pIP40a, a type 1 IncC plasmid from 1969 carries the integrative element GI*sul2* and a novel class II mercury resistance transposon

Christopher J. Harmer*, Mohammad Hamidian, Ruth M. Hall

School of Life and Environmental Sciences, The University of Sydney, NSW 2006, Australia

ARTICLE INFO

Keywords:
IncC
GI*sul2*
Genomic island
Antibiotic resistance
Transposon

ABSTRACT

The 167.5 kb sequence of the conjugative IncC plasmid pIP40a, isolated from a *Pseudomonas aeruginosa* in 1969, was analysed. pIP40a confers resistance to kanamycin, neomycin, ampicillin, sulphonamides and mercuric ions, and several insertions in a type 1 IncC backbone were found, including copies of IS*3*, Tn*1000* and a novel mercury resistance transposon, Tn*6182*. The antibiotic resistance genes were in two locations. Tn*6023*, containing the *aphA1* kanamycin and neomycin resistance gene, is in a partial copy of Tn*1*/Tn*2*/Tn*3* (*bla* TEM, ampicillin resistance) in the *kfrA* gene, and the *sul2* sulphonamide resistance gene is in the integrative element GI*sul2* in the position of ARI-B islands. The 11.5 kb class II transposon Tn*6182* is only distantly related to other class II transposons, with at most 33% identity between the TnpA of Tn*6182* and TnpA of other group members. In addition, the inverted repeats are 37 bp rather than 38 bp, and the likely resolution enzyme is a tyrosine recombinase (TnpI). Re-annotation of GI*sul2* revealed genes predicted to confer resistance to arsenate and arsenite, but resistance was not detected. The location of GI*sul2* confirms it as the progenitor of the ARI-B configurations seen in many IncC plasmids isolated more recently. However, GI*sul2* has integrated at the same site in type 1 and type 2 IncC plasmids, indicating that it targets this site. Analysis of the distribution of GI*sul2* revealed that it in addition to its chromosomal integration site at the 3′-end of the *guaA* gene, it has also integrated into other plasmids, increasing its mobility.

1. Introduction

IncC plasmids have long been known to play an important role in the dissemination of antibiotic and mercuric ion resistance determinants. They were first recovered in Paris in the late 1960s (Chabbert et al., 1972; Witchitz and Chabbert, 1971). However, it was not until 2007 that an IncC plasmid was sequenced (Welch et al., 2007). Plasmid pIP40a was originally recovered from a clinical *Pseudomonas aeruginosa* isolate in Paris in 1969 and classified as IncP-3. It was transferred by conjugation into an *Escherichia coli* host and was assigned to incompatibility group IncC using R55, also recovered in 1969 in Paris (Witchitz and Chabbert, 1971), as the IncC reference plasmid (Chabbert et al., 1972). For a period of time, IncC plasmids were combined with the IncA plasmids as IncA/C (Datta and Hedges, 1972). Later, they were split into A/C_1 and A/C_2 (Carattoli et al., 2006), corresponding to the IncA and IncC groups respectively (Harmer et al., 2016b). However, as IncA and IncC plasmids were known to be compatible (Datta and Hedges, 1973), the designation IncA/C is incorrect and they have been separated again into IncA, of which RA1 is the only sequenced example, and the more common and more important IncC group (Harmer and Hall, 2015).

A comparison of the 29 IncC plasmid sequences available in the public domain in 2014 revealed that there were two distinct lineages, type 1 and type 2 (Harmer and Hall, 2014). The backbones of IncC plasmids from the two lineages differ by ~1% and are also distinguished by two regions where alternate sequences are found, orf1832 in type 1 or orf1847 in type 2, and *rhs1* in type 1 or *rhs2* in type 2, and two small additional segments (i1 and i2) in the type 2 plasmid backbone (Harmer and Hall, 2014). It is clear that the type 1 and type 2 IncC plasmids have separate evolutionary histories, and we recently developed a PCR typing scheme to distinguish type 1 and type 2 plasmids (Harmer and Hall, 2016). However, sequences provide more detailed information.

One of the key stages in the evolution of type 1 and type 2 IncC plasmids has been the acquisition of antibiotic resistance islands, and sub-lineages can be identified based on the location of resistance islands, which can vary internally. The majority of recently isolated type 1 plasmids have the ARI-A island, and are likely to be descended from a common ancestor (Harmer and Hall, 2015). However, surprisingly the ARI-B antibiotic resistance island, containing *sul2* and all or part of the

small mobile element CR2, is found in the same location in both type 1 and type 2 IncC plasmids (Harmer and Hall, 2015; Harmer and Hall, 2014). A variety of ARI-B configurations carrying additional antibiotic resistance genes have been observed (see Table 5 and Fig. 7 in (Harmer and Hall, 2015)). Recently, it has been shown that the *sul2* gene is part of a 15 kb integrative element GI*sul2* (Nigro and Hall, 2011; Hamidian and Hall, 2017). In three sequenced IncC plasmids (R55, pEA1509 and pCFSAN001921), ARI-B is bounded by the two outer ends of GI*sul2* (Harmer and Hall, 2015; Nigro and Hall, 2011), but the central portion of the genomic island has been lost. Hence, ARI-B appears to have originally formed via the integration of GI*sul2*. However, no IncC plasmid carrying a complete copy of GI*sul2* had been seen until recently. Whether GI*sul2* was acquired independently by each lineage, or whether recombination between the two lineages has occurred remains to be resolved.

R55, the oldest IncC plasmid known to be type 2 (Harmer and Hall, 2014), has been sequenced and it had already acquired a complex antibiotic and mercury resistance region (Doublet et al., 2012). However, the type 1 plasmids are the most important, as members of this group have acquired the bla_{CMY} cephalosporin resistance gene and more recently the bla_{NDM} carbapenem resistance gene which has been acquired by the ARI-A resistance island. Among the type 1 plasmids, pDGO100, which was recovered in 1981 (Groot Obbink et al., 1985), was one of the oldest that had been sequenced so far (Harmer et al., 2016b). pDGO100 had acquired the ARI-A resistance island but lacks an ARI-B island. The next oldest, p199061_160, wasn't recovered until 1991 (Fernandez-Alarcon et al., 2011). Though over 100 IncC plasmids have been sequenced since 2007 (last searched 1st February 2017), they mostly originate from the contemporary period and many are identical to one another or very closely related (Harmer and Hall, 2015). Hence, they do not add a great deal to our understanding of the longer-term evolutionary history of the type 1 IncC plasmids, or their role in the emergence of antibiotic and heavy metal resistance.

Recently, we showed that pIP40a is type 1 (Harmer and Hall, 2016). Hence, both type 1 and type 2 IncC plasmids were circulating in Paris in the late 1960s and contributing to the development of antibiotic resistance. We also used PCR to show that ARI-B in pIP40a contains both ends of GI*sul2*, and confirmed by sequencing (GenBank accession number KX709966) that it contains a complete copy of GI*sul2* (Harmer and Hall, 2016). Here, we present a detailed analysis of the complete sequence of pIP40a, aimed at increasing our understanding of the evolution of IncC plasmids and of the role of genomic island GI*sul2* in spreading the *sul2* sulphonamide resistance gene. A novel class II transposon was also identified.

2. Methods

2.1. pIP40a

Plasmid 40a was originally recovered in 1969 in Paris, France, from a *P. aeruginosa* clinical isolate that had caused a urinary tract infection (Chabbert et al., 1972) and later assigned an Institute Pasteur (IP) number. pIP40a was recorded as conferring resistance to kanamycin, neomycin, ampicillin, sulphonamides, and mercuric ions. An *E. coli* containing pIP40a was obtained from the UK National Type Culture Collection via Professor David Gordon (Australian National University).

To determine the frequency of transfer, pIP40a was transferred overnight (16 h) on solid media from nalidixic acid resistant (Nx^R) *E. coli* DH5-α to streptomycin resistant (Sm^R) *E. coli* UB1637 as described previously (Harmer et al., 2016b).

2.2. Antibiotic and heavy metal susceptibility testing

Antibiotic susceptibility profiles were determined as described previously (Bailey et al., 2010). Resistance to mercuric ions was tested by growing strains overnight at 37 °C on LB agar supplemented with 25 μg/ml $HgCl_2$. Resistance to arsenate and arsenite was tested by growth on LB agar containing 5 mM $NaAsO_2$ or 20 mM Na_2HAsO_4, respectively.

2.3. Analysis of pIP40a sequence

The complete sequence of pIP40a assembled previously (Harmer and Hall, 2016), and submitted to GenBank under accession number KX709966, was analysed. Here, additional sequence data that increased the read depth from 12 to 29-fold was assembled and no differences were found. The copy number of pIP40a was assessed using contig coverage relative to that of seven chromosomal genes. Reading frames of > 300 bp (100 aa) were predicted using Gene Construction Kit (Textco BioSoftware, Raleigh, NC, USA) and were annotated using the annotations of the IncC plasmids pRMH760 (GenBank accession number KF976462) and pR148 (GenBank accession number CP000603) as a guide. Insertion sequences were identified using IS Finder (https://www-is.biotoul.fr) (Siguier et al., 2006), and antibiotic resistance genes were identified using ResFinder 2.1 (https://cge.cbs.dtu.dk/services/ResFinder/). Pfam (http://pfam.xfam.org/) was used to assist in the prediction of the function of proteins encoded by genes whose function has not been experimentally determined.

The backbone of pIP40a was derived by removing the transposons, insertion sequences and antibiotic resistance islands, and any duplication created when they were inserted. The backbone was compared to other complete IncC plasmids present in the Genbank non-redundant database. To facilitate comparisons, the backbones of these plasmids were generated in the same manner by removing all insertions, then circularising the sequences and reopening them at the same location, namely 1139 bp upstream of *repA*.

2.4. Detection of excised GI*sul2*

To test for excision of GI*sul2* from pIP40a, *E. coli* UB1637 (Sm^R) containing pIP40a was grown overnight in Luria Broth without selection. Whole cell genomic DNA was extracted and used as the template in a PCR with primers RH1879 (5′-ATCGTGCCAAACACAAACAA-3′) and R3R (5′-GAATGCATAACGACGAGTTTGG-3′). These primers have been utilised previously in different combinations (Harmer and Hall, 2016), to detect the excised circular product. PCR conditions were as published previously (Harmer and Hall, 2016). To test for low frequency excision, 1 μl of the PCR reaction was used as a template for a second round of amplification using the same primers.

3. Results

3.1. Plasmid pIP40a

pIP40a was previously reported to contain genes conferring resistance to ampicillin, kanamycin and sulphonamides (Chabbert et al., 1972). *E. coli* containing pIP40a were also able to grow on plates containing 25 μg/ml $HgCl_2$, indicating the presence of a *mer* module. pIP40a had a unit copy number. The efficiency of transfer of pIP40a (*E. coli* DH5-α to *E. coli* UB1637) was 4.38×10^{-5} $Ap^RKm^RSu^R$ transconjugants per donor (average of three independent determinantions). This is between 100- and 1000-fold lower than the conjugation frequencies of other well-studied type 1 IncC plasmids such as pRMH760 (6.4×10^{-2} (Harmer and Hall, 2014)), pDGO100 (3×10^{-3} (Harmer et al., 2016b)) or the type 2 pVCR94ΔX (3.5×10^{-3}–1×10^{-2} (Carraro et al., 2014b, Carraro et al., 2014a)).

The pIP40a sequence (GenBank accession number KX709966) was 167,554 bp, and the sequence of the backbone confirmed recent findings (Harmer and Hall, 2016) that pIP40a is a type 1 IncC plasmid. Several insertions were identified in the backbone of pIP40a, and their locations are shown in Fig. 1. An IS3, flanked by a 3 bp target site duplication (TSD), has interrupted the *rhs1* gene as noted previously

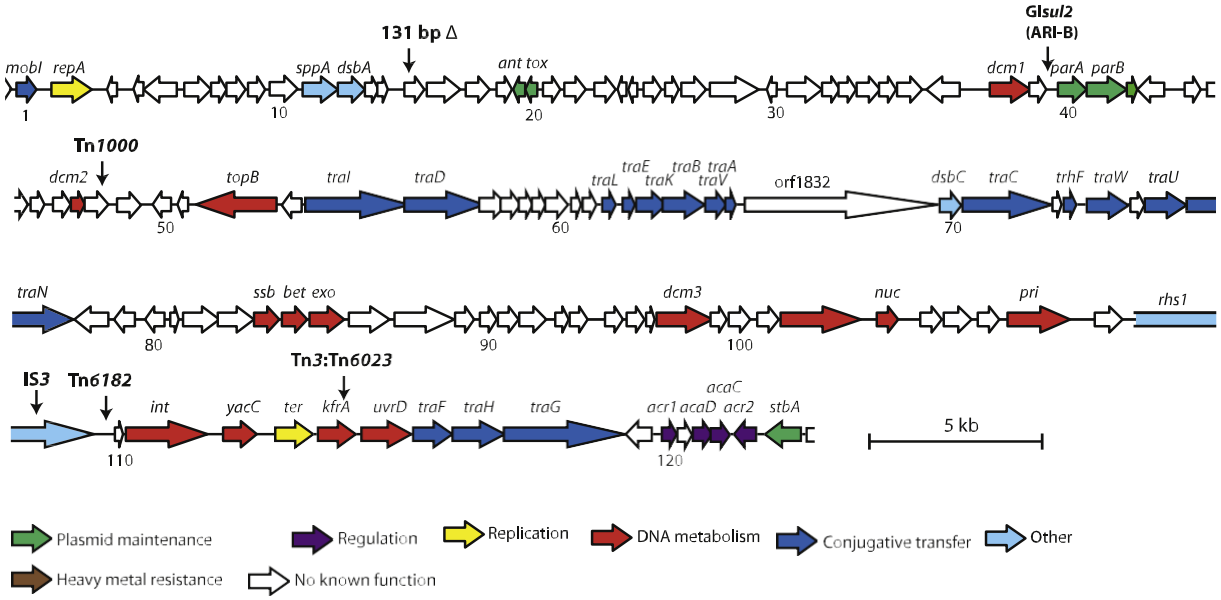

Fig. 1. Genetic organization of pIP40a. The backbone sequence of pIP40a was linearized 1139 bp upstream of *repA* and is drawn to scale from GenBank accession number KX709966. Horizontal arrows indicate the location, size and orientation of ORFs, and every tenth reading frame is numbered below the arrows. The positions of GI*sul2*, transposons, IS3, and a deletion within orf15 are indicated by vertical arrows. All ORFs > 100 aa are shown, and select ORFs < 100 aa with known function are shown. Genes coding for proteins of known function are named above and colored according to the key. Figure was generated using Adobe Illustrator CS6. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

(Harmer and Hall, 2016). pIP40a also includes a partial copy of Tn*1*/Tn*2*/Tn*3* interrupted by the kanamycin resistance transposon Tn*6023* (Cain and Hall, 2011), which includes the *aphA1b* gene conferring resistance to kanamycin and neomycin ($Km^R Nm^R$), a complete copy of the class II transposon Tn*1000* (also known as gamma delta)(Broom et al., 1995), and a novel mercury-resistance transposon which we named Tn*6182*. pIP40a includes GI*sul2*, containing *sul2* (Su^R), as the ARI-B resistance island. Whilst most known type 1 IncC plasmids carry an ARI-A resistance region (Harmer and Hall, 2015), pIP40a does not, and the plasmid backbone is continuous at the ARI-A location. Hence pIP40a, only the second sequenced type 1 IncC plasmid that does not contain an ARI-A island (Harmer and Hall, 2015), belongs to a type 1 lineage that predates the entry of the ancestral form of ARI-A.

3.2. The backbone of pIP40

Removal of ARI-B, the transposons, the IS and the duplications they created, generated a plasmid backbone of 127,670 bp (Fig. 1) that includes a replication initiation gene (*repA*), genes for conjugative transfer (*tra*), plasmid partitioning (*par*), a variety of other genes with predicted functions, and open reading frames with no known function that have been described elsewhere (Harmer and Hall, 2015; Hancock et al., 2017). The pIP40a backbone is 133 bp smaller than the typical type 1 IncC backbone of 127,803 bp (Harmer and Hall, 2014), due to a 131 bp deletion internal to an open reading frame with no known function (see Fig. 1), and the absence of two single base-pairs in intergenic regions of the backbone, that are present in all other type 1 IncC plasmids. A BLAST search using the pIP40a backbone as a query showed that it is most closely related to the backbone of pRMH760 (GenBank accession number KF976462). In addition to the 131 bp deletion in pIP40a and a 2358 bp deletion in the *rhs1* gene of pRMH760, the two backbones differ by a total of 63 single nucleotide polymorphisms (SNPs) distributed throughout the backbone. The majority of the SNPs are intergenic (21/63) or in ORFs of no known function (28/63). Seventeen are in genes of known function, of which nine are silent substitutions. The remaining eight SNPs result in single amino acid changes in TraI, TraL, TraC, TraN, Int, and Yac, and two substitutions in Dcm3. pR148 (GenBank accession number CP000603), which includes the 2358 bp missing from pRMH760, is the next closest relative to pIP40a, differing by 65 SNPs, most shared by pRMH760. The non-synonymous substitutions in the *tra* genes could be responsible for the low transfer frequency.

3.3. Tn*1*/Tn*2*/Tn*3*::Tn*6023*

The bla_{TEM-1} ampicillin resistance gene is found in a remnant of a Tn*1*/Tn*2*/Tn*3* transposon (Fig. 2a) that has inserted into the *kfrA* gene of the pIP40a backbone (Fig. 1) and is flanked by a 5 bp TSD. The 2641 bp *aphA1b*-containing IS*26*-bounded transposon Tn*6023* (GenBank accession number GU562437) has subsequently inserted in this Tn (Fig. 2a), and an IS*26*-mediated deletion arising from either one or both of the IS*26* bounding Tn*6023* has removed 1935 bp of the Tn, removing *res* and truncating the *tnpA* gene. As this region includes the features that distinguish Tn*1*, Tn*2* and Tn*3* from each other (Partridge and Hall, 2005; Bailey et al., 2011), it is not possible to identify the specific variant. Tn*6023* is not flanked by a TSD due to this deletion.

Tn*6023* is unusual because the two IS*26* flanking *aphA1b* are inversely oriented, whereas in most IS*26*-bounded transposons the IS are in the same orientation. Tn*6023* was first described as a component of a larger transposon (GenBank accession number GU562437) in an IncHI2 plasmid, pSRC125, recovered from a *S. enterica* of bovine origin in Australia in the year 2000 (Cain and Hall, 2011). However, the presence of Tn*6023* in pIP40a shows clearly that it arose at least three decades earlier. In pSRC125, Tn*6023* is within Tn*5393* and is flanked by an 8 bp TSD. A BLAST search of the GenBank nucleotide database using Tn*6023* as a query revealed that it has since been found in the chromosomes or plasmids of a number of bacterial species, including *Corynebacterium singular*, *Pseudomonas putida*, *P. aeruginosa*, *Serratia marcescens*, *Acinetobacter baumannii* and *Proteus mirabilis*. In most of these sequences, Tn*6023* is not flanked by a TSD. However, in five *Corynebacterium* genome sequences (GenBank accession numbers CP010827, CP009211, AM942444, FN825254 and CP004085), Tn*6023* is in the same position in Tn*5393* as in pSRC125 flanked by the same 8 bp duplication.

3.4. Tn*6182*, a new class II mercury resistance transposon

A novel 11,514 bp transposon, Tn*6182* (Fig. 2b), is inserted in the

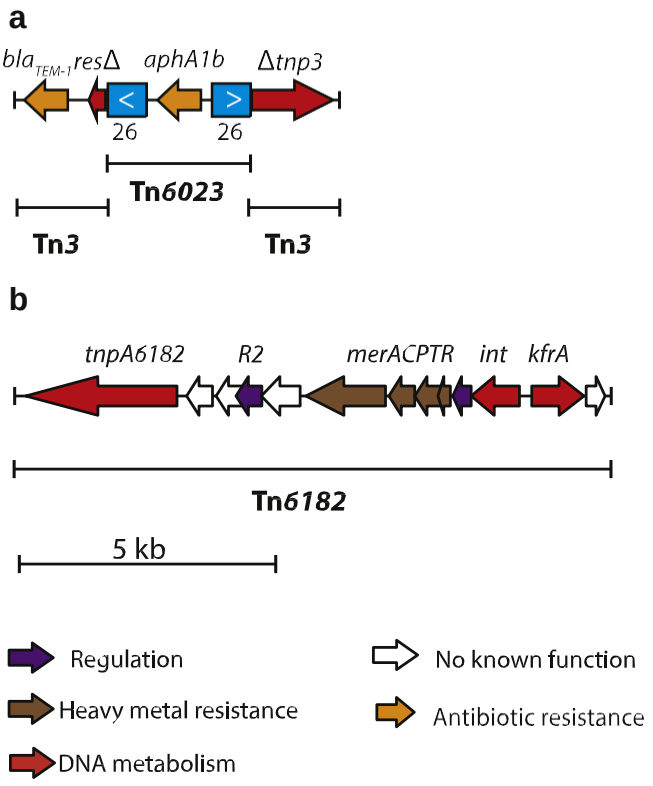

Fig. 2. Resistance transposons in pIP40a. a) Tn*3*::Tn*6023*. b) Novel mercury resistance transposon Tn*6182*. The transposons are drawn to scale from GenBank accession no. KX709966, and are drawn in the orientation consistent with Fig. 1. IS*26* are shown as open boxes with > or < inside indicating the orientation. Genes coding for proteins of known function are named above and colored according to the key. Figure was generated using Adobe Illustrator CS6. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

intergenic region downstream of *rhs1* in pIP40a (Fig. 1). Tn*6182* is bounded by 37 bp terminal inverted repeats and is flanked by a 5 bp TSD. Annotation of all open reading frames encoding proteins > 50 amino acids (see Table 1 and GenBank accession number KX709966 for full annotations), uncovered a gene encoding a 995 aa protein, identified by Pfam as a class II (Tn*3* family) TnpA transposase (PF01526), at one end of Tn*6182*. TnpA6182 is co-linear with the unusually long transposases found in class II transposons, and phylogenetic analysis (Fig. 3a) with one or more representative transposases from each of the clades of the class II transposons defined recently (Nicolas et al., 2015)

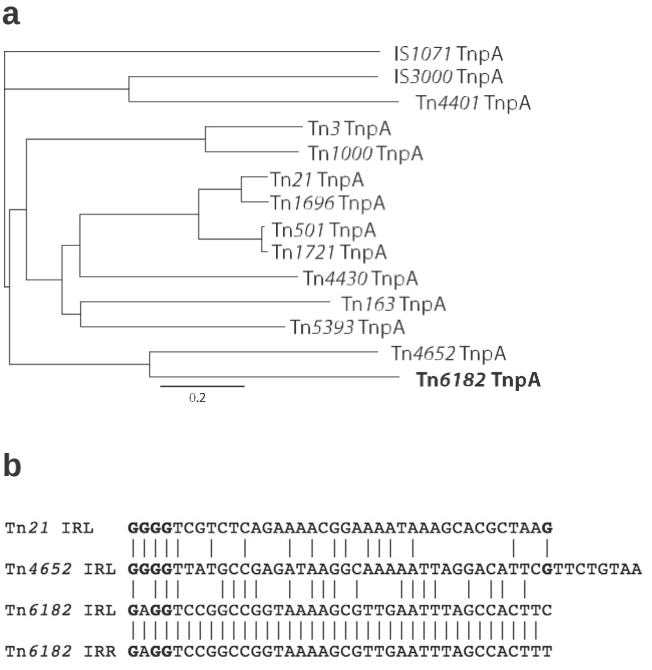

Fig. 3. Comparison of features of Tn*6182* and other class II transposons. a) Phylogenetic tree showing relatedness between the transposase proteins of selected class II transposons. The tree was generated using Geneious R7 using a neighbour-joining algorithm. b) Alignment of the left and right inverted repeats of Tn*6182* with the inverted repeats of Tn*21* and Tn*4652*. Bold letters indicate highly conserved G nucleotides. Vertical lines indicate shared bases.

showed that TnpA6182 is only distantly related to other class II transposases. The closest relative is the TnpA of Tn*4652* (GenBank accession number X83686), with which it shares 33% aa identity. The other class II transposases all share < 26% aa identity with the TnpA of Tn*6182*. Hence, Tn*6182* likely represents a new clade.

The inverted repeats (IR) of Tn*6182* also exhibit differences from those of other class II transposons. The four terminal bases (GGGG) of the 38 bp IRs of all other class II transposons are completely conserved (see alignment in (Nicolas et al., 2015)). However, one of these bases is not conserved in the IRs of Tn*6182* (GAGG) (Fig. 3b). The 38th base of the inverted repeats in all other class II transposons, a highly-conserved G nucleotide, is not present in Tn*6182* (Fig. 3b). Hence, the IRs are one base pair shorter than the more usual 38 bp IRs for other class II transposons. Many class II transposons include a *tnpR* gene encoding a

Table 1
Features of Tn*6182*.

Feature/gene	Position[a]	Size (bp)	Product	Size (aa)[b]	Pfam	Description
IRL	1–37	37	–	–	–	Inverted repeat
tnpA6182	68–3052c	2985	TnpA6182	995	01526	Class II transposase
orf[c]	3172–3735c	564	Orf	188	–	Hypothetical protein
orf	3735–4127c	393	Orf	131	–	Hypothetical protein
R2	4154–4684c	531	R2	177	13411	MerR family regulator
orf	4608–5405c	798	Orf	266	00563	EAL domain family protein
merA	5420–7084c	1665	MerA	555	07992	Mercuric reductase
merC	7125–7541c	417	MerC	139	03203	Mercury resistance protein
merP	7548–7823c	276	MerP	92	00403	Mercury transport protein
merT	7876–8223c	348	MerT	116	02411	Mercury transport protein
merR	8320–8709c	390	MerR	130	13411	MerR family regulator
tnpI	8709–9728c	1020	TnpI	340	00589	Tyrosine recombinase
kfrA	9879–10,979	1101	KfrA	367	11740	Site-specific DNA binding protein
orf	11,102–11,431	330	Orf	110	–	Hypothetical protein
IRR	11,478–11,514	37	–	–	–	Inverted repeat

[a] Position 1 corresponds to position 131,414 of GenBank accession number KX709966. The letter "c" after the base range indicates the gene/orf is on the complimentary strand.
[b] Length of product in amino acids (aa).
[c] Open reading frame.

Table 2
Amino acid identities of Mer proteins encoded in Tn*6182* compared to Tn*21*, Tn*501*, and Tn*1696*.

Tn	Accession no.	MerA % (555aa)	MerC % (139aa)	MerP % (89aa)	MerT % (116aa)	MerR % (130aa)
Tn*6182*	KX709966	100	100	100	100	100
Tn*21*	AF071413	57 (325/567)	56 (73/131)	61 (43/70)	50 (52/105)	46 (58/125)
Tn*501*	Z00027	57 (323/564)	–[a]	64 (45/70)	46 (51/111)	47 (59/125)
Tn*1696*	U12338	56 (219/566)	59 (78/132)	63 (47/70)	53 (53/100)	47 (59/125)

[a] No MerC in Tn*501*.

TnpR serine recombinase (resolvase) that acts to resolve cointegrates, whilst in others, including Tn*6182*, a *tnpI* or *int* gene encoding a tyrosine recombinase (Pfam PF00589) replaces *tnpR* (Nicolas et al., 2015). Another gene, *kfrA*, encodes a DNA binding protein associated with plasmid replication (Pfam PF11740). Though there are similarly named genes in the IncC backbone that encode proteins sharing the same Pfam, the Tn*6182* *tnpI* and *kfrA* genes do not share any significant sequence similarity with those from IncC plasmids.

Like many other class II transposons, Tn*6182* carries a mercuric ion resistance operon, and pIP40a confers resistance to mercuric ions. The *mer* module is novel, but the MerA, MerC, MerP, and MerT proteins share 46–64% amino acid identity with the corresponding Mer proteins in the well-characterised Tn*21* (GenBank accession number AF071413), Tn*501* (GenBank accession number Z00027) and Tn*1696* (GenBank accession number U12338) *mer* operons (Table 2), allowing their functions to be deduced. However, unlike most other transposons that carry the *mer* operon at an outer end (Liebert et al., 1999; Partridge et al., 2001), the *mer* operon of Tn*6182* (*merACPTR*) is located in the central portion of the transposon (Fig. 2b).

3.5. GI*sul2*, the progenitor of ARI-B islands

We recently proposed that the various ARI-B resistance islands seen in many type 1 and type 2 IncC plasmids arose from GI*sul2* integrated into the IncC backbone (Harmer and Hall, 2015; Harmer and Hall, 2016). The pIP40a sequence confirmed this as the ARI-B island of pIP40a is a complete copy of the 15,460 bp integrative element GI*sul2* (Fig. 4a). *Re*-examination of all GI*sul2* open reading frames predicting proteins longer than 50 amino acids (see Table 3 and GenBank accession number KX709966 for full annotations) has enhanced the earlier annotations of GI*sul2* (Nigro and Hall, 2011). In addition to the previously described *int*, *repA*, *repC*, *trbJ*, *trbL*, *resG*, *rcr2*, and *sul2* genes (Nigro and Hall, 2011), a number of genes that had been overlooked were identified.

GI*sul2* includes an apparently complete arsenate/arsenite resistance operon (*arsBCHR*, Fig. 4a) that encodes proteins that are distantly related to those from other previously characterised *ars* operons such as those found in *S. enterica* R46 (GenBank accession number NC003292), *E. coli* R773 (GenBank accession number J02591), and *P. aeruginosa* PAO1 (GenBank accession number NC002516). The module in pIP40a encodes ArsC and ArsB homologues that are over 40% identical to those of *P. aeruginosa* PAO1 (Cai et al., 1998), and an ArsH protein that is 63% identical to ArsH from R64 (Sampei et al., 2010). Hence, all the genes required to confer resistance to arsenite and arsenate are present (Kaur et al., 2011). However, the *E. coli* transconjugants containing pIP40a did not grow on LB agar supplemented with 5 mM sodium arsenite or 20 mM sodium arsenate. This susceptibility to arsenate and arsenite indicates that there may be a defect in the operon or that the ArsR regulator of the operon may be inactive, and this will require further investigation.

GI*sul2* also includes three genes encoding proteins that may be important for the stable maintenance of the genomic island (Fig. 4a), including a putative type 1 toxin (Pfam 12,703), the putative anti-toxin (PemI) sharing 42% amino acid identity with GenPept EXI82236, and a protein that belongs to the StbC superfamily of proteins (NCBI Conserved Domain cl01921) that is usually associated with plasmid stability.

A small 237 bp orf was found between the genes for the conjugative transfer proteins TrbJ (GenPept AOE47832) and TrbL (GenPept AOX98907) described previously. The predicted 78 aa protein (GenPept AOX98906) was identified as a putative entry exclusion lipoprotein TrbK (Fig. 4a), as it shares 33% amino acid identity with the TrbK entry exclusion protein of the IncP1-α plasmid RP4 (RP1, RK2, R18, R68). TrbJ and TrbL of GI*sul2* share 58% and 52% amino acid identity with TrbJ and TrbL from RP4, respectively. RP4 TrbL is believed to be involved in mating pair stabilization, whilst TrbJ is an inner membrane pore-forming protein (Pansegrau et al., 1994). TrbK in RP4 prevents other IncP1-α plasmids from using conjugation to enter the same cell (Haase et al., 1996). Due to the similarity between the Trb proteins encoded by genes on GI*sul2* and IncP1-α plasmids, it is possible that GI*sul2* interacts with IncP1-α or related plasmids. The similarity between the IncP1-α *trbJ*, *trbK* and *trbL* genes and the genes now known to belong to GI*sul2* had been noted previously (Chiu and Thomas, 2004). However, at the time, the context of the genes had not been examined and GI*sul2* had not yet been discovered.

3.6. Stability of GI*sul2* in pIP40a

The *int* gene of GI*sul2* encodes an Int (integrase, Pfam PF00589) protein that shares 89% amino acid identity with the SlpA integrase of the CP4-57 phage (Kirby et al., 1994). Interestingly, GI*sul2* also includes a gene, *alpA*, encoding an AlpA transcriptional regulator (Pfam PF05930, GenPept AOX98904), which shares 84% aa identity with the AlpA (MerR-family) DNA binding protein of the CP4-57 prophage (Trempy et al., 1994). In the CP4-57 phage, AlpA positively regulates the SlpA integrase (Trempy et al., 1994), and increased SlpA leads to excision and loss of the prophage (Kirby et al., 1994). Hence, it is likely that Int of GI*sul2* is regulated by AlpA in the same manner. Related Int and AlpA proteins have been found in a group of integrative elements that are found at the end of the *guaA* guanine monophosphate (GMP) synthase gene (Song et al., 2012).

As GI*sul2* is an integrative element, under appropriate conditions it should be able to excise from the plasmid or chromosome in which it resides. To examine the stability of GI*sul2* (Su^R) in pIP40a, *E. coli* cells containing pIP40a were grown overnight (~22 generations) without sulphonamide selection. No sulphonamide-susceptible cells were detected among 600 colonies screened (200 colonies from each of three independent experiments), indicating that all cells tested had retained GI*sul2*. To test for low-frequency excision of the GI, whole cell DNA extracted after a single cycle of growth without antibiotic selection of cells containing pIP40a was used as a template in a PCR using primers RH1879 and R3R that face outwards from the *int* (RH1879) and *sul2* (R3R) ends of GI*sul2*, and would be expected to generate an amplicon of 788 bp if the excised circular form of GI*sul2* (or a tandem duplication) is present. No amplicon was produced from whole cell DNA from three independent cultures of cells containing pIP40a, and re-amplification using the same primers also failed to produce an amplicon. Hence, GI*sul2* is stable in pIP40a and specific conditions may be required to

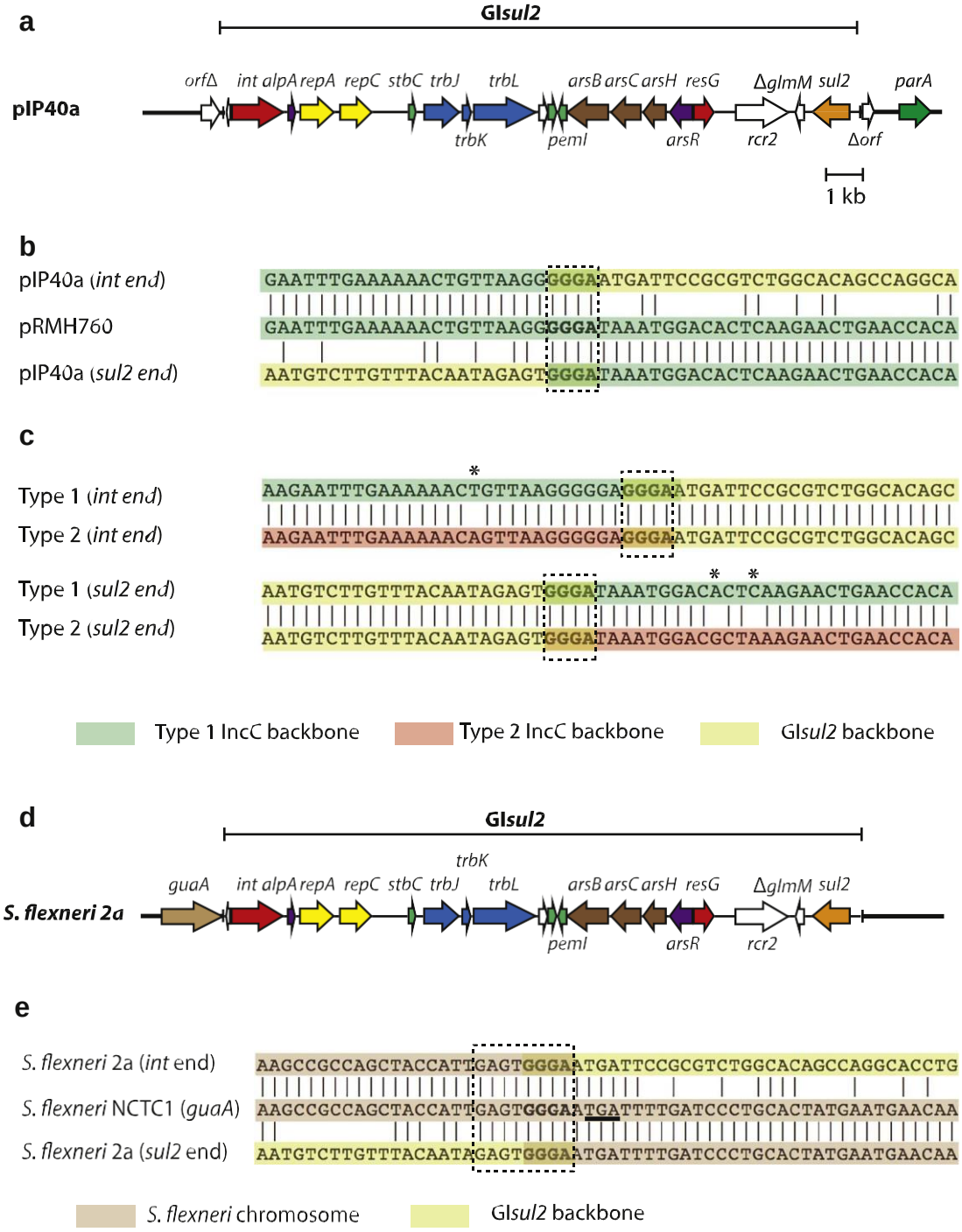

Fig. 4. GI*sul2*. a) Structure of GI*sul2* in pIP40a. A horizontal bar indicates the extent of GI*sul2*. Genes coding for proteins of known function are named above and colored according to following key: green, plasmid maintenance; purple, regulation; yellow, replication; brown, heavy metal resistance; white, no known function; orange, antibiotic resistance; red, DNA metabolism; blue, conjugative transfer. GI*sul2* in pIP40a was drawn to scale from GenBank accession no. KX709966. b) Alignment of the *int* and *sul2* ends of GI*sul2* with the uninterrupted IncC backbone (pRMH760). Shared bases are denoted by a vertical line. The origin of each segment is denoted by shading according to the key below. A dotted box surrounding bold letters indicates the bases involved in the integration of GI*sul2*. pRMH760 sequence was retrieved from GenBank accession nos. KF976462. c) Alignment of the type 1 and type 2 IncC backbone sequences surrounding GI*sul2*. Asterisks denote SNPs between the type 1 and type 2 backbones. Type 1 and type 2 sequences were retrieved from GenBank accession nos. JX709966 and JQ010984, respectively. d) Structure of GI*sul2* in the *S. flexneri* 2a chromosome. GI*sul2* in *S. flexneri* 2a was drawn to scale from GenBank accession no. AW014073. e) Alignment of the *int* and *sul2* ends of GI*sul2* with the *guaA* gene. The uninterrupted *S. flexneri* NCTC1 chromosome was retrieved from GenBank accession no. LM651928. The thick horizontal bar under the *guaA* sequence denotes the position of the *guaA* stop codon. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

induce the excision of GI*sul2*.

3.7. GIsul2 was acquired independently by type 1 and type 2 IncC plasmids

We have previously defined the four bases (GGGA) that are found at both the left and right boundaries of GI*sul2* in the hybrid IncC plasmid p39R861-4 (Anantham et al., 2015), and the same bases are present at each end of GI*sul2* in pIP40a (Fig. 4b). Only one GGGA is derived from the IncC backbone and the second is from GI*sul2*. The backbones of type 1 and type 2 IncC plasmids differ by approximately 1 bp/100 bp

Table 3
Features of GI*sul2*.

Feature/Gene	Position[a]	Size (bp)	Product	Size (aa)[b]	GenPept[c]	Pfam	Description
orf[d]	43–219	177	orf	58	AOX98903	–	Hypothetical protein
int	210–1508	1302	Int	433	AOE47829	00589	Integrase
alpA	1659–1880	222	AlpA	73	AOX98904	05930	Predicted DNA-binding transcriptional regulator
repA	1885–2712	828	RepA	275	AOE47830	13481	Replication initiation protein
repC	2843–3580	738	RepC	245	AOE47831	06504	Replication initiation protein
stbC	4613–4801	189	StbC	62	AOX98905	–	Plasmid stability protein
trbJ	4870–5718	852	TrbJ	283	AOE47832	07996	Conjugative transfer protein
trbK	5734–5970	237	TrbK	78	AOX98906	10907	Entry exclusion protein
trbL	5981–7408	1428	TrbL	475	AOX98907	04610	Conjugative transfer protein
orf	7413–7661	249	orf	82	AOX98908	–	Hypothetical protein
orf	7725–7949	225	orf	74	AOX98909	12703	Putative toxin of toxin-antitoxin type 1 system
pemI	7986–8189c	204	PemI	67	AOX98910	–	Putative programmed cell death antitotin
arsB	8304–9371c	1068	ArsB	355	AOE47834	01758	Arsenate/arsenite transporter protein
arsC	9368–9874c	507	ArsC	168	AOE47835	01451	Arsenate/arsenite reductase protein
arsH	9871–10638c	768	ArsH	255	AOE47836	03358	Arsenate/arsenite resistance protein
arsR	10,635–10967c	333	ArsR	110	AOE47837	12840	Arsenical resistance operon repressor
resG	11,115–11,852	738	ResG	245	AOE47838	02796	Resolvase
CR2	12,166–14,012	1847	–	–	–	–	Mobile element CR2
rcr2	12,284–13,777	1494	Rcr2	497	AOE47839	14319	Mobilization protein
sul2	14,342–15157c	816	Sul2	271	AOE47841	00809	Confers resistance to sulfonamide antibiotics

[a] Position 1 corresponds to position 29940 of GenBank accession number KX709966. The letter "c" after the base range indicates the gene/orf is on the complimentary strand.
[b] Length of product in amino acids (aa).
[c] GenPept from GenBank accession number JX709966.
[d] Open reading frame.

(Harmer and Hall, 2014), and the presence of GI*sul2* or a GI*sul2* remnant in the same position and orientation in both type 1 and type 2 plasmids suggested that this location may be a site recognised by the GI*sul2* Int. Examination of 1 kb of plasmid backbone sequence on either side of GI*sul2* in pIP40a (type 1) and the GI*sul2* remnant in R55 (type 2) revealed that it corresponds to the type 1 and type 2 backbones, respectively. In fact, the closest single basepair differences are 8 away from the conserved GGGA on one side and 10 bp away on the other (Fig.4c). If GI*sul2* had been acquired once then spread via homologous recombination the crossovers would have been within these 8 and 10 bp spans. Hence, this element has been independently acquired by the two lineages.

3.8. Distribution of GI*sul2*

We previously reported that complete copies of GI*sul2* were found in bacterial chromosomes of four species and in two plasmids (Anantham et al., 2015; Nigro and Hall, 2011). In the chromosomes, GI*sul2* was located in the end of the core *guaA* GMP synthetase gene (Anantham et al., 2015; Nigro and Hall, 2011). An updated search for sequences in the GenBank non-redundant database (October 2016), and a recent report of a complete GI*sul2* in the draft genome of *Acinetobacter baumannii* ATCC 19606 (Hamidian and Hall, 2017), revealed that complete copies of GI*sul2* are now found in a total of 11 chromosomes from a variety of Gram negative bacterial species (Table 4). In nine out of the 11 chromosomes, GI*sul2* is located in the end of the core *guaA* gene (Fig. 4d), always in the same orientation, and this is clearly the preferred target site. GI*sul2* inserts 4 bp from the 3'-end of *guaA* within an 8-bp concensus sequence (GAGTGGGA, Fig. 4e), which was found at both the left and right boundaries of the element and this sequence was also noted for related elements that target this site (Song et al., 2012). The next four bases at the *int* end of GI*sul2* (ATGA) restore the stop codon in *guaA*.

In *A. baumannii* strains 3207 (GenBank accession number CP015364) and ATCC19606 (Hamidian and Hall, 2017), the sequence surrounding GI*sul2* is not normally part of the *A. baumannii* chromosome, and GI*sul2* appears to be in a larger genomic island. Complete copies of GI*sul2* were also found in five plasmid sequences in addition to pIP40a (Table 4). The plasmids belong to incompatibility groups IncB/O, and IncX. Two plasmids from *Acinetobacter* species are of an undetermined type. The same four bases are also conserved surrounding GI*sul2* in the plasmids of different incompatibility types listed in Table 4, and these too may represent Int-specific sites.

4. Discussion

pIP40a is one of the oldest known IncC plasmids recovered to-date. However, pIP40a had already acquired a novel mercury resistance transposon, Tn*6182*, as well as Tn*6023* (conferring kanamycin and neomycin resistance), a partial copy of Tn*3* (conferring resistance to ampicillin), Tn*1000*, and GI*sul2*. pIP40a represents a lineage of type 1 IncC plasmids that is yet to acquire the ARI-A resistance island that plays such a key role in resistance gene accumulation, and pDGO100, from 1981 (Harmer et al., 2016b), is currently the earliest known example of a type 1 plasmid containing ARI-A. However, pDGO100 does not contain ARI-B and therefore cannot be directly descended from pIP40a. As GI*sul2* has been introduced on at least two separate occasions into the same site in a type 1 and type 2 IncC plasmid, it may have been acquired by the type 1 lineage on more than one occasion. The ancestor of the type 1 lineage that contains neither the ARI-A nor the ARI-B islands is still to be discovered. The Murray Collection of pre-antibiotic era strains recovered as early as 1917 was recently sequenced (Baker et al., 2015). Though it may have been a source of ancestral IncC plasmids, a search of the sequencing metadata showed that no IncC plasmids were detected in the collection.

Subsequent to the release of our pIP40a sequence (Harmer and Hall, 2016), another pIP40a sequence was reported (Szabo et al., 2016) (GenBank accession number KX156772). Interestingly, whilst the sequences are very similar, there are some differences that are likely to have accumulated over time, during storage and passage of the plasmid in different laboratories. The second sequence lacks the IS*3*, present in the *rhs1* gene of our sequence, and contains three additional IS not present in our version of pIP40a; an IS*2* inserted in a gene encoding a hypothetical protein (orf 21 in Fig. 1), an IS*186* in an intergenic region in the backbone, between orf106 and *pri*, and an IS*150* in the *int* gene of GI*sul2*. The two backbones also differ by six single base pair differences. Comparison to other sequenced IncC plasmids showed that the six differences are unique to our pIP40a sequence, with all other sequences sharing the same bases as GenBank accession number KX156772. One of these single base substitutions occurs in an intergenic region, and one

Table 4
Examples of complete GI*sul2* in sequenced strains.

Organism	Host/Plasmid name	Year of isolation	GI*sul2* location[a]	Insertions	GenBank accession
S. flexneri	2a str 2457T	1954	C - *guaA*	–	AE014073
S. flexneri	G1663	–	C - *guaA*	–	CP007037
P. stuartii	ATCC 33672	–	C - *guaA*	–	CP008920
E. cloacae	ATCC 13047	1890	C - *guaA*	Tn*5393*	CP001918
S. granuli	TFA	1995	C - *guaA*	–	CP012199
M. morganii	FDAARGOS_172	2014	C - *guaA*	–	CP014026
A. xylosoxidans	NCTC10807	1971	C - *guaA*	–	LN831029
E. coli	6409	2012	C - *guaA*	–	CP010371
Pseudomonadaceae spp	E5571	1979	C - *guaA*	–	CP012365
A. baumannii	3207	2008	C[b]	–	CP015364
A. baumannii	ATCC19606	1948	C[c]	–	–[d]
P. aeruginosa	pIP40a	1969	P (IncC)	–	KX709966[e]
E. coli	R16a	1966	P (IncC)	–	KX156773
E. coli	pHUSEC411-like	–	P (IncB/O)	–	HG428756
A. calcoaceticus	NCTC7364	2014	P	ISAba1	LT605060
A. baumannii	ATCC17978-mff	1951	P	ISAba1	CP012005
M. morganii	R485	1972	P (IncX)	–	HE577112

[a] C - *guaA* denotes insertion into the end of the *guaA* gene in the chromosome. P denotes plasmid. Name in brackets indicates plasmid incompatibility group, if known.
[b] Sequence surrounding GI*sul2* is not *guaA* and is unique to this strain.
[c] GI*sul2* is located within a larger genomic island in the chromosome.
[d] Draft genome. See Hamidian and Hall, 2017.
[e] See also GenBank accession number KX156772.

is a missing base pair in *pri* that introduces a frame shift, extending the 576 aa Pri to a 586 aa protein. Two other single base substitutions are silent mutations in *stbA* and orf1832. The remaining two base pair differences result in the leucine usually found at position 113 of Int being replaced with an arginine, and the lysine at position 293 of Dcm3 being replaced with an arginine.

The same study also reported the sequence of R16a (Szabo et al., 2016), recovered three years earlier in 1966, from the St-Antoine Hospital in Paris (Chabbert et al., 1972). This type 1 IncC plasmid also includes a complete copy of GI*sul2* and does not contain an ARI-A island. It carries Tn*6182* in the same position as seen in pIP40a, but the Tn*6182* is interrupted by additional transposons that have introduced genes conferring resistance to ampicillin (Tn*1*, bla_{TEM-1}), and kanamycin and neomycin (Tn*6020*, *aphA1b*). Hence, Tn*6020*, which was first detected in 2009 (Post and Hall, 2009) had been circulating much earlier, likely in response to the use of kanamycin, as had Tn*6023* seen in pIP40a. These transposons continue to be found in modern day plasmids of various types. In contrast, this is the only report of Tn*6182*.

The analysis of the pIP40a sequence also uncovered previously unnoticed functions encoded by genes in the genomic island GI*sul2*. An apparently complete arsenate and arsenite resistance operon does not appear to confer resistance to these heavy metals. The similarity between the Int protein of GI*sul2* and the SlpA integrase of the CP4-57 phage, which is known to be active (Kirby et al., 1994), indicates the role that *int*, and the adjacent *alpA*, are likely to play in the biology of GI*sul2*. However, we were unable to detect excision and further experiments will be required to confirm the role of AlpA in the regulation of Int, and the subsequent Int-mediated excision of GI*sul2*.

The significant levels of amino acid identity between the TrbJ, TrbK and TrbL proteins of GI*sul2* and those found in IncP plasmids suggests that GI*sul2* may be mobilized by IncP plasmids in much the same way that IncA and IncC plasmids are able to mobilize SGI1 and SGI2 type *Salmonella* genomic islands (Hall, 2010; Harmer et al., 2016a; Douard et al., 2010; Doublet et al., 2005; Kiss et al., 2015; Carraro et al., 2014a). This warrants further investigation. GI*sul2* also includes three genes encoding toxin-antitoxin proteins and a stability protein, which may be involved in the stable maintenance of the island once it has integrated into the chromosome or a plasmid.

Abbreviations

GI	genomic island
IR	inverted repeats
IS	insertion sequence
ORF	open reading frame
SNP	single nucleotide polymorphism
TSD	target site duplications

Funding information

This work and CJH were supported by NHMRC grant 1086267. MH was supported by NHMRC grant 1079616.

References

Anantham, S., Harmer, C.J., Hall, R.M., 2015. p39R861-4, a type 2 A/C plasmid carrying a segment from the A/C plasmid RA1. Microb. Drug Resist. 21, 571–576.
Bailey, J.K., Pinyon, J.L., Anantham, S., Hall, R.M., 2010. Commensal *Escherichia coli* of healthy humans: a reservoir for antibiotic-resistance determinants. J. Med. Microbiol. 59, 1331–1339.
Bailey, J.K., Pinyon, J.L., Anantham, S., Hall, R.M., 2011. Distribution of the blaTEM gene and blaTEM-containing transposons in commensal *Escherichia coli*. J. Antimicrob. Chemother. 66, 745–751.
Baker, K.S., Burnett, E., Mcgregor, H., Deheer-Graham, A., Boinett, C., Langridge, G.C., Wailan, A.M., Cain, A.K., Thomson, N.R., Russell, J.E., Parkhill, J., 2015. The Murray collection of pre-antibiotic era Enterobacteriacae: a unique research resource. Genome Med. 7, 97.
Broom, J.E., Hill, D.F., Hughes, G., Jones, W.A., Mcnaughton, J.C., Stockwell, P.A., Petersen, G.B., 1995. Sequence of a transposon identified as Tn*1000* (gamma delta). DNA Seq. 5, 185–189.
Cai, J., Salmon, K., Dubow, M.S., 1998. A chromosomal *ars* operon homologue of *Pseudomonas aeruginosa* confers increased resistance to arsenic and antimony in *Escherichia coli*. Microbiology 144, 2705–2713.
Cain, A.K., Hall, R.M., 2011. Transposon Tn*5393*e carrying the *aphA1*-containing transposon Tn*6023* upstream of *strAB* does not confer resistance to streptomycin. Microb. Drug Resist. 17, 389–394.
Carattoli, A., Miriagou, V., Bertini, A., Loli, A., Colinon, C., Villa, L., Whichard, J.M., Rossolini, G.M., 2006. Replicon typing of plasmids encoding resistance to newer beta-lactams. Emerg. Infect. Dis. 12, 1145–1148.
Carraro, N., Matteau, C., Luo, P., Rodrigue, S., Burrus, V., 2014a. The master activator of IncA/C conjugative plasmids stimulates genomic islands and multidrug resistance dissemination. PLoS Genet. 10, e1004714.
Carraro, N., Sauve, M., Matteau, D., Lauzon, G., Rodrigue, S., Burrus, V., 2014b. Development of pVCR94DeltaX from *Vibrio cholerae*, a prototype for studying multidrug resistant IncA/C conjugative plasmids. Front. Microbiol. 5, 44.
Chabbert, Y.A., Scavizzi, M.R., Witchitz, J.L., Gerbaud, G.R., Bouanchaud, D.H., 1972. Incompatibility groups and the classification of *fi* - resistance factors. J. Bacteriol. 112, 666–675.

Chiu, C.M., Thomas, C.M., 2004. Evidence for past integration of IncP-1 plasmids into bacterial chromosomes. FEMS Microbiol. Lett. 241, 163–169.
Datta, N., Hedges, R.W., 1972. R factors identified in Paris, some conferring gentamicin resistance, constitute a new compatibility group. Ann. Inst. Pasteur (Paris) 123, 849–852.
Datta, N., Hedges, R.W., 1973. R factors of compatibility group A. J. Gen. Microbiol. 74, 335–337.
Douard, G., Praud, K., Cloeckaert, A., Doublet, B., 2010. The *Salmonella* genomic island 1 is specifically mobilized *in trans* by the IncA/C multidrug resistance plasmid family. PLoS One 5, e15302.
Doublet, B., Boyd, D., Mulvey, M.R., Cloeckaert, A., 2005. The *Salmonella* genomic island 1 is an integrative mobilizable element. Mol. Microbiol. 55, 1911–1924.
Doublet, B., Boyd, D., Douard, G., Praud, K., Cloeckaert, A., Mulvey, M.R., 2012. Complete nucleotide sequence of the multidrug resistance IncA/C plasmid pR55 from *Klebsiella pneumoniae* isolated in 1969. J. Antimicrob. Chemother. 67, 2354–2360.
Fernandez-Alarcon, C., Singer, R.S., Johnson, T.J., 2011. Comparative genomics of multidrug resistance-encoding IncA/C plasmids from commensal and pathogenic *Escherichia coli* from multiple animal sources. PLoS One 6, e23415.
Groot Obbink, D.J., Ritchie, L.J., Cameron, F.H., Mattick, J.S., Ackerman, V.P., 1985. Construction of a gentamicin resistance gene probe for epidemiological studies. Antimicrob. Agents Chemother. 28, 96–102.
Haase, J., Kalkum, M., Lanka, E., 1996. TrbK, a small cytoplasmic membrane lipoprotein, functions in entry exclusion of the IncP alpha plasmid RP4. J. Bacteriol. 178, 6720–6729.
Hall, R.M., 2010. *Salmonella* genomic islands and antibiotic resistance in *Salmonella enterica*. Future Microbiol 5, 1525–1538.
Hamidian, M., Hall, R.M., 2017. *Acinetobacter baumannii* ATCC 19606 carries GI*sul2* in a genomic island located in the chromosome. Antimicrob. Agents Chemother. 61 (e01991-16).
Hancock, S.J., Phan, M.D., Peters, K.M., Forde, B.M., Chong, T.M., Yin, W.F., Chan, K.G., Paterson, D.L., Walsh, T.R., Beatson, S.A., Schembri, M.A., 2017. Identification of IncA/C plasmid replication and maintenance genes and development of a plasmid Multilocus sequence typing scheme. Antimicrob. Agents Chemother.. http://dx.doi.org/10.1128/AAC.01740-16. (In press).
Harmer, C.J., Hall, R.M., 2014. pRMH760, a precursor of A/C_2 plasmids carrying bla_{CMY} and bla_{NDM} genes. Microb. Drug Resist. 20, 416–423.
Harmer, C.J., Hall, R.M., 2015. The A to Z of A/C plasmids. Plasmid 80, 63–82.
Harmer, C.J., Hall, R.M., 2016. PCR-based typing of IncC plasmids. Plasmid 87-88, 37–42.
Harmer, C.J., Hamidian, M., Ambrose, S.J., Hall, R.M., 2016a. Destabilization of IncA and IncC plasmids by SGI1 and SGI2 type *Salmonella* genomic islands. Plasmid 87-88, 51–57.
Harmer, C.J., Partridge, S.R., Hall, R.M., 2016b. pDGO100, a type 1 IncC plasmid from 1981 carrying ARI-A and a Tn*1696*-like transposon in a novel integrating element. Plasmid 86, 38–45.
Kaur, S., Kamli, M.R., Ali, A., 2011. Role of arsenic and its resistance in nature. Can. J. Microbiol. 57, 769–774.
Kirby, J.E., Trempy, J.E., Gottesman, S., 1994. Excision of a P4-like cryptic prophage leads to Alp protease expression in *Escherichia coli*. J. Bacteriol. 176, 2068–2081.
Kiss, J., Papp, P.P., Szabo, M., Farkas, T., Muranyi, G., Szakallas, E., Olasz, F., 2015. The master regulator of IncA/C plasmids is recognized by the *Salmonella* genomic island SGI1 as a signal for excision and conjugal transfer. Nucleic Acids Res. 43, 8735–8745.
Liebert, C.A., Hall, R.M., Summers, A.O., 1999. Transposon Tn*21*, flagship of the floating genome. Microbiol. Mol. Biol. Rev. 63, 507–522.
Nicolas, E., Lambin, M., Dandoy, D., Galloy, C., Nguyen, N., Oger, C.A., Hallet, B., 2015. The Tn3-family of replicative transposons. Microbiol Spectr. 3.
Nigro, S.J., Hall, R.M., 2011. GI*sul2*, a genomic island carrying the *sul2* sulphonamide resistance gene and the small mobile element CR2 found in the *Enterobacter cloacae* subspecies cloacae type strain ATCC 13047 from 1890, *Shigella flexneri* ATCC 700930 from 1954 and *Acinetobacter baumannii* ATCC 17978 from 1951. J. Antimicrob. Chemother. 66, 2175–2176.
Pansegrau, W., Lanka, E., Barth, P.T., Figurski, D.H., Guiney, D.G., Haas, D., Helinski, D.R., Schwab, H., Stanisich, V.A., Thomas, C.M., 1994. Complete nucleotide sequence of Birmingham IncP alpha plasmids. Compilation and comparative analysis. J. Mol. Biol. 239, 623–663.
Partridge, S.R., Hall, R.M., 2005. Evolution of transposons containing blaTEM genes. Antimicrob. Agents Chemother. 49, 1267–1268.
Partridge, S.R., Brown, H.J., Stokes, H.W., Hall, R.M., 2001. Transposons Tn1696 and Tn21 and their integrons In4 and In2 have independent origins. Antimicrob. Agents Chemother. 45, 1263–1270.
Post, V., Hall, R.M., 2009. AbaR5, a large multiple-antibiotic resistance region found in *Acinetobacter baumannii*. Antimicrob. Agents Chemother. 53, 2667–2671.
Sampei, G., Furuya, N., Tachibana, K., Saitou, Y., Suzuki, T., Mizobuchi, K., Komano, T., 2010. Complete genome sequence of the incompatibility group I1 plasmid R64. Plasmid 64, 92–103.
Siguier, P., Perochon, J., Lestrade, L., Mahillon, J., Chandler, M., 2006. ISfinder: the reference centre for bacterial insertion sequences. Nucleic Acids Res. 34, D32–D36.
Song, L., Pan, Y., Chen, S., Zhang, X., 2012. Structural characteristics of genomic islands associated with GMP synthases as integration hotspot among sequenced microbial genomes. Comput. Biol. Chem. 36, 62–70.
Szabo, M., Nagy, T., Wilk, T., Farkas, T., Hegyi, A., Olasz, F., Kiss, J., 2016. Characterization of Two Multidrug-resistant IncA/C Plasmids from the 1960s by Using Oxford Nanopore MinION Sequencer Device. (Antimicrob Agents Chemother).
Trempy, J.E., Kirby, J.E., Gottesman, S., 1994. Alp suppression of Lon: dependence on the *slpA* gene. J. Bacteriol. 176, 2061–2067.
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., Ravel, J., 2007. Multiple antimicrobial resistance in plague: an emerging public health risk. PLoS One 2, e309.
Witchitz, J.L., Chabbert, Y.A., 1971. High level transferable resistance to gentamicin. J. Antibiot. (Tokyo) 24, 137–139.