p39R861-4, a type 2 A/C$_2$ plasmid carrying a segment from the A/C$_1$ RA1

Sashindran Anantham, Christopher J. Harmer, Ruth M. Hall*

School of Molecular Bioscience, The University of Sydney, NSW 2006, Australia

Running title: p39R861-4, a hybrid A/C plasmid

*Corresponding author.

Mailing address: School of Molecular Bioscience, Molecular Bioscience Building G08, The University of Sydney, NSW 2006, Australia

Phone: 61-2-9351-3465

Fax: 61-2-9351-5858

E-mail: ruth.hall@sydney.edu.au

Abstract

The largest plasmid in the strain 39R861, which is used as a plasmid size standard, was recovered by conjugation and sequenced to determine its exact size. Plasmid p39R861-4 transferred at high frequency. Though reported to be the A/C_1 plasmid RA1, p39R861-4 is a 155794 bp Type 2 A/C_2 plasmid in which a 39 kb segment derived from RA1 that includes a relative of the RA1 resistance island replaces 26.5 kb of the Type 2 backbone. p39R861-4 includes a single copy of IS*10* and two resistance islands with a CR2-*sul2* region in each of them. The 84 kb of backbone between the resistance islands is inverted relative to other known A/C plasmids and this inversion has arisen via recombination between the CR2-*sul2* regions that are inversely oriented. The resistance islands prior to inversion were one related to but longer than that found in RA1, and a form of the ARI-B island identical to one found in the A/C_2 plasmid R55. They contain genes conferring resistance to tetracycline (*tetA*(D)), sulphonamides (*sul2*) and florfenicol and chloramphenicol (*floR*). The *tet*(D) determinant is flanked by two IS*26* in a transposon-like structure named Tn*tet*(D). Both resistance islands contain remnants of the two ends of the integrative element GI*sul2*, consistent with the *sul2* gene being mobilized by GI*sul2* rather than by CR2.

Introduction

Escherichia coli strain 39R861 was constructed prior to 1986 as a plasmid size standard[20] and is routinely used to measure the size of closed circular plasmids. The four plasmids it carries are reported to be the IncA/C reference plasmid RA1, plasmid X, the IncW plasmid pSa and NTP168, with sizes of 154 kb, 66 kb, 38 kb and 7 kb, respectively, calculated from original size estimates of 98, 42, 23.9 and 4.6 MDa, respectively.[21] Despite ongoing use as a standard, the size of each plasmid has never been precisely determined. RA1 is the only plasmid said to be present in 39R861 that has been sequenced (GenBank accession number FJ705807).[10] The size of pSa can be estimated to be 40.1 kb from the IncW R388 sequence (GenBank accession number BR000038) and the sequence of the pSa integron (GenBank accession number NG_036033).

Plasmids of the A/C group were among the earliest plasmids associated with antibiotic resistance in Gram-negative bacteria. Plasmid RA1 was recovered in 1971 from the fish pathogen *Aeromonas liquafaciens*[1] and is the IncA reference.[5] RA1, carrying *sul2* and *tet*(D), and a deletion derivative RAx were sequenced in 2009.[10] The sequenced version of RA1 appears to have an IS*26*-mediated deletion within its resistance island relative to RAx (see Figure 5 in reference 12). RAx also appears to have acquired the *aphA1* kanamycin and neomycin resistance determinant in a 2.9 kb insertion (bases 32533-35414 in GenBank accession number FJ705806) that is bounded by 24 bp perfect inverted repeats (IR) and flanked by 9 bp direct repeats (DR).[12] This region is not found in RA1. Both of these differences are likely to have arisen after the initial isolation.

A/C plasmids are now designated A/C_1 or A/C_2 based on the sequence of their backbones which share only 85-94% identity at the DNA level.[10] RA1 is the only

sequenced native A/C_1 plasmid, whilst A/C_2 is represented by over 65 complete plasmid sequences in GenBank (4th April 2015). The A/C_2 plasmids can be further divided into two distinct types, Type 1 and Type 2 that have diverged via the accumulation of SNPs, via two regions of replacement internal to large open reading frames, and via the accumulation of insertions or deletions within the backbone.[11] Many A/C_2 plasmids include one or more islands made up of multiple genes conferring resistance to antibiotics. Some A/C_2 plasmids of both types carry a resistance island, ARI-B, which always contains the sulphonamide resistance determinant *sul2* and part or all of the small mobile element CR2, and can contain additional IS*26*-associated resistance genes.[11-13,15,22]

Here, we have recovered the A/C plasmid from 39R861 by conjugation, determined its sequence and compared it to the available RA1 and A/C_2 sequences.

Materials and methods

Bacterial strains and conjugation

Plasmid p39R861-4 was transferred from *E. coli* 39R861[20] to a rifampicin resistant derivative of *E. coli* UB5201 (*pro met recA56 gyrA*)[14] by mixing equal amounts of stationary phase cultures on L-agar and growth overnight at 37ºC. Cells were resuspended in 0.9% (w/v) sterile saline, serially diluted, and transconjugants recovered on plates containing rifampicin to select against the donor and for the recipient, and sulfamethoxazole (Su) to select for transfer of RA1 and pSa. Su^R transconjugants were screened for resistance to tetracycline (Tc), chloramphenicol (Cm), florfenicol (Fl), kanamycin (Km), streptomycin (Sm) and spectinomycin (Sp), and for the presence of plasmid replicons using PCR-based Replicon Typing (PBRT) as described previously.[16] A transconjugant containing both an A/C and an F replicon (plasmid X) was further

mated with *E. coli* UB1637 (Sm^R) to separate the two plasmids. The resulting transconjugants were screened using PBRT and their resistance phenotypes determined by patching onto plates containing Cm, Fl, Tc and Su. A $Cm^RFl^RSu^RTc^R$ transconjugant containing only the A/C replicon was mated with UB5201 (Rif^R) to determine its transfer frequency. Conjugation frequencies were calculated as the number of transconjugants per donor from three independent determinations.

Polymerase chain reaction

Polymerase chain reactions (PCR) were performed on approximately 20 ng of diluted plasmid DNA. PCR conditions were as described previously.[24] Published primers[2] were used to detect antibiotic resistance genes. Amplicons were resolved by electrophoresis on 1% (w/v) agarose gels with molecular weight standards, stained with ethidium bromide, and visualized using a GelDoc1000 image analysis station (BioRad).

Plasmid recovery and sequencing

Total plasmid DNA was extracted and sequenced on an Ion Torrent platform (Life Technologies) as previously described.[11] The sequencing reads (62-fold coverage) were assembled *de novo* into contigs using Geneious version 6.1.6 (Biomatters). PCR and sequencing was used to join contigs that had separated on repeated sequences, namely IS*4321*, IS*26*, *sul2* and CR2. Primers are listed in Table S1. Overlapping contigs were assembled with Sequencher version 5.1 (Gene Codes). The final assembly was confirmed by comparing the fragments obtained from an *in silico* BamHI or EcoRI digest to the fragments obtained when p39R861-4 was digested using BamHI or EcoRI (New England Biolabs). Residues missing from the majority of reads but present in some were assumed to be due to the use of the Ion Torrent platform, and corrected.

Sequence analysis

The backbone of p39R861-4 was derived by removing the antibiotic resistance regions and the insertion sequence (IS*10*) from the plasmid sequence and compared to the R55 (GenBank accession number JQ010984) and RA1 (GenBank accession number FJ705807) backbones, which were derived in the same way. Sequences were circularized and reopened at the same location, 1,139 bp upstream of *repA*. Pairwise comparisons using the BLAST paired alignment facility (http://blast.ncbi.nlm.nih.gov) were visualized using Artemis Comparison Tool.[4]

Potential coding regions in p39R861-4 were identified and annotated based on the available annotations for the A/C plasmids R55 (GenBank accession number JQ010984), RA1 (GenBank accession number FJ705807), pRMH760 (GenBank accession number KF976462), and others.[9,12] Gene Construction Kit version 2.5 (Textco) was used to create figures to scale.

Nucleotide sequence accession number

The 155,794 bp nucleotide sequence of p39R861-4 was submitted to GenBank under accession no. KP276584.

Results

p39R861-4 is not RA1

The A/C plasmid p39R861-4 was separated from the other plasmids carried in 39R861 by conjugation (see Methods). Transconjugants that carried only an A/C plasmid had acquired resistance to sulfamethoxazole and tetracycline and contained *sul2* and *tet*(D) as expected for RA1. However, these transconjugants were also resistant to chloramphenicol and florfenicol, and the *floR* gene conferring resistance to these

antibiotics is also carried on p39R861-4, though it is not present in RA1. Hence, the A/C plasmid in 39R861 is not RA1 as previously reported.[20] The average frequency of transfer of p39R861-4 from an *E. coli* donor to an *E. coli* recipient was 4.21×10^{-2} ($2.81 \times 10^{-2} – 6.88 \times 10^{-2}$; three independent determinations) transconjugants per donor.

Sequence of p39R861-4

The complete p39R861-4 sequence determined here (GenBank accession no. KP276584) was 155,794 bp, and p39R861-4 contains two resistance regions (RI-1 and RI-2 in Figure 1A) and a single copy of IS*10*. Removal of these insertions generated a backbone of 129,998 bp (Fig. 1A) that includes a replication gene (*repA*), genes for conjugative transfer (*tra*), partitioning (*par*), and a variety of other genes that have been described elsewhere[12] or open reading frames with no known function. However, only 26,559 bp of the p39R861-4 backbone was derived from RA1. Bases 44888 to 67672 and 139973 to 143746 of pR39861-4 (GenBank accession number KP276584) share 99.9% nucleotide identity (two SNPs) with bases 90443 to 113228 and 126346 to 130120 of RA1 (GenBank accession number FJ705807). In RA1, these two segments flank the region that includes the resistance genes *sul2* and *tet*(D) (RI-RA1 in Figure 1B). The remainder of p39R861-4 includes the *repA* gene found in A/C_2 plasmids, and shares 99.9% nucleotide identity (eight SNPs) with bases 1 to 30502, 39489 to 100362, and 158764 to 170810 of the type 2 A/C_2 plasmid R55 (GenBank accession number JQ010984). Of the eight SNPs identified between p39R861-4 and R55, four are unique to R55 (i.e. not found in any other Type 2 A/C_2 plasmid) and four are unique to p39R861-4.

The RA1-derived portion of p39R861-4 is continuous in RA1 (Fig. 1B), interrupted only by RI-RA1, and the R55-derived portion is continuous in R55. An inversion in p39R861-4 has split and separated these continuous regions (Fig. 1A).

Resistance islands in p39R861-4

The first resistance island in p39R861-4 (RI-1 in Fig. 1) is 14,379 bp long and carries genes conferring resistance to chloramphenicol and florfenicol (*floR*), and sulphamethoxazole (*sul2*) (Fig. 2A). The second island (RI-2) is 10,088 bp and contains tetracycline (*tetA*(D)) and sulphonamide (*sul2*) resistance genes (Fig. 2A). The inversion described above appears to have occurred via homologous recombination between inversely oriented copies of CR2-*sul2*, one in ARI-B of the R55-derived segment and one in the RI in the RA1-derived segment (Fig. 2). The resistance islands in the configuration presumed to have existed prior to the inversion are shown in Fig. 2B. The island labeled ARI-B is identical to ARI-B in the Type 2 A/C_2 plasmid R55 (GenBank Accession no. JQ010984)[7] which contains the florfenicol resistance determinant *floR*. The second island is closely related to RI-RA1* (Fig. 2C), differing only by the inversion of the segment located between IS*26*-1 and IS*26*-3 that placed the two *sul2* genes in inverse orientation. We believe RI-RA1* is the likely progenitor of the islands seen in the sequenced RA1 and RAx plasmids. The RI-RA1* island (Fig. 2C) contains IS*4321* and IS*5075* flanking a copy of IS*26* and a partial copy of $tnpA_{5051}$. An IS*26*-mediated deletion has removed 11 bp from the one end of IS*4321*. RI-RA1* contains a 4,372 bp fragment of the *sul2* end of GI*sul2* and two copies of IS*26* flanking *tetA*(D) in a 5,166 bp transposon-like structure we have named Tn*tet*(D). A 550 bp fragment between *resG* and IS*26*-2 does not share identity with any other sequence in GenBank.

Origin of ARI-B

The extent of ARI-B in the p39R861-4 precursor and in R55 was determined by comparison with pRMH760 (GenBank accession number KF976462) which has a complete backbone. The plasmid backbone surrounding ARI-B in R55 and in the p39R861-4 precursor was complete, and the island had inserted precisely into the A/C backbone using a conservative mechanism (Fig. 3). ARI-B contains both the *int* and *sul2* ends of the integrative element GI*sul2*[17] at its extremities (Fig. 3A). Alignment of the *sul2* and *int* ends of the island with the intact backbone of pRMH760 (Fig. 3B) revealed that only three bases from the backbone (GGGA, bases 30067-30070 in GenBank accession number KF976462) were found at both the left and right boundaries of the island. However, as GI*sul2* is an integrating element, one GGGA is derived from the A/C backbone and the second from GI*sul2*, and the recombination crossover has occurred within this span. A subsequent event has presumably led to loss of the internal portion of GI*sul2*.

In addition, ARI-B in p39R861-4 also includes an insertion containing the florfenicol resistance determinant *floR* and a truncated copy of the small mobile element CR2. This segment appears to have been incorporated via homologous recombination with a 3,482 bp circle containing Δ2CR2. The two partial copies of CR2, ΔCR2 (802 bp) and Δ2CR2 (640 bp), have only retained bases from the previously defined[19,24] *ori* end of CR2.

Discussion

Here, sequencing of p39R861-4 has shown that the A/C plasmid in the sizing standard 39R861 is not RA1 as previously believed. Instead, it is a hybrid that has formed initially via the integration of a segment of RA1 into a Type 2 A/C_2 plasmid

identical or closely related to R55. This was followed by large inversions in and between the two resistance regions derived from R55 and RA1. Thus, p39R861-4 is the first sequenced example of hybrid formation between an A/C_1 and an A/C_2 plasmid.

The precise origin of the hybrid p39R861-4 plasmid is unknown. However, early incompatibility and entry-exclusion experiments reported the recovery of recombinant plasmids formed between RA1 and R57b.[5] R57b was isolated in Paris around the same time as R55,[23] and is likely to be closely related or even identical to R55. A subsequent study reported the isolation of a spontaneous RA1 recombinant, named RA1-1, that was resistant to tetracycline, sulphonamides and chloramphenicol.[6] This resistance phenotype can be accounted for by the *floR*, *sul2* and *tet*(D) resistance genes identified here in the resistance islands of p39R861-4. It seems possible that p39R861-4 originated from these early experiments.

We propose that ARI-B in A/C_2 plasmids formed first by incorporation of a complete GI*sul2* into the A/C_2 backbone, followed by loss of the central region and then acquisition of other variable segments (Fig. 3A) or loss of the *int* end and adjacent DNA. Some A/C_2 plasmids contain a segment from the *sul2* end of GI*sul2* with a full copy of CR2 and the adjacent *resG* gene,[13,17] whilst others retain a smaller segment with only a partial copy of CR2. In the case of ARI-B in R55 and the p39R861-4 precursor, *floR* in ARI-B appears to have been acquired via homologous recombination with a circle containing CR2Δ, rather than as a transposon as previously reported.[8]

The close association between CR2 and *sul2* led to the belief that CR2 may be responsible for the mobilization of *sul2*.[18,19] However, it is now known that *sul2*-CR2 can be found close to one end of a larger 15 kb structure called GI*sul2*, an integrating element that targets the 3'-end of the *guaA* gene found in bacterial chromosomes.[17] GI*sul2* has been found in a number of bacterial species, including an *Enterobacter*

cloacae strain recovered in the 1890's.[17] In the course of this analysis, we have found a complete copy of GI*sul2* in additional genomes of *Providencia stuartii* strain ATCC 33672 (GenBank accession number number CP008920) and *Achromobacter xylosoxidans* strain ATCC 27061 (GenBank accession CP006958). We also found a complete copy of GI*sul2* in the IncX plasmid R485 (GenBank accession number HE577112)[3] and in an unnamed B/O plasmid (GenBank accession number HG428756), though it was not identified at the time. Hence, it is now evident that *sul2* is being mobilized by GI*sul2*, rather than by the small mobile element CR2.

References

1. **Aoki T., Egusa S., Ogata Y., and Watanabe T.** 1971. Detection of resistance factors in fish pathogen *Aeromonas liquefaciens*. J Gen Microbiol 65: 343-9.
2. **Bailey J.K., Pinyon J.L., Anantham S., and Hall R.M.** 2010. Commensal *Escherichia coli* of healthy humans: a reservoir for antibiotic-resistance determinants. J Med Microbiol 59: 1331-9.
3. **Burmolle M., Norman A., Sorensen S.J., and Hansen L.H.** 2012. Sequencing of IncX-plasmids suggests ubiquity of mobile forms of a biofilm-promoting gene cassette recruited from *Klebsiella pneumoniae*. PLoS One 7: e41259.
4. **Carver T.J., Rutherford K.M., Berriman M., Rajandream M.A., Barrell B.G., and Parkhill J.** 2005. ACT: the Artemis Comparison Tool. Bioinformatics 21: 3422-3.
5. **Datta N., and Hedges R.W.** 1972. R factors identified in Paris, some conferring gentamicin resistance, constitute a new compatibility group. Ann Inst Pasteur (Paris) 123: 849-52.
6. **Datta N., and Hedges R.W.** 1973. R factors of compatibility group A. J Gen Microbiol 74: 335-7.
7. **Doublet B., Boyd D., Douard G., Praud K., Cloeckaert A., and 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: 2354-60.
8. **Doublet B., Schwarz S., Kehrenberg C., and Cloeckaert A.** 2005. Florfenicol resistance gene *floR* is part of a novel transposon. Antimicrob Agents Chemother 49: 2106-8.
9. **Fernandez-Alarcon C., Singer R.S., and 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.

10. **Fricke W.F., Welch T.J., McDermott P.F., Mammel M.K., LeClerc J.E., White D.G., Cebula T.A., and Ravel J.** 2009. Comparative genomics of the IncA/C multidrug resistance plasmid family. J Bacteriol 191: 4750-7.
11. **Harmer C.J., and Hall R.M.** 2014. pRMH760, a Precursor of A/C_2 Plasmids Carrying *bla*$_{CMY}$ and *bla*$_{NDM}$ Genes. Microb Drug Resist 20: 416-23.
12. **Harmer C.J., and Hall R.M.** 2015. The A to Z of A/C plasmids. Plasmid. In Press.
13. **Harmer C.J., Holt K.E., and Hall R.M.** 2015. A type 2 A/C2 plasmid carrying the *aacC4* apramycin resistance gene and the *erm*(42) erythromycin resistance gene recovered from two *Salmonella enterica* serovars. J Antimicrob Chemother 70: 1021-1025.
14. **Kholodii G., Yurieva O., Mindlin S., Gorlenko Z., Rybochkin V., and Nikiforov V.** 2000. Tn*5044*, a novel Tn*3* family transposon coding for temperature-sensitive mercury resistance. Res Microbiol 151: 291-302.
15. **Kim M.J., Hirono I., Kurokawa K., Maki T., Hawke J., Kondo H., Santos M.D., and Aoki T.** 2008. Complete DNA sequence and analysis of the transferable multiple-drug resistance plasmids (R Plasmids) from *Photobacterium damselae subsp. piscicida* isolates collected in Japan and the United States. Antimicrob Agents Chemother 52: 606-11.
16. **Moran R.A., Anantham S., Pinyon J.L., and Hall R.M.** 2015. Plasmids in antibiotic susceptible and antibiotic resistant commensal *Escherichia coli* from healthy Australian adults. Plasmid.
17. **Nigro S.J., and 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: 2175-6.
18. **Partridge S.R.** 2011. Analysis of antibiotic resistance regions in Gram-negative bacteria. FEMS Microbiol Rev 35: 820-55.
19. **Partridge S.R., and Hall R.M.** 2003. In*34*, a complex In*5* family class 1 integron containing orf513 and *dfrA10*. Antimicrob Agents Chemother 47: 342-9.
20. **Threlfall E.J., Rowe B., Ferguson J.L., and Ward L.R.** 1986. Characterization of plasmids conferring resistance to gentamicin and apramycin in strains of *Salmonella typhimurium* phage type 204c isolated in Britain. J Hyg (Lond) 97: 419-26.
21. **Vivian A.** 1994. Plasmid expansion? Microbiology 140 (Pt 2): 213-4.
22. **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. and others**. 2007. Multiple antimicrobial resistance in plague: an emerging public health risk. PLoS One 2: e309.
23. **Witchitz J.L., and Chabbert Y.A.** 1971. High level transferable resistance to gentamicin. J Antibiot (Tokyo) 24: 137-9.
24. **Yau S., Liu X., Djordjevic S.P., and Hall R.M.** 2010. RSF1010-like plasmids in Australian *Salmonella enterica serovar Typhimurium* and origin of their *sul2-strA-strB* antibiotic resistance gene cluster. Microb Drug Resist 16: 249-52.

Figure legends

FIG. 1 Comparison of the p39R861-4 backbone with RA1 and R55. **(A)** p39R861-4 backbone structure with regions derived from RA1 and R55 indicated above. Regions shared with RA1 are denoted by shading. Vertical arrows indicate the location of the RI-1 and RI-2 resistance islands and a copy of IS*10*. The backbone is drawn to scale from GenBank accession number KP276584, with lengths in kb shown above the line. **(B)** RA1 and R55 backbones drawn to scale from GenBank accession numbers FJ705807 and JQ010984, respectively. Vertical arrows indicate the location of the RI-RA1 resistance island in RA1, and the ARI-B and RI-R55 resistance islands in R55. The region of the R55 backbone replaced in p39R861-4 is shown below. Regions containing genes involved in plasmid replication (*rep*), methylation (*dcm*), partitioning (*par*), transfer (*tra*) and the *ybaA*, *uvrD*, *rhs2* genes are indicated by arrows below the line.

FIG. 2. Evolution of p39R861-4. **(A)** Final p39R861-4 structure, **(B)** structure prior to inversion between two oppositely-oriented copies of *sul2*-CR2, and **(C)** original structure of RI-RA1* prior to inversion between two oppositely-oriented copies of IS*26*. Dashed lines denote backbone derived from R55, and dotted lines denote backbone derived from RA1. Genes and ORFs are indicated by horizontal arrows showing the direction of transcription. Resistance islands, denoted by RI-1, RI-2, ARI-B and RI-RA1*, are drawn to scale. ISs and the small mobile element CR2 are shown as open boxes with a vertical bar indicating the ori end of CR2. IS numbers or names are indicated inside the box. The three copies of IS*26* are numbered. Segments derived from GI*sul2* and Tn*tet(D)* are marked.

FIG. 3. Mobilization of *sul2* into the A/C_2 backbone. **(A)** Insertion of ARI-B into the A/C_2 backbone. The backbone is drawn to scale from pRMH760 (GenBank accession number KF976462), ARI-B from R55 (JQ010984) and GI*sul2* from GenBank accession number CP001918. The small mobile element CR2 is shown as an open box with a vertical bar indicating the ori end of CR2. Genes and ORFs are shown as horizontal arrows indicating the direction of transcription. Regions in ARI-B derived from GI*sul2* are indicated by shading. **(B)** Precise bases associated with the mobilization of *sul2*. Bases at the *int* (bases 30477 to 30546 from GenBank accession number KP276584) and *sul2* (bases 129913 to 129844 from GenBank accession number KP276584) ends of ARI-B are aligned with the intact backbone of pRMH760 (bases 30039 to 30108 from GenBank accession number KF976462). Bases found at both boundaries of the island are shown in bold. Shared bases are denoted by vertical lines.