In silico characterization of bla<sub>NDM</sub>-harboring plasmids in Klebsiella pneumoniae

Zhu Zeng<sup>1</sup>, Lei Lei<sup>2</sup>, Linman Li<sup>3</sup>, Shengni Hua<sup>4</sup>, Wenting Li<sup>5</sup>, Limei Zhang<sup>6</sup>, Qiuping Lin<sup>5</sup>, Zhixiong Zheng<sup>5</sup>, Jing Yang<sup>5</sup>, Xiaohui Dou<sup>7</sup>*, Luan Li<sup>8</sup>* and Xiaobin Li<sup>1</sup>*

<sup>1</sup>Department of Respiratory Diseases, The First Affiliated Hospital of College of Medicine, Zhejiang University, Hangzhou, China; <sup>2</sup>Department of Cadre Health Care, Guizhou Provincial People’s Hospital, Guiyang, China; <sup>3</sup>Health Management Center, The First Affiliated Hospital of Chongqing Medical University, Chongqing, China; <sup>4</sup>Department of Pulmonary and Critical Care Medicine, Zhejiang University, Hangzhou, China; <sup>5</sup>Department of Radiation Oncology, Zhejiang University, Hangzhou, China; <sup>6</sup>Department of Radiation Oncology, Jiangsu Cancer Hospital & Jiangsu Institute of Cancer Research & The Affiliated Cancer Hospital of Nanjing Medical University, Nanjing, China; <sup>7</sup>Department of Oncology, Jiangsu Cancer Hospital & Jiangsu Institute of Cancer Research & The Affiliated Cancer Hospital of Nanjing Medical University, Nanjing, China; <sup>8</sup>Department of Oncology, Zhejiang Provincial People’s Hospital, Zhejiang, China

Klebsiella pneumoniae is a primary culprit of antibiotic-resistant nosocomial infections worldwide, and infections caused by NDM-producing strains are a major threat due to limited therapeutic options. The majority of bla<sub>NDM</sub> cases occur on plasmids; therefore, we explored the relationships between plasmids and bla<sub>NDM</sub> genes in K. pneumoniae by analyzing the variants of bla<sub>NDM</sub>, replicon types, conjugative transfer regions of 171 bla<sub>NDM</sub>-harboring plasmids from 4,451 K. pneumoniae plasmids. Of the nine identified bla<sub>NDM</sub> variants, bla<sub>NDM-1</sub> (73.68%) and bla<sub>NDM-5</sub> (16.37%) were the most dominant. Over half of the bla<sub>NDM</sub>-harboring plasmids of K. pneumoniae were classified into Inc<sup>+</sup> plasmids. IncX3 single-replicon plasmids (46–57kb) carried genes encoding relaxases of the MOB<sub>3</sub> family, T4CP genes of the VirD4/TraG subfamily, and VirB-like T4SS gene clusters, which were mainly geographically distributed in China. We found 10 bla<sub>NDM</sub>-harboring IncN plasmids (38.38–63.05kb) carrying the NW-type origin of transfer (oriT<sup>NW</sup>) regions, genes coding for relaxases of MOB<sub>3</sub> family, genes encoding T4CPs of the TrwB/TraD subfamily, and Trw-like T4SS gene clusters, which were also mainly geographically distributed in China. Moreover, we identified 21 IncC plasmids carrying bla<sub>NDM-1</sub> (140.1–329.2kb), containing the A/C-type oriT<sup>C</sup> genes, genes encoding relaxases of MOB<sub>3</sub> family, genes encoding T4CPs belonging to TrwB/TraD subfamily, and Tra<sub>F</sub>-like T4SS gene clusters. The bla<sub>NDM</sub>-harboring IncC plasmids were widely geographically distributed all over the world, mainly in the United States, China and Viet Nam. These findings enhance our understanding of the diversity of bla<sub>NDM</sub>-harboring plasmids in K. pneumoniae.

**KEYWORDS**

Klebsiella pneumoniae, plasmid, bla<sub>NDM</sub>, replicon types, conjugative transfer region
Introduction

*Klebsiella pneumoniae* is a significant cause of nosocomial infections such as pneumonia, bloodstream infections, urinary tract infections, and septicemias (Pitout et al., 2015; Bengoechea and Sa Pessoa, 2019). *Klebsiella pneumoniae* represents one of the most concerning pathogens known for its high frequency and diversity of antimicrobial resistance (AMR) genes (Navon-Venezia et al., 2017; Wyres and Holt, 2018), and it has been classified as an ESKAPE organism (De Oliveira et al., 2020). The emergence and spread of carbapenem-resistant *K. pneumoniae* have become severe medical problems worldwide (Navon-Venezia et al., 2017). Resistance to carbapenems in *K. pneumoniae* involves diverse mechanisms, e.g., production of carbapenemases (e.g., KPC, NDM, and OXA-48-like), alterations in outer membrane permeability and the upregulation of efflux systems (Pitout et al., 2015).

New Delhi metallo-β-lactamase (NDM), belonging to Ambler class B β-lactamase, has the ability to hydrolyze all β-lactam antibiotics (including carbapenems) except the monobactam aztreonam (Nordmann et al., 2011). NDM-1 was first reported in a *K. pneumoniae* isolate recovered in a Swedish patient who traveled to New Delhi in 2008 (Yong et al., 2009). According to the records of the Beta-lactamase database (BLDB; Naas et al., 2017) on September 8th, 2022, more than 40 variants of NDM have been identified so far. A variety of infections caused by NDM-producing Enterobacterales strains are associated with inferior prognosis and high mortality, especially in high-risk immunocompromised patients (Guducuoglu et al., 2018). NDM-producing Enterobacterales clinical isolates, mainly *K. pneumoniae* and *Escherichia coli*, have been found worldwide, with a higher prevalence in the Indian subcontinent, the Balkans, and the Middle East (Albiger et al., 2015; Wu and Feng, 2019).

Antimicrobial resistance (AMR) in carbapenem-resistant Enterobacterales (CRE) strains is often encoded by the plasmid-borne genes (Rozwandowicz et al., 2018). Plasmids, especially conjugative plasmids, play an essential role in mediating horizontal gene transfer (HGT) and dissemination of AMR (Jiang et al., 2020). The conjugal transfer regions of conjugative plasmids typically comprise four key modules, including origin of transfer (oriT) region, gene encoding relaxase, gene encoding type IV coupling protein (T4CP), and gene cluster for the bacterial type IV secretion system (T4SS) apparatus (de la Cruz et al., 2010). The relaxase initiates the bacterial conjugation by recognizing and cleaving the oriT of the plasmid in a site-specific manner, forming a relaxosome (Llosa et al., 2002; Carballeira et al., 2014). Currently, nine types of plasmid-borne oriT2 and eight main relaxase families1 have been identified (Li et al., 2018). Conjugation requires a pilus, which is assembled by T4SS, to connect the donor and the recipient strains (de la Cruz et al., 2010). Currently, five main types of T4SS gene clusters are defined, including 18 different kinds of systems3 (Bi et al., 2013). The T4CP connects the relaxosome to T4SS, which is required for conjugation, and currently, two main subfamilies of T4CPs4 exist (Li et al., 2018).

Studies on the comprehensive analysis of blaNDM-harboring plasmids and their conjugal transfer regions in *K. pneumoniae* are scarce. In this work, we executed *in silico* typing and comparative analysis of blaNDM-harboring plasmids of *K. pneumoniae* using the bacterial plasmids available in the NCBI GenBank database. We systematically analyzed the variants of blaNDM, replicon types, phylogenetic patterns, and conjugal transfer regions of the blaNDM-positive plasmids of *K. pneumoniae*. This study provides deep insights into the characteristics and diversity of blaNDM-harboring plasmids in *K. pneumoniae* and further emphasizes their role in dissemination of resistance genes.

Materials and methods

Plasmid sequences from the NCBI database

The GenBank Genome database (Benson et al., 2018) collect all the plasmids belonging to *K. pneumoniae*. A total of 4,451 plasmids (without duplicates) of *K. pneumoniae* (Supplementary Table S1) were downloaded on April 26th, 2022. Files in FASTA DNA format of the 4,451 plasmids were downloaded in batches into our Linux-based server.

Identification of the blaNDM-harboring plasmids of *Klebsiella pneumoniae*

The β-lactamase genes of the plasmids of *K. pneumoniae* were identified applying the ResFinder software, standalone version 4.1 (Bortolaia et al., 2020), with the minimum coverage of 60%, minimum identity of 90%, and species of “*Klebsiella*.” The term “blaNDM” was used to search in the “Resistance gene” list of the ResFinder results in order to judge the blaNDM-harboring plasmids of *K. pneumoniae* and identify the variants of the blaNDM genes. For some blaNDM-harboring plasmids, the variants of blaNDM were not determined by the ResFinder software; instead, they were submitted to the CARD database5 (Alcock et al., 2020) and the Beta-lactamase database (BLDB; Naas et al., 2017) for further analysis.

---

1. https://bioinfo.mml.sjtu.edu.cn/oriTDB/browse_ort_type_p.php
2. https://bioinfo.mml.sjtu.edu.cn/oriTDB/browse_relaxase.php
3. https://bioinfo.mml.sjtu.edu.cn/SecReT4/browse_type.php
4. https://bioinfo.mml.sjtu.edu.cn/oriTDB/browse_t4cp.php
5. https://www.ncbi.nlm.nih.gov/genome/browse/#/plasmids/815/
6. https://card.mcmaster.ca
Replicon typing of the $bla_{\text{NDM}}$-harboring plasmids of \textit{Klebsiella pneumoniae}

Replicon typing of the $bla_{\text{NDM}}$-harboring plasmids was executed via the PlasmidFinder software (Carattoli and Hasman, 2020). Then, selecting the database “Enterobacteriales,” the FASTA-formatted DNA files were analyzed and classified in batches by using the PlasmidFinder tool version 2.0.1, with a minimum coverage cut-off of 60% and minimum identity cut-off of 95%. The database version was updated on November 29th, 2021.

Phylogenetic cladogram of the $bla_{\text{NDM}}$-harboring plasmids of \textit{Klebsiella pneumoniae}

The files of the $bla_{\text{NDM}}$-harboring plasmids of \textit{K. pneumoniae}, in GenBank format, were downloaded in batches using two Bioperl modules (Bio::SeqIO and Bio::DB::GenBank). Plasmid files containing protein sequences were compiled from the plasmid files in GenBank format through the Bioperl/Bio::SeqIO module. Phylogenetic cladogram based on the presence/absence of orthologous gene families of all the $bla_{\text{NDM}}$-harboring plasmids of \textit{K. pneumoniae} were constructed. First, a binary gene presence/absence matrix was built using OrthoFinder software (Emms and Kelly, 2019), and subsequently a hierarchical cluster result was generated by PAST3 (Hammer et al., 2001) and eventually displayed by iTOL (Letunic and Bork, 2016).

Geographic location and host ST types of the $bla_{\text{NDM}}$-harboring plasmids in \textit{Klebsiella pneumoniae} strains

Information about geographic location of $bla_{\text{NDM}}$-harboring plasmids and its host strains were extracted from the files of the $bla_{\text{NDM}}$-harboring plasmids in GenBank format. Table containing the correspondence between strains and plasmids of \textit{K. pneumoniae} were downloaded from the GenBank. The $bla_{\text{NDM}}$-harboring plasmid-matched host \textit{K. pneumoniae} strains were collected, and their DNA FASTA sequences were downloaded in batch using the Bioperl. The MLST software (Larsen et al., 2012) version 2.0.9 was downloaded from the website and installed on the Linux platform. The genomes of \textit{K. pneumoniae} strains were analyzed in batch using MLST software.

Characterization of the conjugative transfer regions of $bla_{\text{NDM}}$-harboring plasmids

Files in GenBank format of the $bla_{\text{NDM}}$-harboring plasmids in \textit{K. pneumoniae} were analyzed in batches using oriTfinder software (local version; Li et al., 2018) to identify the presence/absence of oriTs, relaxase-coding genes, T4CP-coding genes, and gene clusters for T4SS. Furthermore, the types of oriTs, relaxases, T4CPs, and T4SSs toward the plasmids were determined based on the exhibition of the oriTDB database (Li et al., 2018). In addition, the types of gene clusters for T4SS were classified based on the SecReT4 database (Bi et al., 2013).

Bipartite network construction, clustering and visualization of the $bla_{\text{NDM}}$-harboring plasmids of \textit{Klebsiella pneumoniae}

The bipartite network was constructed based on all the $bla_{\text{NDM}}$-harboring plasmids of \textit{K. pneumoniae} using the AccNet software using default parameters (Lanza et al., 2017). The obtained network files including nodes, edges and clusters were then imported into the Cytoscape software (Shannon et al., 2003) for visualization. We displayed the relative genomic content of each plasmid by making the diameter of each node proportional to its degree.

Results

Variants of $bla_{\text{NDM}}$ genes in the $bla_{\text{NDM}}$-harboring plasmids of \textit{Klebsiella pneumoniae}

Based on the results analyzed by ResFinder, 171 (3.84%) $bla_{\text{NDM}}$-harboring plasmids (Supplementary Table S2) were identified from 4,451 plasmids of \textit{K. pneumoniae}, which were downloaded from the GenBank Genome database. Among the 171 $bla_{\text{NDM}}$-harboring plasmids of \textit{K. pneumoniae}, nine different variants of $bla_{\text{NDM}}$ were identified (Figure 1A). Among the nine variants of $bla_{\text{NDM}}$, $bla_{\text{NDM}-1}$ was found to be the predominant variant, accounting for 73.68% (126 $bla_{\text{NDM}-1}$-harboring plasmids), followed by $bla_{\text{NDM-5}}$ accounting for 16.37% (28 $bla_{\text{NDM-5}}$-harboring plasmids) (Figure 1A).
Replicon types of bla\textsubscript{NDM}\textsuperscript{*}-harboring plasmids of \textit{Klebsiella pneumoniae}

Replicon typing of the 171 \textit{bla}\textsubscript{NDM}\textsuperscript{*}-harboring plasmids of \textit{K. pneumoniae} was executed using PlasmidFinder. Of the 171 plasmids, 154 were successfully identified with their replicon types, including 103 single-replicon plasmids and 51 multi-replicon plasmids (44 plasmids with two replicons, 33 plasmids with three replicons, and four plasmids with four replicons; Figure 1B; Supplementary Figure S1). For the 103 single-replicon plasmids harboring \textit{bla}\textsubscript{NDM}\textsuperscript{*} in \textit{K. pneumoniae}, the TOP5 prevalent replicons (in descending order) were IncX3 (29 plasmids), IncC (15 plasmids), IncFIB(pQil) (11 plasmids), IncFII (11 plasmids), and IncFII(Yp) (11 plasmids). Of the 44 \textit{bla}\textsubscript{NDM}\textsuperscript{*}-harboring plasmids with two replicons, 25 contained replicons IncFIB(pNDM-Mar) and IncHII(B)pNDM-MAR), which were the most prevalent two-replicon plasmids harboring \textit{bla}\textsubscript{NDM}\textsuperscript{*} in \textit{K. pneumoniae} (Figure 1B; Supplementary Figure S1).

In summary, 21 of the 171 \textit{bla}\textsubscript{NDM}\textsuperscript{*}-harboring plasmids of \textit{K. pneumoniae} were found to carry the replicon of IncC, accounting for 12.28% of all the \textit{bla}\textsubscript{NDM}\textsuperscript{*}-harboring plasmids of \textit{K. pneumoniae} in this study (Figure 1B; Supplementary Figure S1). Notably, 91 of the 171 \textit{bla}\textsubscript{NDM}\textsuperscript{*}-harboring plasmids in our study were found to be the IncF plasmids, including IncFI and IncFII plasmids, accounting for 53.22% of all the \textit{bla}\textsubscript{NDM}\textsuperscript{*}-harboring plasmids of \textit{K. pneumoniae} in this study (Figure 1B; Supplementary Figure S1).

Diversity of the \textit{bla}\textsubscript{NDM}\textsuperscript{*}-harboring plasmids in \textit{Klebsiella pneumoniae}

To get the comprehensive overview of \textit{bla}\textsubscript{NDM}\textsuperscript{*}-harboring plasmids in \textit{K. pneumoniae}, we created a phylogenetic cladogram of the 171 \textit{bla}\textsubscript{NDM}\textsuperscript{*}-harboring plasmids (Figure 2). Based on phylogenetic patterns of the 171 plasmids, combined with the replicon types, conjugative transfer regions, and genome sizes of the \textit{bla}\textsubscript{NDM}\textsuperscript{*}-harboring plasmids, most of the 171 \textit{bla}\textsubscript{NDM}\textsuperscript{*}-harboring plasmids were clustered into 10 main clades (clades I–X), representing 10 plasmid patterns carrying \textit{bla}\textsubscript{NDM}\textsuperscript{*} genes (Table 1).

Clade I: A total of 29 IncX3 plasmids were found in the clade I cluster, mainly \textit{bla}\textsubscript{NDM}\textsuperscript{1} and \textit{bla}\textsubscript{NDM}\textsuperscript{5} (Figure 2). For the 29 IncX3 plasmids harboring \textit{bla}\textsubscript{NDM}\textsuperscript{*}, their genome sizes varied from 45.1 to 159.3 kb (25th percentile = 46.2 kb; 75th percentile = 57.3 kb), with a median size of 53.1 kb (Supplementary Figure S2). For the conjugative transfer regions, all the plasmids belonging to clade I were found to carry genes encoding relaxases of the MOB family characterized by the domain “Relaxase (Pfam: PF03432),” T4CPs of the VirD4/TraG subfamily characterized by the domain “T4SS-DNA_transf (Pfam: PF02534),” and VirB-like T4SS gene clusters (Figure 2; Supplementary Figure S3). Members of clade I were mainly geographically distributed in China (Figure 3; Table 1; Supplementary Table S3).

Clade II: Ten \textit{bla}\textsubscript{NDM}\textsuperscript{*}-positive IncN plasmids were clustered into clade II, mainly carrying \textit{bla}\textsubscript{NDM}\textsuperscript{1} (Figure 2). The genome sizes of the 10 \textit{bla}\textsubscript{NDM}\textsuperscript{*}-harboring IncN plasmids varied from 38.4 to 63.1 kb (25th percentile = 47.2 kb; 75th percentile = 59.8 kb), with a median size of 52.0 kb (Supplementary Figure S1). Almost all the IncN plasmids carried the NW-type oriTs and genes encoding relaxases of the MOB family characterized by the domain “TrwC (PF08751).” All the 10 \textit{bla}\textsubscript{NDM}\textsuperscript{*}-positive IncN plasmids carried the genes encoding T4CPs of the TrwB/TraD subfamily characterized by the domain “TrwB_AAD_bind (PF10412)” and VirB-like T4SS gene clusters (Figure 2; Supplementary Figure S4). The members of clade II were mainly geographically distributed in China (Figure 3; Table 1; Supplementary Table S3). The 10 \textit{bla}\textsubscript{NDM}\textsuperscript{*}-positive IncN plasmids were distributed in seven ST types of
Clade III: Eleven \( \text{bla}_{\text{NDM}} \)-positive IncF plasmids with the IncFIB(pQil) replicon were grouped into clade III, and most were

\( \text{K. pneumoniae} \) strains, and four plasmids were distributed in \( \text{K. pneumoniae} \) ST15 (Table 1; Supplementary Table S3).
54-kb plasmids (Figure 2). Moreover, no conjugative transfer regions were identified in the 11 plasmids of clade III, indicating that the 11 plasmids should be non-transferable. Plasmids belonging to clade III were mainly geographically distributed in Italy and United States (Figure 3; Table 1; Supplementary Table S3). All the members of clade III were harbored by the strains of *K. pneumoniae*. The genome sizes of the four IncFII(pKPX1) plasmids varied from 75.3 to 86.0 kb, all carried the F-type replicon were classified into clade VII (Table 1; Supplementary Table S3).

Clade IV: Five *bla*~NDM~-positive untypeable plasmids were clustered into clade IV, involving three *bla*~NDM~-Positive plasmids and two *bla*~NDM~-Positive plasmids and two *bla*~NDM~-Positive plasmids (Figure 2). These five untyped plasmids, with lengths ranging from 75.3 to 86.0 kb, all carried the genes encoding relaxases of MOB<sub>F</sub> family, genes encoding T4CPs of TrwB/TraD subfamily, and Tra<sub>F</sub>-like T4SS gene clusters (Figure 2; Supplementary Figure S5). For the five plasmids, three were found in Viet Nam, one was found in China, and one was found in South Korea (Figure 3; Table 1; Supplementary Table S3). The STs of *K. pneumoniae* host strains containing the clade IV plasmids were distributed into ST395 and ST16 (Table 1; Supplementary Table S3).

Clade V: Thirteen plasmids with the IncFII replicon, mainly carrying *bla*~NDM~, were classified into the clade V (Figure 2). For the 13 IncFII plasmids harboring *bla*~NDM~, genome sizes varied from 75.3 to 140.6 kb (25th percentile = 88.8 kb; 75th percentile = 101.4 kb), with a median size of 96.2 kb (Supplementary Figure S1). They all carried the F-type oriT and Tra<sub>F</sub>-like T4SS gene clusters (Figure 2; Supplementary Figure S6).

Most of the plasmids clustered into clade V were found to carry genes encoding relaxases of the MOB<sub>F</sub> family and genes encoding T4CPs of the TrwB/TraD subfamily (Figure 2). The members of clade V were widely distributed in India, Southeast Asia, North America, East Asia, and Europe, with the highest prevalence in India (Figure 3; Table 1; Supplementary Table S3). The STs of *K. pneumoniae* host strains containing all Clade V plasmids were mainly distributed in ST16, ST147, and ST2096 (Table 1; Supplementary Table S3).

Clade VI: Four *bla*~NDM~-positive plasmids, including one IncM2 plasmid and three untyped plasmids, were classified into a small cluster named clade VI in our study (Figure 2). These four plasmids, with lengths ranging from 75.6 to 100.2 kb, all carried the L/M-type oriT, genes encoding relaxases of MOB<sub>F</sub> family, genes encoding T4CPs of TrwB/TraD subfamily, and Tra<sub>I</sub>-like T4SS gene clusters (Figure 2; Supplementary Figure S7). The four plasmids were sporadically discovered in Canada, Germany, United Kingdom, and China (Figure 3; Table 1; Supplementary Table S3). No prevalent STs of *K. pneumoniae* host strains containing all Clade VI plasmids were found (Table 1; Supplementary Table S3).

Clade VII: Four *bla*~NDM~-positive plasmids with the IncFII(pKPX1) replicon were classified into clade VII (Figure 2). The genome sizes of the four IncFII(pKPX1) plasmids varied from 96.8 to 250.4 kb. Similar to the conjugative transfer regions of plasmids belonging to clade VI, they all carried the L/M-type

### Table 1: Summary of the 171 *bla*~NDM~-harboring plasmids of *Klebsiella pneumoniae*.

| Clade | Plasmid numbers | Main replicon types | Main *bla*~NDM~ sizes (kb) | Main geographic distribution | Main ST types of hosts | Conjugative transfer region |
|-------|-----------------|---------------------|-----------------------------|-------------------------------|------------------------|-----------------------------|
| I     | 29              | IncX3              | *bla*~NDM~<sub>C</sub>,<sub>I</sub> | 45.1–159.3                   | China                   | –                           | MOB<sub>F</sub>/VirD<sub>A</sub>/VirB-like |
| II    | 10              | IncN               | *bla*~NDM~<sub>C</sub>,<sub>I</sub> | 38.4–63.1                    | China                   | ST15                        | MOB<sub>F</sub>/VirG/TraG |
| III   | 11              | IncFIB(pQil)       | *bla*~NDM~<sub>C</sub>,<sub>I</sub> | 54                           | Italy, United States     | ST147                       | TrwB/TraD/TraD |
| IV    | 5               | untypeable         | *bla*~NDM~<sub>C</sub>,<sub>I</sub> | 75.3–86.0                    | Viet Nam                | ST16                        | MOB<sub>F</sub> |
| V     | 13              | IncFI2             | *bla*~NDM~<sub>C</sub>,<sub>I</sub> | 75.3–140.6                   | India                   | F-type                      | MOB<sub>F</sub>/VirG/TraD |
| VI    | 4               | untypeable         | *bla*~NDM~<sub>C</sub>,<sub>I</sub> | 75.6–100.2                   | –                       | L/M-type                    | MOB<sub>F</sub>/TraD/TraI-like |
| VII   | 4               | IncFIIP(KPX1)      | *bla*~NDM~<sub>C</sub>,<sub>I</sub> | 96.8–250.4                   | –                       | L/M-type                    | MOB<sub>F</sub>/TraD/TraI-like |
| VIII  | 18              | IncF               | *bla*~NDM~<sub>C</sub>,<sub>I</sub> | 94.4–316.2                   | –                       | F-type                      | MOB<sub>F</sub>/TraD/TraF-like |
| IX    | 21              | IncC               | *bla*~NDM~<sub>C</sub>,<sub>I</sub> | 140.1–329.2                  | United States, China, Viet Nam | A/C-type                   | MOB<sub>F</sub>, TrwB/TraD/TraI-like |
| X     | 40              | IncF               | *bla*~NDM~<sub>C</sub>,<sub>I</sub> | 238.0–401.6                  | China, United States     | –                           | MOB<sub>F</sub>, TrwB/TraD/TraI-like |
oriTs, genes encoding relaxases of the MOB₅ family, genes encoding T₄CPs of the TrwB/TraD subfamily, and Tra₁-like T₄SS gene clusters (Figure 2; Supplementary Figure S8). The four plasmids were sporadically discovered in United States, India, Switzerland, and China (Figure 3; Table 1; Supplementary Table S3). No obvious common STs of strains were found (Table 1; Supplementary Table S3).

Clade VIII: Eighteen IncF plasmids, mainly carrying blaNDM-1, were grouped into the clade VIII cluster (Figure 2). Most of the IncF plasmids contained the IncFII(Yp) or IncFII(K) replicon in their genomes. For the 18 blaNDM-1-harboring plasmids of clade VIII, genome sizes varied from 94.4 to 316.2 kb (25th percentile = 106.8 kb, 75th percentile = 150.1 kb), with a median size of 110.6 kb (Supplementary Figure S1). Most of the plasmids of clade VIII were found to contain the F-type oriTs. They all carried the genes encoding relaxases of the MOB₅ family, genes encoding T₄CPs of the TrwB/TraD subfamily, and Tra₁-like T₄SS gene clusters (Figure 2; Supplementary Figure S9). Notably, K. pneumoniae strain JNQH116 plasmid pJNQH116-2 (NZ_CP070900), belonging to the clade VIII cluster, was found to contain both Tra₁-like and Trw-like T₄SS gene clusters in its genome. For clade VIII, its members were widely geographically distributed all over the world, including China, India, Southeast Asia, Middle East, North America (Canada and United States), South Africa, and some European countries (e.g., Germany, Romania, and the United Kingdom; Figure 3; Table 1; Supplementary Table S3). No prevalent STs of K. pneumoniae host strains containing all clade VIII plasmids were found (Table 1; Supplementary Table S3).

Clade IX: A total of 21 IncC plasmids carrying blaNDM-1, were grouped into the clade IX cluster of the phylogenetic cladogram (Figure 2). Their genome sizes varied from 140.1 to 329.2 kb, with the 25th percentile, median size, and 75th percentile being 144.3, 147.9, and 178.7 kb, respectively (Supplementary Figure S1). For the conjugative transfer modules, all the plasmids belonging to clade IX carried the A/C-type oriTs, genes encoding relaxases of the MOB₉ family characterized by the domain “Tra₁₂ (Pfam: PF07514),” mostly genes encoding T₄CPs of TrwB/TraD subfamily, and Tra₁-like T₄SS gene clusters (Figure 2; Supplementary Figure S10). The IncC plasmids harboring blaNDM-1 were widely geographically distributed all over the world, mainly in the United States, Viet Nam, and China (Figure 3; Table 1; Supplementary Table S3). ST11 and ST1967 were the common STs strains containing blaNDM-1-harboring IncC plasmids (Table 1; Supplementary Table S3).

Clade X: A total of 40 mega plasmids, where the length range varied from 238.0 to 401.6 kb (25th percentile = 293.5 kb; median = 327.3 kb; 75th percentile = 355.1 kb), mainly carrying blaNDM-1, were grouped into a large cluster, named clade X in our study (Figure 2). Of the plasmids belonging to clade X, 27
(67.5%) were found to contain both replicons IncFIβ (pNDM-Mar) and IncHIIB (pNDM-MAR), seven (17.5%) were unable to be typed, and four (10.0%) were IncQ1 plasmids. Moreover, all the plasmids of clade X carried genes encoding T4CPs of the TrwB/TraD subfamily and Tra_F-like T4SS gene clusters (Figure 2; Supplementary Figure S11). Most of the plasmids belonging to clade X were found to have oriT and harbored the genes encoding relaxases of the MOBβ family. For the clade with the largest number, clade X, its members were widely distributed all over the world, mainly in China and the United States (Figure 3; Table 1; Supplementary Table S3). ST14, ST11, and ST147 were the common STs strains containing the plasmids of clade X (Table 1; Supplementary Table S3).

We also perform a bipartite network analysis with the 171 blaNDM-harboring plasmids of K. pneumoniae. The bipartite network consisted of two classes of nodes: 171 plasmid units (PUs) and 2,502 homologous protein clusters (HPCs), protein families according to amino acid sequence identity, coverage, and E-value; Figure 4). Edges connected every PU with the HPC that it contained. The PUs of the bipartite network clearly showed distinct clustering phenomena. Overall, one homologous PU clusters (PUCs) contained the almost the same members of Clade X in the analysis above (Figure 4), which was clearly distinct from other PUs. One large region including clades III, VI – IX was also identified, which were mostly IncC and IncF plasmids. In addition, clades I, II, IV, and V were also found their corresponding PUCs in the PU–HPC bipartite network.

Discussion

Global spread of the NDM-type carbapenemases can be partly attributed to the dissemination of various blaNDM-harboring plasmids (Lee et al., 2016; Dong et al., 2022). Therefore, to characterize plasmids harboring blaNDM in K. pneumoniae, we systematically analyzed the variants of blaNDM replicon types, and conjugative transfer regions of 4,451 plasmids belonging to K. pneumoniae from the NCBI GenBank database. Overall, 171 blaNDM-harboring plasmids of K. pneumoniae were identified.

In our study, nine different variants of blaNDM were identified from the 171 blaNDM-harboring plasmids in K. pneumoniae, with blaNDM-1 and blaNDM-5 being highly prevalent; blaNDM-1-carrying plasmids were the most prevalent and accounted for 73.68% of the 171 blaNDM-harboring plasmids. NDM-1 was first reported in 2008 on a 180-kb plasmid of K. pneumoniae strain isolated from a Swedish patient hospitalized in New Delhi, India (Yong et al., 2009). After the first report, NDM-1 was reported in many clinical isolates, mainly K. pneumoniae and E. coli (Kumarasamy et al., 2010). In addition, blaNDM-5 was another common variant in our work, accounting for 16.37% of the 171 blaNDM-harboring plasmids. The variant NDM-5 was first reported on an IncF plasmid of E. coli EC405, isolated from a 41-year-old British patient who had a travel history to India (Horne et al., 2011).

Notably, blaNDM-5 was reported to be the predominant variant in blaNDM-harboring plasmids of E. coli (Zhang et al., 2021).

Our results showed that IncX3 single-replicon plasmids were important carriers of blaNDM in K. pneumoniae, mainly blaNDM-1 and blaNDM-5. IncX3 plasmid is narrow-host range plasmids in Enterobacteriales (Johnson et al., 2012), which has been reported to harbor diverse carbapenemase genes in CRE worldwide (Mouftah et al., 2019). Of the 29 blaNDM-harboring IncX3 plasmids grouped into clade I, most were relatively small with lengths of 46–57 kb (25th percentile = 46.2 kb; median size = 53.1 kb; 75th percentile = 57.3 kb). Based on the results analyzed by the oriTfinder software, all the 29 blaNDM-harboring IncX3 plasmids of clade I contained genes encoding for relaxases belonging to the MOBβ family, with TraI protein encoded by the IncPα plasmid RP4 (Pansegrau et al., 1993) as a representative. T4CPs encoded by the 29 blaNDM-harboring IncX3 plasmids of clade I belonged to the VirD4/TraG subfamily, with the TraG protein of plasmid RP4 and the VirD4 protein of Ti plasmids as representatives (Gomis-Ruth et al., 2004). The blaNDM-harboring IncX3 plasmids classified into clade I contained VirB-like T4SS gene clusters, which are the best-characterized T4SS (Guglielmini et al., 2014). However, no known oriT site was found in most of the IncX3 plasmids harboring blaNDM belonging to clade I of the phylogenetic cladogram, indicating a new type of oriT site, different from the nine oriT families collected in the oriTDB database (Li et al., 2018).

We found 10 blaNDM-harboring IncN plasmids, with IncN or IncN2 replicons, clustered into clade II of the phylogenetic cladogram. They were also relatively small plasmids, with genome sizes varying from 38.38 to 63.05 kb. These blaNDM-harboring IncN plasmids carried the NW-type oriTs, which were characterized by the conserved nick region KGTTS[TATAGC (”[ ]” refers to the nic site of oriT), with oriT sites of IncN plasmid R46 (Hall and Vocker, 1987) and IncW plasmid R388 (Revilla et al., 2008) as representatives. Almost all the plasmids of clade II contained genes coding for relaxases of the MOBβ family, which was characterized by the domain “TrwC (PF08751),” with R388 TraC and F TraI as representatives (de la Cruz et al., 2010). The T4CPs encoded by the blaNDM-positive IncN plasmids belonged to the TrwB/TraD subfamily, which was characterized by the domain “TrwB_AAD_bind (PF10412),” with the TrwB encoded by plasmid R388 from E. coli as a representative (Gomis-Ruth et al., 2004). In addition, the blaNDM-positive IncN plasmids carried Trw-like T4SS gene clusters. The Trw T4SS clusters were regarded as the bacterial conjugation machines that mediate the spread of plasmids among bacterial populations (e.g., the trw locus of broad-host-range IncW plasmid R388; Seubert et al., 2003) while also mediating host-specific erythrocyte infection (e.g., the pathogenesis-related Trw system of Bartonella; Vayssier-Taussat et al., 2010).

Our work showed that 21 IncC plasmids carrying blaNDM-1, with genome sizes from 140.1 kb to 329.2 kb, were clustered into the clade IX of the phylogenetic cladogram constructed by the 171 blaNDM-harboring plasmids in K. pneumoniae. The broad-host-range IncC mega plasmids are essential contributors to the
dissemination of antibiotic resistance genes, and more than 200 fully sequenced IncC plasmids have been reported (Ambrose et al., 2018). The blaNDM-1-harboring IncC plasmids of clade IX contained the A/C-type oriT, with the oriT site of IncA/C conjugative pVCR94ΔX from Vibrio cholera as the prototype (Carraro et al., 2014). Furthermore, these blaNDM-1-harboring IncC plasmids carried genes encoding relaxases of the MOB_H family, characterized by the domain "TraI_2 (Pfam: PF07514)," with TraI encoded by IncHI plasmid R27, TraI encoded by IncA/C plasmid pIP1202, TraI encoded by IncI plasmid R391, and TraI encoded by IncT plasmid Rts1 as representatives (de la Cruz et al., 2010). In addition, most of the IncC plasmids clustered into clade IX contained genes encoding T4CPs of the TrwB/TraD subfamily and Tra_F-like T4SS gene clusters.

In our work, 53.22% (91 out of 171 plasmids) of the blaNDM-1-harboring plasmids of K. pneumoniae were found to be IncF plasmids, and most were multi-replicon IncF plasmids, especially IncFI-type plasmids. IncF plasmids are commonly low-copy-number plasmids, >100kb in size (Villa et al., 2010); however, in our study, the blaNDM-1-harboring IncF plasmids in K. pneumoniae were heterogeneous in size. For example, the blaNDM-1-positive IncF plasmids, with the IncFIB(pQil) replicon, clustered into clade III were mostly 54-kb plasmids; the genome sizes of the IncFII plasmids grouped into clade V varied from 75.31 to 140.6 kb (25th percentile = 88.81 kb; 75th percentile = 101.4 kb); and the 27 plasmids with replicon IncFIB(pNDM-Mar) belonging to clade X were > 250 kb in size. The IncF plasmids comprise a diverse set of conjugative plasmids frequently found in Enterobacterales, which contribute to spreading AMR genes (Villa et al., 2010; Carattoli, 2011). The blaNDM-1-harboring IncF plasmids in K. pneumoniae were also heterogeneous in types of conjugative transfer regions. The IncFII-type plasmids, including clades V and VIII, carried F-type oriTs, genes encoding relaxases of the MOB_F family, genes encoding T4CPs of the TrwB/TraD subfamily, and Tra_F-like...
Mega plasmids with replicons IncFIB(pNDM-Mar):IncHI1B(pNDM-MAR) belonging to clade X mostly harbored the genes encoding relaxases of the MOB family. In our study, we found 11 blaNDM-1-positive IncFIB(pQpl) plasmids classified into clade III without any classical conjugative transfer regions, which were predicted as non-transferable plasmids.

### Conclusion

In this study, we analyzed the variants of blaNDM replicon types, conjugative transfer regions, host STs, and geographical distributions of 171 blaNDM-harboring plasmids from 4,451 K. pneumoniae plasmids, which were downloaded from the GenBank database. Nine variants of blaNDM were found among the 171 blaNDM-positive plasmids, with blaNDM-1 (73.68%) and blaNDM-5 (16.37%) as the most dominant. Over half of the blaNDM-harboring plasmids of K. pneumoniae were classified into IncF plasmids. In addition, IncX3 single-replicon plasmids (46–57 kb), IncN plasmids (38.4–63.1 kb), IncC plasmids (140.1–329.2 kb) were also the common carriers of blaNDM in K. pneumoniae. The blaNDM-harboring IncX3 and IncN plasmids were mainly geographically distributed in China. The IncC plasmids harboring blaNDM-1 were widely geographically distributed all over the world, mainly in the United States, China, and Viet Nam. This study provides important insights into the diversity of blaNDM-harboring plasmids in K. pneumoniae and further addresses their role in the acquisition and spread of resistance genes. However, the genetic diversity and characteristics of blaNDM-harboring plasmids in other Gram-negative species need further study in the future.

### Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary material.

### References

Albiger, R., Glasner, C., Struelens, M. J., Grundmann, H., and Monnet, D. L. (2015). Carbapenem-producing Enterobacteriaceae in Europe: assessment by national experts from 38 countries, May 2015. *Euro Surveill.* 20:30062. doi: 10.2807/1560-7917.es.2015.20.45.30062

Alcock, B. P., Raphenya, A. R., Lau, T. Y., Tsang, K. K., Bouchard, M., Edelharmat, A., et al. (2020). CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. *Nucleic Acids Res.* 48, D517–D525. doi: 10.1093/nar/gkz935

Ambrose, S. J., Harmer, C. J., and Hall, R. M. (2018). Evolution and typing of IncC plasmids contributing to antibiotic resistance in gram-negative bacteria. *Plasmid* 99, 40–55. doi: 10.1016/j.plasmid.2018.08.001

Bengoechea, J. A., and Sa Pessoa, J. (2019). Klebsiella pneumoniae infection biology: living to counteract host defences. *FEMS Microbiol. Rev.* 43, 123–144. doi: 10.1093/femmicro/fyz043

Benson, D. A., Cavanaugh, M., Clark, K., Karisch-Mizrachi, I., Ostell, J., Pruitt, K. D., et al. (2018). GenBank. *Nucleic Acids Res.* 46, D41–D47. doi: 10.1093/nar/gkx1094

Bi, D., Liu, L., Tai, C., Deng, Z., Rajakumar, K., and Ou, H. Y. (2013). SecReT4: a web-based bacterial type IV secretion system resource. *Nucleic Acids Res.* 41, D660–D665. doi: 10.1093/nar/gks1248

Bortolaia, V., Kaas, R. S., Ruppe, E., Roberts, M. C., Schwarz, S., Cattoir, V., et al. (2020). ResFinder 4.0 for predictions of phenotypes from genotypes. *J. Antimicrob. Chemother.* 75, 3491–3500. doi: 10.1093/jac/dkaa345

Carattoli, A. (2011). Plasmids in gram negatives: molecular typing of resistance plasmids. *Int. J. Med. Microbiol.* 301, 654–658. doi: 10.1016/j.ijmm.2011.09.003

Carattoli, A., and Hasman, H. (2020). PlasmidFinder and in Silico pMLST: identification and typing of plasmid replicons in whole-genome sequencing (WGS). *Methods Mol. Biol.* 2075, 285–294. doi: 10.1007/978-1-4939-9877-7_20

### Author contributions

XL, LuL, and XD: conceptualization. ZhuZ: methodology. LeL and LiL: software. SH and JY: validation. WL, LZ, QL, and ZhiZ: formal analysis. ZhuZ and XL: writing—original draft preparation. XL and LuL: writing—review and editing. XL: supervision. All authors contributed to the article and approved the submitted version.

### Funding

This work was supported financially by the grants from the National Natural Science Foundation of China (grant nos. 81902460 and 82002170), the Xiangshan Talent Project of Zhuhai People’s Hospital (grant no. 2020XSYC-02), and the Cultivation Project of Zhuhai People’s Hospital (2019PY-19).

### Conflict of interest

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.

### Publisher’s note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

### Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2022.1008905/full#supplementary-material
Carballería, J. D., González-Pérez, B., Moncalhán, G., and la Cruz, F. D. (2014). A high security double lock and key mechanism in HUH relaxases controls oriT-processing for plasmid conjugation. *Nucleic Acids Res.* 42, 10632–10643. doi: 10.1093/nar/gku741

Carraro, N., Sauvè, M., Matteau, D., Laurzon, G., Rodrigue, S., and Burrus, V. (2014). Development of pVCNR94ΔX from vibrio cholerae, a prototype for studying multidrug resistant IncA/C conjugative plasmids. *Front. Microbiol.* 5:44. doi: 10.3389/fmicb.2014.00044

de la Cruz, F., Frost, L. S., Meyer, R. J., and Zeichner, E. L. (2010). Conjugative DNA metabolism in gram-negative bacteria. *FEBS Microbiol. Rev.* 34, 18–40. doi: 10.1111/j.1571-5797.2009.00195.x

De Oliveira, D. M., Forde, B. M., Kidd, T. J., Harris, P. N., Schermbrü, M. A., Beaton, S. A., et al. (2020). Antimicrobial resistance in ESKAPE pathogens. *Clin. Microbiol. Rev.* 33, e00181–e00119. doi: 10.1128/CMR.00181-19

Dong, H., Li, Y., Cheng, J., Xia, Z., Liu, W., Yan, T., et al. (2022). Genomic epidemiology insights on NDM-producing pathogens revealed the pivotal role of plasmids on *bla_(NDM)* transmission. *Microbiol. Spectrum* 10, e02156–e02121. doi: 10.1128/spectrum.02156-21

Emms, D. M., and Kelly, S. (2019). OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol.* 20,238. doi: 10.1186/s13059-019-1832-y

Gomis-Ruth, F. X., Sola, M., de la Cruz, F., and Coll, M. (2004). Coupling factors involved in bacterial conjugation or protein secretion. *Mol. Microbiol.* 52, 5715–5727. doi: 10.1046/j.1365-2958.2004.038487

Guducuoglu, H., Gursoy, N. C., Yakupogullari, Y., Parlak, M., Karasgin, G., Sunnetcioglu, M., et al. (2018). Hospital outbreak of a Colistin-resistant, NDM-1- and OXA-48-producing *Klebsiella pneumoniae*: high mortality from Pandrug resistance. *Microb. Drug Resist.* 24, 966–972. doi: 10.1097/mdr.0000000000000173

Guglielmini, J., Néron, B., Abby, S. S., García-Ruiz-Bacar, M., de la Cruz, F., and Rocha, E. P. (2014). Key components of the eight classes of type IV secretion systems involved in bacterial conjugation or protein secretion. *Nucleic Acids Res.* 42, 5715–5727. doi: 10.1093/nar/gku194

Hall, R. M., and Vockler, C. (1987). The region of the *inc* plasmid R64 coding for resistance to β-lactam antibiotics, streptomycin/spectinomycin and sulphonamides is closely related to antibiotic resistance segments found in *IncW* plasmids and in *IncF*I-like transposons. *Nucleic Acids Res.* 15, 7491–7501. doi: 10.1093/nar/15.18.7491

Hammer, Ø., Harper, D. A., and Ryan, P. D. (2001). PAST: paleontological statistics software package for education and data analysis. *Palaeontol. Electron.* 4:9.

Horsney, M., Phee, L., and Wareham, D. W. (2011). A novel variant, NDM-5, of *NDM* carbapenemases. *Antimicrob. Agents Chemother.* 55, 5952–5954. doi: 10.1128/AAC.00115-12

Jiang, Y., Wang, Y., Hua, X., Qu, Y., Pelég, A. Y., and Yu, Y. (2020). Pooled plasmid sequencing reveals the relationship between Mobile genetic elements and antimicrobial resistance genes in clinically isolated *Klebsiella pneumoniae*. *Genomics Proteomics Bioinformatics* 18, 539–548. doi: 10.1016/j.gpb.2020.12.002

Johnson, T. J., Bielak, E. M., Fortini, D., Hansen, L. H., Hasman, H., Debroy, C., et al. (2012). Expansion of the IncN plasmid family for improved identification and typing of novel plasmids in drug-resistant Enterobacteriaceae. *Plasmid* 68, 43–50. doi: 10.1016/j.plasmid.2012.03.001

Kumarasamy, K. K., Tolemen, M. A., Walsh, T. R., Bagaria, J., Butt, F., Balakrishna, R., et al. (2010). Emergence of a new antibiotic resistance mechanism in India, Pakistan, and the UK: a molecular, biological, and epidemiological study. *Lancet Infect. Dis.* 10, 597–602. doi: 10.1016/S1473-3099(10)70143-2

Lanza, V. F., Baquero, F., de la Cruz, F., and Coque, T. M. J. B. (2017). AcCNET (accessory genome constellation network): comparative genomics software for accessory genome analysis using bipartite networks. *Bioinformatics* 33, 283–285. doi: 10.1093/bioinformatics/btw601

Larsen, M. V., Cosentino, S., Rasmussen, S., Frits, C., Hasman, H., Marvig, R. L., et al. (2012). Multilocus sequence typing of total-genome-sequenced bacteria. *J. Clin. Microbiol.* 50, 1355–1361. doi: 10.1128/JCM.00694-11

Lee, C.-R., Lee, J. H., Park, K. S., Kim, Y. R., Jeong, B. C., and Lee, S. H. (2016). Global dissemination of carbapenemase-producing *Klebsiella pneumoniae*.