32-1 and TnaphA6 in pS21-1 were detected (Table 9.3).

Table 9.3 Transformation of AB307_0294^{Rif} with pS32-1 and pS21-1

Plasmid/s transformed	Phenotype	pS32-1 Tn2006 location		pS21-1 TnaphA6 location	
		RH2505- RH743	oxa23R- RH2506	RH1348- aphA6F	RH1322- aphA6R
N/A	Rif, Nx	-	-	-	-
pS32-1	TIC, Rif, Nx	+	+	-	-
pS32-1, pS21-1	TIC, Km, Rif, Nx	+	+	+	+

Kanamycin, Km; Nalidixic acid, Nx; Rifampicin, Rif; Ticarcillin, TIC.

Plasmid DNA of an AB307_0294 (pS21-1 and pS32-1) transformant was extracted and visualised on a TAE gel (Fig 9.6A). There was only one band present in the lane of the transformant (Fig 9.6A), but as there is less than 1 kb difference between the size of pS21-1 (12,643 bp) and pS32-1 (13,545 bp) it is possible that they would not separate on a standard

TAE gel. The differences in the brightness of the plasmid bands indicates that pS32-1 may have a higher copy number than pS21-1 (Fig 9.6A).

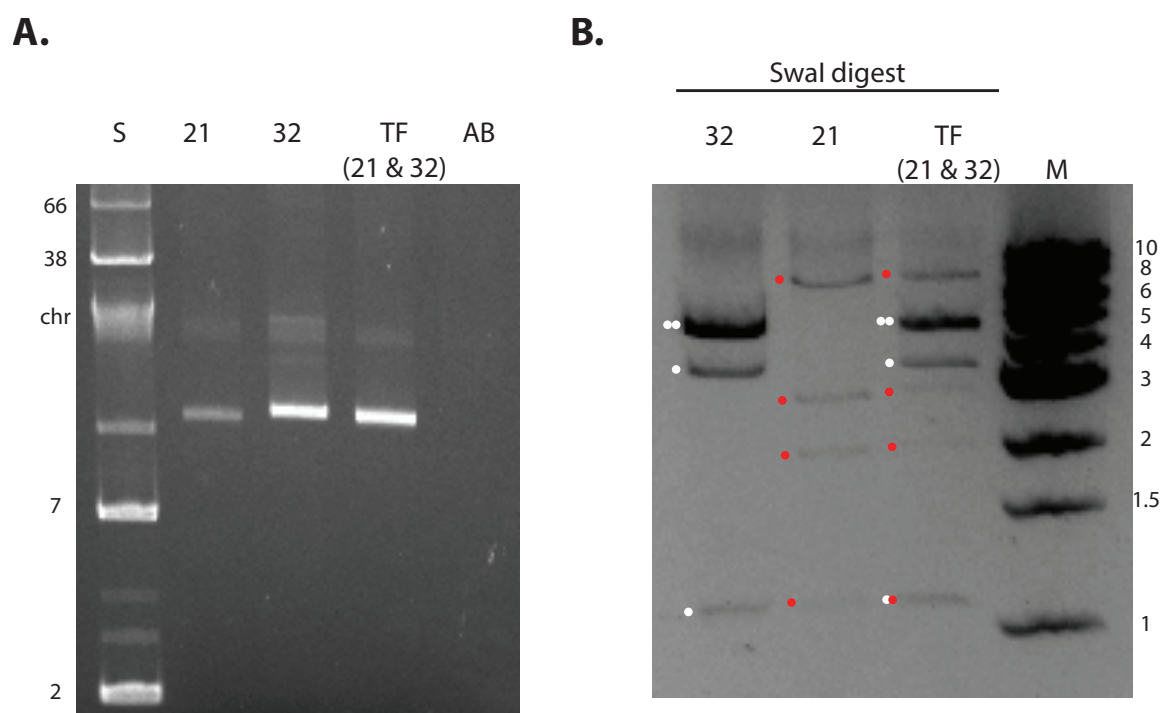


Figure 9.6 AB307_0294 containing pS21-1 and pS32-1. A. Plasmid DNA of a pS21-1 containing AB307_0294 (21), a pS32-1 containing AB307_0294 (32), a pS21-1 and pS32-1 containing AB307_0294 transformant (TF 21 & 32) and empty recipient AB307_0294 (AB). These were run alongside plasmid DNA from 39R861+ (S) in a 0.7% agarose tris-acetate-EDTA gel. Sizes of plasmids in 39R861+ and the position of chromosomal DNA (chr) are indicated to the left. B. Plasmid DNA of a pS21-1 containing AB307_0294 (21), a pS32-1 containing AB307_0294 (32) and a pS21-1 and pS32-1 containing AB307_0294 transformant (TF 21 & 32) digested with *Swa*I. The sizes in kb of the molecular marker are shown to the right. The red and white dots indicate bands visible in the original gel, white for pS21-1 fragments and red for pS32-1 fragments, respectively.

To confirm the presence of both plasmids, plasmid DNA of the AB307_0294 (pS21-1, pS32-1) transformant as well as plasmid DNA of each individual plasmid was digested with *Swa*I. The digests of the pS21-1 and pS32-1 matched the *in silico* digests of each plasmid (Table 9.4, Fig 9.6B). The digest pattern of the transformant matched a combination of the individual plasmids confirming that both pS21-1 and pS32-1 are present in the transformant (Fig 9.6B).

Table 9.4 *In silico* SmaI digest

pS32-1	pS21-1
4,663	6,830
4,483	2,703
3,298	1,961
1,101	1,101
	24
	24

AB307_0294 (pS21-1, pS32-1) conferring resistance to ticarcillin and kanamycin, was grown without antibiotic selection over five cycles of serial subculture. Following each cycle of overnight culture (approximately 22 generations) the total number of cells, as well as the number of cells growing in kanamycin, ticarcillin, and both ticarcillin and kanamycin were determined. The results from a single experiment showed that over the 5 passages, approximately 110 generations, the proportion of antibiotic resistant cells did not decrease (Table 9.5). Fifty of the colonies selected on the L-agar at Day 5 were then patched onto L-agar supplemented with kanamycin or ticarcillin. All were still resistant to both antibiotics and the junctions of *TnaphA6* in pS21-1 and *Tn2006* in pS32-1 were screened and confirmed in five of these (screening PCRs shown in Table 9.4). Hence, confirming that pS21-1 and pS32-1 are compatible with one another over 110 generations.

Table 9.5 Cell counts over 5 serial passages (110 generations)

Overnight culture	L-agar (cfu/mL)	Km (cfu/mL)	TIC (cfu/mL)	KmTIC (cfu/mL)
1	4.7x10 ⁸	7.9x10 ⁸	5.5x10 ⁸	6.3x10 ⁸
2	6.1x10 ⁸	4.3x10 ⁸	6.6x10 ⁸	5.8x10 ⁸
3	7.3x10 ⁸	8.0x10 ⁸	8.0x10 ⁸	7.2x10 ⁸
4	10.2x10 ⁸	10.3x10 ⁸	11.2x10 ⁸	11.1x10 ⁸
5	11.8x10 ⁸	14.0x10 ⁸	12.7x10 ⁸	14.7x10 ⁸

Kanamycin, Km; Ticarcillin, TIC.

9.5 Plasmid mobilisation

9.5.1 Comparison of mobilisation genes

The sizes of the *mob* genes and hence the Mob proteins of RSF1010, pS21-1, pMAC and pRAY are quite varied (Table 9.6). Despite both having some similarity to RSF1010, the *mob* genes of pS21-1 and pMAC showed no sequence identity to one another. When the MobA protein of each plasmid was compared, they matched with 29% identity over 46% of the protein. This match was concentrated in the N-terminus region. Hence, the MobA proteins encoded by each of these plasmids are quite different and may be mobilised by different conjugative plasmids. The MobA protein of pS21-1 matched with 30% identity over only 5% of the MobA protein of pRAY. The MobC proteins of pS21-1 and pRAY are also quite different with a match of 83% identity over only 5% of the protein.

Table 9.6 Mobilisation gene and protein lengths

	<i>mobA</i> gene (bp)	MobA (aa)	<i>mobC</i> gene (bp)	MobC (aa)
RSF1010	2130	710	285	95
pS21-1	1437	479	306	102
pMAC	1170	390	765	255
pRAY	1815	605	408	136

A number of plasmids containing *mobA* and *mobC* genes closely related to those in pS21-1 were identified following a search of the nucleotide database (Table 9.7). However, no experimental work has been performed on plasmids containing *mob* genes of this type to determine if they are functional and which conjugative plasmids are capable of mobilising them.

Table 9.7 Mob proteins related to those of pS21-1

Plasmid	<i>Acinetobacter</i> species	Size (bp)	Rep	Identity to pS21-1		GenBank Acc. No.	Ref
				MobA	MobC		
pD72-1, pC2 ¹	<i>baumannii</i>	10,967	RepAci2	100% [478/478]	100% [101/101]	KU549174	(193)
p6AsACE	<i>schindleri</i>	7,943	RepAci3 ³	91.42% [437/478]	94.06% [95/101]	CP015621	
pM131-5	<i>sp.</i>	7,641	RepAci3 ³	89.92% [437/486]	95.05% [96/101]	JX101644	
CIP70.10 plasmid II	<i>baumannii</i>	7,742	RepAci5	91.21% [436/478]	95.05% [96/101]	LN865144	
pALWVS1.1	<i>lwoffii</i>	134,767 ²	New Rep ⁴	91.42% [437/478]	98.02% [99/101]	KX426232	(175)
p5AsACE	<i>schindleri</i>	10,085	New Rep ⁵	81.59% [390/478]	95.05% [96/101]	CP015620	

¹Is a 2.5 kb deletion variant of pD72-1.

²Is a combination of two plasmids.

³The *rep* genes of these two plasmids are 99.9% identical.

⁴75.1% DNA identity to *repAci1*.

⁵82% identical to the *rep* of pABa3207a (GenBank accession number CP015365).

9.5.2 pS21-1 is not mobilised by a RepAci6 plasmid

The co-residing RepAci6 plasmid in pS21 was found to be non-conjugative (Chapter 8; Section 8.5.1). A different donor strain was constructed by first transforming A25, which carries no other plasmids (Chapter 2; Table 2.5), with pS21-1 (RepAci1, containing the *aphA6* gene and conferring kanamycin resistance), then conjugating in the self-transmissible RepAci6 plasmid pS27-2 (containing *oxa23*, conferring ticarcillin and clavulanic acid resistance) (Chapter 8; Section 8.4.1). The presence of both plasmids in A25 was confirmed by screening for the junctions of *TnaphA6* in pS21-1 and the presence of the *oxa23* gene and a region of the RepAci6 backbone for pS27-2, as summarised in Table 9.8.

Table 9.8 Construction of A25 (pS21-1 and pS27-2)

Plasmid/s	Phenotype	pS21-1 specific PCRs		pS27-2 specific PCRs	
		RH1348- aphA6F	RH1322- aphA6R	RH1348- aphA6F	RH1322- aphA6R
pS21-1	Km	+	+	NA	NA
pS21-1, pS27-2	Km, TIC/CLAV	+	+	+	+

Kanamycin, Km; Ticarcillin/clavulanic acid, TIC/CLAV.

A24 (pS27-2, pS21-1) was mated with AB307_0294^{Nx} (nalidixic acid resistant) and all possible outcomes were tested (Table 9.9). This included selection for just pS27-2 (ticarcillin/clavulanic acid and nalidixic acid resistance), just pS21-1 (kanamycin and nalidixic acid resistance) or both pS27-2 and pS21-1 (kanamycin, ticarcillin/clavulanic acid and nalidixic acid resistance) moving.

Table 9.9 A25 (pS21-1, pS27-2) and AB307_2094^{Nx} mating

Plasmid/s	Selection	Observed transconjugants (transconjugants/donor) ¹
pS27-2	TIC/CLAVNx	6.37x10 ⁻⁴
pS21-1	KmNx	<3.7x10 ⁻⁸
pS21-1, pS27-2	KmTIC/CLAVNx	<3.7x10 ⁻⁸

Kanamycin, Km; Nalidixic acid, Nx; Ticarcillin/clavulanic acid, TIC/CLAV.

¹Transfer frequency is only calculated from one experiment.

Transconjugants were only present on the ticarcillin/clavulanic acid and nalidixic acid selection plates (Table 9.9) and 50 of these were screened for resistance to kanamycin but none harboured this additional resistance phenotype. Therefore, only the conjugative pS27-2 plasmid was moving and if pS21-1 was mobilised it was below the limit of detection (Table 9.9). Hence, it is likely that pS21-1 cannot be mobilised by pS27-2, and probably cannot be transferred by other RepAci6 plasmids.

9.5.3 Mobilisation of pS32-2 by RepAci6 plasmids

S32 contains two plasmids, firstly a RepAci1 plasmid pS32-1 carrying the *oxa23* gene and no mobilisation genes, and second, pS32-2, a potentially conjugative RepAci6 plasmid (Chapter 8; Section 8.4.2). However, pS32-2 does not contain a selectable marker preventing tracking of its movement. S32 was mated with AB307_0294^{Rif} (rifampicin resistant) and potential transconjugants were detected using selection for movement of pS32-1 (ticarcillin and

rifampicin resistance). From one experiment, pS32-1 transferred at a frequency of 4.62×10^{-6} transconjugants/donor. A selection of transconjugants (numbering five) were screened by PCR for pS32-1 and pS32-2 (detail shown in Table 9.10) and both plasmids were detected.

Table 9.10 Transconjugant screening

Phenotype	pS32-1 specific		pS32-2	Plasmid/s present
	PCRs		specific PCR	
	oxa23F- oxa23R	RH2506- oxa23R	RH1397- RH1503	
TICRif	+	+	+	pS32-1, pS32-2

Rifampicin, Rif; Ticarcillin, TIC.

As the donor S32 is resistant to kanamycin, 100 potential transconjugants were additionally screened for resistance to this antibiotic to confirm that they were not rifampicin resistant mutants of the donor. All but one was susceptible to kanamycin indicating that they were transconjugants. To confirm that both of the plasmids were intact and no recombination had occurred to assist in the transfer of pS32-1, plasmid DNA was isolated from the donor, an empty recipient and a transconjugant. In both the donor and the transconjugant, bands at approximately the right size for pS32-1 (13,545 bp) and pS32-2 (70,833 bp) were detected (Figure 9.7). Hence, the plasmids had not fused. Plasmid DNA was digested with *Swa*I and compared to fragment sizes predicted by the *in silico* digest of the plasmid sequences (Figure 9.8A and B). The digest of the donor and the transconjugant matched each other and the predicted product sizes. However, the 3 bands between 8-9 kb and the 3 bands between 4-5 kb did not separate, and the bands below 4 kb from pS32-2 are too faint to be detected (Figure 9.8A and B). Therefore, pS32-1 is able to be mobilised by pS32-2 and likely by other RepAci6 plasmids.

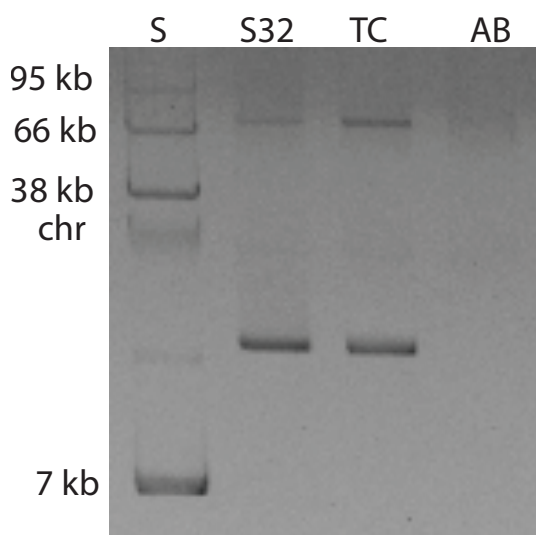


Figure 9.7 Plasmid visualisation in transconjugant. Plasmid DNA from S32 and from a pS32-1 containing AB307_0294 transconjugant (TC) and the empty recipient (AB) AB307_0294. These were run alongside plasmid DNA from 39R861+ (S) in a 0.7% agarose tris-acetate-EDTA gel. Sizes of plasmids in 39R861+ and the position of chromosomal DNA (chr) are indicated to the left.

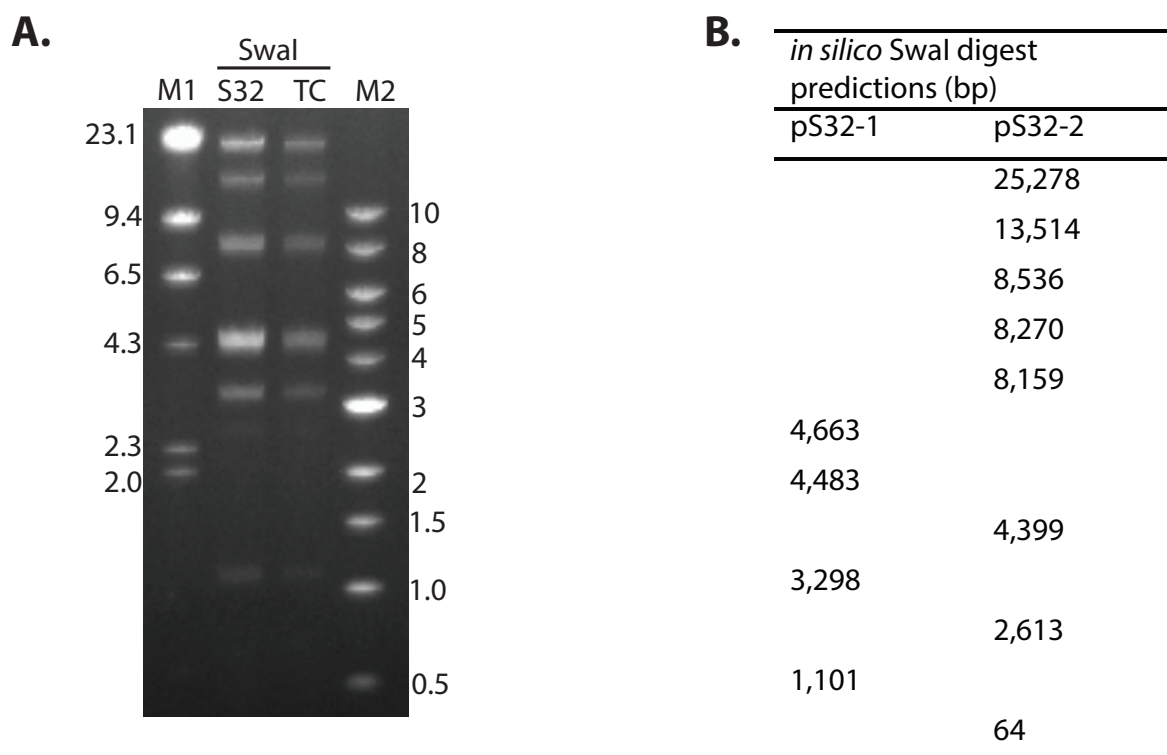


Figure 9.8 Restriction digest of transconjugant. A. Plasmid DNA of S32 and of the AB307_0294 transconjugant (TC) containing pS32-1 digested with SwaI. HindIII λ ladder (M1) and 1 kb DNA Ladder (M2) was used as markers and band sizes in kb are shown to the left and right. B. Predicted SwaI digestion pattern of pS32-1 and pS32-2.

Mobilisation frequencies could not be calculated using S32 as a donor as the frequency of transfer of the cryptic pS32-2 could not be determined due to the absence of a selectable

marker. A different donor of ATCC 17978^{Rif}-A (sulphonamide susceptible and rifampicin resistant) containing pS32-1 (conferring resistance to ticarcillin) and the conjugative *repAci6* plasmid pAbG7-2 (containing *TnaphA6* and conferring resistance to kanamycin) was constructed. First, G7 was mated with ATCC 17978^{Rif}-A and kanamycin resistant ATCC 17978^{Rif}-A were recovered and the of boundaries of *TnaphA6* in pAb-G7-2 were detected (Table 9.11). pS32-1 was then transformed into ATCC 17978^{Rif}-A (pAb-G7-2), and the presence of the transposon boundaries of *Tn2006* and *TnaphA6* in each plasmid was confirmed by PCR screening (Table 9.11).

Table 9.11 Construction of ATCC 17978^{Rif}-A (pS32-1, pAb-G7-2)

Plasmid/s	Phenotype	pS32-1 <i>Tn2006</i> location		pAb-G7-2 <i>TnaphA6</i> location	
		RH2505- RH743	oxa23R- RH2506	RH1501- aphA6R	RH1502- aphA6F
N/A	RfNx	-	-	-	-
pAb-G7-2	RfNxKm	-	-	+	+
pS32-1, pAb-G7-2	RfNxKmTIC	+	+	+	+

Kanamycin, Km; Nalidixic acid, Nx; Rifampicin, Rif; Ticarcillin, TIC.

ATCC 17978^{Rif}-A containing pS32-1 and pAb-G7-2 was mated with AB307_0294^{Nx} (nalidixic acid resistant) and all outcomes of the mating were selected for (Table 9.12). The conjugative plasmid pAb-G7-2 transferred at frequencies ranging from 2.34×10^{-3} to 4.06×10^{-4} transconjugants per donor, though the movement of pS32-1 was much less frequent (Table 9.12). In fact, following the removal of nalidixic-acid resistant mutants of the donor, over the three independent experiments, 82 ticarcillin and nalidixic acid resistant colonies were recovered. Two nalidixic acid and ticarcillin resistant transconjugants from each experiment were screened using PCRs shown in Table 9.11 and the junctions of both transposons were present. A small number of the nalidixic acid and ticarcillin resistant transconjugants (3/82) were kanamycin susceptible and while the junctions of *Tn2006* in pS32-1 were present, neither

the junctions of *TnaphA6* in pAb-G7-2 or the *repAci6* gene was present in these transconjugants.

Table 9.12 Frequency of pS32-1 mobilisation

Plasmid/s selected	Phenotype	Observed transconjugants (transconjugants/donor)			
		Round 1	Round 2	Round 3	Average
pAb-G7-2	KmNx	2.34×10^{-3}	4.06×10^{-4}	6.16×10^{-4}	1.12×10^{-3}
pS32-1	TICNx	2.90×10^{-8}	4.51×10^{-8}	1.45×10^{-8}	2.95×10^{-8}
pS32-1, pAb-G7-2	TICKmNx	3.18×10^{-8}	1.31×10^{-8}	9.68×10^{-9}	1.12×10^{-8}
% mobilisation of pS32-1 ¹		0.0012	0.0111	0.0025	0.0049

Kanamycin, Km; Nalidixic acid, Nx; Ticarcillin, TIC.

¹Frequency of pS32-1 (TICNx) divided by frequency of pAb-G7-2 (KmNx).

Each independent experiment provided a low, but consistent, transfer frequency of pS32-1 averaging at 2.95×10^{-8} transconjugants per donor (Table 9.12). The plasmid pS32-1 is able to be mobilised by *repAci6* plasmids, at a proportion averaging 0.0049% of the transfer of pAb-G7-2 (Table 9.12).

9.6 Putative *oriT* region of pS32-1

As pS32-1 does not contain genes matching any known relaxases (*mob* genes) (see Figure 9.1) it could be moving by “*relaxase in trans*” mobilisation. If this is the case both pS32-1 and pS32-2, and hence other RepAci6 plasmids, must have a closely-related *oriT* (origin of transfer) sequence. The *oriT* of RepAci6 plasmids has not been located nor has one been experimentally identified for any other *Acinetobacter* plasmid. The mobilisation genes of RepAci6 plasmids have previously been called *trwB* (an accessory protein) and *trwC* (the putative relaxase) the products of these genes are 27% and 30% aa identical to those of R388, an exemplar of the IncW plasmids. The nick region of the *oriT* of R388 is an F type (149).

The nucleotide sequence of pS32-1 was compared to the sequence of pS32-2 using pairwise BLAST to find any regions of identity. There was a strong match of 29/32 bp, that in pS32-2 was located upstream of the *trwB* gene (Figure 9.9A). Just downstream of the *trwB* gene in pS32-2 is the *trwC* gene which encodes the predicted relaxase of the RepAci6 plasmid. In pS32-1, this match was located in the intergenic region between the last two open reading frames (Figure 9.9B). The 32 bp matching nucleotides of pS32-1 and pS32-2 as well as some additional surrounding sequence were compared to the *oriT* sequence of R388 (GenBank accession number X51505) (Figure 9.9C) and 7 of the 8 conserved bases of the F-type nick site (149) were present in pS32-1 and pS32-2 (green in Figure 9.9C).

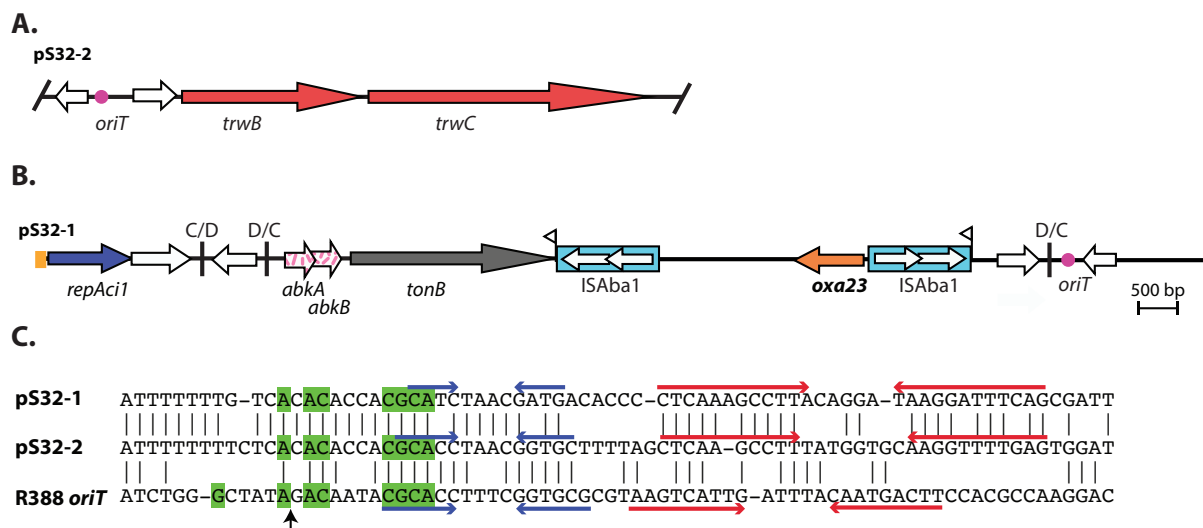


Figure 9.9 Putative *oriT* region of pS32-1 and pS32-2. A. *trwB* and *trwC* region of pS32-2. A pink circle indicates the location of the putative *oriT* region. Red arrows indicate direction and extent of the mobilisation genes. B. Linear structure of pS32-1. Genes encoding putative toxin/antitoxin modules are pink patterned and the *rep* genes is shown in blue. The *tonB* and *oxa23* genes are coloured grey and orange. The light blue boxes represent ISAbal1 and the internal arrow indicates the transposase orientation. The orange box indicates the iteron sequence and flags indicate matching TSD. Vertical lines indicate XerC-XerD binding sites and the orientation of these sites are indicated above. C. Comparison of the putative *oriT* region of pS32-1 and pS32-2, compared to the *oriT* region of R388. Bases highlighted in green are conserved in F type nick sites and a vertical arrow indicates the *nic*-site. Coloured horizontal arrows indicate matching inverted repeats. The *oriT* sequence of R388 is taken from GenBank accession number X51505.

A common feature of *oriT* sequences is the presence of direct and/or inverted repeats which produce a secondary structure that function as recognition sites for DNA-binding proteins

(149). In the three plasmids compared above there are two sets of inverted repeats located in approximately the same region of DNA (indicated by horizontal coloured arrows in Figure 9.9C). There is a short-inverted repeat (4-6 bp) located seven nucleotides away from the *nic* site (vertical arrow in Figure 9.9C). Further away from the *nic* site is the second set of inverted repeats (Blue arrows in Figure 9.9C) and in R388 this is a perfect 9 bp inverted repeat, while in pS32-1 and pS32-2 they are imperfect and are 12 and 11 nucleotides, respectively.

The regions of pS32-1 and pS32-2 identified above, appear to be good candidates for the nick site and surrounding *oriT* region for of these plasmids. It is likely that the variation of the *oriT* site in each of these plasmids accounts for the low co-transfer frequency. Further work such as cloning the putative *oriT* region is still required to experimentally prove this assignment and to further characterise these regions.

9.9 Discussion

The small RepAci1 and RepAci2 plasmids analysed here are closely related to plasmids detected in isolates from Australia, and in the case of RepAci1 plasmids, from around the world. Unlike GC2 isolates in Australia, the most common small plasmid type was the RepAci1 pAb-G7-1 type (also known as the pAB0057 type). Only S21 contained a RepAci2 plasmid that was closely related to the RepAci2 plasmid pD72-1 found in Australian isolates (193). The distribution of these small plasmids also correlated with the features of the resistance islands of the isolates. A RepAci1 plasmid was found in each of the 15 isolates that also contained AbGRI3. The isolates that do not have AbGRI3 and contained AbGRI2-1, either had a RepAci2 plasmid in the case of S21, or do not include a version of either of these small plasmids.

The carriage of *Tn2006* and *TnaphA6* on pS32-1 and pS21-1 respectively, is the first time these transposons have been identified in RepAci1 or RepAci2 plasmids. Previously, *Tn2006* had been seen either in the chromosome or in RepAci6 plasmids (194) and *TnaphA6* has also been detected in RepAci6 plasmids (97, 192, 277). Resistant versions of the RepAci1 and RepAci2 plasmids enabled incompatibility studies to be performed and showed for the first time that these plasmids were compatible.

The plasmid pS21-1, a selectable RepAci2 plasmid containing mobilisation genes, was not mobilised by the conjugative RepAci6 plasmid, pS27-2. Perhaps other conjugative plasmids found in *Acinetobacter*, such as those of the pA297-3 type (87) or the conjugative plasmid types containing the NDM-1 gene (123, 284) would be able to mobilise pS21-1 and this should be tested. The only difference between pS21-1 and pD72-1 is the presence of *TnaphA6* (ISAbal25-*aphA6*-ISAbal25) where only a single copy of ISAbal25 is present in pD72-1 and these insertions are in exactly the same location. Hence, either the pS21-1 version came first and recombination removed a copy of ISAbal25 and the region containing *aphA6* to form pD72-1, or pD72-1 came first and the region containing *aphA6* and ISAbal25 inserted adjacent to the copy of ISAbal25 already present in pD72-1 to form pS21-1. The first scenario is plausible but unusual as this involves the loss of a useful resistance gene. The second scenario is interesting as it suggests that ISAbal25 may function like IS26 and hence, be able to target itself (107).

The plasmid pS32-1, the resistant version of the pAb-G7-1 RepAci1 plasmid, was shown to be mobilised at low frequency by two RepAci6 plasmids, firstly by the co-residing plasmid pS32-2 and secondly by pAb-G7-2. Hence, it is likely that other plasmids resembling pAb-G7-1 are also able to be mobilised by RepAci6 plasmids. A putative *oriT* sequence of the F-type was

identified in pS32-1 and pS32-2, and it appears to be a good candidate due to conservation of bases at the *nic* site and the presence of inverted repeats in the adjacent sequence. However, further work will be required to confirm this. As pS32-1 does not contain any mobilisation genes, this is the first case of mobilisation *in trans* in *Acinetobacter*. This mode of mobilisation where the *oriT* of the small non-conjugative plasmid mimics the *oriT* of the conjugative plasmids was first detected in plasmids from *Staphylococcus* (200). Mobilisation *in trans* has also been detected in Gram-negatives with plasmids from *E.coli* (personal communication, Robert Moran). The detection of *in trans* mobilisation across Gram-negative and Gram-positive species suggests that this form of mobilisation may be quite common and that small plasmids without mobilisation genes are likely to be playing a greater role in horizontal gene transfer than initially thought.

CHAPTER

10

pS30-1

10.1 Introduction

The XerC and XerD recombination system is used by most genera of bacteria to resolve dimers of circular chromosomal DNA into monomers (173). This system involves two tyrosine recombinases XerC and XerD which bind to a specific target site called a *dif* site located in the replication termination region of the chromosome (173). A *dif* site is 28 bp long and contains the 11 bp binding sites for the XerC and XerD recombinases, separated by a 6 bp spacer (173). Some mobile elements exploit the XerC-XerD recombination system of their host. Certain plasmids hijack this system to resolve plasmid dimers such as those containing a *cer* (39, 251) or *psi* (41) site. These sites contain a core *dif*-like site with the XerC and XerD binding site as well as additional sequence for the binding of necessary accessory proteins (173). Some integrative mobile elements do not encode their own recombinase but utilise the Xer recombination system of their host to integrate at *dif* (173). These are known as integrative mobile elements exploiting Xer (IMEXs) and a well-studied example is the CTX ϕ phage which recombines into one of the *dif* sites of its host *Vibrio cholerae* (264).

The *oxa24* carbapenemase gene and its single base pair variant encoding OXA-72, have been found in plasmids from various *Acinetobacter* species. However, unlike other resistance genes such as the aminoglycoside resistance gene *aphA6* or the different carbapenemase resistance gene *oxa23*, the *oxa24* gene is not located in any known transposon. In the plasmid pABVA01, the authors identified conserved inverted repeats of 28 bp homologous to XerC-XerD binding sites (46) (Fig 10.1A, top line) surrounding the *oxa24* gene. These sites contained the 11 bp binding sites for the XerC and XerD recombinase separated by a 6 bp spacer (46). In sequences of related plasmids available at the time, the authors also identified XerC-XerD binding sites in the same position of the backbone relative to *rep*, but they flanked different segments of DNA (46), suggesting that these modules are mobile. Since then, the *oxa24* module has been

located in plasmids of different *rep* types, with different sequence on either side and in one case it is found twice in the same plasmid (46, 85, 172, 220), increasing the likelihood that this module is mobile (Fig 10.1A).

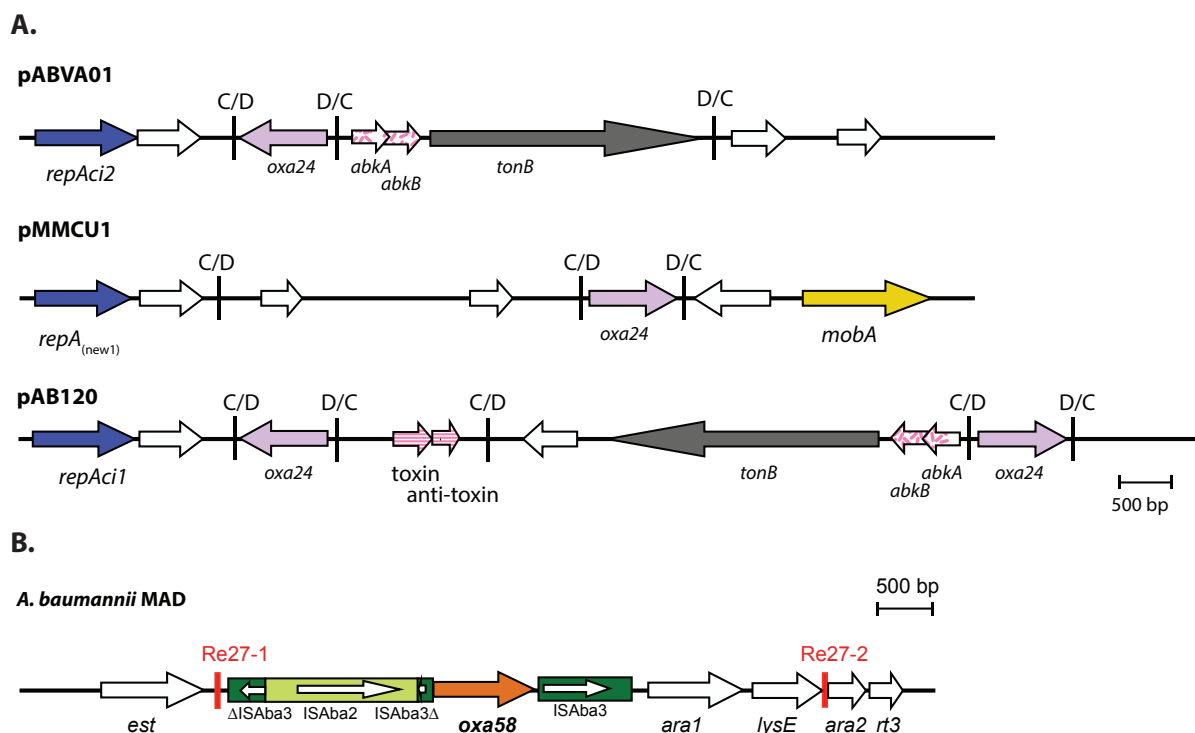


Figure 10.1 Sequences containing the *oxa24* and *oxa58* carbapenemase genes. Arrows indicate the extent and orientation of genes and open reading frames. A. Plasmids containing the *oxa24* *dif* module. The *rep* genes are dark blue and *oxa24* is purple. Genes encoding putative toxin/antitoxin modules are pink patterned. Yellow arrows indicate genes encoding proteins showing homology to mobilisation proteins and the *tonB* dependent transporter is in grey. Vertical black bars indicate *pdif* sites and the orientation of the site is indicated above. B. Re27-1 elements surrounding the *oxa58* gene in *A. baumannii* MAD. The light green box represents ISAbas2, the dark green box is ISAbas3 and the arrows inside indicates the transposase orientation. Re27 elements are indicated as red bars. Figures made to scale from GenBank accession numbers FM210331 (pABVA01), GQ342610 (pMMCU1), JX069966 (pAB120) and AY665723 (MAD).

As the XerC and XerD binding sites do not appear to have any additional accessory sequence, similar to *dif* sites in the chromosome, we will refer to them as plasmid borne *dif* sites or *pdif* from now on. Several further modules surrounded by inversely-oriented *pdif* sites (46), hereafter referred to as *dif* modules, have been identified. A putative toxin/antitoxin system in pAB120 (indicated by the pink striped arrows in Fig 10.1A) is located between two inversely-oriented *pdif* sites and also found in unrelated plasmids may also be mobile (220). Examination

of the sequence surrounding the *oxa58* carbapenemase-encoding resistance gene revealed direct repeats of 27 bp, also called Re-27 elements (red lines in Fig 10.1B) (216). Later these repeat elements were also recognised to be *pdif* sites (46).

Recently, the small plasmid pRCH52-1 recovered from a non GC1 or GC2 clinical isolate from Brisbane was found to contain the *tet39* tetracycline resistance determinant (98). However, identification of clear boundaries for a module containing these genes using the two additional sequences available in GenBank at the time was unsuccessful (98).

The work in this chapter focuses on a small plasmid from the Singapore GC2 collection carrying the *tet39* determinant and the *msrE* and *mphE* macrolide resistance genes. Analysis revealed that both of these resistance determinants are located in *dif* modules. Several further *dif* modules were also identified.

10.2 pS30-1

During the Stand-alone BLAST search of the S30 genome using the pAb-G7-1 backbone as query, two contigs were retrieved (Table 10.1). A small segment of 237 bp of contig 39 (7.5 kb) matched with 85% identity to a region prior to the *rep* gene in pAb-G7-1. Contig 38 (7.7 kb) matched pAb-G7-1 with a 100% identity over a short region of 34 nucleotides. Both contig 38 and 39 ended with fragments of the same IS, which is related to ISAjo2 (Fig 10.2A). Another search of the draft genome of S30 recovered the internal ISAjo2-like IS sequence (Table 10.1).

Table 10.1 Contigs of pS30-1

Contig	Length	Coverage	Features
38	7,498	285.2070	<i>tet39</i> determinant, <i>msrE</i> , <i>mphE</i>
39	7,770	306.8410	-
44	1,482	526.6480	ISAjo2-like IS

ISAjo2 is 1,482 bp in length and is reported to produce a 5-6 bp target site duplication upon insertion into a DNA sequence (64) (<https://www-is.biotoul.fr>). Primers were designed to link the ends of contig 38 and 39 that had matching 5 bp duplications to assemble pS30-1 (Fig 10.2A). RH2502 and tet39R were paired, predicted to amplify a product 1,918 bp in size, while RH2503 and RH2504 should produce an amplicon of 2,058 bp. In both cases, the amplicon appeared to be of the predicted size and was then sequenced to complete the plasmid, which was named pS30-1. This plasmid was 18,234 bp in size and the copy number relative to the chromosome calculated from sequence coverage is approximately 10 to 11.

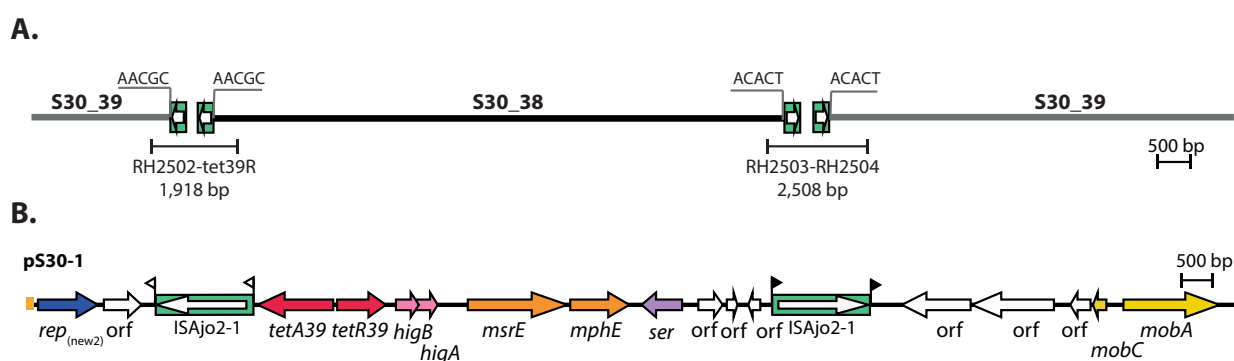


Figure 10.2 Assembly and structure of pS30-1. A. Schematic of contig 38 and 39. Boxes indicate fragments of an ISAjo2-like IS and the arrows inside indicate the direction of the transposase gene. Matching TSD are shown above the line. The bars below the line show primers used for assembly along with the size of the amplicons. B. Annotation of pS30-1. Arrows indicate extent and orientation of genes and open reading frames. Genes encoding putative toxin/antitoxin modules are pink patterned and *rep* genes are shown in blue. The *tet39* determinant is red while the *msrE* and *mphE* genes are orange. Mobilisation genes are highlighted yellow and *ser* is purple. The box represents ISAjo2-1 and the internal arrow indicates the transposase orientation. The orange box indicates the iteron sequence and matching flags indicate target site duplication.

The two IS copies inserted in pS30-1 are identical to one another and share 94% DNA identity with ISAjo2. They are referred to as ISAjo2-1 from here on. ISAjo2-1 contains a single open reading frame that produces a 458 aa putative transposase that shares 97% aa identity with that of ISAjo2.

Several open reading frames were identified in pS30-1 and pBLAST was used to determine their putative functions where possible (Table 10.2 and annotated on Fig 10.2B).

Table 10.2 Open reading frames and predicted products encoded by pS30-1

orf	orf length (bp)	Protein size (aa)	Product	Pfam
1	924	307	RepA (new-2) ¹	Rep_3, pfam01051
2	537	178	OrfQ	YlxM, COG2739
3	1,326	441	TnpA, ISAjo2-1	Rve, pfam00665
4	1,122	374	TetA39	MFS_1, pfam07690
5	720	240	TetR39	TetR_C, pfam02909
6	321	106	HigB, toxin	RelE, pfam062960
7	273	90	HigA, anti-toxin	HTH_3, pfam01381
8	1,476	492	MsrE	Uup, COG0488
9	885	295	MphE	MPH2', cd05152
10	609	203	Ser ²	Resolvase, pfam00239
11	333	111	Orf	
12	171	56	Orf	
13	222	73	Orf	
14	1,326	441	TnpA, ISAjo2-1	Rve, pfam00665
15	1,026	341	Orf	
16	1,248	415	Orf	
17	309	102	Orf	
18	306	101	MobC	Fam199X, pfam15814
19	1,437	479	MobA	MobA/mobL, pfam03387

¹new-1 is in pMMCUI in Fig XA.

²Short for Serine Recombinase.

pS30-1 encodes a replication initiation protein Rep that belongs to the Rep_3 superfamily (pfam01051). The Rep of pS30-1 is 99.7% (306/307 aa) identical to the Rep protein of pM131-5, (GenBank accession number JX101644) (126) from an *A. soli*. To date only Rep types from *Acinetobacter baumannii* plasmids have been characterised and allocated Aci numbers (21), hence we have called the Rep of pM131-5 and pS30-1 Rep (new 2). Thereafter, the closest numbered Rep is RepAci3 (GenBank accession number GU978997), with 91% aa identity (278/307 aa). A putative iteron sequence of four identical repeats of 22 bp is located approximately 50 bp upstream of the *rep* gene (5' AACGAGGTTTACCTTGCAATTA 3') in pS30-1 and pM131-5.

The orf directly following the *rep* gene in pS30-1 contains a domain matching YlxM, which is a predicted DNA binding protein (Table 10.2). Most small plasmids in *Acinetobacter* contain a second reading frame directly downstream of the *rep* gene. This reading frame is most often also annotated as *rep*, implying it has some role in replication, though the reasoning for this is not explained (46). Furthermore, during analysis of pMAC (56), it was determined that the sequence containing the iterons and the region encoding the Rep protein, RepAci9 (178/308 aa, 58% identical to Rep of pS30-1), was all that was required for replication. Henceforth, we will refer to the gene following the *rep* of pS30-1 as orfQ.

Table 10.3 Comparison of OrfQ of pS30-1 to the product of the orf following the *rep* gene in selected plasmids

Plasmid	Rep type	Length of Orf following Rep (aa)	Pfam of Orf	% identity to OrfQ of pS30-1 (aa)
pAb-G7-1	RepAci1	191 ¹	HTH_17, pfam12728	28% (56/200 aa)
pD72-1	RepAci2	191 ¹	HTH_17, pfam12728	28% (57/199 aa)
pBAL_204-1	RepAci3	181	-	79% (144/182 aa)
pMAC	RepAci9	172	-	51% (94/185 aa)
pM131-5	Rep (new 2)	178	YlxM, COG2739	99% (176/178 aa)

¹These two protein as 190/191 aa identical to each other.

The sequence of the orf following the *rep* gene is variable in *Acinetobacter* plasmids as not all of the proteins match the same Pfam or even have a recognisable conserved domain (Table 10.3). Furthermore, the RepAci1 and RepAci2 plasmids pAb-G7-1 and pD72-1 both produce proteins of 99.5% identity (190/191 aa) (Table 10.3) and this homology is clear in the DNA comparison in Chapter 9; Fig 9.1. OrfQ in pS30-1 is most closely related to that produced pM131-5, followed by the Orf translated from the RepAci3 plasmid pBAL_204-1, while only distantly related to the Orf encoded by pAb-G7-1 and pD72-1 (Table 10.3).

Plasmid pS30-1 (Fig 10.2B) contains the *tet39* tetracycline determinant composed of the *tetA39* and *tetR39* genes (6) and the macrolide resistance genes *msrE* and *mphE* (132). In addition, there is also a putative toxin/antitoxin system (indicated by the pink arrows in Fig 10.2B), the first orf of which produces a 107 aa protein of the HigB toxin family (256), while the second translates to a 91 aa protein with a helix-turn-helix DNA binding domain (274), likely performing the function of the counterpart antitoxin HigA. A small serine recombinase, here named *ser* (serine recombinase), is also present (purple arrow in Fig 10.2B). The last orf encodes a 479 aa protein that shares 31% aa identity over 49% of the 709 aa MobA protein of RSF1010 (GenBank accession number NC_001740), the prototype for the MOB_Q family of *mob* genes (68). The reading frame upstream of *mobA* in pS30-1 translates to an 88 aa protein that shares 42% identity with the 94 aa MobC protein of RSF1010. A number of other open reading frames are also present; however, they produce hypothetical proteins of unknown function (Table 10.2).

10.3 pS30-1 transformation

pS30-1 was transformed into ATCC 17978-A and tetracycline resistant transformants were screened to detect specific genes and junctions of the plasmid (Table 10.4). The CDS profile of one transformant was determined and was consistent with the presence of pS30-1 and showed the absence of the second plasmid, pS30-2 (described in detail in Chapter 8; Section 8.6), which would confer resistance to kanamycin (Table 10.4). The *tet39* determinant conferred resistance to tetracycline, reduced susceptibility to doxycycline but did not affect minocycline (Table 10.4), which is consistent with the findings in the original report of this resistance determinant (6).

Table 10.4. Transformation of pS30-1

Strain	Plasmid	pS30-1 specific PCRs		Annular radius (mm)				
		tetA39F- tetA39R	tetA39R- RH2502	Tc	Mc	Do	Km	Su
S30	pS30-1	+	+	0	5	1	0	0
ATCC17978-A	-	-	-	7	10	10.5	8.5	12
ATCC 17978-A TF-1	pS30-1	+	+	0	9	6	9	13

Tetracycline, Tc; Minocycline, Mc; Doxycycline, Dc; Kanamycin, Km; Sulphamethoxazole, Su.

The *msrE* and *mphE* genes confer resistance to macrolide antibiotics (132), a class of antibiotic not commonly used clinically for treatment of Gram-negative bacteria due to a high level of intrinsic resistance (151). However, initial characterisation of these genes involved transforming a region containing both genes into an *E. coli*, which resulted in an increase in erythromycin resistance from 0.5 µg/L to 64 µg/L (132). The ability of pS30-1 to confer resistance to erythromycin was tested by patching ATCC 17978-A and two pS30-1 containing transformants on agar containing increasing concentrations of erythromycin; 2, 8, 16, 32, 64, 128 mg/L (Fig 10.3).

ATCC 17978-A had faint growth on 16 mg/L of erythromycin and could not grow on concentrations above this (Fig 10.3). However, even the highest concentration tested (128 mg/L) was unable to inhibit growth of transformants when pS30-1 was present (Fig 10.3). Although these concentrations of erythromycin are much higher than what would be used clinically it still enables the resistance genes to be detected in the laboratory.

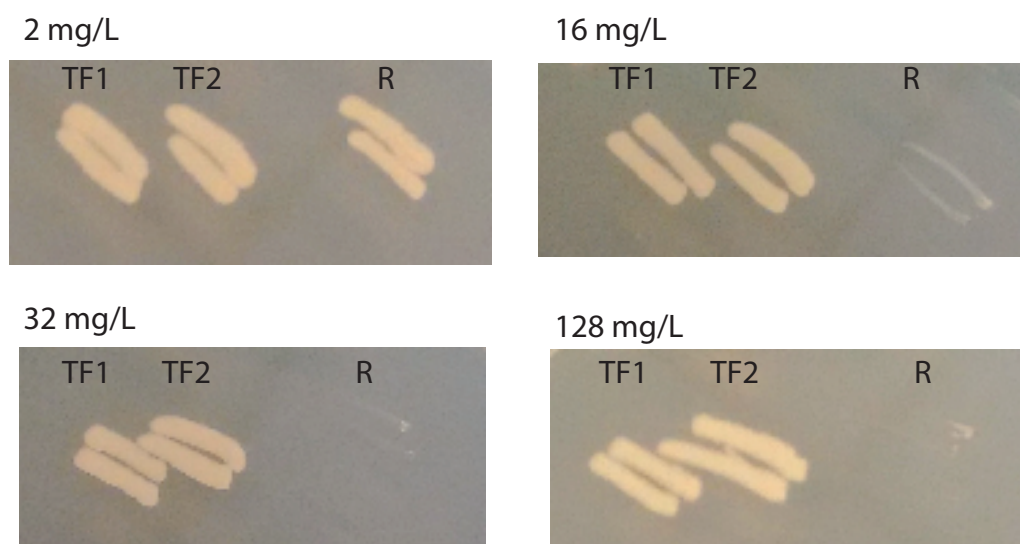


Figure 10.3 Erythromycin resistance. The concentration of erythromycin tested in each frame is shown above. TF1 is ATCC 17978-A(pS30-1) TF-1, TF2 is ATCC 17978-A(pS30-1) TF-2 and R is the empty recipient ATCC 17978-A.

10.4 RepAci6 plasmids does not mobilise pS30-1

As pS30-1 harbours mobilisation genes the ability for it to be mobilised by a RepAci6 plasmid was tested. S30 contains two plasmids, pS30-1 and pS30-2, the second of which is a conjugative RepAci6 plasmid containing the *oxa23* gene in Tn2008 (conferring resistance to ticarcillin) as well as the *aphA6* gene in TnaphA6 (conferring resistance to kanamycin) (Chapter 8; Section 8.2). Transfer experiments were performed using S30 as a donor and ATCC 19606^{Rif} (rifampicin resistant) or AB307_0294^{Rif} (rifampicin resistant) as the recipient and testing all possible outcomes (Table 10.5). With both recipients, only ticarcillin resistance was transferred, attributed to the transferral of only pS30-2, as confirmed via PCR screening (Table 10.5). Hence, pS30-1 is not mobilised by pS30-2 and likely not by other RepAci6 plasmids.

Table 10.5 pS30-1 mobilisation experiments

Strain	Plasmids	Phenotype	Observed transconjugants ¹	pS30-1 specific PCR		pS30-2 specific PCR	
				tet39F-tet39R	tet39R-RH1502	oxa23F-oxa23R	RH1503-RH1397
S30-1 (donor)	pS30-1, pS30-2	TcTIC	N/A	+	+	+	+
ATCC19606 ^{Rif} (recipient)	N/A	Rif	N/A	-	-	-	-
ATCC19606 ^{Rif}	pS30-1	TcRif	<8.7x10 ⁻⁹	N/A	N/A	N/A	N/A
ATCC19606 ^{Rif}	pS30-2	TICRif	Y 2.34x10 ⁻⁵	-	-	+	+
ATCC19606 ^{Rif}	pS30-1, pS30-2	TcTICRif	<8.7x10 ⁻⁹	N/A	N/A	N/A	N/A
AB307_0294 ^{Rif} (recipient)	N/A	Rif	N/A	-	-	-	-
AB307_0294 ^{Rif}	pS30-1	TcRif	<3.9x10 ⁻⁹	N/A	N/A	N/A	N/A
AB307_0294 ^{Rif}	pS30-2	TICRif	Y 3.90x10 ⁻⁵	-	-	+	+
AB307_0294 ^{Rif}	pS30-1, pS30-2	TcTICRif	<3.90x10 ⁻⁹	N/A	N/A	N/A	N/A

Rifampicin, Rif; Tetracycline, Tc; Ticarcillin, TIC.

¹Frequency is expressed as transconjugants per donor and is calculated from one experiment.

10.5 XerC-XerD binding sites in pS30-1

Plasmid pS30-1 did not have any clear relatives in GenBank. Rather, small fragments of this plasmid matched fragments of other *Acinetobacter* plasmids. Some of these matching regions included the sequence encoding the *tet39* tetracycline resistance determinant and the *msrE* and *mphE* macrolide resistance genes. The boundaries of the *tet39* determinant have not been previously defined (98) and the macrolide resistance genes in pS30-1 do not appear to be in any known transposon. As XerC-XerD binding sites had been found surrounding the *oxa24* and the *oxa58* resistance genes in *Acinetobacter* plasmids, a manual search for these sites in pS30-1 was performed, resulting in the identification of 8 *pdif* sites, numbered from left to right (Fig 10.4). The location and orientation of the *pdif* sites are indicated on Fig 10.4A, with sequences of the identified *pdif* sites compared in Fig 10.4B. For these identified *pdif* sites, there is greater variability in the sequence of the XerC binding sites than XerD binding sites. Apart from the sequence of the spacer in the first and third *pdif* site, the rest of the spacers are

unique. A number of these sites surround segments of DNA containing 1 or 2 genes, including the region containing the *tet39* determinant and the *msrE-mphE* genes and the following subsections of Section 10.5 detail these.

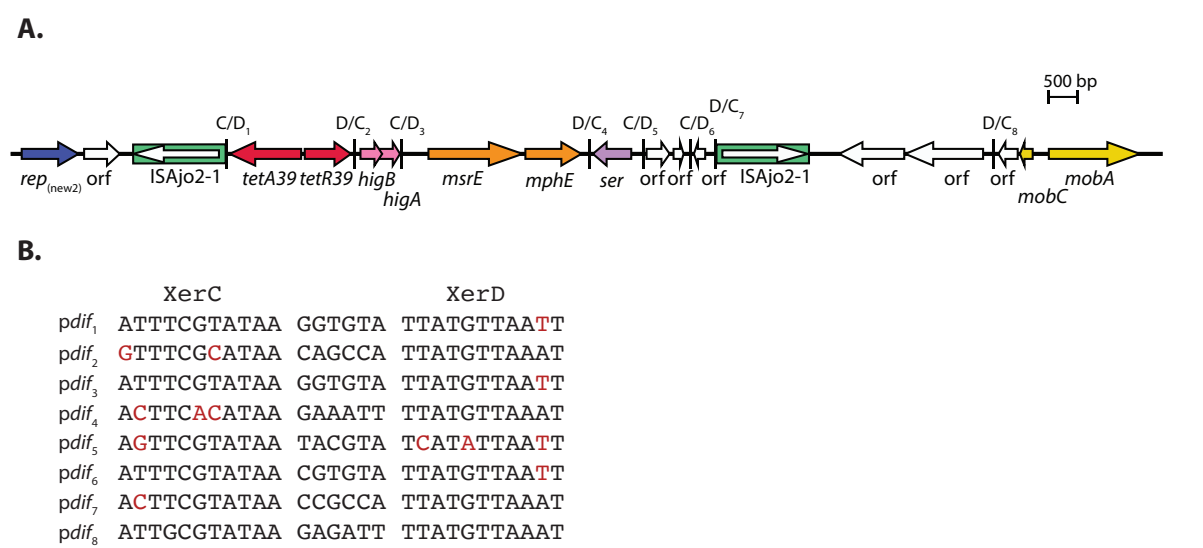


Figure 10.4 pS30-1 *pdif* sites. A. The location of *dif* sites in pS30-1. Arrows indicate the extent and orientation of genes and open reading frames. The *tetA* gene and the *tetR* genes of *tet39* are shown in red and *rep* genes are dark blue. Genes encoding putative toxin/antitoxin modules are pink or pink patterned. The *msrE* and *mphE* resistance genes are orange and *ser*, encoding a serine recombinase is purple. Genes encoding proteins showing homology to mobilisation proteins are coloured yellow. The light green box represents ISAjo2-1 and the internal arrow indicates the transposase orientation. Vertical bars indicate *pdif* sites and the orientation of the site is shown above. B. Comparison of the sequence of the 8 *pdif* sites from left to right. The bases highlighted in red in the XerC and XerD sites vary.

10.5.1 *pdif* sites bound *tet39*

The first two *pdif* sites (*pdif*₁ and *pdif*₂) identified in pS30-1 surround the *tetA39* and *tetR39* genes. Like the *oxa24* module, these two sites are inversely-oriented with XerC-XerD on the left and XerD-XerC on the right. These sites surround a 2,001 bp DNA segment containing the *tet39* determinant that together likely form a *dif* module. Using this putative *dif* module as a query in GenBank, five sequences were retrieved (Table 10.6). A manual search for *pdif* sites was performed in each of these five sequences as no *pdif* sites had been annotated previously (Table 10.6).

Table 10.6 Sequences containing the *tet39 dif* module

Sequence	Sequence Length (bp)	<i>pdif</i> sites	Source	Accession no.	Reference
pS30-1	18,234	8	<i>A. baumannii</i>	KY617771	This study
pRCH52-1	11,164	3	<i>A. baumannii</i>	KT346360	(98)
pAB1H8	10,246	3	<i>A. baumannii</i>	ANNC01000048	(118)
pLU5605 ¹	3,727	1	<i>Acinetobacter</i>	AY743590	(6)
TE_01Z_000_contig 11 ¹	2,446	2	Uncultured bacterium	KU545777	(206)
TE_IN_B_contig_8 ¹	3,953	3	Uncultured bacterium	KU544458	(206)

¹Partial sequences

The same 2,001 bp segment was also found surrounded by inversely-oriented *pdif* sites in pRCH52-1, pAB1H8, TE_01_00_contig 11 and TE_IN_B_contig8 (Fig 10.5). The original sequence for the *tet39* determinant in pLUH5605 (GenBank accession number AY743590) (6) has a *pdif* site on the left but does not extend far enough to include the *pdif* site to the right. Beyond the bounding *pdif* sites there is little similarity between each of the sequences (Fig 10.5), allowing the boundaries of this *dif* module to be defined. Hence, the *tet39* determinant is located in a *dif* module that has been found in 6 different contexts, which demonstrates that it is mobile.

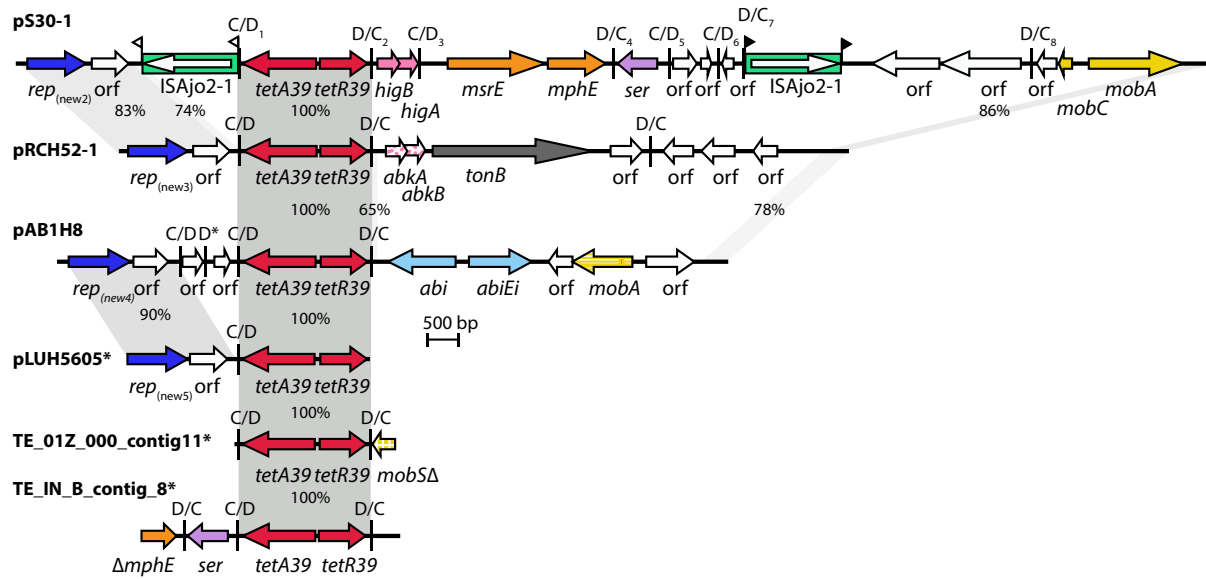


Figure 10.5 The *tet39* *dif* module. Comparison of regions containing the *tet39* resistance determinant. Partial sequences are marked with an asterisk. Arrows indicate the extent and orientation of genes and open reading frames. The *tetA* gene and the *tetR* genes of *tet39* are shown in red and *rep* genes are dark blue. Genes encoding putative toxin/antitoxin modules are pink or pink patterned. The *msrE* and *mphE* resistance genes are orange and *ser*, encoding a serine recombinase is purple. Genes encoding proteins showing homology to mobilisation proteins are coloured yellow. The *tonB* gene encodes a TonB-dependent transporter homologue. The light green box represents ISAjo2-1 and the internal arrow represents the transposase orientation. The extents of regions with significant DNA identities are shown in grey. Vertical bars indicate *pdif* sites and the orientation of the site is shown above, and flags indicate target site duplications. The picture is drawn to scale from the GenBank entries listed in Table 10.6.

The sequences of the *pdif* sites flanking the *tet39* module were compared (Table 10.7) and both the left and the right *pdif* sites showed variation. Consequently, neither the left or right *pdif* site could be clearly be assigned to the module. If variation in the XerC-XerD sequence was present, it was located in the outermost sites (in this case XerC) and the 6 bp spacer region (Table 10.7), while the innermost XerD sites were always conserved.

Table 10.7 Comparison of *pdif* sites surrounding the *tet39* determinant

Sequence	<i>tet39</i> fragment size (bp)	Left <i>pdif</i> site: XerC-XerD	Right <i>pdif</i> site: XerD-XerC
pS30-1	2,001	<u>ATTTTCGTATAAGGTGTATTATGTTAATT</u>	<u>ATTTAACATAATGGCTGTTATGCGAAAC</u>
pRCH52-1	2,001	<u>ATTTTCGTATAAGGTGTATTATGTTAATT</u>	<u>ATTTAACATAATGGATATTATACGAACT</u>
pAB1H8	2,001	<u>ATTTTCGTATAAGGTGTATTATGTTAATT</u>	<u>ATTTAACATAATAGAGCTTATGCGAAAT</u>
pLUH5605	1,986	<u>ATTTTCGTATAAGGTGTATTATGTTAATT</u>	–
TE_01Z_000_ contig11	2,001	<u>ATTTTCGTATAAGGTGTATTATGTTAATT</u>	<u>ATTTAACATAATAGAGCTTATACGAAAT</u>
TE_IN_B_ contig_8	2,001	<u>AATTCGTATAACGCGTATTATGTTAATT</u>	<u>ATTTAACATAATAGAGCTTATACGAAAT</u>

Red bases vary.

10.5.2 *pdif* sites bound the *msrE-mphE* gene pair

The *pdif* sites 3 and 4 are in inverse orientation with XerD sites internal (XerC-XerD and XerD-XerC) and flank a 2,950 bp segment containing the *msrE* and *mphE* macrolide resistance genes (Fig 10.4A). A search of GenBank with this segment revealed that five other plasmids and partial sequences contain the same segment (Table 10.8). However, none these five plasmids also included the *tet39* determinant (Table 10.8).

Table 10.8 Plasmids containing the *msrE* and *mphE dif* module

Sequence	Sequence Length (bp)	<i>pdif</i> sites ¹	Source	Accession no.	Reference
EF102240 ²	13,122	5	<i>A. baumannii</i>	EF102240	(213)
pABIR	29,823	13 (2)	<i>A. baumannii</i>	EU294228	(281)
pOXA58_882	36,862	16 (2)	<i>A. pittii</i>	CP014479	(7)
pXBB1-9	398,857	12 (9)	<i>A. johnsonii</i>	CP010351	(64)
p255n-1	92,939	9	<i>A. baumannii</i>	KT852971	(235)

¹Number of *pdif* sites identified in this study. If any *pdif* or Re27 sites were identified previously, the number found is in brackets.

²This sequence is of a cloned insert of a larger plasmid from *A. baumannii* A1.

No XerC-XerD sites in the sequence of EF102240 and p255n-1 had been annotated. pABIR (281) and pOXA58_882 (7) had previously been reported to contain two Re27 sites while pXBB1-9 had been annotated with both Re27 and XerC-XerD sites (totalling 9 sites) (64). All sequences were manually searched for *pdif* sites and in all cases additional sites were identified (Table 10.8). The region containing the *msrE-mphE* genes were compared (Fig 10.6) to attempt to define the module containing these genes.

A.

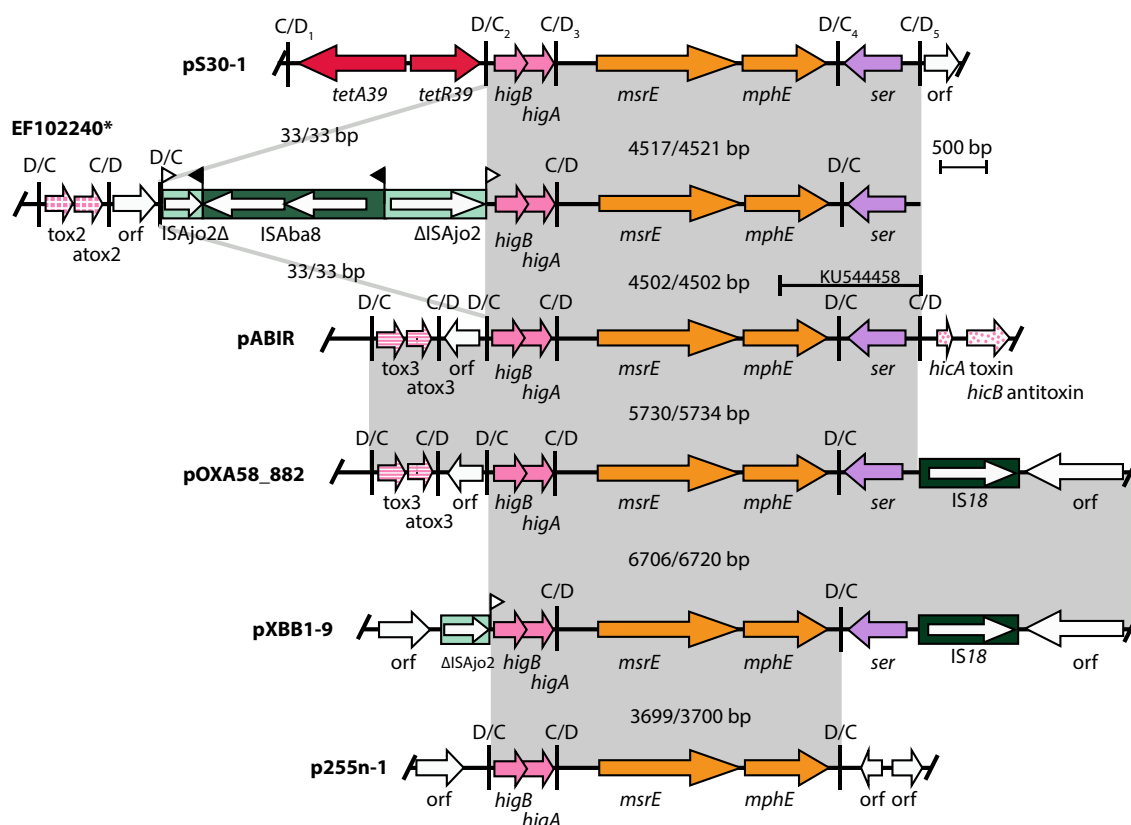


Figure 10.6 *msrE*-*mphE* *dif* module. Comparison of sequences containing the *msrE* and *mphE* resistance genes. Partial sequences are marked with an asterisk. Arrows indicate the extent and orientation of genes and open reading frames. The *tetA* gene and the *tetR* genes of *tet39* are shown in red. Genes encoding putative toxin/antitoxin modules are pink or pink patterned. The *msrE* and *mphE* resistance genes are orange and *ser*, encoding a serine recombinase is purple. Genes encoding proteins showing homology to mobilisation proteins are coloured yellow. The light green box indicates *ISAjo2-1*, while the dark green box indicates *IS18*, and the internal arrows indicate the transposase orientation. Vertical bars indicate *pdif* sites and the direction of the site is shown above, and flags indicate TSD. The extents of regions with significant DNA identities are shown in grey and the numbers are the DNA identities. The picture is drawn to scale from the following GenBank accession numbers listed in Table 10.8.

Unlike the *tet39* modules, all sequences had the same putative toxin/antitoxin sequence to the left of the *msrE*-*mphE* genes and all apart from p255n-1 had the *ser* gene on the right (Fig 10.6). As a consequence, when the sequences of the *pdif* sites surrounding the *msrE*-*mphE* genes were compared there were only two variants (Table 10.9) and the second was present in p255n-1, which is consistent with the different set of genes adjacent to the right *pdif* site. Again, variation was present in the outermost binding site. In the case of the *msrE*-*mphE* genes, the

dif module was defined as the central 2,950 bp containing these resistance genes surrounded by inversely-oriented *pdif* sites, which was present in the 6 sequences examined here (Fig 10.6).

Table 10.9 *pdif* sites surrounding the *msrE-mphE* determinant

Sequence	<i>msrE-mphE</i> fragment size (bp)	Left <i>pdif</i> site: XerC-XerD	Right <i>pdif</i> site: XerD-XerC
pS30-1 [†]	2,950	<u>ATTTCGTATAAGGTGTATTATGTTAATT</u>	<u>ATTTAACATAAAAATTTCTTATGTGAAGT</u>
p255n-1	2,950	<u>ATTTCGTATAAGGTGTATTATGTTAATT</u>	<u>ATTTAACATAAAAATTTCTATATACGAAAT</u>
Tn1548	2,891	–	<u>ATTTAACATAAAAATTTCTTATACGAAAT</u>

[†]EF102240, pABIR, pOXA58_882 and pXBB1-9 have the same sequence *pdif* sequence.

The *msrE* and *mphE* resistance genes are also found in Tn6180, located in AbGRI3 in some GC2 isolates (Chapter 6) (23), and in the closely related transposon Tn1548 (83). These two transposons were examined for *pdif* sites and were found contained the right *pdif* site of the *msrE-mphE dif* module, but an ISEc29 has replaced the left *pdif* site and 120 bp of the internal sequence (Fig 10.7). The sequence of the right *pdif* site is a third variation (Table 10.9).

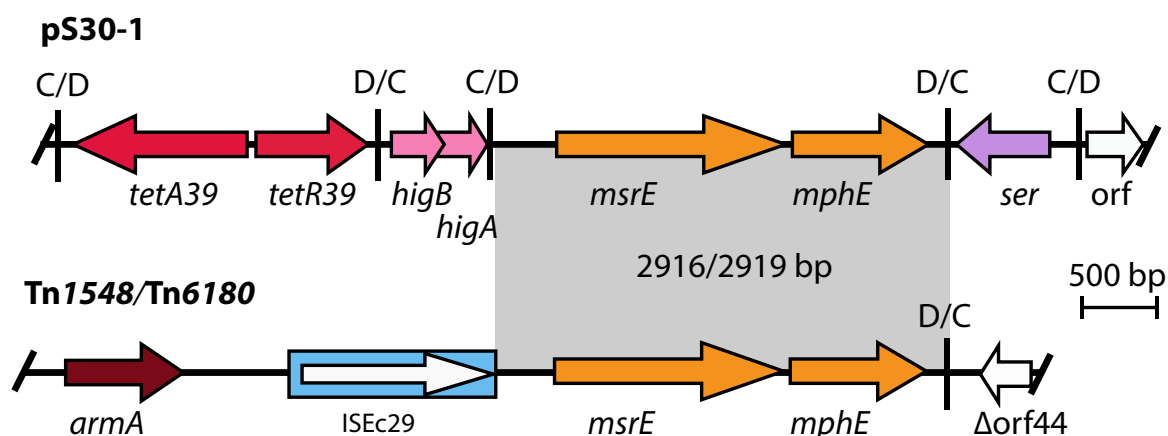


Fig 10.7 Comparison of the region containing *msrE-mphE* in pS30-1 and Tn1548/Tn6180. Arrows indicate the extent and orientation of genes and open reading frames. The *armA*, *tet39* determinants, *msrE-mphE* are coloured maroon, red and orange, respectively. Putative toxin-antitoxin genes are pink and the *ser* gene is purple. The light blue box is ISEc29 and the internal arrow indicate the direction transposase orientation. Vertical bars indicate *pdif* sites and the direction of the site is shown above. The extents of regions with significant DNA identities are shown in grey.

10.5.3 Additional *dif* modules in pS30-1

In addition to the two modules described above, four other potential *dif* modules bounded by inversely-oriented *pdif* sites were identified in pS30-1. One, located between the *tet39* and the *msrE-mphE* modules, contains two overlapping open reading frames that encode a putative HigB toxin and HigA antitoxin system (Fig 10.4A), and are flanked by inversely-oriented *pdif* sites (XerD-XerC and XerC-XerD). As mentioned in Section 10.5.2, this module was located to the left of the *msrE-mphE* module in the six plasmids examined (Fig 10.6). However, in most cases the plasmids differed in the sequence to the left of the *higB/higA* module (Fig 10.6 and Table 10.10). Where the left and right *pdif* sites were detected, the sequences were identical, including the sequence of the 6 bp spacer (Table 10.10).

Table 10.10 The *higB* and *higA* toxin/antitoxin *dif* module

Sequence	<i>higB/higA</i> fragment size (bp) ¹	Sequence left of module	Left <i>dif</i> site: XerD-XerC	Right <i>dif</i> site: XerC-XerD	Sequence right of module
pS30-1	701	<i>tet39</i>	<u>ATTTAACATAATGGCTGTTATGCGAAAC</u>	<u>ATTTCGTATAAGGTGTATTATGTTAATT</u>	<i>msrE-mphE</i>
p255n-1	701	orf (META domain)	<u>ATTTAACATAATGGCTGTTATGCGAAAC</u>	<u>ATTTCGTATAAGGTGTATTATGTTAATT</u>	<i>msrE-mphE</i>
pOXA58_882	703	orf (SMI1_KNR4 domain)	<u>ATTTAACATAATGGCTGTTATGCGAAAC</u>	<u>ATTTCGTATAAGGTGTATTATGTTAATT</u>	<i>msrE-mphE</i>
pABIR	701	orf (SMI1_KNR4 domain)	<u>ATTTAACATAATGGCTGTTATGCGAAAC</u>	<u>ATTTCGTATAAGGTGTATTATGTTAATT</u>	<i>msrE-mphE</i>
EF102240	4,147 ²	orf (no known domains)	<u>ATTTAACATAATGGCTGTTATGCGAAAC</u>	<u>ATTTCGTATAAGGTGTATTATGTTAATT</u>	<i>msrE-mphE</i>
pXBB1-9	>701 ³	orf (HTH domain)	–	<u>ATTTCGTATAAGGTGTATTATGTTAATT</u>	<i>msrE-mphE</i>

¹Sequence between the left and right *dif* site.

²ISAjo2 has inserted adjacent to the left *dif* site and an ISAb8 has interrupted the ISAjo2.

³Has a fragment of ISAjo2 in the same location as EF102240 but then goes into different sequence and there is no discernible *dif* site in this sequence.

Downstream of the *msrE-mphE* module in pS30-1 is a region encoding a small serine recombinase (*ser*) surrounded by *pdif*₄ and *pdif*₅ which are inversely-oriented (XerD-XerC and XerC-XerD). With the exception of p255n-1, the *ser* gene is also found following the *msrE-mphE* module (Fig 10.6), however, there are two variants of the *ser* module based on the presence of a *pdif* site on the right (Table 10.11).

Table 10.11 The *ser dif* module

Sequence	<i>ser</i> fragment size (bp) ¹	Sequence left of module	Left <i>dif</i> site: XerD-XerC	Right <i>dif</i> site: XerC-XerD	Sequence right of module
pS30-1	809	<i>msrE-mphE</i>	<u>ATTTAACATAAAATTTCTTATGTGAAGT</u>	<u>AGTTCGTATAATACGTATCATATTAATT</u>	orf (no known domains)
pABIR	809	<i>msrE-mphE</i>	<u>ATTTAACATAAAATTTCTTATGTGAAGT</u>	<u>AATTCGTATAACGTGCATTATGTTAATT</u>	<i>hicA/hicB</i>
TE_IN_B_ contig_8	809	<i>msrE-mphE</i>	<u>ATTTAACATAAAATTTCTTATGTGAAGT</u>	<u>AATTCGTATAACGCGTATTATGTTAATT</u>	<i>tet39</i>
EF102240	809	<i>msrE-mphE</i>	<u>ATTTAACATAAAATTTCTTATGTGAAGT</u>	<u>AATTC-</u>	-
pOXA58_882	780 ²	<i>msrE-mphE</i>	<u>ATTTAACATAAAATTTCTTATGTGAAGT</u>		IS18, orf
pXBB1-9	780 ²	<i>msrE-mphE</i>	<u>ATTTAACATAAAATTTCTTATGTGAAGT</u>		IS18, orf

¹Sequence between the left and right *dif* site.

²After 780 bp the sequence diverges from the sequence in pS30-1 and the other plasmids.

The variant with the *pdif* site on the right includes pS30-1, TE_IN_B_contig_8 (KU54458) and pABIR. EF102240 is also a member of this group though the available sequence only extends 4 bp into the XerC binding site of the right *dif*. The right *pdif* sequence and 29 bp of the *ser* module directly prior to it have been replaced 33 bp of unrelated sequence including an *IS18* in pOXA58_882 and pXBB1-9 (Fig 10.6), with the absence of any site resembling a XerC-XerD binding site found in the sequence directly following the *IS18*. Serine recombinases bind to three subsites within an approximately 114 bp *res* site, *resI*, *II* and *III* (243). They then cut the DNA at the centre of the active site, located within *resI*, with a 2 bp ‘stagger’ (248). It is reasonable to suggest that a *res* site for *ser* is located in the sequence between the start codon of *ser* and the *dif* site on the right (these sites are underlined and boxed Fig 10.8, respectively). The divergence in sequences between groups could be explained by the *ser* gene facilitating a recombination event within *resI* in pOXA58_882 and pXBB1-9, accounting for the divergence in sequence between the two groups. The putative 2 bp stagger is in bold in Fig 10.8.

```

pS30-1      GTTTAGATCCTGAAGAATCTGCAAAGCCCTTTCTGCATCTTGATCTTTAGTTGAAGCTCG
KU544458    GTTTAGATCCTGAAGAATCTGCAAAGCCCTTTCTGCATCTTGATCTTTAGTTGAAGCTCG
EF102240    GTTTAGATCCTGAAGAATCTGCAAAGCCCTTTCTGCATCTTGATCTTTAGTTGAAGCTCG
pABIR       GTTTAGATCCTGAAGAATCTGCAAAGCCCTTTCTGCATCTTGATCTTTAGTTGAAGCTCG
pOXA58_882  GTTTAGATCCTGAAGAATCTGCAAAGCCCTTTCTGCATCTTGATCTTTAGTTGAAGCTCG
pXBB1-9     GTTTAGATCCTGAAGAGTCTGCAAAGCCCTTTCTGCATCTTGATCTTTAGTTGAAGCTCG
*****

pS30-1      TAGATAAATACGTGTATTTCATTTCTATCATTTAATCTTAATTCATTAGATTAATGATAA
KU544458    TAGATAAATACGTGTATTTCATTTCTATCATTTAATCTTAATTCATTAGATTAATGATAA
EF102240    TAGATAAATACGTGTATTTCATTTCTATCATTTAATCTTAATTCATTAGATTAATGATAA
pABIR       TAGATAAATACGTGTATTTCATTTCTATCATTTAATCTTAATTCATTAGATTAATGATAA
pOXA58_882  TAGATAAATACGTGTATTTCATTTCTATCATTTAATCTTAATTCATTAGATTAATGATAA
pXBB1-9     TAGATAAATACGTGTATTTCATTTCTATCATTTAATCTTAATTCATTAGATTAATGATAA
*****

pS30-1      TACTATCATTTAATTATTGCATTAAGTCTAATGATAGATATTGTATGCGATTTGGCTCGC
KU544458    TACTATCATTTAATTATTGCATTAAGTCTAATGATAGATATTGTATGCGATTTGGCTCGC
EF102240    TACTATCATTTAATTATTGCATTAAGTCTAATGATAGATATTGTATGCGATTTGGCTCGC
pABIR       TACTATCATTTAATTATTGCATTAAGTCTAATGATAGATATTGTATGCGATTTGGCTCGC
pOXA58_882  TACTATCATTTAATTATTGCATTAAGTCTAATGATAGATATTGTATGCGATTTGGCTCGC
pXBB1-9     CACTATCATTTAATTATTGCATTAAGTCTAATGATAGATATTGTATGCGATTTGGCTCGC
*****

pS30-1      TATCATTTCAGGTATACTCTTTTGATAGAAAAAGCATGACTGAAAGCCATGCCTTTGGGAGT
KU544458    TATCATTTCAGGTATAGTCTTTTGATAGAAAAAGCATGACTAAAAGCCATGCCTTTGGGAAT
EF102240    TATCATTTCAGGTATACTCTTTTGATAGAAAAAGCATGACTAAAAGCCATGCCTTTGGGAAT
pABIR       TATCATTTCAGGTATAGTCTTTTGATAGAAAAAGCATGACTAAAAGCCATGCCTTTGGGAAT
pOXA58_882  TATCATTTCAGGTATAGTCTTTTGATAGACTGCTTGTTTGCGGAATGTTCCGATTTGAACC
pXBB1-9     TATCATTTCAGGTATAGTCTTTTGATAGACTGCTTGTTTGCGGAATGTTCCGATTTGAACC
*****
XerC      XerD
pS30-1      TCGTATAATACGTATCATATTAATTTTGGAAAAATTAGATATTGATTGGGAGTGA
KU544458    TCGTATAACGCGTATTATGTTAATTTTAGATAAACTATAGAATCCATTAGTTTTT
EF102240    TC-----
pABIR       TCGTATAACGTGCATTATGTTAATTTTAGAAAAAGCTTAAGAAATATCATTAATAA
pOXA58_882  ACGCAGATTCAAATCCCAAGTGCAACACATTTGTAGTCAGTTGAAAAAGTTAGGT
pXBB1-9     ACGCAGATTCAAATCCCAAGTGCAACACATTTGTAGTCAGTTGAAAAAGTTAGGT
*

```

Figure 10.8 Comparison of the sequence at the right end of the *ser* module. The nucleotides of the *ser* gene are purple and the start codon is underlined. Sequence of *IS18* is green. The XerC and XerD binding sites are boxed in red and orange, respectively. The putative two base pair stagger is in bold. Asterisks indicate conserved base pairs.

The *pdif* sites 6 and 7 in pS30-1 are also in inverse orientation to one another, bounding 368 bp of DNA with a single open reading frame that encodes a protein of no known function (Fig 10.4A). This same 368 bp fragment of DNA is also found in pALWVS1.1 (GenBank accession number KX426232) (175) and a search of the surrounding sequence revealed two *pdif* sites, the sequence of which are compared to those of pS30-1 in Table 10.12. As this module was in different contexts in pS30-1 and pALWVS1.1 (Table 10.12) there is evidence that it too is a mobile *dif* module.

Table 10.12 *pdif* sites 6 and 7 bound a *dif* module

Sequence	orf fragment size (bp)	Sequence left of module ¹	Left <i>dif</i> site: XerC-XerD	Right <i>dif</i> site: XerD-XerC	Sequence right of module ¹
pS30-1	368	orf, orf	<u>ATTTCGTATAACGTGTATTATGTTAATT</u>	<u>ATTTAACATAATGGCGGTTATACGAAGT</u>	ISAjo2-1, orf
pALWVS1.1	368	antitoxin ² / toxin ³	<u>ACTTCGTATAAGGTGTATTATGTTAATT</u>	<u>ATTTAACATAATGGCGGTTATACGAAGT</u>	orf

¹ All genes called orf are different.

² RelE/DinJ family.

³ RelE/StbE family.

The *pdif*₈ site upstream of the *mobC* gene and *pdif*₁ are also in inverse-orientation and could potentially move as a module (Fig 10.4A). However, when comparing this region to available sequences in GenBank this does not appear to be the case. A comparison of pS30-1 to pM131-5, which contains a *rep* almost identical to pS30-1, is shown in Fig 10.9. The region of high similarity (98.9 % DNA identical) over the *rep* gene ends approximately 100 bp prior to the first *pdif* site of pM131-5. In addition, the match over the region containing the *mob* genes starts just prior to these genes and does not encompass *dif*₈ (Fig 10.9).

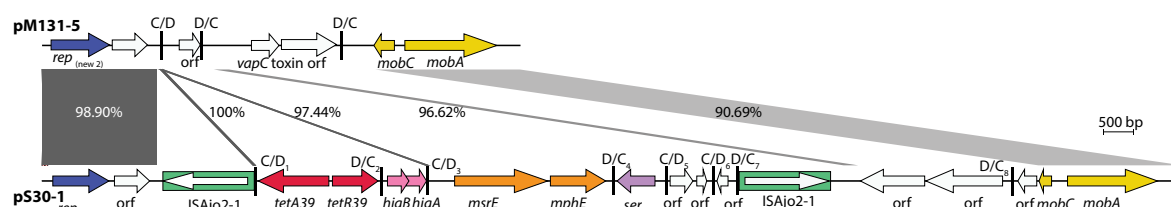


Figure 10.9 Comparison of pM131-5 and pS30-1. Arrows indicate the extent and orientation of genes and open reading frames. The *tetA* gene and the *tetR* genes of *tet39* are shown in red and *rep* genes are dark blue. Genes encoding putative toxin/antitoxin modules are pink or pink patterned. The *msrE* and *mphE* resistance genes are orange and *ser*, encoding a serine recombinase is purple. Genes encoding proteins showing homology to mobilisation proteins are coloured yellow. The *tonB* gene encodes a TonB-dependent transporter homologue. The light green box represents ISAjo2-1 and the internal arrow indicates the transposase orientation. Vertical bars indicate *pdif* sites and the direction of the site is shown above. The extents of regions with significant DNA identities are shown in grey and the numbers are the DNA identities.

The *pdif* sites 5 and 6, and 7 and 8, are in direct orientation with one another (Fig 10.4A). These two pairs of sites flank open reading frames that produce proteins of unknown functions (Table 10.2). When these segments of DNA were used as queries against GenBank, no matches were

detected. Hence, it is unclear at this point if these two segments are potentially mobile *dif* modules.

10.7 Two types of *dif* modules

An emerging feature from the analysis of the *pdif* sites in pS30-1 and in several other plasmids is that, in most cases, the *pdif* sites surrounding a module are inversely-oriented. This was noticed earlier using only a few sequences (46) and is largely true in the cases described above. This inevitably leads to two types of *dif* modules. One variation of the *dif* module is flanked by *pdif* sites which are oriented as XerC-XerD and a XerD-XerC site respectively, such that the XerD binding sites are internal to the segment. I will henceforth refer to such *dif* modules as D modules, which include the *tet39* and *msrE-mphE* module (Fig 10.4A) as well as the *oxa24* module (Fig 10.1) (46). The second variation of *dif* modules are instead bounded with *pdif* in the XerD-XerC/XerC-XerD arrangement (XerC sites are internal) are referred to as C modules. Examples of these include the *higA/higB*, *ser* and *orf* module in pS30-1 (Fig 10.4A).

10.6.1 C and D modules alternate

When a region contains multiple *dif* modules, the D and C modules should alternate due to the nature of inversely-oriented *pdif* sites. The first four modules in pS30-1 alternate (Fig 10.10).

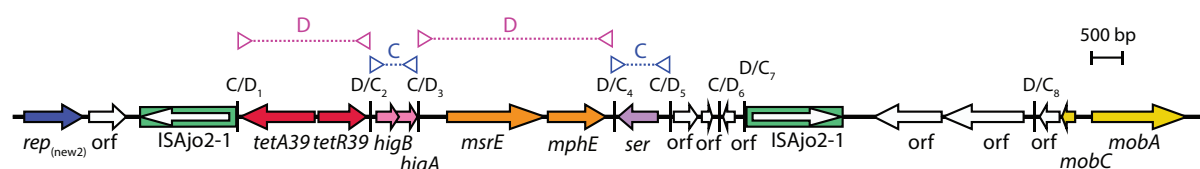


Figure 10.10 Alternating C and D modules in pS30-1. Arrows indicate the extent and orientation of genes and open reading frames. The *tetA* gene and the *tetR* genes of *tet39* are shown in red and *rep* genes are dark blue. Genes encoding putative toxin/antitoxin modules are pink or pink patterned. The *msrE* and *mphE* resistance genes are orange and *ser*, encoding a serine recombinase is purple. Genes encoding proteins showing homology to mobilisation proteins are coloured yellow. The light green box represents ISAjo2-1 and the internal indicates the transposase orientation. Vertical bars indicate *pdif* sites and the orientation of these sites are shown above. D and C modules are indicated above by pink and blue dashed lines bounded by arrows, respectively.

As the Re-27 sites surrounding the *oxa58* gene were identified as direct repeats (64, 216, 281), the regions surrounding this gene in *A. baumannii* MAD were examined. The partial 7.8 kb of sequence supplied in GenBank accession number AY665723) (214, 216), which is a part of an approximately a 30 kb plasmid, was examined for additional *pdif* sites. As shown in Fig 10.11, two additional *pdif* sites were found, and both are in the opposite orientation to the Re-27 sites. One creates a smaller C type *dif* module containing *oxa58* and an adjacent D module containing genes usually annotated *araC* and *lysE*. The second additional *pdif* site discloses an additional adjacent C *dif* module (Fig 10.11). and with the additional *pdif* sites in *A. baumannii* MAD the identified modules in this sequence follow the trend of C and D alternation (Fig 10.11).

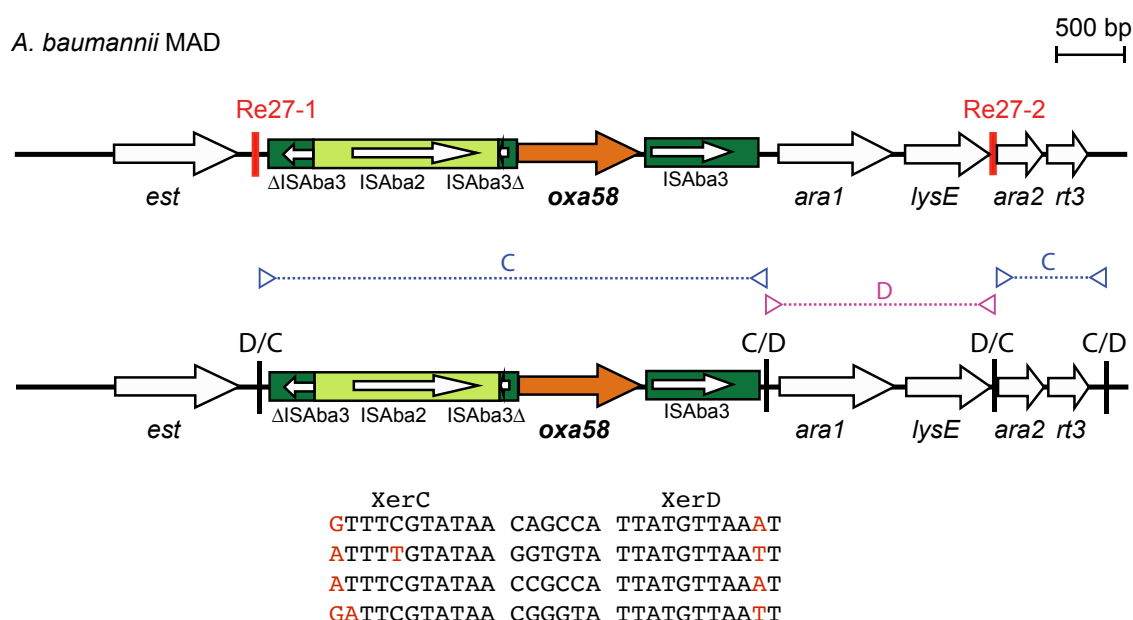


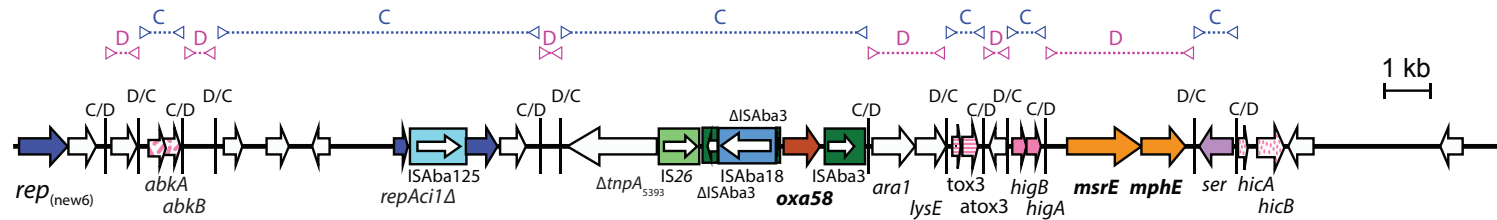
Figure 10.11 The *oxa58* gene is located in a *dif* module in pMAD. Arrows indicate the extent and orientation of genes and open reading frames. The *oxa58* gene is dark orange. The light green box represents ISAbA2, the dark green box is ISAbA3 and the internal arrows indicate the transposase orientation. Vertical red and black bars represent Re27 and *pdif* sites, respectively, and the orientation of the *pdif* sites are shown above. D and C modules are indicated above by pink and blue dashed lines bounded by arrows.

To further test whether alternation of the C and D *dif* modules is a general feature of *dif* module regions, pABIR (29.8 kb) and pOXA58_882 (36.8 kb) were examined. These two plasmids were identified in the *msrE-mphE* module comparison and when more of the adjacent sequence was examined, a total of 13 and 16 *pdif* sites, respectively (Table 10.8) were identified. The 12

modules in pABIR all alternate, as do the 15 modules in pOXA58_882 (Fig 10.12A and B). Both of these plasmids include the *oxa58* resistance gene, and although this gene is located between two inversely-oriented *dif* sites in both plasmids the internal sequence of this module appears to vary.

The overall trend when there are multiple *dif* modules in the same sequence appears to be that the two types of *dif* module should alternate. However, as seen in pS30-1 (Fig 10.10), this is not always the case. Two regions of this plasmid are surrounded by directly-oriented *pdif* sites and careful examination did not reveal an additional site within them. Another region bounded by directly-orientated *dif* sites containing the *abkA-abkB* toxin/antitoxin system and the *tonB* gene, is found in pD72-1 and pAb-G7-2 (Chapter 9; Fig 9.1) and the same region is also present in pRCH51-1 (Fig 10.13A). Whether these regions are actually fusions of two *dif* units may become apparent as further sequences become available but a clear fusion can be seen in the case of the *ser* module (Fig 10.6) and the incomplete *mphE-msrE* gene module in Tn1548 (Fig 10.7), suggesting that these events do occur.

A. pABIR



B. pOXA58_882

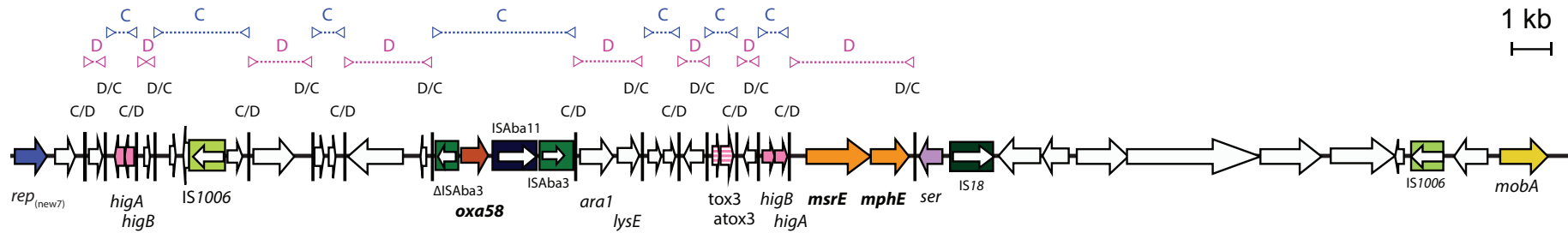


Figure 10.12 Runs of C and D modules. Alternating C and D modules in A. pABIR and B. pOXA58_882. Arrows indicate the extent and orientation of genes and open reading frames. The *oxa58* and *msrE-mphE* genes are in dark orange and orange. The purple arrow is *ser*, and genes encoding proteins showing homology to mobilisation proteins are coloured yellow. Putative toxin/antitoxin genes are pink or pink patterned. Boxes indicate various IS and the identity of each is shown below, and the internal arrows indicates the transposase orientation. *pdif* sites are shown as vertical bars and their orientation is shown above. D and C modules are indicated above by pink and blue dashed lines bounded by arrows. Figure made to scale from GenBank accession numbers EU29428 (pABIR) and CP014479 (pOXA58_882).

10.7 XerC-XerD experiments

In order to begin to understand how these *dif* modules are moving, an experiment was designed to attempt to detect recombination. Two plasmids that have different selectable marker located within *dif* modules were required. The sequence of the 6 bp spacer may be a factor in recombination, so the plasmids were also chosen to share an identical spacer region in at least one of the *pdif* sites present. To make the possible recombination events easier to discern the plasmids used were also selected to contain as few *pdif* sites as possible.

10.7.1 pRCH51-1 and pBAL_204-1

Within the collection of sequenced isolates available in the Hall laboratory, two plasmids fit the above requirements. These were pRCH51-1 (102) (GenBank accession number KT34360) and pBAL_204-1 (235) (GenBank accession number KT946773), and the structures of these two plasmids are shown in Fig 10.13. pRCH51-1 is resistant to tetracycline as it contains the *tet39* determinant in a *dif* module while pBAL_204-1 has the *oxa24* gene in a *dif* module and is resistant to ticarcillin. pRCH51-1 is identical to pRCH52-1 (mentioned in section 10.6.1). The *repA* genes of pRCH51-1 and pBAL_204-1 are 86% identical, with the RepA proteins sharing 90% aa identity. However, the iteron sequences of the two plasmids are identical (5' TAAAACGAGGTTTACCTTGCAT 3'). This indicates that the two plasmids may be incompatible with one another. Utilising the regions of variability, primers were designed to independently detect the *rep* region of each plasmid (Fig 10.13). The combination of RH2530 and RH2531 will amplify an 860 bp region of pRCH51-1 while RH2528 and RH2529 will amplify an 833 bp region of pBAL_204-1. The sequence of the *pdif* sites in both plasmids are shown in Fig 10.13 with identical 6 bp spacer sequences in matching colours. Additionally, plasmids can be distinguished by a unique restriction enzyme site, PstI in pRCH51-1 and DraIII in pBAL_204-1 (indicated by the vertical arrows in Fig 10.13A and B).

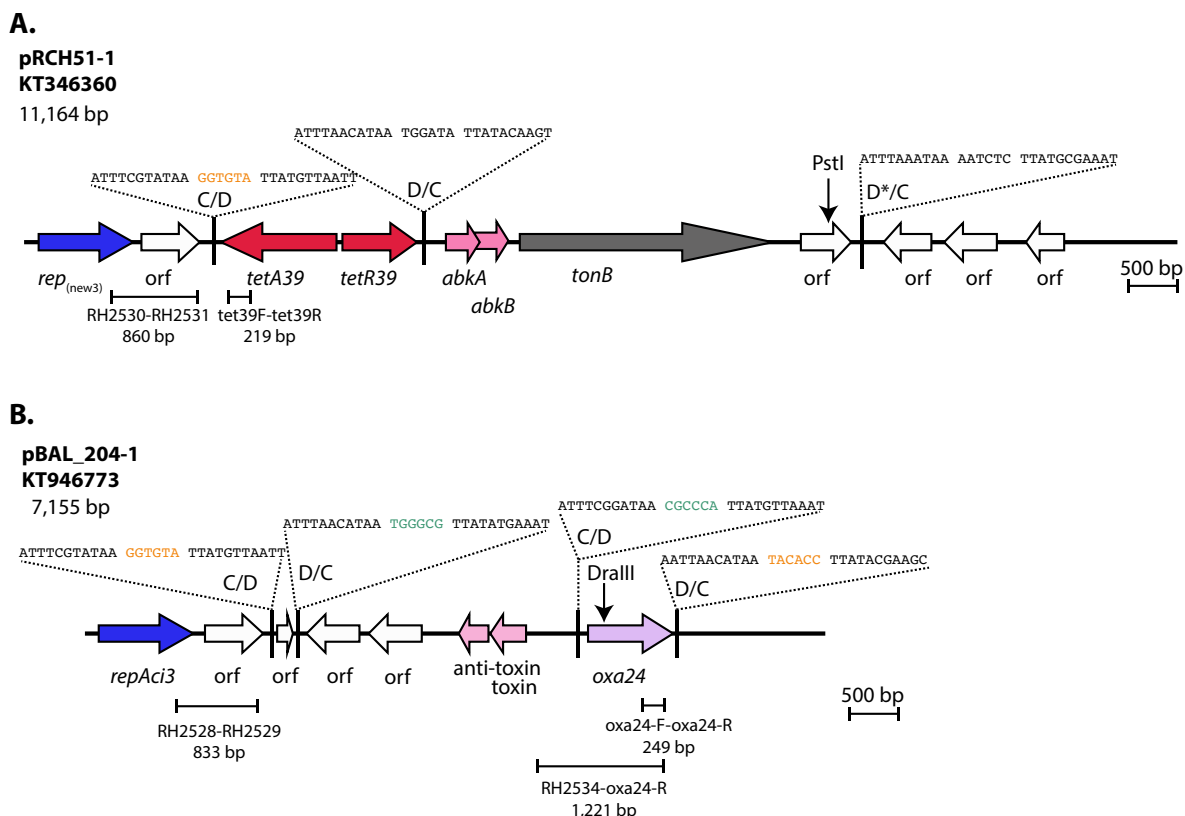


Figure 10.13 Structure of pRCH51-1 and pBAL_204-1. Arrows indicate the extent and orientation of genes and open reading frames. The *rep* genes are blue and the *tetA39* and *tetR39* genes are shown in red. Genes encoding putative toxin/antitoxin modules are shades of pink. The *oxa24* gene is purple. Vertical bars indicate *dif* sites and the sequence and the direction of the sites are shown above. Vertical arrows show enzyme restriction sites. Bars indicate different PCRs and the primers and product sizes are provided. The asterisk associated with a XerD site indicates that it is 10 bp long rather than 11 bp.

10.7.2 pRCH51-1 and pBAL_204-1 in AB307_0294

Electrocompetent AB307_0294 was transformed with RCH51 genomic DNA (2 μ L of 1,585.8 ng/ μ L). Transformants were selected for using tetracycline and 100 potential transformants were present on the 10^{-1} dilution. Five were screened for the presence of the *tetA39* gene (using primers tet39F and tet39R, amplicon size of 219 bp) (Fig 10.13A) to confirm the presence of pRCH51-1. Additionally, AB307_0294 was also transformed with BAL_204 genomic DNA (2 μ L of 1,431 ng/ μ L). Only 30 potential transformants were recovered on plates supplemented with ticarcillin and 5 were screened for the presence of the *oxa24* gene (utilising the primers oxa24-F and oxa24-R to amplify a 249 bp product) (Fig 10.13B) to confirm the presence of

pBAL_204-1. The lower transformation frequency reveals pBAL_204-1 is either less readily transformed into AB307_0294 or is less stable than pRCH51-1.

AB307_0294 containing pBAL_204-1 was then transformed with pRCH51-1 and tetracycline resistant transformants containing pRCH51-1 were selected. 419 tetracycline resistant colonies were detected on the 10^{-1} dilution and 50 of these were screened for additional resistance to ticarcillin. All were susceptible to ticarcillin indicating that they had lost pBAL_204-1, supporting the predicted incompatibility of pBAL_204-1 with pRCH51-1. The reciprocal experiment was performed where AB307_0294 containing pRCH51-1 was transformed with pBAL_204-1, selecting for ticarcillin resistant transformants containing pBAL_204-1. Only three colonies were recovered from the undiluted transformation mix. Upon screening, all colonies were susceptible to tetracycline, indicating that they had not retained pRCH51-1. The lowered yield of transformants recovered from the second experiment could be due to replication incompatibility, but it may not be the only constraint. Indeed, competition between different toxin/anti-toxin systems identified on each of these plasmids (pink arrows in Fig. 10.13A and B) may also be a factor.

The transformation mix of AB307_0294 containing pRCH51-1 (Tc^R) transformed with pBAL_204-1 (TIC^R) was then tested using plates supplemented with tetracycline and ticarcillin to select transformants resistant to both antibiotics. However, no transformants were recovered and the experiment was not repeated due to time limitations.

10.8 ISAjo2 and ISAb32 target *dif* sites

10.8.1 ISAjo2 and relatives

Both copies of ISAjo2-1 (97% aa identity to ISAjo2) in pS30-1 are inserted 5 bp away from a XerC binding site of a *pdif* site (Fig 10.14). The IS also faces in the same direction with respect to the orientation of the *pdif* site in both cases. When the copy of ISAjo2 found in EF102240 (Fig 10.6) was examined here, it was also found to be adjacent to a XerC binding site (Fig 10.14).

pS30-1		XerC		XerD	
CTATTTTGGAGTTACA	AACGC	<u>ATTTCGTATAAGGTGTATTATGTTAAAT</u>		TTAGATAAACTATAGAATC	
←					
CTATTTTGGAGTTACA	AGTGT	<u>ACTTCGTATAACCGCCATTATGTTAAAT</u>		GGGGAATAAAAATACGATCA	
←					
EF102240					
CTATTTTGGAGTTACA	ATCTC	<u>GTTTCGCATAACAGCCATTATGTTAAAT</u>		GCAGGGTATCTAATTTCTA	
←					

Figure 10.14 Alignment of ISAjo2 insertion sites in pS30-1 and EF102240. The bases in green are part of the left IR of ISAjo2 and the arrows indicate the direction of the transposase. *pdif* sites are highlighted in orange and the XerC and XerD binding sites are underlined. Bases in bold represent the target site which is duplicated.

A search of the nucleotide database using ISAjo2 as a query was performed to determine if other relatives of ISAjo2 were also located adjacent to putative *pdif* sites. A number of IS were identified (Table 10.13) with DNA identity to ISAjo2 ranging from 90% to 100%. Most of these IS would be considered to be isoforms of ISAjo2 by ISFinder as the amino acid similarity of their transposase to that of ISAjo2 is 98% or higher. Therefore, they were numbered as variants of ISAjo2 (Table 10.13). The context of these IS were examined and they were always found to be facing the same orientation, 5 or 6 bp away from a putative XerC binding site of a *dif* site (Table 10.13). This indicates that ISAjo2 and its relatives target a specific location in an orientation specific manner and generate a 5 or occasionally 6 bp TSD.

ISAjo2 is currently in the *IS1202* group of the ISNCY (not characterised yet) category in ISFinder. The gene internal to this IS encodes a 441 aa protein (Fig 10.15A) that matches Pfam00665, which is a member of the retroviral integrase superfamily of transposases i.e. a DDE transposase. The IR of this transposon is 24 nucleotides long as shown in Fig 10.15B with 21 bp identical and the bounding ends are 5' TGT 3'. The size of this IR and the sequence of its terminal bases, as well as causing a 5 bp duplication, is similar to those of Tn7 (210), Tn5053 and Tn402 (138) and Tn6019, Tn6021 and Tn6022 (88). A feature of these transposons is the repetition of the IR sequence just within each end of the transposon which delineate the right and left end of the transposon. Almost complete direct duplication and a smaller inverse duplication of the IR sequence can also be found at each end of ISAjo2-1 (black triangles and arrows in Fig 10.15A and B) and their different spacing could differentiate each end.

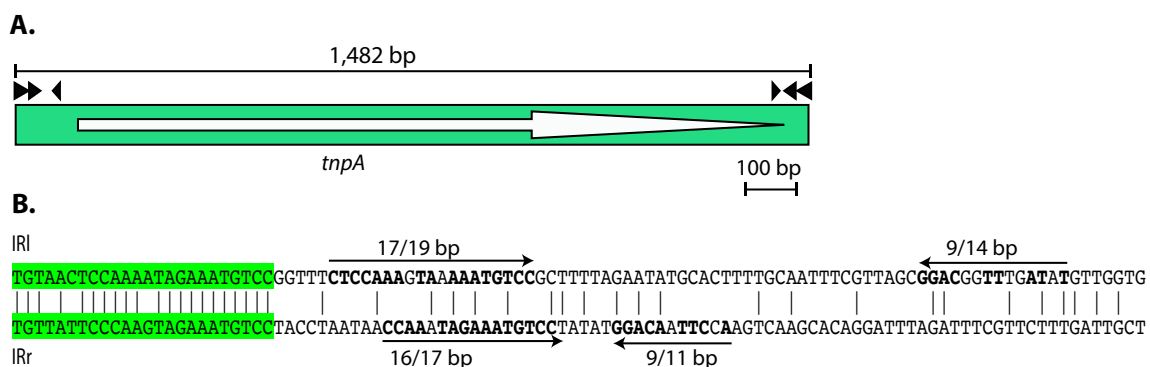


Figure 10.15 IR of ISAjo2-1. A. Schematic of ISAjo2-1. The green box is ISAjo2 and the internal arrow indicates the transposase orientation. The black triangles above indicate the location and the orientation of IR or partial IR repeats. B. The IR of ISAjo2-1 is repeated within each end. The bases highlighted in green are the IR sequences. The arrows above and below indicate the location and extent of repeated sequence matching the IRs. The numbers indicate the identity of the repeat sequence to the IR and the bases in bold are the same as the IR.

The closest relative to ISAjo2 in the *IS1202* group is ISAba32, was also discovered in *A. baumannii* (96), and the location of these IS were examined to determine if they are also found adjacent to *dif* sites.

Table 10.13 Relatives of ISAjo2 in GenBank

Sequence	Number of IS copies	Identity to ISAjo2	ISAjo2 variant	Distance from XerC binding site (bp)	Source	GenBank accession number	Reference
pXBB1-9	1 ¹	100% [1482/1482]	ISAjo2	5 ²	<i>A. johnsonii</i>	CP010351	(64)
pM131-2	1	100% [1482/1482]	ISAjo2	5	<i>Acinetobacter sp.</i>	JX101647	(213)
EF102240	1	96.6% [1432/1482]	ISAjo2-2	5	<i>A. baumannii</i>	EF102240	
pNDM-32	2 ³	94.4% [1399/1482]	ISAjo2-1	5	<i>A. baumannii</i>	LN833432	
pAF-401	3 ³	94.1% [1395/1482]	ISAjo2-1	5	<i>A. baumannii</i>	CP018255	
AF-401	1	94.1% [1395/1482]	ISAjo2-1	5	<i>A. baumannii</i>	CP018254	This study (172)
pS30-1	2 ³	94.1% [1394/1482]	ISAjo2-1	5	<i>A. baumannii</i>	KY617771	
pMMCU2	1	94.1% [1394/1482]	ISAjo2-1	6	<i>A. baumannii</i>	NC_013506	
pNCu2D-2	1 ⁴	92.6% [1373/1483]	ISAjo2-3	6	<i>Acinetobacter sp.</i>	CP015595	
p1AsACE	1 ¹	91.6% [1357/1482]	ISAjo2-4	6	<i>A. schindleri</i>	CP015616	
pAV2	2 ³	90.9% [1349/1484]	ISAjo2-5	5	<i>A. venetianus</i>	DQ278486	
pAHTJS2	1	90.2% [1337/1484]	ISAjo2-6	5	<i>A. haemolyticus</i>	CP018873	

¹Partial copy also present.

²Reported as 6 in ISFinder.

³Copies are within the same sequence and are identical to each other.

⁴SNPs truncate the *tnpA* gene.

Table 10.14 Relatives of ISAb32 in GenBank.

Sequence	Number of IS copies	Identity to ISAb32	ISAb32 variant	Distance from XerC binding site (bp)	Source	GenBank accession number	Reference
pD36-4	1	100% [1482/1482]	ISAb32	5	<i>A. baumannii</i>	CP012956	(96)
pALWED2.1	1	94.0% [1393/1482]	ISAb32-1	5	<i>A. Iwoffii</i>	KX426229	(175)
pALWED3.1	1	92.8% [1375/1482]	ISAb32-2	5	<i>A. Iwoffii</i>	KX528687	(175)
TTH0-4	2	91.8% [1360/1482]	ISAb32-3	5	<i>Acinetobacter sp.</i>	CP012608	
		91.5% [1356/1482]	ISAb32-4	5			
pXBB1-8	1	91.2% [1351/1482]	ISAb32-5	5	<i>A. johnsonii</i>	CP010358	(64)
pHH1107	1 ¹	89.6% [1327/1482]	ISAb32-6	5	Uncultured bacterium	FJ012881	(112)
pHHV216	1 ¹	89.6% [1327/1482]	ISAb32-6	5	Uncultured bacterium	FJ012880	(112)
pM131-2	1 ¹	89.3% [1323/1482]	ISAb32-7	5	<i>Acinetobacter sp.</i>	JX101647	

¹1,481 bp in length.

10.8.2 ISAb32 and close relatives

The sequence of ISAb32 is 73.3% identical to ISAjo2 and the encoded 445 aa transposase is 72% identical to that of ISAjo2. Interestingly, examination of the sequence surrounding the copy of ISAb32 in pD36-4 (GenBank accession number CP012956) (96), revealed that it was also inserted 5 bp to the XerC binding site of a *pdif* site that was not noted previously. ISAb32 is also 1,482 bp in size and created a 5 bp direct duplication upon insertion. The 26 bp inverted repeats with 20 bp identical of ISAb32 are longer than those of ISAjo2 but share homology with those of ISAjo2 at the inner and outermost ends (Fig 10.16).

```
ISAb32 (20/26 bp)
TGTTAGAACCCATTTAAAGTGTCTATATTCTCAC
|||||          ||| ||| ||| ||| ||| |||
TGTTATTCACCAATTAAAATGTCTATGCGAAAAA

ISAjo2-1 vs ISAb32
17/22 bp
TGTTATTCCTAAGTAGAAATGTCTACCTAATAA
||||| ||| ||| ||| ||| ||| ||| |||
TGTTATTCACCAATTAAAATGTCTATGCGAAAAA
```

Figure 10.16 IR sequence of ISAb32 and its relationship to that of ISAjo2-1. The blue highlighted nucleotides are IR sequence of ISAb32 while the green bases are the IR of ISAjo2-1.

Several entries in GenBank contained unnamed IS with sequence identity of 89% or higher to ISAb32 (Table 10.14). In each of these sequences, careful inspection of the surrounding regions revealed that the IS was again located 5 bp away from a previously unidentified XerC binding site.

It appears that the variants of ISAjo2 and ISAb32 both target the same sequence, a XerC binding site, though how this targeting is achieved is unclear. Transposons Tn7 (210), Tn402 (176) and the related transposons Tn6019, Tn6021 and Tn6022 in *A. baumannii* have also been shown to target a specific sequence (88, 217). Target specificity in Tn7 (210) and Tn402 (176)

is provided by an additional protein encoded in the transposon. As the only open reading frame in ISAjo2 and ISAb32 is the transposase, this protein is likely to be performing this role as well.

10.9 Discussion

The finding that the *tet39* determinant and the *msrE-mphE* gene pair are each part of mobile *dif* modules adds to those carrying the *oxa24* and *oxa58* genes (46, 64, 85, 172, 216, 220, 281) bringing the number of antibiotic resistance determinants found in *dif* modules in *Acinetobacter* plasmids to four (bold in Table 10.15). However, the *oxa58* module in pMAD was redefined here (Top line in Fig 10.17).

Table 10.15 *dif* modules in *Acinetobacter* plasmids

<i>dif</i> module	Module size (bp)¹	Module type
<i>tet39A, tet39R</i>	2,001	XerD
<i>msrE-mphE</i>	2,950	XerD
<i>higB, higA</i>	701	XerC
<i>ser</i>	809 ²	XerC
<i>orf</i>	368	XerD
<i>oxa58</i>	3,568 ³	XerC
<i>oxa24</i>	911	XerD
Toxin/antitoxin ⁴	1,369	XerC

dif modules containing antibiotic resistance determinants.

¹Nucleotides between the flanking *pdif* sites.

²When the full module is present.

³This is the distance present in pMAD and this distance varies in other plasmids.

⁴Present in pAB120.

During the analysis of pMAD, pOXA58_882 and pABIR it was noted that the content of the *oxa58* module was quite variable (Fig 10.17 and Table 10.15). In fact, the *oxa58* module was a different size in each of these three plasmids (Table 10.15). Most of this variation can be accounted for by the addition of various IS (location of insertion indicated by the dotted lines in Fig 10.17). However, the left end of the module is missing in pABIR. The *pdif* site on the left and the 27 bp next to it in pMAD and pOXA58_882 have been replaced by IS26 and a

fragment of Tn5393 (Fig 10.17). It is possible that an IS26 located within another module, caused an adjacent deletion, leading to a fusion of two modules. The sequence of *pdif* site on the left in pABIR is consistent with this as it is the first time that there has been variability in the sequence of the internal *dif* site (Table 10.15). Further work is required to analyse all the variants of the *oxa58* module available in GenBank.

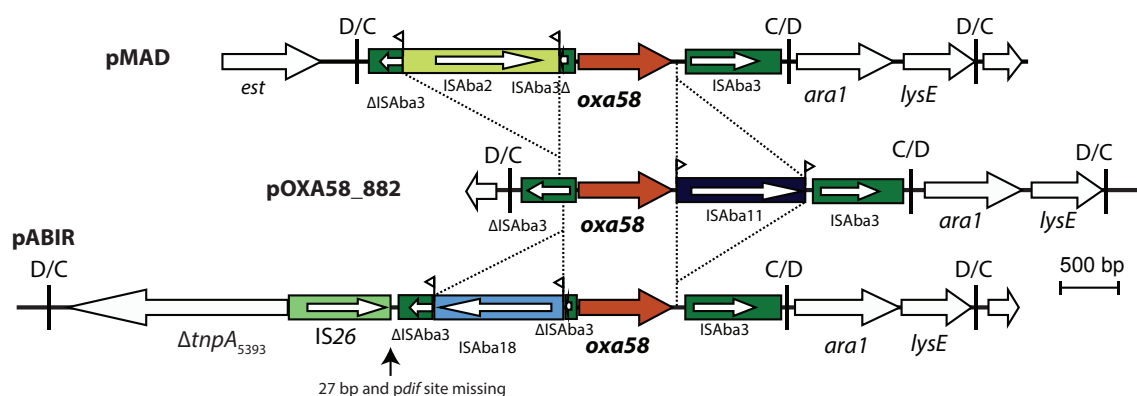


Figure 10.17 Variation in the *oxa58* module. Arrows indicate the extent and orientation of genes and open reading frames. The *oxa58* gene is dark orange. Coloured boxes represent IS, the identity of each is shown below and the internal arrows indicate the transposase orientation. Vertical red and black bars represent Re27 and *pdif* sites, respectively and the orientation of the *pdif* sites are shown above.

Interestingly, S30 is one of the isolates that does not contain a version of AbGRI3 and hence, does not contain the macrolide resistance gene pair *msrE-mphE* in this island. Despite this, S30, like many other *Acinetobacter* isolates, has acquired these genes on plasmids. As the class of antibiotics, macrolides, that the *msrE-mphE* genes confer resistance to are not commonly used to treat infections caused by Gram-negative microorganisms, the presence of these genes, in many different locations and contexts in the *Acinetobacter* genome is curious. There may be selective pressure to maintain these genes in the hospital environment or they may be moving with other, more relevant antibiotic resistance genes.

In addition to the antibiotic resistance genes, a toxin-antitoxin system (*higB*, *higA*) was found in a *dif* module, as was a serine recombinase (Table 10.2). A number of other toxin-antitoxin

gene combinations were identified throughout this investigation (Fig 10.6), as were new modules containing hypothetical proteins, and more work is required to catalogue all of these modules. In most publications and GenBank entries to date, only a limited number of *pdif* sites (or Re27 repeats) have been recorded and presence of further sites in the broader vicinity has not been investigated, likely due to the difficulty in finding them. The thorough analysis of all of the plasmids used in this study enabled some features of this *dif* module mobilisation system to be defined. It was concluded that there are two types of *dif* modules, C and D type modules, which in the general case alternate along the plasmid (Fig 10.12A and B).

Attempts to detect the occurrence of *dif* module recombination was not successful. Due to our increasing understanding of this system and the preference for D and C type modules to alternate, the two plasmids chosen to perform this experiment were not ideal candidates. There is no specific location for the *oxa24* D module to move to in pRCH51-1 that would maintain C and D alternation (Fig 10.13), though perhaps, if two modules moved together, this pattern would be maintained. Ideally, these experiments would be repeated using a host plasmid that only contains a single *pdif* site along with a selectable marker. Experimental investigation of module movement will be needed to understand the mechanism of mobilisation and to demonstrate that the XerC and XerD recombinases are involved.

A pair of *dif* modules were recently found in *Proteus mirabilis* (81) and this represents the first time that *dif* modules have been found outside the *Acinetobacter* genus. One of the modules is related to the smaller *oxa58* module described above and the second is a novel module that includes an *ampC* gene that confers resistance to amoxicillin. In this case, the pair of *dif* modules, *oxa58* with XerD sites internal and *ampC* with XerC sites internal, were reported to be present three times in a single array (81). However, only sequence for one array was

provided for in the GenBank accession KX668205. This configuration clearly should preserve the alternation of a XerC-XerD and a XerD-XerC sites seen in *Acinetobacter* plasmids and again raises the possibility that *dif* modules may usually move in pairs.

It appears that the variants of ISAjo2 and ISAba32 both target the same sequence, a XerC binding site, though how this targeting is achieved is unclear and further investigation of the ISAjo2 group of insertion sequences is warranted. Nonetheless, these IS can serve to pinpoint *pdif* sites that may otherwise be difficult to find.

CHAPTER

11

General discussion

11.1 Overview

In order to explore the diversity of GC2 *A. baumannii* from Singapore a set of 20 XAR GC2 isolates from Singapore General Hospital, isolated over a period of 15 years, were examined. One of these isolates (S35) was also resistant to the last line antibiotic colistin as well as rifampicin, however, as resistance to polymyxin B was not tested, it was unable to be classified as pan resistant. The relationships between the isolates were determined by analysing and assembling all regions containing antibiotic resistance genes in the chromosome or in plasmids, as well as determining ST and KL and the ISAbal profiles of the isolates. The variety of Oxford MLST profiles (Table 11.1) gives an indication of the diversity in these isolates. Seven different STs were detected, two of which were novel. Most of the variation in the ST was due to changes in the *gpi* allele, which was caused by capsule locus replacement (Chapter 3; Table 3.11, Table 11.1). Isolates containing the most recently acquired island AbGRI3 belong to five different STs (Table 11.1), indicating that replacement of the capsule locus has occurred relatively recently. However, the AbGRI3-4 variant was strongly associated with ST195, both in this collection (Table 11.1) and in other sequences isolates (Chapter 7; Fig 7.1).

The majority of the acquired antibiotic resistance genes in this set of GC2 *A. baumannii* were concentrated in one or more of the three known GC2-associated chromosomal resistance islands, AbGRI1, AbGRI2 and AbGRI3. One group of five isolates contained only a version of AbGRI1 and AbGRI2 (Table 11.1). These isolates contain the *aphA6* gene (resistance to amikacin, kanamycin, neomycin) but *TnaphA6* is located on plasmids. The second group of 15 isolates contain a version, or remnant of AbGRI1 and AbGRI2 but also contain AbGRI3, which contains the *armA* gene (resistance to all clinically relevant aminoglycosides) and the presence of this genomic island distinguished this group (Table 11.1). Numerous different structures for variant forms of AbGRI1, AbGRI2 and AbGRI3 were characterised.

Table 11.1 Features of the 20 GC2 isolates from Singapore General Hospital

Isolate	Year	ST (Ox)	KL ¹	OCL ¹	Chromosomal resistance			Plasmids					
					AbGRI			Other <i>oxa23</i> locations	RepAci6				Small plasmids ⁴
									1 ²	2	3	Backbone type	
S21	1996	208	2	1	3	1	-		3	-	3-1	Tn6177, ISAba12-like	RepAci2 (TnaphA6)
S24	2006	208	2	1	5a	1	-		1	+	1-1	-	-
S26	2006	1165	73	1	5b	1	-		1	+	1-2	9.3 kb Tn	-
S29	2007	1165	73	1	5b	1	-		1	+	1-2	9.3 kb Tn	-
S30	2008	369	49	1	1	1	-		4	+	-	Tn2008	pS30-1 (<i>tet39</i> , <i>msrE-mphE</i>)
S31	2008	208	2	1	6	5	5i	AbaR4 (2)	-				RepAci1
S38	2011	208	2	1	1	5	3i	AbaR4 (2)	-				RepAci1
S39	2011	208	2	1	1	5	3i	AbaR4 (2)	-				RepAci1
S27	2007	218	7	1	4	6	2i		2	-	2-1	ISAba1	RepAci1
S22	2004	218	7	1	3	4	2i	AbaR4 (1)	2	-	2-1	-	RepAci1
S23	2004	218	7	1	3	4	2i	AbaR4 (1)	2	-	2-1	-	RepAci1
S25	2006	218	7	1	3	4	6		2	-	2-1	ISAba1	RepAci1
S28	2007	1128	32	1	8	4	1i		2	-	2-1	ISAba1	RepAci1
S34	2009	1166	49	1	4	3	4	Tn2006	2	-	2-2	-	RepAci1
S32	2009	195	3	1	7	2	4		2	-	-	ISAba19	RepAci1 (Tn2006)
S33	2009	195	3	1	7	2	4		-				RepAci1
S40	2011	195	3	1	7	2	4		2	-	2-1	-	RepAci1
S36	2010	195	3	1	7	7	4		2	-	-	ISAba19	RepAci1
S37	2010	195	3	1	7	7	4		-				RepAci1
S35	2010	195	3	1	7	7	4		-				RepAci1

Coloured boxes indicate isolates that share multiple features.

¹KL and OCL provided by Dr J. Kenyon.

²AbGRI1-4, -6, -7 and -8 contain AbaR4, AbaR4Δ, AbaR4 and AbaR4Δ, and Tn2006, respectively.

³The AbaR4 position number is given in respect to the RepAci6 backbone type e.g. 3-1 is backbone 3 and position 1.

⁴The RepAci1 and RepAci2 plasmids are closely related to pAb-G7-1 and pD72-1, respectively.

These variants were often shared by multiple isolates, further subdividing the groups (coloured boxes in Table 11.1). For example, the yellow box around S32, S33 and S40 in Table 11.1 highlights that these isolates contain the same island variants and are likely more closely related to one another than to the other isolates that contain AbGRI3. This relationship is also evident in the phylogeny shown in Chapter 7, Fig 7.5. Most of the variation within AbGRI1 was due to recombination between closely related transposons (Tn6022, Tn6022Δ1 and AbaR4), while in AbGRI2 and AbGRI3 the majority was caused by IS26.

The *oxa23* gene was present in all isolates, but found in multiple locations in the chromosome including in AbGRI1, or in plasmids in the isolates in the collection. As *oxa23* does not have a fixed location, it is likely that the acquisition of this gene is a more recent adaption by antibiotic resistant GC2 *A. baumannii*. This is consistent with the more recent introduction of carbapenems for therapy. In all but one isolate, S30, which had Tn2008A, the *oxa23* gene was located in Tn2006. However, Tn2006 was only found by itself in three isolates (Table 11.2). Most often, Tn2006 was located within AbaR4 or AbaR4Δ (Table 11.2). These two transposons were present in many of the AbGRI1 variants characterised in the isolates that contain AbGRI3 (Table 11.2). However, a subset these AbGRI3-containing isolates also contained an *oxa23*-containing transposon in other chromosomal locations (Table 11.2). New chromosomal positions for AbaR4 and Tn2006 were also detected and positions shared between different isolates indicated their relatedness (orange and green boxes in Table 11.1), in a way that is consistent with the phylogeny, shared IS and other indicators. Assays for these locations may help in tracking related isolates that may carry them.

Table 11.2 Location of *oxa23* containing transposons

Isolate	Context of <i>oxa23</i> in AbGRII	Other chromosomal locations	AbaR4 or Tn2008 on RepAci6 plasmid ¹	Small plasmid
S21			AbaR4, 3-1	
S24			AbaR4, 1-1	
S26, S29			AbaR4, 1-2	
S30			Tn2008	
S25			AbaR4, 2-1	
S22, S23		AbaR4 (1)	AbaR4, 2-1	
S27	AbaR4		AbaR4, 2-1	
S34	AbaR4	Tn2006	AbaR4, 2-2	
S28	Tn2006		AbaR4, 2-1	
S31	AbaR4Δ1	AbaR4 (2)		
S38, S39		AbaR4 (2)		
S32	AbaR4, AbaR4Δ1			Tn2006
S33, S35, S36, S37	AbaR4, AbaR4Δ1			
S40	AbaR4, AbaR4Δ1		AbaR4, 2-1	

¹ The AbaR4 position number is given in respect to the RepAci6 backbone type e.g. 3-1 is backbone 3 and position 1.

Plasmids also played an important role in the carriage of antibiotic resistance genes in this collection. The *oxa23* and *aphA6* genes were both found on large conjugative RepAci6 plasmids. Fourteen of the 20 isolates examined had a RepAci6 plasmid. Twelve of these carried AbaR4 or Tn2008A transposons containing *oxa23* and, of these, four also included Tn*aphA6* (Table 11.1). Surprisingly, the AbaR4 transposon was in five different positions in the RepAci6 plasmid backbone indicating it had been acquired repeatedly rather than once generating a single plasmid that spread through the population (Table 11.1 and Table 11.2). In one isolate, Tn2006 was found in a small RepAci1 plasmid, and Tn*aphA6* was located in a small RepAci2 plasmid in another isolate.

A novel small plasmid pS30-1 containing additional antibiotic resistance genes was found in a single isolate. However, these genes were not located in transposons or known mobile elements. Instead, the *tet39* resistance determinant along with the *msrE* and *mphE* macrolide resistance genes were found to be located within *dif* modules that were defined as part of this work and discussed below. Isolate S30 lies between the two main groups in the phylogeny,

appearing intermediate between them. It also includes a unique RepAci6 plasmid carrying Tn2008.

11.2 Acquisition and evolution of antibiotic resistance islands in GC2 *A. baumannii*

Over the last 30-40 years, the GC2 global clone has accrued sets of genes involved in antibiotic resistance in three blocks. The oldest GC2 isolate for which data is available, A320 from 1982, already contained AbGRI1 (190) and AbGRI2 (24). These two resistance islands conferred resistance to most older antibiotics (to the left of the bold line in Fig 11.1). Additional work has shown that these regions were brought into the GC2 chromosome from plasmids. The source of AbGRI1 has been shown recently to be an *A. baumannii* plasmid (94), while the ancestral version of AbGRI2 was identified on an M1 plasmid from *S. marcescens* (22).

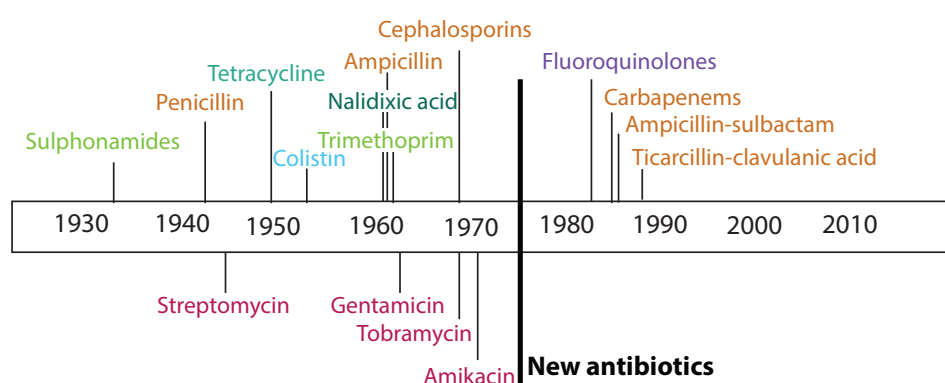


Figure 11.1 Timeline of when antibiotics suitable for Gram-negative pathogens were first used clinically. The black bar separates the new and old antibiotics.

The AbGRI3 island was acquired much later. The first report of this island (277) was published at the commencement of this project but a complete analysis of this region was performed here. Analysis of the set of isolates from Singapore General Hospital revealed that AbGRI3 had been acquired by 2004 (Table 11.1). However, further examination of *A. baumannii* from the early 2000's may allow this date to be further refined. Like AbGRI2, the surrounding TSD suggests

that the AbGRI3 region was acquired in a single-step. There are a number of related *armA*-carrying transposons in Enterobacterial plasmids, again indicating a potential source of the genes in this transposon from outside *Acinetobacter*.

The AbGRI1 and AbGRI2 islands have not remained unchanged over the last three decades. Although versions or remnants of these islands are always present in antibiotic resistant GC2 isolates, many different variants have been detected. In particular, AbGRI2 has been subject to a large amount of change due the presence of multiple copies of IS26. The structure of AbGRI2 in A320 is much larger than variants detected in most recent isolates. In fact, modern variants of AbGRI2, particularly those in isolates from Asia and Australia, have evolved from an important variant, AbGRI2-1(188), a structure unlikely to arise on two independent occasions. Hence, there appears to have been a constriction of the GC2 population but then a clade has expanded from the cell containing the AbGRI2-1 island structure and thereafter AbGRI2-1 has diverged. AbGRI2-1 is smaller than AbGRI2-0 (A320) as the chloramphenicol resistance gene *catA1* and mercury resistance region are missing (24). As *Acinetobacter* are intrinsically resistant to chloramphenicol, this loss is without consequence. AbGRI2-1 was first detected in isolates from the east coast of Australia in 1999 and, in this project, AbGRI2-1 was found in a 1996 isolate from Singapore General Hospital, indicating that it arose even earlier.

The AbGRI2 region of the isolates containing AbGRI3 studied here, was always smaller than AbGRI2-1 but based on the deviations in all the variants, the cell that gained AbGRI3 must have originally contained AbGRI2-1. Deletions in AbGRI2 have removed antibiotic resistance genes and in some cases, a single IS26 in the AbGRI2 region of the chromosome is all that remains. The loss of resistance genes from AbGRI2, particularly loss of the aminoglycoside resistance gene *aacC1*, has had no impact on the overall phenotype of an AbGRI3 containing,

and hence *armA* containing isolate. In addition, due to the resistance spectrum of *armA*, there is no need for isolates containing this gene to acquire *aphA6*, and this is reflected in the gene distribution in the Singapore isolates (Table 11.1).

The AbGRI1 island had acquired the *oxa23* carbapenem resistance gene in Tn2006 at a number of different positions. Tn2006 located within Tn6022 as AbaR4 as well as in Tn6022Δ1 as AbaR4Δ1 was seen (Table 11.2). It is likely that recombination between a copy of AbaR4 and the pre-existing Tn6022 or Tn6022Δ1 has facilitated the acquisition of the *oxa23* transposon by AbGRI1. Tn2006 was also detected by itself in one isolate (Table 11.2) and in addition to the Tn2006 position in AbGRI1-2 seen in Australian isolates (191), is the second site of this transposon in AbGRI1. Thus, the AbGRI1 island appears to be playing an important role in carbapenem resistance as well as conferring resistance to tetracycline, sulphonamides and streptomycin.

11.3 Plasmids with resistance genes

The plasmids in clinical *A. baumannii* isolates are contributing significantly to antibiotic resistance and its spread. The larger conjugative RepAci6 plasmids are playing a key role in resistance to carbapenems and amikacin which became available for clinical use later (Fig 11.1) as many carry *oxa23* and/or *aphA6*. These antibiotics are important in the treatment of antibiotic resistant *A. baumannii*. Both the Tn*aphA6* and the transposons harbouring *oxa23*, Tn2006, Tn2008 and AbaR4, have been acquired multiple times by RepAci6 plasmids.

Large conjugative plasmids were not the only ones to contribute to antibiotic resistance in these isolates. Within the collection, a small RepAci1 plasmid resembling pAb-G7-1 (97) had acquired Tn2006, while a RepAci2 plasmid related to pD72-1 (192) had gained Tn*aphA6*

(Table 11.1). A small plasmid from isolate S30 contained two sets of antibiotic resistance genes, the *tet39* tetracycline resistance determinant and *msrE-mphE* macrolide resistance genes (Table 11.1). These genes were located in *dif* modules, a novel type of mobile element.

11.4 Plasmid evolution and movement

The backbones of the RepAci6 plasmids analysed in this project were quite closely related but harboured some differences. Among the 13 RepAci6 plasmids assembled in this project, four different backbone types were detected, doubling the known number of backbone types (Table 11.1). Examination of plasmids in the public database increased this to 10 backbone types. The backbones of the variants have large regions that share high levels of homology, particularly over the transfer and replication region, but contain short segments of diverged sequence (as little as 80% DNA identity) or have gained additional short regions of DNA including a fourth repeat or a *parB* gene. As more and more isolates are sequenced this number is likely to continue to increase. In addition to antibiotic resistance genes, some of the RepAci6 plasmids detected contain genes predicted to be involved in copper resistance (Tn6177, Table 11.1).

The RepAci6 plasmid backbone has all the required genes for self-transmission. However, when tested, three plasmids that appeared to have all transfer genes intact failed to conjugate. Further work is required to understand the transfer deficiency of these plasmids.

Some RepAci6 plasmids are cryptic, and two of the plasmids characterised from the isolates from Singapore contained no antibiotic resistance genes. However, these conjugative plasmids can still contribute to the spread of antibiotic resistance via their role in mobilising small plasmids that contain antibiotic resistance genes and lack any conjugation or mobilisation genes. The small plasmid characterised here, pS32-1 containing Tn2006, was mobilised by a

co-resident conjugative RepAci6 plasmid and this is the first example of *in trans* mobilisation in *Acinetobacter*. This enabled the identification of a putative *oriT* in both plasmids though these will need to be experimentally confirmed. A number of small non-conjugative plasmids have been found to harbour antibiotic resistance genes in *Acinetobacter*, including two others characterised in this project and the importance of these small plasmids in the spread of antibiotic resistance is likely much greater than first-thought.

11.5 *dif* modules

One small plasmid from a Singapore isolate contains the *tet39* determinant and the *msrE-mphE* genes in novel mobile *dif* modules. The *oxa24* gene had already been detected in a *dif* module but in this work another carbapenem resistance gene, *oxa58*, was also shown to be in a *dif* module. Analysis here expanded the number of different toxin/antitoxin systems located within *dif* modules and many hypothetical proteins were also found. The mechanism of movement of these modules and how they are exploiting the chromosomal recombinases XerC and XerD has yet to be elucidated, but features of this system have now been highlighted. In particular, there are two types of *dif* modules, C and D, and, where multiple modules are present in a sequence, these two module types usually alternate. So far, this system appears to be specific to *Acinetobacter* species.

11.6 Globalisation

As each resistance island, AbGRI1, AbGRI2 and AbGRI3, is defined by its specific location in the chromosome, they were each initially acquired by a single cell. The progeny of these ancestor have been successful in spreading throughout the globe, though the islands have also evolved *in situ*. With the current lack of strains available from prior to 1982, currently there is no way of knowing if AbGRI1 or AbGRI2 entered the chromosome first.

The third island was acquired by a cell that already contained both AbGRI1 and AbGRI2. Evidence so far suggests that it was acquired in the early 2000's, likely in Asia. AbGRI3 containing isolates have since been detected throughout Asia, America and sporadically in Sweden and Australia. In addition, the AbGRI3-4 variant is unlikely have to occurred in more than one cell as it is an unusual deletion event. Here, this particular variant of AbGRI3 was observed in isolates from Singapore, South Korea, India and Sweden, indicating the globalisation of this new lineage of the GC2 clone.

11.7 Links to Australian isolates

Specific features linked some the isolates from Singapore General Hospital to isolates from the east coast of Australia as well as to those from Adelaide and Melbourne. The isolate S21 had AbGRI2-1, as seen in isolates from the east coast of Australia, and shared 11 of the 12 common ISAbal positions present in all the east coast isolates. This indicated that S21 and isolates from this area of Australia have a common origin. Another set of isolates from Singapore General Hospital had the same deletion in AbGRI1 as isolates from Melbourne and Adelaide. This link was further supported by their additional shared set of ISAbal positions. In future, it should be possible to compare a larger group, such as including the positions of ISAbal from Wright *et al.* (277), and this will assist in tracking specific GC2 lineages as they spread.

11.8 Future directions

This in-depth analysis of antibiotic resistance regions and plasmids will underpin larger studies of GC2 *A. baumannii* from Singapore and across South-east Asia. The knowledge gained through this project about the contents, location and forms of the three chromosomal resistance islands as well as plasmid variants can now be fed into bioinformatic pipelines that should

quicken future analysis. The GC2 populations in Australia, Asia and America appear to be mainly of the AbGRI2-1 lineage. It would be interesting to examine *A. baumannii* isolates from Europe and Africa in detail to determine if they are also part of this lineage or if other lineages have been successful in these regions. However, at present genomes from these regions are scarce, a situation that should be rectified.

As the number of *A. baumannii* isolates sequenced has increased, so has the number and diversity of plasmid sequences. There is a real need for the typing system for *Acinetobacter* plasmids to be extended to include all of these new sequence types. As the biology of plasmids from *Acinetobacter* is so different from that of plasmids from Enterobacteriaceae, experimental studies determining what genes and regions are absolutely required for plasmid maintenance and/or transfer are needed. In addition, mobilisation of small plasmids needs to be further studied. Further research into understanding *dif* modules, particularly into their mechanism of movement, would also be important.

REFERENCES

1. Adams MD, Bishop B, Wright MS. 2016. Quantitative assessment of insertion sequence impact on bacterial genome architecture. *Microbial Genomics* 2:e000062.
2. Adams MD, Chan ER, Molyneaux ND, Bonomo RA. 2010. Genomewide analysis of divergence of antibiotic resistance determinants in closely related isolates of *Acinetobacter baumannii*. *Antimicrob Agents Chemother* 54:3569-77.
3. Adams MD, Goglin K, Molyneaux N, Hujer KM, Lavender H, Jamison JJ, MacDonald IJ, Martin KM, Russo T, Campagnari AA, Hujer AM, Bonomo RA, Gill SR. 2008. Comparative genome sequence analysis of multidrug-resistant *Acinetobacter baumannii*. *J Bacteriol* 190:8053-64.
4. Adams MD, Nickel GC, Bajaksouzian S, Lavender H, Murthy AR, Jacobs MR, Bonomo RA. 2009. Resistance to colistin in *Acinetobacter baumannii* associated with mutations in the PmrAB two-component system. *Antimicrob Agents Chemother* 53:3628-34.
5. Adelowo OO, Fagade OE. 2009. The tetracycline resistance gene *tet39* is present in both Gram-negative and Gram-positive bacteria from a polluted river, Southwestern Nigeria. *Lett Appl Microbiol* 48:167-72.
6. Agerso Y, Guardabassi L. 2005. Identification of Tet 39, a novel class of tetracycline resistance determinant in *Acinetobacter* spp. of environmental and clinical origin. *J Antimicrob Chemother* 55:566-9.
7. Ang GY, Yu CY, Cheong YM, Yin WF, Chan KG. 2016. Emergence of ST119 *Acinetobacter pittii* co-harboring NDM-1 and OXA-58 in Malaysia. *Int J Antimicrob Agents* 47:168-9.
8. Anstey NM, Currie BJ, Hassell M, Palmer D, Dwyer B, Seifert H. 2002. Community-acquired bacteremic *Acinetobacter* pneumonia in tropical Australia is caused by diverse strains of *Acinetobacter baumannii*, with carriage in the throat in at-risk groups. *J Clin Microbiol* 40:685-6.
9. Antunes LC, Visca P, Towner KJ. 2014. *Acinetobacter baumannii*: evolution of a global pathogen. *Pathog Dis* 71:292-301.
10. Arthur A, Sherratt D. 1979. Dissection of the transposition process: a transposon-encoded site-specific recombination system. *Mol Gen Genet* 175:267-74.
11. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455-77.
12. Bartual SG, Seifert H, Hippler C, Luzon MA, Wisplinghoff H, Rodriguez-Valera F. 2005. Development of a multilocus sequence typing scheme for characterization of clinical isolates of *Acinetobacter baumannii*. *J Clin Microbiol* 43:4382-90.
13. Baumann P. 1968. Isolation of *Acinetobacter* from soil and water. *Journal of Bacteriology* 96:39-42.
14. Bayuga S, Zeana C, Sahni J, Della-Latta P, el-Sadr W, Larson E. 2002. Prevalence and antimicrobial patterns of *Acinetobacter baumannii* on hands and nares of hospital personnel and patients: the iceberg phenomenon again. *Heart Lung* 31:382-90.
15. Beijerinck M. 1911. Pigmenten als oxydatieproducten gevormd door bacterien. *Versl Koninklijke Akad Wetensch Amsterdam* 19:1092-1103.
16. Bennett PM. 2008. Plasmid encoded antibiotic resistance: acquisition and transfer of antibiotic resistance genes in bacteria. *Br J Pharmacol* 153 Suppl 1:S347-57.

17. Bentley DR, Balasubramanian S, Swerdlow HP, Smith GP, Milton J, Brown CG, Hall KP, Evers DJ, Barnes CL, Bignell HR, Boutell JM, Bryant J, Carter RJ, Keira Cheetham R, Cox AJ, Ellis DJ, Flatbush MR, Gormley NA, Humphray SJ, Irving LJ, Karbelashvili MS, Kirk SM, Li H, Liu X, Maisinger KS, Murray LJ, Obradovic B, Ost T, Parkinson ML, Pratt MR, Rasolonjatovo IM, Reed MT, Rigatti R, Rodighiero C, Ross MT, Sabot A, Sankar SV, Scally A, Schroth GP, Smith ME, Smith VP, Spiridou A, Torrance PE, Tzonev SS, Vermaas EH, Walter K, Wu X, Zhang L, Alam MD, Anastasi C, et al. 2008. Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* 456:53-9.
18. Berger B, Haas D. 2001. Transposase and cointegrase: specialized transposition proteins of the bacterial insertion sequence IS21 and related elements. *Cell Mol Life Sci* 58:403-19.
19. Bergogne-Berezin E, Towner KJ. 1996. *Acinetobacter* spp. as nosocomial pathogens: microbiological, clinical, and epidemiological features. *Clin Microbiol Rev* 9:148-65.
20. Bertini A, Poirel L, Bernabeu S, Fortini D, Villa L, Nordmann P, Carattoli A. 2007. Multicopy *bla*_{OXA-58} gene as a source of high-level resistance to carbapenems in *Acinetobacter baumannii*. *Antimicrob Agents Chemother* 51:2324-8.
21. Bertini A, Poirel L, Mugnier PD, Villa L, Nordmann P, Carattoli A. 2010. Characterization and PCR-based replicon typing of resistance plasmids in *Acinetobacter baumannii*. *Antimicrob Agents Chemother* 54:4168-77.
22. Blackwell GA, Hamidian M, Hall RM. 2016. IncM plasmid R1215 is the source of chromosomally located regions containing multiple antibiotic resistance genes in the globally disseminated *Acinetobacter baumannii* GC1 and GC2 clones. *mSphere* 1:e00117-16.
23. Blackwell GA, Holt KE, Bentley SD, Hsu LY, Hall RM. 2017. Variants of AbGRI3 carrying the *armA* gene in extensively antibiotic-resistant *Acinetobacter baumannii* from Singapore. *J Antimicrob Chemother* 72:1031-1039.
24. Blackwell GA, Nigro SJ, Hall RM. 2015. Evolution of AbGRI2-0, the progenitor of the AbGRI2 resistance island in global clone 2 of *Acinetobacter baumannii*. *Antimicrob Agents Chemother* 60:1421-9.
25. Bonnin RA, Cuzon G, Poirel L, Nordmann P. 2013. Multidrug-resistant *Acinetobacter baumannii* clone, France. *Emerg Infect Dis* 19:822-3.
26. Bonnin RA, Rotimi VO, Al Hubail M, Gasiorowski E, Al Sweih N, Nordmann P, Poirel L. 2013. Wide dissemination of GES-type carbapenemases in *Acinetobacter baumannii* isolates in Kuwait. *Antimicrob Agents Chemother* 57:183-8.
27. Bou G, Martinez-Beltran J. 2000. Cloning, nucleotide sequencing, and analysis of the gene encoding an AmpC β -lactamase in *Acinetobacter baumannii*. *Antimicrob Agents Chemother* 44:428-32.
28. Bou G, Oliver A, Martinez-Beltran J. 2000. OXA-24, a novel class D beta-lactamase with carbapenemase activity in an *Acinetobacter baumannii* clinical strain. *Antimicrob Agents Chemother* 44:1556-61.
29. Boucher HW, Talbot GH, Bradley JS, Edwards JE, Gilbert D, Rice LB, Scheld M, Spellberg B, Bartlett J. 2009. Bad Bugs, No Drugs: No ESKAPE! An Update from the Infectious Diseases Society of America. *Clinical Infectious Diseases* 48:1-12.
30. Bouvet PJM, Grimont PAD. 1986. Taxonomy of the genus *Acinetobacter* with the recognition of *Acinetobacter baumannii* sp. nov., *Acinetobacter haemolyticus* sp. nov., *Acinetobacter johnsonii* sp. nov., and *Acinetobacter junii* sp. nov. and emended descriptions of *Acinetobacter calcoaceticus* and *Acinetobacter lwoffii*. *International Journal of Systematic and Evolutionary Microbiology* 36:228-240.

31. Brown S, Young HK, Amyes SG. 2005. Characterisation of OXA-51, a novel class D carbapenemase found in genetically unrelated clinical strains of *Acinetobacter baumannii* from Argentina. Clin Microbiol Infect 11:15-23.
32. Cain AK, Liu X, Djordjevic SP, Hall RM. 2010. Transposons related to Tn1696 in IncHI2 plasmids in multiply antibiotic resistant *Salmonella enterica* serovar Typhimurium from Australian animals. Microb Drug Resist 16:197-202.
33. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: architecture and applications. BMC Bioinformatics 10:421.
34. Carver TJ, Rutherford KM, Berriman M, Rajandream M-A, Barrell BG, Parkhill J. 2005. ACT: the Artemis comparison tool. Bioinformatics 21:3422-3423.
35. Castro-Jaimes S, Salgado-Camargo AD, Grana-Miraglia L, Lozano L, Bocanegra-Ibarias P, Volkow-Fernandez P, Silva-Sanchez J, Castillo-Ramirez S, Cevallos MA. 2016. Complete Genome Sequence of a Multidrug-Resistant *Acinetobacter baumannii* Isolate Obtained from a Mexican Hospital (Sequence Type 422). Genome Announc 4.
36. Chen CH, Young TG, Huang CC. 2006. Predictive biomarkers for drug-resistant *Acinetobacter baumannii* isolates with *bla*(TEM-1), AmpC-type *bla* and integrase 1 genotypes. J Microbiol Immunol Infect 39:372-9.
37. Chin CS, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J. 2013. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. Nat Methods 10:563-9.
38. Chung DR, Song JH, Kim SH, Thamlikitkul V, Huang SG, Wang H, So TM, Yasin RM, Hsueh PR, Carlos CC, Hsu LY, Buntaran L, Lalitha MK, Kim MJ, Choi JY, Kim SI, Ko KS, Kang CI, Peck KR. 2011. High prevalence of multidrug-resistant nonfermenters in hospital-acquired pneumonia in Asia. Am J Respir Crit Care Med 184:1409-17.
39. Colloms SD, Sykora P, Szatmari G, Sherratt DJ. 1990. Recombination at *ColE1 cer* requires the *Escherichia coli xerC* gene product, a member of the lambda integrase family of site-specific recombinases. J Bacteriol 172:6973-80.
40. Compeau PE, Pevzner PA, Tesler G. 2011. How to apply de Bruijn graphs to genome assembly. Nat Biotechnol 29:987-91.
41. Cornet F, Mortier I, Patte J, Louarn JM. 1994. Plasmid pSC101 harbors a recombination site, *psi*, which is able to resolve plasmid multimers and to substitute for the analogous chromosomal *Escherichia coli* site *dif*. J Bacteriol 176:3188-95.
42. Corvec S, Caroff N, Espaze E, Giraudeau C, Drugeon H, Reynaud A. 2003. AmpC cephalosporinase hyperproduction in *Acinetobacter baumannii* clinical strains. J Antimicrob Chemother 52:629-35.
43. Cosgaya C, Mari-Almirall M, Van Assche A, Fernandez-Orth D, Mosqueda N, Telli M, Huys G, Higgins PG, Seifert H, Lievens B, Roca I, Vila J. 2016. *Acinetobacter dijkschoorniae* sp. nov., a member of the *Acinetobacter calcoaceticus*-*Acinetobacter baumannii* complex mainly recovered from clinical samples in different countries. Int J Syst Evol Microbiol 66:4105-4111.
44. Coyne S, Courvalin P, Perichon B. 2011. Efflux-mediated antibiotic resistance in *Acinetobacter* spp. Antimicrob Agents Chemother 55:947-53.
45. Craig NL. 1991. Tn7: a target site-specific transposon. Molecular Microbiology 5:2569-2573.
46. D'Andrea MM, Giani T, D'Arezzo S, Capone A, Petrosillo N, Visca P, Luzzaro F, Rossolini GM. 2009. Characterization of pABVA01, a plasmid encoding the OXA-24 carbapenemase from Italian isolates of *Acinetobacter baumannii*. Antimicrob Agents Chemother 53:3528-33.

47. Damier-Piolle L, Magnet S, Bremont S, Lambert T, Courvalin P. 2008. AdeIJK, a resistance-nodulation-cell division pump effluxing multiple antibiotics in *Acinetobacter baumannii*. *Antimicrob Agents Chemother* 52:557-62.
48. Dark MJ. 2013. Whole-genome sequencing in bacteriology: state of the art. *Infect Drug Resist* 6:115-23.
49. del Solar G, Giraldo R, Ruiz-Echevarria MJ, Espinosa M, Diaz-Orejas R. 1998. Replication and control of circular bacterial plasmids. *Microbiol Mol Biol Rev* 62:434-64.
50. Di Nocera PP, Rocco F, Giannouli M, Triassi M, Zarrilli R. 2011. Genome organization of epidemic *Acinetobacter baumannii* strains. *BMC Microbiol* 11:224.
51. Diancourt L, Passet V, Nemec A, Dijkshoorn L, Brisse S. 2010. The population structure of *Acinetobacter baumannii*: expanding multiresistant clones from an ancestral susceptible genetic pool. *PLoS One* 5:e10034.
52. Dijkshoorn L, Aucken H, Gerner-Smidt P, Janssen P, Kaufmann ME, Garaizar J, Ursing J, Pitt TL. 1996. Comparison of outbreak and nonoutbreak *Acinetobacter baumannii* strains by genotypic and phenotypic methods. *J Clin Microbiol* 34:1519-25.
53. Doi Y, Adams JM, Yamane K, Paterson DL. 2007. Identification of 16S rRNA methylase-producing *Acinetobacter baumannii* clinical strains in North America. *Antimicrob Agents Chemother* 51:4209-10.
54. Doi Y, Arakawa Y. 2007. 16S ribosomal RNA methylation: emerging resistance mechanism against aminoglycosides. *Clin Infect Dis* 45:88-94.
55. Dolejska M, Villa L, Poirel L, Nordmann P, Carattoli A. 2013. Complete sequencing of an IncHI1 plasmid encoding the carbapenemase NDM-1, the ArmA 16S RNA methylase and a resistance-nodulation-cell division/multidrug efflux pump. *J Antimicrob Chemother* 68:34-9.
56. Dorsey CW, Tomaras AP, Actis LA. 2006. Sequence and organization of pMAC, an *Acinetobacter baumannii* plasmid harboring genes involved in organic peroxide resistance. *Plasmid* 56:112-23.
57. Dortet L, Bonnin RA, Bernabeu S, Escaut L, Vittecoq D, Girlich D, Imanci D, Fortineau N, Naas T. 2016. First Occurrence of OXA-72-Producing *Acinetobacter baumannii* in Serbia. *Antimicrob Agents Chemother* 60:5724-30.
58. Dortet L, Poirel L, Nordmann P. 2014. Worldwide dissemination of the NDM-type carbapenemases in Gram-negative bacteria. *Biomed Res Int* 2014:249856.
59. Edwards DJ, Holt KE. 2013. Beginner's guide to comparative bacterial genome analysis using next-generation sequence data. *Microbial Informatics and Experimentation* 3:2.
60. Eid J, Fehr A, Gray J, Luong K, Lyle J, Otto G, Peluso P, Rank D, Baybayan P, Bettman B, Bibillo A, Bjornson K, Chaudhuri B, Christians F, Cicero R, Clark S, Dalal R, deWinter A, Dixon J, Foquet M, Gaertner A, Hardenbol P, Heiner C, Hester K, Holden D, Kearns G, Kong X, Kuse R, Lacroix Y, Lin S, Lundquist P, Ma C, Marks P, Maxham M, Murphy D, Park I, Pham T, Phillips M, Roy J, Sebra R, Shen G, Sorenson J, Tomaney A, Travers K, Trulson M, Vieceli J, Wegener J, Wu D, Yang A, Zaccarin D, et al. 2009. Real-Time DNA Sequencing from Single Polymerase Molecules. *Science* 323:133-138.
61. Evans BA, Amyes SG. 2014. OXA β -lactamases. *Clin Microbiol Rev* 27:241-63.
62. Fang Y, Quan J, Hua X, Feng Y, Li X, Wang J, Ruan Z, Shang S, Yu Y. 2016. Complete genome sequence of *Acinetobacter baumannii* XH386 (ST208), a multi-drug resistant bacteria isolated from pediatric hospital in China. *Genom Data* 7:269-74.

63. Farrugia DN, Elbourne LD, Hassan KA, Eijkelkamp BA, Tetu SG, Brown MH, Shah BS, Peleg AY, Mabbutt BC, Paulsen IT. 2013. The complete genome and phenome of a community-acquired *Acinetobacter baumannii*. PLoS One 8:e58628.
64. Feng Y, Yang P, Wang X, Zong Z. 2016. Characterization of *Acinetobacter johnsonii* isolate XBB1 carrying nine plasmids and encoding NDM-1, OXA-58 and PER-1 by genome sequencing. J Antimicrob Chemother 71:71-5.
65. Fleischmann RD, Adams MD, White O, Clayton RA, Kirkness EF, Kerlavage AR, Bult CJ, Tomb JF, Dougherty BA, Merrick JM, et al. 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. Science 269:496-512.
66. Flusberg BA, Webster D, Lee J, Travers K, Olivares E, Clark TA, Korlach J, Turner SW. 2010. Direct detection of DNA methylation during single-molecule, real-time sequencing. Nat Methods 7:461-5.
67. Fournier PE, 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 JM. 2006. Comparative genomics of multidrug resistance in *Acinetobacter baumannii*. PLoS Genet 2:e7.
68. Francia MV, Varsaki A, Garcillan-Barcia MP, Latorre A, Drainas C, de la Cruz F. 2004. A classification scheme for mobilization regions of bacterial plasmids. FEMS Microbiol Rev 28:79-100.
69. Frost LS, Leplae R, Summers AO, Toussaint A. 2005. Mobile genetic elements: the agents of open source evolution. Nat Rev Microbiol 3:722-32.
70. Fu Y, Zhou J, Zhou H, Yang Q, Wei Z, Yu Y, Li L. 2010. Wide dissemination of OXA-23-producing carbapenem-resistant *Acinetobacter baumannii* clonal complex 22 in multiple cities of China. J Antimicrob Chemother 65:644-50.
71. Galani I, Kontopidou F, Souli M, Rekatsina PD, Koratzanis E, Deliolas J, Giamarellou H. 2008. Colistin susceptibility testing by Etest and disk diffusion methods. Int J Antimicrob Agents 31:434-9.
72. Galimand M, Courvalin P, Lambert T. 2003. Plasmid-mediated high-level resistance to aminoglycosides in *Enterobacteriaceae* due to 16S rRNA methylation. Antimicrob Agents Chemother 47:2565-71.
73. Galimand M, Sabtcheva S, Courvalin P, Lambert T. 2005. Worldwide disseminated *armA* aminoglycoside resistance methylase gene is borne by composite transposon Tn1548. Antimicrob Agents Chemother 49:2949-53.
74. Gallagher LA, Ramage E, Weiss EJ, Radey M, Hayden HS, Held KG, Huse HK, Zurawski DV, Brittnacher MJ, Manoil C. 2015. Resources for genetic and genomic analysis of emerging pathogen *Acinetobacter baumannii*. J Bacteriol 197:2027-35.
75. Garcia-Garmendia JL, Ortiz-Leyba C, Garnacho-Montero J, Jimenez-Jimenez FJ, Perez-Paredes C, Barrero-Almodovar AE, Gili-Miner M. 2001. Risk factors for *Acinetobacter baumannii* nosocomial bacteremia in critically ill patients: a cohort study. Clin Infect Dis 33:939-46.
76. Gebreyes WA, Altier C. 2002. Molecular characterization of multidrug-resistant *Salmonella enterica* subsp. *enterica* serovar Typhimurium isolates from swine. J Clin Microbiol 40:2813-22.
77. Gerner-Smidt P. 1992. Ribotyping of the *Acinetobacter calcoaceticus*-*Acinetobacter baumannii* complex. Journal of Clinical Microbiology 30:2680-2685.
78. Gerner-Smidt P, Tjernberg I, Ursing J. 1991. Reliability of phenotypic tests for identification of *Acinetobacter* species. Journal of Clinical Microbiology 29:277-282.
79. Giannouli M, Di Popolo A, Durante-Mangoni E, Bernardo M, Cuccurullo S, Amato G, Tripodi M-F, Triassi M, Utili R, Zarrilli R. 2012. Molecular epidemiology and

- mechanisms of rifampicin resistance in *Acinetobacter baumannii* isolates from Italy. International Journal of Antimicrobial Agents 39:58-63.
80. Giannouli M, Di Popolo A, Durante-Mangoni E, Bernardo M, Cuccurullo S, Amato G, Tripodi MF, Triassi M, Utili R, Zarrilli R. 2012. Molecular epidemiology and mechanisms of rifampicin resistance in *Acinetobacter baumannii* isolates from Italy. Int J Antimicrob Agents 39:58-63.
 81. Girlich D, Bonnin RA, Bogaerts P, De Laveleye M, Huang DT, Dortet L, Glaser P, Glupczynski Y, Naas T. 2017. Chromosomal Amplification of the *bla*_{OXA-58} Carbapenemase Gene in a *Proteus mirabilis* Clinical Isolate. Antimicrob Agents Chemother 61.
 82. Golebiewski M, Kern-Zdanowicz I, Zienkiewicz M, Adamczyk M, Zylinska J, Baraniak A, Gniadkowski M, Bardowski J, Ceglowski P. 2007. Complete nucleotide sequence of the pCTX-M3 plasmid and its involvement in spread of the extended-spectrum β -lactamase gene *bla*_{CTX-M-3}. Antimicrob Agents Chemother 51:3789-95.
 83. Gonzalez-Zorn B, Catalan A, Escudero JA, Dominguez L, Teshager T, Porrero C, Moreno MA. 2005. Genetic basis for dissemination of *armA*. J Antimicrob Chemother 56:583-5.
 84. Gottig S, Gruber TM, Higgins PG, Wachsmuth M, Seifert H, Kempf VA. 2014. Detection of pan drug-resistant *Acinetobacter baumannii* in Germany. J Antimicrob Chemother 69:2578-9.
 85. Grosso F, Quinteira S, Poirel L, Novais A, Peixe L. 2012. Role of common *bla*_{OXA-24/OXA-40}-carrying platforms and plasmids in the spread of OXA-24/OXA-40 among *Acinetobacter* species clinical isolates. Antimicrob Agents Chemother 56:3969-72.
 86. Gu B, Tong M, Zhao W, Liu G, Ning M, Pan S, Zhao W. 2007. Prevalence and characterization of class 1 integrons among *Pseudomonas aeruginosa* and *Acinetobacter baumannii* isolates from patients in Nanjing, China. J Clin Microbiol 45:241-3.
 87. Hamidian M, Ambrose SJ, Hall RM. 2016. A large conjugative *Acinetobacter baumannii* plasmid carrying the *sul2* sulphonamide and *strAB* streptomycin resistance genes. Plasmid 87-88:43-50.
 88. Hamidian M, Hall RM. 2011. AbaR4 replaces AbaR3 in a carbapenem-resistant *Acinetobacter baumannii* isolate belonging to global clone 1 from an Australian hospital. J Antimicrob Chemother 66:2484-91.
 89. Hamidian M, Hall RM. 2013. ISAbal targets a specific position upstream of the intrinsic *ampC* gene of *Acinetobacter baumannii* leading to cephalosporin resistance. J Antimicrob Chemother 68:2682-3.
 90. Hamidian M, Hall RM. 2014. pACICU2 is a conjugative plasmid of *Acinetobacter* carrying the aminoglycoside resistance transposon TnaphA6. J Antimicrob Chemother 69:1146-8.
 91. Hamidian M, Hall RM. 2014. Resistance to third-generation cephalosporins in *Acinetobacter baumannii* due to horizontal transfer of a chromosomal segment containing ISAbal-*ampC*. J Antimicrob Chemother 69:2865-6.
 92. Hamidian M, Hall RM. 2014. Tn6168, a transposon carrying an ISAbal-activated *ampC* gene and conferring cephalosporin resistance in *Acinetobacter baumannii*. J Antimicrob Chemother 69:77-80.
 93. Hamidian M, Hall RM. 2016. The resistance gene complement of D4, a multiply antibiotic-resistant ST25 *Acinetobacter baumannii* isolate, resides in two genomic islands and a plasmid. Journal of Antimicrobial Chemotherapy 71:1730-1732.

94. Hamidian M, Hall RM. 2017. Origin of the AbGRI1 antibiotic resistance island found in the *comM* gene of *Acinetobacter baumannii* GC2 isolates. J Antimicrob Chemother doi:10.1093/jac/dkx206.
95. Hamidian M, Hancock DP, Hall RM. 2013. Horizontal transfer of an ISAbal25-activated *ampC* gene between *Acinetobacter baumannii* strains leading to cephalosporin resistance. J Antimicrob Chemother 68:244-5.
96. Hamidian M, Hawkey J, Holt KE, Hall RM. 2015. Genome sequence of *Acinetobacter baumannii* strain D36, an antibiotic-resistant isolate from lineage 2 of global clone 1. Genome Announc 3:e01478-15.
97. Hamidian M, Holt KE, Pickard D, Dougan G, Hall RM. 2014. A GC1 *Acinetobacter baumannii* isolate carrying AbaR3 and the aminoglycoside resistance transposon *TnaphA6* in a conjugative plasmid. J Antimicrob Chemother 69:955-8.
98. Hamidian M, Holt KE, Pickard D, Hall RM. 2016. A small *Acinetobacter* plasmid carrying the *tet39* tetracycline resistance determinant. J Antimicrob Chemother 71:269-71.
99. Hamidian M, Kenyon JJ, Holt KE, Pickard D, Hall RM. 2014. A conjugative plasmid carrying the carbapenem resistance gene *bla*_{OXA-23} in AbaR4 in an extensively resistant GC1 *Acinetobacter baumannii* isolate. J Antimicrob Chemother 69:2625-8.
100. Hamidian M, Nigro SJ, Hall RM. 2012. Variants of the gentamicin and tobramycin resistance plasmid pRAY are widely distributed in *Acinetobacter*. J Antimicrob Chemother 67:2833-6.
101. Hamidian M, Nigro SJ, Hall RM. 2017. Problems with the Oxford multilocus sequence typing scheme for *Acinetobacter baumannii*: do sequence type 92 (ST92) and ST109 exist? J Clin Microbiol 55:2287-2289.
102. Hamidian M, Nigro SJ, Hartstein RM, Hall RM. 2017. RCH51, a multiply antibiotic-resistant *Acinetobacter baumannii* ST103IP isolate, carries resistance genes in three plasmids, including a novel potentially conjugative plasmid carrying *oxa235* in transposon Tn6252. J Antimicrob Chemother doi:10.1093/jac/dkx069.
103. Hamidian M, Wynn M, Holt KE, Pickard D, Dougan G, Hall RM. 2014. Identification of a marker for two lineages within the GC1 clone of *Acinetobacter baumannii*. J Antimicrob Chemother 69:557-8.
104. Hamouda A, Evans BA, Towner KJ, Amyes SG. 2010. Characterization of epidemiologically unrelated *Acinetobacter baumannii* isolates from four continents by use of multilocus sequence typing, pulsed-field gel electrophoresis, and sequence-based typing of (OXA-51-like) genes. J Clin Microbiol 48:2476-83.
105. Han HL, Jang SJ, Park G, Kook JK, Shin JH, Shin SH, Kim DM, Cheon JS, Moon DS, Park YJ. 2008. Identification of an atypical integron carrying an IS26-disrupted *aadA1* gene cassette in *Acinetobacter baumannii*. Int J Antimicrob Agents 32:165-9.
106. Harmer CJ, Hall RM. 2016. IS26-Mediated formation of transposons carrying antibiotic resistance genes. mSphere 1:e00038-16.
107. Harmer CJ, Moran RA, Hall RM. 2014. Movement of IS26-associated antibiotic resistance genes occurs via a translocatable unit that includes a single IS26 and preferentially inserts adjacent to another IS26. MBio 5:e01801-14.
108. Hawkey J, Hamidian M, Wick RR, Edwards DJ, Billman-Jacobe H, Hall RM, Holt KE. 2015. ISMapper: identifying transposase insertion sites in bacterial genomes from short read sequence data. BMC Genomics 16:667.
109. He S, Hickman AB, Varani AM, Siguier P, Chandler M, Dekker JP, Dyda F. 2015. Insertion sequence IS26 reorganizes plasmids in clinically isolated multidrug-resistant bacteria by replicative transposition. MBio 6.

110. Heather JM, Chain B. 2016. The sequence of sequencers: The history of sequencing DNA. *Genomics* 107:1-8.
111. Heritier C, Poirel L, Nordmann P. 2006. Cephalosporinase over-expression resulting from insertion of ISAbal in *Acinetobacter baumannii*. *Clin Microbiol Infect* 12:123-30.
112. Heuer H, Kopmann C, Binh CT, Top EM, Smalla K. 2009. Spreading antibiotic resistance through spread manure: characteristics of a novel plasmid type with low %G+C content. *Environ Microbiol* 11:937-49.
113. Higgins PG, Dammhayn C, Hackel M, Seifert H. 2010. Global spread of carbapenem-resistant *Acinetobacter baumannii*. *J Antimicrob Chemother* 65:233-8.
114. Higgins PG, Lehmann M, Wisplinghoff H, Seifert H. 2010. *gyrB* Multiplex PCR To Differentiate between *Acinetobacter calcoaceticus* and *Acinetobacter* Genomic Species 3. *Journal of Clinical Microbiology* 48:4592-4594.
115. Higgins PG, Perez-Llarena FJ, Zander E, Fernandez A, Bou G, Seifert H. 2013. OXA-235, a novel class D β -lactamase involved in resistance to carbapenems in *Acinetobacter baumannii*. *Antimicrob Agents Chemother* 57:2121-6.
116. Higgins PG, Poirel L, Lehmann M, Nordmann P, Seifert H. 2009. OXA-143, a novel carbapenem-hydrolyzing class D β -lactamase in *Acinetobacter baumannii*. *Antimicrob Agents Chemother* 53:5035-8.
117. Higgins PG, Wisplinghoff H, Krut O, Seifert H. 2007. A PCR-based method to differentiate between *Acinetobacter baumannii* and *Acinetobacter* genomic species 13TU. *Clin Microbiol Infect* 13:1199-201.
118. Ho AY, Chow KH, Law PY, Tse H, Ho PL. 2013. Draft genome sequences of two multidrug-resistant *Acinetobacter baumannii* strains of sequence type ST92 and ST96. *Genome Announc* 1:e00296-13.
119. Holt K, Kenyon JJ, Hamidian M, Schultz MB, Pickard DJ, Dougan G, Hall R. 2016. Five decades of genome evolution in the globally distributed, extensively antibiotic-resistant *Acinetobacter baumannii* global clone 1. *Microbial Genomics* 2:e000052.
120. Holt KE, Hamidian M, Kenyon JJ, Wynn MT, Hawkey J, Pickard D, Hall RM. 2015. Genome sequence of *Acinetobacter baumannii* strain A1, an early example of antibiotic-resistant global clone 1. *Genome Announc* 3.
121. Houang ET, Chu YW, Lo WS, Chu KY, Cheng AF. 2003. Epidemiology of rifampin ADP-ribosyltransferase (*arr-2*) and metallo- β -lactamase (*bla*IMP-4) gene cassettes in class 1 integrons in *Acinetobacter* strains isolated from blood cultures in 1997 to 2000. *Antimicrob Agents Chemother* 47:1382-90.
122. Howard A, O'Donoghue M, Feeney A, Sleator RD. 2012. *Acinetobacter baumannii*: an emerging opportunistic pathogen. *Virulence* 3:243-50.
123. Hu H, Hu Y, Pan Y, Liang H, Wang H, Wang X, Hao Q, Yang X, Yang X, Xiao X, Luan C, Yang Y, Cui Y, Yang R, Gao GF, Song Y, Zhu B. 2012. Novel plasmid and its variant harboring both a *bla*_{NDM-1} gene and type IV secretion system in clinical isolates of *Acinetobacter lwoffii*. *Antimicrob Agents Chemother* 56:1698-702.
124. Hua X, Shu J, Ruan Z, Yu Y, Feng Y. 2016. Multiplication of *bla*_{OXA-23} is common in clinical *Acinetobacter baumannii*, but does not enhance carbapenem resistance. *J Antimicrob Chemother* 71:3381-3385.
125. Huang H, Yang ZL, Wu XM, Wang Y, Liu YJ, Luo H, Lv X, Gan YR, Song SD, Gao F. 2012. Complete genome sequence of *Acinetobacter baumannii* MDR-TJ and insights into its mechanism of antibiotic resistance. *J Antimicrob Chemother* 67:2825-32.
126. Huang TW, Lauderdale TL, Liao TL, Hsu MC, Chang FY, Chang SC, Khong WX, Ng OT, Chen YT, Kuo SC, Chen TL, Mu JJ, Tsai SF. 2015. Effective transfer of a 47

- kb NDM-1-positive plasmid among *Acinetobacter* species. J Antimicrob Chemother 70:2734-8.
127. Hujer KM, Hujer AM, Hulten EA, Bajaksouzian S, Adams JM, Donskey CJ, Ecker DJ, Massire C, Eshoo MW, Sampath R, Thomson JM, Rather PN, Craft DW, Fishbain JT, Ewell AJ, Jacobs MR, Paterson DL, Bonomo RA. 2006. Analysis of antibiotic resistance genes in multidrug-resistant *Acinetobacter* sp. isolates from military and civilian patients treated at the Walter Reed Army Medical Center. Antimicrob Agents Chemother 50:4114-23.
 128. Hwa WE, Subramaniam G, Navaratnam P, Sekaran SD. 2009. Detection and characterization of class 1 integrons among carbapenem-resistant isolates of *Acinetobacter* spp. in Malaysia. J Microbiol Immunol Infect 42:54-62.
 129. Iacono M, Villa L, Fortini D, Bordoni R, Imperi F, Bonnal RJ, Sicheritz-Ponten T, De Bellis G, Visca P, Cassone A, Carattoli A. 2008. Whole-genome pyrosequencing of an epidemic multidrug-resistant *Acinetobacter baumannii* strain belonging to the European clone II group. Antimicrob Agents Chemother 52:2616-25.
 130. Imperi F, Antunes LC, Blom J, Villa L, Iacono M, Visca P, Carattoli A. 2011. The genomics of *Acinetobacter baumannii*: insights into genome plasticity, antimicrobial resistance and pathogenicity. IUBMB Life 63:1068-74.
 131. Jeannot K, Diancourt L, Vaux S, Thouverez M, Ribeiro A, Coignard B, Courvalin P, Brisse S. 2014. Molecular epidemiology of carbapenem non-susceptible *Acinetobacter baumannii* in France. PLoS One 9:e115452.
 132. Kadlec K, Brenner Michael G, Sweeney MT, Brzuszkiewicz E, Liesegang H, Daniel R, Watts JL, Schwarz S. 2011. Molecular basis of macrolide, triamylide, and lincosamide resistance in *Pasteurella multocida* from bovine respiratory disease. Antimicrob Agents Chemother 55:2475-7.
 133. Kakinuma Y, Maeda Y, Mason C, Goldsmith CE, Coulter WA, Matsuda M, Dooley JS, Lowery CJ, Moore JE. 2012. Molecular characterisation of the quinolone resistance-determining regions (QRDR) including *gyrA*, *gyrB*, *parC* and *parE* genes in *Streptococcus pneumoniae*. Br J Biomed Sci 69:123-5.
 134. Kamolvit W, Sidjabat HE, Paterson DL. 2015. Molecular epidemiology and mechanisms of carbapenem resistance of *Acinetobacter* spp. in Asia and Oceania. Microb Drug Resist 21:424-34.
 135. Karah N, Dwibedi CK, Sjostrom K, Edquist P, Johansson A, Wai SN, Uhlin BE. 2016. Novel aminoglycoside resistance transposons and transposon-derived circular forms detected in carbapenem-resistant *Acinetobacter baumannii* clinical isolates. Antimicrob Agents Chemother 60:1801-18.
 136. Karah N, Haldorsen B, Hermansen NO, Tveten Y, Ragnhildstveit E, Skutlaberg DH, Tofteland S, Sundsfjord A, Samuelsen O. 2011. Emergence of OXA-carbapenemase- and 16S rRNA methylase-producing international clones of *Acinetobacter baumannii* in Norway. J Med Microbiol 60:515-21.
 137. Kenyon JJ, Hall RM. 2013. Variation in the complex carbohydrate biosynthesis loci of *Acinetobacter baumannii* genomes. PLoS One 8:e62160.
 138. Kholodii GY, Mindlin SZ, Bass IA, Yurieva OV, Minakhina SV, Nikiforov VG. 1995. Four genes, two ends, and a *res* region are involved in transposition of Tn5053: a paradigm for a novel family of transposons carrying either a mer operon or an integron. Mol Microbiol 17:1189-200.
 139. Kim DH, Choi JY, Kim HW, Kim SH, Chung DR, Peck KR, Thamlikitkul V, So TM, Yasin RM, Hsueh PR, Carlos CC, Hsu LY, Buntaran L, Lalitha MK, Song JH, Ko KS. 2013. Spread of carbapenem-resistant *Acinetobacter baumannii* global clone 2 in Asia and AbaR-type resistance islands. Antimicrob Agents Chemother 57:5239-46.

140. Kim DH, Park YK, Ko KS. 2012. Variations of AbaR4-type resistance islands in *Acinetobacter baumannii* Isolates from South Korea. *Antimicrob Agents Chemother* 56:4544-4547.
141. Kim MS, Jun LJ, Shin SB, Park MA, Jung SH, Kim K, Moon KH, Jeong HD. 2010. Mutations in the *gyrB*, *parC*, and *parE* genes of quinolone-resistant isolates and mutants of *Edwardsiella tarda*. *J Microbiol Biotechnol* 20:1735-43.
142. Kircher M, Kelso J. 2010. High-throughput DNA sequencing--concepts and limitations. *Bioessays* 32:524-36.
143. Koh TH, Sng LH, Wang GC, Hsu LY, Zhao Y. 2007. IMP-4 and OXA β -lactamases in *Acinetobacter baumannii* from Singapore. *J Antimicrob Chemother* 59:627-32.
144. Kotecha B, Richardson GP. 1994. Ototoxicity in vitro: effects of neomycin, gentamicin, dihydrostreptomycin, amikacin, spectinomycin, neamine, spermine and poly-L-lysine. *Hearing Research* 73:173-184.
145. Krizova L, Dijkshoorn L, Nemec A. 2011. Diversity and evolution of AbaR genomic resistance islands in *Acinetobacter baumannii* strains of European clone I. *Antimicrob Agents Chemother* 55:3201-6.
146. Krizova L, Nemec A. 2010. A 63 kb genomic resistance island found in a multidrug-resistant *Acinetobacter baumannii* isolate of European clone I from 1977. *J Antimicrob Chemother* 65:1915-8.
147. Lambert T, Gerbaud G, Bouvet P, Vieu JF, Courvalin P. 1990. Dissemination of amikacin resistance gene *aphA6* in *Acinetobacter* spp. *Antimicrob Agents Chemother* 34:1244-8.
148. Lambert T, Gerbaud G, Courvalin P. 1988. Transferable amikacin resistance in *Acinetobacter* spp. due to a new type of 3'-aminoglycoside phosphotransferase. *Antimicrob Agents Chemother* 32:15-9.
149. Lanka E, Wilkins BM. 1995. DNA processing reactions in bacterial conjugation. *Annu Rev Biochem* 64:141-69.
150. Lean SS, Yeo CC, Suhaili Z, Thong KL. 2015. Whole-genome analysis of an extensively drug-resistant clinical isolate of *Acinetobacter baumannii* AC12: insights into the mechanisms of resistance of an ST195 clone from Malaysia. *Int J Antimicrob Agents* 45:178-82.
151. Leclercq R. 2002. Mechanisms of resistance to macrolides and lincosamides: nature of the resistance elements and their clinical implications. *Clin Infect Dis* 34:482-92.
152. Lee H, Gurtowski J, Yoo S, Nattestad M, Marcus S, Goodwin S, McCombie WR, Schatz M. 2016. Third-generation sequencing and the future of genomics. *bioRxiv* doi:10.1101/048603.
153. Lee H, Yong D, Yum JH, Roh KH, Lee K, Yamane K, Arakawa Y, Chong Y. 2006. Dissemination of 16S rRNA methylase-mediated highly amikacin-resistant isolates of *Klebsiella pneumoniae* and *Acinetobacter baumannii* in Korea. *Diagn Microbiol Infect Dis* 56:305-12.
154. Lee K, Yum JH, Yong D, Lee HM, Kim HD, Docquier JD, Rossolini GM, Chong Y. 2005. Novel acquired metallo-beta-lactamase gene, *bla*_(SIM-1), in a class 1 integron from *Acinetobacter baumannii* clinical isolates from Korea. *Antimicrob Agents Chemother* 49:4485-91.
155. Lee SY, Yun SH, Lee YG, Choi CW, Leem SH, Park EC, Kim GH, Lee JC, Kim SI. 2014. Proteogenomic characterization of antimicrobial resistance in extensively drug-resistant *Acinetobacter baumannii* DU202. *J Antimicrob Chemother* 69:1483-91.
156. Lee Y, D'Souza R, Yong D, Lee K. 2016. Prediction of putative resistance islands in a carbapenem-resistant *Acinetobacter baumannii* global clone 2 clinical isolate. *Ann Lab Med* 36:320-4.

157. Lee YW, Choe H, Lee SH, Kim KM, Kam S, Kim BK, Lee WK. 2016. Complete genome of the multidrug-resistant *Acinetobacter baumannii* strain KBN10P02143 isolated from Korea. *Mem Inst Oswaldo Cruz* 111:355-8.
158. Leverstein-van Hall MA, Paauw A, Box ATA, Blok HEM, Verhoef J, Fluit AC. 2002. Presence of integron-associated resistance in the community is widespread and contributes to multidrug resistance in the hospital. *J Clin Microbiol* 40:3038-3040.
159. Lévesque C, Piché L, Larose C, Roy PH. 1995. PCR mapping of integrons reveals several novel combinations of resistance genes. *Antimicrob Agents and Chemother* 39:185-191.
160. Liou ML, Liu CC, Lu CW, Hsieh MF, Chang KC, Kuo HY, Lee CC, Chang CT, Yang CY, Tang CY. 2012. Genome sequence of *Acinetobacter baumannii* TYTH-1. *J Bacteriol* 194:6974.
161. Liu CC, Tang CY, Kuo HY, Lu CW, Chang KC, Liou ML. 2013. The origin of *Acinetobacter baumannii* TYTH-1: a comparative genomics study. *Int J Antimicrob Agents* 41:318-24.
162. Liu F, Zhu Y, Yi Y, Lu N, Zhu B, Hu Y. 2014. Comparative genomic analysis of *Acinetobacter baumannii* clinical isolates reveals extensive genomic variation and diverse antibiotic resistance determinants. *BMC Genomics* 15:1163.
163. Loli A, Tzouvelekis LS, Tzelepi E, Carattoli A, Vatopoulos AC, Tassios PT, Miriagou V. 2006. Sources of diversity of carbapenem resistance levels in *Klebsiella pneumoniae* carrying *bla_{VIM-1}*. *J Antimicrob Chemother* 58:669-72.
164. Loman NJ, Constantinidou C, Chan JZ, Halachev M, Sergeant M, Penn CW, Robinson ER, Pallen MJ. 2012. High-throughput bacterial genome sequencing: an embarrassment of choice, a world of opportunity. *Nat Rev Microbiol* 10:599-606.
165. Lopes BS, Amyes SG. 2012. Role of ISAbal and ISAbal25 in governing the expression of *bla_{ADC}* in clinically relevant *Acinetobacter baumannii* strains resistant to cephalosporins. *J Med Microbiol* 61:1103-8.
166. Magiorakos AP, Srinivasan A, Carey RB, Carmeli Y, Falagas ME, Giske CG, Harbarth S, Hindler JF, Kahlmeter G, Olsson-Liljequist B, Paterson DL, Rice LB, Stelling J, Struelens MJ, Vatopoulos A, Weber JT, Monnet DL. 2012. Multidrug-resistant, extensively drug-resistant and pandrug-resistant bacteria: an international expert proposal for interim standard definitions for acquired resistance. *Clin Microbiol Infect* 18:268-81.
167. Mahillon J, Chandler M. 1998. Insertion sequences. *Microbiol Mol Biol Rev* 62:725-74.
168. Marque S, Poirel L, Heritier C, Brisse S, Blasco MD, Filip R, Coman G, Naas T, Nordmann P. 2005. Regional occurrence of plasmid-mediated carbapenem-hydrolyzing oxacillinase OXA-58 in *Acinetobacter* spp. in Europe. *J Clin Microbiol* 43:4885-8.
169. Marti S, Fernandez-Cuenca F, Pascual A, Ribera A, Rodriguez-Bano J, Bou G, Miguel Cisneros J, Pachon J, Martinez-Martinez L, Vila J. 2006. Prevalence of the *tetA* and *tetB* genes as mechanisms of resistance to tetracycline and minocycline in *Acinetobacter baumannii* clinical isolates. *Enferm Infecc Microbiol Clin* 24:77-80.
170. McDermott PF, Walker RD, White DG. 2003. Antimicrobials: modes of action and mechanisms of resistance. *Int J Toxicol* 22:135-43.
171. McGann P, Courvalin P, Snesrud E, Clifford RJ, Yoon EJ, Onmus-Leone F, Ong AC, Kwak YI, Grillot-Courvalin C, Lesho E, Waterman PE. 2014. Amplification of aminoglycoside resistance gene *aphA1* in *Acinetobacter baumannii* results in tobramycin therapy failure. *MBio* 5:e00915.

172. Merino M, Acosta J, Poza M, Sanz F, Beceiro A, Chaves F, Bou G. 2010. OXA-24 carbapenemase gene flanked by XerC/XerD-like recombination sites in different plasmids from different *Acinetobacter* species isolated during a nosocomial outbreak. *Antimicrob Agents Chemother* 54:2724-7.
173. Midonet C, Barre FX. 2014. Xer site-specific recombination: promoting vertical and horizontal transmission of genetic information. *Microbiol Spectr* 2.
174. Mierzejewska J, Kulinska A, Jagura-Burdzy G. 2007. Functional analysis of replication and stability regions of broad-host-range conjugative plasmid CTX-M3 from the IncL/M incompatibility group. *Plasmid* 57:95-107.
175. Mindlin S, Petrenko A, Kurakov A, Beletsky A, Mardanov A, Petrova M. 2016. Resistance of permafrost and modern *Acinetobacter iwoffii* strains to heavy metals and arsenic revealed by genome analysis. *Biomed Res Int* 2016:3970831.
176. Mindlin SZ, Minakhina SV, Kholodii G, Kopteva AV, Nikiforov VG. 1996. Incorporation of Tn5053 and Tn402 into various plasmids. *Genetika* 32:1426-30.
177. Mollet B, Iida S, Shepherd J, Arber W. 1983. Nucleotide sequence of IS26, a new prokaryotic mobile genetic element. *Nucleic Acids Res* 11:6319-30.
178. Moran RA, Hall RM. 2017. Analysis of pCERC7, a small antibiotic resistance plasmid from a commensal ST131 *Escherichia coli*, defines a diverse group of plasmids that include various segments adjacent to a multimer resolution site and encode the same NikA relaxase accessory protein enabling mobilisation. *Plasmid* 89:42-48.
179. Mosqueda N, Gato E, Roca I, Lopez M, de Alegria CR, Fernandez Cuenca F, Martinez-Martinez L, Pachon J, Cisneros JM, Rodriguez-Bano J, Pascual A, Vila J, Bou G, Tomas M. 2014. Characterization of plasmids carrying the *bla*_{OXA-24/40} carbapenemase gene and the genes encoding the AbkA/AbkB proteins of a toxin/antitoxin system. *J Antimicrob Chemother* 69:2629-33.
180. Mugnier PD, Poirel L, Naas T, Nordmann P. 2010. Worldwide dissemination of the *bla*_{OXA-23} carbapenemase gene of *Acinetobacter baumannii*. *Emerg Infect Dis* 16:35-40.
181. Mugnier PD, Poirel L, Nordmann P. 2009. Functional analysis of insertion sequence ISAbal, responsible for genomic plasticity of *Acinetobacter baumannii*. *J Bacteriol* 191:2414-8.
182. Nemec A, Dijkshoorn L, van der Reijden TJ. 2004. Long-term predominance of two pan-European clones among multi-resistant *Acinetobacter baumannii* strains in the Czech Republic. *J Med Microbiol* 53:147-53.
183. 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. *J Med Microbiol* 53:1233-40.
184. Nemec A, Janda L, Melter O, Dijkshoorn L. 1999. Genotypic and phenotypic similarity of multiresistant *Acinetobacter baumannii* isolates in the Czech Republic. *J Med Microbiol* 48:287-96.
185. Nemec A, Krizova L, Maixnerova M, van der Reijden TJ, Deschaght P, Passet V, Vaneechoutte M, Brisse S, Dijkshoorn L. 2011. Genotypic and phenotypic characterization of the *Acinetobacter calcoaceticus*-*Acinetobacter baumannii* complex with the proposal of *Acinetobacter pittii* sp. nov. (formerly *Acinetobacter* genomic species 3) and *Acinetobacter nosocomialis* sp. nov. (formerly *Acinetobacter* genomic species 13TU). *Res Microbiol* 162:393-404.
186. Ng LK, Martin I, Alfa M, Mulvey M. 2001. Multiplex PCR for the detection of tetracycline resistant genes. *Mol Cell Probes* 15:209-15.

187. Nhu NTK, Lan NPH, Campbell JI, Parry CM, Thompson C, Tuyen HT, Hoang NVM, Tam PTT, Le VM, Nga TVT, Nhu TDH, Van Minh P, Nga NTT, Thuy CT, Dung LT, Yen NTT, Van Hao N, Loan HT, Yen LM, Nghia HDT, Hien TT, Thwaites L, Thwaites G, Chau NVV, Baker S. 2014. Emergence of carbapenem-resistant *Acinetobacter baumannii* as the major cause of ventilator-associated pneumonia in intensive care unit patients at an infectious disease hospital in southern Vietnam. *Journal of Medical Microbiology* 63:1386-1394.
188. Nigro SJ, Farrugia DN, Paulsen IT, Hall RM. 2013. A novel family of genomic resistance islands, AbGRI2, contributing to aminoglycoside resistance in *Acinetobacter baumannii* isolates belonging to global clone 2. *J Antimicrob Chemother* 68:554-7.
189. Nigro SJ, Hall RM. 2011. GI*sul2*, a genomic island carrying the *sul2* sulphonamide resistance gene and the small mobile element CR2 found in the *Enterobacter cloacae* subspecies *cloacae* type strain ATCC 13047 from 1890, *Shigella flexneri* ATCC 700930 from 1954 and *Acinetobacter baumannii* ATCC 17978 from 1951. *J Antimicrob Chemother* 66:2175-6.
190. Nigro SJ, Hall RM. 2012. Antibiotic resistance islands in A320 (RUH134), the reference strain for *Acinetobacter baumannii* global clone 2. *J Antimicrob Chemother* 67:335-8.
191. Nigro SJ, Hall RM. 2012. Tn*6167*, an antibiotic resistance island in an Australian carbapenem-resistant *Acinetobacter baumannii* GC2, ST92 isolate. *J Antimicrob Chemother* 67:1342-6.
192. Nigro SJ, Hall RM. 2014. Amikacin resistance plasmids in extensively antibiotic-resistant GC2 *Acinetobacter baumannii* from two Australian hospitals. *J Antimicrob Chemother* 69:3435-7.
193. Nigro SJ, Hall RM. 2016. Loss and gain of aminoglycoside resistance in global clone 2 *Acinetobacter baumannii* Australia via modification of genomic resistance islands and acquisition of plasmids. *J Antimicrob Chemother* 71:2432-40.
194. Nigro SJ, Hall RM. 2016. Structure and context of *Acinetobacter* transposons carrying the *oxa23* carbapenemase gene. *J Antimicrob Chemother* 71:1135-47.
195. Nigro SJ, Holt KE, Pickard D, Hall RM. 2014. Carbapenem and amikacin resistance on a large conjugative *Acinetobacter baumannii* plasmid. *J Antimicrob Chemother* doi:10.1093/jac/dku486.
196. Nigro SJ, Holt KE, Pickard D, Hall RM. 2015. Carbapenem and amikacin resistance on a large conjugative *Acinetobacter baumannii* plasmid. *J Antimicrob Chemother* 70:1259-61.
197. Nigro SJ, Post V, Hall RM. 2011. Aminoglycoside resistance in multiply antibiotic-resistant *Acinetobacter baumannii* belonging to global clone 2 from Australian hospitals. *J Antimicrob Chemother* 66:1504-9.
198. Nigro SJ, Post V, Hall RM. 2011. The multiresistant *Acinetobacter baumannii* European clone I type strain RUH875 (A297) carries a genomic antibiotic resistance island AbaR21, plasmid pRAY and a cluster containing IS*Aba1-sul2-CR2-strB-strA*. *J Antimicrob Chemother* 66:1928-30.
199. Nyren P, Lundin A. 1985. Enzymatic method for continuous monitoring of inorganic pyrophosphate synthesis. *Anal Biochem* 151:504-9.
200. O'Brien FG, Ramsay JP, Monecke S, Coombs GW, Robinson OJ, Htet Z, Alshaikh FA, Grubb WB. 2015. *Staphylococcus aureus* plasmids without mobilization genes are mobilized by a novel conjugative plasmid from community isolates. *J Antimicrob Chemother* 70:649-52.

201. Palmer EL, Teviotdale BL, Jones AL. 1997. A relative of the broad-host-range plasmid RSF1010 detected in *Erwinia amylovora*. *Appl Environ Microbiol* 63:4604-7.
202. Park JY, Kim S, Kim SM, Cha SH, Lim SK, Kim J. 2011. Complete genome sequence of multidrug-resistant *Acinetobacter baumannii* strain 1656-2, which forms sturdy biofilm. *J Bacteriol* 193:6393-4.
203. Partridge SR, Ginn AN, Paulsen IT, Iredell JR. 2012. pEI1573 Carrying *bla*IMP-4, from Sydney, Australia, is closely related to other IncL/M plasmids. *Antimicrob Agents Chemother* 56:6029-32.
204. Partridge SR, Recchia GD, Stokes HW, Hall RM. 2001. Family of class 1 integrons related to In4 from Tn1696. *Antimicrob Agents Chemother* 45:3014-20.
205. Paton R, Miles RS, Hood J, Amyes SG, Miles RS, Amyes SG. 1993. ARI 1: β -lactamase-mediated imipenem resistance in *Acinetobacter baumannii*. *Int J Antimicrob Agents* 2:81-7.
206. Pehrsson EC, Tsukayama P, Patel S, Mejia-Bautista M, Sosa-Soto G, Navarrete KM, Calderon M, Cabrera L, Hoyos-Arango W, Bertoli MT, Berg DE, Gilman RH, Dantas G. 2016. Interconnected microbiomes and resistomes in low-income human habitats. *Nature* 533:212-6.
207. Peleg AY, Seifert H, Paterson DL. 2008. *Acinetobacter baumannii*: emergence of a successful pathogen. *Clin Microbiol Rev* 21:538-82.
208. Perez F, Endimiani A, Ray AJ, Decker BK, Wallace CJ, Hujer KM, Ecker DJ, Adams MD, Toltzis P, Dul MJ, Windau A, Bajaksouzian S, Jacobs MR, Salata RA, Bonomo RA. 2010. Carbapenem-resistant *Acinetobacter baumannii* and *Klebsiella pneumoniae* across a hospital system: impact of post-acute care facilities on dissemination. *J Antimicrob Chemother* 65:1807-18.
209. Perez F, Hujer AM, Hujer KM, Decker BK, Rather PN, Bonomo RA. 2007. Global challenge of multidrug-resistant *Acinetobacter baumannii*. *Antimicrob Agents Chemother* 51:3471-84.
210. Peters JE. 2014. Tn7. *Microbiol Spectr* 2.
211. Poirel L, Bonnin RA, Nordmann P. 2011. Genetic basis of antibiotic resistance in pathogenic *Acinetobacter* species. *IUBMB Life* 63:1061-7.
212. Poirel L, Figueiredo S, Cattoir V, Carattoli A, Nordmann P. 2008. *Acinetobacter radioresistens* as a silent source of carbapenem resistance for *Acinetobacter* spp. *Antimicrob Agents Chemother* 52:1252-6.
213. Poirel L, Mansour W, Bouallegue O, Nordmann P. 2008. Carbapenem-resistant *Acinetobacter baumannii* isolates from Tunisia producing the OXA-58-like carbapenem-hydrolyzing oxacillinase OXA-97. *Antimicrob Agents Chemother* 52:1613-7.
214. Poirel L, Marque S, Heritier C, Segonds C, Chabanon G, Nordmann P. 2005. OXA-58, a novel class D β -lactamase involved in resistance to carbapenems in *Acinetobacter baumannii*. *Antimicrob Agents Chemother* 49:202-8.
215. Poirel L, Naas T, Nordmann P. 2010. Diversity, epidemiology, and genetics of class D β -lactamases. *Antimicrob Agents Chemother* 54:24-38.
216. Poirel L, Nordmann P. 2006. Genetic structures at the origin of acquisition and expression of the carbapenem-hydrolyzing oxacillinase gene *bla*_{OXA-58} in *Acinetobacter baumannii*. *Antimicrob Agents Chemother* 50:1442-8.
217. Post V, Hall RM. 2009. AbaR5, a large multiple-antibiotic resistance region found in *Acinetobacter baumannii*. *Antimicrob Agents Chemother* 53:2667-71.
218. Post V, Hamidian M, Hall RM. 2012. Antibiotic-resistant *Acinetobacter baumannii* variants belonging to global clone 1. *J Antimicrob Chemother* 67:1039-40.

219. Post V, White PA, Hall RM. 2010. Evolution of AbaR-type genomic resistance islands in multiply antibiotic-resistant *Acinetobacter baumannii*. J Antimicrob Chemother 65:1162-70.
220. Povilonis J, Seputiene V, Krasauskas R, Juskaite R, Miskinyte M, Suziedelis K, Suziedeliene E. 2013. Spread of carbapenem-resistant *Acinetobacter baumannii* carrying a plasmid with two genes encoding OXA-72 carbapenemase in Lithuanian hospitals. J Antimicrob Chemother 68:1000-6.
221. Qureshi ZA, Hittle LE, O'Hara JA, Rivera JI, Syed A, Shields RK, Pasculle AW, Ernst RK, Doi Y. 2015. Colistin-resistant *Acinetobacter baumannii*: beyond carbapenem resistance. Clin Infect Dis 60:1295-1303.
222. Radstrom P, Skold O, Swedberg G, Flensburg J, Roy PH, Sundstrom L. 1994. Transposon Tn5090 of plasmid R751, which carries an integron, is related to Tn7, Mu, and the retroelements. J Bacteriol 176:3257-68.
223. Ribera A, Ruiz J, Vila J. 2003. Presence of the TetM determinant in a clinical isolate of *Acinetobacter baumannii*. Antimicrob Agents Chemother 47:2310-2.
224. Rice LB. 2008. Federal funding for the study of antimicrobial resistance in nosocomial pathogens: no ESKAPE. J Infect Dis 197:1079-81.
225. Roberts MC. 2005. Update on acquired tetracycline resistance genes. FEMS Microbiol Lett 245:195-203.
226. Robledo IE, Aquino EE, Sante MI, Santana JL, Otero DM, Leon CF, Vazquez GJ. 2010. Detection of KPC in *Acinetobacter* spp. in Puerto Rico. Antimicrob Agents Chemother 54:1354-7.
227. Roca I, Marti S, Espinal P, Martinez P, Gibert I, Vila J. 2009. CraA, a major facilitator superfamily efflux pump associated with chloramphenicol resistance in *Acinetobacter baumannii*. Antimicrob Agents Chemother 53:4013-4.
228. Sahl JW, Johnson JK, Harris AD, Phillippy AM, Hsiao WW, Thom KA, Rasko DA. 2011. Genomic comparison of multi-drug resistant invasive and colonizing *Acinetobacter baumannii* isolated from diverse human body sites reveals genomic plasticity. BMC Genomics 12:291.
229. Sambrook J, Fritsch EF, T. M. 1989. Molecular Cloning: A laboratory manual, 2nd edition. . Cold Spring Harbor Laboratory Press New York.
230. Sanger F, Nicklen S, Coulson AR. 1977. DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci U S A 74:5463-7.
231. Saranathan R, Tomar A, Sudhakar P, Arunkumar KP, Prashanth K. 2014. Draft genome sequence of a multidrug-resistant *Acinetobacter baumannii* PKAB07 clinical strain from India belonging to sequence type 195. Genome Announc 2:e00184.
232. Saule M, Samuelsen Ø, Dumpis U, Sundsfjord A, Karlson A, Balode A, Miklasevics E, Karah N. 2013. Dissemination of a carbapenem-resistant *Acinetobacter baumannii* strain belonging to international clone II/sequence type 2 and harboring a novel AbaR4-like resistance island in Latvia. Antimicrobial Agents and Chemotherapy 57:1069-1072.
233. Schneider GF, Dekker C. 2012. DNA sequencing with nanopores. Nat Biotechnol 30:326-8.
234. SCHUBERT RHW. 1968. The taxonomy and nomenclature of the genus *Aeromonas* Kluver and van Niel 1936. International Journal of Systematic and Evolutionary Microbiology 18:1-7.
235. Schultz MB, Pham Thanh D, Tran Do Hoan N, Wick RR, Ingle DJ, Hawkey J, Edwards DJ, Kenyon JJ, Phu Huong Lan N, Campbell JI, Thwaites G, Thi Khanh Nhu N, Hall RM, Fournier-Level A, Baker S, Holt KE. 2016. Repeated local

- emergence of carbapenem-resistant *Acinetobacter baumannii* in a single hospital ward. *Microbial Genomics* 2:e000050.
236. Segal H, Elisha BG. 1997. Identification and characterization of an *aadB* gene cassette at a secondary site in a plasmid from *Acinetobacter*. *FEMS Microbiol Lett* 153:321-6.
 237. Seifert H, Baginski R, Schulze A, Pulverer G. 1993. Antimicrobial susceptibility of *Acinetobacter* species. *Antimicrob Agents Chemother* 37:750-3.
 238. Seputiene V, Povilonis J, Suziedeliene E. 2012. Novel variants of AbaR resistance islands with a common backbone in *Acinetobacter baumannii* isolates of European clone II. *Antimicrob Agents Chemother* 56:1969-73.
 239. Shaw KJ, Rather PN, Hare RS, Miller GH. 1993. Molecular genetics of aminoglycoside resistance genes and familial relationships of the aminoglycoside-modifying enzymes. *Microbiol Rev* 57:138-63.
 240. Shrestha S, Tada T, Miyoshi-Akiyama T, Ohara H, Shimada K, Satou K, Teruya K, Nakano K, Shiroma A, Sherchand JB, Rijal BP, Hirano T, Kirikae T, Pokhrel BM. 2015. Molecular epidemiology of multidrug-resistant *Acinetobacter baumannii* isolates in a university hospital in Nepal reveals the emergence of a novel epidemic clonal lineage. *Int J Antimicrob Agents* 46:526-31.
 241. Silver S, Phung LT. 1996. Bacterial heavy metal resistance: new surprises. *Annu Rev Microbiol* 50:753-89.
 242. Smillie C, Garcillan-Barcia MP, Francia MV, Rocha EP, de la Cruz F. 2010. Mobility of plasmids. *Microbiol Mol Biol Rev* 74:434-52.
 243. Smith MC, Thorpe HM. 2002. Diversity in the serine recombinases. *Mol Microbiol* 44:299-307.
 244. Smith MG, Gianoulis TA, Pukatzki S, Mekalanos JJ, Ornston LN, Gerstein M, Snyder M. 2007. New insights into *Acinetobacter baumannii* pathogenesis revealed by high-density pyrosequencing and transposon mutagenesis. *Genes Dev* 21:601-14.
 245. Soon RL, Nation RL, Cockram S, Moffatt JH, Harper M, Adler B, Boyce JD, Larson I, Li J. 2011. Different surface charge of colistin-susceptible and -resistant *Acinetobacter baumannii* cells measured with zeta potential as a function of growth phase and colistin treatment. *J Antimicrob Chemother* 66:126-33.
 246. Spratt BG. 1994. Resistance to antibiotics mediated by target alterations. *Science* 264:388-93.
 247. Srinivasan VB, Rajamohan G, Gebreyes WA. 2009. Role of AbeS, a novel efflux pump of the SMR family of transporters, in resistance to antimicrobial agents in *Acinetobacter baumannii*. *Antimicrob Agents Chemother* 53:5312-6.
 248. Stark WM. 2014. The serine recombinases. *Microbiol Spectr* 2.
 249. Stokes HW, Hall RM. 1989. A novel family of potentially mobile DNA elements encoding site-specific gene-integration functions: integrons. *Mol Microbiol* 3:1669-83.
 250. Su XZ, Chen J, Mizushima T, Kuroda T, Tsuchiya T. 2005. AbeM, an H⁺-coupled *Acinetobacter baumannii* multidrug efflux pump belonging to the MATE family of transporters. *Antimicrob Agents Chemother* 49:4362-4.
 251. Summers DK, Sherratt DJ. 1988. Resolution of ColE1 dimers requires a DNA sequence implicated in the three-dimensional organization of the *cer* site. *Embo j* 7:851-8.
 252. Sung JY, Koo SH, Kim S, Kwon KC. 2014. Epidemiological characterizations of class 1 integrons from multidrug-resistant *Acinetobacter* isolates in Daejeon, Korea. *Ann Lab Med* 34:293-9.

253. Tada T, Miyoshi-Akiyama T, Shimada K, Shimojima M, Kirikae T. 2014. Dissemination of 16S rRNA methylase ArmA-producing *Acinetobacter baumannii* and emergence of OXA-72 carbapenemase coproducers in Japan. *Antimicrob Agents Chemother* 58:2916-20.
254. Tan SY, Chua SL, Liu Y, Hoiby N, Andersen LP, Givskov M, Song Z, Yang L. 2013. Comparative genomic analysis of rapid evolution of an extreme-drug-resistant *Acinetobacter baumannii* clone. *Genome Biol Evol* 5:807-18.
255. Testing TECoAS. 2017. Breakpoint tables for interpretation of MICs and zone diameters. Version 7.1. <http://www.weu.cast.org/>.
256. Tian QB, Ohnishi M, Tabuchi A, Terawaki Y. 1996. A new plasmid-encoded proteic killer gene system: cloning, sequencing, and analyzing *hig* locus of plasmid Rts1. *Biochem Biophys Res Commun* 220:280-4.
257. Toh BEW, Paterson DL, Kamolvit W, Zowawi H, Kvaskoff D, Sidjabat H, Wailan A, Peleg AY, Huber CA. 2015. Species identification within *Acinetobacter calcoaceticus*–*baumannii* complex using MALDI-TOF MS. *Journal of Microbiological Methods* 118:128-132.
258. Towner KJ. 2009. *Acinetobacter*: an old friend, but a new enemy. *J Hosp Infect* 73:355-63.
259. Tsakris A, Ikonomidis A, Pournaras S, Tzouvelekis LS, Sofianou D, Legakis NJ, Maniatis AN. 2006. VIM-1 metallo- β -lactamase in *Acinetobacter baumannii*. *Emerg Infect Dis* 12:981-3.
260. Turton JF, Baddal B, Perry C. 2011. Use of the accessory genome for characterization and typing of *Acinetobacter baumannii*. *J Clin Microbiol* 49:1260-6.
261. Turton JF, Gabriel SN, Valderrey C, Kaufmann ME, Pitt TL. 2007. Use of sequence-based typing and multiplex PCR to identify clonal lineages of outbreak strains of *Acinetobacter baumannii*. *Clin Microbiol Infect* 13:807-15.
262. Turton JF, Kaufmann ME, Glover J, Coelho JM, Warner M, Pike R, Pitt TL. 2005. Detection and typing of integrons in epidemic strains of *Acinetobacter baumannii* found in the United Kingdom. *J Clin Microbiol* 43:3074-82.
263. Turton JF, Woodford N, Glover J, Yarde S, Kaufmann ME, Pitt TL. 2006. Identification of *Acinetobacter baumannii* by detection of the *bla*_{OXA-51-like} carbapenemase gene intrinsic to this species. *J Clin Microbiol* 44:2974-6.
264. Val ME, Bouvier M, Campos J, Sherratt D, Cornet F, Mazel D, Barre FX. 2005. The single-stranded genome of phage CTX is the form used for integration into the genome of *Vibrio cholerae*. *Mol Cell* 19:559-66.
265. Valenzuela JK, Thomas L, Partridge SR, van der Reijden T, Dijkshoorn L, Iredell J. 2007. Horizontal gene transfer in a polyclonal outbreak of carbapenem-resistant *Acinetobacter baumannii*. *J Clin Microbiol* 45:453-460.
266. Vila J, Marcos A, Marco F, Abdalla S, Vergara Y, Reig R, Gomez-Lus R, Jimenez de Anta T. 1993. In vitro antimicrobial production of β -lactamases, aminoglycoside-modifying enzymes, and chloramphenicol acetyltransferase by and susceptibility of clinical isolates of *Acinetobacter baumannii*. *Antimicrob Agents Chemother* 37:138-41.
267. Vila J, Ruiz J, Goni P, Jimenez de Anta T. 1997. Quinolone-resistance mutations in the topoisomerase IV *parC* gene of *Acinetobacter baumannii*. *J Antimicrob Chemother* 39:757-62.
268. Vila J, Ruiz J, Goni P, Marcos A, Jimenez de Anta T. 1995. Mutation in the *gyrA* gene of quinolone-resistant clinical isolates of *Acinetobacter baumannii*. *Antimicrob Agents Chemother* 39:1201-3.

269. Vila J, Ruiz J, Navia M, Becerril B, Garcia I, Perea S, Lopez-Hernandez I, Alamo I, Ballester F, Planes AM, Martinez-Beltran J, De Anta TJ. 1999. Spread of amikacin resistance in *Acinetobacter baumannii* strains isolated in Spain due to an epidemic strain. *J Clin Microbiol* 37:758-761.
270. Vila-Farres X, Ferrer-Navarro M, Callarisa AE, Marti S, Espinal P, Gupta S, Rolain JM, Giralt E, Vila J. 2015. Loss of LPS is involved in the virulence and resistance to colistin of colistin-resistant *Acinetobacter nosocomialis* mutants selected in vitro. *J Antimicrob Chemother* 70:2981-6.
271. Visca P, Seifert H, Towner KJ. 2011. *Acinetobacter* infection – an emerging threat to human health. *IUBMB Life* 63:1048-1054.
272. Weber BS, Ly PM, Irwin JN, Pukatzki S, Feldman MF. 2015. A multidrug resistance plasmid contains the molecular switch for type VI secretion in *Acinetobacter baumannii*. *Proc Natl Acad Sci U S A* 112:9442-7.
273. Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. *PLOS Comput Biol* 13:e1005595.
274. Wintjens R, Rooman M. 1996. Structural classification of HTH DNA-binding domains and protein-DNA interaction modes. *J Mol Biol* 262:294-313.
275. Woodford N, Ellington MJ, Coelho JM, Turton JF, Ward ME, Brown S, Amyes SG, Livermore DM. 2006. Multiplex PCR for genes encoding prevalent OXA carbapenemases in *Acinetobacter* spp. *Int J Antimicrob Agents* 27:351-3.
276. Wright GD. 2005. Bacterial resistance to antibiotics: enzymatic degradation and modification. *Adv Drug Deliv Rev* 57:1451-70.
277. Wright MS, Haft DH, Harkins DM, Perez F, Hujer KM, Bajaksouzian S, Benard MF, Jacobs MR, Bonomo RA, Adams MD. 2014. New insights into dissemination and variation of the health care-associated pathogen *Acinetobacter baumannii* from genomic analysis. *MBio* 5:e00963-13.
278. Yoon EJ, Courvalin P, Grillot-Courvalin C. 2013. RND-type efflux pumps in multidrug-resistant clinical isolates of *Acinetobacter baumannii*: major role for AdeABC overexpression and AdeRS mutations. *Antimicrob Agents Chemother* 57:2989-95.
279. Yoon EJ, Goussard S, Touchon M, Krizova L, Cerqueira G, Murphy C, Lambert T, Grillot-Courvalin C, Nemec A, Courvalin P. 2014. Origin in *Acinetobacter guillouiae* and dissemination of the aminoglycoside-modifying enzyme Aph(3')-VI. *MBio* 5:e01972-14.
280. Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O, Aarestrup FM, Larsen MV. 2012. Identification of acquired antimicrobial resistance genes. *J Antimicrob Chemother* 67:2640-4.
281. Zarrilli R, Vitale D, Di Popolo A, Bagattini M, Daoud Z, Khan AU, Afif C, Triassi M. 2008. A plasmid-borne *bla*_{OXA-58} gene confers imipenem resistance to *Acinetobacter baumannii* isolates from a Lebanese hospital. *Antimicrob Agents Chemother* 52:4115-20.
282. Zerbino DR, Birney E. 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res* 18:821-9.
283. Zhang G, Leclercq SO, Tian J, Wang C, Yahara K, Ai G, Liu S, Feng J. 2017. A new subclass of intrinsic aminoglycoside nucleotidyltransferases, ANT(3'')-II, is horizontally transferred among *Acinetobacter* spp. by homologous recombination. *PLoS Genet* 13:e1006602.

284. Zhang WJ, Lu Z, Schwarz S, Zhang RM, Wang XM, Si W, Yu S, Chen L, Liu S. 2013. Complete sequence of the *bla*_{NDM-1} carrying plasmid pNDM-AB from *Acinetobacter baumannii* of food animal origin. J Antimicrob Chemother 68:1681-2.
285. Zhou H, Zhang T, Yu D, Pi B, Yang Q, Zhou J, Hu S, Yu Y. 2011. Genomic analysis of the multidrug-resistant *Acinetobacter baumannii* strain MDR-ZJ06 widely spread in China. Antimicrob Agents Chemother 55:4506-12.
286. Zhu L, Yan Z, Zhang Z, Zhou Q, Zhou J, Wakeland EK, Fang X, Xuan Z, Shen D, Li QZ. 2013. Complete genome analysis of three *Acinetobacter baumannii* clinical isolates in China for insight into the diversification of drug resistance elements. PLoS One 8:e66584.
287. Zienkiewicz M, Kern-Zdanowicz I, Carattoli A, Gniadkowski M, Ceglowski P. 2013. Tandem multiplication of the IS26-flanked amplicon with the *bla*_(SHV-5) gene within plasmid p1658/97. FEMS Microbiol Lett 341:27-36.

APPENDIX I

Supplementary Tables and Figures

Table A1.1 Completed *A. baumannii* genomes as of July 2017

Isolate	Accession No.	GC	ST (IP)	Country	Submission year	Size (Mb)	G+C (%)	Genes	Plasmids	Sequencing technology	Isolation Date
ATCC 17978	CP000521	-	437	France	2007	4.00	38.9	3,469	2	454	1995
AB0057	CP001182	1	1	USA	2008	4.06	39.2	3,892	1	454/Sanger	2004
AB307-0294	CP001172	1	1	USA	2008	3.76	39	3,510	-	454	1994
ACICU	CP000863	2	2	Italy	2008	4.00	38.9	3,802	2	454	2005
AYE	CU459141	1	1	France	2008	4.05	39.3	3,879	4	Sanger	2001
SDF	CU468230	-	17	France	2008	3.48	39.1	3,664	3	Sanger	<2004
1656-2	CP001921	2	2	South Korea	2011	4.02	39.1	3,887	2	454	<2010
TCDC-AB0715	CP002522	2	2	Taiwan	2011	4.22	38.9	4,010	2	Illumina	2007-9
MDR-ZJ06	CP001937	2	2	China	2011	4.01	39	3,847	1	454	2006
MDR-TJ	CP003500	2	2	China	2012	4.15	39.1	4,014	2	454	<2011
TYTH-1	CP003856	2	2	Taiwan	2012	3.96	39	3,758	-	Illumina	2008
D1279779	CP003967	-	267	Australia	2013	3.71	39	3,523	1	454	2009
BJAB07104	CP003846	2	2	China	2013	4.04	39	3,882	2	Illumina HiSeq	2007
BJAB0715	CP003847	-	23	China	2013	4.05	38.9	3,899	1	Illumina HiSeq	2007
BJAB0868	CP003849	2	2	China	2013	4.01	38.9	3,822	3	Illumina HiSeq	2008
ZW85-1	CP006768	-	639	China	2013	3.93	39.2	3684	2	454	?
NCGM 237	AP013357	2	2	Japan	2013	4.02	39.1	3894	-	PacBio	2012
AC29	CP007535	2	2	Malaysia	2014	3.85	39.0	3659	-	Illumina	2011
AC30	CP007577	2	2	Malaysia	2014	3.92	38.8	3729	3	Illumina	2011
AB031	CP009256	-	638	Canada	2014	3.80	39.2	3599	-	PacBio	2010
AB030	CP009257	-	79	Canada	2014	4.34	39.00	4339	-	PacBio	2010
LAC-4	CP007712	-	10	USA	2014	3.97	38.9	3821	2	454/Illumina MiSeq/HiSeq	1997
IOMTU 433	AP014649	-	622	Nepal	2014	4.19	39.10	3999	1	PacBio	2013
A1	CP010781	1	1	UK	2015	3.91	39.30	3735	1	Illumina HiSeq/PacBio	1982
6411	CP010368	-	322	USA	2015	4.0	38.80	3773	2	PacBio	<2014
6200	CP010397	-	464	Colombia	2015	4.07	39.20	3876	3	PacBio	2012
AB5075-UW	CP008706	1	1	USA	2015	4.07	39.10	3929	3	PacBio	2008
XH386	CP010779	2	2	China	2014	4.08	39.16	4062	-	Illumina/PacBio	2015
B8300	CP021347	-	- ¹	India	2015	3.82	39.30	3711	1	PacBio	2014
Ab04-mff	CP012006	-	10	Canada	2015	4.19	38.92	4067	2	Illumina/PacBio	2012
ATCC17978-mff	CP012004	-	437	-	2015	4.06	38.94	3796	1	Illumina/PacBio	-
AbH12O-A2	CP009534	-	79	Spain	2014	3.88	39.10	3673	-	Illumina HiSeq	<2014
D36	CP012952	1	81	Australia	2015	4.13	39.02	3974	4	PacBio	2008
KPN10P02143	CP013924	2	2	South Korea	2016	4.13	39.07	4006	1	PacBio	2012
YU-R612	CP014215	2	2	South Korea	2016	4.16	38.99	4035	2	PacBio	2014

XH860	CP014538	2	2	China	2016	3.85	39.00	3673	-	PacBio	2009
XH859	CP014539	2	2	China	2016	3.98	39.10	3786	-	PacBio	2009
XH857	CP014540	-	215	China	2016	3.88	39.00	3702		PacBio	2010
XH856	CP014541	2	2	China	2016	3.91	39.10	3716	-	PacBio	2010
XH858	CP014528	-	23	China	2016	4.17	38.90	4020	-	PacBio	2010
3207	CP015364		422	Mexico	2016	4.09	38.99	3788	2	Illumina/PacBio	2008
KAB01	CP017642	2	2	South Korea	2016	4.03	38.90	3894	1	PacBio	2015
KAB02	CP017644	2	2	South Korea	2016	4.09	38.99	3980	1	PacBio	2015
KAB03	CP017646	2	2	South Korea	2016	4.06	38.90	3920	1	PacBio	2015
KAB04	CP017648	2	2	South Korea	2016	4.13	38.96	4018	1	PacBio	2015
KAB05	CP017650	2	2	South Korea	2016	4.04	39.00	3925	1	PacBio	2015
KAB06	CP017652	2	2	South Korea	2016	4.04	39.00	3925	1	PacBio	2015
KAB07	CP017654	2	2	South Korea	2016	4.05	39.00	3944	1	PacBio	2015
KAB08	CP017656	2	2	South Korea	2016	4.07	38.89	3963	1	PacBio	2015
HRAB-85	CP018143	2	2	China	2016	4.10	39.01	3977	1	PacBio	2014
AF-637	CP018256	2	2	USA	2016	4.09	39.07	3930	1	PacBio	2008
AF-401	CP018254	-	79	Mexico	2016	4.27	39.09	4153	1	PacBio	2009
XDR-BJ83	CP018421	2	2	China	2016	4.08	38.91	3942	1	PacBio	2007
CMC-CR-MDR-Ab4	CP016295	2	2	USA	2017	4.16	38.97	4021	2	PacBio	<2016
CMC-MDR-Ab59	CP016298	2	2	USA	2017	4.08	39.07	3910	1	PacBio	<2016
CMC-CR-MDR- Ab66	CP016300	2	2	USA	2017	4.17	38.98	4045	2	PacBio	<2016
AB042 ²	CP019034	-	437	-	2017	4.00	39.00	3838	-	PacBio	-
JBA13	CP020584	2	2	South Korea	2017	4.07	39.00	3943	2	PacBio	2013
CBA7	CP020586	2	2	South Korea	2017	4.11	39.17	4030	1	PacBio	2013
15A34	CP020590	2	2	South Korea	2017	3.94	39.00	3799	1	PacBio	2014
15A5	CP020574	2	2	South Korea	2017	4.05	38.89	3916	2	PacBio	2013
USA2	CP020592	2	2	South Korea	2017	3.99	38.90	3825	1	PacBio	2013
SSA6	CP020591	2	2	South Korea	2017	3.88	39.00	3703	-	PacBio	2013
HWBA8	CP020597	-	25	South Korea	2017	4.27	39.06	4115	1	PacBio	2013
WKA02	CP020598	-	10	South Korea	2017	3.87	38.90	3734	-	PacBio	2013
USA15	CP020595	1	1	South Korea	2017	4.03	39.08	3896	1	PacBio	2013
SAA14	CP020579	1	1	South Korea	2017	3.97	39.10	3823	-	PacBio	2013
SSA12	CP020578	2	2	South Korea	2017	4.06	38.99	3957	2	PacBio	2013
SSMA17	CP020581	2	2	South Korea	2017	4.02	38.99	3906	1	PacBio	2013
Ab736	CP015121	-	52	USA	2017	3.99	39.19	3820	1	PacBio	2015
A85	CP021782	1	1	Australia	2017	4.14	39.18	4068	3	PacBio/Illumina HiSeq	2003

¹*fusA-3, gltA-5, pyrG-26, recA-6, rplB-2, rpoB-3*

²Lab mutant of ATCC 17978 that overexpresses the AdelJK efflux pump

Table A1.2 Results of CDS disk diffusion for GC2 isolates from Singapore General Hospital

Strain	Year	Annular radius of zone of inhibition																													
		IPM	MER	Dor	Ap	SAM	TIM	TZP	CTX	CAZ	CRO	FEP	Sm	Sp	Su	Tc	Mc	Dc	Tp	Cl	Fl	Km	Nm	Gm	Ne	Ak	Tm	Nx	CIP	LVX	Rif
S21	1996	4	0		0	3	0		0	1			0	0	0	0			0	0	0	0	1	3	7	1	8	0	0		7.0
		4	3	0		2	0	2			0	0					5	1												2.5	5.0
		3	0		0	2.5	0		0	1			0	0	0	0			0	0	0	0	0	2	6	0	7	0	0		7.0
		3.5	0.5	0		3	0	2			0	0					4	1												2	6.0
S22	2004	1	1		0	2.5	0		0	0			0	0	0	0			0	0	0	0	2	0	0	0	0	0	0		6.0
		2	1	0		2	0	1			0	0					4.5	1												0	6.5
		1	1		0	2.5	0		0	0			0	0	0	0			0	1	0	0	2	0	0	0	0	0	0		6.5
		2	1	0		2.5	0	1			0	0					5	1												0	7.0
S23	2004	3	3		0	3	0		0	5			0	0	0	0			0	0	0	0	3	0	0	0	0	0	0		8.0
		2	1.5	0		3	0	2			0	0					4.5	0.5												0	8.0
		3	3		0	3.5	0		0	5			0	0	0	0			0	0	0	0	3	0	0	0	0	0	0		8.0
		1.5	1.5	0		2	0	2			0	0					5	0.5												0	8.5
S24	2006	2	1		0	2.5	0		0	1			0	0	0	0			0	1	0	0	2	3	7	3	7	0	0		7.0
		1	1	0		2	0	1			0	0					4.5	1												0	6.5
		1	1		0	2.5	0		0	2			0	0	0	0			0	0	0	0	2	2	6	2	7	0	0		7.0
		2	1.5	0		2	0	1			0	0					4.5	1												0	6.0
S25	2006	1	2		0	2.5	0		0	1			0	0	0	0			0	0	0	0	7	0	0	0	0	0	0		6.0
		2.5	2	0		1.5	0	2			0	0					4	1												0.5	5.0
		2	2		0	2.5	0		0	1			0	0	0	0			0	0	0	0	7	0	0	0	0	0	0		6.0
		2.5	2	0		2	0	2			0	0					4.5	1												0.5	6.0
S26	2006	3	3		0	2.5	0		0	1			0	0	0	0			0	1	0	0	2	2	6	2	7	0	0		7.0
		3.5	2	0		2	0	2			0	2					5.5	1												2.5	6.0
		2	2		0	2.5	0		0	1			0	0	0	0			0	1	0	0	2	2	6	2	7	0	0		8.5
		3.5	2	1		2	0	2			0	0					4.5	1												3	6.5

Black performed by M. Wynn and Red performed in this project

Strain	Year	Annular radius of zone of inhibition																													
		IPM	MER	Dor	Ap	SAM	TIM	TZP	CTX	CAZ	CRO	FEP	Sm	Sp	Su	Tc	Mc	Dc	Tp	Cl	Fl	Km	Nm	Gm	Ne	Ak	Tm	Nx	CIP	LVX	Rif
S27	2007	1	1		0	3	0		0	0			0	0	0	0			0	0	0	0	1	0	0	0	0	0	0		6.0
		1	1	1.5		3	0	1			0	1.5					5	1												0.5	6.5
		1	1		0	4	0		0	0			0	0	0	0			0	0	0	0	2	0	0	0	0	0	0		8.0
		1	0.5	1		3	0	1.5			0	0					4	1												0.5	6.0
S28	2007	1	1		0	0	0		0	1			0	0	0	0			0	0	0	0	6	0	0	0	0	0	0		6.0
		1.5	1	1		2	0	2			0	1.5					5.5	2												0	6.0
		1	1		0	2	0		0	2			0	0	0	0			0	0	0	0	6	0	0	0	0	0	0		6.0
		1	0	0		2.5	0	1.5			0	0					4	1												0.5	6.5
S29	2007	3	2		0	2.5	0		0	2			0	0	0	0			0	0	0	0	2	2	6	2	7	0	0		7.0
		3	2	0		2	0	2			0	0					4	1												2.5	6.0
		3	2		0	2.5	0		0	1			0	0	0	0			0	0	0	0	2	2	6	2	7	0	0		7.0
		3	2	0		2.5	0	2			0	2					5	1												2.5	6.0
S30	2008	2	2		0	2.5	0		0	2			0	0	0	0			0	0	0	0	3	3	6	3	8	0	0		6.5
		1	1	1.5		2	0	2			0	1.5					5	1												1	5.5
		1	1		0	2	0		0	2			0	0	0	0			0	0	0	0	3	3	6	2	8	0	0		6.0
		1.5	1	0		2	0	2			0	2.5					5	1												1.5	6.0
S31	2008	1	1		0	3.5	0		0	2			0	0	0	0			0	0	0	0	1	0	0	0	0	0	0		6.0
		0.5	0.5	1		2.5	0	1.5			0	1.5					3.5	1												2	6.0
		1	1		0	3.5	0		0	3			0	0	0	0			0	0	0	1	0	1	0	0	0	0	0		6.0
		1	0.5	0		3	0	1			0	1.5					4	1												2.5	6.0
S32	2009	0	0		0	0	0		0	0			0	0	0	0			0	0	1	0	2	0	0	0	0	0	0		6.0
		0	0	0		1	0	0			0	0					5	1												2	7.0
		0	0		0	0	0		0	0			0	0	0	0			0	0	0	0	2	0	0	0	0	0	0		6.0
		0	0	0		1	0	0			0	0					4	1												2	6.0

Black performed by M. Wynn and Red performed in this project

Strain	Year	Annular radius of zone of inhibition																														
		IPM	MER	Dor	Ap	SAM	TIM	TZP	CTX	CAZ	CRO	FEP	Sm	Sp	Su	Tc	Mc	Dc	Tp	Cl	Fl	Km	Nm	Gm	Ne	Ak	Tm	Nx	CIP	LVX	Rif	
S33	2009	1	0		0	0	0		0	0			0	0	0	0			0	1	1	0	1	0	0	0	0	0	0		6.0	
		0.5	1	0.5		1	0	1			0	0.5					4.5	1.5												3	6.0	
		1	1		0	0	0		0	0			0	0	0	0			0	1	0	0	1	0	0	0	0	0	0		6.0	
		0.5	0.5	1		1.5	0	1			0	1.5					5	1.5													2.5	6.0
S34	2009	1	1		0	0	0		0	1			1	6	0	0			0	1	0	0	3	0	0	0	0	0	0	1		7.0
		0.5	0	0.5		1	0	1			0	2					6	2.5													4	6.0
		1	1		0	0	0		0	1			2	7	0	0			0	1	1	0	4	0	0	0	0	0	0	1		7.0
		1	1	1		1	0	1			0	1.5					5.5	1.5													4	6.5
S35	2010	2	2		0	0	0		0	0			0	0	0	0			0	1	1	0	8	0	0	0	0	0	0		0.0	
		1	1	1		2	0	2			0	1					5	0.5													1.5	0.0
		2	2		0	4	0		0	0			0	0	0	0			0	1	1	0	9	0	0	0	0	0	0		0.0	
		1	1	1		3.5	0	2			0	2					4	1													2	0.0
S36	2010	1	1		0	0	0		0	0			0	0	0	0			0	1	0	0	9	0	0	0	0	0	0		6.0	
		0.5	0.5	1		3	0	2			0	1.5					4	1													2	6.0
		2	2		0	3.5	0		0	0			0	0	0	0			0	0	0	0	8	0	0	0	0	0	0		7.0	
		1	0.5	2		4	0	2			0	1					5	1													1.5	7.0
S37	2010	4	3		0	3.5	0		0	0			0	0	0	0			0	1	0	0	12	0	0	0	0	0	0		7.0	
		2	1	1		4.5	0	3			0	4					5.5	1													3	8.0
		4	4		0	4.5	0		0	1			0	0	0	0			0	1	0	0	12	0	0	0	0	0	0		7.5	
		2	1.5	1		4	0	2.5			0	5					4.5	1													3	6.5
S38	2011	2	2		0	4.5	0		0	2			0	0	0	0			0	1	0	0	8	0	0	0	0	0	0		6.0	
		2	2	2		4	0	2			0	1.5					5	2													2	5.0
		2	2		0	4.5	0		0	2			0	0	0	0			0	0	0	0	8	0	0	0	0	0	0		6.0	
		1.5	1	2		3.5	0	2			0	1.5					4	1													2	5.5

Black performed by M. Wynn and Red performed in this project

Strain	Year	Annular radius of zone of inhibition																													
		IPM	MER	Dor	Ap	SAM	TIM	TZP	CTX	CAZ	CRO	FEP	Sm	Sp	Su	Tc	Mc	Dc	Tp	Cl	Fl	Km	Nm	Gm	Ne	Ak	Tm	Nx	CIP	LVX	Rif
S39	2011	2	2		0	5	0		0	2			0	0	0	0			0	0	0	0	8	0	0	0	0	0	0		7.0
		2	1.5	1		4	0	2			0	2					5	1											2	5.5	
		2	2		0	4	0		0	2			0	0	0	0			0	0	0	0	8	0	0	0	0	0	0		6.0
		2	1.5	1.5		4	0	2			0	2					4	1											2.5	6.0	
S40	2011	1	0		0	0	0		0	0			0	0	0	0			0	0	0	0	2	0	0	0	0	0	0		5.5
		1	0.5	1		2	0	2			0	0.5					5	1.5											0.5	5.5	
		1	0		0	0	0		0	0			0	0	0	0			0	0	0	0	2	0	0	0	0	0	0		5.5
		0.5	0	1		1.5	0	1.5			0	0.5					3	1.5										0.5	5.0		

Black performed by M. Wynn and Red performed in this project

Table A1.3. Isolates sequenced for this study by HiSeq

Isolate name	Source	Year	Mean read depth	No. contigs	Assembly size (bp)	N50 (bp)	Accession No. (reads)
S21	Unknown	1996	93.5	110	4,170,306	173,010	ERR246920
S24	Trach	2006	85.9	113	4,092,060	184,961	ERR246931
S27	Blood	2007	80.7	64	4,061,936	220,457	ERR246908
S29	Blood	2007	69.7	108	4,079,807	173,031	ERR246924
S34	Blood	2009	88.2	90	3,969,270	122,571	ERR246957
S38	Sputum	2011	87.1	66	4,001,907	220,529	ERR246919
S39	Sputum	2011	89	66	4,001,565	291,078	ERR246911

Table A 1.4. Isolates sequenced for this study by MiSeq

Isolate name	Source	Year	Mean read depth	No. contigs	Assembly size (bp)	N50 (bp)	Accession No. (assemblies)
S22	ETTA	2004	72.850	71	4,006,455	220,589	PYEO00000000
S23	ETTA	2004	62.895	68	4,003,659	220,582	PYEN00000000
S25	ETTA	2006	64.600	66	4,077,806	205,040	PYEM00000000
S26	ETTA	2006	71.885	79	4,090,882	162,244	PYEL00000000
S28	Blood	2007	65.094	76	4,125,741	205,040	PYEK00000000
S30	ETTA	2008	67.568	62	4,055,552	262,554	PYEJ00000000
S31	Trach	2008	50.688	56	4,008,413	220,565	PYEI00000000
S32	Blood	2009	71.146	75	3,959,824	162,580	PYEH00000000
S33	Blood	2009	65.782	62	3,923,847	168,007	PYEG00000000
S35	Sputum	2010	72.473	55	3,830,627	177,675	PYEF00000000
S36	Blood	2010	81.878	58	3,949,724	168,007	PYEE00000000
S37	ETTA	2010	93.141	73	3,895,702	172,149	PYED00000000
S38	Blood	2011	67.435	76	4,014,692	162,581	PYEC00000000

Isolates	Year	MLST (Oxford)	ISAbal positions																												Total																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																														
			15	16	17	30	213	215	218	222	224	227	297	234	235	236	115	244	218	299	300	301	302	303	304	317	298	212	216	220		233	306	223	228	307	311	312	316	82	103	133	135	149	168	169	195	257	263	265	266	214																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																									
WM99c	1999	208																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																											

Figure A1.1. Unique positions of ISAbal1. The coloured boxes indicate ISAbal1 positions. The Singapore isolates highlighted in pink lack *armA* and contain *aphA6*, while those highlighted in green have *armA* and lack *aphA6*.

Table A1.5 Positions of ISAbal in XH386

Position	DR	Location ¹	Left ID	Right ID
1	gattagctc	1152001-1152009 and 1153190- 1153198	TE32_05280	TE32_05295
2	aatttttg	1968340-1968348 and 1969529- 1969537	TE32_09325	TE32_09340
3	aaaacaatt	2012130-2012138 and 2013319- 2013327	TE32_09525	TE32_09540
4	ctaattttt	2679800-2679808 and 2680989- 2680997	TE32_12810	TE32_12825
5	aataataat	3827583-3827591	TE32_18435	TE32_18450
6	ataattatt	1725127-1725135	TE32_08240	TE32_08245
7	tatttacta	2209351-2209359	TE32_10555	TE32_10555
8	ttaattcta	2888103-2888111	TE32_13960	TE32_13965
9	caaaattga	3496575-3496583	TE32_16880	TE32_16880
10	attgtttgg	3733796-3733804	TE32_17980	TE32_17980
11	aaaaatgtc	2352487-2352495	TE32_11285	TE32_11285
12	aaaaggatg	2516673-2516681 and 2517862- 2517870	TE32_12045	TE32_12060
13	aaaaaataa	3145756-3145764	TE32_15250	TE32_15255
14	ttaaaaata	1342037-1342045	TE32_06310	TE32_06315
15	gaataattt	1431965-1431973	TE32_06745	TE32_06750
16	caactcaat	1798506-1798514	TE32_08585	TE32_08585
17	ttagaattt	2464309-2464317	TE32_11790	TE32_11795
30	tcttttaaa	1662070-1662078 and 1663259- 1663267	TE32_07900	TE32_07925
58	aaataatgc	639171-639179 and 640360-640368	TE32_03070	TE32_03095
62	catcttaag	1527476-1527484 and 1528665- 1528673	TE32_07175	TE32_07190
82	aaaataaac	173320-173328 and 174509-174517	TE32_00770	TE32_00790
97	tattttttt	688354-688362	TE32_03320	TE32_03320
103	atttcattt	921116-921124 and 922305-922313	TE32_04240	TE32_04265
124	ctcaaattt	1515941-1515949	TE32_07115	TE32_07115
133	ataagcttc	1921632-1921640	TE32_09115	TE32_09115
135	ataattttc		TE32_09245	TE32_09265
148	ttctttttt	2242741-2242749	TE32_10745	TE32_10750

149	ttaatcttt	2266415-2266423 and 2267604- 2267612	TE32_10855	TE32_10870
165	aaaaaccat	2917716-2917724	TE32_14110	TE32_14115
168	aaaatattt	3022078-3022086	TE32_14575	TE32_14590
169	aaaatattt	3030511-3030519	TE32_14625	TE32_14640
178	ctattcata	3394522-3394530	TE32_16405	TE32_16405
182	aaaaaattt	3524825-3524833	TE32_17015	TE32_17020
195	acctttaaa	3998431-3998439 and 3999620- 3999628	TE32_19235	TE32_19250
206	atttaatga	1148871-1148879	TE32_05270	TE32_05275
207	aaataaaat	763919-763927	TE32_03680	TE32_03685
208	caaagattt	1908965-1908973	TE32_09040	TE32_09040
209	attactttt	2060729-2060737	TE32_09750	TE32_09755
210	tttttttc	2516016-2516023	TE32_12040	TE32_12040
211	tactttttt	2752368-2752376	TE32_13150	TE32_13155
212	ttaaaactc	83132-83140	TE32_00380	TE32_00380
213	gctataaat	353989-353997	TE32_01640	TE32_01645
214	acaaaagac	572546-572554	TE32_02765	TE32_02770
215	attgatttt	57796-577934	TE32_02795	TE32_02800
216	atatttttt	970487-970495	TE32_04520	TE32_04525
217	aaaaattat	1017606-1017614	TE32_04735	TE32_04735
218	attgaaaat	1156758-1156766	TE32_05305	TE32_05305
219	gtggttttt	1157455-1157463	TE32_05305	TE32_05305
220	attttggca	1186155-1186163	TE32_05425	TE32_05425
221	cataatttt	1680070-1680078	TE32_07995	TE32_08000
222	taccgcaa	2094052-2094060	TE32_09935	TE32_09935
223	aactgcttt	2098502-2098510	TE32_09960	TE32_09960
224	aacgttttt	2161845-2161853	TE32_10235	TE32_10235
225	aagtcttat	2268556-2268564	TE32_10870	TE32_10875
226	gaaaattag	2889722-2889730	TE32_13970	TE32_13975
227	ttaatattt	3154882-3154890	TE32_15290	TE32_15295
228	aaaaaatta	3994561-3994569	TE32_19225	TE32_19225
229	taccgattg	2947339-2947347	TE32_14240	TE32_14240
230	aaaactttc	3535574-3535582	TE32_17055	TE32_17060
231	aattaatat	3807230-3807238	TE32_18340	TE32_18345
232	atatctttt	3897581-3897589	TE32_18775	TE32_18780
233 ²	aaaatttat			
234	ttaatttta	595524-595532	TE32_02875	TE32_02875
235	aagaattac	1148296-1148304	TE32_05265	TE32_05270
236	atttggttt	2730078-2730086	TE32_13065	TE32_13065
244	atttttttt	374168-374176	TE32_01740	TE32_01745
257	aacaatttt	250397-250405	TE32_01130	TE32_01130

263	tgatactag	1739594-1739602	TE32_08300	TE32_08300
265	tacatat	1919375-1919383	TE32_09090	TE32_09095
294	aaaatttat	299616-299624	TE32_01395	TE32_01400
295	aaaaattaa	1334849-1334857	TE32_06265	TE32_06265
296	attcttaac	1984284-1984292	TE32_0935	TE32_09400
297	aacttactt	552833-552841	TE32_02685	TE32_02685
298	ttcgggatg	1165893-1165901	TE32_05335	TE32_05335
299	agcttttta	1201188-1201196	TE32_05480	TE32_05485
300	aaaatactg	2373350-2373358	TE32_11385	TE32_11385
301	aaagaattt	2507052-2507060	TE32_11995	TE32_11995
302	ctaatttaa	2870306-2870314	TE32_13890	TE32_13890
303	aaacgtcca	2930084-2930092	TE32_14170	TE32_14170
304	taaaaattt	3051823-3051831	TE32_14725	TE32_14725
305	ataatttca	3359535-3359543	TE32_16245	TE32_16245
306	ctatttctt	209269-209277	TE32_00930	TE32_00935
307	tttaattat	552617-552625	TE32_02685	TE32_02685
308	aaatttga	1018337-1018344	TE32_04735	TE32_04740
309	aattttaag	1342703-1342711	TE32_06315	TE32_06315
310	aataaattt	3798732-3798740	TE32_18295	TE32_18300
311	ctaaatttc	118374-118382	TE32_00535	TE32_00535
312	gtagtcact	721842-721850	TE32_03480	TE32_03485
313	aataatcta	876395-876403	TE32_04055	TE32_04060
314	aattttttt	3119804-3119812	TE32_15115	TE32_15115
315	taataaaa	3364074-3364081	TE32_16265	TE32_16265
316	tgtaaaaaat	2358980-2358988	TE32_11305	TE32_11310
317	aattgtttt	2542732-2542740	TE32_12160	TE32_12160

¹In XH386 (GenBank accession number CP010779)

²Region not in XH386

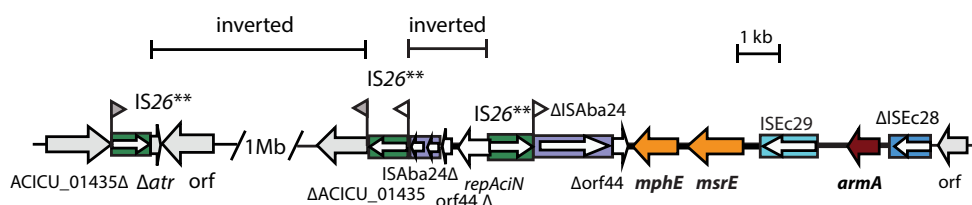
APPENDIX II

Additional assemblies

A2.1 AbGRI3-7 in D86

The *armA* region of D86 was assembled using a similar strategy to that used in Chapter 6, by first retrieving IS26 bounded contigs and using combinations of primers to order the contigs. D86 contains the same characteristic deletion caused by recombination as AbGRI3-4. However, the left end of AbGRI3 in D86 is different (Fig A2.1A) and this variant has been named AbGRI3-7. Firstly, the fragment of the *atr* gene usually bounding the left junction of AbGRI3 was found to be located approximately 1 Mb away, adjacent to an IS26** that has a fragment of an orf corresponding to ACICU_01435 in the genome of ACICU2 on the other side (Fig A2.1A). The remainder of ACICU_01435 was found to be adjacent to the left end of AbGRI3 in D86 (Fig A2.1). It is likely that the IS26** at the left of the island targeted ACICU_01435 and through intramolecular transposition inverted 1 Mb of the chromosome (red segment in Fig A2.1B), duplicating itself and causing an 8 bp TSD (grey flags in Fig A2.1A and B). A small inversion is also present in the left end of AbGRI3 in D86, where again the IS26 on the left has targeted ISAb24, inverting the 1.9 kb of intervening sequence, duplicating IS26 and producing a TSD (green flags in Fig A2.1B).

A.



B.

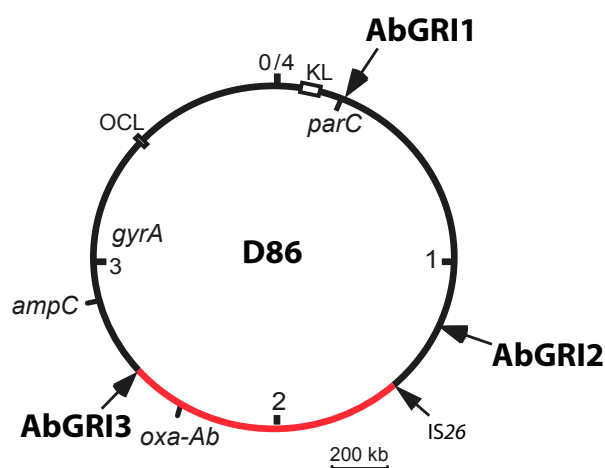


Figure A2.1 AbGRI3 in D86 and chromosomal inversion. A. Structure of AbGRI3 in D86. B. Schematic of the chromosome of D86 showing the location of the inversion. For A. arrows indicate the extent and orientation of genes and open reading frames. The *mphE* and *msrE* genes are orange, the *armA* gene is maroon chromosomal genes are shown in grey. Boxes represent insertion sequences and the internal arrow indicates the orientation of the transposase or the rolling circle replicase. IS26 is dark green, ISAbA24 is purple, ISEc29 and ISEc28 are light and dark blue and CR1 is yellow. IS26** is 2 bp different to the standard sequence of IS26. Green and grey flags indicate matching TSD. In B., the red segment indicates the 1 Mb inversion and black arrows indicate the approximate locations of the resistance islands and the extra IS26.

A2.2 RepAci6 plasmid assembly

A2.2.1 Type 1 plasmid assembly of pS29-1 and pS26-1.

The contigs matching RepAci6 plasmids in the draft genomes of S26 and S29 were retrieved, analysed (Table A2.1) and tiled to the type 1 backbone.

Table A2.1 Contigs matching RepAci6 plasmids in S26 and S29

Isolate	Contig number	Contig length (bp)	Description	Contig edges ²	
				Left	Right
S29 ¹	12	4,464	F3	R2	R3
	23	11,441	F2	R1	R2
	25	11,970	F1, <i>trw</i> genes	R3	ISAbal25
	29	45,852	F1 (extra 10 kb), <i>tra</i> region, <i>repAci6</i>	IRr of Tn6022	R1
	39	2,235	F1, <i>res</i>	ISAbal25	IRl of Tn6022
S26 ²	20	62,163	F1 (<i>tra</i> region, <i>repAci6</i>), F2, F3	IRr of Tn6022	repeat 3
	38	17,312	F1 (<i>trw</i> genes, TnaphA6)	Repeat 3	IRl of Tn6022
	87	424	R2 and R3 sequence		

¹Sequenced by Illumina HiSeq.²R1, R2 and R3 stand for repeat 1, 2 and 3, respectively.

The pS29-1 contigs broke at repeat 2, 3 and the insertion site of TnaphA6 which have already been mapped (Chapter 8; Section 8.2). The reason for the lack of amplicon of repeat 1 is due to an additional 9.3 kb of sequence inserted between the location of annealing of RH1398 and the first repeat. The primer RH2030 was designed closer to the end of contig 29 and used with RH1399 to amplify a 1,069 bp region. This new combination of primers produced an amplicon of the expected size (Fig A2.2).

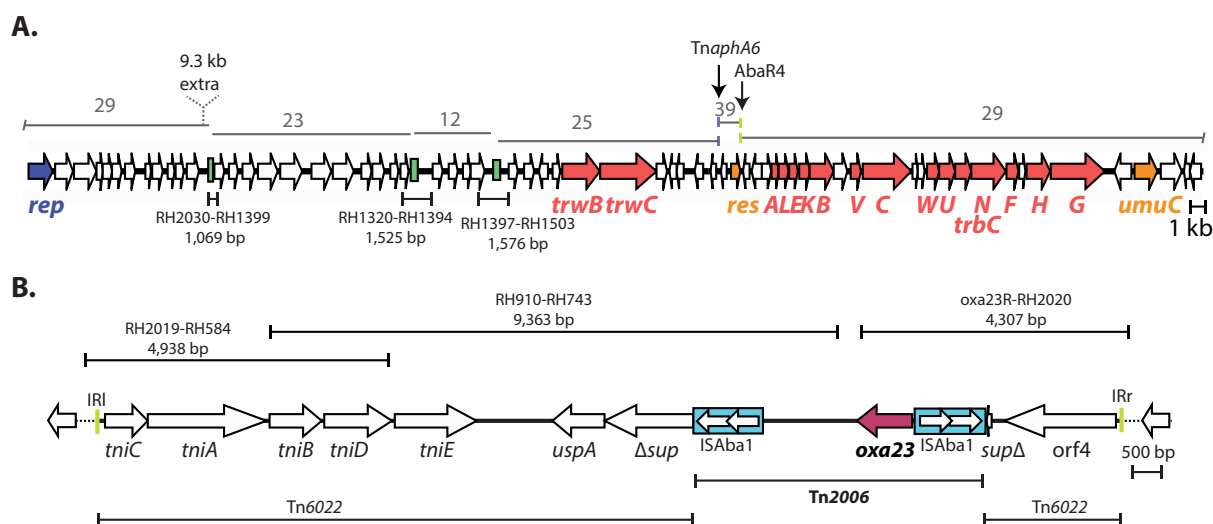


Figure A2.2 Assembly of pS29-1 contigs. A. pS29-1 contigs tiled to the type 1 backbone. Arrows indicate the extent and orientation of genes and open reading frames. The *repAci6* gene is blue, genes involved in conjugal transfer are red and the *res* and *umuC* genes are orange. Green boxes indicate the three repeat sequences. Grey lines above indicate the extent of pS26-1 contigs. The black bars underneath show the primers used for assembly along with PCR amplicon size. Vertical arrows indicate the insertion location of *AbaR4* and *TnaphA6*. B. Assembly of *AbaR4* in pS29-1. The *oxa23* gene is magenta and the IR sequences of Tn6022 are shown in olive green. The blue box is IS_{Aba1} and the internal arrows indicate the transposase orientation. The primers used to span *AbaR4* are shown above by the black bars, along with PCR amplicon size. The lines below indicate the transposons.

An additional break in the assembly was located just downstream to *TnaphA6*. Contig 39 and 29 had the left and right IR of Tn6022 on one edge and a 5 bp target site duplication (5' AATTG 3') was adjacent to the IR sequence on each contig. Again, this may be a new insertion site of *AbaR4*. New primers RH2019 and RH2020 were used in combination with RH583 or *oxa23R* to amplify a 4,983 or 4,307 bp product respectively (Fig A2.2B). The amplicons produced were of the expected sizes and were sequenced. The internal *AbaR4* PCR, RH910 and RH743, which should produce an amplicon 9,363 bp in size was performed and a product of the expected size was generated. pS29-1 is 96,201 bp in size and contains *TnaphA6*, *AbaR4* in a new location as well as an extra 9.3 kb of unknown sequence (Fig A2.2, bottom line).

The assembly of pS26-1 fragmented at repeat 3 which has been previously screened (Chapter 8, Section 8.2), and at the IR sequences of Tn6022. The backbone sequences flanking these IR sequences are identical to that of the *AbaR4* location in pS29-1. Hence, the same primer

combination of RH2019-RH584 and oxa23R and RH2020 were used to confirm that AbaR4 was present in this location (Fig A2.2). Contig 38 also included *TnaphA6* in the pAb-G7-2 location. A brief look at the sequence prior to repeat 1 revealed that pS26-1, like S29-1, also contained 9.3 kb of extra sequence, accounting for the lack of amplification of PCR screening over repeat 1 (Chapter 8, Section 8.1).

A2.2.2 Assembly of Type 2 plasmids

A2.2.2.1 Assembly of pS27-2 and pS34-2 from HiSeq sequencing

The contigs matching RepAci6 plasmids in the draft genomes of S27 and S34 were retrieved, analysed (Table A2.2) and tiled to the type 2 backbone.

Table A2.2 Contigs matching RepAci6 plasmids in S27 and S34

Isolate	Contig number	Contig length (bp)	Kmer coverage	Description	Contig edges ¹	
					Left	Right
S27	9	5,220	29.2	F2	R1	ISAbal
	19	40,921	28.9	F1, <i>tra</i> region, <i>repAci6</i>	IRr of Tn6022	R1
	26	4,675	29.1	F3	R2	R3
	32	8,619	29.3	F2	ISAbal25	R2
	50	9,882	29.8	F1, <i>trw</i> genes	R3	IRl of Tn6022
	78	278	61.8	Repeat sequence		
S34	8	13,024	29.8	F1, <i>repAci6</i>	IRr of Tn6022	R1
	22	37,778	30.1	F1, <i>trw</i> genes, <i>res</i> , <i>tra</i> region	R3	IRl of Tn6022
	23	4,672	29.7	F3	R2	R3
	32	13,669	30.1	F2	R1	R2
	86	277	60.5	Repeat sequence		

²R1, R2 and R3 stand for repeat 1, 2 and 3, respectively.

The contigs retrieved from the draft genome of S27 were mapped to the type 2 backbone (grey lines in Figure A2.3A). In addition to breaking at the repeat sequences (each of which were screened by PCR earlier) the assembly of pS27-2 also fragmented on a copy of ISAbal. Contig

9 and 32 each had a fragment of ISAbal at one edge (Fig A2.3A) and examination of the sequence adjacent to the ISAbal fragments revealed a matching 9 bp duplication (5' AAGTAGTGA 3'). New primers, RH2026 and RH2027, were designed to amplify this region and were predicted to amplify a 1,766 bp product. The amplicon generated was of the predicted size and was sequenced.

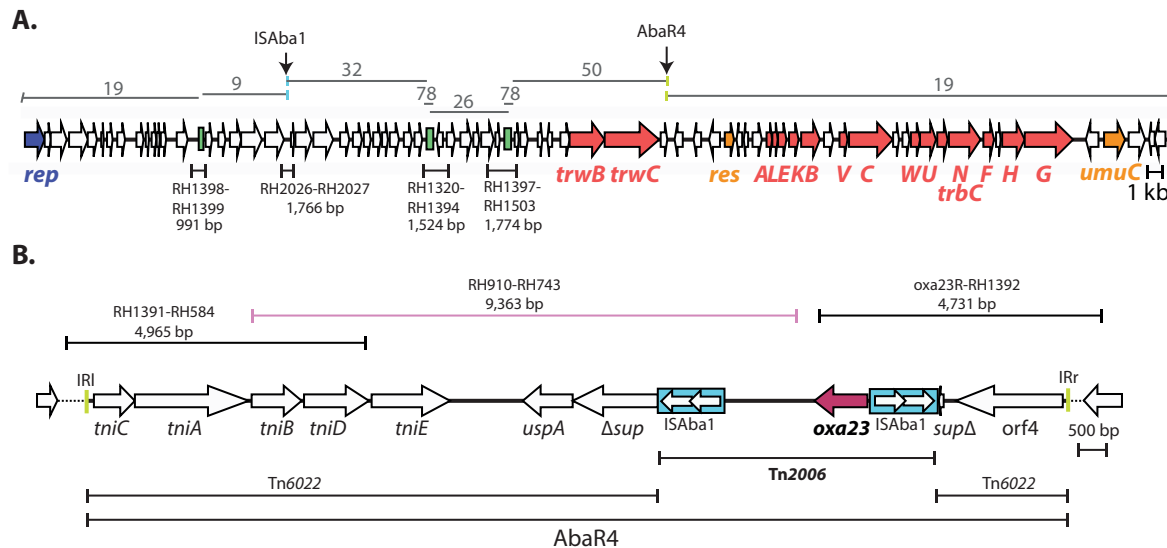


Figure A2.3 Assembly of pS27-2 contigs. A. pS34-2 contigs tiled to the type 2 backbone. Arrows indicate the extent and orientation of genes and open reading frames. The *repAci6* gene is blue, genes involved in conjugal transfer are red and the *res* and *umuC* genes are orange. Green boxes indicate the three repeat sequences. Grey lines above indicate the extent of different pS34-2 contigs. The black bars underneath show the primers used for assembly along with PCR amplicon size. Vertical arrows indicate the insertion location of AbaR4 and ISAbal. B. Assembly of AbaR4 in pS34-2. The *oxa23* gene is magenta orange and the IR sequences of Tn6022 are shown in olive green. The blue box is ISAbal and the internal arrows indicate the transposase orientation. The primers used to span AbaR4 are shown above by the black and pink bars, along with PCR amplicon size. The lines below indicate the transposons.

The assembly also broke at the region following the *trwC* gene. Contig 50 and 16 both ended on the inverted repeats that bound Tn6022 and share a 5 bp target site duplication (5' CCATT 3') in the sequence adjacent to these repeats. This is the same target site duplication and location of AbaR4 in pA85-3. Hence, it was likely that pS27-2 also contained AbaR4 in this position. To confirm this, previously designed primers RH1391 and RH584 were paired and predicted to amplify a product 4,965 bp in size, while oxa23R and RH1392 used in the same reaction should produce an amplicon of 4,731 bp in size. In both cases, the amplicons generated

from S27 DNA appeared to be of the predicted sizes and were sequenced. As this is the second copy of AbaR4 in the genome of S27 (Chapter 4, Section 4.3.1), the internal AbaR4 PCR, RH910 and RH743 (pink in Fig A2.3B), was instead performed on a transconjugant. pS27-2 is 87,530 bp in size.

The contigs of S34 listed in Table A2.2 were tiled to the type 2 backbone and only one contig break in addition to the three repeats was present (Fig A2.4). Contig 22 and 8 both ended with the inverted repeats of Tn6022 and a 5 bp target site duplication (5' TCCAC 3') was present in the sequence adjacent to the IRs. The primers, RH2028 and RH2029, were designed in the flanking plasmid backbone and were combined with internal AbaR4 primers RH584 and oxa23R, to produce an amplicon 5,034 bp and 4,569 bp in size, respectively (Fig A2.4B). The amplicons produced by both of these reactions using S34 DNA appeared to be on the correct sizes and were sequenced. S34 also contains another copy of AbaR4 in its genome (Chapter 4, Section 4.3.1) and therefore the internal PCR (pink in Fig A2.4B) was performed on a transconjugant. pS34-2 contains AbaR4 in a new location and is 86,366 bp in size.

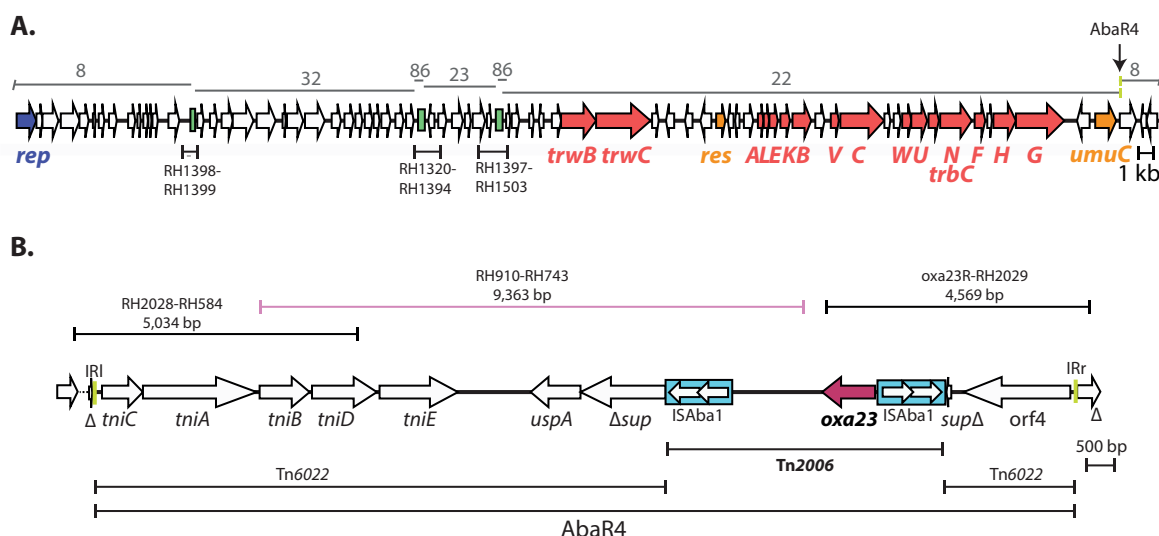


Figure A2.4 Assembly of pS34-2 contigs. A. pS34-2 contigs tiled to the type 2 backbone. Arrows indicate the extent and orientations of genes and open reading frames. The *repAci6* gene is blue, genes involved in conjugal transfer are red and the *res* and *umuC* genes are orange. Green boxes indicate the three repeat sequences. Grey lines above indicate the extent of different pS34-2 contigs. The black bars underneath show the primers used for assembly along with PCR amplicon size. Vertical arrows indicate the insertion location of AbaR4. B. Assembly of AbaR4 in pS34-2. The *oxa23* gene is magenta orange and the IR sequences of Tn6022 are shown in olive green. The blue box indicates ISAbal and the internal arrows are the transposases. The primers used to span AbaR4 are shown above by the black and pink bars, along with PCR amplicon size. The lines below indicate the transposons.

A2.2.2.2 Assembly of pA85-3 plasmids using MiSeq sequencing data

The contigs matching RepAci6 backbone in S22, S23, S25, S28, S32, S36 and S40 were retrieved, tiled on pA85-3 and were mostly completed by the same assembly PCRs as pS27-2 (Table A2.3). The RepAci6 plasmids in S22, S23 and S40 were in one contig that had IRI and IRr of Tn6022 on the edges of the contig. The sequence adjacent to these two IRs was identical to that surrounding the AbaR4 position in pA85-3. PCR screening, as indicated in Table A2.3, revealed that pS22-2, pS23-2 and pS40-2 contain AbaR4 in the same position as pA85-3. S25 and S28 had two contigs matching RepAci6 backbone (Table A2.3). One split was caused by the presence of AbaR4 in the same location as pA85-3, while the other was due to a copy of ISAbal in the same position as pS27-2. The three screening PCRs listed in Table A2.3 were performed and successfully completed the assembly of pS25-2 and pS28-2.

Table A2.3 Contig/s matching RepAci6 backbone and assembly PCRs

Isolate	Contig/s ¹	AbaR4 position (pA85-3)		ISAbal (pS27-2 location)
		RH1391-RH584 (4,965 bp)	oxa23R-RH1392 (4,852 bp)	RH2026-RH2027 (1,776 bp)
S22	1 (22)	+	+	-
S23	1 (20)	+	+	-
S25	2 (21, 31)	+	+	+
S28	2 (24, 33)	+	+	+
S32	2 (23, 44)	-	-	-
S36	2 (24, 43)	-	-	-
S40	1 (20)	+	+	-

¹The contig numbers are in the parentheses.

Contig 23 and 44 in S32 and contig 24 and 43 in S36 matched RepAci6 backbone, but they had not fragmented at a copy of AbaR4 or ISAbal. Instead, there appeared to be an error in the assembly as left of repeat 2 was sequence from F2 and to the right of this repeat was sequence from F1 (Fig A2.5). The sequence of F3 was in a separate contig in both isolates. This issue was easily resolved by sequencing the products previously amplified over repeat 1, 2 and 3 from Chapter 8; Section 8.1.

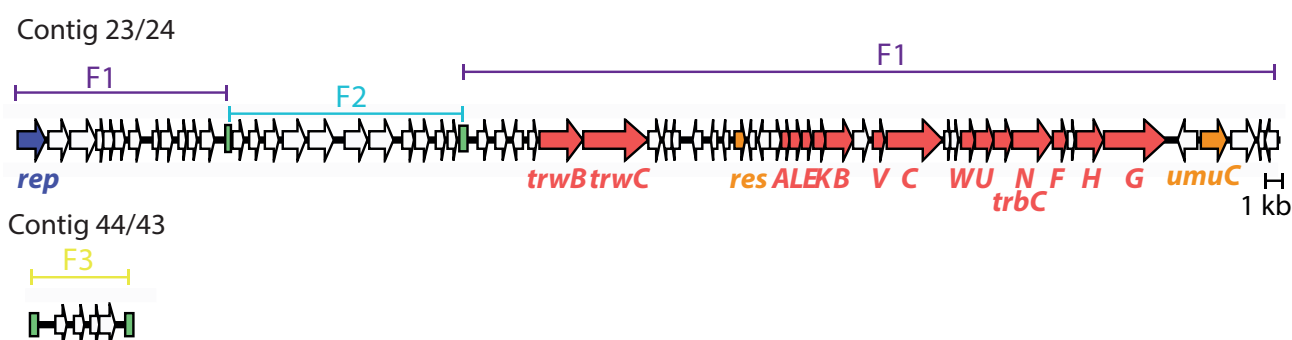


Figure A2.5 Contig 23 and 44 of pS32-2 and contig 24 and 43 of pS36-2. Arrows indicate the extent and orientation of genes and open reading frames. The *repAci6* gene is blue, genes involved in conjugal transfer are red and the *res* and *umuC* genes are orange. Green boxes indicate the three repeat sequences. The lines of three different colours above indicate fragment 1 (purple), fragment 2 (blue) and fragment 3 (yellow).

APPENDIX III

Primer tables

Table A3.1 Primer pairs used for the detection of resistance genes.

Target	Primer	Sequence (5'-3')	Amplicon Size (bp)	Reference
Sulphonamide resistance genes				
<i>sul1</i>	HS549	ACTAAGCTTGCCCCCTTCCGC	1,100	(249)
	HS550	CTAGGCATGATCTAACCCTCG		(249)
<i>sul2</i>	sul2F	GGCAGATGTGATCGACCTCG	407	(158)
	sul2R	ATGCCGGGATCAAGGACAAG		(158)
Aminoglycoside resistance genes				
<i>aacC1</i>	RH935	GCAGTCGCCCTAAAACAAAG	457	Hall lab
	RH936	CCCGTATGCCCAACTTTGTA		Hall lab
<i>aadA1/aadA2</i>	RH522	GTGGATGGCGGCCTGAAGCCA	516	Hall lab
	RH531	GGCAGCGACATCCTTCGGCGC		Hall lab
<i>aacA4</i>	RH532	GTTAGGCATCACAAAGTACAGC	518	Hall lab
	RH533	CATCTGGGGTGGTTACGGTACC		Hall lab
<i>aphA1b</i>	RH880	CAACGGGAAACGTCTTGCTC	454	(188)
	RH881	ATTCGTGATTGCGCCTGAG		(188)
<i>strA-strB</i>	strA-F	CTTGGTGATAACGGCAATTC	1,190	(76)
	strB-R	GGATCGTAGAACATATTGGC		(76)
<i>aphA6</i>	aphA6F	ATACAGAGACCACCATACAGT	234	(269)
	aphA6R	GGACAATCAATAATAGCAAT		(269)
<i>armA</i>	RH2012	TCCATTCCCTTCTCCTTTCC	508	This study
	RH2013	GGGGGTCTTACTATTCTGCCTA		This study
Tetracycline resistance gene				
<i>tetA(B)</i>	tetBF	TTGGTTAGGGGCAAGTTTTG	658	(186)
	tetBR	GTAATGGGCCAATAACACCG		(186)
β-lactamase genes				
<i>bla_{TEM}</i>	RH605	TTTCGTGTCGCCCTTATTCC	690	Hall lab
	RH606	CCGGCTCCAGATTTATCAGC		Hall lab
<i>oxa23</i>	oxa23F	GATCGGATTGGAGAACCAGA	501	(275)
	oxa23R	ATTTCTGACCGCATTTCAT		(275)
Others				
ISAbal upstream of <i>ampC</i>	RH581	TGATGATGCGTGGTGTAGGT	2,583	(95)

Tn2006	RH582	GAGCTGCCATATTGGGAAGA		(95)
	ISAbalB	CATGTAAACCAATGCTCACC	1,369	(42)
	oxa23R	ATTTCTGACCGCATTTCAT		(275)
	oxa23F	GATCGGATTGGAGAACCAGA	2,725	(275)
TnaphA6	ISAbalB	CATGTAAACCAATGCTCACC		(42)
	RH577	ACATTCACCCTGGCAAAAAG	1,427	Hall lab
	aphA6R	GGACAATCAATAATAGCAAT		(269)
	aphA6F	ATACAGAGACCACCATACAGT	1,745	(269)
	RH576	AAACAACGGATCGCTTCAAC		Hall lab

Table A3.2 Primer pairs used to detect Tn6022, AbaR4 and AbaR4Δ

PCR	Primer	Sequence (5'-3')	Amplicon size (bp)	Reference
<i>tniB-tniE</i>	RH910	GCGATAGTGAACGGATTGAGA	3,410 ¹	Hall lab
	RH587	TTGCCCATTAAGCACAAACAG		Hall lab
<i>tniB</i> -Tn2006	RH910	GCGATAGTGAACGGATTGAGA	9,363 ²	Hall lab
	RH743	GGCTTCTTGTGGATGCAACT		Hall lab
<i>sup</i> gene	RH2523	CCCACCTTtaggATCAACGCC	209	This project
	RH2509	GTGGTGTAGTCGCTTGTGTG		This project

¹Predicted amplicon size based on Tn6022. For Tn6022Δ1 it is 560 bp.

²Predicted amplicon size based on AbaR4. For AbaR4Δ1 it is 6,513 bp.

Table A3.3 Primer pairs used for mapping AbGRI1-1

PCR	Primer	Sequence (5'-3')	Amplicon size (bp)	Reference
<i>comM-sup</i> ¹	RH791	TGCTGCAATGAGCTGAAAGT	6,288	Hall lab
	RH771	TGTAAAATCTGGTGGTCGTAC		Hall lab
<i>tniB-tetA(B)</i> ¹	RH910	GCGATAGTGAACGGATTGAGA	7,120	Hall lab
	tetBF	TTGGTTAGGGGCAAGTTTTG		(186)
<i>tetA(B)-strB</i>	tetBR	GTAATGGGCCAATAACACCG	4,616	(186)
	strBout	AGAGGAGCAACGCGATCTAGC		(201)
<i>strA-comM</i>	strAout	AACCAGTCAGAATGCGATTCA	2,942	(201)
	RH1366	CAGATTTGCCCAAAGATGGT		Hall lab

¹Predicted size based on Tn6022Δ1.

Table A3.4 Primer pairs used for mapping AbGRI1-3

PCR	Primer	Sequence (5'-3')	Amplicon size (bp)	Reference
<i>comM-sup</i> ¹	RH791	TGCTGCAATGAGCTGAAAGT	6,288	Hall lab
	RH771	TGTAAAATCTGGTGGTCGTAC		Hall lab
<i>sup-orf2</i>	RH772	GCAGCCATAGGAATGACTTTTA	4,768	Hall lab
	RH596	GGGCGGCAAATATAAAGGTT		Hall lab
<i>orf5-sul2</i>	RH597	TTTGAAGAAATTGAGCATGAGG	5,586	Hall lab
	RH955	GAAACTGTCCGAGGTTATGTTGAC		Hall lab
<i>sul2-uspA</i> ²	RH603	ACTTCATCCGCACACACGAG	8,423	Hall lab
	RH934	TCGGAAATCCGAATAGTGAAA		Hall lab
<i>tniD-tetA(B)</i>	RH583	TCCTGTCTCTCGTGTAGCAAT	8,074	Hall lab
	tetBF	TTGGTTAGGGGCAAGTTTTG		(186)
<i>tetA(B)-strB</i>	tetBR	GTAATGGGCCAATAACACCG	4,616	(186)
	strBout	AGAGGAGCAACGCGATCTAGC		(201)
<i>strA-comM</i>	strAout	AACCAGTCAGAATGCGATTCA	2,942	(201)
	RH1366	CAGATTTGCCCAAAGATGGT		Hall lab

¹Predicted size based on Tn6022Δ1.

²Predicted size based on Tn6022.

Table A3.5 Primer pairs used for mapping AbGRI1-4

PCR	Primer	Sequence (5'-3')	Amplicon size (bp)	Reference
<i>comM-sup</i> ¹	RH791	TGCTGCAATGAGCTGAAAGT	6,288	Hall lab
	RH771	TGTAAAATCTGGTGGTCGTAC		Hall lab
<i>sup-orf2</i>	RH772	GCAGCCATAGGAATGACTTTTA	4,768	Hall lab
	RH596	GGGCGGCAAATATAAAGGTT		Hall lab
<i>orf5-sul2</i>	RH597	TTTGAAGAAATTGAGCATGAGG	5,586	Hall lab
	RH955	GAAACTGTCCGAGGTTATGTTGAC		Hall lab
<i>sul2-uspA</i> ²	RH603	ACTTCATCCGCACACACGAG	8,423	Hall lab
	RH934	TCGGAAATCCGAATAGTGAAA		Hall lab
<i>oxa23-tetA(B)</i>	oxa23R	ATTTCTGACCGCATTTCCAT	5,205	(275)
	tetBF	TTGGTTAGGGGCAAGTTTTG		(186)
<i>tetA(B)-strB</i>	tetBR	GTAATGGGCCAATAACACCG	4,616	(186)
	strBout	AGAGGAGCAACGCGATCTAGC		(201)
<i>strA-comM</i>	strAout	AACCAGTCAGAATGCGATTCA	2,942	(201)
	RH1366	CAGATTTGCCCAAAGATGGT		Hall lab

¹Predicted size based on Tn6022Δ1.²Predicted size based on Tn6022.**Table A3.6 Primer pairs used for mapping AbGRI1-5**

PCR	Primer	Sequence (5'-3')	Amplicon size (bp)	Reference
<i>comM-sul2</i> ¹	RH791	TGCTGCAATGAGCTGAAAGT	5,099 (<i>tniA</i> recombination)	Hall lab
	sul2R	ATGCCGGGATCAAGGACAAG		(158)
<i>sul2-uspA</i> ²	RH603	ACTTCATCCGCACACACGAG	8,423	Hall lab
	RH934	TCGGAAATCCGAATAGTGAAA		Hall lab
<i>tniD-tetA(B)</i>	RH583	TCCTGTCTCTCGTGTAGCAAT	8,074	Hall lab
	tetBF	TTGGTTAGGGGCAAGTTTTG		(186)
<i>tetA(B)-strB</i>	tetBR	GTAATGGGCCAATAACACCG	4,616	(186)
	strBout	AGAGGAGCAACGCGATCTAGC		(201)
<i>strA-comM</i>	strAout	AACCAGTCAGAATGCGATTCA	2,942 ³	(201)
	RH1366	CAGATTTGCCCAAAGATGGT		Hall lab

¹*tniA-tniAb* recombination.

²Predicted size based on Tn6022.

³4,130 bp in AbGRI1-5b.

Table 3.7 Primer pairs used for mapping AbGRI1-6

PCR	Primer	Sequence (5'-3')	Amplicon size (bp)	Reference
<i>comM-sup</i> ¹	RH791	TGCTGCAATGAGCTGAAAGT	6,288 (Δ)	Hall lab
	RH771	TGTAAAATCTGGTGGTCGTAC		Hall lab
Tn2006- <i>tetA</i> (B)	oxa23R	ATTTCTGACCGCATTTCCAT	5,016	Hall lab
	tetBF	TTGGTTAGGGGCAAGTTTTG		(186)
<i>tetA</i> (B)- <i>strB</i>	tetBR	GTAATGGGCCAATAACACCG	4,616	(186)
	strBout	AGAGGAGCAACGCGATCTAGC		(201)
<i>strA-comM</i>	strAout	AACCAGTCAGAATGCGATTCA	2,942	(201)
	RH1366	CAGATTTGCCCAAAGATGGT		Hall lab

¹Predicted size based on Tn6022Δ1.

Table A3.8 Primer pairs used for mapping AbGRI1-7

PCR	Primer	Sequence (5'-3')	Amplicon size (bp)	Reference
<i>comM-sup</i> ¹	RH791	TGCTGCAATGAGCTGAAAGT	6,288	Hall lab
	RH771	TGTAAAATCTGGTGGTCGTAC		Hall lab
Tn2006- <i>orf2</i>	oxa23R	ATTTCTGACCGCATTTCCAT	7,720	(275)
	RH596	GGGCGGCAAATATAAAGGTT		Hall lab
<i>orf5-sul2</i>	RH597	TTTGAAGAAATTGAGCATGAGG	5,586	Hall lab
	RH955	GAAACTGTCCGAGGTTATGTTGAC		Hall lab
<i>sul2-uspA</i> ²	RH603	ACTTCATCCGCACACACGAG	8,423	Hall lab
	RH934	TCGGAAATCCGAATAGTGAAA		Hall lab
<i>oxa23-tetA</i> (B)	oxa23R	ATTTCTGACCGCATTTCCAT	5,205	(275)
	tetBF	TTGGTTAGGGGCAAGTTTTG		(186)
<i>tetA</i> (B)- <i>strB</i>	tetBR	GTAATGGGCCAATAACACCG	4,616	(186)
	strBout	AGAGGAGCAACGCGATCTAGC		(201)
<i>strA-comM</i>	strAout	AACCAGTCAGAATGCGATTCA	2,942	(201)
	RH1366	CAGATTTGCCCAAAGATGGT		Hall lab

¹Predicted size based on Tn6022Δ1.

²Predicted size based on Tn6022.

Table A3.9 Primer pairs used for mapping AbGRI1-8

PCR	Primer	Sequence (5'-3')	Amplicon size (bp)	Reference
<i>comM-sup</i> ¹	RH791	TGCTGCAATGAGCTGAAAGT	6,288	Hall lab
	RH771	TGTAAAATCTGGTGGTCGTAC		Hall lab
<i>sup-orf2</i>	RH772	GCAGCCATAGGAATGACTTTTA	4,768	Hall lab
	RH596	GGGCGGCAAATATAAAGGTT		Hall lab
<i>orf5-sul2</i>	RH597	TTTGAAGAAATTGAGCATGAGG	5,586	Hall lab
	RH955	GAAACTGTCCGAGGTTATGTTGAC		Hall lab
<i>sul2-uspA</i> ²	RH603	ACTTCATCCGCACACACGAG	8,423	Hall lab
	RH934	TCGGAAATCCGAATAGTGAAA		Hall lab
<i>tniD-tetA(B)</i>	RH1310	TTGCTCAAGTGTACCCCTTTG	5,823	Hall lab
	tetBF	TTGGTTAGGGGCAAGTTTTG		(186)
<i>tetA(B)-strB</i>	tetBR	GTAATGGGCCAATAACACCG	4,616	(186)
	strBout	AGAGGAGCAACGCGATCTAGC		(201)
<i>strA-Tn2006</i>	strAout	AACCAGTCAGAATGCGATTCA	2,517	(201)
	oxa23R	ATTTCTGACCGCATTTCCAT		(275)
Tn2006- <i>comM</i>	RH743	GGCTTCTTGTGGATGCAACT	5,023	Hall lab
	RH1366	CAGATTTGCCCAAAGATGGT		Hall lab

¹Predicted size based on Tn6022Δ1.

²Predicted size based on Tn6022.

Table 3.10 Primer pairs used for mapping AbGRI2-1

PCR	Primer	Sequence (5'-3')	Amplicon size (bp)	Reference
ABA1_01191 ¹ - <i>tnpR</i> ₁	RH1315	AGGAGATCTTCTTGGCAGTCA	1,051	(188)
	RH539	CCAGCCCTTCCCGATCTGTTG		(188)
<i>bla</i> _{TEM} - <i>tnpA</i> ₁₀₀₀	RH605	TTTCGTGTCGCCCTTATTCC	2,650	(188)
	RH759	GCCAGCTCATTTACCTTGCCGA		(188)
<i>tnpR</i> _{5393c} - <i>aphA1b</i>	RH520	CATGGCCCAGCGCGATACTTCAG	2,297	(188)
	RH880	CAACGGGAAACGTCTTGCTC		(188)
<i>aphA1b</i> -sul1	RH881	ATTCGTGATTGCGCCTGAG	2,712	(188)
	RH751	GCGGAACTTCACGCGATC		(188)
<i>tnpA</i> -ABA1_01225	RH668	CACCAGAACCGCCTGCTCAA	1,219	(188)
	RH1316	CATCTGCCATCCAGTTTGTG		(188)
5'-CS-3'CS	RH754	CGGCATCGAGCGGGATTTCCTT	4,531 ²	Hall lab
	R1	AAGCAGACTTGACCTGAT		(159)

¹GenBank accession number CP010781.

²Predicted size based on the cassette array *aadA1*-orfQ-orfP-*aacC1*. For the *aadA1*-orfQ-orfP-*aacC1* cassette array it is 3,993.

Table 3.11 Primer pairs used for mapping AbGRI2-12

PCR	Primer	Sequence (5'-3')	Amplicon size (bp)	Reference
TE32_13140 ¹ - <i>tnpR</i> ₁	RH1563	ATAGATCGGCTTCGGACTCA	1,046	Hall lab
	RH539	CCAGCCCTTCCCGATCTGTTG		(188)
<i>bla</i> _{TEM} - <i>tnpA</i> ₁₀₀₀	RH605	TTTCGTGTCGCCCTTATTCC	2,650	(188)
	RH759	GCCAGCTCATTTACCTTGCCGA		(188)
<i>tnpR</i> _{5393c} - <i>aphA1b</i>	RH520	CATGGCCCAGCGCGATACTTCAG	2,297 ²	(188)
	RH880	CAACGGGAAACGTCTTGCTC		(188)
<i>aphA1b</i> -ABA1_01228 ³	RH881	ATTCGTGATTGCGCCTGAG	1,581 ⁴	(188)
	RH2008	TGATGACTTCCATTAAAGCCTGT		This project

¹GenBank accession number CP010779.

²Predicted size without IS26Δ.

³CP010781

⁴Predicted size based on AbGRI2-12a. For AbGRI2-12b it is 1,772.

Table 3.12 Primer pairs used for mapping AbGRI2-13

PCR	Primer	Sequence (5'-3')	Amplicon size (bp)	Reference
ABA1_01189 ¹ - <i>tnpA</i> ₁₀₀₀	RH2009	GATAAGTTCGCACCGCATTT	1,223	This project
	RH781	TCACCAGACCCAGAGCACCGA		Hall lab
<i>tnpR</i> _{5393c} - <i>intI1</i>	RH520	CATGGCCCAGCGCGATACTTCAG	1,881	(188)
	RH935	GCAGTCGCCCTAAAACAAAG		Hall lab
<i>tnpA</i> -ABA1_01225	RH668	CACCAGAACCGCCTGCTCAA	1,219	(188)
	RH1316	CATCTGCCATCCAGTTTGTG		(188)

Table 3.13 Primer pairs used for mapping AbGRI2-14

PCR	Primer	Sequence (5'-3')	Amplicon size (bp)	Reference
ABA1_01191 ¹ - <i>tnpR</i> ₁	RH1315	AGGAGATCTTCTTGGCAGTCA	1,051	(188)
	RH539	CCAGCCCTTCCCGATCTGTTG		(188)
<i>bla</i> _{TEM} - <i>tnpA</i> ₁₀₀₀	RH605	TTTCGTGTCGCCCTTATTCC	2,650	(188)
	RH759	GCCAGCTCATTTACCTTGCCGA		(188)
<i>tnpR</i> _{5393c} - <i>intI1</i>	RH520	CATGGCCCAGCGCGATACTTCAG	1,881	(188)
	RH935	GCAGTCGCCCTAAAACAAAG		Hall lab
<i>tnpA</i> -ABA1_01225 ¹	RH668	CACCAGAACCGCCTGCTCAA	1,219	(188)
	RH1316	CATCTGCCATCCAGTTTGTG		(188)

¹GenBank accession number CP010781.

Table 3.14 Primer pairs used for mapping those isolates with only IS26 in the AbGRI2 region

PCR	Primer	Sequence (5`-3`)	Amplicon size (bp)	Reference
IS26 (1)				
ABA1_01191 ¹ -	RH1315	AGGAGATCTTCTTGGCAGTCA	1,048	(188)
ABA1_01230	RH2007	ACTTTTGGCATCGTCGGTAT		This project
IS26 (2)				
TE32_13140 ² -	RH1563	ATAGATCGGCTTCGGACTCA	1,850	Hall lab
TE32_12895	RH2548	GGATTACGTTCGATGGTCTGC		This project

¹GenBank accession number CP010781.

²GenBank accession number CP010779.

Table 3.15 Primer pairs used for AbGRI3 assembly

PCR	Primer	Sequence (5'-3')	Amplicon size (bp)	Reference
AbGRI3-1i				
<i>Δatr</i>	RH2001	GGAGTTGGTTTTGGTACAGCA	1,323	This project
<i>repAciN</i>	RH2002	TATAAGCCACCTCGCTCACC		This project
<i>intI1</i>	RH2003	GCCTTGATGTTACCCGAGAG	1,193	This project
<i>Δasr</i>	RH2005	CACTGATCTGCTGGCTTTCA		This project
<i>atrΔ</i>	RH2015	CCCAGCAATCCATTCGTAGT	1,524	This project
<i>asrΔ</i>	RH2006	TGACGAGCTTTGTTTAGGTGTG		This project
AbGRI3-1				
<i>intI1</i>	RH2003	GCCTTGATGTTACCCGAGAG	1,566	This project
<i>atrΔ</i>	RH2015	CCCAGCAATCCATTCGTAGT		This project
<i>Δasr</i>	RH2005	CACTGATCTGCTGGCTTTCA	1,151	This project
<i>asrΔ</i>	RH2006	TGACGAGCTTTGTTTAGGTGTG		This project
AbGRI3-2i				
<i>Δatr</i>	RH2001	GGAGTTGGTTTTGGTACAGCA	1,323	This project
<i>repAciN</i>	RH2002	TATAAGCCACCTCGCTCACC		This project
<i>intI1</i>	RH2003	GCCTTGATGTTACCCGAGAG	1,943	This project
<i>aphA1b</i>	RH880	CAACGGGAAACGTCTTGCTC		(188)
<i>aphA1b</i>	RH831	TATACCCATATAAAATCAGCATCC	1,203	Hall lab
<i>Δasr</i>	RH2005	CACTGATCTGCTGGCTTTCA		This project
<i>atrΔ</i>	RH2015	CCCAGCAATCCATTCGTAGT	1,524	This project
<i>asrΔ</i>	RH2006	TGACGAGCTTTGTTTAGGTGTG		This project
AbGRI3-2				
<i>aphA1b</i>	RH831	TATACCCATATAAAATCAGCATCC	1,576	Hall lab
<i>atrΔ</i>	RH2015	CCCAGCAATCCATTCGTAGT		This project
<i>Δasr</i>	RH2005	CACTGATCTGCTGGCTTTCA	1,151	This project
<i>asrΔ</i>	RH2006	TGACGAGCTTTGTTTAGGTGTG		This project
AbGRI3-3i				
<i>Δatr</i>	RH2001	GGAGTTGGTTTTGGTACAGCA	1,605	This project

ISAb24	RH2010	TTTCGTGACACTCTCGCTTG		This project
<i>aacA4</i>	RH2011	CATAGAGCATCGCAAGGTCA	1,007	This project
Δ <i>asr</i>	RH2005	CACTGATCTGCTGGCTTTCA		This project
<i>atr</i> Δ	RH2015	CCCAGCAATCCATTCGTAGT	1,398	This project
<i>asr</i> Δ	RH2006	TGACGAGCTTTGTTTAGGTGTG		This project
AbGRI3-3				
<i>aacA4</i>	RH2011	CATAGAGCATCGCAAGGTCA	1,380	This project
<i>atr</i> Δ	RH2015	CCCAGCAATCCATTCGTAGT		This project
Δ <i>asr</i>	RH2005	CACTGATCTGCTGGCTTTCA	1,025	This project
<i>asr</i> Δ	RH2006	TGACGAGCTTTGTTTAGGTGTG		This project
AbGRI3-4				
Δ <i>atr</i>	RH2001	GGAGTTGGTTTTGGTACAGCA	1,323	This project
<i>repAciN</i>	RH2002	TATAAGCCACCTCGCTCACC		This project
<i>armA</i>	RH2012	TCCATTCCCTTCTCCTTTCC	1,934	This project
<i>asr</i> Δ	RH2014	CCAAATACCGCCCACTCAAC		This project
AbGRI3-5i				
Δ <i>atr</i>	RH2001	GGAGTTGGTTTTGGTACAGCA	2,146	This project
ISAb24	RH2010	TTTCGTGACACTCTCGCTTG		This project
<i>intI1</i>	RH2003	GCCTTGATGTTACCCGAGAG	1,943	This project
<i>aphA1b</i>	RH880	CAACGGGAAACGTCTTGCTC		(188)
<i>aphA1b</i>	RH831	TATACCCATATAAATCAGCATCC	1,203	Hall lab
Δ <i>asr</i>	RH2005	CACTGATCTGCTGGCTTTCA		This project
<i>orf</i> Δ	RH2016	GCAGCCTCAAAGTGGAAC	1,173	This project
<i>asr</i> Δ	RH2006	TGACGAGCTTTGTTTAGGTGTG		This project
AbGRI3-5				
<i>aphA1b</i>	RH831	TATACCCATATAAATCAGCATCC	1,351	Hall lab
<i>orf</i> Δ	RH2016	GCAGCCTCAAAGTGGAAC		This project
Δ <i>asr</i>	RH2005	CACTGATCTGCTGGCTTTCA	1,025	This project
<i>asr</i> Δ	RH2006	TGACGAGCTTTGTTTAGGTGTG		This project
AbGRI3-6				
Δ <i>atr</i>	RH2001	GGAGTTGGTTTTGGTACAGCA	1,232	This project

<i>repAciN</i>	RH2547	AATCAGCCACGCATTCAGTC		This project
<i>repAciN</i>	RH2002	TATAAGCCACCTCGCTCACC	2,638	This project
ISAb24	RH2048	TCTGTCGCTACTGACTCGTC		This project
<i>aacA4</i>	RH2011	CATAGAGCATCGCAAGGTCA	1,147	This project
<i>asrΔ</i>	RH2006	TGACGAGCTTTGTTTAGGTGTG		This project
AbGRI3-7 (D86)				
ACICU_01435 ¹	RH2049	CTTGCTGATATGGCGGGTTG	1,333	This project
<i>Δatr</i>	RH2001	GGAGTTGGTTTTGGTACAGCA		This project
ISAb24	RH2047	CTTGATGTGCCCCAGGTCTA	1,568	This project
ΔACICU_01435 ¹	RH2050	CCGTTCTGCTGCTATTCTGG		This project
<i>repAciN</i>	RH2002	TATAAGCCACCTCGCTCACC	1,293	This project
ISAb24	RH2048	TCTGTCGCTACTGACTCGTC		This project

¹GenBank accession number CP000863.

Table 3.16 Primer pairs used assembly of RepAci6 plasmids

PCR	Primer	Sequence (5'-3')	Amplicon size (bp)	Reference
RepAci6				
<i>repAci6</i>	RH2045	AAGGTGACAGCAAGTACGTG	482	This project
<i>repAci6</i>	RH2046	CGTGCGCCTCATTCAACAT		This project
Repeats				
CBI29_04421 ¹	RH1398	TTTGACGTTGCTCTTGTTGC	987 ² /991 ³ /3,378 ⁴	(97)
CBI29_04422 ¹	RH1399	TTCTCCCAAGTGGTCAGGTC		(97)
CBI29_04441 ¹	RH1320	TTTACGATAATTGCATGGATGT	1,525 ² /1,524 ³	(192)
CBI29_04443 ¹	RH1394	TGGTTGGCAGAACAAGATGA		(97)
CBI29_04450 ¹	RH1397	CCATGTTCTTTTCCACATGC	1,576 ² /1,774 ³	(97)
CBI29_04448 ¹	RH1503	GAAGATCCAGAAGCGGGATA		(97)
TnaphA6⁵ assembly				
CBI29_04485 ¹	RH1502	TTGCTTTAATCGGTGGTTCC	2,540	(97)
<i>aphA6</i>	aphA6F	ATACAGAGACCACCATACAGT		(269)
<i>aphA6</i>	aphA6R	GGACAATCAATAATAGCAAT	1,930	(269)

CBI29_04483 ¹ pS24-1	RH1501	CTTGAGGAAGGGATGGTTGA		(97)
CBI29_04434 ¹ <i>tniD</i>	RH2024	TGTTGTTAGCTTGGATCGTGTC	4,831	This project
<i>oxa23</i>	RH584	TCAATATGCCTCGCTCCACT		Hall lab
CBI29_04440 ¹ pS26-1/pS29-1	oxa23R	ATTTCTGACCGCATTTCCAT	4,683	(275)
	RH2025	TGCAAAAGCAAAATCTTCGTCA		This project
CBI29_04487 ¹ <i>tniD</i>	RH2019	CGTGTAGCAACCATCCAAAA	4,938	This project
<i>oxa23</i>	RH584	TCAATATGCCTCGCTCCACT		Hall lab
CBI29_04487 ¹	oxa23R	ATTTCTGACCGCATTTCCAT	4,307	(275)
orf	RH2020	CATGAATGCCGTAGAGATCG		This project
CBI29_04422 ¹ pS27-2, pS25-2, pS28-2	RH2030	TCATCATAGCACATCGTCCAG	1,069	This project
	RH1399	TTCTCCCAAGTGGTCAGGTC		(97)
CBI29_04426 ¹	RH2026	GAAACAAAACCTTCACTGCTATGC	1,766	This project
CBI29_04428 ¹	RH2027	AGACATACCGAAAAGTCCGAA		This project
CBI29_04460 <i>tniD</i>	RH1391	ATGCGAGGCTTCTTGGTCTA	4,965	(99)
<i>oxa23</i>	RH584	TCAATATGCCTCGCTCCACT		Hall lab
CBI29_04460	oxa23R	ATTTCTGACCGCATTTCCAT	4,731	(275)
pS22-2, pS23-2, pS40-2	RH1392	CCTTAACGCTCAGTCTTTTGG		(99)
CBI29_04460 ¹ <i>tniD</i>	RH1391	ATGCGAGGCTTCTTGGTCTA	4,965	(99)
<i>oxa23</i>	RH584	TCAATATGCCTCGCTCCACT		Hall lab
CBI29_04460 ¹	oxa23R	ATTTCTGACCGCATTTCCAT	4,731	(275)
pS34-2	RH1392	CCTTAACGCTCAGTCTTTTGG		(99)
CBI29_04511 ¹ <i>tniD</i>	RH2028	ACCATGTGCCGTAATAGCCT	5,034	This project
<i>oxa23</i>	RH584	TCAATATGCCTCGCTCCACT		Hall lab
CBI29_04512 ¹ pS21-2	oxa23R	ATTTCTGACCGCATTTCCAT	4,569	(275)
	RH2029	ACAAATGAAGATACGTGCCTGA		This project

CBI29_04460 ¹	RH1391	ATGCGAGGCTTCTTGGTCTA	2,367	(99)
CBI29_04481 ¹	RH2021	TGTGTGCAGTTGGAATTAAAGC		This project
CBI29_04512 ¹	RH2022	CCGACTCTTCAGGTGCTTCA	5,171	This project
<i>tniD</i>	RH584	TCAATATGCCTCGCTCCACT		Hall lab
<i>oxa23</i>	oxa23R	ATTTCTGACCGCATTTCCAT	4,757	(275)
CBI29_04513 ¹	RH2023			This project
pS30-2				
<i>oxa23</i>	oxa23R	ATTTCTGACCGCATTTCCAT	2,101	(275)
CBI29_04441	RH1320	TTTACGATAATTGCATGGATGT		(192)

¹GenBank accession number CP021787.

²Size for the Type 1 backbone.

³Size for the Type 2 backbone.

⁴Product size for Type 3 and 4 backbones.

⁵G7 location.

Table A3.17 Small plasmid assembly

PCR	Primer	Sequence (5'-3')	Amplicon size (bp)	Reference
pAb-G7-1 backbone	RH2035	CCTGACAACCTTCCACACAGC	328	This project
pAb-G7-1 backbone	RH2036	ACAGCAAGCCGATCGTATTG		This project
pD72-1 backbone	RH2037	CTTTGGTACGCTCTCTTGCC	263	This project
pD72-1 backbone	RH2038	GTTGTCGGGGTATTGGGGA		This project
pS32-1				
<i>tonB</i>	RH2505	TTACCTCCTGCTCGTTTGGG	2,728	This project
Tn2006 internal	RH743	GGCTTCTTGTGGATGCAACT		Hall lab
<i>oxa23</i>	Oxa23R	ATTTCTGACCGCATTTCCAT	2,317	(275)
pAb-G7-1 backbone	RH2506	AGAGTTACTGCCACACTGCT		This project
pS21-1				
pD72-1 backbone	RH1348	CACATTGTGGACAGGTTGGA	1,847	Hall lab
<i>aphA6</i>	aphA6F	ATACAGAGACCACCATAACAGT		(269)
<i>aphA6</i>	aphA6R	GGACAATCAATAATAGCAAT	1,624	(269)
pD72-1 backbone	RH1322	ATTGCCGATGGAATGAAGAT		Hall lab
pS30-1				

<i>tet39</i>	RH2502	AACAGGGATGTTTCGGGCTAT	1,918	This project
	Tet39R	GCCTTTTGCGTTGTTACCAT		(98)
	RH2503	TGTCGCTTTGAGAGTTAGGC	2,508	This project
	RH2504	TCGTATTGGTTCGCTCGGTA		This project
