The role of the accessory genome in nosocomial and extranosocomial isolates of *Acinetobacter baumannii*

Daniel J. Farrugia

This thesis is presented as a partial fulfilment to the requirements for a Doctorate of Philosophy

Department of Chemistry and Biomolecular Sciences
Macquarie University (2014)

Statement of Candidature

I certify that the work presented in this thesis only consists of research generated through my own efforts, and has not been previously submitted to fulfil requirements for any other degree nor at any other institution. Any assistance obtained in the preparation of this thesis and the reported research has been acknowledged accordingly. All literature and other sources of information have been cited appropriately. Biosafety approval for this work was obtained through the Macquarie University Ethics Review Committee reference numbers 08/03/EX (19 June 2008) and 5201100898 (2011-present).

Daniel J. Farrugia
July 29 2014

Table of Contents

Acknowledgments

Firstly, I would like to thank my supervisor, Professor Ian Paulsen, for the providing me the opportunity to do my PhD, and for providing both freedom and counsel in the research I had conducted, and my associate supervisor, Associate Professor Bridget Mabbutt, for helping me to better my writing and for her supportive guidance. I would also like to thank the postdoctoral staff for their assistance throughout, particularly Karl Hassan, Sasha Tetu, Liam Elbourne and Amanda Mackie. Without their input, completing this PhD would not have been possible. I would also like to acknowledge past and present fellow postgraduates of the Paulsen Lab: Kent Chee Lim, Prasanth Subramani, Bhumika Shah, Katy Breakwell, Liping Li, Deepa Varkey, Sheemal Kumar, Ali Kameneh, Qi ‘Snow’ Liu, Hua Yu Qin, Taotao Huang, Hasinika Ariyaratne, Brodie Sutcliffe and Silas Vick, for their support and comradery. Finally, I would like to thank my family and friends for their continued support and encouragement.

Co-contributions of manuscripts/chapters

This thesis is presented in a thesis-by-publication format, with a blend of published and unpublished work. Each chapter constitutes a single body of work that consists of research generated by several co-authors. The contribution of each co-author is outlined below.

Chapter 2: The complete genome and phenome of a community-acquired *Acinetobacter baumannii*

Daniel N. Farrugia, Liam D. H. Elbourne, Karl A. Hassan, Bart A. Eijkelkamp, Sasha G. Tetu, Melissa H. Brown, Bhumika S. Shah, Anton Y. Peleg, Bridget C. Mabbutt and Ian T. Paulsen

The experiments outlined in this manuscript were conceived and designed by DNF, KAH, LDHE, MHB, and ITP. The genome data was assembled by LDHE, while gap closure and annotation was conducted by DNF. Analysis of genome data was conducted mainly by DNF, with contributions from KAH, LDHE, BAE, SGT, BSS and ITP, in regards to accessory genome, genome assembly, and virulence analyses. Phenotype microarrays, phenotype confirmation and analysis of phenome data was conducted by DNF. This manuscript was written mainly by DNF with essential input from LDHE, KAH, MHB, AYP, BCM and ITP.

Chapter 3: A novel and mosaic family of mobile genomic islands inserted in the tRNA-dihydrouridine synthase A (*dusA*) gene

Daniel N. Farrugia, Liam D. H. Elbourne, Bridget C. Mabbutt and Ian T. Paulsen

The experiments outlined in this manuscript were conceived and designed by DNF, LDHE and ITP. PCRs, genomic deletions and phenotype microarrays were conducted by DNF. Data was analysed primarily by DNF, with contributions from LDHE, BCM and ITP, regarding phylogenetics and protein function. This manuscript was written mainly by DNF, with input from LDHE, BCM and ITP.

Chapter 4: A novel family of genomic resistance islands, AbGRI2, contributing to aminoglycoside resistance in *Acinetobacter baumannii* isolates belonging to global clone

Steven J. Nigro, Daniel N. Farrugia, Ian T. Paulsen and Ruth M. Hall

SJN and RMH conceived and designed the experiments outlined in this manuscript. Experimental work was conducted by SJN. DNF, LDHE and ITP had annotated and assembled the genome of *A. baumannii* WM99c. Analysis of resistance islands was conducted by SJN and RMH, with contributions of genomic data and bioinformatic analyses by DNF, LDHE and ITP. This manuscript was written mainly by SJN, with added input from DNF, ITP and RMH.

Chapter 5

Chapter 5.1: The accessory genome and phenome of *Acinetobacter baumannii* WM99c

Daniel N. Farrugia, Liam D. H. Elbourne, Karl A. Hassan and Ian T. Paulsen

Genome assembly was conducted by LDHE and genome analysis was performed by DNF. Phenotype microarrays and analysis were performed by DNF. This chapter was written by DNF, with contributions from KAH and ITP.

Chapter 5.2: Virulome comparisons of *Acinetobacter baumannii* originating from nosocomial and community-acquired infections in the Asia-Pacific

Daniel N. Farrugia, Mark D. Adams, and Ian T. Paulsen, Anton Y. Peleg

AYP conceived the analysis outlined in this chapter. Genome assembly and annotation were conducted by MDA. Genome analysis for the presence and absence of virulence traits was primarily conducted by DNF, with contributions from ITP and AYP. This chapter was written by DNF, with contributions from ITP and AYP.

Chapter 5.3: Comparative genomics of *Acinetobacter baumannii* originating from varying extranosocomial environments

Daniel N. Farrugia and Ian T. Paulsen

DNF and ITP conceived the work outlined in this chapter. Genome assembly and all analyses were performed primarily by DNF, with input from ITP. This chapter was written by DNF, with added contributions by ITP.

Abbreviations

AFLP - amplified fragment length polymorphism

AHL - N-acyl homoserine lactone

AME - aminoglycoside-modifying enzyme

CA-AB - community acquired *Acinetobacter baumannii*

cccDNA - covalently closed circular DNA

CDS - coding sequence

DAI - *dusA*-associated integrase

DAMGE - *dusA*-specific mobile genetic element

DusA - dihydrouridine synthase A

ESKAPE - *Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, and *Enterobacter* sp.

FRT - FLP recognition target

GEI - genomic island

IC - international clone/clonal

IS - insertion sequence

KEGG - Kyoto Encyclopedia of Genes and Genomes

LB/LBA - lysogeny broth/ lysogeny broth agar

MFS - major facilitator superfamily

MGE - mobile genetic element

MIC - minimum inhibitory concentration

MLST - multilocus sequence typing

MMC - mitomycin C

ORF - open reading frame

OMP - outer membrane protein

PFGE - pulsed-field gel electrophoresis

PM - phenotype microarray

PNAG - poly-β-1-6-N-acetylglucosamine

PQQ - pyrroloquinoline quinone

QS - quorum sensing

R2A - Reasoner's 2A medium

RAST - Rapid Annotations using Subsystems Technology

Rci - shufflon-specific DNA recombinase

RGP - region of genomic plasticity

RND - resistance nodulation division

Tn*AbaR* - *Acinetobacter baumannii* antibiotic resistance island

WGS - whole genome sequencing

XN-AB - extranosocomial *Acinetobacter baumannii*

Abstract

Acinetobacter baumannii is an opportunistic nosocomial pathogen that is frequently able to resist antimicrobial treatment and persist in the clinical milieu. This organism is also able to cause infection outside the hospital environment as the clinically and epidemiologically distinct community-acquired *A. baumannii* (CA-AB). CA-AB forms part of a larger collective of extranosocomial *A. baumannii* (XN-AB) originating outside the clinical milieu, including *A. baumannii* isolated from industrial and environmental sites. Whole genome sequencing has provided useful insight into the clonal population structure, the influence of lateral gene transfer, and the virulence and drug resistance capabilities of nosocomial *A. baumannii*. Prior to the commencement of this project, complete genomes of several nosocomial isolates were available, whilst genomic data regarding XN-AB isolates was limited to only that of a louse isolate, *A. baumannii* SDF.

In this work, the first complete genome sequence of a CA-AB isolate, *A. baumannii* D1279779, was determined. In addition to the genomic sequence, the phenotypic profile of this organism, and several nosocomial isolates, were assayed using phenotype microarrays. Comparative genomics and phenomics confirmed the distinctive nature of CA-AB, revealing D1279779 to be both metabolically diverse and antimicrobial susceptible compared to nosocomial *A. baumannii*. Additionally, the genome of *A. baumannii* D1279779 encoded several novel gene clusters, presumably acquired through lateral gene transfer.

One putative laterally acquired gene cluster, identified during the D1279779 study above, became the focus for a second project. This genomic region encoded a putative integrase gene that was adjacent to the 5' end of the tRNA-dihydrouridine synthase (*dusA*) gene in several *A. baumannii* and *Pseudomonas* spp. genomes sequenced by our research group. Homologues of the gene encoding this integrase were discovered in the genomes of over two hundred Proteobacterial organisms, and was established to be a component of a novel family of genomic islands in these organisms. These genomic islands were found to be highly variable in both size and gene composition, and capable of chromosomal excision as a circularised intermediate. On that basis, this family of genomic islands was dubbed the ***d****us****A***-specific **M**obile **G**enetic **E**lements (DAMGEs).

The genomes of various other *A. baumannii* isolates were sequenced, amongst these an Australian multidrug resistant outbreak isolate, *A. baumannii* WM99c. The genome of this isolate was found to encode several putatively laterally transferred regions, including the AbaR and AbGRI2 genomic islands, which encoded resistance to various antimicrobials such as tetracyclines and aminoglycosides. Another putative genomic island, known as G41, encoded a single aminoglycoside resistance gene and several other genes putatively involved in fatty acid metabolism. This island was discovered to be exclusive to complete genomes of international clone II lineage *A. baumannii*, a phylogenetic grouping of *A. baumannii* often associated with multidrug resistant nosocomial infection.

The genome sequences of several other nosocomial and community-acquired *A. baumannii* isolates, originating from the Asia-Pacific region, were bioinformatically analysed as part of a collaborative project. Previously conducted phenotypic testing of these isolates had revealed that the CA-AB generally outperformed the nosocomial *A. baumannii* isolates with respect to virulence, including biofilm formation, phospholipase D production, murine mortality, and *in vitro* growth in LB, sera and 1% ethanol. Bioinformatic analysis of the CA-AB genomes revealed that many genes known to be involved in virulence in nosocomial *A. baumannii* were present, but there were few (if any) novel virulence traits that could conclusively account for the observed differences in virulence for nosocomial and CA-AB isolates.

The comparative genomic analyses of nosocomial and CA-AB had somewhat demarcated the clinical idiosyncrasies of the latter, including its antimicrobial susceptibility. Data resultant from the analysis of the various CA-AB genomes sequences, including that of D1279779, had suggested that the current focus on nosocomial *A. baumannii* genomics did not reflect the true diversity of this species. To that end, the genomes of several XN-AB isolates originating from industrial and environmental sites were sequenced for the purpose of comparison to previously sequenced nosocomial *A. baumannii* strains. It was discovered that the genomes of XN-AB isolates were generally smaller than those of nosocomial isolates, presumably owing to the lack of laterally acquired regions, such as prophages and the AbaR antimicrobial resistance island, and extensive insertion sequence-mediated genome decay in the case of one isolate. While accessory elements present in the genomes of nosocomial *A. baumannii* tended to confer resistance to antimicrobials,

those present in the genomes of XN-AB isolates were devoted to the metabolism of varying compounds, including sarcosine and monocyclic aromatic compounds.

Overall, the genomic analyses conducted during this PhD project indicate that *A. baumannii* isolates share a common gene pool quintessential to this organism, regardless of their nosocomial, community or environmental origin. This core gene set includes various genes that were hitherto associated with virulence and intrinsic antimicrobial resistance in nosocomial *A. baumannii*. These analyses have also identified an expanding pool of accessory genetic material, and lateral gene transfer has likely been a significant factor in the acquisition of these accessory genes. The differences in the accessory genome is likely a key factor in the differentiation between nosocomial and extranosocomial isolates of *A. baumannii*.

The accessory genome of nosocomial *A. baumannii* is biased towards antimicrobial resistance, which would presumably facilitate its survival in both the human host and within the clinical milieu. Conversely, the accessory genome of XN-AB is geared towards utilisation of novel carbonaceous and nitrogenous compounds, which would presumably allow these isolates to thrive in their respective natural and industrial environments.

Chapter 1: Introduction

1.0 *Acinetobacter* spp.

The *Acinetobacter* genus is a collective of Gram-negative bacteria that are strictly aerobic, non-fermentative, non-fastidious, aflagellate, catalase-positive and oxidase-negative [1]. There are currently thirty validated species of *Acinetobacter* [2] that have been enriched from soil, water [3] and the clinical environment [4]. These include *A. baylyi*, known for its metabolic versatility and high transformability [5], and *A. calcoaceticus*, known to be variably capable of degrading industrially important hydrocarbons, such as phenol [6] and catechin [7]. While these particular organisms are often associated with environmental settings, there are reports that these *Acinetobacter* species can potentially be a cause of nosocomial infection [4, 8]. Other *Acinetobacter* species have also been reported as causes of nosocomial infection, including *A. lwoffii*, *A. parvus*, *A. ursingii* and *A. nosocomialis* [4, 9, 10], but *A. baumannii* is the most important cause of infection within the *Acinetobacter* genus [4, 11].

1.1 *Acinetobacter baumannii*

1.1.1 Nosocomial *A. baumannii*

Acinetobacter baumannii in recent decades has emerged as a troublesome opportunistic pathogen in both the clinical milieu [12] and repatriated war casualties [13], particularly in the recent Iraq conflict, earning this organism both public notoriety and the moniker 'Iraqibacter' [14]. This organism is also notable for its repertoire of innate and laterally acquired mechanisms of resistance to various classes of antimicrobials, including aminoglycosides, β-lactams, sulphonamides and heavy metals [15, 16]. The antimicrobial resistance of ***A****. baumannii*, along with ***E****nterococcus faecium*, ***S****taphylococcus aureus*, ***K****lebsiella pneumoniae*, ***P****seudomonas aeruginosa*, and ***E****nterobacter* sp., is of continual concern to the medical community, and a source of great clinical and financial burden. This coterie of highly multidrug resistant organisms have fittingly been dubbed the ESKAPE pathogens by the Infectious Diseases Society of America [17, 18], and it is the ability of *A. baumannii* to continually 'escape' antimicrobial therapy and survive on assorted abiotic surfaces, including ceramic, polyvinyl chloride, rubber, stainless steel and glass [19, 20], that contributes to the clinical persistence of this organism.

A. baumannii is able to cause a variety of diseases including (but not limited to): ventilator-associated pneumonia, bacteraemia, secondary meningitis, urinary tract infection [21], endophthalmitis [22], necrotising fasciitis [23, 24], and complicated skin/soft tissue, abdominal, and central nervous system infections [25]. These diseases accompany physical distresses such as thermal and ballistic traumas [26, 27], physical transmission via invasive devices (gastric tubes, breathing tubes and catheters) [21], neurosurgical procedures (lumbar punctures, myelography and ventriculography) [28], or contaminated surfaces, such as computer keyboards [29].

1.1.1.1 Population structure of *A. baumannii*

Elucidation of the population structure of bacteria is crucial to the accurate identification and epidemiological surveillance of pathogenic organisms, ultimately influencing decisions regarding the implementation of healthcare controls [30]. Investigation of pathogen population structure can be conducted through a number of molecular methods, varying in resolution and effectiveness.

Nosocomial *A. baumannii* infections in Europe were discovered to fall within two major clonal lineages [31], and a subsequent third [32], based on the discovery of three distinct ribogroups generated through ribotyping [31, 33], the hybridisation of a labelled *E. coli rrnB* rRNA operon probe against digested DNA to generate a genetic fingerprint [34]. Due to its relatively low discriminatory power, ribotyping was insufficient to conclusively delineate the clonality of *A. baumannii* infections. Data obtained from higher resolution genetic fingerprinting methodologies, including pulsed-field gel electrophoresis (PFGE) and amplified fragment length polymorphism (AFLP), the selective amplification of restriction fragments to generate a genetic fingerprint [35], were congruent with the existence of three major clonal lineages [32]. These three clonal lineages of nosocomial *A. baumannii* were dubbed European clones I-III (ECI-III), now defunct given the existence of these clonal lineages outside of Europe. These three major *A. baumannii* clones have been since referred to as international clones I-III (ICI-ICIII) or global clones I-III (GCI-III) [36-38].

Genetic fingerprinting as a means of identification and epidemiological surveillance has the disadvantage of being laborious and expensive to perform, whilst data generated from these methods can make inter-laboratory comparison difficult [1, 30]. Multilocus sequence typing (MLST), the amplification and sequencing of short

internal fragments of multiple housekeeping genes [30], has largely replaced genetic fingerprinting as a typing method, due to the ability to perform rapid and global inter-laboratory comparisons, coupled with the continually reducing cost of high throughput sequencing [39].

Over 400 MLST sequence types have been discovered in *A. baumannii*, suggesting a highly divergent population structure, though the clinical setting is still predominated by clones I-III [36]. There are currently two MLST schemes used for the typing of *A. baumannii*, colloquially known as the Oxford [40] and Pasteur [37] schemes. Both schemes use a total of seven housekeeping genes, with *gpi*, *gyrB*, *gdhB*, and *rpoD* unique to the Oxford scheme; *fusA*, *pyrG*, *rplB*, *rpoB* unique to the Pasteur scheme, and the *cpn60*, *gltA*, and *recA* genes common to both. Though a head-to-head comparison of two methods has not been conducted in this organism, it has been found previously that the *gpi* allele of the Oxford scheme is prone to significant variation and recombination relative to other housekeeping genes [41], thus making it unsuitable as an epidemiological marker [42].

Whole genome sequencing (WGS) has recently been employed as a high resolution method to track epidemiological changes in *A. baumannii*, even amongst isolates otherwise indistinguishable through other molecular methods, with the ability to detect single nucleotide polymorphisms, recombination events and gene gain/loss [41, 43]. This technique has been applied in the tracking of evolutionary changes in nosocomial *A. baumannii* during polyclonal outbreaks [41, 43, 44] and tigecycline therapy [45]. Despite the decreasing cost and turnaround time, WGS cannot currently be applied to routine diagnostic microbiology, as it would require the technology to be redesigned to suit needs of diagnostic laboratories, including the minimisation of batch sizes, simplification of sample preparation, and the sensitivity to sequence a single colony, obviating the need for subculturing or DNA pre-amplification [46].

1.1.1.2 Molecular mechanisms of antimicrobial resistance in *A. baumannii*

Acinetobacter baumannii, like other bacterial pathogens, has developed various molecular defences to mitigate the deleterious effects of antimicrobials. The primary mechanisms of antimicrobial resistance in bacteria are increased efflux, decreased influx, target modification, target amplification, target repair, target bypass, intracellular localisation, sequestration, biofilm formation, and enzymatic inactivation [47], with some mechanisms relevant to *A. baumannii* expounded below.

1.1.1.2.1 Multidrug efflux transporters

All living cells express efflux transporters that can function in the discharge of potentially lethal compounds into the extracellular space, thereby diminishing their cellular toxicity. Multidrug efflux transporters in some Gram-negative bacteria are adapted to expel a broad range of antimicrobial substrates, currently falling into five key protein families: the ATP binding cassette (ABC) family, the major facilitator superfamily (MFS), the resistance-nodulation-division (RND) family, the multidrug and toxic compound extrusion (MATE) family and the small multidrug resistance (SMR) family [16].

Members of each transporter family are encoded within the *A. baumannii* genome, with the presence of many being common to clinical and non-clinical isolates alike. However, the absolute numbers of transporters in *A. baumannii* do vary per isolate, with the presence of some influenced by mobile genetic elements (MGEs). Transporters of the MFS protein family are predominant in this organism, which alongside the RND transporter family, couple proton translocation to antimicrobial expulsion [48, 49].

One of the foremost efflux-based resistance determinants in *A. baumannii* is AdeABC, a tripartite RND family transporter able to bestow resistance to a range of antimicrobial substrates including aminoglycosides (gentamycin, tobramycin, netilmicin), fluoroquinolones (sparfloxacin, pefloxacin, norfloxacin), cefotaxime, erythromycin, tetracycline, chloramphenicol, trimethoprim [50] and tigecycline [51, 52], a glycylcycline antibiotic often prescribed as a final resort treatment for multidrug resistant Gram-negative infections [53]. This transport system is tightly regulated by the *adeRS* two-component regulator; point mutations in *adeR* (Pro116→Leu) and *adeS* (Thr153→Met) [54], or insertion of IS*Aba1* [52], have been linked to constitutive expression of *adeABC*, bestowing the multidrug resistance phenotype observed in clinical isolates.

AdeFGH and AdeIJK are two other RND family transporters that have been demonstrated to bestow multidrug resistance in *A. baumannii*. Substrates of AdeFGH include chloramphenicol, clindamycin, fluoroquinolones (ciprofloxacin, moxifloxacin, norfloxacin), trimethoprim, tetracyclines (minocycline, tetracycline), tigecycline, and sulphonamides (sulphamethoxazole, co-trimoxazole) [55], while AdeIJK recognises

β-lactams (ticarcillin, cefotaxime), chloramphenicol, tetracyclines (minocycline, tetracycline), erythromycin, fluoroquinolones (levofloxacin, moxifloxacin, perfloxacin) fusidic acid, novobiocin, and trimethoprim as substrates [56]. Like *adeABC*, both *adeFGH* and *adeIJK* are under tight regulator control, though rather than a two-component system, these transporters are controlled by the single-component regulators AdeL [55] and AdeN [57], respectively.

A. baumannii also encodes several characterised MFS transporters, which in contrast to the abovementioned RND transporters, usually consist of a single protein subunit, of a narrower substrate specificity. AmvA [58] (also known as AdeF [49]) and CraA [59] are two MFS family transporters which are known to significantly confer resistance to erythromycin and chloramphenicol, respectively. AdeF was also known to export the antiseptic chlorhexidine, but only when heterologously expressed in *Escherichia coli*; no transcriptional response could be elicited in its *A. baumannii* host [49]. While all three of the above transporters are common to various *A. baumannii* isolates, the TetA and TetB MFS transporters, which confer resistance to tetracyclines, are known only to be carried within MGEs such as transposons [16] or plasmids [60].

There are also a handful of other transporters characterised in *A. baumannii* not of the above two families, including the MATE family transporter AbeM, which through heterologous expression was discovered to extrude aminoglycosides, fluoroquinolones, chloramphenicol and trimethoprim [61], and the SMR family transporter AbeS, which bestows low-level resistance to chloramphenicol, fluoroquinolones, erythromycin and novobiocin [62]. Recently a transporter not part of any of the aforementioned transporter families was discovered to contribute to chlorhexidine resistance in *A. baumannii* [63]. This chlorhexidine efflux pump, dubbed AceI, is conserved in *A. baumannii* and a number of Proteobacterial organisms, resulting in the foundation of a new drug resistance efflux protein family known as the Proteobacterial chlorhexidine efflux (PCE) family [63].

1.1.1.2.2 Decreased influx due to disruption of outer membrane proteins

Outer membrane proteins (OMPs) function in the passive diffusion of small solutes into the bacterial cell, the permissiveness of which varies due to transcriptional changes induced by nutrient starvation or the presence of toxins [64]. CarO, previously known as the 29-kDa OM protein [65], is the best characterised OMP in *A.*

baumannii, and is essential for the passive uptake of ornithine [66]. Loss of this OMP through insertion sequence inactivation is known to contribute to carbapenem resistance in this organism [65, 67]. Furthermore, the *carO* gene also appears to have itself been subject to lateral gene transfer, as polymorphic variants of CarO have arisen in *A. baumannii* as a result of lateral transfer and recombination of DNA encoding CarO [68]. The contribution of this genetic exchange to carbapenem resistance in *A. baumannii* awaits further examination.

The decreased expression of other OMPs in the membrane proteome of *A. baumannii* have also been implicated in carbapenem resistance, including proteins encoding the 22 kDa, 33 kDa [69], 33 to 36 kDa [70], 37 kDa, 44 kDa and 47 kDa [71] OMPs. However, a 43kDa OMP known as OprD, was concluded not to be involved in carbapenem resistance of *A. baumannii* [72]. This observation was confirmed in a separate study, where it was suggested that other studies, that examined the involvement of OprD in carbapenem resistance, had failed to take into account the contribution of other porins that may have been lost [73].

Fluctuations in efflux and influx have been noted to synergistically increase antimicrobial resistance in various Gram-negative bacteria [74]. In the case of *A. baumannii* however, extrusion of β-lactams through efflux transport leads to a minimal increase to β-lactam resistance, with β-lactamases being the major mechanism of resistance in this organism [75]. These and other enzymatic mechanisms of resistance are described below.

1.1.1.2.3 Enzymatic inactivation of antimicrobials

A. baumannii is capable of enzymatically inactivating antimicrobials of the aminoglycoside, rifamycin and β-lactam classes [76]. Resistance to aminoglycosides in *A. baumannii* is chiefly mediated by the aminoglycoside-modifying enzymes (AMEs) [77], aminoglycoside phosphotransferase, acetyltransferase, and adenyltransferase, which covalently attach phosphate, acetate and adenosine moieties to aminoglycosides, respectively [78]. Multidrug resistant *A. baumannii* generally encodes at least one of the AME types, and are carried exclusively within MGEs. In this organism, these resistance genes are predominately extant as integron cassettes [79], and to a lesser extent are carried within transposons [80, 81] and integrons [82] present within plasmids.

A. baumannii is also able to covalently modify rifamycins through the action of the rifamycin ADP-ribosyltransferase enzyme (*arr-2*) [76], which catalyses attachment of ADP-ribose to a free hydroxyl group on a rifamycin [83]. As with the AMEs, the carriage of the *arr-2* gene is mediated exclusively by MGEs, in this instance as an integron cassette. While AMEs are found within the majority of *A. baumannii* integrons [79], the presence of the *arr-2* gene within integrons of this organism is relatively sporadic [76].

β-lactam antibiotics can also be enzymatically inactivated by *A. baumannii*, through the action of β-lactamases, which rather than covalently modify the antibiotic, as in the case of *arr-2* and the AMEs, catalyse cleavage of the four-membered β-lactam ring [77], from which this antibiotic class derives its name. Another thing that sets *A. baumannii* β-lactamases apart from *arr-2* and the AMEs, is that while genes encoding β-lactamases are also encoded on various MGEs, such as integron cassettes, plasmids and transposons [75, 84], there are instances of β-lactamases that are encoded in the core genome of *A. baumannii*. AmpC [85] and OXA-69 [86] are two characterised chromosomal β-lactamases that are common to various isolates of *A. baumannii*. However these chromosomal β-lactamases have a low level of constitutive expression, and are not inducible by β-lactams, thus only bestowing a minimal increase to antibiotic resistance [85, 86]. Nevertheless, when insertion sequences, such as IS*Aba1*, are integrated upstream of these chromosomal β-lactamases, the promoter encoded within the insertion sequence causes overexpression of chromosomal β-lactamases, resulting in increased resistance to β-lactams [85, 87].

1.1.1.2.4 Modification or loss of antimicrobial targets

A. baumannii is capable of rendering certain antibiotics ineffective through modification of their target sites. Fluoroquinolones are, as outlined above, able to be extruded through efflux transporters in *A. baumannii*, but resistance can be further augmented by mutations in the *gyrA* gene, which encodes the target of this antibiotic class, DNA gyrase A [88]. Fluoroquinolone resistance in *A. baumannii* has also been linked to mutations in the *parC* gene, encoding a topoisomerase IV subunit. ParC is suggested to be a secondary target of fluoroquinolones, as it has been observed that *parC* mutations only occur simultaneously with *gyrA* mutations [89]. Mutations in *gyrA* and *parC* primarily occur within the serine codons of both genes, as well as in their respective glycine and glutamate codons [90].

Like fluoroquinolone resistance, resistance to aminoglycosides in *A. baumannii* is mediated via efflux transporter extrusion, in addition to enzymatic modification. However, a third mechanism of aminoglycoside resistance has been discovered in this organism, which involves methylation of the 16S rRNA subunit by the 16S rRNA methylase, ArmA [91]. The *armA* gene encoding this methylase is encoded on several plasmids of *A. baumannii*.

The decreasing effectiveness of the abovementioned antimicrobial treatments has driven the re-emergence of polymyxin antimicrobials as treatments for *A. baumannii* and other Gram-negative pathogens. This is despite having previously fallen into disuse during the 1970s, due to its apparent nephrotoxic and neurotoxic effects [92]. Polymyxins, including colistin and polymyxin B, are amphiphilic cations that bind the amphiphilic and anionic lipid A molecule of bacterial lipopolysaccharide, causing disruption of the bacterial cell membrane [93]. Unfortunately, *A. baumannii* has also demonstrated resistance to polymyxins, be it through mutation of genes involved in lipid A biosynthesis, resulting in loss of lipopolysaccharide [94], or structural modification of lipid A, altering the electrostatic charge of the bacterial membrane [95, 96].

1.1.1.3 Molecular basis of virulence in *A. baumannii*

Until recently, *A. baumannii* was deemed not to be a considerable cause of clinical mortality [97], and as result was often classified as a low-virulence pathogen by the medical community [98]. While *A. baumannii* is now acknowledged as a significant human pathogen, still relatively little is known regarding the virulence mechanisms of this organism [75], with those best characterised outlined below.

1.1.1.3.1 Twitching and surface-associated motility

Members of the *Acinetobacter* genus have historically been described as non-motile due to the absence of flagella [3]. Nonetheless, *A. baumannii* was recently found to be capable of aflagellate motility, categorised as either surface-associated motility or twitching motility, based on whether surface translocation occurred on a semi-solid agar surface or within the agar-plastic interface of solid media, respectively [99, 100]. This motility was thought to be associated with a gene cluster putatively encoding type IV pili, with a positive correlation demonstrated between the sequence conservation of the PilA subunit and the motility phenotype exhibited [99]. Type IV pili

were confirmed to be required for twitching motility and natural transformation, but not for surface-associated motility in *A. nosocomialis* M2 (previously *A. baumannii* [101]) [100]. In contrast to twitching motility, surface-associated motility is a multifactorial process partially dependant on 1,3-diaminopropane synthesis [102], lipopolysaccharide composition [103] and quorum sensing [104].

1.1.1.3.2 Biofilm formation

Adhesion to or colonisation of host cells by bacteria is the first step to the eventual formation of a highly structured microbial community, consisting of a polymeric matrix of bacterial cells and organic macromolecules, known as a biofilm [105]. Biofilm formation in *A. baumannii* contributes to its ability to resist antimicrobial therapies, dehydration and nutrient starvation, and thus represents an important virulence factor for this organism and others [106]. Biofilm formation, like surface-associated motility, is also multifactorial and dependant on multiple molecular determinants, including genes encoding a biofilm associated protein (*bap*) [107, 108], outer membrane protein A (*ompA*) [109], synthesis of poly-β-1-6-N-acetylglucosamine (PNAG) (*pgaABCD*) [110] and a chaperone-usher pili assembly system (*csuA/BABCDE*) [111]. The multifactorial nature of biofilm formation was further elucidated through comparative transcriptomics of *A. baumannii* cells in biofilm and planktonic (free living) stages, revealing several other genes involved in biofilm formation [112].

1.1.1.3.3 Quorum sensing

Bacterial cells are able to respond to external stimuli through the synthesis and excretion of small organic compounds, including amino acids and fatty acid derivatives, that facilitate cell-to-cell signalling [113]. The concentration of these signalling molecules are related to the population density of the producer organism, which upon reaching a critical extracellular concentration results in a coordinated phenotypic response of the whole cell population [113]. The process taken as a whole is known as quorum sensing (QS).

As mentioned above, QS has an influence on both motility and biofilm formation in *A. baumannii*, and is therefore also an important virulence factor [114]. A gene encoding an autoinducer synthase (*abaI*) in *A. baumannii* was discovered in *A. nosocomialis* M2 (previously *A. baumannii*) and facilitated the synthesis of the N-acyl homoserine lactone (AHL) analogue N-(3-hydroxydodecanoyl)-L-HSL, thought to be recognised by its putative cognate receptor, encoded by the *abaR* gene [115]. Loss of *abaI* is

associated with reduced motility and biofilm formation, which are restored upon supplementation of N-(3-hydroxydodecanoyl)-L-HSL [104, 115]. Another study examined the effect of various synthetic exogenous AHLs in an Δ*abaI A. baumannii* reporter strain, which confirmed AHLs resembling N-(3-hydroxydodecanoyl)-L-HSL acted as agonists of AbaR, while synthetic AHLs containing aromatic acyl groups were found to be highly antagonistic for motility and biofilm formation [116].

1.1.1.3.4 Iron acquisition

Iron is an essential micronutrient for almost all living organisms. Iron is of low bioavailability in the human host, as it is often locked within iron-binding molecules such as heme, lactoferrin, and transferrin [75]. Thus, iron acquisition is an important factor in the success of bacterial pathogens. *A. baumannii* encodes a number of mechanisms involved in the acquisition of iron, and are outlined below.

During iron deprivation, *A. baumannii* responds with the overexpression of numerous genes involved in the biosynthesis and transport of siderophores, organic compounds that function as high-affinity iron chelators [117]. Several siderophore biosynthesis genes clusters have been identified in the *A. baumannii* genome [117]. The cluster encoding the biosynthesis of the siderophore acinetobactin [118] is the most extensively studied, and is one of two *A. baumannii* siderophore types structurally determined to date [119]. Impairment of acinetobactin biosynthesis reduced the ability of *A. baumannii* to persist and cause death in murine and *Galleria mellonella* (greater wax moth) infection models, but has no impact on its ability to interact with human lung epithelial cells [120]. A recently synthesised compound (6-phenyl-1-(pyridin-4-ylmethyl)-1H-pyrazolo[3,4-b]pyridine-4-carboxylic acid) was found to act as a blocker of the BasE protein involved in acinetobactin biosynthesis, signifying it as a potential candidate for antimicrobial treatment [121]. Fimsbactins A-F are another siderophore family able to be synthesised by *A. baumannii*, and have recently been structurally determined [122]. However, the biosynthesis pathway for these siderophores is known to be encoded only in the genome of *A. baumannii* ATCC 17978 [122], and its involvement in virulence has yet to demonstrated.

A. baumannii can also acquire iron from haem through the action of a chromosomally encoded gene cluster involved in the degradation of haem, which includes a haem oxygenase (*hemO*) [123]. Like the fimsbactin biosynthesis gene cluster, the haem oxygenase gene cluster is not present in the core genome of *A. baumannii*. The

louse isolate, *A. baumannii* SDF, is thought to rely on the acquisition of iron from haem, as it is the only completely sequenced *A. baumannii* isolate to lack the conserved acinetobactin biosynthesis pathway [123].

1.1.1.3.5 Capsule polysaccharide biosynthesis and glycosylation

Surface polysaccharides, such as capsule, have been long recognised as an important virulence trait in Gram-negative bacterial pathogens. *A. baumannii* was found to be no exception in this regard, as disruption of two genes involved in capsule export and polymerisation resulted in a complete loss of capsule polysaccharide production, and demonstrated reduced virulence in various infection models [124]. These genes form part of a large and highly variable gene locus involved in the biosynthesis of complex carbohydrates, and form the structural basis of capsule [125]. This capsule biosynthesis gene cluster also encodes the *plgC* gene involved in a separate, but otherwise mutual, O-linked protein glycosylation pathway [126]. Inactivation of *pglC* prevented the synthesis of both glycoproteins and capsule, resulting in deficiencies in biofilm formation and virulence in a murine model [126]. Another gene involved in O-linked protein glycosylation (*pglL*) is encoded elsewhere in the genome of *A. baumannii*, however PglL is involved in a general glycosylation pathway of multiple proteins [127]. Loss of *pglL* also resulted in reduced biofilm formation and overall loss of virulence [127].

1.1.2 Extranosocomial *Acinetobacter baumannii*

Whilst *A. baumannii* is relatively well studied as both a human pathogen and a coloniser of the clinical milieu, little is known regarding the niche distribution of extranosocomial *A. baumannii* (XN-AB), or its association with nosocomial infection [128]. Investigations into potential external sources of nosocomial *A. baumannii*, including meat and vegetables [129, 130], soil [130, 131], anthropogenic surfaces [132] and the community [133], revealed that *A. baumannii* isolated from these environments were generally unrelated to nosocomial isolates. Despite the apparent absence of a causal link between XN-AB and nosocomial infections, the presence of antimicrobial resistance genes in some extranosocomial isolates [134-136] may nonetheless serve as a reservoir of resistance traits, with exchange mediated through the acquisition of MGEs. Furthermore, there have been reports of XN-AB capable of pathogenicity in both animals [137] and sugarcane [138], with community-acquired *A. baumannii* representing the most prevalent cause of extranosocomial infections in humans.

1.1.2.1 Community-acquired *Acinetobacter baumannii*

Community-acquired *A. baumannii* (CA-AB) are pathogenic agents originating from an unknown reservoir outside the clinical milieu, thought to be transmitted between residents of a community [139]. Like nosocomial *A. baumannii*, CA-AB have been known to cause multiple diseases in humans, including pneumonia [140], bacteraemia [141], meningitis [142] and endocarditis [143]. However, CA-AB infections are clinically and epidemiologically idiosyncratic compared to their nosocomial counterparts, with CA-AB isolates not only being unrelated to nosocomial isolates [144], but also lacking their clonal population structure [139, 141]. CA-AB infections only comprise less than 10% of all *A. baumannii* infections [144, 145], but are often fulminant and fatal, resulting in mortalities ranging from 30-62% [144, 146, 147]. Infections caused by CA-AB are also relatively susceptible to antimicrobial treatments otherwise ineffective on nosocomial infections, including ciprofloxacin, cotrimoxazole and tobramycin [133, 147]. There has however been a single report of a multidrug resistant CA-AB, causing a fatal case of fulminating septicaemia [148].

CA-AB infections predominantly occur in individuals with underlying comorbidities, including alcoholism, chronic obstructive pulmonary disease, diabetes mellitus and cigarette smoking [144]. Incidences of CA-AB infections tend to be limited to tropical and subtropical climates, often reported within regions of the Asia-Pacific, including Taiwan [141], Hong Kong [147], Singapore [149], Korea [150] and Australia [151]. To a lesser extent though, CA-AB infections have also been observed in non-tropical regions [145], during cooler seasons [152], as well as in otherwise healthy adults and children [142, 153-155].

1.1.3 Genomics of *Acinetobacter baumannii*

Whole genome sequencing via next-generation technologies has been instrumental in providing insight into both the genetic architecture and evolutionary trajectory of *A. baumannii*, particularly in isolates of diverse niches or exceptional clinical persistence. To date, there are fourteen complete genome sequences of *A. baumannii* available, the features of which are summarised in Table 1.1. These genomes vary in both size (3.24-4.14 Mbp) and number of plasmids (0-4), but all have a GC% content of roughly 39%.

Table 1.1: Names, origins and features of the currently complete genomes of fourteen *Acinetobacter baumannii* isolates (alphabetically arranged).

Strain	Country of origin	Infection Type	Size (bp)	GC%	Plasmids	Reference
1656-2	South Korea	Nosocomial	3940614	39.1	2	[156]
AB0057	USA	Nosocomial	4050513	39.2	2	[157]
AB307-0294	USA	Nosocomial	3760981	39.0	0	[157]
ACICU	Italy	Nosocomial	3904116	38.9	2	[158]
ATCC 17978	N/A	Nosocomial	3976747	38.9	2	[159]
AYE	France	Nosocomial	3936291	39.3	4	[160]
BJAB07104	China	Nosocomial	3951920	39.0	2	[161]
BJAB0715	China	Nosocomial	4001621	38.9	1	[161]
BJAB0868	China	Nosocomial	3906795	38.9	3	[161]
MDR-TJ	China	Nosocomial	3964912	39.1	2	[162]
MDR-ZJ06	China	Nosocomial	3991133	39.0	1	[163]
SDF	France	Body louse gut	3421954	39.1	3	[160]
TCDC-AB0715	Taiwan	Nosocomial	4138388	38.9	2	[164]
TYTH-1	Taiwan	Nosocomial	3957368	39.0	0	[165]

Publically available *A. baumannii* genome sequences are limited to those of nosocomial isolates, with the exception of the body louse gut isolate, *A. baumannii* SDF, representing the only completely sequenced extranosocomial isolate available at the time. It is of note though that other louse borne isolates, akin to *A. baumannii* SDF, have been experimentally demonstrated to cause infection in a rabbit model [137]. The genome of a presumably non-pathogenic mangrove soil isolate, *A. baumannii* MSP4-16, has recently been sequenced [166], though this genome is of draft status and has not been examined in detail.

Multiple comparative analyses conducted with several of the above *A. baumannii* genomes have demonstrated a highly syntenic genome organisation [161, 167], despite the geographic variation of these isolates. There are notable exceptions to this conserved genome architecture, including an 800 kb inversion in the genome of *A. baumannii* BJAB07104 [161], and the extensive genomic reduction and rearrangement of the *A. baumannii* SDF genome [160]. The core genome of *A. baumannii*, consisting of a set of homologous coding sequences (CDSs), varies between 1455 and 2688 CDSs, depending on the number and identity of strains analysed [36]. The majority of these core CDSs are involved in metabolic and general cellular processes, followed by hypothetical proteins, which are proteins having no characterised homologues [168]. In contrast, the accessory genome, which will be further outlined below, is enriched in transport and transcription regulation functions [157, 168], but also encompasses various regions of plasticity and laterally acquired regions that encode proteins involved in metabolism and antimicrobial resistance [167].

1.2. The accessory genome

The coding content of an organism, determined through genome sequencing, is often examined in the context of the pangenome, the set of all genes present in the genomes of a group of sequenced organisms [169], typically on a genus/species level. Genes contained in the pangenome can be divided into the core genome, the genes common to all member genomes of the group, and the accessory genome, an assortment of genes limited to one or few strains of the member genomes [170]. The diversity of the accessory genome can be attributed to regions of genomic plasticity and the accretion of MGEs.

1.2.1 Mobile genetic elements

Mobile genetic elements (MGEs) are genomic entities that facilitate the intracellular or intercellular translocation of DNA [171]. Examples of MGEs include conjugative transposons, integrated plasmids, integrative conjugative elements, transposons, insertion sequences, integron cassettes, prophages and plasmids [171]; only the last five MGE types are known to exist in *A. baumannii* and will be explicated further.

1.2.1.1 Insertion sequences and transposons

The most rudimentary of the MGEs, insertion sequences (ISs), are small (<2.5 kb) and genetically streamlined elements. ISs consist of a transposase gene that constitutes the bulk of these elements, and facilitates its mobilisation, acting on a pair of small (10-40 bp) terminal sequences that flank ISs called inverted repeats [172] (Figure 1.1).

ISs are capable of inserting into multiple sites within a host chromosome, including on occasion protein coding genes, or adjacent to them. This can potentially result in gene inactivation or alterations in gene expression, due to the influence of outward facing promoters present in the IS [173] (Figure 1.1). IS*Aba1* is one of the most widely distributed ISs in nosocomial *A. baumannii* [174], and as mentioned previously, is responsible for overexpression of both *adeABC* [52] and latent chromosomal β-lactamases [85, 87, 175], leading to increased tigecycline and β-lactam resistance, respectively. Multiple IS copies can exist within the genome of an organism and lead to genome decay mediated through homologous recombination between ISs [173]. This IS-mediated genome decay is thought to be responsible for the reduction in genome size of various pathogenic agents, including *Mycobacterium*

leprae [176], *Leptospira borgpetersenii* [177] and *A. baumannii* SDF [160], as well as numerous other chromosomal gene loss events in other *A. baumannii* isolates [44].

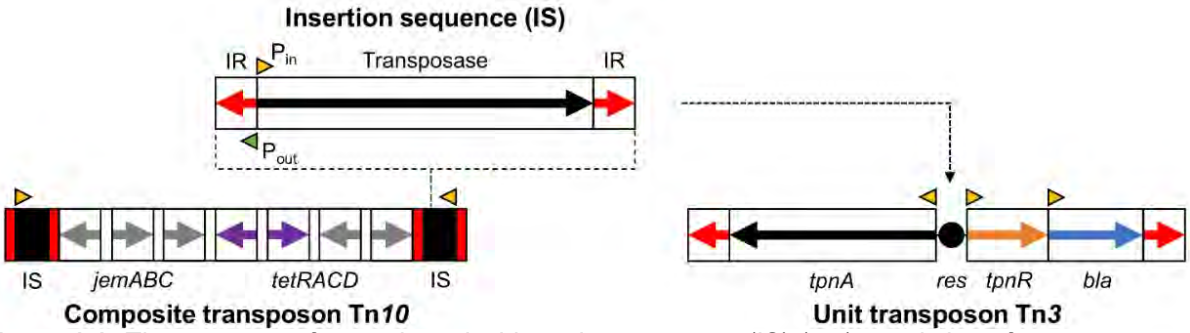


Figure 1.1: The structure of an archetypical insertion sequence (IS) (top) consisting of a transposase gene (black), a pair of inverted repeats (IR) (red), and a set of outward (P_{out}) (green) and inward (P_{in}) (gold) facing promoters. ISs can form increasingly complex structures called unit transposons (right), which consist of the same components as an IS, except for the presence of additional protein coding genes. Like IS elements, the Tn*3* unit transposon consists of a transposase gene (*tpnA*) and a pair of IRs, but also encodes a resolvase (*tpnR*) (orange) and a resolution site (*res*) (black circle) required for transposition, and a β-lactamase (*bla*) (blue) for β-lactam resistance. Two or more ISs can form a composite transposon (left), encapsulating and mobilising a large segment of DNA. In this instance, transposon Tn*10* consists of a pair of ISs flanking genes involved in tetracycline resistance (*tetRA*) (purple), and genes of uncharacterised function (*jemABC* and *tetCD*) (grey). Arrows indicate direction of transcription and figures were adapted from [178-180], but not to scale.

IS elements can form increasingly complex MGEs called transposons, and are divisible into two subcategories based on their genetic architecture. Unit (or non-composite) transposons, are architecturally identical to ISs, except for the presence of one or more genes typically involved in antimicrobial resistance [181]. The 4,957 bp transposon Tn*3* is an example of an archetypical unit transposon, encoding only three genes; a Tn*3* transposase (*tnpA*), a Tn*3* resolvase (*tnpR*) and a β-lactamase (*bla*) [180] (Figure 1.1).

Conversely, composite transposons consist of a larger segment of DNA flanked by two IS elements, in an inverted orientation, and is mobilised as a single discrete unit by transposase catalysis [181]; Tn*10* is an example of a prominent and archetypical composite transposon. Tn*10* is a 9,147 bp element encoding nine genes, including two transposition genes, two tetracycline resistance genes (*tetRA*), encoding a regulator and a tetracycline efflux pump, and five genes (*jemABC* and *tetCD*) of uncharacterised function [179] (Figure 1.1).

A. baumannii can harbour various composite transposons in its genome, typically containing antimicrobial genes, such as β-lactamase genes [182]. The largest and most significant composite transposon identified in this organism, though often denoted as a genomic island [183], is the *Acinetobacter baumannii* antibiotic resistance island (AbaR) (Figure 1.1), recently redubbed Tn*AbaR*, as it is a distant

relative of the Tn*7* transposon [184]. Tn*AbaR* typically disrupts the *comM* gene in this organism [185, 186], and is comprised of numerous antimicrobial resistance genes and an amalgamation of various MGEs including ISs and integrons. *A. baumannii* AYE Tn*AbaR1* (Figure 1.2) is considered to be the archetypical Tn*AbaR*, and is the largest to date [187], consisting of an ~86 kb stretch of sequence that encodes eighty-eight putative genes, including three class 1 integron arrays, several ISs, and resistance to various antibiotics and heavy metals [15]. All other Tn*AbaR* islands have components in common with Tn*AbaR1*, including genes encoding cadmium and arsenic resistance, and one of the three integron arrays present in Tn*AbaR1*, which encodes resistance to aminoglycosides and sulphonamides [187]. Tn*AbaR4*-type islands, initially identified in *A. baumannii* AB0057 [157], also encode several antimicrobial genes but are distinctive from the archetypical Tn*AbaR1* island in that they are integrated into the *pho* gene of *A. baumannii*, rather than the *comM* gene [188]. Conversely, AbaG1 in *A. baumannii* SDF is inserted within the *comM* gene, however it does not encode any genes annotated as mediating antimicrobial resistance [15, 186].

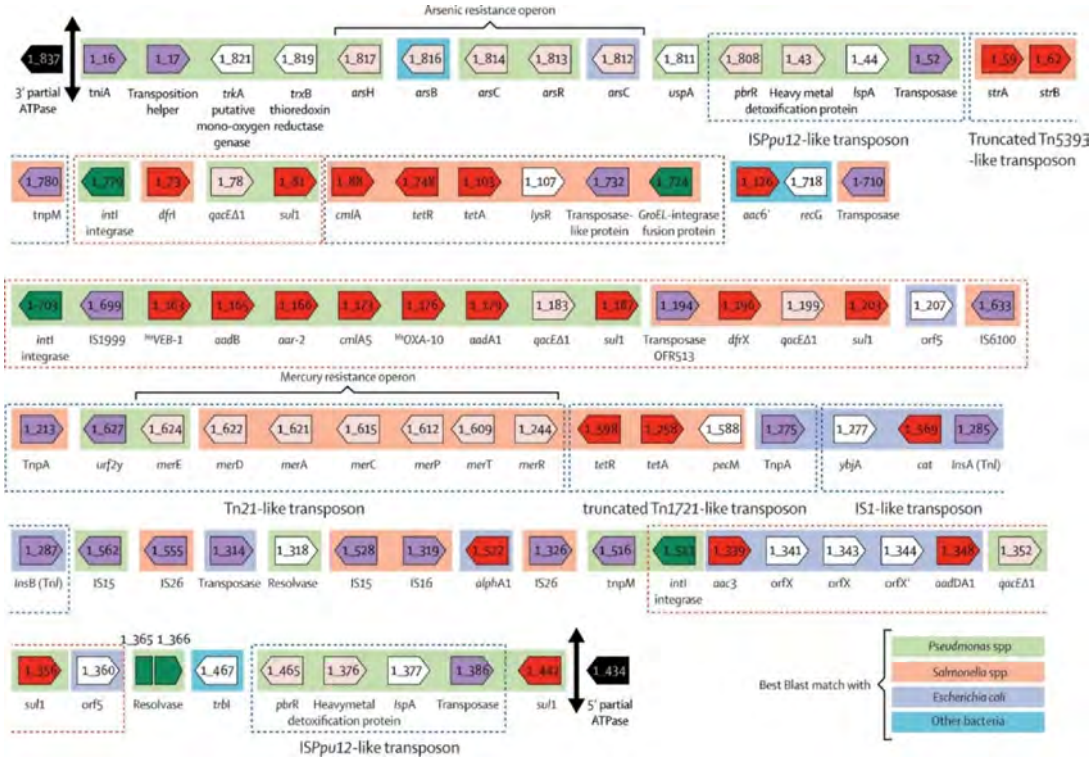


Figure 1.2: The layout of *Acinetobacter baumannii* AYE Tn*AbaR1* integrated into the *comM* gene (partial ATPase). Each open reading frame (ORF) is coloured according to its functional category: antibiotic resistance (red), heavy metal/antiseptic resistance (fawn), transposase (purple), integrase (dark green) and other function (white). The ORFs are shaded according to the best matching genus/species of origin (other than *Acinetobacter* spp.) as determined by BLAST. Complete integrons (red dashed lines), transposons (blue dashed lines) and a gene cluster found in *Salmonella* genomic island 1 (black dashed lines) are indicated. This figure was reproduced from [189] and is not to scale.

1.2.1.2 Integrons

Integrons are DNA assembly platforms consisting of an integron integrase (*intI*), an attachment *attI* site, and gene cassette(s), which are segments of DNA usually encoding an ORF that have been inserted in the *attI* site via site-specific recombination. Also present is an outward facing promoter (P_c) to facilitate the expression of gene cassettes [190, 191] (Figure 1.3). Unlike other MGEs, integrons are immobile unless encapsulated within a transposon [191], as IntI does not catalyse the mobility of the integron, but rather the excision of gene cassettes as circular DNA molecules (covered later), or their site-specific integration, forming an imperfect inverted repeat *(attC*) [192, 193]. Integrons were first discovered due to their frequent carriage of antibiotic resistance gene cassettes, and a large number of antibiotic resistance gene cassettes have been described encoding resistance to β-lactams, aminoglycosides, chloramphenicol, tetracycline, trimethoprim and other antimicrobial agents.

Integrons typically consist of a small number of genes cassettes, though several chromosomally encoded variants have been discovered in *Vibrio* sp. that are exceptionally large, including the 130 kb chromosomal integron array present in *Vibrio cholerae* [194]. Integrons are further subcategorised as classes 1-3 based on integrase sequence homology [195]. Class 1 and 2 integrons have been detected in *A. baumannii*, but not those of class 3 [196, 197]. Class 1 integrons are often carried within the Tn*AbaR* island, but can occasionally be found in plasmids [82] and integrated into other genomic loci [44]. *A. baumannii* integrons often encode resistance to various antimicrobial agents, including aminoglycosides, β-lactams, chloramphenicol and trimethoprim.

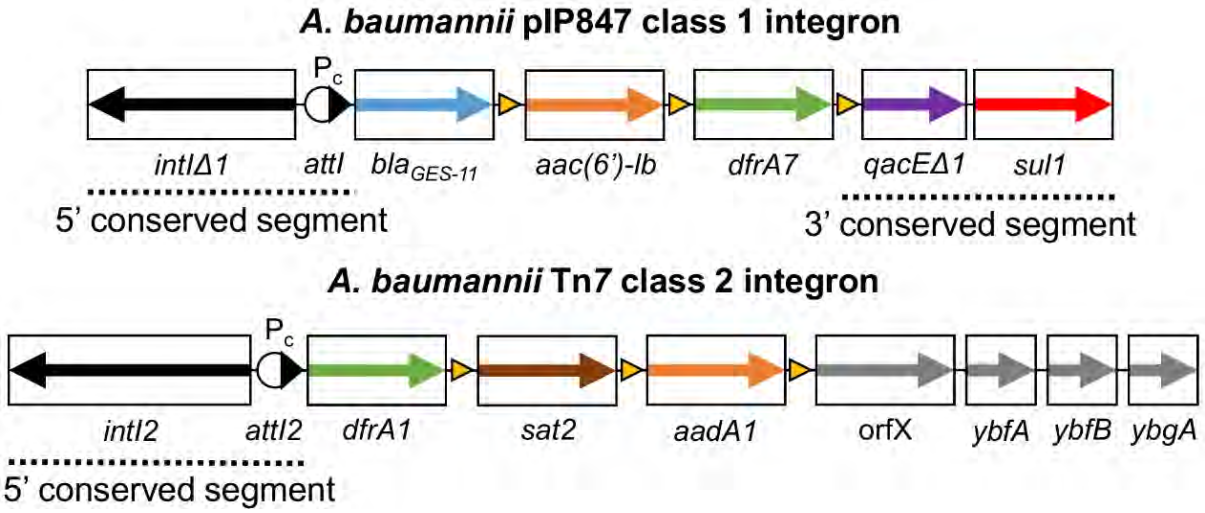


Figure 1.3: Structure of class 1 and 2 integrons in *A. baumannii* which consists of an integron integrase (*intI*), recombination site (*attI*) (white circles), cassette promoter (P_c) (black triangles), *attC* sites (gold triangles) and several cassette genes involved in antimicrobial resistance (coloured arrows) or of unknown function (grey). The *intI* gene of pIP847 has a small deletion, which is otherwise atypical of class 1 integrons in *A. baumannii*. Figures were adapted from [82, 198] and are not drawn to scale.

1.2.1.3 Prophages

Prophages refer to the lysogenic stage of temperate *Caudovirales* bacterial viruses, including the archetypal bacteriophage, *E. coli* phage λ [199] (Figure 1.4). Upon infection, these viruses inject their genetic material into the bacterial cell, integrating into the host chromosome as a dormant lysogen or exist as a linear (e.g. *E. coli* phage N15) or circular (e.g. *E. coli* phage P1) extrachromosomal replicons [199]. Lysogens transition to the lytic cycle when the host cell is subjected to stressors, such as DNA damage, leading to multiplication of the bacteriophage, host cell death, and phage release.

These bacteriophages can also act as agents of lateral gene transfer, transferring host chromosomal DNA to other bacterial hosts through transduction. Occasionally, phages can unintentionally pack its head full of bacterial DNA in the place of viral DNA, potentially incorporating into another bacterial chromosomal upon infection, known as generalised transduction [200]. Conversely, phages can also transfer bacterial DNA through specialised transduction, the inadvertent co-packaging of constrained sections of the host chromosome, due to imprecise excision of the prophage, followed by the infection of a new host [200]. Bacterial genes contained in phage lysogens are often uninvolved in phage function and are capable of autonomous transcription, regardless of prophage repression; these genetic elements have been dubbed morons ('more DNA') [201]. Phage morons, other than acting as a source of novel genetic material, can result in lysogenic conversion, the

induction of phenotypic changes in the host bacterium that are not part of the phage cycle, including imperviousness to further bacteriophage attack or augmented virulence [202]. However, negative lysogenic conversion is possible when bacteriophage integration results in disruption of a protein coding gene, with a subsequent loss of phenotype [202].

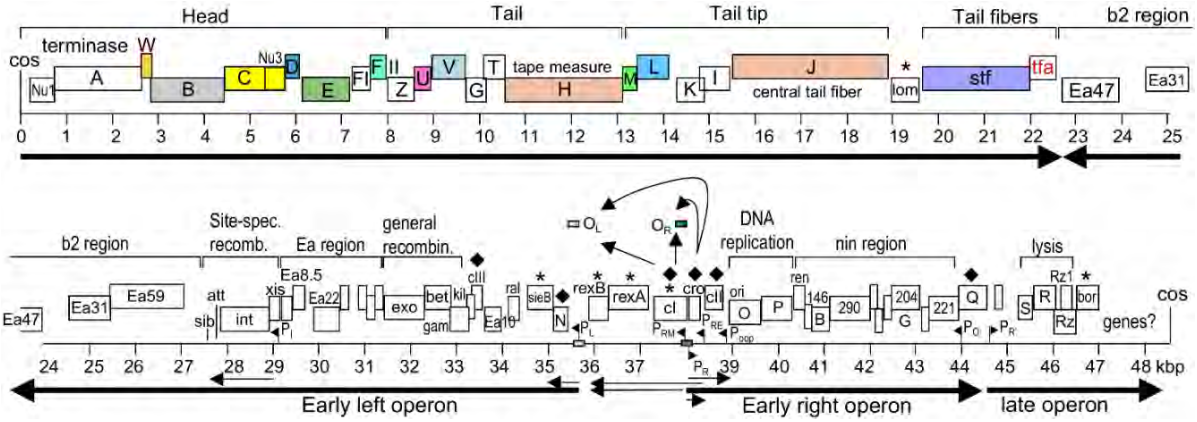


Figure 1.4: Genome map of enterobacteria phage λ, displaying a polycistronic gene organisation divisible in functional roles such as formation of the phage viron (coloured genes), as well as integration, replication and host lysis (white genes). This figure was reproduced from [203].

Genome sequencing of *A. baumannii* has revealed the presence of various prophages [157, 158, 162], some predicted to be cryptic, being unable to function as phages. Prophages in *A. baumannii* are currently uncharacterised, though some are known to harbour phage morons including genes encoding a 3'-phosphoadenosine 5'-phosphosulfate sulfotransferase (*papS*) and an error-prone DNA polymerase V (*umuDC*). While debatable, these particular morons could theoretically confer advantages to the host in iron limited conditions or through bestowing a mutator phenotype, respectively [167]. The only *A. baumannii* bacteriophages that have been characterised so far are lytic phages, which are unable to form lysogens. Such lytic phages may be desirable for the creation of phage cocktails to control *A. baumannii* infection and contamination [204, 205].

1.2.1.4 Plasmids

Prokaryotic plasmids are autonomous extrachromosomal DNA molecules that are commonly circular and double stranded. Plasmids are of great utility to their bacterial hosts, as their flexible architecture can facilitate the augmentation of host metabolic functions and accumulation of genetic traits without altering the gene content of the host chromosome [206]. Conjugative plasmids are one example; a self-transmissible genetic entity that encodes all the genes required to transfer to another host through bacterial conjugation. Alternatively, plasmids can provide the basis for chromosomal

rearrangements through homologous recombination, resulting in the acquisition or loss of genetic content, and occasionally integration of the plasmid itself. Plasmids are also of continued interest for their use as cloning vectors, allowing the propagation and/or expression of foreign genetic material, with one such example being the classical plasmid vector pBR322 [207].

Plasmids in *A. baumannii* are not only further examples of MGEs harbouring resistance genes (Figure 1.5), but can also carry other MGEs, including ISs and integrons, into new hosts [82]. Some *A. baumannii* plasmids are also conjugative plasmids, facilitating the spread of these resistance traits and MGEs [80, 81, 84]. While not yet investigated in this organism, conjugative plasmids also have a disconcerting capability to also mobilise genomic islands into new hosts through conjugative transfer [208, 209].

Figure 1.5: Map of the conjugative drug resistance plasmid *Acinetobacter baumannii* plasmid pAB-G7-2. The open boxes represent the insertion sequence IS*Aba125* containing the aminoglycoside gene *aphA6* (grey). Open reading frames involved in conjugative transfer (*tra*) and DNA replication (*rep*) are shown, but those of unknown function are omitted. This figure was reproduced from [81].

1.2.2 Genomic islands

GEIs refer to any discrete gene cluster of varying genetic organisation and functionality with features suggestive of lateral gene transfer, such as atypical nucleotide content and the presence of mobility genes (e.g. integrases/recombinases and transposases) [210, 211]. They are often integrated within a tRNA or tmRNA locus [212] (Figure 1.6), but can also be integrated within protein coding genes [213]. GEIs can be dubbed pathogenicity, symbiosis, fitness, metabolic or resistance

islands, based on the predicted or characterised functions of cargo genes [214]. The highly flexible nature of the GEI definition allows a chromosomal region to be classified as one, even if it does not possess one or more of the abovementioned features. This definition also tentatively encompasses some MGE classes, including conjugative transposons, integrated plasmids, integrative conjugative elements, integrons and prophages [211] (Figure 1.7).

Recombinases encoded within GEIs fall into two classical protein superfamilies known as the tyrosine and serine recombinases, respectively named after the nucleophilic amino acid residue involved in chromosomal integration of DNA and excision of DNA as circularised intermediates [215]. Integration of GEIs often result in the formation of a pair of direct repeats [209, 216], and a replacement sequence for the target gene may occasionally be provided by the island itself, potentially restoring its function (Figure 1.6). Integron integrases and the vast majority of bacteriophage integrases are members of the tyrosine recombinases superfamily [171], which along with the integrases and recombinases present in various other GEIs covered in later chapters. Previously mentioned MGEs in *A. baumannii* (excluding ISs and plasmids) such as prophages, integrons and Tn*AbaR* are often referred to as GEIs. The presence of other GEIs in this organism have also been identified, be it through PCR detection of GEIs integrated within tRNA loci [217], or bioinformatic prediction [218].

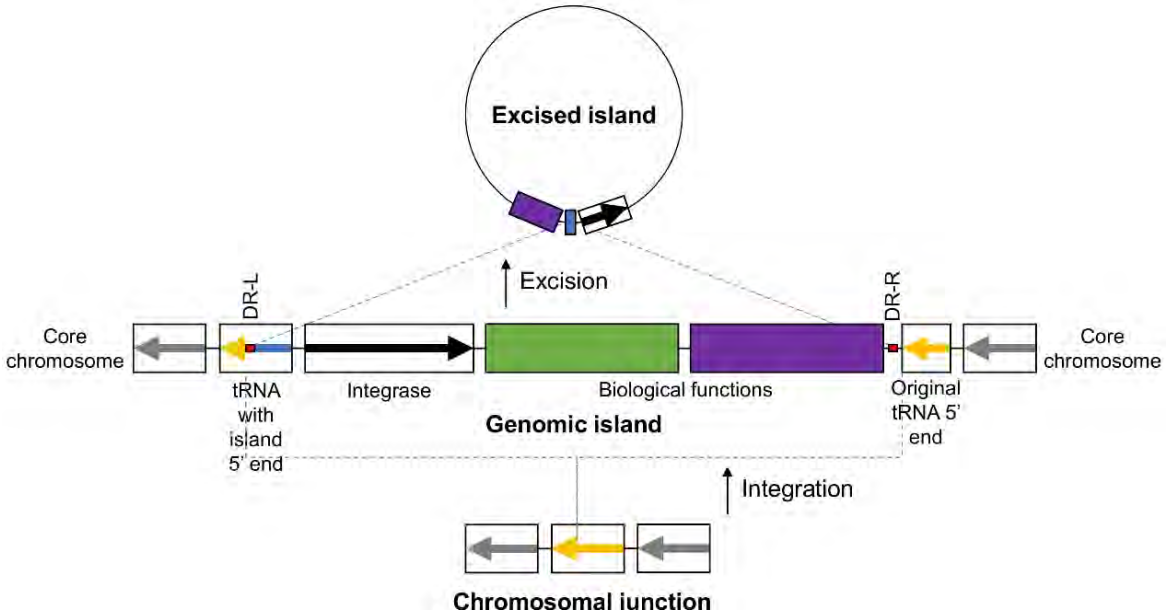


Figure 1.6: An archetypical genomic island (middle) integrated into a tRNA locus (yellow), consisting of an integrase gene (black) and numerous other genes of varying function (green and purple). Integration of the circularised genomic island (top) into the chromosome (bottom) is catalysed by the integrase, resulting in recombination between the two regions, and subsequent formation of a pair of direct repeats (DR-L/R) (red). In this example, the integration event results in partial replacement of the tRNA gene 5' end (blue) with sequence provided by the genomic island, and the former 5' end forming the rightmost island terminus. This figure is not drawn to scale and was adapted from [210, 211].

1.2.3 Regions of genomic plasticity

Regions of genomic plasticity (RGPs) are genomic segments that are variable, compared to related organisms, without any assumption about the evolutionary origin or genetic basis of these variable segments [170]. This broad terminology not only encompasses MGEs and GEIs, but also the plasticity within these elements, and variable regions resultant from genomic deletions or recombination events involving intragenic or extracellular DNA [170, 213] (Figure 1.7). Though originally coined in the course of *Pseudomonas aeruginosa* comparative genomics [170], the usage of this terminology has been extended to other bacterial organisms [213, 219], and can essentially be applied to any number of genomes belonging to any group of organisms. The RGP definition is the most far-reaching (Figure 1.7), while hierarchically, the MGE terminology is the most preferred when the properties of a laterally acquired region have been elucidated, including its mechanism of transfer, probable origin, element stability and site of integration [211].

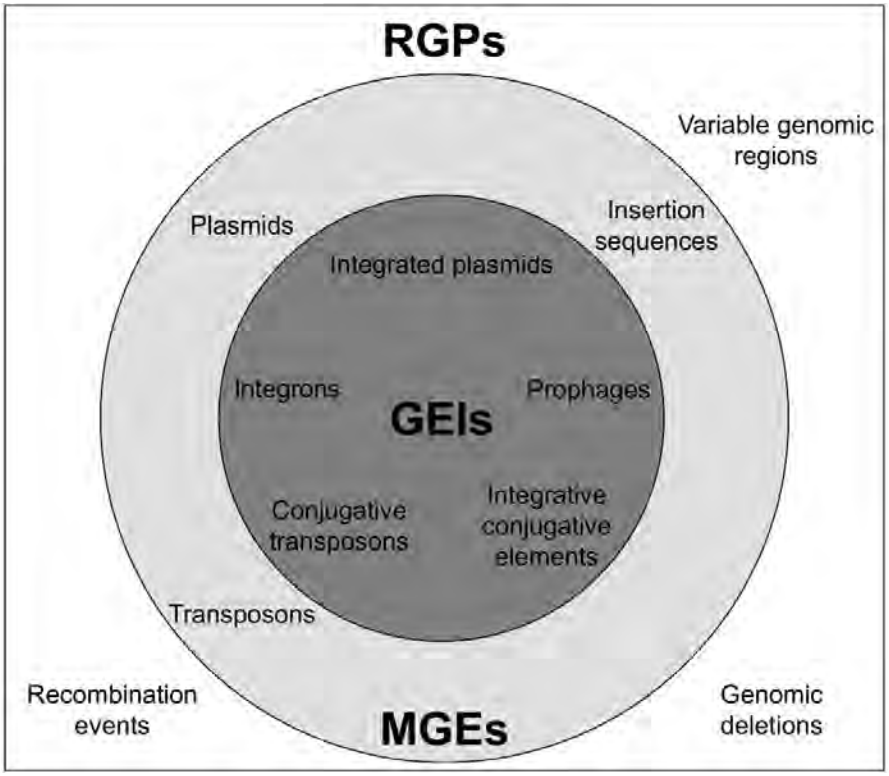


Figure 1.7: The breadth and interrelationship of accessory element classifications: mobile genetic elements (MGEs) (light grey), genomic islands (GEIs) (dark grey) and regions of genomic plasticity (RGPs) (white). "The broad definition of GEIs encompasses several MGEs, and RGPs encompass MGEs, GEIs and variable chromosomal regions not covered by the other two terminologies. Figure adapted from [211].

MGEs in *A. baumannii* are not only instrumental in being RGPs themselves, but also in influencing the creation of them. IS-mediated recombination is thought to be

responsible for considerable genome decay not only in *A. baumannii* SDF [160], but also several nosocomial isolates of *A. baumannii* [44], resulting in losses of genes involved in type VI secretion and glucarate/galactarate metabolism, amongst others. Other recombination events, involving intramolecular or exogenous genetic material, was found to be a significant cause of both loss and acquisition of genetic material in *A. baumannii* [41], including gene clusters putatively involved in iron acquisition and antimicrobial resistance, respectively. Other reported instances of *A. baumannii* RGPs arising from recombination include a gene cluster involved in polysaccharide biosynthesis [125] and the OMP CarO [52], involved in antigenic variation and carbapenem resistance, respectively.

1.3 The phenome

Analogous to the genome and proteome being the respective DNA and protein coding content of an organism, the phenome is the set of all phenotypes expressed by an organism, under the influences of genetic and environmental factors [220]. Phenotype microarrays are a widely utilised technology to assay the phenome of microorganisms.

1.3.1 Microbial phenotype microarrays

The Biolog™ Phenotype Microarray (PM) utilises a series of 96-well microtitre plates (Figure 1.8), with each well containing an individual lyophilised compound [221]. There are currently 1920 assays in the bacterial PM set, encompassing the metabolism of carbonaceous, nitrogenous, phosphorous and sulphurous compounds, biosynthetic pathways, osmotic, ionic and pH stressors, and sensitivity to various toxic and antimicrobial compounds [222]. A bacterial cell suspension, containing a tetrazolium redox dye, is added to each well of the PM plate. In the presence of electrons generated during cellular respiration and growth, the colourless tetrazolium compound is reduced to a purple formazan [221, 222] (Figure 1.8). This colour change is principally irreversible, resulting in signal amplification, and facilitating its ease of detection and quantitation via a charge-coupled device camera inside the OmniLog instrument [221] (Figure 1.8). This technology chiefly enables high-throughput phenotypic testing of various microorganisms, including members of the *Acinetobacter* genus, to generate a phenotypic profile [221]. PMs can also be used to phenotypically assay environmental communities [223], assay gene function of gene-knockouts [221], and to complement *in silico* metabolic modelling [224].

Figure 1.8: The process of phenotypic assaying with the Biolog Phenotype Microarray (PM) system. A 96 well microtitre plate (left), containing various lyophilised compounds, is inoculated with a suspension of cells and a colourless tetrazolium dye. In the presence of electrons generated during cellular respiration, a purple formazan is produced (left). These PM plates are incubated in the OmniLog (middle), a temperature controlled device that performs continuous automated readings. The amount of tetrazolium conversion to formazan is measured and outputted as a timescale kinetic series (right). In this example, phenotypes that are positive for only one strain or the other are indicated in either red or green, while phenotypes shared by both strains are in yellow. The 96 well plate pictured does not correspond with the generated phenotypic profile in this figure, and is for the purpose of demonstration only.

1.3.2 The phenome of *Acinetobacter*

The *Acinetobacter* genus is generally thought to be metabolically diverse, and PMs can go some way to demarcating this diversity. The first published use of PMs on the *Acinetobacter* genus was in 1995 [225], several years before the commercial release of this technology in 2001. In this study, PMs were used on the thirteen known *Acinetobacter* species at the time (including *A. baumannii*), in the hopes of being able to distinguish to the species level by phenotype, though this could not be satisfactorily achieved [225]. Some years later, PMs were utilised to complement *in silico* metabolic modelling derived from the complete genome sequence of the model *Acinetobacter*, *A. baylyi* ADP1 [226]. In addition to confirming aspects of the metabolic modelling, PMs enabled the addition of several carbon sources unaccounted by the modelling [226]. PMs were also used to assay *Acinetobacter* sp. organisms of unique niches, including diesel degrading isolates of *Acinetobacter venetianus* [227], and *Acinetobacter* sp. organisms (including *A. baumannii*) isolated from the tiger mosquito (*Aedes albopictus*) [228]. In both studies, there was some phenotypic variability observed in these isolates, despite originating from common niches.

It should be noted that the use of PMs in the abovementioned studies were limited to carbon source utilisation, and occasionally nitrogen utilisation, whereas PMs are capable of assaying numerous other phenotypes. In a recent study, all twenty PMs were utilised on four isolates of the *Acinetobacter* genus: *A. baumannii*, *A.*

nosocomialis, *A. pittii* and *A. calcoaceticus* [229]. Out of these four *Acinetobacter* species, *A. baumannii* could most effectively utilise peptide nitrogen sources and was the most tolerant to pH stress [229]. Prior to the work outlined in this thesis, efforts to examine the phenome across *A. baumannii* isolates have been limited, and have not kept pace with the release of new genome data.

1.4 Generation of gene knockouts and deletions in *A. baumannii*

Mutation of a gene is vital to characterising the function or phenotype bestowed by its translated protein. Gene mutation can be achieved through single-crossover or double-crossover recombination methods, resulting in either marked (antibiotic resistance cassette) or unmarked mutations [230]. These methods can be used to delete genes or chromosomal regions in *A. baumannii*, each with distinct advantages and disadvantages.

Gene disruption in *A. baumannii* can be achieved by two marked mutation methods, with the first entailing the electroporation of a non-replicative plasmid containing an internal fragment of a target gene, which integrates into the chromosome through a single-crossover recombination event [231] (Figure 1.9). Though simpler and more efficient than double-crossover recombination [232], single-crossover mutants are unstable in the absence of selection, as the plasmid can spontaneously excise through a second crossover event [231], resulting in reversion to the wild-type isogenic form.

The second method involves the creation of double-crossover recombination mutants, which are more stable than those generated through single-crossover recombination, as they result in replacement of the target site, rather than merely disrupting it [231, 232]. These mutants are generated with constructs consisting of an antibiotic resistance cassette, flanked by DNA homologous to the upstream and downstream regions of the gene to be deleted. These constructs are electroporated into an organism as either a linear PCR product [231] or a suicide plasmid [233] (Figure 1.9). Suicide plasmids encode a counterselection marker gene that is fatal to the host organism when expressed, effectively selecting against plasmid integration. Typically the levansucrase gene (*sacB*) of *Bacillus subtilis* [231] is employed as a counterselection marker, and its expression is fatal to most Gram-negative bacteria when cultured in the presence of sucrose; though the mechanism of this lethality is poorly understood [234].

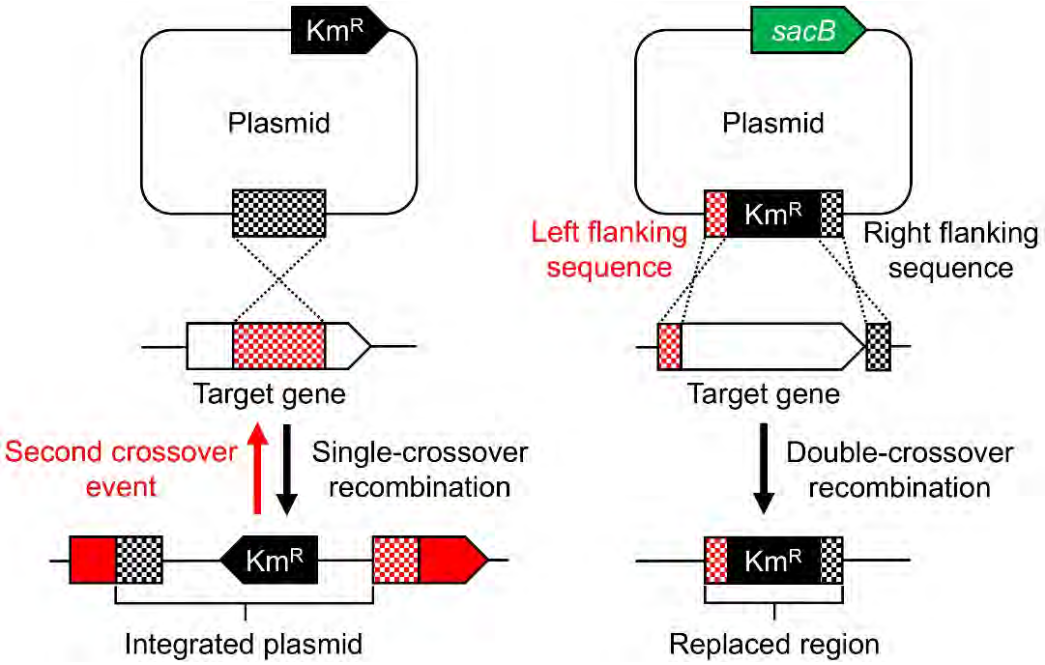


Figure 1.9: Schema of two types of marked mutation methods utilised to generate gene deletions in *Acinetobacter baumannii*. The method on the left involves the electroporation of a non-replicative plasmid containing an antibiotic resistance marker (Km^R) and a DNA fragment (black checkerboard) homologous to an internal region of the target gene (red checkerboard). Single-crossover recombination takes place and results in chromosomal integration of the whole plasmid. In the absence of selection this mutant is quite unstable, as the plasmid is able to spontaneously excise through a second recombination event. The method on the right also uses a non-replicative plasmid, containing a cloned resistance gene cassette (Km^R) (black) flanked by DNA homologous to regions flanking the target gene to be replaced (red and black checkerboards). Double-crossover recombination of the plasmid results in replacement of the target gene (white) with the resistance cassette. This plasmid encodes a counterselection marker (*sacB*) (green) to select against plasmid integration. Figures are adapted from [231] and [232], and are not drawn to scale.

A third mutagenesis method is available, which utilises the EZ-Tn*5*™ transposome (Epicentre) to generate a large pool of marked mutants, consisting of hundreds of isolates with a single transposon inserted into a random chromosomal loci. This approach has been utilised in *A. baumannii* [108, 159], though it not only suffers from a lack of precision, but also several disadvantages common to the above two methods.

The generation of marked mutations bear a number of disadvantages, with the first being the characteristically multiresistant nature of nosocomial *A. baumannii* limits the availability of antimicrobial resistance cassettes that can be utilised. Secondly, because electroporation is required to generate knockouts in these methods, there is the requirement for the organism to not only be recombinogenic but also be competent; deficiency in one or both limits the effectiveness of these methods [230]. Thirdly, these methods are empirically ineffective for deleting large genes or chromosomal regions, especially those containing repetitive regions [230]. Lastly, the introduction of a marked mutation can cause operon disruption or affect the

expression of neighbouring genes through influence of the resistance cassette promoter, causing polar effects [230].

Recently though, two publications have been released detailing the creation of unmarked mutations, which counter most of the above issues. The first article outlined the creation and application of a conjugative suicide vector (pMo130-Tel^R) encoding tellurite resistance (*telAB*) and a *xylE* reporter gene, causing colonies to turn yellow in the presence of pyrocatechol [235] (Figure 1.10). A gene fragment consisting of DNA upstream and downstream of the gene to be deleted is cloned into the plasmid. Following conjugative transfer, the plasmid undergoes an initial single-crossover recombination to integrate within the chromosome (Figure 1.10). A subsequent recombination event takes place, resulting in either reversion to the wildtype sequence, or an unmarked deletion of the target gene [235] (Figure 1.10). This approach was successfully used to disrupt two gene clusters encoding multidrug efflux proteins in a multidrug resistant isolate of *A. baumannii* [235]. While its effectiveness was demonstrated, it was apparent that this method could only delete a fairly small chromosomal region.

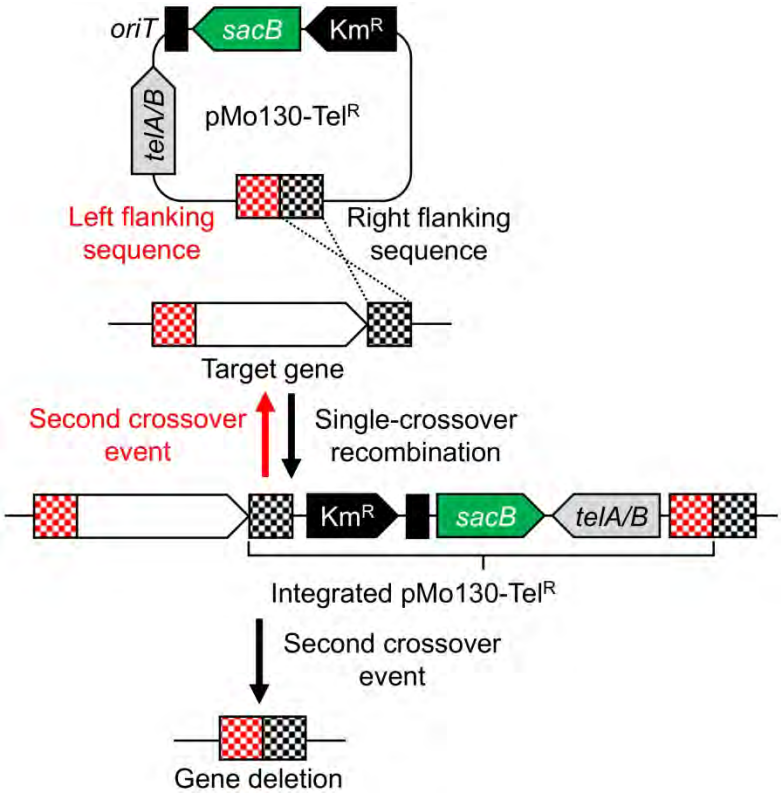


Figure 1.10: A strategy for the generation of unmarked mutations in multidrug resistant *Acinetobacter baumannii*. The suicide vector pMo130-Tel^R encodes tellurium (*telA/B*) (grey) and kanamycin resistance (Km^R) (black) genes, and is used to clone DNA homologous to regions flanking the target region to be deleted (red and black checkerboards). Following conjugative transfer of pMo130-Tel^R, single-crossover recombination occurs, resulting in integration of the plasmid into the host chromosome. A second crossover recombination takes place, resulting in either reversion to the wild type form or deletion of the target gene. Figure was adapted from [235] and not drawn to scale.

The second article outlined the construction of a pair of conjugative plasmid vectors, pJQFRT and pKFRT/FLP, each encoding a FLP recognition target (FRT) site. DNA homologous to upstream and downstream of the target chromosomal region are cloned into each respective plasmid. Following two separate conjugative transfers, each plasmid undergoes single-crossover recombination, leading to chromosomal integration of both plasmids, and sandwiching the target site between two FRT sites (Figure 1.11) [230]. Subsequent induction with anhydrotetracycline results in expression of the FLP recombinase of pKFRT/FLP, which binds the two FRT sites and catalyses excision of the entire sandwiched region, leaving only a remnant FRT site (Figure 1.11) [230]. This method was used to successfully delete a ~10 kb gene encoding a trimeric autotransporter adhesin in *Acinetobacter* sp. Tol-5, though regions in excess of 100 kb can theoretically be deleted in various Gram-negative organism [230]. The use of aminoglycoside resistance cassettes in the two plasmids would be expected to limit the use of this system to antimicrobial susceptible isolates of *A. baumannii*.

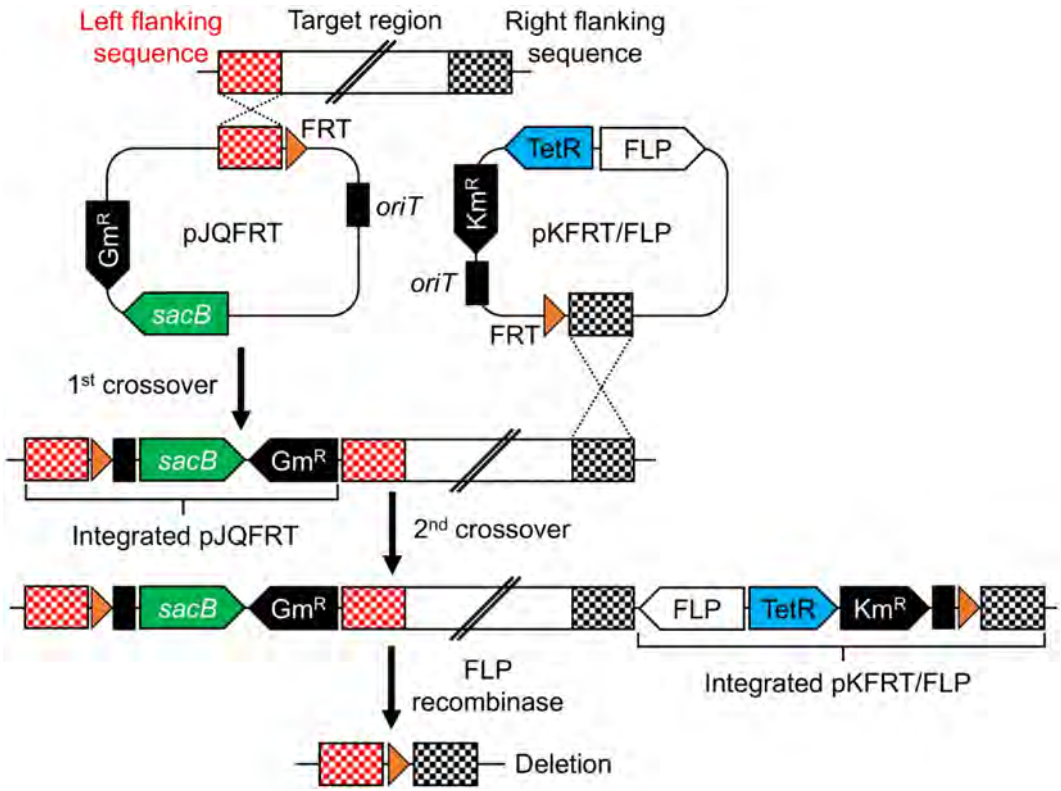


Figure 1.11: A strategy for the generation of unmarked mutations using FLP recombination. The plasmids pJQFRT and pKFRT/FLP are cloned with DNA homologous to the upstream and downstream regions of the target site, respectively. These two plasmids are consecutively transferred into the host through conjugation, and eventually undergo two separate single-crossover recombinations, resulting in chromosomal integration of both plasmids and sandwiching the target region with two FLP recognition target (FRT) sites (orange). The FLP recombinase (white), under control of the TetR repressor (blue), is induced with anhydrotetracycline, and catalyses excision of the DNA between the two FRT sites, leaving only a remnant FRT site. The figure was adapted from [230] and is not drawn to scale.

Upon successful generation of any mutation, complementation is essential for confirmation of gene function. In bacteria, this is usually achieved through cloning and expression of the wild type gene on a plasmid vector. Complementation in *A. baumannii* is currently limited to the use of the shuttle vector pWH1266, originally constructed from a cryptic *A. baylyi* BD413 (formerly *A. calcoaceticus*) plasmid and pBR322 [236]. pWH1266 encodes resistance to tetracycline and ampicillin, and is capable of replication in both *E. coli* and *A. baylyi* [236]. The apparent lack of suitable complementation plasmids further restricts the ability to generate mutations in antimicrobial resistant isolates of *A. baumannii*, which would especially be a problem with mutants generated with pMo130-Tel^R. Construction of unmarked mutations using FLP recombination would be expected to also be hampered by the lack of suitable complementation vectors, as large chromosomal deletions could not feasibly be restored.

1.5 Scope of thesis

Acinetobacter baumannii is an opportunistic pathogen that represents a significant public health threat due to the emergence of multidrug resistant strains. Due to its recent emergence, its virulence and persistence capabilities remain relatively poorly characterised compared with other pathogenic bacteria, including members of the ESKAPE group.

A. baumannii is also known to exist outside the clinical milieu within a number of natural, industrial and anthropogenic niches, including soil, water, arthropods and chemical spills, as well as being a cause of fulminant and fatal CA-AB infections. However, the current knowledge concerning *A. baumannii* does not reflect the diversity of this organism, with nosocomial *A. baumannii* being the primary focus of research. Information regarding XN-AB is limited to some epidemiological and phenotypic studies, and the genome sequence of the body louse isolate, *A. baumannii* SDF.

We sought to address the lack of knowledge concerning XN-AB through WGS of various isolates, including those originating from environmental sources, industrial sources and Asia-Pacific CA-AB infections. In the case of CA-AB, it was anticipated that it would be possible to determine the genetic basis of its fulminant nature, as well as its susceptibility to antimicrobials. In regards to other XN-AB isolates, it was

expected that these isolates would also be susceptible to antimicrobials, but possess a number of novel metabolic genes. Taken together, the genomes of these XN-AB isolates would be bioinformatically compared to previously sequenced nosocomial isolates of *A. baumannii* to identify potential differences regarding virulence, antimicrobial resistance and metabolic capabilities, and determine what proportion of these differences can be attributed to the acquisition of MGEs and GEIs.

Chapter 2: The complete genome and phenome of a community-acquired *Acinetobacter baumannii*

2.1 Non-technical summary

Investigations into the genome sequences of *Acinetobacter baumannii* have primarily focussed on multidrug resistant nosocomial isolates. Consequently, little is known genetically about pathogenic variants outside of the hospital environment. Thus, the genome of a community-acquired *A. baumannii* (strain D1279779), isolated from a bacteraemic infection of an Indigenous Australian, was sequenced. All gaps in the genome sequence were closed iteratively by numerous rounds of PCR and dideoxysequencing, resulting in a complete circular chromosomal sequence. Furthermore, the catabolic and antimicrobial resistance phenotypic profile of D1279779 and several previously sequenced isolates of *A. baumannii* were determined through phenotype microarrays. Both the genome and phenome of D1279779 were compared to several previously sequenced nosocomial isolates to better understand the defining characteristics of nosocomial and community-acquired *A. baumannii* isolates. It was found that D1279779 possessed broad catabolic capabilities, whilst being susceptible to antimicrobials. The opposite relationship was observed for multidrug resistant nosocomial isolates.

OPEN ACCESS Freely available online

PLOS ONE

The Complete Genome and Phenome of a Community-Acquired *Acinetobacter baumannii*

Daniel N. Farrugia[1], Liam D. H. Elbourne[1], Karl A. Hassan[1], Bart A. Eijkelkamp[2], Sasha G. Tetu[1], Melissa H. Brown[2], Bhumika S. Shah[1], Anton Y. Peleg[3,4], Bridget C. Mabbutt[1], Ian T. Paulsen[1]*

1 Department of Chemistry and Biomolecular Sciences, Macquarie University, Sydney, New South Wales, Australia, **2** School of Biological Sciences, Flinders University, Adelaide, South Australia, Australia, **3** Department of Microbiology, Monash University, Clayton, Victoria, Australia, **4** Department of Infectious Diseases, The Alfred Hospital, Melbourne, Victoria, Australia

Abstract

Many sequenced strains of *Acinetobacter baumannii* are established nosocomial pathogens capable of resistance to multiple antimicrobials. Community-acquired *A. baumannii* in contrast, comprise a minor proportion of all *A. baumannii* infections and are highly susceptible to antimicrobial treatment. However, these infections also present acute clinical manifestations associated with high reported rates of mortality. We report the complete 3.70 Mbp genome of *A. baumannii* D1279779, previously isolated from the bacteraemic infection of an Indigenous Australian; this strain represents the first community-acquired *A. baumannii* to be sequenced. Comparative analysis of currently published *A. baumannii* genomes identified twenty-four accessory gene clusters present in D1279779. These accessory elements were predicted to encode a range of functions including polysaccharide biosynthesis, type I DNA restriction-modification, and the metabolism of novel carbonaceous and nitrogenous compounds. Conversely, twenty genomic regions present in previously sequenced *A. baumannii* strains were absent in D1279779, including gene clusters involved in the catabolism of 4-hydroxybenzoate and glucarate, and the *A. baumannii* antibiotic resistance island, known to bestow resistance to multiple antimicrobials in nosocomial strains. Phenomic analysis utilising the Biolog Phenotype Microarray system indicated that *A. baumannii* D1279779 can utilise a broader range of carbon and nitrogen sources than international clone I and clone II nosocomial isolates. However, D1279779 was more sensitive to antimicrobial compounds, particularly beta-lactams, tetracyclines and sulphonamides. The combined genomic and phenomic analyses have provided insight into the features distinguishing *A. baumannii* isolated from community-acquired and nosocomial infections.

Citation: Farrugia DN, Elbourne LDH, Hassan KA, Eijkelkamp BA, Tetu SG, et al. (2013) The Complete Genome and Phenome of a Community-Acquired *Acinetobacter baumannii*. PLoS ONE 8(3): e58628. doi:10.1371/journal.pone.0058628

Editor: Valerie de Crécy-Lagard, University of Florida, United States of America

Received November 6, 2012; **Accepted** February 5, 2013; **Published** March 19, 2013

Funding: This work was supported in part by the National Health and Medical Research Council of Australia Grant No. 535053 awarded to MHB and ITP, and by Macquarie University. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: ian.paulsen@mq.edu.au

Introduction

Acinetobacter baumannii is a significant nosocomial pathogen [1], known for its high intrinsic and laterally acquired resistance to antimicrobials [2,3] as well as its persistence on various abiotic surfaces [4-6]. The complete genome sequences of ten *A. baumannii* strains have been determined to date: 1656-2 [7], AB0057 [8], AB307-0294 [8], ACICU [9], ATCC 17978 [10], AYE [11], MDR-TJ [12], MDR-ZJ06 [13], SDF [11] and TCDC-AB0715 [14]. Nine of these are nosocomial isolates, whereas *A. baumannii* SDF was isolated from a human body louse [11]. These genome sequences have demonstrated extensive divergence due to the acquisition and accretion of various mobile genetic elements, particularly those contributing to antimicrobial resistance [15]. One mobile element of clinical import is the *A. baumannii* antibiotic resistance island (AbaR), that encodes resistance to a multitude of antibiotics and heavy metals [15].

Research regarding *A. baumannii* has occurred primarily within the context of the clinical milieu, with little known about potential environmental reservoirs of this organism. Several non-nosocomial niches of *A. baumannii* have been identified, including human lice [16,17], hydrocarbon contaminated soils [18,19], the plant rhizosphere [20,21] and estuaries [22,23]. *A. baumannii* is also known to exist outside the hospital environment as a commensal of the skin [24] and nasopharynx [25] of humans. This organism is also a public health issue outside of the hospital setting, in the form of community-acquired *A. baumannii* (CA-AB) infections.

Infections caused by CA-AB are clinically and epidemiologically distinct from their nosocomial counterparts [26]. CA-AB infections are uncommon and highly fatal, comprising less than 10% of all *A. baumannii* infections [27,28] but resulting in mortalities ranging from 30–62% [27–30]. These infections are also antimicrobial susceptible [26,30] and present a more acute clinical manifestation [30], but are thought not to be reservoirs of nosocomial outbreaks [26].

The majority of CA-AB infections occur in individuals with underlying comorbidities, who reside in tropical and subtropical climates [28]. Incidences of CA-AB infection have been reported within various regions of the Asia Pacific such as Taiwan [31], Hong-Kong [30], Singapore [32], Korea [33] and Australia [34]. To a lesser extent, CA-AB infections have also been observed in non-tropical regions [27] and in otherwise healthy children and adults [35–39]. Indigenous Australians in the Northern Territory

are overrepresented relative to the general population in rates of community-acquired bacteraemic pneumonia caused by *A. baumannii* and other pathogens [29,34]. This disparity has been attributed to the interaction of both monsoonal climate and a high prevalence of comorbidities in the indigenous Australian population including alcoholism, diabetes mellitus, chronic obstructive pulmonary disease and cigarette smoking [25,34,40].

To explore the underlying basis of epidemiological and phenotypic differences between nosocomial and community-acquired strains of *A. baumannii*, we determined the complete genome sequence of the CA-AB isolate D1279779 and phenotypically profiled this strain using phenotype microarrays. The genome and phenome of D1279779 was subsequently compared to completely sequenced nosocomial *A. baumannii* strains.

Materials and Methods

Bacterial strains, culture conditions and genomic DNA extraction

The *A. baumannii* strain D1279779 was kindly provided by the Menzies School of Health Research (Darwin, Australia). This strain was isolated in 2009 from a bacteraemic infection of an indigenous Australian male at the Royal Darwin Hospital, where the identity and antimicrobial susceptibilities of this isolate were previously determined with a VITEK 2 System (bioMérieux). *A. baumannii* D1279779 was cultured in lysogeny broth (LB) and lysogeny broth agar (LBA) (both without glucose) at 37 °C. Genomic DNA was extracted with the Wizard Genomic DNA Purification Kit (Promega) from 1 mL of overnight culture as per the manufacturer's protocol.

DNA sequencing, genome assembly and annotation

A. baumannii D1279779 genomic DNA was prepared and sequenced by 454 FLX pyrosequencing (Roche Diagnostics) at the Ramaciotti Centre for Gene Function Analysis (University of New South Wales, Sydney). The sequences reads were assembled *de novo* with MIRA [41] using the default parameters. The ninety-four contiguous sequences of D1279779 were reordered relative to the ten currently complete genomes of *A. baumannii* using MAUVE [42]. The highest level of synteny was observed with the *A. baumannii* ACICU genome [9], which was subsequently utilised as a reference in Projector2 [43] to design oligonucleotides for gap closure by PCR and sequencing. Amplicons for gap closure were generated using AccuPrime *Pfx* Mastermix (Invitrogen) or GoTaq DNA Polymerase (Promega) as per the manufacturer's protocol, with variations in the annealing temperature (45 °C to 65 °C) and extension time (2 to 12 min) according to estimated gap sizes. The resultant amplicons were purified using the QIAquick PCR Purification Kit (Qiagen) and cloned into pGEM-T Easy (Promega), or directly sequenced using bidirectional dideoxysequencing performed by the Macquarie University DNA Analysis Facility (Sydney, Australia). The resultant chromatograms were edited and assembled in ChromasPro (Technelysium Pty. Ltd.) and crosschecked against the D1279779 genome by use of the BLASTN [44] application integrated into BioEdit [45]. The presence and directionality of the six rRNA operons in the genome was confirmed by amplification of the gap and sequencing of small junctions flanking this region. The genome assembly was finalised with the aid of CLC Sequence Viewer 6 (CLCbio). Genome annotation was conducted with the RAST automated annotation engine [46] and manual curation was performed with the aid of Artemis [47]. UGENE was routinely utilised for genome browsing and analysis [48]. The nucleotide sequences of the D1279779 chromosome and the plasmid pD1279779 have been deposited into GenBank with the accession numbers CP003967 and CP003968 respectively.

Comparative genomics and accessory element identification

A. baumannii D1279779 was compared to the ten published *A. baumannii* genomes by a reciprocal BLASTP [49] search to identify putative orthologs at an e-value cutoff of 10^{-5}. Trinucleotide composition of the DNA sequence was computed by χ^2 analysis using a 2000 bp sliding window with a 1000 bp overlap [50]; regions containing χ^2 values >500 were suggestive of atypical trinucleotide composition. Evidence derived from BLASTP and χ^2 analysis was used to identify regions of genomic plasticity (RGPs) [51,52], defined as either any putative mobile genetic element or contiguous cluster of genes present in D1279779 and three or less strains. Genes clusters absent in D1279779 but present in more than three other strains were also considered RGPs. The identity of insertion sequences present in the D1279779 genome were elucidated with BLASTP searches conducted in ISFinder [53].

The phylogenetic relationship of D1279779 to sequenced *A. baumannii* strains was inferred with a bootstrapped neighbour-joining analysis in MEGA5 [54] based on the concatenated nucleotide sequences of six of the seven reference genes utilised in the multilocus sequence typing (MLST) of *A. baumannii* [55]. The *fusA* gene was excluded from this analysis as it displayed atypical trinucleotide composition in D1279779, suggesting potential lateral acquisition of this gene.

Phenotype microarray testing

The phenomes of *A. baumannii* D1279779 and three nosocomial strains, ACICU, ATCC 17978 and AYE, were assayed with the Biolog Phenotype MicroArray™ (PM) system [56] to identify compounds that could serve as sole carbon (PM1-2; 190 compounds) or nitrogen sources (PM3; 95 compounds). Additionally, sensitivities to stress conditions (PM9-10; 192 conditions) and various antimicrobials compounds (PM11-20; 240 antimicrobials) were also investigated. All phenotypic tests were performed as per the manufacturer's protocol, except cryogenic stocks of *A. baumannii* were streaked onto either LBA medium (PM1-2 and PM9-20) or Reasoner's 2A agar (Difco) (PM3). The bacterial suspension for nitrogen source testing was supplemented with D-xylose as the sole carbon source at a concentration of 20 mM. Following inoculation, all PM plates were incubated in an OmniLog reader (Biolog) aerobically at 37 °C for 48 h. Reduction of the tetrazolium-based dye (colourless) to formazan (violet) was monitored and recorded at 15 min intervals by an integrated charge-coupled device camera. The resultant data were analysed with the supplied manufacturer's software, resulting in a time-course curve for colorimetric change equating to respiration rate. The phenotypes were classified on the basis of the maximal curve height; a phenotype was considered positive if the height was greater than 115 and 101 OmniLog units for nitrogen sources and all other phenotypes, respectively. Data that exceeded these cutoff values as the result of colouration from certain compounds was excluded from analysis. Observed phenotypic differences between the strains were linked to differences in genotype through a combined analysis of the EcoCyc [57], MetaCyc [58], and KEGG [59] metabolic databases and additional literature searching.

Independent confirmatory testing of phenotype microarray data

Five millilitres of M9 minimal media (Sigma-Aldrich), supplemented with varied carbon source compounds (20 mM), was

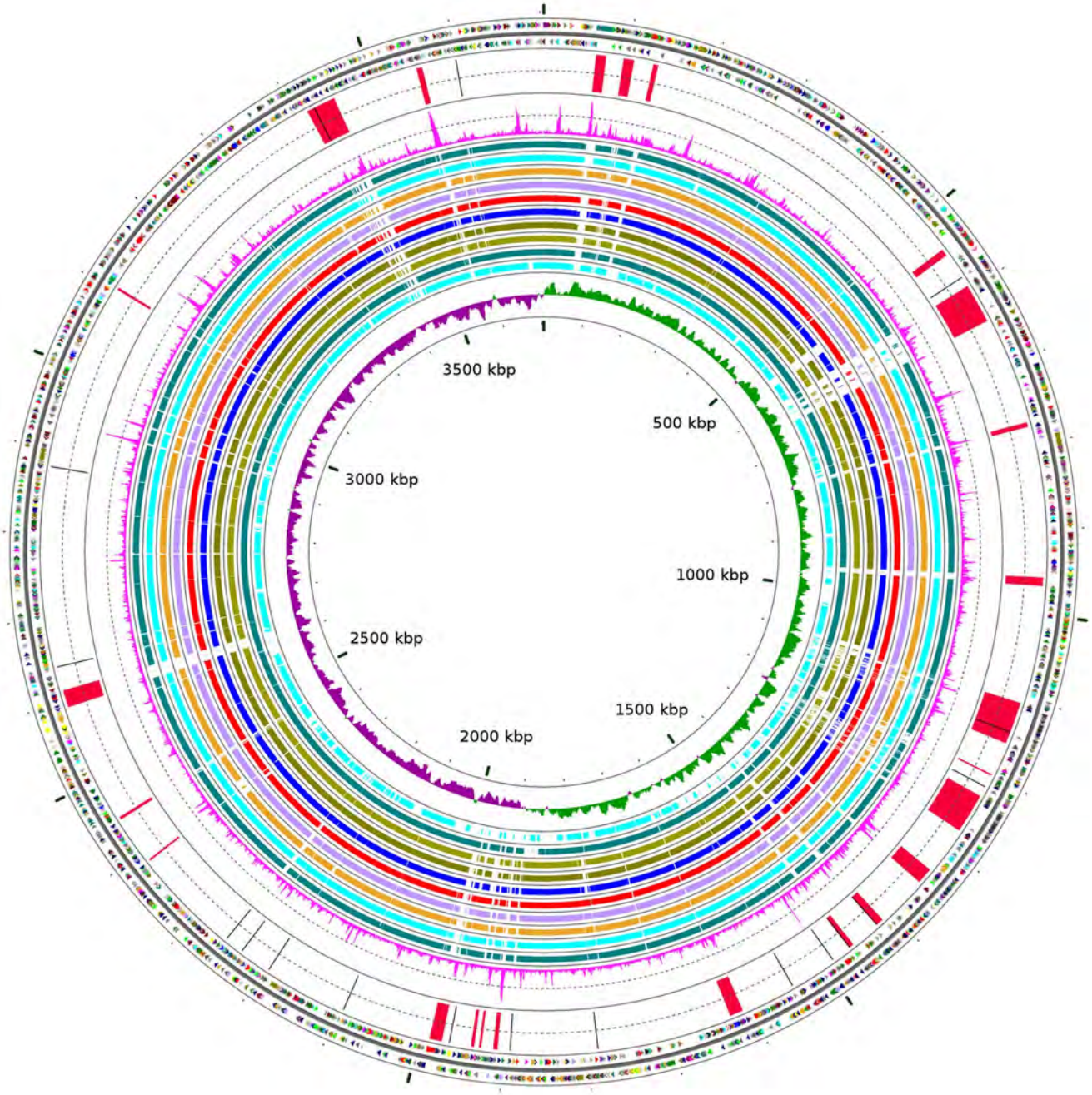


Figure 1. Genome map of *Acinetobacter baumannii* D1279779. The two outermost circles denote positions of protein coding sequences (CDSs) on the positive (circle 1) and negative (circle 2) strands coloured according to clusters of orthologous groups (COGs) [99] functional category: A (lavender), B (apricot), C (olive), D (light brown), E (dark green), F (electric pink), G (electric green), H (peach), I (red), J (dark red), K (midnight blue), L (plum), M (teal), N (blue), O (aquamarine), P (orange), Q (yellow), R (dark grey), S (grey), T (light purple), U (light green), V (light yellow), and unknown COG (black). Circle 3 represents positions of identified regions of genomic plasticity (red) and ISAb13s (black) ordered clockwise from the origin of chromosomal replication as outlined in Table 2 and Table 4, respectively. Circle 4 denotes the calculated chi-squared values based on the trinucleotide composition of the DNA sequence. Circles 5-14 show DNA conservation between D1279779 and other sequenced *A. baumannii* strains based on pairwise BLASTN alignments (evalue threshold 1e-10). Strain comparisons (outermost to innermost): 1656-2 (teal), ACICU (aquamarine), MDR-ZJ06 (orange), MDR-TJ (light purple), TCDC-AB0715 (red), AB307-0294 (blue), AB0057 (olive), AYE (green-brown), ATCC 17978 (teal) and SDF (aquamarine). The innermost circle denotes positive (green) and negative (purple) GC-skew and the scale in kilobase pairs. The CGView software [100] was utilised to construct the genome map.
doi:10.1371/journal.pone.0058628.g001

inoculated with a single colony of *A. baumannii* D1279779 or ACICU previously streaked on LBA medium. Cultures were incubated with shaking at 37 °C for 24 h, with the observation of turbidity from cellular replication deemed to signify a positive phenotype. A minimum of three temporally distinct replicates were performed for each tested compound, in addition to substrate and inoculation negative controls. The carbon sources tested were: L-arabinose, bromosuccinic acid, (±)-carnitine hydrochloride, L-

Table 1. Comparative genome features of *Acinetobacter baumannii* D1279779.

Strain	D1279779	ATCC 17978	ACICU	AYE
Size (base pairs)	3704285	3976747	3904116	3936291
Plasmids	1	2	2	4
G+C content (%)	39.00	38.94	39.03	39.40
Protein-coding sequences (CDSs)	3388	3787	3670	3607
Insertion sequences	18	14	14	33
Average gene length	935	888	929	951
Coding regions (%)	85.60	84.50	84.87	86.13
rRNA operons	6	5	6	6
tRNAs	65	69	64	72

doi:10.1371/journal.pone.0058628.t001

carnitine hydrochloride, disodium fumarate, α-D-glucose, L-histidine monohydrochloride monohydrate, polysorbate 80, putrescine dihydrochloride, L-pyroglutamic acid, quinic acid, sodium 4-hydroxybenzoate, sodium acetate trihydrate, trisodium citrate dihydrate, sodium D-gluconate, and D-xylose. All chemicals were sourced from Sigma-Aldrich (purity ≥96%), dissolved in sterile distilled water and filter-sterilised. Phenotypic testing of α-D-glucose and D-gluconic acid utilisation was further conducted in inoculation fluid zero (IF-0) [56] (without sodium pyruvate and tetrazolium dye). Further experiments with α-D-glucose were supplemented with 10 μM methoxatin disodium salt (pyrroloquinoline quinone) with the corresponding negative control.

Results and Discussion

Genomic features

The complete genome of *A. baumannii* D1279779 was determined and found to consist of a 3704285 bp circular chromosome and a plasmid of 7416 bp, dubbed pD1279779. A total of 3479 genes were annotated on the chromosome including 65 tRNAs, 6 rRNA operons and 3388 predicted protein coding sequences (CDS) (Table 1, Figure 1), which included 1019 annotated CDS (30%) predicted to encode hypothetical proteins. The plasmid pD1279779, unlike many previously sequenced *A. baumannii* plasmids, does not encode any insertion sequences or genes involved in antimicrobial resistance [60]. Plasmid pD1279779 appears to be of mosaic origin, with a replication *repA* gene sharing 100% nucleotide sequence identity with the *A. baumannii* plasmid p203 [60] and a 2609 bp segment sharing 99% nucleotide identity with a region from the otherwise unrelated organic peroxide resistance plasmid pMAC of *A. baumannii* ATCC 19606 [61].

The phylogeny and synteny of *A. baumannii*

The phylogeny of *A. baumannii* D1279779, with respect to other sequenced *A. baumannii* isolates, was inferred using a MLST approach [55]. The allelic profile of this strain is 12-37-2-2-3-2-14, but it does not match any previously assigned STs; the closest allelic profile in the MLST database belonged to an ST117 isolate (12-37-2-2-9-2-14), differing only in the *recA* allele. This MLST-based analysis suggested that the nearest phylogenetic relatives of D1279779 were *A. baumannii* strains of the international clonal (IC) lineage II, though this strain did not fall within either the *A. baumannii* ICI or ICII lineages (Figure 2). This observation was congruent with our own previous PCR typing, which indicated this strain did not belong to any of the three major *A. baumannii* lineages [62]. The phylogenetic relationship of D1279779 to other *A. baumannii* strains was also consistent with the notion that

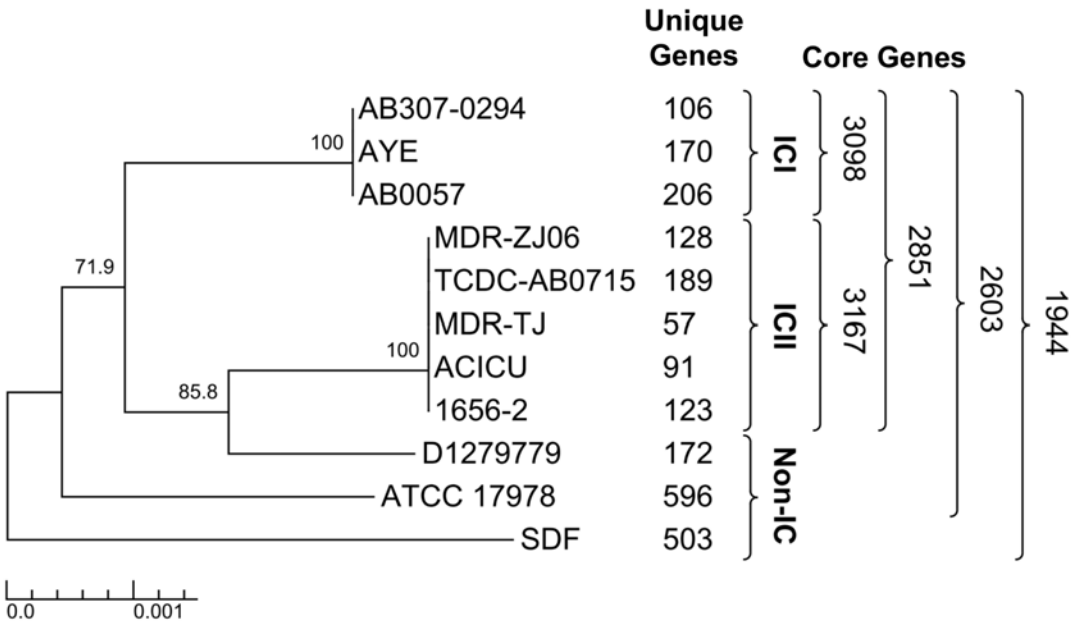


Figure 2. Phylogenetic lineage of *Acinetobacter baumannii*. The phylogenetic relationship of all completely sequenced *A. baumannii* strains was inferred by the neighbour-joining method conducted in MEGA5 [54] using concatenated nucleotide sequences of six reference genes based on the multilocus sequence tag scheme of *A. baumannii* (*cpn60*, *gltA*, *pyrG*, *recA*, *rplB* and *rpoB*). Numerals adjacent to strain names represent the number of unique genes (determined by BLASTP) and the numerals adjacent to the brackets represent the size of the core genome of *A. baumannii* or within clonal lineages. The interior values are bootstrap probabilities based on 1000 replicates, and the tree was drawn using TreeGraph2 [101].
doi:10.1371/journal.pone.0058628.g002

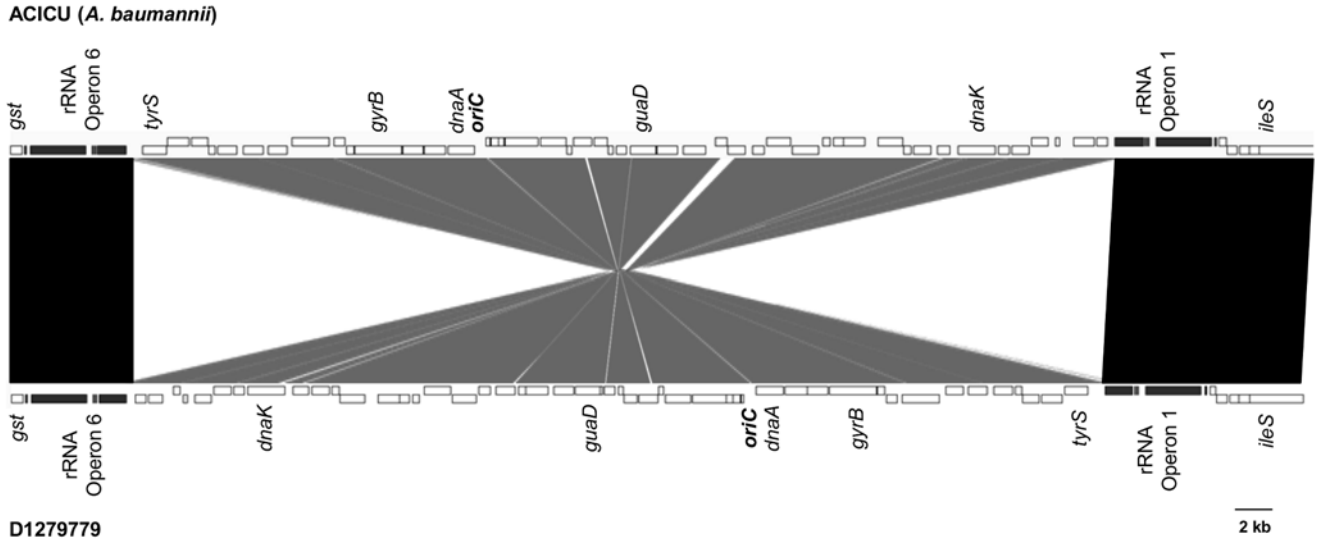


Figure 3. 50kb inversion in the genome of *Acinetobacter baumannii* D1279779. The DNA region between rRNA operons one and six in *A. baumannii* D1279779 is conserved (≥96% nucleotide identity) but is inverted relative to other *A. baumannii* genomes (represented by ACICU), as depicted by the grey shading. Genes on positive and negative strands are depicted on the top and bottom row of rectangles, respectively. Conserved regions (≥99% nucleotide identity) in the same orientation are depicted by black shading. Locations of several conserved genes and the origin of chromosomal replication (*oriC*) are indicated. This figure was generated using the combined outputs of MAUVE [42] and the Artemis Comparison Tool [102].
doi:10.1371/journal.pone.0058628.g003

community-acquired isolates are epidemiologically distinct from nosocomial isolates [26].

Genomic alignments using MAUVE indicated that *A. baumannii* D1279779 shares a high degree of genome synteny with other completely sequenced strains of *A. baumannii*, with the exception of *A. baumannii* SDF (Figure S1), which is known to have undergone both extensive genome reduction and rearrangement [11]. In *A. baumannii* D1279779, a 50.8 kb region of sequence situated between the first and sixth rRNA operons was inverted relative to other *A. baumannii* genomes (Figure 3). We confirmed this rearrangement in D1279779 by PCR analysis conducted on the original cryogenic stock. The inversion of this region resulted in reversing the orientation of a number of critical housekeeping genes as well as the origin of chromosomal replication (*oriC*) (Figure 3). It is of note that the opposing directionality of both rRNA operons flanking the inversion was maintained, which would otherwise potentially result in replicative blockage [63]. Presumably this genomic rearrangement was mediated by homologous recombination between the two oppositely oriented rRNA operons [64]. Such rRNA operon-mediated rearrangements in genome architecture have been known to occur in other organisms and between more distal rRNA operons [64,65], resulting in even larger inversions than observed here.

The core and accessory genome of *A. baumannii*

The predicted proteome for all currently complete genomes of *A. baumannii* was compared by means of a reciprocal BLASTP search, which enabled estimation of the *A. baumannii* core genome size. The numbers overlaid on the phylogenetic tree in Figure 2 (adjacent to the brackets) indicate the various sizes of the core genome. A total of 1944 predicted CDS are shared between all *A. baumannii* genomes in this analysis; this increases to 2603 CDS when strain SDF is excluded from this analysis group (Figure 2). The lineage-specific core genomes of ICI and ICII is larger than the *A. baumannii* core genome, with 3098 and 3167 CDS shared respectively (Figure 2), supporting the notion that ICI and ICII constitute recent clonal lineages [55]. *A. baumannii* strains ATCC 17978 and SDF were the two deepest branching strains in the phylogenetic tree, each displaying over 500 strain-specific genes (Figure 2). *A. baumannii* D1279779 encoded 172 unique genes not present in the other sequenced strains. The majority of these genes were associated with features indicative of recent lateral acquisition, such as atypical trinucleotide composition and the presence of mobile genetic elements.

Eighteen copies of the insertion sequence IS*Aba13* and twenty-four regions of genomic plasticity (RGPs) were identified in the genome of D1279779, including two prophages and four genomic islands (Table 2). Unlike RGPs located in genomes of nosocomial isolates, those in D1279779 do not appear to encode known antibiotic resistance or virulence-associated functions (Table 2). One RGP of interest, D1279779_RGP05, contains genes encoding a type I DNA restriction-modification system and genes associated with the catabolic degradation of nitrogenous compounds. Analysis indicated the gene encoding the specificity subunit (*hsdS*) of the restriction-modification system contained a frameshift mutation. However, a truncated *hsdS* can still potentially translate and dimerize into a functional HsdS subunit, albeit with an altered DNA specificity [66-68]. Nevertheless, the presence of a DNA restriction-modification system in this genomic island may act as a 'selfish' genetic element to ensure its dissemination [69] and/or may function in bacteriophage defence and hindrance of lateral gene transfer [70].

Catabolism-associated genes in this RGP included a gene cluster encoding a glutamate dehydrogenase and a near-complete arginine succinyltransferase (AST) pathway [71]. Complete and partial copies of the AST pathway are encoded elsewhere in the core genome of *A. baumannii* D1279779. Also present in D1279779_RGP05 are genes encoding an acetylpolyamine amidohydrolase (AphA) and an allantoate amidohydrolase (AllC). AphA performs the deacetylation of varied acetylpolyamines [72], and AllC functions in the degradation of allantonate to S-ureidoglycolate [73]. The core genome of *A. baumannii* D1279779

Table 2. Accessory elements present in the *Acinetobacter baumannii* D1279779 genome.

(D1279779_)	CDS (ABD1_)	Size (kb)	G+C (%)	Putative function/features of interest	Integrase (target)	ATC[a]	*A. baumannii* orthologues
RGP01	00530–00630	12.1	31.3	polysaccharide biosynthesis	–	Y	ABTJ_03749–03760
RGP02	00790–00890	12.5	39.0	unknown	–	N	variable region
RGP03	01090–01140	5.9	36.8	metabolism	–	N	ABTJ_03697–03701
RGP04	04780–04960	16.6	41.2	fatty acid biosynthesis	–	N	ACICU_00517– 00535
RGP05	05270–05690	47.3	37.3	genomic island, type I restriction-modification, metabolic augmentation	Y (*dusA*)	Y	unique
RGP06	06860–07920	5.7	30.4	unknown	–	Y	unique
RGP07	08570–08650	10.0	33.8	cryptic genomic island	Y (ND[b])	N	unique
RGP08	09940–10530	50.6	38.4	prophage, phosphoethanolamine transferase, IS*Aba13*_2	Y (ND[b])	Y	$pA_{ICU}9$[c] All ex. AB0057, ATCC 17978, SDF
RGP09	10860–10930	1.7	37.5	lipoproteins	–	N	unique
RGP10	11180–11850	48.0	38.9	prophage, DNA polymerase V	Y (*lysC*)	N	A1S_1142-75 AB57_1310-1224
RGP11	12360–12570	19.3	32.9	temperature shock, metabolism	N	Y	All ex. SDF
RGP12	13160–13230	9.4	37.0	unknown, IS*Aba13*_4	–	N	unique
RGP13	13580–13650	6.2	32.9	unknown, IS*Aba13*_5	–	N	unique
RGP14	15000–15150	16.0	33.7	genomic island, fimbriae biogenesis	Y* (ND[b])	Y	$pA_{ICU}16$[c] In all *A. baumannii*
RGP15	17780–17800	5.0	25.9	unknown	–	Y	unique
RGP16	17930–17970	3.0	32.1	unknown	–	N	unique
RGP17	18030–18070	3.6	40.5	degradative enzymes	–	N	unique
RGP18	18340–18480	14.9	37.1	fatty acid metabolism	–	Y	A1S_1813-22
RGP19	22050–22070	2.1	33.9	unknown	–	N	A1S_2209-11
RGP20	22680–22710	3.4	38.6	RNA modification	–	N	A1S_2271-73
RGP21	23900–24060	20.7	38.7	degradative enzymes	–	N	A1S_2397-2414
RGP22	28390–28410	2.9	30.5	polysaccharide biosynthesis	–	N	A1S_2896-3841
RGP23	31330–31680	35.0	35.5	fosfomycin resistance, IS*Aba13*_17	–	Y	variable region
RGP24	32600–32660	7.0	26.3	unknown	–	Y	A1S_3899-3901

[a]ATC - atypical trinucleotide composition; RGPs with $\chi 2$ values greater than 500.
[b]ND - Undetermined integrase target site.
[c]Accessory element designation in *A. baumannii* ACICU [9].
*Premature stop in integrase protein
doi:10.1371/journal.pone.0058628.t002

Table 3. Accessory elements absent in the *Acinetobacter baumannii* D1279779 genome.

RGP_	D1279779 boundaries	ACICU CDs	Putative function/features of interest	No. of strains with RGP present[c]
D01	02020 (*comM*)	AB57_0243-0306[a]	*A. baumannii* antibiotic resistance island	9/10 strains AbaG1 in SDF
D02	06960–06970	00685–00701	glucose dehydrogenase 2, various insertion sequences	5/10 strains exclusive to ICII
D03	07900–07910	00873–00880	haem degradation, iron acquisition	5/10 strains
D04	11060–11100	AB57_1176–1215[b]	glucarate degradation pathway, vanillate degradation	3/10 strains mediated by IS*Aba13*_3
D05	12530–12580	AB57_1379–1406[b]	degraded prophage	3/10 strains exclusive to ICI
D06	13150–13240	01295–01313	type VI secretion system	9/10 strains replaced with D1279779_RGP12
D07	13570–13660	01354–01403	β-ketoadipate pathway, 2-aminoethylphosphonate transport/ metabolism, carnitine degradation	9/10 strains replaced with D1279779_RGP13
D08	17660–17960	01810–01823	pillin biogenesis, saccharopine dehydrogenase, monoamine oxidase	8/10 strains exclusive to ICI and ICII
D09	17770–17810	01833–01838	unknown	5/10 strains exclusive to ICII replaced with D1279779_RGP15
D10	17920–17980	01852–01863	unknown	5/10 strains Exclusive to ICII replaced with D1279779_RGP16
D11	18260–18270	01892–01897	demethylmenaquinone methyltransferase, phosphoglycerate dehydrogenase	5/10 strains exclusive to ICII mediated by IS*Aba13*_3
D12	18390–18490	01914–01920	metabolism	7/10 strains replaced with D1279779_RGP18
D13	18590–18610	01937–01949	various dehydrogenases	5/10 strains exclusive to ICII
D14	20390–20400	02140–02235	prophage	4/10 strains
D15	21340–21350	ABK1_1354–1403[b]	prophage	3/10 strains
D16	21980–21990	AB57_2539–2553[b]	two component regulation, fatty acid modification	3/10 strains exclusive to ICI
D17	22190–22200	AB57_2575–2579[b]	unknown	3/10 strains exclusive to ICI
D18	23890–24070	02595–02623	transcriptional regulation, transporters/permeases	7/10 strains replaced with D1279779_RGP21
D19	24530–24540	AB57_2901–2906[b]	crispr-associated proteins	3/10 strains exclusive to ICI
D20	28500–28510	03158–03160	hypothetical proteins	4/10 strains

[a]3′ *comM* fragment not present in *A. baumannii* ACICU.
[b]Gene cluster not present in *A. baumannii* ACICU.
[c]Number of currently completely sequenced *A. baumannii* strains other than D1279779.
doi:10.1371/journal.pone.0058628.t003

also encodes an allantoicase (Alc) which functions in a similar manner to AllC, except degradation of allantoate is performed by a single-step catalytic mechanism, rather than a dual-step one [73].

Two other RGPs, D1279779_RGP01 and D1279779_RGP22, encode enzymes involved in capsule polysaccharide biosynthesis. The RGP D1279779_RGP03 encodes additional copies of three enzymes (diaminopimelate decarboxylase, a ribulose-phosphate 3-epimerase, and an uroporphyrinogen decarboxylase) encoded elsewhere in the core genome.

The D1279779 genome encodes two putative prophages, one of which encodes paralogues of the error-prone DNA polymerase V subunits UmuC and UmuD; these genes are frequently associated with prophages and other mobile genetic elements such as genomic islands and plasmids [74]. The second prophage encodes a paralogue of lipid A phosphoethanolamine transferase (EptA) which facilitates the covalent modification of lipid A. Overexpression of *eptA* in *A. baumannii* has previously been associated with increased colistin resistance [75]. However, strain D1279779 appears to exhibit greater colistin sensitivity than nosocomial isolates (discussed below).

Notably, *A. baumannii* D1279779 does not encode the AbaR genomic island (Table 3), a drug resistance island present in the majority of published *A. baumannii* genomes. AbaRs are characteristically found inserted within the *comM* gene [76], and associated with the accretion of multiple insertion sequences and genomic islands [77,78]. When present, AbaRs are capable of encoding resistance to a multitude of antibiotics (aminoglycosides, beta-lactams, sulphonamides, tetracyclines) and heavy metals (arsenic and mercury) [2]. *A. baumannii* D1279779 has an intact *comM* gene and lacks an AbaR island, which may partially explain the antimicrobial susceptibility phenotype observed for this CA-AB isolate.

Nineteen other gene clusters, present in the genomes of other *A. baumannii* isolates, are also absent from D1279779. Ten of the missing gene clusters are unique to *A. baumannii* strains of the ICI and ICII lineages. Six other gene clusters absent in D1279779 were replaced with RGPs unique to this strain or additionally present only in ATCC 17978 (Table 2, Table 3). Some of the absent genes included those involved in the β-ketoadipate pathway [79], the D-glucarate degradation pathway [80] and potentially, the catabolism of carnitine and vanillate (Table 3). The loss of these metabolic genes very likely correlates with the observed inability of this strain to utilise glactarate, glucarate and 4-hydroxybenoazte as carbon sources but does not prevent the utilisation of carnitine (see below).

Table 4. Genomic coordinates of IS*Aba13* copies and orthologous genes disrupted.

IS*Aba*13_	Coordinates	Gene(s) disrupted	Annotation/putative function	*A. baumannii* orthologues(s)
1	583114–584152	ABD1_05170	TetR family transcriptional regulator	ACICU_00556
2	1142807–1143845	ABD1_10420	prophage-associated hypothetical	ACICU_01061
3	1214764–1213726	Loss of ~44 kb	glucarate/galactarate/vanillate catabolism	AB57_1176-1212[b]
4	1421673–1422711	ABD1_13230; 13240	hypothetical protein; ankryin repeat-containing protein	no orthologue[c]; ACICU_01314
5	1463547–1464585	ABD1_13650	hypothetical protein	no orthologue[c]
6	1487979–1489017	ABD1_13850	competence-damage inducible protein CinA	ACICU_01424
7[a]	1544569–1545607	ABD1_14440	MFS transporter	ACICU_01479
8	1787573–1788611	–	–	–
9	1892223–1891185	ABD1_17670	fimbrial family protein	ACICU_01810
10	1963590–1964628	ABD1_18260	type 1 secretion C-terminal target domain	ACICU_01891
11	2095736–2094698	–	–	–
12	2195861–2194823	–	–	–
13	2233207–2234245	–	–	–
14	2256425–2257463	–	–	–
15	2640973–2642011	ABD1_24300	hypothetical protein	ACICU_02649
16	2879635–2878597	ABD1_26440	bacterial capsule synthesis protein	ACICU_02936
17	3422518–3423556	ABD1_31460	hypothetical protein	A1S_3893[b]
18	3599602–3600640	ABD1_33030	hypothetical protein	ACICU_03597

[a]IS*Aba13* sequence isoform 2.
[b]gene(s) not present in *A. baumannii* ACICU.
[c]Not present in any currently complete *A. baumannii* genome.
doi:10.1371/journal.pone.0058628.t004

Comparison of known or predicted virulence genes amongst *A. baumannii* sequences indicated strain D1279779 lacks several genes conserved amongst previously sequenced *A. baumannii* strains. These include a type VI secretion system gene cluster [81] (Table 3), a gene encoding the *Acinetobacter* trimeric autotransporter protein (Ata) [82] (which is truncated in D1279779), haem acquisition (Table 3) and a gene encoding for the biofilm-associated protein (Bap) [83,84]. The loss of these genes may correlate with the observation that D1279779 has only a modest capacity for biofilm formation and adherence to nasopharyngeal cells [62]. Nevertheless, this strain still carries other potential virulence-associated genes, including those coding for acinetobactin biosynthesis [85], capsular polysaccharide polymerisation/export [86], type I and type IV pili biogeneses [87] and phospholipases C and D [88,89].

Multiple copies of IS*Aba13* are present in the genome

Eighteen copies of the transposon IS*Aba13*, previously identified in *A. baumannii* AB0057 [8], are present throughout the D1279779 chromosome (Figure 1). All insertion sequences, with the exception of one, are of an identical isoform (Table 4). The two insertion sequence isoforms have nucleotide identities of 99% and 96% to the IS*Aba13* in AB0057. Thirteen copies of IS*Aba13* are inserted within annotated genes, including some potentially encoding virulence and competence functions such as a fimbrial adhesin, a type I secretion domain protein, bacterial capsule synthesis protein and competence-damaged induced protein CinA (Table 4). One of the copies of IS*Aba13* replaced an approximately 44 kb region, present in some other *A. baumannii* strains, that carries genes for the catabolism of glucarate, galactarate and vanillate (Table 3, Table 4).

The catabolic phenome of *A. baumannii* D1279779

Biolog Phenotype MicroArrays are a respiration-based assay system that can test up to 2000 phenotypic traits simultaneously [56,90]. This system uses 96 well plates with each well testing a separate phenotype using a tetrazolium dye that produces a colour change in response to cellular respiration. The phenome of *A. baumannii* D1279779 was investigated with the Biolog Phenotype MicroArray System and compared with an ICI strain (AYE), an ICII strain (ACICU) and ATCC 17978, a nosocomial isolate from 1951, predating the emergence of the major global clonal lineages as the dominant nosocomial strains.

The four *A. baumannii* strains tested were similar in their utilization of sole carbon (Figure 4) and nitrogen sources (Figure S2). They utilized a combined total of 80 carbon sources out of the 190 tested, encompassing a range of amino acids, carboxylic acids, saccharides and miscellaneous compounds (Figure 4). Strains D1279779 and ATCC 17978 were able to utilise a greater breadth of sole carbon and nitrogen sources compared to ICI and ICII strains, particularly in relation to amino acids including alaninamide, asparagine, isoleucine, glutamate and homoserine (Figure 4, Figure S2). The observed phenotypic profiles suggest the emergence of ICI and ICII lineages in nosocomial settings has coincided with a narrowing of their substrate utilisation capabilities. Furthermore, *A. baumannii* ACICU displayed a higher respiration rate on substrates arginine, ornithine, phenylalanine, pyroglutamic acid, quinic acid and ribonolactone (Figure 4), suggesting a possible specialization in terms of carbon utilization preferences.

In order to independently confirm the Biolog respiration data, the ability of *A. baumannii* D1279779 and ACICU to grow on minimal media in the presence of fifteen sole carbon compounds

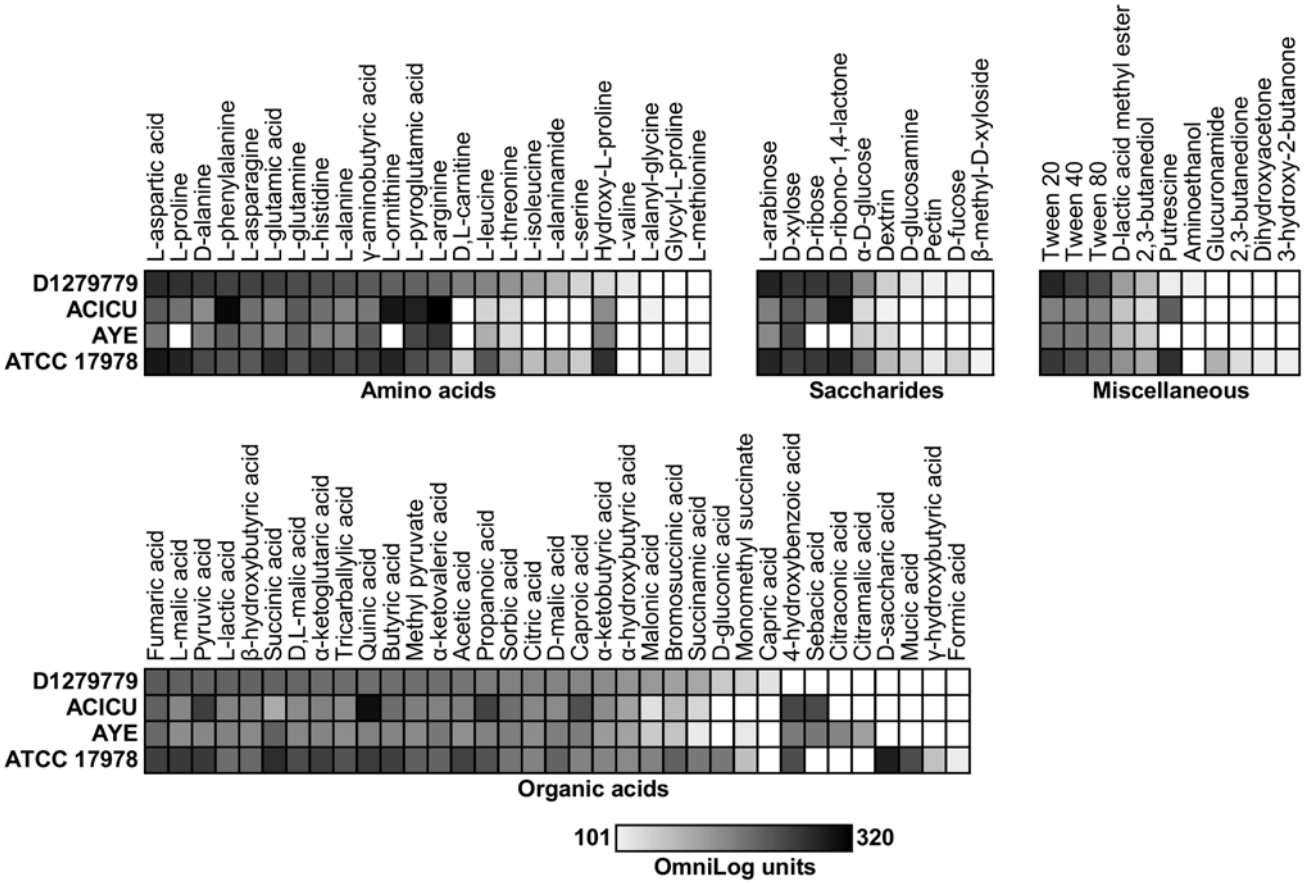


Figure 4. The catabolic phenome of *Acinetobacter baumannii*. Strengths of carbon utilisation phenotypes of *A. baumannii* strains D1279779, ACICU, AYE and ATCC 17978 were determined using Biolog Phenotype Microarray plates PM1 and PM2. The maximal kinetic curve height was expressed as a greyscale ranging from 101 (light grey) to 320 OmniLog units (black). Phenotypes are arranged from strongest to weakest relative to *A. baumannii* D1279779. Phenotypes <101 OmniLog units (white) were considered negative.
doi:10.1371/journal.pone.0058628.g004

was tested. Turbidity consistent with cellular replication was observed in all tested carbon sources except for D-gluconic acid and α-D-glucose (see below) (Table S1), which was concordant with the phenotype microarray data.

Metabolic reconstructions for each of the four strains were undertaken to analyse whether the phenotypic differences detected could be ascribed to the presence or absence of specific genes. *A. baumannii* D1279779 was unable to utilise 4-hydroxybenzoic acid as a sole source of carbon in either the phenotype microarray (Figure 4) or minimal media (Table S1). Genomic analysis indicated that this strain lacked the *pobA* gene encoding a 4-hydroxybenzoate 3-hydroxylase (Table 3), required for the conversion of 4-hydroxybenzoic acid to protocatechuic acid [79]. Both this strain and *A. baumannii* SDF were the only genomes examined that lacked this gene.

A. baumannii ATCC 17978 was the only strain in the test group able to utilise the diastereomers saccharic acid (glucarate) and mucic acid (galactarate) as sole sources of carbon (Figure 4). This was attributable to the presence of a gene cluster involved in the glucarate degradation pathway [80] conserved in some other *A. baumannii* strains, but absent in D1279779, ACICU and AYE (Table 3, Table 4).

Both strains D1279779 and ATCC 17978 showed a positive result for respiration on both D-gluconic acid and α-D-glucose. Our results were concordant with the results of a previous phenotype microarray study [91], which demonstrated the capability of various *Acinetobacter* sp. (including *A. baumannii*) to respire in the presence of both these carbon sources. This is curious, since the majority of *Acinetobacter* species including *A. baumannii* have repeatedly been reported as incapable of utilising D-glucose and D-gluconate as sole carbon sources [92,93]. All four strains encode the Entner-Doudoroff (ED) pathway, an alternative glucose assimilation route that requires the cofactor pyrroloquinoline quinone (PQQ) [94]; a PQQ biosynthetic pathway is evident in three of the strains (but not in ATCC 17978). Growth experiments in both M9 and IF-0 minimal media using D-gluconic acid and α-D-glucose as sole carbon sources, with or without PQQ supplementation, were all negative. The apparent contradiction between the genome, phenotype microarray data and growth assays may indicate these substrates are not assimilated, but rather act as energy donors [95].

A number of the differences observed in carbon and nitrogen utilization between the strains could not be accounted for at the genetic level. This could be due to differences in regulation, membrane transport activity, or the presence of novel uncharacterized catabolic pathways. For instance, strains D1279779 and ATCC 17978 were able to utilize the branched-chain amino acids leucine and isoleucine, while ACICU and AYE were only able to respire weakly on leucine. All of the strains encoded a putative branched-chain amino acid aminotransferase (IlvE) for the

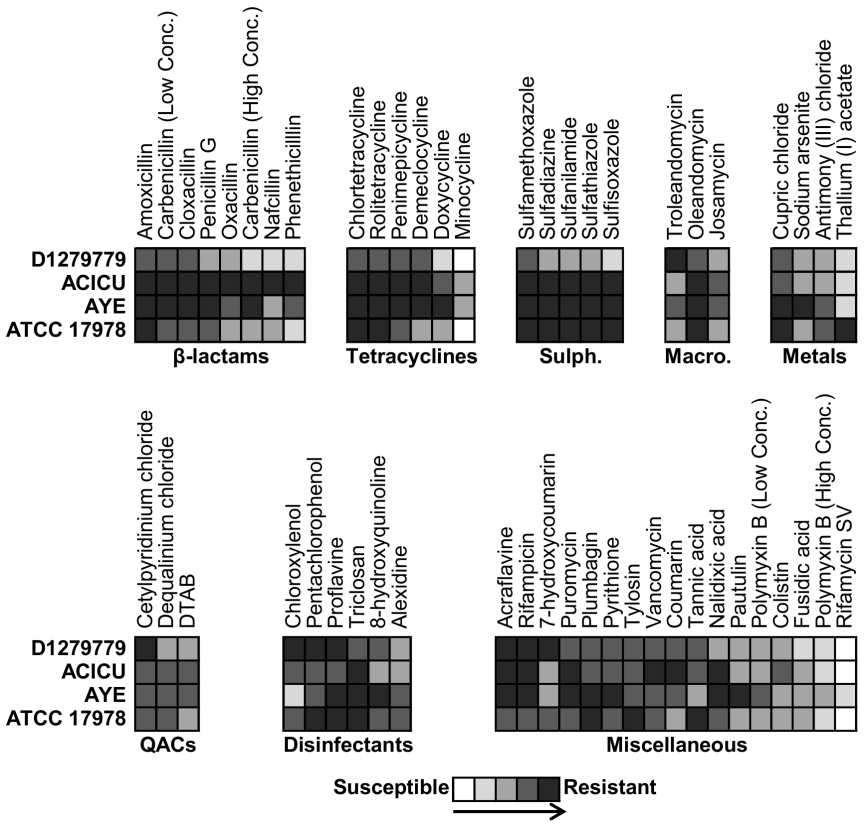


Figure 5. The resistance phenome of *Acinetobacter baumannii*. Select antimicrobial resistance phenotypes of *A. baumannii* strains D1279779, ACICU, AYE and ATCC 17978 are displayed as a five-coloured grey scale ranging from no resistance (white) to maximal resistance (dark grey). Phenotypes are arranged from strongest to weakest relative to *A. baumannii* D1279779. Abbreviations: Macro, macrolides; QACs, quaternary ammonium compounds; and Sulph, sulphonamides.
doi:10.1371/journal.pone.0058628.g005

reversible transamination of isoleucine, leucine and valine [96], it seems likely the phenotypic differences are due to altered regulation or transport factors. It is possible the residual leucine utilisation in AYE and ACICU is due to a tyrosine aminotransferase (TyrB), which overlaps IlvE in specificity to leucine [97]. In another instance, *A. baumannii* AYE was found to be unable to utilise proline, ornithine and putrescine as carbon sources (Figure 4) or citrulline, ornithine and putrescine as nitrogen sources (Figure S2). This suggests potential defects in proline and arginine catabolism, but these deficiencies cannot currently be accounted for at the genetic level.

The resistance phenome of *A. baumannii* D1279779

The osmotolerance (Table S1), pH tolerance (Figure S2) and antimicrobial resistance (Figure 5, Table S1) of the four *A. baumannii* strains was also examined with Biolog Phenotype MicroArrays. *A. baumannii* ACICU and ATCC 17978 were found to be more sensitive to acidic pH and were only able to deaminate a limited number of compounds at pH 4.5 (Figure S2).

The four *A. baumannii* strains displayed high intrinsic resistance to many antimicrobial compounds. Respiration in all four strains was observed, at all concentrations, for 94 of the 240 antimicrobials tested (Table S1). No respiration was observed in any strain for five compounds: the antibiotic novobiocin and the heavy metals salts potassium chromate, cadmium chloride, sodium orthovanadate and sodium metavanadate. Differential susceptibility to a further 106 compounds was observed in the four strains (Figure 5, Table S1). Strains D1279779 and ATCC 17978 were noticeably more susceptible to a range of clinically important antibiotics, including the beta-lactams and tetracyclines. The higher levels of resistance of *A. baumannii* ACICU and AYE towards beta-lactams and sulphonamides is likely due to resistance determinants encoded within their respective AbaR elements [2,9], as well as the carbapenem resistance plasmid pACICU1 [9]. Although, ATCC 17978 was also resistant to sulphonamides, it is likely that this resistance is encoded on a separate genomic island [98]. The strain AYE encodes two TetA tetracycline resistance efflux pumps [2], though in the case of ACICU, there are no characterized tetracycline resistance genes; the observed resistance may be due to the function of other efflux pumps. Resistance to the quinolone nalidixic acid can be accounted for by mutations both in *gyrA* and *parC* in strain AYE, and a mutation in *gyrA* for strain ACICU [8]; neither mutation is present in ATCC 17978 or D1279779. *A. baumannii* AYE was found to be relatively more resistant to arsenic and rifamycin SV, attributable to the respective presence of an arsenic resistance cluster and a rifampin ADP-ribosyltransferase (*arr-2*) in its AbaR island [2]. Strains AYE and ATCC 17978 had increased resistance to copper, very likely due to the presence of a copper resistance cluster (*copABCDRS*) [8], in addition to the copper resistance genes (*pcoAB*) present in the *A. baumannii* core genome.

Conclusions

The CA-AB isolate D1279779, while phylogenetically related to the ICII *A. baumannii* global clonal lineage, phenotypically

resembles ATCC 17978 in terms of carbon and nitrogen utilization, and drug susceptibility profile. Phenotypic testing of nosocomial *A. baumannii* suggests that the narrowing of substrate utilisation capabilities and expansion of drug resistance profiles in both ICI and ICII global clonal lineages has contributed to their success in the nosocomial milieu. Our genomic analysis of the CA-AB isolate D1279779 reveals the absence of the AbaR island common to nosocomial isolates. D1279779 does however comprise of 24 novel RGPs that encode catabolic functions, polysaccharide biosynthesis and many hypothetical proteins of unknown function. Reports in the literature have suggested that CA-AB is associated with higher mortality rates than nosocomial *A. baumannii* strains. Whilst there were no obvious virulence-associated genes unique to D1279779, there was however, the apparent loss of several genes associated with virulence, particularly with respect to biofilm formation and eukaryotic cell adhesion. The characteristics of D1279779 may be more representative of an environmental or pre-antibiotic era clinical *A. baumannii* isolate, and appears quite distinct from the current dominant lineages of nosocomial isolates.

Supporting Information

Figure S1 Synteny of *Acinetobacter baumannii*. Chromosomal alignments of the *A. baumannii* D1279779 genome against the ten currently complete *A. baumannii* genomes were generated using progressive MAUVE [103]. Regions of significant synteny between the strains are shown as coloured blocks and unshared regions are seen as white gaps.
(PNG)

Figure S2 Phenotypic analysis of nitrogen utilisation and pH stress tolerance. Strengths of nitrogen utilisation (A) and the pH tolerance phenotypes (B) of *A. baumannii* strains D1279779, ACICU, AYE and ATCC 17978 were determined were determined using Biolog Phenotype Microarray plates PM3 and PM10, respectively. The maximal kinetic curve height was expressed as a greyscale ranging from 101 OmniLog units (light grey) to 310 and 360 OmniLog units (black) for nitrogen and pH tolerance phenotypes, respectively. Phenotypes are arranged from strongest to weakest relative to *A. baumannii* D1279779. Phenotypes <115 OmniLog units for nitrogen phenotypes and <101 OmniLog units for pH tolerance phenotypes were considered negative phenotypes and are represented in white.
(TIF)

Table S1 Raw phenotype microarray data. The maximal kinetic curve height for all phenotypes obtained from plates PM1-3 (carbon and nitrogen utilisation) and PM9-20 (osmotolerance, pH tolerance and antimicrobial exposure) expressed in OmniLog units.
(XLS)

Acknowledgments

We would like to thank Matthias Maiwald for providing *A. baumannii* D1279779, Alessandra Carattoli for providing *A. baumannii* ACICU and Patrice Nordmann for providing *A. baumannii* AYE.

Author Contributions

Conceived and designed the experiments: DNF KAH MHB ITP. Performed the experiments: DNF LDHE KAH BAE SGT BSS. Analyzed the data: DNF LDHE BAE BSS ITP. Wrote the paper: DNF LDHE KAH MHB AYP BCM ITP.

References

1. Roca I, Espinal P, Vila-Farres X, Vila J (2012) The *Acinetobacter baumannii* oxymoron: commensal hospital dweller turned pan-drug-resistant menace. Front Microbiol 3: 148.
2. Fournier PE, Vallenet D, Barbe V, Audic S, Ogata H, et al. (2006) Comparative genomics of multidrug resistance in *Acinetobacter baumannii*. PLoS Genet 2: 62–72.
3. Vila J, Marti S, Sanchez-Cespedes J (2007) Porins, efflux pumps and multidrug resistance in *Acinetobacter baumannii*. J Antimicrob Chemoth 59: 1210–1215.
4. Wendt C, Dietze B, Dietz E, Ruden H (1997) Survival of *Acinetobacter baumannii* on dry surfaces. J Clin Microbiol 35: 1394–1397.
5. Jawad A, Seifert H, Snelling AM, Heritage J, Hawkey PM (1998) Survival of *Acinetobacter baumannii* on dry surfaces: comparison of outbreak and sporadic isolates. J Clin Microbiol 36: 1938–1941.
6. Neely AN, Maley MP, Warden GD (1999) Computer keyboards as reservoirs for *Acinetobacter baumannii* in a burn hospital. Clin Infect Dis 29: 1358–1359.
7. Park JY, Kim S, Kim S-M, Cha SH, Lim S-K, et al. (2011) Complete genome sequence of multidrug-resistant *Acinetobacter baumannii* strain 1656-2, which forms sturdy biofilm. J Bacteriol 193: 6393–6394.
8. Adams MD, Goglin K, Molyneaux N, Hujer KM, Lavender H, et al. (2008) Comparative genome sequence analysis of multidrug-resistant *Acinetobacter baumannii*. J Bacteriol 190: 8053–8064.
9. Iacono M, Villa L, Fortini D, Bordoni R, Imperi F, et al. (2008) Whole-genome pyrosequencing of an epidemic multidrug-resistant *Acinetobacter baumannii* strain belonging to the European clone II group. Antimicrob Agents Chemother 52: 2616–2625.
10. Smith MG, Gianoulis TA, Pukatzki S, Mekalanos JJ, Ornston LN, et al. (2007) New insights into *Acinetobacter baumannii* pathogenesis revealed by high-density pyrosequencing and transposon mutagenesis. Gene Dev 21: 601–614.
11. Vallenet D, Nordmann P, Barbe V, Poirel L, Mangenot S, et al. (2008) Comparative analysis of *Acinetobacters*: three genomes for three lifestyles. PLoS One 3: e1805.
12. Gao F, Wang Y, Liu Y-J, Wu X-M, Lv X, et al. (2011) Genome sequence of *Acinetobacter baumannii* MDR-TJ. J Bacteriol 193: 2365–2366.
13. Zhou H, Zhang T, Yu D, Pi B, Yang Q, et al. (2011) Genomic analysis of the multidrug-resistant *Acinetobacter baumannii* strain MDR-ZJ06 widely spread in China. Antimicrob Agents Chemother 55: 4506–4512.
14. Chen C-C, Lin Y-C, Sheng W-H, Chen Y-C, Chang S-C, et al. (2011) Genome sequence of a dominant, multidrug-resistant *Acinetobacter baumannii* strain, TCDC-AB0715. J Bacteriol 193: 2361–2362.
15. 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: 3569–3577.
16. La Scola B, Raoult D (2004) *Acinetobacter baumannii* in human body louse. Emerg Infect Dis 10: 1671–1673.
17. Bouvresse S, Socolovshi C, Berdjane Z, Durand R, Izri A, et al. (2011) No evidence of *Bartonella quintana* but detection of *Acinetobacter baumannii* in head lice from elementary schoolchildren in Paris. Comp Immunol Microb 34: 475–477.
18. Sarma PM, Bhattacharya D, Krishnan S, Lal B (2004) Assessment of intra-species diversity among strains of *Acinetobacter baumannii* isolated from sites contaminated with petroleum hydrocarbons. Can J Microbiol 50: 405–414.
19. Chang LK, Ibrahim D, Omar IC (2011) A laboratory scale bioremediation of Tapis crude oil contaminated soil by bioaugmentation of *Acinetobacter baumannii* T30C. Afr J Microbiol Res 5: 2609–2615.
20. Berg G, Roskot N, Steidle A, Eberl L, Zock A, et al. (2002) Plant-dependent genotypic and phenotypic diversity of antagonistic rhizobacteria isolated from different *Verticillium* host plants. Appl Environ Microb 68: 3328–3338.
21. Sachdev D, Nema P, Dhakephalkar P, Zinjarde S, Chopade B (2010) Assessment of 16S rRNA gene-based phylogenetic diversity and promising plant growth-promoting traits of *Acinetobacter* community from the rhizosphere of wheat. Microbiol Res 165: 627–638.
22. Girlich D, Poirel L, Nordmann P (2010) First isolation of the blaOXA-23 carbapenemase gene from an environmental *Acinetobacter baumannii* isolate. Antimicrob Agents Chemother 54: 578–579.
23. Ghai R, Rodriguez-Valera F, McMahon KD, Toyama D, Rinke R, et al. (2011) Metagenomics of the water column in the pristine upper course of the Amazon River. PLoS One 6: e23785.
24. Chu YW, Leung CM, Houang ET, Ng KC, Leung CB, et al. (1999) Skin carriage of Acinetobacters in Hong Kong. J Clin Microbiol 37: 2962–2967.
25. Anstey NM, Currie BJ, Hassell M, Palmer D, Dwyer B, et al. (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: 685–686.

26. Zeana C, Larson E, Sahni J, Bayuga SJ, Wu F, et al. (2003) The epidemiology of multidrug-resistant *Acinetobacter baumannii*: Does the community represent a reservoir? Infect Cont Hosp Ep 24: 275–279.
27. Chen HP, Chen TL, Lai CH, Fung CP, Wong WW, et al. (2005) Predictors of mortality in *Acinetobacter baumannii* bacteremia. J Microbiol Immunol Infect 38: 127–136.
28. Falagas ME, Karveli EA, Kelesidis I, Kelesidis T (2007) Community-acquired *Acinetobacter* infections. Eur J Clin Microbiol 26: 857–868.
29. Elliott JH, Anstey NM, Jacups SP, Fisher DA, Currie BJ (2005) Community-acquired pneumonia in northern Australia: low mortality in a tropical region using locally-developed treatment guidelines. Int J Infect Dis 9: 15–20.
30. Leung AS, Chu CM, Tsang KY, Lo FH, Lo KF, et al. (2006) Fulminant community-acquired *Acinetobacter baumannii* pneumonia as a distinct clinical syndrome. Chest 129: 102–109.
31. Wang JT, McDonald LC, Chang SC, Ho M (2002) Community-acquired *Acinetobacter baumannii* bacteremia in adult patients in Taiwan. J Clin Microbiol 40: 1526–1529.
32. Ong CWM, Lye DCB, Khoo KL, Chua GSW, Yeoh SF, et al. (2009) Severe community–acquired *Acinetobacter baumannii* pneumonia: An emerging highly lethal infectious disease in the Asia-Pacific. Respirology 14: 1200–1205.
33. Chong YP, Jung K-S, Lee KH, Kim M-N, Moon SM, et al. (2010) The bacterial etiology of community-acquired pneumonia in Korea: a nationwide prospective multicenter study. Infect Chemother 42: 397–403.
34. Anstey NM, Currie BJ, Withnall KM (1992) Community-acquired *Acinetobacter* pneumonia in the Northern Territory of Australia. Clin Infect Dis 14: 83–91.
35. Gradon JD, Chapnick EK, Lutwick LI (1992) Infective endocarditis of a native valve due to *Acinetobacter*: case report and review. Clin Infect Dis 14: 1145–1148.
36. Bick JA, Semel JD (1993) Fulminant community-acquired *Acinetobacter* pneumonia in a healthy woman. Clin Infect Dis 17: 820–821.
37. Reindersma P, Nohlmans L, Korten JJ (1993) *Acinetobacter*, an infrequent cause of community acquired bacterial meningitis. Clin Neurol Neurosurg 95: 71–73.
38. Ozaki T, Nishimura N, Arakawa Y, Suzuki M, Narita A, et al. (2009) Community-acquired *Acinetobacter baumannii* meningitis in a previously healthy 14-month-old boy. J Infect Chemother 15: 322–324.
39. Moreira Silva G, Morais L, Marques L, Senra V (2012) *Acinetobacter* community-acquired pneumonia in a healthy child. Rev Port Pneumol 18: 96–98.
40. Einsiedel LJ, Woodman RJ (2010) Two nations: racial disparities in bloodstream infections recorded at Alice Springs Hospital, central Australia, 2001–2005. Med J Aust 192: 567–571.
41. Chevreux B, Wetter T, Suhai S (1999) Genome sequence assembly using trace signals and additional sequence information. Computer Science and Biology: Proceedings of the German Conference on Bioinformatics (GCB) 99: 45–56.
42. Darling ACE, Mau B, Blattner FR, Perna NT (2004) Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res 14: 1394–1403.
43. van Hijum SAFT, Zomer AL, Kuipers OP, Kok J (2005) Projector 2: contig mapping for efficient gap-closure of prokaryotic genome sequence assemblies. Nucleic Acids Res 33: W560–W566.
44. Altschul SF, Madden TL, Schaffer AA, Zhang JH, Zhang Z, et al. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25: 3389–3402.
45. Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl Acids Symp Ser 41: 95–98.
46. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, et al. (2008) The RAST server: Rapid Annotations using Subsystems Technology. BMC Genomics 9: 75.
47. Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, et al. (2000) Artemis: sequence visualization and annotation. Bioinformatics 16: 944–945.
48. Okonechnikov K, Golosova O, Fursov M (2012) Unipro UGENE: a unified bioinformatics toolkit. Bioinformatics 28: 1166–1167.
49. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, et al. (2009) BLAST+: architecture and applications. BMC Bioinformatics 10: 421.
50. Paulsen IT, Seshadri R, Nelson KE, Eisen JA, Heidelberg JF, et al. (2002) The *Brucella suis* genome reveals fundamental similarities between animal and plant pathogens and symbionts. Proc Natl Acad Sci USA 99: 13148–13153.
51. Mathee K, Narasimhan G, Valdes C, Qiu X, Matewish JM, et al. (2008) Dynamics of *Pseudomonas aeruginosa* genome evolution. Proc Natl Acad Sci USA 105: 3100–3105.
52. Ogier J-C, Calteau A, Forst S, Goodrich-Blair H, Roche D, et al. (2010) Units of plasticity in bacterial genomes: new insight from the comparative genomics of two bacteria interacting with invertebrates, *Photorhabdus* and *Xenorhabdus*. BMC Genomics 11: 568.
53. Siguier P, Perochon J, Lestrade L, Mahillon J, Chandler M (2006) ISfinder: the reference centre for bacterial insertion sequences. Nucleic Acids Res 34: D32–D36.
54. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, et al. (2011) MEGA5: Molecular Evolutionary Genetics Analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28: 2731–2739.
55. 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.
56. Bochner BR, Gadzinski P, Panomitros E (2001) Phenotype microarrays for high-throughput phenotypic testing and assay of gene function. Genome Res 11: 1246–1255.
57. Keseler IM, Collado-Vides J, Santos-Zavaleta A, Peralta-Gil M, Gama-Castro S, et al. (2011) EcoCyc: a comprehensive database of *Escherichia coli* biology. Nucleic Acids Res 39: D583–D590.
58. Caspi R, Altman T, Dreher K, Fulcher CA, Subhraveti P, et al. (2012) The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. Nucleic Acids Res 40: D742–D753.
59. Kanehisa M, Goto S, Kawashima S, Okuno Y, Hattori M (2004) The KEGG resource for deciphering the genome. Nucleic Acids Res 32: D277–D280.
60. Bertini A, Poirel L, Mugnier PD, Villa L, Nordmann P, et al. (2010) Characterization and PCR-based replicon typing of resistance plasmids in *Acinetobacter baumannii*. Antimicrob Agents Chemother 54: 4168–4177.
61. 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: 112–123.
62. Eijkelkamp BA, Stroeher UH, Hassan KA, Papadimitrious MS, Paulsen IT, et al. (2011) Adherence and motility characteristics of clinical *Acinetobacter baumannii* isolates. FEMS Microbiol Lett 323: 44–51.
63. Srivatsan A, Tehranchi A, MacAlpine DM, Wang JD (2010) Co-orientation of replication and transcription preserves genome integrity. PLoS Genetics 6: e1000810.
64. Hill CW, Harnish BW (1981) Inversions between ribosomal RNA genes of *Escherichia coli*. Proc Natl Acad Sci USA 78: 7069–7072.
65. Klockgether J, Munder A, Neugebauer J, Davenport CF, Stanke F, et al. (2010) Genome diversity of *Pseudomonas aeruginosa* PAO1 laboratory strains. J Bacteriol 192: 1113–1121.
66. Abadjieva A, Patel J, Webb M, Zinkevich V, Firman K (1993) A Deletion Mutant of the Type-IC Restriction-Endonuclease *Eco*R124I Expressing a Novel DNA Specificity. Nucleic Acids Res 21: 4435–4443.
67. Meister J, MacWilliams M, Hubner P, Jutte H, Skrzypek E, et al. (1993) Macroevolution by transposition: drastic modification of DNA recognition by a type I restriction enzyme following Tn5 transposition. EMBO J 12: 4585–4591.
68. Adamczyk-Poplawska M, Lower M, Piekarowicz A (2011) Deletion of one nucleotide within the homonucleotide tract present in the *hsdS* gene alters the DNA sequence specificity of type I restriction-modification system NgoAV. J Bacteriol 193: 6750–6759.
69. Kobayashi I (2004) Restriction-modification systems as minimal forms of life. In: Pingoud AM, editor. Restriction Endonucleases: Springer Berlin Heidelberg. pp. 19–62.
70. Waldron DE, Lindsay JA (2006) Sau1: a novel lineage-specific type I restriction-modification system that blocks horizontal gene transfer into *Staphylococcus aureus* and between *S. aureus* isolates of different lineages. J Bacteriol 188: 5578–5585.
71. Schneider BL, Kiupakis AK, Reitzer LJ (1998) Arginine catabolism and the arginine succinyltransferase pathway in *Escherichia coli*. J Bacteriol 180: 4278–4286.
72. Sakurada K, Ohta T, Fujishiro K, Hasegawa M, Aisaka K (1996) Acetylpolyamine amidohydrolase from *Mycoplana ramosa*: gene cloning and characterization of the metal-substituted enzyme. J Bacteriol 178: 5781–5786.
73. Agarwal R, Burley SK, Swaminathan S (2007) Structural analysis of a ternary complex of allantoate amidohydrolase from *Escherichia coli* reveals its mechanics. J Mol Biol 368: 450–463.
74. Permina EA, Mironov AA, Gelfand MS (2002) Damage-repair error-prone polymerases of eubacteria: association with mobile genome elements. Gene 293: 133–140.
75. Beceiro A, Llobet E, Aranda J, Bengoechea JA, Doumith M, et al. (2011) Phosphoethanolamine modification of lipid A in colistin-resistant variants of *Acinetobacter baumannii* mediated by the *pmrAB* two-component regulatory system. Antimicrob Agents Chemother 55: 3370–3379.
76. Bonnin RA, Poirel L, Nordmann P (2012) AbaR-type transposon structures in *Acinetobacter baumannii*. J Antimicrob Chemoth 67: 234–236.
77. Post V, White PA, Hall RM (2010) Evolution of AbaR-type genomic resistance islands in multiply antibiotic-resistant *Acinetobacter baumannii*. J Antimicrob Chemoth 65: 1162–1170.
78. 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: 3201–3206.
79. Park SH, Kim JW, Yun SH, Leem SH, Kahng HY, et al. (2006) Characterization of β-ketoadipate pathway from multi-drug resistance bacterium, *Acinetobacter baumannii* DU202 by proteomic approach. J Microbiol 44: 632–640.
80. Aghaie A, Lechaplais C, Sirven P, Tricot S, Besnard-Gonnet M, et al. (2008) New insights into the alternative D-glucarate degradation pathway. J Biol Chem 283: 15638–15646.
81. Henry R, Vithanage N, Harrison P, Seemann T, Coutts S, et al. (2012) Colistin-resistant, lipopolysaccharide-deficient *Acinetobacter baumannii* responds to lipopolysaccharide loss through increased expression of genes involved in the synthesis and transport of lipoproteins, phospholipids, and poly-β-1,6-N-acetylglucosamine. Antimicrob Agents Chemother 56: 59–69.

82. Bentancor LV, Camacho-Peiro A, Bozkurt-Guzel C, Pier GB, Maira-Litrán T (2012) Identification of Ata, a multifunctional trimeric autotransporter of *Acinetobacter baumannii*. J Bacteriol 194: 3950–3960.
83. Loehfelm TW, Luke NR, Campagnari AA (2008) Identification and characterization of an *Acinetobacter baumannii* biofilm-associated protein. J Bacteriol 190: 1036–1044.
84. Brossard KA, Campagnari AA (2012) The *Acinetobacter baumannii* biofilm-associated protein plays a role in adherence to human epithelial cells. Infect Immun 80: 228–233.
85. Gaddy JA, Arivett BA, McConnell MJ, López-Rojas R, Pachón J, et al. (2012) Role of acinetobactin-mediated iron acquisition functions in the interaction of *Acinetobacter baumannii* strain ATCC 19606T with human lung epithelial cells, *Galleria mellonella* caterpillars, and mice. Infect Immun 80: 1015–1024.
86. Russo TA, Luke NR, Beanan JM, Olson R, Sauberan SL, et al. (2010) The K1 capsular polysaccharide of A*cinetobacter baumannii* strain 307-0294 is a major virulence factor. Infect Immun 78: 3993–4000.
87. Eijkelkamp B, Hassan K, Paulsen I, Brown M (2011) Investigation of the human pathogen *Acinetobacter baumannii* under iron limiting conditions. BMC Genomics 12: 126.
88. Camarena L, Bruno V, Euskirchen G, Poggio S, Snyder M (2010) Molecular mechanisms of ethanol-induced pathogenesis revealed by RNA-sequencing. PLoS Pathog 6: e1000834.
89. Jacobs AC, Hood I, Boyd KL, Olson PD, Morrison JM, et al. (2010) Inactivation of phospholipase D diminishes *Acinetobacter baumannii* pathogenesis. Infect Immun 78: 1952–1962.
90. Johnson DA, Tetu SG, Phillippy K, Chen J, Ren Q, et al. (2008) High-throughput phenotypic characterization of *Pseudomonas aeruginosa* membrane transport genes. PLoS Genet 4: e1000211.
91. Bernards AT, Dijkshoorn L, Vandertoorn J, Bochner BR, Vanboven CPA (1995) Phenotypic characterization of *Acinetobacter* strains of 13 DNA-DNA hybridization groups by means of the Biolog system. J Med Microbiol 42: 113–119.
92. Baumann P, Doudoroff M, Stanier RY (1968) A study of the Moraxella group. II. oxidative-negative species (genus *Acinetobacter*). J Bacteriol 95: 1520–1541.
93. Nemec A, Krizova L, Maixnerova M, van der Reijden TJK, Deschaght P, et al. (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.
94. Peekhaus N, Conway T (1998) What's for dinner?: Entner-Doudoroff metabolism in *Escherichia coli*. J Bacteriol 180: 3495–3502.
95. Müller RH, Babel W (1986) Glucose as an energy donor in acetate growing *Acinetobacter calcoaceticus*. Arch Microbiol 144: 62–66.
96. Lee-Peng FC, Hermodson MA, Kohlhaw GB (1979) Transaminase B from *Escherichia coli*: quaternary structure, amino-terminal sequence, substrate specificity, and absence of a separate valine-alpha-ketoglutarate activity. J Bacteriol 139: 339–345.
97. Powell JT, Morrison JF (1978) Role of the *Escherichia coli* aromatic amino acid aminotransferase in leucine biosynthesis. J Bacteriol 136: 1–4.
98. 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: 2175–2176.
99. Tatusov RL, Koonin EV, Lipman DJ (1997) A genomic perspective on protein families. Science 278: 631–637.
100. Stothard P, Wishart DS (2005) Circular genome visualization and exploration using CGView. Bioinformatics 21: 537–539.
101. Stover BC, Muller KF (2010) TreeGraph 2: combining and visualizing evidence from different phylogenetic analyses. BMC Bioinformatics 11: 7.
102. Carver TJ, Rutherford KM, Berriman M, Rajandream MA, Barrell BG, et al. (2005) ACT: the Artemis Comparison Tool. Bioinformatics 21: 3422–3423.
103. Darling AE, Mau B, Perna NT (2010) progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. PLoS One 5: e11147.

2.2 Supplementary information

Supplementary information is available online at:
http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0058628

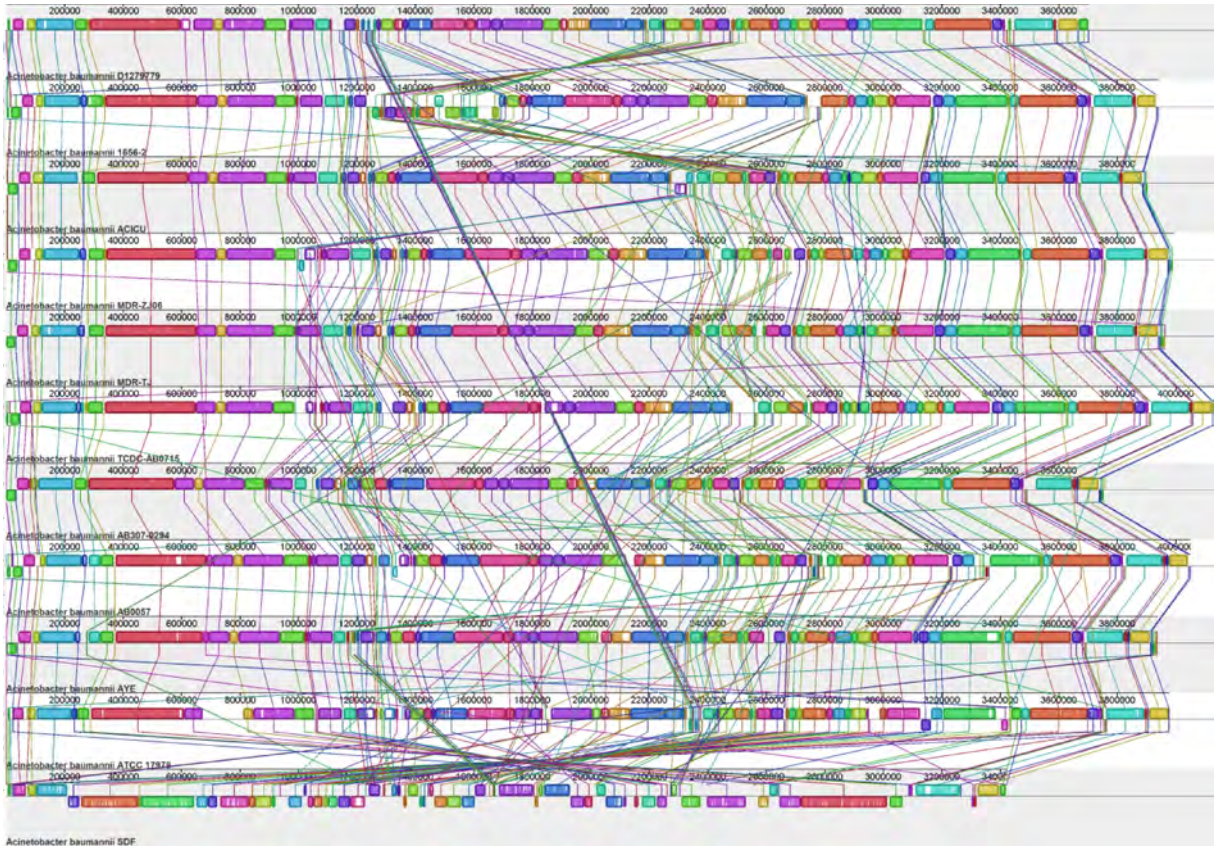


Figure S1: Synteny of *Acinetobacter baumannii*. Chromosomal alignments of the *A. baumannii* D1279779 genome against the ten currently complete *A. baumannii* genomes were generated using progressive MAUVE [237]. Regions of significant synteny between the strains are shown as coloured blocks and unshared regions are seen as white gaps.

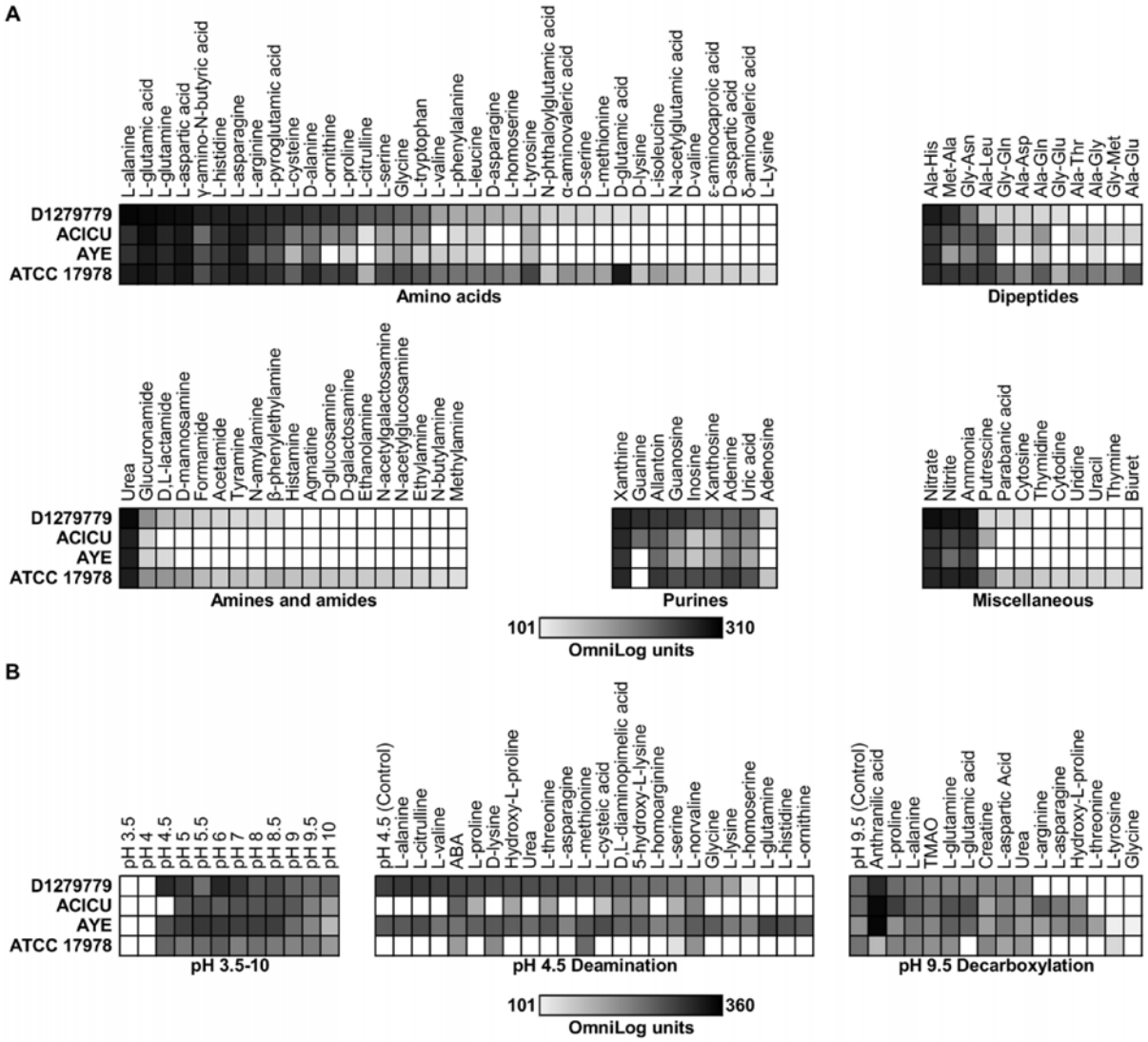


Figure S2: Phenotypic analysis of nitrogen utilisation and pH stress tolerance. Strengths of nitrogen utilisation (A) and the pH tolerance phenotypes (B) of *A. baumannii* strains D1279779, ACICU, AYE and ATCC 17978 were determined were determined using Biolog Phenotype Microarray plates PM3 and PM10, respectively. The maximal kinetic curve height was expressed as a greyscale ranging from 101 OmniLog units (light grey) to 310 and 360 OmniLog units (black) for nitrogen and pH tolerance phenotypes, respectively. Phenotypes are arranged from strongest to weakest relative to *A. baumannii* D1279779. Phenotypes <115 OmniLog units for nitrogen phenotypes and <101 OmniLog units for pH tolerance phenotypes were considered negative phenotypes and are represented in white.

Table S1: Raw phenotype microarray data. The maximal kinetic curve height for all phenotypes obtained from plates PM1-3 (carbon and nitrogen utilisation) and PM9-20 (osmotolerance, pH tolerance and antimicrobial exposure) expressed in OmniLog units.

			Maximum Height (AOUs)				M9 Testing		Notes
Plate	Location	Chemical	D1279779	ACICU	AYE	ATCC 17978	D1279779	ACICU	
PM01	A01	Negative Control	76	45	54	54			
PM01	A02	L-Arabinose	305	266	253	304	+	+	
PM01	A03	N-Acetyl-D-Glucosamine	66	37	47	48			
PM01	A04	D-Saccharic acid	65	34	38	307			
PM01	A05	Succinic acid	278	204	279	301			
PM01	A06	D-Galactose	81	68	82	79			
PM01	A07	L-Aspartic acid	302	282	270	310			
PM01	A08	L-Proline	299	272	49	305			
PM01	A09	D-Alanine	295	250	265	289			
PM01	A10	D-Trehalose	66	38	47	56			
PM01	A11	D-Mannose	90	77	95	81			
PM01	A12	Dulcitol	84	52	68	72			
PM01	B01	D-Serine	73	57	57	71			
PM01	B02	D-Sorbitol	64	37	40	55			
PM01	B03	Glycerol	50	32	36	48			
PM01	B04	L-Fucose	62	35	40	52			
PM01	B05	D-Glucuronic acid	56	31	38	52			
PM01	B06	D-Gluconic acid	160	54	60	269	-	-	
PM01	B07	DL-a-Glycerol Phosphate	68	39	45	57			
PM01	B08	D-Xylose	297	282	287	293	+	+	
PM01	B09	L-Lactic acid	278	261	263	274			
PM01	B10	Formic acid	86	73	70	113			
PM01	B11	D-Mannitol	76	47	56	58			
PM01	B12	L-Glutamic acid	292	260	265	300			
PM01	C01	D-Glucose-6-Phosphate	75	47	53	58			
PM01	C02	D-Galactonic acid-g-Lactone	57	31	35	47			
PM01	C03	DL-Malic acid	277	251	256	288			
PM01	C04	D-Ribose	297	270	88	296			
PM01	C05	Tween 20	304	265	271	297			
PM01	C06	L-Rhamnose	46	28	36	51			
PM01	C07	D-Fructose	71	33	47	67			
PM01	C08	Acetic acid	271	265	243	292	+	+	
PM01	C09	a-D-Glucose	247	137	156	276	-	-	
PM01	C10	Maltose	67	41	45	57			
PM01	C11	D-Melibiose	86	61	71	75			
PM01	C12	Thymidine	88	56	62	66			
PM01	D01	L-Asparagine	292	273	272	283			
PM01	D02	D-Aspartic acid	73	40	37	58			
PM01	D03	D-Glucosaminic acid	67	35	36	49			
PM01	D04	1,2-Propanediol	64	36	42	57			
PM01	D05	Tween 40	294	265	269	288			
PM01	D06	a-Ketoglutaric acid	274	262	260	294			
PM01	D07	a-Ketobutyric acid	249	249	227	259			
PM01	D08	a-Methyl-D-Galactoside	62	33	43	66			
PM01	D09	a-D-Lactose	75	42	52	60			
PM01	D10	Lactulose	67	41	49	56			
PM01	D11	Sucrose	70	50	57	57			
PM01	D12	Uridine	88	68	68	72			
PM01	E01	L-Glutamine	288	281	278	285			
PM01	E02	m-Tartaric acid	60	35	34	58			
PM01	E03	D-Glucose-1-Phosphate	57	33	39	61			
PM01	E04	D-Fructose-6-Phosphate	75	47	45	65			
PM01	E05	Tween 80	288	263	267	276	+	+	
PM01	E06	a-Hydroxyglutaric acid-g-Lactone	82	46	52	69			
PM01	E07	a-Hydroxybutyric acid	237	217	216	252			
PM01	E08	b-Methyl-D-Glucoside	69	37	43	63			
PM01	E09	Adonitol	69	33	40	60			
PM01	E10	Maltotriose	81	47	54	61			
PM01	E11	2`-Deoxyadenosine	66	41	44	52			
PM01	E12	Adenosine	86	64	67	73			
PM01	F01	Gly-Asp	91	62	67	71			
PM01	F02	Citric acid	264	255	247	263	+	+	
PM01	F03	m-Inositol	70	36	41	57			
PM01	F04	D-Threonine	86	39	51	88			
PM01	F05	Fumaric acid	281	279	276	293	+	+	
PM01	F06	Bromosuccinic acid	220	183	167	279	+	+	
PM01	F07	Propionic acid	267	292	263	285			
PM01	F08	Mucic acid	76	47	50	288			
PM01	F09	Glycolic acid	65	39	43	56			
PM01	F10	Glyoxylic acid	64	47	55	42			
PM01	F11	D-Cellobiose	85	58	67	69			
PM01	F12	Inosine	94	64	81	73			
PM01	G01	Gly-Glu	81	68	68	87			
PM01	G02	Tricarballylic acid	274	248	246	285			
PM01	G03	L-Serine	145	92	71	157			
PM01	G04	L-Threonine	240	136	141	234			
PM01	G05	L-Alanine	284	261	261	289			
PM01	G06	Ala-Gly	75	104	79	100			
PM01	G07	Acetoacetic acid	75	52	58	53			

			Maximum Height (AOUs)				M9 Testing		Notes
Plate	Location	Chemical	D1279779	ACICU	AYE	ATCC 17978	D1279779	ACICU	
PM01	G08	N-Acetyl-D-Mannosamine	85	60	64	65			
PM01	G09	Mono-Methylsuccinate	152	97	116	176			
PM01	G10	Methylpyruvate	273	266	253	280			
PM01	G11	D-Malic acid	260	251	266	278			
PM01	G12	L-Malic acid	279	257	247	296			
PM01	H01	Gly-Pro	105	91	75	129			
PM01	H02	p-Hydroxyphenyl Acetic acid	70	60	59	53			
PM01	H03	m-Hydroxyphenyl Acetic acid	82	62	62	66			
PM01	H04	Tyramine	90	69	76	77			
PM01	H05	D-Psicose	100	69	74	77			
PM01	H06	L-Lyxose	96	86	95	72			
PM01	H07	Glucuronamide	99	77	90	198			
PM01	H08	Pyruvic acid	278	294	256	296			
PM01	H09	L-Galactonic acid-g-Lactone	93	62	71	80			
PM01	H10	D-Galacturonic acid	97	77	80	79			
PM01	H11	Phenylethylamine	89	73	81	61			
PM01	H12	2-Aminoethanol	103	79	83	88			
PM02A	A01	Negative Control	75	46	63	67			
PM02A	A02	Chondroitin Sulfate C	53	38	45	55			
PM02A	A03	a-Cyclodextrin	69	38	53	65			
PM02A	A04	b-Cyclodextrin	63	41	53	64			
PM02A	A05	g-Cyclodextrin	70	41	56	57			
PM02A	A06	Dextrin	152	103	135	182			
PM02A	A07	Gelatin	71	42	51	76			
PM02A	A08	Glycogen	81	52	62	84			
PM02A	A09	Inulin	75	46	59	73			
PM02A	A10	Laminarin	79	54	66	86			
PM02A	A11	Mannan	88	54	65	96			
PM02A	A12	Pectin	104	85	90	118			
PM02A	B01	N-Acetyl-D-Galactosamine	69	39	53	75			
PM02A	B02	N-Acetyl-Neuraminic acid	55	29	39	54			
PM02A	B03	b-D-Allose	89	65	77	79			
PM02A	B04	Amygdalin	58	30	43	61			
PM02A	B05	D-Arabinose	56	33	40	56			
PM02A	B06	D-Arabitol	61	29	42	58			
PM02A	B07	L-Arabitol	67	30	41	59			
PM02A	B08	Arbutin	59	30	43	61			
PM02A	B09	2-Deoxy-D-Ribose	56	46	52	53			
PM02A	B10	i-Erythritol	68	49	58	83			
PM02A	B11	D-Fucose	102	86	97	151			
PM02A	B12	3-O-b-D-Galactopyranosyl-D-Arabinos	66	74	75	78			
PM02A	C01	Gentiobiose	89	51	78	90			
PM02A	C02	L-Glucose	59	29	45	52			
PM02A	C03	D-Lactitol	61	27	40	58			
PM02A	C04	D-Melezitose	57	22	35	63			
PM02A	C05	Maltitol	56	21	33	66			
PM02A	C06	a-Methyl-D-Glucoside	58	24	37	78			
PM02A	C07	b-Methyl-D-Galactoside	54	26	33	77			
PM02A	C08	3-Methylglucose	68	34	52	85			
PM02A	C09	b-Methyl-D-Glucuronic acid	59	33	40	80			
PM02A	C10	a-Methyl-D-Mannoside	56	32	41	81			
PM02A	C11	b-Methyl-D-Xyloside	61	44	52	101			
PM02A	C12	Palatinose	86	59	68	99			
PM02A	D01	D-Raffinose	70	41	54	82			
PM02A	D02	Salicin	59	30	38	68			
PM02A	D03	Sedoheptulosan	61	23	34	69			
PM02A	D04	L-Sorbose	60	23	39	70			
PM02A	D05	Stachyose	49	21	34	66			
PM02A	D06	D-Tagatose	67	28	41	77			
PM02A	D07	Turanose	69	35	42	79			
PM02A	D08	Xylitol	60	28	38	74			
PM02A	D09	N-Acetyl-D-Glucosaminitol	64	34	44	94			
PM02A	D10	g-Amino-N-Butyric acid	282	269	281	295			
PM02A	D11	d-Amino Valeric acid	64	40	62	82			
PM02A	D12	Butyric acid	273	274	259	293			
PM02A	E01	Capric acid	125	53	52	64			
PM02A	E02	Caproic acid	258	286	267	263			
PM02A	E03	Citraconic acid	49	20	259	62			
PM02A	E04	Citramalic acid	66	23	221	74			
PM02A	E05	D-Glucosamine	120	90	95	164			
PM02A	E06	2-Hydroxybenzoic acid	22	15	22	22			
PM02A	E07	4-Hydroxybenzoic acid	44	290	268	287	-	+	
PM02A	E08	b-Hydroxybutyric acid	278	261	252	276			
PM02A	E09	g-Hydroxybutyric acid	68	41	50	174			
PM02A	E10	a-Keto-Valeric acid	273	264	269	278			
PM02A	E11	Itaconic acid	42	39	42	52			
PM02A	E12	5-Keto-D-Gluconic acid	87	63	74	97			
PM02A	F01	D-Lactic acid Methyl Ester	223	169	184	241			
PM02A	F02	Malonic acid	231	129	161	263			
PM02A	F03	Melibionic acid	60	26	35	80			
PM02A	F04	Oxalic acid	70	24	41	59			

			Maximum Height (AOUs)				M9 Testing		Notes
Plate	Location	Chemical	D1279779	ACICU	AYE	ATCC 17978	D1279779	ACICU	
PM02A	F05	Oxalomalic acid	71	38	44	71			
PM02A	F06	Quinic acid	273	314	262	295	+	+	
PM02A	F07	D-Ribono-1,4-Lactone	296	312	34	304			
PM02A	F08	Sebacic acid	81	289	269	92			
PM02A	F09	Sorbic acid	264	273	266	271			
PM02A	F10	Succinamic acid	211	144	115	269			
PM02A	F11	D-Tartaric acid	85	51	61	89			
PM02A	F12	L-Tartaric acid	80	52	59	92			
PM02A	G01	Acetamide	60	34	46	79			
PM02A	G02	L-Alaninamide	190	52	72	203			
PM02A	G03	N-Acetyl-L-Glutamic acid	56	23	33	77			
PM02A	G04	L-Arginine	275	321	298	281			
PM02A	G05	Glycine	36	17	23	61			
PM02A	G06	L-Histidine	286	273	263	299	+	+	
PM02A	G07	L-Homoserine	46	23	30	81			
PM02A	G08	Hydroxy-L-Proline	138	249	264	300			
PM02A	G09	L-Isoleucine	216	87	55	179			
PM02A	G10	L-Leucine	259	146	199	284			
PM02A	G11	L-Lysine	72	49	54	82			
PM02A	G12	L-Methionine	57	49	55	109			
PM02A	H01	L-Ornithine	279	310	40	304			
PM02A	H02	L-Phenylalanine	292	317	279	284			
PM02A	H03	L-Pyroglutamic acid	278	306	290	288	+	+	
PM02A	H04	L-Valine	115	38	39	96			
PM02A	H05	D,L-Carnitine	265	56	33	157	+	-	Both strains could utilise L-carnitine
PM02A	H06	sec-Butylamine	66	34	44	93			
PM02A	H07	D,L-Octopamine	53	25	33	82			
PM02A	H08	Putrescine	112	280	38	300	-	+	
PM02A	H09	Dihydroxyacetone	91	73	84	114			
PM02A	H10	2,3-Butanediol	181	135	153	207			
PM02A	H11	2,3-Butanedione	60	50	54	135			
PM02A	H12	3-Hydroxy-2-butanone	85	65	72	112			
PM03B	A01	Negative Control	92	114	106	155			
PM03B	A02	Ammonia	286	262	240	281	+	+	
PM03B	A03	Nitrite	289	247	213	278			
PM03B	A04	Nitrate	298	264	262	275			
PM03B	A05	Urea	301	281	281	283			
PM03B	A06	Biuret	88	87	94	121			
PM03B	A07	L-Alanine	305	267	266	287			
PM03B	A08	L-Arginine	270	260	224	269			
PM03B	A09	L-Asparagine	277	279	278	279			
PM03B	A10	L-Aspartic acid	295	285	288	289			
PM03B	A11	L-Cysteine	265	199	149	221			
PM03B	A12	L-Glutamic acid	301	296	285	291			
PM03B	B01	L-Glutamine	298	275	272	277			
PM03B	B02	Glycine	216	159	168	244			
PM03B	B03	L-Histidine	278	264	258	268			
PM03B	B04	L-Isoleucine	114	81	101	172			
PM03B	B05	L-Leucine	159	129	124	201			
PM03B	B06	L-Lysine	111	90	102	118			
PM03B	B07	L-Methionine	119	103	98	166			
PM03B	B08	L-Phenylalanine	160	120	131	205			
PM03B	B09	L-Proline	243	191	127	253			
PM03B	B10	L-Serine	227	163	148	239			
PM03B	B11	L-Threonine	63	84	70	99			
PM03B	B12	L-Tryptophan	203	164	129	230			
PM03B	C01	L-Tyrosine	145	156	148	248			
PM03B	C02	L-Valine	166	95	132	206			
PM03B	C03	D-Alanine	259	207	206	266			
PM03B	C04	D-Asparagine	156	81	101	204			
PM03B	C05	D-Aspartic acid	101	73	73	130			
PM03B	C06	D-Glutamic acid	116	95	105	288			
PM03B	C07	D-Lysine	116	95	96	141			
PM03B	C08	D-Serine	125	114	93	170			
PM03B	C09	D-Valine	114	99	101	137			
PM03B	C10	L-Citrulline	229	116	98	154			
PM03B	C11	L-Homoserine	151	105	107	189			
PM03B	C12	L-Ornithine	259	189	101	248			
PM03B	D01	N-Acetyl-L-Glutamic acid	100	112	83	155			
PM03B	D02	N-Phthaloyl-L-Glutamic acid	128	89	107	139			
PM03B	D03	L-Pyroglutamic acid	267	246	227	259			
PM03B	D04	Hydroxylamine	52	58	22	60			
PM03B	D05	Methylamine	99	86	88	116			
PM03B	D06	N-Amylamine	121	93	88	134			
PM03B	D07	N-Butylamine	100	80	97	119			
PM03B	D08	Ethylamine	96	87	96	124			
PM03B	D09	Ethanolamine	114	101	96	133			
PM03B	D10	Ethylenediamine	67	63	48	84			
PM03B	D11	Putrescine	124	160	78	196			
PM03B	D12	Agmatine	115	110	90	137			
PM03B	E01	Histamine	106	100	96	140			

			Maximum Height (AOUs)				M9 Testing		Notes
Plate	Location	Chemical	D1279779	ACICU	AYE	ATCC 17978	D1279779	ACICU	
PM03B	E02	b-Phenylethylamine	117	93	96	150			
PM03B	E03	Tyramine	122	93	100	143			
PM03B	E04	Acetamide	124	100	100	134			
PM03B	E05	Formamide	130	99	110	146			
PM03B	E06	Glucuronamide	183	129	129	182			
PM03B	E07	DL-Lactamide	146	114	120	176			
PM03B	E08	D-Glucosamine	108	92	96	135			
PM03B	E09	D-Galactosamine	107	94	99	134			
PM03B	E10	D-Mannosamine	136	106	112	169			
PM03B	E11	N-Acetyl-D-Glucosamine	101	100	87	128			
PM03B	E12	N-Acetyl-D-Galactosamine	113	99	80	133			
PM03B	F01	N-Acetyl-D-Mannosamine	91	103	81	115			
PM03B	F02	Adenine	222	193	196	254			
PM03B	F03	Adenosine	127	96	94	138			
PM03B	F04	Cytidine	107	90	96	127			
PM03B	F05	Cytosine	119	95	98	146			
PM03B	F06	Guanine	267	212	13	73			
PM03B	F07	Guanosine	256	180	159	238			
PM03B	F08	Thymine	103	89	86	122			
PM03B	F09	Thymidine	111	96	97	133			
PM03B	F10	Uracil	102	100	80	125			
PM03B	F11	Uridine	101	95	90	127			
PM03B	F12	Inosine	237	138	138	241			
PM03B	G01	Xanthine	280	273	262	274			
PM03B	G02	Xanthosine	223	145	154	238			
PM03B	G03	Uric acid	221	183	161	235			
PM03B	G04	Alloxan	60	66	16	76			
PM03B	G05	Allantoin	259	225	214	260			
PM03B	G06	Parabanic acid	121	109	78	137			
PM03B	G07	DL-a-Amino-N-Butyric acid	106	72	79	103			
PM03B	G08	g-Amino-N-Butyric acid	278	212	236	238			
PM03B	G09	e-Amino-N-Caproic acid	104	95	75	131			
PM03B	G10	DL-a-Amino-Caprylic acid	70	71	23	75			
PM03B	G11	d-Amino-N-Valeric acid	98	102	76	124			
PM03B	G12	a-Amino-N-Valeric acid	127	102	68	180			
PM03B	H01	Ala-Asp	125	136	81	181			
PM03B	H02	Ala-Gln	123	176	133	227			
PM03B	H03	Ala-Glu	113	124	100	224			
PM03B	H04	Ala-Gly	111	133	117	198			
PM03B	H05	Ala-His	285	260	263	269			
PM03B	H06	Ala-Leu	137	231	227	254			
PM03B	H07	Ala-Thr	113	139	103	203			
PM03B	H08	Gly-Asn	209	223	192	254			
PM03B	H09	Gly-Gln	128	140	107	214			
PM03B	H10	Gly-Glu	117	111	91	154			
PM03B	H11	Gly-Met	111	125	85	184			
PM03B	H12	Met-Ala	270	231	170	260			
PM09	A01	1% NaCl	288	297	282	284			
PM09	A02	2% NaCl	255	289	268	293			
PM09	A03	3% NaCl	253	146	154	223			
PM09	A04	4% NaCl	162	52	72	203			
PM09	A05	5% NaCl	16	43	41	61			
PM09	A06	5.5% NaCl	15	34	42	51			
PM09	A07	6% NaCl	43	62	28	67			
PM09	A08	6.5% NaCl	36	32	47	57			
PM09	A09	7% NaCl	66	40	31	68			
PM09	A10	8% NaCl	26	71	66	76			
PM09	A11	9% NaCl	42	44	57	87			
PM09	A12	10% NaCl	71	65	77	101			
PM09	B01	6% NaCl	61	43	29	74			
PM09	B02	6% NaCl + Betaine	24	36	22	79			
PM09	B03	6% NaCl + NN Dimethyl Glycine	63	24	54	64			
PM09	B04	6% NaCl + Sarcosine	16	26	31	73			
PM09	B05	6% NaCl + Dimethyl Sulphonyl Propior	55	29	20	86			
PM09	B06	6% NaCl + MOPS	22	57	19	46			
PM09	B07	6% NaCl + Ectoine	23	32	18	47			
PM09	B08	6% NaCl + Choline	55	21	16	51			
PM09	B09	6% NaCl + Phosphorylcholine	59	60	33	84			
PM09	B10	6% NaCl + Creatine	43	38	36	82			
PM09	B11	6% NaCl + Creatinine	88	78	47	80			
PM09	B12	6% NaCl + L-Carnitine	99	87	56	87			
PM09	C01	6% NaCl + KCl	80	64	53	93			
PM09	C02	6% NaCl + L-Proline	52	46	38	81			
PM09	C03	6% NaCl + N-Acetyl-L-Glutamine	49	30	20	68			
PM09	C04	6% NaCl + b-Glutamic acid	39	37	24	45			
PM09	C05	6% NaCl + g-Amino-N-Butyric acid	51	38	32	60			
PM09	C06	6% NaCl + Glutathione	26	48	33	40			
PM09	C07	6% NaCl + Glycerol	55	45	18	52			
PM09	C08	6% NaCl + Trehalose	17	23	17	62			
PM09	C09	6% NaCl + Trimethylamine-N-Oxide	69	25	20	56			
PM09	C10	6% NaCl + Trimethylamine	39	40	32	56			

			Maximum Height (AOUs)				M9 Testing		Notes
Plate	Location	Chemical	D1279779	ACICU	AYE	ATCC 17978	D1279779	ACICU	
PM09	C11	6% NaCl + Octopine	71	35	30	87			
PM09	C12	6% NaCl + Trigonelline	90	66	44	99			
PM09	D01	3% Potassium Chloride	263	269	227	229			
PM09	D02	4% Potassium chloride	196	222	153	226			
PM09	D03	5% Potassium Chloride	167	69	104	168			
PM09	D04	6% Potassium chloride	155	40	63	142			
PM09	D05	2% Sodium Sulfate	273	269	274	257			
PM09	D06	3% Sodium Sulfate	275	270	202	244			
PM09	D07	4% Sodium Sulfate	241	268	208	249			
PM09	D08	5% Sodium Sulfate	126	148	121	241			
PM09	D09	5% Ethylene Glycol	290	294	287	270			
PM09	D10	10% Ethylene Glycol	302	284	284	272			
PM09	D11	15% Ethylene Glycol	281	273	286	243			
PM09	D12	20% Ethylene Glycol	264	272	283	239			
PM09	E01	1% Sodium Formate	294	302	294	300			
PM09	E02	2% Sodium Formate	278	276	255	271			
PM09	E03	3% Sodium Formate	254	42	37	250			
PM09	E04	4% Sodium Formate	269	30	16	127			
PM09	E05	5% Sodium Formate	13	29	51	75			
PM09	E06	6% Sodium Formate	55	25	13	56			
PM09	E07	2% Urea	287	300	291	259			
PM09	E08	3% Urea	287	304	293	258			
PM09	E09	4% Urea	264	285	287	233			
PM09	E10	5% Urea	86	30	149	242			
PM09	E11	6% Urea	33	39	38	70			
PM09	E12	7% Urea	54	48	47	87			
PM09	F01	1% Sodium Lactate	300	300	307	269			
PM09	F02	2% Sodium Lactate	282	302	13	247			
PM09	F03	3% Sodium Lactate	287	292	11	223			
PM09	F04	4% Sodium Lactate	296	235	42	197			
PM09	F05	5% Sodium Lactate	18	12	11	241			
PM09	F06	6% Sodium Lactate	27	14	11	76			
PM09	F07	7% Sodium Lactate	17	16	11	84			
PM09	F08	8% Sodium Lactate	28	20	12	122			
PM09	F09	9% Sodium Lactate	25	29	21	92			
PM09	F10	10% Sodium Lactate	225	24	25	112			
PM09	F11	11% Sodium Lactate	144	35	34	84			
PM09	F12	12% Sodium Lactate	49	41	39	115			
PM09	G01	20mM Sodium Phosphate pH 7	302	308	307	289			
PM09	G02	50mM Sodium Phosphate pH 7	289	310	297	255			
PM09	G03	100mM Sodium Phosphate pH 7	284	287	278	261			
PM09	G04	200mM Sodium Phosphate pH 7	283	281	239	232			
PM09	G05	20mM Sodium Benzoate pH 5.2	315	316	330	246			
PM09	G06	50mM Sodium Benzoate pH 5.2	38	20	35	71			
PM09	G07	100mM Sodium Benzoate pH 5.2	24	22	11	74			
PM09	G08	200mM Sodium Benzoate pH 5.2	48	27	16	73			
PM09	G09	10mM Ammonium Sulfate pH 8	304	296	299	256			
PM09	G10	20mM Ammonium Sulfate pH 8	294	303	304	252			
PM09	G11	50mM Ammonium Sulfate pH 8	288	296	282	264			
PM09	G12	100mM Ammonium Sulfate pH 8	294	298	286	271			
PM09	H01	10mM Sodium Nitrate	314	317	307	277			
PM09	H02	20mM Sodium Nitrate	301	325	306	275			
PM09	H03	40mM Sodium Nitrate	299	321	299	268			
PM09	H04	60mM Sodium Nitrate	305	312	284	247			
PM09	H05	80mM Sodium Nitrate	289	307	285	251			
PM09	H06	100mM Sodium Nitrate	289	298	286	257			
PM09	H07	10mM Sodium Nitrite	314	322	315	263			
PM09	H08	20mM Sodium Nitrite	319	320	301	254			
PM09	H09	40mM Sodium Nitrite	305	294	269	248			
PM09	H10	60mM Sodium Nitrite	272	243	163	229			
PM09	H11	80mM Sodium Nitrite	222	133	126	185			
PM09	H12	100mM Sodium Nitrite	64	49	49	100			
PM10	A01	pH 3.5	24	46	52	45			
PM10	A02	pH 4	17	40	48	40			
PM10	A03	pH 4.5	334	39	311	302			
PM10	A04	pH 5	326	309	323	291			
PM10	A05	pH 5.5	295	314	328	297			
PM10	A06	pH 6	338	306	326	306			
PM10	A07	pH 7	328	312	323	286			
PM10	A08	pH 8	311	311	324	295			
PM10	A09	pH 8.5	316	312	324	289			
PM10	A10	pH 9	301	309	297	279			
PM10	A11	pH 9.5	297	284	274	287			
PM10	A12	pH 10	301	234	194	244			
PM10	B01	pH 4.5	322	34	307	37			
PM10	B02	pH 4.5 + L-Alanine	327	26	308	28			
PM10	B03	pH 4.5 + L-Arginine	25	20	29	20			
PM10	B04	pH 4.5 + L-Asparagine	312	24	281	25			
PM10	B05	pH 4.5 + L-Aspartic acid	13	22	30	26			
PM10	B06	pH 4.5 + L-Glutamic acid	11	23	30	25			
PM10	B07	pH 4.5 + L-Glutamine	24	28	321	29			

			Maximum Height (AOUs)				M9 Testing		Notes
Plate	Location	Chemical	D1279779	ACICU	AYE	ATCC 17978	D1279779	ACICU	
PM10	B08	pH 4.5 + Glycine	254	22	307	20			
PM10	B09	pH 4.5 + L-Histidine	33	29	313	26			
PM10	B10	pH 4.5 + L-Isoleucine	32	38	61	41			
PM10	B11	pH 4.5 + L-Leucine	56	36	53	41			
PM10	B12	pH 4.5 + L-Lysine	232	47	289	46			
PM10	C01	pH 4.5 + L-Methionine	307	40	308	299			
PM10	C02	pH 4.5 + L-Phenylalanine	17	36	41	36			
PM10	C03	pH 4.5 + L-Proline	317	210	304	38			
PM10	C04	pH 4.5 + L-Serine	291	203	289	128			
PM10	C05	pH 4.5 + L-Threonine	313	231	310	19			
PM10	C06	pH 4.5 + L-Tryptophan	17	22	30	18			
PM10	C07	pH 4.5 + L-Citrulline	327	23	315	17			
PM10	C08	pH 4.5 + L-Valine	321	23	293	21			
PM10	C09	pH 4.5 + Hydroxy-L-Proline	315	226	306	27			
PM10	C10	pH 4.5 + L-Ornithine	61	23	306	21			
PM10	C11	pH 4.5 + L-Homoarginine	296	33	283	29			
PM10	C12	pH 4.5 + L-Homoserine	101	43	279	40			
PM10	D01	pH 4.5 + Anthranilic acid	39	47	54	50			
PM10	D02	pH 4.5 + L-Norleucine	15	24	32	22			
PM10	D03	pH 4.5 + L-Norvaline	289	281	307	258			
PM10	D04	pH 4.5 + a-Amino-N-Butyric acid	318	301	294	241			
PM10	D05	pH 4.5 + p-Aminobenzoate	30	36	50	37			
PM10	D06	pH 4.5 + L-Cysteic acid	305	177	316	24			
PM10	D07	pH 4.5 + D-Lysine	317	32	309	269			
PM10	D08	pH 4.5 + 5-Hydroxy-L-Lysine	297	251	304	20			
PM10	D09	pH 4.5 + 5-Hydroxy-L-Tryptophan	73	68	94	82			
PM10	D10	pH 4.5 + DL-Diamino-a,e-Pimelic acid	304	267	302	31			
PM10	D11	pH 4.5 + Trimethylamine-N-Oxide	65	38	37	32			
PM10	D12	pH 4.5 + Urea	315	42	298	40			
PM10	E01	pH 9.5	296	298	256	294			
PM10	E02	pH 9.5 + L-Alanine	290	304	292	275			
PM10	E03	pH 9.5 + L-Arginine	16	302	238	17			
PM10	E04	pH 9.5 + L-Asparagine	17	291	272	18			
PM10	E05	pH 9.5 + L-Aspartic acid	271	281	289	227			
PM10	E06	pH 9.5 + L-Glutamic acid	272	313	300	43			
PM10	E07	pH 9.5 + L-Glutamine	278	313	306	265			
PM10	E08	pH 9.5 + Glycine	29	31	103	30			
PM10	E09	pH 9.5 + L-Histidine	22	37	46	30			
PM10	E10	pH 9.5 + L-Isoleucine	32	29	79	28			
PM10	E11	pH 9.5 + L-Leucine	27	39	47	37			
PM10	E12	pH 9.5 + L-Lysine	51	85	86	65			
PM10	F01	pH 9.5 + L-Methionine	25	25	45	24			
PM10	F02	pH 9.5 + L-Phenylalanine	20	17	22	17			
PM10	F03	pH 9.5 + L-Proline	298	324	258	282			
PM10	F04	pH 9.5 + L-Serine	18	17	81	13			
PM10	F05	pH 9.5 + L-Threonine	15	76	235	11			
PM10	F06	pH 9.5 + L-Tryptophan	30	27	37	49			
PM10	F07	pH 9.5 + L-Tyrosine	51	55	108	144			
PM10	F08	pH 9.5 + L-Valine	24	18	50	15			
PM10	F09	pH 9.5 + Hydroxy-L-Proline	92	262	289	27			
PM10	F10	pH 9.5 + L-Ornithine	37	28	35	27			
PM10	F11	pH 9.5 + L-Homoarginine	32	42	44	39			
PM10	F12	pH 9.5 + L-Homoserine	44	43	44	43			
PM10	G01	pH 9.5 + Anthranilic acid	336	355	357	194			
PM10	G02	pH 9.5 + L-Norleucine	34	30	37	26			
PM10	G03	pH 9.5 + L-Norvaline	19	17	25	17			
PM10	G04	pH 9.5 + Agmatine	29	20	23	39			
PM10	G05	pH 9.5 + Cadaverine	52	42	42	58			
PM10	G06	pH 9.5 + Putrescine	31	26	63	19			
PM10	G07	pH 9.5 + Histamine	32	26	30	23			
PM10	G08	pH 9.5 + b-Phenylethylamine	47	57	40	27			
PM10	G09	pH 9.5 + Tyramine	70	57	46	53			
PM10	G10	pH 9.5 + Creatine	272	228	235	278			
PM10	G11	pH 9.5 + Trimethylamine-N-Oxide	281	294	277	260			
PM10	G12	pH 9.5 + Urea	270	250	243	248			
PM10	H01	X-Caprylate	298	320	294	309			
PM10	H02	X-a-D-Glucoside	324	327	329	311			
PM10	H03	X-b-D-Glucoside	313	325	323	303			
PM10	H04	X-a-D-Galactoside	320	308	330	305			
PM10	H05	X-b-D-Galactoside	327	330	334	308			
PM10	H06	X-a-D-Glucuronide	321	326	326	308			
PM10	H07	X-b-D-Glucuronide	313	312	321	307			
PM10	H08	X-b-D-Glucosaminide	308	300	310	290			
PM10	H09	X-b-D-Galactosaminide	317	284	306	303			
PM10	H10	X-a-D-Mannoside	318	303	315	299			
PM10	H11	X-PO4	321	316	311	294			
PM10	H12	X-SO4	313	303	258	309			

Chapter 3: A novel and mosaic family of mobile genomic islands inserted in the tRNA-dihydrouridine synthase A (*dusA*) gene

3.1 Non-technical summary

Bioinformatic analysis examining the presence of laterally transferred regions in the genomes of several *Acinetobacter baumannii* and *Pseudomonas* sp. organisms, previously sequenced by our research group, revealed the sporadic presence of a homologous integrase gene associated with the 5' end of a gene known as *dusA*. The *dusA* gene encodes the enzyme tRNA-dihydrouridine synthase A, which is involved in the postrascriptional reduction of uridine to dihydrouridine in tRNA. Follow-up bioinformatic analyses revealed that *dusA*-associated integrases were present in the genomes of over 200 bacterial organisms, and was found to be a component of genomic regions predicted to be genomic islands. These genomic islands were predicted to be capable of excision from the genome as a circular intermediates, which was confirmed by PCR and sequencing. The information derived from this was used to map the termini of the *dusA*-associated genomic islands and revealed that not only was *dusA* the target of these genomic islands, but also that these islands radically varied in both size and gene content. The only common feature encoded by this collective of genomic islands was found to be the *dusA*-associated integrase. On this basis, this family of genomic islands was dubbed the ***d****us****A***-specific **M**obile **G**enetic **E**lements (DAMGEs).

A novel and mosaic family of mobile genomic islands inserted in the tRNA-dihydrouridine synthase A (*dusA*) gene

Daniel N. Farrugia, Liam D. H. Elbourne, Bridget C. Mabbutt and Ian T. Paulsen*

Department of Chemistry and Biomolecular Sciences, Macquarie University, Sydney, NSW 2109, Australia

* To whom correspondence should be addressed

Ian T. Paulsen

Department of Chemistry and Biomolecular Sciences

Macquarie University

University Avenue

North Ryde, NSW 2109, Australia

Tel: +61-2-9850-8152, Fax: +61-2-9850-8313, Email: ian.paulsen@mq.edu.au

Abstract

Genomic islands have been revealed, through an ever increasing number of prokaryotic genome sequences, to be one of the foremost agents of prokaryotic genome plasticity. Genomic islands integrate into chromosomal loci such as transfer RNA genes and protein coding genes, whilst retaining various cargo genes that potentially bestow novel functions to the host organism. A gene encoding a putative integrase was identified in a similar genetic context in the genomes of over 200 bacterial organisms. This integrase was discovered to be a component of a novel family of genomic islands that appeared to preferentially insert within the 5' end of the *dusA* gene, which encoded the tRNA-dihydrouridine synthase A enzyme responsible for the postranscriptional reduction of uridine to dihydrouridine in tRNA. The ability of these *dusA*-associated islands to excise from the genome as circularised intermediates, resulting in restoration of the wild-type *dusA* sequence, was confirmed. This family of genomic islands were found to consist of a mosaic architecture, varying radically in size and gene content, and sharing only the integrase gene. On this basis, this family of genomic islands were dubbed the ***dusA***-specific **M**obile **G**enetic **E**lements (DAMGEs).

Introduction

Since the discovery that the genomes of *Escherichia coli* K12 and O157:H7 differed by more than a megabase of DNA (1), regions of genomic variability have consistently been observed in the genomes of prokaryotic organisms. The acquisition of genes through lateral gene transfer has probably been one of the main drivers of this variability. These variable regions are currently classified by three overlapping terminologies. Mobile genetic elements (MGEs) are a collective of well characterised genetic entities that are capable of intracellular or intercellular translocation, such as prophages and insertion sequences (2). Genomic islands (GEIs) are discreet gene clusters containing features suggestive of lateral gene transfer, such as atypical nucleotide content and mobilisation genes, and includes some MGE classes (3,4). Regions of genomic plasticity (RGPs) are genomic regions that are variable compared to related organisms, without any assumption regarding their evolutionary or genetic basis, which includes both laterally acquired regions and plasticity arising through other recombinative mechanisms (5,6).

GEIs carry genes that may encode advantageous traits, including antimicrobial resistance, virulence and metabolic pathways (7). They also frequently encode a high number of hypothetical proteins of unknown function, suggesting that GEIs act as reservoirs of genetic novelty (8). GEIs are classically inserted within a narrow assortment of transfer RNA and transfer-messenger RNA genes (9-11), but are also known to integrate into various protein coding genes (6,12).

The accelerating release of bacterial genome sequences has fuelled the discovery of novel GEI variants. However, pathogens constitute the bulk of available bacterial genome sequences, particularly those of the Proteobacterial lineage, which is reflective of the GEIs and MGEs that are discovered (2,13). Some of the most prominent GEI families are resistance

or pathogenicity islands, limited to a single genus or species, such as the *Acinetobacter baumannii* antibiotic resistance island (Tn*AbaR*) (14), *Salmonella* genomic island 1 (SGI1) (15), *Vibrio cholera* SXT (16), and the *Yersinia* high-pathogenicity island (HPI) (17).

Analysis conducted on several *Acinetobacter baumannii* and *Pseudomonas* sp. genomes, previously sequenced by our research group (18-21), revealed the presence of a gene encoding a putative integrase adjacent to the 5' end of the *dusA* gene of these organisms. *dusA* encodes the tRNA-dihydrouridine synthase A enzyme and is a member of the dihydrouridine synthase (DUS) family, a highly conserved enzyme family found in many prokaryotes and eukaryotes (22). Dus catalyses the posttranscriptional reduction of uridine in the D-loop of tRNA to 5,6-dihydrouridine (23), one of the most abundant modified nucleobases in prokaryote and eukaryote tRNAs (24). The genome of the model prokaryote, *E. coli*, contains the *dusA*, *dusB*, and *dusC* genes, which encode three DUS family proteins of differing tRNA specificities (25). Concurrent deletion of all three *dus* genes in *E. coli* had revealed these genes to be dispensable, despite the abolition of all dihydrouridine in this organism (25).

In this manuscript, we outline a novel family of GEIs typically found inserted within the *dusA* gene in over two hundred sequenced Proteobacterial organisms. This GEI family encodes a core integrase gene and a highly mosaic cargo gene pool, encoding various functions including bacteriophage lysogeny, heavy metal resistance, metabolic augmentation and conjugative transfer, in addition to numerous uncharacterised hypothetical proteins. Excision of these GEIs was demonstrated through PCR amplification of both the circularised intermediates and the remnant chromosomal junction. Sequencing of these amplicons, and subsequent bioinformatic analyses, had identified a conserved sequence motif likely involved in the targeting and integration of these GEIs into the *dusA* gene. We propose that these *dusA*-

specific mobile elements form the basis of a large and novel family of MGEs we have dubbed the ***dus*A**-specific **M**obile **G**enetic **E**lements (DAMGEs).

Materials and Methods

Bioinformatic parameters and analyses

BLAST (26) searches to identify integrase homologues were conducted using the default parameters and the NCBI 'non-redundant protein sequences' (nr) database, except the maximum target sequences were set to 20000 and an e-value cutoff of 1e-20 was used. Integrases that were either below this cutoff, or belonging to non-redundant records lacking genome sequences, were excluded from this analysis. Integrases were considered *dusA*-associated if the 5' end of the integrase gene was adjoining the 5' end of the *dusA* gene. However, integrases found not to be *dusA*-associated, but greater than the 1e-20 cuttoff, were included in this analysis. Multiple alignments of protein and nucleotide sequences were carried out using ClustalW (27) and ClustalOmega (28). Neighbour-joining phylogenetic analysis and protein identity/similarity matrices were generated from multiple alignments using MEGA5 (29) and MatGAT (30), respectively. TreeGraph2 (31) and UGENE (32) were routinely used for phylogenetic tree annotation and genome browsing, respectively. The gene content of numerous *dusA*-associated GEIs (Supplementary Table 3.1) were compared reciprocally by BLASTP+ (33) to identify putative orthologues using an e-value cutoff of 10^{-5}. Oligonucleotides used in PCR were designed with the aid of Primer3 (34) with the oligonucleotide length, melting temperature and GC% parameters set to 20-25 nt, 55-60 °C and 30-50%, respectively.

Bacterial strains, culture conditions and genomic DNA extraction

Representative *Acinetobacter* and *Pseudomonas* strains encoding *dusA*-associated GEIs, differing in predicted size and gene content (Table 3.1, Supplementary Table 3.1), were

cultured overnight in Müeller-Hinton broth (MH) (Oxoid) at 37 °C and 25 °C, respectively. Broth microdilution minimal inhibitory concentration (MIC) testing was conducted as described previously (35), to determine the resistance level of *Acinetobacter* and *Pseudomonas* against the antibiotic mitomycin C (MMC) (AG Scientific), with a minimum of three temporally distinct replicates. The MMC MICs of each organism were determined to be 50 $\mu g \cdot ml^{-1}$ for *A. baumannii* strains D1279779 and ACICU, and 0.78125 $\mu g \cdot ml^{-1}$ for *Pseudomonas protegens* Pf-5. Induction of *dusA*-associated GEI excision was modified from a previous protocol for excision of integrative elements (36) as follows: 5 mL of mid-exponential (OD_{600nm} 0.6) *A. baumannii* and *P. protegens* subcultures (in MH) were exposed to subinhibitory concentrations of MMC. *A. baumannii* ACICU and *P. protegens* Pf-5 were treated with 0.5x MIC MMC for up to two hours, while *A. baumannii* D1279779 was treated with 0.75x MIC MMC for one hour. Total genomic DNA from ACICU and Pf-5 was isolated using the PureLink Genomic DNA Mini Kit (Invitrogen), whilst total covalently closed circular DNA (cccDNA) from *A. baumannii* D1279779 was isolated using the Wizard® Plus SV Minipreps DNA Purification System (Promega), as per the manufacturers' protocols.

PCR detection of genomic island excision

PCR amplicons confirming the excision of *dusA*-associated GEIs and restoration of the wild-type *dusA* (and *dusB*) sequence (Figure 3.3) were generated with Platinum *Taq* DNA Polymerase High Fidelity (Invitrogen) at 35 cycles, using the primers listed in Supplementary Table 3.2 and 100 ng of total genomic DNA, or 50 ng of total cccDNA in the case of *A. baumannii* D1279779. The amplicons were purified using the Wizard® SV Gel and PCR Clean-Up System (Promega) per the manufacturer's instructions, and were sequenced bidirectionally using BigDye Terminator v3.1 chemistry dideoxysequencing (Applied Biosciences). All sequencing reactions and purifications were conducted at the Australian Genome Research Facility (AGRF) (Westmead, Australia). The resultant chromatograms

were edited and assembled in ChromasPro (Technelysium Pty. Ltd.) and crosschecked against their respective genome sequences.

Deletion of D1279779_RGP05 and *in trans dusA* complementation

The strains and primers used in the following procedures are listed in Table 3.1 and Supplementary Table 3.2, respectively. Deletion of a 48.3 kb region extending from the original *dusA* 5' end to the *dusA* 3' end of *A. baumannii* D1279779 (Figure 3.3), was deleted per a previous method (37) with several alterations. Mating of *A. baumannii* D1279779 and *E. coli* S17-1 harbouring pJQFRT_05270 was conducted on LB agar at 37 °C for 7 h. Cells were collected and resuspended in 1 ml of 10 mM $MgSO_4$, serially diluted, plated onto CHROMagar *Acinetobacter* (Dutec Diagnostics) containing 10 μg/mL gentamycin, and incubated overnight at 37 °C. In brief, CHROMagar *Acinetobacter* is a proprietary chromogenic agar that is selective for *Acinetobacter* spp. organisms, appearing as bright red colonies, whilst other organisms are mostly inhibited; *E. coli* S17-1 uninhibited by the media appear as metallic blue colonies. D1279779 colonies resistant to gentamycin were screened for integration of pJQFRT_05270 using the primers P1_RGP05_FOR and FRT-SP6R. The D1279779 transformant was subsequently mated with *E. coli* S17-1 harbouring pKFRT/FLP_05691 as above, except transformants were selected on CHROMagar *Acinetobacter* containing 15 μg/mL gentamycin and 30 μg/mL kanamycin. D1279779 colonies resistant to both gentamycin and kanamycin were screened for pKFRT/FLP_05691 integration using the primers FRT-leftF and P4_RGP05_REV. Excision of the DNA sandwiched between the two plasmids and screening of desired deletions were conducted as per the original method (37), except all incubation steps were conducted with LB or LB agar at 37 °C. *In trans* complementation of the resultant Δ*dusA* mutant was accomplished using the *E. coli*/*A. calcoaceticus* shuttle vector pWH1266 (38), containing *dusA* and its native promoter originating from either *A. baumannii* D1279779 (DAI present) or *A. baumannii*

ATCC 17978 (DAI absent), cloned into the BamHI site. The resultant constructs were electroporated into *A. baumannii* D1279779Δ*dusA* and selected on LB agar plates containing 50 μg/mL carbencillin.

Phenotype microarray testing and phenotype confirmation

Biolog phenotype microarray (PM) assays of the five isogenic *A. baumannii* D1279779 strains were conducted as previously (21), except where antibiotic selection was required for propagation of pWH1266 and its variants on agar. Plates PM1-3 were used to test carbon and nitrogen utilisation and plates PM9-10 were used to test responses to osmotic, ionic and pH stressors. Confirmatory testing of sole carbon source utilisation was conducted essentially as previously reported (21) with *A. baumannii* D1279779 and D1279779Δ*dusA*, except M9 was substituted for S2 minimal medium (without lactate) (39) and experiments were scaled down to 690 μL in 12 well microtitre plates. The carbon sources tested were: putrescine, monomethyl succinate and N-acetylputrescine hydrochloride (not in PM assay).

Results and Discussion

Identification of a novel family of potentially mobile islands associated with the *dusA* gene

During bioinformatic analysis conducted on *Acinetobacter baumannii* and *Pseudomonas* sp. genomes previously sequenced by our research group (18-21), we serendipitously observed that these two groups of organisms occasionally encoded a homologous integrase gene associated with the 5' end of the *dusA* gene. Alterations in the 5' sequence of *dusA* were evident in organisms that encoded this integrase, compared to related strains lacking it.

This initial analysis led to the subsequent discovery of other *dusA*-associated integrase (DAI) homologues in previously predicted GEIs of *Acinetobacter* and *Pseudomonas* that were not

identified as having integrated into *dusA*. Such GEIs included D1279779_RGP05 of *A. baumannii* D1279779 (21), $pA_{ICU}30$ of *A. baumannii* ACICU (*dusB*) (40), G08 of *A. baumannii* AB0057 and AYE (41), RGP56 of *P. aeruginosa* PA7, and Prophage 03 of *P. protegens* Pf-5 (42). A handful of other GEIs in published genomes were predicted as having been associated with the *dusA* gene in *Escherichia coli* (43,44), *Bradyrhizobium* sp. BTAi1 (45) and *Vibrio parahaemolyticus* AQ3354 (46). However, none of these reports have realised that these islands form a family of elements with a common insertion site.

Further BLAST analysis uncovered the presence of numerous other DAI homologues in over 200 bacterial genomes, typically located near the 5' end of the *dusA* gene. In each case, the DAI-encoding gene was flanked at the 3' end with a gene cluster that had typical hallmarks of laterally-acquired DNA i.e., they were not conserved in the core genome of related strains and they possessed atypical trinucleotide content. These genomic regions appeared to vary greatly in gene content and estimated size, typically in the range of 30-60 kb. Taken together, this data suggested that there may be a novel family of genomic islands encoding a distinct integrase with an affinity for the *dusA* gene.

The *dusA*-associated integrases are novel members of the tyrosine recombinase superfamily

BLASTP-based analysis suggested that the DAIs were uncharacterised members of the tyrosine recombinase superfamily, and their closest characterised relative was the shufflon-specific DNA recombinase (Rci) family. Rci is encoded in incompatibility group I Enterobacteriaceal plasmids, such as R64 (47) and R621a (48), and mediate the site-specific recombination of the mobile shufflon region located at the C-terminus of the *pilV* gene, determining recipient specificity in liquid mating (49). An alignment of protein sequences from representative tyrosine recombinases against several DAIs confirmed that the latter are

members of the tyrosine recombinase superfamily, as they contain the RHRH tetrad, and nucleophilic tyrosine, characteristic of this protein superfamily (50,51) (Supplementary Figure 3.1). Phylogenetic analysis confirmed that the closest relative to the DAI family is the Rci family (Figure 3.1).

Additional sequence alignments conducted with representative DAI and Rci recombinase proteins revealed, in addition to the conserved RHRH/Y pentad, another fifteen amino acid residues were conserved between the two protein families (Supplementary Figure 3.2). This alignment also demonstrated that Rci has an extended C-terminus relative to the DAIs (Supplementary Figure 3.2), previously established to be essential in the recombination of asymmetrical recombination sites in the shufflon (52). Further mutagenesis of the conserved RHR/Y residues in Rci had confirmed their essentiality for recombination activity, but not DNA binding (53). Both phylogenetic analysis and sequence alignments suggested that while the DAIs and Rci are related, the function of these two tyrosine recombinases may be divergent.

The *dusA*-associated genomic islands are exclusive to three Proteobacterial lineages

The DAI protein sequence from *A. baumannii* D1279779 (protein accession: AGH34419) was utilised in BLASTP searches to retrieve other *dusA*-associated integrases from publicly available genome sequences. A total of 258 unique integrase variants (including that of D1279779) were retrieved, with 213 (82.6 %) of these found to be associated with *dusA*. The other 45 integrases were either associated with an unknown gene, due to incomplete genome sequences (7 hits), or were associated with genes other than *dusA* (32 hits), including those encoding a $tRNA^{Arg}$ (*Pseudogulbenkiania ferrooxidans* 2002), a glutamate-ammonia ligase (*Rhodobacter sphaeroides* ATCC 17029), a sugar transporter (*Enterobacter radicincitans* DSM 16656), an aminobenzoyl-glutamate transporter (*Klebsiella pneumonia*) and the related

dusB and *dusC* genes (16 hits). Another 6 hits originated from bacteriophage genomes of *Burkholderia* phages AH2, BcepMigl, BcepIL02, and DC1, Enterobacteria phage mEp460 and *Pseudomonas* phage H66.

There were several instances of organisms harbouring more than one DAI in their genome (Supplementary Figure 3.3). For instance, *Acinetobacter* sp. ANC 3789 harboured three integrase orthologues, associated with the *dusA*, *dusB*, and *rluA* genes, the latter of which also encodes an enzyme involved in tRNA-modification, pseudouridylate synthase. In other instances, the genome of *Alteromonas macleodii* str. 'Deep ecotype' encoded two copies of the *dusA* gene, with an identical *dusA*-associated GEIs associated with each, whilst *Simonsiella muelleri* ATCC 29453 encoded a pair of *dusB* and *dusC*-associated integrases.

The amino acid identities of the 213 DAIs, relative to that of *A. baumannii* D1279779, ranged from 22.8 % (*Pseudogulbenkiania* sp. NH8B) to 97.6 % (*Acinetobacter baumannii* OFIC143), and were identified exclusively in the Alphaproteobacterial, Betaproteobacterial, and Gammaproteobacterial lineages. No DAI representatives were identified in Deltaproteobacteria or Epsilonproteobacteria, which is presumably linked to the lack of a *dusA* gene in these lineages. BLASTP searches against these taxonomic classes confirmed the complete absence of *dusA* from sequenced representatives of the Epsilonproteobacteria (22), and revealed sequenced Myxococcales of the Deltaproteobacterial lineage encoded this gene. There is currently a lack of data concerning the recently discovered Zetaproteobacteria, as *Mariprofundus ferrooxydans* PV-1 (54) is currently the only sequenced and annotated representative, and it does not encode *dusA* in its genome. The *dusA* gene is also known to exist in non-Proteobacterial organisms including Cyanobacteria and Viridiplantae (22).

The phylogeny of the DAI protein family was inferred with a neighbour-joining analysis based on a multiple alignment of the integrase protein sequences. It was found that the majority of these integrases had clustered according to their genus level relationships, particularly in the overrepresented Gammaproteobacterial lineages *Acinetobacter*, *Escherichia* and *Pseudomonas* (Supplementary Figure 3.3). The integrases of *Acinetobacter* and *Neisseria* appear to be collaterally related, in that each genus is of a distinct taxonomic grouping, yet the phylogeny suggests a common ancestral integrase (Supplementary Figure 3.3). Interestingly, both the Enterobacteria and *Pseudomonas* phage integrases were closely associated with the integrases of their presumed host organisms, suggesting a potential bacteriophage origin of these genomic islands (Supplementary Figure 3.3). Integrases that were not associated with *dusA* tended to be disparate from other integrases in closely related organisms (Supplementary Figure 3.3), particularly in the case of the *Acinetobacter dusB*-associated integrases. These non-cognate integrases may have diverged due to mutation of key amino acid residues that affect the site specificity of the integrase (55).

Detection of *dusA*-associated genomic island excision

Various types of genomic islands have been previously demonstrated capable of excision from the chromosomes of bacteria (56,57), hence we investigated the possibility that the putative *dus*-associated GEIs could also excise from the chromosomes of *A. baumannii* strains ACICU and D1279779, as well as *P. protegens* Pf-5. Oligonucleotides were designed to generate PCR amplicons only when the GEIs have excised as a covalently closed circular molecule and the *dusA* or *dusB* chromosomal junction had reformed without a GEI inserted (Figures 3.2-3.4). When genomic DNA was isolated from cultures grown in MH media, neither of these amplicons were detectable unless ‘unconventional’ PCR protocols involving either a two-step amplification process of 35 cycles each, or a single step amplification with

60 cycles (58,59) were employed (data not shown). Similar PCR protocols were required for detection of other circularised genomic islands (56,60).

MMC is a potent DNA-damaging agent which has been previously shown to induce the excision of various mobile elements, including integron cassettes (61), prophages and prophage remnants (62), and other genomic islands (36). We hypothesised that treatment with MMC might increase the rate of excision of the *dus*-associated GEIs and facilitate their detection by PCR. The three strains were cultured with sub-inhibitory concentrations of MMC and total genomic DNA was isolated. For two of the strains, both the excised GEI and remnant *dus* junction were detectable using a "standard" 35 cycle PCR amplification (Figure 3.2). In the case of *A. baumannii* D1279779 however, only the *dus* chromosomal junction was detectable (see below).

The inability to detect the excised D1279779 GEI may be related to its putatively non-replicative nature, in contrast to the prophage GEIs of ACICU and Pf-5. An approach utilising a kit-based derivative of the alkaline lysis method, conventionally used to isolate plasmid DNA, was devised in an attempt to enrich for the circularised D1279779 GEI. PCR amplification using template DNA resultant from this method enabled detection of both the circularised D1279779 GEI and the remnant chromosomal junction (Figure 3.2).

The combination of MMC and alkaline lysis has been used previously to induce and isolate lysogenic prophage cccDNA (63,64). However, analogous attempts to isolate GEI cccDNA in other bacterial organisms have been unsuccessful (56,60), presumably due to an absence of an inducer such as MMC. The enrichment of D1279779 GEI cccDNA through alkaline lysis represents, to the best of our knowledge, the first successful attempt to physically isolate a

non-phage circularised GEI. With further refinement, the isolation of various mobile elements from other bacterial organisms may be possible.

The amplicons obtained were consistent with the occurrence of genomic island excision and the reformation of the wild type *dusA* sequence (Figure 3.2), and this was confirmed by DNA sequencing of these amplicons. The resultant sequence data was used to map both the genomic island coordinates and putative attachment sites.

The *dusA*-associated genomic islands target a conserved region of *dusA*

Sequence data obtained from the amplicons of excised GEIs and reformed *dusA*/*dusB* genes from *A. baumannii* ACICU, D1279779 and *P. protegens* Pf-5 were examined to determine the precise island termini and the putative attachment sites involved in the integration and excision of the *dusA*-associated GEIs. This analysis revealed a semiconserved pair of direct repeats that formed the basis of the chromosomal and island attachment sites (Figure 3.4). The 5' sequence of the *dusA* genes in these organisms were found to be derived from the *dusA*-associated GEIs inserted into this gene, while the original *dusA* 5' end had unified into the wild type *dusA* sequence upon GEI excision (Figures 3.3 and 3.4). On the basis of this sequence data, the *dusA*-associated integrases and *dusA*-associated GEIs will henceforth be referred to as *dusA*-specific integrases and *dusA*-specific GEIs, respectively.

The size of the putative island attachment sites in *A. baumannii* D1279779 and *P. protegens* Pf-5 varied from 26-28 bp, due to the presence of an additional AC dinucleotide in the island attachment site of D1279779 and the chromosome of Pf-5 (Figure 3.4). The 19 bp attachment sites of the *A. baumannii* ACICU *dusB*-specific GEI did share some commonality with attachment sites of the *dusA*-specific islands (Figure 3.4), despite its integrase targeting an

otherwise non-cognate site. Like the *dusA*-specific GEIs, the *dusB*-specific GEI does provide a potential alternative 5' end for the *dusB* gene.

The predicted D1279779 and Pf-5 attachment sites also enabled the identification of direct repeats in the sequences of ninety-two other *dusA*-specific GEIs (Supplementary Table 3.1), originating from genomes sequences with a sufficient level of completeness. It was found that all the examined GEIs were flanked by two 21 bp semi-conserved attachment sequences (Figure 3.5), suggesting that despite the variability of the integrase protein sequence, these *dusA*-specific integrases target a conserved region within the *dusA* gene. The conserved sequence of the attachment sites allowed the genomic coordinates of genomic islands in numerous organisms to be approximated.

The *dusA*-specific islands are genetically and dimensionally variable

Determination of the aforementioned attachment sites facilitated revision of the size and genomic coordinates of *dus*-specific GEIs in *A. baumannii* and *P. protegens* Pf-5 (Table 3.2), as well as the discovery of ninety-two other *dusA*-specific GEIs (Supplementary Table 3.1).

Other than the *A. baumannii* D1279779 *dusA*-specific GEI (Figure 3.3), the size and integration sites of the other two genomic islands in this study were not accurately predicted beforehand. The prophage $pA_{ICU}30$ was discovered to be 293 bp larger than previously predicted, consistent with the *dusB* target site being unidentified. However, Prophage 03 was discovered to be 7903 bp larger than previously predicted, which was consistent with both the *dusA* target site and several cargo genes being overlooked. The *hsdS* gene in D1279779_RGP05 (Figure 3.3), encoding the specificity subunit of a Type I restriction-modification system, was previously thought to contain a frameshift based on the published

sequence, whilst retaining some functionality (21). However, we confirmed through PCR and sequencing that this frameshift was the result of a sequencing error.

Investigation of the ninety-four *dusA*-specific GEI sequences revealed extensive diversity in both the size and cargo gene content of these islands. The sizes of these GEIs ranged from 4.8 kb in *Sinorhizobium fredii* USDA 257 to 184.8 kb in *Caulobacter* sp. K31, and averaged 43.0 kb in size (Supplementary Table 3.1). A 94-way reciprocal BLASTP search, conducted on the predicted proteome of all ninety-four *dusA*-specific islands, revealed that of the combined 4464 protein coding genes (excluding *dusA*), 2300 genes (51.5%) were unique to any one GEI. Only the integrase gene was common to all *dusA*-specific GEIs, though a gene encoding a *prtN* homologue was found in sixty of these islands (63.8%). PrtN regulates the production of pyocin in *Pseudomonas aeruginosa* (65), with the latter thought to be ancestrally derived from bacteriophages (66), though the role this particular gene might play in the regulation of these genomic islands is not known.

Further examination of cargo gene function in *dusA*-specific GEIs revealed that most of these islands were lysogenic or remnant prophages (59/94 islands), whilst the others encoded an assortment of putative functions, including resistance to heavy metals, DNA restriction-modification, DNA replication, conjugative transfer and various metabolic enzymes (Supplementary Table 3.1). Several unique and noteworthy gene clusters were observed in *dusA*-specific GEIs, including those encoding a Type IV secretion system in *Phenylobacterium zucineum* HLK1 (67), cytochrome biogenesis in *Mesorhizobium opportunistum* WSM2075, and the pyrimidine utilisation (*rut*) pathway in *Acinetobacter baumannii* Naval-57. The variability of cargo gene content in the *dusA*-specific islands may be facilitated by the accretion and degradation of other mobile genetic elements, as suggested by the presence of other integrases and insertion sequences in some islands.

The D1279779_RGP05 island does not bestow any detectable metabolic functions

We utilised a recently published protocol, hypothetically capable of deleting very large genomic regions in Gram-negative bacteria (37), to delete a 48.3 kb region in *A. baumannii* D1279779, encompassing the entire *dusA*-specific GEI, the entire *dusA* gene, and the remnant *dusA* 5' end, to generate *A. baumannii* D1279779Δ*dusA*. D1279779_RGP05 could not be selectively deleted to restore the original *dusA* sequence, as the remnant FRT site would otherwise disrupt the *dusA* gene.

Isogenic variants of D1279779Δ*dusA* were subsequently generated, complemented with the D1279779 *dusA* or the *dusA* gene from *A. baumannii* ATCC 17978, which did not encode a *dusA*-specific island. We sought to determine the contribution of D1279779_RGP05 and *dusA* to the metabolic capabilities of *A. baumannii* D1279779 using Biolog phenotype microarray assays, which are conducted in a series of 96 well plates containing various lyophilised compounds, and response through respiration is measured using a tetrazolium dye indicator (68). Loss of *dusA* was not associated with any changes in carbon utilisation (PM1-2), though surprisingly this was also the case for D1279779_RGP05. Given that this genomic island encoded a number of putative metabolic genes (21), we had expected the loss of this island would result in changes to the utilisation of a number of carbon sources, but none were observed. It is possible that the metabolic genes are functionally redundant, with equivalent functions encoded within the core genome, or they may serve niche-specific functions that are not identifiable under the conditions of the Biolog assays.

Integration of the *dusA*-specific islands may restore the functioning of DusA

The *dusA*-associated GEIs were found to be specific for the 5' end of *dusA*, characteristically replacing the 5' end of *dusA* with new sequence provide by the GEI itself. Thus insertion of

the GEI could affect either transcription of the *dusA* gene or the functioning of the DusA enzyme. None of the amino acid residues in the affected region of DusA are known to be involved in either the binding of the flavin mononucleotide cofactor or tRNA in crystallised dihydrouridine synthases of *Thermotoga maritima* (69) and *Thermus thermophilus* (70). However, the affected sequence does contain a structural motif potentially important for the overall functioning of the DusA enzyme.

Dihydrouridine synthases contain a triosephosphate isomerase (TIM) barrel (69), one of the commonest protein folds, consisting of the canonical eightfold repeated beta-alpha unit (71). The wild-type N-terminal amino acid sequence of *dusA*, prior to the insertion of the *dusA*-associated GEI, encodes a beta strand (β-1) that forms part of the TIM barrel; the loss of this beta strand would presumably abolish DusA function. However, sequence-structure comparison revealed the new 5' *dusA* sequence resulting from the GEI insertion provided a suitable replacement beta strand (Supplementary Figure 3.4).

The *dus*-specific islands are members of the novel DAMGE family

GEIs are classically associated with particular tRNAs and tmRNAs genes (9,10) but *dusA* presents an intriguing target of MGEs, as DusA is coincidentally involved in the posttranscriptional modification of tRNAs. Other genes encoding tRNA modification enzymes that have been targeted by mobile elements include the Enterobacterial tRNA modification GTPase (*thdF*/*mnmE*) by SGI1 (56) and the *Polaromonas naphthalenivorans* CJ2 MiaB-like tRNA modifying enzyme (*rimO*) (72). This raises the question whether there is some selective pressure for insertion of GEIs or MGEs in small RNAs or small RNA modification enzymes.

The previously demonstrated dispensability of *dusA* in independent deletion libraries of three different bacteria (73-75), and the simultaneous dispensability of all three *dus* genes in *E. coli*, resulting in abolition of all dihydrouridine in this organism (25), suggested this gene would be a suitable target for MGE insertions. However, our bioinformatic analysis proposed that the functional and sequence integrity of DusA was maintained due to replacement sequence provided by integration of these islands. The simulated laboratory conditions that these deletion libraries were constructed in may conceal the fitness cost that the loss of *dusA* would have on these (and other organisms) in their respective niches.

This supposition is reinforced by the discovery of DUS family enzymes involved in roles other than posttranscriptional reduction of uridine. The *dusA* gene of *Neisseria meningitides* was upregulated during prolonged interaction with human epithelial cells, and when inactivated resulted in significant reduction to epithelial cell adherence (76). A gene encoding a dihydrouridine synthase in *Clostridium botulinum* was downregulated during heat shock stress (77). Inactivation of *dusB* has been implicated in decreased biofilm formation in *Pseudomonas aeruginosa* (78) and *Actinobacillus pleuropneumoniae* (79). Downregulation of *dusC* has been observed during *E. coli* cell death mediated by both bactericidal antibiotics (80) and *Bdellovibrio bacteriovorus* predation (81).

The presence of dihydrouridine itself in RNA has been demonstrated to alter its thermodynamic properties, resulting in greater conformational flexibility and dynamic motion (82). This adjustment in thermodynamic property is thought to be required for the maintenance of tRNA conformational flexibility in cryogenic conditions, as supported by the overabundance of this modified nucleobase in the tRNAs of psychrophilic microorganisms, compared to those present in psychrotrophic, mesophilic and thermophilic environments (83,84).

The *dusA*-specific integrases are phylogenetically distinct from previously characterised tyrosine recombinases present in other MGEs, with the Rci protein family being the closest characterised relative. On the basis of tyrosine recombinase phylogeny, the *dusA*-specific GEIs can be considered distinct integrative elements (11). However, these islands are evolutionarily derived from bacteriophages, and probably share a similar mechanism of insertion/excision, as evidenced by the incidence of bacteriophage integrases in the *dusA*-specific integrase phylogeny, as well as the presence of bacteriophage-related genes in many *dusA*-specific GEIs. The *dusA*-specific GEIs are capable of excision as a circularised intermediate and consist of a highly mosaic cargo gene pool, encoding genes devoted to varying functions, such as genomic island maintenance, propagation and replication, heavy metal resistance and augmentation of metabolic pathways. Taking the above into account, we propose that these *dusA*-specific GEIs form the basis for a novel genomic island family, which we christen the ***d*usA**-specific **M**obile **G**enetic **E**lements (DAMGEs). In addition to being capable of mobility and having an insertion site preference for *dusA*, DAMGEs appear to not be limited to any particular organism, environment or lifestyle, so long as a suitable *dusA* target is available. The highly mosaic architecture of DAMGEs would potentially prove beneficial to any host bacteria, and as more prokaryotic genome sequences become available, there will inevitably be more DAMGE variants discovered.

Funding

This work was supported by Macquarie University. DNF is supported by a Macquarie University Research Excellence Scholarship (MQRES).

Acknowledgements

We thank Dr. Karl Hassan and Dr. Sasha Tetu for fruitful discussions regarding experimental design. We would also like to thank Katsutoshi Hori for providing the FLP/FRT recombination vectors. The majority of the genomes utilised in this study were sourced from unpublished works made publically available by various genome sequencing consortia, including the US Department of Energy Joint Genome Institute (DOEJGI), the Broad Institute Genome Sequencing Platform, and the J. Craig Venter Institute (JCVI).

References

1. Perna, N.T., Plunkett, G., Burland, V., Mau, B., Glasner, J.D., Rose, D.J., Mayhew, G.F., Evans, P.S., Gregor, J., Kirkpatrick, H.A. *et al.* (2001) Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. *Nature*, **409**, 529-533.
2. Frost, L.S., Leplae, R., Summers, A.O. and Toussaint, A. (2005) Mobile genetic elements: the agents of open source evolution. *Nat. Rev. Microbiol.*, **3**, 722-732.
3. Langille, M.G.I., Hsiao, W.W.L. and Brinkman, F.S.L. (2010) Detecting genomic islands using bioinformatics approaches. *Nat. Rev. Microbiol.*, **8**, 373-382.
4. Juhas, M., Van Der Meer, J.R., Gaillard, M., Harding, R.M., Hood, D.W. and Crook, D.W. (2009) Genomic islands: tools of bacterial horizontal gene transfer and evolution. *FEMS Microbiol. Rev.*, **33**, 376-393.
5. Mathee, K., Narasimhan, G., Valdes, C., Qiu, X., Matewish, J.M., Koehrsen, M., Rokas, A., Yandava, C.N., Engels, R., Zeng, E. *et al.* (2008) Dynamics of *Pseudomonas aeruginosa* genome evolution. *Proc. Natl. Acad. Sci. USA*, **105**, 3100-3105.
6. Ogier, J.C., Calteau, A., Forst, S., Goodrich-Blair, H., Roche, D., Rouy, Z., Suen, G., Zumbihl, R., Givaudan, A., Tailliez, P. *et al.* (2010) Units of plasticity in bacterial genomes: new insight from the comparative genomics of two bacteria interacting with invertebrates, *Photorhabdus* and *Xenorhabdus*. *BMC Genomics*, **11**, 568.
7. Dobrindt, U., Hochhut, B., Hentschel, U. and Hacker, J. (2004) Genomic islands in pathogenic and environmental microorganisms. *Nat. Rev. Micro.*, **2**, 414-424.
8. Hsiao, W.W.L., Ung, K., Aeschliman, D., Bryan, J., Finlay, B.B. and Brinkman, F.S.L. (2005) Evidence of a large novel gene pool associated with prokaryotic genomic islands. *PLoS Genet.*, **1**, e62.
9. Williams, K.P. (2002) Integration sites for genetic elements in prokaryotic tRNA and tmRNA genes: sublocation preference of integrase subfamilies. *Nucleic Acids Res.*, **30**, 866-875.
10. Williams, K.P. (2003) Traffic at the tmRNA gene. *J. Bacteriol.*, **185**, 1059-1070.
11. Boyd, E.F., Almagro-Moreno, S. and Parent, M.A. (2009) Genomic islands are dynamic, ancient integrative elements in bacterial evolution. *Trends Microbiol.*, **17**, 47-53.
12. Song, L., Pan, Y., Chen, S. and Zhang, X. (2012) Structural characteristics of genomic islands associated with GMP synthases as integration hotspot among sequenced microbial genomes. *Comput. Biol. Chem.*, **36**, 62-70.

13. Wassenaar, T.M., Bohlin, J., Binnewies, T.T. and Ussery, D.W. (2009) Genome comparison of bacterial pathogens. *Genome Dyn.*, **6**, 1-20.
14. Kochar, M., Crosatti, M., Harrison, E.M., Rieck, B., Chan, J., Constantinidou, C., Pallen, M., Ou, H.-Y. and Rajakumar, K. (2012) Deletion of Tn*AbaR23* results in both expected and unexpected antibiogram changes in a multidrug-resistant *Acinetobacter baumannii* strain. *Antimicrob. Agents Chemother.*, **56**, 1845-1853.
15. Boyd, D., Cloeckaert, A., Chaslus-Dancla, E. and Mulvey, M.R. (2002) Characterization of variant *Salmonella* genomic island 1 multidrug resistance regions from serovars Typhimurium DT104 and Agona. *Antimicrob. Agents Chemother.*, **46**, 1714-1722.
16. Beaber, J.W., Hochhut, B. and Waldor, M.K. (2002) Genomic and functional analyses of SXT, an integrating antibiotic resistance gene transfer element derived from *Vibrio cholerae*. *J. Bacteriol.*, **184**, 4259-4269.
17. Schubert, S., Rakin, A. and Heesemann, J. (2004) The *Yersinia* high-pathogenicity island (HPI): evolutionary and functional aspects. *Int. J. Med. Microbiol.*, **294**, 83-94.
18. Paulsen, I.T., Press, C.M., Ravel, J., Kobayashi, D.Y., Myers, G.S., Mavrodi, D.V., DeBoy, R.T., Seshadri, R., Ren, Q., Madupu, R. *et al.* (2005) Complete genome sequence of the plant commensal *Pseudomonas fluorescens* Pf-5. *Nat. Biotechnol.*, **23**, 873-878.
19. Roy, P.H., Tetu, S.G., Larouche, A., Elbourne, L., Tremblay, S., Ren, Q., Dodson, R., Harkins, D., Shay, R., Watkins, K. *et al.* (2010) Complete genome sequence of the multiresistant taxonomic outlier *Pseudomonas aeruginosa* PA7. *PLoS ONE*, **5**, e8842.
20. Loper, J.E., Hassan, K.A., Mavrodi, D.V., Davis, E.W., II, Lim, C.K., Shaffer, B.T., Elbourne, L.D.H., Stockwell, V.O., Hartney, S.L., Breakwell, K. *et al.* (2012) Comparative genomics of plant-associated *Pseudomonas* spp.: insights into diversity and inheritance of traits involved in multitrophic interactions. *PLoS Genet.*, **8**, e1002784.
21. Farrugia, D.N., Elbourne, L.D.H., Hassan, K.A., Eijkelkamp, B.A., Tetu, S.G., Brown, M.H., Shah, B.S., Peleg, A.Y., Mabbutt, B.C. and Paulsen, I.T. (2013) The complete genome and phenome of a community-acquired *Acinetobacter baumannii*. *PLoS ONE*, **8**, e58628.
22. Kasprzak, J.M., Czerwoniec, A. and Bujnicki, J.M. (2012) Molecular evolution of dihydrouridine synthases. *BMC Bioinformatics*, **13**, 153.
23. Savage, D.F., de Crecy-Lagard, V. and Bishop, A.C. (2006) Molecular determinants of dihydrouridine synthase activity. *FEBS Lett.*, **580**, 5198-5202.
24. Sprinzl, M. and Vassilenko, K.S. (2005) Compilation of tRNA sequences and sequences of tRNA genes. *Nucleic Acids Res.*, **33**, D139-140.
25. Bishop, A.C., Xu, J., Johnson, R.C., Schimmel, P. and de Crecy-Lagard, V. (2002) Identification of the tRNA-dihydrouridine synthase family. *J Biol Chem*, **277**, 25090-25095.
26. Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.*, **25**, 3389-3402.
27. Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.*, **22**, 4673-4680.
28. Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W.Z., Lopez, R., McWilliam, H., Remmert, M., Soding, J. *et al.* (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.*, **7**.

29. Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. and Kumar, S. (2011) MEGA5: Molecular Evolutionary Genetics Analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.*, **28**, 2731-2739.
30. Campanella, J., Bitincka, L. and Smalley, J. (2003) MatGAT: An application that generates similarity/identity matrices using protein or DNA sequences. *BMC Bioinformatics*, **4**, 29.
31. Stover, B.C. and Muller, K.F. (2010) TreeGraph 2: combining and visualizing evidence from different phylogenetic analyses. *BMC Bioinformatics*, **11**, 7.
32. Okonechnikov, K., Golosova, O. and Fursov, M. (2012) Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics*, **28**, 1166-1167.
33. Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K. and Madden, T.L. (2009) BLAST+: architecture and applications. *BMC Bioinformatics*, **10**, 421.
34. Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M. and Rozen, S.G. (2012) Primer3—new capabilities and interfaces. *Nucleic Acids Res.*, **40**, e115.
35. Wiegand, I., Hilpert, K. and Hancock, R.E. (2008) Agar and broth dilution methods to determine the minimal inhibitory concentration (MIC) of antimicrobial substances. *Nat. Protoc.*, **3**, 163-175.
36. Bellanger, X., Morel, C., Decaris, B. and Guédon, G. (2007) Derepression of excision of integrative and potentially conjugative elements from *Streptococcus thermophilus* by DNA damage response: implication of a cI-related repressor. *J. Bacteriol.*, **189**, 1478-1481.
37. Ishikawa, M. and Hori, K. (2013) A new simple method for introducing an unmarked mutation into a large gene of non-competent Gram-negative bacteria by FLP/FRT recombination. *BMC Microbiology*, **13**, 86.
38. Hunger, M., Schmucker, R., Kishan, V. and Hillen, W. (1990) Analysis and nucleotide sequence of an origin of an origin of DNA replication in *Acinetobacter calcoaceticus* and its use for *Escherichia coli* shuttle plasmids. *Gene*, **87**, 45-51.
39. Juni, E. (1974) Simple genetic transformation assay for rapid diagnosis of *Moraxella osloensis*. *Applied Microbiology*, **27**, 16-24.
40. Iacono, M., Villa, L., Fortini, D., Bordoni, R., Imperi, F., Bonnal, R.J.P., Sicheritz-Ponten, T., De Bellis, G., Visca, P., Cassone, A. *et al.* (2008) Whole-genome pyrosequencing of an epidemic multidrug-resistant *Acinetobacter baumannii* strain belonging to the European clone II group. *Antimicrob. Agents Chemother.*, **52**, 2616-2625.
41. Di Nocera, P., Rocco, F., Giannouli, M., Triassi, M. and Zarrilli, R. (2011) Genome organization of epidemic *Acinetobacter baumannii* strains. *BMC Microbiol.*, **11**, 224.
42. Mavrodi, D.V., Loper, J.E., Paulsen, I.T. and Thomashow, L.S. (2009) Mobile genetic elements in the genome of the beneficial rhizobacterium *Pseudomonas fluorescens* Pf-5. *BMC Microbiology*, **9**, 8.
43. Ogura, Y., Abe, H., Katsura, K., Kurokawa, K., Asadulghani, M., Iguchi, A., Ooka, T., Nakayama, K., Yamashita, A., Hattori, M. *et al.* (2008) Systematic identification and sequence analysis of the genomic islands of the enteropathogenic *Escherichia coli* strain B171-8 by the combined use of whole-genome PCR scanning and fosmid mapping. *J. Bacteriol.*, **190**, 6948-6960.
44. Eppinger, M., Mammel, M.K., Leclerc, J.E., Ravel, J. and Cebula, T.A. (2011) Genomic anatomy of *Escherichia coli* O157:H7 outbreaks. *Proc. Natl. Acad. Sci. USA*, **108**, 20142-20147.

45. Giraud, E., Moulin, L., Vallenet, D., Barbe, V., Cytryn, E., Avarre, J.C., Jaubert, M., Simon, D., Cartieaux, F., Prin, Y. *et al.* (2007) Legumes symbioses: absence of Nod genes in photosynthetic bradyrhizobia. *Science*, **316**, 1307-1312.
46. Shimamoto, T., Ikeda, S. and Shimamoto, T. (2006) Association of a retron with a prophage-like element integrated into the tRNA dihydrouridine synthase gene of *Vibrio parahaemolyticus*. *Seikagaku*, **A11475**, 5P-B-140.
47. Sampei, G., Furuya, N., Tachibana, K., Saitou, Y., Suzuki, T., Mizobuchi, K. and Komano, T. (2010) Complete genome sequence of the incompatibility group I1 plasmid R64. *Plasmid*, **64**, 92-103.
48. Takahashi, H., Shao, M., Furuya, N. and Komano, T. (2011) The genome sequence of the incompatibility group Iγ plasmid R621a: evolution of IncI plasmids. *Plasmid*, **66**, 112-121.
49. Gyohda, A., Funayama, N. and Komano, T. (1997) Analysis of DNA inversions in the shufflon of plasmid R64. *J. Bacteriol.*, **179**, 1867-1871.
50. Nunes-Düby, S.E., Kwon, H.J., Tirumalai, R.S., Ellenberger, T. and Landy, A. (1998) Similarities and differences among 105 members of the Int family of site-specific recombinases. *Nucleic Acids Res.*, **26**, 391-406.
51. Grainge, I. and Jayaram, M. (1999) The integrase family of recombinase: organization and function of the active site. *Mol. Microbiol.*, **33**, 449-456.
52. Gyohda, A., Zhu, S., Furuya, N. and Komano, T. (2006) Asymmetry of shufflon-specific recombination sites in plasmid R64 inhibits recombination between direct sfx sequences. *J. Biol. Chem.*, **281**, 20772-20779.
53. Gyohda, A. and Komano, T. (2000) Purification and characterization of the R64 shufflon-specific recombinase. *J. Bacteriol.*, **182**, 2787-2792.
54. Singer, E., Emerson, D., Webb, E.A., Barco, R.A., Kuenen, J.G., Nelson, W.C., Chan, C.S., Comolli, L.R., Ferriera, S., Johnson, J. *et al.* (2011) *Mariprofundus ferrooxydans* PV-1 the first genome of a marine Fe(II) oxidizing Zetaproteobacterium. *PLoS ONE*, **6**, e25386.
55. Groth, A.C. and Calos, M.P. (2004) Phage integrases: biology and applications. *J. Mol. Biol.*, **335**, 667-678.
56. Doublet, B., Boyd, D., Mulvey, M.R. and Cloeckaert, A. (2005) The *Salmonella* genomic island 1 is an integrative mobilizable element. *Mol. Microbiol.*, **55**, 1911-1924.
57. Bille, E., Zahar, J.R., Perrin, A., Morelle, S., Kriz, P., Jolley, K.A., Maiden, M.C., Dervin, C., Nassif, X. and Tinsley, C.R. (2005) A chromosomally integrated bacteriophage in invasive meningococci. *J. Exp. Med.*, **201**, 1905-1913.
58. Rameckers, J., Hummel, S. and Herrmann, B. (1997) How many cycles does a PCR need? determinations of cycle numbers depending on the number of targets and the reaction efficiency factor. *Naturwissenschaften*, **84**, 259-262.
59. Rose, A. (2010) Tn*AbaR1*: a novel Tn7-related transposon in *Acinetobacter baumannii* that contributes to the accumulation and dissemination of large repertoires of resistance genes. *Biosci. Horiz.*, **3**, 40-48.
60. Domínguez, N.M., Hackett, K.T. and Dillard, J.P. (2011) XerCD-mediated site-specific recombination leads to loss of the 57-kilobase gonococcal genetic island. *J. Bacteriol.*, **193**, 377-388.
61. Guerin, É., Cambray, G., Sanchez-Alberola, N., Campoy, S., Erill, I., Da Re, S., Gonzalez-Zorn, B., Barbé, J., Ploy, M.-C. and Mazel, D. (2009) The SOS response controls integron recombination. *Science*, **324**, 1034.
62. Canchaya, C., Desiere, F., McShan, W.M., Ferretti, J.J., Parkhill, J. and Brüssow, H. (2002) Genome analysis of an inducible prophage and prophage remnants integrated in the *Streptococcus pyogenes* strain SF370. *Virology*, **302**, 245-258.

63. Geslin, C., Le Romancer, M., Erauso, G., Gaillard, M., Perrot, G. and Prieur, D. (2003) PAV1, the first virus-like particle isolated from a hyperthermophilic euryarchaeote, "*Pyrococcus abyssi*". *Journal of Bacteriology*, **185**, 3888-3894.
64. Lehours, P., Vale, F.F., Bjursell, M.K., Melefors, O., Advani, R., Glavas, S., Guegueniat, J., Gontier, E., Lacomme, S., Alves Matos, A. *et al.* (2011) Genome sequencing reveals a phage in *Helicobacter pylori*. *mBio*, **2**.
65. Matsui, H., Sano, Y., Ishihara, H. and Shinomiya, T. (1993) Regulation of pyocin genes in *Pseudomonas aeruginosa* by positive (*prtN*) and negative (*prtR*) regulatory genes. *J. Bacteriol.*, **175**, 1257-1263.
66. Michel-Briand, Y. and Baysse, C. (2002) The pyocins of *Pseudomonas aeruginosa*. *Biochimie*, **84**, 499-510.
67. Luo, Y., Xu, X., Ding, Z., Liu, Z., Zhang, B., Yan, Z., Sun, J., Hu, S. and Hu, X. (2008) Complete genome of *Phenylobacterium zucineum* - a novel facultative intracellular bacterium isolated from human erythroleukemia cell line K562. *BMC Genomics*, **9**, 386.
68. Bochner, B.R., Gadzinski, P. and Panomitros, E. (2001) Phenotype microarrays for high-throughput phenotypic testing and assay of gene function. *Genome Res.*, **11**, 1246-1255.
69. Park, F., Gajiwala, K., Noland, B., Wu, L., He, D., Molinari, J., Loomis, K., Pagarigan, B., Kearins, P., Christopher, J. *et al.* (2004) The 1.59 A resolution crystal structure of TM0096, a flavin mononucleotide binding protein from *Thermotoga maritima*. *Proteins*, **55**, 772-774.
70. Yu, F., Tanaka, Y., Yamashita, K., Suzuki, T., Nakamura, A., Hirano, N., Yao, M. and Tanaka, I. (2011) Molecular basis of dihydrouridine formation on tRNA. *Proc. Natl. Acad. Sci. USA*, **108**, 19593-19598.
71. Wierenga, R.K. (2001) The TIM-barrel fold: a versatile framework for efficient enzymes. *FEBS Lett.*, **492**, 193-198.
72. Yagi, J.M., Sims, D., Brettin, T., Bruce, D. and Madsen, E.L. (2009) The genome of *Polaromonas naphthalenivorans* strain CJ2, isolated from coal tar-contaminated sediment, reveals physiological and metabolic versatility and evolution through extensive horizontal gene transfer. *Environ. Microbiol.*, **11**, 2253-2270.
73. Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K.A., Tomita, M., Wanner, B.L. and Mori, H. (2006) Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.*, **2**, 2006 0008.
74. Liberati, N.T., Urbach, J.M., Miyata, S., Lee, D.G., Drenkard, E., Wu, G., Villanueva, J., Wei, T. and Ausubel, F.M. (2006) An ordered, nonredundant library of *Pseudomonas aeruginosa* strain PA14 transposon insertion mutants. *Proc. Natl. Acad. Sci. USA*, **103**, 2833-2838.
75. de Berardinis, V., Vallenet, D., Castelli, V., Besnard, M., Pinet, A., Cruaud, C., Samair, S., Lechaplais, C., Gyapay, G., Richez, C. *et al.* (2008) A complete collection of single-gene deletion mutants of *Acinetobacter baylyi* ADP1. *Mol. Syst. Biol.*, **4**.
76. Hey, A., Li, M.-S., Hudson, M.J., Langford, P.R. and Kroll, J.S. (2013) Transcriptional profiling of *Neisseria meningitidis* interacting with human epithelial cells in a long-term *in vitro* colonization model. *Infect. Immun.*, **81**, 4149-4159.
77. Liang, W.-d., Bi, Y.-t., Wang, H.-y., Dong, S., Li, K.-s. and Li, J.-s. (2013) Gene expression profiling of *Clostridium botulinum* under heat shock stress. *Biomed. Res. Int.*, **2013**, 8.
78. Musken, M., Di Fiore, S., Dotsch, A., Fischer, R. and Haussler, S. (2010) Genetic determinants of *Pseudomonas aeruginosa* biofilm establishment. *Microbiology*, **156**, 431-441.

79. Grasteau, A., Tremblay, Y.D., Labrie, J. and Jacques, M. (2011) Novel genes associated with biofilm formation of *Actinobacillus pleuropneumoniae*. *Vet. Microbiol.*, **153**, 134-143.
80. Kohanski, M.A., Dwyer, D.J., Hayete, B., Lawrence, C.A. and Collins, J.J. (2007) A common mechanism of cellular death induced by bactericidal antibiotics. *Cell*, **130**, 797-810.
81. Lambert, C., Ivanov, P. and Sockett, R. (2010) A transcriptional "scream" early response of *E. coli* prey to predatory invasion by bdellovibrio. *Curr Microbiol*, **60**, 419-427.
82. Dalluge, J.J., Hashizume, T., Sopchik, A.E., McCloskey, J.A. and Davis, D.R. (1996) Conformational flexibility in RNA: the role of dihydrouridine. *Nucleic Acids Res.*, **24**, 1073-1079.
83. Dalluge, J.J., Hamamoto, T., Horikoshi, K., Morita, R.Y., Stetter, K.O. and McCloskey, J.A. (1997) Posttranscriptional modification of tRNA in psychrophilic bacteria. *J. Bacteriol.*, **179**, 1918-1923.
84. Noon, K.R., Guymon, R., Crain, P.F., McCloskey, J.A., Thomm, M., Lim, J. and Cavicchioli, R. (2003) Influence of temperature on tRNA modification in Archaea: *Methanococcoides burtonii* (optimum growth temperature [T_{opt}], 23°C) and *Stetteria hydrogenophila* (T_{opt}, 95°C). *Journal of Bacteriology*, **185**, 5483-5490.
85. Smith, M.G., Gianoulis, T.A., Pukatzki, S., Mekalanos, J.J., Ornston, L.N., Gerstein, M. and Snyder, M. (2007) New insights into *Acinetobacter baumannii* pathogenesis revealed by high-density pyrosequencing and transposon mutagenesis. *Genes & Development*, **21**, 601-614.
86. Simon, R., Priefer, U. and Puhler, A. (1983) A broad host range mobilization system for *in vivo* genetic engineering: transposon mutagenesis in Gram negative bacteria. *Nat. Biotech.*, **1**, 784-791.
87. Crooks, G.E., Hon, G., Chandonia, J.M. and Brenner, S.E. (2004) WebLogo: a sequence logo generator. *Genome Res.*, **14**, 1188-1190.
88. Gille, C. and Frömmel, C. (2001) STRAP: editor for STRuctural Alignments of Proteins. *Bioinformatics*, **17**, 377-378.

Tables

Table 3.1: Bacterial strains and plasmids utilised in this study with relevant genotypes indicated.

Bacterial strains	Description and relevant genotype	Reference
Acinetobacter baumannii ACICU	Intact *dusA*, *dusB*::$pA_{ICU}30$[1]	(40)
Acinetobacter baumannii ATCC 17978	Intact *dusA*	(85)
Acinetobacter baumannii D1279779	Wild type strain, *dusA*::D1279779_RGP05	(21)
A. baumannii D1279779Δ*dusA*	ΔABD1_05270-05691[2]	This study
A. baumannii D1279779Δ*dusA* (Empty vector)	D1279779Δ*dusA* + pWH1266	This study
A. baumannii D1279779Δ*dusA* (Island)	D1279779Δ*dusA* + pWH5270	This study
A. baumannii D1279779Δ*dusA* (Wild-type)	D1279779Δ*dusA* + pWH0562	This study
Pseudomonas protegens Pf-5	*dusA*::Prophage 03	(18,42)
Escherichia coli S17-1	Donor strain for conjugation	(86)

Plasmids		
pJQFRT	Gene replacement vector, FRT, *sacB*, Gm^R	(37)
pKFRT/FLP	Gene replacement vector, FRT, *flp*, *tetR*, Km^R	(37)
pJQFRT_05270	1 kbp region, upstream of *dusA*, cloned into BamHI and SacI pJQFRT sites	This study
pKFRT/FLP_05691	1 kbp region, downstream of *dusA* 5', cloned into BamHI and BspEI pJQFRT sites	This study
pWH1266	Ap^R, Tc^R	(38)
pWH5270	*A. baumannii* D1279779 *dusA* and native promoter cloned into pWH1266 BamHI site in *tet* gene, Ap^R	This study
pWH0562	*A. baumannii* ATCC 17978 *dusA* and native promoter cloned into pWH1266 BamHI site in *tet* gene, Ap^R	This study

1: The *dusA/B*::IslandX genotypes denote disruption of the *dusA*/*dusB* nucleotide sequence by the insertion of the *dus*-associated islands.
2: Deletion of a 48.3 kb region consisting of *dusA*, D1279779_RGP05 and the original *dusA* 5' end.
Abbreviations: Ap, ampicillin; Gm, gentamycin; Km, kanamycin; Tc, tetracycline.

Table 3.2: The size, genomic coordinates and phenotypes of *dus*-associated genomic islands in *Acinetobacter* and *Pseudomonas* strains used in this study. Abbreviations: AST, arginine succinyltransferase.

Organism	Island Name	Genomic Coordinates (Size)	GC%	Predicted phenotype
A. baumannii D1279779	D1279779_RGP05	595655..642997 (47343 bp)	37.3	Type I restriction-modification, AST pathway
A. baumannii ACICU	$pA_{ICU}30$ (Phage 04)	2878594..2932506 (53913 bp)	39.7	Phage lysogen
P. protegens Pf-5	Prophage 03	2206875..2248336 (41462 bp)	58.9	P2-like phage lysogen

Figures

Figure 3.1. The inferred phylogenetic relationship of the *dusA*-associated integrases (DAI) in relation to representative tyrosine recombinases of the shufflon-specific DNA recombinase (Rci), phage integrase (Phage), integron integrase (IntI) and site-specific recombinase (XerDC) families. Protein accession numbers of sequences used to generate this phylogenetic tree are in brackets. The interior values are the bootstrap probabilities after 1000 replicates.

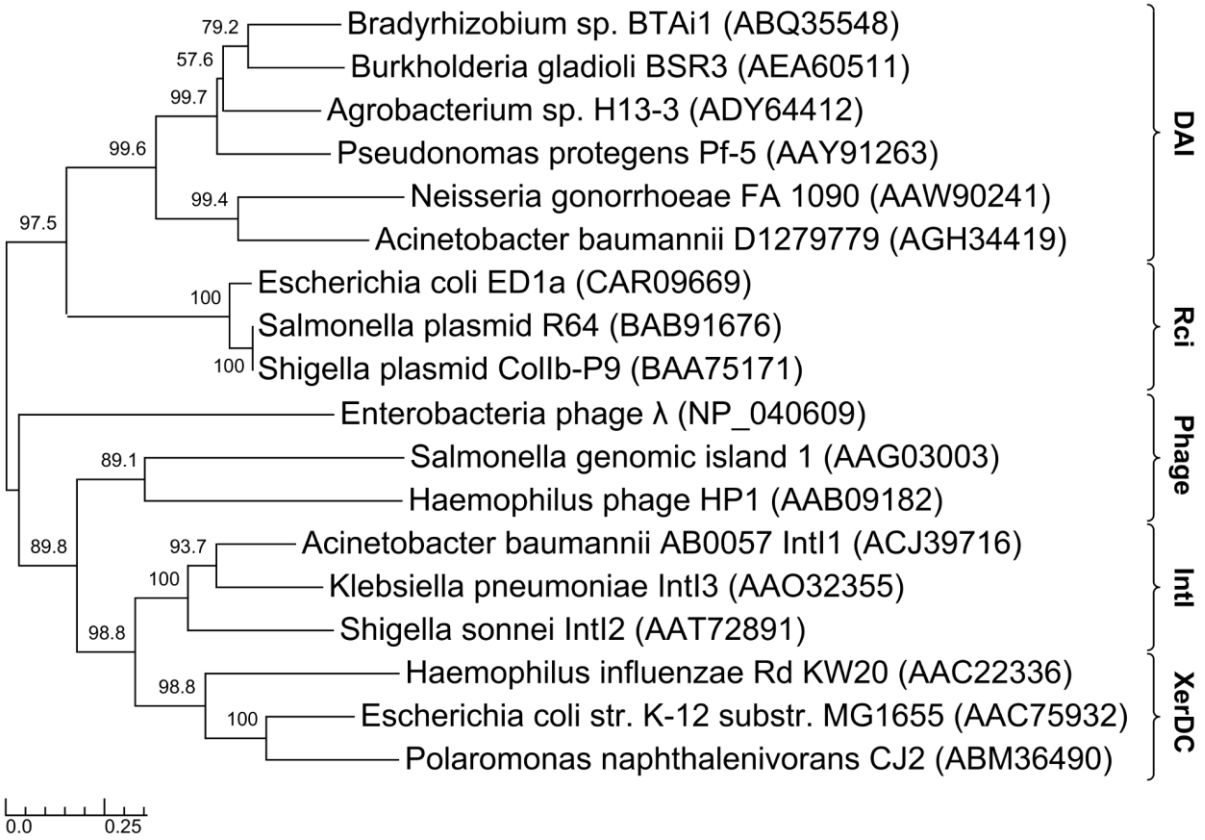


Figure 3.2. PCR detection of excised *dus*-associated genomic islands (X) and restored *dusA*/*dusB* chromosomal junctions (J) in *Acinetobacter baumannii* D1279779, *A. baumannii* ACICU and *Pseudomonas protegens* Pf-5. The faint high molecular weight bands originate from genomic DNA added to the PCR reactions (black arrow).

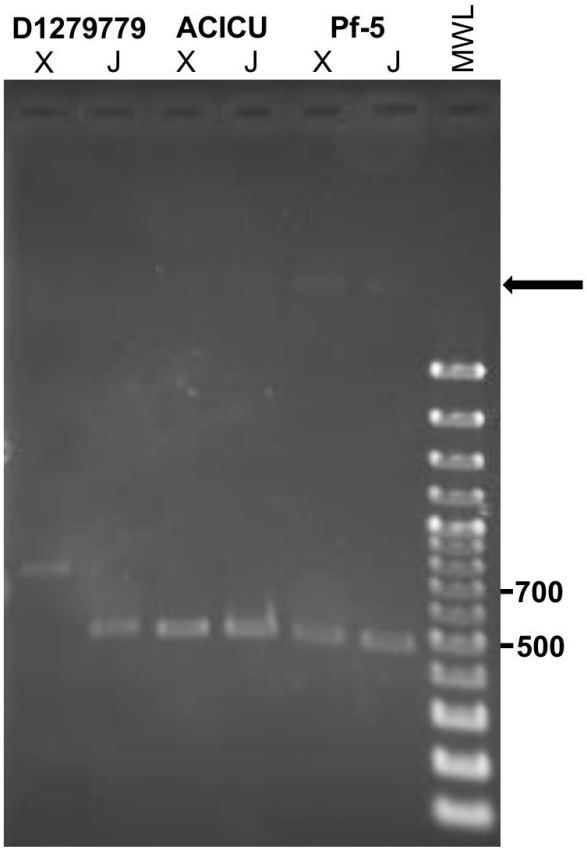


Figure 3.3. Excision and integration of the *Acinetobacter baumannii* D1279779 *dusA*-specific genomic island. The *dusA*-specific genomic islands consist of a *dusA*-specific integrase (*int)* (plaid) that catalyses excision of the genomic island as a circular intermediate as well as its integration into the 5' end of the chromosomal tRNA-dihydrouridine synthase A (*dusA*) gene (black). The 5' portion of the *dusA* gene affected by the integration is replaced by a new 5' end (white) provided by the genomic island, with the original *dusA* 5' end (dark grey) forming the outer boundary of the island. This particular genomic island variant encodes several hypothetical proteins (light grey) as well as proteins putatively involved in metabolism (diamonds), transport (gridlines), transcriptional regulation (checkerboard), and Type I restriction-modification (zig-zags). As is true of other *A. baumannii* strains, the *dusA*-specific genomic islands are flanked by a gene encoding a putative major facilitator superfamily transporter (*MFS*) (black). The binding sites of oligonucleotides used to detect excision of the genomic island (black triangles) and restoration of the *dusA* sequence (white triangles) are indicated. This figure is drawn to scale with the exception of the excised island, which is displayed at 60% scale.

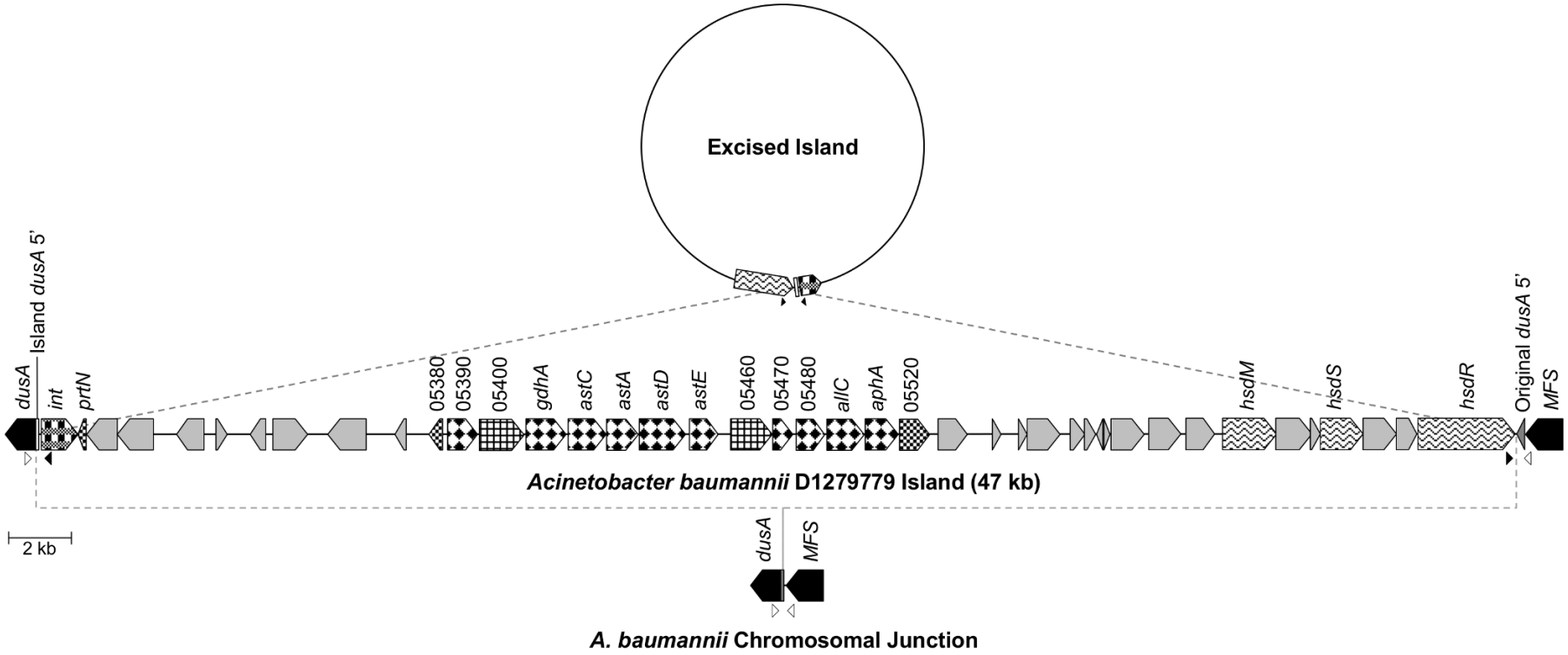


Figure 3.4. Predicted mechanism of genomic island excision, in *Acinetobacter baumannii* D1279779, *A. baumannii* ACICU and *Pseudomonas protegens* Pf-5, as a closed circular molecule (2) from the chromosome (1), resulting in the reformation of the native *dusA/B* sequence (3). The shading of the nucleotide sequences represent their origin as either from the circularised genomic island (black and grey) or chromosome (white). The underlined sequences represent the putative attachment sites involved in genomic island excision and integration (*attP* and *attB*); upon the latter occurring, a pair of imperfect direct repeats (*attL* and *attR*) are generated. Nucleotides in bold are identical in the *attP*/*attB* and *attL*/*attR* pairs.

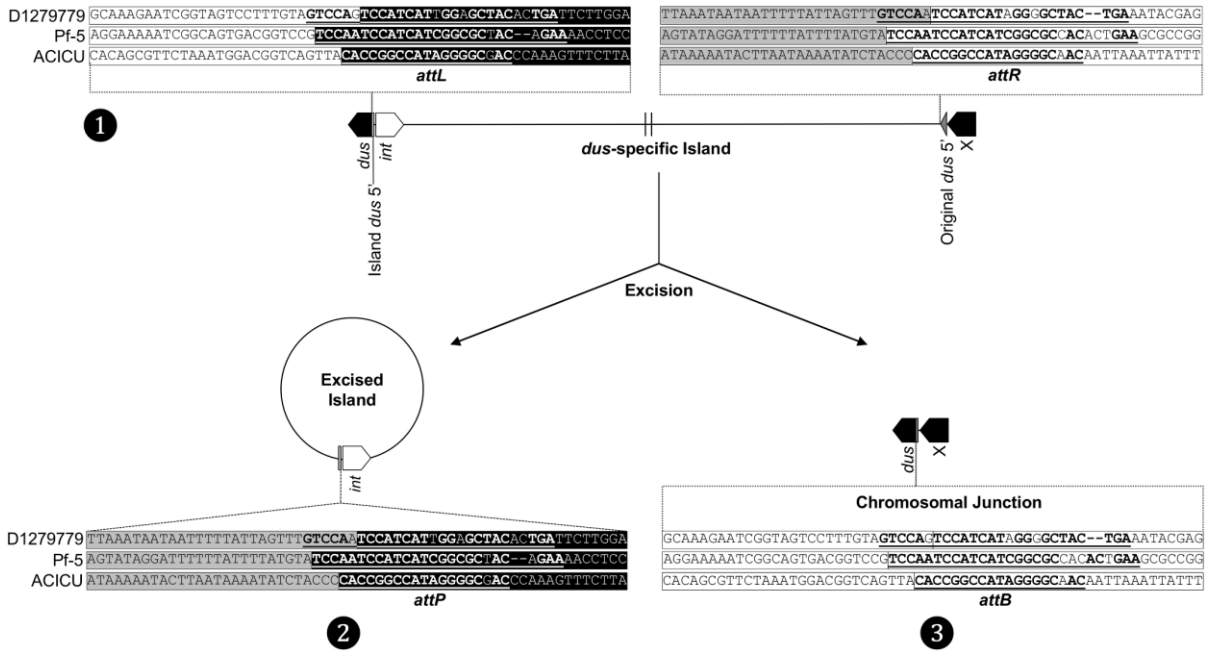


Figure 3.5. Logo consensus sequence of *attL* and *attR* direct repeats (as defined in Figure 3.4) generated with WebLogo (87), using alignments of flanking sequences from the ninety-four *dusA*-specific genome island sequences outlined in Supplementary Table 3.1. The thymidines marked by arrows in both sequences designate the island termini predicted from the sequence of *Acinetobacter baumannii* D1279779.

Supplementary Figures

Supplementary Figure 3.1. Multiple protein sequence alignment of representative tyrosine recombinase superfamily proteins outlined in Figure 3.1. RHRH tetrad and nucleophilic tyrosine amino acids common to the tyrosine recombinase superfamily are highlighted black, and conservatively substituted residues are highlighted grey.

Supplementary Figure 3.2. Multiple protein sequence alignments of shufflon-specific DNA recombinases (Rci) against representative *dusA*-associated integrases (DAIs) outlined in Figure 3.1, with the addition of DAIs from *Shewanella baltica* OS155 (ABN63061), *Variovorax paradoxus* S110 (ACS18756), *Paracoccus denitrificans* PD1222 (ABL69522). Amino acid residues common to the representative DAIs are highlighted in dark blue, whilst conservatively substituted amino acid residues are highlighted dark green. Amino acids common (black), conservatively substituted amino acids (grey), and semi-conservatively substituted amino acids (.) in both Rci and DAI proteins are indicated. Previously mutagenised amino acid residues (52,53) in R64 Rci (red) are also indicated.

Supplementary Figure 3.3. Phylogenetic tree of unique variants of *dusA*-associated integrase proteins. Organism names in bold and italicised are integrases that have an unknown target site, due to incomplete genome sequences. Organism names in bold and underlined are integrases that are not associated with *dusA*, with their associated gene in brackets. Integrases originating from bacteriophages are indicated with underlining and italics. Unless otherwise indicated, all other integrase sequences are *dusA*-associated. The interior values are the bootstrap probabilities after 1000 replicates.

Supplementary Figure 3.4. Alignment of N-terminal DusA protein sequences from *Acinetobacter baumannii* D1279779, prior to (D1279779_DusA_Island) and after excision

(D1279779_DusA_WildType) of its *dusA*-specific genomic island, against crystallised dihydrouridine synthases from *Thermotoga maritima* (1VHN) and *Thermus thermophilus* (3B0U). Amino acid residues that are similar are highlighted pink, residues that are conserved are highlighted blue, and non-conserved residues are white. Yellow blocks and red spirals indicate beta strands and alpha helices, respectively. The top secondary structure map is that of 3B0U, and the bottom is that of 1VHN. The arrow indicates the predicted terminus of the *dusA*-specific genomic island termini of *A. baumannii* D1279779. This figure was generated with STRAP (88).

Supplementary Table 3.1: List of organisms encoding complete *dusA* -associated islands, extracted from publicly available genome sequences. Information on the genomic coordinates, size, G+C percentage and putative function of these genomic islands are listed.

Organism	Island Name	Locus Tag	CDS	Size (kb)	G+C (%)	*dusA*	*int*	*prtN*	Putative genomic island function(s)
Acidovorax sp. KKS102		C380	12130-12220	11.3	61.6	12220	12195		Phage lysogen
Acinetobacter baumannii AB0057		AB57	0641-0666	23.5	37.7	0641	0644	0643	Copper resistance
Acinetobacter baumannii AYE		ABAYE	3198-3233	23.7	37.7	3223	3221	3220	Copper resistance
Acinetobacter baumannii D1279779	D1279779_RGP05	ABD1	05270-05690	47.3	37.3	05270	05280	05290	Type I restriction-modification, AST pathway
Acinetobacter baumannii Naval-57		ACINNAV57	0588-0630	46.6	39.2	0588	0590	0589	Type I restriction-modification, pyrimidine utilisation pathway
Acinetobacter baumannii OIFC137		ACIN3137_A	3137-3319	41.4	37.8	3319	3320	3321	Copper and arsenic resistance
Acinetobacter guillouiae MSP4-18		L291	2552-2589	38.2	39.1	2552	2553	2554	Arsenic, chromium and heavy metal resistance
Acinetobacter gyllenbergii MTCC 11365		L293	1364-1388	22.7	39.2	1364	1365		Arsenic and heavy metal resistance, ferrous iron transport
Acinetobacter lwoffii SH145		HMPREF0017	02541-02567	32.9	40.0	02541	02542	02543	Type I restriction-modification, potassium transport
Acinetobacter radioresistens SK82		ACIRA0001	0075-0114	29.8	40.4	0114	0113	0109	Phage lysogen
Acinetobacter sp. NBRC100985		ACT4_006	00690-00890	27.2	36.2	00690	00700	00710	Type II restriction-modification, DNA helicase
Acinetobacter sp. WC-743		ACINWC743	3553-3576	25.0	38.3	3553	3554		Arsenic resistance, ferric iron uptake, metabolism
Agrobacterium sp. H13-3		AGROH133	06152-06219	41.1	56.6	06152	06153		Phage lysogen
Aliivibrio salmonicida LFI1238	ϕ VS1	VSALI	0388-0399	8.9	33.9	0399	0398		Phage remnant
Alteromonas macleodii Balearic Sea AD45		AMBAS45	16335	19.9	38.6	16335	16340		Phage remnant
Alteromonas macleodii str. 'Deep ecotype' (Island 1)		MADE	1016775-1016910	25.7	41.0	1016775	1016780		Type I restriction-modification, IS*Ama4* transposase
Alteromonas macleodii str. 'Deep ecotype' (Island 2)		MADE	1017135-1017275	25.7	41.0	1017275	1017270		Type I restriction-modification, IS*Ama4* transposase
Bradyrhizobium sp. BTAi1	Island 10	BBta	3310-3456	140.4	62.7	3456	3455	3310	Conjugal transfer, cationic metal resistance, DNA replication
Burkholderia cenocepacia H111		I35	1589-1653	47.0	61.3	1589	1590		Phage lysogen
Burkholderia gladioli BSR3		bgla_1g	18670-18900	22.5	57.0	18670	18690	18680	Phage lysogen
Caulobacter sp. K31		Caul	2043-1864	184.8	66.5	2043	2042		Conjugal transfer, metabolism, multiple integrases
Colwellia psychrerythraea 34H		CPS	0353-0383	20.4	36.1	0353	0354	0369	Phage remnant, IS*Cps1* transposase
Escherichia coli 101-1		EC1011	3356-3393	27.8	50.8	3393	3392	3358	Phage lysogen
Escherichia coli 3003		EC3003	4528-4592	47.1	50.9	4592	4591	4530	Phage lysogen
Escherichia coli B171		EcB171	0688-0730	43.9	50.5	0730	0729	0670	Phage lysogen
Escherichia coli ETEC H10407		ETEC	4359-4304	39.6	50.0	4359	4358	4308	Phage lysogen
Escherichia coli KTE41		WGC	00045-00108	40.7	51.0	00108	00107	00048	Phage lysogen
Escherichia coli O157H7 str. EC4115		ECH74115	5530-5551	10.9	46.1	5551	5550	5532	Phage remnant
Escherichia coli PA40		ECPA40	5484-5503	10.9	46.1	5503	5502	5485	Phage remnant
Escherichia coli STEC_H.1.8		ECSTECH18	5069-5153	52.4	52.8	5153	5152	5071	Phage lysogen, IS3 transposase, 4 transfer RNAs
Glaciecola mesophila KMM 241		GMES	1194-1222	28.7	40.2	1222	1221		Type I restriction-modification
Herbaspirillium sp. CF444		PMI16	00553-00564	8.7	50.7	00564	00563		Phage remnant
Hoeflea phototrophica DFL-43		HPDFL43	10257-10387	32.7	57.4	10257	10262		Metabolism
Hyphomonas neptunium ATCC 15444		HNE	0283-0293	14.2	55.4	0293	0292		Unknown
Kingella oralis ATCC 51147		GCWU000324	01205-01259	35.1	53.7	01259	01258	01207	Phage lysogen
Labrenzia alexandrii DFL-11		SADFL11	1138-2029	5.6	47.6	1138	2278		Degraded genomic island
Mesorhizobium opportunistum WSM2075		Mesop	4302-4443	153.6	59.9	4443	4442		Conjugal transfer, cytochrome biogenesis
Neisseria meningitidis 73696		NM73696	0415-0353	43.8	51.9	0415	0414		Phage lysogen
Neisseria meningitidis ATCC 13091		HMPREF0602	0052-0118	41.7	51.1	0118	0117		Phage lysogen
Pandoraea sp. SD6-2		C266	17541-17581	9.0	56.5	17541	17546	17581	Unknown
Paracoccus denitrificans PD1222		Pden	1420-1548	134.6	61.8	1420	1421		Conjugal transfer
Parvibaculum lavamentivorans DS-1		Plav	1004-1050	58.0	58.9	1050	1049		Pseudoazurin production, nitrogen oxide metabolism
Phaeobacter gallaeciensis DSM 17395		RGBS107	16878-16743	37.1	54.4	16743	16748		Metabolism
Phenylobacterium zucineum HLK1		PHZ_c	2343-2442	135.0	67.7	2343	2344		Type IV secretion system, various metabolic enzymes
Pseudoalteromonas tunicata D2		PTD2	08004-08264	38.5	43.9	08264	08259	08009	Phage lysogen
Pseudogulbenkiania sp. NH8B		NH8B	0534-0598	42.6	62.8	0598	0597	0589	Phage lysogen
Pseudomonas aeruginosa MSH-10		LE46	1516-1588	55.1	62.1	1516	1517	01518	Phage lysogen
Pseudomonas aeruginosa PA7	RGP56	PSPA7	2362-2435	57.0	62.8	2362	2363	2364	Phage lysogen
Pseudomonas chlororaphis O6	Prophage 3	PchlO6	2091-2179	58.4	59.0	2091	2093	2092	Phage lysogen
Pseudomonas fluorescens BRIP34879		A986	15166-15361	35.4	56.3	15166	15171	15176	Phage lysogen
Pseudomonas fluorescens Q2-87	Prophage 2	PflQ2	3552-3598	38.6	58.7	3598	3596	3597	Phage lysogen
Pseudomonas fulva 12-X		Psefu	2114-2205	105.5	58.8	2114	2115	2116	Conjugal transfer, DNA replication and maintenance
Pseudomonas protegens CHA0		PFLCHA0_c	20300-20720	36.9	58.4	20300	20310	20320	Phage lysogen
Pseudomonas protegens Pf-5	Prophage 03	PFL	1975-2023	41.5	58.9	1975	1976		Phage lysogen
Pseudomonas pseudoalcaligenes KF707		ppKF707	4047-4141	113.6	61.8	4047	4048		Conjugal transfer, amino acid metabolism, DNA replication
Pseudomonas putida GB-1		PputGB1	1710-1763	44.8	60.9	1710	1711	1712	Phage lysogen
Pseudomonas putida H8234		L483	08090-08470	53.4	60.2	08090	08095	08100	Phage lysogen
Pseudomonas putida S16		PPS	1736-1780	42.2	59.2	1736	1737	1739	Phage lysogen
Pseudomonas resinovorans NBRC 106553		PCA10	35880-35170	71.0	62.1	35880	35870	35850	Arsenic resistance, metabolism
Pseudomonas sp. BG33R		PseBG33	1689-1699	11.1	52.2	1689	1691	1690	Unknown
Pseudomonas sp. GM60		PM132	02372-02396	17.3	52.8	02396	02395	02394	Phage remnant
Pseudomonas sp. GM74		PMI34	01891-01945	46.3	57.0	01891	01892	01894	Phage lysogen
Pseudomonas sp. GM80		PMI37	05674-05731	46.9	58.1	05731	05730	05729	Phage lysogen
Pseudomonas sp. P179		HMPREF1224	06117-06055	52.6	61.5	06117	06116	06115	Phage lysogen
Pseudomonas syringae pv. avellanae str. ISPaVe037		Pav037	1760-1772	16.2	50.7	1772	1771		Phage remnant
Rhizobium lupini HPC(L)		C241	02694-02994	43.0	56.3	02694	02699		Phage lysogen, tRNA-Met
Rhodobacteraceae bacterium KLH11		RKLH11	310-1236	48.2	55.4	310	1807	3025	Phage lysogen
Rhodomicrobium vannielii ATCC 17100		Rvan	1748-1799	49.3	60.3	1748	1749	1797	Conjugal transfer, various transposases
Ruegeria sp. R11		RR11	404-1812	49.3	56.7	1812	3187		Phage lysogen
Shewanella baltica OS155		Sbal	3585-3633	42.5	46.0	3585	3586	3620	Phage lysogen
Shewanella baltica OS185		Shew185	0722-0774	42.3	46.2	0774	0773	0734	Phage lysogen
Shewanella baltica OS195		Sbal195	0751-0806	43.8	45.3	0806	0805	0763	Phage lysogen
Shewanella baltica OS223		Sba1223	0742-0799	42.3	46.1	0799	0798	0759	Phage lysogen, IS*Sba2* tranposase
Shewanella denitrificans OS217		Sden	3232-3281	42.5	45.3	3232	3233	3273	Phage lysogen
Shewanella sp. MR-7		Shewmr7	0699-0746	33.2	48.6	0746	0745	0703	Phage lysogen
Sinorhizobium fredii HH103		SFHH103	01636-01689	42.2	62.0	01636	01637	01689	Phage lysogen
Sinorhizobium fredii USDA 257		USDA257_c	41740-41790	4.8	57.5	41740	41750		Degraded genomic island, DNA ligase
Sinorhizobium medicae WSM419		Smed	1621-1683	50.6	61.2	1621	1622	1680	Phage lysogen
Sinorhizobium meliloti AK83		Sinme	1811-1876	49.2	60.8	1811	1812	1873	Phage lysogen
Sinorhizobium meliloti BL225C		SinmeB	1656-1716	48.5	61.2	1656	1657	1714	Phage lysogen
Sinorhizobium meliloti SM11		SM11_chr	1426-1491	48.4	60.9	1491	1490	1426	Phage lysogen
Starkeya novella DSM 506		Snov	0848-0872	24.9	66.5	0872	0871		Alkanesulphonate metabolism
Variovorax paradoxus S110		Vapar	2059-2122	63.0	64.1	2122	2121		DNA replication
Vibrio cholerae O1 str. EM-1676A		VCEM1676A	000425-000456	36.5	42.7	000456	000455		Fimbrial biogenesis
Vibrio cholerae TMA 21		VCB	000225-000242	28.2	43.9	000225	000226	000241	Conjugal transfer
Vibrio mimicus VM223		VMA	000316-000303	12.1	46.6	000316	000315		Phage remnant, DNA replication
Vibrio parahaemolyticus BB22OP		VPBB	2545-2586	33.0	49.8	2545	2546	2550	Phage lysogen
Vibrio shilonii AK1		VSAK1	13606-13831	35.7	43.5	13606	13611	13636	Phage lysogen
Vibrio sinaloensis DSM 21326		VISI1226	20525-20665	14.4	45.9	20665	20660	20540	Phage lysogen
Vibrio sp. 16		VPMS16	1054-1081	14.9	47.0	1081	1080	1074	Unknown
Vibrio sp. RC341		VCJ	000612-000632	12.1	46.3	000612	000613		Phage remnant, DNA replication
Vibrio splendidus 12B01		V12B01	10660-10775	14.1	44.0	10660	10665	10760	Phage remnant, DNA replication
Xanthomonas albilineans GPE PC73		XALc	0171-0242	52.1	59.9	0242	0241		Phage lysogen
Xenorhabdus bovienii SS-2004		XBJ1	3925-3991	42.3	48.3	3925	3926	3991	Phage lysogen, tRNA-Phe

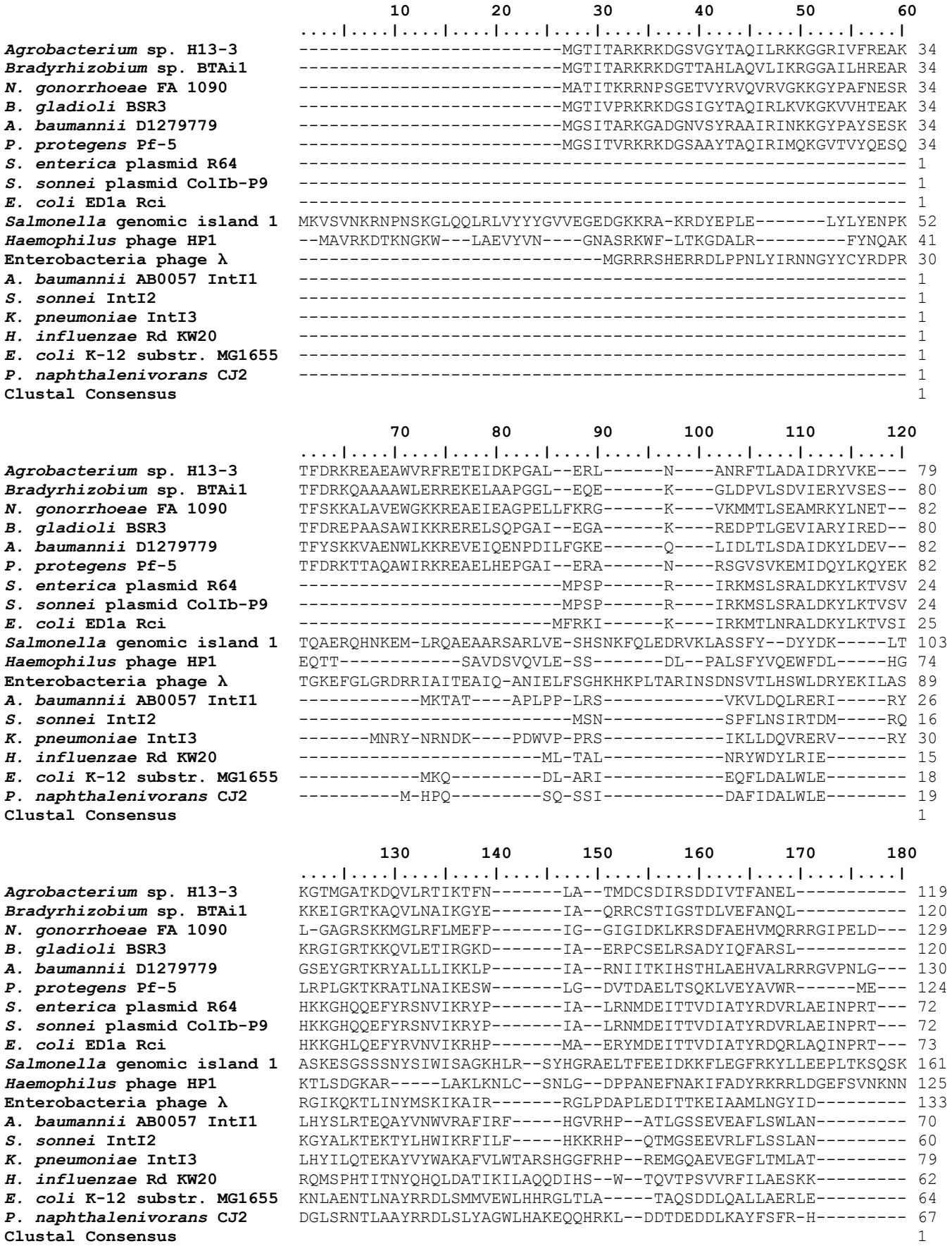
Supplementary Table 3.2: Oligonucleotide sequences used in the detection of excised *dusA*-specific genomic islands and generation of markerless gene deletions. Relevant restriction sites are underlined.

Oligonucleotide	Sequence (5'→3')	Size (bp)
D1279779_Excison_FOR	GAGAAGCTTAGGAAATCAACGACAG	740
D1279779_Excison_REV	TCTATGGCATCTGACAAAGTTAGGTC	
D1279779_Junction_FOR	ACTTCGTCATAACCCCAGTCTT	494
D1279779_Junction_REV	GTGGGACTCGTTGTATTTTGTA	
ACICU_Excison_FOR	ATGTAGCATTATGGCTGAAGAG	534
ACICU_Excison_REV	GCTAGTGATTTCTTACTGAATGTCT	
ACICU_Junction_FOR	TAACTCTCCATCAAAATTAGCAC	534
ACICU_Junction_REV	TGTCTAAGTCTTTCATAATGTCATC	
PF5_Excison_FOR	TATCTGAAGAGAGGGCAGAATA	472
PF5_Excison_REV	CTTTCCTGATAAACTGTGACG	
PF5_Junction_FOR	CACCGATCATGTTGTTCTGCAC	485
PF5_Junction_REV	ATACTCGGACAACCTCGCTAGAC	

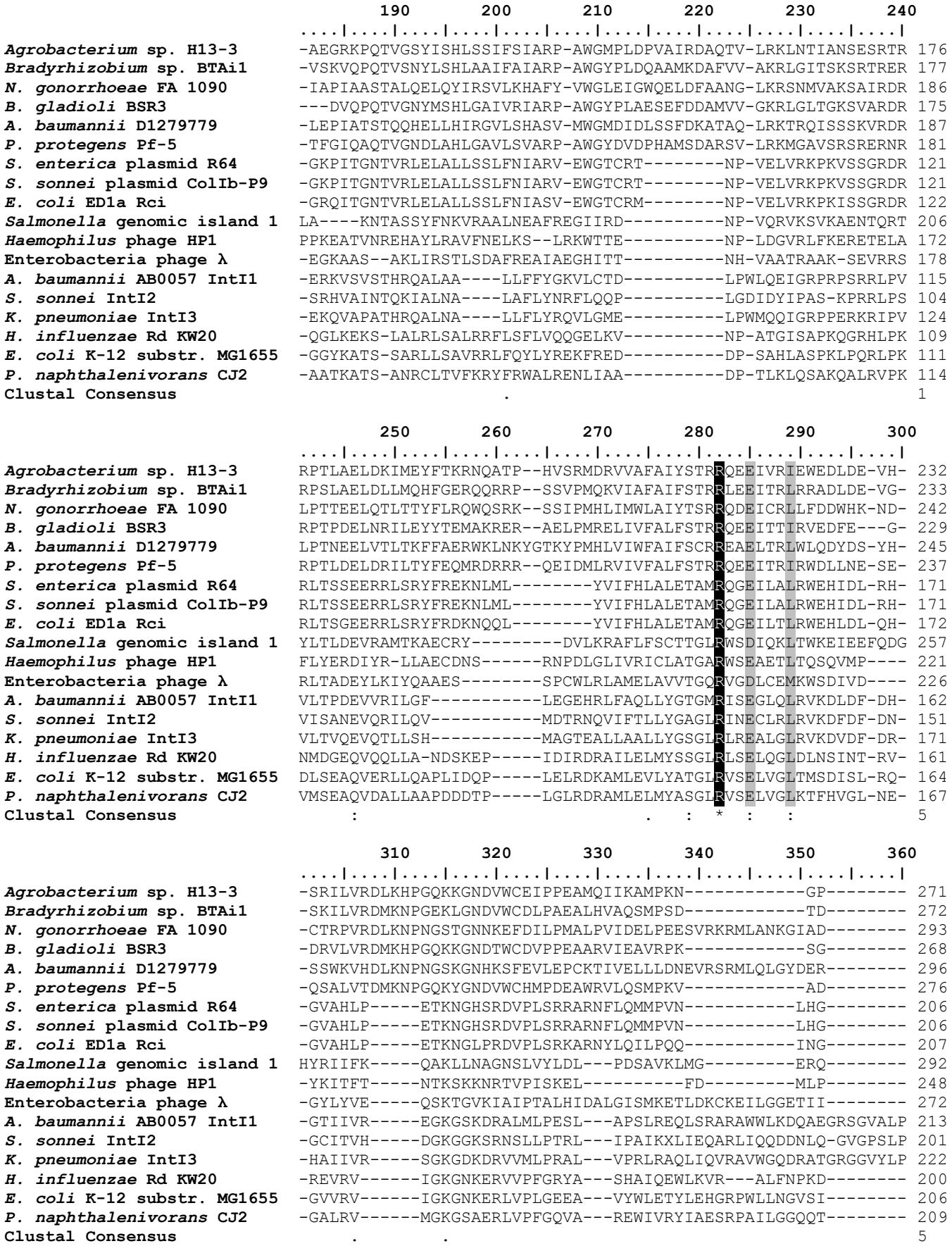
RGP05_UP_FOR	AGATTATT**GGATCC**TTAATGAGTGGC	947
RGP05_UP_REV	CGATTGAAGTGA**GAGCTC**AACCGATA	
RGP05_DWN_FOR	GAACACCAA**GGATCC**GCCAGCCCATG	965
RGP05_DWN_REV	CTCTATCTTT**TCCGGA**AGTTTTACTGG	
P1_RGP05_FOR	AGCTCAAGTGCTTATTCCTCT	50324
P4_ RGP05_REV	ATCTTACCTGCCAATGTAAAAC	
FRT-T7F	AAATTAATACGACTCACTATAGG	229
FRT-SP6R	TACGATTTAGGTGACACTATAG	
FRT-leftF	AATCCATCTTGTTCAATCATGC	N/A
FRT-rightR	AATTCGAGCTCGGGAAGATC	

Abaumannii_dusA_FOR	ATC**GGATCC**TAATTTTGAATATTTCGAGCC	1253
D1279779_dusA_REV	CCATT**GGATCC**ATTATGGTGTAAATTTTGG	
ATCC17978_dusA_REV	TCA**GGATCC**TATATTTAACATAAGGGC	1242
pWH1266_ColonyScreen_FOR	CTTCGCTACTTGGAGCCACTAT	152
pWH1266_ColonyScreen_REV	ATCTTCCCCATCGGTGATGT	

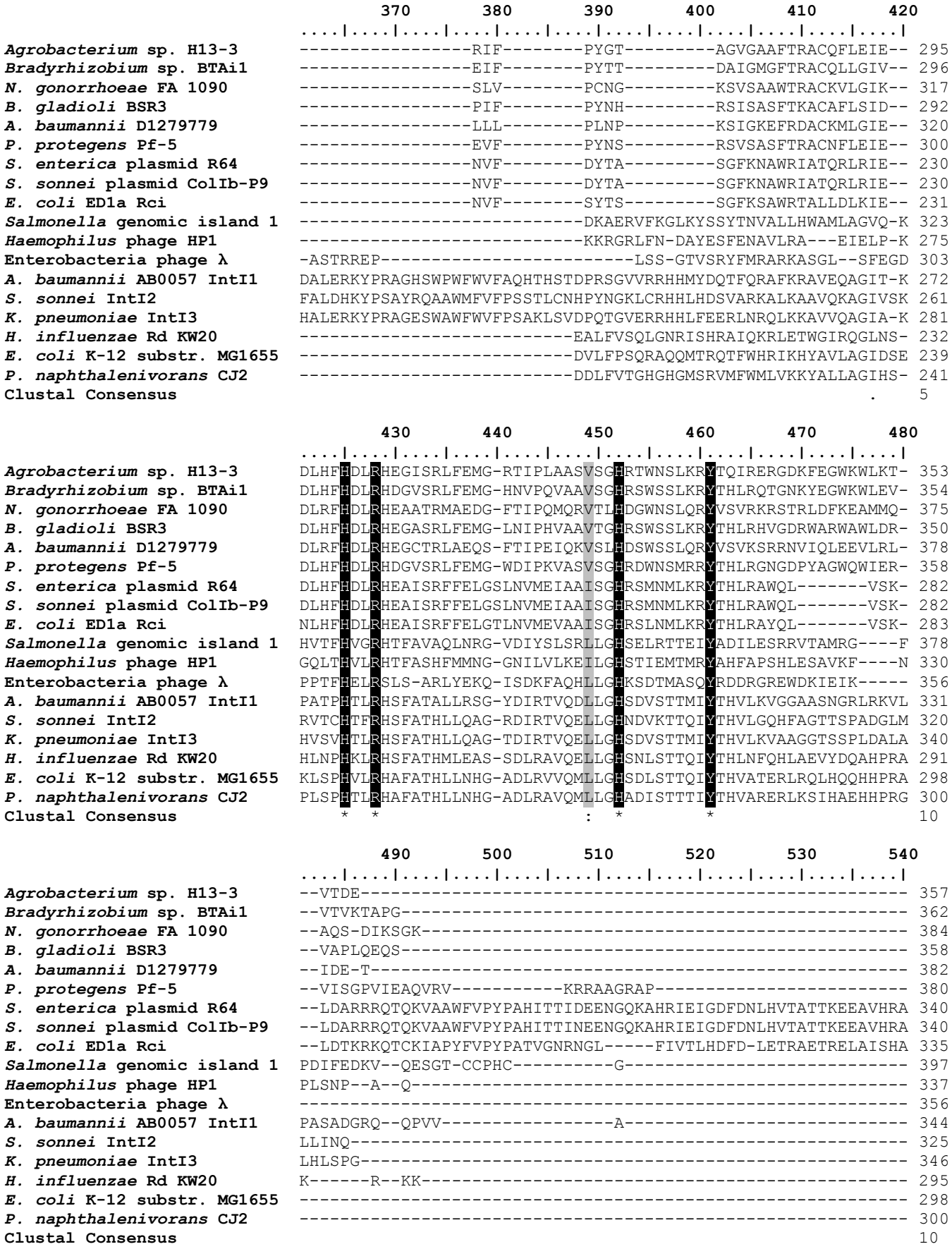
Supplementary Figure 3.1



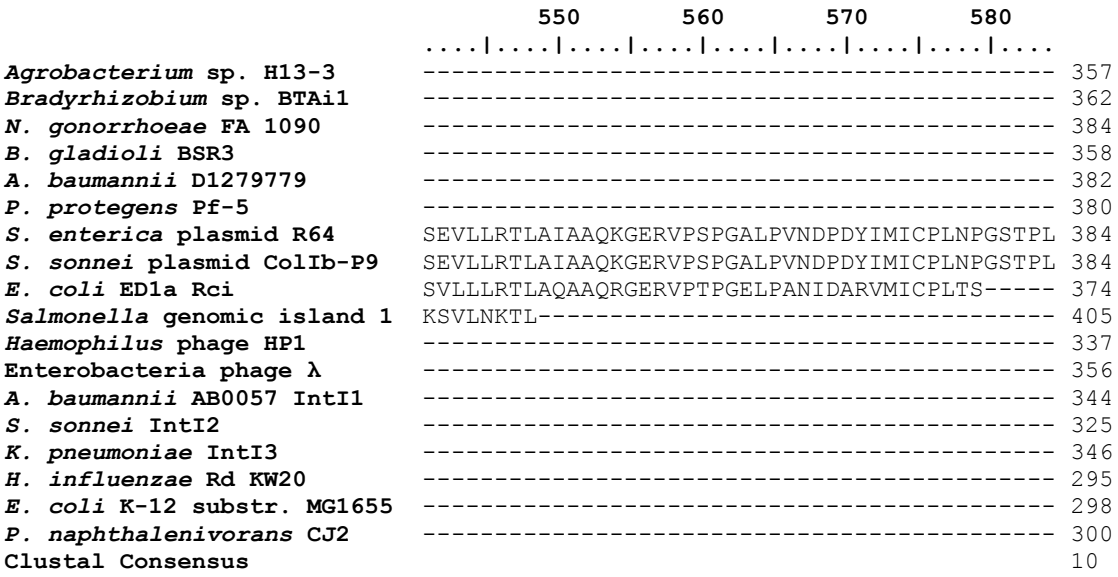
Supplementary Figure 3.1



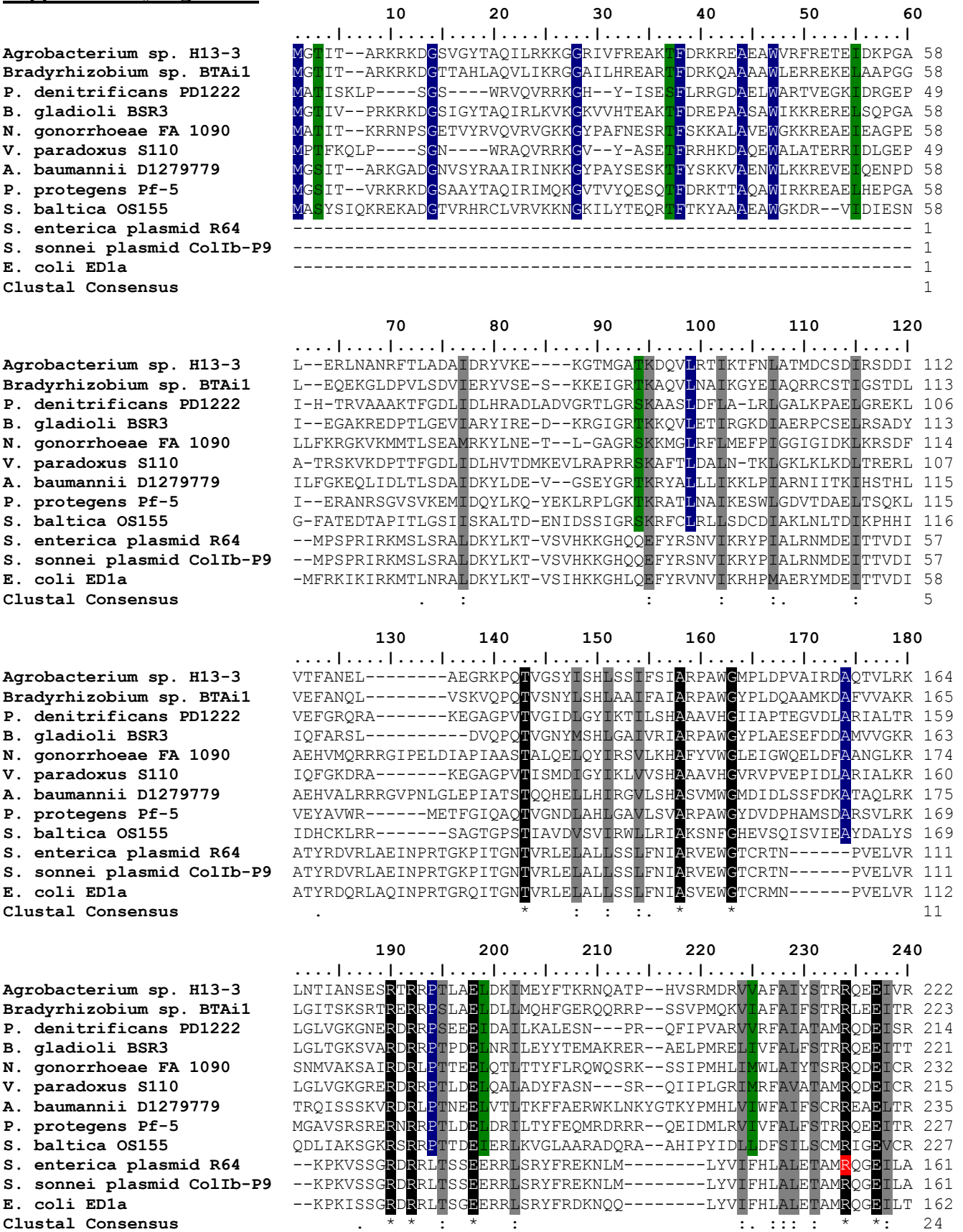
Supplementary Figure 3.1



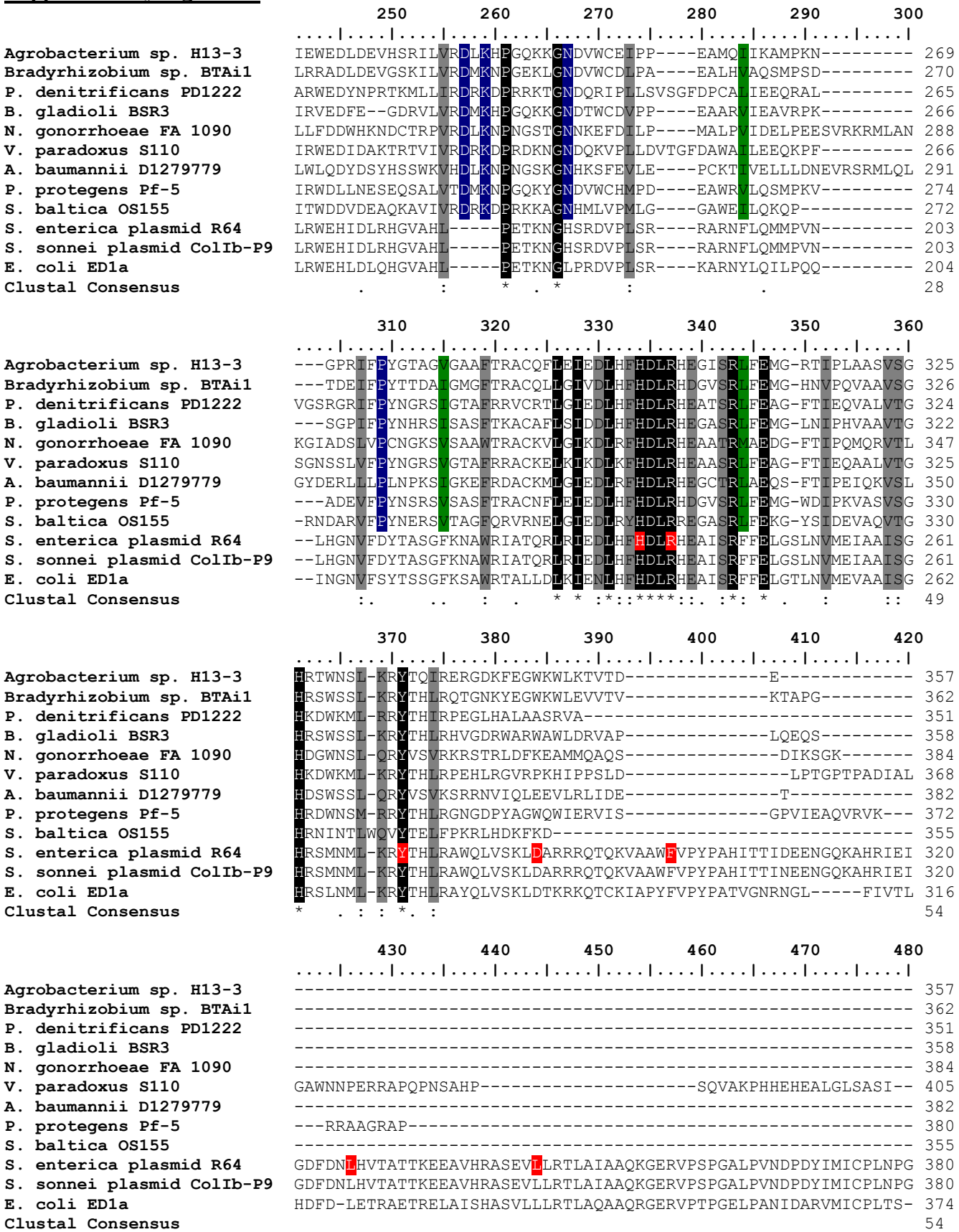
Supplementary Figure 3.1



Supplementary Figure 3.2



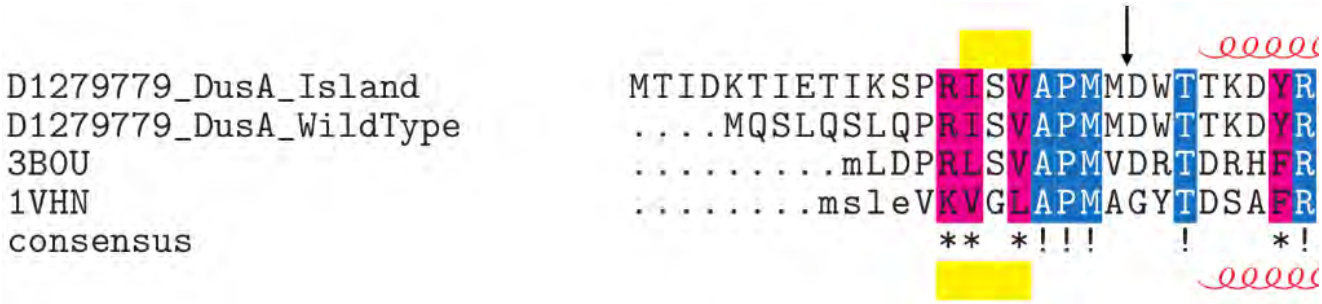
Supplementary Figure 3.2



Supplementary Figure 3.2

```
                                ....
Agrobacterium sp. H13-3         ---- 357
Bradyrhizobium sp. BTAi1        ---- 362
P. denitrificans PD1222         ---- 351
B. gladioli BSR3                ---- 358
N. gonorrhoeae FA 1090          ---- 384
V. paradoxus S110               ---- 405
A. baumannii D1279779           ---- 382
P. protegens Pf-5               ---- 380
S. baltica OS155                ---- 355
S. enterica plasmid R64         STPL 384
S. sonnei plasmid ColIb-P9      STPL 384
E. coli ED1a                    ---- 374
Clustal Consensus                    54
```

Supplementary Figure 3.4



Chapter 4: A novel family of genomic resistance islands, AbGRI2, contributing to aminoglycoside resistance in *Acinetobacter baumannii* isolates belonging to global clone 2

4.1 Non-technical summary

The genome sequences of nosocomial *Acinetobacter baumannii* are found to often possess an AbaR genomic island, inserted into the *comM* gene, which encode several genes involved in antimicrobial resistance and mobilisation, including a class 1 integron array. The AbaR sequences of the ICII isolates *A. baumannii* WM99c and A91 were characterised, but were found to be lacking the class 1 integron array. Instead it was flanked by two IS*26* insertion sequences and had replaced a 40.9 kb loci present in other *A. baumannii* genomes. This genomic island was dubbed AbGRI2 to distinguish it from other previously characterised antimicrobial resistance islands.

Pages 99-102 of this thesis have been removed as they contain published material. Please refer to the following citation for details of the article contained in these pages.

Steven J. Nigro, Daniel N. Farrugia, Ian T. Paulsen, Ruth M. Hall (2013) A novel family of genomic resistance islands, AbGRI2, contributing to aminoglycoside resistance in *Acinetobacter baumannii* isolates belonging to global clone 2, *Journal of Antimicrobial Chemotherapy*, 68(3), pp. 554–557.

DOI: 10.1093/jac/dks459

Chapter 5: Genome sequencing and analysis of nosocomial and extranosocomial *Acinetobacter baumannii*

5.1 The accessory genome and phenome of *Acinetobacter baumannii* WM99c

5.1.1 Background

Acinetobacter baumannii WM99c is a multiply drug resistant sputum isolate, originating from a 1999 outbreak at Westmead Hospital (Sydney, Australia) [238]. It was discovered that this organism is of the ICII lineage, a dominant genomic class of carbapenem-resistant, epidemic-associated *A. baumannii* [239]. The unannotated whole shotgun sequence of this organism was recently determined and deposited by our laboratory group [117] (GenBank accession: AERY00000000), and is in 162 contiguous sequences, totalling 4,042,799 bp. Automated annotation of the genome through the RAST sever [240] uncovered 3809 CDSs and 56 tRNAs. The genome of this organism was examined for accessory elements, with some of further interest outlined below.

5.1.2 Materials and Methods

Materials and methods regarding DNA manipulation, whole genome sequencing, assembly, annotation, analysis, comparative genomics, accessory element identification, and phenotype microarrays were conducted in an identical fashion to a previous study [241] (Chapter 2).

5.1.3 Results and Discussion

5.1.3.1 Plasmid pWM99c

One of the contiguous sequences of this genome (Genbank accession: AERY01000162) contained some features that were suggestive of it being a plasmid, including a gene encoding a DNA replication protein (*repB*) and the relatively high coverage level of this contig compared to the rest of the genome (43.7x vs. ~17x average). PCR amplification using outward facing primers confirmed that this contig was indeed a plasmid, which was named pWM99c. pWM99c is 10,967 bp in size and harboured ten putative protein coding genes (≥300 bp) (Figure 5.1.1), including *repB*, and genes encoding a TonB-dependent receptor, a DNA relaxase, and an insertion sequence (IS*Aba25*).

TonB-dependent receptors rely on TonB-mediated energization to transport iron bound with various compounds, including siderophores, heme [242] or small and simple molecules such as citrate. TonB-dependent receptors are known to be encoded in other *A. baumannii* plasmids [243], and contribute to virulence in this pathogen [244] and other organisms [242, 245]. DNA relaxases initiate bacterial conjugation by binding the origin of transfer (*oriT*) sequence, followed by cleavage of the T-strand, the DNA strand destined for transfer [246]. This plasmid was previously found incapable of conjugal transfer [238], though pWM99c could potentially be mobilised *in trans* by conjugative transfer genes encoded either chromosomally or on another plasmid. The IS*Aba125* insertion sequence present in pWM99c interrupted a gene encoding a hypothetical protein (ABWM_p00070) (Figure 5.1.1). In other *A. baumannii* isolates, this IS has been known to activate expression of latent chromosomal β-lactams [175, 247] and may have been instrumental in the genesis and mobilisation of bla_{NDM} genes [248, 249]. A 6,402 bp segment of pWM99c is conserved (99% nucleotide identity) in the *A. baumannii* plasmids pAB0057 and p1ABTCDC0715 (Figure 5.1.1).

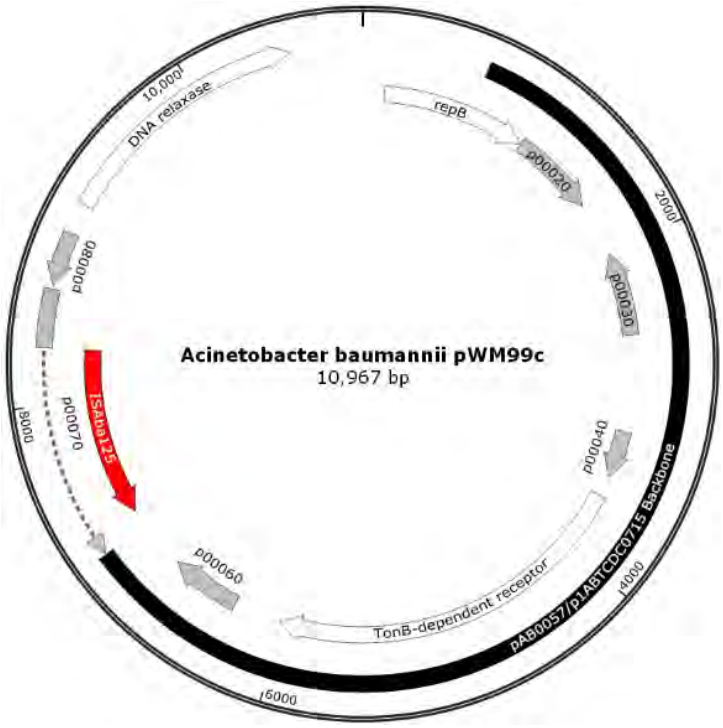


Figure 5.1.1: *Acinetobacter baumannii* WM99c plasmid pWM99c. Genes are numbered clockwise from the origin, with its locus tag prefix (ABWM_) omitted for brevity. Genes are coloured according to annotation: hypothetical genes (light grey), putative function (white) and insertion sequences (red). The region in black is the plasmid backbone common to *A. baumannii* plasmids pAB0057 and p1ABTCDC0715, at a nucleotide identity of 99%. This image was generated with SnapGene Viewer.

5.1.3.2 Variable region G41

The genome of *A. baumannii* WM99c was examined for the presence of additional accessory elements, such as genomic islands. A *dusB*-specific DAMGE island (Chapter 3) variant was identified in this organism that encoded a putative prophage, and was nearly identical in sequence to the island in *A. baumannii* ACICU (nucleotide identity of 94%). The difference in the two islands was due to variation in the sequence of a large gene annotated as a phage tail component (ACICU_02717), thought to be involved in phage host specificity [250]. Two other genomic islands, Tn*AbaR* and ABGRI2, involved in antimicrobial resistance, were identified in this organism and encoded resistance to antibiotics of the aminoglycoside, β-lactam, sulphonamide and tetracycline classes (Chapter 4), [251]. However, there was one putative genomic island that was of interest to our laboratory group, as it was the target of a structural genomics study that examined genes identified as having been laterally transferred.

A variable region ($G41_{abc}$) previously identified in the genome of another ICII strain, *A. baumannii* ACICU [167], was discovered in the genome of *A. baumannii* WM99c. Though dubbed a genomic island [167], $G41_{abc}$ could more aptly be described as a region of genomic plasticity, as there was an absence of any features associated with mobilisation, including genes encoding integrases or transposases, nor were there any discrete breakpoints associated with integration or recombination, such as direct or inverted repeats.

$G41_{abc}$ spans 12.3 kb and contains eleven genes (Figure 5.1.2) including an aminoglycoside phosphotransferase (*aph*), a fructose-2,6-bisphosphatase (*gpm*), and several genes putatively involved in fatty acid metabolism (including *actP*, *calB*, *prpR*), including a short chain dehydrogenase/reductase (*SDR*). The structure of SDR (locus tag: ABWM_00001) has recently been determined (PDB Accession: 4IUY) by our collaborators, Dr. Bhumika Shah and A/Prof. Bridget Mabbutt. *A. baumannii* WM99c G41 was identical in nucleotide sequence to *A. baumannii* ACICU $G41_{abc}$, with the exception of an apparent frameshift in a gene encoding acyl-CoA dehydrogenase (ACICU_1944/1945). Amplification and resequencing of this genomic region in *A. baumannii* ACICU revealed the frameshift mutation to be the result of a sequencing error in the original published sequence. Examination for further instances of G41 within the fifteen currently compete genomes of *A. baumannii* (Table 1.1 and *A. baumannii* D1279779) revealed that it was exclusive to *A.*

baumannii isolates from the ICII lineage, which aside from WM99c and ACICU included the strains 1656-2, BJAB07104, BJAB0868, MDR-TJ, MDR-ZJ06, TCDC-AB0715 and TYTH-1.

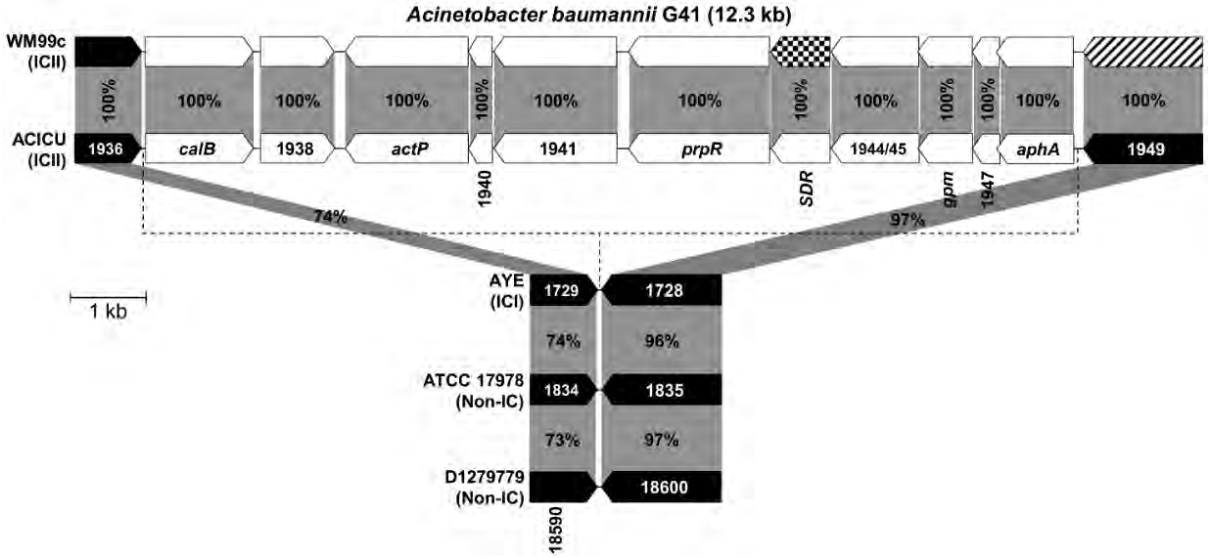


Figure 5.1.2: The genetic organisation of *Acinetobacter baumannii* variable region G41. G41 is present only in the complete genome sequences of international clone two (ICII) lineage *A. baumannii*. Chromosomal genes that flank the variable region are indicated in black. However, the rightmost gene in WM99c is truncated (striped), as it lies at the end of the contiguous sequence. The gene encoding the crystalized WM99c-SDR is shaded as checkerboards. The percentages denote the protein identity of homologous genes relative to WM99c; the other numbers represent the gene locus tags from each respective genome sequence.

G41 is flanked by two chromosomal genes, ACICU_01949, a gene encoding a NAD-dependent aldehyde dehydrogenase that is highly conserved in all analysed *A. baumannii* (Figure 5.1.2). The second gene is ACICU_01936, which encodes a hypothetical protein that has a lower level of sequence identity in ICII *A. baumannii* compared to non-ICII isolates (Figure 5.1.2). This may suggest that this RGP was acquired through a homologous recombination event involving this gene. *A. baumannii* SDF like other non-ICII isolates lacks G41, however it is the only completely sequenced isolate to lack the ACICU_01936 and ACICU_01949 genes, which along with several other genes have been replaced by an IS*Aba7* insertion sequence, a presumed contributor to the extensive genome decay observed in this organism [160].

5.1.3.3 The catabolic phenome of *A. baumannii* WM99c

The carbon utilisation phenotypic profile of *A. baumannii* WM99c was assayed with the Biolog Phenotype MicroArray System to observe phenotypes that were differential compared to previously assayed *A. baumannii* strains [241] (Chapter 2), and to potentially link any phenotypic differences to the WM99c genome sequence.

There were a total of fifty-four compounds potentially utilisable by WM99c; its phenotypic profile most resembled that of *A. baumannii* ACICU, another ICII strain, whilst being similar overall to other *A. baumannii* previously assayed [241] (Chapter 2). The only major phenotypic differences between WM99c and ACICU observed were in regards to the utilisation of L-ornithine and D,L-carnitine (Figure 5.1.3).

A comparison of the WM99c and ACICU genomes could not account for differences in the two strains to utilise these two carbon sources. In the case of L-ornithine, alterations in regulation or membrane transport activity could be responsible for the inability to utilise this compound. The *carO* gene is known to encode an outer membrane protein essential for the passive uptake of L-ornithine in *A. baumannii* [66], however this gene was not disrupted in WM99c. In the study that examined the role of CarO in ornithine uptake, it was found that lower growth temperatures reduced CarO production and impaired utilisation of ornithine, suggesting a role of an as yet unidentified temperature sensitive regulator [66] in *carO* regulation. Changes in this unknown regulator may be related to the difference in phenotype that we observed.

As for D,L-carnitine, it is unknown why WM99c is able to utilise the racemic form, whilst ACICU is not, as both genomes encode a carnitine degradative pathway (ACICU_00801-00807) [252]. We previously confirmed through independent phenotypic testing that ACICU was only able to utilise L-carnitine, which is not present in the PM assay, whilst another *A. baumannii* strain (D1279779), was able to utilise both the laevorotated and racemic forms of carnitine [241] (Chapter 2). It is unknown what the molecular basis for non-stereospecific carnitine utilisation in *A. baumannii* D1279779 and WM99c.

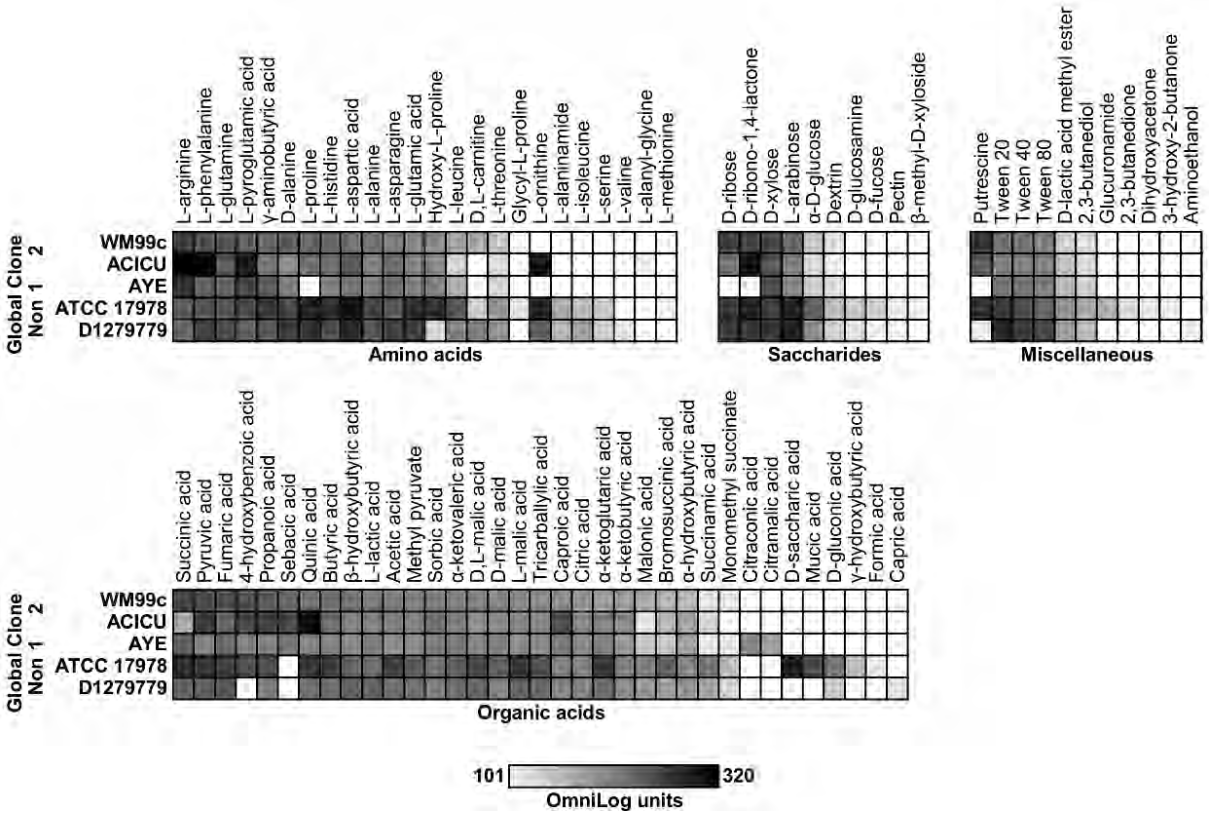


Figure 5.1.3: Carbon utilisation phenotypic profile of *A. baumannii* WM99c generated using Biolog Phenotype Microarray plates PM1 and PM2 compared to previously determined phenotypic profiles of other *A. baumannii* [241] (Chapter 2). Phenotype strengths are represented as a greyscale of the maximal kinetic curve height (expressed as OmniLog units), and are arranged in descending order of strength, relative to WM99c. Phenotypes <101 OmniLog units (white) were considered to be negative.

5.1.3.4 The resistance phenome of *A. baumannii* WM99c

Further to profiling sole carbon source utilisation, PMs were used to assay the antimicrobial resistance profile of *A. baumannii* WM99c. Respiration was observed, at all four concentrations, for 144 of the 240 antimicrobial compounds. No respiration at any of the four concentrations was observed for nine compounds: 2,2`-dipyridyl, the antibiotics novobiocin and rifamycin SV, the antifungals dichlofluanid and tolylfluanid, and the heavy metal salts cadmium chloride, potassium chromate, sodium metavanadate and sodium orthovanadate. *A. baumannii* WM99c demonstrated a phenotypic profile typical of a multidrug resistant *A. baumannii*, however, it appeared to be notably more resistant to β-lactams and tetracyclines compared to other multidrug resistant isolates, such as *A. baumannii* ACICU and AYE (Figure 5.1.4). The phenotypic profile of WM99c can primarily be attributed to the previously outlined AbaR and AbGRI2 resistance islands (Chapter 4), which encoded resistance to antibiotics of the aminoglycoside, β-lactam, sulphonamide and tetracycline classes [251].

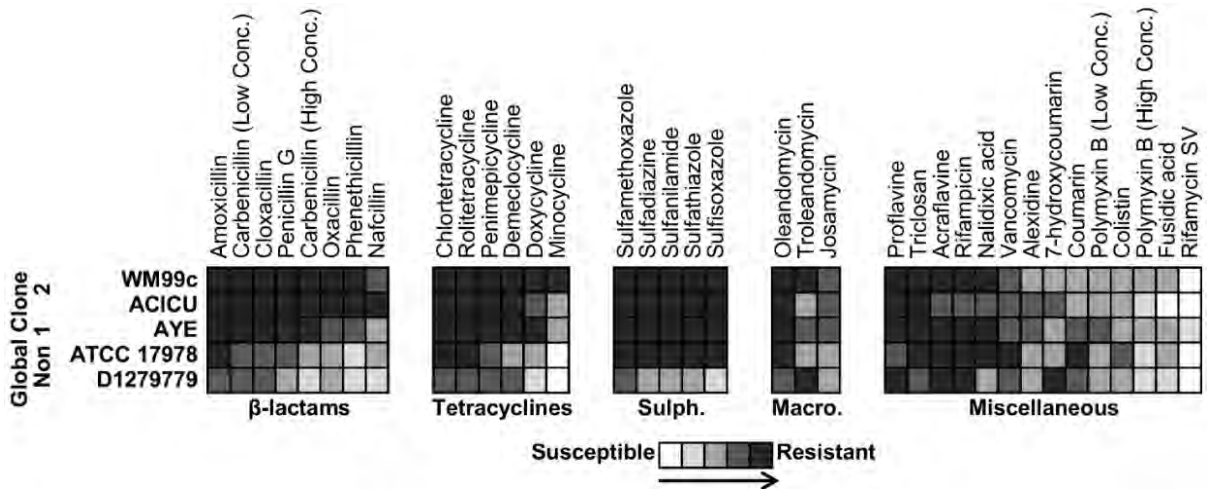


Figure 5.1.4: Highlights of antimicrobial resistance profiling (PM11-20) of *A. baumannii* WM99c compared to previously profiled *A. baumannii* strains [241] (Chapter 2), with emphasis on clinically important antimicrobials. Resistance is expressed as a five-coloured greyscale from white (no resistance) to dark grey (maximal resistance), arranged in descending order of strength, relative to WM99c. Abbreviations: Sulph., sulphonamides.; Macro., macrolides.

5.2: Virulome comparisons of *Acinetobacter baumannii* originating from nosocomial and community-acquired infections in the Asia-Pacific

5.2.1 Background

The genome sequences of six Asia-Pacific *A. baumannii* isolates (CA1, CA7 CA17, HA17, HK4620 and HK4751), originating from either nosocomial or extranosocomial infections, were provided by our collaborator, A/Prof. Anton Peleg, in order to perform bioinformatic analysis. All of the strains prefixed with 'CA' originated from community-acquired infections in Townsville (Queensland, Australia), whilst HA17 is a Queensland nosocomial isolate that was sequenced for comparison. *A. baumannii* HK4620 and HK4751, originated in Hong Kong from a community-acquired infection and a nosocomial infection, respectively. The Peleg laboratory group conducted several virulence and phenotypic assays on these isolates, as well as ATCC 17978 and AB307-0294, two previously sequenced nosocomial *A. baumannii* isolates. They had discovered that the CA-AB isolates, particularly CA17, generally outperformed the nosocomial isolates in respect to biofilm formation, phospholipase D production, murine virulence, and *in vitro* growth in LB, sera and 1% ethanol. The genome sequences of these six *A. baumannii* isolates were examined to determine whether the occurrence of particular virulence genes could be responsible for the observed differences in phenotype.

5.2.2 Materials and Methods

Methods regarding phylogenetics, genome analysis, comparative genomics, and accessory element identification were conducted in an identical fashion to a previous study [241] (Chapter 2).

5.2.3 Results & Discussion

5.2.3.1 Genomic features

The genome features of the four CA-AB and two nosocomial isolates sequenced are summarised in Table 5.2.1. The genome sizes of the CA-AB isolates are all smaller compared to the genomes of nosocomial isolates sequenced in this study and previously (Table 1.1). This was also congruent with the genome sequence of *A. baumannii* D1279779, the only CA-AB isolate completely sequenced to date [241] (Chapter 2). There were an absence of both ISs and plasmids in two CA-AB isolates, CA7 and CA17.

Table 5.2.1: Comparative genome features of community-acquired and nosocomial *Acinetobacter baumannii* isolates.

	CA1	**CA7**	**CA17**	**HA17**	**HK4620**	**HK4751**
Infection type	CA[1]	CA	CA	N[2]	CA	N
Total genome size (bp)[3]	3676426	3864731	3690077	3975982	3878314	4023015
# Plasmids	2	0	0	2	2	1
G+C content (%)	38.8	38.8	38.9	39.0	38.9	38.9
# Protein coding sequences (CDSs)	3442	3410	3409	3739	3653	3804
# Insertion sequences (ISs)	66	0	0	51	14	47
# tRNAs	55	38	44	38	53	38
Number of contigs (≥1 kbp)	99	25	20	71	45	90

1. Community-acquired
2. Nosocomial
3. Genome size including putative plasmids

5.2.3.2 The phylogeny of *Acinetobacter baumannii*

The phylogeny of the six sequenced *A. baumannii* isolates was inferred, with respect to the fifteen previously completed *A. baumannii* genomes, using a multilocus sequence typing (MLST) approach [37]. None of the isolates analysed in this work clustered with any of the three major clonal lineages of nosocomial *A. baumannii*, with the exception of *A. baumannii* HA17, which was found to be a member of the international clone 2 (ICII) lineage (Figure 5.2.1). Nor did any of the CA-AB isolates strongly cluster with a previously sequenced CA-AB isolate, *A. baumannii* D1279779 (Chapter 2).

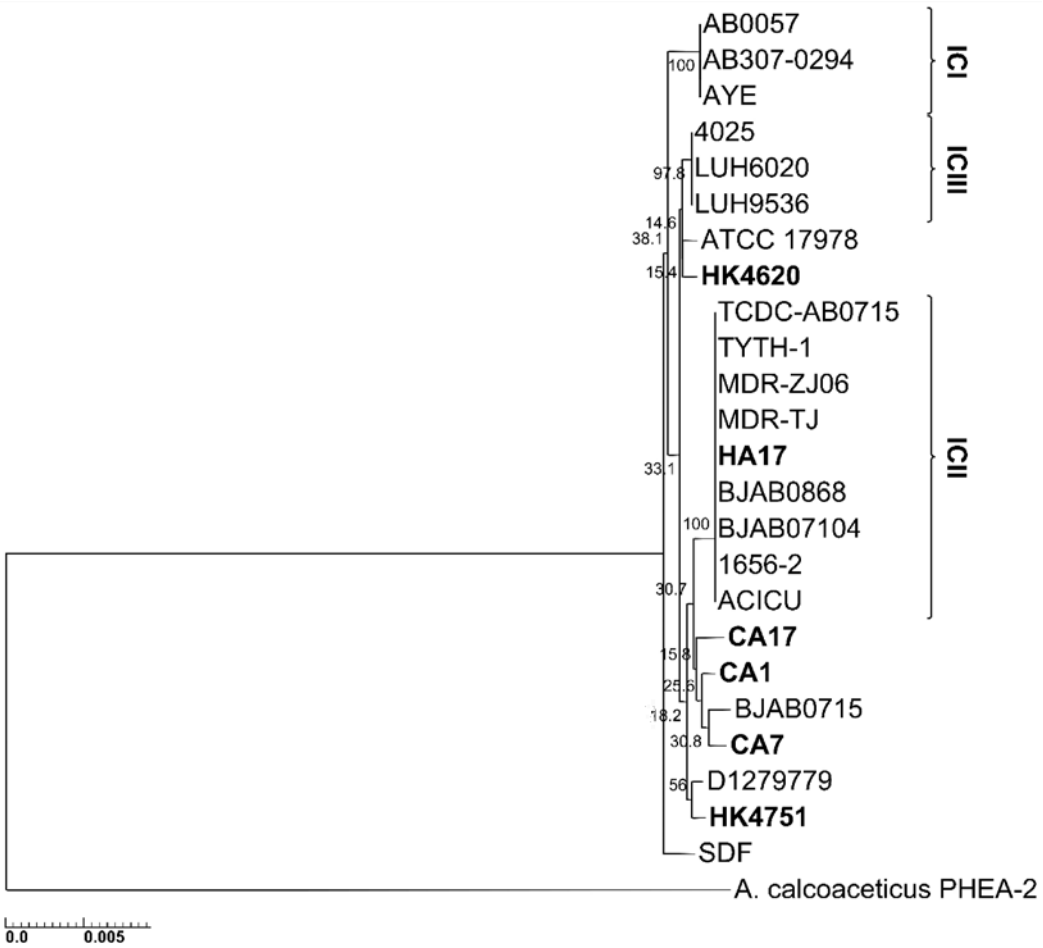


Figure 5.2.1: The phylogenetic relationship of community-acquired and nosocomial *Acinetobacter baumannii* sequenced in this study (bold text) compared to completely sequenced reference strains and international clone III (ICIII) isolates. The phylogeny was inferred using the concatenated partial nucleotide sequences of seven reference genes (*cpn60*, *fusA*, *gltA*, *pyrG*, *recA*, *rplB* and *rpoB*) based on the multilocus sequence typing scheme of *A. baumannii* [37]. The phylogeny was calculated in MEGA5 [253] using the neighbour-joining method (bootstrapped; 1000 replicates) with the Jukes and Cantor substitution model. The tree was drawn and annotated in TreeGraph2 [254].

5.2.3.3 The virulome of community-acquired *Acinetobacter baumannii*

The predicted proteome of the six Asia-Pacific *A. baumannii* isolates, and fifteen completely sequenced *A. baumannii* isolates, including ATCC 17978 and AB307-0294, were compared using BLASTP+ [255] ($e10^{-5}$) to determine the presence of known virulence genes. A total of 1820 CDSs were shared between all *A. baumannii* genomes, which increased to 2342 CDSs when SDF was excluded. The clone III isolates used in the phylogenetic tree (Figure 5.2.1) have yet to be sequenced, so they were not included in this analysis.

Through this analysis it was found that the majority of genes known to be involved in nosocomial *A. baumannii* virulence were present in all analysed CA-AB isolates. However, *A. baumannii* CA1 was notably lacking several virulence genes involved in quorum sensing, motility and biofilm formation (Table 5.2.2). The differences

between the CA-AB isolates and the nosocomial isolates in the phenotypic and virulence assays could not definitively be attributed to the presence or absence of known virulence genes.

Table 5.2.2: The presence (Y), partial presence (P) or absence (N) of known and putative *Acinetobacter baumannii* virulence factors in the genome sequences of this study.

Gene	A1S_[1]	Product	Virulence role	CA1	CA7	CA17	HA17	HK4620	HK4751	AB307-0294	Reference
plc1	0043	Phospholipase C	Eukaryotic cell damage	Y	Y	Y	Y	Y	Y	Y	[256]
abaI	0109	AHL lactone synthetase	Quorum sensing	N	Y	Y	Y	Y	Y	Y	[115][3]
abaR	0112	LuxR-family receptor		N	Y	Y	Y	Y	Y	Y	
	0112-119	Lipopeptide and polyketide synthesis	Surface-associated motility and biofilm formation	N	Y	Y	Y	Y	Y	Y	[104, 112]
pbpG	0237	Penicillin-binding protein 7/8	Serum resistance	Y	Y	Y	Y	Y	Y	Y	[257]
filA-F	0690-95	P pilli biogenesis	Biofilm formation	N	Y	Y	P[6]	Y	P	Y	[258]
nfuA	0979	Fe-S scaffold protein	Iron acquisition	Y	Y	Y	Y	Y	Y	Y	[259]
ata	1032	*Acinetobacter* trimeric autotransporter	Adhesion and biofilm formation	Y[7]	N	N	Y[7]	Y[7]	Y[7]	N	[260]
	1292-312	Type VI secretion system	Bacterial competition	P[6]	Y	Y	Y	Y	Y	Y	[261][3]
acuABCDE	1507-10	Type I pili biogenesis	Biofilm formation	Y	Y	Y	Y	Y	Y	Y	[117]
	1647-57	Siderophore biosynthesis cluster	Iron acquisition	Y	Y	Y	Y	Y	Y	Y	[117]
pgaABCD	2160-62	PNAG biosynthesis	Biofilm formation	Y	Y	Y	Y	Y	Y	Y	[110]
csuA/BABCDE	2213-21	Type I pili biogenesis	Biofilm formation and attachment	Y	Y	Y	Y	Y	Y	Y	[111]
bauA-E; basA-J	2372-92	Acinetobactin biosynthesis	Iron acquisition	Y	Y	Y	Y	Y	Y	Y	[117, 120]
ddc, *dat*	2453-54	1,3-diaminopropane production	Surface-associated motility	Y	Y	Y	Y	Y	Y	Y	[102]
fbsA-Q	2562-82	Fimsbactin A–F biosynthesis	Iron acquisition	N	N	N	N	N	N	N	[122]
bap	2696	Biofilm associated protein	Biofilm formation and adherence	N	Y	Y	Y	Y	Y	Y	[107]
pepO	2789	M13 metalloendopeptidase	Unknown	Y	Y	Y	N	Y	N	N	Unc.[2]
ompA	2840	Outer membrane protein A	Abiotic and eukaryotic cell attachment	Y	Y	Y	Y	Y	Y	Y	[109]
pld	2989	Phospholipase D	Serum resistance, epithelial cell invasion	Y	Y	Y	Y	Y	Y	Y	[262]
pglL	3176	O-oligosaccharyltransferase	Protein glycosylation and biofilm formation	Y	Y	Y	Y	Y	Y	Y	[127]
hemO	ABSDF2280-88[4]	Haem oxygenase cluster	Haem consumption	N	N	N	N	N	Y	N	[123]
katE	3381	Catalase	H_2O_2 resistance	N	N	N	N	N	Y	N	Unc.
bap2	AB57_3081[4]	Biofilm associated protein	Biofilm formation	N	N	Y	Y	N	Y	Y	Unc.
pilE	N/A[5]	Type IV fimbrial biogenesis	Unknown	N	N	Y	N	Y	N	N	Unc.
fimB				N	N	Y	N	N	N	N	
pglL2	ABD1_30650[4]	O-oligosaccharyltransferase	Unknown	N	Y	N	Y	N	N	N	Unc.

1: *A. baumannii* ATCC 17978 locus tag number(s)
2: Uncharacterised in *A. baumannii*
3: *A. baumannii* M2 was recently rechristened *A. nosocomialis* M2 [101]
4: Not present in *A. baumannii* ATCC 17978
5: Not present in the genome of any completed *A. baumannii* genome
6: One or more genes are missing from the virulence gene cluster

5.2.3.4 The accessory genome of community-acquired *Acinetobacter baumannii*

In addition to determining the presence of core virulence genes in the sequenced CA-AB and nosocomial isolates, comparative BLAST analysis had identified a number of genes in these isolates that were not present in the genome of any other isolate. *A. baumannii* CA1 had 210 unique genes, the majority of which were hypothetical proteins or putative metabolic enzymes. There was however a DAMGE island (Chapter 3) present that encoded the uptake, efflux and detoxification of various heavy metals such as copper, iron, and arsenic.

A. baumannii CA7 had 192 unique genes, including several gene clusters that may contribute to virulence, putatively encoding fimbrial biogenesis, secretion system proteins, and the uptake of exogenous siderophores. This organism also encoded two novel metabolic pathways associated with the degradation of phenol and N-methylhydantoin.

A. baumannii CA17 had 141 unique genes; amongst what could be discerned functionally was a gene cluster putatively involved in the transport and degradation of monocyclic aromatic compounds via the 4,5-extradiol ring fission pathway [263, 264]. This degradation pathway is separate from the β-ketoadipate pathway, which also degrades monocyclic aromatic compounds [265] and is encoded in the core genome of *A. baumannii*.

A. baumannii HA17 had 194 unique genes, virtually all of which were localised to a single contiguous sequence, corresponding to a 111 kb plasmid encoding numerous phage-like elements, analogous to the *A. baumannii* plasmids of two other ICII isolates, pABTJ2 and ZW85p2.

A. baumannii HK4620 had 236 unique genes, many of which corresponded to hypothetical proteins or putative metabolic enzymes. However, some unique genes were localised to a 62 kb prophage region that actually comprised of two separate mobile elements - a prophage and a genomic island. The phage had targeted the 3' end of a $tRNA^{Val}$ gene, with the genomic island having subsequently targeted the former 3' end of this gene.

A. baumannii HK4751 had 271 unique genes, including a number of hypothetical proteins and a gene cluster putatively encoding an ammonia monooxygenase enzyme and a tricarboxylate transpoter. There was also a 29 kb gene cluster unique to this organism that encoded a hydroxyphenylacetate and homoprotocatechuate degradation pathway, and a tyrosyl-tRNA synthetase.

Through this comparative approach, we also sought to determine whether there were any genes that were common to CA-AB isolates, and absent in nosocomial isolates, that may be responsible for the differences in phenotypic and virulence of the two different types of isolates. Unfortunately, there were very few genes that could be identified in this way, though one did stand out as a strong virulence candidate.

An uncharacterised gene encoding a putative M13 metalloendopeptidase (*pepO*) was identified as being present only in the genomes of CA-AB isolates (including *A. baumannii* D1279779), as well as ATCC 17978 and BJAB0715, two nosocomial isolates not of the ICI-ICIII lineages (Table 5.2.1 and Figure 5.2.1). Though the role of PepO in *A. baumannii* virulence is unclear, homologues in other organisms are known to promote virulence in *Aeromonas hydrophila* [266], *Porphyromonas gingivalis* [267], and *Streptococcus pneumoniae* [268]. Interestingly, loss of *pepO* in *Francisella novicida* was associated with increased virulence, as isolates lacking this gene were able to cause infection in systemic sites, leading to the proposal that secretion of PepO resulted in increased vasoconstriction, thus limiting systemic spread of this organism [269]. It would be of great interest to elucidate the contribution of *pepO* to *A. baumannii* virulence, which can be conducted through deletion of this gene, in conjunction with the use of virulence models, such as mice or *Galleria mellonella*.

5.3 Comparative genomics of *Acinetobacter baumannii* originating from varying extranosocomial environments

5.3.1 Introduction

Acinetobacter baumannii is a Gram-negative bacterium that has garnered interest as an important nosocomial pathogen [75], with its intrinsic [16] and laterally acquired [187] resistance to antimicrobials being of chief concern [17]. The complete genome sequences of fifteen *A. baumannii* isolates have so far been determined, with thirteen

of these being nosocomial isolates. The community-acquired *A. baumannii* (CA-AB) isolate, D1279779 [241] (Chapter 2), and the human body louse isolate, SDF [160], are the only completely sequenced *A. baumannii* isolates that did not originate from a nosocomial infection. Though genotypically and phenotypically divergent from their nosocomial counterparts [241] (Chapter 2), both CA-AB [144, 145] and louse isolates akin to SDF [137] are capable of infection. Additionally, the genome of a (presumably) non-pathogenic mangrove soil isolate, *A. baumannii* MSP4-16, has been determined recently. However, this genome is neither complete, nor has any extensive bioinformatic analysis been performed [166]. To differentiate nosocomial isolates of *A. baumannii* from CA-AB and other non-nosocomial isolates of this organism, we have elected to refer to the latter two as extranosocomial *A. baumannii* (XN-AB).

There has been speculative debate regarding the existence of XN-AB, its reservoirs, and their role in infection [1, 11, 128, 270, 271], despite the paucity of supporting epidemiological studies. Potential sources of nosocomial *A. baumannii* outside the clinical milieu have been investigated, including meat and vegetables [129, 130], soil [130, 131], anthropogenic surfaces [132] and the community [133]; these extranosocomial isolates were found to be unrelated to nosocomial isolates. Various other extranosocomial niches of *A. baumannii* have been identified, often inadvertently, including various arthropods [228, 272-277], hydrocarbon contaminated soils [278, 279], estuaries [135, 280] and agrarian sites [134, 281], in addition to its carriage in dolphins [136] and humans [139, 282]. Whilst the above epidemiological studies found XN-AB isolates to be generally antimicrobial susceptible, very little else is known regarding the genetic or phenotypic traits of these isolates, as reflected in the bias of available genome sequences.

We sought to sequence the genomes of five XN-AB organisms, isolated from a variety of extranosocomial environments, to determine the features that are present and compare them to previously sequenced XN-AB and nosocomial *A. baumannii* isolates.

5.3.2 Materials and Methods

5.3.2.1 Bacterial strains, species identification, culture conditions

A. baumannii strains 10-3 and 10-4 were previously isolated from Alabamian field soil demonstrating plant growth promotion, and were kindly provided by John McInroy

and Joseph Kloepper (Auburn University, AL, USA). *A. baumannii* B9 was incidentally isolated during sampling of the Seine River for multidrug resistant bacteria [135], and was provided by Patrice Nordmann (Paris-Sud University, Paris, France). The *A. baumannii* B9 in this study is not to be confused with the Indian waste wastewater isolate *A. baumannii* B9 (MTCC 10506) [283]. *A. baumannii* E-011922 and E-072658 [284] were procured from the VTT Culture Collection (Finland), and were previously isolated from the kaolin and recycled fibre pulp of Finnish paper mills, respectively. All five strains were previously identified as *A. baumannii* on the basis of partial 16S rRNA sequences, which we further confirmed through amplification and sequencing of variable zone 2 within the *rpoB* gene [285]. All *A. baumannii* strains were routinely cultured in lysogeny broth (LB) (without glucose) at 37 °C.

5.3.2.2 Genomic DNA isolation, DNA sequencing, genome assembly and annotation

Total genomic DNA was isolated from 1 mL of overnight culture with either the DNeasy Blood & Tissue Kit (Qiagen) or the ISOLATE II Genomic DNA Kit (Bioline), per the manufacturers' protocol. Genomic DNA for all five *A. baumannii* strains were prepared as Nextera XT (Illumina) libraries and sequenced with Illumina MiSeq (2 x 250 bp paired-end reads) at the Ramaciotti Centre for Gene Function Analysis (University of New South Wales, Sydney). The sequencing reads were assembled with the SPAdes algorithm [286] using the recommended settings for long paired-end Illumina reads. The subsequently generated contiguous sequences were annotated using the RAST annotation server [240].

5.3.2.3 Comparative analyses and accessory element identification

The contiguous sequences of the five XN-AB genomes were reorganized relative to the fifteen currently complete genomes of *A. baumannii* using Mauve [237]. Comparative genomic analyses, region of genomic plasticity (RGP) identification, and phylogenetics were all conducted as previously described [241] (Chapter 2), with a few exceptions. RGPs in this study were classified more along the lines of the previously established definition [170, 213], referring to any gene cluster that is variable compared to other *A. baumannii*, including GEIs, MGEs, and genomic deletions. For simplicity, portions of the *A. baumannii* core genome absent in SDF [160] and E-072658 were generally exempted as RGPs, due to extensive reduction and rearrangement in the genomes of both organisms. The phylogenetic relationship

of all *A. baumannii* strains was inferred with bootstrapped neighbour-joining analysis, using all seven reference genes previously utilised in the multilocus sequence typing (MLST) of *A. baumannii* [37].

5.3.3 Results and Discussion

5.3.3.1 Genomic features

The genome features for both newly and previously sequenced XN-AB isolates are summarised in Table 5.3.1. The genome sequences of *A. baumannii* 10-3 and 10-4, both isolated from field soil, were near identical with an overall nucleotide identity of ≥99%. The genomic features of all XN-AB were found to be atypical compared to the genomes of previously sequenced nosocomial *A. baumannii*. The genome sizes of XN-AB isolates were generally smaller than those of nosocomial *A. baumannii* (< 4 Mbp), with the exception of the paper mill kaolin isolate *A. baumannii* E-011922. The number of insertion sequences (ISs) in XN-AB genomes varied greatly, ranging from zero in *A. baumannii* 10-3 and 10-4, to over four hundred in strains SDF and E-072658 (Table 5.3.1). The relatively small genome size of the recycled fibre pulp isolate, *A. baumannii* E-072658, is likely to have arisen from extensive IS-mediated genome decay, previously thought to have occurred in *A. baumannii* SDF [160]. The drug susceptible ICI isolate, *A. baumannii* AB307-0294, is the only nosocomial isolate to share both genome size and number of ISs in common with XN-AB isolates, with a genome size of 3,761,981 bp and an absence of ISs [157]. The number of plasmids present in XN-AB isolates were similar to that observed in previously sequenced nosocomial isolates. However, E-072658 was remarkable in that it had eight putative plasmids in total. This would make it the greatest number of plasmids in an *A. baumannii* isolate, followed by the six present in *A. baumannii* Naval-18 (GenBank accession: NZ_AFDA00000000.2).

Table 5.3.1: Comparative genome features of extranosocomial *Acinetobacter baumannii* (XN-AB) isolates.

	10-3	**10-4**	**B9**	**E-011922**	**E-072658**	**D1279779**	**SDF**
Total genome size (Mbp)	3791351	3781833	3745434	4051111	3466131	3704285	32421954
# Plasmids	0	0	1	3	8	1	4
G+C content (%)	38.9	38.9	39.0	39.1	39.2	39.0	39.1
# Protein coding sequences (CDSs)	3473	3471	3447	3829	3310	3388	2913
# Insertion sequences (ISs)	0	0	82	37	440	18	428
# rRNA operons	6	6	6	6	N. D.	6	5
# tRNAs	63	65	63	66	63	65	64
Number of contigs (≥1 kbp)	16	29	57	40	267	1	1

1. Genome size including putative plasmids
2. Chromosome size only (excluding plasmids)
3. N. D.: not determined due to the fragmented state of the genome assembly

5.3.3.2 The phylogeny of XN-AB and discovery of a 'clinical' XN-AB isolate

The phylogenetic relationship of the five sequenced XN-AB isolates was inferred using a MLST approach [37], with respect to the fifteen completely sequenced isolates of *A. baumannii*. As expected, the XN-AB isolates were generally unrelated to nosocomial isolates of *A. baumannii* (Figure 5.3.1). However, the Seine River isolate, *A. baumannii* B9, had grouped with *A. baumannii* of the international clone II (ICII) lineage, often associated with nosocomial infections (Figure 5.3.1). To the best of our knowledge, *A. baumannii* B9 represents the first ICII lineage strain known to have been isolated outside of the clinical milieu.

It is not clear if *A. baumannii* B9 had originated from the clinical milieu, despite being the most obvious point of origin. The sampling site in downtown Paris was not in proximity to any hospital wastewater discharge sites [135] or wastewater treatment plants, which correlated with a relatively low concentration of antimicrobial contaminants at that site [287]. Other possible sources of clinical *A. baumannii* in the environment also include improperly managed [288] or illegally dumped clinical waste [289]. The Seine River watershed, comprising of the Seine River and its tributaries, is characterised by low microbiological quality stemming from constant agricultural and anthropogenic input [290], which may be the most likely source of this isolate. The presence of multidrug resistant *A. baumannii* in the Seine River watershed warrants some concern, given that its surface waters constitute a major source of drinking water [287].

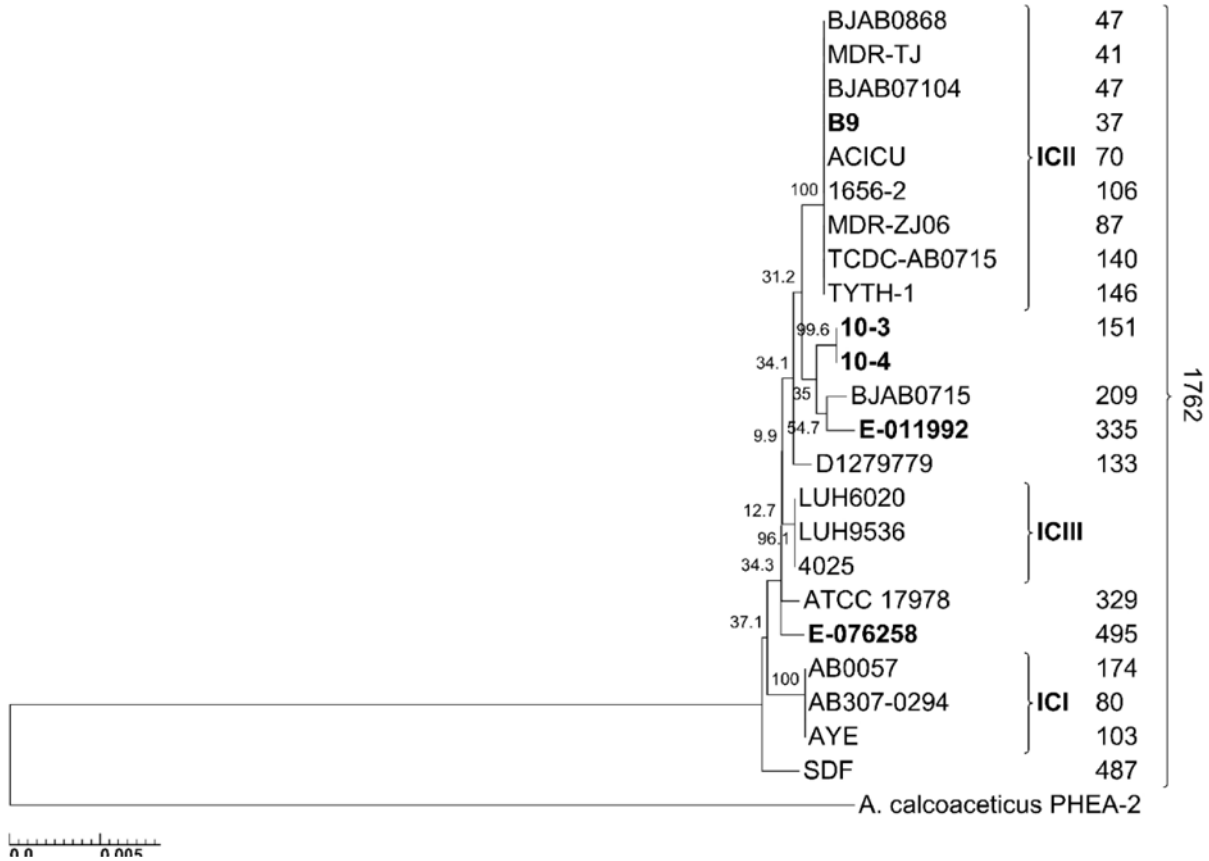


Figure 5.3.1: The phylogenetic relationship of sequenced extranosocomial *Acinetobacter baumannii* (bold text) compared to completely sequenced reference strains. The phylogeny was inferred by the neighbour-joining method (bootstrapped; 1000 replicates) in MEGA5 [253], using the concatenated nucleotide sequences of seven reference genes based on the multilocus sequence typing scheme of *A. baumannii* (*cpn60*, *fusA*, *gltA*, *pyrG*, *recA*, *rplB* and *rpoB*) [37]. The numbers adjacent to each strain name denote the number of genes unique to the genome of that strain. The number adjacent to the bracket represents the size of the core genome of all *A. baumannii* strains (excluding ICIII strains). The tree was drawn and annotated in TreeGraph2 [254].

5.3.3.3 The genome architecture of extranosocomial *A. baumannii*

Whole genome alignments of the five sequenced XN-AB isolates were conducted against representative XN-AB and nosocomial isolates using Mauve [237] to investigate the level of genome synteny. The genome architecture of *A. baumannii* AB307-0294 was discovered to be most suitable for its use as a reference genome, as it shared various genetic features (Table 5.3.1) and the genomes of all XN-AB isolates were highly syntenous with the genome of this organism (Figure 5.3.2). Though the genome of *A. baumannii* E-076258 did display genome synteny with AB307-0294, ascertaining the true genomic architecture of this organism is problematic due to the relatively high number of contigs in the assembly of this genome (Table 5.3.1 and Figure 5.3.2).

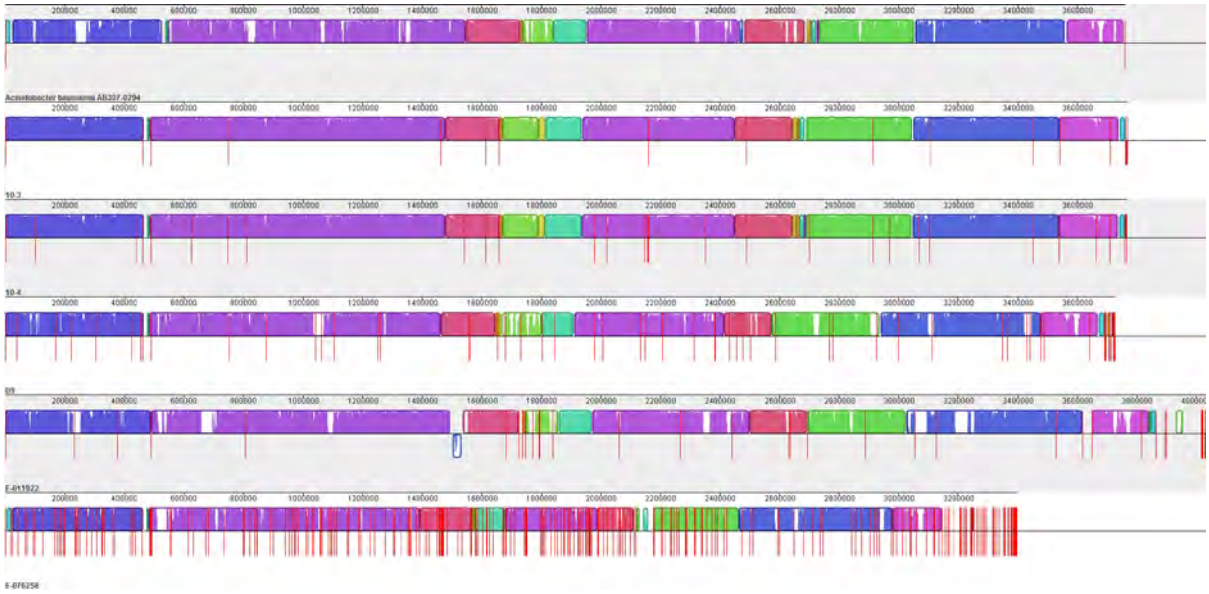


Figure 5.3.2: Synteny of extranosocomial *Acinetobacter baumannii* compared to representative nosocomial *A. baumannii*. Regions of significant synteny between the strains are shown as coloured blocks and unshared regions are seen as white gaps. Coloured blocks that are above and below the black lines are on the positive and reverse strands, respectively. The red lines indicate contiguous sequence boundaries. This figure was generated using Mauve [237].

5.3.3.4 The core genome of extranosocomial *A. baumannii*

The predicted proteomes of the five XN-AB isolates were compared to the proteomes of the fifteen previously sequenced *A. baumannii* with BLASTP+ [255] ($e10^{-5}$) to examine the presence of genes known to be involved in virulence and multidrug resistance in *A. baumannii*. The core genome size was estimated to be 1762 genes (Figure 5.3.1), which increased to 2396 when SDF and E-072658 were excluded from analysis.

A. baumannii 10-3/10-4 and B9 encoded all of the core virulence traits previously characterised in *A. baumannii*, whilst *A. baumannii* E-011922 was only lacking a gene cluster involved in quorum sensing and surfactant biosynthesis (Table 5.3.2). As previously noted, *A. baumannii* D1279779 lacked a few virulence-associated genes encoding a type VI secretion system and a biofilm associated protein [241] (Table 5.3.2). Conversely, the genomes of *A. baumannii* SDF and E-072658 were lacking numerous core virulence genes, including those encoding the *Acinetobacter* trimeric autotransporter (*ata*), quorum sensing, pilin biogenesis and siderophore biosynthesis (Table 5.3.2).

Table 5.3.2: The presence (Y), partial presence (P) or absence (N) of known and putative *Acinetobacter baumannii* virulence factors in extranosocomial isolates.

Gene(s)	ABD1_[1]	Product	Virulence role	10-3	10-4	B9	E-011922	E-072658	SDF	Reference
plc1	00420	Phospholipase C	Eukaryotic cell damage	Y	Y	Y	Y	Y	Y	[256]
abaI	01150	AHL lactone synthetase	Quorum sensing	Y	Y	Y	N	N	N	[115][3]
abaR	01180	LuxR-family receptor		Y	Y	Y	N	N	N	
	01180-01250	Lipopeptide and polyketide synthesis	Surface-associated motility and biofilm formation	Y	Y	Y	N	N	N	[104, 112]
pbpG	02210	Penicillin-binding protein 7/8	Serum resistance	Y	Y	Y	Y	Y	Y	[257]
filA-F	06380-440	P pilli biogenesis	Biofilm formation	Y	Y	Y	Y	N	Y	[258]
nfuA	09330	Fe-S scaffold protein	Iron acquisition	Y	Y	Y	Y	Y	Y	[259]
ata	09870-880*	*Acinetobacter* trimeric autotransporter	Adhesion and biofilm formation	Y*	Y*	Y*	Y*	N	N	[260]
	A1S_1292-1312[4]	Type VI secretion system	Bacterial competition	Y	Y	Y	Y	N	Y	[261][3]
acuABCDE	15100-130	Type I pili biogenesis	Biofilm formation	Y	Y	Y	Y	P	N	[117]
	16320-420	Siderophore biosynthesis cluster	Iron acquisition	Y	Y	Y	Y	N	N	[117]
pgaABCD	21590-610	PNAG biosynthesis	Biofilm formation	Y	Y	Y	Y	Y	N	[110]
csuA/BABCDE	22100-200	Type I pili biogenesis	Biofilm formation and attachment	Y	Y	Y	Y	N	N	[111]
bauA-E; basA-J	23680-860	Acinetobactin biosynthesis	Iron acquisition	Y	Y	Y	Y	Y	N	[117, 120]
ddc, *dat*	24400-410	1,3-diaminopropane production	Surface-associated motility	Y	Y	Y	Y	Y	Y	[102]
fbsA-Q	A1S_2562-2582[4]	Fimsbactin A–F biosynthesis	Iron acquisition	N	N	N	N	N	N	[122]
bap	A1S_2696[4]	Biofilm associated protein	Biofilm formation and adherence	Y	Y	Y	Y	N	N	[107]
ompA	27860	Outer membrane protein A	Abiotic and eukaryotic cell attachment	Y	Y	Y	Y	Y	Y	[109]
pld	28800	Phospholipase D	Serum resistance, epithelial cell invasion	Y	Y	Y	Y	Y	Y	[262]
pglL	30640	O-oligosaccharyltransferase	Protein glycosylation and biofilm formation	Y	Y	Y	Y	Y	Y	[127]
hemO	ACICU_00873-00880[4]	Haem oxygenase cluster	Haem consumption	N	N	N	N	N	Y	[123]

1: *A. baumannii* D1279779 locus tag number(s)
2: Uncharacterised in *A. baumannii*
3: *A. baumannii* M2 was recently rechristened *A. nosocomialis* M2 [101]
4: Not present in *A. baumannii* D1279779
*: The *ata* gene is broken into two halves

The genomes of XN-AB isolates were also examined for the presence of genes and mutations known to be involved in antimicrobial resistance in *A. baumannii*, which are

summarized in Table 5.3.3. The analysed XN-AB isolates encoded the majority, if not all, core efflux proteins known to be involved in the intrinsic multidrug resistance of nosocomial *A. baumannii*. As expected, *A. baumannii* SDF was lacking genes encoding several drug transporters, including *adeABC*, *abeS* and *abeM4* (Table 5.3.3). However, this organism also lacked the *pmrCAB* gene cluster involved in lipid A biosynthesis, the loss of which has been linked to increased resistance to polymxyin antimicrobials [95]. Antimicrobial resistance genes known to be carried within or activated by MGEs were absent from all sequenced XN-AB isolates, except for the Seine River isolate *A. baumannii* B9, which encoded a tetracycline efflux protein (*tetA*), an *ISAba1*-activated chromosomal cephalosporinase (*ampC*), an oxacillinase (bla_{OXA-23}) and an aminoglycoside 3"-phosphotransferase (*strAB*). This organism also had Ser→Leu codon mutations in both the *gyrA* and *parC* genes, which are known to contribute to fluroquinolone resistance [89]. Given that *A. baumannii* B9 was isolated during screening of multidrug resistant bacteria in the Seine [135], and the subsequent revelation of it being an ICII strain, the presence of laterally acquired genes involved in antimicrobial resistance was to be expected, but is nonetheless of great interest.

Table 5.3.3: The occurrence of characterised intrinsic and laterally acquired genes involved in antimicrobial resistance, denoted by their presence (Y) or absence (N).

	Gene(s)	ABD1_[1]	Product	Antimicrobial substrates	10-3	10-4	B9	E-011922	E-072658	SDF	Reference
Efflux pumps	*aedA*	01860	DHA2 family transporter	not determined	Y	Y	Y	Y	Y	Y	[49]
	abeM	03650	MATE family transporter[3]	aminoglycosides, fluoroquinolones, chloramphenicol, trimethoprim and acriflavine	Y	Y	Y	Y	Y	Y	[61]
	aedB	07290	DHA2 family transporter	not determined	Y	Y	Y	Y	Y	Y	[49]
	aedC	08550	DHA2 family transporter	tetracycline and chloramphenicol	Y	Y	Y	Y	Y	N	[49]
	aedD	16340	DHA2 family transporter	not determined	Y	Y	Y	Y	N	N	[49]
	adeABC	17690-700[4]	RND family transporter	aminoglycosides, fluoroquinolones, cefotaxime, erythromycin, tetracycline, chloramphenicol, trimethoprim and tigecycline	Y	Y	Y	Y[5]	Y[5]	N	[50, 52]
	aedE	17910	DHA2 family transporter	not determined	Y	Y	Y	Y	N	N	[49]
	amvA/aedF	20530	DHA2 family transporter	erythromycin and chlorhexidine	Y	Y	Y	Y	Y	Y	[49, 58]
	aceI	20590	Chlorhexidine efflux protein	chlorhexidine	Y	Y	Y	Y	Y	Y	[63]
	abeS	22970	SMR family transporter	chloramphenicol, fluoroquinolones, erythromycin and novobiocin	Y	Y	Y	Y	Y	N	[62]

	adeFGH	23030-50	RND family transporter	chloramphenicol, clindamycin, fluoroquinolones, trimethoprim, tetracyclines, tigecycline, and sulphonamides	Y	Y	Y	Y	Y	Y[6]	[55]
	adeIJK	26820-40	RND family transporter	β-lactams, chloramphenicol, tetracyclines, erythromycin, fluoroquinolones, fusidic acid, novobiocin, and trimethoprim	Y	Y	Y	Y	Y	Y	[56]
	craA	30320	MFS family transporter	chloramphenicol	Y	Y	Y	Y	Y	Y	[59]
	abeM3	32470	MATE family transporter[3]	not determined	Y	Y	Y	Y	Y	Y	[233]
	abeM4	33210	MATE family transporter	ciprofloxacin	Y	Y	Y	Y	Y	N	[233]
	abeM2	A1S_2562[4]	MATE family transporter[3]	ciprofloxacin	N	N	N	N	N	N	[233]
	adeM	A1S_3445-46[4]	RND family transporter	norfloxacin and chlorhexidine	N	N	N	N	N	N	[233]
	tetA	AB57_0275[4]	MFS family transporter	tetracycline	N	N	Y	N	N	N	[291]
β-lactamases	*ampC*-IS*Aba1*	23620[7]	Chromosomal cephalosporinase	third-generation cephalosporins	N[7]	N[7]	Y	N[7]	N[7]	N[8]	[85]
	bla$_{OXA-51/69}$-IS*Aba1*	15230[9]	Chromosomal oxacillinase	carbapenems	N[9]	N[9]	N[9]	N[9]	N[9]	N[9]	[86]
	bla$_{OXA-23}$	AB57_0551[4]	Oxacillinase	carbapenems	N	N	Y	N	N	N	[292]
AMEs[2]	*aac1*	AB57_0291[4]	Aminoglycoside 3'-acetyltransferase	aminoglycosides	N	N	N	N	N	N	[91]
	aacA4	ACICU_002234	Aminoglycoside 6'-acetyltransferase	aminoglycosides	N	N	N	N	N	N	[91]
	aadA1	ABAYE3618[4]	Aminoglycoside 3'-adenyltransferase	aminoglycosides	N	N	N	N	N	N	[91]
	strAB	ABAYE3647-8	Aminoglycoside 3"-phosphotransferase	aminoglycosides	N	N	Y	N	N	N	[91]
	parC	01920[10]	DNA topoisomerase IV	quinolones	N[10]	N	Y	N	N	N	[89]
	gyrA	25790[10]	DNA gyrase subunit A	quinolones	N[10]	N	Y	N	N	N	[89]
	pmrCAB	26970-90[10]	Lipid A biosynthesis	polymxyins	N[10]	N	N	N	N	Y	[95]
	lpxADC	19910,19930, 32260[10]	Lipopolysaccharide biosynthesis	polymxyins	N[10]	N	N	N	N	N	[293]

1: *A. baumannii* D1279779 locus tag number(s)
2: Aminoglycoside modifying enzymes
3: Substrates characterised through heterologous expression
4: Not present in *A. baumannii* D1279779
5: The *adeC* gene is absent
6: The *adeH* gene is absent
7: Chromosomal *ampC* only
8: The chromosomal *ampC* gene is absent
9: Chromosomal *bla*$_{OXA-51/69}$ only
10: Genes are present but lack mutations that bestow resistance phenotype

5.3.3.5 The accessory genome of extranosocomial *A. baumannii*

In addition to elucidating the core genome of *A. baumannii*, comparative BLAST analysis was utilised to identify genes potentially unique to the various XN-AB isolates, followed up with the identification of known and novel regions of genomic plasticity. The numbers adjacent to each strain in the phylogenetic tree (Figure 5.3.1) indicate the number of CDSs unique to each sequenced *A. baumannii* strain.

The two soil isolates, *A. baumannii* 10-3/10-4, were found to possess 151 unique genes, the majority of which appeared to encode various metabolic enzymes, transporters and transcriptional regulators. One such gene cluster that could be discerned functionally was a 10.2 kb gene cluster encoding monocyclic aromatic compound transport and degradation via the 4,5-extradiol ring fission pathway [263, 264]. This degradation pathway is separate from the β-ketoadipate pathway, which also degrades monocyclic aromatic compounds [265] and is encoded in the core genome of *A. baumannii* and strains 10-3/10-4.

The Seine River isolate, *A. baumannii* B9 had a relatively small number of unique genes (37), with most being hypothetical proteins, although this organism did encode the Tn*AbaR* resistance island, commonly present in nosocomial *A. baumannii*. This island encoded resistance to aminoglycosides, β-lactams and tetracyclines (Table 5.3.3), however this variant lacked the class 1 integron array common to most Tn*AbaR* islands [187], and it was not present elsewhere in the chromosome, as had been observed for the AbGRI2 resistance island (Chapter 4) [251].

The paper mill kaolin isolate, *A. baumannii* E-011922, encoded a high number of unique genes (335), including two unique variants of the highly variable K and OC complex carbohydrate biosynthesis loci of *A. baumannii* [125], which form the structural basis of bacterial capsule, an important virulence determinant in this organism [124]. Others genes included a gene cluster putatively encoding an ammonia monooxygenase enzyme and the transport of tricarboxylate. E-011922 also contained several GEIs of interest, including a 7.9 kb genomic island encoding putative DNA repair enzymes, integrated within the *umuC* gene, encoding a subunit of the error-prone DNA polymerase V. Another GEI of interest was a 41.6 kb DAMGE (Chapter 3) variant that encoded the uptake, efflux and detoxification of various heavy metals such as copper, iron, and arsenic.

The paper mill pulp isolate, *A. baumannii* E-076258, was found to have many (495) unique genes, however the majority appear to be located at the termini of each of the contiguous sequences, suggesting they may be artefacts of assembly and annotation, further exacerbated by the very short lengths (<100 kbp) of these contigs. This organism encoded multiple copies of genes involved in sarcosine degradation, which were not present in the other strains.

5.4 Conclusions and future directions

The genome sequence of *A. baumannii* WM99c represents the first genome sequence of an Australian nosocomial isolate of *A. baumannii*. Through combined genomic and phenomic analysis, this isolate was found to be a highly drug resistant member of the dominant ICII lineage, and possessed several accessory elements in common with previously sequenced ICII isolates originating from the Asia-Pacific and Europe. One GEI of continued interest is the Tn*AbaR* resistance island, which was identified and reconstructed through PCR in WM99c. However, the class I integron array common to most Tn*AbaR* islands [187] was found to have relocated to a separate genomic loci in WM99c, as well as several other sequenced ICII isolates [44, 251].

G41 is another RGP previously identified in *A. baumannii*, also discovered to be exclusive to sequenced genomes to the ICII lineage. This gene cluster encoded an aminoglycoside phosphotransferase, and several metabolic genes putatively involved in the transport, regulation and metabolism of fatty acids. One of the genes in this cluster encoded a short chain dehydrogenase/reductase (SDR) which has had its structure determined. Neither bioinformatic prediction nor phenotypic microarrays (Figure 5.1.3) could elucidate the substrate for SDR or the other genes in the G41 cluster. The work regarding G41, SDR and the phenotypic profile of WM99c will be incorporated into a study examining the structure and function of WM99c-SDR, with the further possibility of performing gene deletions in *A. baumannii* WM99c using a recently published methodology [235].

Further analyses or closure of the *A. baumannii* WM99c genome is unlikely, as there are already several complete genome sequences of ICII *A. baumannii* available. Furthermore, WM99c has recently been resequenced by another laboratory group as part of a wider study on multidrug resistant *A. baumannii* originating from Australian hospitals. The raw Illumina sequence for this organism is currently available online (SRA Accession: ERX087518). To avoid further duplicative work on this organism, we sought to refocus on the relatively neglected research area of extranosocomial *A. baumannii* isolates.

The sequencing and phylogenetic analysis of several XN-AB isolates had mostly validated the prior assumption that these isolates are epidemiologically and

genotypically divergent from nosocomial *A. baumannii*. However, the Seine River isolate *A. baumannii* B9, proved to be an interesting exception, as it was revealed not only to be an ICII lineage strain, but it also encoded the Tn*AbaR* resistance island. The epidemiological origin of this organism was quite abstruse, as there was no definitive link between the clinical milieu and its presence in the Seine. *A. baumannii* B9 was not the only multidrug resistant isolate to be discovered in the environment. Very recently a multidrug-resistant isolate of *A. baumannii* was discovered within the palaeosol of an abandoned Croatian quarry, and was genetically and phenotypically similar to a Croatian nosocomial isolate of the ICI lineage [289].

It is of further curiosity that despite our inclusion of the various XN-AB isolates in the comparative genomics of nosocomial *A. baumannii*, what has previously constituted as the core genome of *A. baumannii* [36] is still valid, particularly in regards to the presence of genes encoding virulence traits (Table 5.3.2) and multidrug resistance efflux proteins (Table 5.3.3), otherwise associated with disease and resistance in nosocomial *A. baumannii*.

The genomes of two XN-AB isolates, the Finish paper mill isolate, *A. baumannii* E-072658, and the louse borne isolate, *A. baumannii* SDF, are exceptions to the above, as the genomes of both organisms are characterised by extensive genome decay and rearrangement, presumably mediated by ISs. The aberrant genome architecture of E-072658 had complicated the analysis of this isolate, but nonetheless demonstrated that the genome architecture of *A. baumannii* SDF was not a solitary event.

The two fundamental differences that appear to separate the genomes of XN-AB and nosocomial isolates are genome size and the genes encoded in their respective accessory genomes. Typically the genomes of XN-AB appear to be smaller than those of nosocomial isolates, resultant from either considerable gene loss, as in the case of E-072658, or absence of mobile genetic elements. The genes encoded within the accessory genomes of XN-AB isolates are primarily devoted to the acquisition and metabolism of novel compounds, whilst those in nosocomial *A. baumannii* are primarily involved in antimicrobial resistance, including the Tn*AbaR* and AbGRI2 resistance islands.

In future, complete or high-quality draft genomes of these XN-AB isolates will be produced, and those not known to be pathogenic will be phenotypically compared to nosocomial *A. baumannii* by means of phenotype microarrays, antimicrobial testing, and virulence assays, including biofilm, quorum sensing and motility assays.

Chapter 6: General discussion and conclusions

6.1 *Acinetobacter baumannii* is more conserved than initially assumed

Acinetobacter baumannii is known to have remarkable plasticity in its genomic architecture, be it through the acquisition and integration of exogenous genetic material and MGEs, or genome rearrangement and/or genome decay mediated by homologous recombination. This organism is also known to have varying phenotypic capabilities such as multidrug resistance, survival on abiotic surfaces, fluctuating virulence, or utilisation of numerous carbonaceous and nitrogenous compounds. Furthermore, *A. baumannii* has been reported to be a cause of nosocomial and community-acquired human infections, as well as be present in varying non-nosocomial niches, such as soil, water and arthropods.

However, some of the above aspects of *A. baumannii*, such as mechanisms of virulence, metabolic pathways, and niche distribution, are poorly understood compared to other members of the ESKAPE coterie. Other members of the ESKAPE group have also been isolated from nosocomial and extranosocomial sources, and in some instances have been sequenced, resulting in genetic and phenotypic differences being identified in these isolates. Examples include the discovery of virulence and resistance genes unique to community-acquired *Staphylococcus aureus* isolates [294, 295] and a rhizospheric isolate of *Pseudomonas aeruginosa* that was found to be genetically and phenotypically antimicrobial susceptible, but encoded biosynthetic pathways for several novel biocontrol agents [296]. Prior to the work covered in this thesis, knowledge regarding the genome and phenome of *A. baumannii* was chiefly focused on nosocomial isolates.

The sequencing of the *A. baumannii* D1279779 genome, an isolate originating from a bacteraemic infection of an Indigenous Australian, represents the first CA-AB isolate to be sequenced, and the only completed example to date. CA-AB infections are known to be antimicrobial susceptible compared to nosocomial infections [144], which we had confirmed in D1279779 through phenotypic microarrays. The genetic basis of antimicrobial susceptibility in this isolate was almost certainly linked to the absence of the Tn*AbaR* antimicrobial resistance island. Phenotypic microarrays also demonstrated that D1279779 was able to catabolise a greater breadth of carbon and

nitrogen compounds than nosocomial isolates of *A. baumannii*. Further genomic analysis of D1279779 revealed that the overall genomic architecture was not dissimilar to those of previously sequenced nosocomial *A. baumannii*, albeit its relatively smaller genome size, and encoded the majority of genes known to be involved in the virulence of this organism [241] (Chapter 2).

The genome analysis of several CA-AB isolates sequenced by our collaborator, A/Prof. Anton Peleg, displayed similar features to the genome sequenced of D1279779, including relatively smaller genome size, the absence of Tn*AbaR*, and the conservation of many genes known to be required for virulence in nosocomial *A. baumannii*. Since infections caused by CA-AB are known to be more fulminant and fatal compared to nosocomial infections [144, 146, 147], it is possible that there may be some genes involved in virulence unique to CA-AB isolates. Prior virulence and phenotypic testing conducted by the Peleg laboratory group demonstrated that the CA-AB isolates generally outperformed nosocomial isolates in respect to biofilm formation, phospholipase D production, murine virulence, and in vitro growth in LB, sera and 1% ethanol. However, genomic analysis conducted by us could not definitively account for these phenotypic differences. Although, there was one interesting instance, regarding the presence of a gene encoding a putative PepO M13 metalloendopeptidase, which is known to contribute to virulence in a number of bacterial pathogens. This gene was only present in the genomes of sequenced CA-AB isolates and a few non-IC lineage nosocomial pathogens.

The sequencing and analysis of the abovementioned CA-AB genomes had fostered the notion that perhaps the current knowledge concerning *A. baumannii* does not reflect the true diversity of this species, as evidenced by the existence of over 400 known MLST sequence types that make up *A. baumannii* [36], as well as the lack of clonality observed in the CA-AB isolate phylogeny, compared to their nosocomial counterparts. The abovementioned work concerning CA-AB had prompted us to focus on other extranosocomial *A. baumannii* originating from a diversity of environments.

We had sequenced the genomes of five XN-AB isolates originating from a mixture of industrial and environmental locations. There was a lack of phylogenetic relatedness between these XN-AB isolates and nosocomial *A. baumannii*, which was similar to that observed in other studies examining XN-AB isolates [129-133]. Analysis of the

XN-AB genome sequences revealed that they had several features in common with those of CA-AB isolates, which included relatively small genome size, absence of the Tn*AbaR* resistance island, and presence of most or all known *A. baumannii* virulence determinants. It was also discovered that most of the core characterised efflux proteins, associated with multidrug resistance in *A. baumannii*, were also encoded in the genomes of XN-AB isolates (including CA-AB).

There were two notable exceptions to the above with the first being *A. baumannii* B9, which was found to have phylogenetically clustered with nosocomial *A. baumannii* of the ICII lineage and encoded a number of antimicrobial resistance determinants, including Tn*AbaR*. The second was *A. baumannii* E-072658, which not only had undergone extensive genome decay, analogous to *A. baumannii* SDF, but was discovered to possess eight putative plasmids, the largest number seen in a sequenced *A. baumannii*.

The major difference that was observed between the genome sequences of nosocomial *A. baumannii* and those of XN-AB isolates was the types of genes encoded within their respective accessory genomes. The accessory genomes of XN-AB isolates were dominated by genes involved in transport and transcriptional regulation, as had been observed in nosocomial isolates [157, 168], but were associated with the acquisition and metabolism of novel compounds in the former, and antimicrobial resistance in the latter.

6.2 The evolution of *Acinetobacter baumannii*

The phylogenetic analyses conducted on all *A. baumannii* isolates in this thesis, when taken together (Figure 6.1), suggested that the emergence of ICI, ICII and ICIII as dominant nosocomial lineages had occurred in independent incidents. It appears that these clonal lineages represent a ‘genetic bottleneck’, in which these isolates have adapted to flourish exclusively within a clinical environment. A clear exception to this however would be *A. baumannii* B9, as its existence demonstrates that a clinical isolate is not precluded from returning to the environment, though seemingly at the cost of some of its accessory elements, including the class 1 integron array normally present in the Tn*AbaR* resistance island.

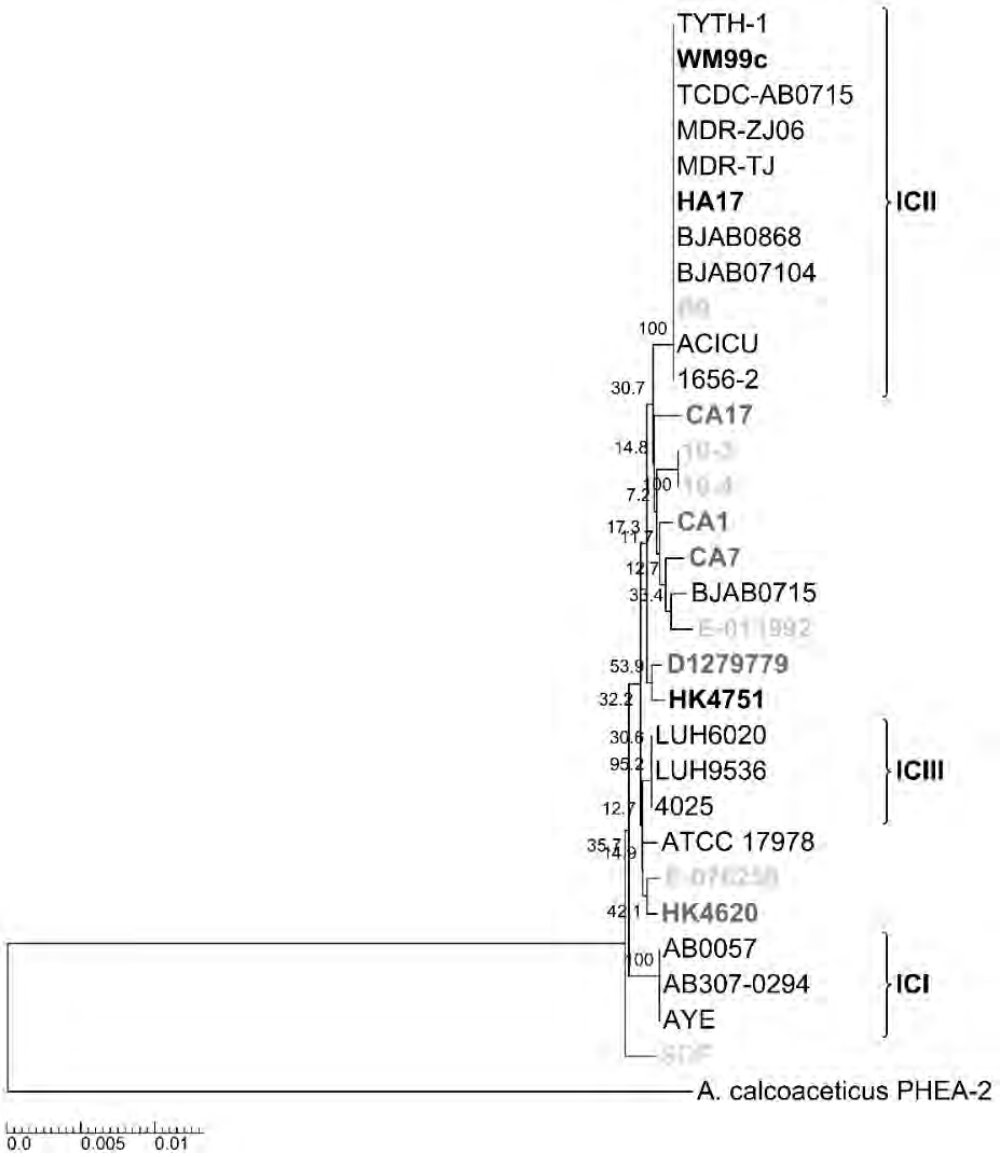


Figure 6.1: The phylogenetic relationship of *Acinetobacter baumannii* analysed in this thesis (bold text) compared to completely sequenced reference strains. *A. baumannii* strains names are coloured based on whether they were isolated from nosocomial infections (black), community-acquired infections (dark grey), or from other extranosocomial sites (light grey). The phylogeny was inferred by the neighbour-joining method (bootstrapped; 1000 replicates) in MEGA5 [253], using the concatenated nucleotide sequences of seven reference genes based on the multilocus sequence typing scheme of *A. baumannii* (*cpn60*, *fusA*, *gltA*, *pyrG*, *recA*, *rplB* and *rpoB*) [37]. The tree was drawn and annotated in TreeGraph2 [254].

While the phylogeny of *A. baumannii* currently cannot determine the origin of nosocomial *A. baumannii*, it may shed light on the origin of CA-AB, the other pathogenic form of this organism. Given the lack of clonality common to both CA-AB and environmental XN-AB isolates, it is possible that the former may have emerged from the latter on multiple, independent occasions. The incidence of GEIs and RGPs common to both these isolates, the presence of other large metabolic gene clusters in CA-AB isolates, and the absence of the Tn*AbaR* antibiotic resistance island, are suggestive of this. Examples of common RGPs include the heavy metal resistance DAMGE variant present in the genomes of strains CA1 and E-011922, and the

presence of a metabolic gene cluster encoding the 4,5-extradiol ring fission pathway in strains CA17 and 10-3/10-4.

Comparative genomics was not able to identify any genes unique to CA-AB isolates that may potentially be responsible for the reported fulminant nature of these isolates. The only exception was the presence of a *pepO* homologue encoding a putative M13 metalloendopeptidase, present only in CA-AB and some non-IC nosocomial isolates. It is possible that genes encoding transcriptional regulators, present in the accessory genome of *A. baumannii*, are responsible for reported observed differences in the virulence of CA-AB and nosocomial isolates.

In this thesis, MLST-based phylogenies were used as a means to examine the evolutionary relationship of sequenced *A. baumannii* isolates. This method was suitable for determining whether an isolate fell into one of the three major IC lineages. *A. baumannii* phylogeny inferred from MLST has been shown to share a similar tree topology to that of single nucleotide polymorphism-based phylogeny, derived from whole genome sequences [297]. However, phylogeny derived from MLST is a poor predictor of genome architecture, and cannot provide the level of detail regarding organism evolution that can be obtained from whole genome phylogeny [297]. In future, whole genome phylogeny can be employed in conjunction with MLST-based phylogeny to explore key evolutionary differences between nosocomial and extranosocomial *A. baumannii*.

6.3 The DAMGEs represent the second major RGP of *Acinetobacter baumannii*

The strong research emphasis on *A. baumannii* originating from the clinical milieu is reflected in the search for RGPs that may contribute to the survival and success of *A. baumannii* as a human pathogen. The most important RGP in nosocomial isolates of this organism would unquestionably be the *Acinetobacter baumannii* antibiotic resistance island (Tn*AbaR*), a genomic island family characterised by the accretion of numerous MGEs encoding antimicrobial resistance determinants [187]. Genomic islands such as these are typified by the presence of recombinases, atypical nucleotide content and direct/inverted repeats.

Bioinformatic analyses conducted to locate genomic regions containing these features resulted in the identification a mobile GEI integrated within the *dusA* gene of not only *A. baumannii*, but also numerous members of the Proteobacteria. These GEIs, which we had dubbed the *dusA*-specific mobile genetic elements (DAMGEs), possessed a highly mosaic cargo gene composition, which in *A. baumannii* were putatively involved in heavy metal resistance (arsenic and copper), metabolism of nitrogenous compounds (arginine and pyrimidine nucleobases), bacteriophage lysogeny, and Type I restriction-modification. The above cargo genes, coupled with the mosaicism of the DAMGEs, have the potential to function in clinical persistence, albeit in a differing capacity to Tn*AbaR*.

For instance, metallic copper (99%) is a highly effective antimicrobial against various Gram-negative pathogens (including *A. baumannii*), and has been proposed as a replacement material for high-contact hospital touch surfaces, in conjunction with established hygiene and cleaning protocols [298]. However, the DAMGEs of a number of *A. baumannii* isolates, including strains AYE, AB0057, CA1 and E-011922, encoded a cluster of genes (*copABCDRS*) known to involved in resistance to copper ions [299], which are released on contact with metallic copper [298]. While the clinical impact of these particular DAMGE variants are not known, we were able to investigate the phenotype of a DAMGE present in *A. baumannii* D1279779 through targeted deletion of the entire island, a feat not previously achieved in this organism. However, phenotype microarray profiling could not definitively elucidate the function of this GEI.

6.4 Creation of large genomic deletions in *Acinetobacter baumannii* is possible

The successful deletion of the 47.3 kb DAMGE genomic island in *A. baumannii* D1279779, as well as the availability of two methodologies to generate unmarked deletions in *A. baumannii* [230, 235], allow for a number of possible research directions.

A feasible research direction might involve not only the identification of laterally acquired regions in the genomes of *A. baumannii*, but also the deletion of them using FLP recombination, so that their phenotypic function may be elucidated. However, this approach would be limited to drug susceptible isolates of *A. baumannii*, such as

D1279779, as the FLP recombination system utilises the aminoglycosides, kanamycin and gentamycin, as resistance markers [230]. Modifications to the vectors used in this deletion system would be necessary to facilitate its use in multidrug resistant *A. baumannii*, and therefore to study resistance islands such as Tn*AbaR* and AbGRI2. One previous study however was able to successfully delete the Tn*AbaR* resistance island though an approach not too dissimilar to the FLP recombination method, allowing its phenotypic characterisation, but it had resulted in the inadvertent deletion of other chromosomal regions [187]. One foreseeable drawback to the deletion of any large chromosomal region would be the difficulty in complementing them.

Another possible research direction involves the use of pMo130-Tel^R, which encodes tellurite resistance and facilitates the generation of gene deletions in *A. baumannii*, including multidrug resistant isolates [235]. One could foreseeably generate a library of single gene deletion mutants in any isolate of *A. baumannii* using this system, which would provide a resource for the study of gene function in this organism. The advantage this system would have over previously constructed mutant libraries in *E. coli* K-12 BW25113 [300], *P. aeruginosa* PA14 [301], and *A. baylyi* ADP1 [302], would be that it would be completely unmarked. Though it is possible to convert individual mutants in the *E. coli* library from marked to unmarked, this requires additional DNA manipulation [300], whereas pMo130-Tel^R can generate unmarked mutants in a single step [235]. Furthermore, the creation of an *A. baumannii* mutant library would obviate the need to use *A. baylyi* ADP1 as a genetic proxy for this organism.

As with the first suggested research direction, difficulty in mutant complementation would be the main issue for the above. Although, this issue would only be applicable if a mutant library was constructed using a multidrug resistant isolate of *A. baumannii*, as the pWH1266 vector commonly used for complementation would presumably be ineffective. The most expedient solution to this issue would be to modify pWH1266 with more suitable antimicrobial resistance cassettes, or possibly construct an entirely new vector.

6.5 Conclusions and future directions

In light of the genome sequences of various XN-AB isolates generated through this work, it was demonstrated that the genetic diversity of *A. baumannii* is much greater

than previous genomic studies of nosocomial isolates has revealed. This work had also revealed that genes previously confirmed to play a role in virulence and multidrug resistance in nosocomial *A. baumannii* are also conserved in the genomes of XN-AB isolates. It would be of further interest to decipher the roles of these genes in environmental XN-AB isolates of this organism. This study, to the best of our knowledge, represents the first effort to simultaneously examine the genomes of nosocomial, CA-AB and other extranosocomial isolates of *A. baumannii*, or any other pathogen for that matter.

It seems likely that the XN-AB isolates sequenced as part of this thesis represent only a small fraction of the diversity of *A. baumannii* extant in the world. A future direction of this work would be to greatly extend the genomic surveying of XN-AB strains, and accompany that with characterisation of catabolic, resistance and virulence capabilities. This should provide a more global view of the biogeography of these strains and perhaps provide clues as to the emergence of *A. baumannii* as a serious nosocomial pathogen.

In the course of studying the DAMGEs, a method was devised to isolate circularised GEIs and MGEs, using the classical methods of mitomycin C induction and alkaline lysis. We had also discovered that it was possible to enrich these circularised GEIs/MGEs using an exonuclease known as Plasmid-Safe ATP-Dependent DNase (Epicentre), which selectively degrades linearized dsDNA, whilst leaving circularised and nicked dsDNA intact (unpublished data). There is a potential to experimentally characterise the mobilome of an organism through sequencing of this enriched circular DNA. It is our intention to formulate a methodology that combines classical and contemporary molecular biological techniques to inexpensively map the mobilome of any bacterial organism, and to experimentally validate the bioinformatically predicted loci of MGEs and GEIs.

The investigations we have conducted into the accessory genome of *A. baumannii* have suggested an untapped importance in the antimicrobial resistance or metabolism of this organism. This has also raised the unmet need to better characterise said role of the accessory genome, which we feel can be initially accomplished through a combination of whole GEI/MGE deletion, and biological validation of bioinformatically predicted MGEs and GEIs.

References

1. Peleg, A.Y., H. Seifert, and D.L. Paterson, *Acinetobacter baumannii: emergence of a successful pathogen.* Clinical Microbiology Reviews, 2008. **21**(3): p. 538-582.
2. LSPN. *Genus Acinetobacter*. 2013 [30 July 2013]; Available from: http://www.bacterio.net/a/acinetobacter.html.
3. Baumann, P., *Isolation of Acinetobacter from soil and water.* Journal of Bacteriology, 1968. **96**(1): p. 39-42.
4. Turton, J.F., et al., *Incidence of Acinetobacter species other than A. baumannii among clinical isolates of Acinetobacter: evidence for emerging species.* Journal of Clinical Microbiology, 2010. **48**(4): p. 1445-1449.
5. Vaneechoutte, M., et al., *Naturally transformable Acinetobacter sp. strain ADP1 belongs to the newly described species Acinetobacter baylyi.* Applied and Environmental Microbiology, 2006. **72**(1): p. 932-936.
6. Zhan, Y., et al., *Comparative analysis of the complete genome of an Acinetobacter calcoaceticus strain adapted to a phenol-polluted environment.* Research in Microbiology, 2012. **163**(1): p. 36-43.
7. Arunachalam, M., et al., *Degradation of (+)-catechin by Acinetobacter calcoaceticus MTC 127.* Biochimica et Biophysica Acta (BBA) - General Subjects, 2003. **1621**(3): p. 261-265.
8. Chen, T.-L., et al., *Acinetobacter baylyi as a pathogen for opportunistic infection.* Journal of Clinical Microbiology, 2008. **46**(9): p. 2938-2944.
9. Seifert, H., et al., *The distribution of Acinetobacter species in clinical culture materials.* Zentralblatt für Bakteriologie, 1993. **279**(4): p. 544-552.
10. Nemec, A., et al., *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).* Research in Microbiology, 2011. **162**(4): p. 393-404.
11. Towner, K.J., *Acinetobacter: an old friend, but a new enemy.* Journal of Hospital Infection, 2009. **73**(4): p. 355-363.
12. Roca, I., et al., *The Acinetobacter baumannii oxymoron: commensal hospital dweller turned pan-drug-resistant menace.* Frontiers in Microbiology, 2012. **3**: p. 148.
13. Turton, J.F., et al., *Comparison of Acinetobacter baumannii isolates from the United Kingdom and the United States that were associated with repatriated casualties of the Iraq conflict.* Journal of Clinical Microbiology, 2006. **44**(7): p. 2630-2634.
14. Howard, A., et al., *Acinetobacter baumannii: an emerging opportunistic pathogen.* Virulence, 2012. **3**(3): p. 243-250.
15. Fournier, P.E., et al., *Comparative genomics of multidrug resistance in Acinetobacter baumannii.* PLoS Genetics, 2006. **2**(1): p. 62-72.
16. Vila, J., S. Martí, and J. Sánchez-Céspedes, *Porins, efflux pumps and multidrug resistance in Acinetobacter baumannii.* Journal of Antimicrobial Chemotherapy, 2007. **59**(6): p. 1210-1215.
17. Rice, L.B., *Federal funding for the study of antimicrobial resistance in nosocomial pathogens: no ESKAPE.* Journal of Infectious Diseases, 2008. **197**(8): p. 1079-1081.
18. Boucher, H.W., et al., *Bad bugs, no drugs: no ESKAPE! An update from the Infectious Diseases Society of America.* Clinical Infectious Diseases, 2009. **48**(1): p. 1-12.
19. Wendt, C., et al., *Survival of Acinetobacter baumannii on dry surfaces.* Journal of Clinical Microbiology, 1997. **35**(6): p. 1394-7.

20. Jawad, A., et al., *Survival of Acinetobacter baumannii on dry surfaces: comparison of outbreak and sporadic isolates.* Journal of Clinical Microbiology, 1998. **36**(7): p. 1938-41.
21. Bergogne-Bérézin, E. and K.J. Towner, *Acinetobacter spp. as nosocomial pathogens: microbiological, clinical, and epidemiological features.* Clinical Microbiology Reviews, 1996. **9**(2): p. 148-65.
22. Chen, K.-J., et al., *Endophthalmitis caused by Acinetobacter baumannii: report of two cases.* Journal of Clinical Microbiology, 2008. **46**(3): p. 1148-1150.
23. Charnot-Katsikas, A., et al., *Two cases of necrotizing fasciitis due to Acinetobacter baumannii.* Journal of Clinical Microbiology, 2009. **47**(1): p. 258-263.
24. Corradino, B., et al., *A difficult case of necrotizing fasciitis caused by Acinetobacter baumannii.* The International Journal of Lower Extremity Wounds, 2010. **9**(4): p. 152-154.
25. Fournier, P.E., H. Richet, and R.A. Weinstein, *The epidemiology and control of Acinetobacter baumannii in health care facilities.* Clinical Infectious Diseases, 2006. **42**(5): p. 692-699.
26. Wisplinghoff, H., W. Perbix, and H. Seifert, *Risk factors for nosocomial bloodstream infections due to Acinetobacter baumannii: a case-control study of adult burn patients.* Clinical Infectious Diseases, 1999. **28**(1): p. 59-66.
27. Sebeny, P.J., M.S. Riddle, and K. Petersen, *Acinetobacter baumannii skin and soft-tissue infection associated with war trauma.* Clinical Infectious Diseases, 2008. **47**(4): p. 444-449.
28. Siegman-Igra, Y., et al., *Nosocomial Acinetobacter meningitis secondary to invasive procedures: report of 25 cases and review.* Clinical Infectious Diseases, 1993. **17**(5): p. 843-849.
29. Neely, A.N., M.P. Maley, and G.D. Warden, *Computer keyboards as reservoirs for Acinetobacter baumannii in a burn hospital.* Clinical Infectious Diseases, 1999. **29**(5): p. 1358-1359.
30. Maiden, M.C.J., et al., *Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms.* Proceedings of the National Academy of Sciences, 1998. **95**(6): p. 3140-3145.
31. Dijkshoorn, L., et al., *Comparison of outbreak and nonoutbreak Acinetobacter baumannii strains by genotypic and phenotypic methods.* Journal of Clinical Microbiology, 1996. **34**(6): p. 1519-25.
32. van Dessel, H., et al., *Identification of a new geographically widespread multiresistant Acinetobacter baumannii clone from European hospitals.* Research in Microbiology, 2004. **155**(2): p. 105-112.
33. Brisse, S., et al., *Molecular surveillance of European quinolone-resistant clinical isolates of Pseudomonas aeruginosa and Acinetobacter spp. using automated ribotyping.* Journal of Clinical Microbiology, 2000. **38**(10): p. 3636-3645.
34. Stull, T.L., J.J. LiPuma, and T.D. Edlind, *A broad-spectrum probe for molecular epidemiology of bacteria: ribosomal RNA.* Journal of Infectious Diseases, 1988. **157**(2): p. 280-286.
35. Vos, P., et al., *AFLP: a new technique for DNA fingerprinting.* Nucleic Acids Research, 1995. **23**(21): p. 4407-4414.
36. Antunes, L.C.S., P. Visca, and K.J. Towner, *Acinetobacter baumannii: evolution of a global pathogen.* Pathogens and Disease, 2014. **0**: p. 1-10.
37. Diancourt, L., et al., *The population structure of Acinetobacter baumannii: expanding multiresistant clones from an ancestral susceptible genetic pool.* PLoS ONE, 2010. **5**(4): p. e10034.

38. Karah, N., et al., *Insights into the global molecular epidemiology of carbapenem non-susceptible clones of Acinetobacter baumannii.* Drug Resistance Updates, 2012. **15**(4): p. 237-247.
39. Cooper, J.E. and E.J. Feil, *Multilocus sequence typing – what is resolved?* Trends in Microbiology, 2004. **12**(8): p. 373-377.
40. Bartual, S.G., et al., *Development of a multilocus sequence typing scheme for characterization of clinical isolates of Acinetobacter baumannii.* Journal of Clinical Microbiology, 2005. **43**(9): p. 4382-4390.
41. Snitkin, E.S., et al., *Genome-wide recombination drives diversification of epidemic strains of Acinetobacter baumannii.* Proceedings of the National Academy of Sciences, 2011. **108**(33): p. 13758-13763.
42. Chen, W., *Comment on: AbaR4 replaces AbaR3 in a carbapenem-resistant Acinetobacter baumannii isolate belonging to global clone 1 from an Australian hospital.* Journal of Antimicrobial Chemotherapy, 2012. **67**(2): p. 512.
43. Lewis, T., et al., *High-throughput whole-genome sequencing to dissect the epidemiology of Acinetobacter baumannii isolates from a hospital outbreak.* Journal of Hospital Infection, 2010. **75**(1): p. 37-41.
44. Wright, M.S., et al., *New insights into dissemination and variation of the health care-associated pathogen Acinetobacter baumannii from genomic analysis.* mBio, 2014. **5**(1): p. e00963-13.
45. Hornsey, M., et al., *Whole-genome comparison of two Acinetobacter baumannii isolates from a single patient, where resistance developed during tigecycline therapy.* Journal of Antimicrobial Chemotherapy, 2011. **66**(7): p. 1499-1503.
46. Köser, C.U., et al., *Routine use of microbial whole genome sequencing in diagnostic and public health microbiology.* PLoS Pathogens, 2012. **8**(8): p. e1002824.
47. Davies, J., *Where have all the antibiotics gone?* The Canadian Journal of Infectious Diseases & Medical Microbiology, 2006. **17**(5): p. 287-290.
48. Coyne, S., P. Courvalin, and B. Périchon, *Efflux-mediated antibiotic resistance in Acinetobacter spp.* Antimicrobial Agents and Chemotherapy, 2011. **55**(3): p. 947-953.
49. Hassan, K.A., et al., *Roles of DHA2 family transporters in drug resistance and iron homeostasis in Acinetobacter spp.* Journal of Molecular Microbiology and Biotechnology, 2011. **20**(2): p. 116-124.
50. Magnet, S., P. Courvalin, and T. Lambert, *Resistance-nodulation-cell division-type efflux pump involved in aminoglycoside resistance in Acinetobacter baumannii strain BM4454.* Antimicrobial Agents and Chemotherapy, 2001. **45**(12): p. 3375-3380.
51. Peleg, A.Y., J. Adams, and D.L. Paterson, *Tigecycline efflux as a mechanism for nonsusceptibility in Acinetobacter baumannii.* Antimicrobial Agents and Chemotherapy, 2007. **51**(6): p. 2065-2069.
52. Sun, J.-R., et al., *A truncated AdeS kinase protein generated by ISAba1 insertion correlates with tigecycline resistance in Acinetobacter baumannii.* PLoS ONE, 2012. **7**(11): p. e49534.
53. Yahav, D., et al., *Efficacy and safety of tigecycline: a systematic review and meta-analysis.* Journal of Antimicrobial Chemotherapy, 2011. **66**(9): p. 1963-1971.
54. Marchand, I., et al., *Expression of the RND-type efflux pump AdeABC in Acinetobacter baumannii is regulated by the AdeRS two-component system.* Antimicrobial Agents and Chemotherapy, 2004. **48**(9): p. 3298-3304.

55. Coyne, S., et al., *Overexpression of resistance-nodulation-cell division pump AdeFGH confers multidrug resistance in Acinetobacter baumannii.* Antimicrobial Agents and Chemotherapy, 2010. **54**(10): p. 4389-4393.
56. Damier-Piolle, L., et al., *AdeIJK, a resistance-nodulation-cell division pump effluxing multiple antibiotics in Acinetobacter baumannii.* Antimicrobial Agents and Chemotherapy, 2008. **52**(2): p. 557-562.
57. Rosenfeld, N., et al., *Expression of the resistance-nodulation-cell division pump AdeIJK in Acinetobacter baumannii is regulated by AdeN, a TetR-type regulator.* Antimicrobial Agents and Chemotherapy, 2012. **56**(5): p. 2504-2510.
58. Rajamohan, G., V.B. Srinivasan, and W.A. Gebreyes, *Molecular and functional characterization of a novel efflux pump, AmvA, mediating antimicrobial and disinfectant resistance in Acinetobacter baumannii.* Journal of Antimicrobial Chemotherapy, 2010. **65**(9): p. 1919-1925.
59. Roca, I., et al., *CraA, a major facilitator superfamily efflux pump associated with chloramphenicol resistance in Acinetobacter baumannii.* Antimicrobial Agents and Chemotherapy, 2009. **53**(9): p. 4013-4014.
60. Srinivasan, V., et al., *Genetic relatedness and molecular characterization of multidrug resistant Acinetobacter baumannii isolated in central Ohio, USA.* Annals of Clinical Microbiology and Antimicrobials, 2009. **8**(1): p. 21.
61. Su, X.-Z., et al., *AbeM, an H+-coupled Acinetobacter baumannii multidrug efflux pump belonging to the MATE family of transporters.* Antimicrobial Agents and Chemotherapy, 2005. **49**(10): p. 4362-4364.
62. Srinivasan, V.B., G. Rajamohan, and W.A. Gebreyes, *Role of AbeS, a novel efflux pump of the SMR family of transporters, in resistance to antimicrobial agents in Acinetobacter baumannii.* Antimicrobial Agents and Chemotherapy, 2009. **53**(12): p. 5312-5316.
63. Hassan, K.A., et al., *Transcriptomic and biochemical analyses identify a family of chlorhexidine efflux proteins.* Proceedings of the National Academy of Sciences, 2013. **110**(50): p. 20254-20259.
64. Papamichail, D. and N. Delihas, *Outer membrane protein genes and their small non-coding RNA regulator genes in Photorhabdus luminescens.* Biology Direct, 2006. **1**(1): p. 12.
65. Mussi, M.A., A.S. Limansky, and A.M. Viale, *Acquisition of resistance to carbapenems in multidrug-resistant clinical strains of Acinetobacter baumannii: natural insertional inactivation of a gene encoding a member of a novel family of β-barrel outer membrane proteins.* Antimicrobial Agents and Chemotherapy, 2005. **49**(4): p. 1432-1440.
66. Mussi, M.A., et al., *CarO, an Acinetobacter baumannii outer membrane protein involved in carbapenem resistance, is essential for L-ornithine uptake.* FEBS Letters, 2007. **581**(29): p. 5573-5578.
67. Lee, Y., et al., *A novel insertion sequence, ISAba10, inserted into ISAba1 adjacent to the blaOXA-23 gene and disrupting the outer membrane protein gene carO in Acinetobacter baumannii.* Antimicrobial Agents and Chemotherapy, 2011. **55**(1): p. 361-363.
68. Mussi, M.A., et al., *Horizontal gene transfer and assortative recombination within the Acinetobacter baumannii clinical population provide genetic diversity at the single carO gene, encoding a major outer membrane protein channel.* Journal of Bacteriology, 2011. **193**(18): p. 4736-4748.
69. Bou, G., et al., *Characterization of a nosocomial outbreak caused by a multiresistant Acinetobacter baumannii strain with a carbapenem-hydrolyzing enzyme: high-level carbapenem resistance in A. baumannii is not due solely to*

the presence of β-lactamases. Journal of Clinical Microbiology, 2000. **38**(9): p. 3299-3305.
70. Tomás, M.d.M., et al., *Cloning and functional analysis of the gene encoding the 33- to 36-kilodalton outer membrane protein associated with carbapenem resistance in Acinetobacter baumannii.* Antimicrobial Agents and Chemotherapy, 2005. **49**(12): p. 5172-5175.
71. Quale, J., et al., *Molecular epidemiology and mechanisms of carbapenem resistance in Acinetobacter baumannii endemic in New York City.* Clinical Infectious Diseases, 2003. **37**(2): p. 214-220.
72. Catel-Ferreira, M., et al., *Deciphering the function of the outer membrane protein OprD homologue of Acinetobacter baumannii.* Antimicrobial Agents and Chemotherapy, 2012. **56**(7): p. 3826-3832.
73. Smani, Y. and J. Pachón, *Loss of the OprD homologue protein in Acinetobacter baumannii: impact on carbapenem susceptibility.* Antimicrobial Agents and Chemotherapy, 2013. **57**(1): p. 677.
74. Poole, K., *Efflux-mediated multiresistance in Gram-negative bacteria.* Clinical Microbiology and Infection, 2004. **10**(1): p. 12-26.
75. Roca Subirà, I., et al., *The Acinetobacter baumannii oxymoron: commensal hospital dweller turned pan-drug resistant menace.* Frontiers in Microbiology, 2012. **3**(148): p. 1-30.
76. Houang, E.T.S., et al., *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.* Antimicrobial Agents and Chemotherapy, 2003. **47**(4): p. 1382-1390.
77. Perez, F., et al., *Global challenge of multidrug-resistant Acinetobacter baumannii.* Antimicrobial Agents and Chemotherapy, 2007. **51**(10): p. 3471-3484.
78. Wright, G.D., *Aminoglycoside-modifying enzymes.* Current Opinion in Microbiology, 1999. **2**(5): p. 499-503.
79. Nemec, A., et al., *Diversity of aminoglycoside-resistance genes and their association with class 1 integrons among strains of pan-European Acinetobacter baumannii clones.* Journal of Medical Microbiology, 2004. **53**(12): p. 1233-1240.
80. Hamidian, M. and R.M. Hall, *pACICU2 is a conjugative plasmid of Acinetobacter carrying the aminoglycoside resistance transposon TnaphA6.* Journal of Antimicrobial Chemotherapy, 2013.
81. Hamidian, M., et al., *A GC1 Acinetobacter baumannii isolate carrying AbaR3 and the aminoglycoside resistance transposon TnaphA6 in a conjugative plasmid.* Journal of Antimicrobial Chemotherapy, 2013.
82. Moubareck, C., et al., *GES-11, a novel integron-associated GES variant in Acinetobacter baumannii.* Antimicrobial Agents and Chemotherapy, 2009. **53**(8): p. 3579-3581.
83. Baysarowich, J., et al., *Rifamycin antibiotic resistance by ADP-ribosylation: Structure and diversity of Arr.* Proceedings of the National Academy of Sciences, 2008. **105**(12): p. 4886-4891.
84. Bertini, A., et al., *Characterization and PCR-based replicon typing of resistance plasmids in Acinetobacter baumannii.* Antimicrobial Agents and Chemotherapy, 2010. **54**(10): p. 4168-4177.
85. Héritier, C., L. Poirel, and P. Nordmann, *Cephalosporinase over-expression resulting from insertion of ISAba1 in Acinetobacter baumannii.* Clinical Microbiology and Infection, 2006. **12**(2): p. 123-130.

86. Héritier, C., et al., *Characterization of the naturally occurring oxacillinase of Acinetobacter baumannii.* Antimicrobial Agents and Chemotherapy, 2005. **49**(10): p. 4174-4179.
87. Turton, J.F., et al., *The role of ISAba1 in expression of OXA carbapenemase genes in Acinetobacter baumannii.* FEMS Microbiology Letters, 2006. **258**(1): p. 72-77.
88. Vila, J., et al., *Mutation in the gyrA gene of quinolone-resistant clinical isolates of Acinetobacter baumannii.* Antimicrobial Agents and Chemotherapy, 1995. **39**(5): p. 1201-3.
89. Vila, J., et al., *Quinolone-resistance mutations in the topoisomerase IV parC gene of Acinetobacter baumannii.* Journal of Antimicrobial Chemotherapy, 1997. **39**(6): p. 757-762.
90. Hamouda, A. and S.G.B. Amyes, *Novel gyrA and parC point mutations in two strains of Acinetobacter baumannii resistant to ciprofloxacin.* Journal of Antimicrobial Chemotherapy, 2004. **54**(3): p. 695-696.
91. Cho, Y.J., et al., *Genetic basis of resistance to aminoglycosides in Acinetobacter spp. and spread of armA in Acinetobacter baumannii sequence group 1 in Korean hospitals.* Diagnostic Microbiology and Infectious Disease, 2009. **64**(2): p. 185-190.
92. Li, J., et al., *Colistin: the re-emerging antibiotic for multidrug-resistant Gram-negative bacterial infections.* The Lancet Infectious Diseases, 2006. **6**(9): p. 589-601.
93. Pristovšek, P. and J. Kidrič, *Solution structure of polymyxins B and E and effect of binding to lipopolysaccharide: an NMR and molecular modeling study.* Journal of Medicinal Chemistry, 1999. **42**(22): p. 4604-4613.
94. Moffatt, J.H., et al., *Colistin resistance in Acinetobacter baumannii is mediated by complete loss of lipopolysaccharide production.* Antimicrobial Agents and Chemotherapy, 2010. **54**(12): p. 4971-4977.
95. Arroyo, L.A., et al., *The pmrCAB operon mediates polymyxin resistance in Acinetobacter baumannii ATCC 17978 and clinical isolates through phosphoethanolamine modification of lipid A.* Antimicrobial Agents and Chemotherapy, 2011. **55**(8): p. 3743-3751.
96. Pelletier, M.R., et al., *Unique structural modifications are present in the lipopolysaccharide from colistin-resistant strains of Acinetobacter baumannii.* Antimicrobial Agents and Chemotherapy, 2013. **57**(10): p. 4831-4840.
97. Falagas, M. and P. Rafailidis, *Attributable mortality of Acinetobacter baumannii: no longer a controversial issue.* Critical Care, 2007. **11**(3): p. 134.
98. Harbarth, S., V. Nobr, and D. Pittet, *Does antibiotic selection impact patient outcome?* Clinical Infectious Diseases, 2007. **44**(1): p. 87-93.
99. Eijkelkamp, B.A., et al., *Adherence and motility characteristics of clinical Acinetobacter baumannii isolates.* FEMS Microbiology Letters, 2011. **323**(1): p. 44-51.
100. Harding, C.M., et al., *Acinetobacter baumannii strain M2 produces type IV pili which play a role in natural transformation and twitching motility but not surface-associated motility.* mBio, 2013. **4**(4): p. e00360-13.
101. Carruthers, M.D., et al., *Draft genome sequence of the clinical isolate Acinetobacter nosocomialis strain M2.* Genome Announcements, 2013. **1**(6).
102. Skiebe, E., et al., *Surface-associated motility, a common trait of clinical isolates of Acinetobacter baumannii, depends on 1,3-diaminopropane.* International Journal of Medical Microbiology, 2012. **302**(3): p. 117-128.
103. McQueary, C., et al., *Extracellular stress and lipopolysaccharide modulate Acinetobacter baumannii surface-associated motility.* Journal of Microbiology, 2012. **50**(3): p. 434-443.

104. Clemmer, K.M., R.A. Bonomo, and P.N. Rather, *Genetic analysis of surface motility in Acinetobacter baumannii.* Microbiology, 2011. **157**(9): p. 2534-2544.
105. Costerton, J.W., *Overview of microbial biofilms.* Journal of Industrial Microbiology, 1995. **15**(3): p. 137-140.
106. Gaddy, J.A. and L.A. Actis, *Regulation of Acinetobacter baumannii biofilm formation.* Future Microbiology, 2009. **4**(3): p. 273-278.
107. Brossard, K.A. and A.A. Campagnari, *The Acinetobacter baumannii biofilm-associated protein plays a role in adherence to human epithelial cells.* Infection and Immunity, 2012. **80**(1): p. 228-233.
108. Loehfelm, T.W., N.R. Luke, and A.A. Campagnari, *Identification and characterization of an Acinetobacter baumannii biofilm-associated protein.* Journal of Bacteriology, 2008. **190**(3): p. 1036-1044.
109. Gaddy, J.A., A.P. Tomaras, and L.A. Actis, *The Acinetobacter baumannii 19606 OmpA protein plays a role in biofilm formation on abiotic surfaces and in the interaction of this pathogen with eukaryotic cells.* Infection and Immunity, 2009. **77**(8): p. 3150-3160.
110. Choi, A.H.K., et al., *The pgaABCD locus of Acinetobacter baumannii encodes the production of poly-β-1-6-N-acetylglucosamine, which is critical for biofilm formation.* Journal of Bacteriology, 2009. **191**(19): p. 5953-5963.
111. Tomaras, A.P., et al., *Attachment to and biofilm formation on abiotic surfaces by Acinetobacter baumannii: involvement of a novel chaperone-usher pili assembly system.* Microbiology, 2003. **149**(12): p. 3473-3484.
112. Rumbo-Feal, S., et al., *Whole transcriptome analysis of Acinetobacter baumannii assessed by RNA-sequencing reveals different mRNA expression profiles in biofilm compared to planktonic cells.* PLoS ONE, 2013. **8**(8): p. e72968.
113. Whitehead, N.A., et al., *Quorum-sensing in Gram-negative bacteria.* FEMS Microbiology Reviews, 2001. **25**(4): p. 365-404.
114. Bhargava, N., P. Sharma, and N. Capalash, *Quorum sensing in Acinetobacter: an emerging pathogen.* Critical Reviews in Microbiology, 2010. **36**(4): p. 349-360.
115. Niu, C., et al., *Isolation and characterization of an autoinducer synthase from Acinetobacter baumannii.* Journal of Bacteriology, 2008. **190**(9): p. 3386-3392.
116. Stacy, D.M., et al., *Attenuation of quorum sensing in the pathogen Acinetobacter baumannii using non-native N-acyl homoserine lactones.* ACS Chemical Biology, 2012. **7**(10): p. 1719-1728.
117. Eijkelkamp, B., et al., *Investigation of the human pathogen Acinetobacter baumannii under iron limiting conditions.* BMC Genomics, 2011. **12**(1): p. 126.
118. Dorsey, C.W., et al., *The siderophore-mediated iron acquisition systems of Acinetobacter baumannii ATCC 19606 and Vibrio anguillarum 775 are structurally and functionally related.* Microbiology, 2004. **150**(11): p. 3657-3667.
119. Yamamoto, S., N. Okujo, and Y. Sakakibara, *Isolation and structure elucidation of acinetobactin, a novel siderophore from Acinetobacter baumannii.* Archives of Microbiology, 1994. **162**(4): p. 249-254.
120. Gaddy, J.A., et al., *Role of acinetobactin-mediated iron acquisition functions in the interaction of Acinetobacter baumannii strain ATCC 19606T with human lung epithelial cells, Galleria mellonella caterpillars, and mice.* Infection and Immunity, 2012. **80**(3): p. 1015-1024.
121. Neres, J., et al., *Non-nucleoside inhibitors of BasE, an adenylating enzyme in the siderophore biosynthetic pathway of the opportunistic pathogen Acinetobacter baumannii.* Journal of Medicinal Chemistry, 2013. **56**(6): p. 2385-2405.

122. Proschak, A., et al., *Structure and biosynthesis of fimsbactins A–F, siderophores from Acinetobacter baumannii and Acinetobacter baylyi.* ChemBioChem, 2013. **14**(5): p. 633-638.
123. de Léséleuc, L., et al., *Serum resistance, gallium nitrate tolerance and extrapulmonary dissemination are linked to heme consumption in a bacteremic strain of Acinetobacter baumannii.* International Journal of Medical Microbiology. **304**(3-4): p. 360–369.
124. Russo, T.A., et al., *The K1 capsular polysaccharide of Acinetobacter baumannii strain 307-0294 is a major virulence factor.* Infection and Immunity, 2010. **78**(9): p. 3993-4000.
125. Kenyon, J.J. and R.M. Hall, *Variation in the complex carbohydrate biosynthesis loci of Acinetobacter baumannii genomes.* PLoS ONE, 2013. **8**(4): p. e62160.
126. Lees-Miller, R.G., et al., *A common pathway for O-linked protein-glycosylation and synthesis of capsule in Acinetobacter baumannii.* Molecular Microbiology, 2013. **89**(5): p. 816-830.
127. Iwashkiw, J.A., et al., *Identification of a general O-linked protein glycosylation system in Acinetobacter baumannii and its role in virulence and biofilm formation.* PLoS Pathogens, 2012. **8**(6): p. e1002758.
128. Eveillard, M., et al., *Reservoirs of Acinetobacter baumannii outside the hospital and potential involvement in emerging human community-acquired infections.* International Journal of Infectious Diseases, 2013. **17**(10).
129. Berlau, J., et al., *Isolation of Acinetobacter spp including A. baumannii from vegetables: implications for hospital-acquired infections.* Journal of Hospital Infection, 1999. **42**(3): p. 201-204.
130. Houang, E.T.S., et al., *Epidemiology and infection control implications of Acinetobacter spp. in Hong Kong.* Journal of Clinical Microbiology, 2001. **39**(1): p. 228-234.
131. Scott, P., et al., *An outbreak of multidrug-resistant Acinetobacter baumannii-calcoaceticus complex infection in the US military health care system associated with military operations in Iraq.* Clinical Infectious Diseases, 2007. **44**(12): p. 1577-1584.
132. Choi, J.-Y., et al., *Acinetobacter species isolates from a range of environments: species survey and observations of antimicrobial resistance.* Diagnostic Microbiology and Infectious Disease, 2012. **74**(2): p. 177-180.
133. Zeana, C., et al., *The epidemiology of multidrug-resistant Acinetobacter baumannii: does the community represent a reservoir?* Infection Control and Hospital Epidemiology, 2003. **24**(4): p. 275-279.
134. Byrne-Bailey, K.G., et al., *Prevalence of sulfonamide resistance genes in bacterial isolates from manured agricultural soils and pig slurry in the United Kingdom.* Antimicrobial Agents and Chemotherapy, 2009. **53**(2): p. 696-702.
135. Girlich, D., L. Poirel, and P. Nordmann, *First isolation of the bla$_{OXA-23}$ carbapenemase gene from an environmental Acinetobacter baumannii isolate.* Antimicrobial Agents and Chemotherapy, 2010. **54**(1): p. 578-579.
136. Schaefer, A., et al., *Antibiotic-resistant organisms cultured from Atlantic bottlenose dolphins (Tursiops truncatus) inhabiting estuarine waters of Charleston, SC and Indian River Lagoon, FL.* EcoHealth, 2009. **6**(1): p. 33-41.
137. Houhamdi, L. and D. Raoult, *Experimental infection of human body lice with Acinetobacter baumannii.* The American Journal of Tropical Medicine and Hygiene, 2006. **74**(4): p. 526-531.
138. Patro, T., G.N. Rao, and J. Gopalakrishnan, *Association of Acinetobacter baumannii with a top rot phase of sugarcane redstripe disease in India.* Indian Phytopathology, 2006. **59**(4): p. 501-502.

139. Anstey, N.M., et al., *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.* Journal of Clinical Microbiology, 2002. **40**(2): p. 685-686.
140. Chen, M.-Z., et al., *Severe community-acquired pneumonia due to Acinetobacter baumannii.* Chest, 2001. **120**(4): p. 1072-1077.
141. Wang, J.-T., et al., *Community-acquired Acinetobacter baumannii bacteremia in adult patients in Taiwan.* Journal of Clinical Microbiology, 2002. **40**(4): p. 1526-1529.
142. Ozaki, T., et al., *Community-acquired Acinetobacter baumannii meningitis in a previously healthy 14-month-old boy.* Journal of Infection and Chemotherapy, 2009. **15**(5): p. 322-324.
143. Kim, M.H., et al., *A case of infective endocarditis caused by community-acquired Acinetobacter baumannii.* Korean J Med, 2013. **84**(2): p. 313-317.
144. Falagas, M.E., et al., *Community-acquired Acinetobacter infections.* European Journal of Clinical Microbiology & Infectious Diseases, 2007. **26**(12): p. 857-868.
145. Chen, H.P., et al., *Predictors of mortality in Acinetobacter baumannii bacteremia.* Journal of Microbiology, Immunology and Infection, 2005. **38**(2): p. 127-36.
146. Elliott, J.H., et al., *Community-acquired pneumonia in northern Australia: low mortality in a tropical region using locally-developed treatment guidelines.* International Journal of Infectious Diseases, 2005. **9**(1): p. 15-20.
147. Leung, W.-S., et al., *Fulminant community-acquired Acinetobacter baumannii pneumonia as a distinct clinical syndrome.* Chest, 2006. **129**(1): p. 102-109.
148. Niphadkar, K., et al., *Fulminating septicemia due to persistent pan-resistant community-acquired metallo-β-lactamase (IMP-1)-positive Acinetobacter baumannii.* Indian Journal of Pathology & Microbiology, 2011. **54**(1): p. 180-182.
149. Ong, C.W.M., et al., *Severe community-acquired Acinetobacter baumannii pneumonia: An emerging highly lethal infectious disease in the Asia–Pacific.* Respirology, 2009. **14**(8): p. 1200-1205.
150. Chong, Y.P., et al., *The bacterial etiology of community-acquired pneumonia in Korea: A nationwide prospective multicenter study.* Infection & Chemotherapy, 2010. **42**(6): p. 397-403.
151. Anstey, N.M., B.J. Currie, and K.M. Withnall, *Community-acquired Acinetobacter pneumonia in the Northern Territory of Australia.* Clinical Infectious Diseases, 1992. **14**(1): p. 83-91.
152. Studemeister, A., *Community-acquired Acinetobacter baumannii infections in Northern California.* Clinical Microbiology: Open Access, 2013. **2**(6).
153. Bick, J.A. and J.D. Semel, *Fulminant community-acquired Acinetobacter pneumonia in a healthy woman.* Clinical Infectious Diseases, 1993. **17**(4): p. 820-821.
154. Moreira Silva, G., et al., *Acinetobacter community-acquired pneumonia in a healthy child.* Revista Portuguesa de Pneumologia (English Edition), 2012. **18**(2): p. 96-98.
155. Reindersma, P., L. Nohlmans, and J.J. Korten, *Acinetobacter, an infrequent cause of community acquired bacterial meningitis.* Clinical Neurology and Neurosurgery, 1993. **95**(1): p. 71-73.
156. Park, J.Y., et al., *Complete genome sequence of multidrug-resistant Acinetobacter baumannii strain 1656-2, which forms sturdy biofilm.* Journal of Bacteriology, 2011. **193**(22): p. 6393-6394.

157. Adams, M.D., et al., *Comparative genome sequence analysis of multidrug-resistant Acinetobacter baumannii.* Journal of Bacteriology, 2008. **190**(24): p. 8053-8064.
158. Iacono, M., et al., *Whole-genome pyrosequencing of an epidemic multidrug-resistant Acinetobacter baumannii strain belonging to the European clone II group.* Antimicrobial Agents and Chemotherapy, 2008. **52**(7): p. 2616-2625.
159. Smith, M.G., et al., *New insights into Acinetobacter baumannii pathogenesis revealed by high-density pyrosequencing and transposon mutagenesis.* Genes & Development, 2007. **21**(5): p. 601-614.
160. Vallenet, D., et al., *Comparative Analysis of Acinetobacters: three genomes for three lifestyles.* PLoS ONE, 2008. **3**(3): p. e1805.
161. Zhu, L., et al., *Complete genome analysis of three Acinetobacter baumannii clinical isolates in China for insight into the diversification of drug resistance elements.* PLoS ONE, 2013. **8**(6): p. e66584.
162. Huang, H., et al., *Complete genome sequence of Acinetobacter baumannii MDR-TJ and insights into its mechanism of antibiotic resistance.* Journal of Antimicrobial Chemotherapy, 2012. **67**(12): p. 2825-2832.
163. Zhou, H., et al., *Genomic aqnalysis of the multidrug-resistant Acinetobacter baumannii strain MDR-ZJ06 widely spread in China.* Antimicrobial Agents and Chemotherapy, 2011. **55**(10): p. 4506-4512.
164. Chen, C.-C., et al., *Genome sequence of a dominant, multidrug-resistant Acinetobacter baumannii strain, TCDC-AB0715.* Journal of Bacteriology, 2011. **193**(9): p. 2361-2362.
165. Liu, C.-C., et al., *The origin of Acinetobacter baumannii TYTH-1: a comparative genomics study.* International Journal of Antimicrobial Agents, 2013. **41**(4): p. 318-324.
166. Singh, N.K., et al., *Draft genome sequence of Acinetobacter baumannii strain MSP4-16.* Genome Announcements, 2013. **1**(2).
167. Di Nocera, P., et al., *Genome organization of epidemic Acinetobacter baumannii strains.* BMC Microbiology, 2011. **11**(1): p. 224.
168. Imperi, F., et al., *The genomics of Acinetobacter baumannii: Insights into genome plasticity, antimicrobial resistance and pathogenicity.* IUBMB Life, 2011. **63**(12): p. 1068-1074.
169. Lapierre, P. and J.P. Gogarten, *Estimating the size of the bacterial pan-genome.* Trends in Genetics, 2009. **25**(3): p. 107-110.
170. Mathee, K., et al., *Dynamics of Pseudomonas aeruginosa genome evolution.* Proceedings of the National Academy of Sciences, 2008. **105**(8): p. 3100-3105.
171. Frost, L.S., et al., *Mobile genetic elements: the agents of open source evolution.* Nature Reviews Microbiology, 2005. **3**(9): p. 722-732.
172. Mahillon, J. and M. Chandler, *Insertion sequences.* Microbiology and Molecular Biology Reviews, 1998. **62**(3): p. 725-774.
173. Toleman, M.A. and T.R. Walsh, *Combinatorial events of insertion sequences and ICE in Gram-negative bacteria.* FEMS Microbiology Reviews, 2011. **35**(5): p. 912-935.
174. Mugnier, P.D., L. Poirel, and P. Nordmann, *Functional analysis of insertion sequence ISAba1, responsible for genomic plasticity of Acinetobacter baumannii.* Journal of Bacteriology, 2009. **191**(7): p. 2414-2418.
175. Lopes, B.S. and S.G.B. Amyes, *Role of ISAba1 and ISAba125 in governing the expression of blaADC in clinically relevant Acinetobacter baumannii strains resistant to cephalosporins.* Journal of Medical Microbiology, 2012. **61**(Pt 8): p. 1103-1108.

176. Cole, S.T., et al., *Massive gene decay in the leprosy bacillus.* Nature, 2001. **409**(6823): p. 1007-1011.
177. Bulach, D.M., et al., *Genome reduction in Leptospira borgpetersenii reflects limited transmission potential.* Proceedings of the National Academy of Sciences, 2006. **103**(39): p. 14560-14565.
178. Wishart, W.L., et al., *Escherichia coli RNA polymerase binding sites and transcription initiation sites in the transposon Tn3.* Gene, 1983. **24**(1): p. 99-113.
179. Chalmers, R., et al., *Complete nucleotide sequence of Tn10.* Journal of Bacteriology, 2000. **182**(10): p. 2970-2972.
180. Heffron, F., et al., *DNA sequence analysis of the transposon Tn3: three genes and three sites involved in transposition of Tn3.* Cell, 1979. **18**(4): p. 1153-1163.
181. Roberts, A.P., et al., *Revised nomenclature for transposable genetic elements.* Plasmid, 2008. **60**(3): p. 167-173.
182. Bonnin, R.A., L. Poirel, and P. Nordmann, *A novel and hybrid composite transposon at the origin of acquisition of blaRTG-5 in Acinetobacter baumannii.* International Journal of Antimicrobial Agents, 2012. **40**(3): p. 257-259.
183. Post, V. and R.M. Hall, *AbaR5, a large multiple-antibiotic resistance region found in Acinetobacter baumannii.* Antimicrobial Agents and Chemotherapy, 2009. **53**(6): p. 2667-2671.
184. Rose, A., *TnAbaR1: a novel Tn7-related transposon in Acinetobacter baumannii that contributes to the accumulation and dissemination of large repertoires of resistance genes.* Bioscience Horizons, 2010. **3**(1): p. 40-48.
185. Shaikh, F., et al., *ATPase genes of diverse multidrug-resistant Acinetobacter baumannii isolates frequently harbour integrated DNA.* Journal of Antimicrobial Chemotherapy, 2009. **63**(2): p. 260-264.
186. Bonnin, R.A., L. Poirel, and P. Nordmann, *AbaR-type transposon structures in Acinetobacter baumannii.* Journal of Antimicrobial Chemotherapy, 2012. **67**(1): p. 234-236.
187. Kochar, M., et al., *Deletion of TnAbaR23 results in both expected and unexpected antibiogram changes in a multidrug-resistant Acinetobacter baumannii strain.* Antimicrobial Agents and Chemotherapy, 2012. **56**(4): p. 1845-1853.
188. Šeputienė, V., J. Povilonis, and E. Sužiedėlienė, *Novel variants of AbaR resistance islands with a common backbone in Acinetobacter baumannii isolates of European clone II.* Antimicrobial Agents and Chemotherapy, 2012. **56**(4): p. 1969-1973.
189. Fournier, P.-E., M. Drancourt, and D. Raoult, *Bacterial genome sequencing and its use in infectious diseases.* The Lancet Infectious Diseases, 2007. **7**(11): p. 711-723.
190. Hall, R.M. and H.W. Stokes, *Integrons: novel DNA elements which capture genes by site-specific recombination.* Genetica, 1993. **90**(2-3): p. 115-132.
191. Mazel, D., *Integrons: agents of bacterial evolution.* Nature Reviews Microbiology, 2006. **4**(8): p. 608-620.
192. Collis, C.M., et al., *Site-specific insertion of gene cassettes into integrons.* Molecular Microbiology, 1993. **9**(1): p. 41-52.
193. Collis, C.M. and R.M. Hall, *Site-specific deletion and rearrangement of integron insert genes catalyzed by the integron DNA integrase.* Journal of Bacteriology, 1992. **174**(5): p. 1574-1585.
194. Mazel, D., et al., *A distinctive class of integron in the Vibrio cholerae genome.* Science, 1998. **280**(5363): p. 605-608.

195. Collis, C.M., et al., *Characterization of the class 3 integron and the site-specific recombination system it determines.* Journal of Bacteriology, 2002. **184**(11): p. 3017-3026.
196. Taherikalani, M., et al., *Dissemination of class 1, 2 and 3 integrons among different multidrug resistant isolates of Acinetobacter baumannii in Tehran hospitals, Iran.* Polish Journal of Microbiology, 2011. **60**(2): p. 169-174.
197. Ploy, M.-C., et al., *Molecular characterization of integrons in Acinetobacter baumannii: description of a hybrid class 2 integron.* Antimicrobial Agents and Chemotherapy, 2000. **44**(10): p. 2684-2688.
198. Ramírez, M.S., et al., *Novel insights about class 2 integrons from experimental and genomic epidemiology.* Antimicrobial Agents and Chemotherapy, 2010. **54**(2): p. 699-706.
199. Krupovic, M., et al., *Genomics of bacterial and archaeal viruses: dynamics within the prokaryotic virosphere.* Microbiology and Molecular Biology Reviews, 2011. **75**(4): p. 610-635.
200. Canchaya, C., et al., *Phage as agents of lateral gene transfer.* Current Opinion in Microbiology, 2003. **6**(4): p. 417-424.
201. Hendrix, R.W., et al., *The origins and ongoing evolution of viruses.* Trends in Microbiology, 2000. **8**(11): p. 504-508.
202. Brüssow, H., C. Canchaya, and W.-D. Hardt, *Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion.* Microbiology and Molecular Biology Reviews, 2004. **68**(3): p. 560-602.
203. Rajagopala, S., S. Casjens, and P. Uetz, *The protein interaction map of bacteriophage lambda.* BMC Microbiology, 2011. **11**(1): p. 213.
204. Popova, A.V., et al., *Isolation and characterization of wide host range lytic bacteriophage AP22 infecting Acinetobacter baumannii.* FEMS Microbiology Letters, 2012. **332**(1): p. 40-46.
205. Turton, J.F., C. Perry, and M.J. Hannah, *Isolation of bacteriophage against currently circulating strains of Acinetobacter baumannii.* Journal of Medical Microbiology & Diagnosis, 2012. **1**(3): p. 1000109.
206. Fondi, M., et al., *Exploring the evolutionary dynamics of plasmids: the Acinetobacter pan-plasmidome.* BMC Evolutionary Biology, 2010. **10**(1): p. 59.
207. Balbás, P., et al., *Plasmid vector pBR322 and its special-purpose derivatives — a review.* Gene, 1986. **50**(1–3): p. 3-40.
208. Douard, G., et al., *The Salmonella genomic island 1 is specifically mobilized in trans by the IncA/C multidrug resistance plasmid family.* PLoS ONE, 2010. **5**(12): p. e15302.
209. Doublet, B., et al., *The Salmonella genomic island 1 is an integrative mobilizable element.* Molecular Microbiology, 2005. **55**(6): p. 1911-1924.
210. Boyd, E.F., S. Almagro-Moreno, and M.A. Parent, *Genomic islands are dynamic, ancient integrative elements in bacterial evolution.* Trends in Microbiology, 2009. **17**(2): p. 47-53.
211. Langille, M.G.I., W.W.L. Hsiao, and F.S.L. Brinkman, *Detecting genomic islands using bioinformatics approaches.* Nature Reviews Microbiology, 2010. **8**(5): p. 373-382.
212. Williams, K.P., *Integration sites for genetic elements in prokaryotic tRNA and tmRNA genes: sublocation preference of integrase subfamilies.* Nucleic Acids Research, 2002. **30**(4): p. 866-875.
213. Ogier, J.-C., et al., *Units of plasticity in bacterial genomes: new insight from the comparative genomics of two bacteria interacting with invertebrates, Photorhabdus and Xenorhabdus.* BMC Genomics, 2010. **11**(1): p. 1-21.
214. Dobrindt, U., et al., *Genomic islands in pathogenic and environmental microorganisms.* Nature Reviews Microbiology, 2004. **2**(5): p. 414-424.

215. Grindley, N.D.F., K.L. Whiteson, and P.A. Rice, *Mechanisms of site-specific recombination.* Annual Review of Biochemistry, 2006. **75**(1): p. 567-605.
216. Murphy, R.A. and E.F. Boyd, *Three pathogenicity islands of Vibrio cholerae can excise from the chromosome and form circular intermediates.* Journal of Bacteriology, 2008. **190**(2): p. 636-647.
217. Suhadsaad, A.K.M., A.H. Ad'hiah, and K. Rajakumar, *Genomic island location in Acinetobacter baumannii strains by tRIP-PCR technique.* Current Research Journal of Biological Sciences, 2013. **5**(6): p. 296-301.
218. Nigro, S.J. and R.M. Hall, *GIsul2, 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.* Journal of Antimicrobial Chemotherapy, 2011. **66**(9): p. 2175-2176.
219. Acosta-Cruz, E., et al., *Insights into the 1.59-Mbp largest plasmid of Azospirillum brasilense CBG497.* Archives of Microbiology, 2012. **194**(9): p. 725-736.
220. Varki, A. and T.K. Altheide, *Comparing the human and chimpanzee genomes: searching for needles in a haystack.* Genome Research, 2005. **15**(12): p. 1746-1758.
221. Bochner, B.R., P. Gadzinski, and E. Panomitros, *Phenotype microarrays for high-throughput phenotypic testing and assay of gene function.* Genome Research, 2001. **11**(7): p. 1246-1255.
222. Bochner, B.R., *Global phenotypic characterization of bacteria.* FEMS Microbiology Reviews, 2009. **33**(1): p. 191-205.
223. Borglin, S., et al., *Application of phenotypic microarrays to environmental microbiology.* Current Opinion in Biotechnology, 2012. **23**(1): p. 41-48.
224. Oberhardt, M.A., et al., *Genome-scale metabolic network analysis of the opportunistic pathogen Pseudomonas aeruginosa PAO1.* Journal of Bacteriology, 2008. **190**(8): p. 2790-2803.
225. Bernards, A.T., et al., *Phenotypic characterisation of Acinetobacter strains of 13 DNA-DNA hybridisation groups by means of the Biolog system.* Journal of Medical Microbiology, 1995. **42**(2): p. 113-119.
226. Durot, M., et al., *Iterative reconstruction of a global metabolic model of Acinetobacter baylyi ADP1 using high-throughput growth phenotype and gene essentiality data.* BMC Systems Biology, 2008. **2**(1): p. 85.
227. Mara, K., et al., *Molecular and phenotypic characterization of Acinetobacter strains able to degrade diesel fuel.* Research in Microbiology, 2012. **163**(3): p. 161-172.
228. Minard, G., et al., *Prevalence, genomic and metabolic profiles of Acinetobacter and Asaia associated with field-caught Aedes albopictus from Madagascar.* FEMS Microbiology Ecology, 2013. **83**(1): p. 63-73.
229. Peleg, A.Y., et al., *The success of Acinetobacter species; genetic, metabolic and virulence attributes.* PLoS ONE, 2012. **7**(10): p. e46984.
230. Ishikawa, M. and K. Hori, *A new simple method for introducing an unmarked mutation into a large gene of non-competent Gram-negative bacteria by FLP/FRT recombination.* BMC Microbiology, 2013. **13**(1): p. 86.
231. Aranda, J., et al., *A rapid and simple method for constructing stable mutants of Acinetobacter baumannii.* BMC Microbiology, 2010. **10**(1): p. 279.
232. Sakaguchi, K., et al., *A targeted gene knockout method using a newly constructed temperature-sensitive plasmid mediated homologous recombination in Bifidobacterium longum.* Applied Microbiology and Biotechnology, 2012. **95**(2): p. 499-509.

233. Eijkelkamp, B.A., et al., *Development of a high-throughput cloning strategy for characterization of Acinetobacter baumannii drug transporter proteins.* Journal of Molecular Microbiology and Biotechnology, 2011. **20**(4): p. 211-219.
234. Reyrat, J.-M., et al., *Counterselectable markers: untapped tools for bacterial genetics and pathogenesis.* Infection and Immunity, 1998. **66**(9): p. 4011-4017.
235. Amin, I., et al., *A method for generating marker-less gene deletions in multidrug-resistant Acinetobacter baumannii.* BMC Microbiology, 2013. **13**(1): p. 158.
236. Hunger, M., et al., *Analysis and nucleotide sequence of an origin of an origin of DNA replication in Acinetobacter calcoaceticus and its use for Escherichia coli shuttle plasmids.* Gene, 1990. **87**(1): p. 45-51.
237. Darling, A.C.E., et al., *Mauve: multiple alignment of conserved genomic sequence with rearrangements.* Genome Research, 2004. **14**(7): p. 1394-1403.
238. Valenzuela, J.K., et al., *Horizontal gene transfer in a polyclonal outbreak of carbapenem-resistant Acinetobacter baumannii.* Journal of Clinical Microbiology, 2007. **45**(2): p. 453-460.
239. Nemec, A., L. Dijkshoorn, and T.J.K. van der Reijden, *Long-term predominance of two pan-European clones among multi-resistant Acinetobacter baumannii strains in the Czech Republic.* Journal of Medical Microbiology, 2004. **53**(2): p. 147-153.
240. Aziz, R., et al., *The RAST Server: Rapid Annotations using Subsystems Technology.* BMC Genomics, 2008. **9**(1): p. 75.
241. Farrugia, D.N., et al., *The complete genome and phenome of a community-acquired Acinetobacter baumannii.* PLoS ONE, 2013. **8**(3): p. e58628.
242. Torres, A.G., et al., *TonB-dependent systems of uropathogenic Escherichia coli: aerobactin and heme transport and TonB are required for virulence in the mouse.* Infection and Immunity, 2001. **69**(10): p. 6179-6185.
243. Acosta, J., et al., *Multidrug-resistant Acinetobacter baumannii harboring OXA-24 carbapenemase, Spain.* Emerging Infectious Diseases, 2011. **17**(6): p. 1064.
244. Dorsey, C.W., et al., *Genetic organization of an Acinetobacter baumannii chromosomal region harbouring genes related to siderophore biosynthesis and transport.* Microbiology, 2003. **149**(5): p. 1227-1238.
245. Reeves, S.A., A.G. Torres, and S.M. Payne, *TonB Is required for intracellular growth and virulence of Shigella dysenteriae.* Infection and Immunity, 2000. **68**(11): p. 6329-6336.
246. Chen, I., P.J. Christie, and D. Dubnau, *The ins and outs of DNA transfer in bacteria.* Science, 2005. **310**(5753): p. 1456-1460.
247. Hamidian, M., D.P. Hancock, and R.M. Hall, *Horizontal transfer of an ISAba125-activated ampC gene between Acinetobacter baumannii strains leading to cephalosporin resistance.* Journal of Antimicrobial Chemotherapy, 2013. **68**(1): p. 244-245.
248. Toleman, M.A., et al., *bla_{NDM-1} is a chimera likely constructed in Acinetobacter baumannii.* Antimicrobial Agents and Chemotherapy, 2012. **56**(5): p. 2773-2776.
249. Poirel, L., et al., *Tn125-related acquisition of bla_{NDM}-like genes in Acinetobacter baumannii.* Antimicrobial Agents and Chemotherapy, 2012. **56**(2): p. 1087-1089.
250. Lee, C.-N., et al., *Lytic myophage Abp53 encodes several proteins similar to those encoded by host Acinetobacter baumannii and phage phiKO2.* Applied and Environmental Microbiology, 2011. **77**(19): p. 6755-6762.

251. Nigro, S.J., et al., *A novel family of genomic resistance islands, AbGRI2, contributing to aminoglycoside resistance in Acinetobacter baumannii isolates belonging to global clone 2.* Journal of Antimicrobial Chemotherapy, 2013. **68**(3): p. 554-557.
252. Zhu, Y., et al., *Carnitine metabolism to trimethylamine by an unusual Rieske-type oxygenase from human microbiota.* Proceedings of the National Academy of Sciences, 2014. **111**(11): p. 4268-4273.
253. Tamura, K., et al., *MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods.* Molecular biology and evolution, 2011. **28**(10): p. 2731-2739.
254. Stover, B. and K. Muller, *TreeGraph 2: combining and visualizing evidence from different phylogenetic analyses.* BMC Bioinformatics, 2010. **11**(1): p. 7.
255. Camacho, C., et al., *BLAST+: architecture and applications.* BMC Bioinformatics, 2009. **10**(1): p. 421.
256. Camarena, L., et al., *Molecular mechanisms of ethanol-induced pathogenesis revealed by RNA-sequencing.* PLoS Pathogens, 2010. **6**(4): p. e1000834.
257. Russo, T.A., et al., *Penicillin-binding protein 7/8 contributes to the survival of Acinetobacter baumannii in vitro and in vivo.* Journal of Infectious Diseases, 2009. **199**(4): p. 513-521.
258. Marti, S., et al., *Growth of Acinetobacter baumannii in pellicle enhanced the expression of potential virulence factors.* PLoS ONE, 2011. **6**(10): p. e26030.
259. Zimbler, D.L., et al., *Iron acquisition functions expressed by the human pathogen Acinetobacter baumannii.* BioMetals, 2009. **22**(1): p. 23-32.
260. Bentancor, L.V., et al., *Identification of Ata, a multifunctional trimeric autotransporter of Acinetobacter baumannii.* Journal of Bacteriology, 2012. **194**(15): p. 3950-3960.
261. Carruthers, M.D., et al., *Acinetobacter baumannii utilizes a type VI secretion system for bacterial competition.* PLoS ONE, 2013. **8**(3): p. e59388.
262. Jacobs, A.C., et al., *Inactivation of phospholipase D diminishes Acinetobacter baumannii pathogenesis.* Infection and Immunity, 2010. **78**(5): p. 1952-1962.
263. Providenti, M.A., et al., *Comamonas testosteroni BR6020 possesses a single genetic locus for extradiol cleavage of protocatechuate.* Microbiology, 2001. **147**(8): p. 2157-2167.
264. Hara, H., et al., *The 4-oxalomesaconate hydratase gene, involved in the protocatechuate 4,5-cleavage pathway, is essential to vanillate and syringate degradation in Sphingomonas paucimobilis SYK-6.* Journal of Bacteriology, 2000. **182**(24): p. 6950-6957.
265. Park, S., et al., *Characterization of β-ketoadipate pathway from multi-drug resistance bacterium, Acinetobacter baumannii DU202 by proteomic approach.* The Journal of Microbiology, 2006. **44**(6): p. 632.
266. Abolghait, S.K., et al., *Aeromonas hydrophila PepO outer membrane endopeptidase activates human big endothelin-3 in vitro and induces skin ulcer in goldfish (Carassius auratus).* Veterinary Microbiology, 2010. **145**(1–2): p. 113-121.
267. Ansai, T., et al., *Construction of a pepO gene-deficient mutant of Porphyromonas gingivalis: potential role of endopeptidase O in the invasion of host cells.* Oral Microbiology and Immunology, 2003. **18**(6): p. 398-400.
268. Agarwal, V., et al., *Streptococcus pneumoniae endopeptidase O (PepO) is a multifunctional plasminogen- and fibronectin-binding protein, facilitating evasion of innate immunity and invasion of host cells.* Journal of Biological Chemistry, 2013. **288**(10): p. 6849-6863.
269. Hager, A.J., et al., *Type IV pili-mediated secretion modulates Francisella virulence.* Molecular Microbiology, 2006. **62**(1): p. 227-237.

270. Dijkshoorn, L., A. Nemec, and H. Seifert, *An increasing threat in hospitals: multidrug-resistant Acinetobacter baumannii.* Nature Reviews Microbiology, 2007. **5**(12): p. 939-951.
271. Camp, C. and O.L. Tatum, *A review of Acinetobacter baumannii as a highly successful pathogen in times of war.* Lab Medicine, 2010. **41**(11): p. 649-657.
272. La Scola, B. and D. Raoult, *Acinetobacter baumannii in human body louse.* Emerging Infectious Diseases, 2004. **10**(9): p. 1671-3.
273. Bouvresse, S., et al., *No evidence of Bartonella quintana but detection of Acinetobacter baumannii in head lice from elementary schoolchildren in Paris.* Comparative Immunology, Microbiology and Infectious Diseases, 2011. **34**(6): p. 475-477.
274. Kempf, M., et al., *Detection of Acinetobacter baumannii in human head and body lice from Ethiopia and identification of new genotypes.* International Journal of Infectious Diseases, 2012. **16**(9): p. e680-e683.
275. Sukontason, K.L., et al., *Comparison between Musca domestica and Chrysomya megacephala as carriers of bacteria in northern Thailand.* Southeast Asian Journal of Tropical Medicine and Public Health, 2007. **38**(1): p. 38-44.
276. Gouveia, C., et al., *Study on the bacterial midgut microbiota associated to different Brazilian populations of Lutzomyia longipalpis (Lutz & Neiva) (Diptera: Psychodidae).* Neotropical Entomology, 2008. **37**: p. 597-601.
277. McCarthy, C.B., L.A. Diambra, and R.V. Rivera Pomar, *Metagenomic analysis of taxa associated with Lutzomyia longipalpis, vector of visceral leishmaniasis, using an unbiased high-throughput approach.* PLoS Negl Trop Dis, 2011. **5**(9): p. e1304.
278. Vangnai, A.S. and W. Petchkroh, *Biodegradation of 4-chloroaniline by bacteria enriched from soil.* FEMS Microbiology Letters, 2007. **268**(2): p. 209-216.
279. Lima, T.M.S., et al., *Evaluation of bacterial surfactant toxicity towards petroleum degrading microorganisms.* Bioresource Technology, 2011. **102**(3): p. 2957-2964.
280. Ghai, R., et al., *Metagenomics of the water column in the pristine upper course of the Amazon River.* PLoS ONE, 2011. **6**(8): p. e23785.
281. Huys, G., et al., *Biodiversity of chloramphenicol-resistant mesophilic heterotrophs from Southeast Asian aquaculture environments.* Research in Microbiology, 2007. **158**(3): p. 228-235.
282. Chu, Y.W., et al., *Skin carriage of Acinetobacters in Hong Kong.* Journal of Clinical Microbiology, 1999. **37**(9): p. 2962-2967.
283. Bhattacharya, A. and A. Gupta, *Evaluation of Acinetobacter sp. B9 for Cr (VI) resistance and detoxification with potential application in bioremediation of heavy-metals-rich industrial wastewater.* Environmental Science and Pollution Research, 2013. **20**(9): p. 6628-6637.
284. Suihko, M.-L. and E. Skyttä, *Characterisation of aerobically grown non-spore-forming bacteria from paper mill pulps containing recycled fibres.* Journal of Industrial Microbiology & Biotechnology, 2009. **36**(1): p. 53-64.
285. La Scola, B., et al., *Sequencing of the rpoB gene and flanking spacers for molecular identification of Acinetobacter species.* Journal of Clinical Microbiology, 2006. **44**(3): p. 827-832.
286. Bankevich, A., et al., *SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing.* Journal of Computational Biology, 2012. **19**(5): p. 455-477.
287. Tamtam, F., et al., *Occurrence and fate of antibiotics in the Seine River in various hydrological conditions.* Science of The Total Environment, 2008. **393**(1): p. 84-95.

288. Hossain, M., et al., *Infectious risk assessment of unsafe handling practices and management of clinical solid waste.* International Journal of Environmental Research and Public Health, 2013. **10**(2): p. 556-567.
289. Hrenovic, J., et al., *Occurrence of an environmental Acinetobacter baumannii strain similar to a clinical isolate in paleosol from Croatia.* Applied and Environmental Microbiology, 2014. **80**(9): p. 2860-2866.
290. Servais, P., et al., *Modelling microbiological water quality in the Seine river drainage network: past, present and future situations.* Hydrology and Earth System Sciences, 2007. **11**(5): p. 1581-1592.
291. Ribera, A., et al., *Partial characterization of a transposon containing the tet(A) determinant in a clinical isolate of Acinetobacter baumannii.* Journal of Antimicrobial Chemotherapy, 2003. **52**(3): p. 477-480.
292. Corvec, S., et al., *Genetics and expression of the carbapenem-hydrolyzing oxacillinase gene bla$_{OXA-23}$ in Acinetobacter baumannii.* Antimicrobial Agents and Chemotherapy, 2007. **51**(4): p. 1530-1533.
293. Moffatt, J.H., et al., *Insertion sequence ISAba11 is involved in colistin resistance and loss of lipopolysaccharide in Acinetobacter baumannii.* Antimicrobial Agents and Chemotherapy, 2011. **55**(6): p. 3022-3024.
294. Baba, T., et al., *Genome and virulence determinants of high virulence community-acquired MRSA.* The Lancet, 2002. **359**(9320): p. 1819-1827.
295. Diep, B.A., et al., *Complete genome sequence of USA300, an epidemic clone of community-acquired meticillin-resistant Staphylococcus aureus.* The Lancet. **367**(9512): p. 731-739.
296. Wu, D.-Q., et al., *Genomic analysis and temperature-dependent transcriptome profiles of the rhizosphere originating strain Pseudomonas aeruginosa M18.* BMC Genomics, 2011. **12**(1): p. 438.
297. Sahl, J.W., et al., *Evolution of a pathogen: a comparative genomics analysis identifies a genetic pathway to pathogenesis in Acinetobacter.* PLoS ONE, 2013. **8**(1): p. e54287.
298. Souli, M., et al., *Antimicrobial activity of copper surfaces against carbapenemase-producing contemporary Gram-negative clinical isolates.* Journal of Antimicrobial Chemotherapy, 2012. **68**(4): p. 852-7.
299. Lim, C.K. and D.A. Cooksey, *Characterization of chromosomal homologs of the plasmid-borne copper resistance operon of Pseudomonas syringae.* Journal of Bacteriology, 1993. **175**(14): p. 4492-4498.
300. Baba, T., et al., *Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection.* Molecular Systems Biology, 2006. **2**(1): p. 1-11.
301. Liberati, N.T., et al., *An ordered, nonredundant library of Pseudomonas aeruginosa strain PA14 transposon insertion mutants.* Proceedings of the National Academy of Sciences of the United States of America, 2006. **103**(8): p. 2833-2838.
302. de Berardinis, V., et al., *A complete collection of single-gene deletion mutants of Acinetobacter baylyi ADP1.* Molecular Systems Biology, 2008. **4**(1): p. 1-15.

“Final Approval – EXEMPT” has removed from Open Access version as they may contain sensitive/confidential content.