Mobile Genetic Elements Driving the Evolution of IncHI Plasmids that Confer Antibiotic and Heavy Metal Resistance

Amy Katherine Cain

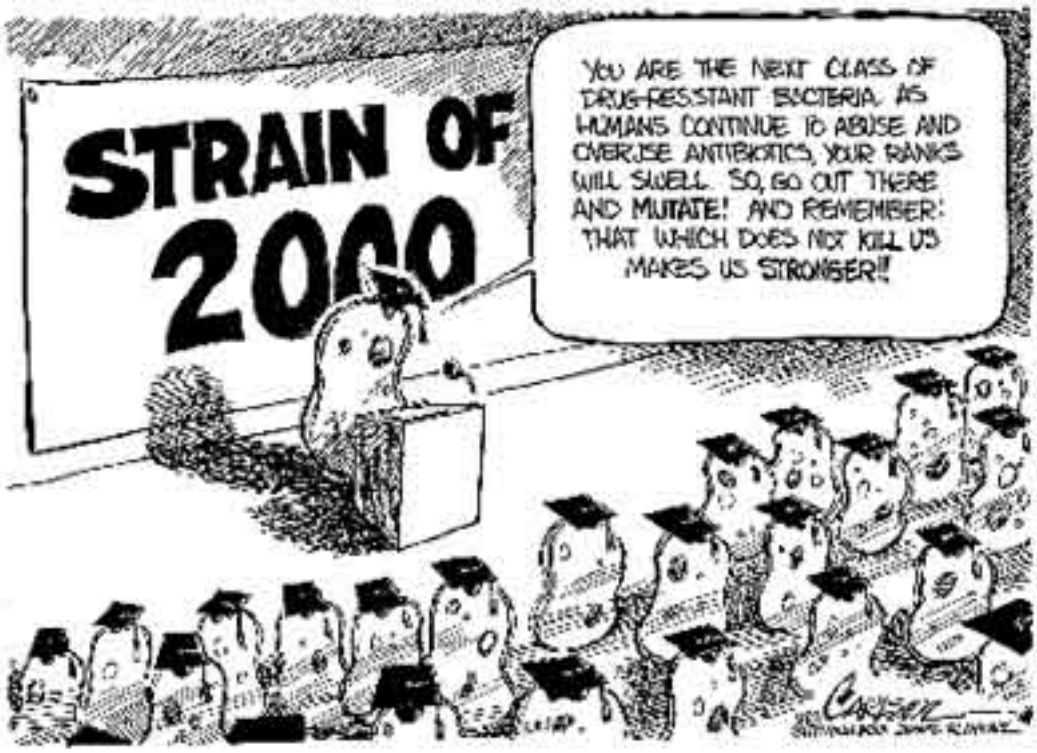
A thesis submitted in fulfillment of the
requirements for the degree of **Doctor of Philosophy**

School of Molecular Bioscience
The University of Sydney
NSW, Australia
2012

Declaration

The thesis presented here contains no material which has been submitted or accepted for any other degree or qualification at any university or institution. To the best of my knowledge this thesis is original and contains no material previously published or written by another person, unless an appropriate reference is stated in the text.

Amy K. Cain
29-2-12


STRAIN OF
2000
YOU ARE THE NEXT CLASS OF DRUG-RESISTANT BACTERIA. AS HUMANS CONTINUE TO ABUSE AND OVERUSE ANTIBIOTICS, YOUR RANKS WILL SWELL. SO, GO OUT THERE AND MUTATE! AND REMEMBER: THAT WHICH DOES NOT KILL US MAKES US STRONGER!!

ACKNOWLEDGEMENTS

First and foremost, I wish to acknowledge my supervisor Prof Ruth Hall, who has been an amazing inspiration as a researcher and as a strong woman. The many hours she spent helping me interpret my data, improving my writing and thinking skills, and the science chats over coffee, have no doubt made me a better scientist. I would also like to acknowledge my co-supervisor Prof. Tom Ferenci, who has given me some invaluable ideas and much help when I needed it. I would also like to express my gratitude to collaborators Dr. Steve Djordjevic and Diane Lightfoot for providing strains used during my PhD, and to Dr. Renee Levings and Jenny Liu for some of the initial characterisation of them. I would like to acknowledge all members of the Hall lab, past and present, for their scientific collaboration, friendship and support: Nick, Neil, Jeremy, Sashi, Steven, Mohammad and Ila. A special thanks to Jannine, who helped with the proof reading of this thesis, and for her kind and encouraging words. A big thanks to Virginia who helped and supported me countless times and shared this PhD journey with me. Thanks to my other fellow PhD students, especially Nick, Sarah, Ali and Will, for being there from the beginning and making this experience as fun as it could be. The Microbiology department have been great for lending me many a reagent or protocol and especially thanks to Nick Coleman and Andy Holmes for their support, advice and wealth of knowledge. Thanks also to Deborah, Helen, Vanessa who have sparked my love for teaching and kept me in a job throughout my PhD. Many thanks to my girls, Sheree, Laura and Jules who have put up with my frequent absences, bad moods and for their encouragement whilst writing this thesis. A huge thanks to Rob who kept me company for some of those long nights, kept my spirits high and got me through some tough times. I could not have done it without your help. Lastly, a big thanks to my parents and to my sister for their love and support, both emotionally and financially. You helped make this mountain seem less steep and this thesis would not have happened without you.

PUBLICATIONS

These publications contain work presented in this thesis:

Manuscripts:

Amy K. Cain, Xiulan Liu, Steven P. Djordjevic, Ruth M. Hall. Transposons Related to Tn*1696* in IncHI2 Plasmids in Multiply Antibiotic Resistant Salmonella enterica Serovar Typhimurium from Australian Animals, *Microbial Drug Resistance* **16**(3): 197-202 (2010)

Amy K. Cain and Ruth M. Hall. Transposon Tn*5393*e carrying the *aphA1*-containing transposon Tn*6023* upstream of *strAB* does not confer resistance to streptomycin, *Microbial Drug Resistance*, ahead of print, 24 June, 2011

Amy K. Cain and Ruth M. Hall, Evolution of IncHI2 Plasmids via Acquisition of Transposons Carrying Antibiotic Resistance Determinants, *Journal of Antimicrobial Chemotherapy*, Feb 14 2012, [Epub ahead of print].

GenBank nucleotide accession numbers:

Cain, A.K. and Hall, R.M., *Salmonella enterica* subsp. enterica serovar Typhimurium IncHI2 plasmid pSRC26 transposon Tn*6026* complete sequence, including transposons Tn*6029* and Tn*4352*, complete sequence, GenBank Accession Number GQ150541, 25,215 bp (Submitted August 2009).

Cain, A.K. and Hall, R.M., *Salmonella enterica* subsp. enterica serovar Typhimurium plasmid pSRC125 transposon Tn*6025*, complete sequence, GenBank Accession Number GU562437.1, 26,072 bp (Submitted September 2010 and updated February 2011).

Cain, A.K. and Hall, R.M., A mosaic multiple antibiotic resistance region from pSRC27-H, an IncHI1 plasmid from *Salmonella enterica*, complete sequence, GenBank Accession Number HQ840942, 45,188 bp (Submitted August 2010).

ABSTRACT

IncHI plasmids are a significant force facilitating the spread of antibiotic and heavy metal resistance among Gram-negative bacteria. The IncHI1 subgroup are often associated with multiply antibiotic resistant typhoid infections and IncHI2 plasmids are not only cited as one of the most common plasmid types in Enterobacteriaceae, but are increasingly being found to carry and spread emerging resistance genes. Both plasmid groups are large and conjugative, and often harbour mosaic and complex antibiotic resistance regions. The work presented in this thesis examined the structure, derivation and position of resistance regions in IncHI plasmids, from *S. enterica* from Australian animals, which were compared to those from sequenced plasmids recovered from a range of sources, in order to investigate the evolution and dissemination of these plasmids.

The IncHI2 plasmids pSRC26 and pSRC125, isolated from *S.* Typhimurium bovine infections harboured the tetracycline resistance transposon Tn*10* and a transposon closely related to the mercuric ion and antibiotic resistance transposon Tn*1696*. These were found in the same position as the equivalent structures in the first sequenced IncHI2 plasmid, R478, and a Tn*1696* remnant was found in the ESBL-encoding plasmid, pK29, indicating a shared history. The configuration around Tn*1696* in pSRC26 and pSRC125 proved to be ancestral, as additional regions had been gained in R478 and pK29, mostly effected by IS*26*. In addition, 3 novel antibiotic resistance transposons were identified: one, a variant of the *strA*/*B*-containing transposon Tn*5393*, that did not confer streptomycin resistance, and 2 transposons bounded by the insertion sequence IS*26*. The backbones of these plasmids were examined using mapping PCRs and RFLP analysis. From this, 2 novel heavy metal resistance transposons were identified in R478, constituting ~15% of its backbone, which were absent in pSRC125 and pSRC26. This analysis allowed the IncHI2 backbone to be more clearly defined. The larger transposon, which confers resistance to copper and silver was found to share features

with Tn*7* and may be a member of a large group of as yet undefined transposons. Two further sets of IncHI2 plasmids were analysed. A set of IncHI2 plasmids from clonal *S*. Infantis chicken isolates contained plasmids related to pSRC125 and pSRC26, yet differences were observed between them. Plasmids from a set of older plasmids (isolated circa 1970), including the reference plasmid TP116, were identified as potential precursors to other IncHI2 plasmids, as they did not confer heavy metal resistances.

The IncHI1 plasmid pSRC27-H, that conferred resistance to a multitude of antibiotics and was isolated from an equine *S*. Typhimurium isolate, formed the basis of one arm of this study. It was found to harbour a large and complex MARR, primarily of Tn*2670* and Tn*10* origin. Similar resistance regions were found in some sequenced IncHI1 plasmids, such as pHCM1, indicating a shared ancestor. Differences in the IncHI1 backbone, some of which were identified as possible mobile genetic elements, allowed 2 major plasmid lineages to be defined. One sequenced IncHI1 was found to be missing the IncF-derived replicon, which is the first report of this occurrence. The stability of antibiotic resistance genes from pSRC27-H was examined by culturing *E. coli* transconjugants without antibiotic selection, over a number of generations. A number pSRC27-H variants, that conferred a different resistance phenotype, were obtained and characterised. IS*1* and IS*26* were found to be the major forces driving the resistance gene loss. Additionally, an apparent stabilising effect of a co-transferred IncI1 plasmid on pSRC27-H was observed when the transconjugants were grown at 37^{O}C.

For both IncHI1 and IncHI2 plasmids, the evolutionary lineages defined during this work demonstrated that these plasmids are clonally spread, that they are distributed globally, that they cross animal and human barriers, and that they persist over time.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

Terms

3'-CS	3'-Conserved Segment
5'-CS	5'-Conserved Segment
Ab^R	Resistant to antibiotics
DNA	Deoxyribonucleic acid
ICE	Integrative and conjugative elements
In	Integron
Inc	Incompatibility
IS	Insertion Sequence
IR	Inverted repeats
Fig.	Figure
HGT	Horizontal Gene Transfer
LHS	Left Hand Side
ME	Mobilisable element
MDU	Microbiological Diagnostic Unit
MGE	Mobile genetic element
MAR	Multiple Antibiotic Resistance
MARR	Multiple Antibiotic Resistance Region
MGE	Mobile Genetic Elements
NCBI	National Centre for Biotechnology information
No.	Number
ORF	Open Reading Frame
RHS	Right Hand Side
Tn	Transposon
TE	Transposable element
VRT	Variable Region Type
ST	Sequence type
S. X	*Salmonella enterica* serovar "X"
X^R	Resistant to the antibiotic "X"
X^S	Sensitive to the antibiotic "X"

Units

aa	Amino Acid
mins	Minutes
secs	Seconds
bp	Base Pairs
kb	Kilobase Pairs
Tm	Melting Temperature
pM	Picomolar
nM	Nanomolar
μM	Micromolar
g	G force
U	Units
v/v	Volume/volume
w/v	weight/volume

Materials and Methods

CDS	Calibrated Dichotomous Sensitivity
DIG	Digoxygenin

dNTPs	Deoxynucleotide Triphosphates mix
EDTA	Ethylenediamine tetraacetic acid
EtBr	Ethidium bromide
HBA	Horse Blood Agar
LA	Luria-Bertani media with agar
LB	Luria-Bertani liquid media
MQ	Mili-Q Water
MHA	Mueller-Hinton Agar
PCR	Polymerase Chain Reaction
RE	Restriction Enzyme
PBRT	PCR-Based Replicon Typing
PDLST	Plasmid Double Locus Sequence Typing
PMLST	Plasmid Multi-Locus Sequence Typing
RAPD	Random amplification of polymorphic DNA
SDS	Sodium dodecyl sulfate

Antibiotics and heavy metals

Ap	Ampicillin
Cm	Chloramphenicol
Fl	Florfenicol
Gm	Gentamycin
Km	Kanamycin
Nm	Neomycin
Nx	Nalidixic Acid
Rf	Rifampicin
Sm	Streptomycin
Sp	Spectinomycin
Su	Sulfamethoxazole
Tb	Tobramycin
Tc	Tetracycline
Tp	Trimethoprim

Ars	Arsenic
Cop	Copper
Hg	Mercury
Pb	Lead
Sil	Silver
Ter	Tellurite

CHAPTER ONE

Introduction

1.1 Overview

Antibiotic resistance can be spread throughout bacterial populations via plasmids. IncHI plasmids are large, conjugative plasmids that have been increasingly associated with antibiotic and heavy metal resistant Enterobacteriaece, particularly *Salmonella*. Recent studies have begun to identify evolutionary lineages within the relatively homogeneous IncHI plasmid subgroups, IncHI1 and IncHI2, allowing the distribution and spread of these plasmids to be monitored. IncHI plasmids often harbour complex and mosaic resistance regions that contain a variety of mobile genetic elements, such as transposable and integrative elements, derived from a range of sources. The evolutionary processes by which these plasmids pick up MGEs and their associated resistance genes are to be discussed, as they are integral to understanding how these regions have arisen and how they will continue to evolve.

1.2 Antimicrobial Resistance

Ever since the use of antibiotics to treat infectious diseases began, bacteria have been developing resistance to them. The use of sulpha drugs in the late 1930s and penicillin in the mid 1940s initiated the antibiotic era. Since then, a number of antibiotics have been discovered and developed, although in the last 20 years only a handful of new antibiotics have been introduced into the market (reviewed in [2]). After each new class of antibiotic is introduced, bacteria develop resistance to it. Bacterial infections that display multiple antibiotic resistance (MAR), defined here as resistance to more than one class of antibiotic, can be extremely difficult to treat. Antibiotic resistance is an increasingly relevant and prevalent health concern and if the problem of antibiotic resistance is not addressed, bacterial infections will be left untreatable.

1.2.1 Antibiotics and antibiotic use in society

The discovery of antibiotics, and their use in clinical contexts, has made an invaluable contribution towards combating bacterial infections in modern society. The copious use of antibiotics, particularly in agricultural settings, increases selective pressure on bacterial populations and encourages the dominance of antibiotic resistant phenotypes. In 2002, it was estimated that global antibiotic sales were $US25 billion [3]. Around 60% of these antibiotics were used for therapeutic purposes in humans and the remaining proportion were used in animals, primarily as growth promoters (which are continuous low levels of antibiotics given to enhance growth of an animal) in animal feed, but also for therapeutic use [4]. In 2006, the European Union banned the use of antibiotics as growth promoters, in an effort to prevent antibiotic resistance from developing in animal-associated bacteria, such as *Salmonella*, which can work their way up the food chain. In Australia, antibiotics continue to be employed as growth promoters in cattle, pig and poultry feed and their use in animals outweighs that for humans [5]. Generally, food-producing farm animals are not administered antibiotics reserved for human use, such as aminoglycosides. However, for therapeutic purposes, gentamicin is the preferred antibiotic used to treat companion animals, such as domestic dogs, cats and even horses. Other antibiotics used to treat human infections, like ampicillin, chloramphenicol and trimethoprim are also administered to these animals therapeutically [6].

1.2.2 Heavy metals as biocides

Heavy metals have been used as an alternative to antibiotics for the control and prevention of bacterial infections. In this context, "heavy metals" refers to metal cations, toxic metals and oxyanions used in clinical and agricultural settings. The heavy metals focussed on in this thesis are silver, copper, tellurite, arsenic and mercury, but commonly the term also includes cadmium, nickel, lead and zinc. Some are non-essential and toxic to the cell (like Ag^{+}, Hg^{2+}) and some are essential trace elements (like Cu^{2+}, Zn^{2+}) that are toxic at high doses. An

example of a clinical use for a heavy metal, is silver sulfadiazine used in burn dressings, to help prevent nosocomial infections [7], or silver salts used in hospital antiseptics [8]. However, bacteria have been developing resistance to silver treatments and the genes conferring this resistance are found both on bacterial plasmids and chromosomes (see for example [9-11]). Heavy metals are also used as biocides and growth promoters in animal feed. In Australia, pig, poultry and cattle farmers use arsenical compounds as growth promoters [4]. Similarly, copper sulfate has been widely used in pig feed ever since Barber and co-workers described its growth-promoting ability in 1955 [12]. Resistance to copper has been described in *E. coli* plasmids from swine gut [13] and in sewage [14], and plasmid-borne arsenic resistance genes have also been identified [15]. Heavy metals are also found in the environment, either as naturally occurring deposits or as by-products from industrial processes (reviewed in [16, 17]). Mercuric ions are a potent example of this, and mercury resistant bacteria have been isolated from a number of environmental sources, for example the soil of industrial plants [18]. However, humans most commonly encounter mercury in dental amalgam fillings, which provide a continuous source of toxic, organic mercury, which is absorbed by the human intestine [19].

1.2.3 Mechanisms of antibiotic resistance

Antibiotic resistance genes encode proteins that allow the bacteria to evade the effect of the antibiotic in a number of ways (Fig 1.1). For example, a range of genes, including bla_{TEM} and bla_{SHV}, encode β-lactamase enzymes that hydrolyse bonds in the β-lactam ring of β-lactam antibiotics, inactivating them [20]. Overall, Gram-negative and Gram-positive bacteria harbour different types of resistance genes. For example, the family of *tet* resistance determinants encode efflux proteins that pump tetracycline out of the cell, enzymes that inactivate the antibiotic, or proteins that protect the bacterial ribosomes (reviewed in [21]).

The most common *tet* genes in Gram-negative bacteria are *tet*(A), -(B), -(C), -(D) and -(G), and in Gram-positives are *tet*(M), -(L), -(K), -(Q) and -(O) [22].

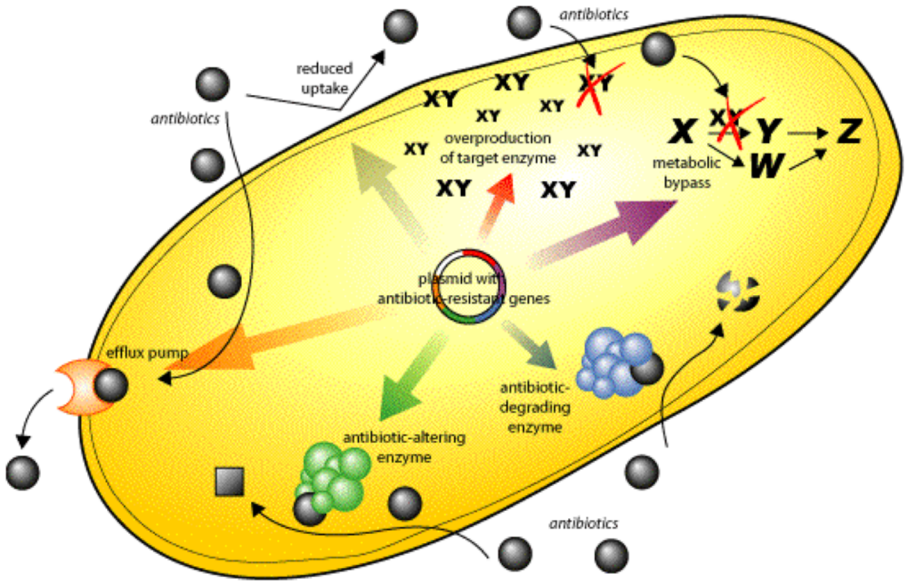


Fig 1.1 Mechanisms of antibiotic resistance. Plasmids can encode efflux proteins that pump antibiotics out of the cell, metabolic bypass systems or proteins that allow antibiotic modification, overproduction or degradation to evade the action of the antibiotic. Chromosomal mutation or intrinsic resistance can also produce these effects. From [23].

Resistance to an antibiotic by an organism can be "intrinsic" where genes determining resistance are already part of the chromosome, for example all *Klebsiella pneumoniae* are resistant to aminopenicillins due to intrinsic β-lactamases [24]. Resistance can also be "acquired" either as a result of chromosomal mutation or when resistance genes that are not usually found in the organism have moved into the cell from an external source. This intercellular movement can occur via a molecular vehicle, such as a plasmid. Resistance genes can be translocated onto these vehicles via intracellular processes, such as transposition.

1.3 Intracellular movement of resistance genes

Mobile genetic elements (MGE), which are discrete DNA segments that are capable of moving themselves and potentially additional DNA within the genome of a cell, play an important role in resistance gene dissemination [25]. Intracellular movement can occur

between different sites within the chromosome or plasmid, between plasmids that co-inhabit a cell or from the chromosome to a plasmid, or vice versa (Fig 1.2). Transposable elements (TE), integrons and gene cassettes are capable of this movement. Another process involved in shaping resistance regions is homologous recombination.

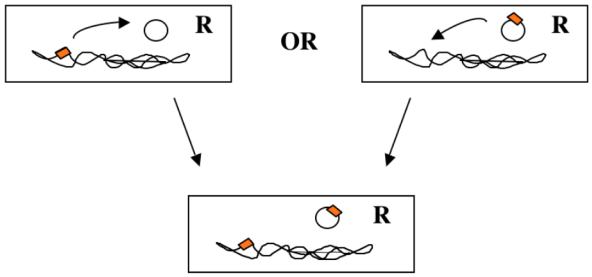


Fig 1.2 Intracellular movement of DNA. The bacterial cell is represented by the rectangle, the chromosome by the wiggly line, and the plasmid a circle. The antibiotic resistance gene is shown as a small orange rectangle, which confers antibiotic resistance to the cell, represented as "R". Figure from Ruth Hall.

1.3.1 Transposable elements

TEs, which include transposons (Tn)s and insertion sequences (IS)s, are MGEs that include elements necessary for transposition, namely a transposase enzyme with a DDE (Asp, Asp(35),Glu) amino acid motif [26, 27] and inverted repeat (IR) sequences, which the transposase recognises as the TE boundaries. TEs can move via conservative transposition, a simple "cut and paste" mechanism, or by replicative transposition which involves a cointegrate intermediate ([28]; see Fig 1.3). The majority of TEs move via replicative transposition, where the end product of their transposition is a cointegrate, which consists of the target and donor DNA molecules joined by 2 copies of the TE (Fig 1.3). The cointegrate then has to be resolved via homologous recombination to produce the final resolution products, which each contain a single copy of the TE (Fig 1.3).

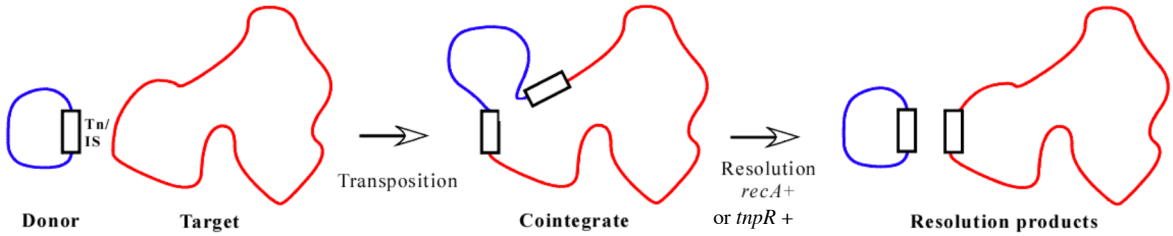


Fig 1.3 TE transposition via cointegrate formation. Target DNA is shown in red and donor DNA is shown in blue. The transposons (Tn) or insertion sequence (IS) is shown as a black box. Each arrow represents the direction of each step via the process named below the arrow. *recA*+ indicates the process of homologous recombination and *tnpR* + the process of resolvase resolution.

The transposition process usually duplicates a short segment of the target DNA, resulting in a repeat on either side of the TE [29]. These duplications are formed during the insertion process, because the target DNA is cleaved in a staggered fashion, the single-stranded ends are joined to the double stranded TE DNA and the ends are filled in by host repair synthesis ([30]; Fig 1.4). If these short direct repeats are not seen flanking a TE, then further events, such as insertion or adjacent deletion, are likely to have occurred since transposition. However, a handful of TE do not create direct repeats upon insertion, for example IS*4321* and IS*5075* [31].

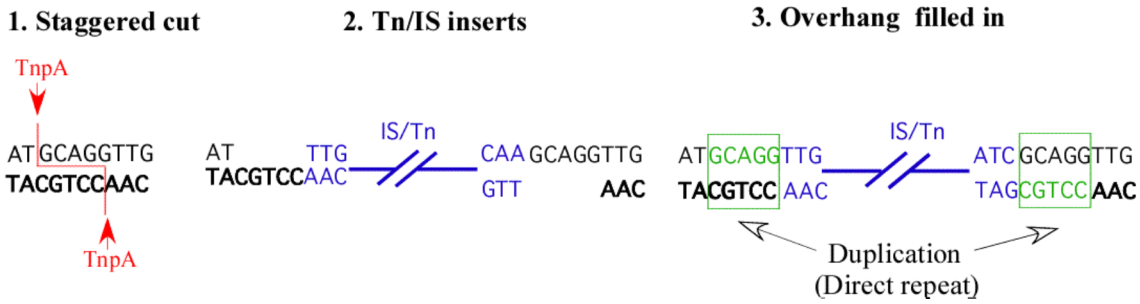


Fig 1.4 Formation of direct repeats. The process involves 3 steps. 1. The transposase (shown as TnpA in red type) makes a staggered cut (shown as a red line) in the target DNA. The target DNA 5' – 3' strand is in normal type, and the 3'- 5' strand is in bold-faced type. 2. The Tn or IS, which shown in blue type, inserts itself. 3. The host systems fill in the single stranded DNA (shown in green type), which results in a duplication or direct repeats (boxed).

Insertion sequences (IS) consist of ~0.5-1.5 kb segment of DNA, that is bounded by IRs, and encode a transposase which catalyses movement (Fig 1.5). The transposase is often encoded by 1 open reading frame (ORF), named *tnp*, which produces a TnpA transposase (Fig 1.5), but can also be encoded by 2 ORFs, named *insA* and *insB* or *orfA* and *orfB*, that produce the

InsA/B or OrfA/B transposase via translational frameshifting [29], as in IS*1* [32]. The direction of the transposase gene/s indicates the orientation of an IS. IS have been expertly reviewed by Chandler and others [29, 33, 34].

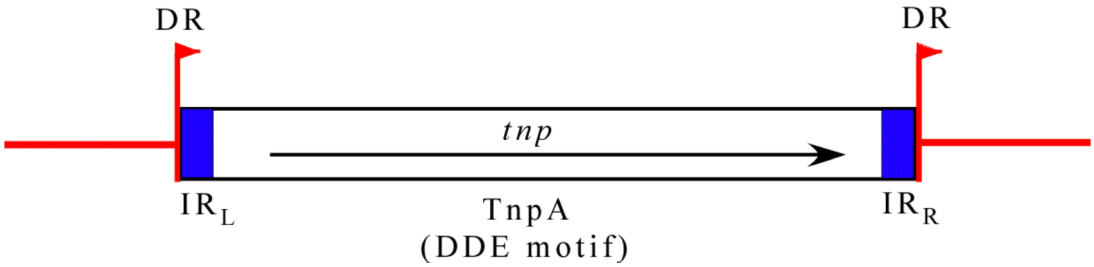


Fig 1.5 Schematic of a generic IS structure. Target DNA is shown in red and the position and direction of direct repeats (DRs) are shown as red flags. The IS is shown as a black box with the direction of the *tnp* open reading frame that encodes the TnpA transpososse is shown as an arrow within. The inverted repeat (IR) sequences at the right- and left-hand boundaries of the IS are marked as blue boxes.

1.3.1.1 Class I transposons

Class I transposons, also termed "compound transposons", have a generic structure that consists of a central DNA segment, flanked by 2 copies of an insertion sequence (IS) [35]. This system of transposon formation is very versatile, with potentially infinite combinations of IS and central regions. Three examples are given in Fig 1.6: Tn*9*, Tn*10* and Tn*4352*. The IS can either be in direct orientation, for example as in Tn*4352* and Tn*9* or opposite orientation, as in Tn*10* (Fig 1.6). The IS encode the transposase and the 2 necessary IRs are the IRs of the 2 outmost IS. The right- and left-hand IS in Class I Tns can vary by nucleotide changes, but their IRs remain highly conserved. For example, the 2 IS*10* in Tn*10* differ by 16 bp and this inactivates the TnpA in one [36]. The central segment can potentially carry any type of gene, for example, the transposon Tn*2555* contains sucrose utilisation genes surrounded by 2 IS*26* [37], but transposons carrying antibiotic resistance genes are most well characterised. The examples in Fig 1.6 carry resistance determinants. Tn*9*, which contains the chloramphenicol resistance gene *catA1* and is bounded by IS*1*, Tn*10* harbours the *tet*(B) tetracycline resistance determinant and is bounded by 2 IS*10* and Tn*4352* carries the *aphA1* kanamycin and neomycin resistance gene flanked by IS*26* (Fig 1.6).

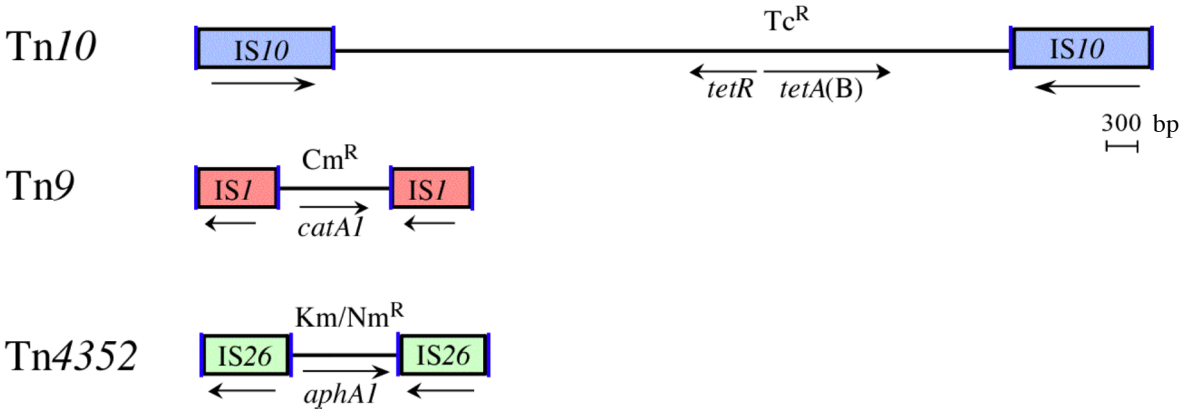


Fig 1.6 Class I transposons Tn*10*, Tn*9* and Tn*4352*. Transposon names as shown on the left. The direction and extent of the antibiotic resistance genes are shown below with the antibiotic resistance they confer above, where Tc is tetracycline, Cm is chloramphenicol and Km/Nm is kanamycin and neomycin. The ISs are shown as coloured boxes with their names inside and the arrow below represents the direction of the transposition gene or genes and orientation of the IS. The IR of the IS are shown as blue, vertical lines. Figs were drawn to scale using GenBank accession nos AP000342 for Tn*10*, AY123253 for Tn*4352* and V00622 for Tn*9*.

1.3.1.2 Class II transposons

Class II transposons move by replicative transposition, via a cointegrate intermediate (see Fig 1.3), the formation of which is catalysed by a transposase and then is resolved via site-specific recombination. Class II Tns contain a transposition module consisting of *tnpA*, that encodes a large transposase, TnpA, of around 1000 amino acids, and the resolution components, most commonly *tnpR*, which encodes a TnpR serine recombinase, and a resolution (*res*) site, which is recognised by the TnpR [38]. TnpR also regulates transcription of *tnpR* and *tnpA*, as it acts as a repressor of the operon [39]. Class II transposons are bounded by IRs, usually of 38 bp in length [38]. In most cases a 5 bp duplication results from transposition [38].

Although all Class II Tns are classified as Tn*3*-family transposons, there are 2 main groups, best represented by Tn*3* and Tn*501*. These differ by the orientation of the *tnpA* and *tnpR* components, as shown in Fig 1.7. The *res* site contains the promoter/s of the *tnpA* and *tnpR* genes. These can be transcribed separately, as in Tn*3* and Tn*3*-like Tns, or co-transcribed, as in Tn*501* and Tn*501*-like Tns (Fig 1.7A) [40]. Tn*5393* is an example of a widely distributed resistance Tn, which is in the same configuration as Tn*3* (Fig 1.7B). The *tnpR* promoter in

res, runs through to allow the co-transcription of the *strA/B* genes [41], which confer streptomycin resistance. Tn*5393* has unusually long, 81 bp inverted repeats, the outer ends of which share 21/38 bp of the 38 bp IR of Tn*3*.

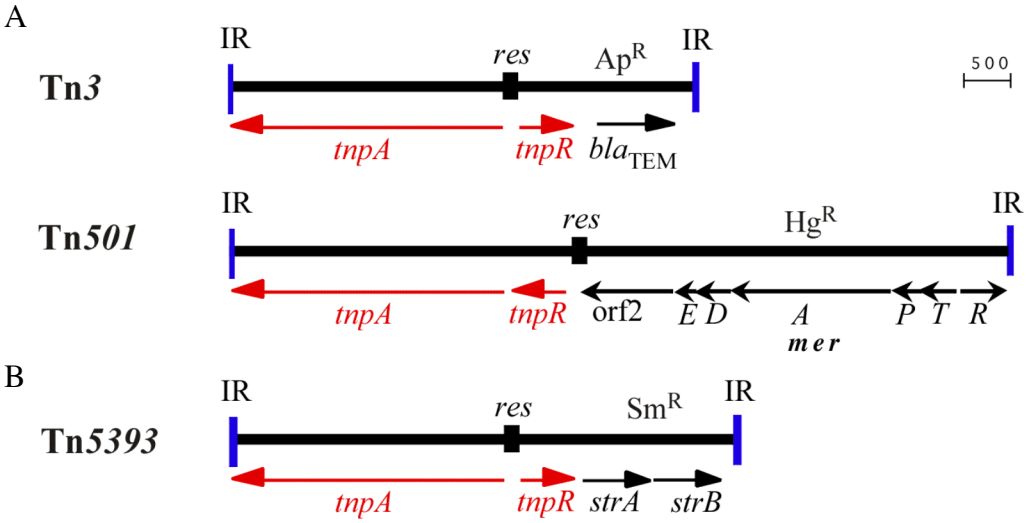


Fig 1.7 Class II transposons A. Tn*3* and **Tn*501*** and **B. Tn*5393*.** The uninterrupted version of Tn*5393* is usually called Tn*5393*c. The *res* site is shown as box, the *tnpA* and *tnpR* genes are highlighted by red lines below with the direction represented by the direction of the arrow. The inverted repeats (IRs) are shown as blue lines. Figs were drawn to scale using GenBank accession nos AY123253 for Tn*3*, AF313472 for Tn*5393*c, and Z00027 for Tn*501*. Other features are as in Fig 1.6.

1.3.1.3 Other transposons - Class III

A number of transposons have been identified and characterised that do not fit the definition of Class I or II transposons. The first one of these recognised was Tn*7*, hence these unclassified transposons often are referred to as Tn*7*-like, although their nucleotide sequence similarity may be very low. Another well-characterised transposon that falls into this potential 3rd class of transposons is Tn*402*. Tn*402* is both a class 1 integron and a transposon. Its transposition involves a set of several *tni* genes, some of which are equivalent to the *tns* genes in Tn*7*, which also encode proteins for transposition. Recently further members of this Tn group, Tn*6021* and Tn*6019*, were described in *Acinetobacter baumanii* [42, 43]. These harbour genes analogous to *tniA* and *tniB*, whose proteins are 33-34% identical to the equivalent TniA and TniB in Tn*402* [43, 44]. A number of features appear to be characteristic

of these "Class III" transposons: they start with the dinucleotide "TG" and end in "AC", they contain ~25 bp IRs, they create a 5 bp duplication upon insertion, and can have target site preference (Ruth Hall, unpublished observations).

1.3.1.3.1 Tn7

Tn*7* is a complicated 14 kb transposon that achieves transposition through both replicative and conservative mechanisms [45, 46]. Transposition can occur in a targeted or a non-specific manner [47]. Tn*7* contains a transposition module consisting of the *tnsA, -B, -C, -D* and *-E* genes ([48]; shown in Fig 1.8). The essential *tnsA* and *tnsB* genes encode transposition proteins: the TnsA component resembles a restriction enzyme and the TnsB component, a transposase with a DDE motif [49, 50]. TnsB also acts as a repressor for *tnsA*/*tnsB* transcription [51]. *tnsC* encodes TnsC, an ATP-binding protein, which regulates the activity of the transposase by interacting with TnsD or TnsE [52]. Either *tnsD* or *tnsE* is expressed and this determines if the insertion is site-specific [53]. The protein combination TnsABC+D results in site- and orientation-specific transposition of Tn*7* into *att*Tn*7*, a site in the transcriptional terminator for the glucosamine-6-phosphate synthase *glmS* gene, which was first identified in *E. coli* and has subsequently been found in a range of Gram-negative bacterial chromosomes [53-55]. Intriguingly, the TnsABC+E protein arrangement results in Tn*7* insertion into a seemingly unspecific sequence, yet preferentially into conjugative plasmids [56]. Tn*7* contains 25 bp IRs, which start with a 5'-TG motif [57] and a 22 bp sequence within the IR, which acts as binding sites for TnsB, is repeated multiple times at both ends of Tn*7* (marked with open-headed vertical arrows in Fig 1.8; [50, 51]). Tn*7* contains a Class 2 integron with 4 integrated gene cassettes (Fig 1.8). Relatives of Tn*7* are also being reported, which encode proteins with amino acid similarity to Tn*7* proteins. For example, *tnsA, B, C, D* genes have been identified on the conjugative plasmid R478 and their

products are between 26-34% identical to the equivalent proteins from Tn*7*, but the boundaries of this element have not been defined ([58]; see section 1.7.2).

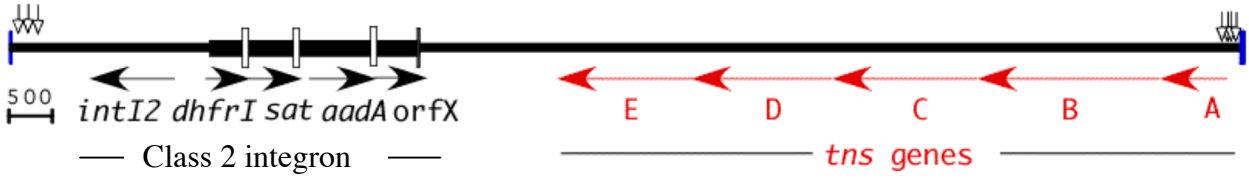


Fig 1.8 Tn*7*. The transposition genes *tnsA-E* are shown in red type. The integron gene cassettes are shown as thicker black lines with their *attC* site as an open box. The open headed arrows mark the approximate positions of the 22 bp repeated sequence for TnsB binding. Drawn to scale using GenBank accession no. AP002527. Other features are as in Fig 1.7.

*1.3.1.3.2 Tn*402 *and Tn*5053

Tn*402* (also named Tn*5090*) is an active transposon, but also a class 1 integron [59]. Tn*402* contains a *tni* transposition module consisting of the genes *tniA*, *B*, *Q* and *R* (Fig 1.9). A similar transposon, named Tn*5053*, contains *tni* genes that are ~90% identical to those of Tn*402* and has been functionally characterised [60]. However, Tn*5053* contains a mercury resistance module, similar to that of Tn*501* (Fig 1.7A), instead of an integron (Fig 1.9; [61]). Both are bounded by 25 bp inverted repeats named IRi (at the *intI* or *merR* end) and IRt (at the *tniA* end) which start with a 5'-TG motif [59]. Within the transposition module, *tniA* encodes a transposase with a DDE motif, TniA, which is 25% identical to TnsB encoded by Tn*7* [57]. The TniB protein, encoded by *tniB*, was found to share 21% identity to TnsC in Tn*7* and is also involved in regulation of transposition [57]. A TniQ protein, which is essential for transposition is also expressed but its function is unknown [60]. The *tniR*-encoded TniR protein resolves the cointegrate intermediate formed after transposition, at the *res* site [60]. Three copies of a 19 bp repeat were detected on the left end of Tn*402* and four on the right, presumably to facilitate transposase binding as in Tn*7* [57, 62].

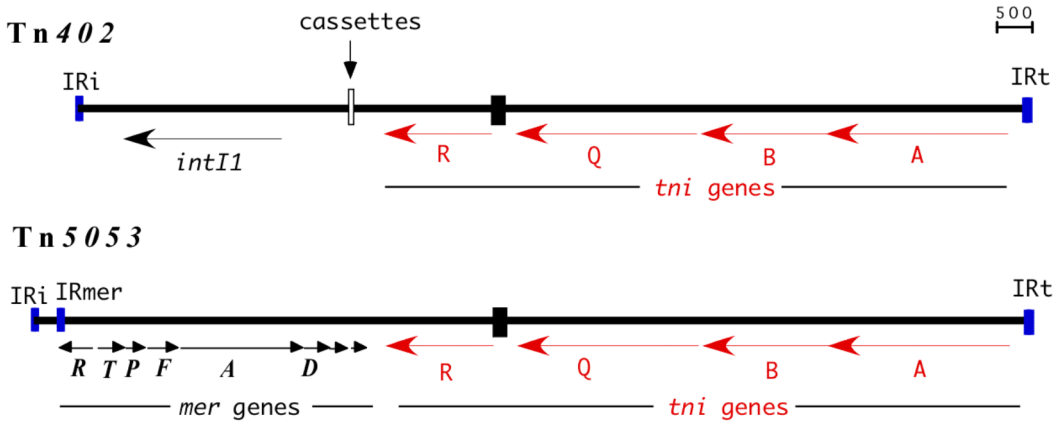


Fig 1.9 Tn*402* and Tn*5053*. The position where the cassettes are integrated is indicated by a vertical arrow. Other features are as in Figs 1.8 and 1.9.

Like Tn*7*, these transposons can transpose in a site-specific manner. They target the resolvase (*res*) regions from Tn*21*-like Class II transposons and the resolvase-encoding partitioning (*par*) regions of plasmids [60, 63-65]. Interestingly, even if the *res* or *par* target sites are present, a functional *tnpR* also has to be provided, either *in cis* or *in trans*, for transposition of Tn*402* or Tn*5053* to occur [63]. Targeted insertion does not occur at a single position within these sites, but rather a short region, which can include a couple of hundred nucleotides of surrounding sequence [63]. For example, from 27 independent experiments examining the transposition of Tn*402* into the *par* region of IncP1α plasmid RP1, 8 unique target sites were found within a 144 bp region [66]. Insertion was also found to occur in an orientation specific manner where the *intI* end is always closest to *res*. In the absence of *par*, Tns related to Tn*5053* have recently been shown to transpose non-specifically at low frequencies and via a separate mechanism involving host-dependent resolution [65].

1.3.1.3.3 Class 1 integrons - defective transposon derivatives

In addition to its transposition module, Tn*402* has a class 1 integron embedded in it, which contains gene cassettes that confer antibiotic resistance. The entire Tn*402* is also considered an integron, and has been speculated to be an ancestor of class 1 integrons that carry an additional 3'-CS [67]. Class 1 integrons represent a major force driving the movement of

antibiotic resistance genes, in the form of mobilisable genes cassettes (see 1.3.2), which are integrated into the integron backbone. Generally, the class 1 integron backbone consists of 2 conserved segments that flank the gene cassette and are bounded by inverted repeats, IRi and IRt [68]. The 5'-CS contains the *intI1* gene and the 3'-CS contains *qacE*Δ which may confer low level resistance to quaternary ammonium compounds [69] and *sul1*, a sulphonamide resistance gene [70]. They fall into 2 major types: In5 type, which can contain IS*1326* includes the integrons In2 and In0 [71], and In4 type, which contains IS*6100* [67], as shown in Fig 1.10. Although these are defective transposons as only part of the *tni* module remains and they cannot move themselves, they can be mobilised *in trans* [72]. In fact, when the transposition genes are provided by an active Tn, the transposition levels can be almost completely restored [65].

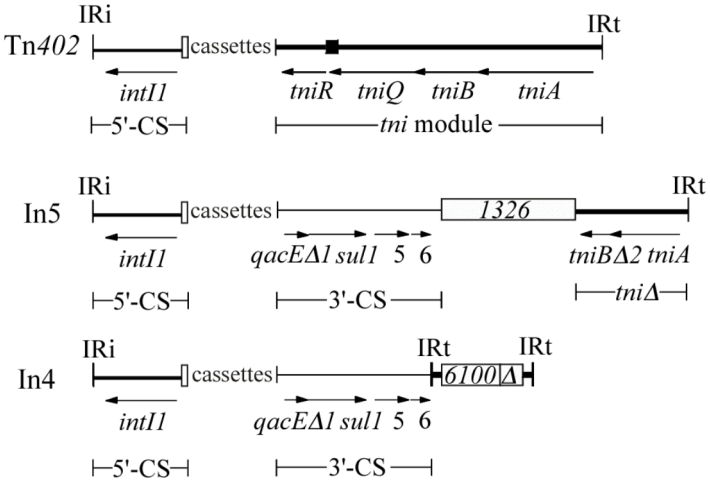


Fig 1.10 In5 and In4. The extent and position of the 5'- and 3'-conserved regions (CS) and the *tni* modules are shown below. The positions of the gene cassettes are represented by "cassettes". The *att*I site is shown as an open box. IS are shown as an open box containing their number and Δ in In4 represents part of IS*6100*. Modified from [73].

1.3.2 Integron gene cassettes

Integrons are genetic elements that encode gene capture systems that integrate discrete mobile units, known as gene cassettes, via site-specific recombination [74-76]. Integration of cassettes is catalysed by an integrase, an enzyme of the tyrosine recombinase family which is

encoded by an *intI* gene ([75, 77]; Fig 1.11). Three major types of integrons have been defined, class 1, class 2 and class 3, based on differences in the amino acid sequences of their integrases named IntI1, IntI2 and IntI3, which are encoded by *intI1*, *intI2* and *intI3*, respectively [76]. Gene cassettes are MGEs and consist usually of a single open reading frame and a short sequence (of 57-141 bp) called an *attC* site, or 59-base element, that acts as specific recombination site [78]. The gene cassettes are integrated at the *attI* site in the integron, but also exist as free, circular molecules ([74]; Fig 1.11). Presently, at least 132 distinct gene cassettes, known to encode antibiotic resistance, have been identified [79].

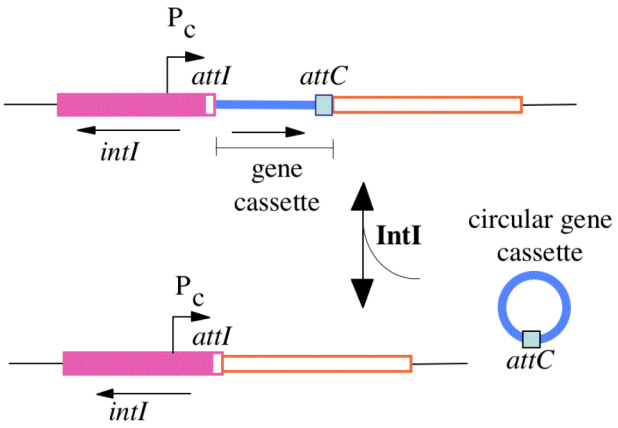


Fig 1.11 Integron and gene cassette interaction. The gene cassette is shown in blue, with the *attC* site as a light blue box. The position and direction of the promoter Pc is shown as a bent arrow. The open orange box represents the 3'-CS and the pink the 5'-CS, with the *attI* site as an open pink box. The integrase protein catalyses the reversible integration of cassettes.

1.3.2.1 Expression of gene cassettes in class 1 integrons

Most gene cassettes do not contain a promoter and thus are only expressed when integrated, via Pc (previously called P_{ant}), the integron promoter. Four types of promoter have been identified, each of different strength, named Pc strong, hybrid 1, hybrid 2 and weak, which can lead to a 20 fold difference in expression levels [80]. The differences are due to sequence variations observed between the -10 and -35 regions (Fig 1.12; [68];[81]). More recently, a TGN motif preceding the -10 hexamer (Fig 1.12) has been shown to increase cassette expression up to 15 fold and is often associated with Pc weak, making it a strong promoter

[82]. A second promoter P2 is also seen in some class 1 integrons, and is usually associated with the weak version of Pc [83].

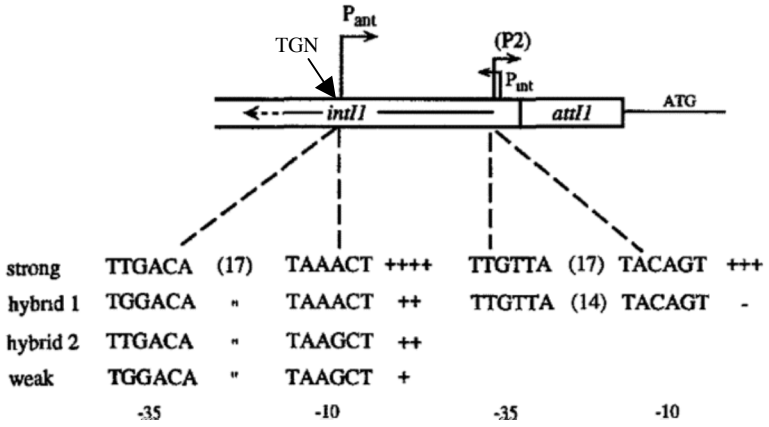


Fig 1.12 Class 1 integron promoter types. The top line shows sequences in *intI1* of the –35 and –10 regions that constitute a strong Pc (or P_{ant}) promoter, the positions of which are shown above, the 2nd line hybrid 1, the 3rd, hybrid 2 and the bottom line, the weak promoter. The (17) indicates a 17 bp spacing in between the –10 and –35 regions. The filled-in arrow marks the position that the TGN motif can be found, preceding the –10 region of Pc. Modified from [68].

1.3.3 Mosaic structures - complex transposons

Transposons can become complex, if other mobile elements, such as integrons or other transposons have inserted within them, often bringing additional resistance genes, without disrupting transposition function. In some cases, the integrated MGE becomes non-mobile, thus its dissemination relies on movement of the entire complex transposon.

1.3.3.1 Class 1 integrons inside Class II transposons

Class II resistance transposons, similar to Tn*501* (Fig 1.7A) have been found that contain a Class 1 integron in or near their *res* site. The most widely distributed example of this is mercury resistance transposons, such as Tn*1403*, Tn*5051*, Tn*1696* and Tn*21*. Two well-studied examples of these, Tn*21* and Tn*1696*, share 82% sequence identity across their backbone (analysis undertaken in this study) and their integrons, In4 and In2, are found in different positions in or near *res*, indicating that they have separate evolutionary origins [67]. Tn*1696*, shown in Fig 1.13A, consists of the In4 integron containing an IS*6100*, within a

Class II mercury resistance transposon equivalent to Tn*5036* [84]. Tn*1696* was first identified in the IncP plasmid R1033 [85], which was isolated in Spain in 1975 [86]. Recently, a relative of Tn*1696*, named Tn*6005*, was identified and it contained an In4-type integron in a Tn*5036* backbone located 6 bp away from where In4 was in Tn*1696* [87], which indicates that In4 has been acquired by Tn*5036* on more than one occasion.

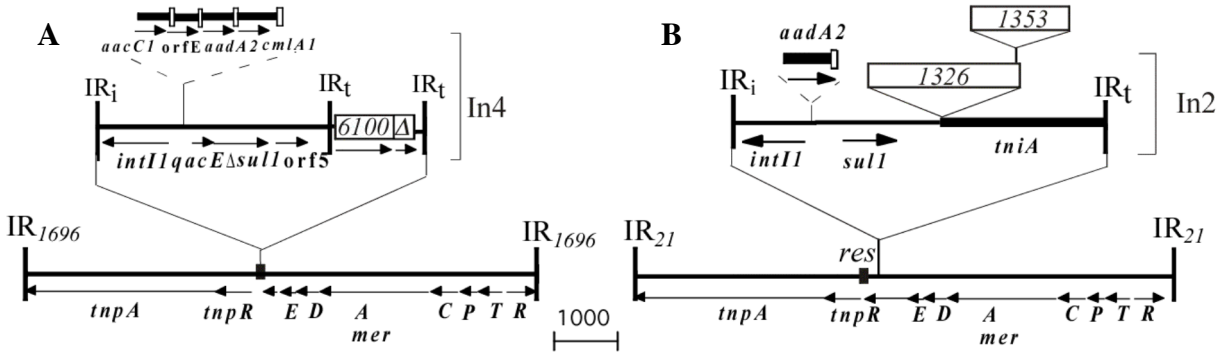


Fig 1.13 Complex mercury II transposons A. Tn*1696* **B.** Tn*21*. The integrons are shown above the mercury II backbone transposons. All other features are as in Figs 1.7, 1.8, 1.11.

Tn*21*, shown in Fig 1.13B, which was originally isolated from the IncFII plasmid NR1, is widely disseminated among both clinical and environmentally isolated Gram-negative bacteria [88]. It contains a mercury II transposon backbone and the integron In2 near the *res* site ([71]; Fig 1.13B). In2 also contains 2 insertion sequences IS*1326* and IS*1353* at the start of the *tni* module. Mercury (II) transposons provide an example of the interlaced dissemination of antibiotic resistance and heavy metal resistance genes.

*1.3.3.2 Tn*2670

A prime example of a transposon inside a transposon, creating a large, functional transposon, is Tn*2670* (Fig 1.14). Tn*2670* consists of Tn*21* (Fig 1.14B) inside a structure similar to Tn*9* (Fig 1.7) and the whole entity can transpose [89]. Transposition of Tn*2670* is effected by the IS*1* elements bounding it and thus its transposition results in an 8 bp duplication, as is consistent with IS*1* transposition [30].

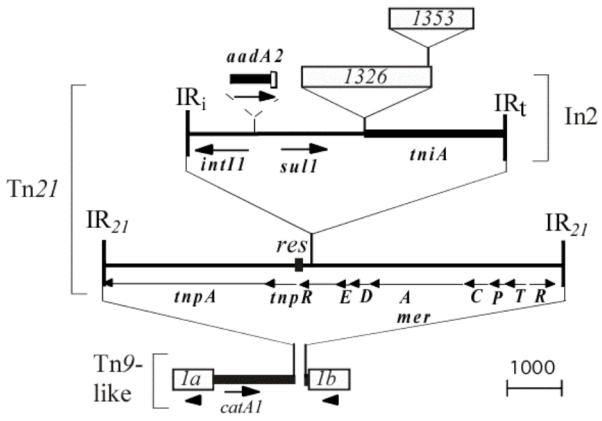


Fig 1.14 Complex transposon Tn*2670*. Transposon Tn*21* is shown above the Tn*9*-like transposon. Other features are as in Figs 1.7, 1.8 and 1.11.

1.3.4 Transposons are ancient structures

The ancestor of the mercury and antibiotic resistance transposon Tn*21*, named Tn*5060*, was identified in *Pseudomonas* isolated from 8,000-10,000 year old Siberian permafrost [90]. Unlike Tn*21*, it does not contain a class 1 integron in its mercury II backbone transposon. Furthermore, plasmids that contain close relatives of the *strA/B*-containing transposon Tn*5393,* that confer streptomycin resistance, have also been isolated from permafrost [91]. Hence, it is clear that transposons, even those encoding antibiotic resistance, are ancient structures that have been circulating for a very long time.

1.3.5 Other forces shaping movement of resistance genes

Movement of resistance genes is commonly attributed to entities described above, however, other forces play a role in rearranging, truncating and swapping segments of host DNA. These include CR elements, which move themselves and adjacent DNA via replicative transposition [84, 92], insertion sequences and host homologous recombination systems.

1.3.5.1 Insertion sequences

Insertion sequences (IS) can insert themselves into host DNA and cause deletion of adjacent DNA [93, 94], which can occur more frequently than transposition ([29, 95]; see Fig 1.15).

Homologous recombination can also act on 2 copies of IS in opposite or direct orientation, resulting inversion or deletion, respectively, of the DNA in between them (Fig 1.15).

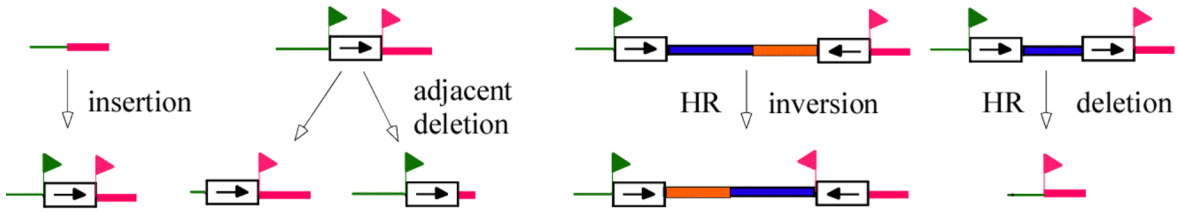


Fig 1.15 Rearrangements to host DNA involving IS. Rearrangements include IS-mediated processes of insertion and adjacent deletion and HR-effected inversion. Green/pink flags are right/left direct repeats, open boxes are insertion sequences and the arrows within them show their orientation.

The properties of some IS commonly encountered in Gram-negatives [33] are shown in Table 1.1. IS can be grouped into IS families, based on shared properties. For example, IS*6100* and IS*26* are both from the IS*6* family (Table 1.1) and share the size of the IRs, DR and have a similar length.

Table 1.1 Properties of important insertion sequences

Name	Length (bp)	IS Family	IR (bp)	DR (bp)	Example Location	Other features	Reference
IS*26*	820	6	14	8	Tn*4352*[1]	Forms tandem arrays	[96]
IS*6100*	880	6	14	8	In4/Tn*1696*		[97]
IS*1*	768	1	23	8[2]	Tn*9*[1]	2 ORFs encode TnpA	[98]
IS*10*	1,329	4	22	9	Tn*10*[1]	Conservative transposition	[99]
IS*4321*	1,326	110	11	None	Tn*21*	Target 38 bp IR of Tns	[31]

[1]a Class 1 Tn, which is bounded by 2 copies of the IS
[2]IS*1* isoforms can create DRs of 7-14 bp; 8 bp or 9 bp are most common

*1.3.5.1.1 IS*26

IS*26* was first identified in Rts1 from *P. vulgaris* in 1983 [96]. It is part of the IS*6* family of insertion sequences. It has many other names: IS*140*, IS*15*Δ, IS*46*, IS*160*, IS*176* and IS*1936* and probably IS*6*. IS*26* contains a –35 region in its IR and can form a strong hybrid promoter if placed with the correct spacing to the –10 region of a resistance gene promoter [100, 101]. The impact that IS*26* has on the rearrangement of genomes and plasmids is being increasingly

noticed [42, 102-105]. IS*257* is also a member of the IS*6* family and has been instrumental in the mobilisation and rearrangement of DNA in *Staphylococcus aureus* [106, 107].

*1.3.5.1.2 IS*1

IS*1* is one of the first characterised bacterial insertion sequences, from R100 [108]. It is also one of the smallest IS in length, measuring 768 bp (Table 1.1). It has a large number of variants or isoforms, which diverge by 0.05-10% at the nucleotide level [29]. IS*1* creates 9 bp direct repeats after insertion. However, the differences within the IS*1* isoform sequences yield altered transposases that result in different sized direct repeats, ranging from 7-14 bp [29]. IS*1* has an insertion site preference for AT-rich sequence [30]. IS*1* contains 2 ORF, *insA* and *insB*, which are needed for transposition: *insA* encodes a protein which binds to the IR of IS*1*. The *insA/B* genes are expressed via translational frameshifting and generate the InsA/B product which catalyses transposition [32, 109].

*1.3.5.1.3 IS*10

IS*10* is unusual as it is part of the only known family of IS that do not move via a cointegrate intermediate, but instead use conservative transposition [110]. IS*10* also has target site preference of AT-rich sequence [29].

*1.3.5.1.4 IS*4321/*IS*5075

IS*4321* and IS*5075* are members of the IS*110* family and their properties are in Table 1.1. They are unusual IS, as they target the 38 bp IRs of Class II mercury resistance transposons [31], such as Tn*1696* or Tn*21* and do not create a direct duplication upon insertion [31].

1.3.5.2 Homologous recombination

Bacterial hosts have intrinsic homologous recombination (HR) machinery, effected by Rec recombination systems. These enable the crossover of identical, or nearly identical sequences.

HR can result in the insertion, deletion or intracellular rearrangement of DNA, which can include resistance genes (Fig 1.16A).

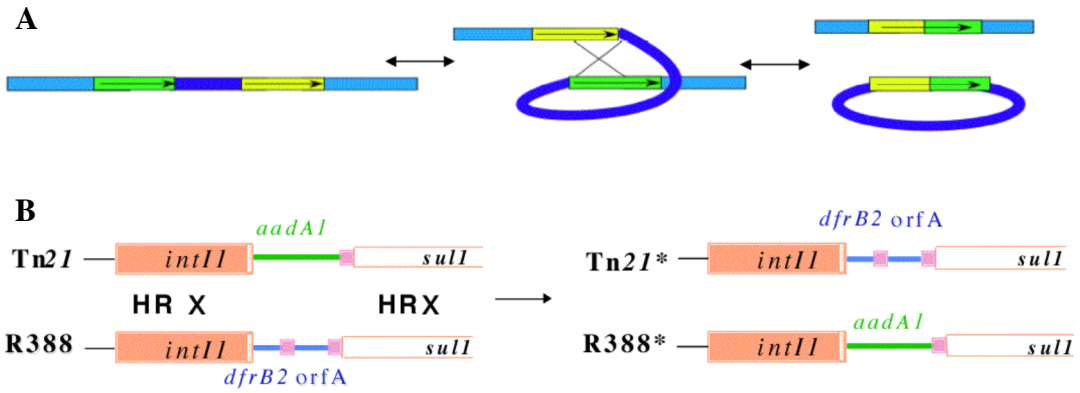


Fig 1.16 Homologous recombination. A HR causing insertion or deletion of host DNA. Host DNA is shown in blue (dark blue highlights the segment that is lost). The yellow and green boxes represent regions of homology and the arrows inside indicate their direction. The black cross indicates the cross over process via RecA. The double-headed arrows indicate that these processes can occur in either direction. **B** HR causing the swapping of gene cassettes. The filled orange box indicates that almost identical sequences of the 5'-CS and the open orange box, the 3'-CS and HR represents the crossover that occurs between them in 2 different class 1 integrons. The cassettes originating from R388, in this example, as shown in blue, whilst those from Tn*21* are in green and the *attC* sites in pink. The * indicates that the structures now contain different cassettes to the original. Fig 1.16B from Ruth Hall.

HR has been observed to occur between class 1 integrons (Fig 1.16B). Recombination occurs between the identical, or near identical, 5'-CS and 3'-CS of different class 1 integrons, which results in their gene cassette arrays being swapped [76, 111], and this has been proposed to be the predominant way in which cassettes are exchanged [72]. HR has also been proposed to be behind the formation of Salmonella Genomic Island 1 (SGI1) variants, where recombination occurs between the 3'-CS regions of a duplicated class 1 integron inside SGI1 to form SGI1-C and between the 5'-CS to yield SGI1-B [112]. Deletions and rearrangements can also occur via very short regions of homology. For example, this has been observed within a class 1 integron, where a 9 bp region of homology between the 5'-CS and orfF appeared to be a site of HR [113]. Similarly, a 13 bp sequence was identified as a HR crossover site, involved in the formation of the streptomycin and sulphonamide resistance region of the IncQ plasmid pSRC15 [114].

1.4 Spread of bacterial resistance

The spread of antibiotic resistant bacteria can be due to the spread of a successful bacterial clone, which either harbours a chromosomal mutation that confers resistance, or which carries antibiotic genes on the chromosome or on an endogenous plasmid. However, antibiotic resistance genes can be disseminated between different bacterial strains or species in a process known as horizontal gene transfer (HGT). There are 3 main ways that DNA can move between cells: transformation, transduction and conjugation (Fig 1.17).

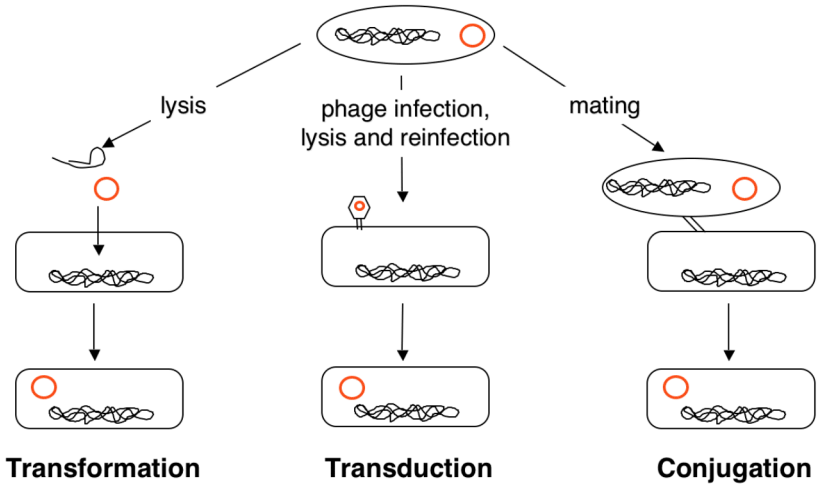


Fig 1.17 Mechanisms of intercellular movement of DNA. DNA can move via transformation, where the cell takes up extracellular DNA, transduction where a bacteriophage infects a cell and injects foreign DNA into the host and conjugation, where direct cell-to-cell contact occurs and DNA is transferred through a conjugation pilus. The orange circle represents plasmid DNA and the larger swiggly line represents chromosome. Figure from Ruth Hall.

First, transformation involves the uptake, integration and expression of extracellular DNA from the environment [25]. In order to take up DNA in this way, cells need to be in a state of competency, which can occur naturally [115], often as a result of a stressed environment, such as starvation or limited growth conditions [116]. Second, transduction involves injection of DNA into a cell by bacteriophage or phage. Phage target specific bacterial hosts and can accidentally package segments of genomic DNA into their capsid, then inject the DNA into a new host cell. If there is substantial homology, a transposon, or an integrative element present, this DNA can integrate with host DNA and be stably inherited [117]. If an entire plasmid was captured (as in Fig 1.17), it can replicate and be maintained. Lastly, DNA

transfer can occur via conjugation which requires 2 cells to be in close contact. The required conjugation apparatus is encoded on specific elements, such as conjugative plasmids or integrative conjugative elements (ICE), once known as conjugative transposons. ICE can move DNA between cells, via conjugation, as well as within a cell via transposition and are rare in Gram-negatives [118]. Conjugation has been implicated as, quantitatively, the primary mode of DNA transfer for many environments [119]. Large, conjugative plasmids are described in detail below.

1.5 Plasmids

1.5.1 Getting to know plasmids

Plasmids are double stranded molecules of DNA that are capable of autonomous replication; they are generally circular, although examples of linear plasmids have been seen in some species [25, 120]. Plasmids contain essential modules, which are responsible for plasmid replication, maintenance and propagation, and many also carry conjugation or mobilisation regions. They can also harbour accessory elements, which provide additional traits that may be beneficial for the host, such as antibiotic and heavy metal resistance. Most plasmids contain 3 essential features: the origin/s of replication (*ori*) which is unique to each replicon, a *rep* region which encodes a replication initiation protein (Rep), and genes controlling the replication process [121, 122]. However, some small plasmids have a minimal origin and do not contain a *rep* gene. For example ColE3 simply contains a 33 bp *ori* [123]. Plasmid replication occurs via one of the theta type, rolling circle, or strand displacement mechanisms [121]. The circular plasmid is either opened followed by RNA priming, as in theta and strand displacement replication, or one of the DNA strands is cleaved to generate a 3'-OH end, as in rolling-circle replication. The most well characterised replication system is theta type, which is often associated with large plasmids. A series of directly repeated sequences, termed iterons, are often found surrounding the *ori* in theta type systems. These facilitate the binding

of the Rep protein and are part of the essential replication control [121]. Plasmids also contain diverse systems that control their copy number and partitioning, to ensure stable inheritance during host cell division [124, 125]. Some plasmids even carry toxin-antitoxin systems, tailored to kill any daughter cells that do not inherit a copy of the plasmid [126].

1.5.2 Plasmid conjugation

Plasmid conjugation provides an efficient process to facilitate the dissemination of resistance genes and occurs in a number of steps. First, the mating pair formation (MPF) complex, which consists of a membrane-spanning protein complex and a surface-exposed sex pilus, brings a recipient cell within close proximity to the donor and when contact is established a mating pair is formed [25, 127, 128]. Conjugation will be inhibited at this stage via entry exclusion systems if a similar plasmid already exists in a recipient cell [127]. The conjugative pilus is a form of a type 4 secretion system (T4SS). Two major lineages of pilus have been identified [129]: F-like pilus (produced by IncF, -H, -T and -J plasmids; Fig 1.18), which are long and flexible, and P-like pilus (produced by IncP, -N, -W and -I plasmids), which are short and rigid [130]. The DNA is prepared for single stranded linear transfer by a relaxase that nicks the circular dsDNA at an *oriT* site and the covalently linked relaxase/ssDNA complex is actively secreted through a mating pore, which is composed of the MPF and a coupling protein called T4CP [131, 132]. In the recipient, the DNA is restored to a double stranded circular form via a relaxasome complex [129]. Recent analysis has shown that there are 6 distinct families of relaxasome proteins, or MOB proteins, ($MOB_{F,}$ $MOB_{-H,}$ $MOB_{-Q,}$ $MOB_{-C,}$ MOB_{-P}, and MOB_{-V}) [133]. The reason conjugation is such an effective route of gene spread is that transfer can occur between such a broad range of organisms, extending from prokaryotes to eukaryotes

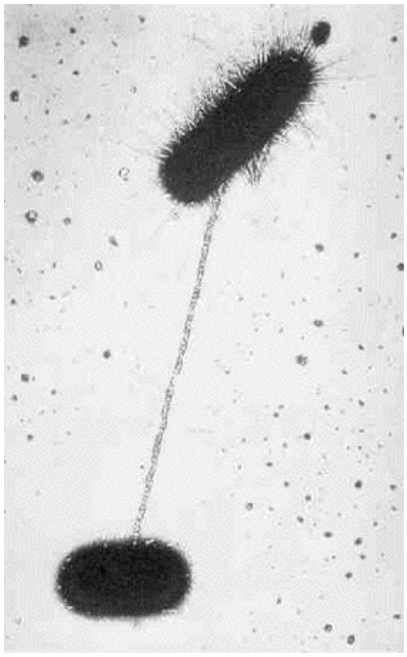


Fig 1.18 *E. coli* to *E. coli* conjugation via an F-pilus.
Taken from [1]

[134]. For example, conjugation of an IncP plasmid from *E. coli* to *Saccharomyces cerevisiae* has been observed [135].

1.5.2.1 Plasmid mobilization

Mobilisable plasmids are typically small and contain only *ori*T and *mob* genes, which encode MOB proteins for relaxasome formation. A non-conjugative (mobilisable) plasmid is able to move when another conjugative plasmid is present, if the remaining conjugation machinery is supplied *in trans*. IncQ plasmids, such as RSF1010 are the most efficient co-mobilised plasmids [136]. Mobilization can also occur in this way for some genomic islands, such as SGI1, which was found to be mobilisable in the presence of an IncC plasmid [137].

1.5.3 Plasmid incompatibility

Plasmids that share identical or highly similar replication or partitioning systems, by a number of complex mechanisms, often result in one plasmid being left unable to replicate leading to a "watering down" effect as the bacterial culture divides [138]. This is known as plasmid incompatibility. Put another way, if 2 normally stable plasmids are not stably maintained when together in a host cell, they are said to be incompatible. The primary plasmid classification system is based on incompatibility (Inc) typing [139, 140] and over 30 distinct Inc groups have been identified [138]. One-way incompatibility can also occur, which is incompatibility only in one direction. Such is the case with IncHI1 and IncF plasmids, where, if an IncHI1 plasmid is introduced into a cell containing an IncF plasmid, the IncF plasmid is lost. Whereas, if an F plasmid enters the IncHI1-containing cell, the 2 are compatible [141].

Plasmids within the same incompatibility group share a number of features that depend on replication functions. For instance plasmid host range varies greatly between incompatibility groups. For example IncF plasmids are found almost exclusively within Enterobacteriacae [142], yet IncQ plasmids are promiscuous and able to replicate in almost all Gram-negative

[143] and some Gram-positive species [144]. In a recent review, the IncA/C, -F, -L/M, -I1, -HI2 and -N were identified as the major plasmid types responsible for the spread antibiotic resistance genes within Enterobactericae [142]. A multiplex PCR scheme, named PCR-based replicon typing (PBRT), targets the replication regions from 18 incompatibility groups in order to rapidly and efficiently distinguish Inc groups [145]. This set has since been extended and refined [146, 147]. This development has prompted a resurgence of the typing of plasmids, as previous incompatibility testing was labour intensive and time consuming. If this trend continues, PBRT will help to build a more comprehensive picture of plasmid distribution in the environment and in clinical contexts, a better understanding of their relationship to each other and host cells, and their evolutionary origins.

1.6 Plasmid evolution

Although incompatibility typing allows grouping of plasmids that may share characteristics, differing degrees of variability are observed with different Inc groups. This variability can be utilised to define plasmid sub-lineages, to extrapolate how plasmid groups have been assembled from the available modular components and to identify potential plasmid ancestors. Analysing evolutionary paths by which plasmids acquire resistance genes is essential, as it can provide insight into epidemiological relationships of plasmid lineages, how plasmids may evolve in the future and how to prevent further development of resistance.

1.6.1 Evolution of plasmids to carry resistance genes

The Murray collection contains plasmids that were isolated between 1917 and 1954, in the pre-antibiotic and early antibiotic eras. One study examining this collection found that of the ~450 isolates, 19% contained conjugative plasmids, none of which conferred antibiotic resistance; however, one isolate was found to confer resistance to the heavy metal tellurite [148]. The conjugative plasmids were found to belong to Inc groups B, F, I, N and X, and

plasmids of the same Inc groups are encountered today. This means the same plasmid types are circulating now as in the pre-antibiotic era but they have evolved by picking up antibiotic resistance genes. Another study furthering this work on the Murray collection, using restriction length polymorphism (RFLP) analysis showed that the cryptic plasmids were similar to recent ones of the same Inc group that conferred antibiotic resistance [149]. Bacterial resistance to mercury has also been observed in the Murray collection of pre-antibiotic era plasmids [150]. The mercury resistance genes were located on antibiotic sensitive ancestors of the complex mercury resistance transposons Tn*21* and Tn*1696*, which have since acquired class 1 integrons that carry antibiotic resistance gene cassettes, within their *res* sites (see Fig 1.13).

1.6.2 Methods used to study plasmid evolution

Historically, plasmids from the same Inc groups have been studied and compared using a number of methods. The application of electron microscopy (EM) in the 1970s to study heteroduplex formation between plasmids from an Inc group that carry different antibiotic resistance genes allowed great advances in the field [151]. It revealed that resistance genes could be located in the same position, indicating a shared evolutionary history, or in different positions, showing that separate acquisition events had occurred within a shared backbone. For example, IncW plasmids [152] and IncP-α plasmids [153] were tested in this way. Around the same time, the power of RFLP analysis was harnessed to compare plasmids and locate fragments that had increased in size, indicating insertions of resistance genes. For instance, the small, promiscuous IncW plasmids [154] were analysed using RFLP with multiple restriction enzymes [155, 156].

Recently, bio-informatic analysis of the increasing number of complete plasmid sequences available has allowed more in-depth analysis of lineages within plasmid groups. Using this approach, IncW plasmids were found to have fairly high sequence identity (>95%) and based

on single nucleotide polymorphisms (SNPs), 2 lineages were found [157]. Similarly, IncA/C plasmids, which are large and often associated with MAR phenotype, have been shown to be a quite diverse group [158, 159]. Some IncA/C plasmids have similar resistance regions, for example ones of Tn*21*-derivation, but these are located in different positions in each plasmid, indicating separate insertion events [159].

1.6.2.1 Analysing large MAR plasmids

Large plasmids (>150 kb), such as those belonging to the IncHI1, IncHI2 and IncA/C groups, can be difficult to analyse due to difficulties in extracting high quality DNA (as described in [160]). This hinders standard laboratory techniques, such as RFLP and transformation. A number of approaches have been utilised to analyse large plasmids. First, the sequencing of an entire large plasmid obviously yields a large amount of information and allows in-depth analysis, for example, by SNP-typing. Sequencing allows the structure of complex resistance regions, which often contain repeated sequences and gene remnants, to be elucidated. However, it has the limitation that only small numbers of plasmids can be studied, due to current costs and because large plasmids, and in particular their variable resistance regions, can be complicated and time-consuming to unravel. Second, sets of PCRs that map the plasmid can be designed, based on the sequences of similar plasmids. This method has been used to analyse IncHI2 plasmids (discussed in section 1.7.2 below; [161]) and IncA/C plasmids. Using a set of 12 PCRs, 70 isolates containing IncA/C plasmids produced amplicons for all PCRs, indicating that their IncA/C plasmids share a well conserved backbone, despite being isolated from a number of species including *Yersinia pestis*, *E. coli*, *S. enterica* and *Klebsiella* [162]. A cost-efficient extension of these mapping PCRs, is multiplex PCRs, which can target 2-3 sites in a single reaction, as was used for a set of 3 genes in IncHI2 plasmids [160]. PCRs to map plasmids, if designed well, can yield a great deal of information on which parts of the backbone are present and can be used to screen large plasmid sets. Third, microarray chips can be used to detect the presence or absence of

genes in plasmids. However, their location is not determined and when studying resistance regions, which often contain repeated elements like IS, they will be detected but no information is gained on their copy number or location. Fourth, plasmid multi-locus sequence typing (PMLST) schemes have been recently created, based on the complete sequence of one, or a few, representatives of an Inc group. PMLST allocates sequence types (ST)s to plasmids based on SNPs within a set of conserved loci. PMLST schemes have been used to analyse IncHI1 plasmids (analysed in section 1.7.1 below; [163]) and IncI1 plasmids [164]. PMLST has the potential to give useful information, but is possibly not powerful enough to differentiate lineages of closely related plasmids. It is currently limited by high costs of sequencing for large plasmid sets and sequencing errors, since a single base change can indicate a different allele and ST. This becomes especially relevant when analysing older plasmid sequences that potentially contain errors. It remains to be seen which of these methods is most effective, but collectively, they certainly allow analysis of large plasmids on a molecular level.

1.7 IncH plasmids

IncH plasmids are large (>150 kb), have an iteron-containing theta based replication system and are conjugative [58, 165, 166]. The conjugation pilus, or H-pilus, produced by IncH plasmids is of the long and flexible F-type [129]. Initially, IncH plasmids were classified as a single incompatibility group on the basis that each produced an H-pilus [167]. The group was subdivided into IncHI and IncHII plasmids, when a plasmid was identified that produced an H-pilus, yet was able to briefly coexist in a host cell with other IncH plasmids [168]. IncHII plasmids are not commonly encountered, have not been greatly studied and none have been sequenced. The IncHI group was further broken down into IncHI1, -HI2 and -HI3 plasmids, based on RFLP analysis and southern hybridisation studies of representative members [167, 169, 170]. Within subgroups, IncHI plasmids were found to share a number of restriction

fragments but a minimal number of fragments were common across groups [171]. To date, only one IncHI3 plasmid has been found, MIP233, [172]. Many examples of the IncHI1 and IncHI2 plasmid groups were found. They were readily distinguished based on a number of differences, such as the one-way incompatibility between IncHI1 and IncF plasmids [173], the resistance that IncHI2 plasmids confer to certain bacteriophages and their surface exclusion properties [174].

IncHI1 and IncHI2 plasmids share some core features, for example, they have a copy number of 1-2 per chromosome [175] and display temperature sensitive conjugation [176]. IncHI plasmids are cited as having optimal transfer at 22-30°C [176], which electron microscopy studies found was due to the slow formation of the H-pilus at higher temperatures [177]. Interestingly, the conjugation frequency of IncHI2 plasmids has been shown to be dramatically decreased when donor cells were grown at 37°C, but largely unaffected by higher mating temperatures [178]. Sequencing of IncHI1 [179] and IncHI2 [58] plasmid representatives has revealed that they share one replication region with 95% identity, which is responsible for their incompatibility, but each also contains a unique replication region [58]. The open reading frames in the sequenced IncHI2 plasmid R478 and the IncHI1 plasmids R27 and pHCM1 were compared and related ORFs (with >35% sequence identity) were numerous (Fig 1.19; [58]).

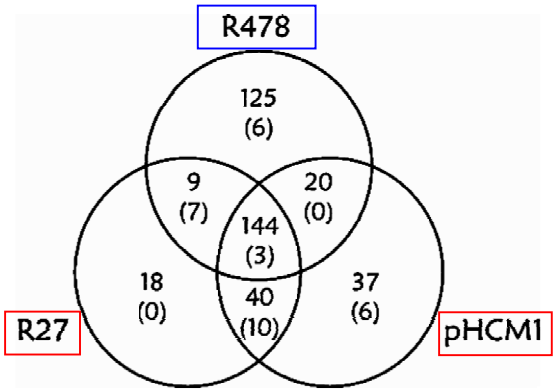


Fig 1.19 Venn diagram of IncHI shared ORFs. The circles represent each plasmid, the number represent the number of ORFs and the overlapping parts of the circles show ORFs that are share at least 35% sequence identity. The number in brackets below shows the number of insertion sequences, not included in the above number. IncHI2 plasmid is boxed in blue and the IncHI1 plasmids are box in red. Modified from [58].

In fact, IncHI1 and IncHI2 plasmids share large, continuous segments. For example, they both have 2 transfer regions, namely Tra1 and Tra2, of ~20 kb and ~35 kb, respectively that share between 71-78% nucleotide identity across approximately 75% of their length (determined in this study using BLASTn analysis). In both plasmid types, the partitioning (Par) region, shown to stabilize the plasmid, is within Tra2 [180].

1.7.1 IncHI1 plasmids

IncHI1 plasmids are commonly associated with MAR in *S.* Typhi, the causative agent of Typhoid fever [163, 179, 181]. However, IncHI1 plasmids have also been shown to transfer across a range of Gram-negative hosts, including Enterobacteriaceae and environmental organisms, at temperatures between 20-30^{O}C [177].

1.7.1.1 Sequenced IncHI1 plasmids and their resistance regions

Currently, five IncHI1 plasmids have been sequenced: R27, pHCM1, pAKU_1, pMAK-1 and pO111_1 (Table 1.2). R27, which was isolated from a tetracycline resistant UK *S.* Typhi isolate in 1961 (Table 1.2), was the first to be sequenced, in 2000 [179]. It is the archetypal IncHI1 plasmid and has been extensively studied (see for example, [180, 182-184]). R27 was found to contain 2 functional replication regions using restriction mapping and cloning [182]. These were named *repHI1A*, a relative of which is shared by IncHI2 plasmids, and *repHI1B,* which is unique to IncHI1 plasmids. When an IncHI1 plasmid was introduced into a cell containing an IncF plasmid, the latter plasmid was eliminated [173]. This phenomenon was traced back to the presence of a partial replication region derived from an IncF plasmid, named *repFIA,* which is flanked by 2 IS*1* elements [141, 185]. Two transfer regions, named Tra1 and Tra2, were identified in R27 and both were found to be necessary for transfer, using mutagenesis and complementation studies, sequence analysis and an H-pilus assay [179, 184, 186]. Two partitioning modules, Par1 and Par2, shown to stabilize the plasmid, were found to be located within Tra2 [180].

Table 1.2 Sequenced IncHI1 plasmids

Plasmid	Isolate				Size (kb)	Year sequenced	GenBank Acc. No.	Reference
	Organism	Source	Year	Location				
R27	*S.* Typhi	Human	1961	UK	180	2000	AF250878	[179]
pHCM1	*S.* Typhi	Human	1993	Vietnam	218	2001	AL513383	[187]
pAKU_1	*S.* Paratyphi A	Human	2002	Pakistan	212	2007	AM412236	[188]
pMAK-1	*S.* Dublin	Porcine	?	?	208	2009	AB366440	Akiba
pO111_1	*E. coli*	Human	2001	Japan	205	2009	AP010961	[189]

pHCM1 was recovered from clinical a *S.* Typhi isolate in Vietnam in 1993 and was sequenced in 2001 as part of a *S.* Typhi genome sequencing project ([190]; Table 1.2), then analysed in more detail in 2003 [187]. Although R27 and pHCM1 share large regions of extremely high identity (>99.97% [188]), the composition and position of their antibiotic resistance regions differ (Fig 1.20).

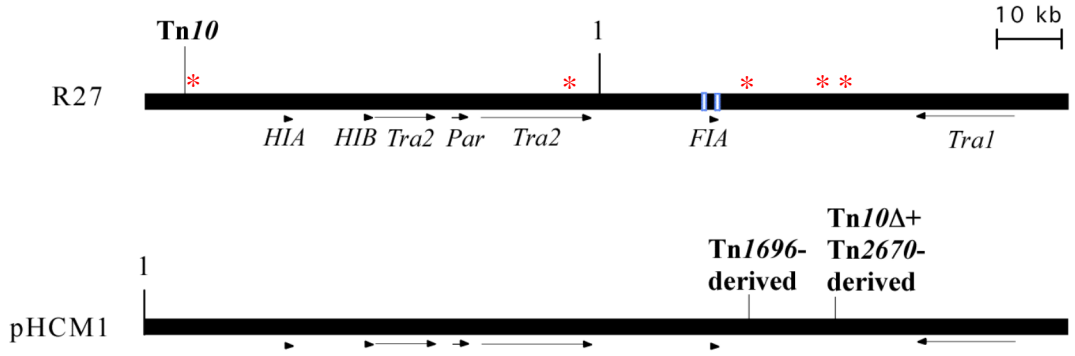


Fig 1.20 Positions of antibiotic resistance regions in R27 and pHCM1. The IncHI1 backbone sequence is represented as a thick black line, starting at position 1 in the pHCM1 sequence. Position 1 in the R27 sequence is shown as "1". The name/derivation of each resistance region is given above the vertical line, which indicates points of insertion. The 3 replication regions are shown as arrows below and named below R27. The red asterisk mark the positions of regions of backbone variation, as described by [188, 163]. Drawn to scale using GenBank Accession Numbers AF250878 and AL513383 for R27 and pHCM1, respectively.

R27 contains only the *tet*(B) tetracycline resistance determinant, located within a complete Tn*10* (see Fig 1.6; Table 1.3), which is flanked by a direct duplication of 9 bp. pHCM1 confers resistance to chloramphenicol, trimethoprim, streptomycin, sulphonamides, and ampicillin (Table 1.3), some of which are used for treatment of typhoid [191]. pHCM1 has 2 antibiotic resistance regions, one consisting of a Tn*2670* relative located next to part of Tn*10*

(Fig 1.21A) and one that is related to Tn*1696* (Fig 1.21B; [31, 67, 187]). A detailed analysis of these regions is shown in Fig 1.21.

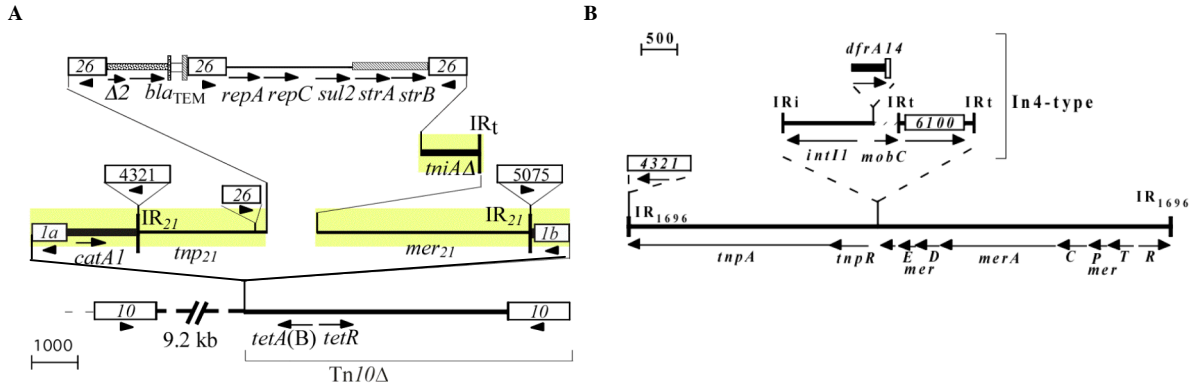


Fig 1.21 MARRs in pHCM1. **A.** Tn*2670*/Tn*10*-derived **B.** T n*1696*-derived. The open, numbered boxes represent insertion sequences of that number, arrows represent direction and extent of genes. Derivation of regions are named and bracketed above or below the genes. Vertical bars are inverted repeats (IR) from either transposons, where the transposon number is subscripted, or from integrons where IRi (inverted initial repeat) denotes the IR from the 5' conserved segment (5'-CS) end of the integron, or IRt (inverted terminal repeat) for the IR from the 3' conserved segment (3'-CS) end. In A, The yellow shading indicates the segments shared by Tn*2670*. The dashed line, which is not drawn to scale, represents a 9.2 kb region that is unique to IncHI1 plasmids and was thought to be part of this resistance region [187], despite not containing resistance genes. In B, the hashed lines represent *mobC*, which is not in Tn*1696* and may be part of a gene cassette. The 3'-CS to the right of IRt is missing compared to Tn*1696*. Figures are drawn to scale based on GenBank Accession Number AL513383. Detailed analysis of pHCM1 was undertaken in this study and by Ruth Hall and Sally Partridge.

The Tn*1696*-derived region, where only a *dfrA14* gene cassette remains, is bounded by 5 bp direct repeats (Fig 1.21B). The second region is made up of part of Tn*10* (7,036 bp of 9,147 bp; Tn*10*Δ in Fig 1.21A) and an IS*1*-bounded transposon, which is derived from Tn*2670* (portion of Tn*2670* remaining in pHCM1 is shaded yellow in Fig 1.21A). Direct repeats were not detected surrounding this MARR, indicating that additional events, such as adjacent deletion, have occurred since its insertion. The Tn*2670*-derived region in pHCM1 harbours a large insertion containing bla_{TEM}, *sul2, strA/B* and 3 IS*26* (Fig 1.21A). This insertion has replaced most of the integron and some of the transposition module of Tn*21* and only IR_{tnp} and bases 1-2072 of $tnpA_{21}$ remain (Fig 1.21). It also contains 3 insertion sequences not present in Tn*2670* (Fig 1.21A). One is an IS*26* in $tnpA_{21}$ that is flanked by 8 bp repeats. The other 2 are IS*4321* and IS*5075* and are found in the IR_{tnp} and IR_{mer}, respectively, of Tn*21* [31].

pAKU_1, was recovered from a *S*. Paratyphi A isolate was sequenced in 2007 ([188]; Table 1.2). pAKU_1 contains a Tn*2670*-derived structure that has undergone inversion events involving IS, in addition to a complete and interrupted Tn*10* ([188]; Table 1.3). However, both of these regions are in different positions to the equivalent structures in either R27 or pHCM1. pAKU_1 also carries the *strA/B*-containing transposon Tn*5393* (Ruth Hall, personal communication; Table 1.3), though this was not identified in the original study.

Table 1.3 Resistance determinants in IncHI1 plasmids

Plasmid	Antibiotic resistance determinants	Transposon/s
pHCM1	*tet*(B), bla_{TEM}, *sul2, strA/B, catA1; dfrA14*	Tn*2670*/Tn*10*Δ; Tn*1696*-like
pAKU_1	*catA1, sul1, dfrA7, strA/B,* bla_{TEM}, *sul2; tet*(B); *strA/B*	Tn*2670*[1]; Tn*10*; Tn*5393*
pMAK-1	*tet*(B), bla_{TEM}, *sul2, strA/B, catA1*	Tn*2670*/Tn*10*Δ[1]
pO111_1	*tet*(B), bla_{TEM}, *sul2, strA/B, catA1*	Tn*2670*/Tn*10*Δ

[1]Tn*2670* had undergone several rearrangements about the IS

Two more recently sequenced IncHI1 plasmids, pO111_1 and pMAK-1 had not previously been analysed in detail. The sequence of pMAK-1 was deposited in GenBank without an associated manuscript. pO111_1 was sequenced as part of a multiple EHEC genome sequencing project and it is described only as R27-like [189]. Their properties are listed in Table 1.2 and Table 1.3.

1.7.1.2 Comparisons of IncHI1 plasmids

The relationships between the sequenced IncHI1 plasmids R27, pHCM1 and pAKU_1, have been studied using a number of techniques: SNP typing, PMLST, and analysing the presence of variable regions and the positions of resistance regions in the IncHI1 backbone. Long-range PCRs and microarray techniques were used to analyse a larger set of IncHI1 plasmids.

When a study conducted by Holt and others in 2007 compared the 164.4 kb shared backbones of R27, pHCM1 and pAKU_1 using SNP-typing, they found that pAKU_1 was 99.89% identical to R27, but only 99.71% identical to pHCM1 [188]. That is, there was at least 3 fold more SNPs unique to pHCM1 than either of the other plasmids. Thus, pAKU_1 and R27

appeared to be more related to each other than to pHCM1. However, they also noted that the resistance regions of the 3 plasmids were in different positions, indicating separate lineages.

In 2009, Phan and others examined a collection of 36 IncHI1 plasmids and created a PMLST typing scheme for IncHI1 plasmids, based on SNPs of 6 loci, with 2-3 alleles per locus [163]. Using the full set of 36 plasmids, 8 STs were identified. A single nucleotide difference defined different alleles for 4 loci and 2 SNPs differentiated alleles for 2 loci. STs that share 5/6 alleles were allocated to 2 overarching groups: "Group 1" which contained pHCM1 and 10 other plasmids tested and "Group 2", which contained pAKU_1 and the 23 remaining plasmids. However, R27 was not grouped with any other plasmid and was classed as a singlet. Despite this, pAKU_1 and R27 shared 4/6 alleles, pHCM1 and R27 2/6, and pHCM1 and pAKU_1 1/6, indicating that R27 and pAKU_1 are more related to each other than to pHCM1. Generally, traditional MLST is based on several differences at each locus, because single nucleotide changes, such as from sequencing errors, can affect sequence types [192]. Thus, the limited number of SNPs at each locus may affect the overall accuracy of this PMLST scheme.

These two previous studies had also collectively identified 5 potentially informative segments of variation in the IncHI1 backbone (positions of which are marked with red asterisk in Fig 1.20), the presence or absence of which was used to shed light onto the evolutionary relationships of these plasmids. When pAKU_1 was compared to both pHCM1 and R27, 2 major sources of variation, Regions A/B (later deduced to be 1 variation) and C, were detected [188]. Using these, pAKU_1 and R27 had identical profiles, illustrating a close relationship, whereas pHCM1 did not share either of these regions. When pHCM1 was compared to R27, 5 regions named Regions A, B, C, D and E (where Regions C and D corresponded to Regions A/B and C, respectively, from Holt et al., (2007) [188]) were identified [163]. Regions A and B included part of the Tn*2670*/Tn*10*- and Tn*1696*-derived

resistance regions, respectively, in addition to some backbone sequence. By analysing Regions A-E, the authors found that R27, the singleton from PMLST data, was more related to Group 2 plasmids (like pAKU_1), than to Group 1 plasmids (like pHCM1).

Microarray analysis was used to detect the presence of all genes in pHCM1, as well as focussing on the presence/absence of the above-mentioned variable Regions A-E, within a subset of 18 IncHI1 plasmids from a range of sources [163]. Outside Regions A-E, not much variation was observed. The presence of Regions C and E did largely reflect the PMLST grouping. However, the remaining regions did not yield cohesive data that could be used to group the plasmids, perhaps due to the limitations of the approach. Microarray cannot detect the context of the genes and this can be problematic, particularly when analysing the resistance regions. For example, the probe for variable Region A, which corresponded to the Tn*2670*- and Tn*10*-derived MARR in pHCM1 (Fig 1.21A), yielded a partial signal for R27, even though R27 does not contain a Tn*2670*-like element, because it contains Tn*10* located elsewhere (Fig 1.6). This study also examined this subset of plasmids using long-range PCR, consisting of 20 PCR reactions that yielded amplicons of ~10 kb, the primers for which were designed off the pHCM1 sequence [163]. However, this technique did not yield any clear information regarding plasmid relatedness and seemed to be effective in mapping only half of the plasmid backbone, across the vast majority of plasmids tested. This may be due to rearrangements in pHCM1 or poor primer placement, like those in regions unique to pHCM1. Although some of these methods produced inconclusive results, both major studies concluded that R27 and pAKU_1 are more related to each other, than either is to pHCM1.

1.7.2 IncHI2 plasmids

IncHI2 plasmids encode resistance to multiple antibiotics and to heavy metals. IncHI2 plasmids have been reported to confer a MAR phenotype of *Salmonella* hospital outbreaks

endemic to particular regions, in Spain in 1996-1999 [193] and again in 2005-2007 [194, 195] and in Chile in 1977-1993 [196]. They have also been associated with providing resistance in food-producing animal isolates, for example, from chickens in 2001-2007 in Belgium [197] and from poultry from various sources [161]. IncHI2 plasmids have been identified as the 5th most common plasmid group in Enterobacteriaceae [160]. They appear to be fairly promiscuous as they have been isolated from a range of Gram-negative bacteria [142] and are capable of transfer to a number of hosts, including *E. coli*, *Vibrio cholerae*, Salmonella spp., *Aeromonas salmonicida, Vibrio anguillarum*, and *Yersinia ruckeri* [177].

1.7.2.1 R478

The most well studied IncHI2 plasmid is R478, a reference plasmid that was recovered from a *Serratia marcescens* isolated from USA in 1969 [198]. R478, like other IncHI2 plasmids associated with *S. marcescens*, was originally classified as belonging to incompatibility group IncS [199], which were later renamed IncHI2s. Sub-cloning fragments of R478 [166] revealed that it contains 2 replication regions, one related to *repHI1A* from IncHI1 plasmids, named *repHI1A*$_{(R478)}$ and the unique *repHI2A* (Fig 1.22). Using mutagenesis studies, R478 was shown to contain two large (17.3 kb and 36.3 kb in R478), transfer regions, Tra1 and Tra2, both essential for conjugation ([200]; Fig 1.22). A partitioning module was found to be located within Tra2. The complete sequence of R478 was reported in 2004 (Genbank accession no. BX664015) and revealed that it was larger than the sequenced IncHI1 plasmids (Table 1.2), at 275 kb [58].

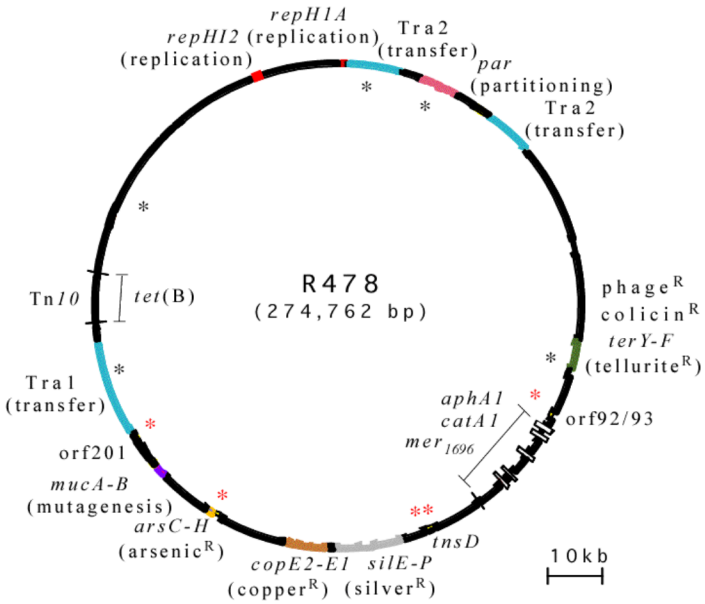


Fig 1.22 Map of the IncHI2 plasmid R478. The circle represents the sequence of R478 (GenBank accession number BX664015). Important features are as follows: 2 coding regions for replication - $repA_{HI2}$ and $repA_{H1A}$ (red); 2 transfer regions Tra1, 2 (blue); partitioning region (pink); mutagenesis (purple); regions that give resistance to the heavy metals tellurite (green), mercury (orange), silver (silver), copper (copper), and arsenic (yellow). 2 antibiotic resistance regions - one adjoining mer_{1696} containing *catA1* (chloramphenicol resistance) and *aphA1* (kanamycin and neomycin resistance) and Tn*10* containing *tetA*(B) (tetracycline resistance) are indicated by bars inside the circle. Insertion sequences IS*26* and IS*10* from the former and latter resistance regions are represented by open boxes. The asterisk represent the positions of PCR primers previously used to examine the IncHI2 backbone and those in red are the ones that vary [161].

It carried regions conferring resistance to the heavy metals tellurite, copper, arsenate, arsenite and silver [201-203]. The heavy metal resistance determinants are included in an ~80 kb segment in between Tra1 and Tra2, which makes up around a 3rd of R478 (Fig 1.22). Because of the elevated GC content, the arsenite/arsenate resistance module was predicted to have been acquired [202]. Other features of R478 include resistance to colicins and phage, *muc* genes for mutagenesis and *tnsA-D* genes, the protein products from which share 21-39% aa sequence identity to the equivalent products from Tn*7* [58].

1.7.2.1.1 Antibiotic resistance regions in R478

Two antibiotic resistance regions, that collectively confer resistance to chloramphenicol, kanamycin, neomycin and tetracycline, have been identified in R478 [58]. The first is the transposon Tn*10* which contains the *tet*(B) tetracycline resistance determinant (see Fig 1.6).

The second is a mosaic MARR, which consists of an IS*26*-bounded region containing the *aphA1* and *catA1* resistance genes (marked red in Fig 1.23). It is next to the mercuric ion resistance (*mer*) module of Tn*1696* (Ruth Hall, unpublished observations), which was incorrectly annotated as Tn*21* [58]. R478 lacks the transposition (*tnp*) module and the majority of the integron of Tn*1696*. Only a part of the Tn*1696*-derived integron remains (boxed blue in Fig 1.23). However, another integron remnant was identified nearby (boxed green in Fig 1.23).

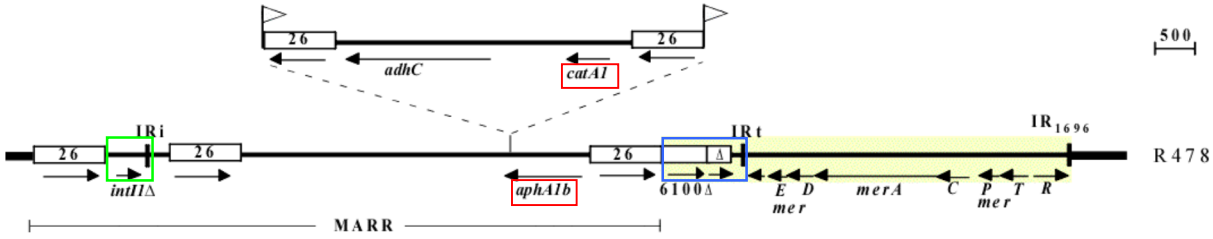


Fig 1.23 MARR region of R478. The segment derived from Tn*1696* is highlighted in yellow. Names of antibiotic resistance genes are boxed in red. Part of the integron derived from Tn*1696* that remains is boxed in blue and another integron remnant is boxed in green. Other features are as in Fig 1.18. Figure drawn to scale using GenBank accession no. BX664015. Detailed analysis of this structure from this study, and from Ruth Hall and Virginia Post.

1.7.2.2 Other IncHI2 plasmids

Recently, attention has been turned to IncHI2 plasmids, as they are increasingly found to be associated with isolates that produce extended spectrum β-lactamases (ESBLs), metalo-β-lactamases and are resistant to 3rd generation cephalosporins [20]. IncHI2 plasmids have been found to contain a range of clinically relevant genes. These include, $bla_{\text{CTX-M9}}$ [193];[194];[161];[204];[205], $bla_{\text{CTX-M2}}$, [161];[206] and $bla_{\text{CTX-M-14}}$ [207]; bla_{CMY} and $bla_{\text{CTX-M3}}$[208]; *qnrB2*, $bla_{\text{IMP-8}}$ and $bla_{\text{SHV-12}}$ [209]; $bla_{\text{IMP-4}}$ [210] and bla_{VIM} [195].

There are currently 3 sequenced IncHI2 plasmids that are closely related to R478, pK29 [208], pEC-IMP and pEC-IMPQ [209], all of which encode ESBLs. Their properties are listed in Table 1.4. One further plasmid, pAPEC-O1-R [211], closely related to these IncHI2s, was

sequenced and found to contain an IS*26*-bounded antibiotic resistance region derived from Tn*1696* and Tn*1404* [212]. It is discussed separately in section 1.7.2.3.

Table 1.4 Sequenced IncHI2 plasmids

Plasmid	Isolate				Size (kb)	GenBank Acc. No.	Reference
	Organism	Source	Year	Location			
R478	*S. marcescens*	Human	1969	USA	275	BX664015	[58]
pK29	*K. pneumoniae*	Human	2001	Taiwan	269	EF382672	[208]
pEC-IMP	*E. cloacae*	Human	2004	Taiwan	319	EU855787	[209]
pEC-IMPQ	*E. cloacae*	Human	2004	Taiwan	325	EU855788	[209]
pAPEC-O1-R	*E. coli*	Avian	2007	USA	241	DQ517526	[211]

pK29 was recovered from a *Klebsiella pneumoniae* isolate in Taiwan in 2001 (Table 1.4). It is 269 kb (GenBank accession no. EF382672; [208] and sequence comparisons with R478, showed that they shared much of their backbone and the major regions that varied were antibiotic resistance regions (boxed red in Fig 1.24). Also, a number of additional single IS can be seen in pK29, such as IS*903*, IS*30*, IS*5*, IS*1* and IS*630*. pK29 was found to contain the bla_{CMY8} and $bla_{CTX\text{-}M3}$ in 2 separate regions, "Region 1" and "Region 2" (Fig 1.24). pK29 did not contain Tn*10* or the Tn*1696*-like MARR of R478.

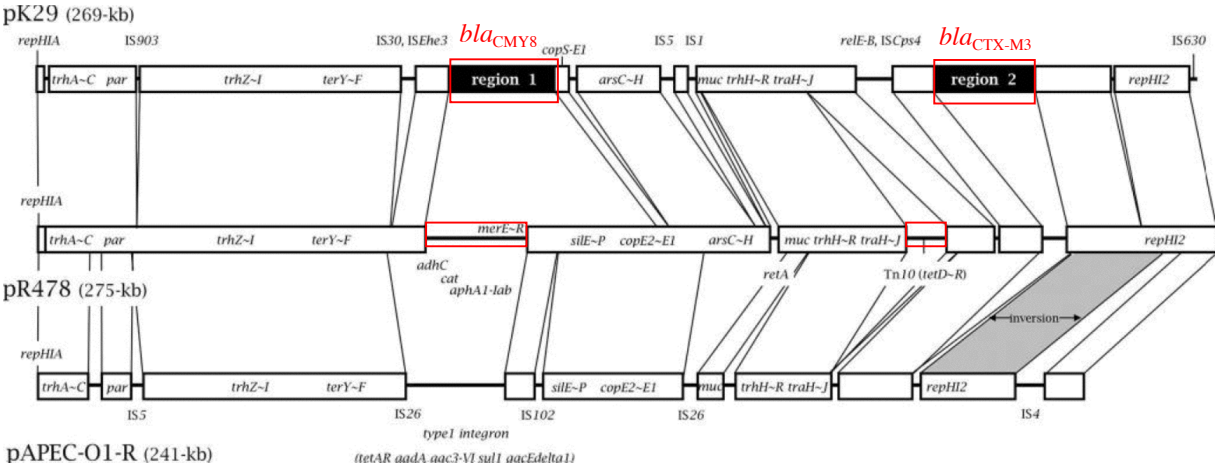


Fig 1.24 Map of pK29 and comparison of pK29 with R478 and pAPEC-O1-R. The lines between the three plasmids indicate the homologous sequence blocks which are >99% identical (white squares). Inversion between the homologous regions is shaded. The two antibiotic resistance regions unique to pK29 (regions 1 and 2) are shown in black and boxed in red. The relevant resistance genes are written above in red. Modified from [209].

Recently, 2 more IncHI2 plasmids, isolated in Taiwan in 2004 from *Enterobacter cloacae*, were sequenced, named pEC-IMP and pEC-IMPQ (Table 1.4; [209, 213]). These plasmids

carry bla_{IMP-8} and bla_{SHV-12} and are extremely closely related, differing only by an additional region in pEC-IMPQ, which carries the *qnrB* resistance gene (here named pEC-IMP/Q when referring to both plasmids). At 324 kb, pEC-IMPQ is the largest IncHI2 plasmid sequenced to date (Table 1.4).

1.7.2.3 pAPEC-O1-R

In 2006, a plasmid called pAPEC-O1-R was sequenced and analysed (GenBank accession no. DQ517526; [211]; Table 1.4). pAPEC-O1-R displays >94% identity to R478 and the other sequenced IncHI2 plasmids across the IncHI2-specific PBRT replicon target. However, whether this plasmid is compatible with known IncHI2 plasmids has not been determined. For simplicity, I have designated the R478-type IncHI2 plasmids "IncHI2a" and the pAPEC-O1-R-type "IncHI2b". pAPEC-O1-R shares a proportion of the IncHI2 backbone with IncHI2a plasmids, including the transfer and replication regions (Fig 1.24). The shared backbone segments of the 4 sequenced IncHI2a are >99.91% identical to R478, compared to >97.04% for the IncHI2b plasmid pAPEC-O1-R [160].

Previous studies have distinguished the IncHI2a and IncHI2b lineages using a plasmid double locus sequence typing (PDLST) scheme and a multiplex PCR. PDLST gives 2 sequence types (ST1 for R478-type and ST2 for pAPEC-O1-R-type plasmids) using two conserved loci, named smr0018 and smr0199 in R478, near the 2 transfer regions [160]. However, this differentiation could be more simply achieved using PCR with the IncHI2 PBRT primers, as there is only a 619/647 bp match between IncHI2a and IncHI2b plasmids within the PCR amplicon obtained. Sequencing this PCR product, or by digesting the amplicon with a restriction enzyme (for example with PstI, which only has a site in the IncHI2b sequence) should readily distinguish IncHI2a and IncHI2b plasmid subtypes. To date, no work has described the evolutionary origins of IncHI2a plasmids. This thesis focuses on evolution within IncHI2a plasmids and IncHI2b plasmids will be discussed only peripherally.

1.7.2.4 Variation within IncHI2a plasmids

Previously, 2 complementary studies [161];[193] had developed a set of 10 PCRs (the positions of which are marked with asterisk in Fig 1.22) in order to map and compare sets of IncHI2 plasmids. In the more comprehensive study [161], 11 IncHI2 plasmids from a broad range of sources and isolation dates were examined experimentally. Variation was observed in only 5 of the 10 PCRs (marked with red asterisk in Fig 1.22), positioned in or near heavy metal resistance regions, indicating that this region in IncHI2 plasmids varies.

1.8 Antibiotic resistance in *Salmonella enterica*

Salmonellae are typical *Enterobacteriacae* and non-spore forming Gram-negative rods [214]. *Salmonella enterica* is the most clinically relevant species of *Salmonella* as it accounts for >99% of *Salmonella* isolates involved in disease in mammals [215]. *S. enterica* is divided into serotypes or serovars, based on the agglutination properties of different types of antigens: the somatic O, flagellar H and capsular Vi antigen [216]. These serovars can be further broken down into phage types, based on their susceptibilities to lytic bacteriophage [217]. Specific strains are identified using XbaI chromosomal digestion or IS*200* profiles [218];[219]. Different *S. enterica* serovars cause a number of diseases and display a varied host range. For example, the human-restricted pathogen *S.* Typhi causes the potentially fatal disease typhoid fever. Globally, it was estimated that *S.* Typhi caused 21.6 million typhoid cases resulting in 216,000 deaths, in 2000 [220]. Other serovars, such as *S.* Typhimurium can cause gastroenteritis in both animals and humans [221]. Approximately 95% of salmonellosis cases in humans are caused by ingestion of contaminated food products [222]. The definitive type strain of *S.* Typhimurium, named DT104 is resistant to the antibiotics Ap, Cm, Sm, Sp, Su and Tc [223], attributed to the presence of SGI1, a 43 kb genomic island that contains a class 1 integron and 5 antibiotic resistance genes [224]. DT104 has been isolated from humans and meat-producing animals and in many countries, although not including Australia [221, 225].

1.8.1 Australian S. enterica

In Australia, all *Salmonella* infections are notifiable, which means each case of salmonellosis or typhoid must be reported and a sample collected as part of the National Notifiable Diseases Surveillance System [226]. In 2009, the number of salmonellosis cases reported in Australia was 9,533, and there were 116 reported cases of typhoid. Some of this national sample collection is held at the Microbiological Diagnostic Unit (MDU) in Melbourne, where the antigenic serovar, phage type and antibiotic resistance phenotype are determined. Circa 2000, Dr. Djordjevic at the Elizabeth Macarthur Agricultural Institute initiated a project aiming to characterise a multiply antibiotic resistant subset of the MDU collection, because little information existed on the types and organization of resistance genes in *S. enterica* in Australia. Molecular analysis of these strains was undertaken in collaboration with the Hall Laboratory at the University of Sydney. One arm of this work detected SGI1 and variants of it, in different *S. enterica* serovars [112, 227, 228]. Another focussed on screening for antibiotic resistance gene cassettes in class 1 integrons, which resulted in a number of novel gene cassettes and cassette arrays being identified [113, 229-231]. One project screened for antibiotic resistance gene clusters within a set of *S.* Typhimurium isolates from animal and human origin [232]. During this work, plasmids were detected in some of the *S.* Typhimurium isolates. Analysis was undertaken of the antibiotic resistance gene clusters on IncQ plasmids from this set [114], but little other information on these plasmids was produced. A handful of isolates were found to contain IncHI1 and IncHI2 plasmids (Hall laboratory, unpublished observations). These plasmids were subjected to further analysis, which aimed to provide molecular information on IncHI MAR plasmids circulating in Australian *S. enterica*, and formed the basis of this study.

1.9 Scope and aims of this thesis

Due to the presence of IncHI plasmids in a number of multiply antibiotic resistant Gram-negative bacterial infections, they clearly play a central role in the dissemination of resistance genes among Enterobacteriaceae, particularly *Salmonella*. This study was directed at providing detailed molecular information on the IncHI resistance plasmids circulating in Australian *S. enterica* animal isolates. This entailed determining the structure and derivation of antibiotic and heavy metal resistance genes and gene clusters on these plasmids. By comparing these resistance regions to those on sequenced IncHI plasmids from a range of origins, insight into the evolution of these regions was gained. Also, the forces that effect change, for example MGEs that add or remove resistance genes, could be identified. Features that can distinguish evolutionary lineages of IncHI plasmids, such as the location of resistance region/s or differences with the plasmid backbones, were identified in order to determine if plasmid lineages are globally disseminated, if they cross animal/human boundaries, and if they persist over time. The stability of a single, large resistance region in the absence of antibiotic selection was examined and the insertion sequence/s responsible for facilitating the loss of resistance gene/s identified.

Hence, the aims of this study are as follows:

- To elucidate the composition, origin and evolution of antibiotic and heavy metal resistance regions on IncHI plasmids.
- To identify evolutionary lineages within the IncHI1 and IncHI2 plasmid subgroups using simple screening tools, and to determine whether these plasmids lineages are widely distributed.
- To investigate the role that MGEs, such as IS and Tns, play in the evolution of these plasmids and in the stability of their resistance regions.

CHAPTER TWO

Materials and methods

2.1 Materials

2.1.1 Chemicals and reagents

Basic chemicals were purchased from Sigma-Aldrich (St Louis, MO, USA), Ajax Chemicals (Auburn, NSW, Australia), Bio-Rad (Hercules, CA, USA) or Promega (Madison, MI, USA) and were of molecular biology grade.

2.1.2 Buffers and solutions

Buffers and solutions were made up using H_2O treated by reverse osmosis, according to protocols in Sambrook et al. (1989) [233].

0.5 X TBE buffer	45 mM Tris 45 mM boric acid 1 mM EDTA
TE buffer	10 mM Tris-HCl, pH 7.5 1 mM EDTA, pH 8
EB buffer	10 mM Tris-HCl, pH 8
Loading dye	0.1% (w/v) bromophenol blue 0.2% (w/v) xylene cyanol 20 mg/ml ficoll
Saline	0.9% (w/v) NaCl

2.1.3 Bacterial growth media

Bacto™ Yeast extract, Bacto™ tryptone and Bacto™ agar were purchased from Becton Dickinson (Cockeysville, MD, USA). Sensitest Agar, Mueller Hinton Agar, MacConkey Agar, Blood Agar base and defibrinated horse blood were obtained from Oxoid (Basingstoke, Hampshire, UK). Bacterial growth media was prepared according to the instructions supplied by the manufacturer. Horse blood agar was prepared by suspending 40 g Blood Agar base in 1 L of

distilled water, autoclaving to sterilise and dissolve and upon cooling to 50^{O}C, 7% sterile horse blood was added.

Luria-Bertani Broth (LB)	1% (w/v) tryptone 0.5% (w/v) yeast extract 1% (w/v) NaCl
Luria-Bertani Agar (LA)	LB containing 1.5% (w/v) agar
Mueller-Hinton Agar (MHA)	3.8% (w/v) Mueller-Hinton Agar in H_2O
Horse Blood Agar (HBA)	4% (w/v) Blood Base Agar in H_2O 7% (w/v) defibrinated horse blood
MacConkey Agar (MAC)	5.2% (w/v) MacConkey Agar in H_2O

2.1.4 Sterilisation

All appropriate media, buffers and solutions were autoclaved at 121^{o}C at 103.4 kPa for 20 minutes. Heavy metal stock solutions were sterilised using a 0.2 micron filter. Antibiotics were prepared in sterile tubes and sterile solvents and were not further sterilised.

2.1.5 Enzymes

Enzymes including *Taq* Polymerase, Phusion^{TM} Polymerase, T4 DNA Quick Ligase, T4 DNA Polymerase, Antarctic Phosphatase, all restriction endonucleases and their appropriate buffers were obtained from New England Biolabs (Ipswich, USA). RNase A was purchased from Sigma-Aldrich (St Louis, MO, USA) and was heat treated at 65^{O}C for 20 mins before use.

2.1.6 Antibiotics and heavy metals

All antibiotics and heavy metals were obtained from Sigma-Aldrich (St Louis, MO, USA). Antibiotic stock solutions were made up under the specifications given in Table 2.1 and stored at $4^{O}C$.

Table 2.1 Stock antibiotic solutions

Antibiotic	Abbreviation	Stock Concentration (mg/ml)	Diluent
Ampicillin	Ap	100	MQ H_2O
Chloramphenicol	Cm	25	100% Ethanol
Gentamicin	Gm	8	MQ H_2O
Kanamycin	Km	50	MQ H_2O
Nalidixic acid	Nx	25	0.5 M NaOH
Neomycin	Nm	25	MQ H_2O
Rifampicin[1]	Rf	10	100% Methanol
Spectinomycin	Sp	25	MQ H_2O
Streptomycin	Sm	25	MQ H_2O
Sulfamethoxazole	Su	100	MQ H_2O and NaOH drops[2]
Tetracycline[1]	Tc	10	50% Ethanol
Trimethoprim	Tp	2.5	100% Methanol

[1]Light sensitive; stored in light-proof container to prevent degradation.
[2]NaOH drops added until dissolved. Heat gently applied whilst adding NaOH.

Heavy metal stocks solutions were made up in MQ H_2O to the following concentrations: 1 M sodium arsenite ($NaAsO_2$), 1 M sodium arsenate ($Na_2HAsO_4.7H_20$), 5% (w/v) potassium tellurite (K_2TeO_3) and 1% (w/v) mercuric chloride ($HgCl_2$), then stored at room temperature.

2.2 Bacterial strains and plasmids

2.2.1 Bacterial strains

Experimental work was largely performed on a set of *Salmonella enterica* strains, isolated in Australia from a number of animal sources. These strains, which are listed in Table 2.2, were kindly supplied by Dr. Diane Lightfoot (Microbiological Diagnostic Unit (MDU), Melbourne, Australia). At the MDU, these strains were partially characterised (Table 2.2). Bacterial strains were prepared for long-term storage at $-80^{O}C$ by mixing 500 µl of fresh overnight bacterial culture with an equal volume of 20% glycerol and then were snap frozen with liquid nitrogen

before storage at –80°C. *Escherichia coli* strains used as host strains for plasmids and negative control strains for resistance phenotypes are listed in Table 2.3. *E. coli* strains, which contain *Salmonella* plasmid/s, generated during this study are given in Table 2.4.

Table 2.2 *Salmonella* strains used in this study

Strain	Serovar/ Phage Type[1]	Source	Year	Resistance profile[2]
SRC26	Typhimurium/PT1 var 2	Bovine	1999	Ap Km Nm Sm Su Tc Tp
SRC27	Typhimurium/PT135	Equine	1999	Ap Cm Gm Km Nm Sm Sp Su Tc Tp
SRC70	Infantis	Feline	2001	Sm Sp Su Tc
SRC71	Infantis	Chicken	2001	Sm Sp Su Tc
SRC72	Infantis	Chicken	2001	Sm Sp Su Tc
SRC83	Infantis	Chicken	2000	Sm Sp Su Tc Tp
SRC92	Infantis	Canine	2000	Sm Sp Su Tc Tp
SRC93	Infantis	Chicken	2000	Sp Su Tc
SRC94	Infantis	Chicken	2000	Sm Sp Su Tc
SRC95	Infantis	Chicken mince	2000	Sm Sp Su Tc
SRC96	Infantis	Chicken carcass	2000	Sm Sp Su Tc
SRC125	Typhimurium/PT44	Bovine	2000	Km Nm Su Tc Tp

[1]Where available
[2]Ap=Ampicillin, Cm=Chloramphenicol, Gm=Gentamicin, Km=Kanamycin, Nm=Neomycin, Sm=Streptomycin, Sp=Spectinomycin, Su=Sulfamethoxazole, Tb=Tobramycin, Tc=Tetracycline, Tp=Trimethoprim

Table 2.3 Host *E. coli* strains

Strain	Antibiotic selection	Genotype	Reference
DH5α	Nx	*supE44 DlacU169 (ϕ80lacZDM15) hsdR17 recA1[1] endA1 gyrA96 thi-1 relA1*	[234]
K12 strain 294 (E294)	Rf	*endo I-, B1-, rK-, mK+*	[235]

[1]*recA* mutants lack the ability to perform *recA*-mediated recombination

Table 2.4 *E. coli* strains generated in this study.

Strain	Description	Construction method
E294/pSRC27-H/pSRC27-I	E294 containing pSRC27-H and pSRC27-I	Conjugation at 37°C
E294/pSRC27-H	E294 containing pSRC27-H	Conjugation at 26°C
E294/pSRC27-I	E294 containing pSRC27-I	Conjugation at 37°C

2.2.2 Plasmids

Plasmids used in this study are listed in Table 2.5.

Table 2.5 Plasmids used in this study

Name	Resistance Phenotype[1]	Description	Reference
Control plasmids			
pUC19	Ap	Cloning vector	[236]
R388 ::Tn*1403*	Ap Cm Sm Sp Su Tp Hg	R388 with Tn*1403*	[73]
R1033	Ap Cm Gm Km Sp Su Tc Hg	IncP plasmid that contains Tn*1696*	[86]
pACYC184 ::Tn*21*	Cm Sp Sm Su Tc	pACYC184 with Tn*21*	[237]
R478	Cm Km Nm Tc Hg	Sequenced, conjugative IncHI2 plasmid	[58]
R27	Tc	Sequenced, conjugative IncHI1 plasmid	[179]
TP116	Cm Sm Su Hg	Reference, conjugative IncHI2 plasmid	[139]
Plasmids isolated from S. *enterica* isolates			
pSRC26	Ap Km Nm Sm Su Tc Tp Hg	IncHI2 plasmid from SRC26	This study
pSRC27-H	Ap Cm Gm Km Nm Sm Sp Su Tc Tp	IncHI1 plasmid from SRC27	This study
pSRC27-I	Sm	IncI1 plasmid from SRC27	This study
pSRC70	Sm Sp Su Tc Hg	IncHI2 plasmid from SRC70	[230]
pSRC71	Sm Sp Su Tc Hg	IncHI2 plasmid from SRC71	[230]
pSRC72	Sm Sp Su Tc Hg	IncHI2 plasmid from SRC72	[230]
pSRC83	Sm Sp Su Tc Tp Hg	IncHI2 plasmid from SRC83	This study
pSRC92	Sm Sp Su Tc Tp Hg	IncHI2 plasmid from SRC92	[230]
pSRC93	Sp Su Tc Hg	IncHI2 plasmid from SRC93	[230]
pSRC94	Sm Sp Su Tc Hg	IncHI2 plasmid from SRC94	[230]
pSRC95	Sm Sp Su Tc Hg	IncHI2 plasmid from SRC95	[230]
pSRC96	Sm Sp Su Tc Hg	IncHI2 plasmid from SRC96	[230]
pSRC125	Km Nm Su Tc Tp Hg	IncHI2 plasmid from SRC125	This study
Plasmids generated in this study			
pRMH951	Ap Tp	12 kb PstI clone of pSRC26 in pUC19	This study
pRMH952	Ap Tp	14.5 kb SacI clone of pSRC26 in pUC19	This study
pRMH953	Ap Km Nm	5.8 kb SacI clone of pSRC125 in pUC19	This study
pRMH955	Ap Gm Sm Sp Tc Tp	20 kb SacI clone of pSRC27-H in pUC19	This study
pRMH958	Ap Km Nm Sm	8 kb SacI clone of pSRC27-H in pUC19	This study
pRMH959	Ap Cm Su	13 kb SacI clone of pSRC27-H in pUC19	This study
pRMH960	Ap	4.2 kb BamHI clone of pSRC125 in pUC19	This study
pRMH961	ApHg	5 kb BamHI clone of pSRC125 in pUC19	This study

[1]Ap=Ampicillin, Cm=Chloramphenicol, Gm=Gentamicin, Km=Kanamycin, Nm=Neomycin, Sm=Streptomycin, Sp=Spectinomycin, Su=Sulfamethoxazole, Tb=Tobramycin, Tc=Tetracycline, Tp=Trimethoprim, Hg= Mercuric ions

[2]Royal North Shore Hospital, Sydney, Australia.

2.3 Bacterial characterisation and manipulation

2.3.1 Bacterial growth conditions

Bacterial cultures were routinely set up with 2 ml LB, which was inoculated with a single bacterial colony, and incubated overnight at 37ºC with shaking at 200 rpm. LB was used for all broth cultures containing antibiotics. Mueller-Hinton Agar (MHA) (Oxoid, Basingstoke, UK) was used for plates containing the antibiotics Su and Tp, while LA was used for all other antibiotics. Antibiotics, if any, were added after cooling to 60ºC at the final concentrations provided in Table 2.6.

Table 2.6 Working concentrations of antibiotics and media used

Antibiotic	Working Concentration (µg/mL)	Media used
Ampicillin	100	LA
Chloramphenicol	25	LA
Gentamicin	8	LA
Kanamycin	20	LA
Nalidixic acid	25	LA
Neomycin	50	LA
Rifampicin	100	LA
Spectinomycin	25	LA
Streptomycin	25	LA
Sulphamethoxazole	100	MHA
Tetracycline	10	LA
Trimethoprim	25	MHA

2.3.2 Antibiotic resistance phenotypic determination

2.3.2.1 Picking and patching

To determine which antibiotics an isolate was resistant to, single colonies from dilution or streak plates were patched, using a toothpick, onto different antibiotic plates (some or all of those in Table 2.6) and lastly an LA or MacConkey agar plate, to ensure the colony transferred to all plates in the series. After overnight growth at 37ºC each colony was qualitatively scored for antibiotic resistance and sensitivity in relation to suitable positive and negative controls.

2.3.2.2 Disc diffusion

The Calibrated Dichotomous Sensitivity (CDS) disc diffusion assay was used to quantify the level of resistance an isolate has to a set of antibiotics. A single bacterial colony was resuspended in 4 ml of saline which was used to flood a sensitest agar (Oxoid, Basingstoke, UK) plate to form an even lawn of growth and the excess fluid was collected. This bacterial suspension could be used on multiple plates if desired. Six antibiotic discs were stamped onto each plate, which was incubated for 16 hours at 37^{O}C. The antibiotic discs contained amikacin (30 μg), ampicillin (25 μg), cefotaxime (30 μg), ceftazidime (30 μg), chloramphenicol (30 μg), florfenicol (30 μg), gentamicin (10 μg), kanamycin (50 μg), neomycin (30 μg), streptomycin (25 μg), spectinomycin (25 μg), sulfafurazole (300 μg), tetracycline (30 μg), tobramycin (10 μg) or trimethoprim (5 μg) (Oxoid, New Hampshire, England). The annular radius of the inhibition zone was measured from the edge of each antibiotic disc. A radius of 6 mm or greater was defined as sensitive to that antibiotic, and a radius of less than 6 mm as resistant, as per CDS guidelines (http://web.med.unsw.edu.au/cdstest/).

2.3.3 Heavy metal resistance determination

Some strains were also tested for susceptibility to the heavy metals arsenic, mercury and tellurite. Strains were patched and scored for growth on LA containing the heavy metals sodium arsenate (20 mM), sodium arsenite (5 mM) or mercuric chloride (20 μg/ml) and horse blood agar containing potassium tellurite (100 μg/ml).

2.3.4 Conjugation

Plasmids were conjugated from a donor strain into the *E. coli* recipient strain E294 (see Table 2.3) by mixing 100 μl of overnight LB liquid culture of donor with an equal amount of recipient on an LA plate and incubating for 6 – 24 hours at 37°C. Cells were harvested in 1 ml saline and diluted using a 1 in 10 dilution series, spanning the dilutions 10^{-1} to 10^{-7}. 100 μl of each dilution

was plated onto agar plates containing either rifampicin or nalidixic acid (depending on the recipient) and an antibiotic that the target plasmid confers resistance to, so as to enumerate and isolate transconjugants. To enumerate the donor cells, 100 µl of the more dilute samples (dilutions 10^{-5} to 10^{-7}) was spread onto plates that contain antibiotics that select for only donor cells. The transfer frequency was then calculated as the number of transconjugants per donor. At least 10 transconjugant colonies from each mating were screened for resistance by picking and patching onto plates containing some or all of the antibiotics in Table 2.6. Transconjugants of *Salmonella* and *E. coli* matings were also patched onto MacConkey agar, to confirm the transconjugants were *E. coli* and not antibiotic resistant mutants of *Salmonella*. MacConkey is a differential media, where *Salmonella* colonies (that can not metabolise lactose) are white, and *E. coli* (which break down lactose into lactic acid, causing a lowering of pH and thus a colour change) are pink. RAPD PCR (see 2.4.3.5) was performed on whole cell DNA extractions (see 2.4.1.2) of transconjugants from *E. coli* and *E. coli* matings to differentiate the host and recipient *E. coli* strains.

2.3.5 Extended growth experiments

To investigate the stability of antibiotic resistance/s conferred by plasmid/s, bacterial cultures containing the plasmid/s were continued for more than one night, without antibiotic selection, using serial dilution. A 2 ml starter culture was inoculated with a single colony and grown overnight with appropriate antibiotics to ensure the presence of the plasmid/s to be studied. The starter culture was diluted in saline and 0.1 ml of a 10^{-5} dilution (~1000 cells) was added to 2 ml of fresh LB without antibiotics and incubated overnight with shaking at 37^{O}C. The resulting culture constituted the "Day 1" culture, which was diluted and ~1000 cells were added to fresh LB and grown overnight to yield the "Day 2" culture and so on. At certain intervals, cultures were diluted (dilutions used were approximated to give between 30 and 300 colonies per plate)

and spread in duplicate onto plates with and without antibiotics to determine the proportion of antibiotic resistant cells (if the loss was greater than 10 fold). Plates were incubated at 37ºC overnight and counted the next day. Also, for more subtle detection of antibiotic resistance loss, individual colonies were taken off dilution plates without antibiotics and picked and patched onto plates containing single antibiotics to determine their antibiotic resistance phenotype. Serially diluted cultures with antibiotics were run in parallel and tested in the same way as above, to act as a positive control.

2.4 DNA preparation and manipulation

2.4.1 DNA extraction

2.4.1.1 Plasmid DNA

Plasmid DNA was extracted using an alkaline lysis method [233]. 1.5 ml of overnight culture was pelleted by centrifugation at 12,000 rpm for 30 secs. The pellet was resuspended in 100 µl of ice-cold buffer (Solution 1: 50 mM glucose, 25 mM Tris-HCl and 10 mM EDTA, pH 8) and 200 µl of freshly prepared lysis solution (Solution 2: 0.2 M NaOH, 1% SDS) was mixed in by gentle inversion. The tube was kept on ice whilst 150 µl of ice-cold precipitation solution (Solution 3: potassium acetate; 5 M with respect to K and 3 M with respect to acetate) was added, the tube inverted and then held on ice for 5 mins. After centrifugation at 12,000 rpm for 5 mins the supernatant was transferred to a new tube. An equal volume of phenol:chloroform:isoamyl alcohol (of proportions 25:24:1) was added, mixed by vortexing, then centrifuged for 2 mins at 12,000 rpm. Double-stranded DNA was precipitated from the supernatant using 2 volumes ethanol and mixed by vortexing. This was allowed to stand 2 mins at room temperature, then centrifuged for 5 mins at 12,000 rpm. The supernatant was removed and the pellet washed with 1 ml of ice-cold 70% (w/v) ethanol, spun and centrifuged. The pellet was redissolved in 30 µl TE buffer containing RNase A (20 µg/ml) and stored at –20^{O}C. For

more concentrated plasmid DNA samples, for example those used for restriction endonuclease digestion (see 2.4.2), up to 12 samples from the same culture were extracted separately, resuspended in 8.5 µl and then combined.

2.4.1.2 Whole cell DNA

Whole cell DNA was obtained crudely, using a boiling lysis technique. A single bacterial colony was resuspended in 200 µl MQ water and incubated at 100^{O}C for 5 mins and then centrifuged for 10 mins at 20,000 rpm. The resulting supernatant was used as a template for PCR (see 2.4.3; 2 - 5 µl supernatant used per PCR reaction).

2.4.2 Restriction endonuclease digestion

Between 200 ng and 20 µg of template DNA (PCR products and plasmid DNA) was digested with 5-10 U of restriction endonuclease in 1 x buffer, and 1 x BSA as recommended and supplied by the manufacturer (New England Biolabs, Ipswich, USA) in a total volume of between 20 and 50 µl. The digestion mix was incubated in a water-bath at 37^{O}C for most enzymes, but digestion with the restriction enzyme SwaI was performed at 25^{O}C and BsiWI at 55^{O}C. The digestion time varied depending on the DNA type, for example PCR products were digested for 1-2 hours, but concentrated plasmid DNA was digested for up to 16 hours (with the possibility of re-addition of more enzyme, depending on the enzyme half life). The digestion reaction was separated and visualised via either conventional agarose gel electrophoresis, or pulse field gel electrophoresis (see 2.4.4).

2.4.3 PCR

2.4.3.1 Primers

Oligonucleotides were synthesised by Sigma Genosys, Australia, received in lyophilised form, resuspended in TE at a concentration of 500 µM and stored at –20^{O}C. These were diluted in MQ

H_2O to a working concentration of 50 µM. All primers made during the course of this study, were designed using the program Primer 3 (Rozen and Skaletsky, 2000), specifically, each had an estimated annealing temperature of 65 ^{O}C± 2^{O}C and a 1 bp GC clamp. Sequences and properties of the primers used in PCRs for: 1) Targeting resistance genes, insertion sequences, integrons and other general primers; 2) PCR based replicon typing; 3) Mapping Tn*6029*/Tn*4352*; 4) Mapping Tn*10;* 5) Mapping mercury resistance transposons; 6) Mapping resistance regions in pSRC125; 7) IncHI2 resistance region positions; 8) IncHI2 backbone; 9) IncHI1 variable regions; 10) primers for variants are in Tables A3.1–A3.10, respectively, in Appendix 3.

2.4.3.2 General PCR

PCR amplification reactions contained 160 µM dNTP mix, 1 U *Taq* Polymerase (New England BioLabs, Ipswich, USA) and 50 pmol of each primer and made up to a total volume of 12.5 - 25 *µ*l in 1 x PCR reaction buffer (supplied with *Taq* polymerase; New England BioLabs, Ipswich, USA). Reactions were run in an Eppendorf Thermocycler (Eppendorf) with an initial incubation of 95°C for 5 min, followed by 35 cycles of 95°C for 30 sec, 60°C for 30 sec, and 72°C for 0.5-3 mins (where 1 min was allowed for each kb of sequence expected), then a final incubation at 72°C for 5 min. DNA was diluted to approximately 10 ng/µl and used as a template for PCR. Amplicons were separated and visualised using agarose gel electrophoresis (see 2.4.4) and products were identified using comparisons to standards. Identities were confirmed by either digestion with restriction enzymes (see 2.4.2) or by DNA sequencing (see 2.7).

2.4.3.3 Long-range PCR

For PCR products greater than 3 kb in size, a high fidelity enzyme was used. PCR reactions were carried out in a total volume of 20 µl, in 1 x High Fidelity Buffer (supplied with Phusion^{TM} Polymerase; New England BioLabs, Ipswich, USA) and contained 200 µM dNTP mix, 0.4 U

Phusion™ Polymerase and 20 pmol of each primer. Reactions were run in an Eppendorf Thermocycler (Eppendorf) with an initial incubation of 98°C for 30 secs, followed by 35 cycles of 98°C for 10 sec, 60°C for 30 sec, and 72°C for 2- 5 mins (where 15- 30 secs was allowed for each kb of sequence expected) and a final incubation at 72°C for 5 min. Products were analysed after resolution and visualisation using agarose gel electrophoresis (see 2.4.4).

2.4.3.4 Vectorette PCR

Vectorette PCR (modified from [238]) is a method for determining the unknown location of a particular gene, which is present in multiple copies, such as insertion sequences. 4 μM vectorette primers V-F and V-R (primers in Table A3.1, Appendix 3) were annealed to form the "vectorette unit", by incubating at 65^{O}C for 5 mins, adding 1 mM $MgCl_2$ then cooling to room temperature. Approximately 1 μg of plasmid DNA was then digested at 25^{O}C with a frequently cutting restriction enzyme that does not cut in the gene of interest. 16 nmol of the vectorette unit was added to the digestion mix, supplemented with 0.5 mM ATP and ligated with 1 U T4 DNA ligase. Samples were incubated for 5 cycles of 20^{O}C for 1 hour followed by 37^{O}C for 30 mins in an Eppendorf Thermocycler (Eppendorf), to allow plasmid/vectorette unit complexes to form. These vectorette complexes were then amplified using primers in the vectorette unit (Universal Vectorette Sequencing primer 224; in Table A3.1, Appendix 3) and a primer in the known sequence, using a PCR protocol of an initial incubation at 95°C for 5 min, followed by 35 cycles of 95°C for 30 sec, 60°C for 30 sec, and 72°C for 3 mins, then a final incubation at 72°C for 10 min, run in an Eppendorf Thermocycler (Eppendorf). The product/s of this reaction were run on a 1% (w/v) agarose gel and discrete bands were visualized and extracted from the gel (see 2.4.5). Each band potentially contained a different copy of the target gene, which was sequenced with outwardly facing primer to identify the sequence adjacent to the gene and thus determine the location.

2.4.3.5 Random Amplification Polymorphic DNA (RAPD) PCR

RAPD PCR was performed according to the reaction conditions described in 2.4.3.2, but with the addition of only 1 primer (the volume of the other primer was made up with MQ H_2O). Only 1 primer was used for RAPD PCR, named 1290 (in Table A3.1 in Appendix 1). Amplification conditions were: 5 cycles of 94°C for 5 mins, 37^{O}C for 5 mins and 72^{O}C for 5 mins, followed by 30 cycles of 94°C for 1 min, 37^{O}C for 1 min and 72^{O}C for 2 mins, then a final incubation at 72°C for 10 min, which was run in an Eppendorf Thermocycler (Eppendorf). Products were run on a 1% agarose gel and profiles were interpreted according to the presence or absence of bands, compared to control strains. RAPD PCR was performed on each sample at least twice, to ensure reproducibility.

2.4.3.6 PCR Based Replicon Typing (PBRT)

Classification of plasmids into incompatibility (Inc) groups was achieved using PCR based replicon tying (PBRT), which detects replicons from 18 different Inc groups [145]. Primer pairs are grouped into 3 separate multiplex PCR panels [239] (Panels 1, 2 and 3; primers within each are given in Table A3.2 in Appendix A3) as described previously. Replicon identification was based on unique sizes yielded by primers pairs, which target different replicons, in each panel (sizes in Table A3.2).

2.4.4 Agarose gel electrophoresis

2.4.4.1 Standard agarose gel electrophoresis

0.8%-1.5 % (w/v) agarose gels for electrophoresis were made by boiling analytical grade agarose (Promega, Madison, USA) in 0.5 X TBE for 1 minute. The proportion of agarose in each gel, was dependant on the size of the DNA to be separated: for <0.5 kb 1.2 - 1.5% (w/v) agarose was used; 0.5- 2 kb 1% (w/v) agarose was used; >2 kb 0.8% (w/v) agarose was used. Gels were stained using 0.5 μg/ml Ethidium Bromide (EtBr), which was added to cooled liquid agarose.

Once set, the gel was loaded with sample (approximately 5 µl for PCR sample, 10-20 µl for restriction digests and 4 µl for raw DNA extraction) that was mixed with loading dye. Molecular weight markers 1 kb, 100 bp and HindIII digested lambda (New England BioLabs, Ipswich, USA) were used to estimate the size and concentration of DNA. Agarose gels were run in 0.5 X TBE, in electrophoresis boxes (Bio-Rad, Hercules, USA), at 180 V for 40 mins. DNA bands on agarose gels were visualized with a UV transilluminator (at 280 nm) and imaged using a GelDoc1000 image analysis station (Bio-Rad, Hercules, USA).

2.4.4.2 Pulse Field Gel Electrophoresis (PFGE)

Restriction enzyme digests of plasmid DNA were separated on 0.9 – 1.2 (w/v) % agarose gels in 0.5 X TBE buffer, using the CHEF-DR II pulse field electrophoresis unit (Bio-Rad, Hercules, USA). Gels were run under the following conditions: a total run time of 14 -18 hours (depending on % agarose and sizes of DNA fragments), a voltage of 6 V, an initial switch time of 0.1 sec and a final switch time of 4 secs, all at 14^{O}C. The molecular size standard used was prepared (courtesy of Dr. Neil Wilson, Hall Laboratory) by combining undigested λ(48.5 kb), XhoI digested λ (33.5, 15 kb) and HindIII digested λ (23.1, 9.4, 6.7, 4.4, 2.3 and 2 kb), which was heated to 65ºC for 3 mins before loading to separate sticky ends. To determine the sizes of fragments, a standard curve was constructed based on the migration distances of bands from the molecular size standard.

2.4.5 DNA purification techniques

DNA was routinely purified using ethanol precipitation where 1/10 volume of sodium acetate, then 2 volumes of 100% ethanol were added to the DNA, which was cooled at –20^{O}C for at least 1 hour, and then spun at 12,000 rpm for 20 mins. The pellet was then washed with 1 mL ice-cold 70% ethanol and resuspended in TE or EB, depending on future manipulation. If multiple bands were observed in a PCR product, the products were run on an agarose gel and the desired band

was excised using visualisation under long wave UV (to prevent damage of the DNA). The extracted band was then purified using Qiaquick Gel Extraction Kit (QIAGEN Inc, Valencia, CA, USA) and the spin column protocol was followed according to the manufacturer's instructions and resuspended in a final volume of 30 µl EB. This kit was also used to purify restriction digest fragment/s of plasmid DNA, for later use in cloning (see 2.5). In this case, bands from a range of sizes (depending on the predicted insert size) were extracted from the gel and purified.

2.4.6 Estimation of DNA concentration

Concentration of DNA was estimated using a Thermo Scientific Nanodrop™ 1000 Spectrophotometer (Thermo Fisher Scientific) using optical density (OD) at 260 nm. DNA was also estimated for purity using the absorbance ratio of 260/280 nm, where DNA was considered pure if the ratio was ≥1.8.

2.5 Cloning

2.5.1 Preparation of $CaCl_2$ competent cells

Chemically competent cells were prepared (adapted from Morrison, "Transformation and preservation of competent bacterial cells by freezing" Methods Enzymol, (1979), **68**:326-31) by inoculating 1 L LB (in a 5 L flask) with 1 ml of DH5α starter culture. The flask was shaken at 37°C, and spectrophotometric readings of the culture were taken until an OD_{600} of 0.5 was reached. The cells were cooled on ice and then pelleted at 8,000 rpm for 8 mins at 4°C. Cells were resuspended in 0.1 M $MgCl_2$ and pelleted as above. Cells were then resuspended in ice-cold 0.1 M $CaCl_2$ and re-pelleted. The final pellet was resuspended in 14% glycerol + 86 mM $CaCl_2$. Cell aliquots of 0.5 ml were snap frozen in liquid nitrogen and stored at –80°C. These cells were

tested for competency using the cloning vector pUC19 (see Table 2.5) and yielded at least 10^5 transformants per μg DNA.

2.5.2 Preparation of insert

Restriction digestion (see 2.4.2) was performed on approximately 5 μg plasmid DNA and run on a 0.9% (w/v) agarose gel. Fragments of the desired size, or if unknown a range of sizes, were excised and purified using a QIAquick Gel Extraction Kit (QIAGEN Inc, Valencia, CA, USA), resuspending in 30 μl EB. DNA concentration was also estimated (see 2.4.6).

2.5.3 Preparation of vector

Vector was prepared by digesting 1 μg pUC19 DNA (New England BioLabs, Ipswich, USA) with I U of restriction enzyme and then phosphate groups were removed to prevent self-religation, using Antarctic Alkaline Phosphatase as per the manufacturer's instructions (New England BioLabs, Ipswich, USA). Vector was dephosphorylated by incubating 1 μg of pUC19 DNA in 1 x Antarctic Phosphatase buffer (50 mM Bis-Tris-Propane-HCl, 1 mM $MgCl_2$, 0.1 mM $ZnCl_2$, pH 6.0) and 5 U Antarctic Phosphatase enzyme (New England BioLabs, Ipswich, USA), in a total volume of 20 μl and incubating for 15 mins at 37ºC. The reaction was stopped by heating for 5 mins at 65ºC.

2.5.4 Ligation

Ligation reactions were set up on ice and carried out using the Quick Ligation Kit according to the manufacturer's instructions (New England BioLabs, Ipswich, USA). 1 μl of Quick DNA ligase was incubated for 5 -30 mins at room temperature with insert (~5 μg) and prepared vector (~100 ng) in a molar ratio of 3:1. Reactions were set up in 1 X Quick ligase buffer to a total volume of 10 μl and were used immediately or stored at -20^{O}C.

2.5.5 Transformation

Transformation was performed by thawing 100 µl of $CaCl_2$ chemically competent DH5α (Nx^R) cells on ice and adding 1-10 µl of ligation mix (~1 µg DNA) then incubating on ice for 30 minutes. After heat shocking the cells at 42ºC for 40 secs, the cells were incubated on ice for 2 mins. The cells were diluted 1 in 10 with LB and incubated with shaking at 37ºC for 1 hour. Dilutions of the culture were made, if necessary, and were plated (100 µl/ plate) in duplicate on LA containing Ap (selection for vector; to determine transformation frequency) and LA containing an appropriate antibiotic (selection for insert). If the insert did not carry a resistance gene marker, 100 µl of the transformation mix was plated onto LA supplemented with 0.004% (w/v) X-gal (5-bromo-4-chloro-indolyl-n-galactopyranoside) and 0.1 mM IPTG (isopropyl-B-D-thiogalactopyranoside) per 25 ml plate, to allow for blue/white selection. White colonies were isolated and screened for inserts of interest (see below).

2.5.5.1 Clone analysis and controls

The insert size of each clone was determined by extracting plasmid DNA from at least 2 colonies off each selection plate, which was then digested with the enzyme used for cloning and separated on an agarose gel next to pUC19 control. Transformation frequency was determined by transforming 100 ng of pUC19 DNA and selecting with Ap, which was calculated back to give the number of transformants/ µg DNA.

2.6 Southern hybridisation

2.6.1 Nitrocellulose membrane DNA transfer

Southern hybridisation was performed on 5 µg plasmid DNA digested with the desired restriction enzyme (see 2.4.2) and separated on a 0.9% (w/v) agarose gel in 0.5 X TBE and run next to DIG-labelled DNA Molecular Weight Ladders II and/or VII (Roche Applied Science,

Mannheim, Germany) at 100 V for 2 hours. Gels were depurinated twice in 0.2 M HCl with gentle agitation for 5 min, denatured twice in denaturation solution (0.4 M NaOH and 1 M NaCl) for 15 min and neutralised in neutralisation solution (1 M Tris, 1.5 M NaCl) with 2 steps of 7.5 mins. DNA was transferred to a Hybond N+ nylon membrane (Amersham, Buckinghamshire, UK) via capillary transfer overnight, then fixed by UV crosslinking using a Stratalinker® 1800 at 600 μJ (Stratagene, CA, USA).

2.6.2 Probe preparation

DNA probe labelling was performed using the DIG-labelling mix (Roche Applied Science, Mannheim, Germany), according to the manufacturer's instructions. Digoxygenin (DIG)-labelled dUTPs were incorporated into a PCR reaction targeting the individual genes IS*26* (using primers IS*26*-F/RH601; Table A3.1 in Appendix 3), bla_{TEM} (using primers RH605/RH606; Table A3.1) or *strA* (using primers strA-F/strA-R; Table A3.1). PCR reactions were purified as described in 2.4.6, denatured by boiling at 100^{O}C for 10 mins and stored at -20^{O}C. Hybridisation solutions were prepared by adding 10 μl of DIG-labelled probe to 50 ml of Easy Hyb pre-hybrisation solution (Roche Applied Science, Mannheim, Germany), which was used fresh or stored at -20^{O}C. Once used, the hybridisation solution was collected and stored at -20^{O}C for re-use later.

2.6.3 Hybridisation

Hybridisation was performed using the DIG Easy Hyb Kit (Roche Applied Science, Mannheim, Germany) and sterile hybridisation bottles (Pyrex tubes, Amersham, Buckinghamshire, UK). The membrane was pre-hybridised with Easy Hyb Solution (Roche Applied Science, Mannheim, Germany) at 58^{O}C for 30 min with rotation, followed by an initial hybridisation in Easy Hyb Solution containing 25 pM DIG-labelled probe at 68^{O}C for 30 min with rotation, then a final

hybridisation at 42^{O}C overnight with rotation. The membrane was prepared for detection by undertaking 2 washing steps of 5 min in 2 x SSC (3 M sodium chloride and 0.3 M sodium citrate) and 0.1% SDS at room temperature, followed by 2 more washes in 2 x SSC and 0.1% SDS for 15 min at 68 ^{O}C. Immunological detection was performed at room temperature using DIG High Primer Labelling and Detection Starter Kit II (Roche Diagnostics, Mannheim, Germany) by rinsing for 5 min in 1 x Roche Washing Buffer (Roche Applied Science, Mannheim, Germany), blocking for 30 min in 1 x Roche Blocking Buffer (Roche Applied Science, Mannheim, Germany), incubating for 30 min with 1 x Antibody Solution (Roche Applied Science, Mannheim, Germany), washing twice for 15 mins in 1 x Roche Washing buffer (Roche Applied Science, Mannheim, Germany) and equilibrating for 5 min in 1 x Detection Buffer (Roche Applied Science, Mannheim, Germany).

2.6.4 Detection and visualisation

Luminescence was detected using CSPD chemiluminescence substrate (Roche Applied Science, Mannheim, Germany) and exposure to Hyperfilm^{TE} ECL X-ray film (Amersham, Buckinghamshire, UK), which was allowed to expose for 1- 60 mins, then developed using a CP100 film processor (AGFA, Mortsel, Belgium). The membrane could then be stripped, using 2 washes for 15 min with 0.2 M NaOH and 0.1% (w/v) SDS at 37 ^{O}C, if desired for later use with another probe.

2.7 DNA sequencing and analysis

2.7.1 Preparation of DNA for sequencing

30-90 ng of purified PCR sample or 500-1,500 ng of plasmid DNA purified by ethanol precipitation (see 2.4.5) was sequenced with 3-10 pmol of suitable primer in a final volume of 12 µl.

2.7.2 Sequencing reaction

Automated sequencing was performed at the DNA Analysis Facility at Macquarie University (Sydney, Australia) on an ABI PRISM 377 DNA Sequencer (AME Bioscience), at Supamac at the University of Sydney (Sydney, Australia) on a 3130 Exel Genetic Analyzer (Applied Biosystems, Carlsbad, USA), or at the Australian Genome Research Facility (AGRF) at Westmead Millennium Institute (Sydney, Australia) on an Applied Biosystem 3730, 7900 Genetic Analyser and Transgenomic WAVE-dHLC (Applied Biosystems, Carlsbad, USA), using the Big Dye system.

2.7.3 Sequencing analysis

Sequences were visualized and assembled using Sequencher version 4.9 (Gene Codes Corporation, Ann Arbor, Michigan). Sequenced regions were identified using BLAST (http://blast.ncbi.nlm.nih.gov) or in pair-wise comparisons using the BLAST paired alignment facility. Antibiotic resistance regions were delineated by detecting a series of common features, such as 3'-CS, 5'-CS, IS*1*, IS*26*, IS*10*, IS*6100*, using BLAST pair-wise analysis. This approach often allowed resistance region boundaries to be defined and then analysis of the sequence in between and flanking regions (including detection of direct repeats) could be undertaken. Gene Construction Kit version 2.5 (Textco, West Lebanon, New Hampshire) was used to create figures to scale. Single nucleotide differences between the expected sequence based on published sequences and the generated sequence were confirmed using more than one PCR product as sequencing template.

CHAPTER THREE

Resistance regions in IncHI2 plasmids

3.1 Introduction

3.1.1 S. *Typhimurium strains SRC26 and SRC125*

Salmonella enterica serovar Typhimurium strains SRC26 and SRC125 were isolated in Australia from bovine infections, in 1999 and 2000, respectively. They were obtained from the *Salmonella* collection held at the Microbiological Diagnostic Unit (The University of Melbourne), where the resistance phenotype and phage type (PT) had been determined. Both were resistant to multiple antibiotics (Table 3.1). SRC26 was PT1 and SRC125 was PT44. Both strains had been partially characterised in Dr. Djordjevic's Laboratory at the Elizabeth Macarthur Agricultural Institute (EMAI), where the antibiotic resistance genes were identified and some linkages between these genes were established, during routine screening of a larger *S*. Typhimurium set. The resistance genes *dfrA5* and *sul1* were detected in both strains and were located in a class 1 integron, which was linked to the insertion sequence IS*6100* (see Table 3.1; Xuilan Liu and Renee Levings, personal communication). Both strains also contained *aphA1* and *strA/B,* and SRC26 contained the additional genes bla_{TEM} and *sul2* (see Table 3.1;[232]). These additional genes in SRC26 were linked to each other and to *strA/B*, and *aphA1* was located within the known kanamycin resistance transposon Tn*4352* [84]. The location of the remaining resistance genes had not been determined. A mercuric ion resistance gene from Tn*1696, merA*$_{1696}$, was also identified in both strains.

Table 3.1 Resistance genes and resistances of SRC26 and SRC125.

Strain/ SRC No.	Isolation year	Resistance profile[1]	Resistance genes
125	2000	Sm, Su, Tc, Tp, Km/Nm	*tetA*(B), *dfrA5*[2], *sul1*[2], *aphA1, strA/B*
26	1999	Sm, Su, Tc, Tp, Km/Nm, Ap	*tetA*(B), *dfrA5*[2], *sul1*[2], *aphA1*[3] *strA/B*[3], *sul2*[3], bla_{TEM}[3]

[1]Ap: ampicillin; Km: kanamycin; Nm: neomycin; Sm: streptomycin; Sp: spectinomycin; Su: sulphamethoxazole; Tc: tetracycline; Tp: trimethoprim
[2]linked to a class 1 integron and these strains also contained *intI1* and IS*6100*
[3]these genes were linked together

SRC26 and SRC125 were also shown to contain IncHI2 plasmids (Nick Evershed, Hall Laboratory, personal communication) using PCR-based replicon typing (PBRT) [145]. The sequence of the pSRC26 amplicon was 100% identical to the R478 sequence, the only IncHI2 plasmid published when this work began ([58]; GenBank accession no. BX664015). R478 was one of the first large plasmids to be sequenced and is the exemplar IncHI2 plasmid, as it had been extensively studied (see for example [8],[200],[166],[202],[240]). It contains three resistance genes *catA1, aphA1* and *tetA*(B) and thus confers resistance to the antibiotics chloramphenicol (Cm), kanamycin (Km) and neomycin (Nm), and tetracycline (Tc). R478 also confers resistance to mercuric ions and contains a mercuric ion resistance module.

The aim of this chapter was to elucidate the structure, origin and position of antibiotic resistance regions from IncHI2 plasmids from Australian *S*. Typhimurium isolates and compare them to those in sequenced plasmids.

3.2 IncHI2 plasmids in SRC26 and SRC125

The resistance profiles for SRC26 and SRC125 obtained by the MDU (Table 3.1) were confirmed using the CDS disc diffusion assay (inhibition zone sizes listed in Table 3.2). Conjugation of plasmids from SRC26 and SRC125 into *E. coli* recipient strain E294 was performed at 37^{O}C using tetracycline to select for the plasmid and rifampicin for the recipient. Conjugation occurred at a frequency of 3.8 x 10^{-5} transconjugants per donor for pSRC26 and 5.3 x 10^{-5} transconjugants per donor for pSRC125. Ten individual transconjugants were tested for antibiotic resistances initially by picking and patching onto antibiotic-containing plates (all antibiotics in Table 3.1), and then at least 2 representative colonies were tested with CDS disc diffusion, to confirm the resistance phenotype (Table 3.2). All resistances transferred, except for the streptomycin resistance of SRC125 (highlighted purple in Table 3.2). All antibiotic resistance genes present in the *S*. Typhimurium strains (from Table 3.1) were

detected in these transconjugants using PCR, including the *strA* and *strB* streptomycin resistance genes in E294/pSRC125. This indicated that all the resistance genes in SRC26 and SRC125 were on plasmids. Although the *strA/B* genes were present in the SRC125 transconjugants, they did not express streptomycin resistance (discussed in section 3.5.2). The control plasmid R478 was also tested (Table 3.2).

Table 3.2 CDS on SRC125 and SRC26 and their transconjugants

	Inhibition zone in annular radius (mm)[1, 2]										
Strain	**Ap[3]**	**Cm**	**Sm**	**Sp**	**Su**	**Tc**	**Gm**	**Km**	**Nm**	**Tp**	**CTX**
SRC26	1[4]	13	0	8.5	1	0	12	1	2	0	ND
SRC26	1	13	0	8.5	0	0	12	0	1	0	17
SRC125	11	9.5	0.5[5]	6	0	0	8.5	0	0	0	ND
SRC125	12	10	2	6.5	0	0	8	0	0	0	16
E294/pSRC26 1[6]	1	11	0.5	7	0	0.5	9	0	2	0	16
E294/pSRC26 2[7]	1	10	0	6	0	0	9	0	1	0	16
E294/pSRC125 1	8	12.5	11	8	0	0	13	0	0	0	17
E294/pSRC125 2	9	12	11	7	0	0	12	0	0	0	18
R478 (E294)	10	0	7	7	15	1	11	0	2	14	16
E294	12	13	12	9	16	12	13	11	9.5	15	17
E294	10	11	11	8	17	12	12	10	9	14	16

[1]Inhibition zones of <6mm are regarded as resistant
[2]ceftazidime, florfenicol, tobramycin, amikacin, nalidixic acid and ciprofloxacin also tested: all were sensitive
[3]Ap: ampicillin; Cm: chloramphenicol; Sm: streptomycin; Sp: spectinomycin; Su: sulphamethoxazole; Tc: tetracycline; Gm: gentamicin; Km: kanamycin; Nm: neomycin; Tp: trimethoprim; CTX: cefotaxime
[4]zone sizes that reflect sensitive colonies are highlighted in yellow
[5]zone sizes that differ from *Salmonella* donor strains and transconjugants are highlighted in purple
[6]transconjugant colony 1
[7]transconjugant colony 2

PCR for the IncHI2 replicon performed on transconjugants E294/pSRC26 and E294/pSRC125 showed that they contained the IncHI2 plasmids, here named pSRC26 and pSRC125, respectively (see Fig 3.1). The extra bands evident in Fig 3.1 are due to one of the primers (IncHI2-R) being located in an iteron region [145] associated with the IncHI2 replication region. These iteron regions comprise of a number of short sequences repeated multiple times ([58]; see section 1.5.1). This banding, which is also produced by R478 (Fig 3.1), is characteristic of IncHI2 plasmids.

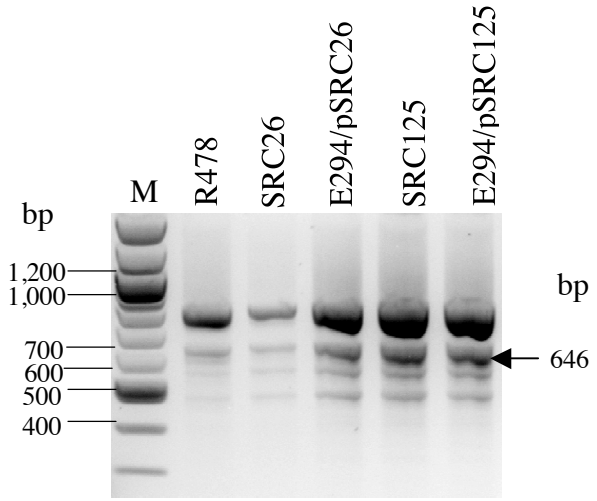
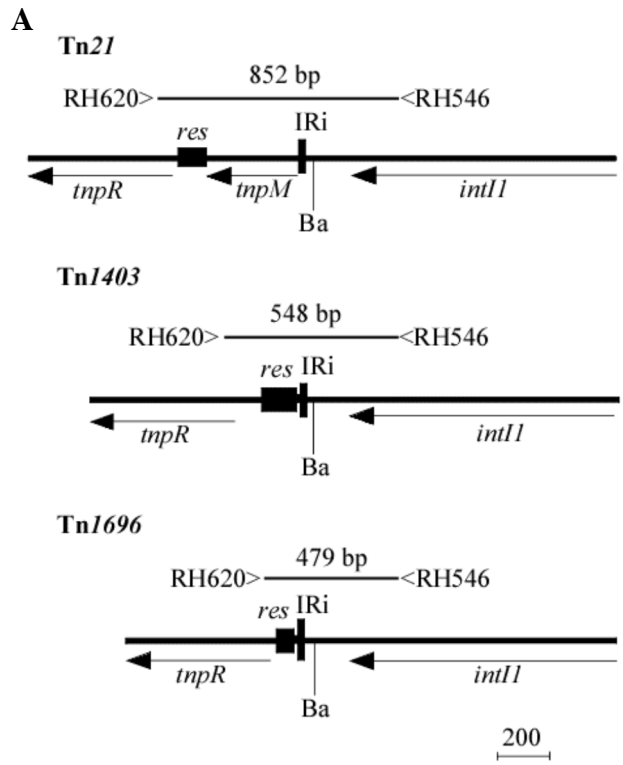


Fig 3.1 PBRT analysis of SRC26, SRC125 and their transconjugants. The horizontal arrows indicate the position of the predicted amplicons for the IncHI2 replicon (a 646 bp product). Lanes are labelled SRC26 or SRC125 for the *Salmonella* strain, and E294/pSRC26 or E294/pSRC125 for the *E. coli* transconjugant and the control R478. 100 bp Molecular Weight Markers are loaded in lane M and selected sizes shown.

3.3 Structure and position of the MARRs

3.3.1 Location of the integron

Both pSRC26 and pSRC125 carried a class 1 integron containing the *dfrA5* gene cassette and *sul1*. However, integrons can be found in many locations. Because these strains contained IS*6100*, an In4-type integron (see Fig 1.10) was likely. In addition, a Tn*1696*-type *merA* mercuric ion resistance gene was detected in both strains (using primers merA1 and merA5; Table A3.5 in Appendix 3), this suggested that the In4-type integrons may lie within a mercuric ion resistance transposon. Additionally, both *E. coli* transconjugants containing pSRC125 and pSRC26 and the *S.* Typhimurium strains SRC125 and SRC26 were determined to be resistant to mercuric ions by patching onto LA plates containing 20 μg/ml mercuric chloride (plate pictures in Fig A2.1 in Appendix A2) using R478 and pACYC::Tn*21* as positive controls. PCR was developed to locate the integrons in Class II transposon backbones. Primer RH620 was designed to detect the *tnpR* gene from a range of Class II transposons including Tn*21*, Tn*1696*, Tn*1403* and Tn*5051* and was used with primer RH546 in *intI1* (Fig 3.2A). The PCR product sizes were unique for each transposon (Fig 3.2B) because the position of the integron differs (Fig 3.2A).



B

Tn	PCR amplicon (bp)	BamHI fragments (bp)
Tn*21*	852	250, 602
Tn*5051* [1]	698	250, 447
Tn*1403*	548	250, 297
Tn*1696*	479	250, 228

[1]no sample containing this was available and thus it was not tested experimentally

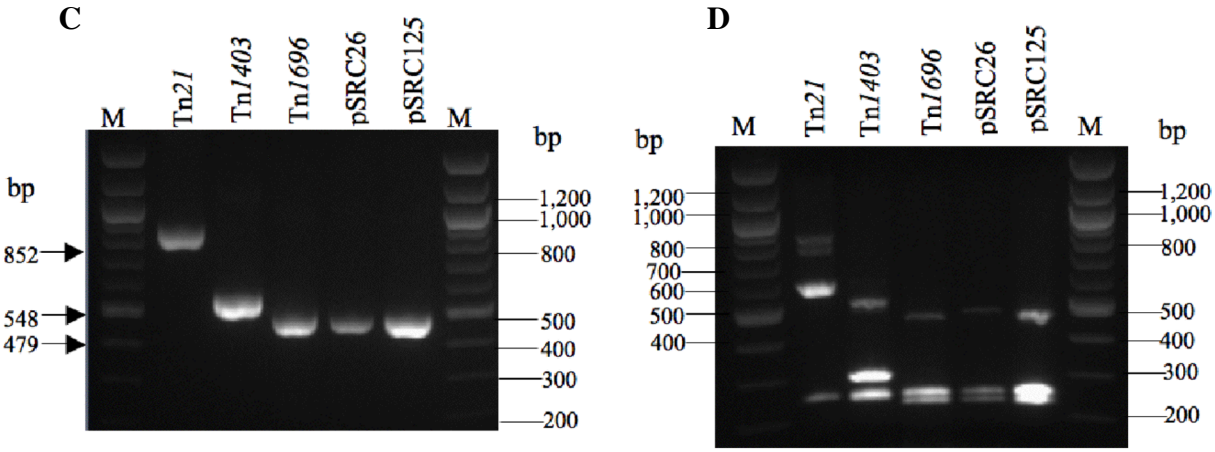
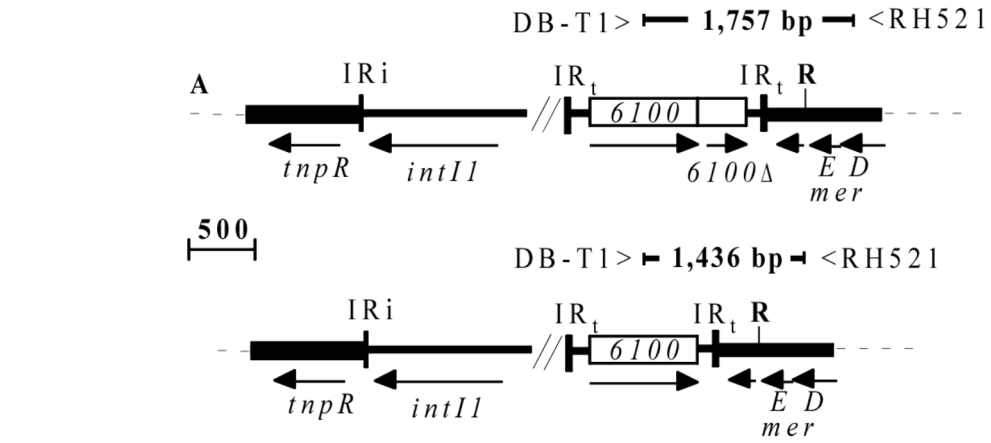


Fig 3.2 *tnpR*/*intI1* linkage by PCR. A Schematic of positions of integrons in different mercury (II) transposons. The uniquely sized PCR products are shown above with the primer names shown to the side with an arrow showing the direction of the primer. Primer sequences in Table A3.5 in Appendix 3. Horizontal arrows represent direction and extent of genes. Vertical bars are the inverted repeats (IRi) from the integrase end of integrons (where IRi denotes the IR from the 5' conserved segment (5'-CS) end of the integron). The filled box represents the *res* site. The positions of the BamHI sites are marked as "Ba". **B**. Table showing predicted PCR product sizes and BamHI digestion fragment sizes **C**. Agarose gel of PCR products. Horizontal arrows represent the positions of the predicted amplicons for Tn*21* (852 bp), Tn*1403* (548 bp) and Tn*1696* (479 bp). **D**. Agarose gel of PCR product digests. Lanes are labelled pSRC26 and pSRC125 for the plasmids, Tn*1696*, Tn*21* and Tn*1403* for the positive controls. 100 bp Molecular Weight Markers are loaded in lane M and selected sizes are as indicated.

For this PCR, R1033 containing Tn*1696*, pACYC184::Tn*21* and R388::Tn*1403* were used as standards, but a structure containing Tn*5051* was not available and could not be tested. Amplicons were digested with BamHI, which removes a fragment of 250 bp of the 5'-CS, to confirm the identity of the product (Fig 3.2B). In both pSRC125 and pSRC26, the IRi end of the integron was linked to the *tnpR* gene as in Tn*1696* and sequencing confirmed that it was in precisely the same position (Fig 3.2C/D).

PCR was also used to determine if pSRC125 and pSRC26 contained the IRt end of an In4-type integron next to a *mer* module, using PCR primers in IS*6100* and *merE* (Fig 3.3). This PCR also serves to detect if a second partial copy of IS*6100,* abutting the complete IS*6100,* was present in the In4-type integron (as in Tn*1696*) or not (as in the standard pXS2) (Fig 3.3A), as different sized amplicons would be produced. PCR products were digested with RsaI to confirm their identities (Fig 3.3B). pSRC125 yielded a band of 1.4 kb and thus linkage was established, and the partial IS*6100* was not present (see Fig 3.3C/D). No amplicon was observed for pSRC26 (see Fig 3.3C/D). However, PCR internal to IS*6100* and *merE* were positive for pSRC26 (data not shown) therefore these genes were present. A possible explanation for this phenomenon is that a large insertion is present in between them and thus linkage could not be achieved (see section 3.5.1 below).

A structure for the resistance region of pSRC125 was proposed, based on Tn*1696,* but containing the *dfrA5* gene cassette and without the IS*6100*Δ. The complete 13,302 bp sequence of this transposon, designated Tn*6025*, was determined using overlapping PCR products as templates (Fig 3.4; primers in Table A3.6 in Appendix A3). The sequence was as predicted and deposited into the GenBank database under accession number GU562437.



B

IS*6100Δ*	PCR (bp)	RsaI digest (bp)
+	1,757	1,150+500 (+58+49)
-	1,436	829+500 (+58+49)

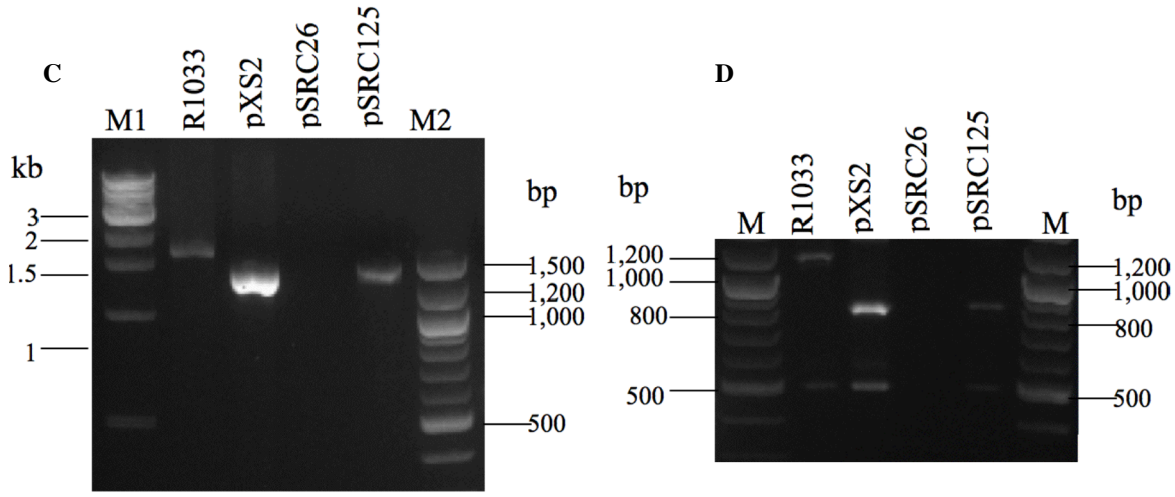


Fig 3.3 IS*6100* and mer_{1696} linkage by PCR. A Schematic of different configurations of IS*6100*-IRt-mer with and without IS*6100*Δ. The centre of the integron is not shown and is represented by 2 slashes. **B**. Table showing predicted PCR sizes and RsaI digestion fragment sizes **C**. Gel of PCR products. Lanes labelled M1 and M2 are the 1 kb and 100 kb molecular weight markers, respectively and selected sizes are shown. **D**. Gel of PCR product digests. Lanes labelled R1033 and pXS2 are the IS*6100*Δ+ and IS*6100*Δ– controls. Lanes marked M are the 100 bp molecular weight marker and relevant sizes are as indicated. Primer sequences in Table A3.5 in Appendix 3. Other features are as in Fig 3.2.

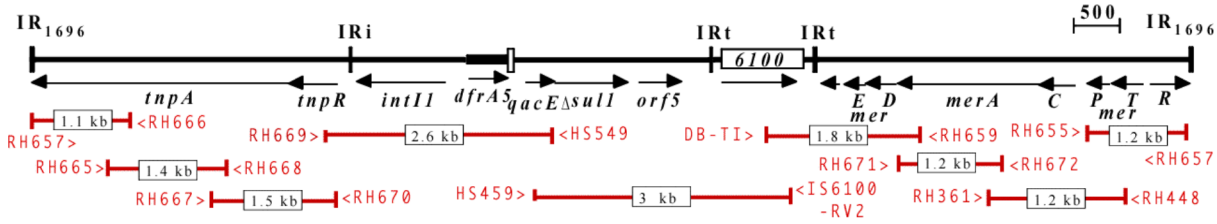


Fig 3.4 PCR for mapping Tn*6025*. Primers are shown as red vertical lines, with the size of the PCR product boxed between them. The open, numbered boxes represent insertion sequences of that number. All other features are as in Fig 3.2. Primers sequences are given in Table A3.6 in Appendix A3.

The Tn*1696*-derived resistance region in pSRC26, named Tn*6026*, was also sequenced using the same PCR mapping scheme as was used to sequence Tn*6025* (Fig 3.4). Unlike Tn*6025*, Tn*6026* contained the IS*6100*Δ present in In4. Tn*6026* also contained a 10.6 kb insertion (Tn*6029*/Tn*4352*; described and analysed in 3.5.1) in IS*6100*. The 25.4 kb sequence of Tn*6026* was deposited into GenBank under accession number GQ150541.

*3.3.2 Comparison to Tn*1696

The shared region of pSRC125, pSRC26 and Tn*1696* were compared using pair-wise Blast searches. Single nucleotide differences between Tn*6026* or Tn*6025* and the published Tn*1696* sequence (GenBank accession no. U12338) were confirmed using more than one PCR product as sequencing template. Differences found in comparison to Tn*1696* were a single base change in *urf2Y* in both pSRC26 and pSRC125 and a silent mutation in *sul1* in pSRC26. Overall, the only major difference between Tn*1696* and Tn*6025*/Tn*6026* was that the gene cassettes in In4 vary from the *dfrA5* gene cassette found in Tn*6025*/Tn*6026* (Fig 3.5A-C). Also, compared to Tn*1696,* Tn*6025* is missing IS*6100*Δ, which could simply be a result of homologous recombination of the identical sequence to the left and right of the IS*6100*Δ.

Both pSRC26 and pSRC125 also lacked a 19 bp duplication present in Tn*1696,* which is located 39 bp to the left of the gene cassettes. The promoter of *dfrA5*, positioned 183 bp to the left of the cassette, is a strong PcS promoter, identical to that of In4 in Tn*1696* [67]. However, all 3 sequenced promoters associated with *dfrA5* (see GenBank accession nos X12868, GQ259888, HM999791) are hybrid 2 (PcH2), with a single base difference in the -10 region, compared to PcS. Exchange of gene cassettes between class 1 integrons can occur via homologous recombination between the 5'- and 3'- conserved segments [72,111]. In this case, homologous recombination may have occurred, on the 5'-CS side, in between the PcS promoter and the 19 bp duplication. This would thus bring in the *dfrA5* gene cassette and the

surrounding sequence (which on the left lacked the 19 bp duplication), but not its original promoter.

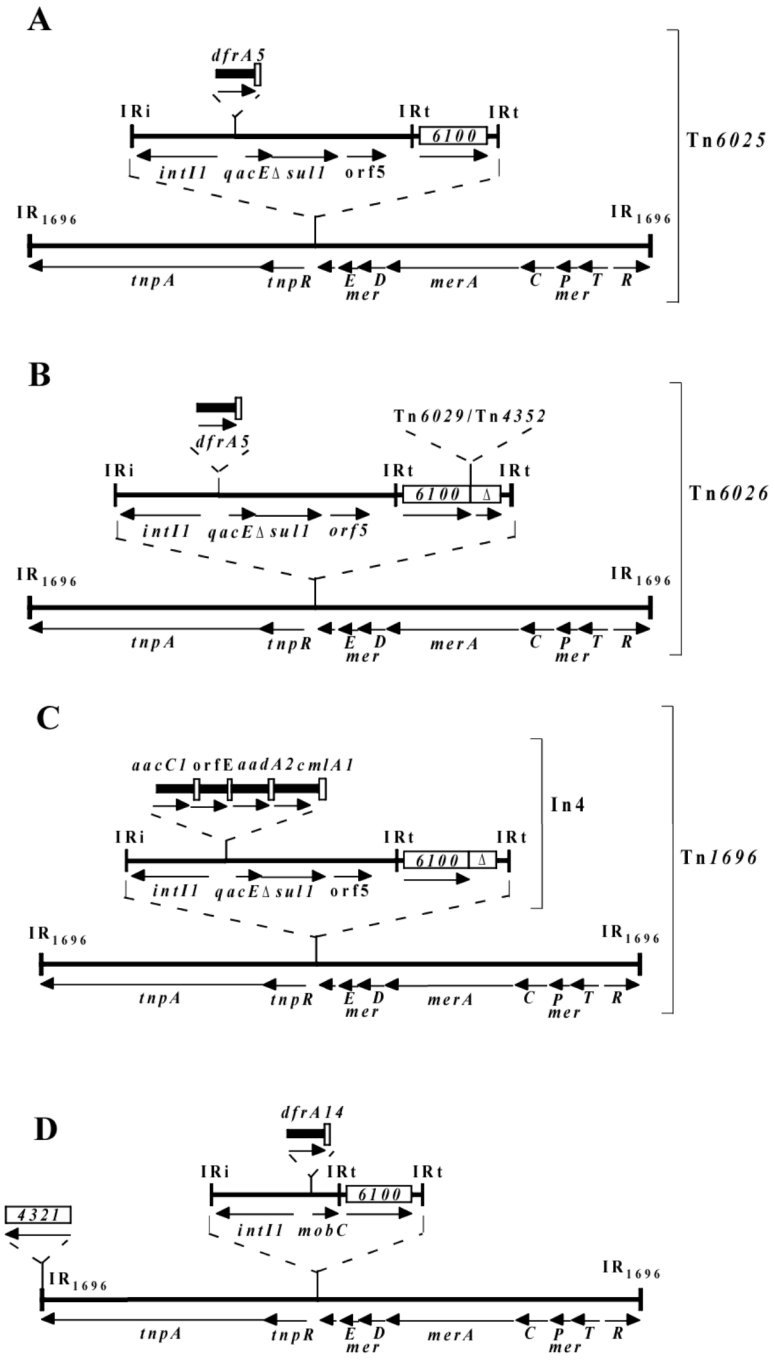


Fig 3.5 Tn*1696* and related Tns with integrons in the same position as In4. A. Tn*6025* **B.** Tn*6026* **C.** Tn*1696* **D.** transposon in pHCM1. The insertion in Tn*6026* is not shown but its name is shown above. Other features are as in Figs 3.2 and 3.3.

Tn*6025* and Tn*6026* were also compared to other Tn*1696* relatives in the GenBank database using the BLASTn analysis tool. Only one other case of an In4-type integron in the same

position as Tn*1696*, Tn*6026* and Tn*6025* in the *res* site of a mercury (II) resistance class 2 transposon, equivalent to Tn*5036* ([241]; Fig 3.5 A-C) was detected. It is found in the IncHI1 plasmid pHCM1 (GenBank accession number AL513383). In this transposon, the integron has a *dfrA14* gene cassette and has lost the 3'-CS (Fig 3.5D). Tn*6026* and Tn*6025* are therefore only the third and fourth reports of intact, close relatives of Tn*1696*.

*3.3.3 Position of Tn*6026 *and Tn*6025

R478 was annotated as carrying the mercury resistance module of the mercury (II) resistance transposon Tn*21* [58]. However, further analysis of this module revealed it was from Tn*1696* and not Tn*21* (Ruth Hall, personal communication). The *mer* modules of Tn*1696* and Tn*21* share only 82% sequence identity. R478 lacks the transposition (*tnp*) module and the majority of the integron of Tn*1696*, as these had been replaced with an IS*26*-bounded multiple antibiotic resistance region (MARR). To determine if the *mer* module of Tn*6026* and Tn*6025* was in the same position in the IncHI2 backbone as it is in R478, primers were designed in $merR_{1696}$ and in the adjacent IncHI2 backbone (Fig 3.6A). R478, pSRC125 and pSRC26 all produced amplicons of 540 bp (Fig 3.6B) and the sequence of the PCR products obtained from pSRC26 and pSRC125 showed that the IR_{mer} of Tn*6026* and Tn*6025* were in precisely the same position as in R478 in the IncHI2 backbone.

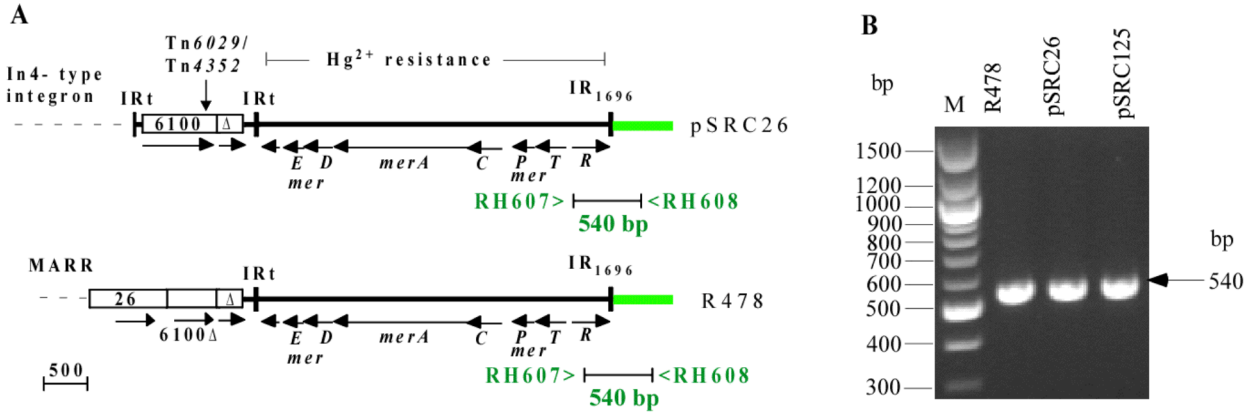


Fig 3.6 Tn*1696* *mer* boundaries as in R478. **A** Schematic of the boundary PCRs of mer_{1696} in R478. Description of regions are bracketed above. **B** Agarose gel of *merR*/backbone linkage PCR for pSRC125 and pSRC26. Other features are as in Figs 3.2, 3.3. Sequences of primers in Table A3.7 in Appendix A3.

As R478 lacked the Tn*1696*/Tn*5036* transposition module, to find the location of the *tnp* end of Tn*6026*, a PstI fragment of pSRC26 was cloned into pUC19. The restriction enzyme PstI was used as there were no sites in between the IR_{tnp} and the *dfrA5* gene cassette, thus trimethoprim resistance could be selected for and the insert would extend beyond IR_{tnp}. The clone recovered, named pRMH951, had a 12 kb insert that contained the *dfrA5* gene cassette, the 5'-CS, the transposition module and 5.6 kb of unknown sequence (Fig 3.7). 0.9 kb of sequence adjacent to IR_{tnp} was obtained using primer RH329 in *tnpA* (Fig 3.7).

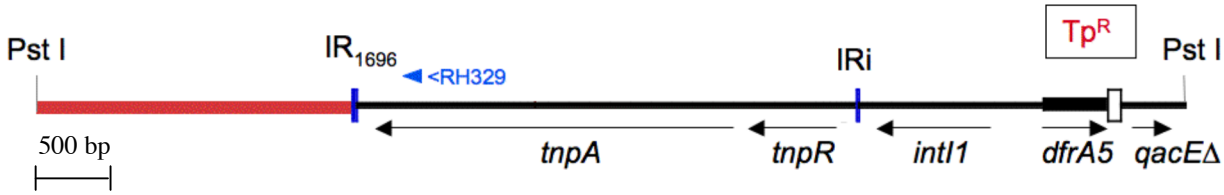


Fig 3.7 Insert of pRMH951. PstI sites are shown above. The IRs are shown by vertical blue lines. The primer used to sequence out of tnp_{1696} is named in blue. RH329 sequence is in Table A3.7 in Appendix 3. Unknown sequence to the left of IR1696 is shown in red. The selection used to recover the clone is boxed above the *dfrA5* trimethoprim resistance gene. Other features are as in Figs 3.2 and 3.3.

Analysis of the sequence adjacent to IR_{tnp} revealed firstly that Tn*6026* was surrounded by a 5 bp direct repeat, indicative of transposition. A PCR primer was designed in the backbone sequence adjacent to IR_{tnp} and used with a primer in $tnpA_{1696}$ (Fig 3.9) in a PCR performed on pSRC125. pSRC125 produced an amplicon of 800 bp, the sequence of which showed that this boundary in pSRC125 was identical to the equivalent region in pSRC26 (Fig 3.8). Thus Tn*6025* was in the same position as Tn*6026*.

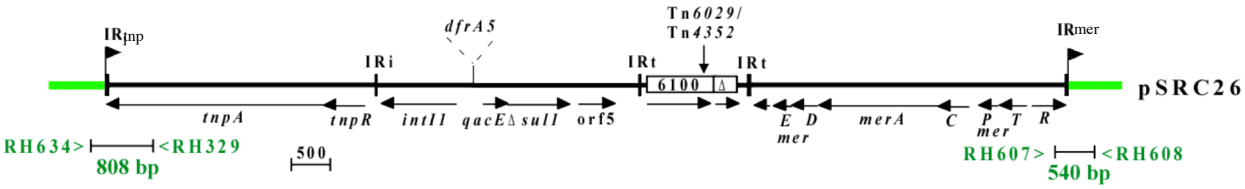


Fig 3.8 Tn*6026* boundary PCRs. Schematic of 2 Tn*6026* in pSRC26 showing the PCRs used to identify the position of Tn*1696*-like in the IncHI2 backbone, which is marked in green. The flags represent the position and direction of direct repeats. Other features are as in Figs 3.2, 3.3. The sequences of primers RH634 and RH329 are in Table 3.7 in Appendix A3.

Though the sequence to the left of IR_{tnp} was not in R478, it matched another IncHI2 plasmid, named pK29. Initial analysis of pK29 showed that it did not appear to contain any of Tn*1696*. However, alignments with the pRMH951 sequence revealed that a 22 bp remnant of the 38 bp IR_{1696} remained in pK29 (Fig 3.9). The IR was truncated by IS*4321*, an insertion sequence that targets 38 bp IRs of mercuric ion resistance transposons [31]. Furthermore, only 184 bp of IS*4321* (Fig 3.9) was left, as it was truncated by an IS*26*. This IS*26* formed part of a MARR, containing the antibiotic resistance gene bla_{CMY-8} and a class 1 integron containing *catB2* and *aadA2* and 2 copies of *sul1*. Thus, pK29 also retains one end of Tn*1696*-like. Hence, pSRC125 and pSRC26, appear to be in the ancestral configuration as their Tn*1696*-like Tns are complete and are flanked by direct repeats whereas in R478 and pK29 only part of the Tn remains.

Fig 3.9 Sequence alignment of pSRC26 and pK29. The colons represent base matches. The colour of the sequences represent the origin of sequence, where black is Tn*1696*, green is backbone and purple is IS*4321*. Regions are also named above or below and the arrows indicate the direction the region continues in. One copy of the 5 bp direct repeat identified in pSRC26 is bracketed and "DR" boxed above.

3.4 Structure and position of Tn*10*

As R478 contains Tn*10* and both pSRC125 and pSRC26 conferred tetracycline resistance due to the *tetA*(B) gene, they were tested for the presence of Tn*10*. Three PCRs that span Tn*10* were designed and were performed on pSRC26 and pSRC125, using R478 as a control (Fig 3.10A). R478, pSRC26 and pSRC125 produced amplicons of sizes 3.1 kb, 3.4 kb and 3.5 kb for the 3 PCRs (Fig 3.10B). Thus, Tn*10* was present in both pSRC26 and pSRC125, like R478.

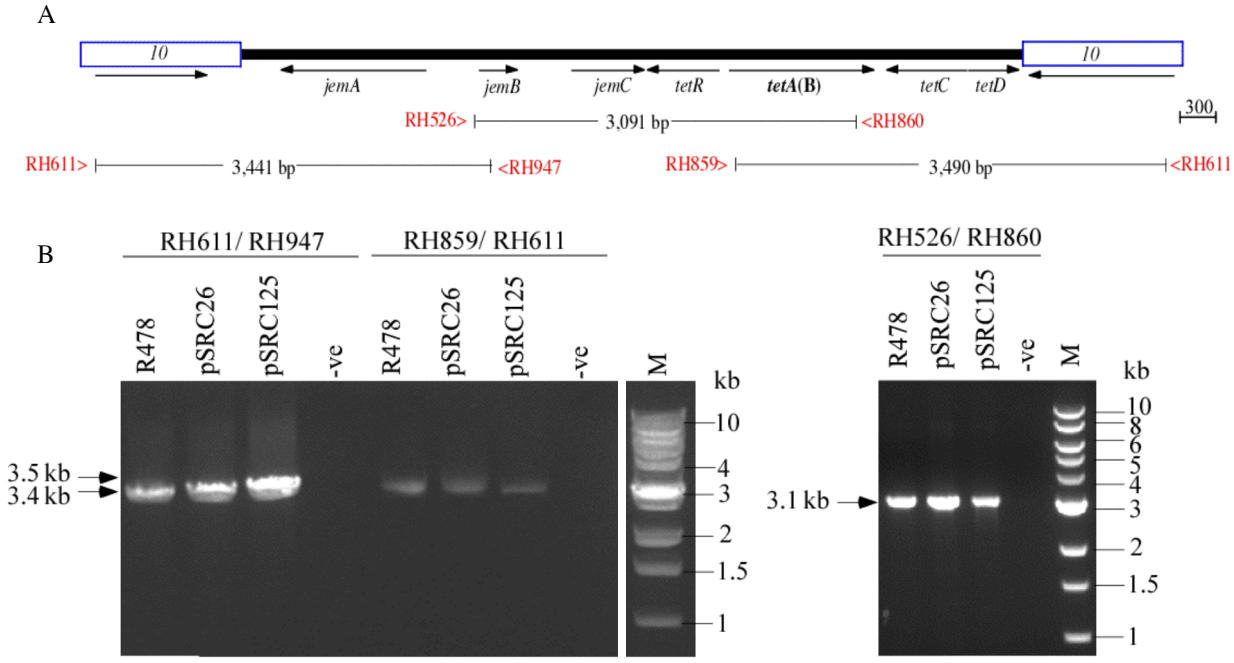


Fig 3.10 Tn*10* in pSRC26 and pSRC125. A Schematic of PCRs used to map Tn*10*. Primers are indicated as vertical lines and primer names are beside in red. The size of each PCR product is given between the 2 primers. Other features are as in Fig 3.2. **B** PCR gel results of Tn*10* in pSRC26 and pSRC125. The horizontal arrows indicate the position of the predicted amplicons. Lanes are labelled pSRC26 and pSRC125 for the plasmids, R478 for the positive control and –ve for the negative water control. 1 kb Molecular Weight Markers are loaded in lane M and selected sizes are as indicated. The sequence of primers are in Table A3.4 in Appendix A3.

To determine if pSRC125 and pSRC26 contained Tn*10* in the same position as in R478, primers were designed to match the sequence of the IncHI2 backbone to the left and right of Tn*10* in R478, and used with primers facing out of Tn*10* (Fig 3.11A). R478, pSRC125 and pSRC26 produced amplicons of 1.7 and 1.6 kb for the left and right boundaries, respectively (Fig 3.11B). Sequencing of all of these PCR products revealed that Tn*10* was in precisely the same position in all 3 plasmids. A 9 bp direct repeat was detected flanking Tn*10*, and this is characteristic of IS*10* transposition [242].

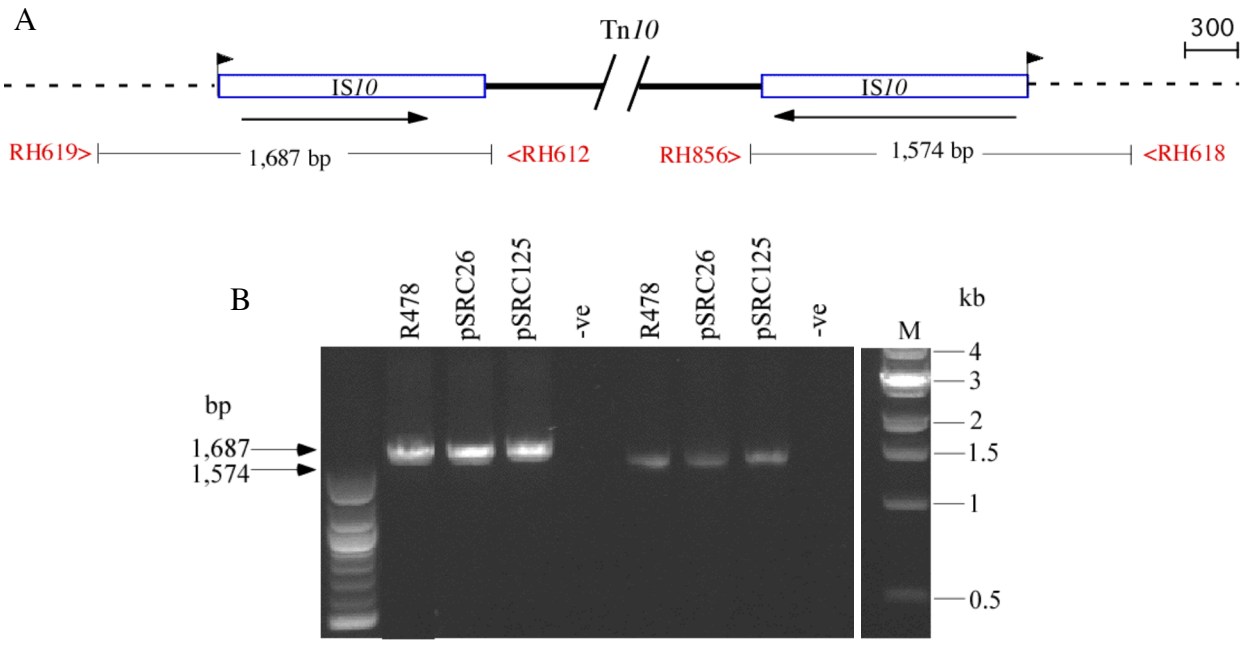


Fig 3.11 PCRs of the Tn*10* boundaries in IncHI2 plasmids. **A** Schematic of the boundary PCRs of the Tn*10* in R478. **B** Agarose gel picture of Tn*10* boundary PCR amplicons. The double slashed lines represent additional sequence and are not to scale. Features are as in Fig 3.10. The sequences of primers RH619, RH612, RH856 and RH618 are in Table A3.7 in Appendix A3.

Plasmid pK29 did not contain Tn*10* and, at the position where Tn*10* was inserted in pSRC26, pSRC125 and R478, only one copy of the 9 bp duplication was detected. Because Tn*10* was found to be in exactly the same position in the IncHI2 backbone for R478, pSRC26 and pSRC125, it is most likely that only a single insertion event occurred. Therefore, all four of these plasmids share a common ancestor. However because pSRC26, pSRC125, R478 and pK29 have a Tn*1696* relative or remnant in the same position, the insertion of Tn*10* may have occurred after Tn*1696* (discussed in 3.6.2).

3.5 Other resistance regions

*3.5.1 Tn*6029/*Tn*4352

In pSRC26, the bla_{TEM}, *sul2*, *strA/B* genes were previously linked together in a potential transposon containing 3 IS*26*, here designated Tn*6029* (Fig 3.12). Overlapping Tn*6029*, was the *aphA1*-containing transposon known as Tn*4352* (Fig 3.12). The IS*26*-bounded

Tn*6029*/Tn*4352* structure was mapped using overlapping PCR primers (shown in red in Fig 3.12) and sequenced. Because the location of this entity had been previously shown to be next to part of IS*6100* [232], it was proposed to be the large insertion mentioned in section 3.3.1 above. PCRs were used to link Tn*6029*/Tn*4352* (primer RH880 in *aphA1*) to *merE* (primer RH521; Fig 3.12). Linkage was established between *aphA1* and $merE_{1696}$ and confirmed between bla_{TEM} and IS*6100* (shown in blue in Fig 3.12). The sequences of these PCR products confirmed that the Tn*6029*/Tn*4352* entity had inserted 25 bp from the end of IS*6100*, within the In4-type integron (see Fig 3.12). The Tn*6029*/Tn*4352* structure was flanked by a direct duplication of 8 bp (shown as flags in Fig 3.12), a size characteristic of an IS*26* insertion.

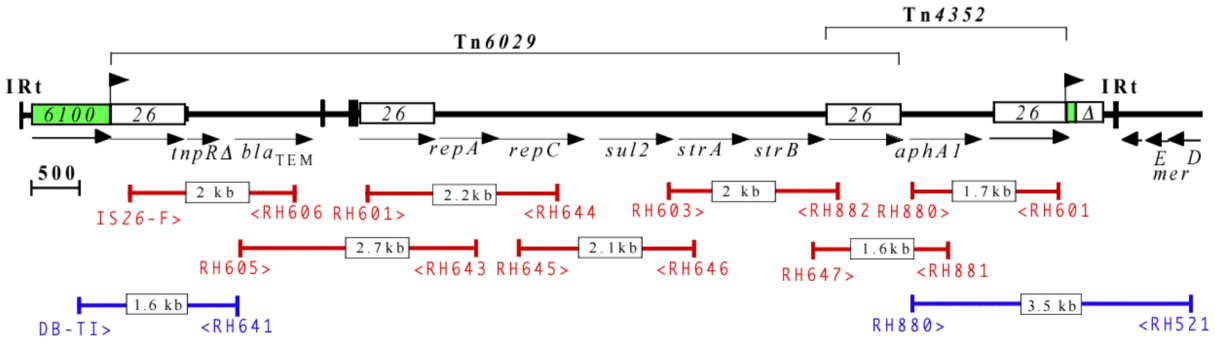


Fig 3.12 PCR mapping of Tn*6029* in pSRC26. Red lines indicate positions of primers and predicted products of PCRs used to map Tn*6029*. Blue lines represent primers and products linking out of Tn*6029*. The box representing the IS*6100* which is interrupted by Tn*6029* is coloured green. Other features are as in Figs 3.1, 3.3, 3.6 and 3.8. Sequences of the red primers are listed in Table A3.3 and of the blue primers in Table A3.5 in Appendix A3.

The segments between the copies of IS*26* in Tn*6029* are derived from two sources, namely Tn*2* ([243]; GenBank accession no. AY123253) and the IncQ plasmids RSF1010 [244] or pSRC15 ([114]; GenBank accession no. GQ379901), which contain *sul2* next to part of the *strA*/*B*-containing transposon Tn*5393*. The route by which Tn*6029* appears to have been generated is illustrated in Fig 3.13. It involves insertion of Tn*2* into RSF1010 and insertion of three copies of IS*26* into the resulting structure, followed by the inversion of one of the RSF1010-derived segments.

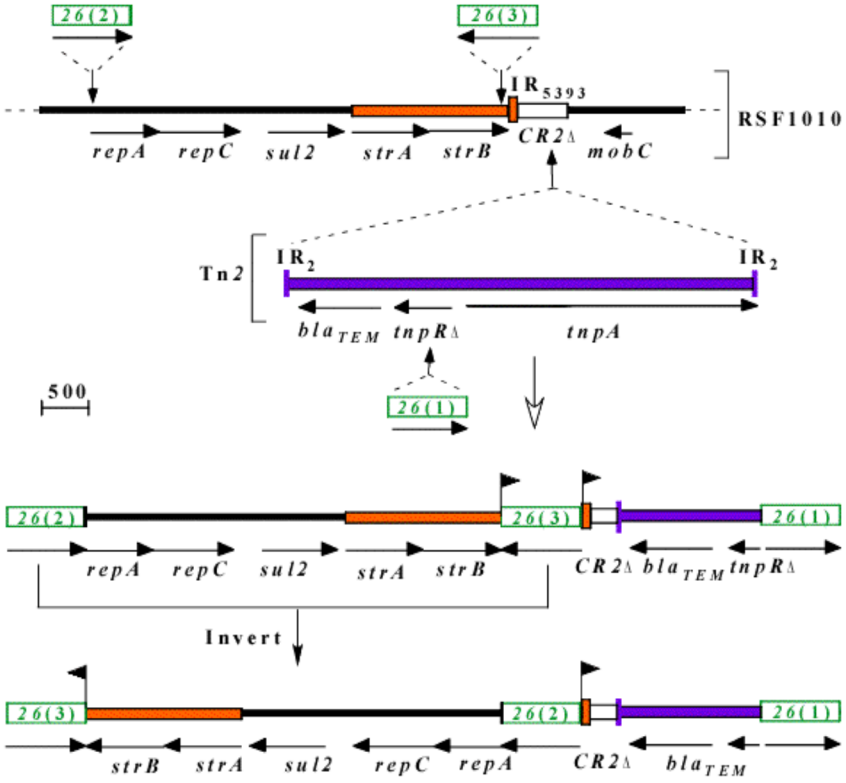


Fig 3.13 Origins of Tn*6029*. Colours represent regions of different origins, purple is Tn*2*, orange is Tn*5393* and IS*26* green. Other features are as in Figs 3.1, 3.3 and 3.8.

In the GenBank database a close relative of Tn*6029* was found, here named Tn*6029*B, within a Tn*21*-derived transposon in the IncHI1 plasmid pHCM1 (GenBank accession no. AL513383; see Table 3.3). Tn*6029*B has suffered an IS*26*-mediated deletion removing 85 bp in the Tn*2*-derived portion, directly adjacent to IS*26*(1), compared to all other published Tn*6029* sequences (Table 3.3). Intriguingly, a structure that is the same as the Tn*6029*/Tn*4352* entity in pSRC26, is found in precisely the same position as Tn*6029*B in pHCM1 in a closely related Tn*21*-derived transposon in the IncHI1 plasmid, pO111_1, from an enterohaemorrhagic *E. coli* strain (GenBank accession no. AP010961; Table 3.3). These two structures have potentially been generated in situ in the IncHI1 plasmids, by gain or loss of the Tn*4352* portion. Two further Tn*6029*/Tn*4352* copies that share the same location as each other and the IncHI1 plasmids on one side (*tniA*; Table 3.3), are found in an IncF plasmid, pRSB107, isolated from sewage (GenBank accession no. AJ851089) and an *E. coli*

virulence plasmid, pO26-CRL (GenBank accession no. GQ259888). However, none of these Tn*6029* or Tn*6029*/Tn*4352* elements are flanked by direct repeats.

Table 3.3 Locations of Tn*6029* and its relatives

Plasmid Name	Inc group	Organism	Structure	Accession no.	Reference
pSRC26	IncHI2	*S.* Typhimurium	IS*6100*-Tn*6029*/Tn*4352*[1]-IS*6100*	GQ150451	This study
PHCM1	IncHI1	*S.* Typhi	$tnpA_{21}\Delta$- Tn*6029*B[2]-*tniA*	AL513383	[190]
pO111_1	IncHI1	*E. coli*	$tnpA_{21}\Delta$- Tn*6029*/Tn*4352*-*tniA*	AP010961	[189]
pRSB107	IncF	Unknown[3]	*mph*(A)- Tn*6029*/Tn*4352*-*tniA*	AJ851089	[245]
pO26-CRL	IncF	*E. coli*	3'-CSΔ- Tn*6029*/Tn*4352*-*tniA*	GQ259888	[246]

[1]surrounded by an 8 bp duplication
[2]$tnpR_2\Delta 2$ has 85 bp deletion relative to others
[3]uncultured bacterium

3.5.2 aphA1 *and* strA/strB *in pSRC125*

pSRC125 also carries the *aphA1* (kanamycin/neomycin resistance) and *strA/B* (streptomycin resistance) genes in SRC125 (Table 3.1) and, although the *strA/B* resistance genes were detected in pSRC125, resistance to streptomycin was not expressed in the transconjugant (Table 3.2). Thus, only *aphA1* could be cloned, using kanamycin resistance for selection. A 5.9 kb SacI fragment of pSRC125 was cloned into vector pUC19. The clone obtained was named pRMH953 (Fig 3.14).

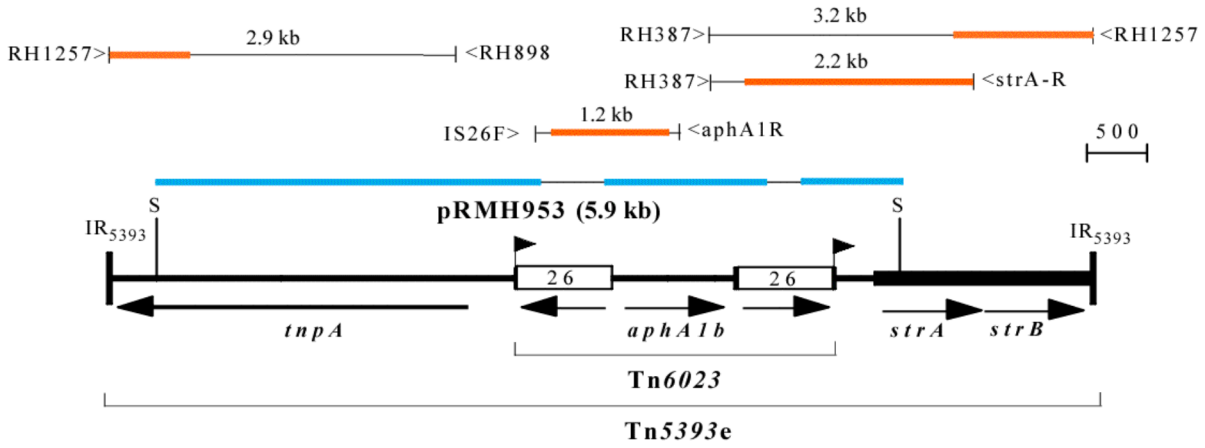


Fig 3.14 The *aphA1*/*strA*/*B*-containing region in pSRC125. The SacI sites are shown by a thin vertical line and an S. The PCR primers are thin vertical lines with the predicted size product between them. The flags show the position and direction of direct repeats. The sequence obtained is shown in orange for PCR products and blue for the clone. Other features are as in Fig 3.2 and 3.3. The sequences of primers are in Table A3.6 in Appendix A3.

The 5.9 kb insert in pRMH953 was sequenced using primers in the vector and *aphA1* (marked blue in Fig 3.14). The insert was found to contain *aphA1b* bounded by 2 IS*26* in opposite orientation and this structure was designated Tn*6023* (Fig 3.14). Tn*6023* was surrounded by an 8 bp duplication and had interrupted the *tnpR* gene from the *strA/B*-containing transposon Tn*5393c* (Fig 1.7; GenBank accession number M95402). pRMH953 also included part of *strA* and part of $tnpA_{5393}$, indicating the presence of a Tn*5393*c-derived structure, which Tn*6023* had interrupted. The products of PCR that linked *aphA1* to *strA* and *aphA1* to IS*26* (see Fig 3.14) were sequenced, in order to obtain sequence between the IS*26*, as pRMH953 contains 2 IS*26* and sequencing with primers in IS*26* produced double peaks. A primer, RH1257, was designed in the IR of Tn*5393* and used with primers in *aphA1* and $tnpA_{5393}$ to obtain sequence of the remainder of Tn*5393*c-like structure, up to the IRs (sequenced regions coloured in Fig 3.14). This structure was designated Tn*5393*e, to distinguish it from Tn*5393*c (analysed in section 3.5.2.2). A further 11.3 kb of sequence, which included Tn*5393*e, was added to the GenBank entry of pSRC125 containing Tn*6025* (GenBank accession no. GU562437; see section 4.4.3).

*3.5.2.1 Analysis of Tn*6023

The sequence of Tn*6023* displayed greater than 99.9% identity to 2 other sequences, each of which contained Tn*6023* in the exact same position in Tn*5393*. They were both from *Corynebacterium* species, the *Corynebacterium urealyticum* genome DSM7109 (GenBank accession number AM942444) and *Corynebacterium resistens* plasmid pJA144188 (GenBank accession number FN825254). *Corynebacterium striatum* plasmid pTP10 (GenBank accession number AF024666) also contained a related structure. However, the Tn*5393e* derivative in pTP10 contained the insertion sequence IS*1250* within *strB* and harboured a deletion of 51 bp within Tn*6023*. This deletion consisted of 26 bp missing from the LH IS*26* and 25 bp from the central piece, just before the start of *aphA1b* (see Fig 3.15). This variant version of Tn*6023* in pTP10 had previously been named Tn*5715* [247].

The possible promoter for the *aphA1b* in Tn*6023* was identified, by searching for matches to the –35 and –10 promoter consensus sequences of "TTGACA" and "TATAAT" (where underlined letters are most conserved; [248]). The putative promoter was a hybrid one, where the –35 region was provided by the IR of the upstream IS*26* (underlined in Fig 3.15). The capability of IS*26* to form efficient hybrid promoters for adjacent genes has previously been described [100];[101]. The promoter of Tn*6023* is likely to be a particularly strong one, due to its high identity to the consensus and the presence of a TGN sequence upstream of the –10 region (overlined in Fig 3.15), which has been shown to increase promoter efficiency [249]. The 51 bp deletion in Tn*6023* in pTP10 (see Fig 3.15) clearly removes the promoter. However, pTP10 still confers kanamycin resistance [247], thus another *aphA1b* promoter must be present, which was not able to be detected, or another kanamycin resistance gene.

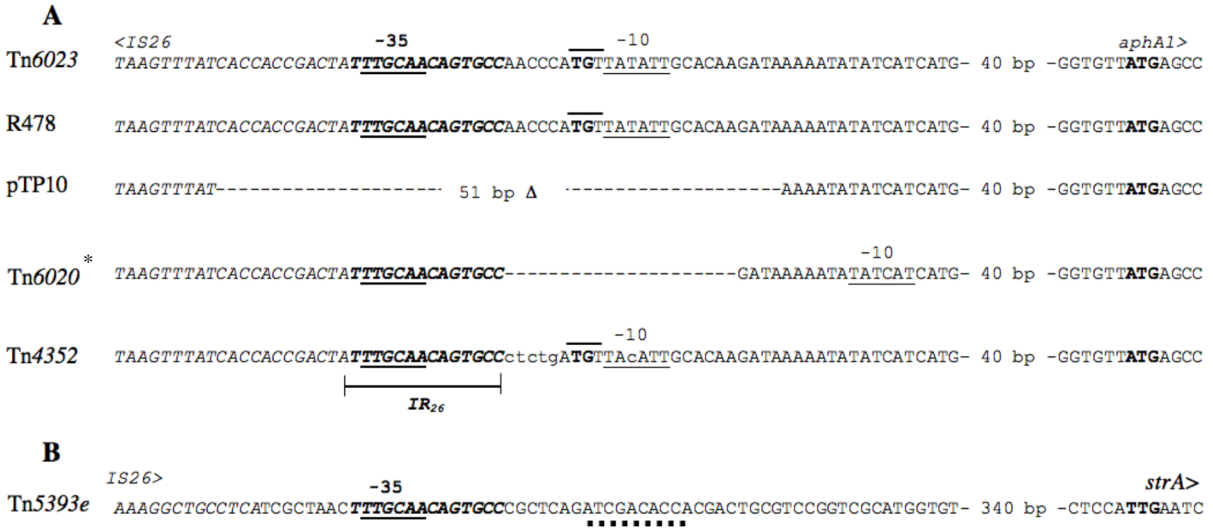


Fig 3.15 Putative promoter regions involving IS*26* A Sequence alignments of *aphA1* promoter regions. The IS*26* sequence is in italics and the IR_{26} is bold and bracketed below. The *aphA1* start codon is in bold face type. The -35 region and -10 region are underlined and named above. The extended -10 motif is overlined and the effective "TG" sequence is in bold. The asterisk indicates this is the structure that has been experimentally proven. 40 bp in between the -10 regions and the *aphA1* start codon is not shown. **B** Lack of a *strA* promoter with the -35 hexamer from IS*26* in Tn*5393*e. The dotted line indicates where the -10 region would be.

The Tn*6023* promoter was compared to a similar *aphA1b*-containing transposon named Tn*6020,* where the promoter had been experimentally determined, in plasmid pBWH77 ([101]; Fig 3.15). It also contained a hybrid promoter with the same -35 region from IS*26*, but

a different -10 motif to Tn*6023*. An additional 21 bp were present in Tn*6023* compared to Tn*6020* adjacent to IS*26*. A similar promoter was also found in the *aphA1a*-containing transposon Tn*4352* (present in pSRC26; Fig 3.16), which has 6 bp differences compared to Tn*6023* (lower case font in Fig 3.15) but retains the same hybrid –35 region and an almost identical extended –10 motif. Other sequences in GenBank were also examined. R478 contains *aphA1b* in an IS*26*-bounded structure, where the IS*26* to the left of *aphA1b* is in the same position and orientation as in Tn*6023* (Fig 3.16) and thus it contains the same hybrid promoter (Fig 3.16). However, the *aphA1b*-containing central piece is longer than in Tn*6023* and the IS*26* are in direct orientation. This configuration is interrupted by an IS*26*-bounded structure containing *catA1* (Fig 3.16).

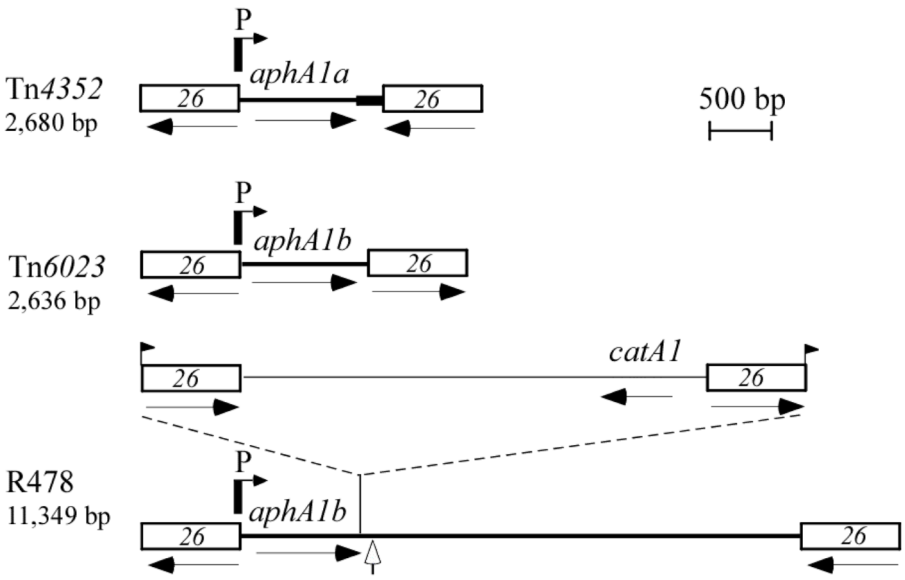


Fig 3.16 IS*26*-bounded transposons containing *aphA1*. The bent arrow and "P" represent the hybrid promoter for *aphA1*. Vertical open headed arrow marks where Tn*6023* matches up to in R478. Lines of different thickness represent segments of different origin. Other features are as in Figs 3.2, 3.3 and 3.8.

3.5.2.2 Tn5393 and its variants

Tn*5393*c is the simplest form of Tn*5393* transposon as it contains only the *strA*/*B* and *tnpA*/*R* genes (Fig 3.17A). However, there are several forms of this transposon. Tn*5393*a contains the insertion sequence IS*1133* directly after the stop codon of *tnpR*, Tn*5393b* contains IS*6100* positioned 26 bp within *tnpR* and finally Tn*5393*e contains Tn*6023* 225 bp into *tnpR* (Fig

3.17B). Tn*5393*d has a complex structure and is not analysed here. Three further derivatives of Tn*5393* were identified in the present study, by examining the GenBank database. These were hereby named Tn*5393*f, Tn*5393*g, and Tn*5393*h and are shown in Fig 3.17D. Tn*5393*f consists of an IS*26* within the *tnpA* and was found in one entry (GenBank accession no. FJ012881). Tn*5393*g consists of a complex structure where an *aacC3*-contining structure flanked by ISKpn12, has interrupted an IS related to ISKpn11 (the amino acid sequence encoded by *tnpA* is ~63% identical), which has inserted just before the start of *strA*. Tn*5393*g was found in one entry (GenBank accession no. FJ012882). Tn*5393*f and Tn*5393*g had been previously been described, but not named [250]. Tn*5393*h is equivalent to Tn*5393*a with Tn*10* within *tnpA* [158]. Although there are 3 incidences of it in the GenBank database (accession nos CP002090, CP002089, CP001122), they are all in the same position, thus Tn*5393*h has only been seen in one position.

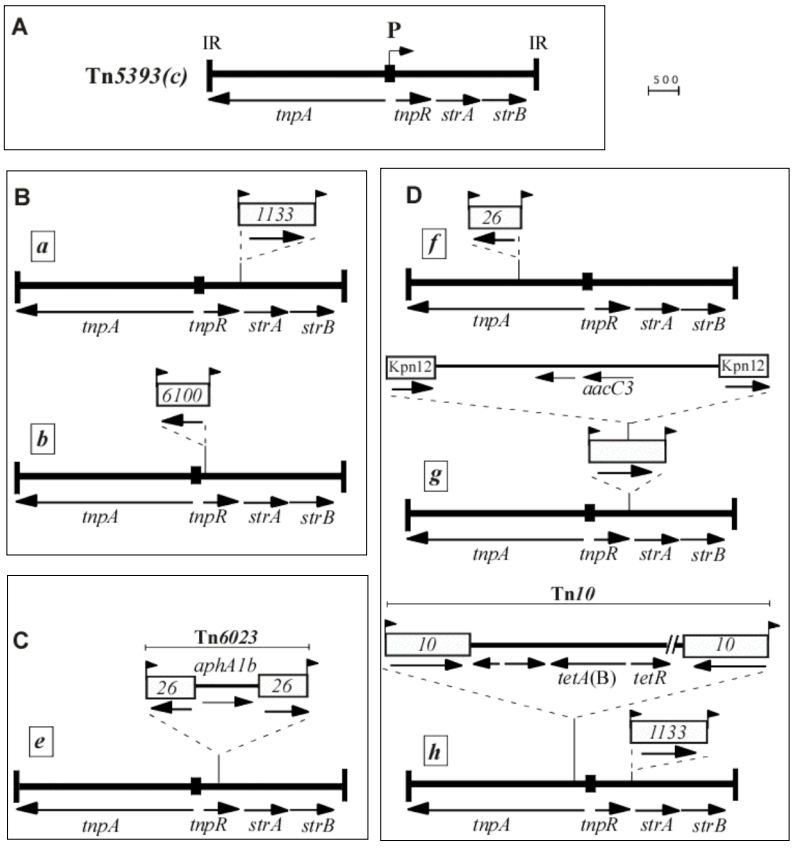


Fig 3.17 Tn*5393*-related transposons. A Tn*5393*c **B** Tn*5393*a and Tn*5393*b **C** Tn*5393*e **D.** Tn*5393*f, Tn*5393*g and Tn*5393*h. Figure drawn to scale based on GenBank accession numbers AF313472, HM371194, CU928144, GU562437 (pSRC125), FJ012881, FJ012882 and CP002090 for Tn*5393*a-h respectively. All other features are as in Figs 3.2, 3.3 and 3.8.

3.5.2.3 Why does Tn5393e not confer streptomycin resistance?

Disc diffusion, using streptomycin antibiotic discs, was performed on different *strA/B*-containing structures, to give insight into the level of streptomycin resistance that each confers. The annular radius of the inhibition zones was between 11-13 mm for both the *E. coli* strain (E294) and the E294/pSRC125 transconjugants (Table 3.2), indicating that Tn*5393*e did not confer streptomycin resistance. Whereas, Tn*5393*a (Fig 3.17) which was tested in the IncI1 plasmid pSRC27-I, in E294 (described in Chapter 6) and Tn*6029*, which was tested in pSRC26, in E294, both displayed an inhibition zone size of zero and thus conferred streptomycin resistance. Also, Tn*5393*c was tested in pRMH536, a pUC19 clone containing a 9.3 kb EcoRI fragment of Tn*1403,* in DH5α and also displayed an inhibition zone of zero. However, these numbers are not directly comparable, because pUC19 based clones have a high copy number and pRMH536 was in a different host *E. coli* strain. Interestingly, the pSRC125-containing *Salmonella* strain SRC125 had a zone size of 0.5-2 mm and so was considered resistant, yet it contained Tn*5393*e. This could only partly be due to the intrinsic streptomycin resistance displayed by *S.* Typhimurium strains, because sensitive control strains (in which streptomycin resistance genes could not be detected) have a zone size of 3-5.5 mm (Hall laboratory, unpublished observations). Thus, this resistance may be due to another copy of a streptomycin resistance gene in the SRC125 chromosome.

A previous study had shown that in Tn*5393c* the *strA/B* genes do not have their own promoter [251], but instead, are co-transcribed with *tnpR* from a promoter in the *res* site next to *tnpR* (Fig 3.17A). In Tn*5393*a (Fig 3.17B), the expression increased and it was speculated that IS*1133* provides a strong promoter for the resistance genes [252]. However, in Tn*5393*b (Fig 3.17B), where expression is also increased, it was shown that the promoter was not provided by IS*6100* [41], thus expression was still driven by the promoter near *tnpR*, as in Tn*5393*c. Because the TnpR protein regulates its own transcription, it is likely that when the IS*6100*

interrupted the *tnpR* gene, transcription can no longer be repressed and expression of *strA/B* is increased. Although Tn*6023* in Tn*5393*e (Fig 6.17D) has also interrupted the *tnpR*, it does not confer streptomycin resistance. This suggests that the *tnpR* promoter is not effective over the large size of the Tn*6023* insertion. Also, the IS*26* in Tn*6023* upstream of *strA* had not formed part of a promoter for *strA/B* as no suitable -10 region could be identified 16-19 bp downstream of the -35 motif in the IR_{26} (shown as a dotted line in Fig 3.15B). Likewise, the promoter for *aphA1b* must fail to reach *strA/B*. Thus the transcript must terminate before the start of *strA*.

3.6 Discussion and summary

*3.6.1 Tn*1696 *evolution*

Class 2 mercury (II) resistance transposons related to Tn*1696* are not often found (Table 3.4). In addition to the complete Tn*1696*-relatives analysed here, namely Tn*6026* and Tn*6025,* the only other intact relative of Tn*1696* found in GenBank was the transposon in pHCM1 (Fig 3.5). Tn*6005* (GenBank accession no. EU591509) has a complete Tn*5036* backbone, and an In4-type integron, but represents a separate transposon lineage because the integron is positioned 16 bp to the left of where In4 is in Tn*1696* [87]. A path detailing how these Tn*1696*-like plasmids have evolved is shown in Fig 3.18.

Table 3.4 Transposons related to Tn*1696* in GenBank[1]

GenBank Ac#	Organism	Plasmid/Tn	In4-like	mer_{1696}	Date Published	Reference
Y09025	*E. cloacae*	Pl/Tn*5036*	-	Y	1997	[241]
U12338	*P. aeruginosa*	R1033/Tn*1696*	+	Y	2001	[67]
GQ150541	*S.* Typhimurium	pSRC26/Tn*6026*	+	Y	2010	This study
GU562437	*S.* Typhimurium	pSRC125/Tn*6025*	+	Y	2010	This study
AL513383	*S.* Typhi	pHCM1[1]/Tn	P[2]	Y	2001	[190]
EU591509	*E. cloacae*	Tn*6005*	+[3]	Y	2008	[87]

[1]IS*4321* in IR_{tnp}
[2]Partial; missing the 3'-CS
[3]In4-type integron inserted 6 bp to the left of In4 in Tn*1696*

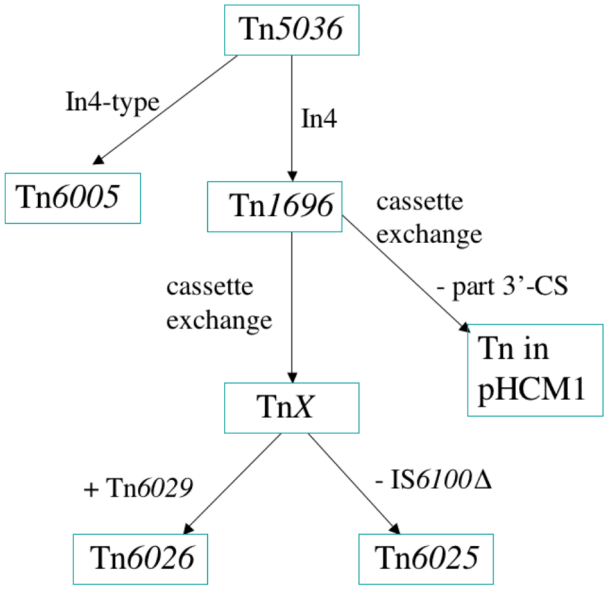


Fig 3.18 Evolution of Tn*1696*. Transposon names are boxed in green, the arrows represent evolutionary events that are named next to the arrows. The "+" means the region was gained, "-" means it was lost. The IS*6100*Δ is the second, partial copy of IS*6100* in In4. This figure is not drawn to scale and the length of each arrow is not indicative of relatedness.

*3.6.2 Are Tn*6029*, Tn*4352 *or Tn*6029*/Tn*4352 *transposons?*

The 8 bp direct repeat found flanking the Tn*4352*/Tn*6029* element in pSRC26 is consistent with a transposition event of the whole entity. However, Tn*4352* and Tn*6029* share one copy of IS*26* and have been seen individually, elsewhere. A variant of Tn*6029* has been identified by itself (Table 3.3) and complete examples of Tn*6029*, have also been seen (Hall Laboratory, unpublished observations). Tn*4352* has also been described alone, surrounded by 8 bp direct repeats, (see for example [243]: GenBank accession no. AY123252). The fact that they are seen individually, but also overlapping and surrounded by direct repeats indicates that their movement is not via a standard transposition mechanism. Standard transposition would bring in both copies of the IS*26* bounding the incoming entity (either Tn*4352* or Tn*6029*). This would result in the presence of an additional IS*26*, as well as create a duplication flanking that entity. One way the Tn*6029*/Tn*4352* entity could form is via homologous recombination between the IS*26* copies, or perhaps it is a specific feature of this particular IS, the

transposition of which has not been studied. A further direction of this work would include transposition assays of IS*26* and IS*26*-flanked entities such as Tn*4352* and Tn*6029*.

3.6.3 IncHI2 plasmid evolution

The IncHI2 plasmids R478, pK29, pSRC26 and pSRC125 share a common ancestor as they all contain a Tn*1696*-like transposon or part thereof, in the same position. Is it clear that the configuration in pSRC26 and pSRC125 is ancestral to R478 and pK29, as their Tn*1696*–family Tns are complete and flanked by a 5 bp direct repeat (Fig 3.19A), whereas those in R478 and pK29 have been truncated by insertion and deletion events whilst within the IncHI2 backbone. R478 is missing 5.3 kb of backbone to the left of the IR_{tnp} in pSRC26 and pSRC125 (Fig 3.19A) which is replaced by a unique, IS*26*-bounded region containing the *aphA1b* and *catA1* resistance genes (Fig 1.23). This suggests that an IS*26*-mediated adjacent deletion has removed the segment in R478. pK29 retained this backbone segment, but lacks 34.7 kb at the other end. In pK29, a large MARR containing 3 IS*26* and the resistance genes bla_{CMY-8}, *catBR2*, *aadA2* and 2 copies of *sul1*, has displaced this sequence (Fig 3.19A). Interestingly, previous analysis of pK29 [209] had defined the resistance region, named "Region 1" (Fig 1.24), as spanning from the bla_{CMY-8}-containing MARR to where it began to match R478 (Fig 3.19A), assuming that R478 included backbone sequence and pK29 contained an insertion. However, from the analysis undertaken here, R478 had suffered a deletion and the segment to the left of IR_{1696} in pK29 is, in fact, IncHI2 backbone sequence. Hence, it is evident that IS*26* is a major force driving the evolution of regions surrounding the Tn*1696*-like transposons in IncHI2 plasmids (Fig 3.19A).

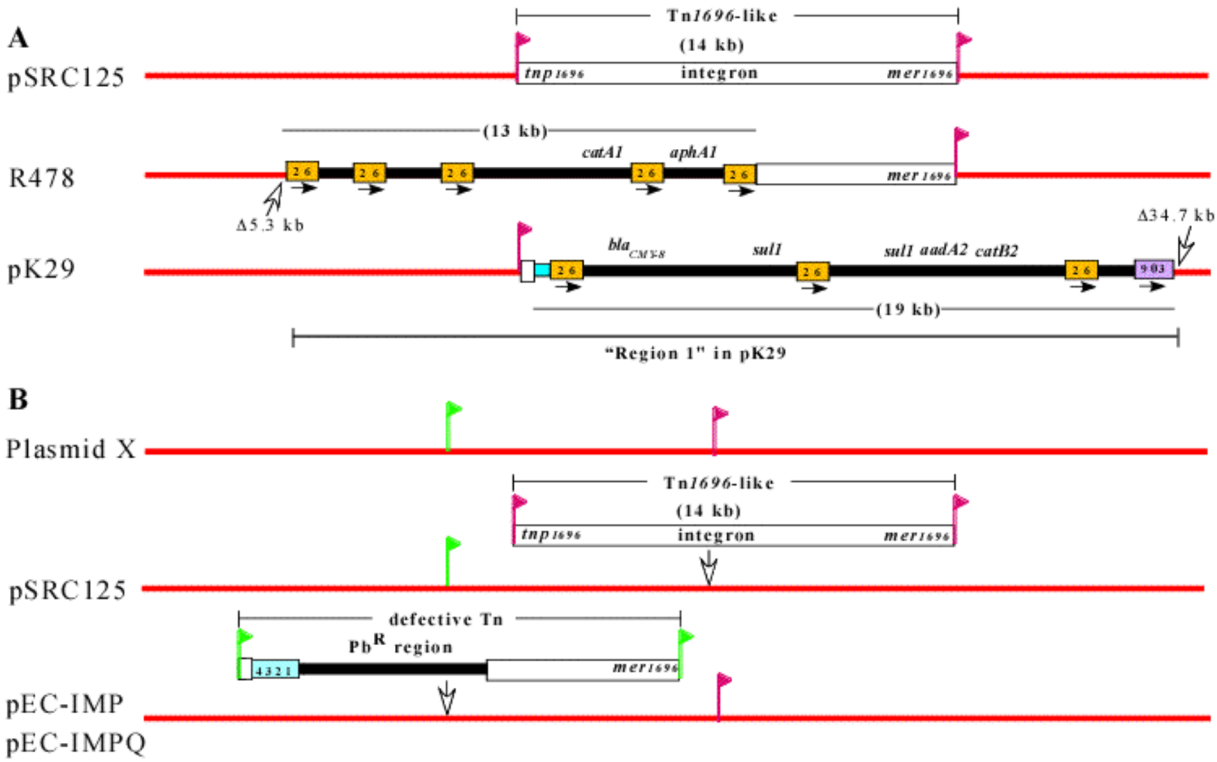


Fig 3.19 Schematic showing the backbone surrounding relatives of Tn*1696*. Configurations in **A** the theoretical IncHI2 ancestor, pSRC125, R478, pK29 and **B** pEC-IMP and pEC-IMP-Q. The red horizontal line is the backbone sequence. The Tn*1696*-like transposon is shown by an open white box and the pink flags represent the position and direction of its direct repeats. The black line shows the additional DNA not in other IncHI2s with the size shown below or above. Yellow boxes represent IS*26* and their direction is reflected by the arrows below, cyan represents IS*4321* or part of it and purple IS*903*. The extent of the backbone deletions in R478 and pK29 is indicated by the dotted line. The previously defined resistance region in pK29 "Region 1" is bracketed below. This figure is not drawn to scale.

Two nearly identical IncHI2 plasmid sequences, pEC-IMP and pEC-IMPQ became available during the course of the current study (GenBank accession nos. EU855787 and EU855788 respectively; [209]). Like pK29, neither contained Tn*10*, and at the site of Tn*10* insertion in R478, pSRC26 and pSRC125, one copy of the direct repeat was detected. Although both harbour most of the *mer* module (only 25 bp from the integron end is missing) and both IRs of Tn*1696*, this structure is located 2.3 kb to the left of where Tn*6025* and Tn*6026* are found in pSRC125 and pSRC26 (Fig 3.19B). At the positions where the transposons lie in pSRC125 and pSRC26, the uninterrupted backbone sequence was found (marked red in Fig 3.19B). Thus there was an ancestral plasmid that did not contain a Tn*1696* relative in either position

(Fig 3.19B). Interestingly, the integron and *tnp* module are replaced by a lead resistance region (see Fig 3.19B) making the transposon defective, referred to here as TnPb. To the right of IR_{tnp}, which is interrupted by an IS*4321*, only 7 bp of the *tnp* module remains. Furthermore, this defective transposon is surrounded by a 5 bp direct repeat (shown as green flags in Fig 3.19) indicating that evolution (namely the addition of the IS*4321* and the Pb^R region) occurred within TnPb, once it was within the IncHI2 backbone. These 2 plasmids represent a different evolutionary lineage of MAR IncHI2 plasmids, as a separate insertion event of a Tn*5036*-derived element has occurred (Fig 3.19). Clearly, pSRC26 and pSRC125 belong to the same lineage as R478, as they share a number of features (Table 3.5).

Table 3.5 Summary of IncHI2 plasmid properties

Plasmid	Organism	Country of origin	Isolation year	Tn*1696*-like	Tn position	Tn*10*[3]	Lineage
PSRC26	*S.* Typhimurium[1]	Australia	1999	+	1	+	1
PSRC125	*S.* Typhimurium[1]	Australia	2000	+	1	+	1
R478	*S. marcescens*[2]	USA	1969	P[4]	1	+	1
pK29	*K. pneumoniae*[2]	Taiwan	2001	P	1	-	1a
pEC-IMP	*E. cloacae*[2]	Taiwan	2004	P	2	-	2
pEC-IMPQ	*E. cloacae*[2]	Taiwan	2004	P	2	-	2

[1]Isolated from a bovine infection
[2]Isolated from a human clinical infection
[3]In the same position within the IncHI2 backbone
[4]P= partial

An IncHI2 progenitor, that does not contain antibiotic resistance genes, or the Tn*10* and Tn*1696*-like transposons (in either location; as in Fig 3.19) is designated "Plasmid X" (Fig 3.20; Fig 3.19B). The possible order in which resistance regions have been gained by the IncHI2 plasmids was proposed as shown in Fig 3.20. For example, although all the IncHI2 plasmids from lineage 1 contain all or some of a Tn*1696*-derived region in the same position, Tn*10* is present (in the same position) only in R478, pSRC125 and pSRC26, indicating that they share a more recent common ancestor than with pK29 (designated as belonging to lineage 1a in Table 3.3). Thus it is likely that the Tn*10* insertion event thus occurred after that of Tn*1696*-like Tn (Fig 3.20). However, the order of some events is designated arbitrarily,

such as the addition of other resistance genes, and other forces that shape plasmid evolution, such as homologous recombination, were not considered here.

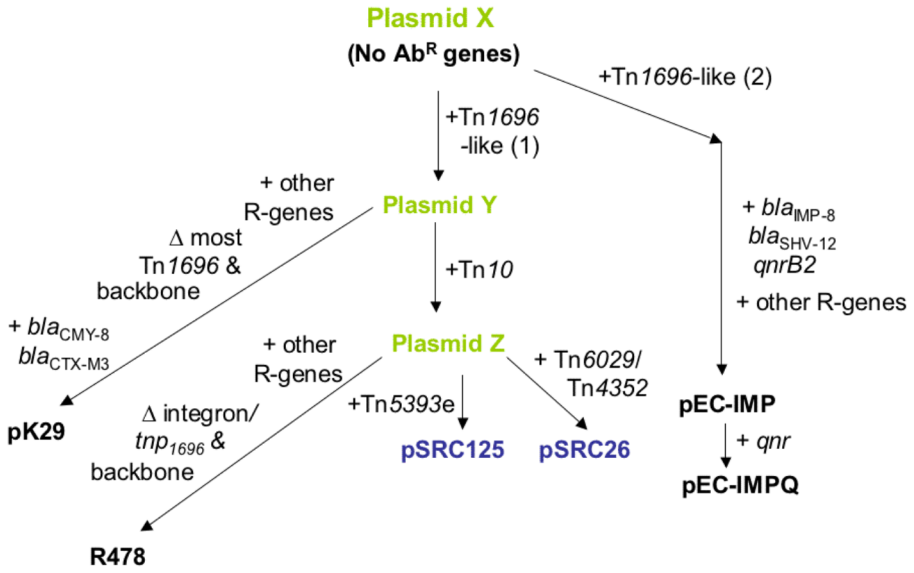


Fig 3.20 Evolution of resistance in IncHI2 plasmids. Sequenced plasmids are named in bold-face type, pSRC125 and pSRC26 are highlighted in blue and theoretical plasmid ancestors in green. Positions of additional regions are noted in brackets, for example (1) denotes position 1 and (2) position 2. Other features are as in Fig 3.19.

The PCRs designed in the current study that test the position of Tn*1696*-like and Tn*10* in the IncHI2 backbone, as in pSRC125 and pSRC26, provide a simple method to identify IncHI2 plasmids from the R478-type evolutionary lineage. IncHI2 plasmids have been isolated from a number of sources and are spreading globally (Table 3.5). Because pSRC26 and pSRC125, which were isolated in circa 2000, are closely related to R478, which was isolated in 1969, this indicates that these IncHI2 plasmids have been circulating, originating from a single clonal type, for more than 30 years. Furthermore, plasmids within the same lineage, have been isolated from animal and human sources (Table 3.5) and this illustrates the flow of resistance genes between animals and humans.

CHAPTER FOUR

Mapping pSRC125 and pSRC26

4.1 Introduction

In Chapter 3, the structures of resistance regions in IncHI2 plasmids pSRC125 and pSRC26 were determined. Both plasmids were found to contain Tn*10* in the same position as R478. pSRC125 and pSRC26 also carried MARRs belonging to the Tn*1696* family, which conferred resistance to mercuric ions, and were in the same position as in R478. Regions conferring resistance to the heavy metals tellurite, copper, arsenic and silver have previously been identified in R478 [58]. The copper, arsenic and silver resistance regions are contained within a 36.6 kb segment to the left of mer_{1696}, whereas the tellurite resistance region is separate, to the right of the R478 MARR (see Fig 1.22). Previous studies [161, 193] used a set of 10 PCRs (the positions of which are marked with asterisk in Fig 1.22) to analyse and map 11 IncHI2 plasmids. Variation was observed in 5 out of 10 PCRs, all of which were positioned within the segment containing the heavy metal resistance regions (marked with red asterisk in Fig 1.22).

The aim of the work described in this chapter, was to determine whether pSRC125 and pSRC26 share features with R478 beyond the antibiotic resistance regions. Also, to map the whole of pSRC125 and pSRC26 and to examine their relationship to other sequenced IncHI2 plasmids. This will enable the IncHI2 backbone to be more clearly defined and will determine whether further MGEs are present. From this, the structure of an IncHI2 ancestor can be extrapolated.

4.2 Mapping pSRC125 and pSRC26

4.2.1 PCR with published primers

To determine which regions were present surrounding the MARRs in pSRC125 and pSRC26, four published PCRs that showed variation (Fig 1.22) in the plasmids tested by Garcia-Fernandez et al., (2007) [161], were used (PCRs 3-6 in Fig 4.1). These targeted the arsenic resistance region, 2 uncharacterised orfs numbered 201 and 92/93, and a gene named "*tnsD*", which encodes a protein that shares 26% identity with TnsD encoded by Tn*7* (Fig 4.1). One

primer from PCR 6 (orf201 in Fig 4.1) was redesigned (RH877; Table A3.8 in Appendix 3) so that the 2 primers had a more similar melting temperature. A 5th PCR that displayed variation from the previously used set [161] was not used here, as it was positioned only 80 bp from the primers in *tnsD*. A PCR targeting the tellurite resistance region was also included (PCR 2 in Fig 4.1) even though this region was present in all plasmids tested previously [161]. The IncHI2 PCR targeting the *repHI2* replicon (PCR 1 in Fig 4.1) is from the PBRT set used in Carattoli et al., (2005) [145] and had previously shown to be positive for pSRC125 and pSRC26 (Fig 3.1).

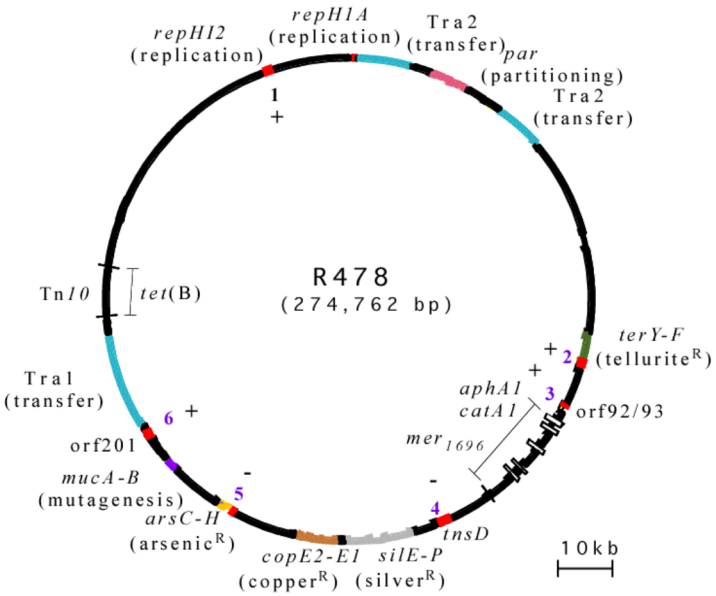


Fig 4.1 Analysis of the IncHI2 backbone using previously designed primers. The positions of primers are marked in red in the circle representing R478 sequence, some features of which are shown. PCRs 2-6, numbered with purple lettering are from [161] and PCR 1 is from [145]. The + or – next to the PCR numbers indicate whether the PCR was positive or negative in pSRC125 and pSRC26.

Table 4.1 Backbone PCRs with published primers[1]

Plasmid	*repHIA* (647 bp) 1.	Tellurium^R (893 bp) 2 .	ORF 92/93 (1,001 bp) 3.	*tnsD* (1,466 bp) 4.	Arsenic^R (1,136 bp) 5.	ORF 201 (1,011 bp) 6. [2]
R478	+	+	+	+	+	+
pSRC125	+	+	+	-	-	+
pSRC26	+	+	+	-	-	+
pK29[3]	+	+	+	-	+	+
pEC-IMP/Q[3,4]	+	+	+	-	+	+

[1]from [145]; all other PCRs from [161]
[2]Modified from original primer set -1 primer was redesigned in this study
[3]Results predicted for pK29, pEC-IMP using GenBank accession numbers EF382672, EU855787
[4]PCR results representing both pEC-IMP and pEC-IMPQ as they were identical across their backbone

When pSRC125 and pSRC26 were tested, PCRs 4 and 5, in *tnsD* and the arsenic resistance region, did not produce an amplicon (highlighted blue in Table 4.1). The remaining PCRs

yielded products of correct size (Table 4.1). This indicated that both these plasmids were missing a region that includes the arsenic resistance region and the *tnsD* gene.

4.2.2 Extending the PCR set

Seven PCRs that amplify segments of 0.5–3 kb and evenly span the remainder of R478 were designed (GenBank accession number BX664015; PCRs 7-9, 11-14 in Fig 4.2). A set of primers was also designed in the *hipA* gene (PCR 10 in Fig 4.2), which lies in the additional 5.4 kb segment present in pK29 (GenBank accession no EF382672) which is missing in R478, located adjacent to the shared Tn*1696*-derived MARR (see Fig 3.19). In order to determine if the entire region between *ars* and *tnsD* was missing, PCRs 11 and 12 were designed to target the copper and silver resistance regions, respectively. The set of PCRs was used on pSRC125 and pSRC26, with R478 as a control and the results of PCRs 1-14 are tabulated in Table 4.2. The results for pSRC125 and pSRC26 are also shown as + or – in Fig 4.2. A region of at least 42 kb (Fig 4.3) containing the silver, copper and arsenate/arsenite resistance regions appears to be missing from pSRC125 and pSRC26 in between the Tn*1696*-like Tns and the *mucA/B* genes.

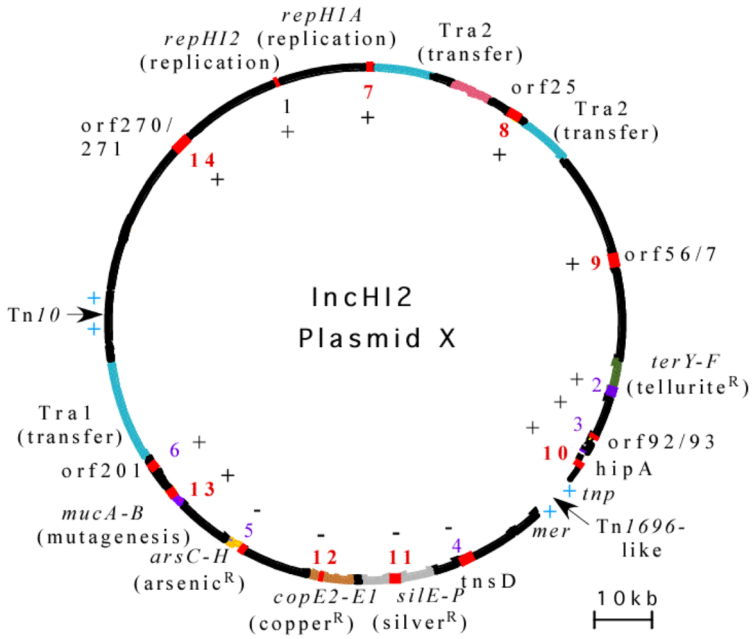


Fig 4.2 PCRs to map IncHI2 plasmids. The circular backbone of an IncHI2 plasmid, based on the R478 sequence (GenBank accession number BX664015), with the 5.4 kb deletion in R478, including the *hipA* gene, added from the pK29 sequence (GenBank accession no EF382672) and the resistance regions removed. The position of PCR primers in the IncHI2 backbone designed in this study are shown as red numbers. The blue +s represent results for PCRs across the resistance regions, as determined in sections 3.3 and 3.4. Other features are as in Fig 4.1. Primers are in Table A3.8 in Appendix A3.

Table 4.2 PCRs to map IncHI2 plasmids

PCR target	*repHIA*	ORF25	ORF56/57	Tel^R	ORF92/93	*hipA*	*tnsD*	Sil^R	Cu^R	Ars^R	mucA	ORF201	ORF270/1	*repHI2*
size (bp)	(1150)	(2421)	(2217)	(893)	(1001)	(1553)	(1466)	(1704)	(512)	(1136)	(1276)	(1011)	(2966)	(647)
number[1]	7.	8.	9.	2.	3.	10.	4.	11.	12.	5.	13.	6. [2]	14.	1.
R478	+	+	+	+	+	-	+	+	+	+	+	+	+	+
pSRC125	+	+	+	+	+	+	-	-	-	-	+	+	+	+
pSRC26	+	+	+	+	+	+	-	-	-	-	+	+	+	+
R478[3]	+	+	+	+	+	-	+	+	+	+	+	+	+	+
pK29[3]	+	+	+	+	+	+	-	-	+	+	+	+	+	+
pEC-IMP/Q[3,4]	+	+	+	+	+	+	-	-	+	+	+	+	+	+
pAPEC-O1-R[3]	+	+	+	+	-	+	+[5]	+	+	-	+	-	+	+

[1]PCRs 2-6 are from [161] and PCRs 7-14 were designed in this study. PCR 1 is used in PBRT, from [145]

[3]Results predicted for R478, pK29, pEC-IMP, pEC-IMPQ and pAPEC-O1-R using GenBank accession nos BX664015, EF382672, EU855787, EU855788 and DQ517526

[4]PCR results representing both pEC-IMP and pEC-IMPQ as they are identical across their backbone

[5][161] identified this PCR as being negative from the GenBank sequence, but analysis in this study shows that it is positive

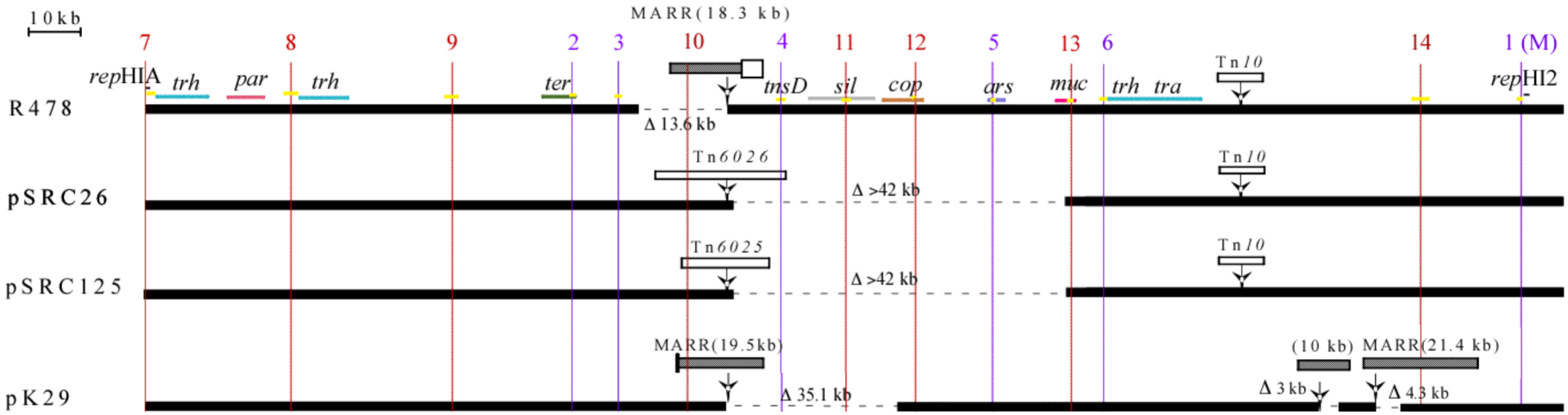


Fig 4.3 Regions present in IncHI2 plasmids. Structures of R478, pSRC26, pSRC125, pK29 and pEC-IMP are shown as a horizontal line, opened at position 1 in R478 (GenBank accession no. BX664015) with resistance regions shown above. Vertical lines indicate the positions of PCRs numbered 1-14, the ones designed in this study are in red and those from previous studies in purple. Dotted lines represent missing or potentially missing segments based on negative PCR results. Insertions and MGE not found in the other plasmids are shown above the backbone as boxes.

4.2.3 Resistance to heavy metals

From the PCR results in Table 4.2 it was expected that pSRC125 and pSRC26 would confer resistance to tellurite but not to copper, silver or arsenic. Colonies of the *Salmonella* strains SRC26 and SRC125 and the *E. coli* transconjugants containing pSRC125 and pSRC26 were picked and patched onto LA plates containing 5 mM sodium arsenite or 20 mM sodium arsenate, to test for arsenic resistance, and onto HBA plates with 25 μg/ml potassium tellurite for tellurite resistance (plate pictures are in Fig A2.1). R478 was used as a positive control. Four different *E. coli* strains (E294, DH5α, UB5201 and UB1637) were used as negative controls. Attempts to accurately measure resistance to copper and silver were unsuccessful, despite using a range of concentrations and different types of media. As expected, pSRC125 and pSRC26 conferred resistance to tellurite, but not to arsenate or arsenite (Table 4.3).

Table 4.3 Heavy metal resistances

Plasmid	Host strain	Arsenite		Arsenate		Tellurite	
		Expected[1]	Tested[2]	E	T	E	T
R478	E294	+	+	+	+	+	+
pSRC125	E294	-	-	-	-	+	+
SRC125	*S.* Typhimurium	-	-	-	-	+	+
pSRC26	E294	-	-	-	-	+	+
SRC26	*S.* Typhimurium	-	-	-	-	+	+
-	E294	-	-	-	-	-	-
-	DH5α	-	-	-	-	-	-
-	UB5201	-	-	-	-	-	-
-	UB1637	-	-	-	-	-	-

[1]Expected from presence of genes detected using PCR
[2]Tested in at least 3 replicates

The regions of pSRC125 and pSRC26 present, based on data from the set of 14 PCRs and heavy metal phenotypic testing, are presented in Fig 4.3, above. Also shown is a comparison of these plasmids with the IncHI2a plasmids R478 and pK29 (analysis from this study).

4.3 Restriction mapping

To determine how much of R478 was missing in pSRC125 and pSRC26, particularly surrounding the copper, silver and arsenic resistance regions, plasmid DNA was digested with

the infrequent cutter SwaI. In order to separate and resolve the larger bands, pulse field gel electrophoresis (PFGE) was used in preference to standard agarose gel electrophoresis. R478 was run as the control.

SwaI fragment sizes were predicted (shown in red in Fig 4.4) for an artificial IncHI2 ancestor, named "Plasmid X" (see Fig 3.20), which was generated as follows. The basis of Plasmid X was the R478 sequence (GenBank accession no BX664015), with the Tn*1696*-derived resistance region removed and the sequence adjacent to it, which is missing in R478 (see Fig 3.19), added in from the pK29 sequence (GenBank accession no EF382672). Tn*10* was included in Plasmid X, as R478, pSRC125 and pSRC26 all contained Tn*10* in the same position (see section 3.4). A restriction map was generated using Gene Construction Kit. Each band was allocated a number in the order they appear in the R478 sequence (shown in blue in Fig 4.4).

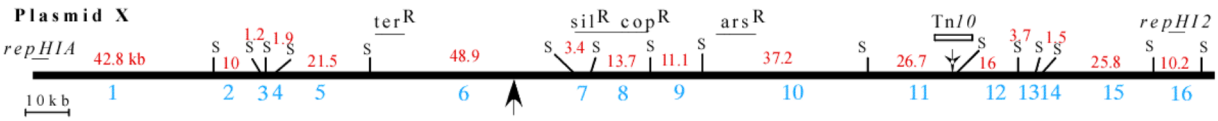


Fig 4.4 Restriction map of Plasmid X. Position 1 in Plasmid X is the same as in the R478 sequence (GenBank accession no BX664015) and is situated at the left of the solid, horizontal line. The SwaI sites are represented by "S" and the fragments in between them have been allocated band numbers from 1-16, shown in blue below. The sizes of the fragments are shown above in red and the positions of relevant regions in R478 are marked above. The arrow below, within fragment 6, marks the position of the Tn*1696*-like resistance region in R478.

The digests were analysed as follows. The fragments predicted to be the same size in pSRC125 and pSRC26 as in Plasmid X, were analysed (bands 1-5 and 11-16 in Fig 4.5). Then the presence of bands 6-10, which collectively contain the antibiotic and heavy metals resistance regions in R478, was analysed as these were expected to differ between plasmids.

4.3.1 Bands common to R478, pSRC125 and pSRC26

Restriction maps of Plasmid X, pSRC125 and pSRC26 showing bands 1-5 and 11-16 (numbered in blue) and their predicted sizes (values in red) are in Fig 4.5A, below. The PFGE

picture is shown in Fig 4.5B and the size of each fragment is listed in Table 4.4. The bands shared by R478, pSRC125 and pSRC26 are marked with a blue arrow in Fig 4.5, these correspond to bands 1, 2, 5, 11, 13-16 (Fig 4.5A). Although bands of less than 2 kb (that is bands 3, 4 and 14) were not detected, as their molecular weight was too low to be visualised, they were assumed to be present as the bands surrounding them were present. A 16 kb band, the correct size for band 12, was seen only in R478 (marked with a red arrow in Fig 4.5). Although pSRC125 and pSRC26 both yielded bands close to this size (+/- <2.5 kb), none aligned with band 12 of R478.

The contents of the corresponding fragment in R478 were analysed in an attempt to explain the apparent difference in the size of fragment 12 in pSRC125 and pSRC26. None of the PCR targets shown in Table 4.2 fall within fragment 12. Initial analysis revealed that the insertion sequence IS*186* was present in fragment 12 of R478. PCR primers (RH888 and RH887; sequences in Table A3.8 in Appendix 3) were designed on either side of IS*186* in R478 and a PCR product of 245 bp indicates that IS*186* is not present and one of 1,594 bp shows that IS*186* is present. Both pSRC125 and pSRC26 contained IS*186* in this position. Further analysis of band 12 in R478 revealed the presence of at least one other insertion sequence, named IS*150*. However, the presence of this or other IS in pSRC125 and pSRC26 was not investigated. It was assumed that the bands closest in size, namely 14.8 kb and 18.6 kb (highlighted with a red box in Fig 4.5), corresponded to band 12 in pSRC125 and pSRC26, respectively. After completion of the analyses described below, the initial assignments of these fragments to band 12 were confirmed. Differences in the size of band 12 remain to be investigated but may be explained by the presence or absence of additional IS.

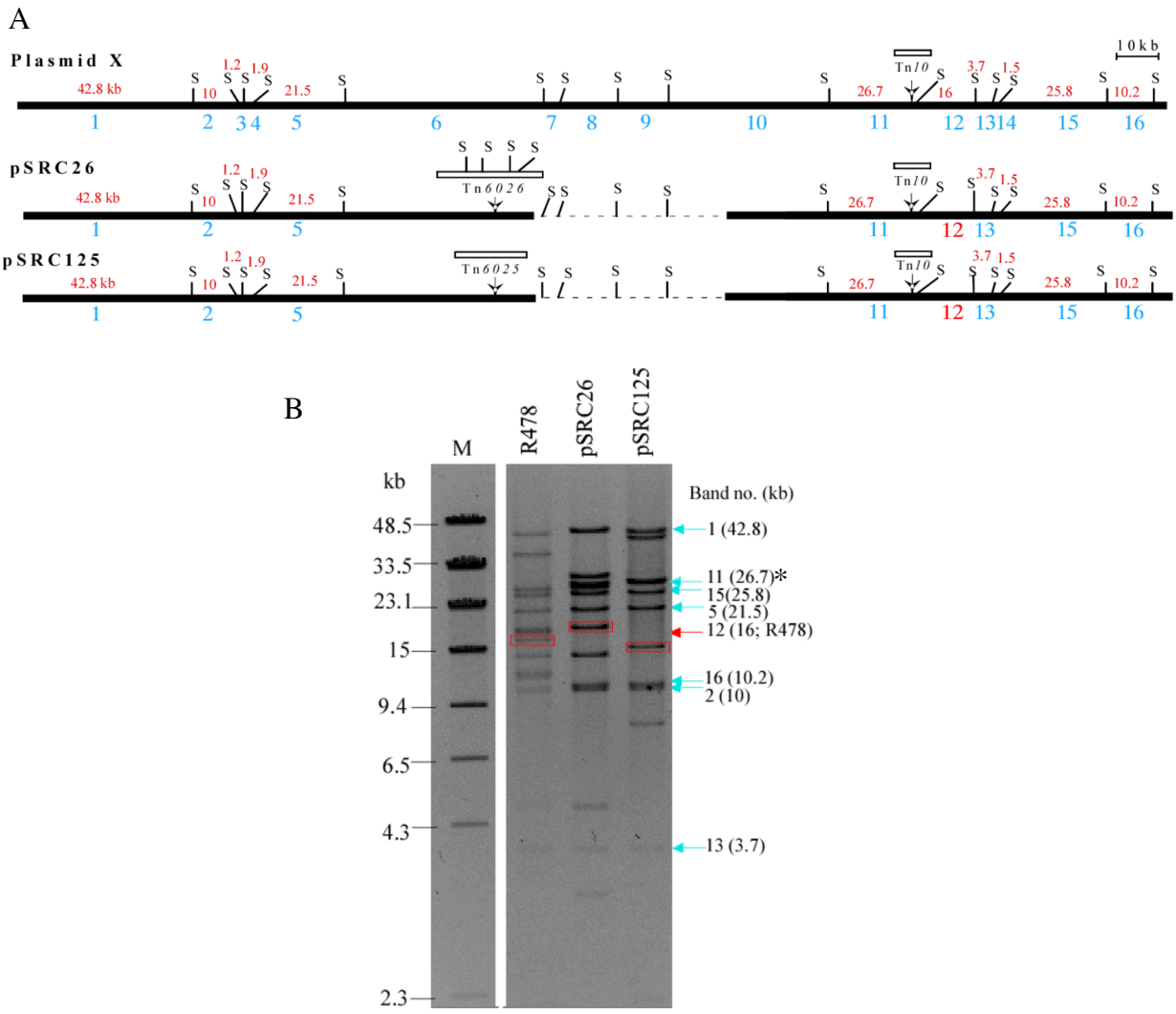


Fig 4.5 Restriction fragments in the conserved backbone A. Predicted SwaI sites in the IncHI2 backbone **B.** PFGE picture of SwaI digested R478, pSRC125 and pSRC26 plasmid DNA. In A the sizes of fragments in (kb) from these SwaI sites are in red above the horizontal line and the band numbers, are numbered 1-16 in the order they appear in the sequence. The asterisk indicates that this band is of double intensity, and represents 2 fragments of this size, one of which is marked as band 11 in this figure. Other features are as in Fig 4.4. In B, bands that are the correct size are marked with blue arrows and position band 12 of the correct size is marked with a red arrow. The bands assumed to be band 12 are boxed in red. Lane M is the PFGE ladder and sizes are as indicated. Smaller sized bands <2.3 kb (bands 3, 4 and 14) are not able to be visualised by this method.

Table 4.4 Observed bands 1-5 and 11-16

Fragment no.	Contents of band (PCR#)	Plasmid X expected size (kb)	Observed bands (kb)		
			R478	pSRC26	pSRC125
1	RepHIA(2), orf25(8)	42.8	43.9	44.5	43.1
11	Tn*10*	26.7	26.2	27.6	28
15	orf270(13)	25.8	25.2	25.6	27.5
5	orf56/57 (9)	21.5	22.1	22.5	22.9
12		16	16.2	18.6	14.8
16	RepHI2 (1)	10.2	12.4	11.3	11.3
2		10	10.7	11.1	11.1
13		3.7	3.9	4.3	4.3
4		1.9[1]			
14		1.5[1]			
3		1.1[1]			

[1]too small to be visualised by this method

Thus, each fragment in the region spanning the fragments 1-5 or 11-16 was shown to be present in pSRC125 and pSRC26 and, except for some minor variation in band 12, of the same size as in Plasmid X. This also indicated that the as yet unlocated resistance region in pSRC125, Tn*5393*e (see section 3.5.2), was not within bands 1-5 or 11-16.

4.3.2 Unravelling the variable region

Bands that were expected to appear in R478, pSRC125 and pSRC26 within the region spanning bands 6-10 from Plasmid X were then analysed. These are shown in Fig 4.6, with their predicted sizes marked in red. The sizes of some fragments within this region from pSRC125 and pSRC26, such as those within Tn*6026* and Tn*5393*e could be predicted (see Fig 4.6) and these were analysed first, then those fragments of unknown size.

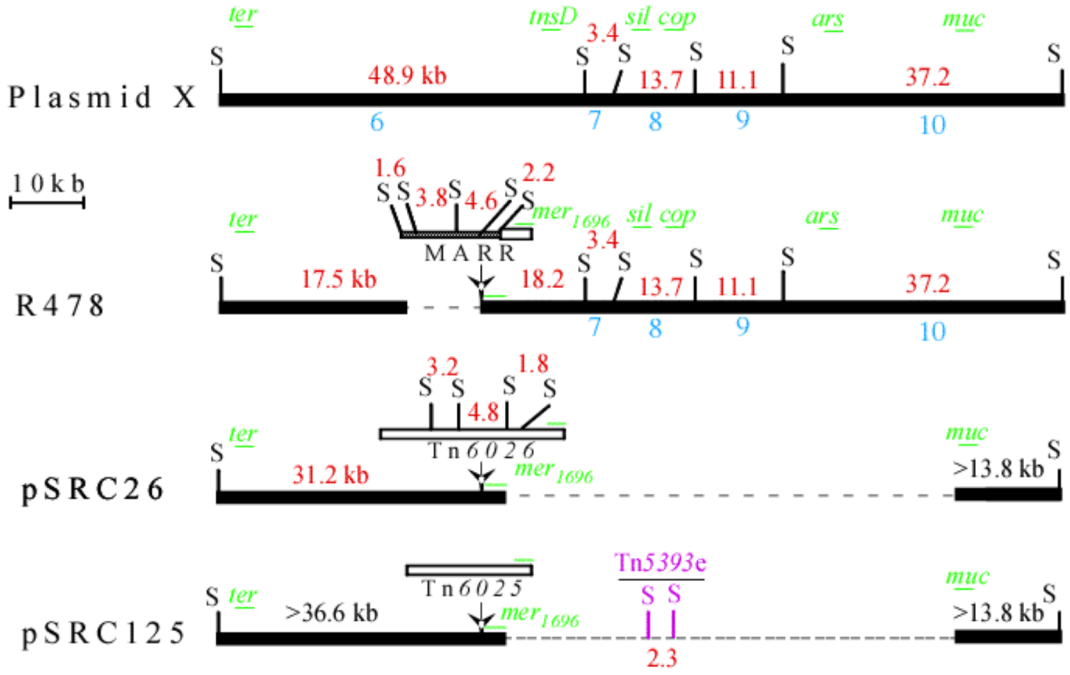


Fig 4.6 Restriction maps of the variable region. The band numbers are shown in blue. The positions of PCRs shown to be positive in this region are shown in green above. The sizes of predicted SwaI fragments are shown in red in kb and the minimum sizes of the outermost fragments, based on the last PCR shown to be positive and the SwaI sites present are shown in black. The location of Tn*5393*e is unknown and therefore its SwaI sites are shown in purple and are placed arbitrarily. Other features are as in Figs 4.4 and 4.5.

All bands observed in R478 corresponded to the expected fragments within this region (marked with purple dots in Fig 4.7A). Bands from Tn*6026* in pSRC26 and Tn*5393*e in

pSRC125 are marked with green dots in Fig 4.7A. Fragments like these, that carry resistance regions, tend to be smaller and more numerous because these regions can contain multiple copies of IS*26* (which contains a SwaI site). In pSRC26, Tn*6026* would be within fragment 6. One fragment that extends from the SwaI site on the LHS of band 6, to the first SwaI site in Tn*6026,* was predicted to be 31.2 kb (Fig 4.6). A band of approximately this size was observed (marked with a blue dot in Fig 4.7A; Table 4.5) indicating that the part of fragment 6 to the left of Tn*6026* is intact.

The remaining bands that had yet to be accounted for, are marked with yellow dots in Fig 4.7B and their sizes are shown in Table 4.5. Two of these bands were observed in pSRC26, of ~30 kb and ~14 kb, and three in pSRC125 of ~30 kb, ~41 kb and ~8 kb. One band common to both plasmids, which was estimated to be 28-30 kb (Tables 4.4 and 4.5), is of double intensity compared to other bands on the gel and thus represents 2 fragments: one of which was identified as band 11 (Fig 4.5) and one unallocated one, which is henceforth referred to as "the ~30 kb band". Both pSRC125 and pSRC26 have the ~30 kb band, which is a shared fragment (confirmed below), that was assumed to include *muc* and part of band 10 (as in Fig 4.8, below). It was smaller than the corresponding 37 kb fragment 10 from R478, which is to be expected, because the arsenic resistance region located within this fragment is missing in these plasmids.

The 2 bands in pSRC26 indicate that there is one SwaI site located in between Tn*6026* and *muc* – the one shared with pSRC125 (shown in purple within the dashed line in Fig 4.8). The other additional band in pSRC26 is unique and 14 kb in size. It must contain part of Tn*6026* and part of band 6 (Fig 4.8). In pSRC26, the segment from the last SwaI in Tn*6026* to the mer_{1696} boundary PCR is ~6.5 kb, a size consistent with the 14 kb band containing this region.

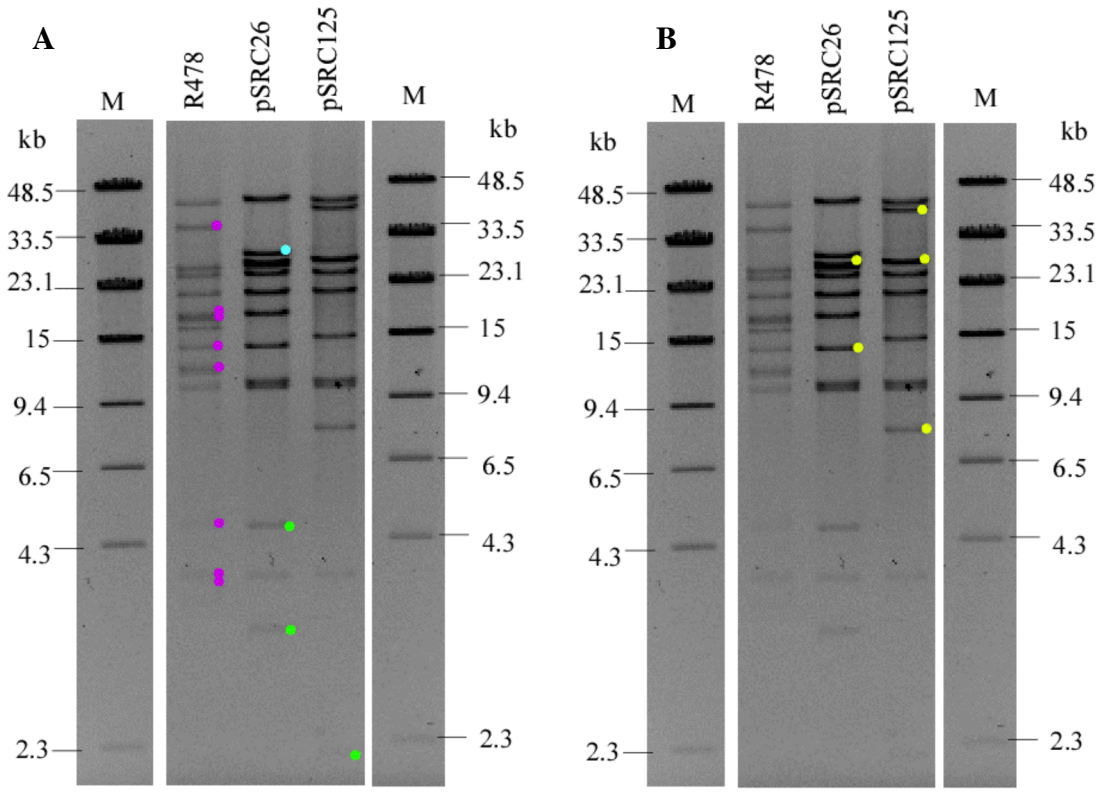


Fig 4.7 PFGE highlighting bands in the variable region A expected bands from backbone in R478 (purple), pSRC26 (in blue) and from Tn*6026* and Tn*5393*e (green). **B** Bands observed in pSRC125 and pSRC26, Tn*5393*e bands in blue for pSRC125 and Tn*6029/4352* in pSRC26 in green. Unexpected bands, assumed to be around deletion site is in yellow. Other features are as in Fig 4.5B.

Table 4.5 Predicted and observed bands in the variable region

Plasmid X band no. [1, 2]	Contents of band (PCR#)[3]	R478 (kb)		pSRC26 (kb)		pSRC125 (kb)	
		expected	observed[4]	expected	observed[1]	expected	observed[1]
10	*ars*(5), *muc*(13), *trh*(6)	37.2	36.4	>13.8	30.1	>13.8[5]	29.2
6	*ter*(2), orf92/3(3), *hipA*(10), *tnsD*(4)[6]						
6*	Tn*6025, ter*(2), orf92/3(3)					>36.6[5]	40.5
6*	tnp_{1696}, *ter*(2), orf92/3(3)			31.2	32.2		
6*	$mer_{1696;}$ *tnsD*(4)	18.2	18.7				
-	?			?	14.2		
6*	*ter*(2), orf92/3(3)	17.5	18.3				
8	*sil* (11), *cop* (12)	13.7	14.7				
9		11.1	12.2				
7		3.4	3.7				
-	from MARR	4.6	5.2	4.8	5.1	?	7.9
-	from MARR	3.8	3.7	3.2	3.4	2.3	2.2
-	from MARR	2.2	-[7]	1.9	-[7]		
-	from MARR	1.6	-[7]				

[1] * indicates part of this band is present
[2] - indicates a band not in Plasmid X but observed or predicted in pSRC125/pSRC26
[3] ? indicates, unknown or not expected
[4] size determined using a standard curve – averaged from triplicate PFGE gels
[5] determined using the distance between the last known SwaI site to the last known PCR to be present
[6] 48.9 kb if no Tn*1696*-like transposon
[7] too small to be observed

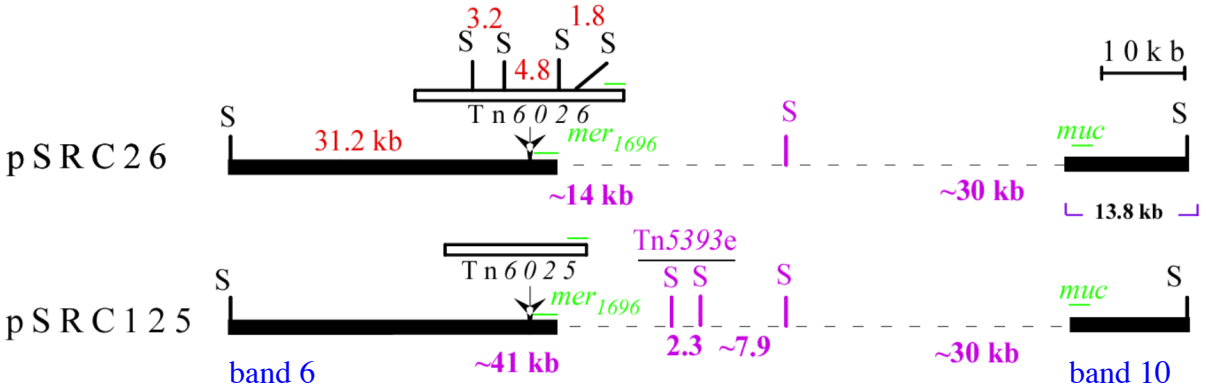


Fig 4.8 Observed SwaI fragments of the variable region. The observed band sizes are shown in purple below. The SwaI sites predicted to be in this region but with unknown location is represented by a purple "S". The 13.8 kb segment reaching from the SwaI site to the *muc* PCR known to be present is bracketed below. Fragments containing parts of bands 6 and 10 from Plasmid X are numbered in blue below. Other features are as in Figs 4.6 and 4.4.

Tn*5393*e in pSRC125 must also be within fragments 6-10 (marked purple in Fig 4.8), as all other bands have been accounted for and if they contained Tn*5393*e, they would have been a different size. The expected 2.3 kb fragment containing the *aphA1b* from Tn*6023* was observed in pSRC125 (marked with a green dot in Fig 4.7A). The three unallocated bands observed in pSRC125 (marked with yellow dots in Fig 4.7B) indicated that a single SwaI site must be present in between Tn*6025* and the ~30 band, in addition to the two sites within Tn*5393*e (Fig 4.8). In pSRC125, the size of the segment that contains Tn*6025*, reaching from the SwaI site on the left of band 6 to the primer on the RHS of Tn*6025* (green in Fig 4.8), would be 36.6 kb (shown in Fig 4.6 in black). Thus, the band containing Tn*6025* must be at least this size. Only one fragment in pSRC125 was unaccounted for and greater than 36 kb, and it was estimated to be ~41 kb (shown in purple in Fig 4.8).

On the other side of the gap, the region known to be present spans from the SwaI site on the RHS of band 10, to *muc*, the last PCR shown to be present on the left of band 10, (green in Fig 4.8) and is 13.8 kb (shown in Fig 4.6 in black and bracketed in Fig 4.8). The only band that was not yet unaccounted for and large enough to contain this region, is the ~30 kb band

(shown in purple in Fig 4.8), which pSRC26 shares. Although the location of Tn*5393*e in pSRC125 had not been determined, the ~41 kb band, and not the ~30kb, must contain part of Tn*5393*e because pSRC125 shares the ~30 kb band with pSRC26, which does not contain Tn*5393*e (Fig 4.8). The last unallocated band in pSRC125 is ~7.9 kb and this must contain the remainder of Tn*5393*e. The configuration of this region in pSRC125 and pSRC26 is shown in Fig 4.8. Therefore, all the SwaI fragments present in pSRC125 and pSRC26 could be accounted for. However, exactly what was in between the mer_{1696} and *muc* PCRs (Fig 4.8) had to be examined using other techniques.

Using RFLP analysis, the sizes of pSRC125 and pSRC26 were estimated to be 242.8 kb and 252.4 kb respectively, by adding up the sizes of the observed bands which were calculated by measuring the gel migration dista nces and using standard curves (sizes shown in Tables 4.4 and 4.5). These sizes are inclusive of 2 fragments of ~28 kb, which appear as a single band of double intensity on the gel. These sizes are similar to other sequenced IncHI2 plasmids, for example, R478 is 275 kb and pK29 is 269 kb (see Table 1.4).

4.3.3 Cloning the backbone

In an attempt to determine how much of the backbone was present to the right of Tn*6025*, a 5.5 kb BamHI fragment of pSRC125 was cloned into pUC19, using mercuric ion selection. BamHI was used as it does not cut in the *mer* module of Tn*6026*. The clone pRMH961 was recovered. Sequencing of this clone revealed that it extended only to a BamHI site, also present in R478, 750 bp to the right of Tn*6025* (Fig 4.9).

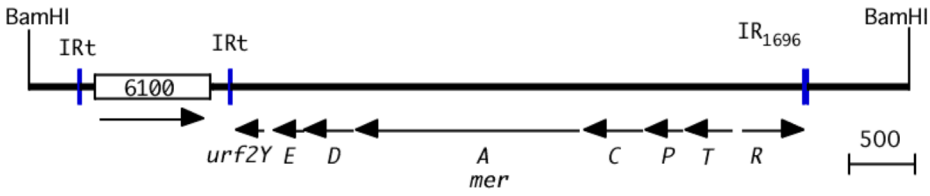


Fig 4.9 pRMH961. The extent of the 5.5 kb insert is shown by a horizontal line between the 2 BamHI sites. The open box numbered with 6100 represents the insertion sequences IS*6100*, arrows represent direction and extent of genes. Vertical bars are inverted repeats (IR) from either transposons, where the transposon number is subscripted, or from integrons where IRi marks the initial IR and IRt denotes the inverted terminal repeat.

Five arbitrarily selected white clones also recovered, using X-gal-supplemented media, were found to contain inserts of 2-6 kb. These clones were sequenced with a primer in pUC19 and only one yielded clear sequence. It contained a 4.6 kb insert and was named pRMH960. Fortuitously, it contained a fragment derived from the region between the arsenic and copper resistance operons of R478, both of which were missing from pSRC125 (Fig 4.10).

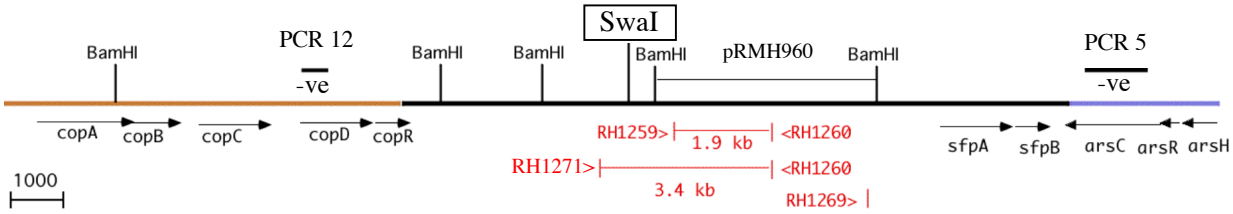


Fig 4.10 pRMH960 insert and its surrounding sequence. The BamHI clone pRMH960 insert, a 4.6 kb fragment of pSRC125 is shown as a line between 2 BamHI sites. BamHI sites are shown as "BamHI" and SwaI sites are shown as "SwaI". Positions of primers are shown as red lines. The sequences of RH1259, RH1260, RH1271 are in Table A3.8 and RH1269 in Table A3.7 in Appendix A3. The copper resistance region is highlighted copper and the arsenic resistance region, purple. Positions of PCR 5 and 12 are shown as a black line above and the "-ve" below indicates that no PCR product was obtained. Other features are as in Fig 4.9.

The presence of this BamHI fragment was confirmed in pSRC125 and tested in pSRC26 using PCR with primers internal to the fragment (RH1259 and RH1260; Fig 4.10). Both plasmids contained the fragment, which indicated that although all PCRs targeting sequences between *tnsD* and *ars* did not produce amplicons (Table 4.3), a piece in between *ars* and *cop* remained. Thus, 2 separate segments are missing in pSRC125 and pSRC26, and not a single region that spans all the negative PCRs. In R478, the left-hand SwaI site of band 10 (Fig 4.6), was positioned 487 bp to the left of where the BamHI fragment (in pRMH961) lies. A primer within the BamHI fragment was used with one designed to the left of the SwaI site, so that the

amplicon included the SwaI site (boxed in Fig 4.10). This site was found to be present in pSRC125 and pSRC26, thus both ends of the 37.2 kb fragment 10 from R478 are present in both. Hence, the ~30 kb band observed in pSRC125 and pSRC26 corresponds to fragment 10, which contains the arsenic resistance region in R478, but it must harbour an internal deletion that makes the fragment ~7 kb smaller.

4.4 Defining the ends of the missing regions

Regions to the left and the right of the LH SwaI site of fragment 10, which is present in pSRC125 and pSRC26 (Fig 4.10), were then analysed separately using long-range PCR.

4.4.1 Right-hand missing segment

The segment in between the SwaI sites of fragment 10, which includes *muc* and *ars*, is 37.2 kb in the R478 sequence, but is only ~30 kb in pSRC125 and pSRC26 (Table 4.5), suggesting that these plasmids are missing ~7 kb from fragment 10 which must include the arsenic resistance region, or part thereof, as these plasmids do not confer arsenic resistance or carry the arsenic resistance gene (see Tables 4.2 and 4.3). PCR primers, named RH900 and RH899, were designed 3.5 kb to the left of *muc*, the last PCR positive on this side (Fig 4.11). Both pSRC125 and pSRC26 gave amplicons of 1.5 kb, indicating that this segment was present.

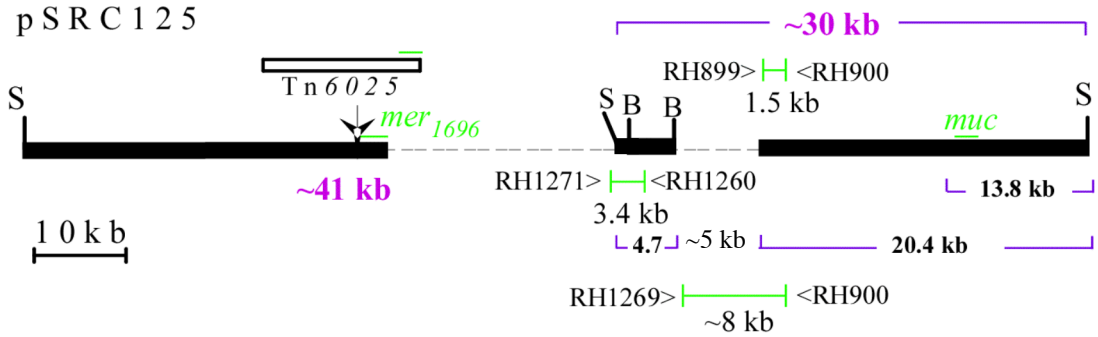


Fig 4.11 Region to the right of pRMH960 in pSRC125. Observed fragment sizes in pSRC125 are shown in purple. The sizes of segments known to be present are bracketed in black. PCR primers are shown as green vertical lines, and PCR amplicons as green horizontal lines. Primer sequences are in Table A3.8 in Appendix A3. Other features are as in Fig 4.4 and 4.6.

Because the size of the SwaI fragment observed in both pSRC125 and pSRC26 was ~30 kb (shown in purple in Fig 4.11), the size of the segments in between the sequence known to remain (of sizes 4.7 kb and 20.4 kb; Fig 4.11), must be ~5 kb. A segment of this size should be able to be amplified in long-range PCR reactions. The primer RH1269 was designed closer to the right hand BamHI site, facing out of the BamHI fragment. RH1269 was used with RH900 in a PCR performed on pSRC125, which yielded an amplicon of ~8kb (Fig 4.11). The 950 bp of sequence obtained from this amplicon, using primer RH1269, revealed that 7.2 kb of sequence present in R478 was missing in pSRC125. The position of this deletion is shown in Fig 4.12 and it was located 778 bp to the right of the BamHI site from pRMH960. A primer named RH895 was designed closer to the site of deletion, to replace RH900 and when used with RH1269, yielded a 3 kb PCR amplicon from pSRC125 (Fig 4.12). This PCR was performed on pSRC26 and the sequence of the amplicon obtained confirmed that pSRC26 contained the same configuration as in pSRC125.

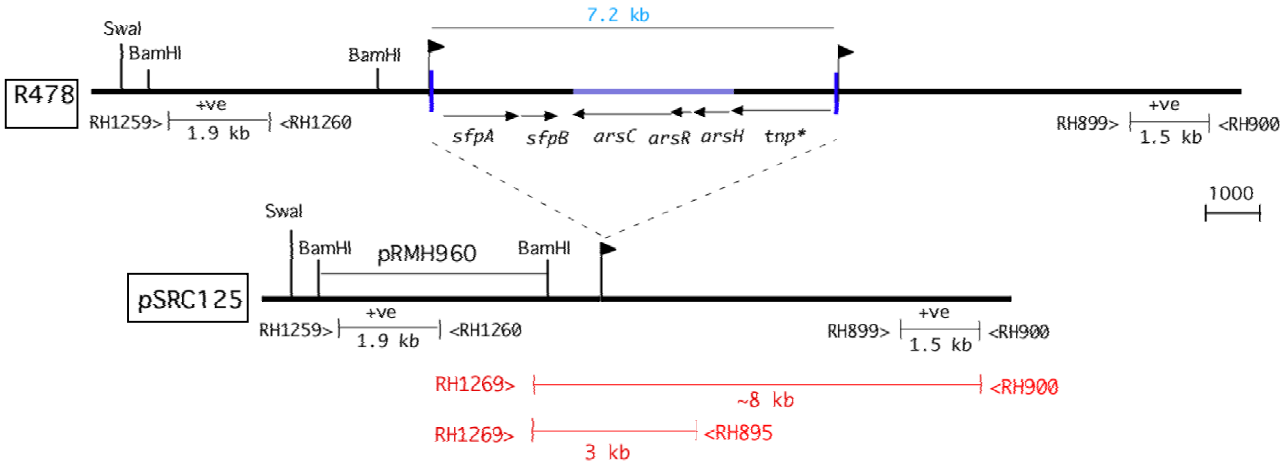


Fig 4.12 The arsenic resistance transposon in R478. The 7.2 kb arsenic resistance transposon and surrounding sequence in R478 is shown above. The flags show the position and orientation of direct repeats. Primer positions are shown as vertical red lines and the product sizes are shown below the horizontal line joining them. The sequence of RH895 is in Table A3.7 in Appendix A3. Other features are as in Fig 4.11.

4.4.1.1 An arsenic resistance transposon in R478

A 7,193 bp segment present in R478 (positions 156,498-163,690 in GenBank accession no. BX664015), containing the arsenic resistance region and *sfp* genes, which encode a sulphate permease, was not in pSRC125 and pSRC26. Detailed analysis revealed that this segment in

R478 was flanked by an "ACTTA" direct repeat (represented as flags in Fig 4.12), whereas pSRC125 and pSRC26 contained only one copy in the position of the 7.2 kb insertion (Fig 4.13A). Furthermore, a putative imperfect (17/25 bp match) inverted repeat was identified at the ends of the insertion in R478 (Fig 4.13B).

Fig 4.13 Sequence alignments of inverted repeats of the arsenic Tn A Sequence alignments of R478 and pSRC26 on the left and right hand boundaries of the insertion. In the RHS the start codon of the transposase is shown with an arrow. The asterisk indicates that the transposase is only putative. **B** R478 right and left ends of the insertion aligned, the RHS is complimented for easy comparison. Single nucleotide matches between the sequence alignments are shown as colons. The direct repeats are underlined and in red. The inverted repeats are in bold.

An open reading frame that encodes a putative transposase was identified adjacent to *arsH* (marked as *tnp** Figs 4.12 and 4.13). This putative transposase is part of the HMM PF00665 rve protein family, which includes bacterial transposases and viral integrases, all of which contain a DDE motif in their catalytic site. BLASTp searches of the GenBank protein database have revealed that this protein was similar to a large number of entries, most of which were annotated "putative transposase" and were from bacterial genomes. Some of these matches were from other plasmids, such as a match with 97% identity to a protein from the IncHI1 plasmids R27 and pAKU_1 (see section 7.2.2). Across the 100 protein entries from the BLASTp search, the protein identity remained above 47%, indicating that this is a large and apparently well-conserved family of transposases. However, no member of this family has been experimentally shown to be an active bacterial transposase. This potential transposase family warrants further analysis, which is outside the scope of this thesis.

The inverted repeats and putative transposase identified and the direct repeats surrounding this element, provide strong evidence that it is a Tn. Thus, this sequence is more likely to be an insertion in R478, rather than a deletion in pSRC125 and pSRC26. The arsenic resistance transposon (ars^R Tn) was also present in other IncHI2a plasmids pK29, pEC-IMP and pEC-IMPQ. IncHI2b plasmid pAPEC-O1-R does not contain the Tn. However, a 15.9 kb region, which would contain the arsenic resistance transposon was replaced by a 2.9 kb additional region in pAPEC-O1-R. At the nucleotide sequence level no other entries contained this ars^R Tn, or even sections of it, larger than 2 adjoining genes, indicating that this arsenic resistance transposon is unique to IncHI2 plasmids and that the arsenic resistance operon had not simply been mobilised from another source. This potential transposon was not investigated further.

4.4.2 Left-hand missing segment

To the left of where the LH SwaI site of fragment 10 is located, amplicons were not obtained for mapping PCRs targeting *tnsD, cop* and *sil* for pSRC125 or pSRC26 (Fig 4.6; Table 4.2). These PCRs span a segment of 26 kb, indicating that the sequence missing on this side must be at least this size. The remaining unidentified band in pSRC26 was ~14 kb, which must contain the last SwaI site from Tn*6029* and thus the RH part of Tn*6026* (Fig 4.8). The segment between the last SwaI site in Tn*6029* and IR_{mer} is 5.5 kb. Because the only SwaI site present in this region is the one on the LHS of fragment 10 (boxed in Fig 4.14A), ~8.5 kb must remain in between the IR_{mer} and this SwaI site (Fig 4.14A). Hence, long-range PCR could be used to map across Tn*6026* and this SwaI site. A primer named RH1251 was designed 3 kb from BamHI fragment in pRMH961, the last sequence known to be present. It was used with primer RH1260 in the pRMH960 BamHI fragment (Fig 4.14A), in a PCR performed on pSRC125 and pSRC26. A ~7.5 kb amplicon was obtained for pSRC26, but no product was observed for pSRC125 (Fig 4.14B). Sequencing of the pSRC26 amplicon with

RH1251 produced 1.2 kb of clear sequence, which revealed that there was 32,414 bp of R478 sequence missing in pSRC26 (Fig 4.14A). From this amplicon, 857 bp matched to the left and 348 bp to the right of this 32.4 kb segment in R478.

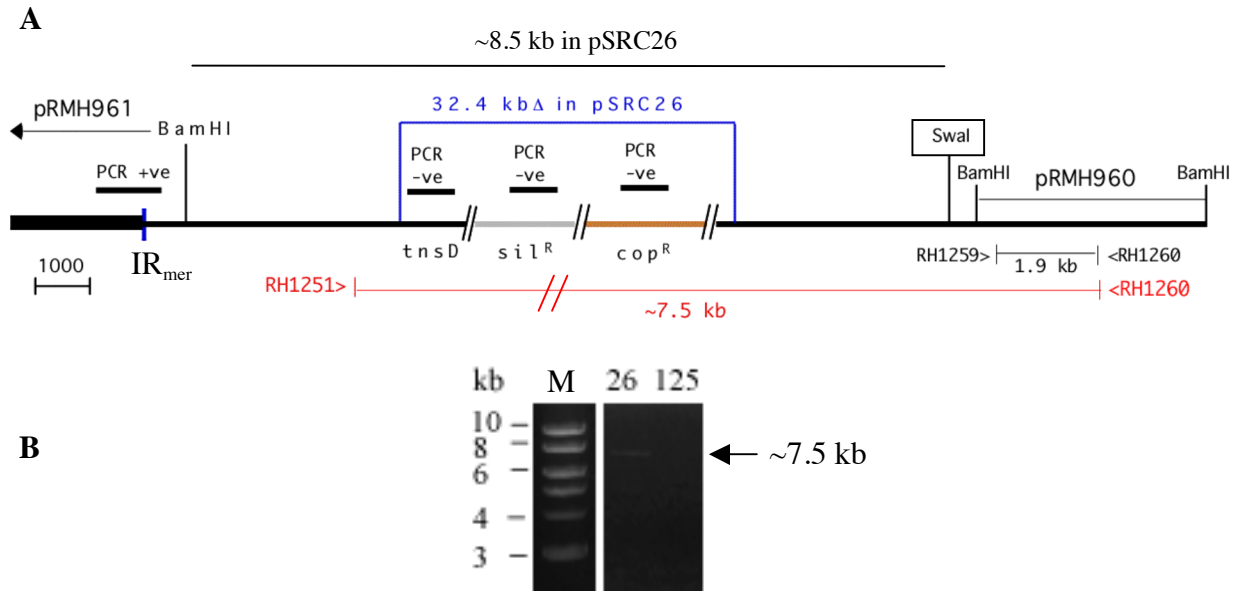


Fig 4.14 Identifying missing sequence across the silver and copper resistance regions. A. Schematic of long-range PCR used to link across the gap in pSRC26. Two slashed lines indicate that there is additional sequence within the line and is not to scale. The blue line above indicates the amount of sequence missing in pSRC26 compared to R478. The SwaI site shared between pSRC125 and pSRC26 is boxed. The sequence of RH1251 is in Table A3.7. Other features are as in Figs 4.10 and 4.12. **B.** Agarose gel of PCR products from the above reaction. The lane 26 represents pSRC26, 125 represents pSRC125 and the 1 kb MW ladder, the lane marked M, with selected sizes shown. The size of the product obtained in pSRC26 is marked by an arrow on the right.

4.4.2.1 A silver/copper resistance transposon in R478

The 32.4 kb region missing in pSRC26, corresponded to positions 114,733-147,146 in the R478 sequence (GenBank accession no. BX664015) and included the *tns* genes, annotated as Tn*7*-like mobility genes, as well as the copper and silver resistance regions. Like the ars^R Tn, this region appeared to be an insertion in R478, rather than a deletion in pSRC26. This putative transposon was named Tn*6024*, as it had many features of a Tn. In R478, Tn*6024* was surrounded by a 5 bp "GGTCC" duplication, of which pSRC26 contained only 1 copy at the site of insertion in R478 (Fig 4.15A). Tn*6024* in R478 also contained imperfect (17/23 match) inverted repeats (Fig 4.15B) at its ends. Class III transposons (defined in section 1.3.1.3) often contain short repeated sequences, that are similar to part of their IRs, near their

ends, which facilitate binding of the transposase. For example, Tn*7* has three 16 bp repeats on the LHS and four on the RHS, within 200 bp of its IRs (Fig 1.8; [49]). Here, a 13 bp imperfect repeat with a consensus CAYTAAAGTTGTA was identified which occurred twice on the right and left hand sides (Fig 4.15C).

A

```
R478 LHS  TTACTATGCGCGTTTCTACGTAAATTCTATTGGTTTTAGAGGGTCCTGTCCGAGGACAATAAAGTTGTACACAAAGCACTAAAG
          ||||||||||||||||||||||||||||||||||||||||||||||||||||  |   |  |||   |         |   || |||
pSRC26    TTACTATGCGCGTTTCTACGTAAATTCTATTGGTTTTAGAGGGTCCTACCTTTGCTCAAATCAAAATGTTCTGACATACCAAAA
           |  |    |        |   || ||   |   | |     ||||||||||||||||||||||||||||||||||||||||||
R478 RHS  CTTTTTCAGGTTAAAGGATACAACTTTAATGTCTCTACACAGGTCCTACCTTTGCTCAAATCAAAATGTTCTGACATACCAAAA
```

B

```
LH-IR  TGTCCGAGGACAATAAAGTTGTACACAA
       :::      ::::  ::::::::::   :
RH-IR  TGTGTAGAGACATTAAAGTTGTATCCTT
```

C

LHS

```
TGTCCGAGGACAATAAAGTTGTACACAAAGCACTAAAGTTGTAACCATTCTCAAAAAACGACGCTAATGTTGTATGCTTACTGAATGTGTGGTTATAGGG
```

RHS

```
TGTGTAGAGACATTAAAGTTGTATCCTTTAACCTGAAAAAGACTTTAAAGTTGTAAAGTTTTAAAAAAACACATACAAAACATGATGCTAA:GTGTGTAT
```

Fig 4.15 Tn*6024* sequence alignments. A Sequence of R478 across the left and right hand sides of Tn*6024* aligned with pSRC26 sequence spanning at the site of insertion. The red font highlights the direct repeats and the bold, the inverted repeats **B** The inverted repeats of Tn*6024*. The left hand IR (LH-IR) is complimented then aligned with RH-IR, which are shown in bold. **C** The transposase binding sequence repeated near the IRs of Tn*6024* which are shown in bold. The consensus binding sequence (CAYTAAAGTTGTA) is highlighted in yellow. Right and left sides are defined as in Fig 4.16, below. Other features are as in Fig 4.13.

Once the boundaries of Tn*6024* were defined, its distribution in other locations was analysed using BLASTn. First, the only IncHI2 plasmid other than R478 that contained an intact copy of Tn*6024*, was pAPEC-O1-R (GenBank accession no. DQ517526). It was found in the same location as in R478 and shared >99.99% sequence identity, but a copy of IS*4* was detected in *tnsD*. Although Tn*6024* appeared to be mostly absent in the other IncHI2s, one or both its ends were identified in the same position as R478, in pK29 and pEC-IMP/Q. Both the left and right boundaries of Tn*6024* (including the inverted repeats and direct repeats) remained in pEC-IMP/Q and only the RH boundary was found in pK29 (Fig 4.16).

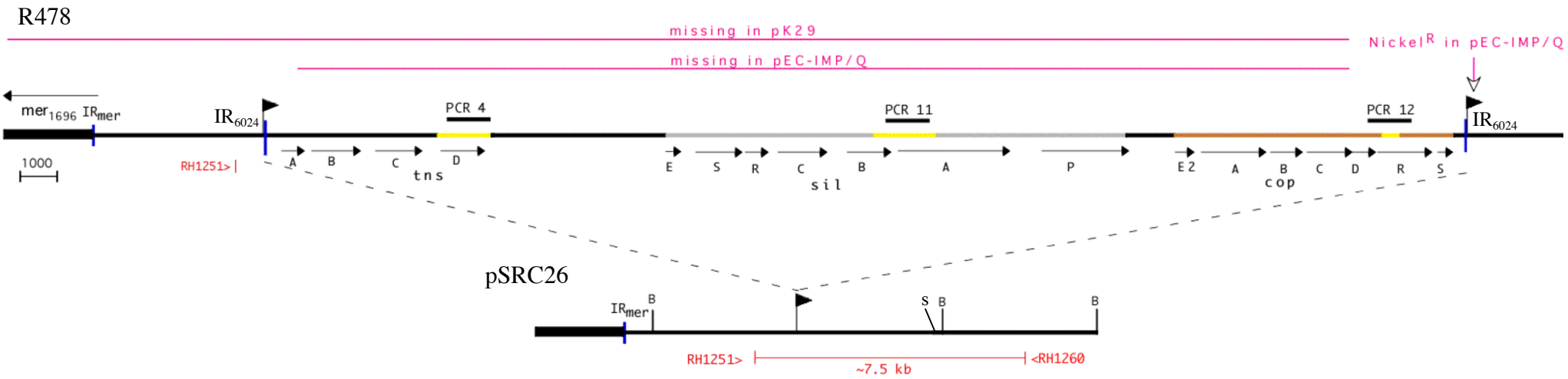


Fig 4.16 Schematic of Tn*6024* in R478. The pink lines above indicate the extent of Tn*6024* in pK29 (top) and pEC-IMP and pEC-IMPQ (bottom), compared to in R478. The position of a region in pEC-IMP and pEC-IMPQ conferring nickel resistance is shown by a pink open headed arrow. The regions are shown below, where "tns" indicates the Tn*7*-like *tns* genes, "sil" the silver resistance region, and "cop", the copper resistance region. The position of BamHI sites are shown as "B" and SwaI as "S". Remaining features are as in Figs 4.11 and 4.13. The sequences of primers are in Table A3.7 in Appendix 3.

In pK29 and pEC-IMP/Q, the majority of the centre of Tn*6024* containing the Tn*7*-like *tns* genes, silver resistance and part of the copper resistance regions, had been replaced by different antibiotic resistance regions, neither of which had obvious features of an MGE (Fig 4.3). Thus, at some point Tn*6024* had been present in pK29 and pEC-IMP/Q, but subsequent insertion and deletion events had occurred once within the plasmid. Interestingly, pEC-IMP and pEC-IMPQ were found to contain an additional nickel resistance region directly adjacent to the RHS of Tn*6024* (Fig 4.16).

Three further GenBank entries matched to Tn*6024* with >99.99% identity and each was found in a different location, surrounded by a 5 bp direct repeat (Fig 4.17A). None were on plasmids. They included an *E. coli* chromosome (GenBank accession no. CU928145), an *E. cloacae* genome (GenBank accession no. CP001918) and a *Cronobacter sakazakii* genome (GenBank accession no. CP000783) (Table 4.6). Each entry had some variation within Tn*6024*, such as the presence of IS (Table 4.6). Thus, R478 contains the only example of an intact and uninterrupted Tn*6024*. In the case of the *E. coli* chromosome (GenBank accession no. CU928145), an otherwise identical sequence lacking Tn*6024* and 1 copy of the 5 bp direct repeat (Fig 4.17B) was found in another *E. coli* chromosome (GenBank accession no. CU928160; Table 4.6), providing further evidence that Tn*6024* is capable of movement. Interestingly, the *E. cloacae* genome was from the Type Strain ATCC 13047, isolated from Jordan in 1890 [253], indicating that the Tn*6024* structure has existed for at least 100 years. Tn*6024*-like was located within different, unique sequences that encode hypothetical proteins of unknown function, except the *E. cloacae* and *C. sakazakii* genomes (GenBank accession nos. CP001918 and CP000783), shared 80% identity across a region on the LHS of Tn*6024* and were annotated as encoding "predicted flavoproteins". Thus, all examples of Tn*6024* are in different positions and not in or near *att*Tn*7*. This indicates, that although Tn*7* can display site-specific insertion at *att*Tn*7* using the protein combination TnsABC+D (see section

1.3.1.3.1), this does not appear to be the case for Tn*6024*. It appears to be inserting randomly into chromosomes, as well as conjugative plasmids, demonstrating that non-specific transposition has been occurring. Further experimental work examining the transposition of Tn*6024* in strains with and without *att*Tn*7*, or analysis of more sequences containing Tn*6024* to detect any similarity between insertion sites, would be needed to determine whether Tn*6024* can target *att*Tn*7* or another specific site.

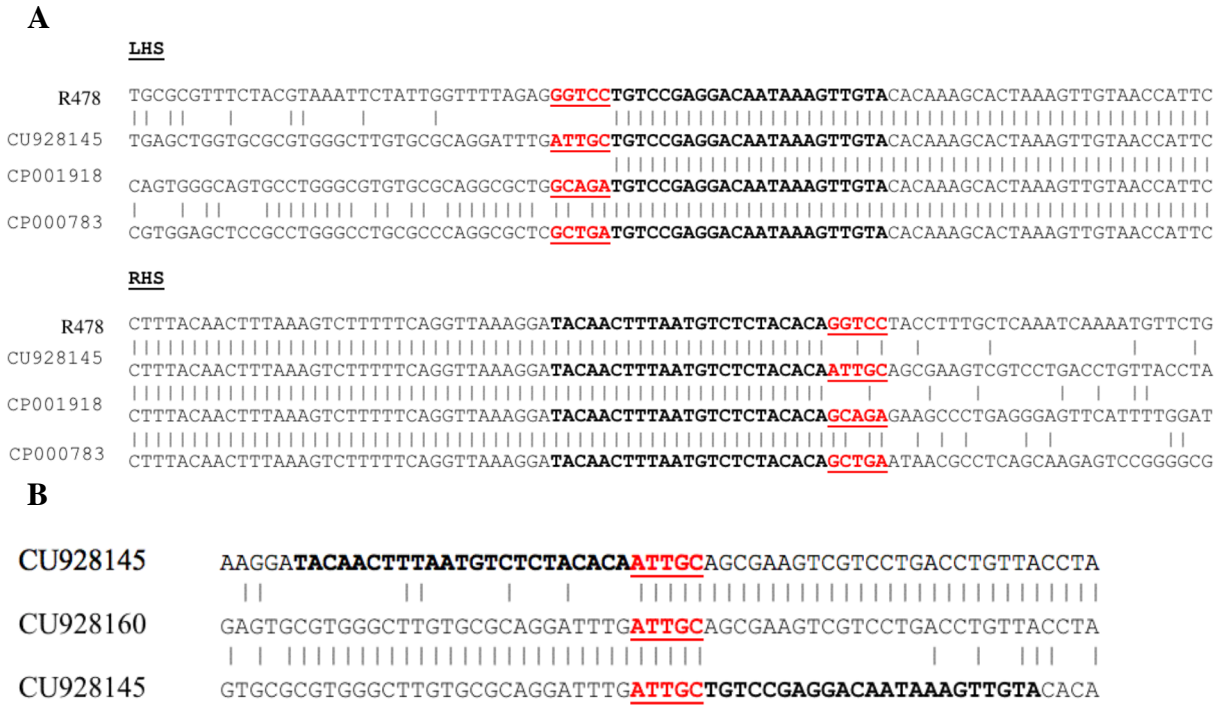


Fig 4.17 Sequences of different locations of Tn*6024*. The GenBank accession nos of the Tn*6024*-containing entries are given on the LHS of the alignments. **A** Sequence alignments of R478 and 3 other entries on the right and left hand side of Tn*6024* **B.** Sequence alignment of the same position in 2 *E. coli* chromosomes, one that contains Tn*6024* (CU928145) and one that does not (CU928160). Other features are as in Fig 4.15.

Table 4.6 GenBank entries associated with Tn*6024*

Source	Comments on Tn*6024*	GenBank acc no.	Position	Date submitted	Reference
E. coli chromosome	IS*1* in *tnsD*	CU928145	3968018-4001203	2008	Genoscope-C.E.A[1]
E. cloacae genome	IS*903* after *tnsD*; 2 unnamed IS in sil^R region	CP001918	4982867-5017694	2010	[253][2]
C. sakazakii genome	5.4 kbΔ including *tnsD* and *tnsC*	CP000783	4195106-4222149	2007	[254]
E. coli chromosome	No Tn*6024*	CU928160	-	2008	Genoscope-C.E.A[1]

[1]No manuscript associated with sequence
[2]Genome announcement

Initial annotation of the *tns* genes in R478 stated they were of Tn*7* origin [58] and annotation of the other copies of Tn*6024*-like elements identified in the GenBank database, all stated that these genes were "Tn*7*-like". Although Tn*6024* shares other features of Tn*7*, such as the essential "TG" motif at the termini of the IR [57] and the similar length of IRs (23 bp in Tn*6024* compared to 25 bp in Tn*7*) and they both create 5 bp duplications, they appear to be distantly related. The nucleotide sequence analysis performed here revealed that Tn*6024* and Tn*7* were not all related at the nucleotide level. Yet at the protein level, the *tns* encoded proteins, TnsA-D, shared between 26% (TnsB) and 34% (TnsA) peptide sequence identity with the equivalent proteins in Tn*7*. In fact, when the amino acid sequence of TnsB from Tn*7* was analysed using BLASTp, over 100 entries matched with greater identity than TnsB from Tn*6024* and only 19 of these were identical to TnsB from Tn*7*. This indicates that there is an extensive set of transposons that are related to Tn*7*. Thus, the "Class III" transposon group, including Tn*7* and Tn*402* (see section 1.3.1.3), are clearly a diverse group of transposons that warrant further study and classification, which is beyond the scope of this study.

4.4.3 Location of Tn5393e

As it was previously predicted that Tn*5393*e was contained in this region in pSRC125 (Fig 4.8), its location was determined using the primers linking across the Tn*6024* insertion site, and ones facing out of Tn*5393*e (Fig 4.18). Two PCRs were performed on pSRC125, using the primer near IR_{mer} (RH1251 in Fig 4.18) and primers in *strB* and $tnpA_{5393}$ (RH647 and RH1258 in Fig 4.18) in order to detect Tn*5393*e in both orientations. Linkage between the *strB* end of Tn*5393*e and the sequence adjacent to IR_{mer} was established via a 1.2 kb amplicon, which was sequenced (Fig 4.18). Then the $tnpA_{5393}$ end of Tn*5393*e was linked to the BamHI fragment in pRMH960 by a 7.5 kb amplicon (primers RH1258/RH1260; Fig 4.18), which was sequenced with RH1258 and 500 bp of sequence to the right of Tn*5393*e was obtained.

These sequences showed firstly that Tn*5393*e was surrounded by a 5 bp direct repeat. Secondly, it revealed that the same 32.4 kb segment of sequence missing in pSRC26, that is Tn*6024*, was also missing in pSRC125. Tn*5393*e was positioned 4,251 bp to the right of Tn*6025* and 387 bp to the left of site of Tn*6024* insertion. The pSRC125 GenBank entry GU562437 was extended 11.8 kb of discontinuous sequence to the right of Tn*6025* to include the whole of Tn*5393*e and the sequence extending across the Tn*6024* and arsenic resistance transposon boundaries (with 3 kb and 7.4 kb gaps between Tn*6025* and Tn*5393*e, and Tn*6024* and the ars^{R} Tn, respectively).

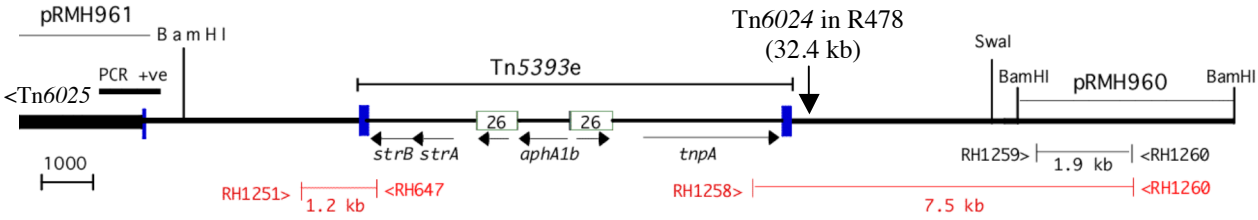


Fig 4.18 Schematic of PCR used to locate Tn*5393*e. The extent of Tn*5393*e is marked above. The position of the Tn*6024* insertion is shown by a vertical arrow. The sequences of primers are listed in Table A3.7 in Appendix A3. All other features are as in Fig 4.11, 4.13 and 4.17.

4.5 Discussion

The IncHI2 plasmids pSRC125 and pSRC26 appear to share many features with R478. Only one SwaI fragment in the non-variable region (bands 1-5, 11-16 in Fig 4.5A) differed in size for pSRC125 and pSRC26 compared to R478 (band 12; Fig 4.5). The remainder was shared with R478. Similarly, when the backbones of the other sequenced IncHI2 plasmids pK29 and pEC-IMP/Q were examined, the majority of the IncHI2 backbone remained constant but 2 regions of variability were apparent (marked green in Fig 4.19). The sizes of both pSRC125 and pSRC26 were determined to be 243 kb and 252 kb, respectively, using RFLP and separation by PFGE. These are comparable to the sizes of other IncHI2 plasmids (Table 1.4). Multiple attempts to size these large plasmids using S1-nuclease treated plasmid DNA separated by PFGE were unsuccessful, but RFLP seemed to determine their sizes efficiently.

Here, 2 novel heavy metal resistance transposons that both contain IRs and are surrounded by direct repeats, were identified. Previously, it had been predicted that the silver resistance region was potentially on a MGE because it was adjacent to the "Tn*7*-like" *tns* genes, yet the boundaries of this element could not be found [58]. Similarly, the arsenic resistance operon had been predicted to be on a MGE because of the elevated GC content and because directly upstream of the *ars* resistance gene *arsH*, a transposase gene similar to that in IS*2501,* from an arsenate resistance transposon Tn*2501* from *Yersinia*, was identified. It was even speculated to contain the sulfate permease [202] now known to be included in the Tn. However, transposibility assays failed to yield transposition and this putative element could not be defined [202]. This is the first report that confirms this arsenic resistance region is indeed on a Tn, the boundaries of which are clearly defined. Furthermore, when these heavy metal resistance transposons were identified as insertions in the other IncHI2 plasmids, this showed that pSRC125 and pSRC26 clearly contained a more "ancestral" backbone, which had not yet acquired these elements. This allows the minimal IncHI2 plasmid backbone to be refined, as IncHI2s do not necessarily contain these heavy metal resistance regions, which had previously been considered a characteristic of IncHI2 plasmids. This minimal backbone, as would be found in the ancestor plasmid, Plasmid X, is shown in Fig 4.19.

The copper and silver resistance transposon Tn*6024* is particularly interesting as it is large (32.4 kb) and is found in 3 other locations, also surrounded by direct repeats. To the best of the author's knowledge, this is the first silver and copper resistance transposon of this size to be identified. Plasmids pSRC125 and pSRC26 contain a sequence equivalent to removing Tn*6024* and one copy of the direct repeat from R478. Another instance of identical sequence with and without Tn*6024* was also found in *E. coli* chromosomes. Collectively, this provides strong evidence that the element had moved and is, in fact, an active transposon.

Transposition assays, detailing the movement of Tn*6024*, would be ideal to perform in the future, to confirm its ability to move. Previously, the mobilisation genes on R478 had been annotated as "*tnsA-D"* from Tn*7*. However, analysis from the current study found the protein sequences encoded by these genes were only ~30% identical to those in Tn*7* and hence were described as "Tn*7*-like". Thus, both Tn*7* and Tn*6024* are part of a large and diverse family of "Class III" transposons, which have similar characteristics: 22-25 bp IRs, create 5 bp direct repeats and have transposase binding sites, (described in 1.3.1.3). The arsenic resistance Tn, does not fit the definition of a Class I, Class II or "Class III" Tn and it has no relatives in GenBank, so may represent yet another class of Tns. It becomes apparent that a more rigorous classification system of transposons is needed, which expands on the restricted definition of Class I and II transposons, to define these novel elements.

Tellurite compounds have been used as therapeutic agents and for selective agents for the isolation of pathogens [255]. Tellurite genes are in a separate region to the other heavy metal resistance genes in R478, and have extensive similarity to chromosomal sequences, indicating that the Te^R region may also be on an MGE. However, when this region was examined, the boundaries of this potential MGE could not be found. In fact, in this study (see Fig 4.19) and across a number of previous studies [161];[193], the tellurite resistance genes in IncHI2 plasmids have been shown to be present in every IncHI2 plasmid analysed to date. This indicates that the Te^R genes are deeply embedded in the IncHI2 backbone. Perhaps this MGE has lost the ability to move and it may have lost the characteristic features of MGEs, such as IR or DRs, making it difficult to detect.

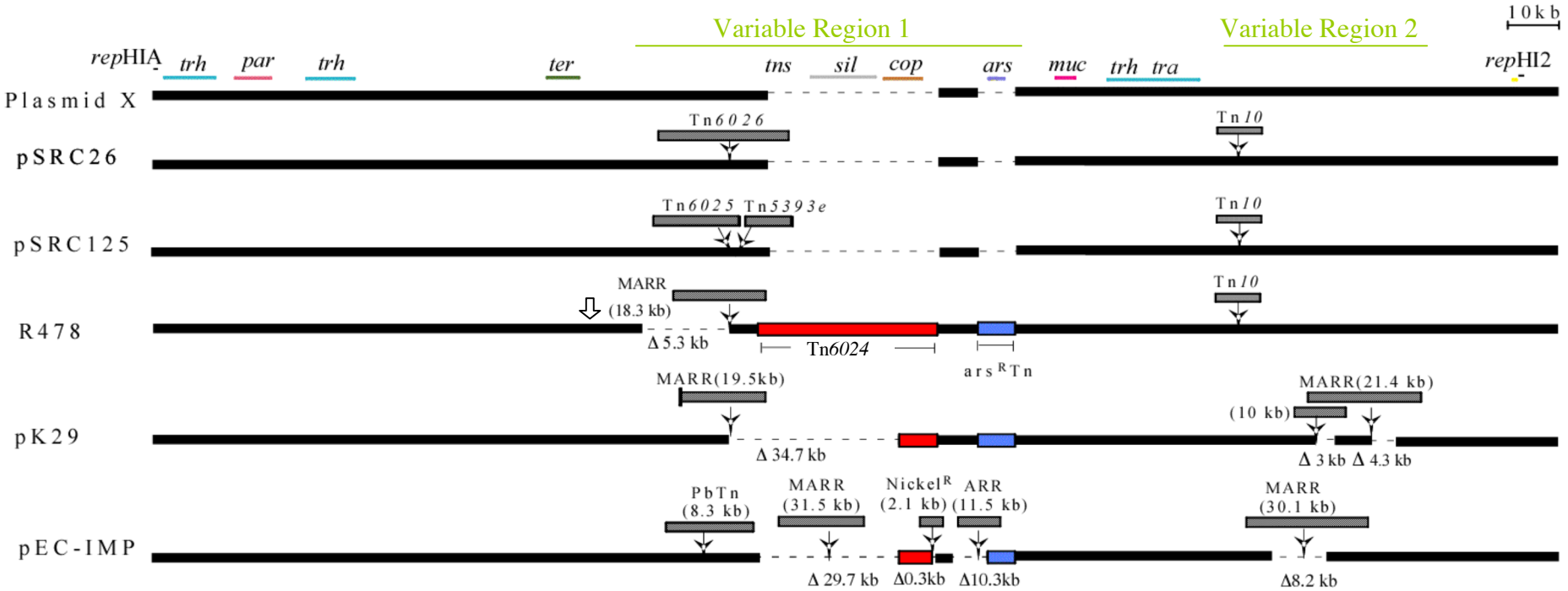


Fig 4.19 Regions present in IncHI2 plasmids. Structures of R478, pSRC26, pSRC125, pK29 and pEC-IMP are shown as a horizontal line, some features of which are shown above the R478 line. Tn*6024* or parts thereof are shown as a red box and the ars^R Tn a blue one. Dotted lines represent missing sequence compared to R478. Insertions of other regions, such as MARRs (muitple antibiotic resistance regions) or ARR (antibiotic resistance region) are shown above the backbone as grey boxes with their sizes in brackets. These variations are contained within 2 major "variable regions", the locations of which are marked above in green. The open arrow indicates the location of a 4.1 kb MGE also present in IncHI1 plasmids.

Although 3 MGEs (Tn*1696*, Tn*6024* and the ars^R Tn) were identified in the backbone of the IncHI2 plasmids, during subsequent analysis undertaken on the IncHI1 plasmid pSRC27-H (see section 7.3.3), a fourth MGE was identified. It is 4.1 kb, contains 23 bp IRs and is surrounded by 5 bp DRs. It is in a different location in the IncHI1 and IncHI2 plasmids. In R478 and pK29 it is present 1.9 kb to the right of the tellurite resistance region. Whether is it is in pSRC125 and pSRC26 was not investigated. In the future, it may be advantageous to develop a PCR across the boundaries of this MGE and use these determine whether it is present in pSRC125 and pSRC26. It may be predicted that, because these plasmids seem to be in an ancestral configuration, they would not contain it.

CHAPTER FIVE

Further IncHI2 plasmids

5.1 Introduction

A set of 8 multiply antibiotic resistant *Salmonella enterica* serovar Infantis stains, isolated from chicken, chicken meat, feline and canine sources in Australia in 2000-2001, was obtained from the Microbiological Diagnostic Unit, via Dr. Steve Djordjevic's laboratory at the EMAI, and examined. Their IS*200* profiles were known to be identical, indicating that these strains are clonally related [230]. Previously, their antibiotic resistance phenotype had been determined and all were found to carry the *tetA*(B) tetracycline resistance gene, as well as a class 1 integron containing the unusual *dfrB6-aadA1* cassette array ([230]; Table 5.1). Another *S*. Infantis strain named SRC83, was subsequently found to share these characteristics (analysed by Dr. Neil Wilson, Hall laboratory) and was added to this set (Table 5.1). SRC83 was originally incorrectly grouped with another set of *S. enterica* strains that were screened for plasmid content using PBRT [145] and were found to contain IncHI2 plasmids. This raised the possibility that the other *S*. Infantis strains carried IncHI2s also.

Table 5.1 Properties of *S*. Infantis strains[1]

Strain	Source	Year	Resistance profile	*intI1*/*sul1*	Cassettes	Other resistance genes
SRC83[2]	Chicken	2000	SmSpSuTcTp	+/+	*dfrB6-aadA1*	*tetA*(B)
SRC70	Feline	2001	SmSpSuTc	+/+	*dfrB6-aadA1*	*tetA*(B)
SRC71	Chicken	2001	SmSpSuTc	+/+	*dfrB6-aadA1*	*tetA*(B)
SRC72	Chicken	2001	SmSpSuTc	+/+	*dfrB6-aadA1*	*tetA*(B)
SRC92	Canine	2000	SmSpSuTcTp	+/+	*dfrB6-aadA1*	*tetA*(B)
SRC93	Chicken	2000	SpSuTc	+/+	*dfrB6-aadA1*	*tetA*(B)
SRC94	Chicken	2000	SmSpSuTc	+/+	*dfrB6-aadA1*	*tetA*(B)
SRC95	Chicken mince	2000	SmSpSuTc	+/+	*dfrB6-aadA1*	*tetA*(B)
SRC96	Chicken carcass	2000	SmSpSuTc	+/+	*dfrB6-aadA1*	*tetA*(B)

[1]from [230]
[2]Information for strain received from Neil Wilson, Hall laboratory, personal communication

Here, the PCR-based tools developed in Chapter 3 were used to detect antibiotic resistance regions in plasmids from the *S*. Infantis strain set and to determine if they are in the same position as in pSRC125 and pSRC26. Further PCR sets designed in Chapter 4, were used to examine the IncHI2 backbone and heavy metal resistance regions. The aim of this chapter was to utilise these tools to efficiently characterise a further set of IncHI2 plasmids and,

additionally, to determine how related these plasmids are to each other and to the IncHI2 plasmids examined previously, in particular pSRC26, pSRC125 and R478.

5.2 *S.* Infantis isolates

Because the *S*. Infantis strain SRC83 had previously been shown to contain an IncHI2 plasmid, the IncHI2 simplex PCR was performed on the other *S*. Infantis strains listed in Table 5.1, revealing that all but one, SRC70, was positive for the IncHI2 replicon and thus contained an IncHI2 plasmid (data not shown). The resistance phenotypes of the *S*. Infantis strains were confirmed by picking and patching colonies onto agar plates containing the antibiotics: Sm (25 μg/mL), Sp (25 μg/mL), Su (100 μg/mL), Tc (10 μg/mL) and Tp (25 μg/mL). All resistance phenotypes were the same as previously determined by Levings et al., (2006) [230], except that SRC70 was no longer tetracycline resistant (see Table 5.2, below). Although all stains contained *dfrB6,* which confers low-level trimethoprim resistance [230], it was not scored as resistant, as growth did not occur at 25 μg/mL. The *tetA*(B) gene was present in all the *S*. Infantis strains except for SRC70, which no longer contained this gene (Table 5.2, below). The tetracycline resistant version of SRC70 could not be recovered from the stock sample originally obtained from EMAI. This suggests that SRC70 had lost the tetracycline resistance gene since the previous study in 2006 and that this must have occurred during storage of the strain (see section 5.4.5). Mercuric ion resistance was also tested by patching onto LA plates containing mercuric chloride (20 μg/ml) and except for SRC72, all strains were resistant, indicating that they contain a mercury resistance region. This also indicated that although SRC70 had lost *tetA*(B) and *repHI2*, it retained mercury resistance, and so perhaps it contained part of an IncHI2 plasmid.

5.3 Structure and location of resistance regions

To characterise the antibiotic resistance regions on the *S*. Infantis IncHI2 plasmids, PCRs targeting Tn*10* and transposons related to Tn*1696*, which are common to pSRC125, pSRC26 and R478, were performed on plasmid DNA from the *S*. Infantis strains.

*5.3.1 Tn*10

The presence of Tn*10* was tested using 3 overlapping mapping PCRs (as described in Fig 3.10A) and except for SRC70, all strains produced 3 amplicons of correct size, indicating that contained Tn*10* (Table 5.2). SRC70 had lost Tn*10* and the resistance gene *tetA*(B), which is consistent with loss of tetracycline resistance. In all other isolates, Tn*10* was in exactly the same position (Table 5.2), as it is in R478, pSRC125 and pSRC26 (see Fig 3.10B).

Table 5.2 Tn*10* resistance genes and linkages

Plasmid/strain	Resistance profile	*tetA*(B)	Tn*10*	Tn*10* LHJ[1]	Tn*10* RHJ[1]
R478	Cm, Km, Nm, Tc	+	+	+	+
SRC83	Sm Sp Su Tc	+	+	+	+
SRC70	Sm Sp Su[2]	-	-	-	-
SRC71	Sm Sp Su Tc	+	+	+	+
SRC72	Sm Sp Su Tc	+	+	+	+
SRC92	Sm Sp Su Tc Tp	+	+	+	+
SRC93	Sp Su Tc	+	+	+	+
SRC94	Sm Sp Su Tc	+	+	+	+
SRC95	Sm Sp Su Tc	+	+	+	+
SRC96	Sm Sp Su Tc	+	+	+	+

[1]PCR that detects no Tn*10* at this position in the backbone was negative for all strains
[2]Phenotype previously determined to be Tc^R [230]

*5.3.2 Tn*1696*-like*

The mercuric ion resistance gene *merA* was shown to be present in all *S*. Infantis strains, except for SRC72, using PCR. Digestion of the amplicons with RsaI confirmed that the *merA* amplicon was of Tn*1696* type, when compared to a Tn*1696* positive control. The location of the class 1 integron was then tested using the PCR that links *intI1* to the *tnpR* gene, in a range of Class II mercury transposons, described in Chapter 3 (see Fig 3.2). Each of the strains

produced amplicons of the size that corresponded to Tn*1696* thus *intI1* was linked to $tnpR_{1696}$. PCR linking the IS*6100* end of the integron to the *mer* module (see Fig 3.3) was used to show that in all strains, except for SRC72, IS*6100* was linked to *merE* and that the second, partial copy of IS*6100* was present, as in Tn*1696* (Table 5.3). The IS*6100* was shown to be present in all samples using PCR internal to IS*6100* (Table 5.3).

Table 5.3 Presence and position of Tn*1696*-like Tns

Plasmid/ strain	$merA_{1696}$	*merE*-IS*6100*/ *6100Δ*	IS*6100*	*intI1*/ $tnpR_{1696}$	Tn*1696*-like[1] *tnp*	*mer*	Tn*1696*-like (*tnp*) LHJ	Tn*1696*-like *(mer)* RHJ
pSRC125	+	+/-	+	+	+	+	+	+
pSRC26	+	+/+	+	+	+	+	+	+
R478	+	+/+	P[2]	+	-	+	-	+
SRC83	+	+/+	+	+	+	+	+	+
SRC70	+	+/+	+	+	+	+	+	+
SRC71	+	+/+	+	+	+	+	+	+
SRC72	-	-/-	+	+	+	-	+	-
SRC92	+	+/+	+	+	+	+	+	+
SRC93	+	+/+	+	+	+	+	+	+
SRC94	+	+/+	+	+	+	+	+	+
SRC95	+	+/+	+	+	+	+	+	+
SRC96	+	+/+	+	+	+	+	+	+

[1]PCR that detects no Tn*1696*-like was also negative for all
[2]P= Partial; IS*6100* truncated by IS*26*

The presence of the *tnp* and *mer* modules of Tn*1696* was tested using the mapping strategy developed for pSRC125 (described in Fig 3.4). The plasmid in SRC72 was missing the entire mer_{1696} module and this was the only strain to be sensitive to mercuric ions (see Fig A2.1 in Appendix 2). However, the *tnp* module, integron and the IS*6100* remained. All other *S*. Infantis plasmids carried the entire Tn*1696*-like region (Table 5.3). The structure of this Tn*1696*-like transposon in the *S*. Infantis IncHI2 plasmids, which contained the *dfrB6-aadA1* cassette array, was generated based on these mapping PCR results and was named Tn*6028* (Fig 5.1). However, Tn*6028* was not sequenced.

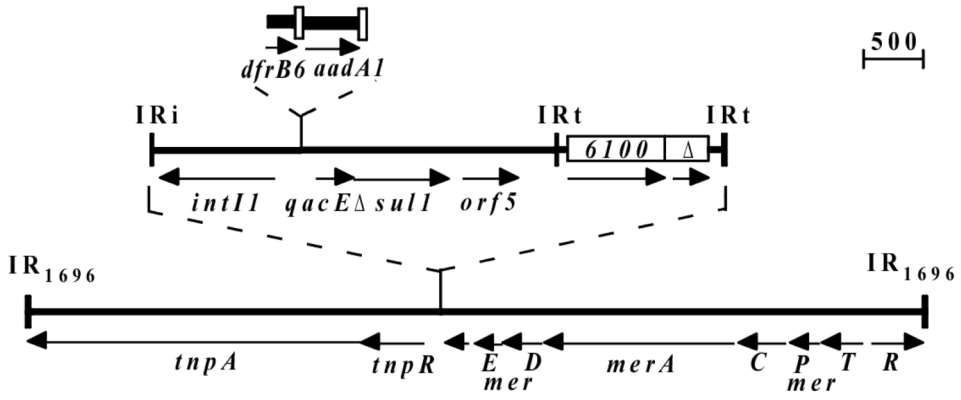


Fig 5.1 Tn*6028* The open, numbered box represents an insertion sequence of that number and the Δ represents a partial copy of IS*6100*, while arrows represent direction and extent of genes. Vertical bars are inverted repeats (IR) from either transposons, where the transposon number is subscripted, or from integrons where IRi (inverted initial repeat) denotes the IR from the 5' conserved segment (5'-CS) end of the integron, or IRt (inverted terminal repeat) for the IR from the 3' conserved segment (3'-CS) end. Figure drawn to scale based on GenBank accession number GQ150541 and sequence obtained in this study. Scale bar given in bp.

Tn*6028* was shown to be in the same position as Tn*6026* and Tn*6025* in pSRC26 and pSRC125 (Table 5.3) using previously described PCRs (Fig 3.8). This suggests that these plasmids share a common ancestor. The plasmid in SRC72 was missing the right-hand junction of Tn*1696*, which is consistent with it having lost the mercuric ion resistance module. Thus, the segment missing in pSRC72 extends from the right-end of IS*6100* to IR_{mer} and possibly beyond. This could possibly be a result of an IS*6100*-mediated adjacent deletion.

Tn*6028* differed from both Tn*6025* and Tn*6026* in its unusual cassette array and from Tn*6025* as it contained IS*6100*Δ. Like Tn*6025,* Tn*6026* and Tn*1696*, Tn*6028* contained a PcS promoter, indicating that it had probably undergone cassette exchange via homologous recombination after insertion into the IncHI2 backbone, as described in section 3.3.2.

5.4 Mapping the remainder of the plasmids

The presence of features in the IncHI2 backbone and the heavy metal resistance regions in the *S*. Infantis plasmids was tested using PCR, heavy metal resistance testing, RFLP and conjugation.

5.4.1 Mapping PCRs

The set of 15 mapping PCRs designed in Chapter 4 (see Fig 4.3) was performed on the *S*. Infantis strains (Table 5.4; Fig 5.2). SRC70, which had been shown to be missing Tn*10* (Table 5.2), appeared to also be missing a large amount of backbone, as only PCRs 2, 3, 5, 9, 10 and 12, produced amplicons. Only SRC94 gave the same PCR results as pSRC125 and pSRC26, where the *tnsD, sil, cop* and *ars* PCRs were negative. The remaining *S*. Infantis plasmids had identical results: all PCRs, except for those targeting the *sil* or *tnsD* genes, produced amplicons of correct size. Interestingly, they did contain copper and arsenic resistance genes. This may indicate that they contain part of Tn*6024* and the arsenic resistance transposon.

5.4.2 Heavy metal resistance

The *S*. Infantis strains were tested for arsenic and tellurite resistance, using the procedure described in section 4.2.3. SRC94 was resistant to tellurite but sensitive to arsenite and arsenate (highlighted blue in Table 5.5), and this was the only strain that did not contain the *ars* resistance gene. The remaining strains were arsenite/arsenate and tellurite resistant (Table 5.5). All phenotypic results reflected the expected results, based on the presence or absence of the resistance genes detected by PCR.

Table 5.4 Backbone PCR on *S*. Infantis strains

PCR target, size (bp), number[1]	*rep*$_{HIA}$ (1150) 7.	ORF25 (2421) 8.	ORF56/57 (2217) 9.	Te^R (893) 2.	ORF92/93 (1001) 3.	*HipA* (1553) 10.	*tnsD* (1466) 4.	Silver^R (1704) 11.	Cu^R (512) 12.	Ars^R (1136) 5.	*muc* (1276) 13.	ORF201 (1011) 6.[2]	IS*186*/ bbone 15.	ORF270/271 (2966) 14.	rep$_{HI2}$ (647) 1.
R478[1]	+	+	+	+	+	-	+	+	+	+	+	+	+	+	+
pSRC125[1]	+	+	+	+	+	+	-	-	-	-	+	+	+	+	+
pSRC26[1]	+	+	+	+	+	+	-	-	-	-	+	+	+	+	+
SRC83	+	+	+	+	+	+	-	-	+	+	+	+	+	+	+
SRC70[2]	-	-	+	+	+	+	-	-	+	+	-	-	-	-	-
SRC71	+	+	+	+	+	+	-	-	+[3]	+[3]	+	+	+	+	+
SRC72[4]	+	+	+	+	+	+	-	-	+[3]	+[3]	+	+	+	+	+
SRC92	+	+	+	+	+	+	-	-	+	+	+	+	+	+	+
SRC93	+	+	+	+	+	+	-	-	+[3]	+[3]	+	+	+	+	+
SRC94	+	+	+	+	+	+	-	-	-	-	+	+	+	+	+
SRC95	+	+	+	+	+	+	-	-	+[3]	+[3]	+	+	+	+	+
SRC96	+	+	+	+	+	+	-	-	+[3]	+[3]	+	+	+	+	+

[1]data from Chapter 4
[2]Plasmid in SRC70 does not contain Tn*10*
[3]PCR also performed on *E. coli* transconjugants and the same result was obtained
[4]Plasmid in SRC72 does not contain the *mer* module of Tn*6028*

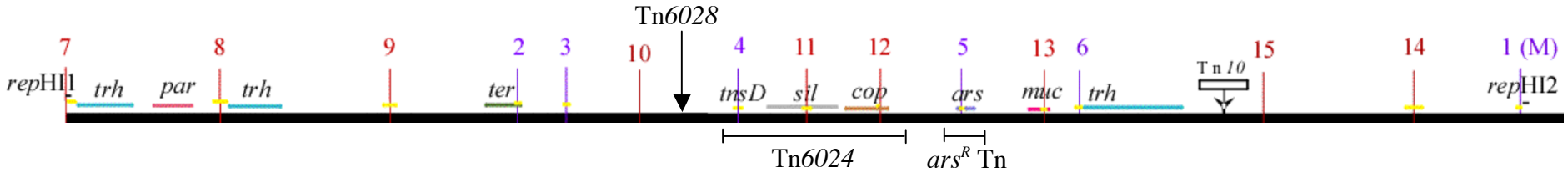


Fig 5.2 Position of IncHI2 backbone PCRs The predicted backbone structure of an IncHI2 plasmid from an *S*. Infantis strain is shown as a horizontal line, some features of which are shown above including the transfer (*trh*) partitioning (*par*) replication (*rep*) and heavy metal resistance regions (*sil*, *cop*, *ars*, *ter*). Vertical lines represent positions of PCRs numbered 1-15, the ones designed in this study are in red and the ones used previously [161] are in purple. Positions of MGEs are marked by vertical arrows above the backbone.

Table 5.5 Heavy metal resistance phenotypes

Plasmid	Host strain	Arsenite		Arsenate		Tellurite	
		Expected[1]	Tested[2]	E	T	E	T
R478	**E294**	**+**	**+**	**+**	**+**	**+**	**+**
pSRC125	**E294**	-	-	-	-	+	+
pSRC83[3]	*S*. Infantis	+	+	+	+	+	+
pSRC70	*S*. Infantis	+	+	+	+	+	+
pSRC71	*S*. Infantis	+	+	+	+	+	+
pSRC72	*S*. Infantis	+	+	+	+	+	+
pSRC92	*S*. Infantis	+	+	+	+	+	+
pSRC93	*S*. Infantis	+	+	+	+	+	+
pSRC94	*S*. Infantis	-	-	-	-	+	+
pSRC95	*S*. Infantis	+	+	+	+	+	+
pSRC96	*S*. Infantis	+	+	+	+	+	+
-	E294[4]	-	-	-	-	-	-
-	DH5α[4]	-	-	-	-	-	-
-	UB5201[4]	-	-	-	-	-	-
-	UB1637[4]	-	-	-	-	-	-

[1]Expected from presence of genes detected using PCR
[2]Tested in at least 2 replicates
[3]Plasmid pSRC83 is the IncHI2 plasmid in SRC83 and so on for the other strains
[4]*E. coli* strains as negative controls

To determine if the copper/silver resistance transposon, Tn*6024*, and the arsenic resistance transposon were present, PCRs across the junctions of where they are inserted in R478 were performed as described in section 4.4.

5.4.2.1 Arsenic resistance transposon

Primers in the *ars* transposon used with primers in the backbone (Fig 5.3) showed that both the left and right hand junctions of the transposon were present for all *S*. Infantis strains, except SRC94 (Table 5.6). Thus, the majority of the *S*. Infantis strains isolated contained the ars^R transposon at the same position as it is in R478. However, the exact position of this Tn could not be determined, as these amplicons were not sequenced.

Table 5.6 Arsenic resistance transposon PCR results

Plasmid	*ars*^{R} gene	*ars*^{R} Tn RHS	*ars*^{R} Tn LHS	Empty *ars*^{R} Tn insertion site
R478	+	+	+	-
pSRC26, 125[1]	-	-	-	+
pSRC70, 72, 92, 93, 95, 96	+	+	+	-
pSRC94	-	-	-	-

[1]data from Chapter 4

PCR across the site where the ars^R Tn had inserted in R478 did not produce amplicons for any of the strains, consistent with the transposon being present (using primers RH1269 and RH895 as in Fig 5.3; Table 5.6). However, this PCR was also negative for SRC94, the strain that does not contain the *ars*^{R} genes (Table 5.6). Thus, this indicates that although SRC94 is missing the arsenic resistance region, it is not simply missing the transposon, but instead seems to be missing the surrounding regions as well.

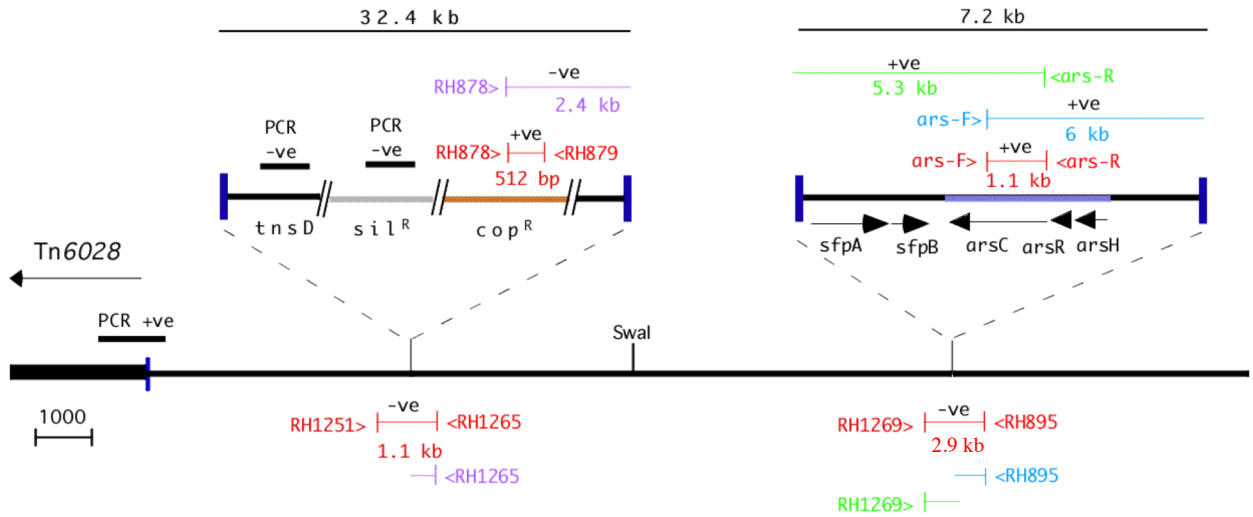


Fig 5.3 PCRs surrounding the heavy metal resistance transposons. The position of PCR primers are shown as vertical coloured lines, with the size of the product and result from the *S*. Infantis plasmids shown below and above respectively. Fig drawn to scale, based on the sequence of R478 (GenBank acc no. BX664015). The scale bar is shown in bp. The double slashed lines represent sequence not shown and are not to scale. Other features are as in Fig 5.1. Primer sequences are given in Tables A3. 7 and A3.8 in Appendix A3.

*5.4.2.2 Tn*6024

The silver resistance gene and *tnsD* in Tn*6024* was absent in all strains. However, the copper resistance gene remained in all strains, except SRC94 (Table 5.4). To detect if Tn*6024* was absent, a PCR was performed across its insertion site in R478, so that a 1.1 kb amplicon would be produced if Tn*6024* was missing (Fig 5.3). This PCR was negative for all strains, indicating that Tn*6024* was not simply absent. Instead either sequence was missing surrounding the insertion site or Tn*6024*, or part thereof, was present. A primer in *copS* (from PCR 12 in Table 5.4) was used with a primer to the right of the Tn*6024* insertion site in R478

to determine if the right hand junction of Tn*6024* was present (RH878 and RH1265 in Fig 5.3). Amplicons were not produced for any of the strains. These results could be explained in a number of ways. First, Tn*6024*, or part thereof, is present, but the region containing the PCR primer/s is missing. Second, Tn*6024*, or part thereof is present, but it is located in a different position to R478 and so would not have the predicted RH junction. A further possibility is that, the region surrounding Tn*6024* is missing in these plasmids and that the copper resistance gene is elsewhere, or a more complex series of events has occurred. Further work is needed to determine exactly what the configuration of this region is in these plasmids, such as PCRs to map the entire Tn*6024* or cloning the heavy metal resistance genes and sequencing their surrounding sequence.

5.4.3 Conjugation

The IncHI2 plasmids were conjugated at 37^{O}C from *S.* Infantis into *E. coli* recipient strain E294, using Sp and Rf selection. Transconjugants were obtained from SRC71, SRC72, SRC93, SRC95 and SRC96, named E294/pSRC71, E294/pSRC72, E294/pSRC93, E294/pSRC95 and E294/pSRC96. They were confirmed to contain IncHI2 plasmids using the IncHI2 simplex PCR. These plasmids conjugated at frequencies (as listed in Table 5.7) of the same order of magnitude as pSRC125 and pSRC26 (~3-5 x 10^{-5} transconjugants per donor).

Table 5.7 Conjugation frequencies

Strain SRC No.	Conjugation frequency (transconjugants per donor)[1]
SRC83	<1 x 10^{-10}
SRC70	<1 x 10^{-10}
SRC71	1.4 x 10^{-5}
SRC72	1.5 x 10^{-5}
SRC92	<1 x 10^{-10}
SRC93	3.1 x 10^{-5}
SRC94	<1 x 10^{-10}
SRC95	2.4 x 10^{-5}
SRC96	5.1 x 10^{-5}

[1]conjugations performed at least in triplicate

These transconjugants were tested for resistance phenotype by picking and patching onto antibiotic-containing plates, revealing that all the resistances of the parent strains (Table 5.1) transferred. Transconjugants from SRC83, SRC70, SRC92 and SRC94 could not be obtained. This suggests that the plasmids from these strains are not conjugative, under the conditions tested, and their regions of transfer may be missing or inactivated. This is perhaps an expected result for SRC70, as its plasmid is missing the majority of the IncHI2 backbone, including regions surrounding the transfer regions (PCRs 6, 7, 8 in Fig 5.2; Table 5.4). However, the PCRs located near the transfer regions (PCRs 6, 7 and 8 in Fig 5.2) were positive for the remaining strains (Table 5.4). PCR that specifically targets the transfer regions were not included within the PCR set. Thus, whether an entire, intact transfer region is present, could not be determined. This was not investigated further, but in the future, these conjugation regions should be mapped.

5.4.4 Restriction mapping

The high quality large plasmid DNA required for RFLP and PFGE could only be extracted from the *E. coli* strain E294, but not from *Salmonella* strains. Thus, plasmid DNA was only extracted from the transconjugants obtained, as shown in Table 5.6. They were digested with the restriction enzyme SwaI and separated by PFGE, as was described in section 4.3. The sizes of the observed bands are provided in Table A4.1 in Appendix 4. The bands shared with R478 are marked with blue arrows in Fig 5.4A and were identified as bands 1-5, 11 and 13-15 (Fig 5.4B). However, there were clearly more bands than expected. Bands 3, 4 and 14 were too small to be visualised using this method, but were assumed to be present, as surrounding bands were detected. Some variation was observed between the *S*. Infantis plasmids within these bands. Band 15, which is usually 25.8 kb, is approximately 2 kb larger in pSRC92 and

pSRC95 (boxed red in Fig 5.4A). Band 16 (10.2 kb), is missing in pSRC72, and band 2 (10 kb), is missing in pSRC95 (boxed green in Fig 5.4A). This data is tabulated in Table 5.8.

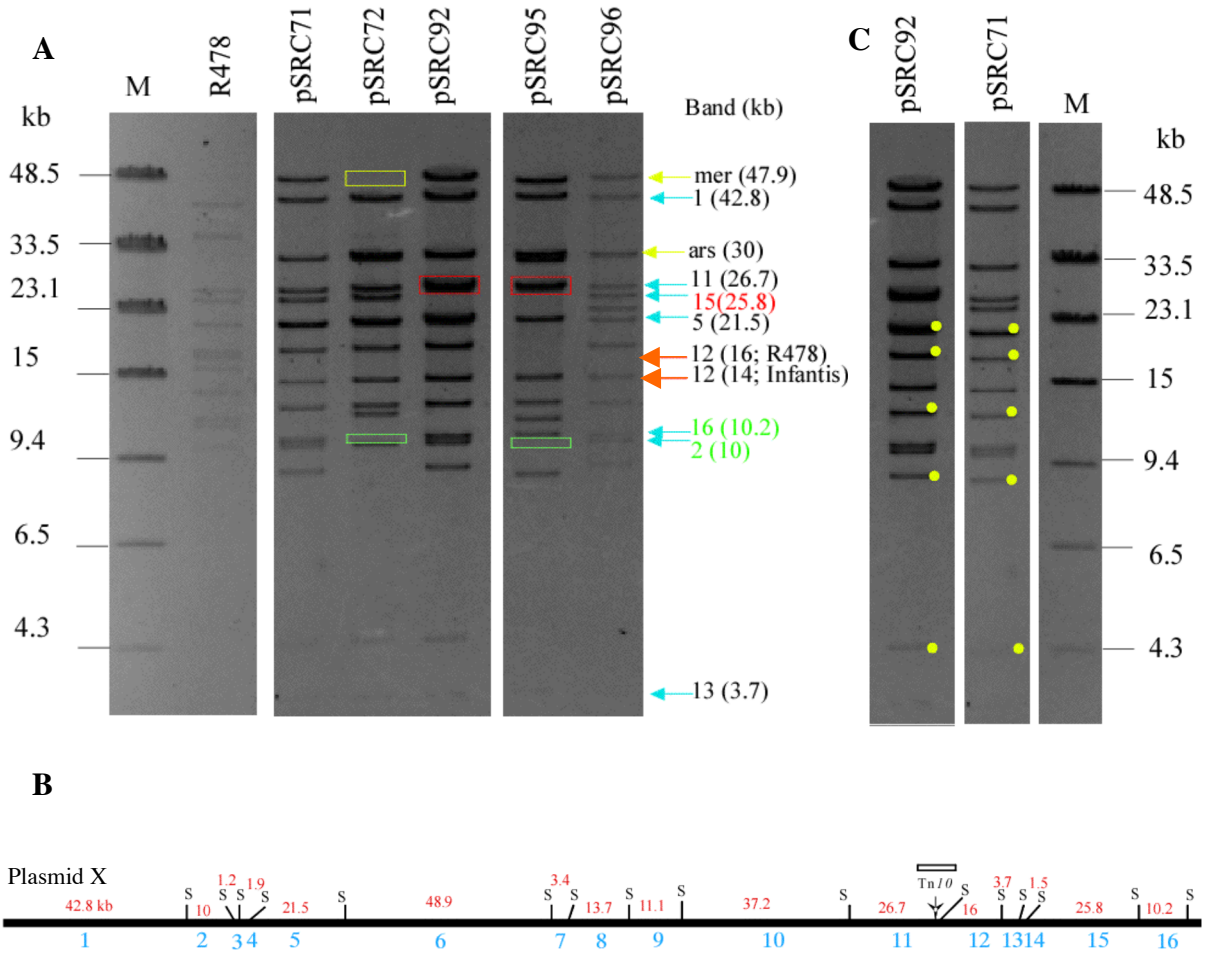


Fig 5.4 PFGE on *S.* Infantis transconjugants. A Expected, shared and unique bands. The horizontal arrows indicate the position of the predicted amplicons for the bands that could be accounted for. Bands 1-5 and 11-16, from the invariable backbone regions, are marked with a blue arrow; band 12, which varies from R478 is marked with a red arrow for both R478 and the *S.* Infantis plasmids; a yellow arrow indicates the variable heavy metal resistance region-containing bands 6-10. Bands that vary between the S. Infantis plasmids are boxed, red for bigger than expected size and green for absent **B.** Plasmid X SwaI restriction map. The band numbers are below in blue and the sizes are on top in red. Tn*10* is shown as an open box above. **C** Unexpected bands in representative plasmids pSRC92 and pSRC71. Bands are marked with yellow dots for pSRC71 and pSRC92. In A and C, lanes are labelled R478, pSRC71, pSRC72, pSRC92, pSRC95 or pSRC96 for the *E. coli* strain E294 transconjugants containing those plasmids. PFGE molecular weight ladders are loaded in lane M and sizes are as indicated.

In Chapter 4, band 12 was found to vary from 16 kb in R478 to 14 kb in pSRC125 and 18 kb in pSRC26, possibly due to the presence of insertion sequences, such as IS*186* and IS*150*.

Because all the *S*. Infantis plasmids share a band of ~14 kb which is the size of band 12 in pSRC125 (marked with an orange arrow in Fig 5.4A), this was allocated as band 12. However, this allocation was rather arbitrary, as the band 12 size of pSRC26 was ~18 kb and pSRC71, pSRC72 and pSRC92 have a band of approximately this size (Table 5.8).

Table 5.8 Summary of bands shared by R478

Band number[1]	**Contents of band**	**PCR no.**	**Observed bands**						
			R478 (kb)	***S*. Infantis (kb)**	**SRC71**	**SRC72**	**SRC92**	**SRC95**	**SRC96**
6 (*mer*)	Tn*6028*		**47.9**		+	-	+	+	+
1	*repHIA; trh*	7, 8	**42.8**	**42.8**	+	+	+	+	+
10 (ars)	*ars; muc; trh*	5, 13, 6	**37.2**	**~30**	+	+	+	+	+
11	Tn*10*		**26.7**	**26.7**	+	+	+	+	+
15	orF270	14	**25.8**	**25.8**	+	+	+ [2]	+ [2]	+
5	orf 56/57	9	**21.5**	**21.5**	+	+	+	+	+
12	IS*186*	15	**16**	**14**	+	+	+	+	+
16	*repHI2*	1	**10.2**	**10.2**	+	-	+	+	+
2			**10**	**10**	+	+	+	-	+
13			**3.7**	**3.7**	+	+	+	+	+

[1]Bands 3, 4, 14 too small to be visualised using this method
[2]size of band increased by ~2 kb in these samples

Two large bands containing the heavy metal resistance regions were expected in the *S*. Infantis plasmids. The predicted size of the fragment from the last known SwaI site, to the IR_{mer} of Tn*6028* is 37 kb, and the only band unaccounted for and greater than 37 kb was estimated to be 47.9 kb (marked with a yellow arrow in Fig 5.4A). This size is ~2 kb bigger than the predicted fragment size if Tn*6024* was not present. SRC72, which is the only strain with a plasmid that does not contain mer_{1696} or confer mercury resistance, is the only strain missing this large band (boxed in yellow in Fig 5.4A), confirming that the 47.9 kb fragment contains Tn*6028*. The only band big enough to contain the *ars* resistance region is estimated to be 33 kb (marked with a yellow arrow in Fig 5.4A). This is approximately 4 kb smaller than the band in R478. Thus further events must have occurred within the region represented by this band, making it smaller. This was not investigated further. These results are summarised in Table 5.8 and unexpected differences in them are highlighted using the same colouring as the boxes or arrows used in Fig 5.4A.

All bands common to R478 and those containing the heavy metal and antibiotic resistance regions had thus been identified (Table 5.8). However, many observed bands were left unaccounted for. These additional bands were examined in pSRC71 and pSRC92 as examples, because they display the same pattern of additional banding (marked with yellow dots in Fig 5.4C). The extra bands were estimated to be 21.9 kb, 18.7 kb, 13.3 kb, 12.2 kb, 8.8 kb, 4.5 kb, which add up to an approximate size of ~80 kb. If this additional DNA is within the IncHI2 plasmids, it must be positioned within band 12, or within the fragments containing the *ars* or *cop/sil* resistance regions, as these fragments yielded bands of sizes different to what was predicted. Because IS*26* contains a SwaI site and is often associated with additional regions, an internal IS*26* PCR was performed on the *S*. Infantis strains. All transconjugants and *Salmonella* strains contained IS*26*, which indicates that there are further SwaI sites that have not been accounted for. Additional IS*26*-containing regions have been shown to reach large sizes (see Fig 3.19). Thus, it is possible that an additional ~80 kb IS*26*-containing region is present. Alternatively, the extra DNA could be a result of another plasmid of ~80 kb co-transferring into the recipient cell, or forming a cointegrate fusion with the IncHI2 plasmid.

5.4.5 pSRC70

pSRC70, the IncHI2 plasmid in SRC70, was found to be missing the majority of its backbone, including both replication regions (Table 5.4). Further investigation was carried out on this strain, to determine how pSRC70 replicates. A previous study had observed that IncHI2 plasmids missing their transfer regions had formed cointegrate fusions with co-resident IncFIB plasmids, which allowed transfer of the entire cointegrate [193]. Further examples of IncHI2 and IncFII cointegrate formation have been described [256]. The presence of other plasmids in SRC70 was investigated, as they could give rise to the formation of a cointegrate fusion with pSRC70, providing the replication region. The additional ~80 kb of DNA

observed in the other *S*. Infantis plasmid transconjugants (Fig 5.4C) is consistent with the presence of a co-resident plasmid. The full set of multiplex PCRs from the PBRT scheme, as well as additional PCRs that target IncF replicons (developed by Sheree Yau, Hall Laboratory), was performed on SRC70, but amplicons were not obtained. Therefore, no additional plasmid could be detected using this method. Long-range PCR, which attempted to link outwardly-facing primers from the last 2 PCRs for which SRC70 produce amplicons (ars^R and orf56/57 in Table 5.2), also failed (Fig 5.5). This indicates that pSRC70 is not simply missing the region in between *ars* and orf56/57, which would result in a PCR product.

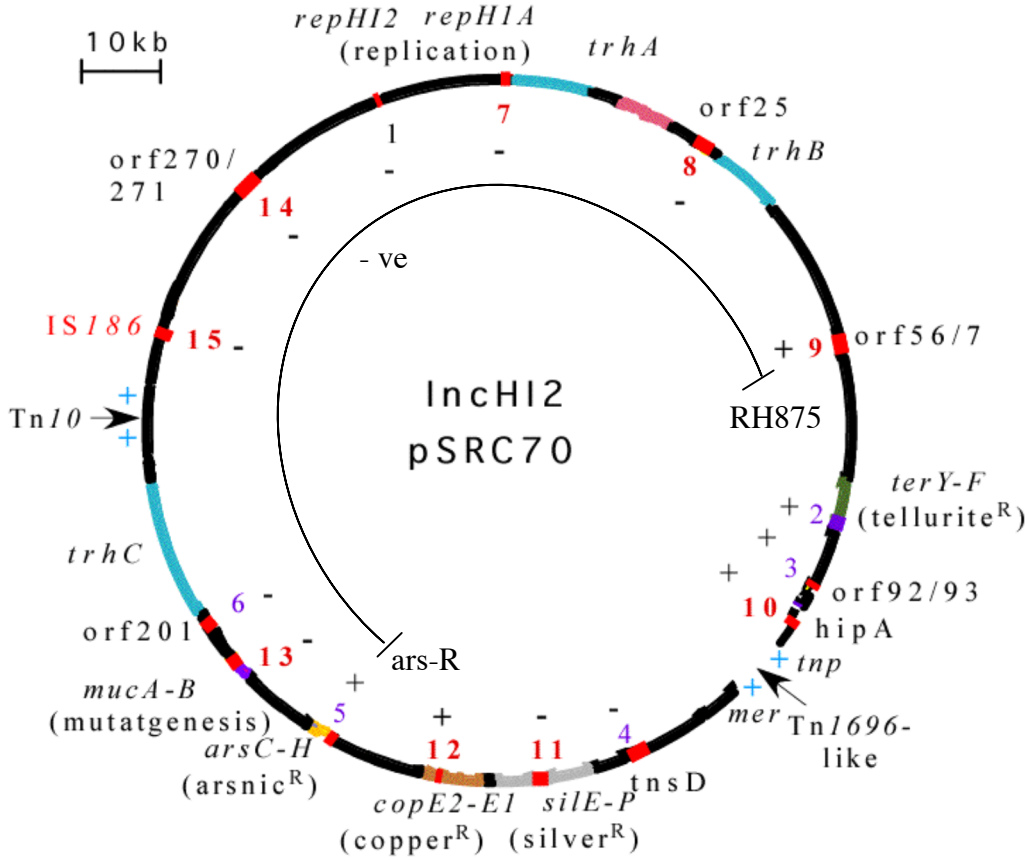


Fig 5.5 pSRC70 The IncHI2 backbone sequence is shown in a circular form. Important backbone features are shown: regions for replication (red), transfer (blue); partitioning (pink), mutagenesis (purple), and regions that give resistance to the heavy metals tellurite (green), mercury (orange), silver (silver), copper (copper), and arsenic (yellow) and the 2 antibiotic resistance regions – Tn*1696*-like and Tn*10*. The numbers represent the positions of PCR primers used to examine the IncHI2 backbone and the PCR results for pSRC70 are indicated with a + or -. PCR primers ars-R and RH875 are shown as lines and the arc between them represents the predicted PCR product, which was negative - indicated by "-ve". Figure is drawn to scale using R478 GenBank accession nos BX664015 and pK29 EU382672.

Because of difficulties extracting high quality plasmid DNA from *Salmonella*, manipulation of this plasmid was difficult and due to time constraints, further analysis of pSRC70 was not undertaken. However, further experiments will be needed to investigate the possibility of the presence of another replication region, as well as experiments to test the presence of additional plasmid/s using other methods, such as S1-nuclease digestion.

5.5 A small set of older IncHI2-containing isolates

A small number of *E. coli* transconjugants containing plasmids from older strains (isolated circa the 1970s) were obtained from the strain collection held at the Health Protection Agency (HPA) in London, kindly provided by Dr. John Wain and Dr. John Threlfall. Detailed information regarding the sources of strains or the exact dates of their isolation was not available. Only 3 strains from the set supplied were found to contain an IncHI2 plasmid, using an IncHI2 simplex PCR. One of these strains contained a reference IncHI2 plasmid, named TP116 [139]. These strains were obtained in a late stage of this project and only preliminary data was gathered. The antibiotic resistance phenotypes the IncHI2 plasmids conferred were determined using CDS (zone sizes in Table A1.1 in Appendix A1), and are listed in Table 5.9.

Because none of these strains were tetracycline resistant, it was likely that Tn*10* was not present. Consistent with this, the strains did not produce amplicons for PCRs spanning Tn*10* (Fig 3.10), nor for PCRs that detect the LH and RH boundaries of Tn*10* within IncHI2 backbone (Fig 3.11). All strains produced an amplicon of correct size for the PCR across the insertion site of Tn*10* in R478, pSRC125 and pSRC26, confirming they did not contain Tn*10* and the sequence surrounding Tn*10* in other IncHI2 plasmids was intact. Similarly, PCRs targeting the LH and RH boundaries of Tn*1696*-like in the R478 location were negative, but PCR across the site when Tn*1696*-like is absent was positive. Both the amplicons obtained

from TP116, which confirm that Tn*1696*-like and Tn*10* were absent, were sequenced. This revealed that sequence in TP116 was equivalent to the sequences of pSRC125 and pSRC26, with the transposon and 1 copy of the direct repeat removed, in both cases.

PCR targeting the *merA* of Class II mercury transposons showed that 2 strains, 28R823 (TP116) and 31R672, contained the *merA* gene and digestion of amplicons showed that both were of Tn*21* type (Table 5.9). Primers in the 3'-CS and 5'-CS of a class 1 integron (L1 and R1; sequences in Table A3.1) were used to detect gene cassette arrays. Only 28R823 and 31R672 (the 2 strains that contained $merA_{21}$) produced an amplicon, of ~1 kb, showing that they contained cassette/s in a class 1 integron. Because both conferred resistance to Sm and Sp, it was hypothesised that they contained one of the Sm/Sp^R gene cassettes *aadA1* or *aadA2*. Digestion of the 1 kb amplicon with RsaI produced fragments of ~490, 310 and 200 bp, which were identical to the *aadA1*-containing control. This indicates that although these strains most likely contain an *aadA* cassette, both *aadA1* and *aadA2* yield identical digestion patterns and so to determine which particular *aadA* gene they contain, sequencing would be needed.

Table 5.9 Properties of 1970s plasmids[1]

Sample	Plasmid	R phenotype[2]	*merA*	*IntI1*/*tnpR*	R genes/ linkages	Integron cassettes
28R823	TP116	Cm Sm Sp Su	Tn*21*	Tn*21*	*catA1*/IS*1*	*aadA*[3]
31R672	?	Cm Sm Sp Su	Tn*21*	Tn*21*	*catA1*/IS*1*	*aadA*[3]
33R948	?	Km Sm	none	none	*aphA1*	none

[1]In an unknown *E. coli* host strain, previously transferred by conjugation
[2]tested by CDS – zone sizes of which are in Table A1.1
[3]*aadA1* or *aadA2*

Furthermore, 28R823 and 31R672 were shown to have *intI1* linked to *tnpR* of Tn*21* (Table 5.9), using the linkage PCR previously described (Fig 3.2). Thus, these strains appeared to contain Tn*21*, or part thereof. Because both strains also conferred Cm^R, and the Tn*21*-containing transposon Tn*2670* contains the *catA1* Cm^R gene next to IS*1* (see Fig 1.14), a

catA1/IS*1* linkage PCR was performed. Both produced amplicons of the same size as the Tn*2670*-like control. Taken together, this evidence is consistent with the presence of a Tn*2670*-like element, but this was not investigated further. The remaining strain, 33R948, was resistant to Km and Nm and, using PCR, the *aphA1* gene was detected (Table 5.8).

Resistance to the heavy metals mercury, arsenic and tellurite was tested only in 28R823, which contained TP116. It was resistant only to mercuric ions, which is presumably conferred by the Tn*21* element detected. It was sensitive to tellurite and arsenic, indicating that perhaps it does not contain the arsenic resistance transposon but, due to time constraints, this was not investigated. Interestingly, TP116 is the only IncHI2 plasmid examined in this study that does not confer tellurite resistance.

5.6 Discussion and future directions

The *S*. Infantis strains described here (except for SRC83) had previously been subjected to IS*200* typing analysis and had produced identical patterns. Thus, they were clonally related [230]. Therefore, we may expect that the plasmids they contain to be similar. However, variation was observed between their RFLP patterns (Fig 5.4) and PCR results. It is not known whether the variations occurred during circulation of these plasmids in the wild, or during storage. For example, in a study performed in 2006 on the same strains [230], the authors found SRC70 to be Tc^R and to contain *tetA*(B), but when it was reanalysed here, it was Tc^S and no longer contained *tetA*(B). This was found to be a result of a large deletion within the IncHI2 backbone (Table 5.4). In this case, variation was most likely to have occurred during storage. To fully map and analyse the *S*. Infantis plasmids and to determine exactly how they differ from one another, sequencing of them, using next generation techniques, such as 454 or Illumina, would be needed. Furthermore, this would elucidate the

structures of their heavy metal resistance regions. For example, it was determined, using PCR, that the majority of plasmids contained the *cop* gene, but not the *tnsD* or *sil* genes. Thus, the part of Tn*6024* encoding copper resistance may be present and the remainder lost. This reflects the configuration observed in pEC-IMP/Q across this region (that is, $tnsD^{-}sil^{-}cop^{+}$; Fig 4.16). However, these *S*. Infantis plasmids share the positions of their antibiotic resistance regions Tn*10* and Tn*6028* with the equivalent regions in pSRC26 and pSRC125 and R478, indicating that they are quite closely related to this branch of IncHI2 plasmids (Fig 3.20) and not to pEC-IMP/Q. Thus, it is more likely that a separate deletion event has occurred in the *S*. Infantis plasmids within the Tn*6024* region. Sequencing of the *S*. Infantis plasmids that were not conjugative (see Table 5.6) also may provide insight into why they did not conjugate. Alternatively, PCR primers directly targeting the IncHI2 transfer regions should be designed to determine if any are missing, truncated or interrupted. Furthermore, conjugation of the plasmids in the *S*. Infantis strains should be repeated at 26^{O}C as IncHI2 plasmids display temperature sensitive transfer [178];[177] and this may increase the frequency enough to recover a transconjugant.

This study highlights limitations of using PCR to map large plasmids and the importance of using multiple techniques. For example, by simply using the backbone PCRs on the *S*. Infantis plasmids, the additional ~80 kb of plasmid DNA detected in the transconjugants by RFLP/PFGE, would have been overlooked. However, there are technical limitations of RFLP/PFGE as a technique, for example, as has been noted elsewhere [160], that it is hard to extract the DNA of these large plasmids. However, in the current work, I extracted high quality plasmid DNA by optimising extraction procedures. For example, a number of *E. coli* strains were tested as host strains for DNA extraction and they were found to give highly variable results, whereas the host strain E294 worked reliably. However, it would be ideal to

develop a method to extract high quality plasmid DNA from *Salmonella*, because plasmids that do not transfer, such as pSRC70, could be examined using this method.

Further analysis of pSRC70 would be useful to answer a number of questions. How is it maintained without any of the known replication regions? Is there a 3rd replication region? Is there another plasmid supplying one *in trans*? Has there been a cointegrate fusion? A number of approaches could be used, in addition to whole plasmid sequencing, to address these questions. For example the replication region could be readily isolated and identified by digesting, religating and transforming plasmid DNA, then selecting for antibiotic resistances conferred by Tn*6028*. Any recovered transformant should contain the replication region and an antibiotic resistance gene. Also, as proposed in section 5.4.5, perhaps another plasmid is present in the form of a co-resident plasmid or cointegrate fusion. In this study, PBRT revealed that plasmids of targeted incompatibility groups were not present in the *S.* Infantis strains. However, plasmids of other incompatibility groups could be tested for by using S1-nuclease treatment of plasmid DNA and separation and visualisation using PFGE, as has been used previously to study IncHI2 cointegrate formation [193].

The work described in this chapter raises a number of hypotheses that would be extremely interesting to investigate, for example, examining the backbone and resistance regions of the older HPA plasmids, to explore the possibility that they represent another IncHI2 plasmid lineage, which contains transposons related to Tn*21* and not Tn*1696*. Furthermore, to the best of the author's knowledge, TP116 is the first tellurite sensitive IncHI2 plasmid to be described. It is possible that the tellurite resistance region, like the arsenic, copper and silver resistance regions, is also on a MGE. Further analysis of TP116, for example by sequencing it, may reveal the boundaries of a novel tellurite resistance transposon in all other IncHI2 plasmids.

Previous studies found that variation across a set of 11 plasmids only occurred between the tellurite resistance region and transfer region (Fig 4.1; [161]). The set of plasmids analysed in this chapter, provides further evidence that this region is indeed the main region where variation is observed (with the exception of SRC70; Table 5.4). In R478, this region is contained in a continuous 87 kb region, which comprises of less than a 3rd of the entire plasmid. With the identification of Tn*1696*-like elements, Tn*6024* and the ars^{R} transposon within this region, (which accounts for~60 kb) it is possible there are further MGE in this region, which have not yet been identified. It is common for MGEs to end up in a "hotspot" region such as this, as other features of the plasmid are essential. This also suggests that large plasmids, like IncHI2s, are comprised of a number of large insertions. Future work to further the current study, would include examining more IncHI2 plasmids. For example, analysing the set already partially characterised in García-Fernández et al., (2007) [161], using the tools developed throughout Chapters 3-5, especially to test for the presence/absence of the heavy metal resistance transposons, would greatly enhance knowledge of these plasmids.

CHAPTER SIX

Resistance regions in IncHI1 plasmids

6.1 Introduction

S. enterica serovar Typhimurium strain SRC27 was isolated from an Australian equine infection in 1999. It was acquired as part of the *Salmonella* collection obtained from the Microbiological Diagnostic Unit (MDU), where the resistance phenotype had been determined (Diane Lightfoot, personal communication; Table 6.1). It was unusual because it was resistant to gentamicin. SRC27 was amongst a smaller set of *Salmonella* Typhimurium strains, that had been partially characterised in Dr. Djordjevic's Laboratory, at the Elizabeth Macarthur Agricultural Institute (EMAI). Some common genes conferring antibiotic resistance were detected in SRC27 using PCR screening (Table 6.1; Xuilan Liu, personal communication). Linkage of some of these resistance genes to others or to insertion sequences had also been established using PCR (Table 6.1). The bla_{TEM}, *sul2* and *strA/B* genes were found to be in the transposon Tn*6029* and *aphA1* was in Tn*4352* adjoining the Tn*6029* (as seen in IncHI2 plasmid pSRC26; see Fig 3.12). SRC27 had been shown to contain part of a class 1 integron, including *intI1* and the *dfrA12*-orfF-*aadA2* gene cassette array, but not the *sul1* gene. The *catA1* gene was linked to an IS*1* (Renee Levings, personal communication).

Table 6.1 Resistances and resistance genes of SRC27.

Antibiotic	**Abbreviation**	**Ab^R gene**	**Location**
Ampicillin	Ap	bla_{TEM}	Tn*6029*
Chloramphenicol	Cm	*catA1*	next to IS*1*
Gentamicin	Gm	**unknown**	**unknown**
Kanamycin	Km	*aphA1*	Tn*4352*
Neomycin	Nm		
Streptomycin	Sm	*strA/strB*	Tn*6029*
Spectinomycin	Sp	*aadA2*	integron
Streptomycin	Sm		
Sulphamethoxazole	Su	*sul2*	Tn*6029*
Tetracycline	Tc	*tetA*(B)	**unknown**
Trimethoprim	Tp	*dfrA12*	integron

Previous work in the Hall Laboratory, using PCR-based replicon typing (PBRT) [145] had shown that IncHI1 and IncI1 replicons were present in SRC27 (Nick Evershed, personal communication).

The aim of this chapter was to analyse the antibiotic resistance region/s from the IncHI1 plasmid in SRC27, which was isolated from an Australian animal *S*. Typhimurium infection, and to examine the evolutionary relationship between this IncHI1 plasmid, and IncHI1 plasmids isolated internationally, from human *S*. Typhi infections. SRC27 was chosen for further characterisation because it was resistant to a range of antibiotics, including the aminoglycosides gentamicin, kanamycin and neomycin.

6.2 Antibiotic resistances of SRC27

The resistance profile for SRC27 obtained by the MDU (ampicillin, chloramphenicol, gentamicin, kanamycin, neomycin, streptomycin, spectinomycin, sulphamethoxazole, tetracycline and trimethoprim resistant), was confirmed using the CDS disc diffusion assay (inhibition zone sizes are listed in Table A1.2 in Appendix 1). A gentamicin resistance gene had not been found previously, but for all other antibiotics at least one gene had been identified to account for the resistance (Table 6.1). The presence of these resistance genes was confirmed using PCR (for properties of primers used see Table A3.1 in Appendix 3).

6.3 Plasmids in SRC27

Plasmids were conjugated from SRC27 into the *E. coli* recipient strain, E294, initially at 37^{O}C and using selection on Ap. All the antibiotic resistances and antibiotic resistance genes in Table 6.1 were detected in the transconjugants, indicating that these resistances in SRC27 were conferred by plasmids. However, both the IncHI1 and IncI1 replicons were detected in a transconjugant, named E294/pSRC27-H/pSRC27-I, using PBRT (Fig 6.1).

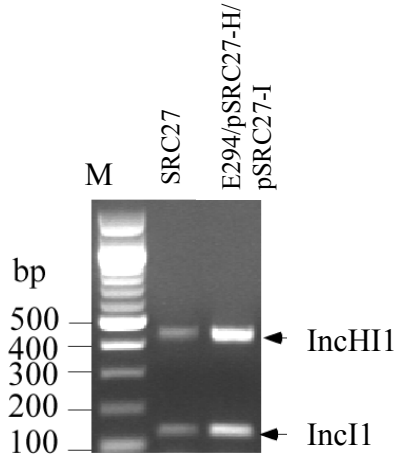


Fig 6.1 PBRT analysis of SRC27 and its transconjugant. The horizontal arrows indicate the position of the predicted amplicons for the IncHI1 replicon (a 471 bp product) and the IncI1 replicon (a 139 bp product). Using primers listed in Table A3.2 in Appendix 3. Lanes are labelled SRC27 for the *Salmonella* strain, and E294/pSRC27-H/pSRC27-I for the *E. coli* strain E294 transconjugant. 100 bp Molecular Weight Markers are loaded in lane M and selected sizes are as indicated.

Transfer of all resistances (Ap, Cm, Gm, Km, Nm, Sm, Sp, Su, Tc and Tp) together also occurred at 37ºC, when using Cm or Km selection, but at a very low frequency of ~10^{-7} transconjugants per donor (Table 6.2). However, when Sm selection was used, the conjugation frequency was 3 orders of magnitude higher and only Sm resistance transferred. Only the IncI1 replicon type was observed in the Sm resistant transconjugants, whereas those that were resistant to all the antibiotics, contained both the IncHI1 and IncI1 replicons (Table 6.2). Hence, these replicons are clearly associated with 2 separate plasmids of IncI1 and IncHI1 incompatibility groups, henceforth designated pSRC27-I and pSRC27-H, respectively. Plasmid pSRC27-I only confers streptomycin resistance and pSRC27-H presumably carries the remainder of the resistance genes.

As pSRC27-H could not be isolated on its own using this standard protocol, the conjugation temperature was lowered. It has been reported that IncHI1 plasmid transfer is temperature sensitive. Transfer occurs efficiently at 15-30^{O}C (at a maximum rate of approximately 10^{-4}

transconjugants/donor), but drops substantially at 37^{O}C, to approximately 10^{-7} transconjugants/donor [176].

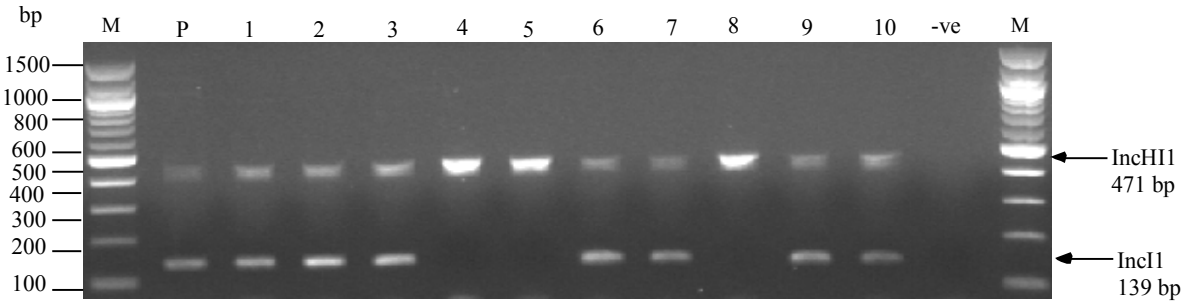
Table 6.2 Conjugation of plasmids from SRC27 to *E. coli* under different conditions.

Mating temp. (°C)	Antibiotic selection[2]	Frequency (transconjugants / donor)	No. colonies tested	Plasmid/s transferred	Resistance phenotype [1, 2]
37	Ap	8.1 x 10^{-7}	23	23x HI1 + I1	ApCmGmKmNmSmSpSuTcTp
	Cm	5.7 x 10^{-7}	2	2x HI1 + I1	ApCmGmKmNmSmSpSuTcTp
	Km	4.3 x 10^{-7}	2	2x HI1 + I1	ApCmGmKmNmSmSpSuTcTp
	Sm	1.4 x 10^{-4}	23	23x I1[3]	Sm
26	Ap	5.2 x 10^{-4}	30	**6x HI1**[4]	ApCmGmKmNmSmSpSuTcTp
				24x HI1 + I1	ApCmGmKmNmSmSpSuTcTp
	Cm	4.8 x 10^{-4}	2	2x HI1 + I1	ApCmGmKmNmSmSpSuTcTp
	Km	5.7 x 10^{-4}	2	2x HI1 + I1	ApCmGmKmNmSmSpSuTcTp
	Sm	7.1 x 10^{-4}	25	**1x HI1**	ApCmGmKmNmSmSpSuTcTp
				11x I1	Sm
				13x HI1 + I1	ApCmGmKmNmSmSpSuTcTp

[1]Transconjuants were also rifampicin resistant from Rif^{R} *E. coli* E294 recipient strain and selected with rifampicin in addition to the named antibiotic
[2]Ap: ampicillin; Cm: chloramphenicol; Gm: gentamicin; Km: kanamycin; Nm: neomycin; Sm: streptomycin; Sp: spectinomycin; Su: sulphamethoxazole; Tc: tetracycline; Tp: trimethoprim
[3]Cells highlighted green used to emphasise when the IncI1 plasmid transferred alone
[4]Cells highlighted purple used to emphasise when the IncHI1 plasmid transferred alone

At 26^{O}C, pSRC27-H transferred at 10^{-4} transconjugants/donor, approximately 1,000 fold more efficiently than at 37^{O}C (Table 6.2) and recipient cells harbouring only pSRC27-H were recovered (highlighted purple in Table 6.2). For example, after selection with Ap, 6 out of 30 transconjugants tested contained pSRC27-H only and the remainder had both plasmids. Using streptomycin selection, pSRC27-I also transferred alone in 11 out of 20 transconjugants analysed, indicating pSRC27-I can transfer efficiently (in the order of 10^{-4} transconjugants/donor) at both 26^{O}C and 37^{O}C. Two colonies containing only the IncHI1 plasmid (lanes 4 and 5 from Fig 6.2A), named E294/pSRC27-H-1 and E294/pSRC27-H-2, and two with the IncI1 plasmid only (lanes 1 and 11 from Fig 6.2B), named E294/pSRC27-I-1 and E294/pSRC27-I-2, were purified and used to characterise the individual plasmids, pSRC27-H and pSRC27-I.

A. Selected with Ap



B. Selected with Sm

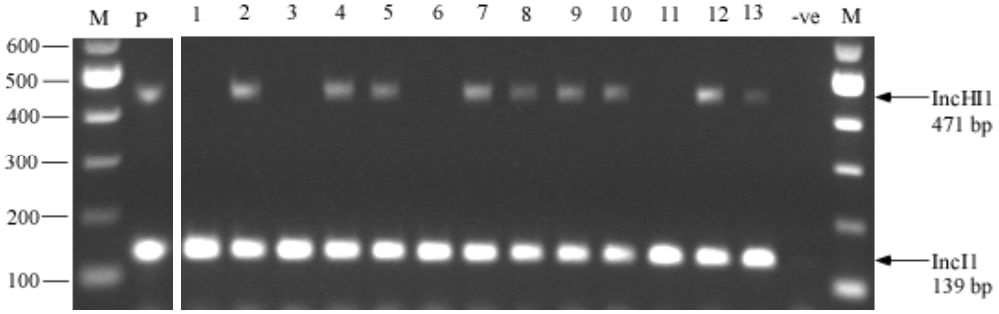


Fig 6.2 Plasmids in *E. coli* transconjugants. PBRT performed on transconjugants from conjugations of SRC27 and E294 at 26^{O}C. Selected with **A.** ampicillin **B.** streptomycin. The 471 bp amplicon indicates the presence of an IncHI1 plasmid and the 139 bp amplicon the IncI1 plasmid, as shown by horizontal arrows. Using primers listed in Table A3.2 in Appendix 3. Lane marked P is the parent *Salmonella* SRC27 strain and the positive control, numbered lanes represent individual transconjugants and the lane marked –ve is the negative water control. 100 bp Molecular Weight Markers are loaded in lane M and selected sizes are as indicated.

6.3.1 pSRC27-I

The *strA*/*strB* genes are most often found in 2 locations, in Tn*5393* and variants of it (particularly Tn*5393*a and Tn*5393*c; see Fig 3.17), or in the MARR from RSF1010 (or RSF1010-derived structures, such as Tn*6029;* see Fig 3.13), which contains a fragment of Tn*5393*. Thus, *strA* is seen most commonly adjacent to *sul2*, IS*1133* or $tnpR_{5393}$. A set of PCRs were designed to determine if *strA* was in one of these locations, using a primer in *strA* with primers in $tnpR_{5393}$, IS*1133* and *sul2* (Fig 6.3).

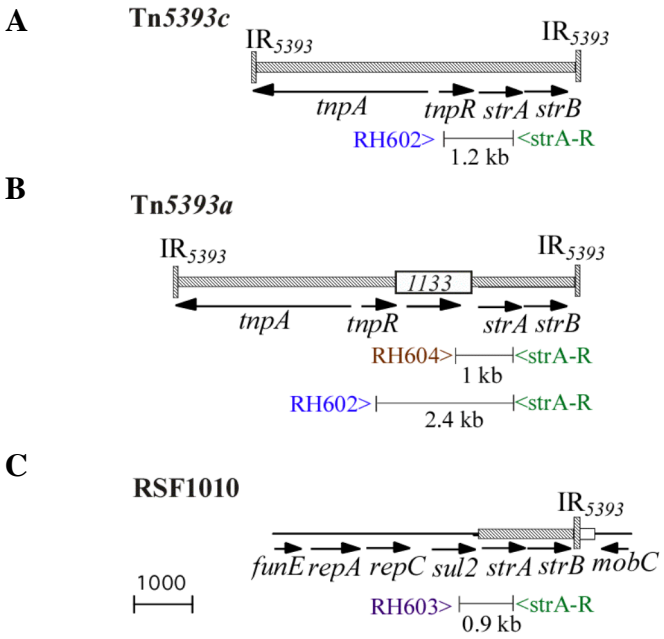


Fig 6.3 PCR to detect common locations of *strA/B*: A Tn*5393c*, **B** Tn*5393* and **C** RSF1010. Lines of different styles represent gene segments with different origins, where striped open lines represent Tn*5393* and thin lines, RSF1010. Inverted repeats (IRs) are vertical bars. The arrows below show the direction and extent of genes. IS*1133* is an open box. Primer positions are shown below as thin vertical lines with the primer names bounding the lines. The predicted amplicon is the thin horizontal line between the primers with the size indicated below this line. Figures are drawn to scale using GenBank accession numbers: AF313472, M95402 and GQ379901 for A-C respectively. Primers sequences of strA-R, RH602, RH604 and RH603 in Table A3.1 are in Appendix 3.

For pSRC27-I, an amplicon of the predicted size (958 bp) for the Tn*5393* configuration where *strA/B* is next to IS*1133,* was obtained (Fig 6.3B). The amplicon sequence was identical to that of Tn*5393* (GenBank accession no. M95402). Linkage of *strA* to $tnpR_{5393}$, using primers RH602 and strA-R yielded a 2.4 kb product, indicating *tnpR* of Tn*5393* was also present (Fig 6.3B; data not shown). In the parent, SRC27, the same set of PCR reactions yielded 2 amplicons, corresponding to 2 different locations of *strA*: one next to *sul2* (Fig 6.3C) and the one from pSRC27-I. This is consistent with previous results that had shown *strA/B* was located within the transposon Tn*6029*, part of which is derived RSF1010 (Table 6.1) and this second copy is not in pSRC27-I. Southern hybridisation confirmed the presence of 2 copies of *strA* in SRC27. SwaI was used to digest plasmid DNA because it is a rare cutting restriction enzyme that has a single site in IS*26* and should yield a 4.8 kb band from the *strA/B*

containing part of Tn*6029* which is flanked by 2 IS*26* (Fig 6.4). However, SwaI does not cut in Tn*5393* and the exact expected fragment size could not be calculated, but it would be >6.8 kb, the size of Tn*5393*.

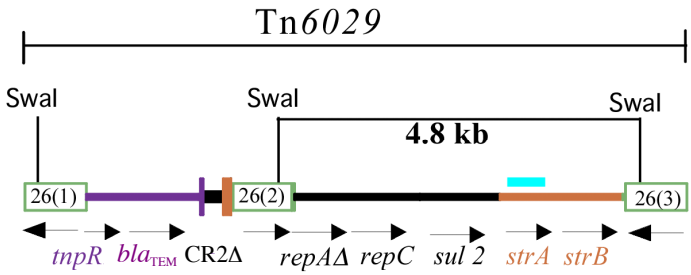


Fig 6.4 Schematic of southern hybridisation of pSRC27-I. Plasmid DNA was digested with SwaI and probed. The DIG-labelled *strA* probe is represented by the cyan line above the *strA* gene in Tn*6029*. Green open boxes with 26 are IS*26*, and the copy number of each is in brackets next to the 26. Purple shows the Tn*2*-derivation, and brown the Tn*5393*-derivation. The transposon Tn*6029* is named and bracketed above. The predicted size of the fragment is given as 4.8 kb between the 2 SwaI sites, which are marked as SwaI.

Hybridisation with a DIG-labelled *strA* probe revealed 2 bands in SRC27, the 4.8 kb band from Tn*6029* and one which was also in the transconjugant pSRC27-I, corresponding to Tn*5393* (Fig 6.5). The Tn*5393* fragment ran at ~23 kb indicating it is a fragment larger than 20 kb. Whether the complete Tn*5393* was present was not examined and no further analysis of pSRC27-I was undertaken.

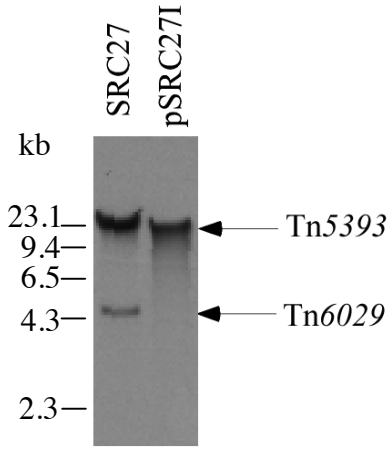


Fig 6.5 Southern hybridisation. SwaI digested plasmid DNA from *Salmonella* strain SRC27 and a transconjugant containing pSRC27-I using a *strA* probe. The band corresponding to Tn*6029* is expected size 4.8 kb and the Tn*5393* band runs at the ~20 kb migration distance. The positions of the molecular weight markers are shown as lines and sizes and were determined by measuring the migration distance and comparing them to distances of DIG-labelled molecular weight markers II and VII on a previous blot of the same membrane.

6.4 Analysis of pSRC27-H

6.4.1 Antibiotic resistances and resistance genes

pSRC27-H was found to confer resistance to all the antibiotics the parent *Salmonella* strain was resistant to: ampicillin, chloramphenicol, gentamicin, kanamycin, neomycin, streptomycin, spectinomycin, sulphamethoxazole, tetracycline and trimethoprim (CDS zone sizes are in Table A1.2 in Appendix 1). pSRC27-H conferred resistance to kanamycin and neomycin (conferred by *aphA1*) and gentamicin, but not to the other aminoglycosides tested: tobramycin, netilmicin and amikacin (CDS zone sizes of which were all zero; data not shown). All of the resistance genes present in SRC27 (Table 6.1) were also identified in pSRC27-H by PCR. These genes could account for the resistance phenotype of the pSRC27-H-containing isolates, with the exception of gentamicin resistance (discussed below in 6.4.3). Resistance to gentamicin and not other aminoglycosides, suggested the presence of an *aacC* gene.

*6.4.2 pSRC27-H contains Tn*6029 *and Tn*4352

Previously, linkage PCR performed on SRC27 (Table 6.1) had indicated the presence of 2 transposons, Tn*6029* and Tn*4352,* collectively containing the resistance genes bla_{TEM}, *sul2* and *strA*/*B*, *aphA1* and 4 copies of the insertion sequence IS*26*. This Tn*6029*/Tn*4352* configuration was confirmed to be present in pSRC27-H by overlapping linkage PCRs completely spanning the region (see Fig 3.12 in Chapter 3). The products of these PCR reactions were sequenced and the sequence of Tn*6029*/Tn*4352* was identical to the equivalent region on the IncHI2 plasmid pSRC26 (see section 3.5.1), except for a silent mutation at position 203 in the *aphA1a* gene, where an A was present instead of a G. This nucleotide difference was observed in multiple PCR products and hence was not a sequencing or PCR reaction error. This mutation in *aphA1* was present in only one GenBank entry, namely for the

IncHI1 plasmid pMAK-1 (GenBank accession no. AB366440) which will be discussed below.

6.4.3 Identifying the gentamicin resistance gene

During the mapping of Tn*6029*, a PCR linking bla_{TEM} and IS*26* using the primers TEM-SRV and IS*26*-F, yielded an extra amplicon of approximately 0.9 kb, in addition to the 1.9 kb fragment expected from Tn*6029* (Fig 6.6). The sequence of the 0.9 kb product matched a gentamicin resistance gene *aacC2* and IS*26* (Fig 6.7A). The same configuration is found in plasmid pCTX-M3 (Genbank accession no. AF550415), which was isolated from a human *Citrobacter freundii* infection.

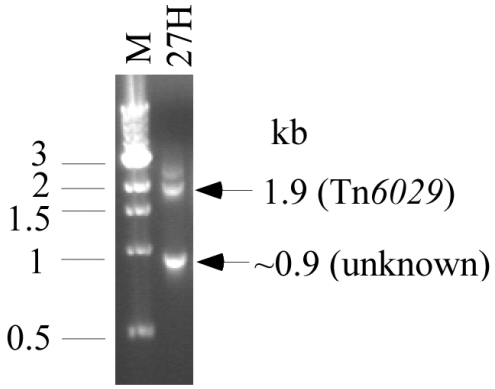


Fig 6.6 PCR products obtained with primers IS26F/TEM-SRV. The top band represents the correct size for the expected from Tn*6029*, which is marked with an arrow. The unexpected band of approximately 0.9 kb of unknown origin is also marked with an arrow. Lane marked 27H presents the pSRC27-H containing transconjugant. Lane M represents a 1 kb Molecular Weight Marker and selected sizes marked are in kb.

When primers were compared to the *aacC2* sequence, TEM-SRV had high similarity (20/23 bp match) to a region in *aacC2* (Fig 6.7B). This allowed mispriming of TEM-SRV to *aacC2* effecting formation of the 0.9 kb product with the primer IS*26*-F seen in Fig 6.6. For subsequent PCR reactions involving bla_{TEM}, and particularly during mapping of Tn*6029*, the TEM-SRV was replaced with primer RH606, which did not anneal elsewhere.

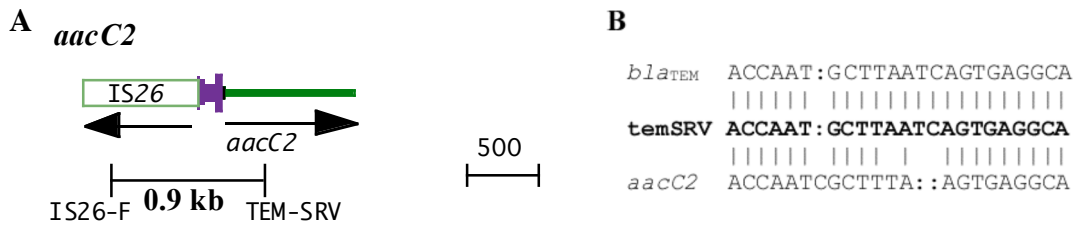


Fig 6.7 Mismatching of primer TEM-SRV to *aacC2*. A Position of the TEM-SRV priming site in *aacC2*. Features used are as in Figure 6.2. **B** Alignment of the TEM-SRV primer its target sequence bla_{TEM} (above) and the priming site in *aacC2*. Sequence of *aacC2* taken from GenBank Accession number AF550415.

The *aacC2* gene, which confers gentamicin resistance [257], accounts for the gentamicin resistance observed in SRC27. Published primers that target the *aacC2* gene [258] were used to confirm the presence of *aacC2* in pSRC27-H (primers aacC2-F and aacC2-R in Table A3.1) and an amplicon of the expected 698 bp was obtained and sequenced. The fortuitous mispriming of TEM-SRV had not only identified *aacC2*, but also shown that it was located next to IS*26*. This linkage was also confirmed by PCR using the primers aacC2-F and IS*26*-F. The sequence of *aacC2* obtained from pSRC27-H was analysed using BLASTn. It was found there were 2 discrete types of *aacC2*-like genes, that each displayed a distinct pattern of single base differences within the sequence. The names of these genes were used inconsistently in the literature and included *aacC2, aacC3, aac(2), aac(3)*, *aac(3)*-II and *aac2* (see for example GenBank accession numbers FJ848785, EU022314, AF466526). The type that matched the pSRC27-H sequence was originally and correctly named *aacC2* [259]. Other *aacC2* genes, with 0-2 mismatches to the pSRC27-H sequence out of the 861 bp gene, occurred 5 times in the GenBank database and 2 of the 3, which had sufficient sequence for analysis, were directly next to an IS*26*. A second type, correctly named *aacC3* [260], that has 26-28 mismatches compared to the *aacC2* sequence from pSRC27-H occurred 11 times. A summary of the distribution of *aacC2* compared to *aacC3* genes is given in Table A4.2 in Appendix 4.

6.4.4 Where is aacC2*?*

To establish where the *aacC2* gene and its adjacent IS*26* were located, southern hybridisation was performed using digestion with the SwaI restriction enzyme and a DIG-labelled IS*26* probe, which binds only to the larger SwaI fragment of the IS*26*. This hybridisation was expected to yield 2 bands, from Tn*6029* and Tn*4352,* of sizes 3.2 kb and 1.9 kb respectively, as well as at least one band of unknown size, a result of the probe binding to the remainder of the 4th copy of IS*26* (IS*26*(4)) in Tn*4352* (Fig 6.8A). A total of 3 bands were observed, 2 of the expected sizes for Tn*6029* and Tn*4352* and one further band was approximately 8 kb (Fig 6.8C). It was deduced that the *aacC2*-associated IS*26* was on this ~ 8 kb band (Fig 6.8B), otherwise, probe would have bound to another SwaI fragment and a 4th band would have been observed. It was possible that *aacC2* is associated with IS*26*(4) in Tn*4352* (Fig 6.8A), or alternatively there is a 5th inversely orientated copy of IS*26* (IS*26*(5)) at the other end of the fragment. In the latter case, we can also surmise that *aacC2* is contained on this ~ 8 kb band, because it is in the opposite orientation to the $tnpA_{26}$ and thus must be located in 1 of 2 positions (Fig 6.8B). Attempts to link *aphA1* and *aacC2* using PCR failed. Therefore, because *aacC2* is not next to Tn*4352*, there must be a 5th copy of IS*26,* which it adjoins (position 2 in Fig 6.8B). When the same membrane (as in Fig 6.8C) was hybridised with a DIG-labelled bla_{TEM} probe, one band of 3.2 kb corresponding to Tn*6029* was expected, however 2 bands were observed. The second band was of approximately 8 kb (Fig 6.8D), which was also observed when probing with IS*26* (Fig 6.8C). This indicated that 2 copies of bla_{TEM} were present in pSRC27-H; one of these copies was in Tn*6029* and the other in the ~8 kb region containing *aacC2* and IS*26*(5). To ascertain if the second bla_{TEM} was next to *aphA1*, PCR using primers in *aphA1* and in bla_{TEM} in both orientations was performed. The forward primer in bla_{TEM} (RH605) with aphA1-F in *aphA1* produced an amplicon of 3.1 kb (Fig 6.9). Thus bla_{TEM} was next to *aphA1* with IS*26* in between (Fig 6.8E).

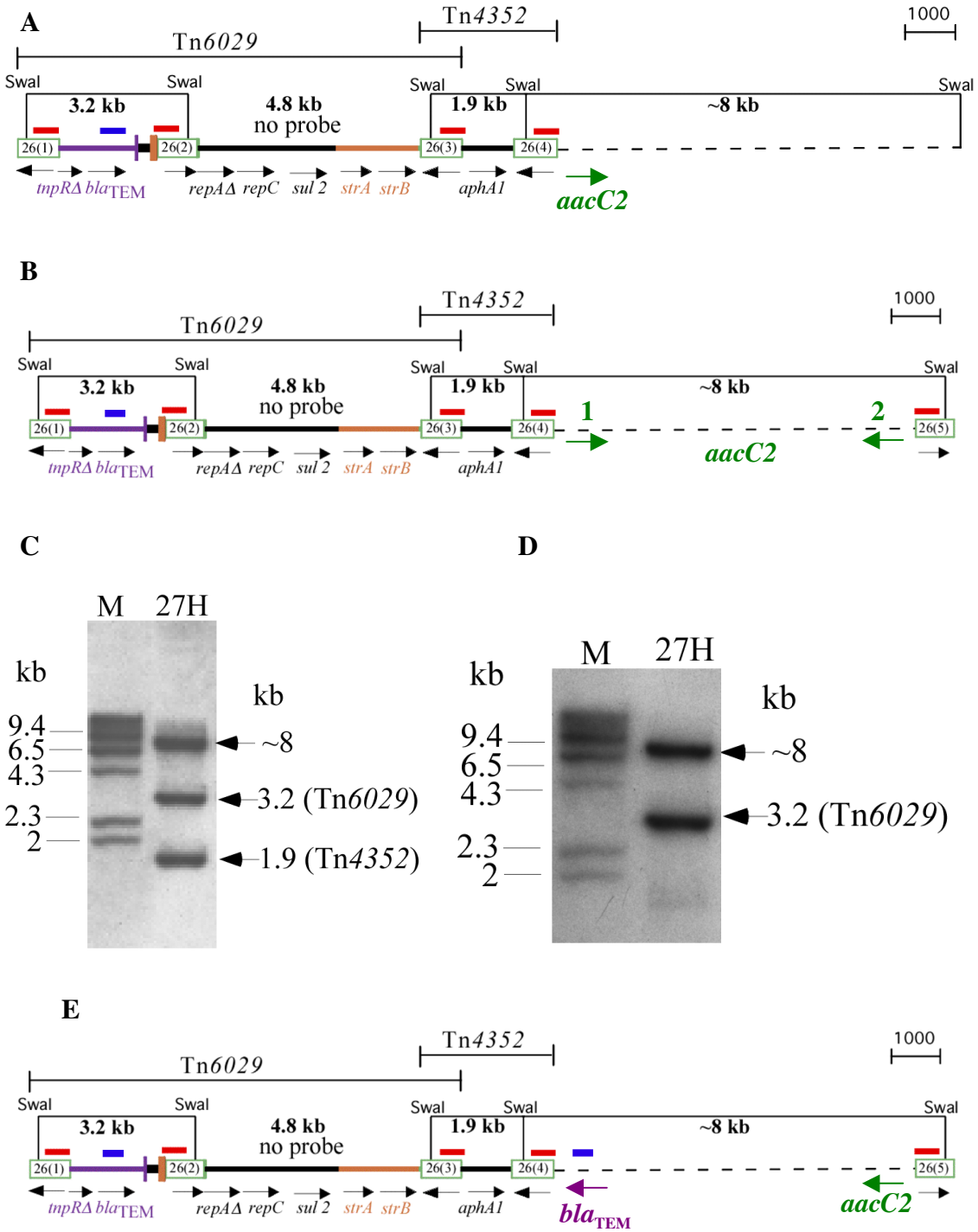


Fig 6.8 Location of *aacC2*. A. Hypothetical configuration if there were 4 copies of IS*26* **B.** 2 possible positions 1, or 2, for the *aacC2* gene if there were 5 copies. A and B are schematics of the SwaI fragments with the unknown region represented by a dotted line. Sites where the probes are predicted to bind are shown as red lines above the IS*26* and blue line above bla_{TEM}. The expected band sizes are given below lines joining the SwaI sites. A green arrow represents direction and possible position of the *aacC2* gene in each case. Other features used are as in Fig 6.4. Southern hybridisation autoradiographs of SwaI digested pSRC27-H using a DIG-labelled **C** IS*26* probe **D** bla_{TEM} probe. Lane 27H is sample pSRC27-H. The bands from Tn*6029* and Tn*4352* are shown as horizontal arrows with their expected sizes marked. A band of unknown origin of approximately 8 kb is shown by a horizontal arrow. Lane M is the DIG-labelled molecular weight marker II and selected sizes are shown in kb. **E.** The confirmed configuration of *aacC2* and bla_{TEM} in pSRC27-H.

When the product of the PCR linking bla_{TEM} and *aphA1* was sequenced, part of Tn*6029*, including the IR_{5393} and CR2Δ (Fig 3.13) was identified between IR_2 and IS*26* (Fig 6.4). Because bla_{TEM} and *aacC2* appeared to be on the same ~8 kb fragment, long-range PCR was performed to detect linkage between them. A faint ~7 kb band was detected with primers RH606 and aacC2-F (Fig 6.9). The sequence of this fragment matched only to the sequence of IncL/M plasmid pCTX-M3 (GenBank accession no. AF550415), with 100% identity. The matching region contained the resistance genes *aacC2* and bla_{TEM} and an insertion sequence ISCfr1. It also included 2 different segments of Tn*2* (purple in Fig 6.9): 1,200 bp including IR_{bla2} and bla_{TEM} to the left of ISCfr1 and 135 bp including IR_{tnp2}, which is from the opposite end of Tn*2* and in the opposite orientation, to the right of *aacC2*. The region was sequenced and found to be identical, from IR_{bla2} to $IR_{tnp2,}$ to pCTX-M3.

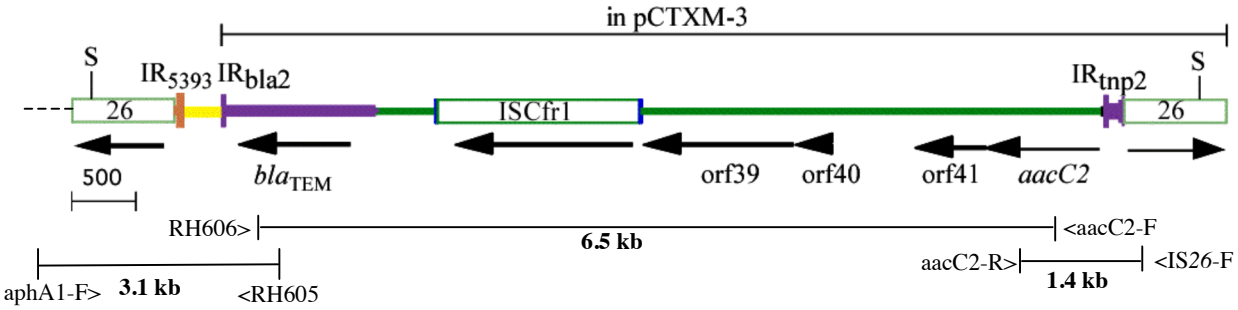


Fig 6.9 Region containing *aacC2*. The part of this region in pCTX-M3 is bracketed above and extends from inverted repeat of the bla_{TEM} end of Tn*2*, represented as IR_{bla2} to an IS*26*. SwaI sites as shown as a vertical line with an "S" above. Yellow represents a CR2 remnant. Green colour is sequence only found in pCTX-M3 (GenBank Accession number AF550415). Primer sequences are in Table A3.1 in Appendix 3. Other features are as in Figs 6.2 and 6.4.

The region in pSRC27-H in between IR_{bla2} and *aphA1* which is identical to part of Tn*6029* was not present in pCTX-M3 (Fig 6.9). It is possible that the configuration in pSRC27-H has formed as a result of 2 homologous recombination events (Fig 6.10) one between 2 IS*26* the other between 2 copies of bla_{TEM}. The outermost IS*26* in Tn*4352* has recombined with another in Tn*6029*. Then recombination has occurred between the copy of bla_{TEM} next to *aphA1* on this molecule in Tn*6029* and bla_{TEM} in the *aacC2*-containing region in pCTX-M3 (Fig 6.10). This yields a Tn*6029*/pCTX-M3-matching region hybrid.

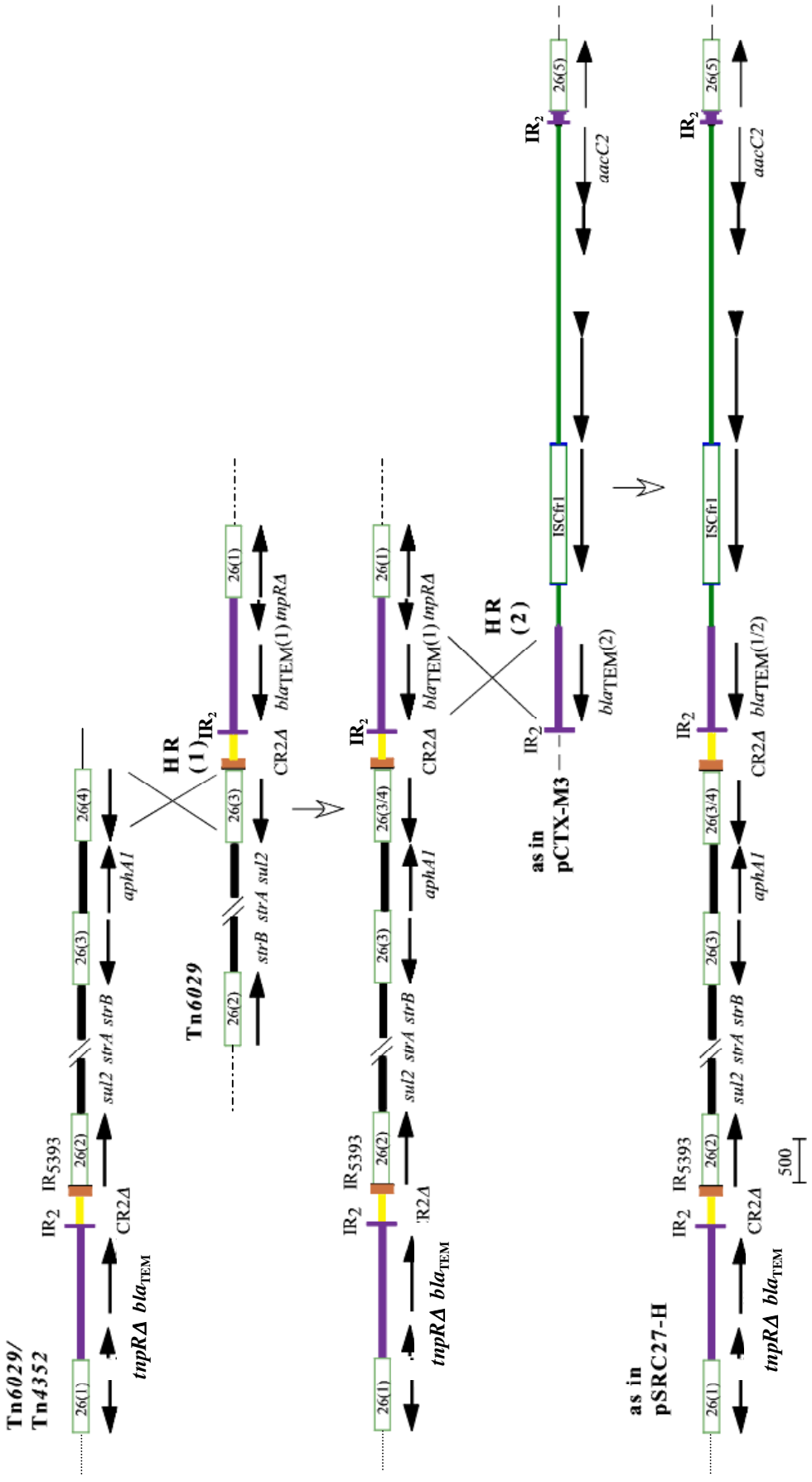


Fig 6.10 Possible mechanism of formation of the *aacC2*-containing region. Homologous recombination signified by a black cross and "HR" which occurs in 2 discrete events (1) and (2) in an unknown order, between the bla_{TEM} copies, and IS*26* copies. Cr2-derived regions are marked in yellow. Other features are as in 6.2, 6.9 and 6.10.

A summary of the IS*26*-containing part of the resistance region in SRC27-H and how they were linked together is given in Fig 6.11.

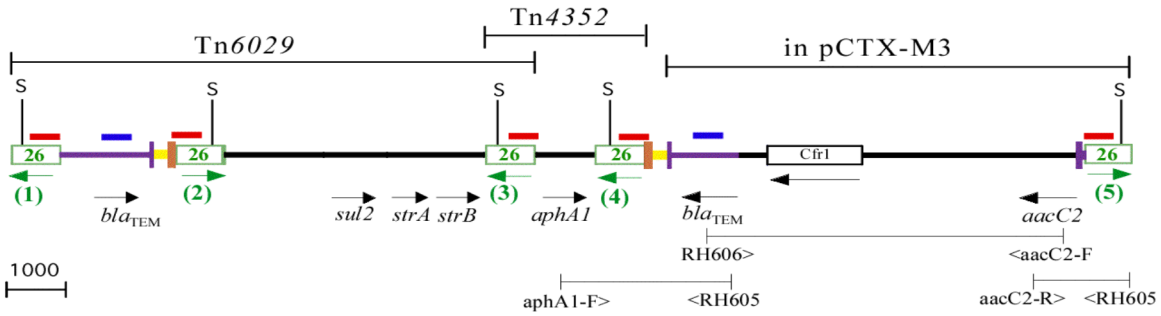


Fig 6.11 Map of the IS*26* bounded resistance region. IS*26* copies are numbered in green. Primer sequences are in Table A3.1 in Appendix 3. Other features used are as in Figs 6.2, 6.6 and 6.8.

6.4.5 Location of the remaining resistance genes.

In order to investigate the location and linkage of the remaining resistance genes, a number of approaches were taken. These included vectorette PCR, cloning and linkage PCR.

6.4.5.1 Vectorette PCR

To determine what was next to the outermost IS*26* in Tn*6029* and the *aacC2*-containing region (copies (1) and (5); Fig 6.11) the vectorette PCR method [238] was employed. Details of this method are given in 2.4.3.4. Vectorette is used to determine the location of genes that are in multiple copies, such as insertion sequences. A number of bands were observed (Fig 6.12) each of which were gel extracted and sequenced with IS*26*-R. Only the ~900 bp band marked in Fig 6.12 gave clear sequence out of IS*26*, which matched to part of *tnpA* of Tn*21*. This sequence was shown to be adjacent to IS*26* copy (1) in Tn*6029* as bla_{TEM} was also linked to $tnpA_{21}$, using PCR. The sequence of the $tnpA_{21}$/IS*26* junction was exactly the same as in the equivalent sequence in pHCM1 (Fig 1.22).

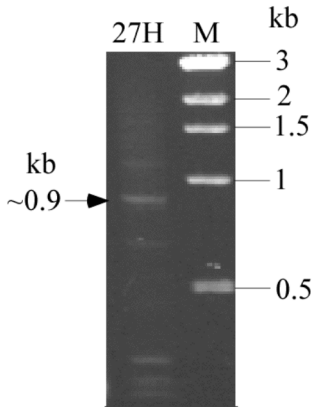


Fig 6.12 Vectorette PCR agarose gel electrophoresis. The products were cut out and sequenced with primer IS*26*R. The extracted and analysed band yielding clear sequence is marked at approximately 0.9 kb. Lanes are 27H representing pSRC27-H and M for the 1 kb Molecular Weight Marker. Selected sizes are shown in kb.

6.4.5.2 Cloning

Cloning was used to locate the resistance determinants *catA1* and *tet*(B) and the class 1 integron (Table 6.1), and to ascertain whether they were linked to one another or to the MARR structure determined so far (Fig 6.11). Plasmid DNA digested with restriction enzyme SacI was cloned into vector pUC19. Using Cm, Km and Sp selection, 3 clones named pRMH959, pRMH958 and pRMH955 respectively, were obtained and their properties are shown in Table 6.3.

Table 6.3 SacI clones of pSRC27-H in pUC19 vector.

Plasmid	Resistant to	Insert resistance determinants	Insert size (kb)
pRMH955	(Ap)[1], Gm, Sp/Sm, Tc, Tp	*aacC2, aadA2, tet*(B), *dfrA12*	20
pRMH958	Ap, Km/Nm, Sm	bla_{TEM}, *aphA1, strA/B*	8.2
pRMH959	Ap, Cm, Su	bla_{TEM}, *catA1, sul2*	13.1

[1]Ap resistance from pUC19 vector

As all resistance genes in pSRC27-H were collectively present in these 3 clones, all the resistance genes in pSRC27-H must be linked in a single resistance region, contained within the 41 kb the clones spanned (Fig 6.13). Where pRMH959 and pRMH958 abut, a 366 bp gap

lies between the SacI sites of pRMH958 and pRMH955, which had previously been sequenced using PCR products (Fig 6.13).

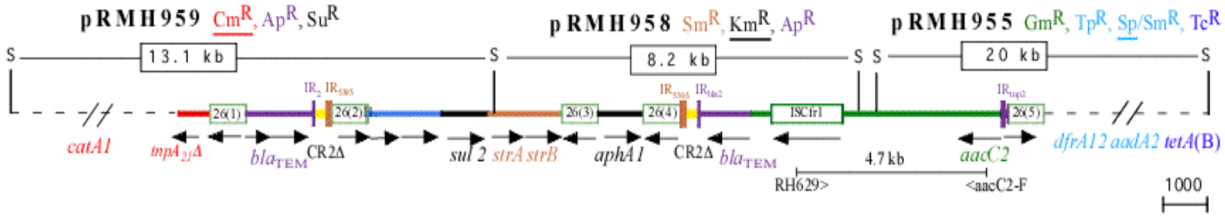


Fig 6.13 Clones that span the MARR. The extent of each clone insert is shown by a horizontal line joining 2 SacI sites, shown as "S", and the size of the insert is given in a box on this line. Clone names are above, followed by the antibiotic resistances each insert confers and the resistance used for selection for each clone is underlined. Ap: ampicillin; Cm: chloramphenicol; Gm: gentamicin; Km: kanamycin; Nm: neomycin; Sm: streptomycin; Sp: spectinomycin; Su: sulphamethoxazole; Tc: tetracycline; Tp: trimethoprim. PCR to link the region not covered by the clones is shown below. Derivations of regions are indicated using colours: red for Tn*2670*; brown for Tn*5393*, purple for Tn*2*. Primer sequences are in Table A3.1 in Appendix 3. Other features are as in Fig 6.2.

pRMH958 contained the resistance genes *strA/B, aphA1* and bla_{TEM} from Tn*6029*/Tn*4352* and part of the ISCfr1-containing region, a configuration which are within the IS*26*-bounded structure described above (Fig 6.11). pRMH959 contained *catA1*, the Cm resistance gene, and the *sul2* and bla_{TEM} resistance genes from Tn*6029*, indicating that the *catA1* gene was linked to Tn*6029*. pRMH955 was recovered with Sp selection and also conferred resistance to Gm, Sm, Tc and Tp. Therefore the integron, containing the gene cassettes *dfrA12* and *aadA2*, as well as the *tet*(B) determinant were all linked to the *aacC2* gene. The sequences of the SacI fragments in pRMH959 and pRMH955 were determined and compared to other plasmids in GenBank to determine their structure and origins (Fig 6.14).

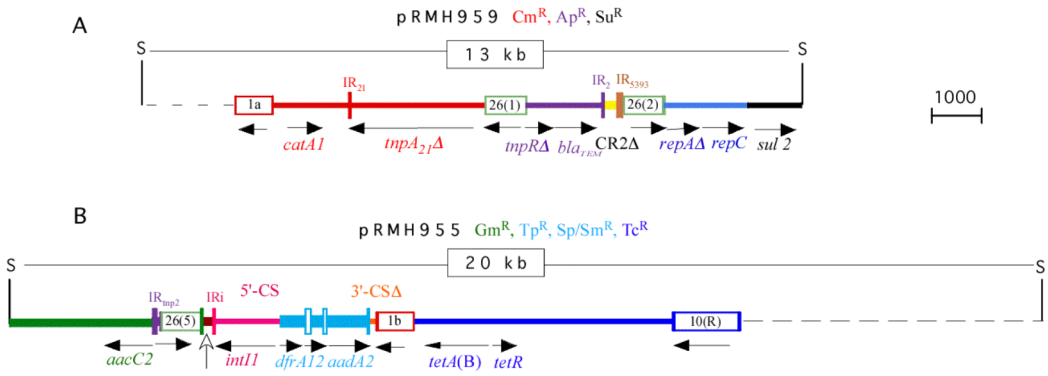


Fig 6.14 Structure of SacI fragments in clones A pRMH959 and **B pRMH955.** Regions derived from Tn*10* are shown in blue. The vertical arrow in B indicates the region derived from IncN plasmids. The cassette orfF is in between *dfrA12* and *aadA2* and is not labelled. Other features are as in Fig 6.13 and 6.2.

6.4.5.2.1 pRMH959

The pRMH959 sequence revealed the region to the left of IS*26*(1), was identical to the corresponding region in pHCM1, except that the 2 additional insertion sequences (IS*26*, and IS*4321*) were not in pSRC27-H. The sequence of the IS*1*/*catA1*/$tnpA_{21}$ segment was 100% identical to parts of Tn*2670* (GenBank accession no. AP000342; Fig 1.14). However, most of Tn*21* was missing, including the remaining 300 bp of the *tnpA* gene, the remainder of the transposition module, the entire integron and the mercuric ion resistance operon. Identity to pHCM1 extended from the right of IS*26*(1) to IS*26*(3), but in pHCM1 85 bp of Tn*6029* directly adjacent to IS*26*(1) was missing (named Tn*6029*B). Hence, Tn*6029* in pSRC27-H was in exactly the same position in $tnpA_{21}$ as Tn*6029*B in pHCM1 (discussed in 6.5). The 1.7 kb of sequence to the left of IS*1a* (dashed lines in Fig 6.14A) also matched pHCM1, which had IS*1*a in the same position (discussed in 6.6).

6.4.5.2.2 pRMH955

The sequence of pRMH955 revealed it contained a class 1 integron with a complete 5'-CS and the *dfrA12*-orfF-*aadA2* gene cassette array, however, only 148 bp of the 3'-CS remained (Fig 6.14B). An IS*1*, possibly the IS*1b* from Tn*2670,* abutted this partial 3'-CS. The cassette

promoter was defined as Pc_{weak} using the sequence of the –10 and –35 regions (see Fig 1.12; [68]). However a "TGN" sequence directly preceding the –10 region, named PcW_{TGN-10}, has recently been shown to increase the efficiency of Pc_{weak} 15 fold [82] thus forming an effectively strong promoter. This promoter differs from the configuration seen in Tn*21*, but is usually associated with the *dfrA12*-orfF-*aadA2* cassette array.

The *intI1* gene was near the *aacC2*-containing region, with IS*26*(5) in between (Fig 6.11) and a similar configuration is seen in plasmid pCTX-M3. However, the 220 bp segment between the IS*26* and IRi in pSRC27-H (marked by a vertical arrow in Fig 6.14B) was not the same as the equivalent sequence of 21 bp in pCTX-M3, indicating that it is not derived from pCTX-M3. This region contained part of the *resP* resolvase gene originally identified in the IncN plasmid R46 [73] and was also found to match with >99% identity to a number of IncN plasmids isolated from *Klebsiella*, listed in Table 6.4. Three of these plasmids, pKP96, Plasmid 12 and pKOX105 contained an integron in exactly the same position as pSRC27-H (Fig 6.15), though the integron of pKOX105 was truncated by an IS*26*. Hence, it is likely that the integron originated from these IncN plasmids.

Table 6.4 IncN plasmids matching sequence in pSRC27-H.

Plasmid	**Organism**	**GenBank Acc. No.**	**Genes LHS**	**Genes RHS**	**Integron**	
					IRi[1]	**Cassettes**
pNL194	*K. pneumoniae*	GU585907	*repA*	EcoRII restriction system	-	-
Plasmid 9[2]	*K. pneumoniae*	FJ223607	*repA*	"	-	-
pKP96	*K. pneumoniae*	EU195449	*repA*	"	+	*aacA4*[3] *qnrA1-ampR*[4]
Plasmid 12	*K. pneumoniae*	FJ223605	*repA*	"	+	*dfrA14*
pKOX105	*K. oxytoca*	HM126016	*repA*	IS*26*	+	bla_{VIM-1}, *aacA4*
pSRC27-H[5]	***S.* Typhimurium**	**HQ840942**	**IS*26***	**IS*1***	**+**	***drfA12*-orfF-*aadA2***

[1] IRi in the same position
[2] 96% match to R46 IncN *repA* region; all others were 99-100%
[3] Partial copy of 3'-CS (Qac, *sul1*, ISCfr) and a 2nd full copy of 3'-CS
[4] bla_{DHA-1} regulator[3]
[5] IncHI1 plasmid

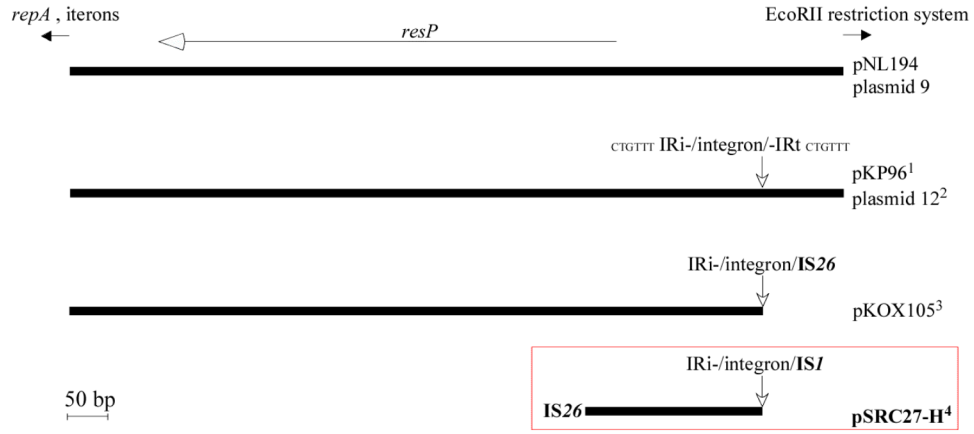


Fig 6.15 Region derived from IncN plasmids. Horizontal lines represent sequences of the *resP*-containing region from different plasmids. All are from IncN plasmids, except IncHI1 plasmid pSRC27-H which is boxed in red. The open headed horizontal arrow represents the *resP* gene and the filled head arrows indicate genes to the left and right of the region on the IncN plasmids. The vertical arrows represent the positions of the integrons, any IRs or direct repeats are named next the integron. Names of insertion sequences are in bold type. The superscript numbers indicate different cassette array for each integron. The figures are drawn to scale based on the GenBank accession nos. GU585907, FJ223607, EU195449, FJ223605 and HM126016.

The sequence of pRMH955 to the right of IS*1b* revealed that part of Tn*10* was present. It included the *tet*(B) resistance determinants *tetA*(B) and *tetR* but was missing the remaining 2,467 bp of Tn*10*, including IS*10*R. This configuration is very similar to the one in pHCM1 (Fig 1.21) however, the position to the right of IS*1*b within Tn*10*, is slightly different for pSRC27-H and pHCM1. In pSRC27-H 356 bp less of Tn*10* remains (named Tn*10*Δ2) than in pHCM1 (named Tn*10*Δ1), possibly as the result of an IS*1*-mediated deletion. The remainder of Tn*10*, including the *tetA*(B)/*tetR* resistance genes and IS*10*L are present in both plasmids. 5.6 kb of sequence beyond IS*10* was obtained (dashed lines in Fig 6.14B) and was found to be present in pHCM1 but not adjoining the Tn*10* fragment. This is explained in 6.4.8.

6.4.6 The complete structure of the MARR

pSRC27-H contained a single 34.6 kb MARR containing 11 resistance genes, 9 insertion sequences and a class 1 integron (Fig 6.16). The structure bounded by the 2 IS*1* that is clearly derived from Tn*2670* was named Tn*6027*. Part of Tn*10* flanks Tn*6027* on the right (Fig 6.16),

but the remainder of Tn*10* was not found on the left. Table 6.5 provides a summary of the origins and composition of the mosaic MARR in pSRC27-H. The sequence of the pSRC27-H MARR, as well as 5.6 kb of flanking sequence on the right and 4.8 kb on the left (discussed in detail below), was deposited into Genbank (for a total size of 45.2 kb) under accession number HQ840942.

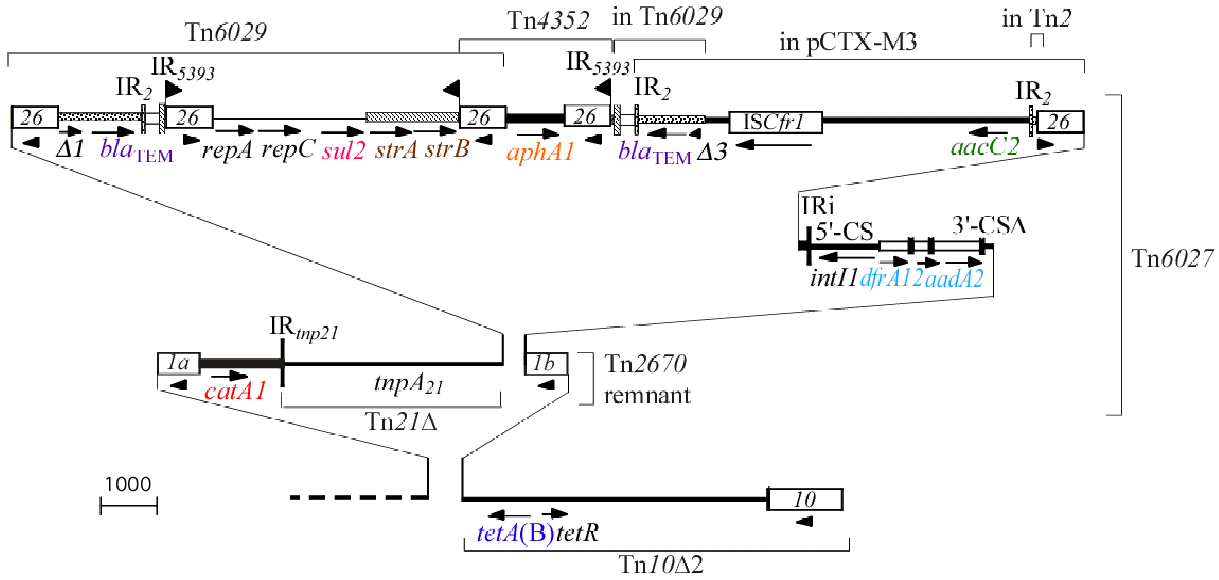


Fig 6.16 The MARR in pSRC27-H. Resistance genes are coloured and the derivations and names of the regions are identified using lines either above or to the side of the region. Flags are the duplications, the orientation of which is represented by the direction of the flag. Δ1 and 3 are different amounts of $tnpR_2$ present. Other features are as in Fig 6.2.

Table 6.5 Composition of MARR in pSRC27-H

MARR position	Regions/Genes	Length (bp)	Matching Region	Part Present	Genbank acc no.	Plasmid name	Identity (%)
1-4,964	IS*1*a/*catA1*/$tnpA_{21}$Δ	4,964	Tn*2670*	1-4,964	AP000342	NR1 (R100)	99.9
4,965-13,726	IS*26*(1)/bla_{TEM}/IS*26*(2)/*sul2*/*strAB*/IS*26*(3)	8,762	Tn*6029*	1-8,762	GQ150541	pSRC26	100
12,907-15,586	IS*26*(3)/*aphA1*/IS*26*(4)	2,680	Tn*4352*	1-2,680	AY123253	pRMH760	99.9
15,587-15,973	Tn*5393Δ*/IR_{5393}/CR2Δ	389	Tn*6029*	2,351-2,737	GQ150541	pSRC26	100
15,974-23,802	IR_2/bla_{TEM}/ISCfr1/orf38/40/41/*aacC2*/IR_2/IS*26*(5)	7,673	*aacC2*-region	1-7,673	AF550415	pCTX-M3	100
23,803-24,022	*resP*Δ/IRi/*intI1*/5'-CSΔ	1,570	IS*26*/5'-CS	3,031-4,600	EU195449	pKP96	100
25,373-27,134	*dfrA12*-orfF-*aadA2*	1,760	gene cassettes	222-1,981	AF284063	Unknown	100
27,135-27,282	3'-CSΔ	148	3'-CSΔ	5,373-5,520	EU195449	pKP96	100
27,283-28,050	IS*1*b	768	Tn*2670*	21,993-22,760	AP000342	NR1	99.9
28,051-34,730	*tetA*(B)/*tetR*/*jemC*/*B*/*A*/IS*10*	6,680	Tn*10*	1-6,680	AP000342	NR1	99.9

[1]Sequence for IS*26*(5) is equalivent to IS*26* found in pSRC26 (acc no. GQ150541)

6.4.7 How related are MARRs in pHCM1 and pSRC27-H?

As mentioned in previous sections, pSRC27-H and pHCM1 share a number of features in their equivalent MARRs (shaded in blue in Fig 6.17). Sequence flanking the left of the MARRs up to IS*26*(3) in Tn*6029* is identical, except for 2 additional insertion sequences and Tn*6029*B, instead of Tn*6029*, in pHCM1. Sequence reaching from the IS*1*b to IS*10* is also identical, except that pSRC27-H contains Tn*10*Δ2 and pHCM1 has Tn*10*Δ1 (Fig 6.17).

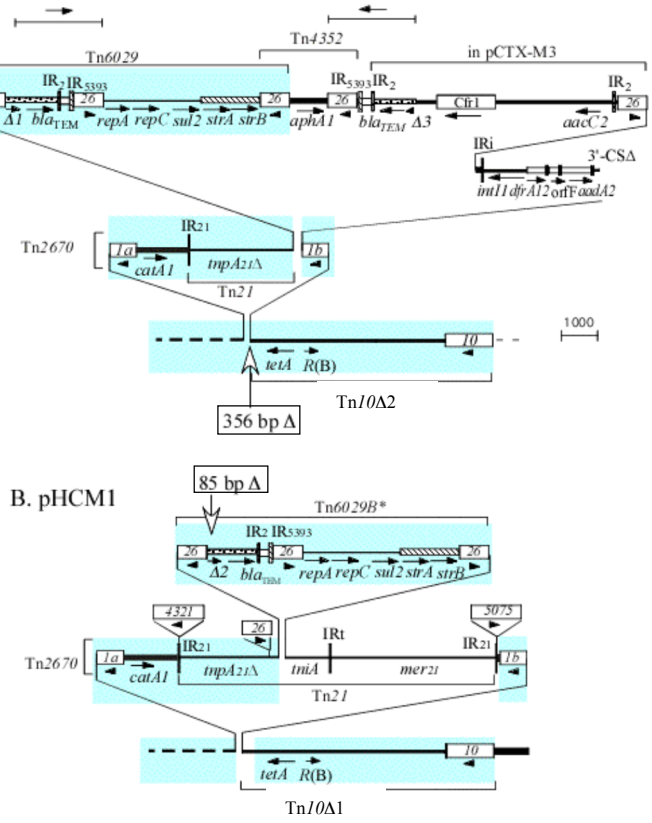


Fig 6.17 Shared regions of MARRs in IncHI1 plasmids A pSRC27-H and **B** pHCM1. Regions these 2 MARRs have in common are shaded. Additional regions are not shaded and regions missing in one, are indicated by vertical arrows that state the size of deletion. The region of duplication of Tn*6029* in A is bracketed and the direction of the duplication is indicated by horizontal arrows above the MARR. Other features are as in Fig 6.2.

The entire segment between the IS*26*(3) of Tn*6029*B and the IS*1*b in pHCM1, is absent in pSRC27-H. This segment includes 1,393 bp of Tn*21*-derived integron, the entire Tn*21* mercuric ion resistance operon and the insertion sequence IS*5075* in the IR. Instead, pSRC27-H contains 2 additions: an *aacC2*-containing region and a class 1 integron (see Fig 6.16).

Although this integron is in a similar position to where the Tn*21*–derived integron is in Tn*2670*, the region was surmised to have originated from an external source, because it contains DNA next to IRi originating from IncN plasmids (region discussed in 6.4.5.2.2).

The MARRs in pHCM1 and pSRC27-H are in same location to the left of IS*1a,* which starts to match to position 153,534 in pHCM1 sequence. The sequence to the right of Tn*10* in pSRC27-H was present in pHCM1, however, it was adjacent to a different, isolated copy of IS*10*. One question that remains to be addressed is why this sequence was not flanking the MARR of pHCM1.

*6.4.8 Sequence adjoining Tn*10 *in pHCM1*

In pHCM1, the sequence adjacent to Tn*10* in pSRC27-H is found adjoining another IS*10*, which is located 9.2 kb to the right of IS*1a* and in opposite orientation to the IS*10* within Tn*10Δ* (Fig 6.18).

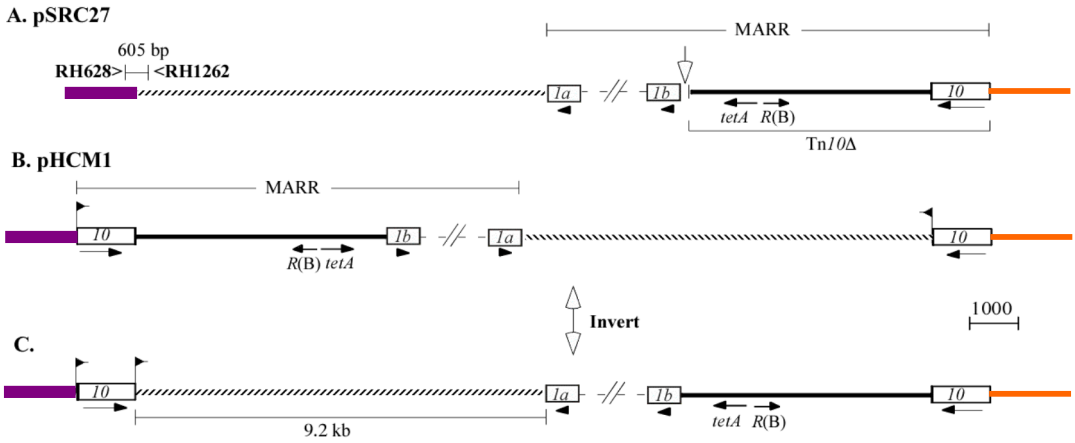


Fig 6.18 IS*10*-mediated inversion in pHCM1. The MARR and surrounding regions in **A** pSRC27-H **B** pHCM1 and **C** predicted structure of pHCM1 prior to inversion. The hashed lines and MARR region comprise the region that had undergone inversion. The hashing shows the orientation of the region relative to the backbone. The region beyond the inversion is shown as thick purple and thin orange lines. The extent of the MARR region is marked above. The slashed lines represent the Tn*2670*-derived part of the MARR and is not to scale. Black flags represent the direct repeats from the IS*10* insertion, the direction of which reflect the direction of the repeats. Primers are shown as vertical lines and the size of the product are shown above the horizontal line connecting the primers. Sequences of primers RH628 and RH1262 are in Table A3.9 in Appendix 3. Figures are drawn to scale using GenBank Acc. Nos HQ840942 and AL513383.

To determine if this additional IS*10* was in pSRC27-H, PCR was performed using primers on either side of where IS*10* appeared to have inserted in pHCM1 (Fig 6.18). A PCR product of 0.6 kb was obtained, which was the correct size when no IS*10* was present (Fig 6.19A) and the sequence of the amplicon confirmed this.

Nine bp repeats (represented as flags Fig 6.18B), which are characteristic of IS*10* insertion, were identified in pHCM1. Their locations, one on the outer edge of the additional IS*10* and one on the inner edge of the Tn*10*-associated IS*10* in addition to their opposite orientation are indicative of an insertion, followed by an inversion event. Therefore, an inversion of the entire 38.9 kb region between the 2 IS*10* is likely to have occurred in pHCM1. If this region was inverted back to its original position relative to the backbone, the partial Tn*10* in pHCM1 would be in exactly the same position as pSRC27-H (Fig 6.18C).

Previously, the MARR of pHCM1 had been defined as the region contained within the 2 IS*10* ([187]; Fig 1.21) however it is clear that an additional IS*10* in pHCM1 has inserted into the backbone sequence. Here, the MARR is redefined as being contained within IS*1*a and IS*10* and the additional IS*10* is unique to pHCM1.

6.5 Comparison with other IncHI1s

6.5.1 Comparison of MARRs

During the course of this work, the sequences of 2 more IncHI1 plasmids became available; pO111_1 (GenBank accession. no. AP010961; [189] and pMAK-1 (GenBank accession. no. AB366440). Neither of the 2 recently sequenced IncHI1 plasmids had been analysed thoroughly: pMAK-1 does not have a manuscript associated with its sequence and pO111_1

was sequenced as part of a multiple *E. coli* genome sequencing project and the manuscript only mentioned the size of the plasmid and that it was similar to R27 [189]. Therefore, the structures of their MARRs were delineated, from their sequences. Both plasmids contained a MARR of Tn*2670* derivation, which adjoined part of Tn*10*. For both plasmids, the MARR was in the same position as pSRC27-H, that is, they had exactly the same sequence to the left of IS*1*a and the right of IS*10* as pSRC27-H.

The MARR of pO111_1 is very similar to that of pHCM1. Some features appear to be unique to pHCM1 (highlighted green in Fig 6.19), namely that pHCM1 contains 3 additional insertion sequences and has Tn*6029*B and not Tn*6029*. Like pSRC27-H, pO111_1 harbours the *aphA1*-containing transposon Tn*4352* overlapping Tn*6029* (highlighted green in Fig 6.19).

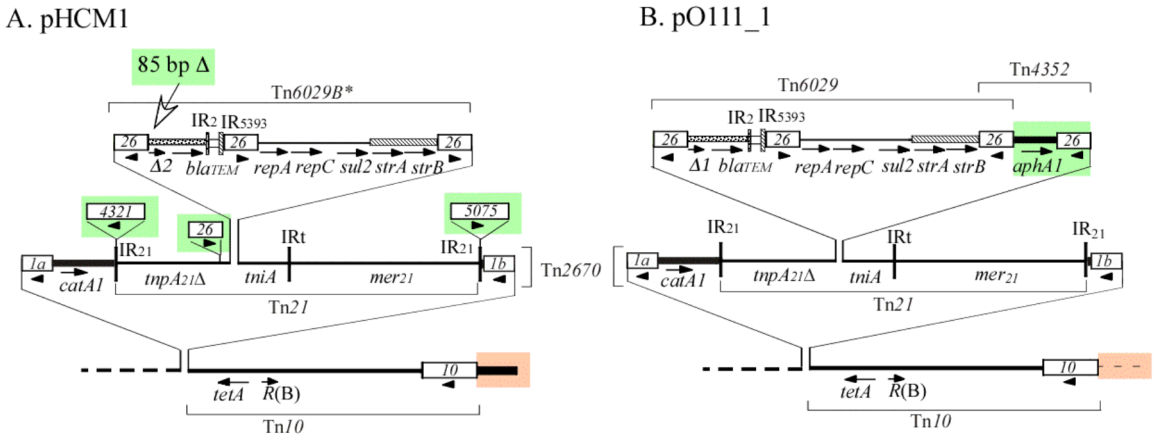


Fig 6.19 Differences between the IncHI1 MARRs A. pHCM1 and **B.** pO111_1. Features that differ between the MARR are highlighted green and differences in the sequence surrounding the MARRs are in orange. Other features are as in Figs 6.17 and 6.2.

The MARR of pO111_1 is in the same position as the MARR of pSRC27-H on both sides, while pHCM1 has a different sequence to the right of IS*10* (highlighted orange in Fig 6.19), due to the IS*10*-mediated inversion described above. pO111_1, like pSRC27-H, does not contain this additional IS*10*, confirming this feature is unique to pHCM1. The MARR of

pMAK-1 is also of similar derivation and in the same position as the MARR of pSRC27-H. However, its MARR though related, appears to have undergone extensive rearrangements (Fig A5.1 Appendix 5) and its unusual structure is not further described here.

Therefore, pHCM1, SRC27-H, pO111_1 and pMAK-1 all contain a Tn*2670*-derived region in the same position and belong to a single plasmid lineage, referred to as the pHCM1-type group of IncHI1 plasmids.

6.6 Summary and discussion

pSRC27-H is the only IncHI1 plasmid described thus far that confers gentamicin resistance. Gentamicin resistance is relatively rare in *Salmonella* isolates, for example, of 396 Australian bovine *Salmonella* isolates (over the period of 1990-1997) 0.6% were gentamicin resistant, which is considerably lower than streptomycin (86%), sulphonamides (70%), tetracycline (47%) or chloramphenicol (18%) resistance [261]. In Australia, gentamicin is not used in farm animals, however, horses are considered domestic pets and consequently are treated with antibiotics that overlap with human treatment. In foals, antibiotics recommended to treat infections include gentamicin, as well as ampicillin, chloramphenicol or trimethoprim [6]. This use may have contributed to selecting for and/or maintaining the gentamicin resistance displayed in SRC27.

This is the first study to analyse an IncHI1 plasmid isolated from *S*. Typhimurium or from Australia. It is also the first detailed description of an IncHI1 plasmid from an animal source. pSRC27-H is closely related to IncHI1 plasmids of the pHCM1-type family, most of which are of human origin, indicating that these plasmids, or the bacteria carrying them, are likely to have transferred between humans and animals. IncHI1 plasmids have the capacity to transfer

into a wide range of gram negative hosts, including species of *Vibrio, Yersinia, Enterobacter, Erwinia, Klebsiella, Proteus, Shigella* and *Aeromonas* [177]. This work by Maher and Taylor (1993) also showed that IncHI1 plasmids are only able to conjugate efficiently at low temperatures. In fact at 24^{O}C, they transferred more efficiently than IncB, C, D, F and K groups and at 14^{O}C, more efficiently than all incompatibility groups tested, including IncP and W groups. The temperature sensitive transfer of IncHI1 plasmids encountered in this study has previously been attributed to the slow and unstable H-pilus formation at higher temperatures [176, 177]. Because IncHI1s are conjugative at ambient temperatures (<30^{O}C), they can be vectors for transfer of antibiotic resistance genes in environmental contexts.

The IncHI1 plasmid pSRC27-H was shown to contain 11 resistance genes and 9 insertion sequences, within in a single 34.6 kb mosaic MARR (Fig 6.15). The MARR contained regions that originated from a number of primary sources: Tn*2670*, Tn*10,* Tn*6029* and Tn*4352* and a class 1 integron but it also contained a number of additional regions from other sources (Table 6.5). It shares features with MARRs in the sequenced IncHI1 plasmids pHCM1 and pO111_1. The MARR structure of another IncHI1 plasmid, pMAK-1, was analysed (Fig A5.1 in Appendix 5) using the sequence in GenBank (acc. no. AB366440), as there is no manuscript that describes this plasmid. However, the structure of pMAK-1 MARR is atypical, despite the fact it shares many features with the MARRs of pSRC27-H, pHCM1 and pO111_1. How its structure arose was difficult to determine and 2 possible discrepancies in the assembly were identified (marked with red arrows in Fig A5.2). For example, bla_{TEM} is not bounded by the usual 2 IS*26,* as in Tn*6029*, but instead is flanked by one IS*26* and the remnant of Tn*9*-like, which is proceeded by the mer_{21} module. Interestingly, pMAK-1 is the only sequence in the GenBank database that contains an identical *aphA1* sequence to

pSRC27-H, which contains a silent mutation (see section 6.4.2), strengthening the possibility of a relationship.

Despite the fundamental similarity of the MARRs on pSRC27-H, pHCM1 and pO111_1, there were a number of insertions and deletions that had occurred within them (see Fig 6.17 and 6.19). Insertions of 2 additional regions were evident in pSRC27-H: the integron-containing region, deduced to have originated from IncN plasmids, and the *aacC2*-containing region, which has also been seen in an IncL/M plasmid from *C. freundii*. Small deletion events were observed multiple times in these plasmids, occurring directly adjacent to the insertion sequences IS*26* and IS*1*. For example, Tn*6029*B from pHCM1 is missing 85 bp of $tnpR_2$, directly adjacent to IS*26*(1), when compared to Tn*6029*. Also, directly adjacent to IS*1*b there is 356 bp from Tn*10* missing in pSRC27-H, compared to pO111_1 and pHCM1. The inference that insertion sequences cause deletions like these within complex MARRs, mainly effected by IS*26* and IS*1*, is examined experimentally in Chapter 8.

Here, the positions of MARRs have primarily been used as indicators of plasmid families, as each IncHI1 plasmid type has resistance region/s that differ in composition and location in the IncHI1 backbone (plasmid type denoted by different colours in Fig 6.20). The MARR of pSRC27-H was found to be in the same position as in pHCM1. The MARRs of 2 more recently sequenced IncHI1 plasmids, pO111_1 and pMAK-1 are also in this position (green in Fig 6.20). It is likely the insertion of these Tn*2670*/Tn*10*-derived MARRs is a product of the same event/s, indicating that pHCM1, pO111_1, pMAK-1 and pSRC27-H share a common ancestor. Therefore they were grouped into a plasmid group, named in this study pHCM1-type IncHI1 plasmids.

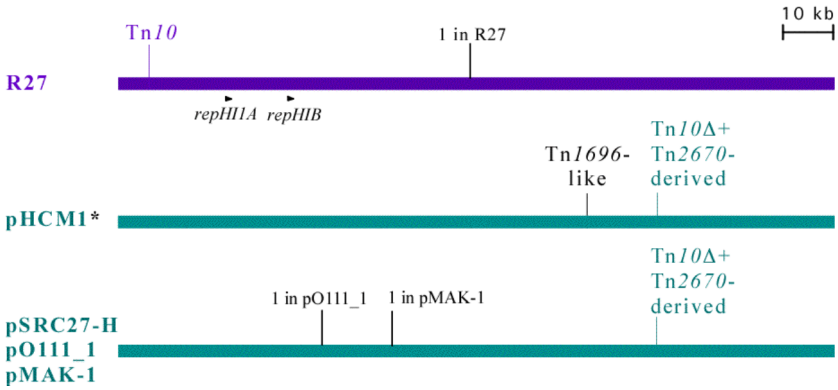


Fig 6.20 Location of resistance regions in IncHI1 plasmids. The horizontal lines represent the IncHI1 backbone, the colour of which shows the plasmid family: purple is R27 and green is pHCM1-type. The plasmid is opened at the start point of pHCM1, the start positions of other plasmids are indicated above the line as "1 in plasmid". The thin vertical lines show the insertion sites of the resistance regions, the names of which are above and those that are analysed in this chapter are emphasised in colour. These figures were drawn to scale using a backbone sequence from pHCM1 (GenBank acc. no. AL513383) and positions of MARRs in pO111_1, pMAK-1 and R27 using GenBank acc. nos. AP010961, AB366440, AF250878.

Tn*2670*, or a derivative of it, has inserted into the pHCM1-type plasmids, however, its original position was unable to be determined (it may have inserted into Tn*10*, or just near Tn*10* and an adjacent deletion subsequently occurred). Tn*2670* is able to move due to the directly orientated IS*1* bounding it, despite the lack of an intact Tn*21* transposition module [89]. It appears that Tn*2670* came in only once and evolution has occurred within this position. By analysing both the composition of MARR regions and insertion positions, the order that evolutionary events could have occurred was deduced and a possible evolutionary tree constructed (Fig 6.21). For example, when considering the pHCM1-type plasmids, we can infer that a hypothetical ancestor exists, "Plasmid X", that contains part of Tn*10,* part of Tn*2670* and Tn*6029*/Tn*4352*. This plasmid has spread and subsequently incurred changes within this framework to give rise to pHCM1, pSRC27-H and pO111_1. In Fig 6.21, although pO111_1 would have identical resistance regions to the most recent common ancestor of pHCM1-type plasmids, it is shown as separate to Plasmid X here because unique additional

events have occurred elsewhere in pO111_1, such as multiple insertions of IS*629* (see section 7.3.3, below).

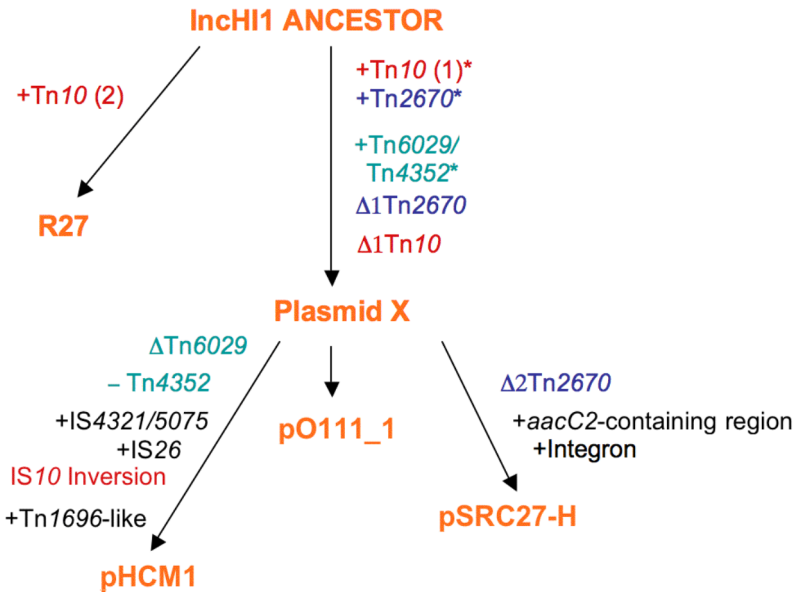


Fig 6.21 Evolution of IncHI1 plasmids via acquisition and loss of resistance regions. Plasmid names are in orange, the arrows represent an evolutionary path of how the plasmid the arrow points to was constructed from its possible ancestor above. Events that have occurred are detailed along the arrows connecting 2 plasmids. Events are colour coded to particular regions, where red is Tn*10*, blue Tn*2670* and green Tn*6029*/Tn*4352* events. Numbers in brackets next to a region indicate insertion position. The + preceding a region name means the region was gained, -lost and Δ means partial. Δ 1 or Δ2 are different amounts lost. The * indicates the order of these events unknown. This figure is not drawn to scale and the length of each arrow is not indicative of relatedness.

Evolutionary relationships of IncHI1 plasmids were further investigated in Chapter 7, by examining other features in their backbone.

CHAPTER SEVEN

Variation in the IncHI1 plasmid backbone

7.1 Introduction

In Chapter 6, the IncHI1 plasmid pSRC27-H was shown to belong to the same IncHI1 plasmid group as the sequenced plasmids pHCM1, pO111_1 and pMAK-1 (the pHCM1-type group) on the basis that they contained similar MARRs in the same position (Fig 6.20). Using this method, the well-characterised IncHI1 plasmid, R27, was allocated to a separate family.

A third IncHI1 plasmid recovered from a *S*. Paratyphi A isolate, named pAKU_1, was sequenced more recently (GenBank Acc. No. AM412236; [188]; see Table 1.2). pAKU_1 contains a Tn*2670*-derived structure that has undergone rearrangements, and a complete and separate Tn*10*. Both these regions are in different positions to the equivalent structures in either R27 or pHCM1. pAKU_1 also carries the *strA/B* in Tn*5393*, which was not identified in the original study (Fig 7.1).

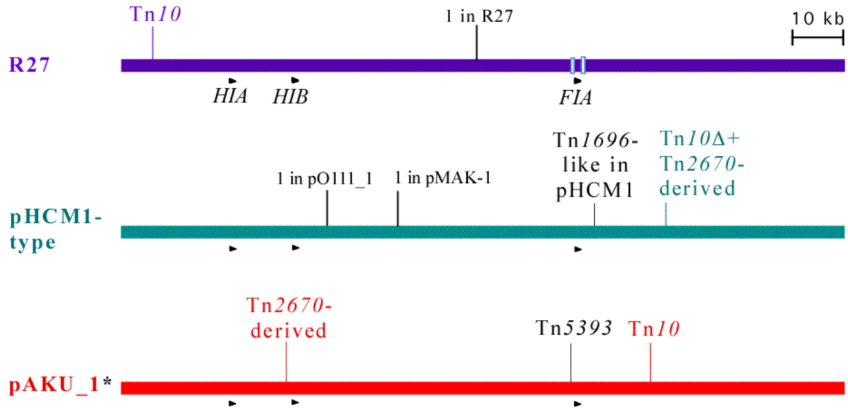


Fig 7.1 The positions resistance regions in IncHI1 plasmid families. Different colour lines represent backbones of different plasmid families. The plasmid is opened at the start point of pHCM1, the start positions of other plasmids are indicated above the line as "1 in plasmid". The thin vertical lines show the insertion sites of the resistance regions, the names of which are above and those that are analysed in this chapter are emphasised in colour. These figures were drawn to scale using a backbone sequence from pHCM1 (GenBank acc. no. AL513383) and positions of MARRs in pO111_1, pMAK-1, R27 and pAKU_1 using GenBank acc. nos. AP010961, AB366440, AF250878 and AM412236.

pAKU_1 had previously been compared to R27 and pHCM1 and they were found to share >99.97% identity [188]. Based on the SNPs observed between R27, pHCM1 and pAKU_1, pHCM1 had 3-fold more unique SNPs than either of the other 2 plasmids. This may suggest that R27 and pAKU_1 can be grouped into one family, and pHCM1 and pHCM1-type plasmids into another. Two previous studies have mentioned that pAKU_1 was more related

to R27 than to pHCM1 [163, 188], but they have also been classified as 3 separate plasmid groups using PMLST [163]. These previous studies had also defined 5 regions of backbone variation between these 3 IncHI1 plasmids (named Regions A-E; see Table 7.1 below) identified by comparing pHCM1 to R27 [163], and pAKU_1 to both pHCM1 and R27 [188]. By this criterion, both studies found that R27 and pAKU_1 were more related to each other than to pHCM1, and for some regions R27 and pAKU_1 had identical profiles (see section 1.7.1.2). However, using the positions of the resistance regions, no 2 IncHI1 plasmids were alike. These studies did not look for other regions of variability nor did they analyse the regions identified.

The aim of the work described in this chapter was to identify unique markers in the IncHI1 backbone and to use them to study the relationship between pSRC27-H and other IncHI1 plasmids. Furthermore, it aimed to give further insight into the evolutionary relationships and lineages of IncHI1 plasmids. Lastly, to develop a clear and efficient classification scheme of IncHI1 plasmids, based on these evolutionary lineages.

7.2 Regions surrounding the pSRC27-H MARR

Additional sequence to the left of IS*1*a and the right of IS*10* was obtained and further investigated, specifically for regions of variation between IncHI1 plasmids.

7.2.1 Right of the MARR

The 5.6 kb of sequence adjacent to IS*10* at the RHS of the MARR in pSRC27-H (Fig 7.2), obtained from the SacI clone pRMH955, was present in all pHCM1-type plasmids and once the IS*10* inversion was reversed in pHCM1 (see section 6.4.8) in the same position in all of them. Most of the sequence was also in the other IncHI1 plasmids R27 and pAKU_1. However, 687 bp of sequence present in the pHCM1-type plasmids was missing in R27 and pAKU_1. The indel was situated 684 bp to the right of IS*10* (shown in Fig 7.2 to the right of

the MARR). This difference in pHCM1, compared to R27, has been described previously as region "E" [163].

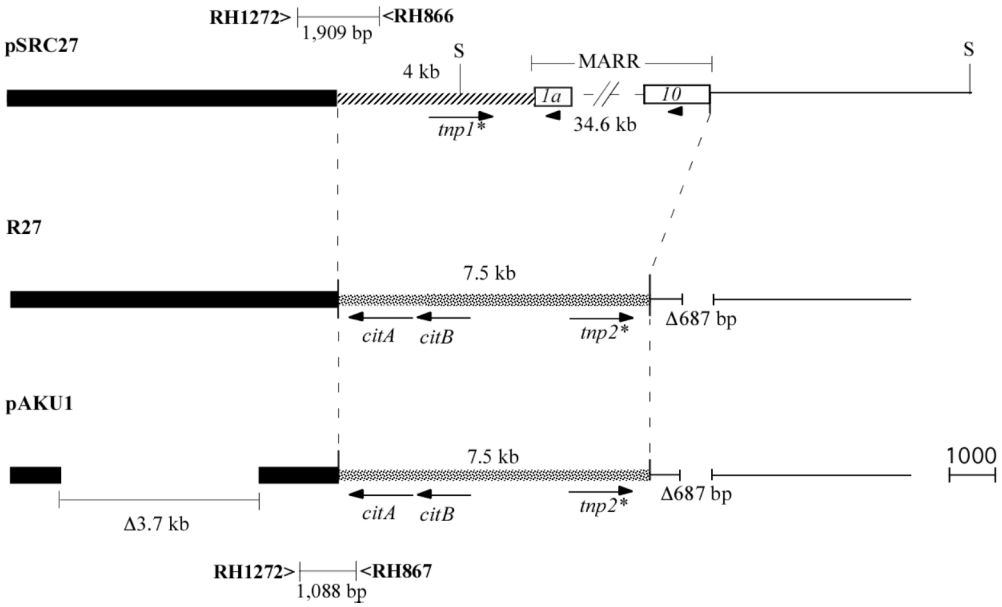


Fig 7.2 Regions surrounding the MARR. Regions are represented by lines of different types: the hatched line represents the region unique to pHCM1/pO111_1/pMAK-1, the spotted line the region unique to R27 and pAKU_1. Sequence missing is shown as a break in the horizontal line and size. Vertical dashed lines align the matching regions between the 3 plasmids."S" represents the SacI sites used for cloning and sequencing this region in pSRC27-H. The 34.6 kb MARR is condensed by 2 diagonal slashes and is not to scale. Primers are shown as thin vertical lines and below the horizontal lines connecting them is the expected size of the PCR product. Primers sequences are shown in Table A3.9 in Appendix 3. The direction and extent of genes are shown by horizontal arrows below. Only genes with predicted or known function are shown. The asterisk indicate that transposases *tnp1* and *tnp2* are 70% identical at the nucleotide level. pAKU_1 and R27 are drawn to scale based on GenBank Accession Numbers AP010961 and AF250878, respectively.

7.2.2 Left of the MARR

To the left of IS*1a*, 1.7 kb of sequence obtained from pRMH959, a SacI clone of pSRC27-H, (Fig 7.2) was present in the pHCM1-type plasmids, but not R27 or pAKU_1. A comparison of the R27 and pHCM1 sequences was used to identify where they began to share sequence again. A 4 kb region, which was unique to pHCM1-type IncHI1s was found to replace a 7.5 kb region present only in R27 and pAKU_1. Previously, this 4 kb region had been included in the MARR region of pHCM1, collectively named region "A" [163], but it had not been identified as a separate region or analysed. Most of this 4 kb sequence did not match anything

else in the GenBank database. However, a predicted ORF of 1,440 bp within the region, matched to 18 GenBank entries, including R27 and pAKU_1 and the IncHI2 plasmids R478 and pK29. In R27, the *tnp2* gene in Fig 7.2, shared 70% nucleotide sequence identity with *tnp1*. Interestingly, *tnp2* shares 97% identity at the protein level with the putative transposase found within an arsenic resistance transposon in IncHI2 plasmid R478 (described in section 4.4.1). Both *tnp1* and *tnp2* were described in the IncHI1 GenBank entries as encoding a "putative transposase". When this sequence was translated, the polypeptide was 68% identical to an experimentally proven "integrase" from the human immunodeficiency virus [262]. This "integrase" protein is actually an insertion sequence transposase, as it contains a D,D-35-E motif in its catalytic region, which is a common feature of many families of transposases (including those encoded by IS*26* and IS*1;* see section 1.3.1*)*. However, inverted or direct repeats could not be found in the surrounding sequence in pSRC27-H, and thus the boundaries of potential insertion sequences could not be defined. The 7.5 kb segment in R27 and pAKU_1 contained genes *citA/B* (Fig 7.2) which encode CitA, a protein responsible for citrate utilisation [263], as well as the putative *tnp2* gene described above.

To examine pSRC27-H, PCRs were designed that spanned the boundaries between the shared region and either the 4 kb or 7.5 kb regions (Fig 7.2). pSRC27-H produced a 1.9 kb amplicon, the correct size for the pHCM1-type boundary PCR and no product for the R27-type boundary (Fig 7.2). The sequence of pSRC27-H (GenBank Accession No. HQ840942) was extended to include the region up to the primer RH1272 and hence the entire 4 kb region (see Fig 7.2). This sequence was identical to the corresponding part of pHCM1, pO111_1 and pMAK-1. Features of MGEs, such as inverted repeats or genes encoding transposases, were not detected within the indel to the right of the MARR, or the 4 kb sequence to the left of the MARR. It is possible that the 4 kb adjacent to the MARR, and the MARR were gained as a single element, however, direct repeats surrounding the 4 kb region combined with the MARR were not found. A homologous recombination cross over site could not be identified

either. Thus, it remains to be seen how these regions were gained or lost and they were not further investigated. These variable regions located near the MARR in pSRC27-H were named. To the left the 4 kb region in pHCM1-type was allocated Region 1a, which replaces the 7 kb in R27 and pAKU_1, named Region 1b, and to the right, the 687 bp indel was named Region 5. Both Region 1a and Region 5 appear to be characteristic of pHCM1-type plasmids, as they were absent in both R27 and pAKU_1.

7.3 Further regions of variation between pHCM1 and R27

Using the criterion of regions that are present in pHCM1, but are absent in R27 (Regions 3, 4, 5 in Table 7.1), a further 3 regions of variation were identified. The properties of these regions are summarised in Table 7.1. A representation of where these regions are positioned in the IncHI1 backbone is shown in Fig 7.3

Table 7.1 Regions in pHCM1 and not R27

Region	Previous names		pHCM1 Position	Size (bp)	Description	GenBank matches	IRs	DR
	Holt[1]	Phan[2]					(bp)	
1a	-	(A)[3]	149,592-153,534	3,953[2]	4 kb segment + MARR **replaces** 7 kb in R27	4 kb has 1 ORF encodes putative transposase; 7kb has *citA/B* & 1 ORF encodes putative transposase	no	no[4]
2	-	(B)[3]	113,599-114,496 + 127,197-128,349	2,051	**insertion**; interrupted by Tn*1696*-like in pHCM1[5]	None	no	3
3 a/b	A/B	C	7,962-8,846; 9,040-10,457	a= 885; b=1,418	2 **indels**; a & b are 193 bp apart	None	no	no[4]
4	C	D	77,104-81,204	4,101	**insertion** in *trh* region in pHCM1; contains 2 ORFs	99.8% to IncHI2 plasmids; ORFs 1 and 2 match 68% to *Shewanella* chrom; ORF2 putat. DNA helicase	27	8
5	-	E	141,604-142,292	689	**indel**; 684 bp from Tn*10*Δ in pHCM1	None	no	no[4]

[1] As described in [188]
[2] As described in [163]
[3] This region was not specifically recognised in this study but was included as part of the resistance region
[4] no regions of homology to account for HR-mediated deletion were identified either
[5]The Tn*1696*-like in pHCM1 is a 12,700 bp insertion surrounded by 5 bp DRs

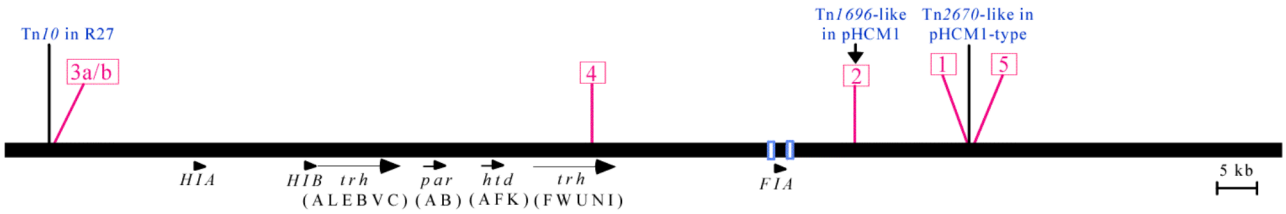


Fig 7.3 Regions present in pHCM1 and absent in R27. The pink lines show the positions of the variable regions 1-5 within the IncHI1 backbone of 157.6 kb. Positions of the resistance regions in pHCM1 and R27 are named in blue. The replication (*rep*), transfer (*trh*), partitioning (*par*) and stability (*htd*) regions are shown as arrows below. The blue boxes in the backbone are the 2 IS*1*s flanking the F-rep replication region.

7.3.1 Region 2

The previously described Region "B" includes the Tn*1696*-like transposon in pHCM1 [163], however it had not been analysed in detail. Here, a 2,051 bp segment consisting of sequence unique to pHCM1-type IncHI1 plasmids, was identified and designated Region 2 (see Fig 7.4, below). In pHCM1, it is interrupted by the insertion of the Tn*1696*-like transposon (Fig 1.21B), which is surrounded by a 5 bp direct repeat (Fig 7.4). This transposon therefore represents a latter addition. A 3 bp direct repeat was identified flanking Region 2, and in both R27 and pAKU_1 only 1 copy of the 3 bp was found (Fig 7.4). Although this is indicative of an insertion into pHCM1-type plasmids, inverted repeats were not detected at the ends of the insertion. The 2 predicted ORFs within Region 2 provide no additional information as they have no homology to any other ORFs in GenBank.

7.3.2 Region 3a/b

Regions 3a and 3b were described previously as regions "A" and "B" in Holt et al., (2007) [188] and collectively as region "C" in Phan et al., (2009)[163] (which used the same 2 sets of PCR primers as in [188]). Regions 3a and 3b are indels that consist of 885 bp and 1,418 bp of sequence, respectively, spaced 193 bp apart (see Table 7.1). Regions 3a and 3b exhibit no homology to other entries in GenBank and have no predicted ORFs. They have no features to indicate they are mobile elements. Region 3a is located 244 bp to the left of the insertion site of Tn*10* in R27 (Fig 7.3).

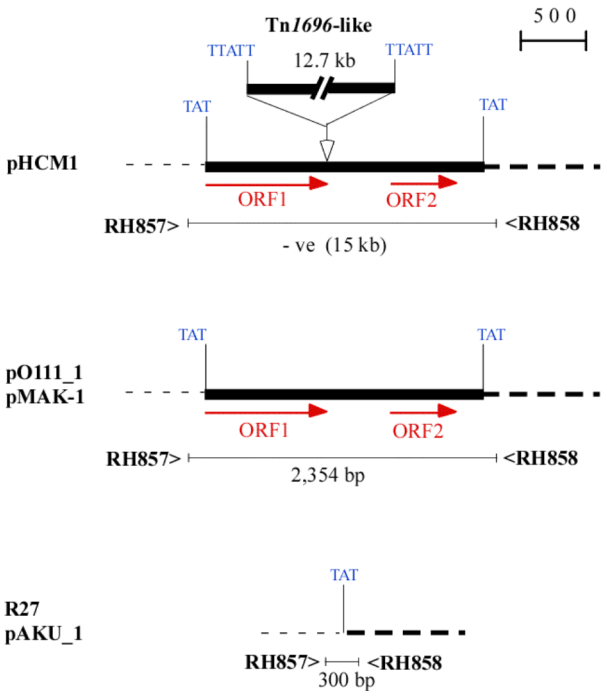


Fig 7.4. Region 2. The IncHI1 backbone is shown by dashed lines and the additional region by a solid black line. The Tn*1696*-like transposon in pHCM1 is shown above without detail. The open reading frames (ORFs) are shown in red below. The direct repeat sequences are shown above in blue lettering. Primers are shown as thin vertical lines below with the product size below. Primer sequences are in Table A3.9 in Appendix 3. Other features are as in 7.2, 7.3.

7.3.3 Region 4

Region 4 is not in R27 or pAKU_1 and has been described previously as region "C" by Holt et al., (2007)[188] and as region "D" by Phan et al., (2009)[163] (Table 7.1). It is a segment of 4,101 bp directly following the stop codon of *trhN*, in the second transfer region in pHCM1. In the present study, evidence of transposition was detected: an 8 bp direct repeat was found surrounding Region 4 (shown in Fig 7.5) and a 27 bp imperfect inverted repeat (3/27 bp mis-matches) was detected at the boundaries of Region 4 (LH IR: GGCTATTTGCAAACTAAGGTTGGAATT, RH IR (complemented and mismatches to the LH IR underlined for comparison): GGCTATTTGCAATCTGAGGTTGG-ATT). Region 4 is also in pMAK-1 and pO111_1. However, in pO111_1 a copy of the insertion sequence IS*629*, surrounded by a 3 bp direct repeat, was identified within Region 4 (Fig 7.5). Previously, IS*629* had been sequenced and its inverted repeats identified [264], but this is the first report

that identifies that it creates a 3 bp direct repeat upon insertion. IS*629* is in the IS*3* family of insertion sequences and members of which, such as IS*600*, have been shown to create 3 bp direct repeats [265].

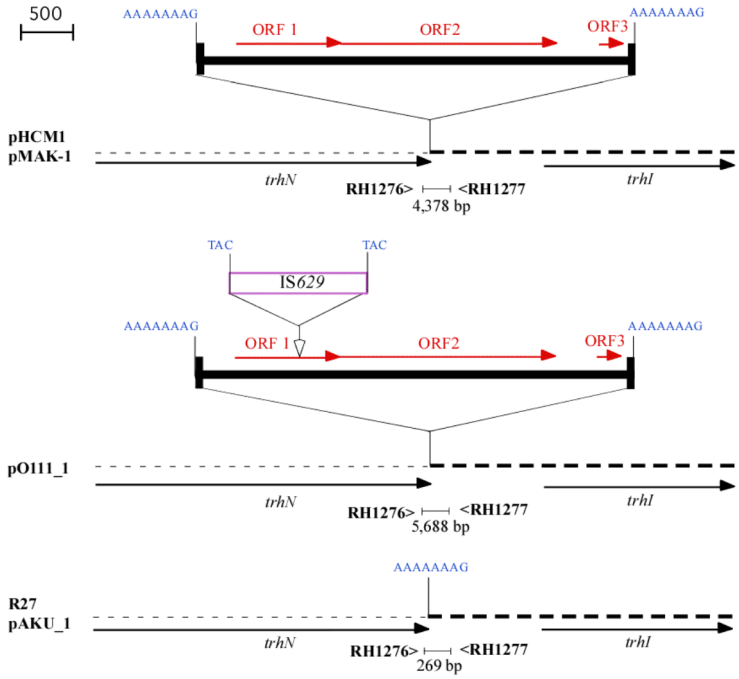


Fig 7.5 Region 4. The predicted open reading frames (ORFs) are in red. The 27 bp inverted repeats are shown as thick vertical lines and the insertion sequence IS*629* in a purple box above. Primer sequences are in Table A3.9 in Appendix 3. Other features are as in Fig 7.4.

Further evidence that Region 4 is a mobile element comes from the fact that it is also found in other locations. The entire Region 4 was also found in the IncHI2 plasmids R478, pK29, pEC-IMP and pEC-IMPQ (see Chapter 4), with a minimum 99.8% sequence identity. Although IncHI2 plasmids also contain a gene identified as *trhN* has 76% nucleotide sequence identity with *trhN*, Region 4 is inserted in pHCM1-type IncHI1s. Region 4 was found in adjacent to the tellurite resistance region and was flanked by a different 8 bp direct repeat. This supports that it is a translocatable element.

Two open reading frames (ORFs) were predicted within the region. From the start codon of orf1 to the stop codon of orf2 (but not the rest of Region 4), matched with 68% identity, only to a piece of the *Shewanella loihica* chromosome (GenBank acc. no. CP000606), which had

annotated ORF1 as having unknown function and ORF2 as being putative "DNA helicase". The predicted amino acid (aa) sequences of each ORF were analysed using BLASTp. The aa sequence of ORF 2 was predicted to encode the UvrD/REP DNA helicase protein (from a 68% match to a *S. loihica* protein; Protein ID, YP_001093393). UvrD has been shown to unwind DNA in a 3' to 5' direction [266]. The aa sequence encoded by ORF 1 did not match anything in the database and so its function could not be predicted.

Segments related to both 27 bp IRs and part of Region 4 adjacent to the IRs (as shown in Fig 7.6) were identified in 3 other nucleotide sequences (*K. pneumoniae* plasmids pKpQIL and pKPN4 and *K. pneumoniae* genome MGH 78578; see Fig 7.6) with >75% identity to pHCM1. These 3 sequences contained the same 8.1 kb (>99.98 % identity between them) within their IRs, but it was different to the sequence in IncHI1s (apart from the ~300 bp of sequence near the 2 IRs from Fig 7.6). This 8.1 kb region contained 4 ORFs, 3 of which encoded a putative restriction modification system (ORFs 2-4 in Fig 7.6) and one a putative transposase from the ISL3 family (ORF 1 in Fig 7.6). It was in the same position in the 2 plasmids, surrounded by an 8 bp duplication, but the genome sequence harboured the region in a different location, surrounded by a different 8 bp duplication. Another sequence, of a *Desulfovibrio desulfuricans* genome, (G20 in Fig 7.6) was 92% identical to these and was missing 2 segments within ORF 3. It was in another location, again surrounded by 8 bp direct repeats. The IRs and part of Region 4 next to the IRs also matched, with 79% identity, to a *Shewanella baltica* plasmid, pS18501, which had 9.4 kb in between the two IRs and had 8 bp direct repeats (as in Fig 7.6). It was 75% identical to the corresponding part of pKpQIL and contained the same ORFs 1-4. It also contained an additional ORF (ORF5; Fig 7.6), predicted to encode a C14 caspase subunit.

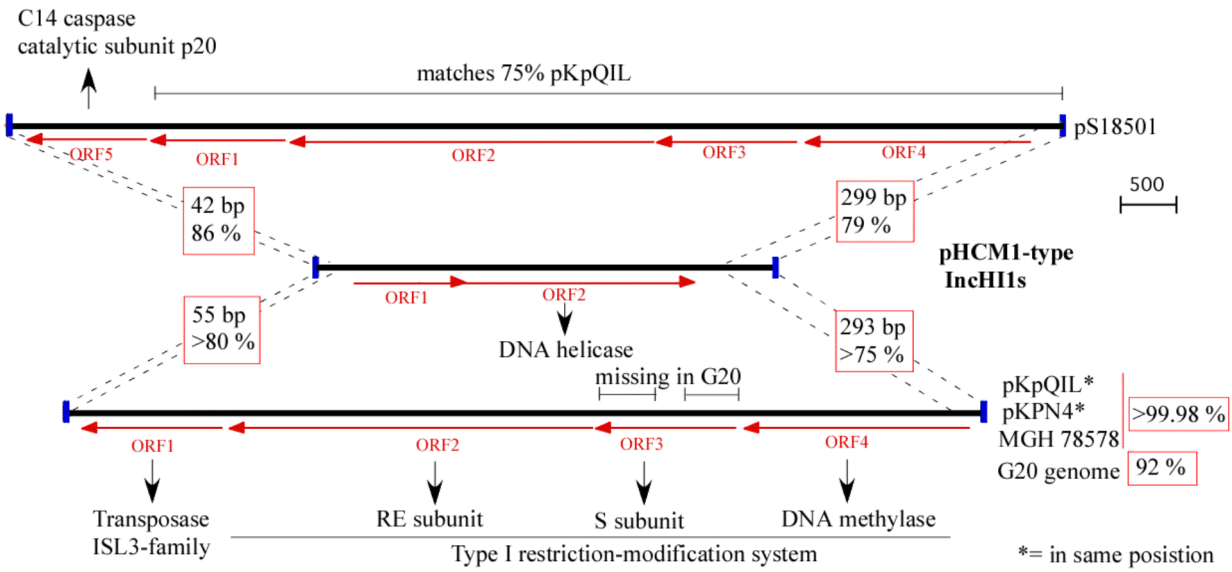


Fig 7.6 Parts of Region 4 in other sequences. Dashed lines are the segments of Region 4 from IncHI1s, that matched to other sequences in GenBank. The sequence identities between them are boxed in red. The predicted ORFs are shown in red and the arrows from them indicate the predicted protein product. The * indicates the region in the same position in these sequences. Drawn to scale using GenBank entries GU595196, CP000649, CP000647, CP000112 and CP000754 for pKpQIL, pKPN4, MGH 78578, G20 and pS18501, respectively. Other features are as in Figs 7.2, 7.3 and 7.4.

Region 4 shares features of transposons, specifically having inverted repeats and being flanked by direct repeats. Region 4 is seen in other plasmids, in a different position, which strongly suggests that this region is capable of movement by transposition. However, there is no sequence homology of the ORFs in Region 4 to a known transposase, either at the nucleotide or protein level. This region was not investigated further, but these preliminary findings warrant a more detailed analysis of this possibly new type of mobile genetic element.

7.4 Defining differences within pHCM1-type backbone

Regions 1-5 are unique to pHCM1-type IncHI1 plasmids, indicating that they have a distinctive backbone different to that of R27 or pAKU_1. To establish if any further variation was seen within the pHCM1-type IncHI1 backbone, the sequences of Regions 1-5 and the resistance regions were extracted from the pHCM1 sequence, and compared to the other pHCM1-type plasmids and R27 using BLASTn (Fig 7.7).

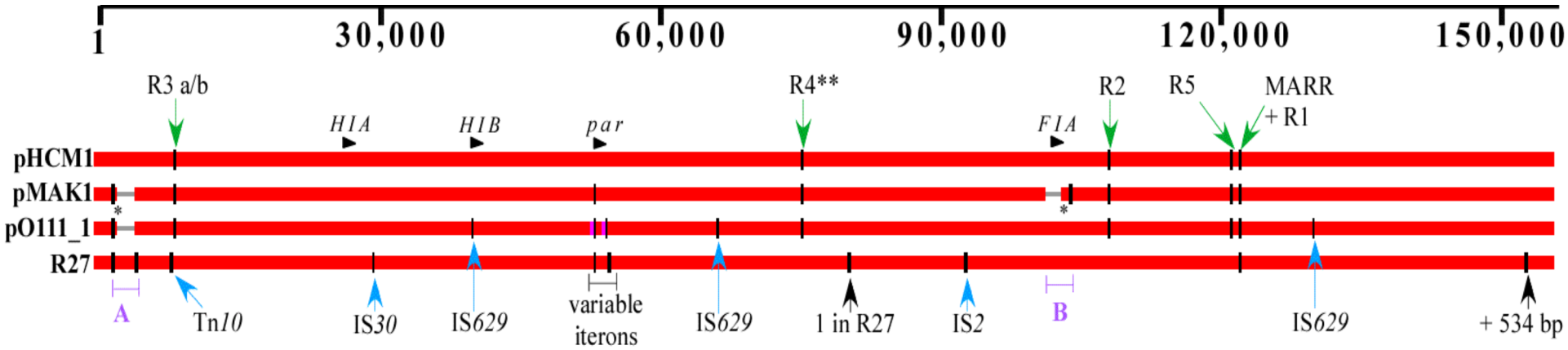


Fig 7.7 BLASTn analysis of the pHCM1 backbone. Matches to IncHI1 plasmids are represented by red lines, which are to scale with the positions (bp) on the horizontal black line above. Black vertical lines within the red lines indicate additional sequence at that position. The horizontal black arrows show the *rep* and *par* regions above. Features that differ between pHCM1-tye IncHI1s are bracketed in purple. The positions of Regions 1-5 and the pHCM1-type MARR are marked with green arrows. Positions of MGEs are marked with blue arrows. Other features are marked with black vertical arrows and brackets: "1 in R27" indicates the start position of R27 sequence; "+ 534 bp" indicates an additional 534 bp. * represents IS*1* that has matched but from a different position in the plasmid. ** denotes Tn*1696*-like is within Region 4 in pHCM1. Other features are as in Figs 7.2, 7.3 and 7.4.

From this analysis, the iteron regions were observed to vary. Iteron regions are highly variable as they contain tandem repeats (see 1.5.1) and the differences observed (Fig 7.7) were due to different numbers of repeats in each plasmid. Three copies of IS*629*, each surrounded by a 3 bp direct repeat, were observed in the pO111_1 backbone, in addition to the one identified in Region 4 (see Fig 7.5), which are not found in the other plasmids (Fig 7.7). Two main regions of variation were identified between the pHCM1-type plasmids (marked as A and B in Fig 7.7). One region (A in Fig 7.7), here named Region 6, is only present in pHCM1, but is seen in R27. The other (B in Fig 7.7), corresponding to the *repFIA* replication region, is missing only in pMAK-1. These regions are analysed below.

This comparison also revealed that 3 regions were present in R27 but not in the pHCM1-type backbone (Fig 7.7). Two insertions were identified as IS*30* and IS*2*. These insertion sequences were not found in pAKU_1. An additional 534 bp present in R27 was also seen in pAKU_1 and had been previously identified as a region of variation [188]. It did not have any features of a mobile element.

7.4.1 repFIA

In addition to the 2 functional replication regions, one unique to IncHI1 plasmids (*repHI1A*), one shared with IncHI2 plasmids (*repHI1B*), IncHI1 plasmids carry part of the *repA* region from incompatibility group F plasmids (*repFIA*) here referred to as "F-rep". F-rep confers the property that when an IncHI1 plasmid was introduced into a cell containing an IncF plasmid, the latter was lost [173]. The F-rep segment, is flanked by directly orientated IS*1* (see Fig 7.2) and therefore has the potential to be lost. The pHCM1-type IncHI1 plasmid pMAK-1 did not contain this region (marked as B in Fig 7.7), but instead contained a single IS*1* at the position where F-rep, flanked by 2 IS*1,* was located in the other IncHI1 plasmids. However, direct repeats were not found flanking the single IS*1* in pMAK-1 or the 2 IS*1* in the other IncHI1 plasmids, indicating subsequent events have occurred.

7.4.2 Region 6

A segment of 2,333 bp found in pHCM1, which was not in the other pHCM1-type plasmids pO111_1 and pMAK-1 was named Region 6 (marked as A in Fig 7.7). It consisted of an IS*1* and 1,697 bp of sequence, which matches to 41 *E. coli* chromosomes (see for example, GenBank accession numbers: CU928160; AP009378) with 99.1- 97.4% sequence identity. None of the *E. coli* chromosomal regions are adjoined by an IS*1*. The region contains one ORF, predicted to encode a 316 amino acid sequence, which is 100% identical to that of a protein named CorA. CorA has been shown to be important in Mg^{2+} transport [267]. Region 6 was found to be surrounded by an 8 bp direct repeat of the sequence found in pO111_1 and pMAK-1 (see Fig 7.8). This size is consistent with IS*1* transposition (Table 1.1) and indicates that the region has probably been mobilized from the *E. coli* chromosome by the IS*1*. Region 6 is found in both R27 and pAKU-1, but is inverted relative to the backbone compared to pHCM1 (Fig 7.7). This inversion had previously been noted [187, 188]. Another difference within this region is that pHCM1 is missing 33 bp within the IS*1* whereas R27 and pAKU_1 have an intact IS*1* (Fig 7.8).

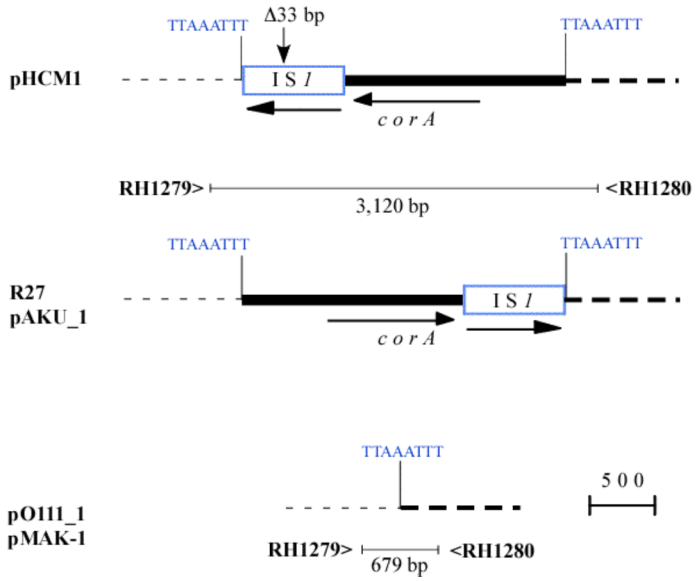


Fig 7.8 Region 6. Blue box is the IS*1*, the direct repeat sequence is shown in blue lettering above. The IncHI1 backbone is shown by dashed lines, of different thickness to orientate Region 6 in comparison the backbone. Primer sequences are in Table A3.9 in Appendix 3. Other features are as in 7.4, 7.2, 7.3.

7.5 The backbone of pSRC27-H

The presence of variable regions 1-6 and F-rep in pSRC27-H was tested using a number of specific PCRs. Each PCR was designed so that the primers were positioned in the backbone sequence (one on either side of the variable region) and the product spanned the region. Each was performed on pSRC27-H, using R27 as a control. The identities of the PCR products were confirmed using digestion with the restriction enzyme RsaI. The results obtained are tabulated in Table 7.2, together with the sizes of predicted fragments from the available sequences. From the PCR results present, pSRC27-H clearly shares all of the features that are characteristic of members of the pHCM1 family.

Table 7.2 PCR of variable regions in the IncHI1 plasmid backbone

Region	Fig	Primers[1]	PCR product size (bp)					
			Experimental		Predicted[2]			
			R27[3]	**pSRC27-H[4]**	**pHCM1[4]**	**pO111_1[4]**	**pMAK-1[4]**	pAKU_1
1a	7.2	RH1272/RH866	-	**1,909**	**1,909**	**1,909**	**1,909**	-
1b	7.2	RH1272/RH867	1,088	**-**	**-**	**-**	**-**	1,088
2	7.4	RH857/RH858	300	**2,354**	**14,751[5]**	**2,354**	**2,354**	300
3a/b	-	RH1278/RH616	-[6]	**3,027**	**3,027**	**3,027**	**3,027**	1,004
4	7.5	RH1276/RH1277	269	**4,378**	**4,378**	**5,688[7]**	**4,378**	269
5	7.2	ND[8]	-	**+**[9]	**+**	**+**	**+**	-
6	7.7	RH1279/RH1280	3,120[10]	679	**3,087**	**679**	**679**	3,120[10]
F-rep	-	RH1281/RH1282	3,354	**3,354**	**3,354**	**3,354**	**942**	3,354

[1]primers sequences are in Table A3.9 in Appendix 3
[2]predicted from GenBank sequences
[3]PCR products were not sequenced - exact size predicted from GenBank sequence
[4]pHCM1 type plasmids are highlighted in bold type
[5]region present but too large to be amplified; interrupted by Tn*1696*-like
[6]region is present but too large to be amplified; PCR product interrupted by Tn*10* which is adjacent to Region 5
[7] size increased due to insertion of IS*629*
[8] determined by sequencing of clone not by PCR
[9] + = region is present, but cloning not PCR was used
[10]orientation 2 relative to the backbone compared to pHCM1. Orientation distinguished by RsaI digestion pattern

7.6 Summary and discussion

The backbones of pHCM1-type plasmids pO111_1 and pMAK-1 had not been previously been described in any detail, and this study is the first to analyse them. Six variations were used to examine evolutionary relationships of IncHI1 plasmids and pSRC27-H is clearly of

the same type as pHCM1, pO111_1 and pMAK-1. This is tabulated in Table 7.3, where common features are highlighted in blue. A single "Variable Region Type" (VRT) was used to summarise the properties of the backbone variable regions for each plasmid (Table 7.3). Plasmids within the same VRT display no more than one difference in the presence or absence of these regions.

Table 7.3 Variable regions between IncHI1 plasmid types

Variable Region Type	Plasmid	Variable region						
		1	2	3a/b	4	5	6	F-rep
1	**pSRC27-H**	**a**	**+**	**+**	**+**	**+**	**-**	**+**
1	pHCM1	a	+[1]	+	+	+	+[2]	+
1	pO111_1	a	+	+	+[3]	+	-	+
1	pMAK-1	a	+	+	+	+	-	-
2	R27	b	-	-	-	-	+[4]	+
2	pAKU_1	b[5]	-	-	-	-	+[4]	+

[1] + Tn*1696*-like insertion
[2]Orientation 2 relative to the backbone after the inversion
[3] + IS*629* insertion; pO111_1 also carries 3 additional IS*629* elsewhere
[4]Orientation 1 relative to the backbone
[5] pAKU_1 has a deletion near Region 1 (Fig 7.2)

7.6.1 Two IncHI1 backbone types

Four IncHI1 plasmids, including pHCM1 are Variable Region Type 1 (VRT1). pSRC27-H appears to represent the progenitor backbone for this group, as other VRT1 plasmids showed minor variations, such as IS insertion, within the 6 regions. R27 and pAKU_1 are both Variable Region Type 2 (VRT2) (highlighted purple in Table 7.3). Hence, there appears to be 2 main IncHI1 plasmid backbones represented by VRT1 and VRT2 type plasmids. The presence of Regions 1- 5 is unique to VRT1 plasmids in all cases. These regions consist of 2 indels, 1 region adjacent to the MARR and 2 possible mobile genetic elements. Sequence obtained surrounding the MARR and PCRs designed in this study that target Regions 1- 5 allowed pSRC27-H to be allocated to VRT1. As these PCRs readily distinguish VRT1 from VRT2 IncHI1 plasmids, they could be used a simple tool to screen for VRT1 IncHI1 plasmids. Analysis of a larger IncHI1 plasmid set should confirm the effectiveness of these tools and could define other VRTs.

One study had previously recognised Regions 3 and 4 and used the presence of these regions, as well as an IS*1*-bounded region unique to pAKU_1, to compare R27, pHCM1, pAKU_1 and 2 other pAKU_1-like plasmids [188]. Using this method, the authors concluded that R27 and pAKU_1 were more related to each other than to pHCM1, and this also correlated with SNP-typing analysis they performed (see below). Another study performed on a larger set of plasmids, had recognised variable Regions 1- 5 (named regions A-E in Phan et al., 2009[163]; see Table 7.1), however from the data presented, it was not possible to discern whether Regions 1 and 2 were present, as they were combined with the pHCM1 resistance regions. Although the primary tool of their classification was PMLST (see below), the presence of Regions 3 and 5 did largely correlate with their 2 main PMLST groups. Interestingly, some older plasmids (isolated in the 1970s) that otherwise share the features of VRT1, did not contain Region 4. Thus Region 4, which has features of an MGE, may have been a more recent acquisition of VRT1 plasmids and it may represent a point of variation, which could be used to further classify IncHI1 plasmids.

Neither of the earlier studies found features of MGEs in any of these regions, and one study claimed that they specifically looked for them [163]. In this study, evidence is presented that suggests that some of these regions may be mobile elements (Table 7.1): Regions 2 and 4 are flanked by direct repeats and Region 4 has inverted repeats and is also seen in other plasmids in different positions. Most of these regions, however, did not contain genes predicted to encode known proteins to effect movement, such as transposases or integrases. However, transposases encoded elsewhere in the genome may be acting *in trans*. Alternatively, these regions may represent novel MGEs, in which case it may be useful to further characterise and study them. For example, does the DNA helicase play a role in effecting movement of Region 4? However, this type of analysis was beyond the scope of this work.

A further region, was previously found to be present in pHCM1, R27 and pAKU_1, but inverted in pHCM1 [188], but there were insufficient plasmid sequences available to define the element. Here, it is named Region 6 and because it was not present in pSRC27-H or pO111_1 or pMAK-1, its boundaries could be defined. It was concluded that this region was not characteristically present in VRT1 plasmids, despite the fact that pHCM1 contained it. Further IncHI1 plasmids need to be analysed to confirm this. Region 6 in both orientations (in pHCM1 and VRT2 plasmids) was flanked by an 8 bp duplication and contains a segment derived from the *E. coli* chromosome, indicating that it has been mobilised. However, there is only 1 copy of IS*1*, thus the mechanism of acquisition or inversion of Region 6 in pHCM1, compared with R27 and pAKU_1, is not obvious. It is theoretically possible that one IS*1* could move or invert a segment of DNA if the end of this sequence displayed high homology to the IR of IS*1*, which the IS*1* could recognise, but no such homology could be found. How this element works also warrants investigation.

7.6.2 Variations within VRTs

It is to be expected that plasmids would accrue other minor variations and no two VRT1 plasmids were identical, even over these defining regions. However, most of these changes were simply insertions, such as the insertion of IS*629* in Region 4 in pO111_1 or the Tn*1696*-like Tn in Region 2 in pHCM1. These presumably represent more recent events. Similarly, the presence and location of resistance regions vary within VRT2 plasmids and have the potential to be used to further classify plasmids within VRTs. For example, plasmids R27 and pAKU_1, which are VRT2, both contain Tn*10* but in different positions; this suggests the acquisition of Tn*10* occurred more recently than these variations in the backbone.

In the VRT1 group, pMAK-1 did not contain the F-rep IS*1*-bounded replication region, in the IncHI1 backbone. pMAK-1 is the first IncHI1 plasmid reported that does not contain an F-rep

region. Previously, the incompatibility of F plasmids with IncHI1 plasmids was used as a phenotypic marker to differentiate IncHI1 from other, particularly from the related IncHI2 plasmids, which do not display this property [141, 173]. Because the F-rep region is clearly not essential to IncHI1 plasmids, it should not be used to classify them. This finding potentially has consequences involving the misclassification of older IncHI1 and IncHI2 plasmids, based on the presence of F-rep and older collections should be re-analysed using PBRT.

7.6.3 Other approaches to classification

In addition to examining the presence of variable regions in the backbone, 3 previous studies have trialled a number of other approaches to elucidate evolutionary relationships of IncHI1 plasmids. The first used restriction fragment length polymorphism (RFLP) of IncHI1 plasmid DNA with BamHI [187]. They showed that certain plasmid RFLP patterns were common to IncHI1 plasmids in outbreaks of *S*. Typhi, which may reflect the clonality of the isolates. They also showed that the RFLP profiles of pHCM1 and STY7 (related to pAKU_1) were more similar to each other than to R27, most likely because they both contain a Tn*2670*-like element and R27 does not. From this, they proposed a stepwise evolution of R27 to pHCM1 to STY7, which they later revised [163]. This does, however, illustrate that RFLP is not an ideal tool to classify plasmids. Antibiotic resistance regions tend to have more restriction sites in them than backbone sequence and often produce lower molecular weight fragments, which are better separated by standard gel electrophoreses than higher molecular weight bands from the backbone.

The second study used SNP typing which showed that there were 3 fold more SNPs in pHCM1 compared to R27 or pAKU_1, across the shared IncHI1 backbone [188]. This data correlates well with the VRT grouping described here, but requires a complete plasmid

sequence and cannot be performed efficiently on a large number of isolates using current technology and resources.

The third study used PMLST, microarray/hybridisation analysis and long range PCR [163]. The long range PCR across the backbone was not effective as it was only able to map approximately half the backbone for the plasmids studied. The pHCM1 sequence was used as a template for primer design and the null results obtained may have been due to the unique features of pHCM1 (such as a large inversion and insertions). In light of the constant and variable regions from a number of IncHI1 plasmids analysed here, a re-assortment of primers may solve this. The microarray/hybridisation approach was able to rapidly identify all genes in the backbone, and indeed Regions 1- 5 described here. However, the data was made not available and so cannot be re-assessed here. However, this method is potentially less accurate than using VRT PCR for variable regions as it shows the presence of a gene, but not the location. This can also be problematic when probing for regions that occur in multiple copies, such as resistance genes. Phan et al., (2009) also set up a PMLST scheme for IncHI1 plasmids, based on sequence types (ST) using single bp changes to define 2-3 alleles over 6 loci [163]. pHCM1 was defined as ST1, pAKU_1 as ST7 and R27 as ST5 (where pHCM1 and R27 shared 2/6 alleles, pHCM1 and pAKU_1 1/6 and pAKU_1 and R27 4/6). In the current study, both pO111_1 and pMAK-1 were identified as ST2, which differed from pHCM1 by only a single nucleotide. Each of the STs were allocated to 2 overarching IncHI1 plasmid groups which were used as an indication of relatedness. STs that have 5/6 identical loci are grouped together: ST1-ST4 are in “Group 1”, ST6-ST8 in “Group 2” and ST5 stands alone. Thus pHCM1, pO111_1 and pMAK-1 would all be in Group 1, which is consistent with data presented in the current study. However, this PMLST scheme separates plasmids based on only 1 bp change within a locus, and this method has the potential to mis-classify plasmids based on single base sequencing errors, especially for older plasmids like R27, which were sequenced over 10 years ago. It was not mentioned whether the studies re-

sequenced PMLST loci or checked for sequencing errors. If R27 had only one less SNP compared to pAKU_1, they would be grouped together in Group 2, and this would be more consistent with the VRT and SNP typing results.

Overall, testing these MARR positions using PCR across their boundaries, is more time efficient and less costly than sequencing 6 PMLST loci or microarray/hybridisation. It seems to be a more robust method of distinguishing IncHI1 plasmid lineages.

7.6.4 Spread of VRT1 IncHI1s

To determine the temporal and spatial spread of IncHI1 plasmids, their origins and properties were tabulated (Table 7.4). These IncHI1 plasmids were isolated from a range of geographical sources (Table 7.4). The VRT1 all occurred within the Asia-Pacific region. A previous study examined some plasmids from a wider geographical range [163] that were grouped with pHCM1 via PMLST and lacked Region 4, but data is not available to determine their full VRT profile. The simple tools designed in this study to track IncHI1 plasmids, could be used in the future to examine isolates from a wider geographical range.

Table 7.4 Properties of IncHI1 plasmids

Plasmid	**Isolate**				**PMLST**		**Variable Region Type**
	Organism	Source	Year	Location	ST	Group	
pSRC27-H	*S.* Typhimurium	Equine	1999	Australia	ND	ND	1
pHCM1[1]	*S.* Typhi	Human	1993	Vietnam	1	1	1
pO111_1	*E. coli*	Human	2001	Japan	2	1	1
pMAK-1[2]	*S.* Dublin	Porcine	?	Japan	2	1	1
R27	*S.* Typhi	Human	1961	UK	5	Singlet	2
pAKU_1[1]	*S.* Paratyphi A	Human	2002	Pakistan	7	2	2

[1]After theoretical re-inversions between 2 oppositely orientated IS to bring them back to predicted original configuration
[2]MAR region sequence may be incorrectly assembled

CHAPTER EIGHT

Loss of antibiotic resistance genes from pSRC27-H

8.1 Introduction

The complex 34.6 kb multiple antibiotic resistance region (MARR) in pSRC27-H described in Chapter 6, is composed of the IS*1*-bounded transposon Tn*6027* and part of Tn*10* and it contained 11 antibiotic resistance genes and 9 insertion sequences (Fig 6.16). pSRC27-H was shown to belong to the VRT1 plasmid family (defined in Chapter 7), along with the IncHI1 plasmids pO111_1, pHCM1 and pMAK-1. These plasmids contained MARRs of similar composition located in the same position as in pSRC27-H. However, a number of differences were identified within the 4 MARRs, including insertions, deletions and inversions. The main forces driving the evolution of these MARRs appeared to be the insertion sequences IS*26*, IS*1* and IS*10*. A simple way to investigate the evolution of MARRs experimentally, is to study the loss of antibiotic resistance genes, by screening for loss of antibiotic resistance. Because of the large number of insertion sequences and resistance genes within the MARR in pSRC27-H, it was an ideal candidate to examine. Three main ways that antibiotic resistance genes could be lost were envisaged. First the entire plasmid could be lost from the cell. The second, homologous recombination between identical sequences within the cell, for example, of 2 directly orientated IS*26* or the almost identical copies of IS*1*. This would result in loss of DNA in between the 2 IS, together with 1 copy of the IS (see Fig 8.1). Third, deletion of adjacent DNA, which is an active process that is effected by the IS-encoded transposase and can occur in either direction from the IS (see Fig 8.1). In addition, more complex routes of loss could occur for example inversion of DNA between 2 IS in inverse orientation (which would rearrange the MARR), followed by adjacent deletion.

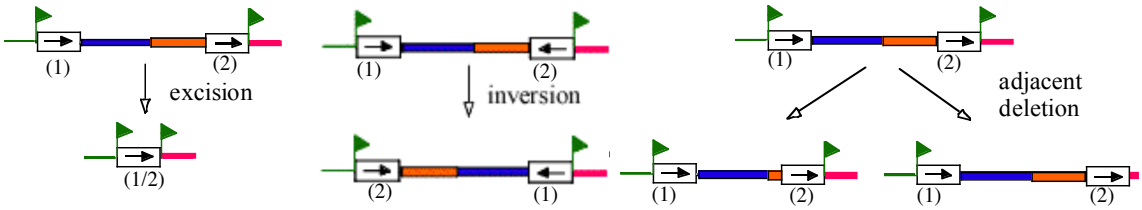


Fig 8.1 Types of IS-mediated loss or rearrangements. Deletion adjacent to IS(2), excision and inversion events. Green flags are direct repeats, open boxes are insertion sequences, each are numbered arbitrarily, and the arrows within them show their orientation.

IS*1* has been shown to create adjacent deletions at relatively high frequencies of approximately 5 x 10^{-4} [93], yet transpose at the relatively low frequency of 10^{-7} in a standard mating assay [29]. Transposition occurs via either simple "cut and paste" transposition or cointegrate formation, where an intermediate composed of 2 directly repeated copies of the IS, one at each junction between target and donor replicons, is generated. The cointegrate is then separated by resolution. IS*10* has been shown to create adjacent deletions and transpose at similar frequencies to IS*1* [94] and to promote loss of DNA in between 2 directly orientated IS*10* via a different mechanism, proposed to involve intrastrand base pairing and replication slippage [29]. Although IS*1* and IS*10* are quite well studied experimentally, IS*26* is not. It has been shown to transpose via a cointegrate at frequencies ranging from 10^{-5} to 10^{-7} [268]; [269], but most of the evidence that IS*26* causes deletion is based upon observational data, where deletions are seen directly adjacent to IS*26*, compared to other equivalent regions. For example, Tn*6029*B in pHCM1 has 85 bp of sequence missing directly adjacent to IS*26*(1), compared to Tn*6029* in other pHCM1-type IncHI1s (see Fig 6.19). Other studies have described evolution of resistance regions, driven by IS*26* (see for example [42]; [243]; [104]; [270]).

The aim of the work described in this chapter was to investigate the loss of resistance genes from the MARR in pSRC27-H in the absence of antibiotic selection. Also, to identify which insertion sequences are the forces driving the evolution of the MARRs in VRT1 IncHI1 plasmids.

8.2 Loss of pSRC27-H

The stability of resistance genes in pSRC27-H was evaluated experimentally. Bacterial cultures were grown at 37^{O}C without antibiotics for an extended period of time, continued by

serial dilution. Initially, *E. coli* strain E294 transconjugants, that carried all resistances, were cultured and tested for loss of resistances and plasmids. Subsequently, it was shown these isolates contained both the IncHI1 and IncI1 plasmids in SRC27, namely E294/pSRC27-H/pSRC27-I (see section 6.3). During these experiments less than 1% of cells lost the pSRC27-H plasmid after 280 generations (see below).

When these experiments were repeated on *E. coli* cells containing pSRC27-H plasmid alone (E294/pSRC27-H; see Fig 6.2), the majority of the cells lost all resistances and the IncHI1 plasmid. For example, after 280 generations, from a culture of E294/pSRC27-H grown without antibiotic selection, 120 out of 200 individual colonies tested were sensitive to all antibiotics and PBRT confirmed that they had lost the IncHI1 plasmid. Thus, when pSRC27-I was present, pSRC27-H remained relatively stable in these experiments (discussed below in section 8.6).

8.3 Loss of resistance genes from pSRC27-H

Duplicate cultures (culture 1 and culture 2) of different E294/pSRC27-H/pSRC27-I isolates were grown overnight at 37^{O}C without antibiotics, diluted and ~1,000 cells were added to fresh LB and grown overnight again. Culturing was continued in this way for 14 days (for complete method see 2.3.5). At various points, after days 1, 2, 7, and 14, (corresponding to 20, 40, 140 and 280 generations), individual colonies from both cultures were isolated on non-selective LA and tested for their resistance phenotype. 500 individual colonies from each culture (total of 1,000 isolates) were screened for loss of antibiotic resistances by patching them onto a series of plates containing a single antibiotic. Antibiotics used were ampicillin, chloramphenicol, kanamycin, streptomycin, and trimethoprim. Colonies that had lost resistance to one or more of these antibiotics were recovered. These colonies were patched

onto plates containing each of the antibiotics that E294/pSRC27-H/pSRC27-I was resistant to (Ap, Cm, Gm, Km, Nm, Sp, Su, Sm, Tc and Tp).

8.3.1 Resistance variants recovered

Out of the 1,000 colonies screened, 17 colonies from culture 1 and 22 from culture 2 displayed antibiotic resistance phenotypes that differed from E294/pSRC27-H/pSRC27-I (Table 8.1). These variants exhibited 9 resistance phenotypes (Table 8.1). Colonies of the same resistance phenotype were allocated a "variant type" number. To determine if the resistance variants still contained the plasmids, PBRT targeting only the IncHI1 and IncI1 replicons was performed on each variant. All variants retained both the IncI1 and IncHI1 replicons except Variant 9, which displayed only streptomycin resistance and had lost the IncHI1 replicon (Table 8.1). Variant 9 had lost the IncHI1 plasmid and thus was equivalent to E294/pSRC27-I. It was included in subsequent analyses as a negative control.

Table 8.1 Resistance variants of SRC27 transconjugants obtained

Type	**Phenotype:**		**Culture**		**Total**	**PBRT**
	Resistant to	**Sensitive to**	**1**	**2**		
Parent	ApCmGmKmNmSmSpSuTcTp	None	483	478	**961**	HI1, I1
1	ApCmGmSmSpSuTcTp	KmNm	1	0	**1**	HI1, I1
2	ApCmGmSmSpTcTp	KmNmSu	9	3	**12**	HI1, I1[1]
3	ApCmGmKmNmSmSu	SpTcTp	2	1	**1**	HI1, I1[2]
4	ApCmKmNmSmSuTc	GmSpTp	0	7	**7**	HI1, I1[3]
5	SmSpTcTp	ApCmGmKmNmSu	1	2	**2**	HI1, I1
6	CmSmTc	ApGmKmNmSpSuTp	0	2	**2**	HI1, I1
7	CmSm	ApGmKmNmSpSuTcTp	0	2	**2**	HI1, I1
8	SmTc	ApCmGmKmNmSpSuTp	0	1	**1**	HI1, I1
9	Sm	ApCmGmKmNmSpSuTcTp	4	4	**8**	I1[1]

[1]Two isolates from each culture tested
[2]One isolate from each culture tested
[3]Three isolates in total tested

Only 3 variant types occurred in both cultures. Because colonies with the variant 2, 3, and 5 resistance phenotypes were isolated from both cultures, they should have arisen from independent events that happened to produce the same resistance phenotype. Therefore, isolates of the same phenotype from different cultures were assigned different names and tested separately in all subsequent analysis. For example, colonies sensitive to Sp, Tc and Tp

(Variant 3 phenotype) were obtained in both cultures and isolates from culture 1 were named Variant 3-1 and those from culture 2, Variant 3-2, and so on for the other variants. All colonies isolated that varied from the parent phenotype were analysed by CDS and PBRT, except Variant 4 where only 3 isolates were tested and Variant 2 where 1 isolate from each culture was tested. These were used as representatives for that variant type, for all further analysis (Table 8.1). The resistance phenotypes of each variant were confirmed using the CDS method (zone sizes of all variants are given in Table A1.3, and the plate pictures of representative colonies are provided in Fig A1.1, in Appendix 1).

8.3.2 Resistance genes in variants

The variants were analysed for the presence of all the resistance genes in E294/pSRC27-H/pSRC27-I (see Table 6.2) using PCR internal to the genes or a PCR across the 3'-CS and 5'-CS regions of the integron to detect the *aadA2* and *dfrA12* gene cassettes (Table 8.2; primer sequences in Table A2.1 in Appendix 3). In all cases, when the resistance phenotype was lost, the PCR for the corresponding resistance gene was negative. Resistance genes that had been lost appeared to occur in groups, where missing resistance genes were originally next to one another and this suggests a single deletion.

Table 8.2 Resistance genes in variants[1]

Type[2]	Resistant to	*catA1*	*bla*$_{TEM}$	*sul2*	*strA*/ IS*26*	*aphA1*	*aacC2*	*dfrA12* / *aadA2*	*tetA* (B)	*strA*/ IS*1133*
Parent	ApCmGmKmNmSmSpSuTcTp	+	+	+	+	+	+	+	+	+
1	ApCmGmSmSpSuTcTp	+	+	+	+	-	+	+	+	+
2-1	ApCmGmSmSpTcTp	+	+	-	-	-	+	+	+	+
2-2	ApCmGmSmSpTcTp	+	+	-	-	-	+	+	+	+
3-1	ApCmGmKmNmSmSu	+	+	+	+	+	+	-	-	+
3-2	ApCmGmKmNmSmSu	+	+	+	+	+	+	-	-	+
4	ApCmKmNmSmSuTc	+	+	+	+	+	-	-	+	+
5-1	SmSpTcTp	-	-	-	-	-	-	+	+	+
5-2	SmSpTcTp	-	-	-	-	-	-	+	+	+
6	CmSmTc	+	-	-	-	-	-	-	+	+
7	CmSm	+	-	-	-	-	-	-	-	+
8	SmTc	-	-	-	-	-	-	-	+	+

[1]genes are listed in the order they appear within the MARR
[2]Variant 9 also included as a negative control

However, in Chapter 6, it was determined that there were two copies of *strA/B* in E294/pSRC27-H/pSRC27-I (Fig 6.5). One was located in Tn*6029* (next to IS*26*(3)) on pSRC27-H, and the second in Tn*5393* (next to IS*1133*) on pSRC27-I. In order to determine which copy/s had been lost, a set of PCRs to detect *strA* in these locations, as previously described (Fig 6.5), was performed (Table 8.2). The *strA/B* genes next to IS*1133*, as in pSRC27-I, were detected in all variants, which is consistent with all variants harbouring the IncI1 plasmid. This also indicates that all of the differences between the variants occur within pSRC27-H and not pSRC27-I. Hence from now on, pSRC27-I will not be considered. Only Variants 1, 3-1, 3-2 and 4 retained the Tn*6029* copy of *strA/B*. These variants are the same ones that retained *sul2*, which is adjacent to *strA/B* in Tn*6029*.

8.3.3 Duplicate copies of $\mathbf{bla}_{TEM}$

pSRC27-H also harboured 2 copies of the bla_{TEM} gene within its MARR (Fig 6.10). One was in Tn*6029* between IS*26*(1) and (2), here named bla_{TEM}-1, and one was adjacent to IS*26*(4), named bla_{TEM}-2. The presence of these individual copies was detected, using PCR linking bla_{TEM}-1 to IS*26*(1) (primers RH605/IS*26*-F; Table A3.1 in Appendix A3) and bla_{TEM}-2 to ISCfr1 (primers RH606/RH630; Table A3.1 in Appendix A3; Table 8.3).

Table 8.3 Distribution of bla_{TEM} in variants.

	Linkage PCR (location)		**bla_{TEM} Southern bands (kb)** [1]		
	bla_{TEM}-1/ IS*26*(1) (Tn*6029*)	bla_{TEM}-2/ ISCfr1 (*aacC2*-containing region)	Predicted 3.2	Predicted 8.5	Additional
Parent	+	+	+	+	
V1[2]	+	+	+	+	-
V2-1	-	+	-	+	-
V2-2	-	+	-	+	-
V3-1[3]	+	+	+	+	-
V3-2	+	+	+	+	-
V4	+	-	+	-	(~10) [4]
V5-1[3]	-	-	-	-	-
V5-2[3]	-	-	-	-	-
V6	-	-	-	-	~10
V7	-	-	-	-	-
V8	-	-	-	-	-
V9[5]	-	-	-	-	-

[1]Band sizes estimated from DIG-labelled markers, averaged over 2 membranes
[2]Data only obtained from replicate i)
[3]Data only obtained from replicate ii)
[4]Band faint compared to the 3.2 kb band from the same sample
[5]Negative control

The presence of the two copies of *bla*$_{TEM}$ was also explored by determining whether the entire IS*26*-bounded fragments containing *bla*$_{TEM}$ were present in the variants, using Southern hybridisation. SwaI digested plasmid DNA extracted from representative isolates of each variant type, was hybridised with a probe internal to *bla*$_{TEM,}$ as previously described in section 6.2. When the entire MARR was present, two bands of expected sizes, 3.2 kb for *bla*$_{TEM}$-1 (in Tn*6029*) and 8.6 kb for *bla*$_{TEM}$-2 (*aacC2*-containing region), were seen (Fig 8.2). The Southern hybridisation was repeated (Fig 8.2 i) and ii)) and the band sizes recorded (Table 8.3). The Southern hybridisation bands correlated exactly with the PCR results obtained (Table 8.3), except for additional bands of ~10 kb observed in Variants 4 and 6, which indicate that part of the gene is present.

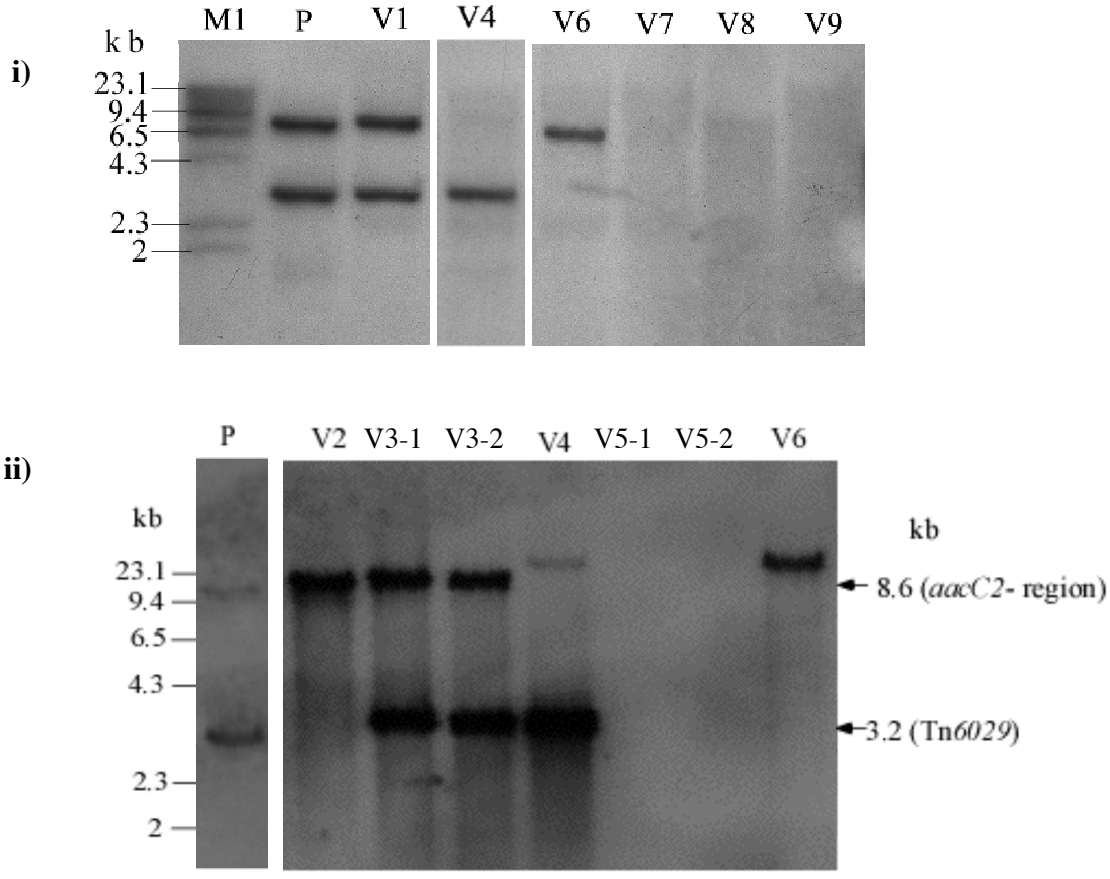


Fig 8.2 Southern hybridisation of variants probed with *bla*$_{TEM}$. Southern hybridisation of the MARR of SwaI digested plasmid DNA from the SRC27 transconjugant parent and its resistance variants. 2 membranes i) and ii) were hybridised with *bla*$_{TEM}$ DIG-labelled probes. Lanes are marked V1 – V9, representing Variants 1 -9. Lane P is the parent SRC27 transconjugant, in ii) the parent is separate from the rest of the samples as it needed a longer exposure and there were blank lanes in between. Gaps in the gels represent blank lanes. Lane M1 is the DIG-labelled molecular weight marker II. In ii) the molecular weight marker is shown as lines and numbers which were determined by measuring the migration distance and comparing them to distances of DIG-labelled molecular weight markers II on a overexposed blot of the same membrane.

Variants 1, 3-1, 3-2 and 4 retained *bla*$_{TEM}$-1 (the Tn*6029* copy), and produced a Southern hybridisation fragment that was the correct size for Tn*6029*. These variants also retained *sul2* and *strA/B*, indicating that they contain Tn*6029*. The rest of the variants have lost this region. Only Variants 1, 2-1, 2-2, 3-1 and 3-2 contained *bla*$_{TEM}$-2 and the same variants retained *aacC2*. These variants also displayed a Southern hybridisation band that was the correct size for the *aacC2*-containing fragment, indicating that this region is intact. A revised summary of the resistance genes in the IncHI1 plasmids present in each of the variants is in Table 8.4.

Table 8.4 Resistance genes in IncHI1 plasmid of variants

Type[1]	Resistant to	*catA1*	*bla*$_{TEM}$ -1[3]	*sul2*	*strA/* IS*26*	*aphA1*	*bla*$_{TEM}$ -2[3]	*aacC2*	*dfrA12-aadA2*	*tetA* (B)
Parent	ApCmGmKmNmSmSpSuTcTp	+	+	+	+	+	+	+	+	+
1	ApCmGmSmSpSuTcTp	+	+	+	+	-	+	+	+	+
2-1	ApCmGmSmSpTcTp	+	-	-	-	-	+	+	+	+
2-2	ApCmGmSmSpTcTp	+	-	-	-	-	+	+	+	+
3-1	ApCmGmKmNmSmSu	+	+	+	+	+	+	+	-	-
3-2	ApCmGmKmNmSmSu	+	+	+	+	+	+	+	-	-
4	ApCmKmNmSmSuTc	+	+	+	+	+	-[2]	-	-	+
5-1	SmSpTcTp	-	-	-	-	-	-	-	+	+
5-2	SmSpTcTp	-	-	-	-	-	-	-	+	+
6	CmSmTc	+	-[2]	-	-	-	-[2]	-	-	+
7	CmSm	+	-	-	-	-	-	-	-	-
8	SmTc	-	-	-	-	-	-	-	-	+

[1]Variant 9 also included as a negative control
[2]Potential partial *bla*$_{TEM}$
[3]Highlighting represents differences compared to Table 8.2

Variant 4 retained the band for *bla*$_{TEM}$ in Tn*6029* as well as an additional faint band that runs slightly higher than the *aacC2*-containing fragment, of ~10 kb which is only on the more highly exposed gel (Fig 8.2ii). However, PCR for *bla*$_{TEM}$-2 and other genes on the *aacC2*-containing fragment, such as *aacC2,* were negative (Table 8.4). A similar sized additional band was also observed for Variant 6, even though neither *bla*$_{TEM}$-1 nor -2 were present (Table 8.2). This indicates that a fragment of *bla*$_{TEM}$ must remain in these variants.

8.4 Mapping resistance variants

In Table 8.4 the resistance genes of pSRC27-H are shown in the order they occur in the MARR and it can be seen that genes lost together are adjacent. To further characterise the

resistance regions of the 11 variant types, linkage PCR and Southern hybridisation targeting the MARR of pSRC27-H, was performed.

8.4.1 Loss of IS26-containing fragments

The Southern hybridisation procedure used was the same as for the mapping of the pSRC27-H MARR (summarised in Fig 6.14). Plasmid DNA from each variant and E294/pSRC27-H/pSRC27-I was digested with SwaI and then hybridised with a probe internal to IS*26* (Fig 8.3). When the entire MARR was present, three predicted bands of sizes 1.9 kb, 3.2 kb and 8.6 kb for Tn*4352*, Tn*6029* and the *aacC2*-containing piece (Fig 6.11), respectively were observed (Lane P in Fig 8.3 i) and ii)).

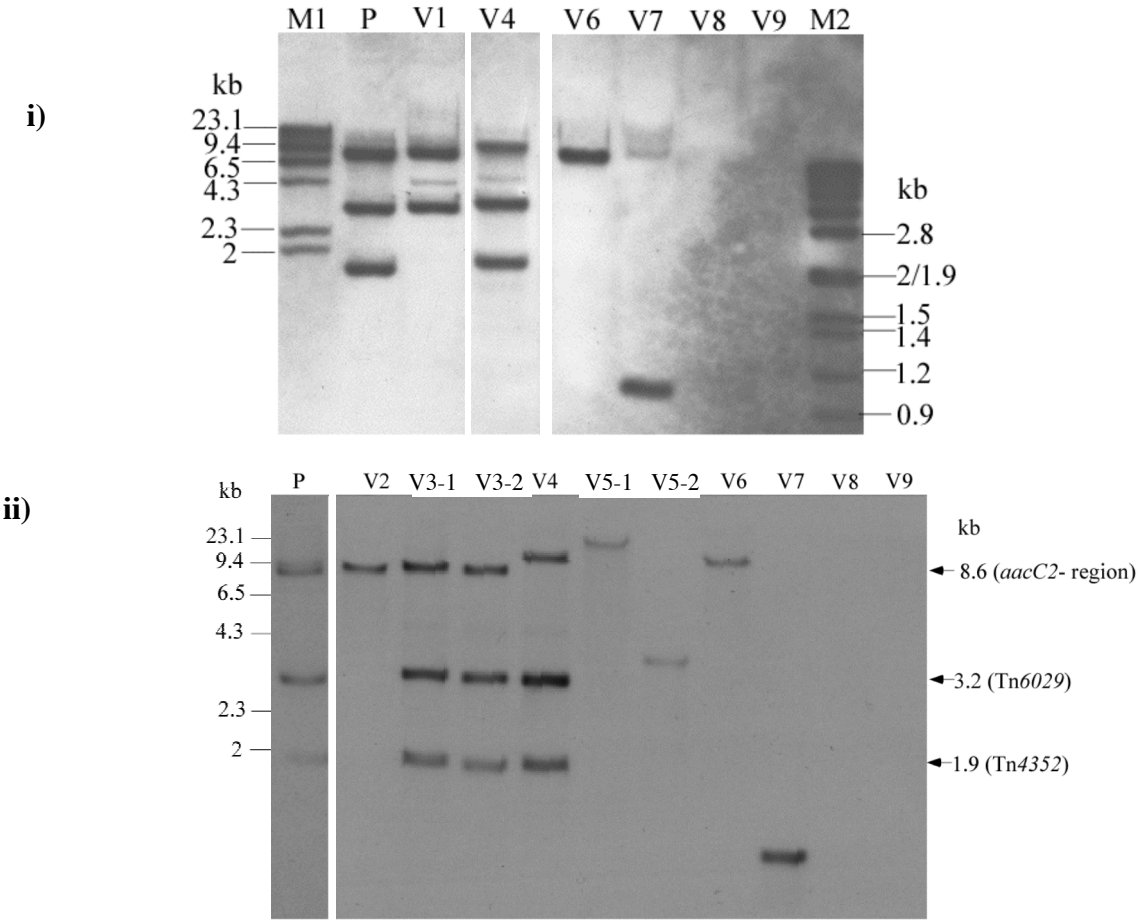


Fig 8.3 Southern hybridisation of variants probed with IS*26*. Southern hybridisation of the MARR of SwaI digested plasmid DNA from the SRC27 transconjugant parent and its resistance variants. 2 membranes i) and ii) were the same as those in Fig 8.2, and were stripped, then re-hybridised with an IS*26* DIG-labelled probe. Lanes are marked V1 – V9, representing Variants 1 -9. Lane P is the parent SRC27 transconjugant, in ii) the parent is separate from the rest of the samples as it needed a longer exposure and there were blank lanes in between. Gaps in the gels represent blank lanes in between. Lane M or M1 is the DIG-labelled molecular weight marker II and M2 is marker VII. In ii) the molecular weight marker is shown as lines and numbers which were determined by measuring the migration distance and comparing them to distances of DIG-labelled molecular weight markers II on a overexposed blot of the same membrane.

Only Variants 3-1 and 3-2 yielded 3 bands of the same size as the parent, indicating that they have retained the IS*26*-bounded region of the MARR and this was reflected by the presence of all the resistance genes in this region (bla_{TEM}-1, *sul2, strA/B,* bla_{TEM}-2, *aacC2*; Table 8.2). All other variants displayed banding patterns different to the parent strain. The bands observed are compared to the bla_{TEM} bands in Table 8.5.

Table 8.5 Bands from Southern hybridisation on variants

	IS*26* probe bands (kb) [1]				***bla*$_{TEM}$ probe bands (kb)** [1]		
	Predicted			Additional	Predicted		Additional
	1.9	3.2	8.5		3.2	8.5	
Parent	+	+	+		+	+	
V1[2]	-	+	+	-	+	+	-
V2	-	-	+	-	-	+	-
V3-1[3]	+	+	+	-	+	+	-
V3-2	+	+	+	-	+	+	-
V4	+	+	-	~10 kb	+	-	(~10) [4]
V5-1[3]	-	-	-	~14	-	-	-
V5-2[3]	-	-	-	~4	-	-	-
V6	-	-	-	~10	-	-	~10
V7	-	-	-	~1.1	-	-	-
V8	-	-	-	-	-	-	-
V9[5]	-	-	-	-	-	-	-

[1]Band sizes estimated from DIG-labelled markers, averaged over 2 membranes
[2]Data only obtained from replicate i)
[3]Data only obtained from replicate ii)
[4]Band faint compared to the 3.2 kb band from the same sample
[5]Negative control

Variants 5-1 to 8 have lost all of the resistance genes within the IS*26*-bounded segment and consistent with this, had none of the Southern hybridisation bands derived from this region. However, Variants 5-1 to 7 each have an additional band, indicating these variants have either 1 or 2 IS*26* copies remaining. The additional bands observed for Variants 5-1 and 5-2 were of different sizes, indicating that despite their identical resistance phenotype, different deletion events have occurred. Variant 8 had no bands and therefore no IS*26*. Thus, Variant 8 had a large deletion, that included the whole IS*26*-bounded region (18.6 kb), consistent with only the *tetA*(B) resistance gene remaining. Variants 1 and 2 have lost only the resistance genes in Tn*4352* and Tn*6029*/Tn*4352*, respectively, and the Southern hybridisation results are consistent with loss of these regions. Variant 4 produced the 2 bands derived from Tn*6029*

and Tn*4352* as well as an additional band (of ~10 kb) for both the IS*26*- and the bla_{TEM}-probed membranes, indicating that at least copies 1 – 4 of IS*26* remain.

A set of 7 linkage PCRs within the IS*26*-bounded region were also performed on each variant type to determine which copy/s of IS*26* remain and to determine the extent of each deletion within the IS*26*-bounded region (Fig 8.4; Table 8.6). For each variant, the results of these linkage PCRs paralleled the Southern hybridisation results obtained. Linkage PCRs were also used to determine if the deletions had extended beyond the IS*26*-bounded region (Fig 8.4). PCRs 9 and 10 in Table 8.6 linked IS*26* to the to the left ($tnpA_{21}$/IS*26*(1)) and the right (IS*26*(5)/*intI1*) of the IS*26*-bounded region, respectively.

Table 8.6. Linkage PCRs of IS*26*-bounded region

Target	**Tn*21*/ *IS26***	**Tn*6029***			**Tn*4352***		***aacC2*-containing region**		**integron/ IS*26***
PCRs	9.	1.	2.	3.	4.	5.	6.	7.	10.
Region	$tnpA_{21}$/ IS26(1)	IS*26*(1)/ bla_{TEM}	*sul2*/ *strAB*	*strB*/ IS*26*(3)	*strB*/ *aphA1*	*aphA1*/ IS*26*(4)	bla_{TEM}/ ISCfr1	*aacC2*/ *IS26*(5)	IS*26*(5)/ *intI1*
Parent	+	+	+	+	+	+	+	+	+
1	+	+	+	+	-	-	+	+	+
2-1	+	-	-	-	-	-	+	+	+
2-2	+	-	-	-	-	-	+	+	+
3-1	+	+	+	+	+	+	+	+	-
3-2	+	+	+	+	+	+	+	+	-
4	+	+	+	+	+	+	-	-	-
5-1	-	-	-	-	-	-	-	-	+
5-2	-	-	-	-	-	-	-	-	+
6	+	-	-	-	-	-	-	-	-
7	+	-	-	-	-	-	-	-	-
8	-	-	-	-	-	-	-	-	-

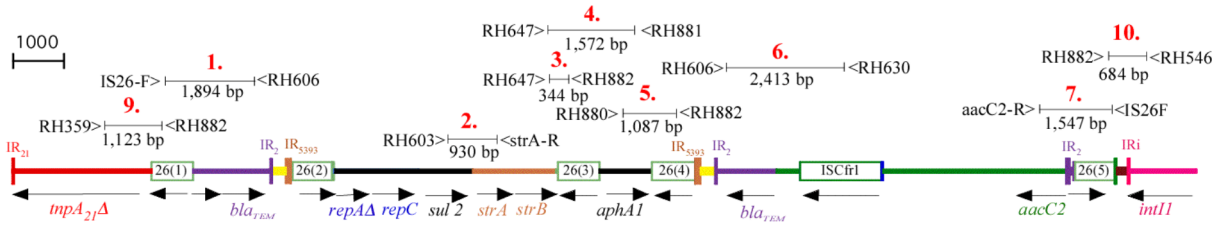


Fig 8.4 Linkage PCRs in the IS*26*-bounded region. Arrows show the extent and direction of genes. Boxes represent insertion sequences with the name inside and the copy number bracketed. Different colours represent different regions: purple for Tn*2*, brown for Tn*5393*, green for region that matches plasmid pCTX-M3 only. The regions flanking the MARR are shown as dashed lines. PCRs spanning this region are numbered 1 -7 in red and are shown above with the primers are vertical lines and the product sizes below. Primer sequences are listed in Tables A3.1, A3.3, A3.5 and A3.10 in Appendix 3.

These results confirmed that Variant 1 lacked only Tn*4352* and Variants 2-1 and 2-2 lacked Tn*6029* and Tn*4352*. Variants 3-1 and 3-2 displayed products for all PCRs, consistent with retaining the entire region. However, Variants 3-1 and 3-2 had lost the *intI1*/IS*26*(5) junction, suggesting that the deletion may have arisen to the left of IS*26*(5). Variant 4 did not produce amplicons for PCRs associated with the *aacC2*-containing region, but retained IS*26*(4) (Table 8.6). The remaining variants did not product any amplicons within the IS*26*-containing region, however, all but Variant 8 did retain one of the bounding IS*26*. Variants 6 and 7 retain the $tnpA_{21}$/IS*26*(1) junction, indicating IS*26*(1) remained. Similarly, Variants 5-1 and 5-2 retained the IS*26*(5)/*intI1* junction thus they contained IS*26*(5). These variants had also lost all resistance genes to the left of IS*26*(5), indicating that the deletion may start from this IS*26*.

8.4.2 Loss of the rest of the MARR

Further linkage PCRs were performed, to determine the extent of the deletions and to map the remainder of the MARRs in each variant type. They were dispersed throughout the remaining regions of the pSRC27-H MARR (PCRs 11-13 Fig 8.5; Table 8.7) and flanking sequence in the VRT1 backbone (PCRs 8 and 14 Fig 8.5; Table 8.7). As expected, the rest of the MARR was present in Variants 1, 2-1 and 2-2. For Variants 3-1 and 3-2, which had given identical results up to this point, these PCRs revealed that they were different to each other, as the right-hand MARR boundary was negative in Variant 3-1, but not 3-2 (PCR 14 in Table 8.7). PCR 14 was also negative in Variant 7. This indicated that the deletion in Variants 3-1 and 7 had reached beyond IS*10*, at the right-hand MARR boundary. Similarly, Variants 5-1 and 5-2 were missing the left-hand boundary of the MARR and thus the deletion had extended into the backbone (PCR 8 in Table 8.7). In Variant 8, which only retained *tetA*(B) both boundaries of the MARR remained (PCRs 8 and 14 in Table 8.7), indicating that the region missing is between IS*1*a and IS*1*b.

Table 8.7 Segments of the MARR present in variants

Target	LH MARR boundary	*tnp*$_{21}$	Tn*6029*[1]	Tn*4352*[1]	*aacC2*-containing region[1]	Integron		Tn*10*		RH MARR boundary
PCRs	8.	9.	1., 2., 3.	4., 5.	6., 7.	10.	11.	12.	13.	14.
Regions	VRT1 backbone /IS*1*	tnpA$_{21}$/ IS26(1)				IS*26*(5)/ *intI1*	*aadA2*/ IS*1*	*aadA2*/ *tetA*(B)	IS*1*/ *tetA*(B)	IS*10*/ VRT1 backbone
Parent	+	+	+	+	+	+	+	+	+	+
1	+	+	+	-	+	+	+	+	+	+
2-1	+	+	-	-	+	+	+	+	+	+
2-2	+	+	-	-	+	+	+	+	+	+
3-1	+	+	+	+	+	-	-	-	-	-
3-2	+	+	+	+	+	-	-	-	-	+
4	+	+	+	+	-	-	-	-	+	+
5-1	-	-	-	-	-	+	+	+	+	+
5-2	-	-	-	-	-	+	+	+	+	+
6	+	+	-	-	-	-	-	-	+	+
7	+	+	-	-	-	-	-	-	-	-
8	+	-	-	-	-	-	-	-	+	+

[1]individual PCRs in Table 8.6

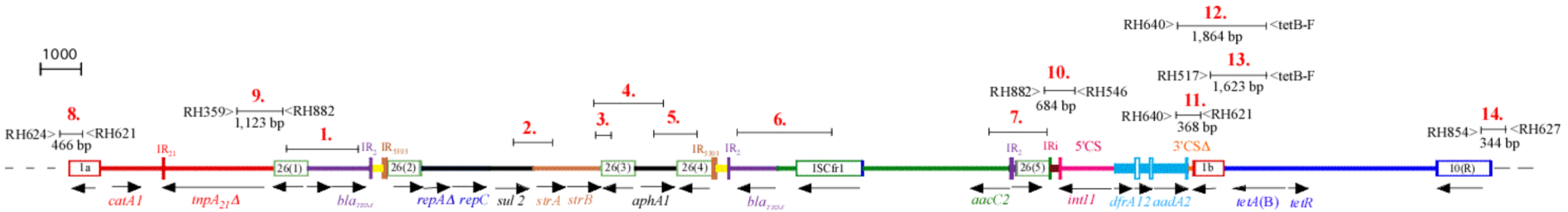


Fig 8.5 Mapping variants with linkage PCR. Arrows show the extent and direction of genes. Boxes represent insertion sequences with the name inside and the copy number bracketed. Different colours represent different regions: red for Tn*2670*, purple for Tn*2*, brown for Tn*5393*, green for the region that matches plasmid pCTX-M3 only, pink for 5'-CS, cyan for gene cassettes, orange for 3'-CS, blue for Tn*10*. The regions flanking the MARR are shown as dashed lines. PCRs spanning this region are numbered 1 -14 in red and are shown above with the positions of primers as vertical lines and the product sizes below. Primer sequences are in Tables A3.1, A3.4 and A3.10 in Appendix 3.

8.5 The deletion boundaries of variants

The extent of deletion for each variant type was deduced from the combined results of the internal resistance gene PCRs (Table 8.2), the linkage PCRs (Table 8.7) and the Southern hybridisation data (Table 8.5). The MARR structures were proposed and for all variants, except for Variant 5-1 (see 8.5.5), the exact position of the deletion was elucidated using additional PCR across the deleted region and sequencing of these PCR product/s. The primers used for linking across deletion sites within the MARRs were the same as in Fig 8.5. If the deletion had extended beyond one boundary of the MARR, which was the case for Variants 5-1 and 5-2 on the left and Variants 3-1 and 7 on the right, (when PCRs 8 or 14 were negative), further linkage PCRs into the flanking sequence were performed. Primers within the last resistance gene that remained next to the deletion, were used in conjunction with primers in the flanking sequence. Primers were made sequentially further out, until a PCR product was obtained. A number of primers were designed in the MARR flanking regions (Fig 8.6). The furthest of them was 7,912 bp to the left of IS*1*a (RH863) and 5,835 bp to the right of IS*10* (RH1271). These 2 primers defined the region surveyed.

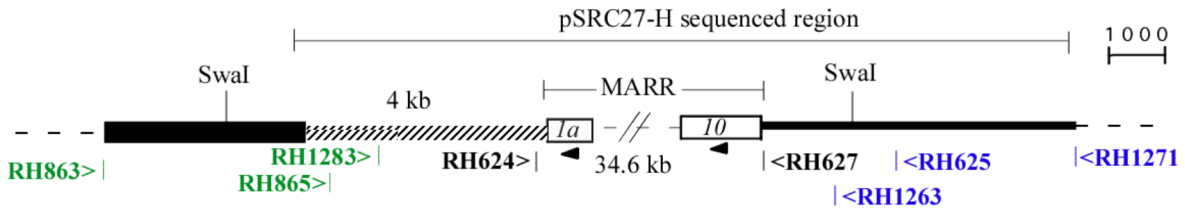


Fig 8.6 Primers surrounding the pSRC27-H MARR. Positions of primers are shown by a vertical line with their name and an arrow indicating their direction beside. Primers to the left of the MARR are shown in green, primers to the right are in blue and primers used to link to the MARR are in black. Primer sequences are listed in Tables A3.8, A3.9 and A3.10 in Appendix 3. The SwaI sites are shown. The MARR region is not shown and the boxes represent insertion sequences with their name inside. The hashed line represents the Region 5 segment unique to pHCM1-type plasmids. The region sequenced in pSRC27-H (GenBank acc no. HQ840942) is bracketed above. Figure is drawn to scale.

8.5.1 Variant 1- ApCmGmSmSpSuTcTp^{R}

Variant 1 had only lost the *aphA1* (Km^{R}) gene and the 1.9 kb IS*26*-probed Southern band (Fig 8.7). It retained all the remaining features of the parent pSRC27-H MARR. To demonstrate that only one of the IS*26* assigned to Tn*4352* remains between *strB* and *bla*$_{TEM}$-2, PCR linking *strB* to *bla*$_{TEM}$ was performed (Fig 8.7). A 2.3 kb amplicon was observed for Variant 1, compared to a 4.1 kb band for the pSRC27-H parent. The sequence of this product revealed *strB* was separated from *bla*$_{TEM}$–2 by IS*26*. Hence a 1.9 kb deletion had occurred consisting of IS*26* and the central piece of Tn*4352*. The IS*26* removed could be either copy 3 or 4, or a combination of both, the latter being the most likely, if the deletion was a product of homolgous recombination.

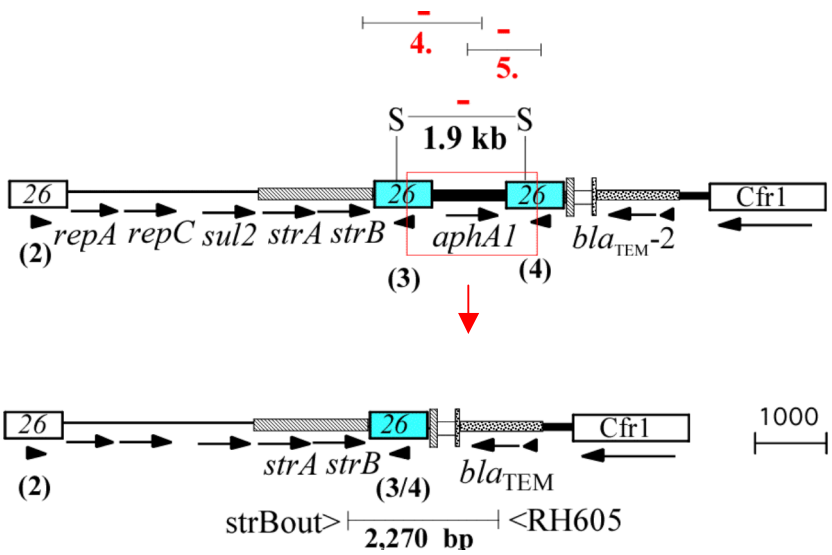


Fig 8.7 Deletion in Variant 1. Only the region between IS*26*(2) and ISCfr1 is shown. The open boxes are insertion sequences with the IS name inside and their copy number bracketed below. The horizontal arrows show the direction and extent of genes. The SwaI sites are shown as "S" and a vertical line and the fragment size between 2 SwaI sites is shown below a horizontal line joining the sites. The genes missing in Variant 1 are boxed in red and the IS effecting the deletion are coloured blue. Relevant mapping PCRs are shown above as their numbers (allocated in Fig 8.5) and a linkage PCR is below, with the positions of primers shown as vertical lines and the product size below a horizontal line joining the primers. Primer sequences are listed in Table A3.1 in Appendix 3. Southern hybridisation and PCR results are in red where + indicates a positive result and – a negative result.

8.5.2 Variant 2- ApCmGmSpTcTp^R

Variant 2 was missing the *bla*$_{TEM}$-1, *sul2*, *strA/B* and *aphA1* resistance genes and all the features associated with Tn*4352* and Tn*6029*, as shown in Fig 8.8. A PCR was used to link *tnpA*$_{21}$ and *bla*$_{TEM}$–2 (Fig 8.8) and an amplicon of 3,175 bp was observed for Variant 2-1 and 2-2 (Fig 8.8). The sequence confirmed that a 9.8 kb segment in between IS*26* (1) and (4), as well as one of the IS*26*, had been lost. As in Variant 1, the excision was precise and left behind exactly 1 IS*26*. This deletion was identical in Variants 2-1 and 2-2 indicating that this deletion arose independently in both cultures.

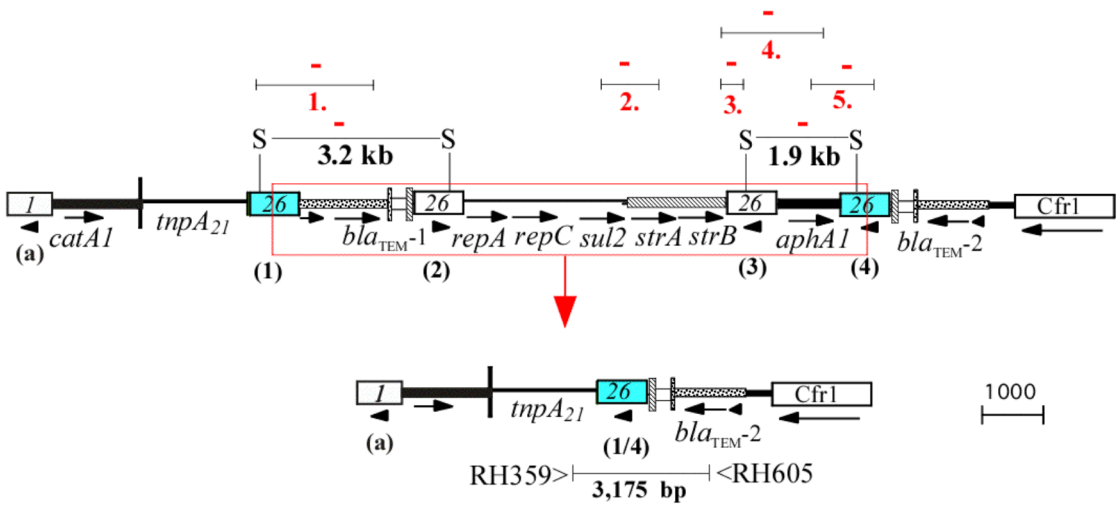


Fig 8.8 Deletion in Variant 2. Only the region between IS*26*(1) and ISCfr1 is shown. Primer sequences are listed in Tables A3.1 and A3.10 in Appendix 3. Features are as in Fig 8.7.

8.5.3 Variants 3-1 and 3-2 – ApCmGmKmNmSmSu^R

Both Variants 3-1 and 3-2 had lost the region immediately to the right of IS*26*(5) including the gene cassettes and *tetA*(B). They were negative for PCRs 10-13, as shown in Fig 8.9. However, they both retained all Southern bands and all PCRs to the left of IS*26*(5) were positive. Variant 3-2 retained the IS*10*/VRT1 backbone junction (PCR 14), but Variant 3-1 did not (Fig 8.9).

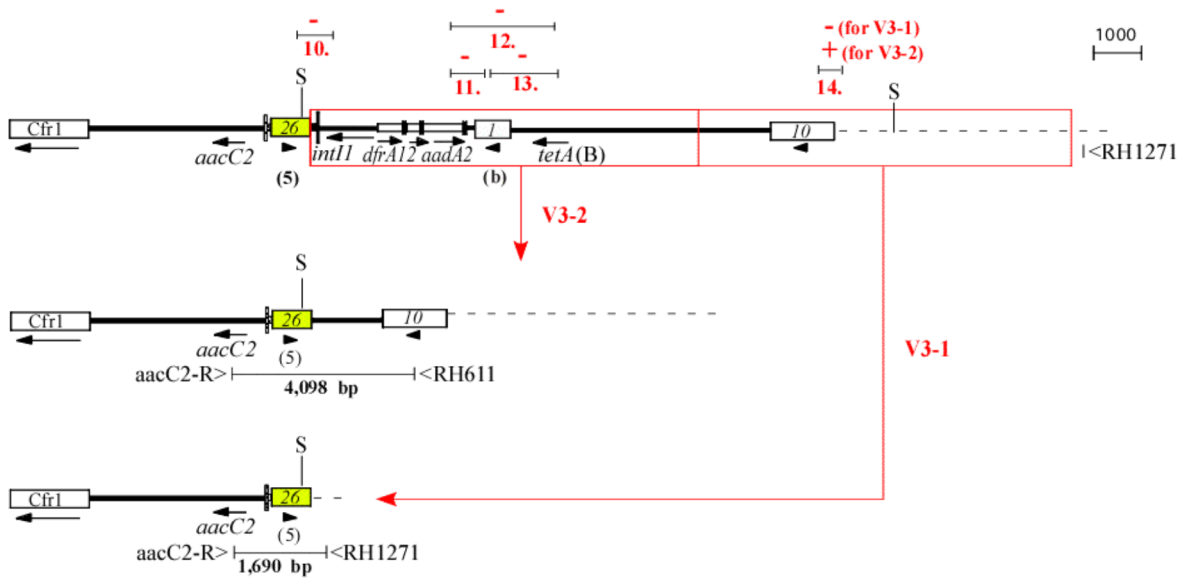


Fig 8.9 Deletion in Variants 3-1 and 3-2. Only the region to the right of ISCfr1 is shown. The arrows indicate the different deletion sites for Variants 3-2 and 3-1. Primer sequences are listed in Tables A3.1, A3.4 and A3.8 in Appendix 3. All other features are as in Fig 8.7.

As the RH MARR boundary PCR (PCR 14) was positive for Variant 3-2, (Fig 8.9), the deletion ended within the MARR. A PCR to link *aacC2* and IS*10* (Fig 8.9) yielded a 4.1 kb amplicon. Sequencing showed that on the right of IS*26*(5), only 1,313 bp of the centre of Tn*10* and the IS*10* remained (bases 1-2,642 as in the Tn*10* standard AF162223; Fig 8.9). On the left the IS*26*(5) to *aacC2* junction was the same sequence as for the equivalent region in the parent strain. Thus, Variant 3-2 was missing 8.3 kb to the right of IS*26*(5), which included the entire integron and part of Tn*10*, but the complete IS26(5) was retained (see Fig A5.2 in Appendix 5).

In Variant 3-1, the deletion had extended beyond the IS*10*. PCR with the primer RH1271, located 5,835 bp to the right of IS*10* (Fig 8.6), and primer aacC2-R in *aacC2* (Fig 8.9) produced a 1.7 kb amplicon. Sequence adjacent to IS*26*(5) obtained from this PCR product (Fig A5.2 in Appendix 5) showed that exactly 5,691 bp to the right of the MARR, which included Variable Region 5 (Table 7.2), had been deleted. Sequence to the left of IS*26*(5)

revealed the *aacC2*/IS*26* boundary was identical to that of the parent. Therefore, a 16.6 kb deletion had occurred in Variant 3-1, to the right of IS*26*(5).

8.5.4 Variant 4 - ApCmKmNmSmSuTc^{R}

Variant 4 had lost the *aacC2* gene and the *aadA2* and *dfrA12* gene cassettes. Linkage PCRs 6, 7, 10-12 were negative, which spanned the region from *bla*$_{TEM}$–2 up to the integron (Fig 8.10). Although linkage PCR of *bla*$_{TEM}$-2 to ISCfr1 were negative, an additional faint ~10 kb Southern hybridisation band on the *bla*$_{TEM}$ probed membrane was observed in addition to the Tn*6029* band, indicating part of *bla*$_{TEM}$-2 may have been present (Table 8.5). Thus it was possible IS*1*b had truncated *bla*$_{TEM}$-2. A PCR with primers in *aphA1* and *tetA*(B) produced a 3.3 kb amplicon (Fig 8.10). When the product was sequenced, IS*1*b was found to be located 39 bp to the right of the stop codon of *bla*$_{TEM}$-2 (see Fig A5.2 in Appendix 5).

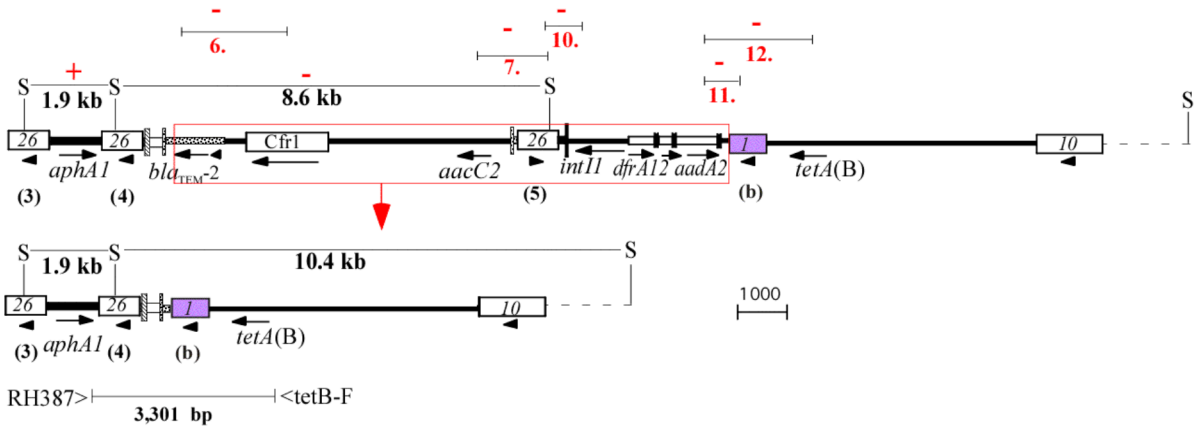


Fig 8.10 Deletion in Variant 4. Only the region to the right of IS26(3) is shown. Primer sequences are listed in Tables A3.6 and A3.4 in Appendix 3. All features are as in Fig 8.7.

The ~10 kb band was faint (relative to the 3.2 kb band) in the *bla*$_{TEM}$-probed membrane (Fig 8.2) because there was only 36 bp left of sequence homologous to the 861 bp probe (probe was made using primers TEM-SFW/TEM-SRV; see section 2.6.2). This 10.4 kb fragment was produced by a SwaI site 1,749 bp beyond IS*10*, in conjunction with the SwaI site in IS*26*(3). A deletion of 11.1 kb was shown to have occurred in Variant 4, to the left of IS*1*(b).

8.5.5 Variants 5-1 and 5-2 –SpTcTp^R

Variant 5-2, had only retained the gene cassettes and *tetA*(B). PCRs 1-9 were negative, which included the LH MARR boundary up to *aacC2* (Fig 8.11). Therefore, because PCR 8 was negative, the deletion reached beyond the MARR left-hand boundary and into the flanking sequence. In Variant 5-2, in the IS*26* probed Southern membrane, only a band of ~4 kb was observed (Table 8.5). A SwaI site in IS*26*(5) and a SwaI site located 5,714 bp to the left of IS*1*a in pSRC27-H would yield the ~4 kb fragment observed in Variant 5-2 (Fig 8.10) if ~2 kb of flanking sequence was missing.

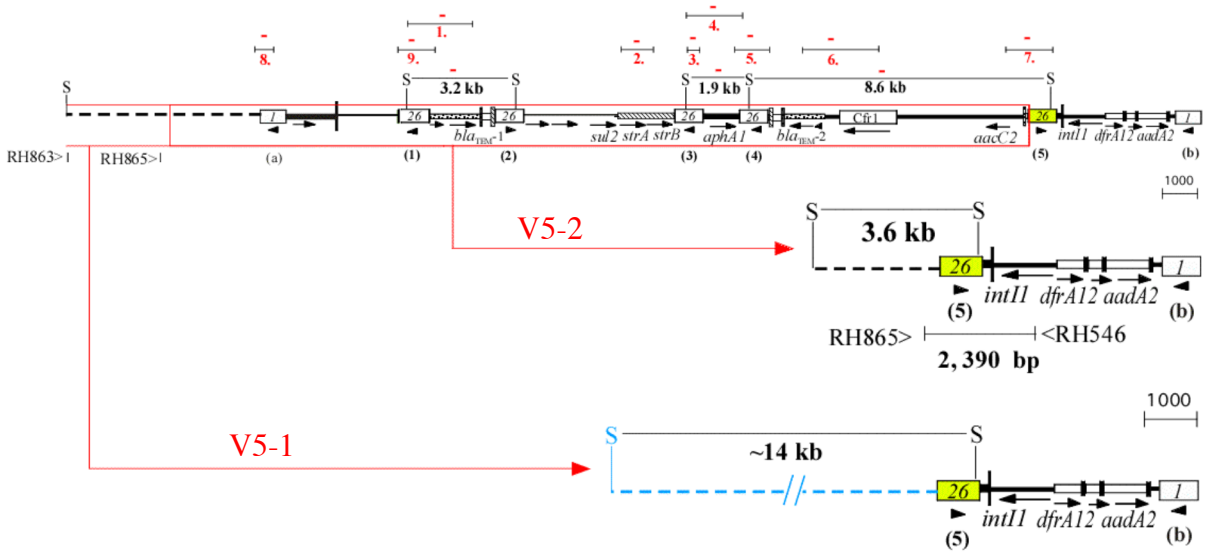


Fig 8.11 Deletion in Variants 5-1 and 5-2. Only the region to the left of IS*1*b is shown. The red arrows indicate the different amounts of the MARR missing in Variants 5-2 and 5-1. The open ended box in Variant 5-1 indicates the deletion site was not found and is beyond the area surveyed. Unknown sequence is shown as a blue dashed line and the 2 slashed lines represent unknown sequence of unknown exact length and is not to scale. Primer sequences are listed in Tables A3.5 and A3.10 in Appendix 3. The unidentified SwaI also in blue. All other features are as in Fig 8.7.

When the primer RH865, located 3,915 bp to the left of IS*1*a in the flanking sequence, (Fig 8.6) was used with a RH546 in *intI1* (Fig 8.11), a 2.4 kb amplicon was obtained. The sequence of this product using a primer facing out of IS*26* revealed the IS*26*(5) was adjacent to the sequence 2,483 bp to the left of IS*1*a (see Fig A5.2 in Appendix 5). Thus, a 25.5 kb deletion to the left of IS*26*(5) had occurred in Variant 5-2.

Variant 5-1 yielded only a ~14 kb SwaI band, a result of a fragment from the IS*26*(5) SwaI site and another outside the MARR (Fig 8.11). This SwaI site was not the first one to the left of IS*1*a, as shown in Fig 8.12, as it is only 7.9 kb to the left of IS*1a* and would not produce a fragment big enough. The PCR with primers RH546 in *intI1* and RH865 in the LH flanking region, failed to yield an amplicon for Variant 5-1. In addition, when primer RH863, positioned 7,912 bp to the left of IS*1* (Fig 8.6) was used with the primer in *intI1*, again, no amplicon was obtained. These results indicated that the deletion had reached beyond 7.9 kb to the left of IS*1*a. Therefore Variant 5-1 had a deletion greater than 33.8 kb, reaching from the left of IS*26*(5) to a point past the primer RH863 and beyond the area surveyed.

8.5.6 Variant 6- $CmTc^R$

Variant 6 was missing all resistance genes, except for *catA1* and *tetA*(B). The linkage PCRs 3 -12 were negative in Variant 6, confirming that Tn*6029*, the *aacC2*-containing piece and the integron were missing (Fig 8.12). Like Variant 4, although the PCR internal to bla_{TEM} was negative, a single additional ~10 kb Southern hybridisation band was observed on the bla_{TEM}-probed membrane, indicating that a bla_{TEM} remnant remained. Therefore, it was likely IS*1*b had truncated bla_{TEM}-1. PCR linking $tnpA_{21}$ to *tetA*(B) (Fig 8.12) produced a 4.2 kb amplicon. Sequence with a primer facing out of IS*1* established that IS*1*b had removed most of bla_{TEM}-1, leaving only 200 bp (see Fig A5.2 in Appendix 5).

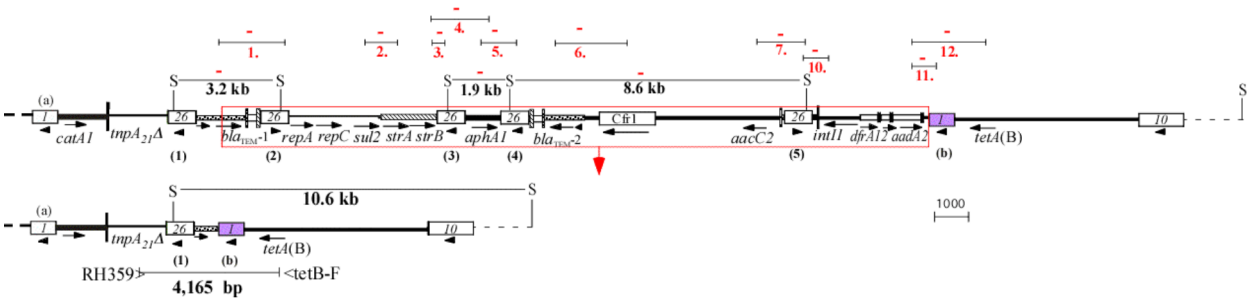


Fig 8.12 Deletion in Variant 6. Primer sequences are listed in Tables A3.10 and A3.4 in Appendix 3. All features are as in Fig 8.7.

The additional 10.6 kb Southern hybridisation band observed with both the IS*26* and bla_{TEM} probe was a result of the SwaI site in IS*26*(1) and one to the right of the MARR. Of the 861 bp bla_{TEM} probe, only 200 bp of bla_{TEM}-2 remained in Variant 6, but this was enough for the probe to hybridise effectively, allowing the bands to be observed. Therefore, a deletion of 20.8 kb to the left of IS*1*(b) had occurred.

8.5.7 Variant 7- Cm^R

Variant 7 was Cm^R, retained only *catA1* and was positive only for PCRs across the LH MARR boundary and the $tnpA_{21}$/IS*26*(1) junction (PCR 8 and 9 in Table 8.7). Thus the deletion had extended beyond the MARR and into the RHS flanking sequence (Fig 8.13). The only Southern hybridisation band observed was a ~1 kb additional band on the IS*26* probed Southern membrane (Table 8.5). This ~1 kb fragment could have arisen from a SwaI site in the RHS flanking sequence (that is normally 1,750 bp to the right of IS*10*) and one in IS*26*(1) (Fig 8.13). With this option, a fragment of this size indicates ~1 kb directly adjacent to IS*10* was missing.

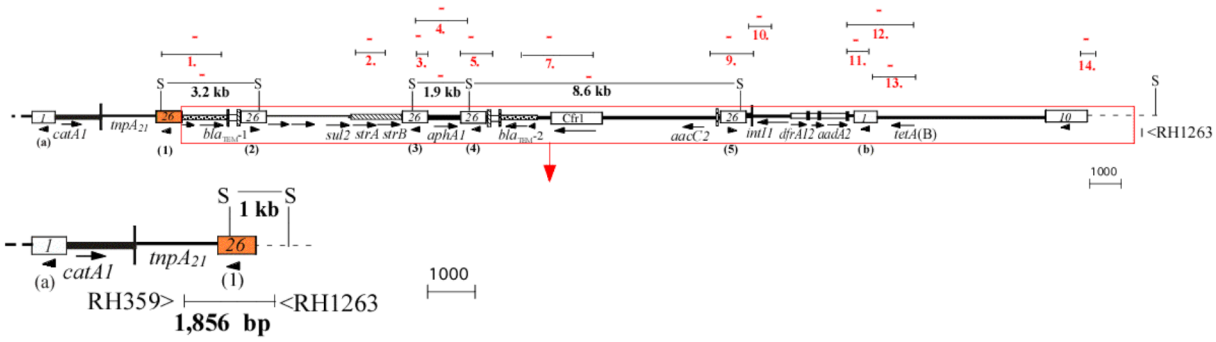


Fig 8.13 Deletion in Variant 7. Only the region to the right of IS*1*a is shown. Primer sequences are listed in Table A3.10 in Appendix 3. All features are as in Fig 8.7.

The primer RH1263 close to the SwaI site and located 1,416 bp to the right of IS*10* (Fig 8.6*)* was used in a PCR with the primer RH359 in $tnpA_{21}$, the last part of the MARR known to

remain, which gave a 1.8 kb product (Fig 8.13). Sequence out of IS*26*(1) revealed the deletion extended 1,556 bp into the RH flanking sequence, beyond the Tn*10* (see Fig A5.2 in Appendix 5). The deletion in Variant 7 spanned 30.3 kb, including most of the MARR, and at to the right end of IS*26*(1).

8.5.8 Variant 8- Tc^R

Variant 8 only retained *tetA*(B), was negative for PCRs 2-12 and no Southern hybridisation bands were observed, indicating most of the MARR and all IS*26* were missing. However, PCRs 1, 13 and 14 were positive, therefore both MARR boundaries were intact as well as the IS*1*a/*tetA*(B) junction. Thus, the deletion was contained within the MARR and at least 1 IS*1* remained (see Fig 8.14). The region to the left of the MARR and *tetA*(B) were linked using PCR. Primers RH624 (Fig 8.6; used in PCR 1) located 315 bp to the left of IS*1*a and one in *tetA*(B), tetB-F, generated a 3 kb amplicon. This amplicon was too large to contain just 1 copy of IS*1*, which would be the product of homologous recombination between the 2 IS*1*. Confirming this, when this product was sequenced with RH621, a primer in IS*1*, double peaks were observed in the sequence file, which was unreadable (Fig 8.15). However, when it was sequenced using the primer RH624 and the primer tetB-F (See Fig 8.14), both the Tn*10*/IS*1* and LH flanking region/IS*1a* boundaries were shown to be identical to the parent.

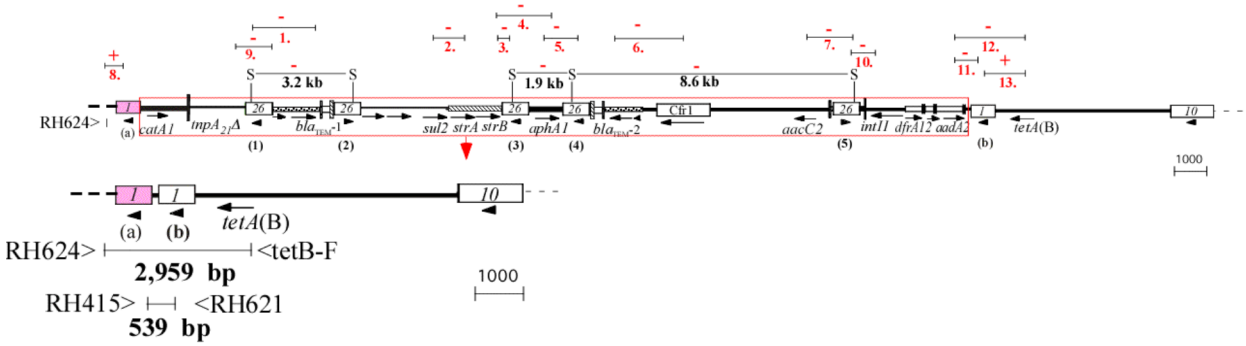


Fig 8.14 Deletion in Variant 8. Primer sequences are listed in Tables A3.4 and A3.10 in Appendix 3. All features are as in Fig 8.7.

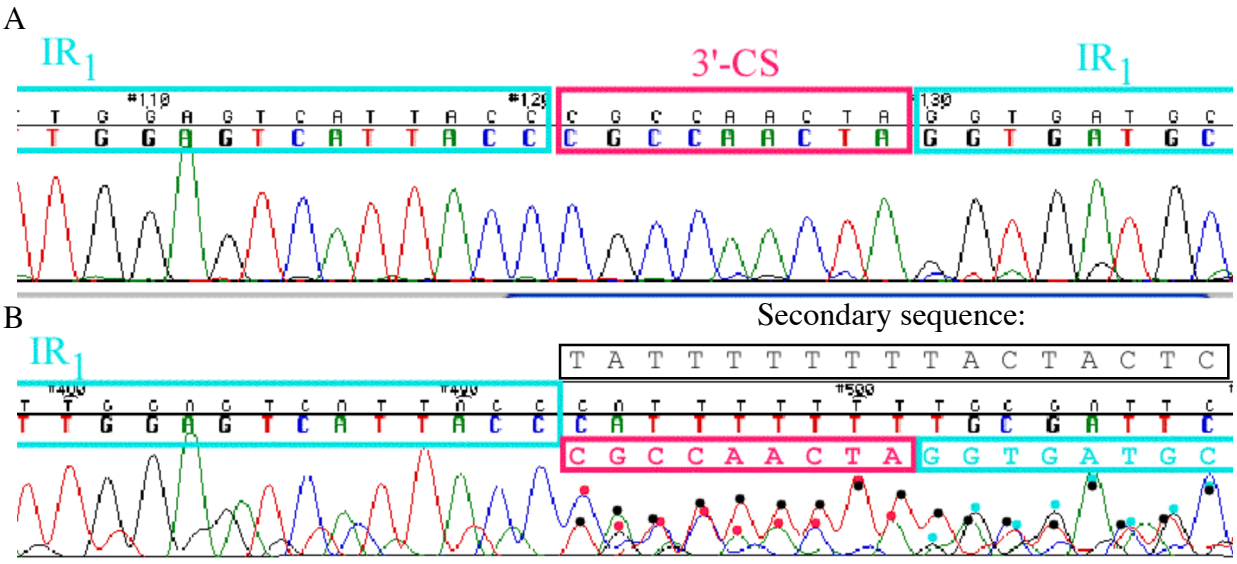


Fig 8.15 Sequence chromatograph of the deletion site in Variant 8. Each peak represents a nucleotide base, where red peaks are T, black are G, blue are C and green are A. **A** sequence from a product from the PCR on Variant 8 linking the 2 IS*1* (primers RH621/RH415) **B** sequence from a PCR linking IS*1* to the backbone (using primers RH624/tetB-F); both A and B were sequenced with the same primer (RH621) out of IS*1*. In A, the IR of the 2 IS*1* are boxed in blue and part of the 3'-CS are boxed in pink. B has 2 sets of peaks, one from part of the integron sequence (marked by pink dots above the peaks) and one from the sequence flanking the MARR (marked as black dots).

From the size of the linkage PCR product, it was deduced that 2 copies of IS*1* had to be present. which would cause double peaks (one peak from each copy of IS*1*; marked black in Fig 8.15B). To determine what was between the 2 IS*1*, a PCR was performed linking IS*1a* and IS*1*b, which yielded a 539 bp amplicon (using primers RH415 and RH621; see Fig 8.14). The sequence of this amplicon was clear (Fig 8.15A) and revealed that there was 9 bp of the 3'-CS of the integron remaining in between 2 copies of IS*1* (Fig 8.16). In pSRC27-H this part of the integron is found adjacent to IS*1*b (see Fig A5.2 in Appendix 5). Therefore, 27.5 kb of sequence was missing to the right of IS*1*a, including most of the MARR.

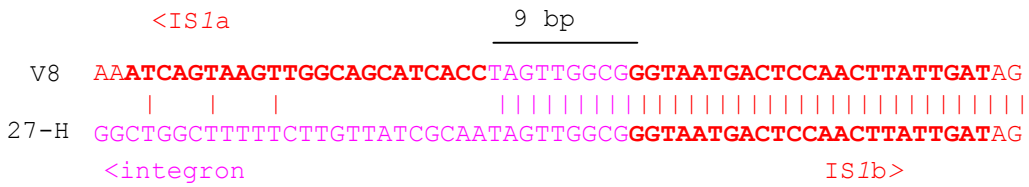


Fig 8.16 Sequence alignment from Variant 8. The variant 8 sequence (on top) is aligned with the sequence to the left of IS*1*b in pSRC27-H (on bottom). The vertical lines represent matches between the sequences. Red lettering is the IS*1* and the bold face type shows IR_1. The purple lettering is sequence from the integron and marked with a horizontal line is the 9 bp left in Variant 8.

8.6 Summary of variants and how they arose

Eleven different variant types were identified among the 9 resistance phenotypes recovered by testing 1,000 isolates, from 2 cultures of E294/pSRC27-H/pSRC27-I grown without antibiotic selection. Variant 9 had lost the IncHI1 plasmid. For the remainder of the variants, a copy of IS*26* or IS*1* was directly adjacent to the deletion site on one or both sides. A linear representation of each of the variant MARRs, compared to the parent is given in Fig 8.17. Sequence alignments of deletion site boundaries are provided in Fig A5.2 in Appendix 5. The mechanism by which each deletion arose was deduced and tabulated in Table 8.8. These were consistent with possible deletion mechanisms proposed in the introduction (Fig 8.1).

Variants 1 and 2 had lost Tn*4352* and Tn*6029*/Tn*4352*, respectively. This may have occurred via homologous recombination between 2 directly orientated IS*26*. Only Variant 2 was recovered from both cultures. Variants 3-1, 3-2, 5-1, 5-2 and 7 were a result of IS*26*-mediated adjacent deletion. Of these, 4/5 were caused by IS*26*(5). Variants 4, 6 and 8 were a result of IS*1*-mediated adjacent deletion. Variants 3-1 and 3-2 as well as 5-1 and 5-2, arose independently but generated identical phenotypes. The deletion sizes ranged from 1.9 kb (Variant 1) to >33.8 kb (Variant 5-1).

Table 8.8 Variants of E294/pSRC27-H/pSRC27-I.

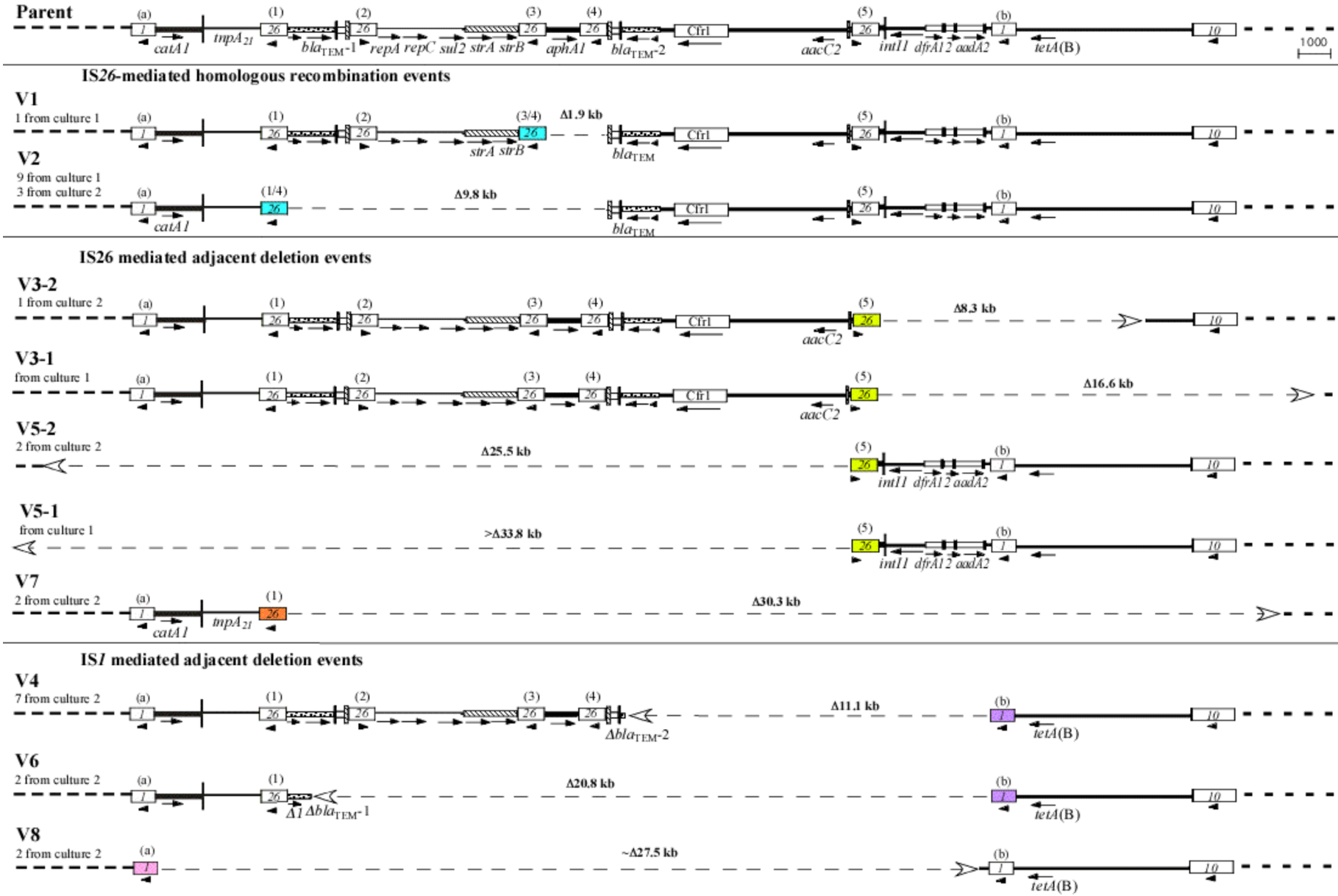
Variant[1]	Resistance gene/s lost	Deletion size (kb)	Causative event/description	Deletion direction[2]
1	*aphA1*	1.9	IS*26*(1)/(2) homologous recombination	-
2	*aphA1, bla(1), strA/B, sul2*	9.8	IS*26*(1)/(4) homologous recombination	-
3-1	*aadA2, dfrA12, tetA*(B)	16.6	IS*26* (5) adjacent deletion	RH
3-2	*aadA2, dfrA12, tetA*(B)	8.3	IS*26* (5) adjacent deletion	RH
4	*aacC2, aadA2, part bla(2), dfrA12*	11.1	IS*1*b adjacent deletion	LH
5-1	*aacC2, aphA1, bla(1+2), catA1, sul2, strA/B*	>33.8	IS*26* (5) adjacent deletion	LH
5-2	*aacC2, aphA1, bla(1+2), catA1, sul2, strA/B*	25.5	IS*26* (5) adjacent deletion	LH
6	*aadA2, part bla(1), bla(2), dfrA12, sul2, strA/B*	20.8	IS*1*b adjacent deletion	LH
7	*aacC2, aadA2, aphA1, bla(1+2), dfrA12, sul2, strA/B, tetA*(B)	30.3	IS*26* (1) adjacent deletion	RH
8	*aacC2, aadA2, aphA1, bla(1+2), catA1, dfrA12, sul2, strA/B*	27.5	IS*1*a adjacent deletion	RH
9	*aacC2, aadA2, aphA1, bla(1+2), catA1, dfrA12, sul2, strA/B, tetA*(B)	pSRC27-H	Plasmid loss	-

[1]All also contain pSRC27-I
[2]direction defined in Fig 8.18

8.7 Discussion

The purpose of this study was to gain an overview of the deletion events that occur in a MARR, in the absence of antibiotic selection. However, it is important to note the limitations of this experimental design. Firstly, in this pilot study, not all types of deletions that could occur would be detected. For example, variants that had lost Tn*6029* alone, would not have been recovered, as *sul2* is the only resistance marker in it not duplicated (there are 2 copies of bla_{TEM} and *strA/B*) and sulphonamide resistance was not screened for. This can be overcome by screening for it, or by removing pSRC27-I in future studies. Secondly, it does not detect deletions that do not remove a resistance gene. Thirdly, inversion of DNA would not be detected by this method. Lastly, the sample size of isolates screened is not large enough to be considered significant, thus observations, such as the bias toward certain IS being more active, are preliminary and need further investigation.

Fig 8.17 Deletions identified in variants. Thin, horizontal dashed lines represent the deletion, and the arrow head, in the case of adjacent deletion represents the direction of deletion. The IS presumed to have caused each deletion is coloured. Other features are as in Fig 8.5


Parent
IS26-mediated homologous recombination events
V1
1 from culture 1
V2
9 from culture 1
3 from culture 2
IS26 mediated adjacent deletion events
V3-2
1 from culture 2
V3-1
from culture 1
V5-2
2 from culture 2
V5-1
from culture 1
V7
2 from culture 2
IS1 mediated adjacent deletion events
V4
7 from culture 2
V6
2 from culture 2
V8
2 from culture 2
Δ1.9 kb
Δ9.8 kb
Δ8.3 kb
Δ16.6 kb
Δ25.5 kb
>Δ33.8 kb
Δ30.3 kb
Δ11.1 kb
Δ20.8 kb
~Δ27.5 kb
1000

Loss of resistance genes was attributed to either deletions involving IS within pSRC27-H, or cells that had lost the entire IncHI1 plasmid. A variant that had lost pSRC27-H, arose in both cultures and represented <1% of the cells tested (8/1000). However when the *E. coli* strain that did not contain the IncI1 plasmid, pSRC27-I, was grown at 37^{O}C, the majority of the cells had lost the IncHI1 plasmid. Previous studies conducted on the stability of IncHI1 plasmids [176] mention that IncHI1 plasmids in *E. coli* or *S.* Typhimurium hosts are not maintained when cultures are grown at temperatures between 30-44^{O}C. However, the IncHI1 plasmids did remain stable when growth was at 26-30 ^{O}C [176]. The variants were generated prior to knowledge of the incompatibility groups of the plasmids in SRC27, hence the experiments were performed at 37^{O}C. It may be advantageous to repeat the loss experiments in this chapter, with cells containing pSRC27-H only, at 30^{O}C. These observations relating to plasmid loss, raised an intriguing question of how does the IncI1 plasmid stabilise the IncHI1 plasmid. This warrants further investigation. Furthermore, how are IncHI1 plasmids maintained in the wild at 37^{O}C, when they are associated with *S.* Typhi, which is a human-specific pathogen?

The remaining 10 resistance variants were produced in 3 main ways. Variants 1 and 2 had lost the DNA in between 2 directly orientated IS*26* and one copy of IS*26*, and these could arise by a result of homologous recombination. Five variants arose via IS*26*-mediated adjacent deletion and 3 via IS*1*-mediated adjacent deletion. No IS*10*-mediated deletions were detected. This was perhaps expected because the only IS*10* MARR of pSRC27-H was spatially separated from the majority of the resistance genes and only large IS*10*-mediated adjacent deletions would be detected using this procedure.

Although the sample size of these experiments is not great enough to be statistically significant, the activity of individual IS did appear to be influenced by its position within the

MARR. In the case of IS*26*, out of the 5 copies of IS*26* present in the region, only 2 were active in causing IS*26*-mediated adjacent deletions. Four variants (3-1, 3-2, 5-1 and 5-2) were generated by IS*26*(5) and only one (Variant 7) by IS*26*(1). IS*26*(2), (3) or (4) did not produce adjacent deletions. However IS*26* (1), (3) and (4) were involved in the loss of Tn*4352* and Tn*6029*/Tn*4352* (Variants 1 and 2). The influence of the position of IS*26* activity within a resistance region is an interesting notions that warrant further experimentation, but were outside the scope of this study. Future experiments would include investigating the influence of genetic context on IS*26* activity, by determining IS*26* transposition frequencies (or cointegrate formation) in a number of different structures.

These experiments demonstrate that resistance regions can accrue deletions over a relatively short time frame, in the absence of antibiotics. The information gained here also supports evidence gathered in Chapter 6, which nominates the VRT1 IncHI1 plasmids pHCM1, pO111_1 and pSRC27-H as being closely related, despite multiple differences within their MARRs, because clearly, these IS-mediated changes readily occur. The deletion events observed here are similar to those proposed to have occurred between the VRT1 IncHI1 plasmid MARRs and this evolution must have occurred *in situ* because the MARR of VRT1 IncHI1 plasmids are in the same position.

CHAPTER NINE

General discussion

9.1 Overview

The work presented in this thesis examined the composition and evolution of 2 clinically important, large plasmid groups: IncHI1 plasmids, which are associated with multiple antibiotic resistance in *S*. Typhi [163], the human-specific pathogen that is responsible for typhoid, and IncHI2 plasmids, which have been identified as a major force responsible for spreading antibiotic resistance in Enterobacteriaceae [142]. The representative IncHI plasmids investigated in this study were recovered from *S. enterica* Australian animal isolates, circa the year 2000. These plasmids were found to share features with sequenced IncHI plasmids, allowing them to be allocated to distinct evolutionary lineages. Plasmids within these lineages were recovered from a range of dates, sources and origins, suggesting that they move freely between animals and humans, that plasmids lineages are globally disseminated and that they have persisted for at least 4 decades (since the earliest plasmids analysed were isolated in the 1960s). The structures of antibiotic resistance regions in these plasmids were elucidated and forces that drive the evolution of these regions were identified. Furthermore, in the case of IncHI2 plasmids, segments containing the antibiotic and heavy metal resistance regions analysed here, appeared to be in ancestral configurations compared to related, sequenced plasmids, which had undergone deletion and insertion events, often bringing in additional resistance genes. Also, the ancestral form of the IncHI2 plasmids pSRC125 and pSRC26 allowed novel heavy metal transposons to be identified, as they did not contain them.

9.2 Using the position of MGEs to track evolution

The work presented in this thesis highlights the power of using the positions of MGEs for establishing relationships between resistance regions and plasmids. The Tn*1696*-derived resistance regions in pSRC125 and pSRC26 and the MARR of Tn*2670* and Tn*10* origin in pSRC27-H, were found in the same positions as related regions in other IncHI2 and IncHI1 plasmids, respectively. This finding served a dual purpose: to allocate plasmids that share the

Tn position to an evolutionary plasmid lineage, and to examine the evolution of resistance regions that occurs *in situ*.

Comparing related resistance regions within the same fixed position is ideal for identifying forces that shape their evolution. This has previously been described in SGI1 [225], and in AbaR-type genomic islands [42], both of which have variant forms located in the same position as the original. This variation could thus be attributed to IS-mediated and homologous recombination events occurring within these elements *in situ*.

In this study, the positions of integrons were examined. In4-type integrons were found in the same position within the Tn*1696*-like Tns in pSRC125 and pSRC26 and in plasmids recovered from the *S*. Infantis isolates. In turn, these transposons were in an identical location in the IncHI2 plasmid backbone. However, the integrons carried different gene cassettes, namely *dfrA5* and *dfrB6-aadA1*. The fact that these integrons had swapped cassettes *in situ* emphasises the likely role of homologous recombination in integron cassette exchange. HR between the conserved segments of class 1 integrons (see Fig 1.16B) has previously been proposed as being the major mechanism of cassette exchange [72, 111], but is often overlooked. Thus, the widely used "cassette PCR" that detects cassette arrays in class 1 integrons, allows the types of cassettes present to be identified, but does not provide information on the integron structure or the context of the integrons. PCR sets designed here that detect class 1 integrons from a range of mercury II resistance transposons (Figs 3.2 and 3.3), provide examples of how simple tools can be used to explore the context of resistance gene cassettes, thereby more accurately examining evolutionary relationships.

Using the positions of resistance regions and MGEs within plasmids, a great deal of information regarding the evolution of IncHI1 and IncHI2 plasmids, was gained. The configuration of the Tn*1696*-like resistance regions in pSRC125 and pSRC26 were ancestral

to those found in R478 and pK29. The Tn*1696*-derived resistance regions were in the same position for R478, pK29, pSRC26, pSRC125 and 7 of the plasmids originating from *S.* Infantis. However, R478 and pK29 were missing not only parts of the resistance region, but segments of what was determined to be IncHI2 backbone. These truncation events were due to the acquisition of IS-flanked regions containing additional resistance genes and multiple copies of IS*26* (Fig 3.19). This reiterates the notion that IS*26* is a major force driving change in resistance regions. The evolutionary path constructed here, based on the composition and position of this shared resistance region, shows that these plasmids belong to a single lineage of IncHI2a plasmids (Fig 3.19). R478, pSRC26 and pSRC125 were found to also harbour Tn*10* in the same position, confirming their close relationship and also providing an order that some of these regions were likely to have been acquired (Fig 3.20). Other sequenced plasmids pEC-IMP and pEC-IMPQ, were found to be of a separate lineage. The IncHI2 reference plasmid, TP116, may represent a 3rd IncHI2a lineage that contains a Tn*2670*-like resistance region instead of a Tn*1696*-derived one. This is the first study that has been able to clearly define lineages of IncHI2a plasmids, which demonstrates the highly discriminatory power of tracking MGE positions, compared to other techniques such as SNP typing or PMLST.

Previously, IncHI1 plasmids had been grouped in 3 families using PMLST and microarray data [163], represented by R27, pAKU_1 and pHCM1. However, using SNP typing and the presence of some variable regions in the backbone, studies concluded that R27 and pAKU_1 were more closely related to each other than to pHCM1 [163, 188]. Here, IncHI1 plasmids were grouped into 2 evolutionary lineages based on the position of resistance regions as well as using the 6 variations occurring between their backbones. Some of these variable regions displayed characteristics of MGEs (Table 7.1). The variable region type (VRT) classification scheme used here defined VRT1 and VRT2 plasmids as representing 2 separate lineages. This method was found to be effective, as it supported the data presented from these previous studies that grouped IncHI1 plasmids using SNP-typing [188], which would find that a VRT1

plasmid has 3 fold more SNPs across the shared backbone than VRT2 plasmids, and a PMLST scheme [163], which would place all VRT1 plasmids in "Group 1" IncHI1 plasmids (analysis undertaken in this study). Furthermore, VRT2 plasmids would be placed into "Group 2" plasmids if one less SNP occurred within the 6 loci in R27, compared to pAKU_1.

9.3 New transposons

Three novel antibiotic resistance transposons and two heavy metal resistance transposons were identified over the course of this work. In addition, two further possible cryptic transposons were also found.

Three antibiotic resistance transposons, that had not been analysed previously, were defined and named Tn*6029*, Tn*5393*e and Tn*6023*. The IS*26*-bounded transposon Tn*6029*, which contains *strA/B, sul2* and bla_{TEM} and was derived from the plasmid RSF1010 and Tn*2* (Fig 3.13), was found within Tn*6026* in pSRC26 (Fig 3.5B) and Tn*6027* in pSRC27 (Fig 6.16). Tn*5393*e, a variant of Tn*5393* that did not confer streptomycin resistance due to the insertion of Tn*6023* (Fig 3.17), was located 4 kb downstream of Tn*6025* in pSRC125. Tn*6023* consists of *aphA1b* flanked by 2 IS*26* in opposite orientation, one of which provides the –35 region for *aphA1b* promoter (Fig 3.15). The IS*26*-bounded resistance transposons described here illustrate that IS*26* is capable of forming novel Tns from a range of different sources. A direct duplication of 8 bp was identified surrounding the overlapping transposons Tn*6029* and Tn*4352* in pSRC26, indicating that the whole entity had transposed. This is the first report of direct repeats surrounding overlapping Tns sharing an IS, that have been seen separately in other locations. Although this finding suggests that entire Tn*6029*/Tn*4352* can move in addition to its components, these Tns may be formed as a result of homologous recombination. To determine whether IS*26* forms these overlapping transposons and by what mechanism, transposition assays will be needed, carried out using host strains that are not capable of homologous recombination (a *recA*- background).

Previously, it had been recognised that R478 contains genes that encode proteins related to those of Tn*7*, however, the boundaries of a transposon or the short, repeated sequences near the ends of the element, that in Tn*7* facilitate transposase binding [51], could not be found [58]. In the current study, Tn*6024*, a 32.4 kb silver and copper resistance transposon that shares features with Tn*7*, was characterised in R478. It contains 23 bp IRs, 2 sets of 13 bp repeats at its ends, starts with "TG" and ends in "CA" and is surrounded by a 5 bp direct repeat. The *tnsA-E* genes in Tn*6024* were predicted to encode proteins that were 26-34% identical to the equivalent ones encoded by Tn*7*. As suggested in the introduction, a third class of transposons, which includes Tn*7* and Tn*402* and named here Class III Tns, can be defined by their shared features (see section 1.3.1.3). In addition, more than 100 entries in the GenBank protein database shared >25% aa identity to TnsB from Tn*7*, indicating that the Class III transposon family has many members. In the future, analysing these large numbers of related Tns, by bioinformatically comparing them to each other and their surrounding sequence to equivalent entries that do no contain Tns (the procedure used in section 4.4), should identify the boundaries of further Class III transposons and determine whether the characteristics proposed in section 1.3.1.3 are common across a large sample of Class III Tns.

An arsenic resistance transposon had been previously speculated to exist in R478 on the basis of high GC content, yet the boundaries of the element could not be identified, genes for transposition function were not found, and transposition experiments failed to mobilise an arsenic resistance transposon [202]. Here, a 7.2 kb arsenic resistance transposon, that is not in pSRC26 or pSRC125, was identified in R478. It contains 25 bp IRs, a putative transposase gene, and is surrounded by a 5 bp direct repeat. However, it was not found in other locations and thus it was not named.

Although both these heavy metal Tns display strong evidence of being active transposons, mobilisation and transposition experiments should be performed in the future, to confirm that they are indeed mobile and to characterise the mechanisms of their movement.

The possible cryptic Tn found VRT1 IncHI1 plasmids, referred to as "Region 4" displayed evidence of being a transposon, capable of movement. It is bounded by 25 bp IR and is found in multiple locations, including in IncHI2 plasmids, where in each case it is surrounded by an 8 bp direct repeat (see Table 7.1). However, none of its ORFs encode a protein that could be identified as a transposase. One ORF encodes an UvrD DNA helicase and this raises a number of questions. Is this a defective transposon, where the transposition proteins were provided *in trans*? Does this DNA helicase have the capacity to promote DNA movement? These questions could be answered by further characterising this potential MGE, for example, by cloning the region and performing transposition assays, in different hosts.

9.4 What are large plasmids made of?

The nature of the large plasmids examined in this work was determined to be mosaic and a significant proportion of the IncHI2 plasmid structure was found to be composed of acquired MGEs. A previous estimation of the conserved IncHI2 backbone by means of an *in silico* comparative analysis of the sequenced IncHI2 plasmids, found that only 174 kb (out of 275 kb for R478) was present in all of them [160]. Here, the large heavy metal resistance transposons Tn*6024* and the arsenic resistance Tn identified in R478, collectively span ~40 kb. These elements alone constitute 15% of the 275 kb R478 backbone. The MARR encompasses a further 7%. The IncHI2 plasmids examined in this study, using mapping PCRs, mainly varied within only a region of ~100 kb - approximately one third of the R478 backbone (Variable Region 1 in Fig 4.19). This is a clear example of how horizontally acquired regions have the potential to increase plasmid size dramatically. Similarly, the variation observed in the 11 $bla_{CTX\text{-}M}$-carrying IncHI2 plasmids previously tested also lies

within this region [161]. This segment contains the antibiotic and heavy metal resistance regions, which have been shown to be on MGEs. The possibility that the tellurite resistance module, as well as other elements known to be associated with virulence and mutagenesis, including the *mucA/B* mutagenesis region, a bacteriophage (k, T1, T5, T7 and ϕ80) resistance region and a pore-forming colicin (types A, B, and K) resistance region ([174, 201]; marked in red in Fig 9.1), may also lie on MGEs needs to be examined. Because of the discovery of the arsenic, and the copper and silver resistance transposons, it is reasonable to suppose that the tellurite resistance region may also be on a transposon. The IncHI2 reference plasmid TP116 was found not to confer tellurite resistance and therefore, by sequencing TP116, perhaps the boundaries of a Ter^R transposon will be revealed. The only other MGE outside this region is Tn*10* (Fig 9.1). If this ~100 kb region is, as hypothesised here, made up primarily of integrated MGEs, then the actual R478 backbone would be reduced to ~160 kb (when Tn*10* is also removed). Furthermore, when the variable region is removed, the "essential" plasmid regions, like the 2 transfer regions, would be brought into closer proximity (Fig 9.1), which may represent the structure a IncHI2 precursor plasmid that does not confer antibiotic or heavy metal resistance.

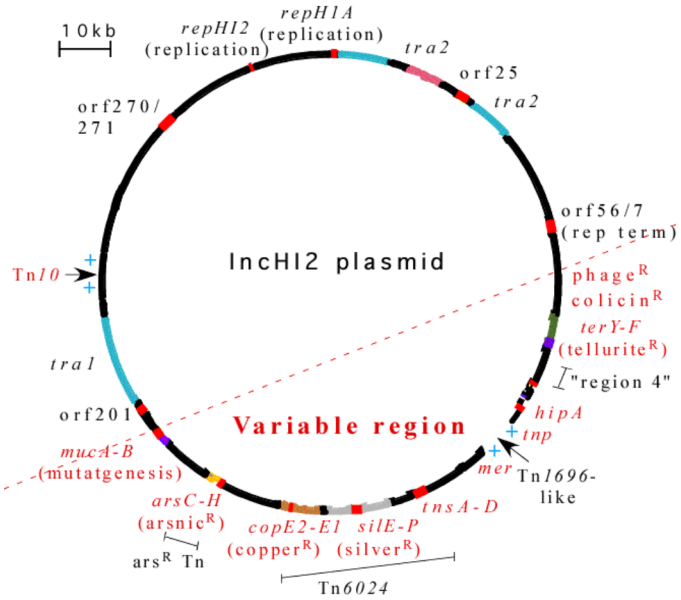


Fig 9.1. Variable region in the IncHI2 backbone. The red dotted line represents the cut off of where IncHI2 plasmids start to vary. MGEs and factors associated with virulence and resistance that are usually found on MGEs are highlighted in red.

A region previously used as a distinguishing feature of IncHI1 plasmids has been shown to be mobile. In this study, the plasmid pMAK-1 was analysed and found to be missing the IS*1*-bounded IncF-derived replicon (F-rep), which is responsible for the one way incompatibility observed between IncHI1 and IncF plasmids [141, 173]. Although this region obviously has the potential to be mobile, due to the IS*1* elements bounding it, this is the first report of an IncHI1 plasmid without this region. F-rep has historically been considered a fundamental component of an IncHI1 plasmid and was even used as a marker to differentiate IncHI1 and IncHI2 plasmids [173]. It is possible that older IncHI1 and IncHI2 plasmids could have been mis-classified when using this phenotypic marker to differentiate them.

9.5 Forces driving the evolution of resistance regions

MARR regions have high plasticity and can change quite rapidly in the absence of selection. In light of the results obtained in Chapter 8, where numerous IS*1* and IS*26*-mediated deletions that removed antibiotic resistance genes from the mosaic 34.6 kb MARR in pSRC27-H, were observed after only 14 days of culturing without antibiotics, it is possible to understand how readily antibiotic resistance genes can be lost. These experimental results reflected the bioinformatic-based observations made when comparing the similarly derived MARRs in VRT1 IncHI1 plasmids. Although these plasmids display high sequence identity and contained the same variable regions in their backbone (Table 7.3), they differed primarily within their antibiotic resistance regions and some of these variations were attributed to the activity of insertion sequences, such as the insertion of IS*26*, IS*4321* and IS*5075* found only in the MARR in pHCM1 (Figs 6.17 and 6.19). Similarly, the IncHI2 plasmids pK29 and R478 were observed to have undergone IS-mediated changes extending from within the resistance regions into the backbone (Fig 3.19). Homologous recombination also plays a role in the evolution of resistance region, for example in the swapping of integron gene cassette arrays (see section 9.2).

9.6 Plasmid stability and co-resident plasmid interaction

Since the event of the simple and rapid identification of plasmid replicon types using PBRT, a better picture of what plasmids are present, the types of co-residency that occur and how these plasmids interact in wild type cells, is being developed. For example, an intriguing observation that warrants further investigation was that an IncI1 plasmid appeared to stabilise an IncHI1 plasmid. A previous study had shown that IncHI1 plasmids were unstable in both *E. coli* and *S.* Typhimurium cultures when grown at 37°C [176] and this raised the question of how IncHI1 plasmids are maintained in the wild, especially as they are found in *S.* Typhi, a human-specific pathogen adapted for life at 37°C. Perhaps co-resident plasmids stabilise them. During the current work, when an *E. coli* transconjugant containing pSRC27-H alone was cultured at 37°C, over time the majority of cells lost the plasmid, but when pSRC27-I was also present, only a small proportion of cells lost pSRC27-H (Chapter 8). Thus pSRC27-I appeared to stabilise the pSRC27-H. This finding was not investigated further. However, it would be of interest to identify the specific regions of pSRC27-I that are involved in this apparent stabilisation, and to examine whether other IncI1 plasmids have this same stabilising effect on pSRC27-H or on other IncHI1 plasmids.

In Chapter 5, RFLP analysis revealed that additional fragments, with a combined size of ~80 kb, were observed in all of the *S.* Infantis transconjugants. One explanation proposed for this unexpected additional DNA was that other plasmid/s had co-transferred or formed cointegrates with the IncHI2 plasmids, although not of a detectable Inc group. This was supported by the fact that a large portion of one plasmid, pSRC70, including both rep regions, has been lost, meaning that its replication function is most likely provided by another plasmid. Co-resident plasmids, particularly cryptic ones can be hard to detect and may occur more commonly than is currently realised. Further work to detect these potential plasmids would include digestion with S1-nuclease and separation with PFGE or whole genome sequencing.

9.7 Concluding remarks

Valuable insights into the architecture and evolution of IncHI1 and IncHI2 plasmids and their resistance regions have been presented in this thesis. Although other methods had been previously used to analyse these plasmids, such as sequencing and SNP-typing, the tools designed in this work are a quicker, cheaper and simpler way to analyse sets of these large and complex plasmids with a high discriminatory power. These tools have been used to identify 2 separate IncHI1 plasmid types and at least 2 IncHI2a plasmid lineages. In the future it would be of interest to test these PCRs, by screening larger sets of IncHI plasmids, which may establish further lineages not detected in this study, and may determine other evolutionary paths by which these plasmids acquire antibiotic and heavy metal resistance regions. The ideal plasmid sets to examine are those that have been partially characterised previously, the IncHI1 plasmids examined by Phan and others (2009) [163] and the set of IncHI2 plasmids analysed by García-Fernández and others in 2007 [161] or in 2010 [160]. Testing the set of older IncHI1 and IncHI2 plasmids, previously analysed using RFLP [171], may also be useful in finding further ancestral plasmids. Furthermore, the mosaic nature of these plasmids has been emphasized here, with segments identified as originating from a number of sources. The novel MGEs that were uncovered indicated that they constitute a large proportion of these large plasmids, which were once much smaller and have accrued these elements over time. Insertion sequences, particularly IS*26*, seem to be extremely active within antibiotic and heavy metal resistance regions, causing deletions and insertions that bring in new resistance gene/s. Further work to study the transposition of IS*26* alone, and Tns formed by flanking insertion of IS*26*, is critical to understanding how and why it is so effective at making these rearrangements. Lastly, sequencing the IncHI2 plasmid TP116 and either pSRC125 or pSRC26 will yield valuable information and further define the IncHI2 plasmid backbone by determining whether other MGEs, particularly one carrying the tellurite resistance cluster, are present.

REFERENCES

1. Turner, P. http://www.yale.edu/turner/projects/ecoli.htm. 2003.
2. Hopwood, D., Levy, S., Wenzel, R. P., Georgopapadakou, N., Baltz, R. H., Bhavnani, S. and Cox, E., *A call to arms*. Nature Reviews Drug Discovery, 2007. **6**(1): p. 8-12.
3. Coates, A., Hu, Y. M., Bax, R. and Page, C., *The future challenges facing the development of new antimicrobial drugs*. Nature Reviews Drug Discovery, 2002. **1**(11): p. 895-910.
4. Hughes P. and Heritage, J., *Antibiotic growth-promoters in food animals*, F. Information division, Editor. 2004, Food and Agriculture Organization of the United Nations: Rome. p. 129-160.
5. JETACAR, *Report of the Joint Expert Advisory Committee On Antibiotic Resistance: The Use of Antibiotics in Food-Producing Animals: Antibiotic-Resistant Bacteria in Animals and Humans*. 1999, Department of Health and Ageing.
6. Sterner, K.E., *Veterinary-medicine - a textbook of the diseases of cattle, sheep, pigs, goats and horses*. Journal of the American Veterinary Medical Association, ed. C.B. D., O., M. Radostits, J., A. Henderson. Vol. 184. 1984. 705-705.
7. George, N., J. Faoagali, and M. Muller, *Silvazine (TM) (silver sulfadiazine and chlorhexidine) activity against 200 clinical isolates*. Burns, 1997. **23**(6): p. 493-495.
8. Gupta, A., D.E. Taylor, and S. Silver, *Molecular diversity of silver resistance genes*. Abstracts of the General Meeting of the American Society for Microbiology, 2001. **101**: p. 399.
9. Ip, M., Lui, S. L., Poon, V. K. M., Lung, I. and Burd, A., *Antimicrobial activities of silver dressings: an in vitro comparison*. Journal of Medical Microbiology, 2006. **55**(1): p. 59-63.
10. Gupta, A. and S. Silver, *Molecular genetics - Silver as a biocide: Will resistance become a problem?* Nature Biotechnology, 1998. **16**(10): p. 888-888.
11. Gupta, A., Matsui, K., Lo, J. F. and Silver, S., *Molecular basis for resistance to silver cations in Salmonella*. Nature Medicine, 1999. **5**(2): p. 183-188.
12. Barber, R.S., R. Braude, and K.G. Mitchell, *Antibiotic and copper supplements for fattening pigs*. British Journal of Nutrition, 1955. **9**(4): p. 378-381.
13. Tetaz, T.J.a.L., R. K., *Plasmid-controlled resistance to copper in Escherichia coli*. Journal of Bacteriology, 1983. **154**(3): p. 1263-1268.
14. Shakoori, A.R. and B. Muneer, *Copper-resistant bacteria from industrial effluents and their role in remediation of heavy metals in wastewater*. Folia Microbiologica, 2002. **47**(1): p. 43-50.
15. Silver, S. and L.T. Phung, *A bacterial view of the periodic table: genes and proteins for toxic inorganic ions*. Journal of Industrial Microbiology & Biotechnology, 2005. **32**(11-12): p. 587-605.
16. Silver, S. and L.T. Phung, *Bacterial heavy metal resistance: New surprises*. Annual Review of Microbiology, 1996. **50**: p. 753-789.
17. Bruins, M.R., S. Kapil, and F.W. Oehme, *Microbial resistance to metals in the environment*. Ecotoxicology and Environmental Safety, 2000. **45**(3): p. 198-207.
18. Rehman, A., Ali, A., Muneer, B. and Shakoori, A. R., *Resistance and biosorption of mercury by bacteria isolated from industrial effluents*. Pakistan Journal of Zoology, 2007. **39**(3): p. 137-146.
19. Barkay, T., S.M. Miller, and A.O. Summers, *Bacterial mercury resistance from atoms to ecosystems*. Fems Microbiology Reviews, 2003. **27**(2-3): p. 355-384.
20. Bradford, P.A., *Extended-spectrum beta-lactamases in the 21st century: Characterization, epidemiology, and detection of this important resistance threat*. Clinical Microbiology Reviews, 2001. **14**(4): p. 933-951.
21. Roberts, M.C., *Tetracycline resistance determinants: Mechanisms of action, regulation of expression, genetic mobility, and distribution*. Fems Microbiology Reviews, 1996. **19**(1): p. 1-24.

22. Chopra, I. and M. Roberts, *Tetracycline antibiotics: Mode of action, applications, molecular biology, and epidemiology of bacterial resistance*. Microbiology and Molecular Biology Reviews, 2001. **65**(2): p. 232-+.
23. Wilcox, S., *Cationic peptides: a new hope*, in *The Science Creative Quarterly*. 2004.
24. Babini, G.S. and D.M. Livermore, *Are SHV beta-lactamases universal in Klebsiella pneumoniae?* Antimicrobial Agents and Chemotherapy, 2000. **44**(8): p. 2230-2230.
25. Frost, L.S., Leplae, R., Summers, A. O. and Toussaint, A., *Mobile genetic elements: The agents of open source evolution*. Nature Reviews Microbiology, 2005. **3**(9): p. 722-732.
26. Kulkosky, J., Jones, K. S., Katz, R. A., Mack, J. P. and Skalka, A. M., *Residues critical for retroviral integrative recombination in a region that is highly conserved among retroviral retrotransposon integrases and bacterial insertion-sequence transposases*. Molecular and Cellular Biology, 1992. **12**(5): p. 2331-2338.
27. Polard, P. and M. Chandler, *Bacterial transposases and retroviral integrases*. Molecular Microbiology, 1995. **15**(1): p. 13-23.
28. Derbyshire, K.M. and N.D. Grindley, *Replicative and Conservative Transposition in Bacteria*. Cell, 1986. **47**(3): p. 325-327.
29. Mahillon, J. and M. Chandler, *Insertion sequences*. Microbiology and Molecular Biology Reviews, 1998. **62**(3): p. 725-+.
30. Grindley, N.D. and D.J. Sherratt, *Sequence-Analysis at IS1 Insertion Sites - Models for Transportation*. Cold Spring Harbor Symposia on Quantitative Biology, 1979. **43**: p. 1257-1261.
31. Partridge, S.R. and R.M. Hall, *The IS1111 family members IS4321 and IS5075 have subterminal inverted repeats and target the terminal inverted repeats of Tn21 family transposons*. Journal of Bacteriology, 2003. **185**(21): p. 6371-6384.
32. Sekine, Y. and E. Ohtsubo, *Frameshifting is required for production of the transposase encoded by insertion sequence-1*. Proceedings of the National Academy of Sciences of the United States of America, 1989. **86**(12): p. 4609-4613.
33. Siguier, P., J. Filee, and M. Chandler, *Insertion sequences in prokaryotic genomes*. Current Opinion in Microbiology, 2006. **9**(5): p. 526-531.
34. Chandler, M. and J. Mahillon, *Insertion Sequences Revisited*, in *Mobile DNA II* N.L. Craig, et al., Editors. 2002, ASM Press. p. 305-366
35. Galas, D.J. and M. Chandler, *On the Molecular Mechanisms of Transposition*. Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences, 1981. **78**(8): p. 4858-4862.
36. Chalmers, R.M. and N. Kleckner, *IS10 Tn10 transposition efficiently accommodates diverse transposon end configurations*. Embo Journal, 1996. **15**(18): p. 5112-5122.
37. Doroshenko, V.G. and V.A. Livshits, *Structure and mode of transposition of Tn2555 carrying sucrose utilization genes*. Fems Microbiology Letters, 2004. **233**(2): p. 353-359.
38. Grindley, N.D., *Movement of Tn3-like elements: transposition and cointegrate resolution*, in *Mobile DNA II*, N.L. Craig, et al., Editors. 2002, ASM press. p. 272-302.
39. Cohen, S.N., Casadaban, M. J., Chou, J. and Tu, C. P., *Studies of the specificity and control of transposition of the Tn3 element*. Cold Spring Harbor Symposia on Quantitative Biology, 1979. **43**: p. 1247-1255.
40. Grinsted, J., F. Delacruz, and R. Schmitt, *The Tn21 Subgroup of Bacterial Transposable Elements*. Plasmid, 1990. **24**(3): p. 163-189.
41. Sundin, G.W. and C.L. Bender, *Expression of the strA-strB streptomycin resistance genes in Pseudomonas-syringae and Xanthomonas-campestris and characterization of IS6100 in Xanthomonas-campestris*. Applied and Environmental Microbiology, 1995. **61**(8): p. 2891-2897.

42. 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.
43. Post, V., P.A. White, and R.M. Hall, *Evolution of AbaR-type genomic resistance islands in multiply antibiotic-resistant Acinetobacter baumannii*. Journal of Antimicrobial Chemotherapy, 2010. **65**(6): p. 1162-1170.
44. Post, V., *PhD Thesis: Analysis of Multiply Antibiotic Resistant Klebsiella pneumoniae and Acinetobacter baumannii Isolates From Australian Hospitals*, in *School of Molecular Bioscience*. 2011, The University of Sydney
45. Bainton, R., P. Gamas, and N.L. Craig, *Tn7 transposition invitro proceeds through an excised transposon intermediate generated by staggered breaks in DNA*. Cell, 1991. **65**(5): p. 805-816.
46. May, E.W. and N.L. Craig, *Switching from cut-and-paste to replicative Tn7 transposition*. Science, 1996. **272**(5260): p. 401-404.
47. Craig, N.L., *Transposon Tn7*, in *Transposable Elements*. 1996. p. 27-48.
48. Flores, C., M.I. Qadri, and C. Lichtenstein, *DNA-sequence analysis of 5 genes - tnsA, tnsB, tnsC, tnsD and tnsE, required for Tn7 transposition*. Nucleic Acids Research, 1990. **18**(4): p. 901-911.
49. Craig, N.L., *Tn7*, in *Mobile DNA II*, N.L. Craig, et al., Editors. 2002, ASM press.
50. Peters, J.E. and N.L. Craig, *Tn7: Smarter than we thought*. Nature Reviews Molecular Cell Biology, 2001. **2**(11): p. 806-814.
51. Tang, Y., C. Lichtenstein, and S. Cotterill, *Purification and characterization of the TnsB protein of Tn7 - a transposition protein that binds to the ends of Tn7*. Nucleic Acids Research, 1991. **19**(12): p. 3395-3402.
52. Gamas, P. and N.L. Craig, *Purification and characterization of TnsC, a Tn7 transposition protein that binds ATP and DNA*. Nucleic Acids Research, 1992. **20**(10): p. 2525-2532.
53. Bainton, R.J., Kubo, K. M., Feng, J. N. and Craig, N. L., *Tn7 Transposition - Target DNA Recognition Is Mediated by Multiple Tn7-Encoded Proteins in a Purified Invitro System*. Cell, 1993. **72**(6): p. 931-943.
54. McKown, R.L., Orle, K. A., Chen, T. and Craig, N. L., *Sequence Requirements of Escherichia-Coli Atttn7, a Specific Site of Transposon Tn7 Insertion*. Journal of Bacteriology, 1988. **170**(1): p. 352-358.
55. Craig, N.L., *Tn7 - a Target Site-Specific Transposon*. Molecular Microbiology, 1991. **5**(11): p. 2569-2573.
56. Wolkow, C.A., R.T. DeBoy, and N.L. Craig, *Conjugating plasmids are preferred targets for Tn7*. Genes & Development, 1996. **10**(17): p. 2145-2157.
57. Radstrom, P., Skold, O., Swedberg, G., Flensburg, J., Roy, P. H. and Sundstrom, L., *Transposon Tn5090 of plasmid R751, which carries an integron, is related to Tn7, Mu, and the retroelements*. Journal of Bacteriology, 1994. **176**(11): p. 3257-3268.
58. Gilmour, M.W., Thomson, N. R., Sanders, M., Parkhill, J. and Taylor, D. E., *The complete nucleotide sequence of the resistance plasmid R478: defining the backbone components of incompatibility group H conjugative plasmids through comparative genomics*. Plasmid, 2004. **52**(3): p. 182-202.
59. Shapiro, J.A. and P. Sporn, *Tn402 - new transposable element determining trimethoprim resistance that inserts in bacteriophage-lambda*. Journal of Bacteriology, 1977. **129**(3): p. 1632-1635.
60. Kholodii, G.Y., Mindlin, S. Z., Bass, I. A., Yurieva, O. V., Minakhina, S. V. and Nikiforov, V. G., *4 Genes, 2 Ends, and a Res Region Are Involved in Transposition of Tn5053 - a Paradigm for a Novel Family of Transposons Carrying Either a Mer Operon or an Integron*. Molecular Microbiology, 1995. **17**(6): p. 1189-1200.

61. Kholodii, G.Y., Yurieva, O. V., Lomovskaya, O. L., Gorlenko, Z. M., Mindlin, S. Z. and Nikiforov, V. G., *Tn5053, a mercury resistance transposon with integrons ends*. Journal of Molecular Biology, 1993. **230**(4): p. 1103-1107.
62. Kamali-Moghaddam, M. and L. Sundstrom, *Arrayed transposase-binding sequences on the ends of transposon Tn5090/Tn402*. Nucleic Acids Research, 2001. **29**(4): p. 1005-1011.
63. Minakhina, S., Kholodii, G., Mindlin, S., Yurieva, O. and Nikiforov, V., *Tn5053 family transposons are res site hunters sensing plasmidal res sites occupied by cognate resolvases*. Molecular Microbiology, 1999. **33**(5): p. 1059-1068.
64. Mindlin, S.Z., Minakhina, S. V., Kholodii, G. Y., Kopteva, A. V. and Nikiforov, V. G., *Integration of Tn5053 and Tn402 into different plasmids*. Genetika, 1996. **32**(10): p. 1426-1430.
65. Petrovski, S. and V.A. Stanisich, *Tn502 and Tn512 Are res Site Hunters That Provide Evidence of Resolvase-Independent Transposition to Random Sites*. Journal of Bacteriology, 2010. **192**(7): p. 1865-1874.
66. Kamali-Moghaddam, M. and L. Sundstrom, *Transposon targeting determined by resolvase*. Fems Microbiology Letters, 2000. **186**(1): p. 55-59.
67. Partridge, S.R., Brown, H.J., Stokes, H.W. and Hall, R.M., *Transposons Tn1696 and Tn21 and their integrons In4 and In2 have independent origins*. Antimicrobial Agents and Chemotherapy, 2001. **45**: p. 1263-1270.
68. Hall, R.M. and C.M. Collis, *Antibiotic resistance in gram-negative bacteria: the role of gene cassettes and integrons*. Drug Resistance Updates, 1998. **1**(2): p. 109-119.
69. Paulsen, I.T., Littlejohn, T. G., Radstrom, P., Sundstrom, L., Skold, O., Swedberg, G. and Skurray, R. A., *The 3' conserved segment of integrons contains a gene associated with multidrug resistance to antiseptics and disinfectants*. Antimicrobial Agents and Chemotherapy, 1993. **37**(4): p. 761-768.
70. Stokes, H.W. and R.M. Hall, *A novel family of potentially mobile DNA elements encoding site-specific gene-integration functions - integrons*. Molecular Microbiology, 1989. **3**(12): p. 1669-1683.
71. Brown, H.J., H.W. Stokes, and R.M. Hall, *The integrons In0, In2, and In5 are defective transposon derivatives*. J Bacteriol, 1996. **178**(15): p. 4429-4437.
72. Partridge, S.R., Brown, H.J. and Hall, R.M., *Characterization and movement of the class 1 integron known as Tn2521 and Tn1405*. Antimicrobial Agents and Chemotherapy, 2002. **46**(5): p. 1288-1294.
73. Partridge, S.R., Recchia, G. D., Stokes, H. W. and Hall, R. M., *Family of class 1 integrons related to In4 from Tn1696*. Antimicrobial Agents and Chemotherapy, 2001. **45**(11): p. 3014-3020.
74. Collis, C.M., Hall, R. M., *Gene Cassettes from the Insert Region of Integrons Are Excised as Covalently Closed Circles*. Molecular Microbiology, 1992. **6**(19): p. 2875-2885.
75. Collis, C.M., Grammaticopoulos, G., Briton, J., Stokes, H. W., Hall, R. M., *Site-Specific Insertion of Gene Cassettes into Integrons*. Molecular Microbiology, 1993. **9**(1): p. 41-52.
76. Hall, R.M. and C.M. Collis, *Mobile Gene Cassettes and Integrons - Capture and Spread of Genes by Site-Specific Recombination*. Molecular Microbiology, 1995. **15**(4): p. 593-600.
77. Collis, C.M., Hall, R. M., *Site-Specific Deletion and Rearrangement of Integron Insert Genes Catalyzed by the Integron DNA Integrase*. Journal of Bacteriology, 1992. **174**(5): p. 1574-1585.
78. Hall, R.M., Collis, C. M., Kim, M. J., Partridge, S. R., Recchia, G. D. and Stokes, H. W., *Mobile gene cassettes and integrons in evolution*, in *Molecular Strategies in Biological Evolution*, L.H. Caporale, Editor. 1999. p. 68-80.

79. Partridge, S.R., Tsafnat, G., Coiera, E. and Iredell, J. R., *Gene cassettes and cassette arrays in mobile resistance integrons*. Fems Microbiology Reviews, 2009. **33**(4): p. 757-784.
80. Collis, C.M. and R.M. Hall, *Expression of antibiotic resistance genes in the integrated cassettes of integrons*. Antimicrobial Agents and Chemotherapy, 1995. **39**(1): p. 155-162.
81. Levesque, C., Brassard, S., Lapointe, J. and Roy, P.H., *Diversity and relative strength of tandem promoters for the antibiotic resistance genes of several integrons*. Gene, 1994. **142**: p. 49-54.
82. Jove, T., Da Re, S., Denis, F., Mazel, D., Ploy, M. C., *Inverse Correlation between Promoter Strength and Excision Activity in Class 1 Integrons*. Plos Genetics, 2010. **6**(1).
83. Collis, C.M. and R.M. Hall, *Expression of Antibiotic-Resistance Genes in the Integrated Cassettes of Integrons*. Antimicrobial Agents and Chemotherapy, 1995. **39**(1): p. 155-162.
84. Partridge, S.R. and R.M. Hall, *In34, a complex In5 family class 1 integron containing orf513 and dfrA10*. Antimicrobial Agents and Chemotherapy, 2003. **47**(1): p. 342-349.
85. Hirsch, P.R., C.L. Wang, and M.J. Woodward, *Construction of a Tn5 derivative determining resistance to gentamicin and spectinomycin using a fragment cloned from R1033*. Gene, 1986. **48**(2-3): p. 203-209.
86. Smith, D.I., Gomezlus, R., Calvo, M. C., Datta, N., Jacob, A. E. and Hedges, R. W., *3rd type of plasmid conferring gentamicin resistance in Pseudomonas-aeruginosa*. Antimicrobial Agents and Chemotherapy, 1975. **8**(3): p. 227-230.
87. Marquez, C., Labbate, M., Raymondo, C., Fernandez, J., Gestal, A. M., Holley, M., Borthagaray, G. and Stokes, H. W., *Urinary tract infections in a South American population: Dynamic spread of class 1 integrons and multidrug resistance by homologous and site-specific recombination*. Journal of Clinical Microbiology, 2008. **46**(10): p. 3417-3425.
88. Liebert, C.A., R.M. Hall, and A.O. Summers, *Transposon Tn21, flagship of the floating genome*. Microbiol Molec Biol Rev, 1999. **63**(3): p. 507-522.
89. Iida, S., Hanni, C., Echarti, C. and Arber, W., *Is the IS1-flanked R-determinant of the R plasmid NR1 a transposon?* Journal of General Microbiology, 1981. **126**(OCT): p. 413-425.
90. Kholodii, G., Mindlin, S., Petrova, M. and Minakhina, S., *Tn5060 from the Siberian permafrost is most closely related to the ancestor of Tn21 prior to integron acquisition*. Fems Microbiology Letters, 2003. **226**(2): p. 251-255.
91. Petrova, M.A., Gorlenko, Z. M., Soina, V. S. and Mindlin, S. Z., *Association of the strA-strB genes with plasmids and transposons in the present-day bacteria and in bacterial strains from permafrost*. Russian Journal of Genetics, 2008. **44**(9): p. 1116-1120.
92. Toleman, M.A., P.M. Bennett, and T.R. Walsh, *ISCR elements: Novel gene-capturing systems of the 21st century?* Microbiology and Molecular Biology Reviews, 2006. **70**(2): p. 296-+.
93. Saedler, H., Cornelis, G., Cullum, J., Schumacher, B. and Sommer, H., *IS1-mediated DNA rearrangements* Cold Spring Harbor Symposia on Quantitative Biology, 1980. **45**: p. 93-98.
94. Roberts, D.E., D. Ascherman, and N. Kleckner, *IS10 promotes adjacent deletions at low-frequency*. Genetics, 1991. **128**(1): p. 37-43.
95. Ohtsubo, E. and Y. Sekine, *Bacterial insertion sequences*, in *Transposable Elements*. 1996. p. 1-26.

96. Mollet, B., Iida, S., Shepherd, J. and Arber, W., *Nucleotide-sequence of IS26, a new prokaryotic mobile genetic element*. Nucleic Acids Research, 1983. **11**(18): p. 6319-6330.
97. Martin, C., Timm, J., Rauzier, J., Gomezlus, R., Davies, J. and Gicquel, B., *Transposition of an antibiotic-resistance element in Mycobacteria*. Nature, 1990. **345**(6277): p. 739-743.
98. Shapiro, J.A., *Mutations caused by insertion of genetic material into galactose operon of Escherichia coli*. Journal of Molecular Biology, 1969. **40**(1): p. 93-&.
99. Kleckner, N., Chan, R. K., Tye, B. K. and Botstein, D., *Mutagenesis by insertion of a drug-resistance element carrying an inverted repetition*. Journal of Molecular Biology, 1975. **97**(4): p. 561-&.
100. Brau, B., U. Pilz, and W. Piepersberg, *Genes for gentamicin-(3)-n-acetyltransferases-iii and gentamicin-(3)-n-acetyltransferase-iv .1. Nucleotide-sequence of the aac(3)-iv gene and possible involvement of an IS140 element in its expression*. Molecular and General Genetics, 1984. **193**(1): p. 179-187.
101. Lee, K.Y., J.D. Hopkins, and M. Syvanen, *Direct Involvement of Is26 in an Antibiotic-Resistance Operon*. Journal of Bacteriology, 1990. **172**(6): p. 3229-3236.
102. Dawes, F.E., Kuzevski, A., Bettelheim, K. A., Hornitzky, M. A., Djordjevic, S. P. and Walker, M. J., *Distribution of Class 1 Integrons with IS26-Mediated Deletions in Their 3 '-Conserved Segments in Escherichia coli of Human and Animal Origin*. Plos One, 2010. **5**(9).
103. Shahid, M., *Citrobacter spp. Simultaneously Harboring bla(CTX-M), bla(TEM), bla(SHV), bla(ampC), and Insertion Sequences IS26 and orf513: an Evolutionary Phenomenon of Recent Concern for Antibiotic Resistance*. Journal of Clinical Microbiology, 2010. **48**(5): p. 1833-1838.
104. Doublet, B., Praud, K., Weill, F. X. and Cloeckaert, A., *Association of IS26-composite transposons and complex In4-type integrons generates novel multidrug resistance loci in Salmonella genomic island 1*. Journal of Antimicrobial Chemotherapy, 2009. **63**(2): p. 282-289.
105. Han, H.L., Jang, S. J., Park, G., Kook, J. K., Shin, J. H., Shin, S. H., Kim, D. M., Cheon, J. S., Moon, D. S. and Park, Y. J., *Identification of an atypical integron carrying an IS26-disrupted aadA1 gene cassette in Acinetobacter baumannii*. International Journal of Antimicrobial Agents, 2008. **32**(2): p. 165-169.
106. Leelaporn, A., Firth, N., Paulsen, I. T. and Skurray, R. A., *IS257-mediated cointegration in the evolution of a family of Staphylococcal trimethoprim resistance plasmids*. Journal of Bacteriology, 1996. **178**(20): p. 6070-6073.
107. Skurray, R.A. and N. Firth, *Molecular evolution of multiply-antibiotic-resistant staphylococci*, in *Antibiotic Resistance: Origins, Evolution, Selection and Spread*, D.J. Chadwick and J. Goode, Editors. 1997. p. 167-183.
108. Ohtsubo, H. and E. Ohtsubo, *Nucleotide sequence of an insertion sequence, IS1*. Proceedings of the National Academy of Sciences of the United States of America, 1978. **75**(2): p. 615-619.
109. Escoubas, J.M., D. Lane, and M. Chandler, *Is the IS1 transposase, InsAB', the only IS1-encoded protein required for efficient transposition*. Journal of Bacteriology, 1994. **176**(18): p. 5864-5867.
110. Kleckner, N., Chalmers, R. M., Kwon, D., Sakai, J. and Bolland, S., *Tn10 and IS10 transposition and chromosome rearrangements: Mechanism and regulation in vivo and in vitro*, in *Transposable Elements*. 1996. p. 49-82.
111. Stokes, H.W. and R.M. Hall, *The integron In1 in plasmid R46 includes 2 copies of the oxa2 gene cassette*. Plasmid, 1992. **28**(3): p. 225-234.
112. Djordjevic, S.P., Cain, A. K., Evershed, N. J., Falconer, L., Levings, R. S., Lightfoot, D. and Hall, R. M., *Emergence and Evolution of Multiply Antibiotic-Resistant*

Salmonella enterica Serovar Paratyphi B D-Tartrate-Utilizing Strains Containing SGI1. Antimicrobial Agents and Chemotherapy, 2009. **53**(6): p. 2319-2326.
113. Evershed, N.J., Levings, R. S., Wilson, N. L., Djordjevic, S. P. and Hall, R. M., *Unusual Class 1 Integron-Associated Gene Cassette Configuration Found in IncA/C Plasmids from Salmonella enterica*. Antimicrobial Agents and Chemotherapy, 2009. **53**(6): p. 2640-2642.
114. Yau, S., Liu, X. L., Djordjevic, S. P. and Hall, R. M., *RSF1010-Like Plasmids in Australian Salmonella enterica Serovar Typhimurium and Origin of Their sul2-strA-strB Antibiotic Resistance Gene Cluster*. Microbial Drug Resistance, 2010. **16**(4): p. 249-252.
115. Lorenz, M.G. and W. Wackernagel, *Bacterial gene-transfer by natural genetic-transformation in the environment*. Microbiological Reviews, 1994. **58**(3): p. 563-602.
116. Chen, I. and D. Dubnau, *DNA uptake during bacterial transformation*. Nature Reviews Microbiology, 2004. **2**(3): p. 241-249.
117. Zinder, N.D. and J. Lederberg, *Genetic exchange in Salmonella*. Journal of Bacteriology, 1952. **64**(5): p. 679-699.
118. Roberts, A.P. and P. Mullany, *A modular master on the move: the Tn916 family of mobile genetic elements*. Trends in Microbiology, 2009. **17**(6): p. 251-258.
119. Halary, S., Leigh, J. W., Cheaib, B., Lopez, P., Bapteste, E., *Network analyses structure genetic diversity in independent genetic worlds*. Proceedings of the National Academy of Sciences of the United States of America, 2010. **107**(1): p. 127-132.
120. Bentley, S.D. and J. Parkhill, *Comparative genomic structure of prokaryotes*. Annual Review of Genetics, 2004. **38**: p. 771-792.
121. del Solar, G., Giraldo, R., Ruiz-Echevarria, M. J., Espinosa, M. and Diaz-Orejas, R., *Replication and control of circular bacterial plasmids*. Microbiology and Molecular Biology Reviews, 1998. **62**(2): p. 434-+.
122. Scott, J.R., *Regulation of plasmid replication*. Microbiological Reviews, 1984. **48**(1): p. 1-23.
123. Yasueda, H., T. Horii, and T. Itoh, *Structural and Functional-Organization of Cole2 and Cole3 Replicons*. Molecular & General Genetics, 1989. **215**(2): p. 209-216.
124. Gerdes, K., J. Moller-Jensen, and R.B. Jensen, *Plasmid and chromosome partitioning: surprises from phylogeny*. Molecular Microbiology, 2000. **37**(3): p. 455-466.
125. Hiraga, S., *Chromosome and plasmid partitioning in Escherichia coli*. Annual Review of Biochemistry, 1992. **61**: p. 283-306.
126. Hayes, F., *Toxins-antitoxins: Plasmid maintenance, programmed cell death, and cell cycle arrest*. Science, 2003. **301**(5639): p. 1496-1499.
127. Schroder, G. and E. Lanka, *The mating pair formation system of conjugative plasmids - A versatile secretion machinery for transfer of proteins and DNA*. Plasmid, 2005. **54**(1): p. 1-25.
128. Smillie, C., Garcillan-Barcia, M. P., Francia, M. V., Rocha, E. P. and de la Cruz, F., *Mobility of Plasmids*. Microbiology and Molecular Biology Reviews, 2010. **74**(3): p. 434-+.
129. Lawley, T.D., Klimke, W. A., Gubbins, M. J. and Frost, L. S., *F factor conjugation is a true type IV secretion system*. Fems Microbiology Letters, 2003. **224**(1): p. 1-15.
130. Novotny, C.P. and K. Lavin, *Some effects of temperature on growth of F-pili*. Journal of Bacteriology, 1971. **107**(3): p. 671-&.
131. Lawley, T., B.M. Wilkins, and L.S. Frost, *Bacterial conjugation in gram-negative bacteria*. Plasmid Biology, 2004: p. 203-226.
132. Willetts, N. and B. Wilkins, *Processing of plasmid DNA during bacterial conjugation*. Microbiological Reviews, 1984. **48**(1): p. 24-41.

133. Garcillan-Barcia, M.P., M.V. Francia, and F. de la Cruz, *The diversity of conjugative relaxases and its application in plasmid classification*. Fems Microbiology Reviews, 2009. **33**(3): p. 657-687.
134. Heinemann, J.A., *Genetics of gene-transfer between species*. Trends in Genetics, 1991. **7**(6): p. 181-185.
135. Bates, S., A.M. Cashmore, and B.M. Wilkins, *IncP plasmids are unusually effective in mediating conjugation of Escherichia coli and Saccharomyces cerevisiae: Involvement of the Tra2 mating system*. Journal of Bacteriology, 1998. **180**(24): p. 6538-6543.
136. Meyer, R., *Replication and conjugative mobilization of broad host-range IncQ plasmids*. Plasmid, 2009. **62**(2): p. 57-70.
137. Doublet, B., Boyd, D., Mulvey, M. R. and Cloeckaert, A., *The Salmonella genomic island 1 is an integrative mobilizable element*. Molecular Microbiology, 2005. **55**(6): p. 1911-1924.
138. Novick, R.P., *Plasmid Incompatibility*. Microbiological Reviews, 1987. **51**(4): p. 381-395.
139. Grindley, N.D., J.N. Grindley, and E.S. Anderson, *R Factor Compatibility Groups*. Molecular & General Genetics, 1972. **119**(4): p. 287-297.
140. Couturier, M., Bex, F., Bergquist, P. L. and Maas, W. K., *Identification and classification of bacterial plasmids*. Microbiological Reviews, 1988. **52**(3): p. 375-395.
141. Taylor, D.E., R.W. Hedges, and P.L. Bergquist, *Molecular Homology and Incompatibility Relationships between F-Plasmids and Inch1-Plasmids*. Journal of General Microbiology, 1985. **131**(JUN): p. 1523-1530.
142. Carattoli, A., *Resistance Plasmid Families in Enterobacteriaceae*. Antimicrobial Agents and Chemotherapy, 2009. **53**(6): p. 2227-2238.
143. Morales, V., Bagdasarian, M. and Bagdasarian, M., *Promiscuous plasmids of the IncQ group - mode of replication and use for gene cloning in Gram-negative bacteria*, in *Pseudomonas : Biotransformations, Pathogenesis, and Evolving Biotechnology*, S. Silver, et al., Editors. 1990. p. 229-241.
144. Gormley, E.P. and J. Davies, *Transfer of plasmid RSF1010 by conjugation from Escherichia-coli to Streptomyces-lividans and Mycobacterium-smegmatis*. Journal of Bacteriology, 1991. **173**(21): p. 6705-6708.
145. Carattoli, A., Bertini, A., Villa, L., Falbo, V., Hopkins, K. L. and Threlfall, E. J., *Identification of plasmids by PCR-based replicon typing*. Journal of Microbiological Methods, 2005. **63**(3): p. 219-228.
146. Villa, L., Garcia-Fernandez, A., Fortini, D. and Carattoli, A., *Replicon sequence typing of IncF plasmids carrying virulence and resistance determinants*. Journal of Antimicrobial Chemotherapy, 2010. **65**(12): p. 2518-2529.
147. Bertini, A., Poirel, L., Mugnier, P. D., Villa, L., Nordmann, P. and Carattoli, A., *Characterization and PCR-Based Replicon Typing of Resistance Plasmids in Acinetobacter baumannii*. Antimicrobial Agents and Chemotherapy, 2010. **54**(10): p. 4168-4177.
148. Datta, N. and V.M. Hughes, *Plasmids of the same inc groups in Enterobacteria before and after the medical use of antibiotics*. Nature, 1983. **306**(5943): p. 616-617.
149. Jones, C. and J. Stanley, *Salmonella plamids of the preantibiotic era*. Journal of General Microbiology, 1992. **138**: p. 189-197.
150. Essa, A.M.M., Julian, D. J., Kidd, S. P., Brown, N. L. and Hobman, J. L., *Mercury resistance determinants related to Tn21, Tn1696, and Tn5053 in enterobacteria from the preantibiotic era*. Antimicrobial Agents and Chemotherapy, 2003. **47**(3): p. 1115-1119.

151. Cohen, S.N., *Transposable genetic elements and plasmid evolution*. Nature, 1976. **263**(5580): p. 731-738.
152. Gorai, A.P., Heffron, F., Falkow, S., Hedges, R. W. and Datta, N., *Electron-Microscope Heteroduplex Studies of Sequence Relationships among Plasmids of the W-Incompatibility Group*. Plasmid, 1979. **2**(3): p. 485-492.
153. Villarroel, R., Hedges, R. W., Maenhaut, R., Leemans, J., Engler, G., Vanmontagu, M. and Schell, J., *Heteroduplex Analysis of P-Plasmid Evolution - the Role of Insertion and Deletion of Transposable Elements*. Molecular & General Genetics, 1983. **189**(3): p. 390-399.
154. Fernandez-Lopez, R., Garcillan-Barcia, M. P., Revilla, C., Lazaro, M., Vielva, L. and de la Cruz, F., *Dynamics of the IncW genetic backbone imply general trends in conjugative plasmid evolution*. Fems Microbiology Reviews, 2006. **30**(6): p. 942-966.
155. Ward, J.M. and J. Grinsted, *Physical and genetic-analysis of the IncW group plasmids R388, SA, and R7K*. Plasmid, 1982. **7**(3): p. 239-250.
156. Valentine, C.R., *One-Kilobase Direct Repeats of Plasmid Psa*. Plasmid, 1985. **14**(2): p. 167-170.
157. Revilla, C., Garcillan-Barcia, M. P., Fernandez-Lopez, R., Thomson, N. R., Sanders, M., Cheung, M., Thomas, C. A. and de la Cruz, F., *Different pathways to acquiring resistance genes illustrated by the recent evolution of IncW plasmids*. Antimicrobial Agents and Chemotherapy, 2008. **52**(4): p. 1472-1480.
158. Fricke, W.F., Welch, T. J., McDermott, P. F., Mammel, M. K., LeClerc, J. E., White, D. G., Cebula, T. A. and Ravel, J., *Comparative Genomics of the IncA/C Multidrug Resistance Plasmid Family*. Journal of Bacteriology, 2009. **191**(15): p. 4750-4757.
159. Call, D.R., Singer, R. S., Meng, D., Broschat, S. L., Orfe, L. H., Anderson, J. M., Herndon, D. R., Kappmeyer, L. S., Daniels, J. B. and Besser, T. E., *bla(CMY-2)-Positive IncA/C Plasmids from Escherichia coli and Salmonella enterica Are a Distinct Component of a Larger Lineage of Plasmids*. Antimicrobial Agents and Chemotherapy, 2010. **54**(2): p. 590-596.
160. Garcia-Fernandez, A. and A. Carattoli, *Plasmid double locus sequence typing for IncHI2 plasmids, a subtyping scheme for the characterization of IncHI2 plasmids carrying extended-spectrum beta-lactamase and quinolone resistance genes*. Journal of Antimicrobial Chemotherapy, 2010. **65**(6): p. 1155-1161.
161. García-Fernández, A., Cloeckaert, A., Bertini, A., Praud, K., Doublet, B., Weill, F. X. and Carattoli, A., *Comparative analysis of IncHI2 plasmids carrying bla(CTX-M-2) or bla(CTX-M-9) from Escherichia coli and salmonella enterica strains isolated from poultry and humans*. Antimicrobial Agents and Chemotherapy, 2007. **51**(11): p. 4177-4180.
162. Welch, T.J., Fricke, W. F., McDermott, P. F., White, D. G., Rosso, M. L., Rasko, D. A., Mammel, M. K., Eppinger, M., Rosovitz, M. J., Wagner, D., Rahalison, L., LeClerc, J. E., Hinshaw, J. M., Lindler, L. E., Cebula, T. A., Carniel, E. and Ravel, J., *Multiple Antimicrobial Resistance in Plague: An Emerging Public Health Risk*. Plos One, 2007. **2**(3).
163. Phan, M.D., Kidgell, C., Nair, S., Holt, K. E., Turner, A. K., Hinds, J., Butcher, P., Cooke, F. J., Thomson, N. R., Titball, R., Bhutta, Z. A., Hasan, R., Dougan, G. and Wain, J., *Variation in Salmonella enterica Serovar Typhi IncHI1 Plasmids during the Global Spread of Resistant Typhoid Fever*. Antimicrobial Agents and Chemotherapy, 2009. **53**(2): p. 716-727.
164. Garcia-Fernandez, A., Chiaretto, G., Bertini, A., Villa, L., Fortini, D., Ricci, A. and Carattoli, A., *Multilocus sequence typing of IncI1 plasmids carrying extended-spectrum beta-lactamases in Escherichia coli and Salmonella of human and animal origin*. Journal of Antimicrobial Chemotherapy, 2008. **61**(6): p. 1229-1233.

165. Newnham, P.J. and D.E. Taylor, *Molecular analysis of repHI1a, a minimal replicon of the IncHI1 plasmid R27*. Molecular Microbiology, 1994. **11**(4): p. 757-768.
166. Page, D.T., K.F. Whelan, and E. Colleran, *Characterization of two autoreplicative regions of the IncHI2 plasmid R478: RepHI2A and RepHI1A((R478))*. Microbiology-Sgm, 2001. **147**: p. 1591-1598.
167. Grindley, N.D., H. G., and E.S. Anderson, *Molecular studies of R-factor compatibility groups*. Journal of Bacteriology, 1973. **115**(1): p. 387-398.
168. Bradley, D.E., et al., *R plasmids of a new incompatibility group determine constitutive production of H-pili* Plasmid, 1982. **7**(3): p. 230-238.
169. Taylor, D.E. and R.B. Grant, *Incompatibility and bacteriophage inhibition properties of N1, a plasmid belonging to H2 incompatibility group*. Molecular & General Genetics, 1977. **153**(1): p. 5-10.
170. Taylor, D.E. and R.B. Grant, *Incompatibility and surface exclusion properties of H1 and H2 plasmids*. Journal of Bacteriology, 1977. **131**(1): p. 174-178.
171. Whiteley, M. and D.E. Taylor, *Identification of DNA homologies among H-incompatibility group plasmids by restriction enzyme digestion and southern transfer hybridization*. Antimicrobial Agents and Chemotherapy, 1983. **24**(2): p. 194-200.
172. Roussel, A.F. and Y.A. Chabbert, *Taxonomy and epidemiology of gram-negative bacterial plasmids studied by DNA-DNA filter hybridization in formamide*. Journal of General Microbiology, 1978. **104**(FEB): p. 269-276.
173. Smith, H.R., Grindley, N. D., Humphrey, G., and Anderson, E. S., *Interactions of group H resistance factors with F factor*. Journal of Bacteriology, 1973. **115**(2): p. 623-628.
174. Taylor, D.E. and R.B. Grant, *Inhibition of bacteriophage-lambda, T1, and T7 development by R plasmids of H incompatibility group*. Antimicrobial Agents and Chemotherapy, 1976. **10**(4): p. 762-764.
175. Taylor, D.E., Brose, E. C., Kwan, S. and Yan, W., *Mapping of Transfer Regions within Incompatibility Group-Hi Plasmid R27*. Journal of Bacteriology, 1985. **162**(3): p. 1221-1226.
176. Taylor, D.E. and J.G. Levine, *Studies of Temperature-Sensitive Transfer and Maintenance of H Incompatibility Group Plasmids*. Journal of General Microbiology, 1980. **116**(FEB): p. 475-484.
177. Maher, D. and D.E. Taylor, *Host-Range and Transfer Efficiency of Incompatibility Group-Hi Plasmids*. Canadian Journal of Microbiology, 1993. **39**(6): p. 581-587.
178. Rodriguezlemoin, V., Jacob, A. E., Hedges, R. W. and Datta, N., *Thermosensitive production of their transfer systems by group S plasmids*. Journal of General Microbiology, 1975. **86**(JAN): p. 111-114.
179. Sherburne, C.K., Lawley, T. D., Gilmour, M. W., Blattner, F. R., Burland, V., Grotbeck, E., Rose, D. J. and Taylor, D. E., *The complete DNA sequence and analysis of R27, a large IncHI plasmid from Salmonella typhi that is temperature sensitive for transfer*. Nucleic Acids Research, 2000. **28**(10): p. 2177-2186.
180. Lawley, T.D. and D.E. Taylor, *Characterization of the double-partitioning modules of R27: Correlating plasmid stability with plasmid localization*. Journal of Bacteriology, 2003. **185**(10): p. 3060-3067.
181. Wain, J. and C. Kidgell, *The emergence of multidrug resistance to antimicrobial agents for the treatment of typhoid fever*. Transactions of the Royal Society of Tropical Medicine and Hygiene, 2004. **98**(7): p. 423-430.
182. Gabant, P., Newnham, P., Taylor, D. and Couturier, M., *Isolation and Location on the R27 Map of 2 Replicons and an Incompatibility Determinant Specific for Inchi1 Plasmids*. Journal of Bacteriology, 1993. **175**(23): p. 7697-7701.

183. Lawley, T.D., Gilmour, M. W., Gunton, J. E., Standeven, L. J. and Taylor, D. E., *Functional and mutational analysis of conjugative transfer region 1 (Tra1) from the IncHI1 plasmid R27*. Journal of Bacteriology, 2002. **184**(8): p. 2173-2180.
184. Rooker, M.M., Sherburne, C., Lawley, T. D. and Taylor, D. E., *Characterization of the Tra2 region of the IncHI1 plasmid R27*. Plasmid, 1999. **41**(3): p. 226-239.
185. Saul, D., D. Lane, and P.L. Bergquist, *A replication region of the inchi plasmid, R27, is highly homologous with the repfia replicon of F*. Molecular Microbiology, 1988. **2**(2): p. 219-225.
186. Lawley, T.D., Gilmour, M. W., Gunton, J. E., Tracz, D. M. and Taylor, D. E., *Functional and mutational analysis of conjugative transfer region 2 (Tra2) from the IncHI1 plasmid R27*. Journal of Bacteriology, 2003. **185**(2): p. 581-591.
187. Wain, J., Nga, L. T. D., Kidgell, C., James, K., Fortune, S., Diep, T. S., Ali, T., Gaora, P. O., Parry, C., Parkhill, J., Farrar, J., White, N. J. and Dougan, G., *Molecular analysis of incHI1 antimicrobial resistance Plasmids from Salmonella serovar typhi strains associated with typhoid fever*. Antimicrobial Agents and Chemotherapy, 2003. **47**(9): p. 2732-2739.
188. Holt, K.E., Thomson, N. R., Wain, J., Phan, M. D., Nair, S., Hasan, R., Bhutta, Z. A., Quail, M. A., Norbertczak, H., Walker, D., Dougan, G. and Parkhill, J., *Multidrug-resistant Salmonella enterica serovar Paratyphi A harbors IncHI1 plasmids similar to those found in serovar Typhi*. Journal of Bacteriology, 2007. **189**(11): p. 4257-4264.
189. Ogura, Y., Ooka, T., Iguchi, A., Toh, H., Asadulghani, M., Oshima, K., Kodama, T., Abe, H., Nakayama, K., Kurokawa, K., Tobe, T., Hattori, M. and Hayashi, T., *Comparative genomics reveal the mechanism of the parallel evolution of O157 and non-O157 enterohemorrhagic Escherichia coli*. Proceedings of the National Academy of Sciences of the United States of America, 2009. **106**(42): p. 17939-17944.
190. Parkhill, J., et al., *Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhi CT18*. Nature, 2001. **413**(6858): p. 848-852.
191. Bhan, M.K., R. Bahl, and S. Bhatnagar, *Typhoid and paratyphoid fever*. Lancet, 2005. **366**(9487): p. 749-762.
192. Urwin, R. and M.C.J. Maiden, *Multi-locus sequence typing: a tool for global epidemiology*. Trends in Microbiology, 2003. **11**(10): p. 479-487.
193. Garcia, A., Navarro, F., Miro, E., Villa, L., Mirelis, B., Coll, P. and Carattoli, A., *Acquisition and diffusion of bla(CTX-M-9) gene by R478-IncHI2 derivative plasmids*. Fems Microbiology Letters, 2007. **271**(1): p. 71-77.
194. Diestra, K., Juan, C., Curiao, T., Moya, B., Miro, E., Oteo, J., Coque, T. M., Perez-Vazquez, M., Campos, J., Canton, R., Oliver, A. and Navarro, F., *Characterization of plasmids encoding bla(ESBL) and surrounding genes in Spanish clinical isolates of Escherichia coli and Klebsiella pneumoniae*. Journal of Antimicrobial Chemotherapy, 2009. **63**(1): p. 60-66.
195. Tato, M., Coque, T. M., Baquero, F. and Canton, R., *Dispersal of Carbapenemase bla(VIM-1) Gene Associated with Different Tn402 Variants, Mercury Transposons, and Conjugative Plasmids in Enterobacteriaceae and Pseudomonas aeruginosa*. Antimicrobial Agents and Chemotherapy, 2010. **54**(1): p. 320-327.
196. Cordano, A.M. and R. Virgilio, *Evolution of drug resistance in Salmonella panama isolates in Chile*. Antimicrobial Agents and Chemotherapy, 1996. **40**(2): p. 336-341.
197. Smet, A., Martel, A., Persoons, D., Dewulf, J., Heyndrickx, M., Cloeckaert, A., Praud, K., Claeys, G., Catry, B., Herman, L., Haesebrouck, F. and Butaye, P., *Comparative analysis of extended-spectrum-beta-lactamase-carrying plasmids from different members of Enterobacteriaceae isolated from poultry, pigs and humans: evidence for a shared beta-lactam resistance gene pool?* Journal of Antimicrobial Chemotherapy, 2009. **63**(6): p. 1286-1288.

198. Hedges, R.W., Rodriguezlemoin, V. and Datta, N., *R-factors from Serratia marcescens* Journal of General Microbiology, 1975. **86**(JAN): p. 88-92.
199. Taylor, D.E. and R.B. Grant, *R-Plasmids of S-Incompatibility Group Belong to H2-Incompatibility Group*. Antimicrobial Agents and Chemotherapy, 1977. **12**(3): p. 431-434.
200. Page, D.T., K.F. Whelan, and E. Colleran, *Mapping studies and genetic analysis of transfer genes of the multi-resistant IncHI2 plasmid, R478*. Fems Microbiology Letters, 1999. **179**(1): p. 21-29.
201. Whelan, K.F., E. Colleran, and D.E. Taylor, *Phage inhibition, colicin resistance, and tellurite resistance are encoded by a single cluster of genes on the incHI2 plasmid R478*. Journal of Bacteriology, 1995. **177**(17): p. 5016-5027.
202. Ryan, D. and E. Colleran, *Arsenical resistance in the IncHI2 plasmids*. Plasmid, 2002. **47**(3): p. 234-240.
203. Gupta, A., Phung, L. T., Taylor, D. E. and Silver, S., *Diversity of silver resistance genes in IncH incompatibility group plasmids*. Microbiology-Sgm, 2001. **147**: p. 3393-3402.
204. Hopkins, K.L., Liebana, E., Villa, L., Batchelor, M., Threlfall, E. J. and Carattoli, A., *Replicon typing of Plasmids carrying CTX-M or CMY beta-lactamases circulating among Salmonella and Escherichia coli isolates*. Antimicrobial Agents and Chemotherapy, 2006. **50**(9): p. 3203-3206.
205. Novais, A., Canton, R., Valverde, A., Machado, E., Galan, J. C., Peixe, L., Carattoli, A., Baquero, F. and Coque, T. M., *Dissemination and persistence of bla(CTX-M-9) are linked to class 1 integrons containing CR1 associated with defective transposon derivatives from Tn402 located in early antibiotic resistance plasmids of IncHI2, IncP1-alpha, and IncFI groups*. Antimicrobial Agents and Chemotherapy, 2006. **50**(8): p. 2741-2750.
206. Marcade, G., Deschamps, C., Boyd, A., Gautier, V., Picard, B., Branger, C., Denamur, E. and Arlet, G., *Replicon typing of plasmids in Escherichia coli producing extended-spectrum beta-lactamases*. Journal of Antimicrobial Chemotherapy, 2009. **63**(1): p. 67-71.
207. Navarro, F., Mesa, R. J., Miro, E., Gomez, L., Mirelis, B. and Coll, P., *Evidence for convergent evolution of CTX-M-14 ESBL in Escherichia coli and its prevalence*. Fems Microbiology Letters, 2007. **273**(1): p. 120-123.
208. Chen, Y.T., Lauderdale, T. L., Liao, T. L., Shiau, Y. R., Shu, H. Y., Wu, K. M., Yan, J. J., Su, I. J. and Tsai, S. F., *Sequencing and comparative genomic analysis of pK29, a 269-kilobase conjugative plasmid encoding CMY-8 and CTX-M-3 beta-lactamases in Klebsiella pneumoniae*. Antimicrobial Agents and Chemotherapy, 2007. **51**(8): p. 3004-3007.
209. Chen, Y.T., Liao, T. L., Liu, Y. M., Lauderdale, T. L., Yan, J. J. and Tsai, S. F., *Mobilization of qnrB2 and ISCR1 in Plasmids*. Antimicrobial Agents and Chemotherapy, 2009. **53**(3): p. 1235-1237.
210. Espedido, B.A., S.R. Partridge, and J.R. Iredell, *bla(IMP-4) in different genetic contexts in Enterobacterzaceae isolates from Australia*. Antimicrobial Agents and Chemotherapy, 2008. **52**(8): p. 2984-2987.
211. Johnson, T.J., Wannemeuhler, Y. M., Scaccianoce, J. A., Johnson, S. J. and Nolan, L. K., *Complete DNA sequence comparative genomics, and prevalence of an IncHI2 plasmid occurring among extraintestinal pathogenic Escherichia coli isolates*. Antimicrobial Agents and Chemotherapy, 2006. **50**(11): p. 3929-3933.
212. Hall, R.M., *Antibiotic resistance gene cluster of pAPEC-O1-R*. Antimicrobial Agents and Chemotherapy, 2007. **51**(9): p. 3461-3462.
213. Wu, J.J., Ko, W. C., Tsai, S. H. and Yan, J. J., *Prevalence of plasmid-mediated quinolone resistance determinants QnrA QnrB, and QnrS among clinical isolates of*

Enterobacter cloacae in a Taiwanese hospital. Antimicrobial Agents and Chemotherapy, 2007. **51**(4): p. 1223-1227.
214. Velge, P., A. Cloeckaert, and P. Barrow, *Emergence of Salmonella epidemics: The problems related to Salmonella enterica serotype Enteritidis and multiple antibiotic resistance in other major serotypes*. Veterinary Research, 2005. **36**(3): p. 267-288.
215. Christensen, H., S. Nordentoft, and J.E. Olsen, *Phylogenetic relationships of Salmonella based on rRNA sequences*. International Journal of Systematic Bacteriology, 1998. **48**: p. 605-610.
216. Popoff, M.Y., J. Bockemuhl, and L.L. Gheesling, *Supplement 2001 (no. 45) to the Kauffmann-White scheme*. Research in Microbiology, 2003. **154**(3): p. 173-174.
217. Ward, L.R., J.D.H. Desa, and B. Rowe, *A phage-typing scheme for Salmonella enteritidis*. Epidemiology and Infection, 1987. **99**(2): p. 291-294.
218. Stanley, J., N. Baquar, and E.J. Threlfall, *Genotypes and phylogenetic-relationships of Salmonella typhimurium are defined by molecular fingerprinting of IS200 and 16s rrn loci*. Journal of General Microbiology, 1993. **139**: p. 1133-1140.
219. Guerra, B., P. Schrors, and M.C. Mendoza, *Application of PFGE performed with XbaI to an epidemiological and phylogenetic study of Salmonella serotype typhimurium. Relations between genetic types and phage types*. Microbiologica, 2000. **23**(1): p. 11-20.
220. Crump, J.A., S.P. Luby, and E.D. Mintz, *The global burden of typhoid fever*. Bulletin of the World Health Organization, 2004. **82**(5): p. 346-353.
221. Threlfall, E.J., Ward, L. R., Skinner, J. A. and Graham, A., *Antimicrobial drug resistance in non-typhoidal salmonellas from humans in England and Wales in 1999: Decrease in multiple resistance in Salmonella enterica serotypes Typhimurium, Virchow, and Hadar*. Microbial Drug Resistance-Mechanisms Epidemiology and Disease, 2000. **6**(4): p. 319-325.
222. Mead, P.S., Slutsker, L., Dietz, V., McCaig, L. F., Bresee, J. S., Shapiro, C., Griffin, P. M. and Tauxe, R. V., *Food-related illness and death in the United States*. Emerging Infectious Diseases, 1999. **5**(5): p. 607-625.
223. Ribot, E.M., Wierzba, R.K., Angulo, F.J. and Barrett, T.J., *Salmonella enterica serotype Typhimurium DT104 Isolated from Humans, United States, 1985, 1990, and 1995*. Emerg Infect Dis, 2002. **8**(4): p. 387-391.
224. Boyd, D.A., et al., *Partial characterization of a genomic island associated with the multidrug resistance region of Salmonella enterica Typhymurium DT104*. FEMS Microbiology Letters, 2000. **189**(2): p. 285-291.
225. Hall, R.M., *Salmonella genomic islands and antibiotic resistance in Salmonella enterica*. Future Microbiology, 2010. **5**(10): p. 1525-1538.
226. NNDSS. *National Notifiable Diseases Surveillance System*. [Website]; Available from: http://www9.health.gov.au/cda/Source/CDA-index.cfm.
227. Levings, R.S., Lightfoot, D., Partridge, S. R., Hall, R. M. and Djordjevic, S. P., *The genomic island SGI1, containing the multiple antibiotic resistance region of Salmonella enterica serovar typhimurium DT104 or variants of it, is widely distributed in other S-enterica serovars*. Journal of Bacteriology, 2005. **187**(13): p. 4401-4409.
228. Levings, R.S., S.P. Djordjevic, and R.M. Hall, *SGI2, a relative of Salmonella genomic island SGI1 with an independent origin*. Antimicrobial Agents and Chemotherapy, 2008. **52**(7): p. 2529-2537.
229. Levings, R.S., Hall, R. M., Lightfoot, D. and Djordjevic, S. P., *linG, a new integron-associated gene cassette encoding a lincosamide nucleotidyltransferase*. Antimicrobial Agents and Chemotherapy, 2006. **50**(10): p. 3514-3515.
230. Levings, R.S., Lightfoot, D., Elbourne, L. D. H., Djordjevic, S. P. and Hall, R. M., *New integron-associated gene cassette encoding a trimethoprim-resistant dfrB-type*

dihydrofolate reductase. Antimicrobial Agents and Chemotherapy, 2006. **50**(8): p. 2863-2865.
231. Levings, R.S., Partridge, S.R., Lightfoot, D., Hall, R.M. and Djordjevic, S.P., *A new integron-associated gene cassette encoding a 3-N-aminoglycoside acetyltransferase*. Antimicrob Agents Chemother, 2005. **49**(3): p. 1238-1241.
232. Liu, X., *PhD Thesis: Characterisation of antibiotic resistance gene clusters and their mobility within a collection of multi-drug resistant Salmonella spp*, in *School of Biological Sciences*. 2009, University of Wollongong.
233. Sambrook, J., E.F. Fritsch, and T. Maniatis, *Molecular cloning, a laboratory manual*. 2nd ed. 1989, Cold Spring Harbor, N. Y.: Cold Spring Harbor Laboratory Press.
234. Hanahan, D., *Studies on Transformation of Escherichia-Coli with Plasmids*. Journal of Molecular Biology, 1983. **166**(4): p. 557-580.
235. Walker, M.J.a.J.M.P., *Construction of a transposon containing a gene for polygalacturonate trans-eliminase from Klebsiella oxytoca*. Archives of Microbiology, 1987. **146**(4): p. 390-395.
236. Yanischperron, C., J. Vieira, and J. Messing, *Improved m13 phage cloning vectors and host strains - nucleotide sequences of the m13mp18 and pUC19 vectors*. Gene, 1985. **33**(1): p. 103-119.
237. Delacruz, F. and J. Grinsted, *Genetic and Molecular Characterization of Tn21, a Multiple Resistance Transposon from R100.1*. Journal of Bacteriology, 1982. **151**(1): p. 222-228.
238. Zhong, S.B. and A.M. Dean, *Rapid identification and mapping of insertion sequences in Escherichia coli genomes using vectorette PCR*. Bmc Microbiology, 2004. **4**:26.
239. Johnson, T.J., Wannemuehler, Y. M., Johnson, S. J., Logue, C. M., White, D. G., Doetkott, C. and Nolan, L. K., *Plasmid replicon typing of commensal and pathogenic Escherichia coli isolates*. Applied and Environmental Microbiology, 2007. **73**(6): p. 1976-1983.
240. Whelan, K.F., Maher, D., Colleran, E. and Taylor, D. E., *Genetic and nucleotide-sequence analysis of the gene htdA, which regulates conjugal transfer of incHI plasmids*. Journal of Bacteriology, 1994. **176**(8): p. 2242-2251.
241. Yurieva, O., Kholodii, G., Minakhin, L., Gorlenko, Z., Kalyaeva, E., Mindlin, S. and Nikiforov, V., *Intercontinental spread of promiscuous mercury-resistance transposons in environmental bacteria*. Molecular Microbiology, 1997. **24**(2): p. 321-329.
242. Benjamin, H.W. and N. Kleckner, *Intramolecular transposition by Tn10*. Cell, 1989. **59**(2): p. 373-383.
243. Partridge, S.R. and R.M. Hall, *Complex multiple antibiotic and mercury resistance region derived from the r-det of NR1 (R100)*. Antimicrob. Agents Chemother., 2004. **48**(11): p. 4250-4255.
244. Scholz, P., Haring, V., Wittmannliebold, B., Ashman, K., Bagdasarian, M. and Scherzinger, E., *Complete nucleotide sequence and gene organization of the broad-host-range plasmid RSF1010*. Gene, 1989. **75**(2): p. 271-288.
245. Szczepanowski, R., Braun, S., Riedel, V., Schneiker, S., Krahn, I., Puhler, A. and Schluter, A., *The 120 592 bp lncF plasmid pRSB107 isolated from a sewage-treatment plant encodes nine different anti biotic-resistance determinants, two iron-acquisition systems and other putative virulence-associated functions*. Microbiology-Sgm, 2005. **151**: p. 1095-1111.
246. Venturini, C., Beatson, S. A., Djordjevic, S. P. and Walker, M. J., *Multiple antibiotic resistance gene recruitment onto the enterohemorrhagic Escherichia coli virulence plasmid*. Faseb Journal, 2010. **24**(4): p. 1160-1166.
247. Tauch, A., Krieft, S., Kalinowski, J. and Puhler, A., *The 51,409-bp R-plasmid pTP10 from the multiresistant clinical isolate Corynebacterium striatum M82B is composed*

of DNA segments initially identified in soil bacteria and in plant, animal, and human pathogens. Molecular and General Genetics, 2000. **263**(1): p. 1-11.
248. Hawley, D.K. and W.R. McClure, *Compilation and analysis of Escherichia coli promoter DNA sequences*. Nucleic Acids Research, 1983. **11**(8): p. 2237-2255.
249. Burr, T., Mitchell, J., Kolb, A., Minchin, S. and Busby, S., *DNA sequence elements located immediately upstream of the -10 hexamer in Escherichia coli promoters: a systematic study*. Nucleic Acids Research, 2000. **28**(9): p. 1864-1870.
250. Heuer, H., Kopmann, C., Binh, C. T. , Top, E. M. and Smalla, K., *Spreading antibiotic resistance through spread manure: characteristics of a novel plasmid type with low %G plus C content*. Environmental Microbiology, 2009. **11**(4): p. 937-949.
251. L'Abee-Lund, T.M. and H. Sorum, *Functional Tn5393-like transposon in the R plasmid pRAS2 from the fish pathogen Aeromonas salmonicida subspecies salmonicida isolated in Norway*. Applied and Environmental Microbiology, 2000. **66**(12): p. 5533-5535.
252. Chiou, C.S. and A.L. Jones, *Nucleotide sequence analysis of a transposon (Tn5393) carrying streptomycin resistance genes in Erwinia-amylovora and other gram negative bacteria*. Journal of Bacteriology, 1993. **175**(3): p. 732-740.
253. Ren, Y., et al., *Complete Genome Sequence of Enterobacter cloacae subsp cloacae Type Strain ATCC 13047*. Journal of Bacteriology, 2010. **192**(9): p. 2463-2464.
254. Kucerova, E., et al., *Genome Sequence of Cronobacter sakazakii BAA-894 and Comparative Genomic Hybridization Analysis with Other Cronobacter Species*. Plos One, 2010. **5**(3): p. A51-A60.
255. Taylor, D.E., *Bacterial tellurite resistance*. Trends in Microbiology, 1999. **7**(3): p. 111-115.
256. Bradley, D.E., Taylor, D. E., Levy, S. B., Cohen, D. R., Brose, E. C. and Whelan, J., *PIN32 - a cointegrate plasmid with incHI2 and incFII components*. Journal of General Microbiology, 1986. **132**: p. 1339-1346.
257. Vliegenthart, J.S., P.A.G. Ketelaarvangaalen, and J.A.M. Vandeklundert, *Nucleotide sequence of the aacC2 gene, a gentamicin resistance determinant involved in a hospital epidemic of multiply resistant members of the family Enterobacteriaceae*. Antimicrobial Agents and Chemotherapy, 1989. **33**(8): p. 1153-1159.
258. Frana, T.S., S.A. Carlson, and R.W. Griffith, *Relative distribution and conservation of genes encoding aminoglycoside-modifying enzymes in Salmonella enterica serotype Typhimurium phage type DT104*. Applied and Environmental Microbiology, 2001. **67**(1): p. 445-448.
259. Brau, B. and W. Piepersberg, *Cointegrational transduction and mobilization of gentamicin resistance plasmid pWP14A is mediated by IS140*. Molecular & General Genetics, 1983. **189**(2): p. 298-303.
260. Allmansberger, R., B. Brau, and W. Piepersberg, *Genes for gentamicin-(3)-n-acetyl-transferase-III and transferase-IV 2 Nucleotide sequences of 3 aac(3)-III genes and evolutionary aspects*. Molecular and General Genetics, 1985. **198**(3): p. 514-520.
261. Barton MD, P.R., Hart WS., *Antibiotic resistance in animals*. Commun Dis Intell, 2003. **27**: p. S121-126.
262. Maignan, S., Guilloteau, J. P., Zhou-Liu, Q., Clement-Mella, C. and Mikol, V., *Crystal structures of the catalytic domain of HIV-1 integrase free and complexed with its metal cofactor: High level of similarity of the active site with other viral integrases*. Journal of Molecular Biology, 1998. **282**(2): p. 359-368.
263. Hall, B.G., *Chromosomal mutation for citrate utilization by Escherichia coli K12*. Journal of Bacteriology, 1982. **151**(1): p. 269-273.
264. Matsutani, S. and E. Ohtsubo, *Complete sequence of IS629*. Nucleic Acids Research, 1990. **18**(7): p. 1899-1899.

265. Matsutani, S., Ohtsubo, H., Maeda, Y. and Ohtsubo, E., *Isolation and characterization of IS-elements repeated in the bacterial chromosome*. Journal of Molecular Biology, 1987. **196**(3): p. 445-455.
266. Matson, S.W., *Escherichia coli dna helicase-II (uvrD gene-product) catalyzes the unwinding of DNA RNA hybrids invitro*. Proceedings of the National Academy of Sciences of the United States of America, 1989. **86**(12): p. 4430-4434.
267. Daniels, D.L., Plunkett, G., Burland, V. and Blattner, F. R., *Analysis of the Escherichia coli genome - DNA sequence of the region from 84.5 to 86.5 minutes*. Science, 1992. **257**(5071): p. 771-778.
268. Iida, S., Mollet, B., Meyer, J. and Arber, W., *Functional-characterization of the prokaryotic mobile genetic element IS26*. Molecular and General Genetics, 1984. **198**(1): p. 84-89.
269. Trieucuot, P. and P. Courvalin, *Transposition behavior of IS15 and its progenitor IS15-delta - are cointegrates exclusive end products*. Plasmid, 1985. **14**(1): p. 80-89.
270. Miriagou, V., Carattoli, A., Tzelepi, E., Villa, L., Tzouvelekis, L. S., *IS26-Associated In4-type integrons forming multiresistance loci in enterobacterial plasmids*. Antimicrobial Agents and Chemotherapy, 2005. **49**(8): p. 3541-3543.
271. Stokes, H.W., Nesbo, C. L., Holley, M., Bahl, M. I., Gillings, M. R., Boucher, Y., *Class 1 integrons potentially predating the association with Tn402-like transposition genes are present in a sediment microbial community*. Journal of Bacteriology, 2006. **188**(16): p. 5722-5730.
272. Leverstein-van Hall, M.A., et al., *Presence of integron-associated resistance in the community is widespread and contributes to multidrug resistance in the hospital*. J Clin Microbiol, 2002. **40**(8): p. 3038-3040.
273. Gebreyes, W.A. and C. Altier, *Molecular characterization of multidrug-resistant Salmonella enterica subsp. enterica serovar Typhimurium isolates from swine*. J Clin Microbiol, 2002. **40**(8): p. 2813-2822.
274. Ng, L.K., Martin, I., Alfa, M., Mulvey, M., *Multiplex PCR for the detection of tetracycline resistant genes*. Molecular and Cellular Probes, 2001. **15**(4): p. 209-215.
275. Boyd, D., Cloeckaert, A., Chaslus-Dancla, E., Mulvey, M. R., *Characterization of variant Salmonella genomic island 1 multidrug resistance regions from serovars Typhimurium DT104 and Agona*. Antimicrobial Agents and Chemotherapy, 2002. **46**(6): p. 1714-1722.
276. Riley, J., et al., *A novel, rapid method for the isolation of terminal sequences from yeast artificial chromosome (YAC) clones*. Nucleic Acids Research, 1990. **18**(10): p. 2887-2890.
277. Pacheco, A.B., Guth, B. E., Soares, K. C., Nishimura, L., DeAlmeida, D., Ferreira, L. C., *Random amplification of polymorphic DNA reveals serotype-specific clonal clusters among enterotoxigenic Escherichia coli strains isolated from humans*. Journal of Clinical Microbiology, 1997. **35**(6): p. 1521-1525.
278. Liebert, C.A., Wireman, J., Smith, T., Summers, A. O., *Phylogeny of mercury resistance (mer) operons of gram-negative bacteria isolated from the fecal flora of primates*. Applied and Environmental Microbiology, 1997. **63**(3): p. 1066-1076.

APPENDIX 1

CDS zone sizes

Table A1.1 CDS zones sizes from IncHI2-containing HPA strains

Isolate	Inhibition zone in annular radius (mm)[1]									
	Ap[2]	**Cm**	**Sm**	**Sp**	**Su**	**Tc**	**Gm**	**Km**	**Nm**	**Tp**
28R823(TP116) (1)	6	0[3]	3	1	0	7	8.5	7	7	8
28R823(TP116) (2)	6.5	0	4	1	0	8	8.5	7.5	7.5	9
31R872 (1)	6	0	1	1	0	6	8	7	7	8.5
31R872 (2)	6.5	0	1	1.5	2	6.5	8.5	7	7	9
33R944 (1)	6	7	3	6	9	7	8	0.5	4	8
33R944 (2)	6.5	7	3.5	6	10	7	8.5	1	4	9
E. coli 294 (1)	13	13	11.5	8	20	12	15	15	13	18
E. coli 294 (2)	12	11	11	6.5	17	11	13	14.5	14	16

[1]Colonies that have an annular radius of <6mm are regarded as resistant to the above antibiotic
[2]Ap: ampicllin; Cm: chloramphenicol; Sm: streptomycin; Sp: spectinomycin; Su: sulphamethoxazole; Tc: tetracycline; Gm: gentamicin; Km: kanamycin; Nm: neomycin; Tp: trimethoprim;
[3]Yellow highlighting denotes a resistant phenotype

Table A1.2 CDS zones sizes from SRC27 and its transconjugants

Isolate	Inhibition zone in annular radius (mm)[1]									
	Ap[2]	**Cm**	**Sm**	**Sp**	**Su**	**Tc**	**Gm**	**Km**	**Nm**	**Tp**
SRC27 (1)	0	0	0	0	0	0	1	0	2	0
SRC27 (2)	0	0	0	0	0	0	0	0	1	0
E294/pSRC27-H/pSRC27-I (1)	0	0	0	2	0	0	4	0	4	0
E294/pSRC27-H/pSRC27-I (2)	0	0	0	3	0	0	3	3	4	0
pSRC27-I (1)	14[3]	14	1	6	13	12	12	13	12	18
pSRC27-I (2)	10.5	10	2	6	9	9	9.5	10	9	14
pSRC27-H (1)	1.5	0	2.5	2	0	0.5	5	1.5	5	0
pSRC27-H (2)	1	0	2.5	2	0	0.5	5	2	5.5	0
E. coli 294 (1)	13	13	11.5	8	20	12	15	15	13	18
E. coli 294 (2)	12	11	11	6.5	17	11	13	14.5	14	16

[1]Colonies that have an annular radius of <6mm are regarded as resistant to the above antibiotic
[2]Ap: ampicllin; Cm: chloramphenicol; Sm: streptomycin; Sp: spectinomycin; Su: sulphamethoxazole; Tc: tetracycline; Gm: gentamicin; Km: kanamycin; Nm: neomycin; Tp: trimethoprim;
[3]Blue highlighting denotes a sensitive phenotype

Table A1.3 CDS zone sizes[1] [2] of SRC27 transconjugants and variants

Strain	Sample	Ap[3]	Cm	Sm	Sp	Su	Tc	Gm	Km	Nm	Tp
S. enterica	27 Sal I	0	0	0	0	0	0	1	0	2	0
Parent I[4]	27 (*E. coli*) (1)	0	0	0	2	0	0	4	0	4	0
Parent II[4]	27 (*E. coli)* (2)	0	0	0	3	0	0	3	3	4	0
E. coli	*E. coli* 294	13	13	11.5	8	20	13	15	15	13	18
Variant 1	27I D14-2[5]	0	0	1	0	0	0	3	9	8.5	0
2	27I D7-1	0	0	2	0	6	0	1	8.5	8	0
2	27II D14-5	0	0	1.5	0	8	0	1.5	8	8	0
2	27I D7-2	0	0	2	0	10	0	1	7.5	7.5	0
2	27I D14-3	0	0	1	0	7.5	0	1	9	8	0
2	27I D14-6	0	0	1	0	10	0	1	8	7	0
2	27I D14-7	0	0	1.5	0	8	0	1	8	8	0
2	27I D14-8	0	0	1	0	10	0	1	7	7	0
2	27I D14-11	0	0	1	0	9	0	1	7.5	7.5	0
2	27I D14-12	0	0	1.5	0	10	0	0.5	8	7	0
2	27I D14-15	0	0	1.5	0	9	0	1	8	7.5	0
2	27II D2-1	0	0	1	0	8	0	1	9	8	0
2	27II D14-6	0	0	1	0	9	0	1	8	8	0
3	27I D14-5	0	0	0	5.5	0	8	1	0	2.5	12
3	27I D14-9	0	0	0.5	5	0	5	1	0	2	11
3	27II D7-1	0	0	0	5	0	5	1	0	2	11
4	27II D7-3	0	0	0	5	0	0	8	0	2	12
4	27II D14-2	0	0	0	5.5	0	0	8	0	2	12
4	27II D7-6	0	0	0	6	0	0	9	0	3	15
4	27II D14-1	0	0	0	5	0	0	9	0	2	13
4	27II D14-7	0	0	0	5	0	0	8	0	2	12
4	27II D14-9	0	0	0	5	0	0	8	0	2.5	12.5
4	27II D14-11	0	0	0	5	0	0	8	0	2	12
5	27I D14-14	11	10	2	0	6	0	8	9	8.5	0
5	27II D14-4	11	9	1.5	0	6	0	8.5	9	8	0
5	27II D14-8	11	10	2	0	8	0	9	9	8.5	0
6	27II D14-13	11.5	0	1.5	5.5	8	0	10.5	9.5	8	13
6	27II D14-14	11.5	0	2	5.5	8	0	9	9.5	8	12
7	27II D7-2	11	0	1.5	5	6	6	8	8	8	12
7	27II D1-1	11	0	1	5	10	7	8	8.5	7.5	12
8	27II D14-3	11	8	1.5	5	8	0	7.5	9	8.5	12
9	27I D14-1	10.5	10	0	6	9	9	9.5	10	9	14
9	27I D14-4	10	9	1	5.5	6	7	8	9	8	12
9	27II D14-10	11	9.5	0.5	5	9	8	9	8	7.5	11
9	27I D14-10	12	10.5	1	5.5	9	8	9	9	8	13.5
9	27I D14-13	14	14	1	6	10	12	12	13	12	18
9	27II D7-4	11	12	0	6	9.5	10	10.5	11	9	14
9	27II D7-5	11	10	1	5	9	8	8.5	9	8	13

[1] Inhibition zone in annular radius (mm)

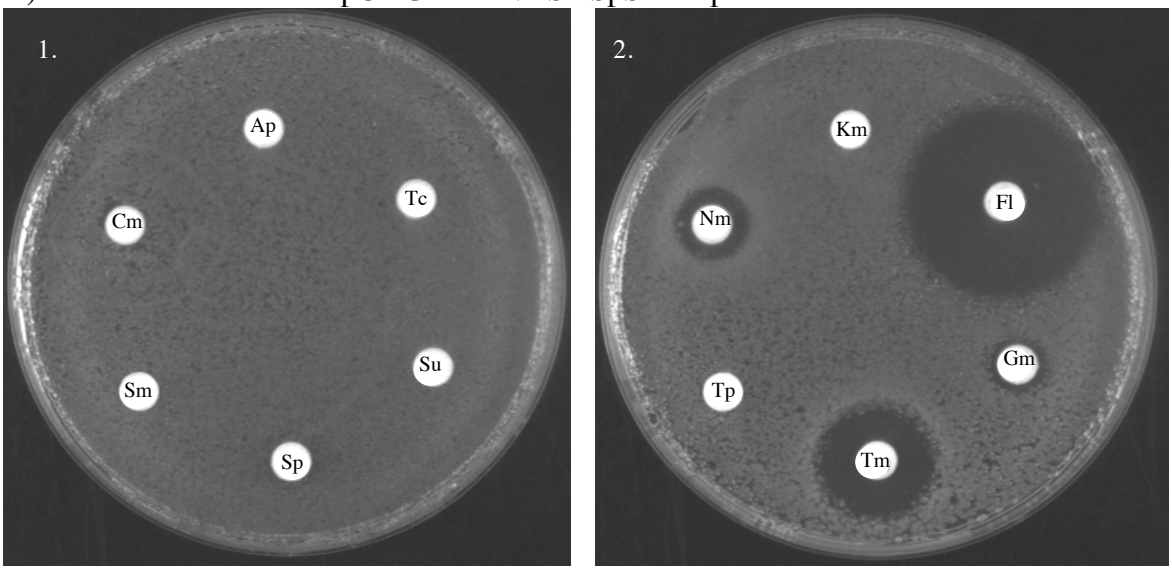
[2] Colonies that have an annular radius of <6mm are regarded as resistant to the above antibiotic

[3] Tobramycin and florfenicol also tested and all were sensitive

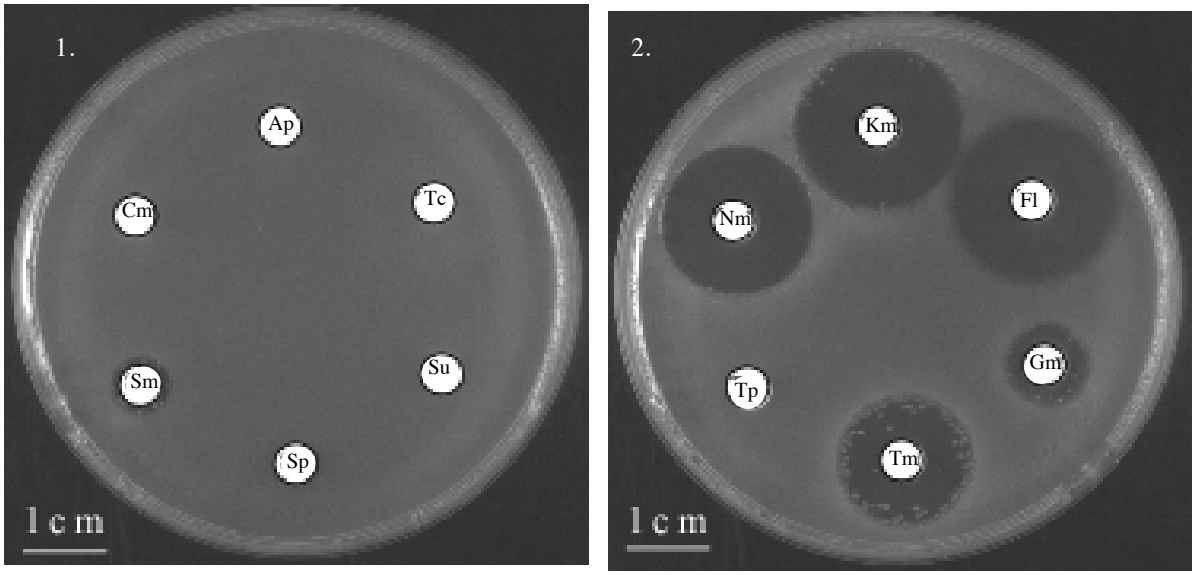
[4]Parent for initial inoculation of cultures I and II

[5]"27I D14-2" denotes a E294/pSRC27-H/pSRC27-I colony recovered from culture I, after 14 days growth and was 2nd colony recovered from that day and culture, to have a varied phenotype, and so on.

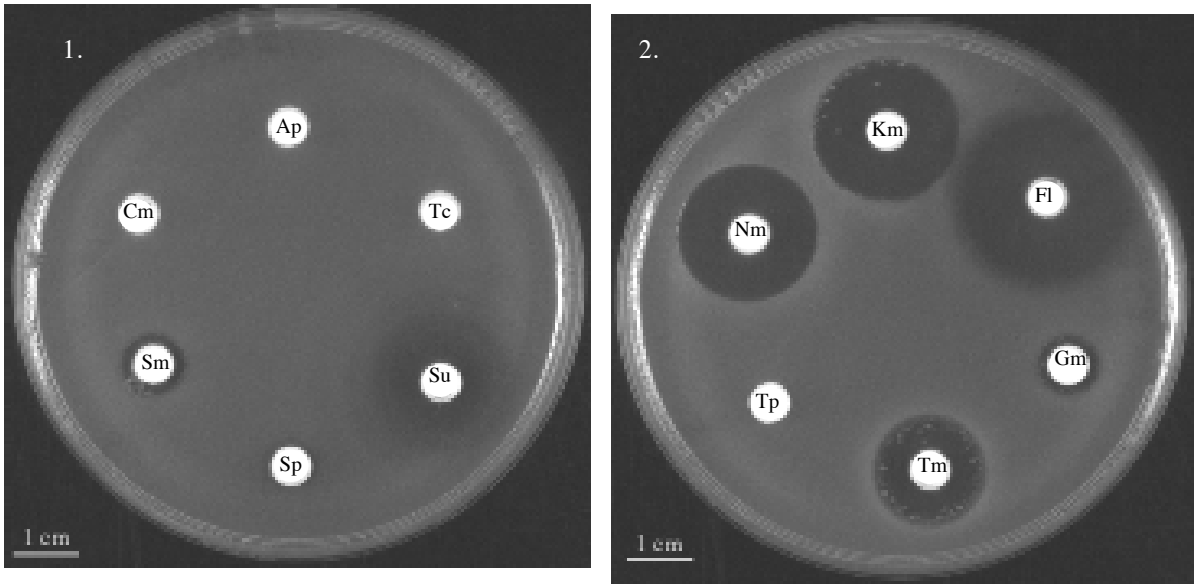
A) Parent – Resistant to ApCmGmKmNmSmSpSuTcTp



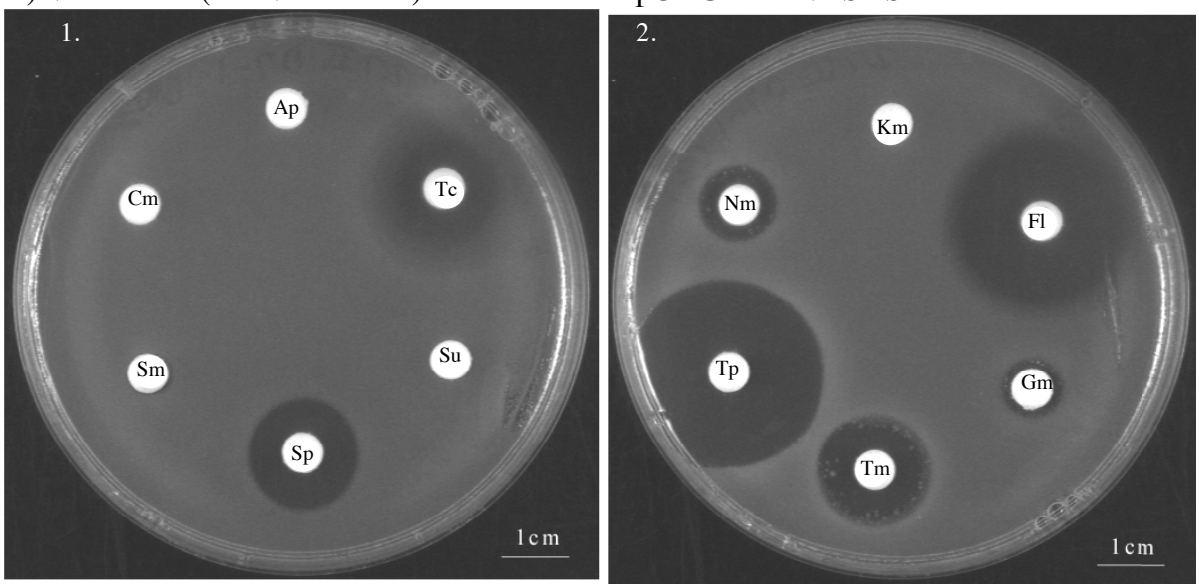
B) Variant 1– Resistant to ApCmGmSmSpSuTcTp



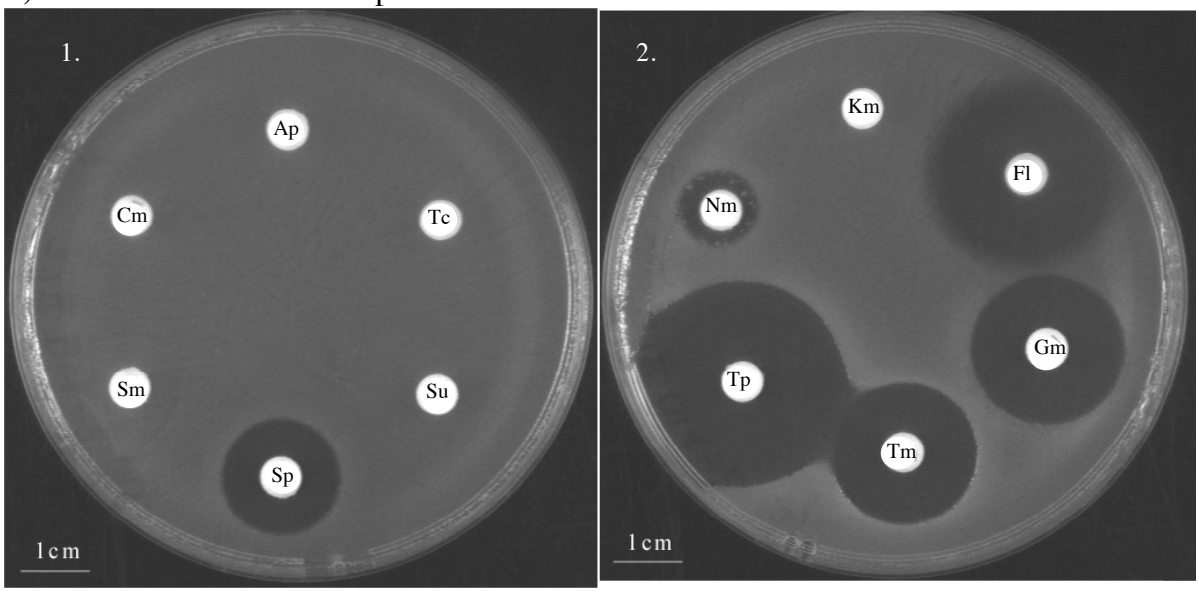
C) Variant 2– Resistant to ApCmGmSmSpTcTp



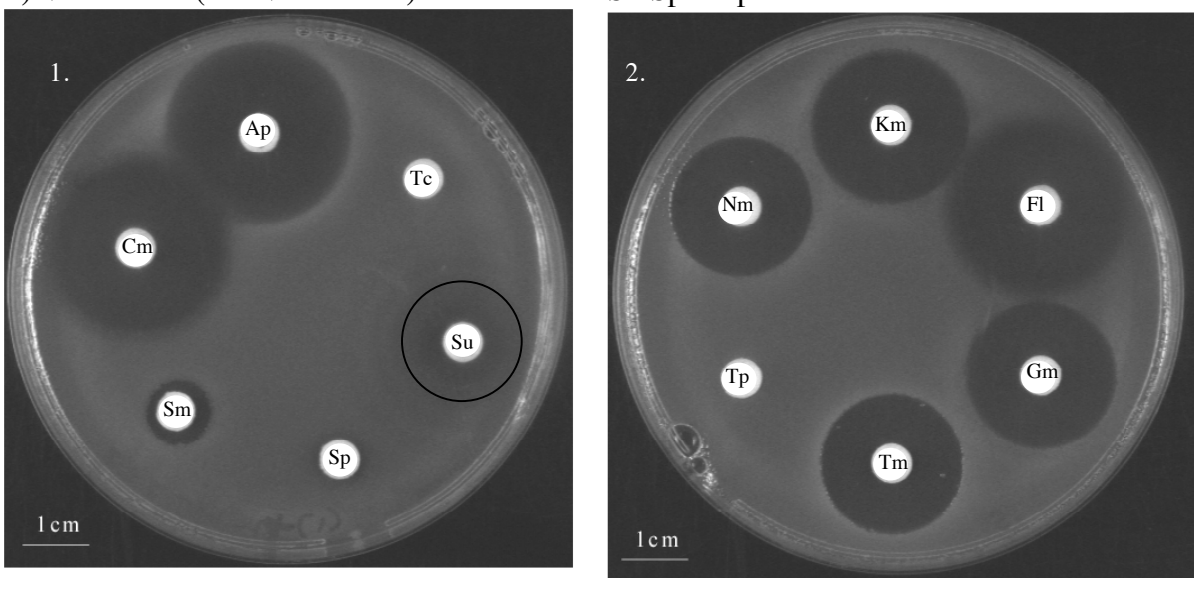
D) Variant 3-2 (and Variant 3-1) – Resistant to ApCmGmKmNmSmSu



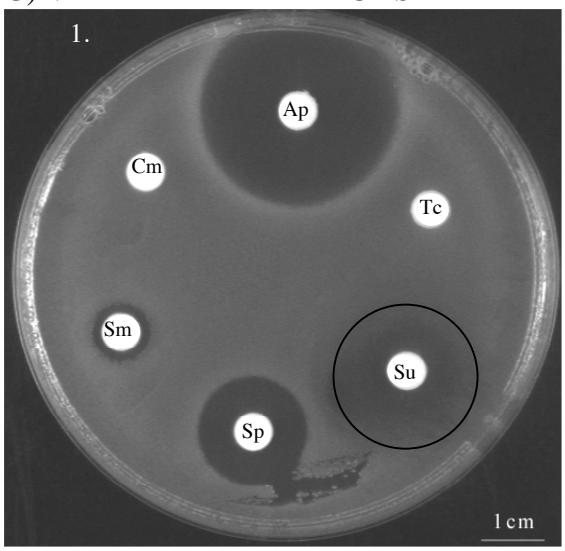
E) Variant 4 – Resistant to ApCmKmNmSmSuTc

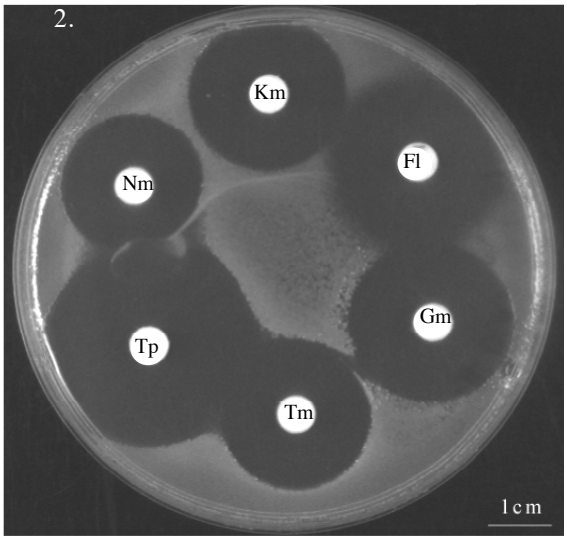


F) Variant 5-2 (and Variant 5-1) – Resistant to SmSpTcTp

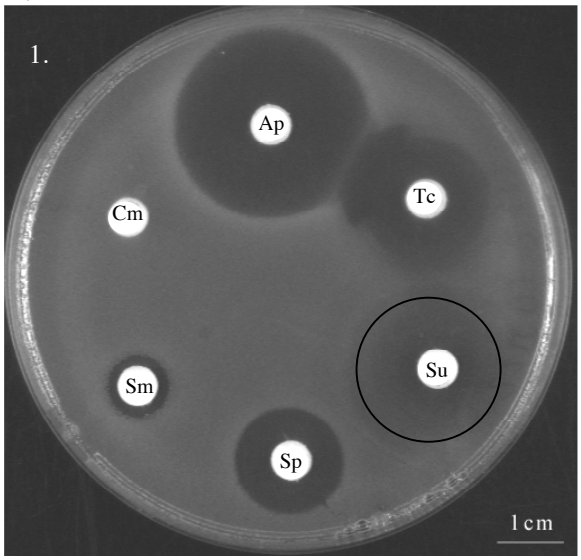


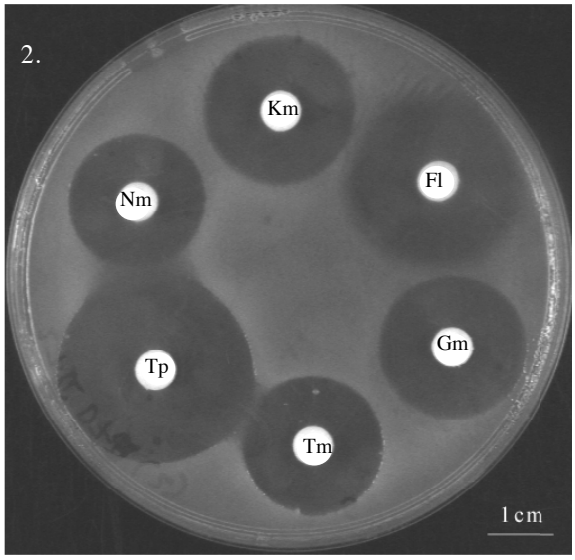
G) Variant 6– Resistant to CmSmTc



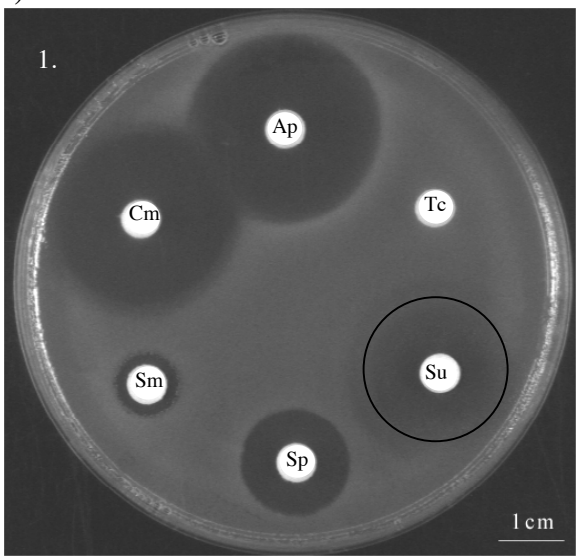


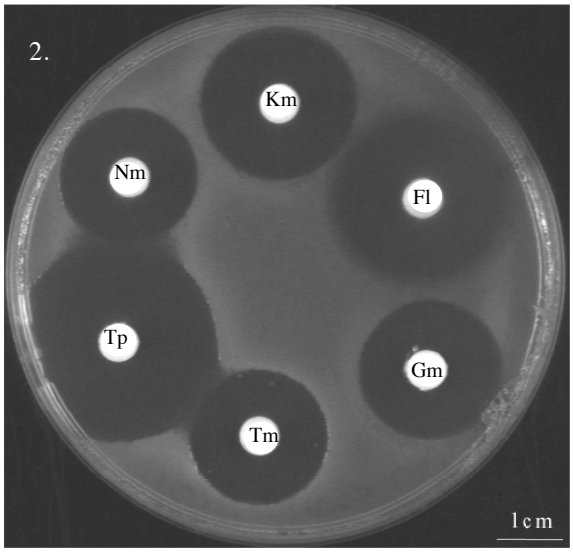
H) Variant 7– Resistant to CmSm



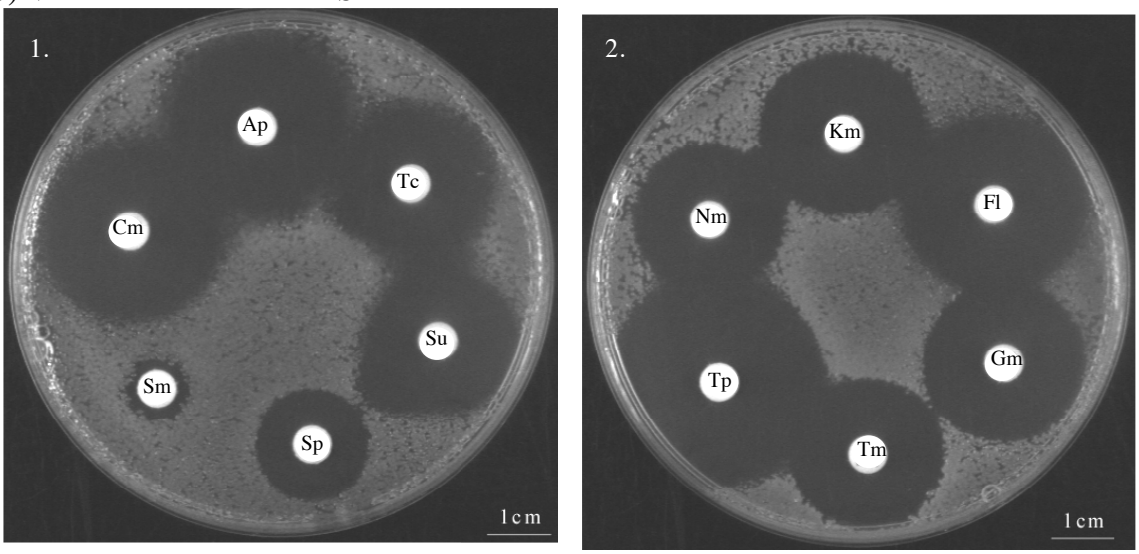


I) Variant 8– Resistant to SmTc





J) Variant 9 – Resistant to Sm



K) *E. coli* strain E294 – sensitive to all (NB Tc not shown but is sensitive)

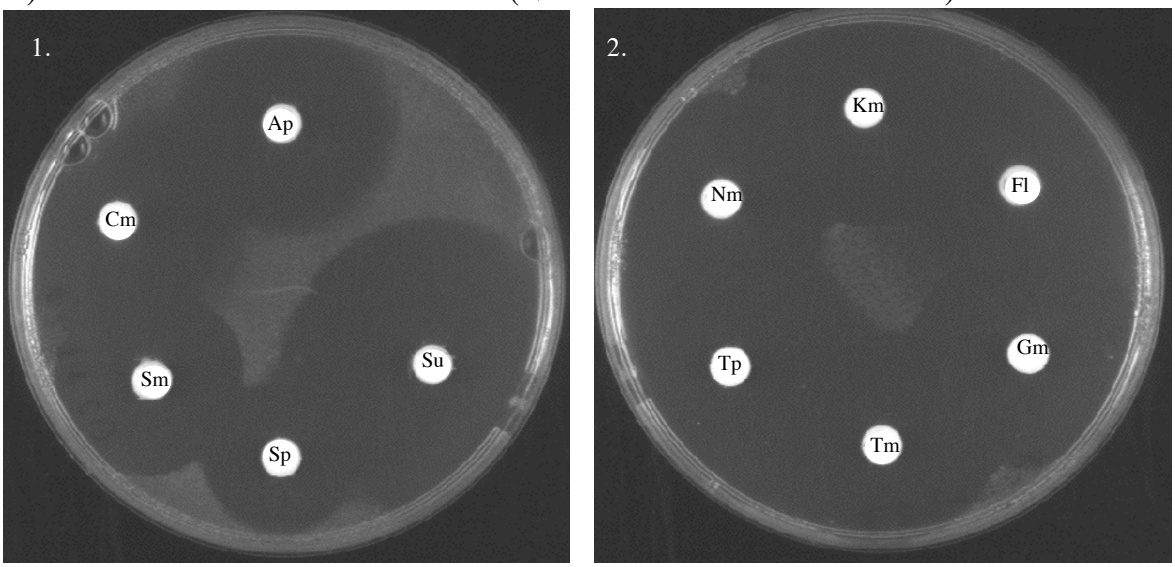
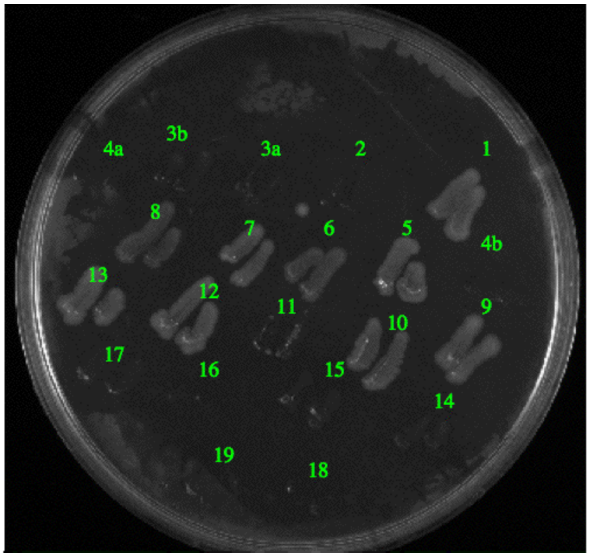


Fig A1.1 CDS plates of E294/SRC27-H/pSRC27-I transconjugant parent and variants. A) is the parent strain and B) –J) represent Variants 1 – 9 respectively and K) the E294 host strain. Plate 1. contains the antibiotics (in order, starting from the top going clockwise): Ap, Tc, Su, Sp, Sm, Cm and Plate 2. contains Km, florfenicol (Fl), Gm, tobramycin (Tm), Tp and Nm. In some cases with the antibiotic Su, the inhibition zone is unclear when printed and so the edge clearing has been emphasised with a black circle.

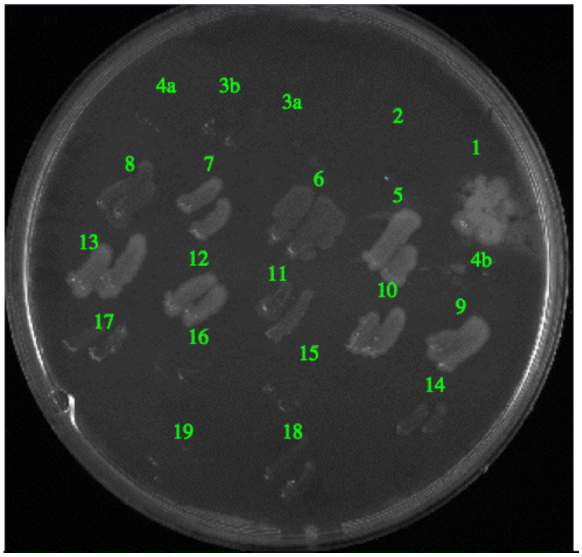
APPENDIX 2

Heavy metal resistance phenotypes

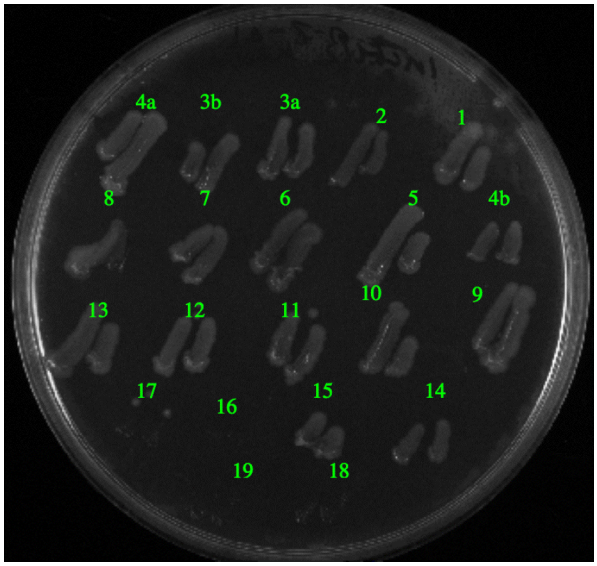
LA + 5 mM Sodium Arsenite



LA + 20 mM Sodium Arsenate



LA + 20 μg/ml Mercuric chloride



HBA + 10 μg/ml potassium tellurate

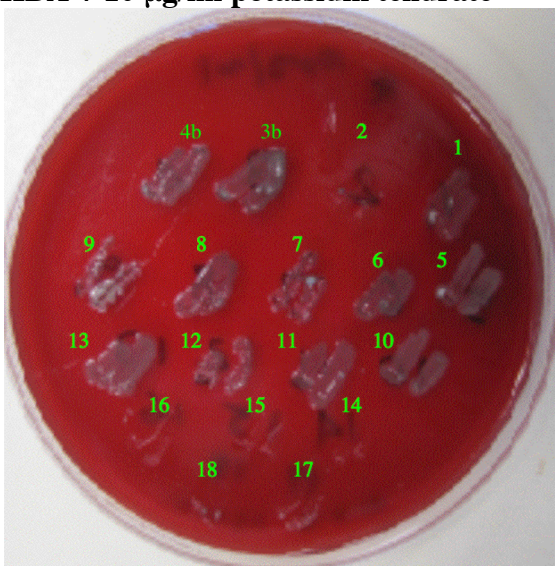


Fig A2.1 Pictures of heavy metal resistance test plates. The composition of the plate is stated above, where LA is Luria broth agar and HBA is horse blood agar. The numbers represent the strains tested and plasmids in the strains are in brackets. 1. E294 (R478) 2. E294 (TP116) 3a. SRC125 (pSRC125) 3b. E294 (pSRC125) 4. SRC26 (pSRC26) 4b. E294 (pSRC26) 5. SRC83 (pSRC83) 6. SRC70 (pSRC70) 7. SRC71 (pSRC71) 8. SRC72 (pSRC72) 9. SRC92 (pSRC92) 10. SRC93 (pSRC93) 11. SRC94 (pSRC94) 12. SRC95 (pSRC95) 13. SRC96 (pSRC96) 14. *E. coli* (pACYC184::Tn21) 15. *E. coli* (R388::Tn*1696)* 16. *E. coli* E294 17. *E. coli* DH5α 18. *E. coli* UB5201 (Nx) 19. *E. coli* UB1637 (Rf)

APPENDIX 3

PCR primers

Table A3.1 Primers targeting resistance genes, IS, integrons and general primers

Location	Name	Sequence 5' - 3'	Product size (bp)	Annealing temp (^{O}C)	GenBank accession no.	Reference[1]
Class 1 integrons						
intI1	HS463a	CTGGATTTCGATCACGGCACG	473	60	U12338	[267]
	HS464	ACATGCGTGTAAATCATCGTCG			U12338	[267]
5'-CS	L1	GGCATCCAAGCAGCAAGC	variable	57	U12338	[268]
3'-CS	R1	AAGCAGACTTGACCTGAT			U12338	[227]
Resistance genes						
sul1	HS549	ACTAAGCTTGCCCCTTCCGC	1103	65	U12338	[267]
	HS550	CTAGGCATGATCTAACCCTCG			U12338	[267]
sul2	sulII-F	GGCAGATGTGATCGACCTCG	405	60	M28829	[269]
	sulII-R	ATGCCGGGATCAAGGACAAG			M28829	[269]
s*trA*	strA-F	CTTGGTGATAACGGCAATTC	548	53	M95402	[270]
	strA-R	CCAATCGCAGATAGAAGGC			M95402	[270]
strB	strB-F	ATCGTCAAGGGATTGAAACC	509	53	M95402	[270]
	strB-R	GGATCGTAGAACATATTGGC			M95402	[270]
aphA1	aphAI-F	AAACGTCTTGCTCGAGGC	460	56	AY123253	[254]
	aphAI-R	CAAACCGTTATTCATTCGTGA			AY123253	[254]
aacC2	aacC2-F	GGCAATAACGGAGGCAATTCGA	680	60	X51534	[254]
	aacC2-R	CTCGATGGCGACCGAGCTTCA			X51534	[254]
tetA(B)	tet(B)	TTGGTTAGGGGCAAGTTTTG	659	55	J01830	[271]
	tet(B)	GTAATGGGCCAATAACACCG			J01830	[271]
catA1	RH513	GTACTGTTGTAATTCATTAAGCAT	595	60	AY123253	V. Post[2]
	RH514	TCCCAATGGCATCGTAAAGAACA			AY123253	V. Post
bla_{TEM}	RH605	TTTCGTGTCGCCCTTATTCC	692	60	AJ851089	this study
	RH606	CCGGCTCCAGATTTATCAGC			AJ851089	this study
Insertion sequences						
IS*26*	IS26-F	ACCTTTGATGGTGGCGTAAG	617	60	AY123253	this study
	RH601	GATGGAGCTGCACATGAACC			AY123253	this study
IS*1*	RH517	GTGTTTTATGTTCAGATAATGCC	493	60	AY123253	V. Post
	RH518	GCCGGGCAACCGCCCGCATTATG			AY123253	V. Post
IS*10*	RH524	CACGACTCTTTACCAATTCTG	1103	60	AF071413	V. Post
	RH525	GTGTAGCCAGAATGCCGCAA			AF071413	V. Post
IS*6100*	IS6100-Rv2	AATGGTGGTTGAGCATGCC	356	60	U12338	[268]
	DB-T1	TGCCACGCTCAATACCGAC			U12338	[272]
ISCfr1	RH629	AGCGGTAAATGACGGTTTTG	563	60	AF550415	this study
	RH630	AGATGAGGTCGATTGGAACG			AF550415	this study
***strA* position**			size with strA-R:			
s*trA*	strA-R	CCAATCGCAGATAGAAGGC			M95402	[270]
RH602	$tnpR_{5393}$	TCAACACGGTGAAGGAGCTG	1,234	60	AF313472	this study
RH603	*sul2*	ACTTCATCCGCACACACGAG	929	60	AJ851089	this study
RH604	IS*1133*	AATCCACGCCGAAGACATTC	958	60	M95402	this study

Table A3.1 cont.

Name	Sequence 5'-3'			Reference[1]
Vectorette				
V-F	GAGAGGGAAGAGAGCAGGCAAGGAATGGAAGCTGTCTGTCGCAGGAGAGGA			[273]
V-R	GACTCTCCCTTCTCGAATCGTAACCGTTCGTACGAGAATCGCTGTCCTCTCCTT			[273]
224	CGAATCGTAACCGTTCGTACGAGAATCGCT		-	[273]
RAPD				
1290	GTGGATGCGA	-	-	[274]

[1]reference for primers
[2]in the Hall laboratory

Table A3.2 Primers used in PBRT[1]

Replicon	Direction	Sequence 5' - 3'	Product size (bp)
Panel 1			
B/O	F	GCGGTCCGGAAAGCCAGAAAAC	159
	R	TCTGCGTTCCGCCAAGTTCGA	
FIC	F	GTGAACTGGCAGATGAGGAAGG	262
	R	TTCTCCTCGTCGCCAAACTAGAT	
A/C	F	GAGAACCAAAGACAAAGACCTGGA	465
	R	ACGACAAACCTGAATTGCCTCCTT	
P	F	CTATGGCCCTGCAAACGCGCCAGAAA	534
	R	TCACGCGCCAGGGCGCAGCC	
T	F	TTGGCCTGTTTGTGCCTAAACCAT	750
	R	CGTTGATTACACTTAGCTTTGGAC	
Panel 2			
K/B	F	GCGGTCCGGAAAGCCAGAAAAC	160
	R	TCTTTCACGAGCCCGCCAAA	
W	F	CCTAAGAACAACAAAGCCCCCG	242
	R	GGTGCGCGGCATAGAACCGT	
FIIA	F	CTGTCGTAAGCTGATGGC	270
	R	CTCTGCCACAAACTTCAGC	
FIA	F	CCATGCTGGTTCTAGAGAAGGTG	462
	R	GTATATCCTTACTGGCTTCCGCAG	
FIB	F	GGAGTTCTGACACACGATTTTCTG	702
	R	CTCCCGTCGCTTCAGGGCATT	
Y	F	AATTCAAACAACACTGTGCAGCCTG	765
	R	GCGAGAATGGACGATTACAAAACTTT	
Panel 3			
I1	F	CGAAAGCCGGACGGCAGAA	139
	R	TCGTCGTTCCGCCAAGTTCGT	
Frep	F	TGATCGTTTAAGGAATTTTG	270
	R	GAAGATCAGTCACACCATCC	
X	F	AACCTTAGAGGCTATTTAAGTTGCTGAT	376
	R	TGAGAGTCAATTTTTATCTCATGTTTTAGC	
HI1	F	GGAGCGATGGATTACTTCAGTAC	471
	R	TGCCGTTTCACCTCGTGAGTA	
N	F	GTCTAACGAGCTTACCGAAG	559
	R	GTTTCAACTCTGCCAAGTTC	
HI2	F	TTTCTCCTGAGTCACCTGTTAACAC	644
	R	GGCTCACTACCGTTGTCATCCT	
L/M	F	GGATGAAAACTATCAGCATCTGAAG	785
	R	CTGCAGGGGCGATTCTTTAGG	

[1] Primer reference: [137]; panel grouping reference:[236]

Table A3.3 Primers for mapping Tn*6029* and Tn*4352*

Linkage	Primer[1,2]	Sequence	Tm (°C)[3]	size (kb)
Linkage out LHS	RH641	ACGAGGCCCTTTCGTCTTC	66	variable
IS*26*/ bla_{TEM}	IS26-F	ACCTTTGATGGTGGCGTAAG	64	2
	RH606	CCGGCTCCAGATTTATCAGC	66	
bla_{TEM}/ *repA*	RH605	TTTCGTGTCGCCCTTATTCC	66	2.7
	RH643	ATCTGTCACCAGTGCGAAGG	66	
IS*26*/ *repC*	RH601	GATGGAGCTGCACATGAACC	66	2.2
	RH644	CTCTGACGGCCAGACATAGC	65	
repC/ *strA*	RH645	CTGGGCCTAATGGCCTAGTG	66	2.1
	RH646	TCCTCCTGCCAGTTGATGAC	65	
sul2/ IS*26*	RH603	ACTTCATCCGCACACACGAG	62	2
	RH882	GATGCGTGCACTACGCAAAG	67	
strB/aphA1	RH647	CGATCCTAGACGCATTGCAC	66	1.6
	RH881	ATTCGTGATTGCGCCTGAG	66	
aphA1/ IS*26*	RH880	CAACGGGAAACGTCTTGCTC	67	1.7
	RH601	GATGGAGCTGCACATGAACC	66	
Linkage out RHS	RH880	CAACGGGAAACGTCTTGCTC	67	variable

[1]primers designed using GenBank accession no. AJ851089
[2]all primers designed in this study
[3]all PCR reaction performed with an annealing temperature of 60°C

Table A3.4 Primers to map Tn*10*

PCR	Primer[1,2]	Sequence	Tm (°C) [3]	Location	size (bp)
1	RH611	AAATGGTTTCGCCAAAAATC	62	IS*10*	3,441
	RH947	CGTATTCGTTTTCCCACCAA	64	*jemB*	
2	RH526	CGTTGCTTATCGAAGTGTTAAT	65	*jemB*	3,091
	RH860	AGGTAAAGCGATCCCACCAC	66	*tetA*(B)	
3	RH859	CACTTTGGCGTATTGCTTGC	65	*tetA*(B)	3,490
	RH611	AAATGGTTTCGCCAAAAATC	62	IS*10*	

[1]Primers designed using GenBank accession no. AF250878
[2]all primers designed in this study
[3]all reactions performed at 60°C annealing temperature

Table A3.5 Primers to map all Tn*501*-like mercury resistance transposons

Primer[1]	Sequence	Location	Tm (°C)	Reference
***tnp* module**				
RH665	GGTCGGAGATRTGGGTGTAG	*tnpA*	60	J. Pinyon[2]
RH666	CTGACCAARATGGCCGAGTC	*tnpA*	60	J. Pinyon
RH667	CTGGAACTGCTGCTGATGC	*tnpA*	60	J. Pinyon
RH668	CACCAGAACCGCCTGCTCAA	*tnpA*	60	J. Pinyon
RH670	CTTCGACCAGAACCCGGAAC	*tnpR*	60	J. Pinyon
intI1/tnpR				
RH546	GGGCATGGTGGCTGAAGGACC	*intI1*	63	V. Post[2]
RH620	GTTCCGGGTTCTGGTCGAAG	*tnpR*	65	this study
IS*6100/merE*				
RH521	CATATCGTGCGCGACTACCTG	*merE*	63	V. Post
DB-T1	TGCCACGCTCAATACCGAC	IS*6100*	64	[272]
***mer* module**				
merA1	ACCATCGGCGGCACCTGCGT	*merA*	58	[275]
merA5	ACCATCGTCAGGTAGGGGAACAA	*merA*	57	[275]
RH659	GGGCTACAGCGARGCGGAAGC	*merA*	60	J. Pinyon
RH660	CGCCCCACARGTAGCCGG	*merE*	60	J. Pinyon
RH655	GTCGCCTTGGTCARCTTCTG	*merP*	60	J. Pinyon
RH657	CGGAAAATAAAGYACGCTAAG	IR_{mer}	60	J. Pinyon
RH671	GTCAGGTAGGGGAACAACTG	*merA*	60	J. Pinyon
RH672	GTGTGCCGTCCAAGATCATG	*merA*	60	J. Pinyon

[1]primers designed using GenBank Accession nos Z00027, X61367, L27758 and U12338
[2]in the Hall laboratory

Table A3.6 Primers to map resistance regions of pSRC125

Linkage	Primer[1]	Sequence	Annealing temp (°C)	size (kb)	Reference[2]
Tn*6025*					
IR_{tnp}/*tnpA*	RH657	CGGAAAATAAAGYACGCTAAG	60	1.1	J. Pinyon[3]
	RH666	CTGACCAARA TGGCCGAGTC			J. Pinyon
tnpA	RH665	GGTCGGAGATRTGGGTGTAG	60	1.4	J. Pinyon
	RH668	CACCAGAACCGCCTGCTCAA			J. Pinyon
tnpA/tnpR	RH667	CTGGAACTGCTGCTGATGC	60	1.5	J. Pinyon
	RH670	CTTCGACCAGAACCCGGAAC			J. Pinyon
tnpR/3'-CS	RH669	CCATGCTRTGCACCACCAC	60	2.6	J. Pinyon
	HS549	ACTAAGCTTGCCCCTTCCGC			[267]
3'-CS/IS*6100*	RH459	CCACAACGAGCTGATGTACTGCC	58	3	[267]
	IS6100-RV2	AATGGTGGTTGAGCATGCC			[268]
IS*6100*/*merA*	DB-TI	TGCCACGCTCAATACCGAC	60	1.8	[272]
	RH659	GGGCTACAGCGARGCGGAAGC			J. Pinyon
merA	RH671	GTCAGGTAGGGGAACAACTG	60	1.2	J. Pinyon
	RH672	GTGTGCCGTCCAAGATCATG			J. Pinyon
merA/merT	RH361	GATGCCTTCGTACTTGG	60	1.2	Hall Lab
	RH448	CGGTGTTGGAACCCTATCG			Hall Lab
merP/IR_{mer}	RH655	GTCGCCTTGGTCARCTTCTG	60	1.2	J. Pinyon
	RH657	CGGAAAATAAAGYACGCTAAG			J. Pinyon
Tn*5393*e					
IR_{5393}/$tnpA_{5393}$	RH1257	GGGCGGAATCCTACGCTAAG	60	2.9	this study
	RH898	CTCAACCAGCTTCATAAACAAAG			this study
IS*26/aphA1*	IS26-F	ACCTTTGATGGTGGCGTAAG	58	1.2	this study
	aphA1-R	CAAACCGTTATTCATTCGTGA			[254]
aphA1/strA	RH387	ATCCTATGGAACTGCCTCGG	57	2.2	Hall Lab
	RH1257	GGGCGGAATCCTACGCTAAG			this study
aphA1/IR_{5393}	RH387	ATCCTATGGAACTGCCTCGG	56	3.2	Hall Lab
	strA-R	CCAATCGCAGATAGAAGGC			[270]

[1]designed using GenBank entries as described in above Tables
[2]reference for the primer
[3]in the Hall laboratory

Table A3.7 Primers targeting the IncHI2 resistance region positions

PCR	Primer[1]	Sequence	Location	size (bp)	Reference
Tn*10*/ backbone					
LHS	RH612	TGCTGAGCACTGAGAGATCC	Tn*10*	1,687	this study
	RH619	AAAGCCGTACCATCCATGAG	IncHI2 bbone		this study
RHS	RH856	TTATGGGAATTGGAGGCAATG	Tn*10*	1,574	this study
	RH618	TGAATTCTTTTTCCGGTTGC	IncHI2 bbone		this study
No Tn	RH619	AAAGCCGTACCATCCATGAG	IncHI2 bbone	785	this study
	RH618	TGAATTCTTTTTCCGGTTGC	IncHI2 bbone		this study
Tn*1696-like* /backbone					
tnp end	RH634	ATCTGTCACCAGTGCGAAGG	IncHI2 bbone	808	this study
	RH329	CAATCTGGTGACGGCGG	*tnpA*		[30]
mer end	RH607	TGGAGACCATCCGGTTCTAC	*merR*	540	this study
	RH608	AGTTTCACATCGCCCATTTC	IncHI2 bbone		this study
No Tn	RH634	ATCTGTCACCAGTGCGAAGG	IncHI2 bbone	658	this study
	RH608	AGTTTCACATCGCCCATTTC	IncHI2 bbone		this study
Tn*6024*/ backbone					
LHS	RH1251	CGTCTGCCCCCAGTAAGATG	IncHI2 bbone	7,026	this study
	tnsD-Rv	CAAAAGCCAGCCATGCCC	*tnsD*-like		[155]
RHS	RH878	ATCAGGACCGCGATTTTCTG	*copS*	2,433	this study
	RH1265	CCGACGCTTCCGAGATTAAC	IncHI2 bbone		this study
No Tn	RH1251	CGTCTGCCCCCAGTAAGATG	IncHI2 bbone	1,104	this study
	RH1265	CCGACGCTTCCGAGATTAAC	IncHI2 bbone		this study
ars^{R} Tn					
LHS	ars-Rv	GGCAGATAGTGTGGAATGCG	*arsC*	5,297	[155]
	RH1269	GCCAGAACAAAAGGGCTGAG	IncHI2 bbone		this study
RHS	ars-Fw	AGTGAAAGACAGACGAAGCG	*arsC*	5,987	[155]
	RH895	GGCATGTGCCTTTCGTCTTC	IncHI2 bbone		this study
No Tn	RH1269	GCCAGAACAAAAGGGCTGAG	IncHI2 bbone	2,955	this study
	RH895	GGCATGTGCCTTTCGTCTTC	IncHI2 bbone		this study

[1]primers designed using GenBank Accession nos BX664015 and EF382672

Table A3.8 Primers targeting the IncHI2 backbone

Mapping PCRs	Primer[1]	Sequence	Location	size (bp)	Reference
Backbone PCR#					
1	HI2-F	TTTCTCCTGAGTCACCTGTTAACAC	*repHI2*	646	[275]
	HI2-R	GGCTCACTACCGTTGTCATCCT	*repHI2*		[275]
2	ter-Fw	ATGCAGGCTCAAGGAATCGC	*terF*	893	[155]
	ter-Rv	TTCATCGATCCACGGTCTG	*terF*		[155]
3	92-Fw	CTATGTAAGCAATGATCCTC	orf92	1,001	[155]
	93-Rv	TATAGAGAGCACCGAAGG	orf93		[155]
4	tnsD-Fw	AATCCTTGTTCAGCCGG	*tnsD*-like	1,466	[155]
	tnsD-Rv	CAAAAGCCAGCCATGCCC	*tnsD*-like		[155]
5	ars-Fw	AGTGAAAGACAGACGAAGCG	*arsC*	1,136	[155]
	ars-Rv	GGCAGATAGTGTGGAATGCG	*arsC*		[155]
6	201-Fw	TGTCAGGCTAAGTCACTGG	orf201	1,011	[155]
	RH877	CCATAGAACTTTGGTGCTTG	orf201		this study
7	RH869	CTTGCATCGGTAGTGCTTGC	*repHI1*	1,150	this study
	RH870	TACACCTCTGGGCGTGAGTG	*repHI1*		this study
8	RH873	GGAATCGCCGTACTCTGACC	orf25	2,421	this study
	RH874	TCACGTGCTGTGTTGCAAAG	orf25		this study
9	RH875	TGGGCATTTTCATTGTCAGC	orf56	2,217	this study
	RH876	GTAAGTTCGCTCCCCGTGTC	orf57		this study
10	RH632	CGGTGCTCAGGAGAAAACAG	*hipA*	1,553	this study
	RH633	TTTTCGTAATGGCGAAATGC	*hipA*		this study
11	RH885	GGTGGTGGCTCAGATTCAGG	*silB*	1,704	this study
	RH886	ACAGACCAAGCGGCAGAGAG	*silA*		this study
12	RH878	ATCAGGACCGCGATTTTCTG	*copS*	512	this study
	RH879	GGTCATGGAGGAAGCCAGAC	*copS*		this study
13	RH883	GAGATATGGGGCGTCGTCAC	*mucA*	1,276	this study
	RH884	CCAGCGGGTCGTATATCTCG	*mucA*		this study
14	RH871	ACCGGTGGCTACGGTAATTG	orf270	2,966	this study
	RH872	TCGCACGACCTTTTATGGTG	orf270		this study
Other PCRs					
IS*186*+/-	RH887	TGAACGGTTTGCGTGAAGTG	Band 12	1,594	this study
	RH888	GCGAGTTGCGGATGAAAGAC	Band 12		this study
BamHI fragment	RH1259	TGATCCTGGCCACATCAAAG	IncHI2 bbone	1,967	this study
	RH1260	TTTTGCCAGGCCATTTTACC	IncHI2 bbone		this study
BamHI fragment to SwaI site	RH1271	CCGATACCGAGCCTTACAGC	IncHI2 bbone	3,367	this study
	RH1260	TTTTGCCAGGCCATTTTACC	IncHI2 bbone		this study
between *mucA* and BamHI site	RH899	TCAATCCCCTGGAAATGTGC	IncHI2 bbone	1,524	this study
	RH900	AGGCGCTTCAGAAACAGCAG	IncHI2 bbone		this study

[1]primers designed using GenBank Accession nos BX664015 and EF382672

Table A3.9 Primers targeting the IncHI1 variable regions

Linkage	Primer[1,2, 3]	Sequence	size (bp)
Variable Regions			
1a	RH1272	ACCAGCAGACTGGCGAACAT	1,909
	RH866	TGTGGCACAATGGATTACCC	
1b	RH1272	ACCAGCAGACTGGCGAACAT	1,088
	RH867	TCCATCTTGGCGTGGTTATG	
2	RH857	CCTGTTAGTCACCGGCTTGC	2,354 (if present) or
	RH858	CCCGGCAATTGACAAACAC	300 (if absent)
3	RH1278	AATGGTCCGGGCAATTCTG	3,027 (if present) or
	RH616	TCCGCAATAATATGCTCTGG	1,004 (if absent)
4	RH1276	CGCTCAGGCCATTAAAGAGG	4,378 (if present) or
	RH1277	TGATAGTCCAAAGTTCACCATGC	269 (if absent)
6	RH1279	CGCCAAGGAGTCGCTTAATG	3,120 (if present) or
	RH1280	AGCGCCTGACTGAACAGGAG	679 (if absent)
F-rep	RH1281	TTCCGCTCCATGGTGGTAAC	3,354 (if present) or
	RH1282	CATGAGCTCCGGGTCACTTC	942 (if absent)
Additional region			
Across additional IS*10* in pHCM1	RH628	CTGTTCATGCCAGCGATAAG	1,814 (if IS*10* present) or
	RH1262	TGACCAGGTCCCCCATAAAC	605 (if IS*10* absent)

[1]primers designed using GenBank accession no. AL513383 and AF250878
[2]all primers designed in this study
[3]all PCR reaction performed with an annealing temperature of 60^{O}C

Table A3.10 Additional primers to map variants

Primer[1,2]	Sequence	Location	Tm (^{o}C)	Reference
RH863	GCAAATTTTCCCGACTGAGC	LHS MARR	65	this study
RH865	AGTACCCGGAACCACACACC	LHS MARR	66	this study
RH1283	AATCACTGTGCGGGGTAATC	LHS MARR	64	this study
RH624	CAGAGGGCAACAGCGAAGTG	LHS MARR	64	this study
RH627	CGATCATGCTCAATTGCATAC	RHS MARR	65	this study
RH1263	ACGGCCAGGGATGTAAACTG	RHS MARR	65	this study
RH625	GCGGTCGACAGTTGAGTACC	RHS MARR	66	this study
RH415	CAGTGAGAGCAGAGATAGC	IS*1*	58	Hall Lab
RH621	GCGAATTGAGCGGCATAACC	IS*1*	65	this study
RH359	GTTGTCGGTGTTCATGC	$tnpA_{21}$	59	Hall Lab

[1]primers designed using GenBank accession no. AL513383 and AF250878
[2]all PCR reaction performed with an annealing temperature of 60^{O}C

APPENDIX 4

Additional tables

Table A4.1 Bands observed in RFLP with SwaI from *S*. Infantis plasmids

Frag Order seq	Comments on/ contents of band (PCR#)	Expected R478	R478 PFGE bands (kb)*	Expected Infantis	PFGE SRC71	PFGE SRC72	PFGE SRC92	PFGE SRC95	PFGE pSRC96
				49.7 Tn1696, ter (5), orf92/3 (6), hipA(14), tnsD (7)	49.3		49.3	49.3	50.9/49.1
1	RepHIA(PCR 2);trh (3)	**42.8**	42.2/43.9	**42.8**	44.9	44.9	44.9	44.9	44.9/44.7
10 6	ars(8);muc (12); trh(9) ter (5), orf92/3 (6), hipA(14), tnsD (7) 34.9 kb (no if Tn)	**37.2**	33.9/36.4	**37.2**	30.1	31.9 31.9	31.9	30.9 30.9	31.9/31.4
11	Tn10	**26.7**	24.8/26.2	**26.7**	25.6	26.4	26.4	26.4	26.4/26.1
15	orF270(10)	**25.8**	23.3/25.2	**25.8**	24.1	24.1	26.4	26.4	24.8/25.2
					21.3	21.9			23.3/22.9
5	orf 56/57 (4)	**21.5**	20.6/22.1	**21.5**	21.3	21.9	21.9	21.9	21.9/21.7
		18.2(*mer*$_{1696}$,*tnsD*(7))	17.6/18.7		18.2	18.2	21.9		18.8/18.7
		17.5(ter, 92/3,tnp)	17.1/18.3				18.7		
12	IS*186*	**16**	16/17.4	**16**	15.1	15.1	15.6	15.6	15.5/15.2
8	sil (11)	**13.7**	13.3/14.7	**13.7**	12.9	12.9	13.3	13.3	13.3/13.2
9	cop (13)	**11.1**	11.7/12.4	**11.1**	10.4	12.2	12.2	12.1	-
16	RepHI2 (1)	**10.2**	11/12.2	**10.2**	10	10	10.4	11	11/10.7
2		**10**	9.7/10.7	**10**	8.9	-	8.8	8.8	10.7/10.5
		4.6 (MDR)	-/5.2						9.2/8.9
		3.8(MDR)			4.5	4.5	4.5		4.6/4.5
13		**3.7**	-/3.9	**3.7**	3.6	3.6	3.6	3.6	3.7/3.7
7		**3.4**	-/3.7	**3.4**					
		2.2(MDR)							
4		**1.9**		**1.9**					
		1.6(MDR)							
14		**1.5**		**1.5**					
3		**1.1**		**1.1**					

Table A4.2 Distribution of location of *aacC* genes in the GenBank database.

***aacC* Type**	**Ac#**	**Organism**	**Plasmid Name**	**Sequence length (bp)**	**Genes within 10kb of *aacC***
2[1]	-	*S.*Typhimurium	pSRC27-H	-	**IS*26***/Tn*2*Δ/***aacC2***/-[3]---/ISCfr1/*bla*$_{TEM}$
2	AF550415	*C. freundii*	pCTX-M3	89 468	"
2	AY333434	*S.* Typhimurium	pU302L	84 514	**IS*26***/***aacC2***/----/ISCfr1/bla$_{TEM}$
2	DQ449578	*K. pneumoniae*	pK245	98 264	bla$_{SHV-2}$/IS15(**IS*26*)**/IS*10*/**aacC2**/---/ **IS*26*** /-/bla$_{TEM}$
2	L22613	*S. marcescens*	-	992	only *aacC2* sequenced
2	EU022315	*E. coli*	-	861	"
3[2]	AY458016	*E. coli*	pC15-1a	92 353	IS*26* /oxa-1/aac6-Ib/**IS*26***/catB3Δ/IS26(c)/***aacC3***/hypo/orfAB/**IS*26***/-----/*bla*$_{TEM}$
3	AY138987	*A. baumannii*	Unnamed	6 080	IS*1133* /***aacC3***//*intI1*Δ/**IS*26*** **(**end of seq)
3	AF466526	*P. aeruginosa*	-	4486	Tn*1*Δ/IS*6100*/bla$_{TEM}$/***aacC3*** (end of seq)
3	X13542	*E. coli*	pWP14a	1 336	only *aacC3* sequenced
3	EU022314	*E. coli*		861	"
3	AY191460	*synthetic from E. coli*	-	1761	"
3	S68058	*E. coli*	pUO055	1014	"
3	X13541	*?*	pWP116a	1336	"
3	X13543	*?*	pWP113a	1336	"
3	M62833	*A. baumannii*	-	1123	"
3	X54723	*E. coli*	Unnamed	2273	"

[1] *aacC2*= >99% similarity to pSRC27-H

[2] *aacC3*= a standard pattern of bp changes and <97% similarity to *aacC2*

[3] "-" denotes an orf with function not known

APPENDIX 5

Additional figures

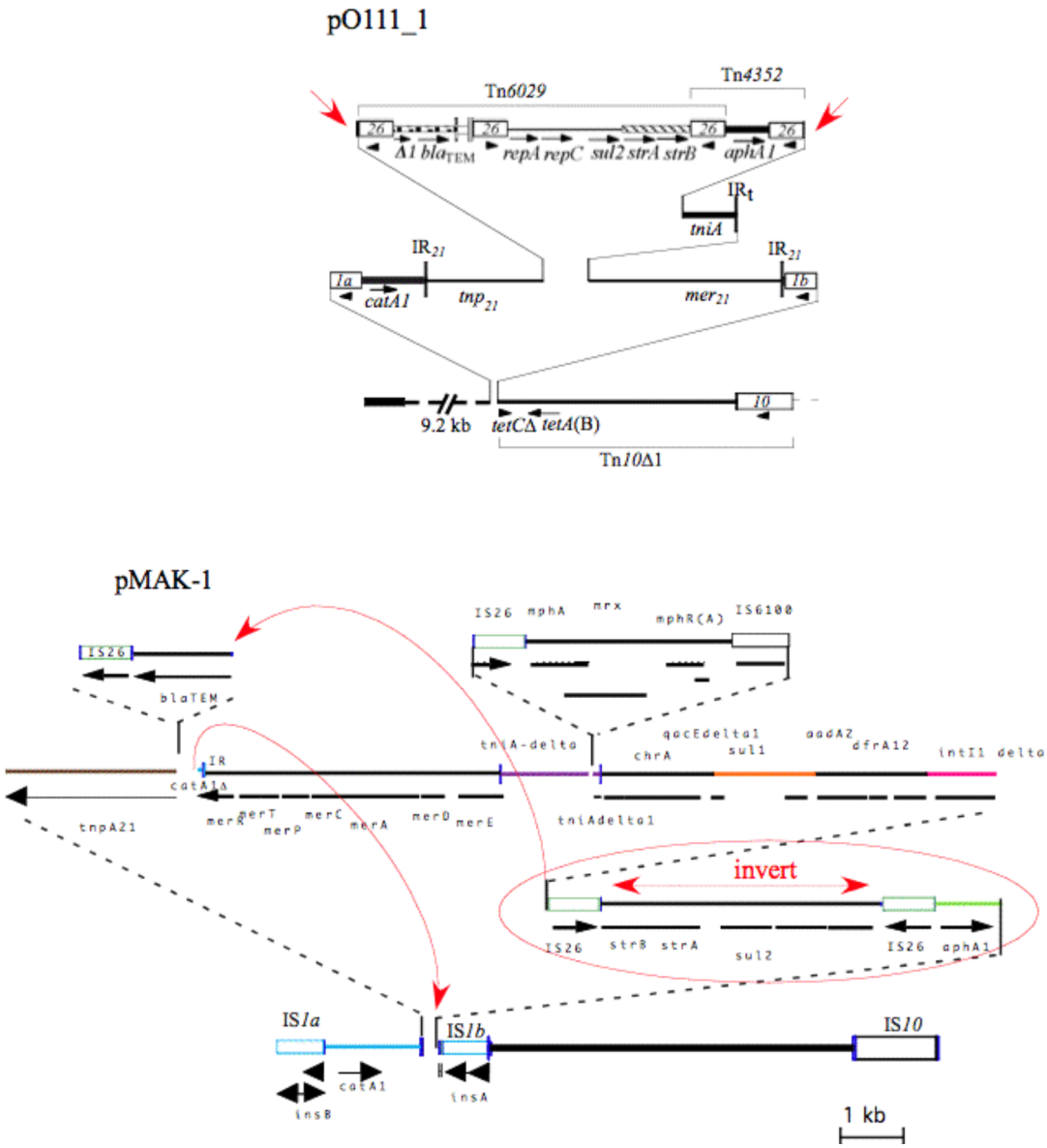


Fig A5.1 pMAK-1 MARR structure. The structure was determined using GenBank acc no AB366440. The red arrows indicate possible sites of misalignment. pO111_1 is included above for comparison.

Fig A5.2 Sequences alignments across deletion sites of variants. The boundaries of Variants 3-1, 3-2, 4, 5-2, 6, 7, and 8 (top sequence) are compared to the parent (bottom sequence). The vertical lines indicate a bp match. Colour denotes the type of element, where red is IS*1*, blue is Tn*10*, pink is the integron, green is IS*26*, purple is Tn*2* and black is backbone. The bold-face type shows the IRs of insertion sequences.