Complete sequence of pJIE143, a pir-type plasmid carrying IS\textit{Ecp}1-\textit{bla}\text{CTX-M-15} from an \textit{Escherichia coli} ST131 isolate.

Sally R. Partridge\textsuperscript{1*}, Justin A. Ellem\textsuperscript{1}, Sasha G. Tetu\textsuperscript{2}, Zhiyong Zong\textsuperscript{1,3}, Ian T. Paulsen\textsuperscript{2}, Jonathan R. Iredell\textsuperscript{1}

Centre for Infectious Diseases and Microbiology, University of Sydney, Westmead Hospital, Sydney, New South Wales 2145\textsuperscript{1}, Department of Chemistry and Biomolecular Sciences, Macquarie University, Sydney, New South Wales, 2052\textsuperscript{2}, Australia and Department of Infectious Diseases, West China Hospital, Sichuan University, Chengdu, China\textsuperscript{3}

Running title: Complete sequence of pJIE143 carrying \textit{bla}\text{CTX-M-15}
Abstract

pJIE143 (34 kb), from an *Escherichia coli* ST131 isolate, carries *bla*$_{CTX-M-15}$ but could not be typed using the standard PCR-based replicon typing primer set. Complete sequencing revealed a backbone with similarity to IncX plasmids, including a *pir*-like gene encoding a π-like replication protein and iterons related to those of other IncX plasmids. The 2.971 kb *ISEcp1-bla*$_{CTX-M-15}$orf477Δ transposition unit often found within Tn2 is inserted just beyond the end of *pir*, flanked by 5-bp direct repeats.
*bla*$_{\text{CTX-M}}$ now appear to be the dominant family of genes conferring resistance to extended spectrum β-lactamase antibiotics worldwide and *bla*$_{\text{CTX-M-15}}$ is most common in many locations (19).

*bla*$_{\text{CTX-M-15}}$ is generally found in a 2.971 kb *ISEcp1*-mediated transposition unit that also includes a truncated orf477, usually inserted in Tn2 e.g. (17, 21, 23) but also found directly inserted in the backbone of an IncA/C$_2$ IncY plasmid from *Salmonella* (8). Surveys (e.g. 9, 13) indicate that in *Escherichia coli* *bla*$_{\text{CTX-M-15}}$ is predominantly carried by IncF or IncI1 plasmids, but some plasmids carrying this gene could not be typed by standard PCR-based replicon typing (PBRT) (4).

We identified *bla*$_{\text{CTX-M-15}}$ as the dominant ESBL gene in *E. coli* in western Sydney, always linked to *ISEcp1* (24). Fourteen of 19 *E. coli* isolates with different pulsed-field gel electrophoresis (PFGE) profiles gave transconjugants in DH5α carrying *bla*$_{\text{CTX-M-15}}$ and PBRT (4) identified replicons (IncF, n=9; IncI1, n=4) in all but one (24). Isolate JIE143 gave transconjugants resistant to ampicillin (80 µg/ml) at a frequency of 6.7x10$^{-7}$/recipient by filter mating but these failed to give amplicons by PBRT (24). JIE143, isolated from blood <48 h after admission of a patient to Blacktown Hospital, Sydney, in June 2006, was considered community acquired and was found to be phylogenetic group B2 by PCR (5), O25b using antisera (Denka Sieken, Coventry, UK) and by PCR (6) and ST131 by MLST (http://mlst.ucc.ie/mlst/dbs/Ecoli; 22). S1 nuclease/PFGE (2, 17) of DNA from a selected transconjugant (Tx143) revealed a single plasmid of ~35 kb (data not shown), designated pJIE143, which was completely sequenced.

DNA extracted from Tx143 using the QIAGEN (Hilden, Germany) HiSpeed Plasmid Purification system was treated with Plasmid Safe ATP-Dependent DNase (Epicentre Biotechnologies, Madison WI) and 10 ng was amplified with a GenomiPhi v2 DNA kit (GE Heathcare, NJ), following the manufacturers’ protocols. After sodium acetate precipitation purified DNA was resuspended in 45 µl of 10 mM Tris (pH 7.5) and quantified using the Quant-iT PicoGreen dsDNA Assay (Invitrogen,
Carlsbad, CA) following the manufacturer’s protocol. A library prepared from 500 ng of plasmid DNA (Roche Rapid Library Preparation Method Manual, GS FLX Titanium Series, October 2009) was bar-coded with Rapid Library MID adaptors (Roche 454 life sciences, Mannheim, Germany) and pooled with libraries from 11 other plasmids prior to emulsion PCR. Sequencing was carried out on one quarter of a 4-region plate in a GS-FLX sequencer (Roche) (14) at the Ramaciotti Centre, Sydney.

Assembly with Newbler version 2.3 (Roche) gave a 34.345 kb contig (180-fold coverage), confirmed to correspond to the entire plasmid by PCR to link the ends (143-1, CAATGCGGATATCTGTCC; 143-2, GCACGAGCGTTCCTCTCC). The typical 2.971 kb IS\textsubscript{Ecp1}-bla\textsubscript{CTX-M-15}-orf477\Delta transposition unit was identified, flanked by direct repeats (GGATA) indicative of insertion by IS\textsubscript{Ecp1}-mediated transposition using the same alternative IR\textsubscript{R} as previously observed. pJIE143 does not carry any other known antibiotic resistance genes.

BLASTn searches revealed that most of the pJIE143 backbone is closely related to pBS512_33 from \textit{Shigella boydii} (33 kb; GenBank accession no. CP001059) and pCROD2 from \textit{Citrobacter rodentium} (39 kb; FN543504; 18) (Fig. 1A), neither of which carry any known antibiotic resistance genes. As noted for pCROD2 (18), pJIE143 is similarly organized to plasmids in the narrow host-range IncX group found in the \textit{Enterobacteriaceae}. The majority of known IncX plasmids belonging to a subgroup designated IncX1, while the archetypal IncX plasmid R6K exemplifies the rarer IncX2 subgroup (10). The complete sequence of R6K (http://www.sanger.ac.uk/Projects/Plasmids/) has not yet been annotated or published but this plasmid has been well studied and pOLA52 (51.6 kb; GenBank accession no. EU370913), classified as IncX1, has also been analyzed in detail (15). The pJIE143 backbone was annotated by reference to these two plasmids.
The R6K and pOLA52 replication regions include a \textit{pir} gene, encoding the $\pi$ protein for initiation of replication, and three vegetative origins, \textit{oriV-\alpha}, \textit{oriV-\beta} and \textit{oriV-\gamma} (see 20 and references therein). pJIE143 encodes a protein related to $\pi$ of R6K (39\% identical/63\% similar) and pOLA52 (35\%/60\%) and more closely related (75\%/86\%) to the putative $\pi$ protein of pJARS35 from \textit{Yersinia pestis} (35 kb; CP002179). pCROD2 has a quite different replication protein (<29\%/<50\%). The \textit{bis} gene found downstream of \textit{pir} in R6K and pOLA52, which has an uncertain role in replication from \textit{oriV-\beta} (16), is missing from pJIE143. A different open reading frame (orf in Fig. 1A) is found in this position, interrupted after 10 bp by \textit{ISEcp1-bla_{CTX-M-15}}-orf477\Delta (Fig. 1A, C), and equivalent orfs are present in pCROD2 (95\% identical) and pJARS35 (69\% identical).

In R6K \textit{oriV-\gamma} is located upstream of \textit{pir} (Fig. 1A) and contains seven 22-bp itersons to which $\pi$ binds. \textit{oriV-\beta}, downstream of \textit{pir}, has a half-iteron and \textit{oriV-\alpha}, found some distance away near a DnaG binding site, has a single iteron (16). Seven \textit{oriV-\gamma} itersons related to those of R6K were identified in pJIE143 in the expected location (Fig. 1B). A putative \textit{oriV-\beta} half iteron was identified close to the end of \textit{pir} in pJIE143 and although a probable DnaG binding site was also found potential $\alpha$-itersons in this region were not convincing (Fig. 1B and C).

The conjugation region of pJIE143 encodes equivalents (31-55\% identical/46-71\% similar) of Pilx1-Pilx11 involved in pilus synthesis and assembly in R6K (GenBank accession no. AJ006342) and pOLA52. pJIE143 also encodes equivalents of TaxA, the TaxB coupling protein and the TaxC relaxase of R6K and pOLA52 (25-40\%/45-59\%) (16). A putative \textit{eex} (entry exclusion) gene encodes a protein 57\% identical/72\% similar to Eex of pOLA52, but is located in a different position (Fig. 1A). Like R6K (1) and pOLA52 (15), pJIE143 appears to have two origins of transfer \textit{(oriT)} situated in perfect inverted repeats (91-bp; Fig. 1C).

pJIE143 and pCROD2 do not include the \textit{parFG} plasmid partitioning locus found in pOLA52,
but both have other putative \textit{par} genes and a resolvase gene (>93\% identical; \textit{parA}, \textit{parB} and \textit{res} in Fig. 1A). pJIE143 lacks the \textit{stbDE} toxin/antitoxin genes found in pOLA52 and pCROD2, but does have the \textit{hicAB} genes also present pCROD2 (99\% identical). These are proposed to encode another toxin/antitoxin system in which HicA inhibits translation, causing mRNA cleavage, while HicB appears to be an unstable repressor that neutralizes HicA (11). pJIE143 also includes a region 97\% identical to both pCROD2 and pOLA52 containing \textit{hns}, \textit{hha} and \textit{topB} genes, putatively encoding a histone-like nucleotide structuring protein (H-NS), a modulator of H-NS activity (12) and a topoisomerase, respectively. These factors may all contribute to plasmid maintenance.

R6K carries Tn3 (\textit{bla}_{TEM-1a}) inserted in Tn5393 (\textit{strAB}), while pOLA52 carries \textit{oqxAB}, encoding an efflux pump conferring resistance to olaquindox, in an IS26-mediated composite transposon inserted in Tn2 (\textit{bla}_{TEM-1b}). Other "IncX-like" resistance plasmids are also beginning to be identified e.g. pE001 (GenBank accession no. JF776874) carries the \textit{bla}_{TEM-52} extended-spectrum \beta-lactamase gene in Tn2 (3). However, these plasmids may be being missed in surveys, as the R6K-derived probe used for hybridization-based replicon typing did not hybridize with all IncX plasmids (7) and the PBRT IncX primers were designed from this probe (4). These primers flank the \textit{oriV-\gamma} iteron region of R6K and have essentially no similarity with this region in other characterized IncX-like plasmids, would not be expected to detect pOLA52 and did not detect pE001 (3) or pJIE143. Similarly, primers in the pOLA52 or pE001 \textit{pir} gene (3) would not be expected to detect pJIE143.

Thus evaluation of different targets and/or development of a set of primers may be required to enable detection of subsets of "IncX-like" plasmids that appear to have similar organization but quite variable sequences.

\textbf{Nucleotide sequence accession number.} The complete nucleotide sequence of pJIE143 has been submitted to GenBank under accession no. JN194214.
We thank Andrew Ginn for technical assistance. SRP was partly supported by Project grant no. 512396 from the Australian National Health and Medical Research Council. JAE was supported by an Australian Postgraduate Award. ZZ was supported by an Endeavour International Postgraduate Research Scholarship from the Australian Government Department of Education, Science and Training. ITP is supported by a Life Science Research Award provided by the New South Wales Office of Science and Medical Research.
REFERENCES


Fig. 1. (A) Comparison of backbones of pBS512_33 (GenBank accession no. CP001059), pJIE143, pCROD2 (FN543504) and pOLA52 EU370913). Darker shading indicates regions with 99% nucleotide identity, lighter shading 97% identity. The extents and directions of various genes are shown as arrows labelled with the gene name (as described in the text, plus dpp3, DNA distortion polypeptide; actX, antiterminator; dnaJ; chaperone), or the number for pilx genes. The insertion point of the IS\textsubscript{Ecp1}-\textit{bla}_{CTX-M-15}-orf477\Delta transposition unit in pJIE143 is indicated. The vertical arrow labeled IS above pCROD2 indicates the site of insertion of IS\textsubscript{I02}. The arrow labeled ME above pOLA52 indicates the region where several mobile elements and resistance genes are inserted. The approximate positions of proposed \textit{oriV} (small vertical bars labeled \textit{V}) and \textit{oriT} (small diamonds labeled \textit{T}) of pJIE143 and pOLA52 are indicated. (B) \textit{oriV} iterons of R6K and proposed \textit{oriV} iterons of pOLA52 (all taken from ref. 15, except \textit{oriV}-\alpha of R6K, which appears to have an extra T in ref. 16) and pJIE143. Bases conserved in the majority of sequences in each set are in bold type face and bases which differ are in lower case. (C) Sequences surrounding the proposed \textit{oriV}/\textit{oriT} \beta (top) and \alpha (bottom, reverse orientation to part A) regions of pJIE143. The sequences of proposed iterons are shown as in part (B) and by doubled-headed arrows. Stop and start codons of various genes are underlined. Vertical arrows indicate \textit{nic} sites (1). The 5 bp duplicated by insertion of IS\textsubscript{Ecp1}-\textit{bla}_{CTX-M-15}-orf477\Delta are indicated by an oval. The DnaG site in the \alpha region is boxed (16) and "74" indicates 74 bp corresponding to the distance between this site and one end of the \textit{oriV}-\alpha iteron in R6K and pOLA52 (15). Sequences are continuous from one line to the next, but have been broken to enable the 94 bp \textit{oriT} repeats (bold) shown on the lower lines to be aligned.