Identification of a Type IV CRISPR-Cas system located exclusively on *IncHI1B/ IncFIB* plasmids in *Enterobacteriaceae*

Enas Newire [1,2,#], Alp Aydin[3,##], Samina Juma[3], Virve I Enne[3], Adam P. Roberts[1,4,5,*]

[1]UCL Eastman Dental Institute, University College London, 256 Gray's Inn Road, London WC1X 8LD, UK. [2]US. Naval Medical Research Unit-3, Cairo, Egypt. [3]Centre for Clinical Microbiology, University College London, Royal Free Hospital Campus, Rowland Hill Street, London NW3 2PF, UK. [4]Department of Tropical Disease Biology and [5]Centre for Drugs and Diagnostics, Liverpool School of Tropical Medicine, Pembroke Place, Liverpool L3 5QA, UK.

*Corresponding Author: Tel: +44(0)151 705 3247; E-mail: adam.roberts@lstmed.ac.uk

Running Title: A plasmid based CRISPR-Cas system.

Keywords: Type IV-B, *IncFIIK, IncFIB(K),* plasmid incompatibility, Mobile genetic element.

[#] Present address; Division of Biosciences, University College London, Gower Street, London, WC1E 6BT, UK.

[##] Present address; Quadram Institute, James Watson Road, Norwich, NR4 7UQ

Abstract [171/250 words]

During an investigation of CRISPR carriage in clinical, multi-drug resistant, *Klebsiella pneumoniae*, a novel CRISPR-Cas system (which we have designated Type IV-B) was detected on plasmids from two *K. pneumoniae* isolates from Egypt (isolated in 2002-2003) and a single *K. pneumoniae* isolate from the UK (isolated in 2017). Sequence analysis of other genomes available in GenBank revealed that this novel Type IV-B CRISPR-Cas system was present on 28 other plasmids from various *Enterobacteriaceae* hosts and was never found on the chromosome. Type IV-B is found exclusively on *IncHI1B/ IncFIB* plasmids and is associated with multiple putative transposable elements. Type IV-B has a single repeat-spacer array (CRISPR1) upstream of the *cas* loci with some spacers matching regions of conjugal transfer genes of *IncFIIK/ IncFIB(K)* plasmids suggesting a role in plasmid incompatibility. Expression of the *cas* loci was confirmed in available clinical isolates by RT-PCR; indicating the system is active. To our knowledge, this is the first report describing a new subtype within Type IV CRISPR-Cas systems exclusively associated with *IncHI1B/ IncFIB* plasmids.

Importance [121/150 words]

Here, we report the identification of a novel subtype of Type IV CRISPR-Cas that is expressed and exclusively carried by *IncHI1B/ IncFIB* plasmids in *Enterobacteriaceae,* demonstrating unique evolutionarily juxtaposed connections between CRISPR-Cas and mobile genetic elements (MGEs). Type IV-B encodes a variety of spacers showing homology to DNA from various sources, including plasmid specific spacers and is therefore thought to provide specific immunity against plasmids of other incompatible groups (*IncFIIK/ IncFIB(K)*).

The relationship between Type IV-B CRISPR-Cas and MGEs that surround and interrupt the system is likely to promote rearrangement and be responsible for the observed variability of this type. Finally, the Type IV-B CRISPR-Cas is likely to co-operate with other *cas* loci within the bacterial host genome during spacer acquisition.

Introduction [247/250 words]

Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) are widespread, bacterial adaptive, RNA-mediated, immune systems that target invading foreign DNA such as bacteriophages and conjugative plasmids [1, 2]. CRISPR functions through a three-stage process: adaptation involving acquisition of foreign DNA molecules spacers, expression and maturation of the short CRISPR RNAs (crRNAs), and the interference with a cognate invading foreign DNA molecule [3]. To date, CRISPR-Cas systems are classified into 2 classes, 5 Types (I-V) and 33 subtypes [4]. The two classes are differentiated based on the effector module; class 1 utilises multi-protein effector Cas complexes, while class 2 utilise a single-protein effector (either Cas9 or Cpf1) [5, 6]. All types are confirmed, or expected, to provide immunity against invading DNA, while Type III CRISPR-Cas systems can target both DNA and RNA [7, 8].

Type IV was previously called the Unknown Type, due to its rare occurrence and lack of the adaptation module, until an updated classification in 2015 [7, 9]. In 2017, Type IV classification was updated to show an associated repeat-spacer array for a *cas* loci that has *csf1* (*cas8-Like*), *csf2* (*cas7*), *csf3* (*cas5*), *csf4* (*dinG*) and *csf5* (*cas6-Like*) genetic arrangement, respectively [4]. Type IV is the only type to possess *csf4* (*dinG*)[4]. Type IV CRISPR–Cas systems were shown to employ crRNA-guided effector complexes in 2019 [8]. It has been hypothesised

that Type IV is similar to an ancestral innate immune system that gained adaptive ability by associating with a transposon-like element containing *cas1* and *cas2*[3].

Results [186/150 words]

A novel CRISPR-cas family, which we have designated Type IV-B due to the presence of *dinG*, containing a repeat-spacer array, leader sequence and *cas* loci was detected on thirty-one (three clinical isolates and twenty-eight genomes in GenBank) *IncHI1B/ IncFIB(Mar)* plasmids from *Enterobacteriaceae* (Figure 1A, Table S1). Type IV-B *cas* loci show homology to each other, while *csf5* (*cas6)* in Type IV-B shows 100% protein identity to Type I-E *cas6*. Type IV-B CRISPR-Cas systems could be grouped according to the presence of an IS element interrupting the *cas* loci and both groups are associated with multiple MGEs (Figure 1A). We also identified partial *cas* loci (*cas6* and *dinG*) on other *IncHI1B/ IncFIB(Mar)* plasmids (Table S1). A single repeat-spacer CRISPR1[10] array was identified upstream of all *cas* loci. The repeats have a predicted stem-loop secondary-structure likely involved in pre-crRNA processing (Figure 1B, 1C). Spacer-1 in CP018720.1 and spacer-20 in CP014776.1 correspond to *IncFIIK* conjugal transfer genes; *traN* and *traL*, respectively. The Protospacer Adjacent Motif (PAM) alignment revealed the leader-proximal repeat signature conservation (TGCC/TTAT). Finally, RT-PCR demonstrated that Type IV-B *cas* loci genes (*csf2*, *dinG* and *cas6*) are expressed.

Discussion [197/200 words]

Type IV CRISPR-Cas systems are the only ones possessing *dinG*[4], therefore we propose the CRISPR-Cas system described here be designated Type IV-B. Unlike classical Type IV, Type IV-

B lacks *cas8-like* and *cas5*, however, it has *dinG* and *cas7* (involved in interference), and *cas6* (involved in expression and maturation of short crRNAs)[11-13]. Type IV-B has a variable repeat-spacer array and a conserved leader sequence. Expression of Type IV-B genes indicates system activity; likely providing immunity to incoming DNA matching the spacers[8]. The spacers demonstrated conservation and polymorphism and cluster into two main groups (Figure 2A, 2B) both matching DNA from a variety of sources. However, the adaptation module is missing, thus adding new spacers will require *cas1* and *cas2* from other CRISPR-Cas that exist within the *Enterobacteriaceae* genomes. Interestingly, some of Type IV-B spacers match conjugal transfer genes *traN* and *traL* of *IncFIIK/ IncFIB(K)* plasmids, suggesting a role in plasmid incompatibility.

Type IV-B demonstrates a complex evolutionary connection with MGEs in terms of parasitism and immunity[14]. The association between Type IV-B and multiple MGEs, plus the identification of partial *cas* loci on other *IncHI1B/ IncFIB(Mar)* plasmids, indicates that dynamic, MGE mediated rearrangement, of CRISPR-Cas Type IV-B is ongoing.

Conclusion [56/75 words]

To our knowledge, this is the first identification of a CRISPR-Cas system, which we have designated Type IV-B, exclusively associated with plasmids. The system demonstrates an evolutionary association and role for MGEs in dissemination and, additionally, the spacer analysis suggests a role in plasmid incompatibility. We propose updating the CRISPR-Cas system classification to include Type IV-B.

Materials and Methods [125/125 words]

Isolate information Clinical *K. pneumoniae-53* and *K. pneumoniae-65* were isolated from Egyptian university teaching hospitals (2002-2003), and *K. pneumoniae-CR5* from University College London Hospital in the UK (2017).

CRISPR-Cas loci expression RT-PCR was performed using LightCycler® RNA Amplification Kit SYBR Green (Roche Diagnostics Ltd., UK). The primers were (*csf2*-fw:AAAATGCGGTCTCAACTTCCG; *csf2*-rev:TGACGAAGAGTTCCCCGAATG), (*dinG*-fw:GAGTCTGCCGGATTGTCGTTA; *dinG*-rev:GTACCAGATAGCCCAGCGTTT) and (*cas6*-fw:AATGCGTTTCGGTTGCGTATC; *cas6*-rev:GAGTACGGCAGCTTCTCTCC).

Bioinformatics analysis DNA sequences were analysed using CRISPRFinder, CRISPRTarget and Snapgene (GSL Biotech) [12, 15, 16]. Multi-Locus Sequence Typing, resistance genes and plasmids were identified using MLST, ResFinder and PlasmidFinder, respectively [17]. Spacer analysis was performed by BLAST and Geneious [18]. A phylogenetic UPGMA-based tree was constructed for CRISPR using MEGA7 [19, 20]. Direct repeats and PAM conservation were assessed using WebLogo, RNA secondary structure was predicted using RNAfold [21-23].

Funding

E. N. was supported by a grant from the Schlumberger Foundation's Faculty for the Future Programme (2012-2016).

Transparency declarations

None to declare.

References

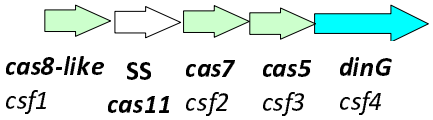
1. Marraffini, L.A. and E.J. Sontheimer, *CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea.* Nature Reviews Genetics, 2010. **11**(3): p. 181-90.
2. Barrangou, R., et al., *CRISPR provides acquired resistance against viruses in prokaryotes.* Science, 2007. **315**(5819): p. 1709-1712.
3. Rath, D., et al., *The CRISPR-Cas immune system: biology, mechanisms and applications.* Biochimie, 2015. **117**(2015): p. 119-28.
4. Koonin, E.V., K.S. Makarova, and F. Zhang, *Diversity, classification and evolution of CRISPR-Cas systems.* Current Opinion in Microbiology, 2017. **37**(1): p. 67-78.
5. Makarova, K.S., F. Zhang, and E.V. Koonin, *SnapShot: Class 1 CRISPR-Cas Systems.* Cell, 2017. **168**(5): p. 946-946.e1.
6. Makarova, K.S., et al., *An updated evolutionary classification of CRISPR-Cas systems.* Nature Reviews Microbiology, 2015. **13**(11): p. 722-36.
7. Zhang, Q. and Y. Ye, *Not all predicted CRISPR-Cas systems are equal: isolated cas genes and classes of CRISPR like elements.* BMC Bioinformatics, 2017. **18**(1): p. 92.
8. Ozcan, A., et al., *Type IV CRISPR RNA processing and effector complex formation in Aromatoleum aromaticum.* Nat Microbiol, 2019. **4**(1): p. 89-96.
9. Makarova, K.S., Y.I. Wolf, and E.V. Koonin, *The basic building blocks and evolution of CRISPR-CAS systems.* Biochem Soc Trans, 2013. **41**(6): p. 1392-400.
10. Touchon, M., et al., *CRISPR distribution within the Escherichia coli species is not suggestive of immunity-associated diversifying selection.* Journal of Bacteriology, 2011. **193**(10): p. 2460-7.
11. Koonin, E.V. and M. Krupovic, *Evolution of adaptive immunity from transposable elements combined with innate immune systems.* Nat Rev Genet, 2015. **16**(3): p. 184-92.
12. Lundgren, M., E. Charpentier, and P.C. Fineran, *CRISPR Methods and Protocols.* Methods in Molecular Biology. Vol. 1311 2015, New York Springer Science Business Media New York
13. Dwarakanath, S., et al., *Interference activity of a minimal Type I CRISPR–Cas system from Shewanella putrefaciens.* Nucleic Acids Res, 2015. **43**(18): p. 8913-23.
14. Koonin, E.V. and K.S. Makarova, *Mobile genetic elements and evolution of CRISPR-Cas systems: all the way there and back.* Genome Biology and Evolution, 2017. **9**(10): p. 2812-2825.
15. Grissa, I., G. Vergnaud, and C. Pourcel, *CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats.* Nucleic Acids Research, 2007b. **35**(2): p. W52-7.
16. GSL, "SnapGene software (from GSL Biotech; available at snapgene.com)", Biotech Chicago, IL.
17. Carattoli, A., et al., *In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing.* Antimicrob Agents Chemother, 2014. **58**(7): p. 3895-903.
18. Kearse, M., et al., *Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data.* Bioinformatics, 2012. **28**(12): p. 1647-9.
19. Kumar, S., G. Stecher, and K. Tamura, *MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets.* Molecular Biology and Evolution, 2016. **33**(7): p. 1870-4.
20. Shen, J., et al., *Comparative analysis of CRISPR-Cas systems in Klebsiella genomes.* Journal of Basic Microbiology, 2017. **57**(4): p. 325-336.
21. Crooks, G.E., et al., *WebLogo: a sequence logo generator.* Genome Research, 2004. **14**(6): p. 1188-90.
22. Mathews, D.H., *Predicting a set of minimal free energy RNA secondary structures common to two sequences.* Bioinformatics, 2005. **21**(10): p. 2246-53.

23. Hofacker, I.L., *Vienna RNA secondary structure server.* Nucleic Acids Research, 2003. **31**(13): p. 3429-31.

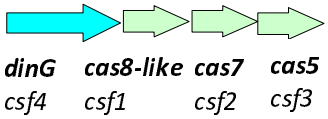
Figures

A

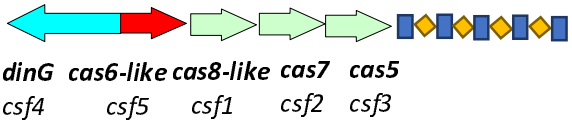
1 Type U



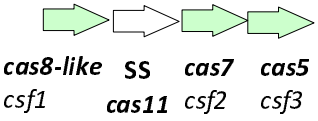
2 Type IV (putative)



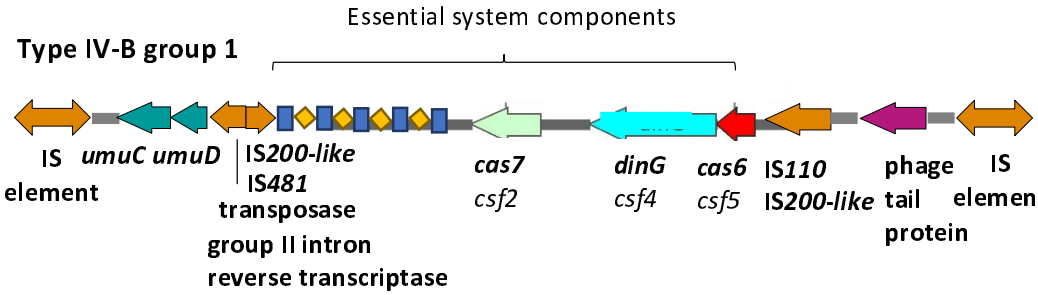
3 Type IV



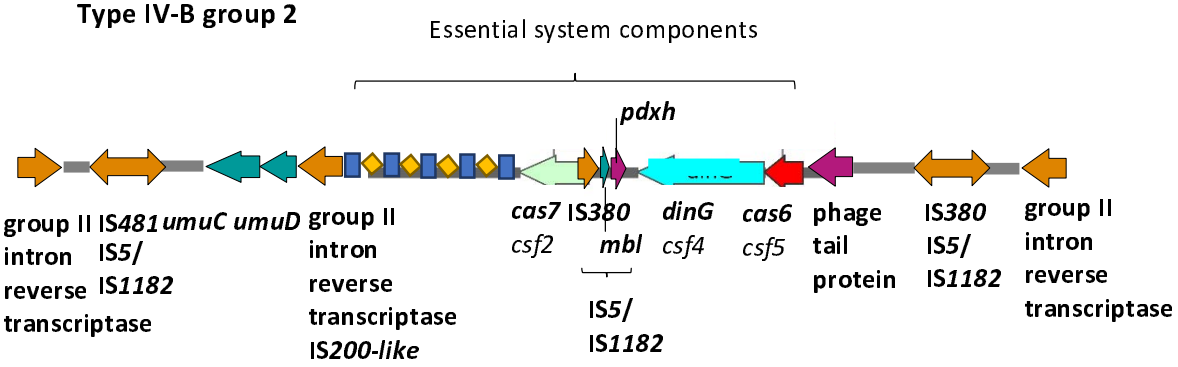
4 Type IV variant



5 Type IV-B group 1



6 Type IV-B group 2



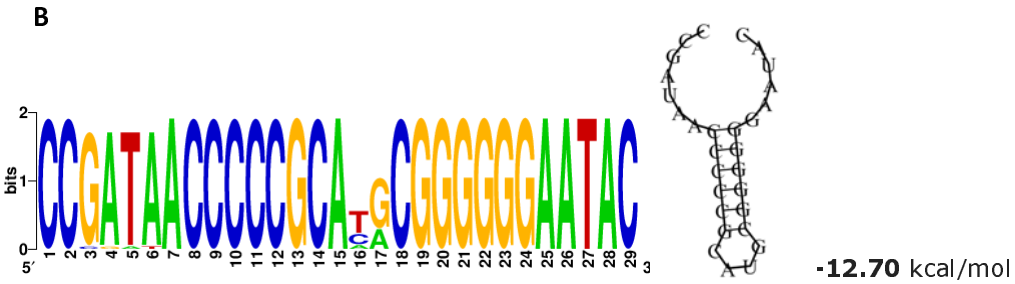
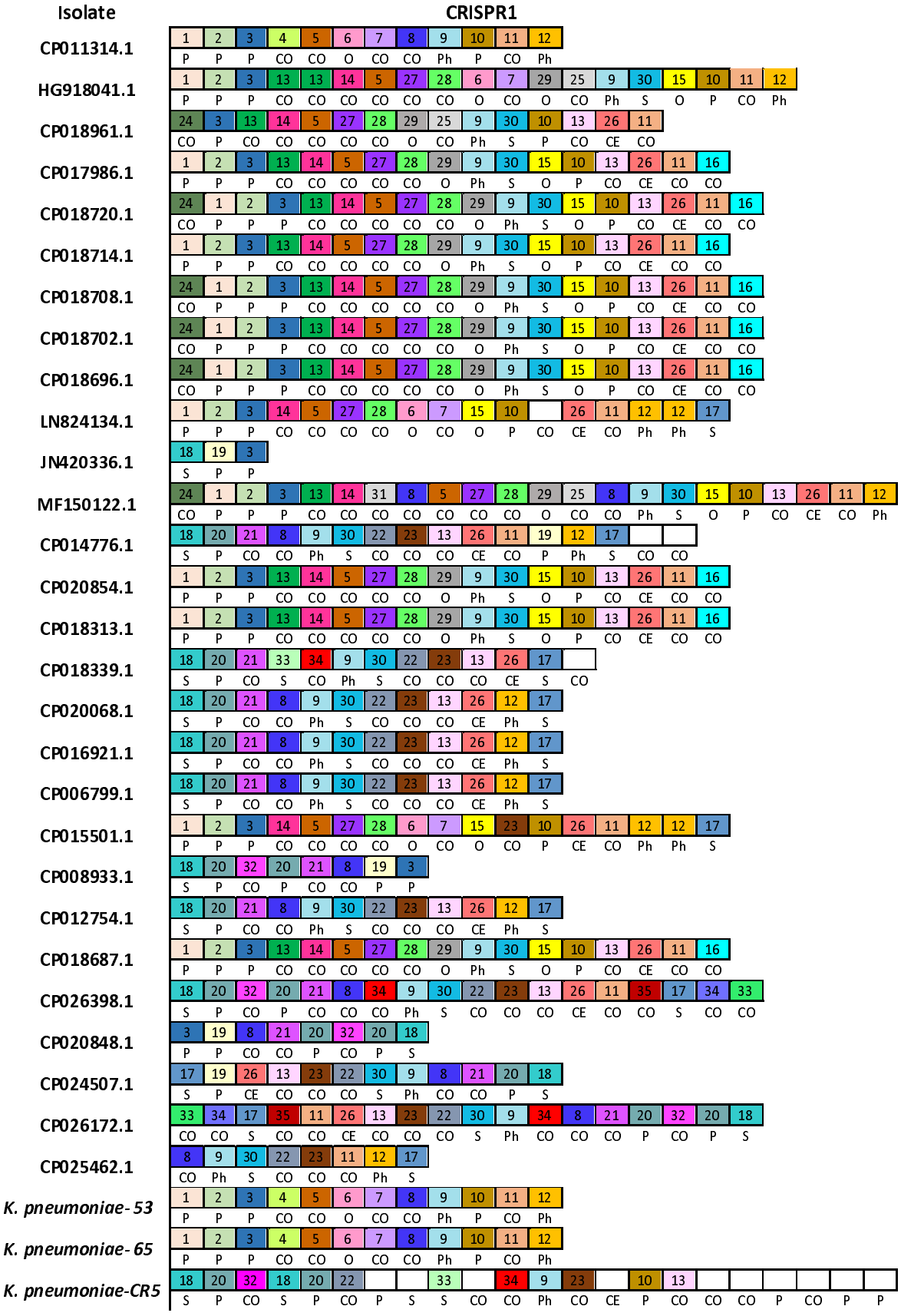
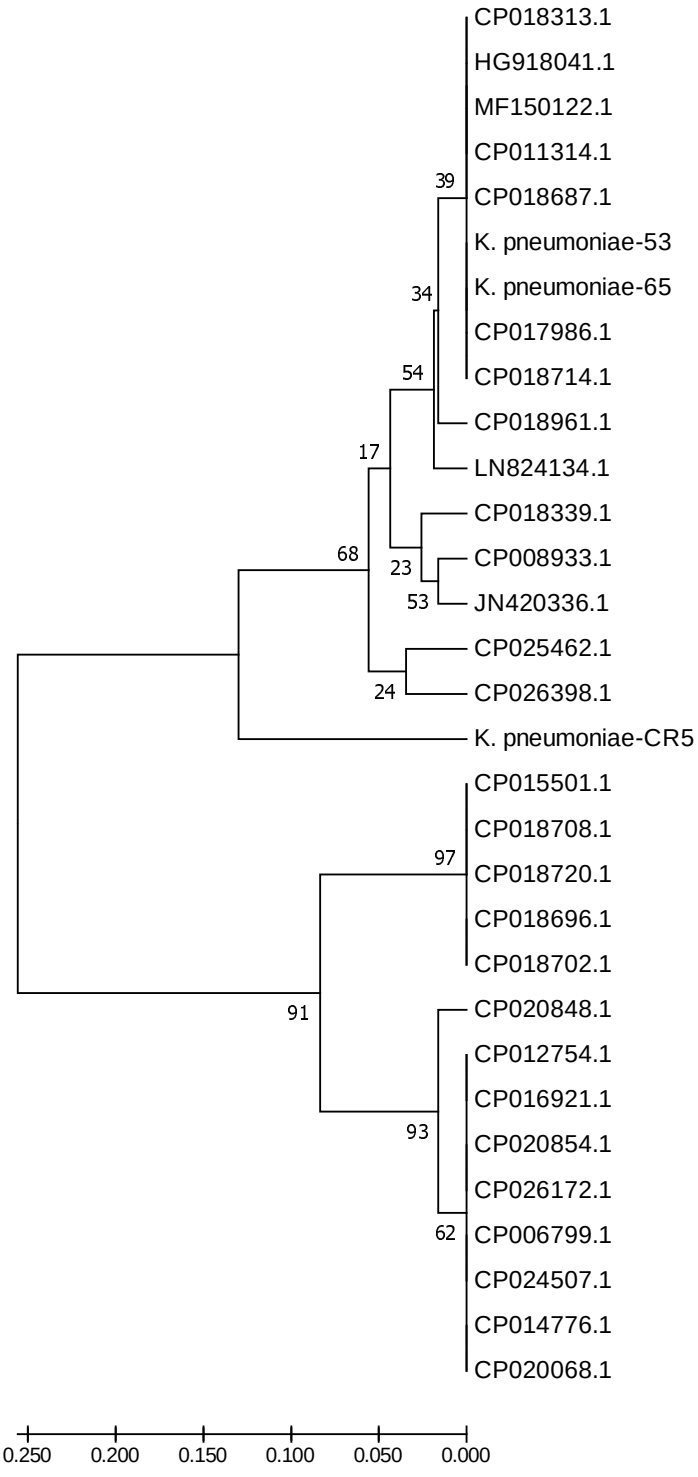


Figure 1: Type IV-B CRISPR-Cas system. (A) Schematic representation of Type IV CRISPR-Cas sytems and the two groups of Type IV-B described here. Panel 1 is Type U (unknown) as identified in 2013 [9]. Panel 2 is Type IV (putative) as identified in *Acidithiobacillus ferroxidans* in 2015 [6]. Panel 3 is Type IV and Panel 4 is Type IV variant identified in *Thioalkalivibrio sp.* K90mix (TK90_2699-TK90_2703) and *Rhodococcus jostii* RHA1 (RHA1_ro10069-RHA1_ro10072), respectively, in 2017 [14]. Panel 5 is Type IV-B group 1 and Panel 6 is Type IV-B group 2 as detected in *Enterobacteriaceae* isolates and genomes in this study. Arrows in different colours represent genes; red represents *cas6*; bright blue represents *dinG*; light green represents *cas7*; white represents *cas11;* blue-yellow pattern represents the direct repeat-spacer loci; orange represents the location of the associated MGEs occurring upstream (IS*5*/ IS*1182,* IS*630,* IS*6,* IS*1,* IS*481* or IS*110*) and downstream (IS*5*/ IS*1182,* IS*630,* IS*6,* IS*5-like,* IS*3000,* IS*Kra4,* IS*10L* or group II intron reverse transcriptase) of the system; green represents resistance genes; purple represents other genes associated with the system. (B) Conservation of the repeats and predicted stable stem-loop secondary structure predicted to be involved in the mechanism of pre-crRNA processing. The height of the letters in the sequence logo shows the relative frequency of their recurrence at that position. Wobbles at position 16 and 17 are within the loop of the predicted stem-loop

structure and are therefore tolerated in the structural prediction shown in C; (C) The predicted secondary structure of direct repeats and the associated Minimum Free Energy (MFE) estimated in (kcal/mol) shown underneath the structure.





279

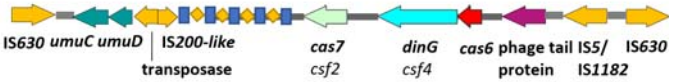
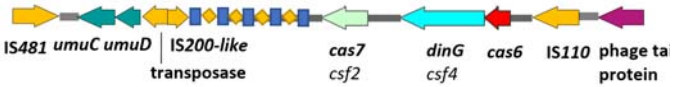
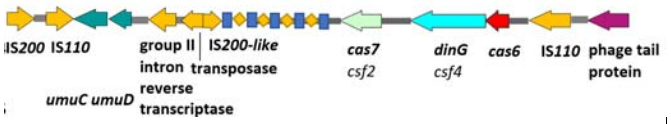
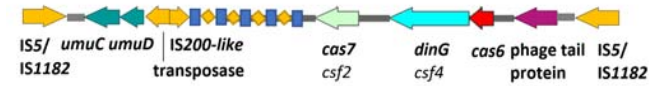
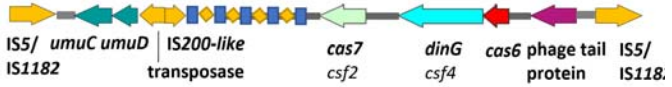
280 **Figure 2: Type IV-B CRISPR spacer polymorphism and evolutionary relationships.** (A) The

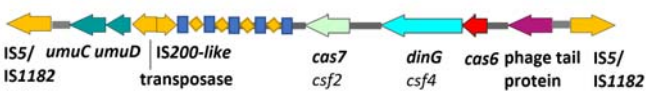
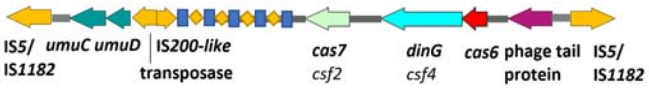
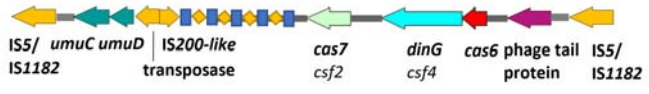
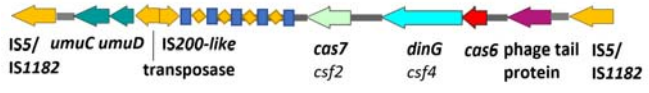
281 spacers map. Only spacers are represented by boxes, and no repeats are included. Identical

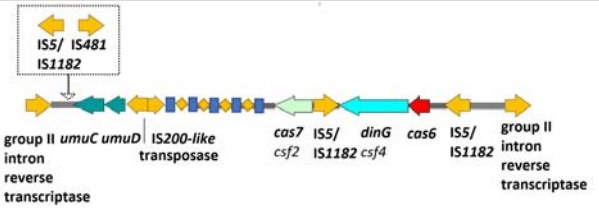
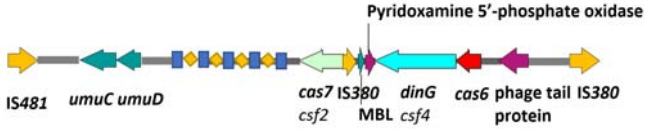
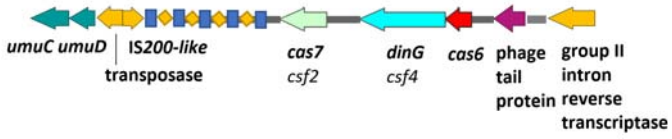
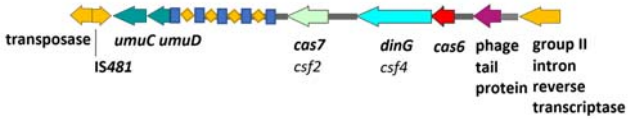
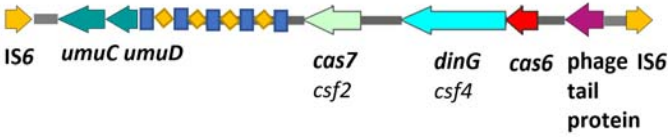
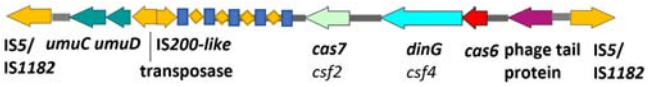
spacers are represented by the same number and colour, while unique spacers are represented by white colour and no number is associated with the box. Self-targeting spacers are indicated by letter (S) and show 100% identity to host DNA, plasmid-targeting spacers are indicated by letter (P), phage targeting spacers are indicated by letters (Ph), other *Enterobacteriaceae* targeting spacers (100% identity) are indicated by letter (O), cryptic spacers with similarity to other bacterial DNA are indicated by letters (CO), and those with similarity to Eukaryotic DNA are indicated by letters (CE) that are positioned underneath the relevant spacer. CE spacers showed at least 57% identity to eukaryotic DNA. CE spacers were confirmed by multiple sequences alignments. (B) The phylogenetic tree illustrating the evolutionary relationships of the Type IV-B repeat-spacer CRISPR loci. Phylogenetic UPGMA tree was constructed using the MUSCLE algorithm of MEGA7. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Maximum Likelihood method, [bootstrap test (1000 replicates), and the rate variation among sites was modelled with a gamma distribution (shape parameter = 1).

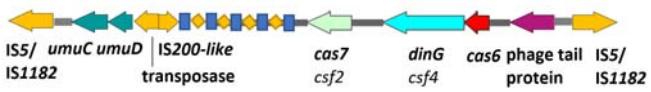
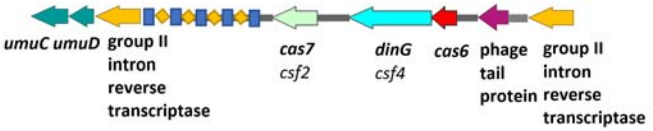
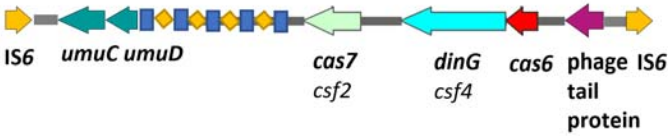
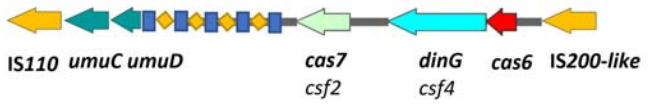
Supplementary Table

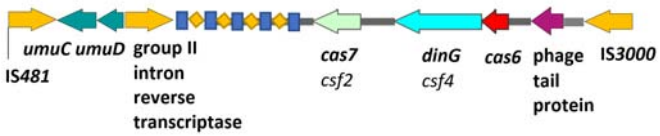
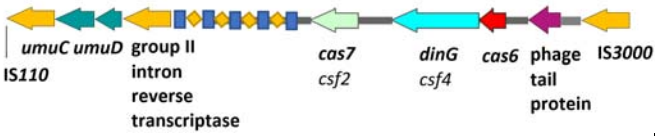
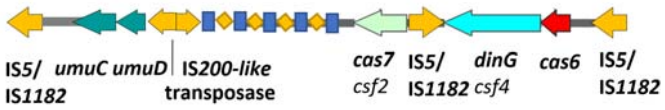
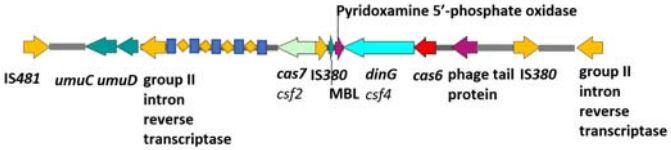
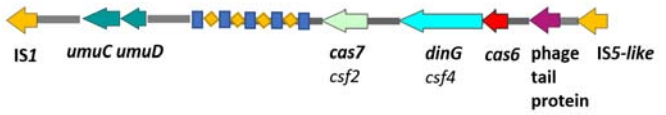
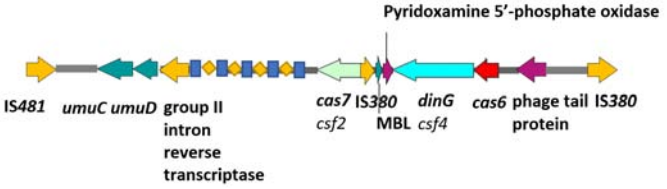
Table S1: Type IV-B *IncHI1B/ IncFIB (Mar)* plasmids information. †

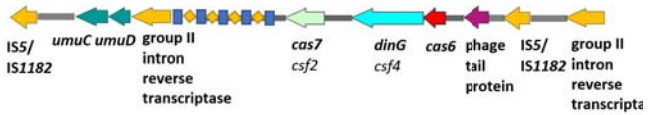
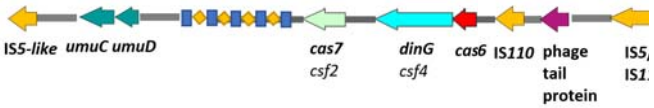
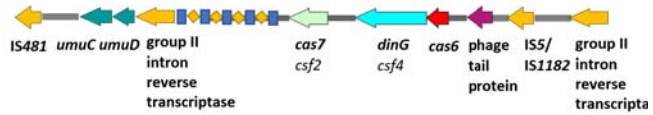
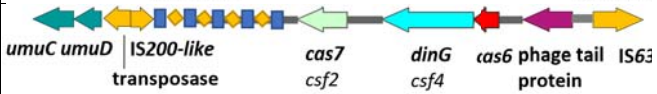
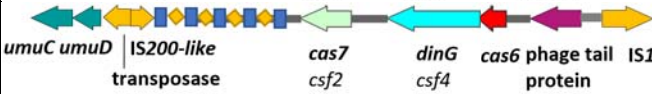
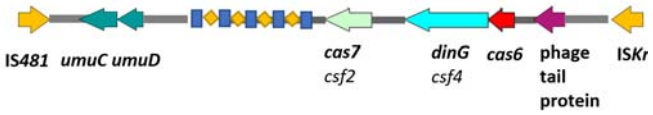
	Strain (Accession number)	ST	location (date)	plasmid	Resistance genes	Repeat-spacer CRISPR loci*					Type IV-B Structure
						Start	End	Spacer number	Length of DR	CRISPR length	
1	*K. pneumoniae* strain 234-12 plasmid pKpn23412-362 (CP011314.1)	ST-514	UK (2015)	*IncHI1B*	*merC, aph(6)-Id, tmrB, terY, terX, terW, terZ, terA, terB, terC,* bla_{TEM-1}, $bla_{CTX-M-15}$, bla_{OXA-1}, *aacA4, aph(3'')-, aac(3)-IIa, aac(6')Ib-cr, strA, strB, catB4, catA1, sul2, tet(A), dfrA1*	141865	142618	12	23	753	
2	*K. pneumoniae* Kp15 plasmid pENVA (HG918041.1)	unknown	Germany (2014)	*IncH*	*merD, merA, merP, qnrb4, aadA, terY, terX, terW, terZ, terA, terB, terC,* $bla_{CTX-M-15}$, bla_{DHA-1}, *qacE,* bla_{TEM-1}, *aadA1, aac(3)-II, sul1, tet(A), dfrA15*	99454	100629	19	23	1175	
3	*E. coli* strain Ecol_422 plasmid pEC422_1 (CP018961.1)	ST-2	UK (2016)	*IncHI1B*	*tetC, sul1, mer, qnrE1, terY, terX, terW, terZ, terA, terB, terC,* $bla_{CTX-M-2}$, *qacE,* bla_{OXA-1}, bla_{TEM-26}, *aac(6')Ib-cr, aac(3)-IIa, mph(A), catB3,arr-3, sul1*	276922	277868	15	29	946	
4	*K. pneumoniae* strain 825795-1 plasmid unnamed1 (CP017986.1)	ST-147	Germany (2016)	*IncHI1B*	*terA*	148524	149585	17	30	1061	
5	*K. pneumoniae* strain KP_Goe_828304 plasmid	ST-147	Germany (2016)	*IncHI1B*	*terY, terX, terW, terZ, terA, terB, terC*	24563	25679	18	24	1116	

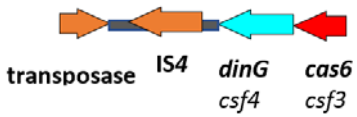
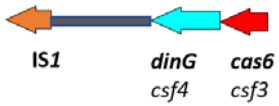
	Strain (Accession number)	ST	location (date)	plasmid	Resistance genes	Repeat-spacer CRISPR loci*					Type IV-B Structure
						Start	End	Spacer number	Length of DR	CRISPR length	
	pKp_Goe_304-1 (CP018720.1)										
6	*K. pneumoniae* strain Kp_Goe_152021 plasmid pKp_Goe_021-1 (CP018714.1)	ST-147	Germany (2016)	*IncHI1B*	*terY, terX, terW, terZ, terA, terB, terC*	7687	8748	17	30	1061	
7	*K. pneumoniae* strain Kp_Goe_827026 plasmid pKp_Goe_026-1 (CP018708.1)	ST-147	Germany (2016)	*IncHI1B*	none	54661	55777	18	24	1116	
8	*K. pneumoniae* strain Kp_Goe_827024 plasmid pKp_Goe_024-1 (CP018702.1)	ST-147	Germany (2016)	*IncHI1B*	*terY, terX, terW, terZ, terZ, terB, terC*	5006	6122	18	24	1116	
9	*K. pneumoniae* strain Kp_Goe_149832 plasmid pKp_Goe_832-1 (CP018696.1)	ST-147	Germany (2016)	*IncHI1B*	*terY, terX, terW, terZ, terA, terB, terC*	102143	103259	18	24	1116	

	Strain (Accession number)	ST	location (date)	plasmid	Resistance genes	Repeat-spacer CRISPR loci*					Type IV-B Structure
						Start	End	Spacer number	Length of DR	CRISPR length	
10	*K. pneumoniae* MS6671.v1, plasmid (LN824134.1)	ST-147	Australia (2015)	*IncHI1B*	none	16764	17823	17	29	1059	
11	*K. pneumoniae* plasmid pNDM-MAR (JN420336.1)	ST-15	Italy (2011)	IncH	bla_{NDM-1}, ble_{MBL}, *qnrB66, merR, terY3, terY1, terW, terZ, terA, terC, terD, terE, terF,* $bla_{CTX-M-15}$, bla_{OXA-1}, *aac(6')Ib-cr, aac(6')Ib-cr, catA1, catB4*	240016	240226	3	28	210	
12	*K. pneumoniae* A64477 plasmid pKP64477b (MF150122.1)		Brazil (2017)	*IncHI1B*	*terZ, terA, terB, terD*	84064	85428	22	29	1364	
13	*P. gergoviae* FB2 plasmid pFB2.1 (CP014776.1)	unknown	Malaysia (2016)	*IncFIB (Mar)*	*terF, terC, terB, terA, terZ, terW, terY, terF*	56849	57852	16	29	1003	
14	*K. pneumoniae* KPN528 plasmid pKPN528-1 (CP020854.1)	ST-14	USA (2013)	*IncHI1B*	bla_{NDM-1}, bla_{OXA-1}, *terF, terC, terB, terA, terZ, terW, terY, terX, sul1, dfrA12, dfrA14, qnrB1, aac(6')Ib-cr, aph(3')-Via, armA, aadA2, aac(6')Ib-cr, mph(E), msr(E), catB4*	251737	252495	12	29	758	
15	*K. pneumoniae* Kp_Goe_149473 plasmid	ST-147	Germany (2016)	*IncHI1B*	*terY, terX, terW, terZ, terA, terB, terC, terF*	116666	117727	17	30	1061	

	Strain (Accession number)	ST	location (date)	plasmid	Resistance genes	Repeat-spacer CRISPR loci*					Type IV-B Structure
						Start	End	Spacer number	Length of DR	CRISPR length	
	pKp_Goe_473-1 (CP018687.1)										
16	*K. pneumoniae* strain Kp_Goe_822579 plasmid pKp_Goe_579-1 (CP018313.1)	ST-147	Germany (2016)	*IncHI1B*	none	94773	95834	17	30	1061	
17	*K. pneumoniae* Kp_Goe_154414 plasmid pKp_Goe_414-1 (CP018339.1)	ST-23	Germany (2016)	*IncFIB (Mar)*	*terY, terX, terW, terZ, terA, terB, terC, terF*	75291	76111	13	29	820	
18	*K. pneumoniae* AR_0068 plasmid unitig_1 (CP020068.1)	ST-14	USA (2017)	*IncHI1B*	*terF, terC, terB, terA, terZ, terW, terY, terX, aac(3)-Iid, sul1, dfrA12, aph(3'')-Ib, aph(6)-Id, aph(3')-VI, aadA2, bla*$_{SHV-11}$*, bla*$_{NDM-1}$*, strA, strB, aac(3)-IId, armA, aadA2, mph(E), msr(E), sul2*	136244	137004	12	29	760	
19	*K. pneumoniae* 11 plasmid pIncHI1B_DHQP1300920 (CP016921.1)	ST-14	USA (2016)	IncHI1B	*terZ, terA, terB, terC, terF, dfrA1, terY, terX, terW, aph(3'), aacA4, ant(3'')-Ia, bla*$_{NDM-1}$*, bla*$_{OXA-1}$*, aac(6')Ib-cr, armA, aadA2, aac(6')Ib-cr, qnrB1, mph(E), msr(E), catB4, sul, dfrA14, dfrA12*	229597	230357	12	29	760	

	Strain (Accession number)	ST	location (date)	plasmid	Resistance genes	Repeat-spacer CRISPR loci*					Type IV-B Structure
						Start	End	Spacer number	Length of DR	CRISPR length	
20	*K. pneumoniae* KP617 plasmid KP-plasmid1 (CP012754.1)	ST-14	Korea (2015)	*IncHI1B*	*terY, terX, terW, terZ, terA, terB, terC, terF, qnrB1, qacEDelta1, ble*$_{MBL}$*, merE, aph(3')-VI, bla*$_{NDM-1}$ *aadA2, armA, msr(E), mph(E), sul1, dfrA12*	114404	115164	12	29	760	
21	*K. pneumoniae* PittNDM01 plasmid1 (CP006799.1)	ST-14	USA (2013)	*IncHI1B*	*terF, terE, terC, terA, terZ, terW, terY, terX, bla*$_{OXA-1}$*, bla*$_{NDM-1}$*, merE, qnrB1, aac(6')-Ib, aph(3')-VI, armA, aadA2, msr(E), mph(E), catB4, sul1, dfrA14, dfrA12*	114132	114892	12	29	760	
22	*K. pneumoniae* SKGH01 plasmid unnamed 1 (CP015501.1)	ST-147	UAE (2016)	*IncHI1B*	*terF, terC, terA, terZ, terW, terY, nreA, chrA, terX*	7545	8597	17	29	1052	
23	*K. pneumoniae* strain PMK1 plasmid pPMK1-NDM (CP008933.1)	ST-15	UK (2014)	*IncHI1B*	*bla*$_{CTX-M-15}$*, aadA2, armA, aac(6')Ib-cr, bla*$_{OXA-1}$*, bla*$_{NDM-1}$*, qnrB66, msr(E), mph(E), catA1, catB4, sul1, dfrA12*	10628	11144	8	30	516	
24	*K. pneumoniae* strain KPNIH48 plasmid pKPN-edaa (CP026398.1)	ST-252	USA (2018)	*IncHI1B / IncFIB (Mar)*	none	7401	8525	18	29	1124	
25	*K. pneumoniae* strain KPN1481 plasmid pKPN1481-1 (CP020848.1)	ST-906	USA (2017)	*IncHI1B / IncFIB (Mar)*	*aac(6')-Ib, aadA1, aac(6')Ib-cr, bla*$_{OXA-9}$*, bla*$_{TEM-1A}$*, bla*$_{NDM-1}$*, bla*$_{OXA-1}$*, bla*$_{CTX-M-15}$*, aac(6')-Ib-cr, qnrB1, aac(6')-Ib-cr, catB4*	325794	326308	8	29	514	

	Strain (Accession number)	ST	location (date)	plasmid	Resistance genes	Repeat-spacer CRISPR loci*					Type IV-B Structure
						Start	End	Spacer number	Length of DR	CRISPR length	
26	*K. pneumoniae* strain KSB2_1B plasmid unnamed1 (CP024507.1)	ST-323	Australia (2017)	*IncFIB (Mar)*	none	41987	42747	12	29	760	
27	*K. pneumoniae* strain KPNIH50 plasmid pKPN-bbef (CP026172.1)	ST-252	USA (2018)	*IncHI1B / IncFIB (Mar)*	none	227283	228407	18	29	1124	
28	*K. pneumoniae* strain F44 plasmid p44-1 (CP025462.1)	ST-11	USA (2017)	*IncHI1B / IncFIB (Mar)*	*aac(3)-IId, bla*$_{TEM-1B}$*, bla*$_{SHV-12}$*, mph(A)*	42604	43120	8	29	516	
29	*K. pneumoniae-53* plasmid1	ST-502	Egypt (2002)	*IncHI1B / IncFIB (Mar)*	*dfrA1*	19059	19812	12	23	753	
30	*K. pneumoniae-65* plasmid 1	ST-15	Egypt (2003)	*IncFIB (Mar)*	none	6012	6765	12	23	753	
31	*K. pneumoniae-CR5* plasmid 1	ST-392	UK (2017)	*IncFIB (Mar)*	*aac(6')Ib-cr, aac(3)-IIa, strA, aac(3)-IId, strB, armA, bla*$_{TEM-1B}$*, bla*$_{CTX-M-15}$*, bla*$_{DHA-1}$*, bla*$_{SHV-11}$*, bla*$_{NDM-1}$*, bla*$_{OXA-1}$*, aac(6')Ib-cr, oqxB, oqxA, qnrB66, fosA, msr(E), mph(E), catB4, sul2, sul1, dfrA14*	5467	6833	22	29	1366	
A	*K. pneumoniae* strain K66-45 plasmid pK66-45-1	ST-11	Norway (2017)	*IncFIB(Mar)/ IncHI1B*	*aph(3')-VI, armA, aadA2, bla*$_{NDM-1}$*, bla*$_{CTX-M-15}$*, qnrS1, mph(E), msr(E), sul1, dfrA12*						

	Strain (Accession number)	ST	location (date)	plasmid	Resistance genes	Repeat-spacer CRISPR loci*					Type IV-B Structure
						Start	End	Spacer number	Length of DR	CRISPR length	
B	(CP020902.1) *K. pneumoniae* strain AR_0158 plasmid tig00000727 (CP021699.1)	ST-163	USA (2017)	*IncFIB(Mar)/ IncHI1B*	*aac(6')Ib-cr, aac(3)-IId,* bla_{OXA-1}, bla_{SHV-2}, bla_{NDM-1}, *aac(6')-Ib-cr, qnrB1, catB4, tet(B), dfrA30*						
C	*K. pneumoniae* strain LS356 plasmid pKP8-2 (CP025638.1)	ST-485	China (2018)	*IncHI1B*	none						

† Items 1-31 are the plasmids that carried complete Type IV-B CRISPR-Cas system. Items A-C represent examples of partial Type IV-B components that were found carried on the same plasmids.

* All detected repeat-spacer CRISPR were CRISPR1.