**Prevalence and Distribution Characteristics of \textit{bla}_KPC-2 and \textit{bla}_NDM-1** Genes in \textit{Klebsiella pneumoniae}

**Background:** Carbapenem-resistant \textit{Klebsiella pneumoniae} infections have caused major concern and posed a global threat to public health. As \textit{bla}_KPC-2 and \textit{bla}_NDM-1 genes are the most widely reported carbapenem resistant genes in \textit{K. pneumoniae}, it is crucial to study the prevalence and geographical distribution of these two genes for further understanding of their transmission mode and mechanism.

**Purpose:** Here, we investigated the prevalence and distribution of \textit{bla}_KPC-2 and \textit{bla}_NDM-1 genes in carbapenem-resistant \textit{K. pneumoniae} strains from a tertiary hospital and from 1579 genomes available in the NCBI database, and further analyzed the possible core structure of \textit{bla}_KPC-2 or \textit{bla}_NDM-1 genes among global genome data.

**Materials and Methods:** \textit{K. pneumoniae} strains from a tertiary hospital in China during 2013–2018 were collected and their antimicrobial susceptibility testing for 28 antibiotics was determined. Whole-genome sequencing of carbapenem-resistant \textit{K. pneumoniae} strains was used to investigate the genetic characterization. The phylogenetic relationships of these strains were investigated through pan-genome analysis. The epidemiology and distribution of \textit{bla}_KPC-2 and \textit{bla}_NDM-1 genes in \textit{K. pneumoniae} based on 1579 global genomes and carbapenem-resistant \textit{K. pneumoniae} strains from hospital were analyzed using bioinformatics. The possible core structure carrying \textit{bla}_KPC-2 or \textit{bla}_NDM-1 genes was investigated among global data.

**Results:** A total of 19 carbapenem-resistant \textit{K. pneumoniae} were isolated in a tertiary hospital. All isolates had a multi-resistant pattern and eight kinds of resistance genes. The phylogenetic analysis showed all isolates in the hospital were dominated by two lineages composed of ST11 and ST25, respectively. ST11 and ST25 were the major ST type carrying \textit{bla}_KPC-2 and \textit{bla}_NDM-1 genes, respectively. Among 1579 global genomes data, 147 known ST types (1195 genomes) have been identified, while ST258 (23.6%) and ST11 (22.1%) were the globally prevalent clones among the known ST types. Genetic environment analysis showed that the ISKpn7-dnaA/ISKpn27-\textit{bla}-ISKpn6 and NDM-1-blc-trpf-nagA may be the core structure in the horizontal transfer of \textit{bla}_KPC-2 and \textit{bla}_NDM-1, respectively. In addition, DNA transferase (\textit{hin}) may be involved in the horizontal transfer or the expression of \textit{bla}_NDM-1.

**Conclusion:** There was clonal transmission of carbapenem-resistant \textit{K. pneumoniae} in the tertiary hospital in China. The prevalence and distribution of \textit{bla}_KPC-2 and \textit{bla}_NDM-1 varied by countries and were driven by different transposons carrying the core structure. This study shed light on the genetic environment of \textit{bla}_KPC-2 and \textit{bla}_NDM-1 and offered basic information about the mechanism of carbapenem-resistant \textit{K. pneumoniae} dissemination.

**Keywords:** \textit{Klebsiella pneumoniae}, \textit{bla}_KPC-2, *\textit{bla}_NDM-1*, bioinformatics

---

**Introduction**

\textit{Klebsiella pneumoniae}, as a member of ESKAPE pathogens (\textit{Enterococcus faecium}, \textit{Staphylococcus aureus}, \textit{Klebsiella pneumoniae}, \textit{Acinetobacter baumannii},
Enterobacter species is the most common factor of nosocomial infections. Infections caused by *K. pneumoniae* are associated with a high risk of mortality and increased economic costs. In recent years, *K. pneumoniae* has been considered as a growing global threat due to its high level of resistance, particularly to those last resort antibiotics, such as carbapenems. In 2017, carbapenem-resistant *K. pneumoniae* (CRKP) has been listed in the critical priority tier of pathogens and the highest priority in new antibiotic development by WHO.

Carbapenem resistance in *K. pneumoniae* involves multiple mechanisms, including the production of carbapenemases, alterations in outer membrane permeability, and the upregulation of efflux systems. The latter two mechanisms are often combined with other types of β-lactamases (eg, AmpC, ESBLs). Carbapenemases are carbapenem-hydrolyzing β-lactamases and the most frequently reported in *K. pneumoniae* are ampler molecular class A (KPC), class B (VIM, IMP, NDM), and class D (OXA-48-like) types. The KPC-type β-lactamases were the most prevalent mechanism for carbapenem resistance in *K. pneumoniae*. Even though there are more than 20 different KPC variants reported, KPC-2 and -3 remain the most commonly identified variants. The class B β-lactamases identified in *K. pneumoniae* have also been found in other *Enterobacteriaceae* worldwide. The majority of NDM-1-producing *K. pneumoniae* strains also carried a diversity of other resistance genes but mostly remained susceptible to agents such as colistin, fosfomycin, and tigecycline. These different carbapenemase genes circulating within *K. pneumoniae* are often associated with mobile structures, including insertion sequences (IS), plasmids, and transposons, which caused serious challenges to clinic treatment.

Nosocomial infection outbreaks caused by CRKP have been rapidly emerging in many countries and they are more often found in critically ill patients dwelling in intensive care units. Recently, genotypic characterization has shown that several high-risk clones or ST types were reported for the global distribution of CRKP. The pandemic of KPC-producing *K. pneumoniae* is primarily driven by clonal complex 258 (CC258). CC258 consists of the predominant clone ST258 and its single-locus variants (ST11, ST340, and ST512). A previous study indicated that ST11 is an emerging high-risk clone in KPC-producing *K. pneumoniae*, while NDM-1-producing *K. pneumoniae* reported widely in ST14, ST25, and ST340. Hence, it is crucial to carry out molecular mechanism research and epidemiological investigation on carbapenem resistance genes among *K. pneumoniae* in the world.

In this study, we conducted a 6-year-long longitudinal study of a tertiary care hospital in China, to study the genetic characterization of carbapenem-resistant *K. pneumoniae* using Whole-genome sequencing (WGS) and bioinformatic analyses. In addition, we combined the NCBI global *K. pneumoniae* genomes data to demonstrate the distribution and spread characteristics of *blaKPC-2* and *blanDM-1* around the world. The study aims to provide a theoretical basis for the rational use of clinical antibiotics and the reduction of the outbreak of nosocomial infection.

### Materials and Methods

#### Strains Collection and Antimicrobial Susceptibility Testing

*K. pneumoniae* strains were collected from patients from Shunde hospital in China between 2013 and 2018. These clinical samples were routinely obtained in the clinical microbiology laboratory. Antimicrobial susceptibility testing was performed by a BD Phoenix™ 100 Automated Identification and Susceptibility Testing system (BD, USA) according to CLSI and EUCAST guidelines. The minimum inhibitory concentrations (MICs) were determined for 28 different antibiotics; amoxicillin, ampicillin, ampicillin/sulbactam, cefazolin, cefepime, cefotaxime, cefoperazone/sulbactam, cefazidime, cefuroxime, aztreonam, piperacillin, piperacillin/tazobactam, imipenem, meropenem, chloramphenicol, florfenicol, ciprofloxacin, levofloxacin, fosfomycin, amikacin, gentamicin, neomycin, streptomycin, polymyxin B, colistin, tigecycline, tetracycline, and trimethoprim/sulfamethoxazole.

For tigecycline, the EUCAST clinical MIC breakpoint for *Enterobacteriaceae* was used as the breakpoint for *K. pneumoniae* (a MIC > 2.0 mg/L is resistant). The reference strains *Escherichia coli* ATCC 25,922 was used as control strain.

#### Whole-Genome Sequencing (WGS)

Bacterial DNA of carbapenem-resistant *K. pneumoniae* (CRKP) strains were extracted using a HiPure Bacterial DNA Kit (MAGEN, China, [http://www.magentec.com.cn/](http://www.magentec.com.cn/)). The libraries were created using the VAHTSTM Universal DNA Library Prep kit for Illumina. The genomes were sequenced using the Illumina Hiseq 2500 system to obtain 2×150 bp reads. Processed reads after quality trimming were...
de novo assembled into contigs with a CLC Genomics Workbench 10.1 (CLC Bio, Aarhus, Denmark) and the genomes were annotated by the NCBI Prokaryotic Annotation Pipeline (PGAP).

Global Data from NCBI
As of November 2019, all K. pneumoniae genomes were downloaded from the NCBI database (https://www.ncbi.nlm.nih.gov/genome/?term=Klebsiella+pneumoniae). Then we used conventional python scripts to screen samples with complete information (including country, host, assembly level) for further analysis. A total of 1579 human samples from 44 countries, including 247 fully assembled (chromosome) level samples and 1332 scaffold level samples, were gathered to subsequent analysis (Supplementary Appendix 1).

Taxonomic Assignment
FastANI v1.1 (https://github.com/ParBLiSS/FastANI) with the core algorithm of BLAST-based ANI (ANIb) solver was used to identify the species of all isolates.\(^\text{19}\) Genome ASM24018v2 (https://www.ncbi.nlm.nih.gov/search/all/?term=%20ASM24018v2%20) was used as K. pneumoniae reference genome for species identification. Species were determined if the genome in question had >95% ANIb compared with reference genome.\(^\text{20}\) Software MLST v2.16.1 was used to perform multi-site sequence typing (MLST) of genomes.\(^\text{21}\) The ResFinder BLAST identification program (https://cge.cbs.dtu.dk/services/ResFinder/) and Isfinder (https://www-is.biotoul.fr/index.php) were used to identify acquired antibiotic resistance genes (ARGs) and mobile elements, respectively.\(^\text{22}\) The VFDB (http://www.mgc.ac.cn/VFs/) virulence database (setB) was conducted for bacterial virulence genes analysis.

Pan-Genome Analysis
Prokka v1.14 was used to produce gff file format for the contigs of K. pneumoniae genomes we collected.\(^\text{23}\) After that, core genome alignment was constructed with Roary v3.8.0 and PRANK v1.0.\(^\text{24}\) Core_genome_alignment.aln, the output file of Roary pipeline, was converted to fastGEAR to identify lineages by hierBAPS.\(^\text{25}\)

\textbf{bla}_{KPC-2} and \textbf{bla}_{NDM-1} Loci Annotation and Comparison
Regular Python scripts and Easyfigv2.2.3 (http://easyfig.sourceforge.net/) were used to extract gene sequences. Seventy-six K. pneumoniae genomes carrying \textbf{bla}_{KPC-2} and 30 K. pneumoniae genomes carrying \textbf{bla}_{NDM-1} with complete chromosomal genomic data were selected for \textbf{bla}_{KPC-2} and \textbf{bla}_{NDM-1} gene environment analysis, respectively. Through python script, we extracted the upstream and downstream 2–6 kb sequences of the target gene for blast alignment and annotation.

Visualization of Data
ArcGIS 10.3, R package ggplot2 (http://had.co.nz/ggplot2), package pheatmap (https://stat.ethz.ch/pipermail/r-help/2012-November/330785.html) and conventional Python scripts were used for visualizing analysis. Statistical analyses were performed in R v3.6.2.

Nucleotide Accession Number
These assemblies sequence data of CRKP isolates from hospital were deposited in the GenBank database under BioProject accession PRJNA564463.

Results
CRKP Isolates of Hospital Had High Genotypic and Phenotypic Resistance
A total of 2846 K. pneumoniae strains were collected from hospital during 2013–2018 and 19 isolates showed resistance to carbapenems. The information of these strains is shown in Supplementary Appendix 2. ANI analysis confirmed these 19 strains (>98% ANIb) are K. pneumoniae. Antibiotic susceptibility testing results revealed that 94.7% (18/19) of CRKP strains showed multi-drug resistant phenotype (Figure 1B). All isolates showed 100% resistance against amoxicillin, ampicillin, ampicillin/subbactam, cefazolin, cefepime, cefuroxime, ceferozeng/subbactam, cefotaxime, ceftriaxone, piperacillin, piperacillin/tazobactam, imipenem, and fosfomycin. None of them were resistant to polymyxins and tigecycline.

WGS analysis demonstrated that CRKP strains carried 29 unique ARGs for eight different classes of antibiotics. At least one kind of ß-lactamase genes, \textit{fosA} gene, was detected in all strains. For carbapenem resistance, 26.3% (n=5) and 68.4%(n=13) of all strains carried \textbf{bla}_{KPC-2} and \textbf{bla}_{NDM-1} genes, respectively. In addition, NDM-producing strains contained more antibiotic resistance genes than KPC-producing isolates (P=0.002) (Figure 1C).

Two Lineages Dominated Hospital CRKP Populations
MLST typing showed that CRKP isolates were divided into four ST types and three unknown types. Most isolates
belonged to ST25 (47.3%, n=9) and ST11 (26.3%, n=5) types, while two isolates were assigned to ST105 and ST1212, respectively. We further identified lineages with FastGEAR/BAPS results which revealed that 19 CRKP isolates contained five BAPS lineages with ST11 and ST25 relating to lineages 4 and 5, respectively. Lineages 4 and 5 represented 78.9% (n=15) of all isolates, and the differences between lineages were consistent with ST types and time collections (Figure 1A).

In addition, bla\textsubscript{KPC-2} was only found in ST11 strains, whereas bla\textsubscript{NDM-1} was found in ST25, ST105, and unknown ST samples (Figure 1A and C and Figure 2C). Hierarchical clustering of isolates based on ARG presence or phenotypic susceptibility indicated lineage was the major predictor of resistance-based clustering patterns (Figure 1B and C).

Geographical Genetic Distribution Characteristics of \textit{K. pneumoniae}

Through MLST classification of 1579 \textit{K. pneumoniae} genomes, a total of 147 known ST types (75.7%, n=1159) were identified and unknown ST types (24.3%) were found in 384 samples. Among the known ST type samples, the distribution of epidemic ST varied by counties (Figure 3). Of these, ST258 (23.6%, n=282) and ST11 (22.1%, n=264) were predominant, followed by ST15 (4.7%, n=56), ST512 (4.2%, n=50), ST147 (3.3%, n=39), and ST231 (3.1%, n=37). In addition, ST11 and ST258 were dominant in China (51.7%, n=231) and the United States (54.1%, n=252), respectively.

Then we detected the presence of bla\textsubscript{KPC-2} and bla\textsubscript{NDM-1} genes in 1579 genomes and found 462 genomes carried the bla\textsubscript{KPC-2} gene and 106 had the bla\textsubscript{NDM-1} gene, respectively. Gene bla\textsubscript{KPC-2} was mostly detected in ST11 (88.2%), ST45 (80.0%), and ST437 (66.7%), while bla\textsubscript{NDM-1} was mostly found in ST1 (66.7%) and ST14 (53.8%) (Figure 2A). Besides, only four genomes contained both genes and they were all isolated from China, with three ST11 isolates and one ST86 strain.

Among the dominant ST types, we found 88% of ST11 \textit{K. pneumoniae} contained bla\textsubscript{KPC-2} and 4% had bla\textsubscript{NDM-1}. 
Figure 2 Distribution of *K. pneumoniae* and carbapenem resistance genes (CRGs) (*bla*KPC-2 and *bla*NDM-1): Types of CRGs are annotated by different colors. All STs are based on the Pubmlst MLST scheme. (A) Distribution of carbapenem resistance genes of CRGs in *K. pneumoniae* genomes from NCBI in the ST type with large abundance or with *bla*KPC-2 or *bla*NDM-1 for display. (B) Geographical distribution of CRGs in *K. pneumoniae* genomes from NCBI (only countries with ≥1 CRG-containing genome are shown). Countries are shown on the y-axis and the numbers on the x-axis indicate the number of CRGs. (C) Distribution of carbapenem resistance genes of CRGs in *K. pneumoniae* genomes from Shunde hospital.
However, in ST258 *K. pneumoniae*, only 28% of genomes carried the *bla*KPC-2 gene and none of them contained *bla*NDM-1. The presence of the *bla*KPC-2 and *bla*NDM-1 genes in China (9% and 62%) was significantly higher than those in the United States (2% and 18%) (Figure 2B).

**Gene Environments of *bla*KPC-2 and *bla*NDM-1**

For the *bla*KPC-2 gene, we found 97.4% (74/76) of the *bla*KPC-2 genes were located in the plasmid. Gene environment analysis found that the *bla*KPC-2 gene is more likely linked to mobile component ISKpn6 and ISKpn27 (46.1%, n=35)/ISKpn7-dnaA (51.3%, n=37), including those strains belonging to ST11 and ST258. Of these, ISKpn6 and ISKpn27/ISKpn7-dnaA are commonly located in *bla*KPC-2 gene 3’ and 5’, respectively. Additionally, ISKpn7-dnaA, instead of ISKpn27, at *bla*KPC-2 gene 5’ was detected in the chromosome and plasmid of ST258 samples from the US and Australia. Moreover, in ST11 samples from China, the IS26 family transposon Tn3 (tnpR) was commonly found in the gene environment of *bla*KPC-2. In addition, the aminoglycoside resistance gene accA4 was also detected in the *bla*KPC-2 gene environment of ST258 samples from the US (Figure 4A).

For the *bla*NDM-1 gene environment, the *bla*NDM-1 gene was all located in plasmids of all genomes (Figure 4B). In most samples, the *bla*NDM-1 was mostly link with IS30 family mobile elements such as *ISaba*125. However, the IS630 family mobile element ISSpu2 and DNA transferase (*hin*) encoding gene fragments were detected in 5’ of *bla*NDM-1 gene of ST11 genomes from Norway, two *hin* were found in 5’ of *bla*NDM-1 gene of ST16 genomes from Thailand. Antibiotic resistance genes were also detected in the *bla*NDM-1 gene environment, such as the tetracycline resistance gene (*tet*), aminoglycoside antibiotic resistance gene (*aph*), and bleomycin resistance (*ble*, *trpF*, and *nagA*).

**Discussion**

*K. pneumoniae* is responsible for human infections and the patient population is the most important reservoir in high-frequency nosocomial *K. pneumoniae* outbreaks. The emergence of *K. pneumoniae* carrying carbapenemases and their worldwide dissemination posed a significant public health crisis. As *bla*KPC-2 and *bla*NDM-1 are the most widely reported among CRKP worldwide, we used genomic and phylogenetic approaches to analyze CRKP isolated from hospital in China and uncover the distribution of these two genes in global genomes from NCBI.

Our findings indicated that the CRKP population from hospital is highly diverse, encompassing five lineages and different STs. All these CRKP isolates were resistant to multiple antibiotics and sensitive to polymyxins and tigecycline, suggesting resistance in CRKP isolates is serious. In particular, ST11 CRKP have been associated with the
global dissemination of *K. pneumoniae*, all the ST types found in hospital were also widely reported in China.\(^{28}\) Interestingly, through core genome phylogenetic analysis, there was no significant genetic difference in the core genome except the sample P38, which indicated there is high homology between samples in different time of the hospitals. Interestingly, P38 did not contain *bla*\(_{KPC-2}\) or *bla*\(_{NDM-1}\), it only carried *bla*OXA-396 and *bla*OXA-494, which may relate to carbapenem resistance.

Hence, based on the actual collection time of the samples (at least in different months of the same year), we can infer that there are certain nosocomial infections and clonal transmission in the time segment of the hospital.

MLST typing results showed that *bla*\(_{KPC-2}\) was only detected in ST11-CRKP (5/19) strains from hospital. According to the global genome data, 88.25% (n=264) of ST11 strains contained *bla*\(_{KPC-2}\) and 4.5% (n=12) carried *bla*\(_{NDM-1}\). These results suggested that ST11 CRKP strains have obvious preference for carrying *bla*\(_{KPC-2}\), which is consistent with previous studies.\(^{29,30}\) In addition, ST11 also has a relatively high level of carrying *bla*\(_{NDM-1}\) in NDM-1 producing CRKP strains (n=106). In addition, the
bla<sub>KPC-2</sub> gene was also found in ST45, ST307, and ST437 <i>K. pneumoniae</i> isolates, among which ST437 was the dominant ST-type in Brazil. These three ST <i>K. pneumoniae</i> were widely distributed in the global scope and had no obvious locality, hence it was worth exploring the underlying spread mechanism.\textsuperscript{31,32}

Notably, CRKP isolates in China carried a relatively high proportion of <i>bla</i><sub>KPC-2</sub> and <i>bla</i><sub>NDM-1</sub>, especially these strains carrying both genes only found in China. Co-occurrence of <i>bla</i><sub>NDM-1</sub> and <i>bla</i><sub>KPC-2</sub> in a clinical <i>K. pneumoniae</i> isolate was considered to result in broad-spectrum antibiotic resistance profiles.\textsuperscript{33} This phenomenon has been reported in India, South Africa, and China.\textsuperscript{34,35} This may indicate that circulating isolates in China are more likely to develop pan drug resistance. Interestingly, all ST25 samples (9/19) from hospital carried <i>bla</i><sub>NDM-1</sub>, while the NDM-1 producing. This preference has not been previously reported and is worth further exploration.

It was noted that <i>bla</i><sub>KPC-2</sub> together with its adjacent mobile component formed three core structures through genetic analysis, including the n3-ISKpn27-bla<sub>KPC-2</sub>-ISKpn6 chain, ISKpn6-bla<sub>KPC-2</sub>-ISKpn27-IS26, and ISKpn6-bla<sub>KPC-2</sub>-ISKpn7-dnaA. This phenomenon means that the core structures ISKpn27-bla<sub>KPC-2</sub>-ISKpn6 and ISKpn6-bla<sub>KPC-2</sub>-ISKpn7-dnaA are likely to carry <i>bla</i><sub>KPC-2</sub> to transfer in the global scope. The first two structures were called Tn1722-based unit transposon and IS26-based composite transposon, respectively.\textsuperscript{36,38} The insertion sequences ISKpn6 and ISKpn7-dnaA were often reported in transposon Tn4401, which was the main genetic structure enhancing the spread of the bla<sub>KPC-2</sub> type genes in different plasmid scaffolds.\textsuperscript{36} In addition, the presence of various resistance genes was detected in the bla<sub>KPC-2</sub> environment of ST11 samples from China. The interaction between these resistance genes and <i>bla</i><sub>KPC-2</sub> is worthy to further investigate.

We found that there were ISA<sub>ba125</sub>, trpF, ISSen4, and IS family elements detected in the environment of <i>bla</i><sub>NDM-1</sub>. In addition, the upstream and downstream of the <i>bla</i><sub>NDM-1</sub> gene often contained transposons (Tn3) or inserted sequence fragments (IS30), which were often involved in horizontal transfer of drug resistance genes.\textsuperscript{39} In most NDM-1 producing samples, ble, trpF, and nagA are closely linked downstream, hence it can be considered that <i>bla</i><sub>NDM-1</sub>-ble-trpF-nagA may be the core structure of horizontal transfer of <i>bla</i><sub>NDM-1</sub>.\textsuperscript{40} It was noted that <i>hin</i> was also detected in 5′ of the <i>bla</i><sub>NDM-1</sub> gene of some ST11 and ST16 genomes. <i>Hin</i> is a special recombinant binding enzyme that can promote the inversion of DNA position, which has been reported to be involved in the inversion control of DNA segment H in the flagellum phase of <i>Salmonella enterica</i>.\textsuperscript{41} This hints that <i>hin</i> may be involved in the horizontal transfer or expression control of <i>bla</i><sub>NDM-1</sub>.

**Conclusion**

In this study, we found all CRKP strains were dominated by two lineages composed of ST25 and ST11. Phylogenetic analysis showed that there was clonal transmission of CRKP in the hospital. Bioinformatic analysis of 1597 NCBI samples revealed that ST11 and ST25 were the most detected clones and both ST types had a preference for carrying the <i>bla</i><sub>KPC-2</sub> gene. The core structure of ISKpn27/ISKpn7-dnaA-bla<sub>KPC-2</sub>-ISKpn6 and bla<sub>NDM-1</sub>-ble-trpF-nagA is highly epidemic in KPC- and NDM-1-producing <i>K. pneumoniae</i>. Our study provides a complete genetic background and geographical distribution for further understanding the transmission mode and mechanism of <i>bla</i><sub>KPC-2</sub> and <i>bla</i><sub>NDM-1</sub>.

**Acknowledgment**

This work was supported by the National Natural Science Foundation of China (Grant No. 31672608) and Guangdong Special Support Program Innovation Team 2019BT02N054.

**Disclosure**

All authors declare that they have no conflict of interest.

**References**


