

# Comparative genomics analysis of pKF3-94 in *Klebsiella pneumoniae* reveals plasmid compatibility and horizontal gene transfer

## OPEN ACCESS

### Edited by:

John R. Battista,  
Louisiana State University and A & M  
College, USA

### Reviewed by:

David L. Bernick,  
University of California, Santa Cruz,  
USA  
Zhang Wang,  
GlaxoSmithKline, USA

### \*Correspondence:

Qiyu Bao,  
Institute of Biomedical  
Informatics/Zhejiang Provincial Key  
Laboratory of Medical Genetics,  
Wenzhou Medical University, Chashan  
University Town, Wenzhou 325035,  
China  
baoqy@genomics.cn;  
Junwan Lu,  
School of Medicine, Lishui College,  
No. 1, Xueyuan Road, Lishui 323000,  
China  
304041474@qq.com

†These authors have contributed  
equally to this work.

### Specialty section:

This article was submitted to  
Evolutionary and Genomic  
Microbiology,  
a section of the journal  
Frontiers in Microbiology

Received: 21 December 2014

Accepted: 28 July 2015

Published: 18 August 2015

### Citation:

Ying J, Wu S, Zhang K, Wang Z, Zhu  
W, Zhu M, Zhang Y, Cheng C, Wang  
H, Tou H, Zhu C, Li P, Ying J, Xu T, Yi  
H, Li J, Ni L, Xu Z, Bao Q and Lu J  
(2015) Comparative genomics  
analysis of pKF3-94 in *Klebsiella  
pneumoniae* reveals plasmid  
compatibility and horizontal gene  
transfer. *Front. Microbiol.* 6:831.  
doi: 10.3389/fmicb.2015.00831

Jianchao Ying<sup>1†</sup>, Songquan Wu<sup>2†</sup>, Kaibo Zhang<sup>2</sup>, Ziqiang Wang<sup>1</sup>, Wen Zhu<sup>1</sup>, Mei Zhu<sup>1</sup>, Ying Zhang<sup>1</sup>, Cong Cheng<sup>2</sup>, Huifeng Wang<sup>3</sup>, Huifen Tou<sup>4</sup>, Chuanxin Zhu<sup>4</sup>, Peizhen Li<sup>1</sup>, Jun Ying<sup>1</sup>, Teng Xu<sup>1</sup>, Huiguang Yi<sup>1</sup>, Jinsong Li<sup>1</sup>, Liyan Ni<sup>5</sup>, Zuyuan Xu<sup>1</sup>, Qiyu Bao<sup>1\*</sup> and Junwan Lu<sup>2\*</sup>

<sup>1</sup> Institute of Biomedical Informatics/Zhejiang Provincial Key Laboratory of Medical Genetics, Wenzhou Medical University, Wenzhou, China, <sup>2</sup> School of Medicine, Lishui College, Lishui, China, <sup>3</sup> National Institute of Biological Sciences, Beijing, China, <sup>4</sup> Wenzhou Center for Disease Control and Prevention, Wenzhou, China, <sup>5</sup> The Second Affiliated Hospital, Wenzhou Medical University, Wenzhou, China

In order to get insights into plasmid evolution and the dissemination of multidrug resistance, we performed extensive comparative genomics analyses of the *Klebsiella pneumoniae* plasmid pKF3-94 and some of its related plasmids. pKF3-94 is one of three plasmids isolated from the *K. pneumoniae* strain KF3. Of the 144 putative genes it harbors, 69 can be functionally assigned to be involved in transfer conjugation, transfer leading, antimicrobial resistance, transposon function, and plasmid replication. Comparison of plasmid replicon sequence types revealed that pKF3-94 carries two replicons that are distinct from those carried on the two sibling *K. pneumoniae* plasmids pKF3-70 and pKF3-140, thereby allowing pKF3-94 to coexist with these latter plasmids in the same host cell. Comparative genomics analyses further showed that pKF3-94 is more similar to plasmids pK1HV and pC15-k, which were isolated from different *K. pneumoniae* strains, than to pKF3-70 and pKF3-140. Interestingly, pK1HV contains a unique 49 kb region rich in mobile genetic elements and drug resistance genes, while pKF3-94 and pC15-k share a 15 kb homology region partitioned into a region rich in drug resistance genes and one containing a replicon. It is conceivable, therefore, that pK1HV and pC15-k have both arisen from a common pKF3-94-like plasmid. The comparisons lend further support for the role horizontal gene transfer plays in genome evolution and in the dissemination of genetic elements including drug resistance genes.

**Keywords:** plasmid compatibility, comparative genomics, horizontal gene transfer, *Klebsiella pneumoniae*, plasmid

## Introduction

Plasmids are versatile genetic elements that participate in various processes including the shuttling of genetic information between different hosts. This process, called horizontal gene transfer (HGT), can rapidly introduce newly evolved donor genes into a variety of host genomes (Jain et al., 2003). HGT mechanisms include transformation, conjugation, transduction, and variations of these as

well as the role of mobile genetic elements (MGE) (Brigulla and Wackernagel, 2010). Conjugation is driven by MGE which can be conjugative plasmids or conjugative transposons (Brigulla and Wackernagel, 2010). MGEs generally carry determinants for their translocation and maintenance (Bahl et al., 2009) and are often beneficial to the host cell by protecting, for instance, against antibiotics, heavy metals, or extreme environmental conditions (Sobecky and Coombs, 2009). In fact, by their capacity to transfer diverse genetic information between bacteria, conjugative plasmids critically contribute to bacterial genome evolution (Périchon et al., 2008).

A limitation to HGT, however, is the phenomenon of plasmid incompatibility. Plasmid incompatibility is generally defined as the failure of two co-resident plasmids to be stably inherited in a single host cell in the absence of external selection (Novick, 1987). It is due to sharing of one or more elements of the plasmid replication or partitioning systems rather than to any specific incompatibility (inc) gene (Novick, 1987). By definition, plasmids incompatible with each other belong to the same incompatibility group (Couturier et al., 1988). In multi-replicon plasmids, potential incompatibility determinants depend on combinations of distinct replicons. Multi-replicon plasmids can allow a host cell to acquire plasmids carrying incompatible replicons when replication is driven by compatible replicons (Villa et al., 2010). The classic multi-replicon IncF plasmid contains the FII replicon regulated by CopA, a constitutively synthesized 90 nt antisense-RNA, which is normally silent, and FIA and FIB replicons, which function only in enteric bacteria and are regulated by iterons, in cis-negative binding sites of the replication protein RepA. In plasmids F and p307, the FII replicon is substituted by a non-functional FIC replicon, while plasmids R1 and R100 of *E. coli* contain only one functional FII replicon (Villa et al., 2010). FII replicons are more divergent than FIA, FIB, and FIC replicons, and consequently can be divided into FII, FII<sub>K</sub> (identified in *K. pneumoniae*), FII<sub>S</sub> (identified in *Salmonella spp.*), FII<sub>Y</sub> (identified in *Y. pestis*, *Y. pseudotuberculosis*, and *Yersinia enterocolitica*), FII<sub>SH</sub> (identified in *Shigella flexneri* and *Shigella sonnei*). It has been demonstrated that, for instance, FII replicons do not participate in the initiation of replication of a plasmid when they are associated with FIA and/or FIB replicons, and hence, the presence of these latter replicons can overcome the incompatibility barrier between IncF plasmids (Villa et al., 2010). Plasmids of the IncF group represent one of the most frequent plasmid types and they often carry more than one replicon promoting the initiation of replication (Villa et al., 2010; Dolejska et al., 2013).

Plasmid-mediated transfer of drug-resistance genes is considered one of the most important mechanisms for the spread of multidrug resistance of pathogens (Zhao et al., 2010; Bai et al., 2013). For instance, from the nosocomial pathogen *Klebsiella pneumoniae*, which causes pneumonia and urinary tract infections particularly in patients in intensive care units (Landman et al., 2007), more than 100 plasmids ranging from 1.3 to 317 kb have been isolated and sequenced. Some of them were isolated from the same strain, such as pKDO1 (127 kb) and pKPN-CZ (207 kb) from *K. pneumoniae* ST416 (Dolejska et al., 2013); or pPKPN1 (283 kb), pPKPN2 (103 kb), pPKPN3

(70 kb), and pPKPN4 (6 kb) from *K. pneumoniae* PittNDM01 (Doi et al., 2014). The conjugative plasmid pKF3-94 of the IncF group was isolated from *K. pneumoniae* strain KF3, which contained in addition pKF3-70 and pKF3-140 (Zhao et al., 2010). Plasmid pKF3-94 exhibits a low sequence identity with either pKF3-70 or pKF3-140 but is highly similar to pK1HV from *K. pneumoniae* strain K1HV and to pC15-k from *K. pneumoniae* strain 997. pKF3-94 not only encodes the antimicrobial resistance genes *bla*<sub>CTX-M-15</sub> and *bla*<sub>TEM</sub>, but also carries two replicons. Here, we analyzed the replicon sequence types of these plasmids (Plasmid RST) to explain the mechanism of plasmid compatibility by using the Plasmid MLST Database (<http://pubmlst.org/plasmid/>). Identifying Plasmid RST contributes to recognizing and sub-categorizing plasmids, and further to analyzing the distribution of plasmid replicon and discovering their evolutionary origin (Villa et al., 2010). Moreover, we determined the genomic structure of pKF3-94 and performed evolutionary and extensively comparative analysis among the plasmid pKF3-94 and the related plasmids in order to evaluate the origin and evolution of these plasmids from *Klebsiella pneumoniae* and the distribution of multidrug resistance genes.

## Results and Discussion

### General Features of pKF3-94

Plasmid pKF3-94 is a circular 94,219 bp plasmid with an average GC content of 51.6% (Figure 1). We identified 144 putative open reading frames (ORFs), with the majority of them (119/144) transcribed from the leading strand. Sixty nine ORFs were predicted to encode proteins with known functions, including transfer conjugation, transfer leading, antimicrobial resistance, transposase function, replicon functions, and other functions (Figure 1).

### Replicon Sequence Types of pKF3-94, pKF3-70, and pKF3-140

pKF3-94 and pKF3-140 are both multi-replicon plasmids, while pKF3-70 only carries one replicon. As shown in Table 1, pKF3-70 belongs to the incompatibility group FII, as it carries the FII<sub>2</sub> replicon allele (F2:A:-B-). pKF3-140 carries three different replicons, FII<sub>1</sub>, FIA<sub>2</sub>, and FIB<sub>2</sub> (F1:A2:B2); pKF3-94 carries two replicons; although the Plasmid RST program identified only one FII<sub>K2</sub> replicon allele, (K2:A:-B-). The replication initiation protein (RepA) of FII<sub>K2</sub> shares 83% sequence similarity with that of FII<sub>2</sub>/FII<sub>1</sub>. The amino acid sequence of the RepA of FII<sub>2</sub>, however, shows 98% similarity with that of FII<sub>1</sub>. The unassigned replicon (unknown replicon type and not belonging to IncF family) of pKF3-94 is only found in 11 other plasmids, including pK1HV, pNJST258C3, pNJST258N3, and CCBH10892 (an unnamed plasmid from *E. hormaechei*, Table 1). The corresponding *repA* genes in these plasmids are >99% identical with each other. It is noteworthy that pNJST258C3 and pNJST258N3 are single-replicon plasmids, and so it is likely that this unassigned replicon has the function of replication initiation.

Although pKF3-94, pKF3-70, and pKF3-140 each carry one replicon of the incompatibility group FII, they belong to different

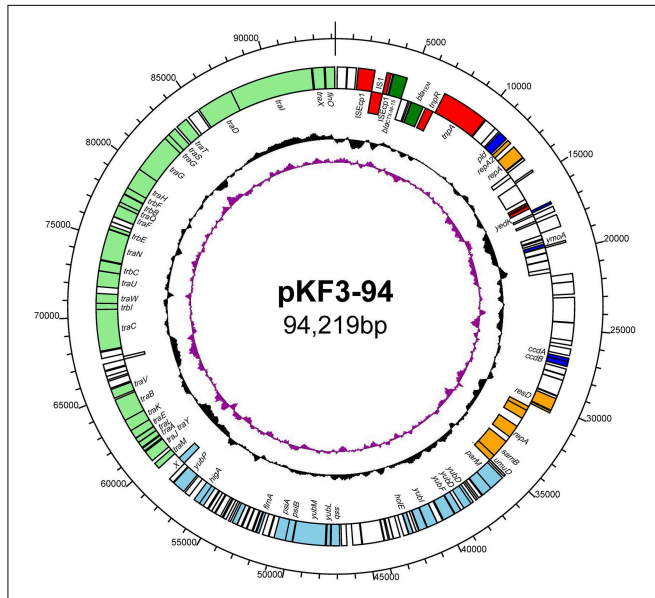
FII replicon alleles and possess different origins. In particular, FII<sub>K2</sub> is different from the other two FII replicons (Villa et al., 2010). FII<sub>K2</sub> is widely distributed in the plasmids from *K. pneumoniae*, while FII<sub>1</sub> and FII<sub>2</sub> are derived from *E. coli*. In addition, because FIA<sub>2</sub> and FIB<sub>2</sub> are also present in pKF3-140, FII<sub>1</sub> does not participate in the initiation of replication in this plasmid (Villa et al., 2010). Therefore, these three

plasmids could coexist in the same bacterial cell, avoiding the potential incompatibility resulting from carrying similar replicons.

### Comparative Genomics Analysis of pKF3-94 and Other Plasmids

To explore the phylogenetic relationship of pKF3-94 to plasmids from different bacteria, we performed a comparative genomics analysis of the corresponding plasmid sequences available in the NCBI database. A phylogenetic tree was constructed based on the six homologous genes concatenated sequences of 56 plasmids including pKF3-94, pKF3-70, and pKF3-140 using the six homologous genes concatenated sequences (Gadagkar et al., 2005, **Figure 2**). The phylogenetic relationship shows that plasmids from the hosts of the same genus cluster together. For instance, the plasmids pKF3-94 clusters together with other *Klebsiella* plasmids (**Figure 2**, C4), but pKF3-70 (Yi et al., 2010) and pKF3-140 (Bai et al., 2013) cluster together with plasmids from *Escherichia* (**Figure 2**, C3), but not with other plasmids from *Klebsiella* such as pKF3-94. This suggests that the origin of pKF3-94 is different from that of either pKF3-70 or pKF3-140. In addition, we found that pKF3-94 clusters closer together with pC15-k and is located on the same branch with pK1HV (**Figure 2**, C4). A collinearity analysis of the plasmid genomes including pKF3-94, pK1HV, pC15-k, pKF3-70, and pKF3-140 was also performed (**Supplementary Figure S1**). This suggests that pKF3-94 might have a closer phylogenetic relationship with pK1HV and pC15-k than with pKF3-70 and pKF3-140.

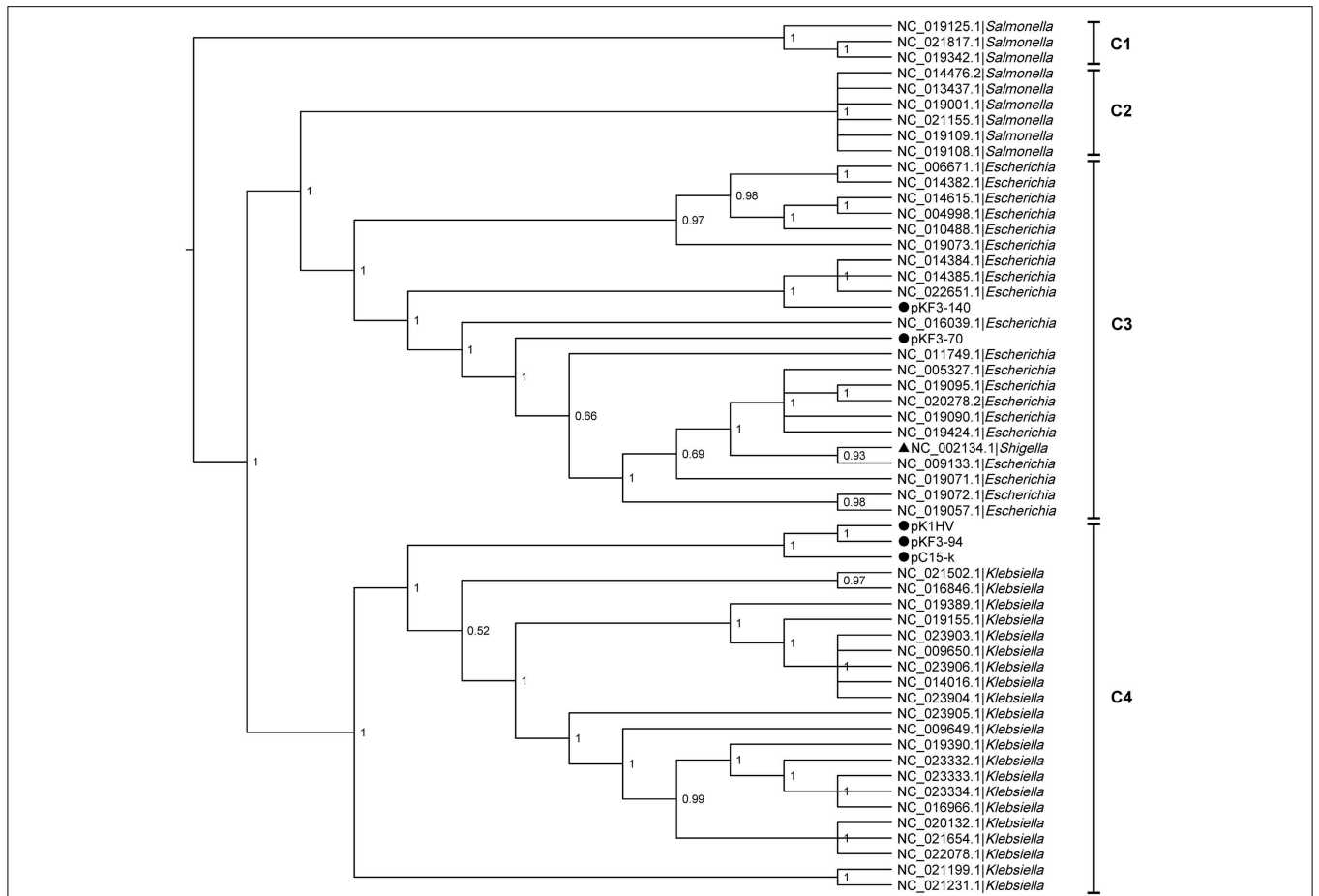
Further comparative genomics analysis showed that pKF3-94 is about 39 kb smaller than pK1HV (133 kb), but of similar size compared to pC15-k (95 kb). Nevertheless, 83.3% (120/144) and 68.8% (99/144) of the genes of pKF3-94 have a high degree of similarity (>90%) with those of pK1HV and pC15-k, respectively. Furthermore, the three plasmids share a conserved backbone sequence, which includes transfer conjugation and transfer leading regions, and they possess their own variable regions, which mainly include antimicrobial resistance genes and insertion sequences (**Figure 3**).



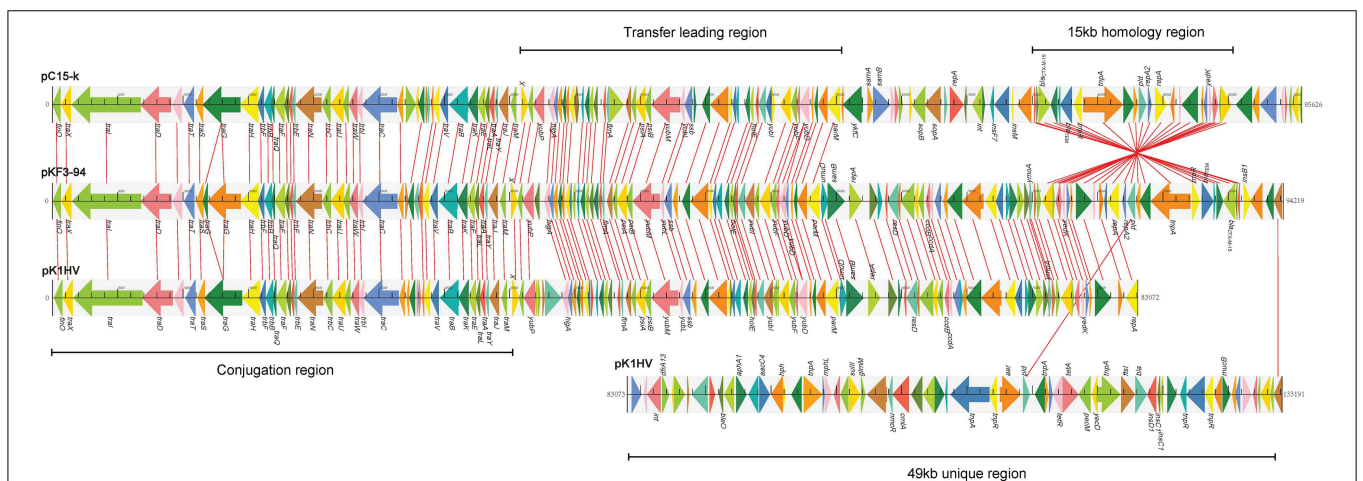
**FIGURE 1 | The circular map of pKF3-94 genome.** Counting from outside toward the center, the first circle refers to the position in bp. The second circle marks genes encoded on the leading strands (outwards) or lagging strands (inwards). The different functional regions are shown in different colors; green, drug-resistance related genes; red, transposase/insertion sequences; orange, replicon related genes; light blue, transfer leading regions; light green, transfer conjugation regions; blue, genes with other functions; blank, genes with unknown functions. The third circle shows GC content with an average of 50%, whereby a G+C content of more than 50% is shown toward the outside and a G+C content of less than 50% toward the inside. The fourth circle shows GC skew (G–C/G+C) with a positive GC skew toward the outside and a negative GC skew toward the inside.

**TABLE 1 | Plasmids/strains and their replicons identified in this study.**

| Plasmid/strain     | Bacterium            | Allele numbers of the replicons          |     |     | FAB formula |
|--------------------|----------------------|--|-----|-----|-------------|
|                    |                      | FII, FII <sub>K</sub> , FII <sub>S</sub> | FIA | FIB |             |
| pKF3-70            | <i>K. pneumoniae</i> | F2                                       | –   | –   | F2:A-:B-    |
| pKF3-94            | <i>K. pneumoniae</i> | K2                                       | –   | –   | K2:A-:B-    |
| pKF3-140           | <i>K. pneumoniae</i> | F1                                       | A2  | B2  | F1:A2.B2    |
| pK1HV              | <i>K. pneumoniae</i> | K2                                       | –   | –   | K2:A-:B-    |
| pC15-k             | <i>K. pneumoniae</i> | K2                                       | –   | –   | K2:A-:B-    |
| pNJST258C3         | <i>K. pneumoniae</i> | –  | –   | –   | –           |
| pNJST258N3         | <i>K. pneumoniae</i> | –  | –   | –   | –           |
| CCBH 10892 plasmid | <i>E. hormaechei</i> | –  | –   | –   | –           |
| pKP1780-kpc        | <i>K. pneumoniae</i> | K2                                       | –   | –   | K2:A-:B-    |
| SA20094177         | <i>S. enterica</i>   | S1                                       | –   | B22 | S1:A-:B22   |



**FIGURE 2 | The phylogenetic tree of 56 plasmids including pKF3-94, pKF3-70 and pKF3-140.** The tree can be divided into four clades: C1, C2, C3, and C4. Every clade was supported by Bayesian posterior probabilities (BPP).



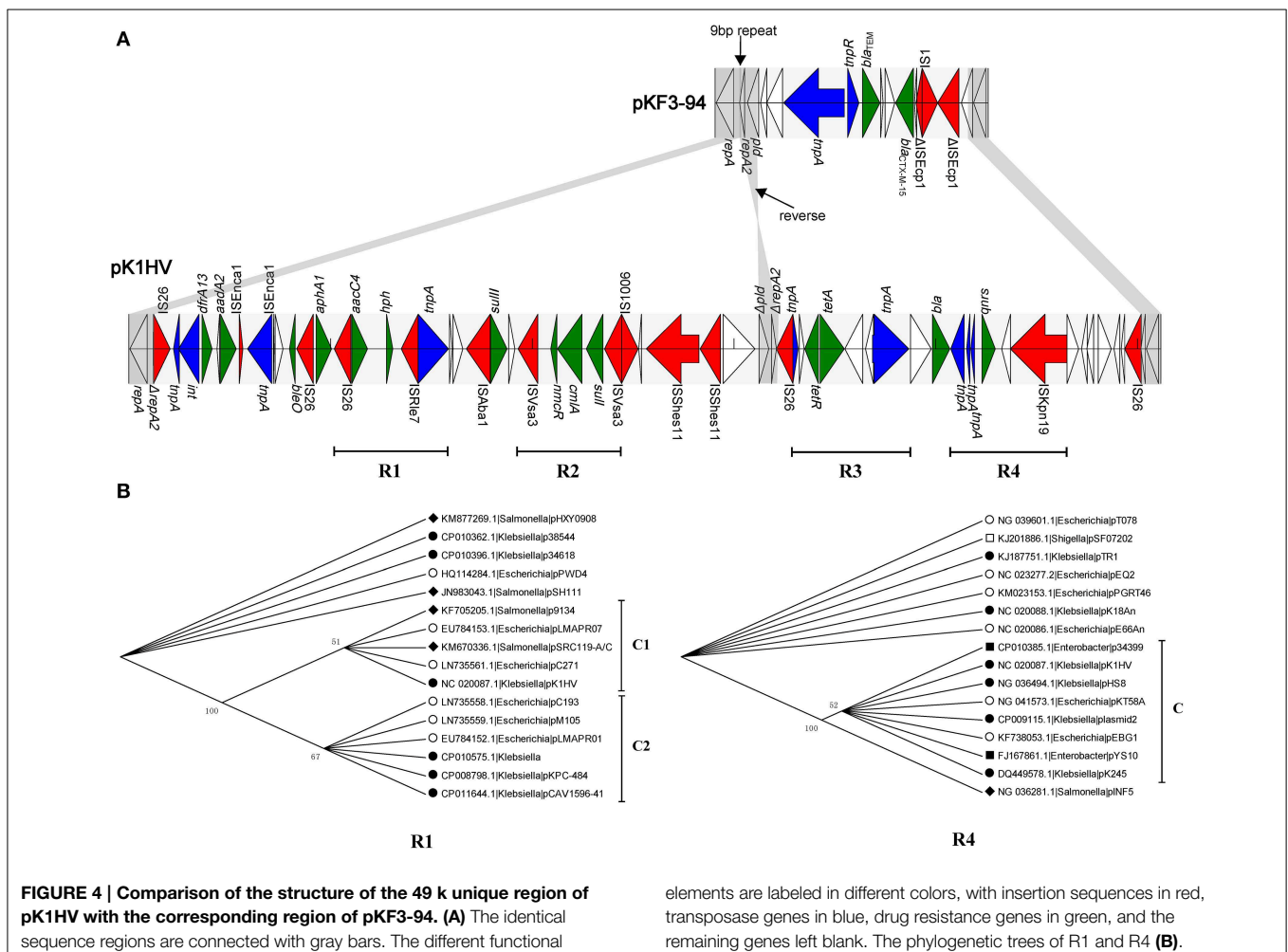
**FIGURE 3 | Comparison of genome structure of the plasmids pKF3-94, pK1HV, and pC15-k.** Homologous genes are marked with the same color and connected by lines, while non-homologous genes are left unconnected.

### Replicon Regions of pKF3-94 and pK1HV and the 49 kb Unique Region of pK1HV

The three plasmids pKF3-94, pK1HV, and pC15-k each carry two replicons, with FII<sub>K2</sub> co-existing in all of them. The unassigned replicon of pKF3-94 is also present in pK1HV, but not in pC15-k. The FII<sub>K2</sub> replicon consists of *repA* (858 bp, encoding RepA) and *repA2* (252 bp, encoding the replication regulatory protein RepA2) in pKF3-94, while the FII<sub>K2</sub> replicon in pK1HV is incomplete and includes *repA* and only 17 bp of a truncated *repA2* gene. Moreover, the *pld* gene (561 bp, encoding phospholipase D) downstream of *repA2* of pKF3-94 also exists in pK1HV, but is incomplete (480 bp, designated  $\Delta pld$ ) and in a reverse orientation (Figure 4A). Comparative analysis of the flanking sequences of  $\Delta pld$  indicates that the sequence downstream of  $\Delta pld$  of pK1HV includes a truncated *repA2* gene (244 bp), with an intervening sequence identical to that of pKF3-94 (a total of 379 bp in length). In addition, there are 9 bp perfect direct repeats (DRs), each lying downstream of the two parts of  $\Delta repA2$  in pK1HV (Figure 4A). This arrangement suggests that recombinations have occurred between the inverted  $\Delta pld$  and  $\Delta repA2$  sequence in pK1HV. The region downstream of

*pld* of pKF3-94 seems to be a hot spot of recombination. The resistance genes (*bla*<sub>CTX-M-15</sub> and *bla*<sub>TEM</sub>) and two transposase genes (*tnpA* and *tnpR*) are clustered in this region. The sequences flanking  $\Delta pld$  in pK1HV are about 49 kb long and are also littered with MGEs and antimicrobial resistance genes.

Further, we identified 15 complete or incomplete insertion sequences, 9 transposase/integrase genes and 14 antimicrobial resistance genes in this region. What's more, we found four relatively complete underlying structure of transposons (Figure 4A, R1, R2, R3, and R4). They have similar characteristics that the resistant genes are flanked by several insertion sequences/transposases. In addition, a phylogenetic analysis was also performed for these regions. For example, *aacC4* (coding aminoglycoside N(3')-acetyltransferase IV) and *hph* (coding hygromycin-B phosphotransferase) are flanked by insertion sequences IS26 and ISRle7 in R1. This region is only present in 16 plasmids covering 3 genera (*Escherichia*, *Klebsiella* and *Salmonella*). The phylogenetic analysis showed that the branch of pK1HV and the other 4 plasmids (Figure 4B, C1) cluster together with that of other 6 plasmids (Figure 4B, C2). It suggests that they probably have the same origin. While in R4, *qnrS*

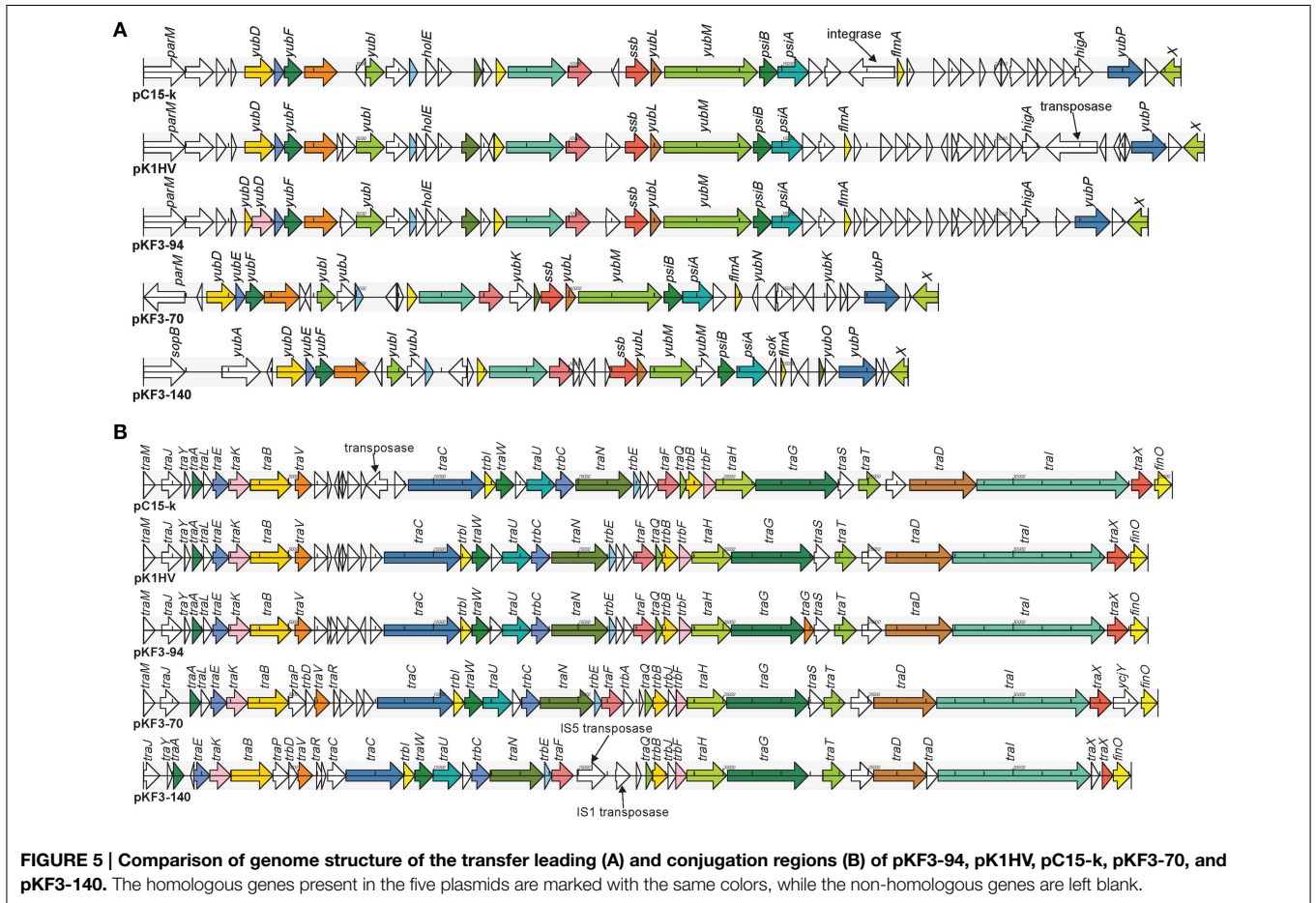


(coding fluoroquinolone resistance protein) is surrounded by ISKpn19 and *tnpA* (IS2 transposase). *qnrS* along with its adjacent insertion sequences are present in 16 plasmids covering 5 genera (*Escherichia*, *Klebsiella*, *Enterobacter*, *Salmonella*, and *Shigella*). R4 of pK1HV clusters together with other 7 plasmids (Figure 4B, C), and they also have a closer phylogenetic relationship with the plasmid pINF5 from *Salmonella* than with other 7 plasmids. It is not difficult to see that these transposons (R1, R4 and including R2 and R3 in Supplementary Figure S2) are uniformly distributed among these genera (*Escherichia*, *Klebsiella*, and so on). Furthermore, it suggests that horizontal gene transfer of these transposons probably happens widely among strains, species or genera, even some phylogenetically remote bacteria. This finding supports the notion that this region corresponds to a hot spot of recombination, whereby the antimicrobial resistance genes have likely been inserted with the help of MGEs.

### Comparative Genomics Analysis of Transfer Leading and Transfer Conjugation Regions between pKF3-94 and the Related Plasmids

The 23.6 kb transfer leading region of pKF3-94 is the fragment of DNA firstly transferred into the recipient bacterium during the conjugation process (Manwaring et al., 1999). It contains a large number of ORFs with unknown function and genes involved

in plasmid stability, including *ssb* (encoding plasmid-derived single-stranded DNA-binding protein), *flmA* (encoding stable plasmid inheritance protein), and *psiA* and *psiB* (genes related to plasmid SOS inhibition) (Figure 5A). Of the 48 ORFs in the transfer leading region of pKF3-94, 47 were found to be similar to those of pK1HV (the similarity of amino acid sequences are between 92.11 and 100%, except for two ORFs where they are 49.23 and 74.14%, respectively). Forty-two of the 48 ORFs are similar to those of pC15-k (the similarity of amino acid sequences are between 88.46 and 100%). In addition, compared with the sequence of pKF3-94, there are some insertions in pK1HV and pC15-k, such as the insertion sequence (including a transposase gene) downstream of *higA* of pK1HV and the insertion sequence (including an integrase gene) upstream of *flmA* of pC15-k. Despite these insertions, the backbone of this region (including 14 genes) remains unchanged. This suggests that the transfer leading regions of pKF3-94, pK1HV, and pC15-k share a common origin. Nevertheless, the transfer leading region of pKF3-94 is more dissimilar to those of pKF3-70/pKF3-140 compared with to those of pC15-k/pK1HV. Of the 48 ORFs encoded on the transfer leading region of pKF3-94, only 21 correspond to homologous genes in pKF3-70 (the similarity of amino acid sequences is between 36.29 and 80.26%), and only 19 correspond to homologous genes in pKF3-140 (the similarity of



**FIGURE 5 | Comparison of genome structure of the transfer leading (A) and conjugation regions (B) of pKF3-94, pK1HV, pC15-k, pKF3-70, and pKF3-140. The homologous genes present in the five plasmids are marked with the same colors, while the non-homologous genes are left blank.**

amino acid sequences is between 36.40 and 86.84%). This further supports the hypothesis that pKF3-94 share a closer phylogenetic relationship with pK1HV and pC15-k than with the plasmids of pKF3-70 or pKF3-140. Moreover, these five plasmids share 18 homologous genes in the transfer leading regions (**Figure 5A**).

In pKF3-94, the 34.6 kb conjugation region lies adjacent to the transfer leading region and encodes 40 CDSs. This region harbors 22 *tra* genes (*traM*, *traJ* etc.), 5 *trb* genes (*trbI*, *trbC* etc.), and a conjugal transfer repressor gene, *finO* (**Figure 5B**). The conjugation region of pK1HV is almost identical with that of pKF3-94, with a nucleotide sequence identity of 99.89% and sharing 38 homologous genes (the similarity of amino acid sequences were all above 99.57%). pC15-k shows a nucleotide sequence similarity of greater than 99% with pKF3-94 and shares 38 homologous genes (amino acid sequence similarities are all above 99.88%), although an insertion (a transposase gene with a size of 826 bp) upstream of *traC* was identified in pC15-k compared with the corresponding regions of pK1HV and pKF3-94. This suggests that the conjugation region of these three plasmids might have a common origin. Unexpectedly, the genome structure of the conjugation regions of pKF3-70 and pKF3-140 is also similar to that of pKF3-94. However, the degree of similarity of the homologous genes in this region of pKF3-94 and pKF3-70 or pKF3-140 is far lower than that of the corresponding regions of pKF3-94 and pK1HV or pC15-k. Twenty-nine of the 40 genes in pKF3-94 share an amino acid sequence similarity above 22.75% with pKF3-70, and only 16 genes show a similarity of 60% or more (with 85.19%, *traT* has the highest similarity). Between pKF3-94 and pKF3-140, there are 26 genes with similarity above 31.15%, and only 14 genes share a similarity above 60% (with 84.21%, *traT* has again the highest similarity). There are 23 homologous genes in these five plasmids (**Figure 5B**). In addition, we found that the conjugation region of pKF3-70 and pKF3-140 are more similar to each other than to that of pKF3-94. For example, we found an insertion of *traP* and *trbD* between *traB* and *traV* in both pKF3-70 and pKF3-140 and an insertion of *trbA* or IS5-transposase and IS1-transposase between *traF* and *traO* in pKF3-70 or pKF3-140, respectively. Nevertheless, pKF3-70 and pKF3-140 also have their distinct characteristics, including the unique genes of *trbA* and *ycjY* in pKF3-70, and the deletion of *traM* and *traS* in pKF3-140 (**Figure 5B**).

### The Homology Region between pKF3-94 and pC15-k

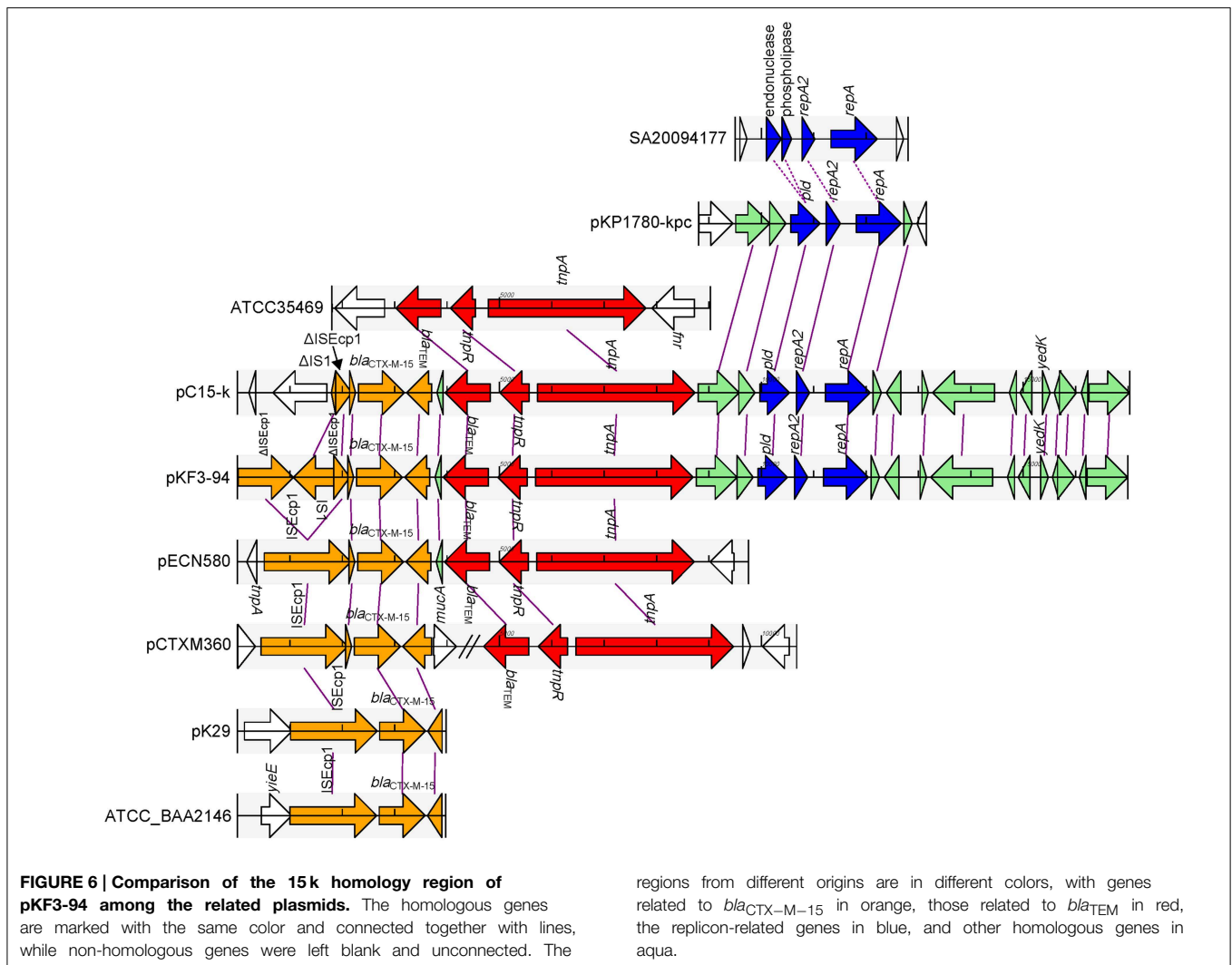
pKF3-94 and pC15-k share a 15 kb homology region with a sequence identity of 99%. This region could be divided into two parts. One part contains resistance genes, including *bla*<sub>CTX-M-15</sub> and *bla*<sub>TEM</sub>, and the other contains the FII<sub>K2</sub> replicon (**Figure 6**). In addition, this latter part also contains a few MGEs that help to trace the formation of this homology region.

In the part containing resistance genes, *bla*<sub>CTX-M-15</sub> is flanked by two truncated ISEcp1s (**Figure 6**). It has been demonstrated that ISEcp1 was able to mobilize an adjacent gene (*bla*<sub>CTX-M-15</sub>) (Zong et al., 2010; Dhanji et al., 2011). However, sequence analysis of this region in pKF3-94 showed that there is a complete IS1 (768 bp) inserted into ISEcp1 upstream of

*bla*<sub>CTX-M-15</sub>, thereby dividing the complete ISEcp1 into two portions of 1057 and 309 bp, respectively. Further analysis showed that a sequence of 290 bp is deleted in the middle of ISEcp1, thereby rendering it non-functional. Interestingly, the ISEcp1 in pC15-k is shorter than that in pKF3-94 and only contains the truncated IS1 (80 bp, identical to the downstream side of IS1 in pKF3-94) and the truncated ISEcp1 (309 bp, identical to downstream part of ISEcp1 in pKF3-94, **Figure 6**). This indicated that this region in pC15-k has been formed later than that in pKF3-94 and has probably arisen from a pKF3-94-like plasmid. Comparative genomics showed that the complete sequence of ISEcp1 and *bla*<sub>CTX-M-15</sub> are present in a few plasmid and bacterial genomes, such as in the plasmids pCTXM360 (from *K. pneumoniae*), pCTX-M3 (from *Citrobacter freundii*), pECL0701 (from *Enterobacter cloacae*), and pECN580 (from *E. coli*), and the chromosomes of *E. coli* JJ1886 and *K. pneumoniae* strains of ATCC BAA-2146, Kp13, and JM45. This finding suggests that the ISEcp1 enabled the spread of *bla*<sub>CTX-M-15</sub> among a variety of plasmids and chromosomes in bacteria.

The other resistance gene, *bla*<sub>TEM</sub>, along with its adjacent genes of *tnpR* and *tnpA*, is associated with a Tn3 transposon (**Figure 6**). The *bla*<sub>TEM</sub> gene encoding TEM β-lactamase and conferring resistance to the penicillin family of antibiotics was found in a group of three closely related transposons, Tn1, Tn2, and Tn3 (Bailey et al., 2011). This fragment of pKF3-94 (about 4.8 kb in length) is almost identical with that of pC15-k. The Tn3 transposon is incomplete, as its left side is shortened by 124 bp compared with the complete Tn3 found in the database. The deletion probably happened during recombination of Tn3 transposition. Complete or incomplete Tn3 transposons are present in a large number of plasmids such as pECN580 (incomplete, from *E. coli*), pKo6 (incomplete, from *K. pneumoniae*), pKOX\_R1 (incomplete, from *K. oxytoca*), pCTXM360 (complete, from *K. pneumoniae*), and pCFSAN007428\_01 (complete, from *Salmonella enterica*). A complete Tn3 transposon (4948 bp) was also found in the chromosomes of *Escherichia fergusonii* ATCC35469 and *Enterobacteria* phage P7, and showed nucleotide sequence identity of 100% compared to the corresponding region in pKF3-94. This suggests that the Tn3 transposon of pKF3-94 and pC15-k probably shares a common origin with those in the genomes of *E. fergusonii*, and *Enterobacteria* phage P7.

The replicon encoding part of the above mentioned 15 kb homology region is about 2.6 kb in length. It mainly encodes a FII<sub>K2</sub> replicon containing *repA* (encoding RepA), *repA2* (encoding RepA2), and *pld* (**Figure 6**). An FII<sub>K2</sub> replicon sequence has also been identified in *K. pneumoniae* plasmids such as pKP1780-kpc or pKpQIL-IT (FII<sub>K2</sub> replicon, see **Table 1**; both with a nucleotide sequence identity of 99% compared to that of pKF3-94). No such FII<sub>K2</sub> replicon is identified in *K. pneumoniae* chromosomes at present. Interestingly, a similar region was found in *Salmonella enterica* chromosomes such as SA20094177 (FII<sub>S1</sub> replicon, see **Table 1**; with nucleotide sequence identity of 72%). Furthermore, the RepA and RepA2 proteins on pKF3-94 and SA20094177 show an amino acid sequence similarity of 91 and 75%, respectively. Sequence comparison shows that



the phospholipase and endonuclease genes located upstream of *repA2* in SA20094177 are also similar to *pld* in pKF3-94 (with amino acid sequence similarity of 78 and 92%, respectively; **Supplementary Figure S3**). We deduce from this that the FII<sub>S1</sub> replicon in SA20094177 and the FII<sub>K2</sub> replicon in pKF3-94 probably evolved from a common origin.

Besides its presence in pKF3-94 and pC15-k, the part containing resistance genes was also found in some other plasmids, including pECN580, pK06, and pK0X\_R1. As shown in pKF3-94 and pC15-k, the two resistance genes (*bla*<sub>CTX-M-15</sub> and *bla*<sub>TEM</sub>) and their related MGEs are situated next to each other while in the plasmid pCTXM360, they are separated by other sequences (**Figure 6**). The plasmid pKP1780-kpc only harbors the 2.6 kb part containing the replicon, suggesting that the two parts of the 15 kb region in pKF3-94 were probably derived from three different origins.

## Conclusion

Using the program Plasmid RST, we identified the replicons of the three IncF plasmids pKF3-70, pKF3-94, and pKF3-140. pKF3-70

is a one-replicon plasmid while pKF3-94 and pKF3-140 are both multi-replicon plasmids. Despite the fact that the three plasmids carry similar FII replicons of the same incompatibility group, the presence of the replicons of the other incompatibility groups in pKF3-94 and pKF3-140 enable them to co-exist in the same host and hence participate in widespread horizontal gene transfer.

Through extensive comparative genomics analysis, our work demonstrates that pKF3-94, pK1HV, and pC15-k share a common backbone structure but also contain their own variable regions. It is conceivable that the backbone structure of pC15-k and pK1HV is derived from that of a pKF3-94-like plasmid, and the variable regions are derived from a variety of plasmid or chromosomal sequences by HGT. The results lend further support for the role horizontal gene transfer plays in plasmid genome evolution and in the dissemination of genetic elements including drug resistance genes.

## Materials and Methods

The host strain *K. pneumoniae* KF3 which harbored plasmid pKF3-94 was isolated from the laboratory of the First Affiliated



Hospital of Wenzhou Medical University, Wenzhou, China in 2006. The plasmid pKF3-94 is one of three plasmids isolated from *K. pneumoniae* KF3. The other two plasmids named pKF3-70 (Yi et al., 2010) and pKF3-140 (Bai et al., 2013) have been reported earlier. Sanger sequencing was used to sequence three pKF3 plasmids. pKF3-94 was extracted using alkaline lysis method (Felicciello and Chinali, 1993). The purified plasmid DNA was sheared by a HydroShear DNA shearing device (volume, 200 ml; cycle number, 20; speed code, 7–8). Fragments of 1.6–3.0 Kb were recovered from agarose gel electrophoresis and ligated into a pUC18 vector. Clones were sequenced using an ABI 3730 automated sequencer. The derived sequences were assembled using the Phred/Phrap/Consed software package (<http://www.phrap.org/phredphrapconsed.html>).

The plasmids and chromosome genome sequences used in this study for comparative analysis and phylogenetic analysis were downloaded from the NCBI Database (<http://www.ncbi.nlm.nih.gov>). The 56 plasmids were selected based on the whole genome sequence (pKF3-94) comparison against the whole plasmids available in NCBI database with a cut-off value (max score) of about 8700. The accession numbers of the related plasmids and the chromosome sequences mainly mentioned in this paper were pKF3-94 (NC\_013950.1/FJ876826.1), pKF3-70 (NC\_013542.1), pKF3-140 (NC\_013951.1), pK1HV (NC\_020087.1/HF545434.1), pC15-k (NC\_015154.1/HQ202266.1), pECN580 (KF914891.1), pCTXM360 (NC\_011641.1), pK29 (EF382672.1), pKP1780-kpc (KF874497.2), CCBH10892 plasmid (KF727591.2), *E. fergusonii* ATCC35469 chromosome (CU928158.2), *K. pneumoniae* ATCC BAA-2146 (CP006659.1), and *S. enterica* subsp. SA20094177 (CP007468.1).

Multiple sequence alignments were performed using MAFFT (Kato and Standley, 2013). Gene concatenation tree was reconstructed by Bayesian method using TOPALi v2 (Milne et al., 2009) and MrBayes (<http://mrbayes.csit.fsu.edu/>), incorporating the GTR+G model of nucleotide substitution. The generation number and sample frequency sampling were fixed to 2,000,000 and 100, respectively. Every clade was supported by Bayesian posterior probabilities (BPP). The best-fitting models [GTR+G model in **Figure 2**, HKY model in **Figure 4B** (R1) and **Supplementary Figure S2** (R2, R3), K80 model in **Figure 4B** (R4)] of the nucleotide substitutions was selected using modelgenerator (Keane et al., 2006). Visualization and annotation of phylogenetic trees were done using Figtree (<http://tree.bio.ed.ac.uk/software/figtree/>). Potential open reading frames (ORFs) were predicted and annotated using Glimmer3 (<http://www.cbc.umd.edu/software/glimmer>) and

BASys (Van Domselaar et al., 2005), respectively. The Plasmid RST was identified with default parameters by Plasmid MLST Database (<http://pubmlst.org/plasmid/>; Updated in June, 2014). Comparisons of the nucleotide sequences were made using BlastN. Insertion sequences were predicted using ISFinder (Siguier et al., 2006). Orthologous groups of genes from plasmids (pKF3-94, pK1HV, pC15-k, etc.) were identified using BLASTp and inparanoid (Remm et al., 2001). Other bioinformatics software was written using Python (<https://www.python.org/>) and Biopython (Cock et al., 2009).

## Author Contributions

JY designed and performed experiments, contributed to the interpretation of results, wrote the manuscript and approved the final version for publication; QB designed experiments, wrote the manuscript and approved the final version for publication. SW, KZ, ZW, WZ, YZ, MZ, CC, HW, HT, CZ, PL, JY, TX, HY, JL, LN, ZX, and JL approved the final version for publication.

## Acknowledgments

This work is supported by the Natural Science Foundation of Zhejiang Province, China (LY14C060005); the Science and Technology Foundation of National Health and Family Planning Commission of China (WKJ2012-2-032); the National Natural Science Foundation of China (81401702); Technology Innovation Team of Zhejiang Province, China (2010R50048-13); Science and Technology Foundation of Wenzhou City, China (Y20110101, Y20140422); Science and Technology Foundation of Wenzhou Municipal Health Bureau (2011A021).

## Supplementary Material

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fmicb.2015.00831>

**Supplementary Figure S1 | Collinearity analysis of the plasmid genomes of pKF3-94, pK1HV, pC15-k, pKF3-70, and pKF3-140.** The four different colors of lines indicate the orthologous genes between each plasmid pair.

**Supplementary Figure S2 | The phylogenetic trees of R2 and R3.**

**Supplementary Figure S3 | Amino acid sequences alignment of *pld* gene of pKF3-94 and phospholipase/endonuclease genes of SA20094177.** The identical amino acids are outlined with gray shadow, while the similar ones are outlined and left blank.

## References

- Bahl, M. I., Hansen, L. H., and Sorensen, S. J. (2009). Persistence mechanisms of conjugative plasmids. *Methods Mol. Biol.* 532, 73–102. doi: 10.1007/978-1-60327-853-9\_5
- Bai, J., Liu, Q., Yang, Y., Wang, J., Li, J., Li, P., et al. (2013). Insights into the evolution of gene organization and multidrug resistance from *Klebsiella pneumoniae* plasmid pKF3-140. *Gene* 519, 60–66. doi: 10.1016/j.gene.2013.01.050
- Bailey, J. K., Pinyon, J. L., Anantham, S., and Hall, R. M. (2011). Distribution of the blaTEM gene and blaTEM-containing transposons in commensal *Escherichia coli*. *J. Antimicrob. Chemother.* 66, 745–751. doi: 10.1093/jac/dkq529
- Brigulla, M., and Wackernagel, W. (2010). Molecular aspects of gene transfer and foreign DNA acquisition in prokaryotes with regard to safety issues.

- Appl. Microbiol. Biotechnol.* 86, 1027–1041. doi: 10.1007/s00253-010-2489-3
- Cock, P. J., Antao, T., Chang, J. T., Chapman, B. A., Cox, C. J., Dalke, A., et al. (2009). Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics* 25, 1422–1423. doi: 10.1093/bioinformatics/btp163
- Couturier, M., Bex, F., Bergquist, P. L., and Maas, W. K. (1988). Identification and classification of bacterial plasmids. *Microbiol. Rev.* 52, 375–395. doi: 10.1186/1471-2148-6-29
- Dhanji, H., Patel, R., Wall, R., Doumith, M., Patel, B., Hope, R., et al. (2011). Variation in the genetic environments of bla(CTX-M-15) in *Escherichia coli* from the faeces of travellers returning to the United Kingdom. *J. Antimicrob. Chemother.* 66, 1005–1012. doi: 10.1093/jac/dkr041
- Doi, Y., Hazen, T. H., Boitano, M., Tsai, Y. C., Clark, T. A., Korch, J., et al. (2014). Whole-genome assembly of *Klebsiella pneumoniae* coproducing NDM-1 and OXA-232 carbapenemases using single-molecule, real-time sequencing. *Antimicrob. Agents Chemother.* 58, 5947–5953. doi: 10.1128/AAC.03180-14
- Dolejska, M., Villa, L., Dobiasova, H., Fortini, D., Feudi, C., and Carattoli, A. (2013). Plasmid content of a clinically relevant *Klebsiella pneumoniae* clone from the Czech Republic producing CTX-M-15 and QnrB1. *Antimicrob. Agents Chemother.* 57, 1073–1076. doi: 10.1128/AAC.01886-12
- Feliciello, I., and Chinali, G. (1993). A modified alkaline lysis method for the preparation of highly purified plasmid DNA from *Escherichia coli*. *Anal. Biochem.* 212, 394–401. doi: 10.1006/abio.1993.1346
- Gadagkar, S. R., Rosenberg, M. S., and Kumar, S. (2005). Inferring species phylogenies from multiple genes: concatenated sequence tree versus consensus gene tree. *J. Exp. Zool. B Mol. Dev. Evol.* 304, 64–74. doi: 10.1002/jez.b.21026
- Jain, R., Rivera, M. C., Moore, J. E., and Lake, J. A. (2003). Horizontal gene transfer accelerates genome innovation and evolution. *Mol. Biol. Evol.* 20, 1598–1602. doi: 10.1093/molbev/msg154
- Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. doi: 10.1093/molbev/mst010
- Keane, T. M., Creevey, C. J., Pentony, M. M., Naughton, T. J., and McInerney, J. O. (2006). Assessment of methods for amino acid matrix selection and their use on empirical data shows that ad hoc assumptions for choice of matrix are not justified. *BMC Evol. Biol.* 6:29. doi: 10.1186/1471-2148-6-29
- Landman, D., Bratu, S., Kochar, S., Panwar, M., Trehan, M., Doymaz, M., et al. (2007). Evolution of antimicrobial resistance among *Pseudomonas aeruginosa*, *Acinetobacter baumannii* and *Klebsiella pneumoniae* in Brooklyn, NY. *J. Antimicrob. Chemother.* 60, 78–82. doi: 10.1093/jac/dkm129
- Manwaring, N. P., Skurray, R. A., and Firth, N. (1999). Nucleotide sequence of the F plasmid leading region. *Plasmid* 41, 219–225. doi: 10.1006/plas.1999.1390
- Milne, I., Lindner, D., Bayer, M., Husmeier, D., McGuire, G., Marshall, D. F., et al. (2009). TOPALi v2: a rich graphical interface for evolutionary analyses of multiple alignments on HPC clusters and multi-core desktops. *Bioinformatics* 25, 126–127. doi: 10.1093/bioinformatics/btn575
- Novick, R. P. (1987). Plasmid incompatibility. *Microbiol. Rev.* 51, 381–395.
- Périchon, B., Bogaerts, P., Lambert, T., Frangeul, L., Courvalin, P., and Galimand, M. (2008). Sequence of conjugative plasmid pIP1206 mediating resistance to aminoglycosides by 16S rRNA methylation and to hydrophilic fluoroquinolones by efflux. *Antimicrob. Agents Chemother.* 52, 2581–2592. doi: 10.1128/AAC.01540-07
- Remm, M., Storm, C. E., and Sonnhammer, E. L. (2001). Automatic clustering of orthologs and in-paralogs from pairwise species comparisons. *J. Mol. Biol.* 314, 1041–1052. doi: 10.1006/jmbi.2000.5197
- Siguier, P., Perochon, J., Lestrade, L., Mahillon, J., and Chandler, M. (2006). ISfinder: the reference centre for bacterial insertion sequences. *Nucleic Acids Res.* 34, D32–D36. doi: 10.1093/nar/gkj014
- Sobecky, P. A., and Coombs, J. M. (2009). Horizontal gene transfer in metal and radionuclide contaminated soils. *Methods Mol. Biol.* 532, 455–472. doi: 10.1007/978-1-60327-853-9\_26
- Van Domselaar, G. H., Stothard, P., Shrivastava, S., Cruz, J. A., Guo, A., Dong, X., et al. (2005). BASys: a web server for automated bacterial genome annotation. *Nucleic Acids Res.* 33, W455–W459. doi: 10.1093/nar/gki593
- Villa, L., García-Fernández, A., Fortini, D., and Carattoli, A. (2010). Replicon sequence typing of IncF plasmids carrying virulence and resistance determinants. *J. Antimicrob. Chemother.* 65, 2518–2529. doi: 10.1093/jac/dkq347
- Yi, H., Xi, Y., Liu, J., Wang, J., Wu, J., Xu, T., et al. (2010). Sequence analysis of pKF3-70 in *Klebsiella pneumoniae*: probable origin from R100-like plasmid of *Escherichia coli*. *PLoS ONE* 5:e8601. doi: 10.1371/journal.pone.0008601
- Zhao, F., Bai, J., Wu, J., Liu, J., Zhou, M., Xia, S., et al. (2010). Sequencing and genetic variation of multidrug resistance plasmids in *Klebsiella pneumoniae*. *PLoS ONE* 5:e10141. doi: 10.1371/journal.pone.0010141
- Zong, Z., Partridge, S. R., and Iredell, J. R. (2010). ISEcp1-mediated transposition and homologous recombination can explain the context of bla(CTX-M-62) linked to qnrB2. *Antimicrob. Agents Chemother.* 54, 3039–3042. doi: 10.1128/AAC.00041-10

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Ying, Wu, Zhang, Wang, Zhu, Zhu, Zhang, Cheng, Wang, Tou, Zhu, Li, Ying, Xu, Yi, Li, Ni, Xu, Bao and Lu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.