



Recombination Drives Evolution of Carbapenem-Resistant *Klebsiella pneumoniae* Sequence Type 11 KL47 to KL64 in China

Tao Chen,^a Yuan Wang,^a Yanzi Zhou,^a Wangxiao Zhou,^a Xiaohui Chi,^a Ping Shen,^{a,b}  Beiwen Zheng,^a Yonghong Xiao^a

^aState Key Laboratory for Diagnosis and Treatment of Infectious Diseases, National Clinical Research Center for Infectious Diseases, Collaborative Innovation Center for Diagnosis and Treatment of Infectious Diseases, the First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, Zhejiang, China

^bJinan Microecological Biomedicine Shandong Laboratory, Jinan, Shandong, China

ABSTRACT Carbapenem-resistant *Enterobacteriaceae*, especially carbapenemase-producing *Klebsiella pneumoniae*, is an urgent problem in health care facilities worldwide. *K. pneumoniae* isolates classified as sequence type 11 (ST11) are largely responsible for the spread of carbapenem-resistant *K. pneumoniae* (CRKP) in China. Our previous phylogenetic reconstruction suggested that CRKP ST11 capsular locus 64 (KL64) was derived from an ST11-KL47 ancestor through recombination. However, the molecular origin of KL64 remains largely unknown, and our understanding of the recombination is incomplete. Here, we screened a global sample of 22,600 *K. pneumoniae* genomes and searched for KL64-harboring STs, which were found to be ST1764, ST3685, ST1764-1LV, ST30, ST505, ST147, and ST11, wherein ST1764, ST3685, ST1764-1LV, and ST30 belonged to a clonal complex, CC1764. We compared the genetic structures of representative strains from ST11-KL47, ST11-KL64, CC1764-KL64, ST505-KL64, and ST147-KL64 and further performed phylogenetic analysis and single-nucleotide polymorphism analysis among 248 isolates from all these STs. The results suggested a recombination event has occurred in a homologous ~154-kb region covering KL and the lipopolysaccharide biosynthesis locus (OL) between a recipient ST11-KL47-OL101 and a donor CC1764 (except ST30), giving rise to ST11-KL64-O2v1 strains (recombination I). Furthermore, we also found an infrequent ST11-KL64-O2v1 subclone which was not produced by recombination I but was hybridized from ST11-KL47-OL101 and ST147-KL64-O2v1 strains through recombination of a homologous ~485-kb region covering KL and OL (recombination II). Our findings provide important insights into the role of recombination in the evolution of clinical strains and the diversity of capsule and lipopolysaccharide of widely distributed KPC-associated ST11 *K. pneumoniae* in China.

IMPORTANCE Chromosomal recombination events are considered to contribute to the genetic diversification and ultimate success of many bacterial pathogens. A previous study unravelled the molecular evolution history of ST258 strains, which have been largely responsible for the spread of KPC in the United States. Here, we used comparative genomic analyses to discover two recombination events in ST11 CRKP strains, which is a predominant KPC-associated CRKP clone in China. Two new ST11-CRKP subclones with altered capsule and lipopolysaccharide, which are two primary determinants of antigenicity and antigenic diversity among *K. pneumoniae*, were produced through these two recombination events, respectively. Horizontal transfer of the KL and OL appears to be a crucial element driving the molecular evolution of *K. pneumoniae* strains. These findings not only extend our understanding of the molecular evolutionary history of ST11 but also are an important step toward the development of preventive, diagnostic, and therapeutic strategies for CRKP.

KEYWORDS *Klebsiella pneumoniae*, ST11, recombination, evolution, capsule, lipopolysaccharide

Editor Chi-Tai Fang, National Taiwan University

Copyright © 2023 Chen et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Yonghong Xiao, xiaoyonghong@zju.edu.cn.

The authors declare no conflict of interest.

Received 18 July 2022

Accepted 9 December 2022

Published 9 January 2023

In recent decades, carbapenem-resistant *Enterobacteriaceae* (CRE) have become an urgent public health concern worldwide with few treatment options. Among these, carbapenem-resistant *Klebsiella pneumoniae* (CRKP) strains account for more than half of clinical CRE infections in the European Union, United States, and China (1–3). In the United States, CRKP sequence type 258 (allelic profile 3-3-1-1-1-79) emerged as a notable clinical problem during the early to mid-2000s and remains the main ST in the United States and elsewhere (4, 5). Phylogenetic analysis of the core genome of ST258 *K. pneumoniae* strains revealed that they are comprised of two distinct genetic clades (ST258 clades I and II), largely due to an ~215-kb region of divergence that includes genes involved in capsular polysaccharide (CPS) biosynthesis (6). Furthermore, ST258 is thought to be a hybrid clone caused by a recombination event—80% of the genome originated from ST11-like strains and 20% from ST442-like strains, and the ST258 clade I strains evolved from a clade II strain as a result of the replacement of an ~52-kb region encoding the CPS biosynthetic machinery (7). Extensive capsule locus variation is believed to be primarily responsible for clonal group 258 (CG258) diversification, and recombination spanning the CPS locus is considered a crucial mechanism by which lineages diversify their antigenic profile against exterior selection pressures (8, 9). These studies consistently indicated that the genomic region involved in CPS biosynthesis in *Klebsiella pneumoniae* is a vital recombination hot spot.

In China, ST11 (allelic profile 3-3-1-1-1-4), a single-locus (*tonB*) variant of ST258, is believed to be the most dominant *Klebsiella pneumoniae* carbapenemase (KPC)-associated CRKP clone, with a prevalence greater than 60% (10). Currently, more concerning is the emergence of carbapenem-resistant hypervirulent *Klebsiella pneumoniae* (CR-hvKP), which may lead to even higher mortality and morbidity (11). Of note, several interhospital outbreaks of ST11 CR-hvKP with capsular locus 64 (KL64) and plasmid-encoded virulence genes were reported recently (12, 13). Among ST11 CR-hvKP in China, the group of KL64 strains is the most widespread, as it has been observed to harbor more plasmid-encoded virulence genes than ST11 non-KL64 strains (14–16). Our previous study found a subclonal shift in the dominant clone ST11 CRKP in which the previously prevalent KL47 had been replaced by KL64 and enhanced virulence of ST11-KL64 CRKP was found by phenotypic tests (17). By using a Bayesian dating method, we found that KL64 isolates probably evolved from a KL47 ancestor, but the molecular donor of KL64 and the hypothetical recombination event were not further explored.

In this study, we sought to find the molecular origin of KL64 and a probable recombination event leading to the capsular switching from KL47 to KL64 among CRKP-ST11 populations in China. In addition, we discovered a special infrequent ST11-KL64 subclone, different from the predominant ST11-KL64 subclone, which might have been produced by another recombination event.

RESULTS

An ~154-kb genomic region covering KL and OL was distinctly different between ST11-KL47 and ST11-KL64 *K. pneumoniae*. The genomes of two ST11 *K. pneumoniae* clinical isolates (KP16932 and KP47434) were obtained from a hybrid assembly (GenBank accession numbers [QVAN00000000](#) [KP16932] and [QURI00000000](#) [KP47434]). The result of genomic analysis showed KP16932 contained CPS biosynthesis locus (K_locus [KL]) KL47 and LPS biosynthesis locus (O_locus [OL]) OL101, while KP47434 contained KL64 and O2v1. We took these two complete genomes as representative genomes for ST11-KL47 and ST11-KL64 isolates, respectively. As KL is close to OL in genomic location, further analysis of the KL, OL, and flanking nucleotide sequences revealed that the differences between ST11-KL47 and ST11-KL64 are expansive, covering an ~154-kb contiguous region corresponding to nucleotide positions ~4,060,800 to ~4,214,550 in strain KP47434. This ~154-kb contiguous region covers the ~52-kb region of which the replacement was believed to cause the evolution of ST258 clade I strains from a clade II strain (7). The ~154-kb region contained 129 open reading frames (ORFs), such as KL, OL, the *hisLGDCBHAFI* operon (which encodes the enzymes of the histidine biosynthesis pathway), *mdtABCDQ* multidrug transporters, the two-component system BaeSR (involved in envelope stress response), and so on. All ORFs annotated by the RAST server in the ~154-kb region are listed in Table S3 in the supplemental

material. We carefully compared the ~154-kb contiguous regions among these isolates. Analysis of single-nucleotide polymorphisms (SNPs) in the genomes of KP16932 and KP47434 indicated that these strains differed by 1,233 SNPs, of which 903 SNPs (~73%) were concentrated in the contiguous ~154-kb region. Furthermore, 79% coverage and 97.94% identities were found in the ~154-kb contiguous region between KP16932 and KP47434. According to these results in combination with the previous results from the Bayesian dating method, which indicated KL64 isolates probably evolved from a KL47 ancestor (17), we speculated that the ~154-kb region covering KL and OL was a recombination region acquired from other unrelated STs by ST11-KL47 to evolve into ST11-KL64.

CPS biosynthesis and LPS biosynthesis loci of *K. pneumoniae* isolates belong to different STs. To seek the possible molecular origin of recombination region, we selected *K. pneumoniae* strains of different STs which harbored KL64 genetic locus from 22,600 genomes as described in Materials and Methods. We found that KL64 could be detected in the genomes of ST1764 (allelic profile 5-3-1-1-9-4-283), ST3685 (5-3-1-232-9-4-283), ST1764-1LV (5-3-1-1-9-4-683), ST30 (5-3-1-1-9-4-26), ST505 (7-1-5-1-1-1-84), ST147 (3-4-6-1-7-4-38), and ST11 (3-3-1-1-1-1-4). Among these, ST1764, ST3685, ST1764-1LV, and ST30 belonged to a clonal complex, CC1764. A complete genome was selected as a representative genome for each ST (or CC): strain WSHvKp (GenBank accession number [GCA_013283935.1](https://www.ncbi.nlm.nih.gov/nuccore/GCA_013283935.1)) for CC1764 except ST30, strain INF357-sc-2280189 (GenBank accession number [GCA_904863125.1](https://www.ncbi.nlm.nih.gov/nuccore/GCA_904863125.1)) for ST30, strain ATCC35657 (GenBank accession number [GCA_001936035.1](https://www.ncbi.nlm.nih.gov/nuccore/GCA_001936035.1)) for ST505, and strain 11305 (GenBank accession number [GCA_009931095.1](https://www.ncbi.nlm.nih.gov/nuccore/GCA_009931095.1)) for ST147. The results of genomic analysis showed that strains WSHvKp, INF357-sc-2280189, and ATCC 35657 contained the same KL and OL (KL64 and O1v1), while strain 11305 harbored KL64 and O2v1. We further compared the nucleotide sequences of the KL and OL in these isolates by using BLAST. As shown in Fig. 1, KP47434, WSHvKp, INF357-sc-2280189, ATCC 35657, and 11305 displayed almost identical nucleotide sequences for KL and OL which were distinctly different from those of KP16932.

The ~154-kb recombination region covering KL and OL in ST11-KL64 was derived from CC1764 (except ST30). To find the possible ST from which the recombination region originated, we compared the contiguous ~154-kb regions among these representative isolates. Interestingly, the ~154-kb regions were nearly identical in KP47434 and WSHvKp, differing by only 8 SNPs, while the genomes of KP47434 and WSHvKp differed by 22,407 SNPs in total. By comparison, the SNP distances between KP47434 and INF357-sc-2280189, ATCC 35657, and 11305 in this region were 35, 720, and 770 SNPs, respectively. Furthermore, 99% coverage and almost 100% identities were found in this region between KP47434 and WSHvKp, indicating the ~154-kb region in ST11-KL64 strains was more closely related to strains of ST1764 than ST30, ST505, and ST147. These results suggested that the ~154-kb region in ST11-KL64 strains might have been derived from CC1764 (except ST30), which led to an assumption that a recombination event (recombination I) probably has occurred in a homologous ~154-kb region covering KL and OL between a recipient ST11-KL47-OL101 and a donor CC1764 (except ST30), giving rise to ST11-KL64-O2v1.

To provide further evidence to the speculation, we performed comparative sequence analysis focusing primarily on mobile genetic elements of the contiguous ~154-kb region and flanking sequences. We found an ISKpn1 (nucleotide positions 4,051,983 to 4,053,427 in strain KP47434) and an *ltrA* (nucleotide positions 4,214,595 to 4,216,504 in strain KP47434) encoding group II intron reverse transcriptase or maturase at the neighboring sequences upstream and downstream from this ~154-kb region in ST11 strains (KP16932 and KP47434), respectively. As expected, they were absent in ST1764 strain WSHvKp (Fig. 2). Both of these findings indicated the ISKpn1 and *ltrA* were probably not a part of the recombination region. Moreover, inside the ~154-kb recombination region we assumed, two ISKpn1 (nucleotide positions 4,079,631 to 4,081,075 and 4,210,135 to 4,211,579 in strain KP47434) at the identical position with the same orientation were found in KP47434 and WSHvKp, which were absent in KP16932 (Fig. 2). All these findings provided further support for the assumed recombination event. Unexpectedly, inside the ~154-kb recombination region, an ISKpn26 was observed in KP47434 (nucleotide positions 4,182,516 to 4,183,711 in strain KP47434) but an ISKpn74 (nucleotide positions 1,724,517 to 1,725,572 in strain WSHvKp) was found in WSHvKp at the

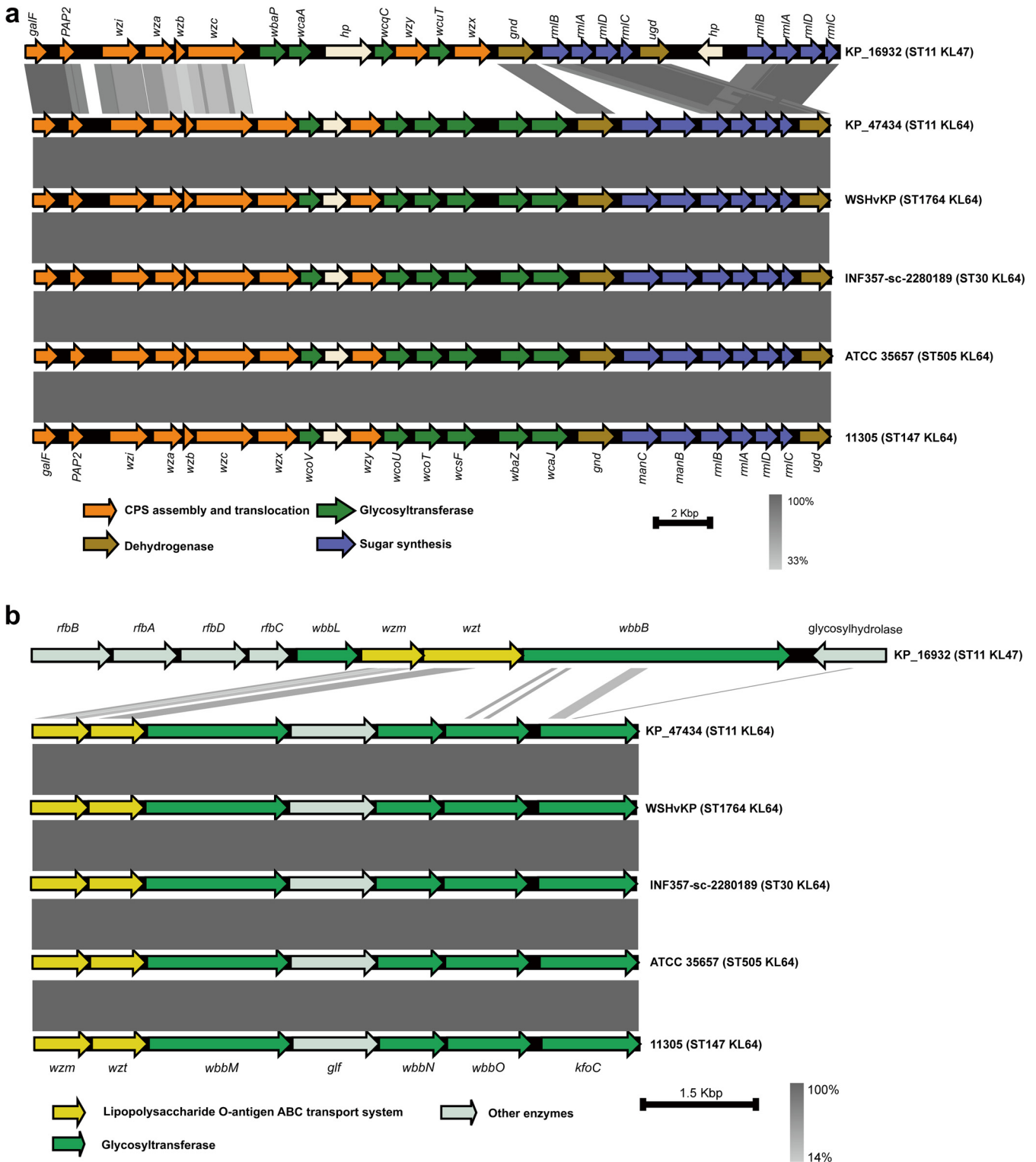


FIG 1 Comparison of the KL (a) and OL (b) in strains KP16932 (ST11-KL47), KP47434 (ST11-KL64), WSHvKp (ST1764-KL64), INF357-sc-2280189 (ST30-KL64), ATCC 35657 (ST505-KL64), and 11305 (ST147-KL64). Arrows indicate the direction, relative length, and function of ORFs. Homologous regions are connected by areas of different colors reflecting the degree of nucleotide identity. Sequence comparisons were performed using BLASTn and visualized using Easyfig.

same genomic location. Therefore, we supposed an insertion sequence (IS) substitution had occurred after the ~154-kb recombination event.

To analyze the hypothetical recombination region at the level of multiple isolates, we performed a phylogenetic analysis, including all of the 248 *K. pneumoniae* genomes used

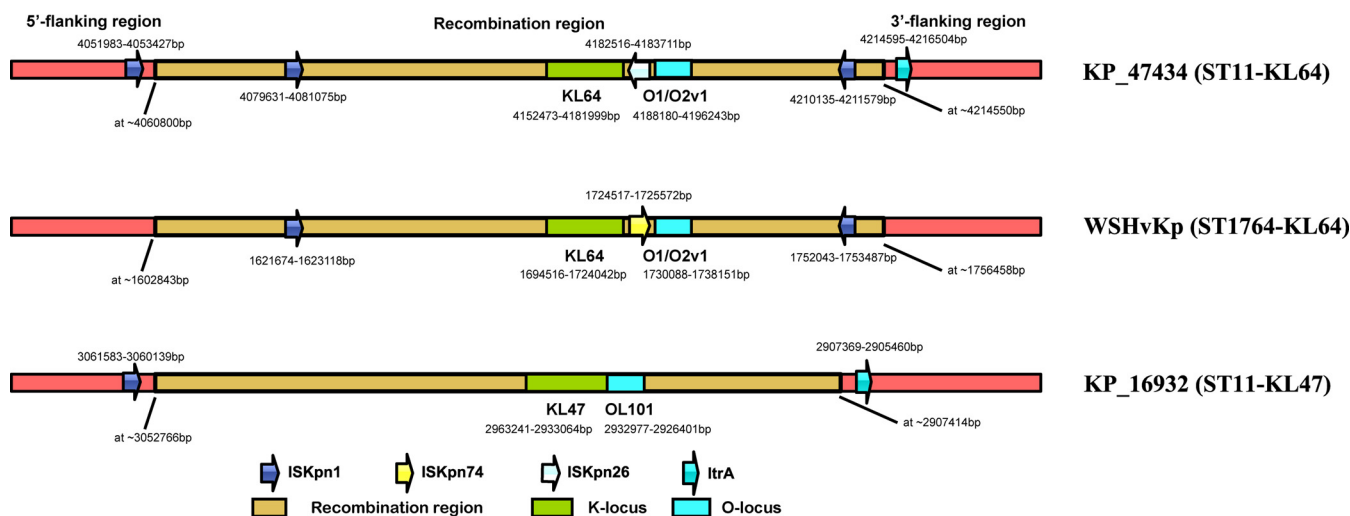


FIG 2 Mobile genetic elements analysis of the contiguous ~154-kb region and flanking sequences in KP16932 (ST11-KL47), KP47434 (ST11-KL64), and WSHvKP (ST1764-KL64) strains. Arrows indicate the direction of an IS. Different colors reflect the different kinds of ISs. The recombination region, KL, and OL are shown in different colors. The image was not drawn to scale.

in this work, on the genomic core SNPs and in parallel on the core SNPs located in the ~154-kb region. The phylogenetic tree of genomic core SNPs showed the 248 strains were clustered into two major groups, which were named clade A and clade B and contained 184 and 64 strains, respectively (Fig. 3 and Fig. S1). All strains in clade A belonged to ST11 with various KL and OL (including KL47-OL101, KL64-O2v1, KL103-O2v1, and KL105-O2v2), while all strains in clade B were other STs (CC1764-KL64-O1v1, ST505-KL64-O1v1, and ST147-KL64-O2v1). Almost all ST11 strains in clade A produced the well-known KPC-2 carbapenemase, while other STs did not. Interestingly, ST11-KL64 strain HB25-1, identified in the Second Affiliated Hospital of Zhejiang University in China (18), was not grouped with other ST11-KL64 strains in clade A, which suggested the emergence of a special ST11-KL64 subclone. In contrast, the resulting phylogenetic tree of the core SNPs located in the ~154-kb region demonstrated that 248 strains were clustered into three major groups, which were named clade I, clade II, and clade III, and contained 116, 86, and 46 strains, respectively (Fig. 4 and Fig. S1). The strains in clade I comprised ST11-KL64, CC1764-KL64, and ST505-KL64 strains. Moreover, CC1764-KL64 (except ST30) strains were more closely related to ST11-KL64 strains than ST30-KL64 and ST505-KL64 strains in clade I. The strains in clade II comprised strains of ST11-KL47, ST11-KL103, and ST11-KL105, while clade III contained all ST147-KL64 strains and the ST11-KL64 strain HB25-1. These results suggested that the ~154-kb region in ST11-KL64 strains (except strain HB25-1) was more closely related to strains of CC1764 (except ST30) than ST30, ST505, and ST147, providing further support to the assumed recombination event. Notably, the ST11-KL64 strain HB25-1 was grouped with ST147-KL64 strains in clade III (see below).

Minimum spanning trees (MSTs) were also generated based on SNP data for the genomes and the ~154-kb regions of 248 strains. The genomic core SNP distances of isolates of different STs suggested that the isolates from different groups were distinctly related and isolates within the same group were highly related (Fig. S2). The ST11-KL64 group showed a significantly closer SNP distance from ST11-KL47 than other STs (only 1,099 SNPs between the central strains of ST11-KL64 and ST11-KL47). However, it should be noted that strain HB25-1 showed a distance of 2,191 SNPs from the central strain of ST11-KL64. By comparison, the core SNP distances within the ~154-kb region of isolates of different STs (Fig. S3) suggested the ~154-kb region in ST11-KL64 group (except strain HB25-1) was more closely related to CC1764 (except ST30) than ST30, ST11-KL47, ST505, and ST147 (only 2 SNPs between the central strains of ST11-KL64 and CC1764 except ST30). All these findings indicated that the donor of the ~154-kb region in ST11-KL64 was closely related to that for CC1764 isolates (except ST30), providing further support to the recombination hypothesis. Interestingly, the ST11-KL64

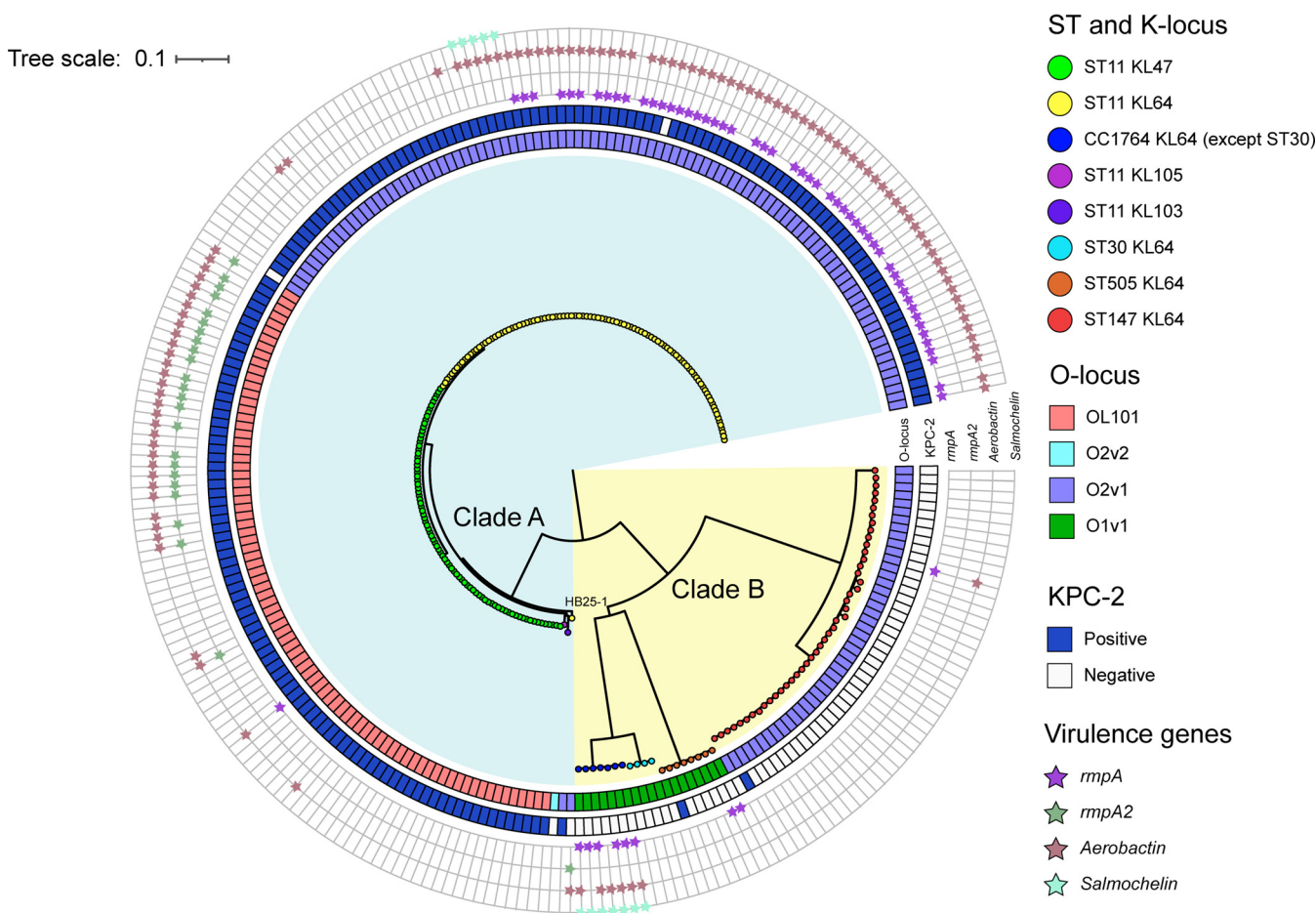


FIG 3 Genomic core SNP-based phylogenetic analysis of 248 *Klebsiella pneumoniae* strains, showing branch lengths. Colored rectangles indicate, from the innermost to the outermost, the lipopolysaccharide biosynthesis locus and presence of the *bla*_{KPC-2} gene. The presence of virulence genes is indicated by small pentacles of diverse colors.

isolate HB25-1 showed a closer SNP distance to ST147-KL64 isolates in the MST based on the ~154-kb region, suggesting novel recombination events or evolutionary convergence were associated with this specific ST11-KL64 isolate.

Given that ST11-KL64 is O2v1 and not O1v1 (as is CC1764-KL64), this aspect was at odds with our hypothesis. Thus, we performed a genomic analysis to find out what led to the different OL variants between ST11-KL64 and CC1764-KL64. It is known that both the O1 and O2 polysaccharide chains are based on a repeat unit named D-galactan I. The difference is that the O1 antigen is capped by a distal D-galactan II unit, whereas O2 is not. D-Galactan II is the only known O-antigen polysaccharide for which biosynthesis is enabled by genes (*wbbY* and *wbbZ*) located well away from the OL in *Klebsiella* genomes (19). As shown in Fig. 5, the ST11-KL47 chromosome lacks these genes (*wbbY* and *wbbZ*). Although the ~154-kb region covering OL in CC1764 (except ST30) was recombined to the chromosome of ST11-KL47, the OL of the recombination offspring could only directly transform from OL101 into O2 (as with ST11-KL64) but not O1 (as with CC1764-KL64) due to the absence of *wbbY* and *wbbZ*.

HB25-1: a rare ST11-KL64 subclone with an ~485-kb contiguous region derived from ST147-KL64. Because the ST11-KL64 isolate HB25-1 was not gathered with other ST11-KL64 strains in the phylogenetic tree that was based on the genomic core SNPs and gathered with ST147-KL64 strains based on the core SNPs located in the ~154-kb region, we speculated that another recombination event was associated with this specific isolate HB25-1. Likewise, we analyzed the KL, OL, and flanking nucleotide sequences to seek the differences between KP16932 and HB25-1, and an ~485-kb contiguous region corresponding to nucleotide positions ~1,676,640 to ~2,161,600 in strain HB25-1 was found to be significantly

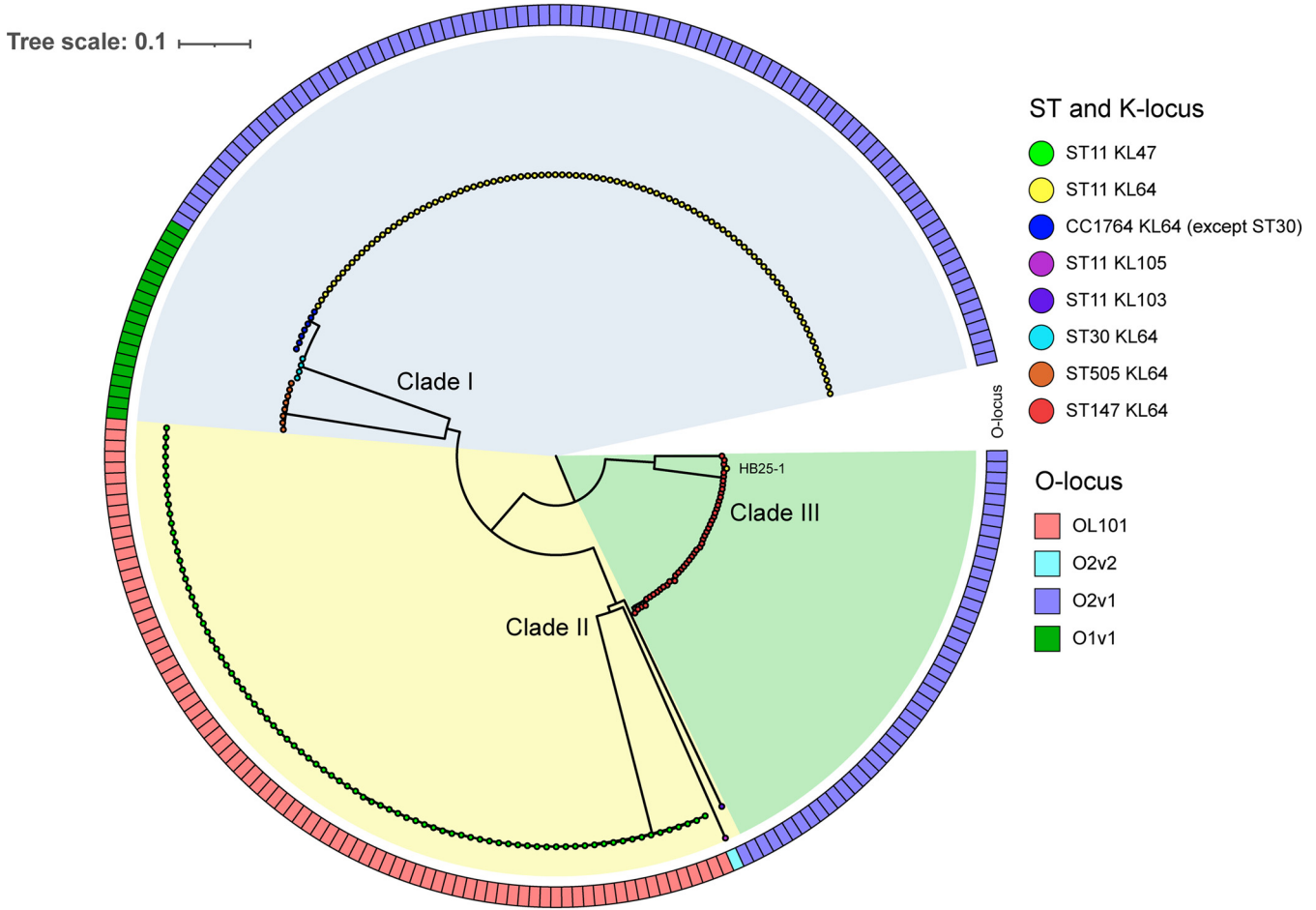


FIG 4 Phylogenetic analysis of 248 *Klebsiella pneumoniae* strains based on core SNPs located in the ~154-kb region. Different STs and KLs are marked with different colors. Rectangles in different colors indicate different lipopolysaccharide biosynthesis loci.

different. Analysis of SNPs in the genomes of KP16932 and HB25-1 indicated that these two strains differed by 3,001 SNPs, of which 2,079 SNPs (~69%) were concentrated in the contiguous ~485-kb region, implying that this region was probably a recombination region acquired from other unrelated STs. The RAST annotation results indicated that the ~485-kb region contained 456 ORFs, as shown in Table S4. Given the results that HB25-1 was gathered with ST147-KL64 strains based on the core SNPs located in the ~154-kb region (Fig. 4), we speculated that ST147-KL64 was the molecular donor of KL64 in strain HB25-1. We compared the SNP distances between HB25-1 and the ST147-KL64 strain 11305 in the ~485-kb region, and we found this region was nearly identical in HB25-1 and 11305, differing by only 26 SNPs. Furthermore, 90% coverage and 99.99% identities were found in this region between HB25-1 and 11305. The reason why the coverage was not nearly 100% was further found to primarily be the presence of mobile genetic elements, which were a 46,482-bp prophage (nucleotide positions 1,837,435 to 1,883,916) in strain 11305 and a 64,545-bp integrative and conjugative element (ICE; nucleotide positions 1,881,956 to 1,946,500) in strain HB25-1 (Fig. 5). These results indicated another recombination event (recombination II) probably occurred in a homologous ~485-kb region covering KL and OL between a recipient ST11-KL47-OL101 strain and a donor ST147-KL64-O2v1 strain, giving rise to a specific ST11-KL64-O2v1 subclone (like strain HB25-1) (Fig. 5).

To investigate the prevalence of the specific ST11-KL64 subclone like strain HB25-1, we performed a phylogenetic analysis of 623 genomes of ST11-KL64 *K. pneumoniae* strains isolated from China (including KP47434 and HB25-1) on the core SNPs located in the ~485-kb region. The results indicated only 1 of 621 genomes was grouped with strain HB25-1 in clade 1, while the other 620 genomes were grouped with KP47434 in clade 2 (see Fig. S4 in

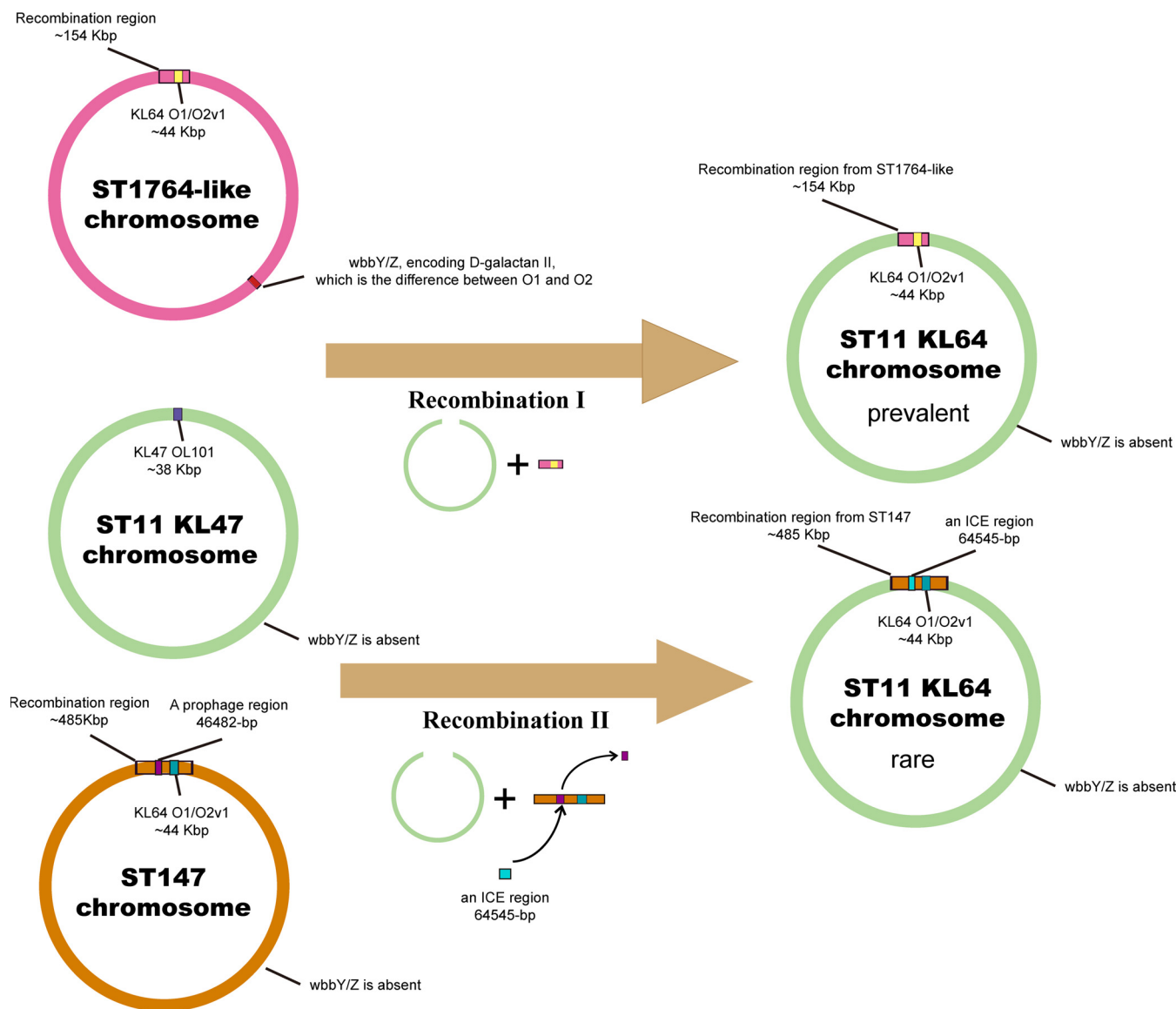


FIG 5 Two recombination events occurred in *Klebsiella pneumoniae* ST11 strains. Schematic representation based on the results of our analyses. Two recombination events (recombination I and recombination II) produced two ST11-KL64 descendants. The donors of KL64 and O2v1 among two ST11-KL64 subclones were CC1764 (except ST30) and ST147, respectively.

the supplemental material). These results suggested the ST11-KL64 subclones produced by recombination I were much more prevalent than those caused by recombination II in China.

DISCUSSION

The current rise of KPC-producing *K. pneumoniae* infections in health care facilities in Asian countries has been overwhelmingly related to strains belonging to ST11, which is a single-locus variant of ST258 strains that are frequently found in the United States (20–22). Control of ST11 CRKP in populations and health care networks is often complicated due to rapid evolution. Therefore, it is crucial to track evolutionary events and understand their clinical importance. Previously, we found a KL shift (the previously prevalent KL47 had been replaced by KL64) in the dominant clone BSI-CRKP-ST11 over a 4-year period and that KL64 isolates probably evolved from a KL47 ancestor (17). In this work, we have provided independent evidence of the recombination events that drove evolution of widely distributed KPC-associated ST11 *K. pneumoniae* and resulted in capsule and LPS switching.

Recombination events and replacement of chromosomal regions have been documented in numerous bacteria, contributing to genetic plasticity, and the evolved strains are associated with epidemiological success. *Staphylococcus aureus* ST239, a pandemic methicillin-resistant *S. aureus* (MRSA) sequence type, is responsible for ~90% of the nosocomial infections throughout mainland Asia and South America (23). ST239 is believed to be comprised of large chromosomal regions from two distinctly related lineages, ST8 and ST30 (24). The replacement of approximately 20% of the ST8 genome by an ~550-kb contiguous chromosomal fragment from an ST30 donor strain generated the strains of novel sequence type ST239. In group B *Streptococcus* (GBS), genetic replacement of the CPS region between unrelated sequence types is the common mechanism by which this species alters its surface antigen composition, which contributes to the diversity of GBS and selection of the dominant clone (25).

Based on recent genome-scale studies on *K. pneumoniae*, there have been numerous putative chromosomal recombination events involving the region related to CPS biosynthesis (6, 26). Chen et al. illustrated ST258 clades I and II have distinct CPS regions and the ST258 clade I strains evolved from a clade II strain caused by the replacement of an ~52-kb region encoding the CPS biosynthetic machinery (7). Here, by using phylogenetic and SNP analyses, we discovered that a recombination event had occurred in a homologous ~154-kb region covering CPS and LPS biosynthesis loci between a recipient, ST11-KL47-OL101, and an unrelated donor from CC1764 (except ST30), giving rise to ST11-KL64-O2v1 (recombination I). This ~154-kb contiguous region covers the previously described ~52-kb region in the *K. pneumoniae* genome. Of note, CPS and LPS are not only two important virulence factors of *K. pneumoniae* but also antigenic targets for vaccine and phage-derived exopolysaccharide depolymerase, which is considered a novel approach for the treatment of CRKP infections (27). The diversity of CPS and LPS caused by the chromosomal recombination is posing a challenge to polysaccharide vaccine development. Thus, continued surveillance of CRKP populations remains of importance during vaccine development. Moreover, ST1764 (allelic profile 5-3-1-1-9-4-283), a three-locus variant of hypervirulent ST23 (allelic profile 2-1-1-1-9-4-12), is a regional ST entirely found in China so far and believed to be a hypervirulent clone (28–30). In some of our ongoing work, we have discovered 7 of 219 *K. pneumoniae* isolates, collected from pyogenic liver abscess patients, that belong to ST1764 (data not shown). The uniformity of epidemic geographical location of ST1764 and ST11 strains further confirms the recombination event.

Comandatore and coworkers proposed that the success of an emerging hybrid strain produced by a large recombination event depended on gene content in the exchanged genomic portion (31). Of note, we previously identified a clonal replacement in the CRKP-ST11 population over a 4-year period in a national key hospital, which indicated ST11-KL47, the previously predominant clone, was by degrees replaced by ST11-KL64 (17). It has been shown that the “hybrid” ST11-KL64 appeared to be more “successful” than its precursor, ST11-KL47, in the hospital. We compared the gene compositions in the exchanged genomic regions between ST11-KL64 and ST11-KL47. The results indicated that only genes involved in CPS and LPS biosynthesis and transposase-encoding genes in ISs were distinctly changed following the recombination. All these findings suggested that the obtained KL64 and/or O2 from the exchanged genomic donor portion was likely to be a critical factor conferring a survival advantage on the hybrid ST11-KL64 in response to external selection pressures (e.g., the immune response).

A previous study showed that among *K. pneumoniae* isolates causing pyogenic liver abscess, the O1 serotype was highly predominant, with a prevalence rate of approximately 90% (32). The O1 antigen contains two domains: D-galactan I and D-galactan II, while the O2 antigen is composed only of D-galactan I and is not recognized by anti-D-galactan II antiserum. Therefore, the D-galactan II structure is the difference between the LPS composition of *K. pneumoniae* serotypes O1 and O2. Previous studies indicated that the presence of D-galactan II was required for the resistance of *K. pneumoniae* to serum killing, and the D-galactan II-deficient mutant was less virulent than the parental strain in mouse intraperitoneal infection models (19, 33). Thus, D-galactan II is a crucial LPS composition in *K. pneumoniae* which enables the bacterium to exhibit a higher virulence potential. In this study, we found that

although the ~154-kb region covering OL in CC1764 (except ST30) was recombined to the chromosome of ST11-KL47, the OL of the recombination offspring could only directly transform from OL101 into O2 (as with ST11-KL64) but not O1 (as with CC1764-KL64) due to the absence of D-galactan II-encoding *wbbY* and *wbbZ* genes in the ST11 chromosome. All these findings implied that the ST11-KL64 clone might not exhibit a similar virulence trait as ST1764 strains.

Nevertheless, we also found genetic diversity within the ST11-KL64 group. A ST11-KL64 isolate HB25-1 (18), identified in China, was found to have a homologous ~485-kb recombination region derived from ST147-KL64 strains (recombination II). Subsequently, we discovered that this recombination II event also produced a specific ST11-KL64 subclone which was significantly more infrequent than the predominant ST11-KL64 subclone produced by recombination I in China, despite showing the same KL64 and O2v1 in these two subclones. ST147-KL64 is a highly prevalent carbapenem-resistant clone in Asian countries such as India, Pakistan, and Thailand. The similarity of geographical distribution of ST147-KL64 and ST11-KL64 also provides support to the assumed recombination II event. However, the reason why the ST11-KL64 subclone produced by recombination II was not as prevalent as the subclone produced by recombination I remains to be elucidated.

There are certain limitations associated with our study. First, we could not find the explicit breakpoint of the recombination event, because there was no apparent difference existing in the large nucleotide fragments adjacent to the terminal of the region which contained a large number of SNPs. Therefore, we provided an approximate recombination region based on the SNP analysis and comparative genomic analysis. Second, a mechanism to explain the large-scale exchange of sequence observed in *K. pneumoniae* was not confirmed in the study. It was proposed that conjugation and transformation were efficient mechanisms of large-scale transfer of sequence in *Streptococcus pneumoniae* and *Streptococcus agalactiae*, but how it has occurred in Gram-negative bacteria remains a difficult issue (7, 34, 35). Third, we could not collect the ST11-KL64 strain produced by recombination II during the study period. Therefore, we were unable to explore the differences in microbiological characteristics between these two subclones.

Taken together, our findings underlined the role of recombination in the evolution of clinical strains of widely distributed KPC-associated ST11 *K. pneumoniae* in China. Understanding the evolutionary genomic history of the microbe is a critical approach to gain new insight into its success as a pathogen, as we have done here. This information will assist our efforts to develop diagnostics, therapeutics, and vaccines against infections caused by CRKP. Notably, CPS and LPS of *K. pneumoniae* contribute to the evasion of the innate host defense and thus are considered crucial for survival in the host. The switching of CPS and LPS in the newly emerging hybrid ST11 descendant altered antigens on the surface of the bacteria and made the production of a polysaccharide-based vaccine (such as phage-derived exopolysaccharide depolymerase) against *K. pneumoniae* infection more complicated, posing a substantial threat to health care networks.

MATERIALS AND METHODS

Isolates and whole-genome sequencing. We previously collected 150 genomes of ST11 carbapenem-resistant *K. pneumoniae* (KL47, $n = 71$; KL64, $n = 77$; KL105, $n = 1$; KL103, $n = 1$) from patients with a blood culture positive for *K. pneumoniae* and a clinical course consistent with bloodstream infection during 2012 to 2017 in the First Affiliated Hospital of Zhejiang University in China (17). One ST1764-KL64 carbapenem-susceptible *K. pneumoniae* strain (KP41076) was identified in the hospital and sequenced by using an Illumina HiSeq2,500 instrument (Illumina) with 2×125 -bp paired-end libraries. We performed *de novo* assembly of the short-read data by using Unicycler 0.4.0 (36). The ST of the isolate was acquired by MLST (<https://cge.food.dtu.dk/services/MLST/>). We annotated the assembly by using the RAST server (<https://rast.nmpdr.org>) (37). The 151 genomes of the above strains and 34 downloaded genomes of *K. pneumoniae* strains isolated from China, which belonged to ST11 (KL47, $n = 13$; KL64, $n = 21$), were added to the genome group for genomic analysis (see Table S1 in the supplemental material). Moreover, we analyzed 22,600 *K. pneumoniae* genomes among global isolates from PATRIC (38) for selecting the STs with the KL64 genetic locus. We found KL64 could be detected in the genomes of ST1764 (allelic profile 5-3-1-1-9-4-283), ST3685 (5-3-1-232-9-4-283), ST1764-1LV (5-3-1-1-9-4-683), ST30 (5-3-1-1-9-4-26), ST505 (7-1-5-1-1-1-84), ST147 (3-4-6-1-7-4-38), and ST11 (3-3-1-1-1-1-4). Using the stringent criteria, where all members share identical alleles at 6 of the 7 loci with at least one other member of the group, we defined ST1764, ST3685, ST1764-1LV, and ST30 as a clonal complex, CC1764. Ten

genomes of CC1764 (ST1764, $n = 4$; ST3685, $n = 1$, ST1764-1LV, $n = 1$; ST30, $n = 4$), 8 ST505, and 45 ST147 were downloaded and used for genomic analysis. The detailed data of all 248 isolates are provided in Table S1.

Furthermore, 1,150 genomes of ST11 *K. pneumoniae* strains isolated from China were downloaded from PATRIC and screened for selecting ST11-KL64 genomes. As a result, 621 Chinese ST11-KL64 *K. pneumoniae* genomes were detected and used in the phylogenetic analysis to investigate the prevalence of the specific ST11-KL64 subclone. The detailed data of these 621 isolates are shown in Table S2.

Genomic analysis. Antimicrobial resistance genes and virulence genes were detected using Abricate (<https://github.com/tseemann/abricate>) with gene databases ResFinder (39) and VFDB (40), respectively. We identified K-type, O-type, and multilocus sequence typing by using Kleborate and Kaptive (41, 42). Insertion sequences were identified using the ISSaga (http://issaga.biotoul.fr/issaga_index.php) and the IS Finder database (<https://www-is.biotoul.fr/>) (43, 44). The complete nucleotide sequence of an ICE was identified by ICEfinder (<https://db-mml.sjtu.edu.cn/ICEfinder/>) with manual modification (45). Prophages were detected by using VRprofile (<https://tool2-mml.sjtu.edu.cn/VRprofile/>). Multiple genome sequence alignments and comparison analysis were then performed with Mauve (46). Sequence comparisons were performed using BLASTn v2.4.0 (47) and visualized using Easyfig v2.2.3 (48).

Phylogenetic analysis. Snippy v4.6.0 (<https://github.com/tseemann/snippy>) was used to perform reference-based mapping and identify SNPs in the core genome, with the KP47434 genome (GenBank accession no. [QURI0000000](https://genbank.ncbi.nlm.nih.gov/GenBank/FASTA/CP039524)) or the HB25-1 genome (GenBank accession no. [CP039524](https://genbank.ncbi.nlm.nih.gov/GenBank/FASTA/CP039524), [CP039525](https://genbank.ncbi.nlm.nih.gov/GenBank/FASTA/CP039525), [CP039526](https://genbank.ncbi.nlm.nih.gov/GenBank/FASTA/CP039526)) used as references. An SNP alignment was then extracted from the core genome alignment using snp-sites v2.5.1 (49), and a phylogenetic tree was constructed using the core SNPs with Fasttree 2.1 with default parameters (50). The phylogenetic tree was visualized using the Interactive Tree of Life software (iTOL v.3) software (51). The minimum spanning trees (MSTs) based on SNP data were generated using PHYLOVIZ v2.0 (52).

Data availability. The whole-genome sequences of all collected *K. pneumoniae* isolates have been deposited in the GenBank database under BioProject numbers [PRJNA588307](https://ncbi.nlm.nih.gov/bioproject/PRJNA588307) and [PRJNA798548](https://ncbi.nlm.nih.gov/bioproject/PRJNA798548). The data that support the findings of this study are available and accession numbers of the genome sequences are listed in Table S1 and Table S2.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

SUPPLEMENTAL FILE 1, XLSX file, 0.04 MB.

SUPPLEMENTAL FILE 2, XLSX file, 0.1 MB.

SUPPLEMENTAL FILE 3, XLSX file, 0.02 MB.

SUPPLEMENTAL FILE 4, XLSX file, 0.04 MB.

SUPPLEMENTAL FILE 5, PDF file, 1.4 MB.

ACKNOWLEDGMENTS

We thank Yunbo Chen, Jinru Ji, Chaoqun Ying, Zhiying Liu, Hao Xu, Hao Fu, Mengmeng Lu, Luying Xiong, and Jing Ye for providing assistance during sample collection and bacterial identification, and Tingting Xiao and Qixia Luo for providing assistance with data analysis.

We 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.

This study was supported by grants from the National Key Research and Development Program of China (2021YFC2300300), the National Natural Science Foundation of China (81971984), and Research Project of Jinan Microecological Biomedicine Shandong Laboratory (JNL2022027C).

REFERENCES

- Guh AY, Bulens SN, Mu Y, Jacob JT, Reno J, Scott J, Wilson LE, Vaeth E, Lynfield R, Shaw KM, Vagnone PM, Bamberg WM, Janelle SJ, Dumyati G, Concannon C, Beldavs Z, Cunningham M, Cassidy PM, Phipps EC, Kenslow N, Travis T, Lonsway D, Rasheed JK, Limbago BM, Kallen AJ. 2015. Epidemiology of carbapenem-resistant Enterobacteriaceae in 7 US communities, 2012-2013. *JAMA* 314:1479–1487. <https://doi.org/10.1001/jama.2015.12480>.
- Grundmann H, Glasner C, Albiger B, Aanensen DM, Tomlinson CT, Andrasevic AT, Canton R, Carmeli Y, Friedrich AW, Giske CG, Glupczynski Y, Gniadkowski M, Livermore DM, Nordmann P, Poirel L, Rossolini GM, Seifert H, Vatopoulos A, Walsh T, Woodford N, Monnet DL, European Survey of Carbapenemase-Producing Enterobacteriaceae Working Group. 2017. Occurrence of carbapenemase-producing *Klebsiella pneumoniae* and *Escherichia coli* in the European survey of carbapenemase-producing Enterobacteriaceae (EuSCAPE): a prospective, multinational study. *Lancet Infect Dis* 17:153–163. [https://doi.org/10.1016/S1473-3099\(16\)30257-2](https://doi.org/10.1016/S1473-3099(16)30257-2).
- Zhang R, Liu L, Zhou H, Chan EW, Li J, Fang Y, Li Y, Liao K, Chen S. 2017. Nationwide surveillance of clinical carbapenem-resistant Enterobacteriaceae (CRE) strains in China. *EBioMedicine* 19:98–106. <https://doi.org/10.1016/j.ebiom.2017.04.032>.
- Munoz-Price LS, Poirel L, Bonomo RA, Schwaber MJ, Daikos GL, Cormican M, Cornaglia G, Garau J, Gniadkowski M, Hayden MK, Kumarasamy K, Livermore DM, Maya JJ, Nordmann P, Patel JB, Paterson DL, Pitout J, Villegas MV, Wang H, Woodford N, Quinn JP. 2013. Clinical epidemiology of the global expansion of *Klebsiella pneumoniae* carbapenemases. *Lancet Infect Dis* 13:785–796. [https://doi.org/10.1016/S1473-3099\(13\)70190-7](https://doi.org/10.1016/S1473-3099(13)70190-7).
- Pitout JD, Nordmann P, Poirel L. 2015. Carbapenemase-producing *Klebsiella pneumoniae*, a key pathogen set for global nosocomial dominance. *Antimicrob Agents Chemother* 59:5873–5884. <https://doi.org/10.1128/AAC.01019-15>.
- Deleo FR, Chen L, Porcella SF, Martens CA, Kobayashi SD, Porter AR, Chavda KD, Jacobs MR, Mathema B, Olsen RJ, Bonomo RA, Musser JM, Kreiswirth BN. 2014. Molecular dissection of the evolution of carbapenem-resistant multilocus sequence type 258 *Klebsiella pneumoniae*. *Proc Natl Acad Sci U S A* 111:4988–4993. <https://doi.org/10.1073/pnas.1321364111>.

7. Chen L, Mathema B, Pitout JD, DeLeo FR, Kreiswirth BN. 2014. Epidemic *Klebsiella pneumoniae* ST258 is a hybrid strain. *mBio* 5:e01355-14. <https://doi.org/10.1128/mBio.01355-14>.
8. Croucher NJ, Klugman KP. 2014. The emergence of bacterial “hopeful monsters.” *mBio* 5:e01550-14. <https://doi.org/10.1128/mBio.01550-14>.
9. Wyres KL, Gorrie C, Edwards DJ, Wertheim HF, Hsu LY, Van Kinh N, Zadoks R, Baker S, Holt KE. 2015. Extensive capsule locus variation and large-scale genomic recombination within the *Klebsiella pneumoniae* clonal group 258. *Genome Biol Evol* 7:1267–1279. <https://doi.org/10.1093/gbe/evv062>.
10. Wang Q, Wang X, Wang J, Ouyang P, Jin C, Wang R, Zhang Y, Jin L, Chen H, Wang Z, Zhang F, Cao B, Xie L, Liao K, Gu B, Yang C, Liu Z, Ma X, Jin L, Zhang X, Man S, Li W, Pei F, Xu X, Jin Y, Ji P, Wang H. 2018. Phenotypic and genotypic characterization of carbapenem-resistant Enterobacteriaceae: data from a longitudinal large-scale CRE study in China (2012–2016). *Clin Infect Dis* 67:S196–S205. <https://doi.org/10.1093/cid/ciy660>.
11. Gu D, Dong N, Zheng Z, Lin D, Huang M, Wang L, Chan EW, Shu L, Yu J, Zhang R, Chen S. 2018. A fatal outbreak of ST11 carbapenem-resistant hypervirulent *Klebsiella pneumoniae* in a Chinese hospital: a molecular epidemiological study. *Lancet Infect Dis* 18:37–46. [https://doi.org/10.1016/S1473-3099\(17\)30489-9](https://doi.org/10.1016/S1473-3099(17)30489-9).
12. Zhang X, Ouyang J, He W, Zeng T, Liu B, Jiang H, Zhang Y, Zhou L, Zhou H, Liu Z, Liu L. 2020. Co-occurrence of rapid gene gain and loss in an inter-hospital outbreak of carbapenem-resistant hypervirulent ST11-K64 *Klebsiella pneumoniae*. *Front Microbiol* 11:579618. <https://doi.org/10.3389/fmicb.2020.579618>.
13. Kong Y, Sun Q, Chen H, Draz MS, Xie X, Zhang J, Ruan Z. 2021. Transmission dynamics of carbapenem-resistant *Klebsiella pneumoniae* sequence type 11 strains carrying capsular loci KL64 and rmpA/rmpA2 genes. *Front Microbiol* 12:736896. <https://doi.org/10.3389/fmicb.2021.736896>.
14. Zhang Y, Jin L, Ouyang P, Wang Q, Wang R, Wang J, Gao H, Wang X, Wang H, China Carbapenem-Resistant Enterobacteriaceae (CRE) Network. 2020. Evolution of hypervirulence in carbapenem-resistant *Klebsiella pneumoniae* in China: a multicentre, molecular epidemiological analysis. *J Antimicrob Chemother* 75:327–336. <https://doi.org/10.1093/jac/dkz446>.
15. Wu X, Shi Q, Shen S, Huang C, Wu H. 2021. Clinical and bacterial characteristics of *Klebsiella pneumoniae* affecting 30-day mortality in patients with bloodstream infection. *Front Cell Infect Microbiol* 11:688989. <https://doi.org/10.3389/fcimb.2021.688989>.
16. Zhao J, Liu C, Liu Y, Zhang Y, Xiong Z, Fan Y, Zou X, Lu B, Cao B. 2020. Genomic characteristics of clinically important ST11 *Klebsiella pneumoniae* strains worldwide. *J Glob Antimicrob Resist* 22:519–526. <https://doi.org/10.1016/j.jgar.2020.03.023>.
17. Zhou K, Xiao T, David S, Wang Q, Zhou Y, Guo L, Aanensen D, Holt KE, Thomson NR, Grundmann H, Shen P, Xiao Y. 2020. Novel subclone of carbapenem-resistant *Klebsiella pneumoniae* sequence type 11 with enhanced virulence and transmissibility, China. *Emerg Infect Dis* 26:289–297. <https://doi.org/10.3201/eid2602.190594>.
18. Xie M, Dong N, Chen K, Yang X, Ye L, Chan EW, Zhang R, Chen S. 2020. A hybrid plasmid formed by recombination of a virulence plasmid and a resistance plasmid in *Klebsiella pneumoniae*. *J Glob Antimicrob Resist* 23:466–470. <https://doi.org/10.1016/j.jgar.2020.10.018>.
19. Hsieh PF, Wu MC, Yang FL, Chen CT, Lou TC, Chen YY, Wu SH, Sheu JC, Wang JT. 2014. D-Galactan II is an immunodominant antigen in O1 lipopolysaccharide and affects virulence in *Klebsiella pneumoniae*: implication in vaccine design. *Front Microbiol* 5:608. <https://doi.org/10.3389/fmicb.2014.00608>.
20. Liao W, Liu Y, Zhang W. 2020. Virulence evolution, molecular mechanisms of resistance and prevalence of ST11 carbapenem-resistant *Klebsiella pneumoniae* in China: a review over the last 10 years. *J Glob Antimicrob Resist* 23:174–180. <https://doi.org/10.1016/j.jgar.2020.09.004>.
21. Chen L, Mathema B, Chavda KD, DeLeo FR, Bonomo RA, Kreiswirth BN. 2014. Carbapenemase-producing *Klebsiella pneumoniae*: molecular and genetic decoding. *Trends Microbiol* 22:686–696. <https://doi.org/10.1016/j.tim.2014.09.003>.
22. Zhang X, Wang W, Yu H, Wang M, Zhang H, Lv J, Tang YW, Kreiswirth BN, Du H, Chen L. 2019. New Delhi metallo- β -lactamase 5-producing *Klebsiella pneumoniae* sequence type 258, southwest China, 2017. *Emerg Infect Dis* 25:1209–1213. <https://doi.org/10.3201/eid2506.181939>.
23. Harris SR, Feil EJ, Holden MT, Quail MA, Nickerson EK, Chantratita N, Gardete S, Tavares A, Day N, Lindsay JA, Edgeworth JD, de Lencastre H, Parkhill J, Peacock SJ, Bentley SD. 2010. Evolution of MRSA during hospital transmission and intercontinental spread. *Science* 327:469–474. <https://doi.org/10.1126/science.1182395>.
24. Robinson DA, Monk AB, Cooper JE, Feil EJ, Enright MC. 2005. Evolutionary genetics of the accessory gene regulator (*agr*) locus in *Staphylococcus aureus*. *J Bacteriol* 187:8312–8321. <https://doi.org/10.1128/JB.187.24.8312-8321.2005>.
25. Bellais S, Six A, Fouet A, Longo M, Dmytruk N, Glaser P, Trieu-Cuot P, Poyart C. 2012. Capsular switching in group B *Streptococcus* CC17 hyper-virulent clone: a future challenge for polysaccharide vaccine development. *J Infect Dis* 206:1745–1752. <https://doi.org/10.1093/infdis/jis605>.
26. Dong N, Zhang R, Liu L, Li R, Lin D, Chan EW, Chen S. 2018. Genome analysis of clinical multilocus sequence type 11 *Klebsiella pneumoniae* from China. *Microb Genom* 4:e000149. <https://doi.org/10.1099/mgen.0.000149>.
27. Dunstan RA, Bamert RS, Belousoff MJ, Short FL, Barlow CK, Pickard DJ, Wilksch JJ, Schittenhelm RB, Strugnelli RA, Dougan G, Lithgow T. 2021. Mechanistic insights into the capsule-targeting depolymerase from a *Klebsiella pneumoniae* bacteriophage. *Microbiol Spectr* 9:e0102321. <https://doi.org/10.1128/Spectrum.01023-21>.
28. Liu Z, Gu Y, Li X, Liu Y, Ye Y, Guan S, Li J. 2019. Identification and characterization of NDM-1-producing hypervirulent (hypermucoviscous) *Klebsiella pneumoniae* in China. *Ann Lab Med* 39:167–175. <https://doi.org/10.3343/alm.2019.39.2.167>.
29. Li Q, Zhu J, Kang J, Song Y, Yin D, Guo Q, Song J, Zhang Y, Wang S, Duan J. 2020. Emergence of NDM-5-producing carbapenem-resistant *Klebsiella pneumoniae* and SIM-producing hypervirulent *Klebsiella pneumoniae* isolated from aseptic body fluid in a large tertiary hospital, 2017–2018: genetic traits of blaNDM-like and blaSIM-like genes as determined by NGS. *Infect Drug Resist* 13:3075–3089. <https://doi.org/10.2147/IDR.S261117>.
30. Zhao B, Hu R, Gong L, Wang X, Zhu Y, Wu G. 2021. Pyogenic liver abscess and endogenous endophthalmitis due to K64-ST1764 hypervirulent *Klebsiella pneumoniae*: a case report. *Infect Drug Resist* 14:71–77. <https://doi.org/10.2147/IDR.S289088>.
31. Comandatore F, Sasser D, Bayliss SC, Scaltriti E, Gaiarsa S, Cao X, Gales AC, Saito R, Pongolini S, Brisse S, Feil EJ, Bandi C. 2019. Gene composition as a potential barrier to large recombinations in the bacterial pathogen *Klebsiella pneumoniae*. *Genome Biol Evol* 11:3240–3251. <https://doi.org/10.1093/gbe/evz236>.
32. Hsieh PF, Lin TL, Yang FL, Wu MC, Pan YJ, Wu SH, Wang JT. 2012. Lipopolysaccharide O1 antigen contributes to the virulence in *Klebsiella pneumoniae* causing pyogenic liver abscess. *PLoS One* 7:e33155. <https://doi.org/10.1371/journal.pone.0033155>.
33. McCallum KL, Schoenhals G, Laakso D, Clarke B, Whitfield C. 1989. A high-molecular-weight fraction of smooth lipopolysaccharide in *Klebsiella* serotype O1:K20 contains a unique O-antigen epitope and determines resistance to nonspecific serum killing. *Infect Immun* 57:3816–3822. <https://doi.org/10.1128/iai.57.12.3816-3822.1989>.
34. Golubchik T, Brueggemann AB, Street T, Gertz RE, Jr, Spencer CC, Ho T, Giannoulata E, Link-Gelles R, Harding RM, Beall B, Peto TE, Moore MR, Donnelly P, Crook DW, Bowden R. 2012. Pneumococcal genome sequencing tracks a vaccine escape variant formed through a multi-fragment recombination event. *Nat Genet* 44:352–355. <https://doi.org/10.1038/ng.1072>.
35. Brochet M, Rusniok C, Couvé E, Dramsi S, Poyart C, Trieu-Cuot P, Kunst F, Glaser P. 2008. Shaping a bacterial genome by large chromosomal replacements: the evolutionary history of *Streptococcus agalactiae*. *Proc Natl Acad Sci U S A* 105:15961–15966. <https://doi.org/10.1073/pnas.0803654105>.
36. Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput Biol* 13:e1005595. <https://doi.org/10.1371/journal.pcbi.1005595>.
37. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST Server: rapid annotations using subsystems technology. *BMC Genomics* 9:75. <https://doi.org/10.1186/1471-2164-9-75>.
38. Davis JJ, Wattam AR, Aziz RK, Brettin T, Butler R, Butler RM, Chlenski P, Conrad N, Dickerman A, Dietrich EM, Gabbard JL, Gerdes S, Guard A, Kenyon RW, Machi D, Mao C, Murphy-Olson D, Nguyen M, Nordberg EK, Olsen GJ, Olson RD, Overbeek JC, Overbeek R, Parrello B, Pusch GD, Shukla M, Thomas C, VanOeffelen M, Vonstein V, Warren AS, Xia F, Xie D, Yoo H, Stevens R. 2020. The PATRIC Bioinformatics Resource Center: expanding data and analysis capabilities. *Nucleic Acids Res* 48:D606–D612. <https://doi.org/10.1093/nar/gkz943>.
39. Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O, Aarestrup FM, Larsen MV. 2012. Identification of acquired antimicrobial resistance genes. *J Antimicrob Chemother* 67:2640–2644. <https://doi.org/10.1093/jac/dks261>.
40. Chen L, Zheng D, Liu B, Yang J, Jin Q. 2016. VFDB 2016: hierarchical and refined dataset for big data analysis—10 years on. *Nucleic Acids Res* 44:D694–D697. <https://doi.org/10.1093/nar/gkv1239>.
41. Lam MMC, Wick RR, Watts SC, Cerdeira LT, Wyres KL, Holt KE. 2021. A genomic surveillance framework and genotyping tool for *Klebsiella pneumoniae*

- and its related species complex. *Nat Commun* 12:4188. <https://doi.org/10.1038/s41467-021-24448-3>.
42. Wyres KL, Wick RR, Gorrie C, Jenney A, Follador R, Thomson NR, Holt KE. 2016. Identification of *Klebsiella* capsule synthesis loci from whole genome data. *Microb Genom* 2:e000102. <https://doi.org/10.1099/mgen.0.000102>.
 43. Siguier P, Perochon J, Lestrade L, Mahillon J, Chandler M. 2006. ISfinder: the reference centre for bacterial insertion sequences. *Nucleic Acids Res* 34:D32–D36. <https://doi.org/10.1093/nar/gkj014>.
 44. Varani AM, Siguier P, Gourbeyre E, Charneau V, Chandler M. 2011. Issaga is an ensemble of web-based methods for high throughput identification and semi-automatic annotation of insertion sequences in prokaryotic genomes. *Genome Biol* 12:R30. <https://doi.org/10.1186/gb-2011-12-3-r30>.
 45. Bi D, Xu Z, Harrison EM, Tai C, Wei Y, He X, Jia S, Deng Z, Rajakumar K, Ou HY. 2012. ICEberg: a web-based resource for integrative and conjugative elements found in Bacteria. *Nucleic Acids Res* 40:D621–D626. <https://doi.org/10.1093/nar/gkr846>.
 46. Darling AE, Mau B, Perna NT. 2010. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS One* 5:e11147. <https://doi.org/10.1371/journal.pone.0011147>.
 47. Zhang Z, Schwartz S, Wagner L, Miller W. 2000. A greedy algorithm for aligning DNA sequences. *J Comput Biol* 7:203–214. <https://doi.org/10.1089/10665270050081478>.
 48. Sullivan MJ, Petty NK, Beatson SA. 2011. Easyfig: a genome comparison visualizer. *Bioinformatics* 27:1009–1010. <https://doi.org/10.1093/bioinformatics/btr039>.
 49. Page AJ, Taylor B, Delaney AJ, Soares J, Seemann T, Keane JA, Harris SR. 2016. SNP-sites: rapid efficient extraction of SNPs from multi-FASTA alignments. *Microb Genom* 2:e000056. <https://doi.org/10.1099/mgen.0.000056>.
 50. Price MN, Dehal PS, Arkin AP. 2009. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Mol Biol Evol* 26:1641–1650. <https://doi.org/10.1093/molbev/msp077>.
 51. Letunic I, Bork P. 2016. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res* 44:W242–W245. <https://doi.org/10.1093/nar/gkw290>.
 52. Nascimento M, Sousa A, Ramirez M, Francisco AP, Carrico JA, Vaz C. 2017. PHYLOViZ 2.0: providing scalable data integration and visualization for multiple phylogenetic inference methods. *Bioinformatics* 33:128–129. <https://doi.org/10.1093/bioinformatics/btw582>.