

# Chapter 4

## Study of antibiotic resistance

## Chapter 4

---

### 4.1 Introduction

The increasing number of hospital and community-acquired infections by *Klebsiella pneumoniae* (*Kp*), especially by extended-spectrum beta-lactamases (ESBL) and carbapenemase-producing *Kp*, led to the declaration of *Kp* as an ‘urgent threat to public health’ and ‘priority pathogen’ by public health agencies (CDC2013, WHO 2014). *bla*<sub>OXA</sub>, *bla*<sub>NDM</sub> and *bla*<sub>CTX-M</sub> are one of the most prevalent carbapenemase- and ESBLs-producing genes (Guducuoglu et al., 2018). IMP (Zhao and Hu, 2011), SIM (Lee et al., 2005), AIM (Yong et al., 2012) and VIM (Zhao and Hu, 2011) are important metallo  $\beta$ -lactamase (MBLs); they exhibit resistance to a broad-spectrum  $\beta$ -lactams. These enzymes can partially hydrolyse carbapenems and overexpression of *bla*<sub>IMP</sub>, *bla*<sub>SIM</sub>, *bla*<sub>AIM</sub>, *bla*<sub>VIM</sub> can lead to carbapenem resistance (Meletis., 2016). For instance, *bla*<sub>SIM</sub> share similar identity with *bla*<sub>IMP-type</sub> MBL and exhibit lower-level resistance to meropenem and imipenem (Lee et al., 2005). Certain mobile genetic elements (MGEs) such as insertional sequences and transposons play important role in dissemination and spread of these resistance genes among the bacterial populations (Partridge et al., 2018). Integrative and conjugative elements of *Kp* (ICEKp) are a type of MGE that carry and mobilize *ybt* locus (responsible for yersiniabactin biosynthesis); of which there are at least 14 distinct variants called *ybt*;ICEKp variants (Russo et al., 2015 and 2018; Lam et al., 2018). Besides carbapenemase genes, colistin resistance mediated either by chromosomal mutations in *mcrB* (most commonly) or by plasmid-encoded genes like *mcr* genes contributed to emergence of pandrug-resistant *Kp* strains (Poirel et al., 2017; Magiorakos et al., 2012); which leave no therapeutic options and are associated with high mortality rates (Lázaro-Perona et al., 2018; deMan et al., 2018; Zowawi et al., 2015). Identification of *Kp* strain based on allelic combination present on seven loci is called MLST. A different allele number was given to each distinct sequence within a locus, and a distinct number was attributed to each distinct combination of alleles is called sequence type (ST) (Diancourt et al., 2005). The rapid emergence of multidrug-resistant (MDR) *Kp* is mainly driven by resistance plasmids dissemination and wide geographic

spread of certain MDR *Kp* lineages like ST15, ST101, ST147, ST258, ST307, ST231 (Wyres et al., 2020). Therefore, genotypic investigation of these clinically and globally dominant lineages is important. Mutations occurring in the quinolone-resistance-determining region (QRDR) of the *gyrA* and *parC* gene are responsible for resistance to quinolones are called QRDR mutations (Ramos et al., 2020). Multiplex PCR is a technique for rapid detection of several number of genes present in an isolate, however more than 2000 antibiotic resistant genes are reported in *Kp* (e.g. 250 variants of *bla<sub>OXA</sub>*, more than 200 variants of *bla<sub>SHV</sub>*); it is impossible to detect these genes and their variants using the conventional methods. Hence, for high-resolution analysis of a bacterial genome, whole genome sequencing based approach is necessary.

## 4.2 Materials and methods

### 4.2.1 Antibiotic susceptibility testing

#### 4.2.1.1 Disk Diffusion method

2-3 colonies of respective *Klebsiella* isolates were inoculated in Luria Bertani Broth (LB) (HiMedia Laboratories Pvt Ltd, Mumbai, India) and incubated at 37°C till OD of the culture tube reached 0.5 McFarland standard. Then the culture with OD 0.5 McFarland standard was spread on Mueller-Hilton Agar (HiMedia Laboratories Pvt Ltd, Mumbai, India) plate using a sterile spreader. Then the antibiotic susceptibility disks (HiMedia Laboratories Pvt Ltd, Mumbai, India) were kept on the uniformly spread culture using sterile forceps. 4 to 5 disks can be kept on a single agar plate. Plates were incubated for 16 to 24 hours at 37 °C. After the incubation, plates were observed for the zone of inhibition. Diameter of the zone of inhibition/clearance was noted, and interpretation was done according to the Clinical and Laboratory Standards Institute guidelines, 2016 (CLSI, 2016). Table 4.1 shows the list of all antibiotics tested using disk diffusion method.

Table 4.1 List of antibiotics used for disk diffusion method.

Antimicrobial categories	Antimicrobial agents
Aminoglycosides	Amikacin, Gentamycin, Tobramycin
Penicillins	Ampicillin, Oxacillin, Cloxacillin, Piperacillin
Cephalosporin	Cefaclor, Cefazolin, Cefalexin, Cefazolin, Cefepime, Cefixime, Cefonicid, Cefoperazone, Cefotaxime, Cefotetan, Ceftazidime, Ceftriaxone, Cefuroxime
Monobactam	Aztreonam
Carbapenem	Doripenem, Imipenem, Meropenem, Ertapenem
Combination Drug	Amoxicillin/Clavulanic acid, Ampicillin-sulbactam, Piperacillin/tazobactam, Ticarcillin/clavulanic acid, Cefoperazone/Sulbactam, Co-Trimoxazole
Phenicol	Chloramphenicol
Fluoroquinolones	Nalidixic Acid, Ciprofloxacin, Levofloxacin, Norfloxacin, Ofloxacin
Peptide Drugs	Tigecycline, colistin

#### 4.2.1.2 Determination of Minimal Inhibitory Concentration (MIC)

Broth and agar dilution methods for MIC detection was performed for colistin, tigecycline and fosfomycin, respectively as according to CLSI and EUCAST guidelines, disc diffusion method is not accurate for determination of resistance phenotype against these drugs.

##### **a) MIC determination using Broth-dilution method**

Broth microdilution method was used for MIC determination of tigecycline and colistin using both CLSI (CLSI, 2016) and European Committee on Antimicrobial Susceptibility Testing (EUCAST, 2018) guidelines. OD<sub>600</sub> of an overnight culture was adjusted with phosphate-buffered saline to a turbidity equal to that of the 0.5 McFarland standard, and then diluted 10-fold. This prepared bacterial suspension is used as inoculum for the test and was further

diluted with antibiotic solution to provide a final inoculum density of approximately  $10^5$  CFU/ml in each well. Antibiotic solution was prepared in sterile glass test-tubes using appropriate solvent. For Tigecycline DMSO was used as a solvent. Tigecycline was dissolved in DMSO while Sterile molecular grade distilled water was used to dissolve colistin. For tigecycline stock solution of 512  $\mu\text{g/ml}$  was prepared in glass tube and 50  $\mu\text{l}$  of the stock solution was added in the wells containing 50  $\mu\text{l}$  of MHB. Then 50  $\mu\text{l}$  of bacterial inoculum was added to the wells containing broth and antibiotics. one set of triplicate wells was kept as control with only MHB and bacterial inoculum without the antibiotics. The microtiter-plate was incubated for 24 hours at 37 °C. After incubation, the MIC was read as the lowest concentration of antibiotic at which there was no visible growth. The turbidity was also measured at 600 nm using a microplate reader (Multiskan Go, Thermo Fisher Scientific, Waltham, MA, USA). MIC determination of colistin was done using the same protocol. Concentration range used for colistin was 0.25–64 mg/L using sterile distilled water as solvent. No specific CLSI breakpoints for determination of MIC for colistin and Tigecycline are available, but breakpoint for *P. aeruginosa* and *Acinetobacter* spp. are the same as for EUCAST. Hence, EUCAST breakpoints for MIC determination for colistin and tigecycline were considered. For colistin: susceptible  $\leq 2 \mu\text{g/ml}$ , resistant  $> 2 \mu\text{g/ml}$  (EUCAST and CLSI, 2016; Matuschek et al., 2018); for tigecycline:  $\leq 1 \mu\text{g/ml}$  is susceptible, 2.0  $\mu\text{g/ml}$  is intermediate, and  $> 2.0 \mu\text{g/ml}$  is resistant (Hentschke et al., 2010; EUCAST and CLSI, 2016).

#### **b) MIC determination using agar-dilution method**

Agar dilution was performed according to CLSI guidelines, 2016 (CLSI, 2016). Mueller-Hinton agar (MHA) (HiMedia Laboratories Pvt Ltd, Mumbai, India) was prepared and autoclaved as per the manufacturer's instructions. 15. 25 mg/liter glucose-6-phosphate (G6P) was filter-sterilized using Whatman membrane filter with 0.2-micron pore size (Sigma-aldrich, USA). Fosfomycin disodium salt (HiMedia, India) was prepared using sterile distilled water in concentrations from 0.25 mg/litre to 1,024 mg/liter. Fosfomycin and G6P was

then added to the MHA after the agar was cooled down to 50 °C. An inoculum of 10<sup>4</sup> CFU was spotted onto the agar plates having different concentration of fosfomycin using micropipette and allowed to dry. The plates were incubated for overnight in at 37°C. Next day the plates were observed for the bacterial growth and the lowest concentration of fosfomycin with inhibition of the bacterial growth was determine as the MIC. According to CLSI breakpoints, susceptible = MIC ≤64 µg/ml; intermediate = MIC 128 µg/ml and resistant = MIC ≥ 256 µg/ml (CLSI, 2016; Fedrigo et al., 2017).

#### 4.2.2 Categorization of isolates

Categorization of isolates was done using the criteria given by Magiorakos et al. (Magiorakos et al., 2012) based on the results of antimicrobial susceptibility testing (using Disk diffusion and MIC). Isolates were categorized into three categories: Pandrug-resistant (PDR), Extreme drug-resistant (XDR) and Multidrug-resistant (MDR). Following criteria for defining MDR, XDR and PDR for *Enterobacteriaceae* were used. MDR: non-susceptible to ≥1 agent in ≥3 antimicrobial categories. XDR: non-susceptible to ≥1 agent in all but ≤2 categories (i.e., bacterial isolates are resistant to ≥1 agents of all the categories, but will remain susceptible to all agents of ≤2 antimicrobial categories). PDR: non-susceptible to all antimicrobial agents listed in table 4.2 (Magiorakos et al., 2012). The antimicrobial categories and antimicrobial agents selected for the categorization are listed in table 4.2. Isolates M25 was susceptible for almost all listed antibiotics except ampicillin; hence, considered as susceptible.

Table 4.2 List of antibiotics used for categorization of isolates (as per Magiorakos et al., 2012).

Antimicrobial category	Antimicrobial agent
Aminoglycosides	Amikacin, Gentamycin, Tobramycin
Non-extended spectrum cephalosporins; 1st and 2nd generation cephalosporins	Cefazolin, Cefuroxime
Extended-spectrum cephalosporins; 3rd and 4th generation cephalosporins	Cefepime, Cefotaxime, Ceftazidime
Cephameycin	Cefotetan, Cefoxitin

Phenicols	Chloramphenicol
Tetracyclines	Tetracycline, Doxycycline, Minocycline
Antipseudomonal penicillins + $\beta$ -lactamase inhibitors	Piperacillin/tazobactam, Ticarcillin/clavulanic acid
Fluoroquinolones	Ciprofloxacin
Folate pathway inhibitors	Trimethoprim / Sulfamethoxazole
Monobactams	Aztreonam
Penicillins	Ampicillin
Penicillins + $\beta$ -lactamase inhibitors	Amoxicillin/Clavulanic acid, Ampicillin-sulbactam
Carbapenem	Doripenem, Imipenem, Meropenem, Ertapenem
Phosphonic acids	Fosfomycin
Polymyxins	Colistin
Glycylcyclines	Tigecycline

## 4.2.3 Phenotypic assay for ESBL and MBL production

### 4.2.3.1 ESBL assay

Combined-disk method was used to perform ESBL phenotypic assay (Drieux et al., 2008). Overnight grown culture (OD at 600 nm ~ 0.5 McFarland standard) of the respective isolate was spread uniformly on the Mueller-Hinton agar (HiMedia Laboratories Pvt Ltd, Mumbai, India) plate using sterile spreader. 30  $\mu$ g disk of each cefotaxime (CTX) and ceftazidime (CAZ) (HiMedia Laboratories Pvt Ltd, Mumbai, India) along with disk containing 30ug of the same cephalosporines plus Clavulanic acid (10ug) [cefotaxime + clavulanic acid (CTX + CA or CEC); ceftazidime + clavulanic acid (CAZ + CA or CAC)] were kept at a distance of 30 mm (centre to centre) using sterile forceps. Plates were incubated overnight at 37 °C. After the incubation, plates were observed for the differences between the diameter of two zones of inhibition. Isolates with the difference between two zones with diameters of  $\geq 5$  mm were considered ESBL-producers; the difference between two zones with diameter of  $< 5$  mm were considered as negative for ESBL production.

#### 4.2.3.2 MBL assay

Combined-disk method was used to detect the production of MBL (Yong et al., 2002). Overnight grown culture (OD at 600 nm ~ 0.5 McFarland standard) of respective isolate was spread uniformly on the Muller-Hinton agar plate using sterile spreader. Two imipenem (IPM) disks (HiMedia Laboratories Pvt Ltd, Mumbai, India) were placed on the surface of the agar plate at a distance of 4 to 5 cm from each other. 5 µL of 750 µg/mL EDTA solution (HiMedia, Mumbai, India) was then added to one of the IPM disks (IMP + EDTA or IE). Plates were incubated at 37 °C for 14 to 16 hours. The diameters of inhibition zones displayed around the IPM disk and the IPM-EDTA (IE) disks were compared and measured. Only the isolates with  $\geq 7$ mm difference between two inhibition zones were considered as MBL-producers.

#### **4.2.4 Multiplex PCR for detection of carbapenemase genes**

Multiplex PCR was performed to detect eleven carbapenemase genes present in the collected isolates. The eleven pairs of primers used to amplify internal fragments of the carbapenemase genes with the amplicon size is shown in table 4.3. The primers used were from previous reports by (Ellington et al., 2007) and (Poirel et al., 2011). Amplicons in one PCR reaction were of different sizes to avoid ambiguity and clear interpretation, Hence, 11 pairs of primers were distributed in three set of multiplex-PCR reactions (shown in table 4.4). Multiplex PCR reaction mixture and PCR conditions are shown in table 4.5 and table 4.6. 70 -100ng/µL concentration of genomic DNA (gDNA) isolated previously was used as a template in the reaction. Concentration of each primers used was 100 pM. The distribution of primers in three PCR reactions is shown in figure 4.1.

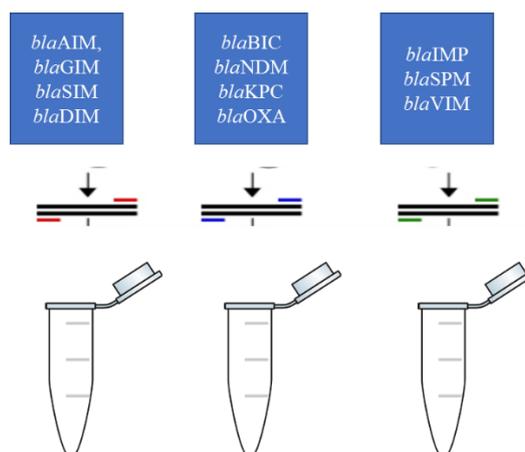
Table 4.3 Primers used for Multiplex PCR and amplicon size.

Primer	Sequence	Amplicon size
<i>bla</i> IMP-F	GGAATAGAGTGGCTTAAYTCTC	232bp
<i>bla</i> IMP-R	GGTTTAAAYAAAACAACCACC	
<i>bla</i> SPM-F	AAAATCTGGGTACGCAAACG	271bp
<i>bla</i> SPM-R	ACATTATCCGCTGGAACAGG	

<i>bla</i> AIM-F	CTGAAGGTGTACGGAAACAC	322bp
<i>bla</i> AIM-R	GTTTCGGCCACCTCGAATTG	
<i>bla</i> VIM-F	GATGGTGTTTGGTTCGCATA	390bp
<i>bla</i> VIM-R	CGAATGCGCAGCACCAG	
<i>bla</i> OXA-F	GCGTGGTTAAGGATGAACAC	438bp
<i>bla</i> OXA-R	CATCAAGTTCAACCCAACCG	
<i>bla</i> GIM-F	TCGACACACCTTGGTCTGAA	477bp
<i>bla</i> GIM-R	AACTTCCAACCTTGGCCATGC	
<i>bla</i> BIC-F	TATGCAGCTCCTTTAAGGGC	537bp
<i>bla</i> BIC-R	TCATTGGCGGTGCCGTACAC	
<i>bla</i> SIM-F	TACAAGGGATTTCGGCATCG	570bp
<i>bla</i> SIM-R	TAATGGCCTGTTCCCATGTG	
<i>bla</i> NDM-F	GGTTTGGCGATCTGGTTTTTC	621bp
<i>bla</i> NDM-R	CGGAATGGCTCATCACGATC	
<i>bla</i> DIM-F	GCTTGTCTTCGCTTGCTAACG	699bp
<i>bla</i> DIM-R	CGTTCGGCTGGATTGATTTG	
<i>bla</i> KPC-F	CGTCTAGTTCTGCTGTCTTG	798/232bp
<i>bla</i> KPC-R	CTTGTCATCCTTGTTAGGCG	

Table 4.4 Distribution of n=11 primer pairs into three sets and their melting temperature (T<sub>m</sub>).

Primer set	Primers	T <sub>m</sub>
AIMset	<i>bla</i> AIM, <i>bla</i> DIM, <i>bla</i> GIM, <i>bla</i> SIM	52°C
KPCset	<i>bla</i> KPC, <i>bla</i> NDM-1, <i>bla</i> OXA, <i>bla</i> BIC	56.5°C
IMPset	<i>bla</i> IMP, <i>bla</i> SPM, <i>bla</i> VIM	57°C



**Figure 4.1 Multiplex-PCR.** Distribution of n=11 primer pairs into three set for performing multiplex PCR.

Table 4.5 Reaction mixture for multiplex PCR.

Reagents	Quantity
Multiplex 2X Fast PCR Master Mix (TaKaRa Bio Inc., Tokyo, Japan)	25 $\mu$ l
Forward primers	1 $\mu$ l each x 4 primers = 4 $\mu$ l (according to the set of primers).
Reverse primers	1 $\mu$ l each x 4 primers=4 $\mu$ l (according to the set of primers).
Sample DNA	2 $\mu$ l
Distilled water	15 $\mu$ l
Total system	50 $\mu$ l

Table 4.6 Cycle/conditions used for multiplex PCR reaction

Step	Temperature ( $^{\circ}$ C)	Time
Initial Denaturation	94	10 seconds
Final Denaturation	94	30 seconds
Primer annealing	52-57 (as per the primer set)	40 seconds
Primer extension	72	50 seconds
Final extension	72	5 minutes

## **4.2.5 Quantitative expression of beta-lactamase genes using quantitative real-time reverse transcription-PCR (qRT-PCR)**

### 4.2.5.1 RNA isolation

Hot phenol method was used for RNA isolation from *Klebsiella* isolates (Sambrook and Russell, 2001). For RNA isolation, molecular grade nuclease-free water was used. Cells from 5ml culture with 0.5 OD<sub>600</sub> (optical density) were harvested and resuspended in 400 µl 50mM sodium acetate (pH 5.3) and 10mM EDTA (pH 8.0). 40µL of 10µL SDS along with 400µL acid phenol (pH 4.3) was added and heated at 65°C for 15mins using water-bath. After the incubation, the tube was transferred to ice for 5 min for rapid chilling. Then the cell suspension was centrifuged at 15000g for 10minutes at 4°C and aqueous phase was transferred to fresh sterile tube. Equal volume (~500 µl) of Acid Phenol: Chloroform: Isoamyl alcohol (in 25:24:1 ratio) was added and incubated on ice for 5 mins followed by centrifugation at 15000g for 10minutes at 4°C. Aqueous phase was transferred to a new sterile tube and the same step was repeated. Equal volume of Chloroform: Isoamyl alcohol (in 24:1 ratio) was added, centrifuged, and aqueous phase was transferred to new tube. The same step was repeated again. Finally, purified RNA was precipitated with 1/10 volume 3M sodium acetate and 2.5 volumes of 100% ethanol and kept on ice for 5 mins, then centrifuged (10minutes, 15000g, 4°C). The pellet is washed in 500µL of 70%ethanol, centrifuged, air dried and resuspended in 20µL of DNase-RNase free water.

### **DNase treatment**

DNase treatment was given using RNase-free DNase-1 recombinant (Roche, Basel, Switzerland), as per the manual instructions. The reaction mixture was made as shown in the table 4.7 and then incubated at 37 °C for 20 minutes. After the incubation, 2µL of 0.2 M EDTA (pH 8.0) was added to stop the reaction and heated for 10 minutes at 75 °C.

Table 4.7 Reaction mixture for DNase treatment of extracted mRNA.

Reaction components	Volume ( $\mu$ l)
RNA	1 $\mu$ g
10X buffer	2.5 $\mu$ l
DNase I	3 (1 Unit/ $\mu$ l)
AMQ water	X $\mu$ l
Total system	25 $\mu$ l

After DNase treatment, 1/10 volume of 3M sodium acetate and 2.5 volumes of 100% ethanol was added for precipitation; and then incubated overnight at -20 °C. After incubation the tube was centrifuged at 10,000g for 10 minutes. Supernatant was removed, and 70% ethanol was added to the pellet followed by centrifugation at 7500g for 5 minutes. Pellet obtained was air-dried and dissolved in 20 $\mu$ L RNase free molecular grade water.

#### 4.2.5.2 cDNA synthesis

cDNA synthesis from total RNA was done using PrimeScript™ cDNA Synthesis Kit (Takara Bio Inc., Shiga, Japan) following the manufacturer's protocol. Table 4.8 shows the reaction system 1 used. The reaction system-I was incubated at 65°C for 5 minutes and then immediately placed on ice. It was then used to prepare reaction system-II as described in table 4.9. The reaction system 2 was then incubated at 30 °C for 10 minutes, followed by 1 hour at 50 °C and finally for 5 minutes at 70 °C. The mixture was then cooled on ice and stored at -20 °C.

Table 4.8 Reaction system-I for cDNA synthesis

Reaction component	Volume/ amount
Rand-6-mers	2 $\mu$ l
dNTP mix	1 $\mu$ l
Template RNA	1.5 $\mu$ g
RNase free water	X $\mu$ l
Total system	10 $\mu$ l

Table 4.9 Reaction system-II for cDNA synthesis

Reaction component	Volume
Reaction system I	10 $\mu$ l
5X Prime Script Buffer	4 $\mu$ l
RNase inhibitor	0.5 $\mu$ l
Reverse Transcriptase	1 $\mu$ l
RNase free water	4.5 $\mu$ l
Total system	20 $\mu$ l

#### 4.2.5.3 Primer designing for qRT-PCR

Primers were designed for *bla*<sub>NDM</sub>, *bla*<sub>OXA-48-like</sub> (carbapenemase-producing genes) and *bla*<sub>SHV</sub> (ESBL-producing gene) genes using Primer-3-Plus online software. Primers were designed from conserved gene sequence available from whole genome sequencing were used for designing primers, hence, presence of variants can't be claimed using these primers. Primer sequences used for qRT-PCR are showed in table 4.10.

Table 4.10 Primers used for qRT-PCR

Primers	Sequence	Amplicon size (bp)	Tm
NDM-1(FP)	TTTGCGGATCTGGTTTTCCG	104bp	59.1
NDM-1(RP)	ACGATCAAACCGTTGGAAGC		59.1
SHV-1(FP)	GGATCTGGTGGACTACTCGC	217bp	58.6
SHV-1(RP)	CGCCTCATTCAAGTCCGTTT		58.8
OXA10-(FP)	GGCGTGGTTAAGGATGAAC	160bp	58.9
OXA-10(RP)	ACTCATACGTGCCTCACCAA		59.3
TEM-1(FP)	GTGCACGAGTGGGTTACATC	177bp	58.9
TEM-1(RP)	GAATAGTGTATGCGGCGACC		59.3

#### 4.2.5.4 qRT-PCR to study gene expression

Relative qRT-PCR was performed to study the expression levels of beta-lactamase genes (*bla<sub>NDM</sub>*, *bla<sub>SHV</sub>*, *bla<sub>OXA-48-like</sub>*, *bla<sub>TEM</sub>*) of *Klebsiella* using SYBR Green (Takara Bio Inc., Shiga, Japan). The *rpoB*, housekeeping gene was used as the internal control. Isolated RNA was used as negative control to check for gDNA contamination. M-25, susceptible isolate with no expression for respective genes was used as reference isolate for calculating fold change. The relative fold change of *bla<sub>NDM</sub>* for each sample was normalized against a housekeeping gene (*rpoB*) by using the comparative Ct method (Schmittgen and Livak, 2008). Table 4.10 shows the sequences of the primers used. Table 4.11 shows the reaction system used for qRT-PCR and table 4.12 shows the qRT-PCR cycle.

Table 4.11 Reaction system used for qRT-PCR

Components for reaction	Volume
SYBR Green Master Mix	10 µl
AMQ water	4 µl
Forward Primer	2 µl
Reverse Primer	2 µl
cDNA Template	2 µl
Total System	20 µl

Table 4.12 Cycle used for qRT-PCR

Step	Temperature (°C)	Time
Initial denaturation	95	3 minutes
Denaturation	94	30 seconds
Annealing	55-59	30 seconds
Primer extension	72	30 seconds
Repeat cycle 35X		
Melt curve	65 to 95 increments 0.5	5 seconds

## **4.2.6 Genomic analysis of PDR, XDR, MDR and susceptible isolates using whole genome sequencing**

### 4.2.6.1 Selection of the isolates

Total n=8 isolates belonging to each resistance category namely, 1 PDR isolate DJ; 3 XDR isolates M2, M6 and M17B; 3 MDR isolates M3, M20 and M27; 1 susceptible isolate M25 were selected for detailed analysis of their whole genomes.

### 4.2.6.2 Whole genome sequencing(WGS) using Illumina and Oxford Nanopore NGS

DNA extraction from n=8 isolates was performed using XpressDNA Bacteria kit (MagGenome Technologies Pvt Ltd., India). Whole-genome sequencing data were generated using an Illumina NextSeq-500 platform with a 2 x 150 nucleotide (nt) paired-end protocol (Nextera XT library, Illumina, San Diego, CA) for all 8 isolates. In case of isolate DJ, long-read Oxford Nanopore sequencing using MinION device was also performed along with Illumina sequencing. MinION device was integrated with a FLO-MIN-106 flow cell and libraries were prepared using a 1D ligation sequencing kit (SQK-LSK109) following the protocol for 1D genomic DNA (gDNA) long reads without BluePippin (Oxford Nanopore Technologies, New York, NY, USA).

### 4.2.6.3 Assembly, annotation and K-serotype determination

*De novo* assemblies of the reads were obtained using SPAdes v3.12.0 (<http://cab.spbu.ru/software/spades/>) for Illumina data and using Unicycler v0.4.4 (Wick et al., 2017) for hybrid assembly. Quality of the assembled genomes was evaluated with QUAST (<http://bioinf.spbau.ru/quast>). Prokka v1.12 (<http://vicbioinformatics.com/>) was used to annotate the assembled sequences. Reads and assembly were deposited at the European Nucleotide Archive database (under the BioProject PRJEB41234).

### 4.2.6.4 Multilocus Sequence Typing (MLST) determination

Multilocus Sequence Typing (MLST) (7 genes) was performed using the international *Klebsiella* sequence typing scheme available on BIGSdb database:

<http://bigsdbs.pasteur.fr/Klebsiella/Klebsiella.html> (Bialek et al., 2014). This data was used to assign the isolates to their respective Sequence Type (ST), which allowed us to determine the respective *Klebsiella* sublineage of each isolate.

#### 4.2.6.5 Phylogenetic analysis of the core genome

##### **a) Phylogenetic tree building of n=8 *Kp* isolates**

Maximum-likelihood phylogenetic tree was built using CSI phylogeny, the online tool available at Centre for Genomic Epidemiology (<https://cge.cbs.dtu.dk/services/CSIPhylogeny/>) (Kaas et al., 2014). The default setting was used to remove SNPs such that no SNPs are within 10 base pairs of each other. SNPs were filtered from the analysis by pruning. Illumina data was used for the phylogenetic tree building. Concatenated alignment of high quality of SNPs were used for the tree building. In the tree, the branch length indicates the divergence (substitution/site) of the genomes and was used to study the most recently diverged genomes.

##### **b) Global analysis of CG147 and ST231 *Kp* genomes**

###### **i) Selection of genomes for CG147 and ST231 global analysis**

We downloaded all publicly available *Kp* genomes belonged to CG147 lineage from NCBI (May 2019). From the 245 CG147 genomes available, we excluded the duplicated ones (n=5), the ones with non-acceptable assembly quality (n=6; genome size and G+C content not matching with *Kp* and/or more than >1000 contigs), and those for which the isolation year was not available (n=17). These filters resulted in a final dataset of 218 genomes (including strain DJ).

Similarly, publicly available *Kp* genomes belonging to ST231 sublineage were downloaded from NCBI (May 2019). From 109 genomes available, the duplicated ones (n=6) were excluded, the ones with non-acceptable assembly quality (n=7) were removed. So, after the filtration, n=97 (including M2 and M6) genomes were used for the ST231 global analysis.

## ii) Phylogenetic analysis of CG147 and ST231 genomes

For phylogenetic analysis of CG147 (n=218) and ST231 genomes (n=97), a core-genome alignment based on the concatenation of 4,529 core genes was obtained using Roary v3.12 (<https://sanger-pathogens.github.io/Roary/>) using a blastP identity cut-off of 90% and core genes defined as those being present in more than 90% of the isolates. Recombination events were removed from the core-genome alignment using Gubbins v2.2.0 (Croucher et al., 2015). The final recombination-free alignment comprised ~8000 single-nucleotide variants (SNVs) and was used to construct a maximum-likelihood phylogenetic tree using IQ-TREE v1.6.11.

These phylogenetic trees were then used as input file to generate the comparative analysis figure in iTOL.

## iii) iTOL map generation for Global analysis of CG147 and ST231

iTOL v3 (<https://itol.embl.de/>) (Letunic and Bork, 2019), is an online tool to display genomic features of the genomes. iTOL was used to represent the genomic features of our isolates (n=8), global analysis of genomes belonging to CG147 lineage (n=218) and ST231 sublineage (n=97). Phylogenetic tree built previously (using method described above) were used as input file to generate figures using iTOL for comparative analysis.

### 4.2.6.6 Detection of antibiotic resistance genes

Genes conferring resistance to different antibiotic families were detected using Kleborate (<https://github.com/katholt/Kleborate>) (Wick et al., 2018) and BIGSdb analytical tools (Jolley and Maiden, 2020). Presence of acquired antimicrobial resistance genes was further confirmed using Resfinder v2.1 (<http://cge.cbs.dtu.dk/services/ResFinder-2.1/>) available at Center for Genetic Epidemiology (CGE). Total, three types of drug-resistance genes were investigated, including acquired drug-resistance genes, genes of efflux pumps, and genes associated with drug resistance by specific mutations for the final data compilation. Geneious Prime v2019.1.1 software (<https://www.geneious.com>) was used for further manual curation of antibiotic resistance genes. EasyFig v2.2.5 (<https://mjsull.github.io/Easyfig/>) (Sullivan

and Beatson, 2011) was used to generate figures for beta-lactamase gene clusters based on nanopore NGS data.

#### a) Detection of *mgrB* mutation

Results from Kleborate indicated the mutation present in *mgrB* gene. Further, we visualized *mgrB* gene sequence and performed pairwise sequence alignment of *mgrB* of DJ and the wild type sequence of *mgrB* (downloaded from NCBI) using Geneious Prime 2019.1.1 software (<https://www.geneious.com>). The insertional sequence observed inside *mgrB* of DJ was then blast using ISFinder. The same was also confirmed by analysing *mgrB* sequence obtained using nanopore NGS of DJ.

#### 4.2.6.7 Detection of mobilome

We investigated mobilome including plasmids, insertional sequences and integrative, conjugative elements (ICE) present in genomes. Plasmid replicons were detected using PlasmidFinder (<https://cge.cbs.dtu.dk/services/PlasmidFinder/>) (Carattoli et al., 2014). ISFinder(<https://isfinder.biotoul.fr>) was used to look for the insertion sequences in the resistance genes or in their genetic context. integrative and conjugative elements (ICEs) were identified using ICEFinder (<https://db-mml.sjtu.edu.cn/ICEfinder/ICEfinder.html>) (Leblond-Bourget et al., 2013).

#### 4.2.6.8 Detection of virulence genes

Genes presumptively associated with virulence such as siderophores ([yersiniabactin](#), aerobactin, colibactin, salmochelin), iron uptake systems and regulators (kfuABC), regulators of mucoid phenotype (*rmpA*, *rmpA2*) and adhesins (*mrkABCD* cluster) were searched using Kleborate (<https://github.com/katholt/Kleborate>) and BIGSdb analytical tools (<http://bigsd.bpasteur.fr/Klebsiella/>) (Jolley and Maiden, 2020). Kleborate was also used for detection of *wzi* type, characterize the capsular gene clusters and ybt;ICEKp variants present in the genomes.

#### 4.2.6.9 Detection of location of the genes

Geneious Prime v2019.1.1 (<https://www.geneious.com>) was used to visualize the contigs and genes present (resistance and virulence genes) on respective contigs. Mplasmid (<https://sarredondo.shinyapps.io/mlplasmids/>) (Arredondo-Alonso et al., 2018) was used to predict whether the contigs were located on a plasmid or chromosome. Mplasmid, Geneious Prime and Kleborate also gave information about number of copies present of respective gene.

Besides Mplasmid and Geneious Prime, location of *bla*<sub>NDM</sub> was studied using a web-based tool, PLACNETw (<https://castillo.dicom.unican.es/upload/>) (Luis Vielva et al; 2017). The same was further confirmed using long-read nanopore sequence of isolate DJ.

## **4.3 Results**

### **4.3.1 Antimicrobial susceptibility testing**

#### 4.3.1.1 Disk diffusion method

Disk diffusion method was used to evaluate antimicrobial susceptibility of the isolate. Table 4.13 shows the results of disk diffusion of all the isolates against various antimicrobial agents. Figure 4.2 shows the representative image for disk diffusion method performed for DJ, M17B, M25 and M27. DJ and M17B did not exhibit zone of inhibition against any antibiotic tested, therefore considered to be resistant against the respective antibiotics. M-25 showed zone of inhibition to all tested antibiotics and considered as susceptible; whereas, M27 was susceptible to very few antibiotics and exhibited resistance against the rest remained susceptible to majority of the antibiotics except resistance against oxacillin, penicillin and ticarcillin/clavulanic acid and considered as susceptible isolates.

Figure 4.3 shows the resistance among isolates against various antimicrobial agents tested. All isolates were resistant against ampicillin, oxacillin and cephalexin (figure 4.3; table 4.13). In case of amikacin, n=10 isolates were found to be resistant and 1 was intermediately resistant. n=5, 13 and 12 isolates

were resistant to chloramphenicol, ciprofloxacin and co-trimoxazole, respectively. Importantly, n=8 and 7 isolates were found to be resistant to imipenem and meropenem, respectively. Colistin and tigecycline were found to be the most effective antimicrobial agents as all isolates were found to be susceptible against them. However, isolate DJ was found to be resistant against colistin and tigecycline (figure 4.3; table 4.13).

Percentage resistance observed in the collected isolates against different antibiotic classes is shown in figure 4.4. 96% and 89% resistance rate were observed against the penicillins and penicillin + Beta-lactamase inhibitors. Resistance to extended-cephalosporines was the highest after penicillins with resistance rate of 82% followed by folate inhibitors (71%), monobactams (64%), Aminoglycosides (53%), Fluoroquinolones (46%) and Phenicol (42%). Importantly, 39% resistance was observed against carbapenems. Low rate of resistance, only 3% (n=1) resistance was observed against glycylyclines (tigecycline) and polymyxin-B (colistin) as only one isolate, DJ was found to be resistant against these drugs.



Antibiotic resistance

ID	Cat	AC	AS	AP	CX	OX	PP	PT	TC	CF	CX	CN	CU	CP	CM	CZ	CE	CC	CR	CT	CO	AZ	DR	ER	IM	MR	CH	AK	GT	TB	NA	CI	LF	NF	OF	CA	TG	CL	TR	DX	MN	
M2 0	MDR	R	R	R	R	R	S	S	R	R	R	R	R	S	R	R	R	R	R	R	R	R	S	S	S	S	S	S	S	S	R	S	R	R	R	S	S	S				
M2 2	MDR	R	R	R	R	R	S	S	R	R	R	R	R	S	R	R	R	R	R	S	R	R	S	R	S	S	R	S	S	S	R	R	R	S	R	R	S	S				
M2 3	MDR	S	S	R	R	R	R	S	R	R	R	R	S	S	R	S	S	R	R	R	R	S	S	S	R	R	S	R	S	S	S	S	S	S	S	S	S	S				
M2 5	Sus	R	S	R	R	R	S	S	R	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	R	S	S	S	S	S	S	S				
M2 7	MDR	R	R	R	R	R	R	R	R	R	R	R	R	S	R	R	R	R	R	R	R	R	R	S	S	S	S	R	S	S	R	S	S	S	S	S	R	S	S			
M2 9	MDR	R	R	R	R	R	R	S	S	R	R	R	R	S	S	R	R	R	R	R	R	R	R	S	S	S	S	S	S	S	R	S	S	S	S	S	S	S	S			
M3 3	MDR	R	S	S	R	R	R	S	R	R	R	R	S	S	S	S	S	S	R	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S			
M3 4	MDR	R	S	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	S	R	R	S	S	S	S	S	S	S	S	S	S	S	S	S	S	R	S	S			
M3 5	MDR	R	S	R	R	R	S	S	R	S	R	R	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S			
M3 6	MDR	R	S	R	R	R	R	S	R	S	R	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S			

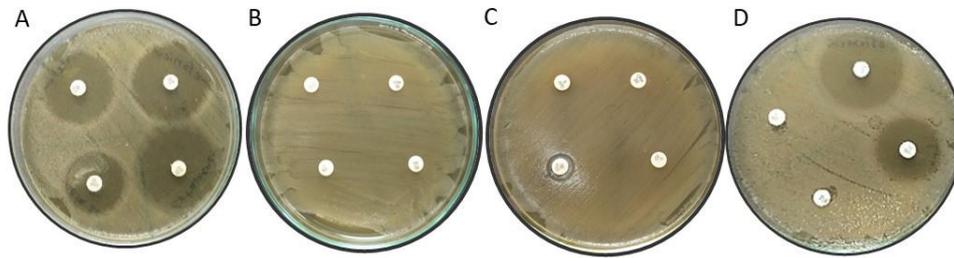
Antibiotic resistance

ID	Cat	AC	AS	AP	CX	OX	PP	PT	TC	CF	CX	CN	CU	CP	CM	CZ	CE	CC	CR	CT	CO	AZ	DR	ER	IM	MR	CH	AK	GT	TB	NA	CI	LF	NF	OF	CA	TG	CL	TR	DX	MN	
M3 7	MDR	R	S	R	R	R	R	S	R	R	R	R	R	S	R	S	S	R	R	R	S	R	S	S	S	S	S	S	S	S	S	S	S	S	R	S	S	S				
M3 9	MDR	R	S	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	S	R	R	S	S	S	S	S	S	R	S	S	S	S	S	S	R	S	S				
M4 0	MDR	R	S	R	R	R	R	S	R	R	R	R	R	R	R	R	R	R	R	S	R	R	S	S	S	S	S	S	S	S	R	R	R	R	R	R	S	S				
M4 1	MDR	R	S	R	R	R	R	R	R	R	R	R	R	S	R	S	R	R	R	R	R	R	S	R	S	R	S	S	R	S	R	R	S	S	S	S	R	S	S			
M4 2	MDR	R	S	R	R	R	R	R	R	R	R	R	R	S	S	S	S	R	R	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S			
M4 3	XDR	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	S	R	R	R	R	R	R	R	R	S	S	R	S	S
M4 4	MDR	R	S	R	R	R	R	S	R	S	R	R	S	R	S	R	R	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S			
M4 6	MDR	R	S	R	R	R	R	R	R	R	R	R	S	S	S	S	S	S	S	S	S	S	S	S	S	S	R	S	S	S	S	S	S	S	S	S	S	S	S			
DJ	PDR	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	S	R	R	R	R	R	R	R	R
ST- 1	XDR	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	S	R	R	R	R	R	R	R	R	R	R	S	S			

Antibiotic resistance

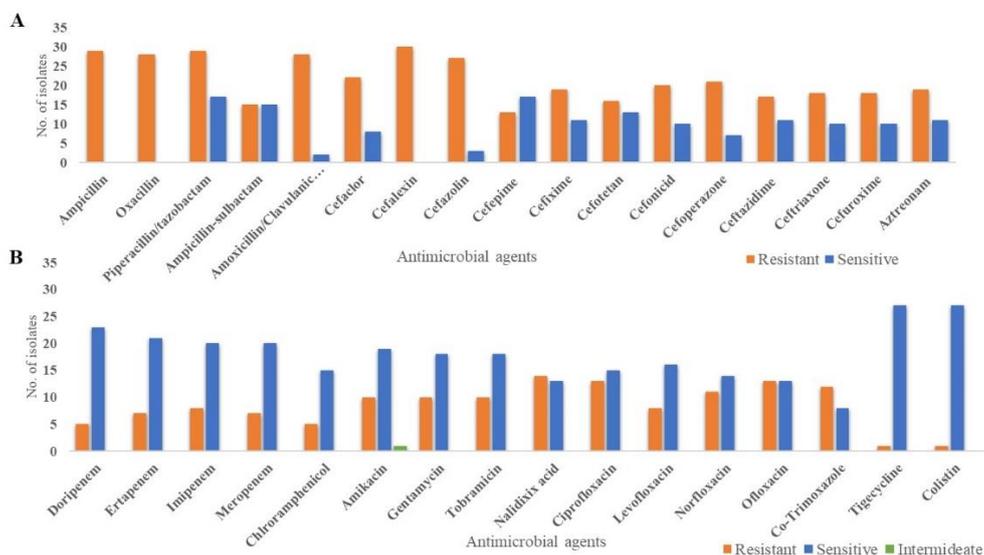
ID	Cat	AC	AS	AP	CX	OX	PP	PT	TC	CF	CX	CN	CU	CP	CM	CZ	CE	CC	CR	CT	CO	AZ	DR	ER	IM	MR	CH	AK	GT	TB	NA	CI	LF	NF	OF	CA	TG	CL	TR	DX	MN
MT CC	MDR	S	R	R	R	R	S	S	R	S	R	R	S	R	S	S	R	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S			

**Footnote:** R= resistant; I=Intermediate; S=sensitive; ID=Isolate ID; Cat=category; Sus=Susceptible; AC=Amoxicillin/Clavulanic acid; AS=Ampicillin-sulbactam; Ap=Ampicillin; Cx=Cloxacillin; Ox=Oxacillin; PP=Piperacillin; PT=Piperacillin/tazobactam; TC=Ticarcilline/clavulinic acid; CF=Cefaclor; CX=Cefalexin; CN=Cefazolin; CU=Cefuroxime; CP=Cefepime; CM=Cefotaxime; CZ=Ceftazidime; CE=Cefixime; CC=Cefonicid; CR=Cefoperazone; CT=Cefotetan; CO=Ceftriaxone; AZ=Aztreonam; DR=Doripenem; ER=Ertapenem; IM=Imipenem; MR=Meropenem; CH=Chlroramphenicol; AK=Amikacin; GT=Gentamycin; TB=Tobramicin; NA=Nalidixic acid; CI=Ciprofloxacin; LF=Levofloxacin; NF=Norfloxacin; OF=Ofloxacin; CA=Co-Trimoxazole; TG=Tigecycline; CL=Colistin; TR=Tetracycline; DX=Doxycycline; MN=Minocycline. Columns highlighted in grey are the antibiotics selected for categorization of isolates as per Magiorakos et al., 2012. Antimicrobial suceptibility to tetracycline, doxycycline and minocycline were performed only for XDR and PDR isolates.

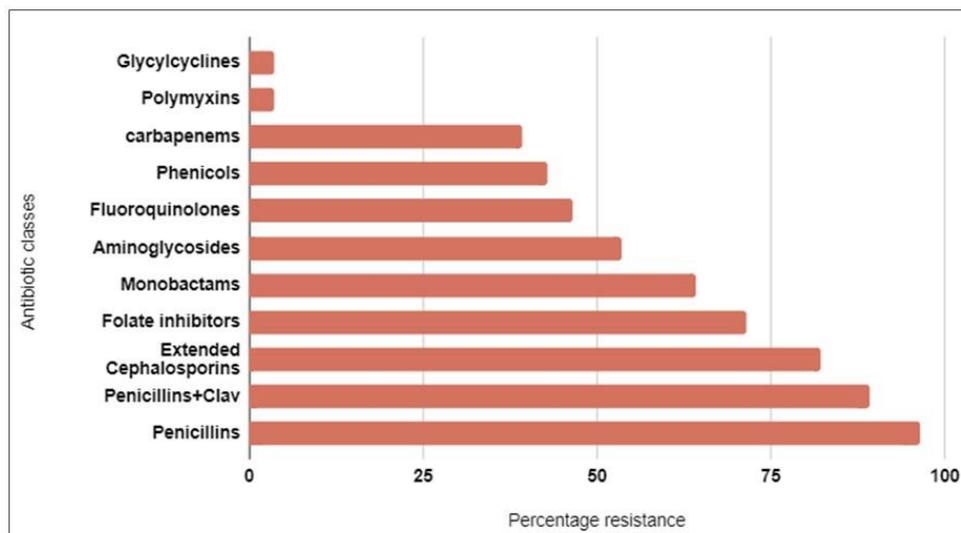


**Figure 4.2 Antimicrobial susceptibility testing using disk diffusion method.**

Representative images of antimicrobial susceptibility testing. Plates containing Himedia antibiotic discs of imipenem, doripenem, meropenem and ertapenem antibiotics are shown. (A) M25 (B) DJ (C) M17B and (D) M27 were spread on respective agar plates. Zone of clearance indicates susceptibility and absence of the clear zone indicates the resistance of the isolate to respective antibiotics.



**Figure 4.3 Resistance against various antimicrobial agents.** Resistance profile of all (n=29) isolates against various antimicrobial categories (A) Antimicrobial agents belong to penicillin and cephalosporins (B)Antimicrobial agents belong to carbapenems, fluoroquinolones, polymyxin and glycytycline categories.



**Figure 4.4 Percentage resistance against various antimicrobial categories.** Percentage of antibiotic resistance observed among the isolates against various antimicrobial categories are shown in the figure.

#### 4.3.1.2 MIC determination

##### **a) MIC determination using agar microdilution method**

MIC of fosfomycin of DJ and few other highly resistant isolates is shown in table 4.14. all 4 isolates were found to be resistant to fosfomycin.

##### **b) MIC determination using broth dilution method**

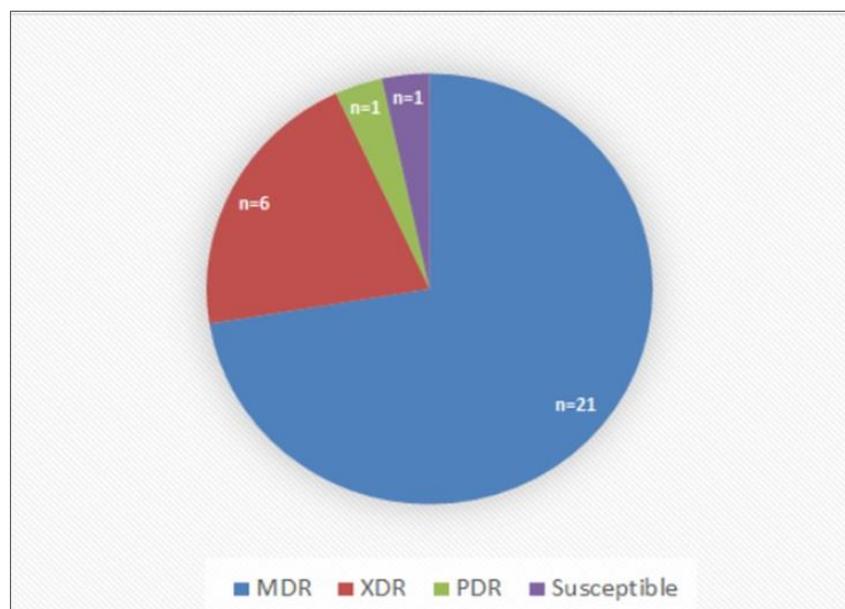
Only DJ remained resistant to colistin and tigecycline during testing using disk diffusion method. However, MIC determination for colistin and tigecycline was performed for DJ and few other highly resistant isolates (according to disk diffusion results). The results of MIC for colistin and tigecycline is shown in the table 4.14. Only DJ was found to be resistant against colistin (MIC = 4  $\mu\text{g/ml}$ ) and tigecycline (MIC= 16  $\mu\text{g/ml}$ ); M17B was intermediate against tigecycline (2  $\mu\text{g/ml}$ ). Rest of the isolates were susceptible to both the drugs.

Table 4.14 MIC determination for tigecycline and colistin.

Isolate ID	Fosfomycin (µg/ml)	MIC of Colistin (µg/ml)	MIC of Tigecycline (µg/ml)
M2	256	0.5	0.25
M6	256	1	1
M17B	512	2	2
DJ	512	4	16

### 4.3.2 Categorization of isolates

Isolates were classified into four categories (criteria for categorization of XDR, PDR and MDR were as per Magiorakos et al., 2012) based on their resistance profile. DJ was categorized as PDR as it was resistant to all antimicrobial agents listed. M25 was susceptible for almost all listed antibiotics except ampicillin; hence, considered as susceptible. From the collected isolates, n=21, n=6, n=1 and n=1 were categorized as MDR, XDR, PDR and susceptible, respectively (figure 4.5 and table 4.13).



**Figure 4.5 Categorization of isolates.** Categorization of the isolates were done in four major categories PDR, XDR, MDR and Susceptible. The categorization was done based on resistance profile of the isolates (criteria used for the categorization were as

per as per Magiorakos et al; 2012). PDR: Pandrug-resistant; XDR: Extreme drug-resistant; MDR: Multidrug-resistant.

### 4.3.3 Phenotypic assay

#### 4.3.3.1 ESBL

Representative images of ESBL producers and non-ESBL producers are shown in figure 4.6. M-34 showed a clear difference of  $\geq 5$  mm between two zones of inhibition (CAC-CAZ  $\geq 5$ mm or CEC-CTX  $\geq 5$ mm) and was considered as ESBL-producer. While M-33 was found to be negative for ESBL production (non-ESBL producers) as the difference between the zones was  $< 5$  mm. Results of rest of the isolates are shown in table 4.15.

#### 4.3.3.2 MBL

Representative images of MBL producers and non-MBL producers are shown in figure 4.7. M-17B was found to be MBL-producer as the difference between two zones of inhibition (IMP and IE) was  $> 7$ mm. While M-33 was found to be negative for MBL production (non-MBL producers) as the difference between the zones was  $< 7$  mm. Results of rest of the isolates is shown in table 4.15.

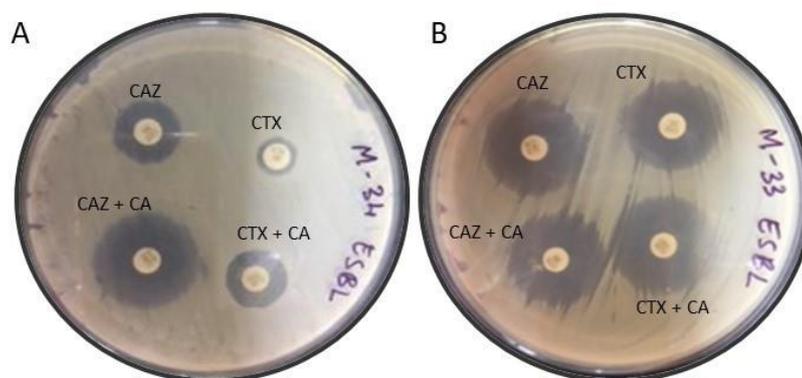
Table 4.15 Results of ESBL and MBL phenotypic assay

Categories	No. of isolates	Sample ID
ESBL-producers	12	M3, M10, M7, M22, M23, M29, M34, M37, M39, M40, ST1
MBL-producers	6	M19, <b>M17B</b> , M18, M27, M33, M36
Both MBL- and ESBL-producers	2	<b>M-6, DJ, M2</b>
Non- ESBL and non-MBL producers	9	M15B, M20, <b>M25</b> , M35, M41, M42, M44, MTCC, M46

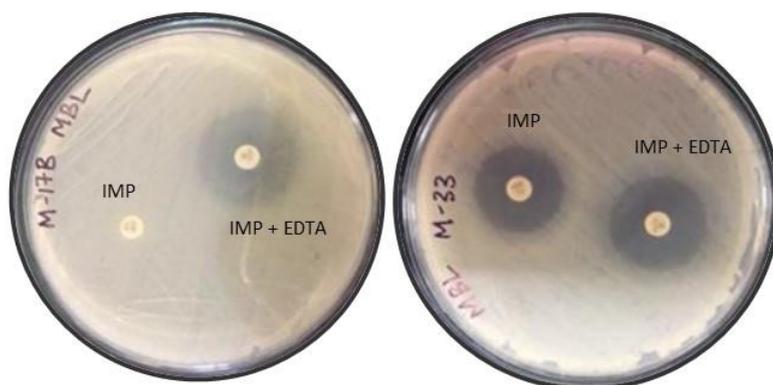
The isolates written in **bold** were used for further investigations.

Table 4.15 and figure 4.8 shows results of phenotypic assay for ESBL and MBL production of all isolates. Among all, n=12 (41.4%) isolates and n=6 (20.7%) isolates were found to be ESBL-producers and MBL-producers, respectively. n=2 (6.9%) of the isolates were producing ESBL and MBL both; 31% (n=9) isolates were non-ESBL and non-MBL producers.

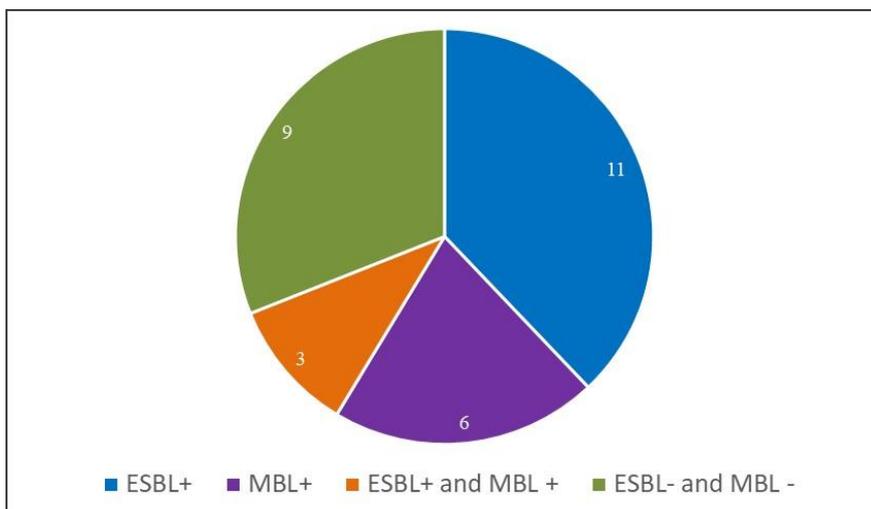
Based on above phenotypic data, five isolates, 1 PDR (DJ), 3 XDR (M-2, M-6, M-17B) and 1 susceptible (M-25) were selected for investigation of the genotype responsible for different resistance phenotype of these isolates.



**Figure 4.6 Phenotypic assay for ESBL.** Representative image of phenotypic assay for ESBL performed for isolate M33 and M34 using Muller Hilton agar plate. Isolates with the difference between the zone diameter of  $\geq 5$ mm were considered as positive for production of ESBL. (A) M34, positive for ESBL production as CAC-CAZ  $\geq 5$ mm or/and CEC-CTX  $\geq 5$ mm (B) M33, negative for ESBL production as CAC-CAZ  $< 5$ mm or/and CEC-CTX  $< 5$ mm. Antibiotic discs CAZ= ceftazidime(30 µg), CAZ+CA(CAC)= ceftazidime(30 µg) + clavulanic acid(10 µg) (B) CTX= cefotaxime(30 µg) and CTX+CA(CEC)=cefotaxime(30 µg)+ clavulanic acid(10 µg).



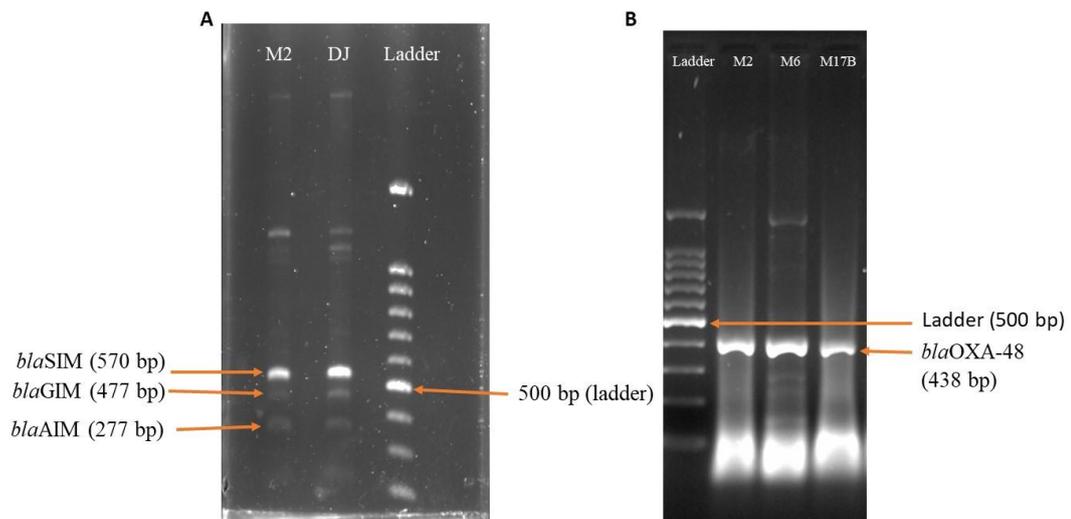
**Figure 4.7 Phenotypic assay for MBL.** Representative image of phenotypic assay for MBL performed for isolate M17B and M33 using Muller Hilton agar plate. Antibiotic discs used were: IMP = imipenem(10ug); IMP+EDTA(IE)= imipenem(10ug) +EDTA(750ug). (A) M-17B positive MBL-producer as IE-IMP $\geq$  7mm. (B) M33, negative for MBL production as IE-IMP < 7mm.

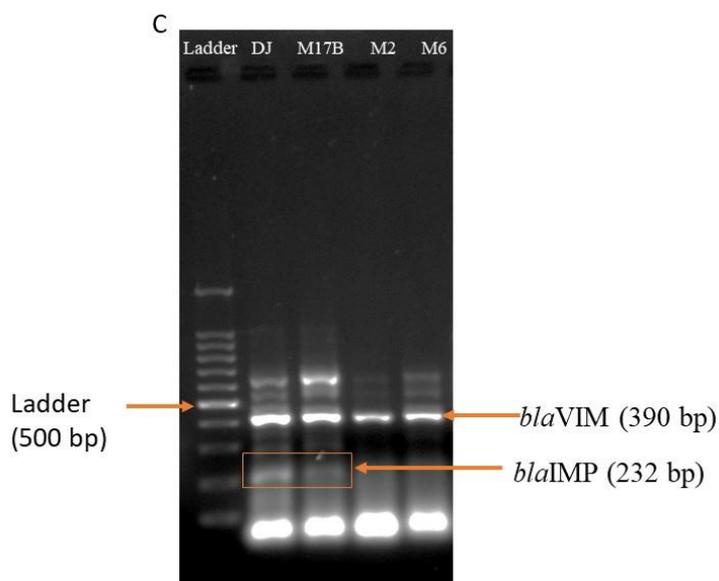


**Figure 4.8 Summary of ESBL MBL phenotypic assay.** n=11 were only ESBL producers (ESBL+), n=6 were only MBL-producers (MBL+), n=3 were both ESBL and MBL producers (ESBL+ and MBL+) and n=9 were non-ESBL, non-MBL producers (ESBL- and MBL-).

#### 4.3.4 Multiplex PCR

Figure 4.9 shows representative images of multiplex PCR amplicons. In case of *bla*<sub>AIM</sub> primer-set, *bla*<sub>GIM</sub> was amplified in M2, M6, M17B (XDR) and DJ (PDR); *bla*<sub>SIM</sub> and *bla*<sub>AIM</sub> were amplified only in M2 and DJ. In case of *bla*<sub>KPC</sub> primer-set, *bla*<sub>OXA</sub> was detected in M2, M6, M17B and DJ; *bla*<sub>NDM</sub> was detected only in DJ. *bla*<sub>KPC</sub> was not detected in any of the isolates. In case of *bla*<sub>IMP</sub> primer-set, *bla*<sub>VIM</sub> was detected in M2, M6, M17B and DJ; *bla*<sub>IMP</sub> was detected only in M17B and DJ. Table 4.16 shows the summary of carbapenemase genes in the selected isolates. More number of carbapenemase genes including the coexistence of *bla*<sub>OXA-48-like</sub> and *bla*<sub>NDM</sub> was observed in DJ (PDR); whereas, no carbapenemase gene was detected in M25 (susceptible). *bla*<sub>NDM-1</sub> and *bla*<sub>OXA-48-like</sub> positive strains were further confirmed by DNA sequencing.





**Figure 4.9 Multiplex PCR of carbapenemase genes.** Representative images of agarose (1%) gel electrophoresis of carbapenemase genes amplified using multiplex PCR. (A) amplicons of *bla<sub>SIM</sub>* (570 bp), *bla<sub>GIM</sub>* (477 bp) and *bla<sub>AIM</sub>* (277 bp) in M2 and DJ. (B) Amplicon of *bla<sub>OXA-48</sub>* (438 bp) in M2, M6 and M17B. (C) Amplicons of *bla<sub>VIM</sub>* (390 bp) amplified in DJ, M17B, M2 and M6; amplicons of *bla<sub>IMP</sub>* (232 bp) in DJ and M17B. 1Kb DNA ladder was used for each run along with the amplicons also seen in the images. Name of the isolates are shown on the top of each lane in all the three images. Orange arrows show the amplicons of respective carbapenemase genes of different length. were amplified in isolates M2 and M3.

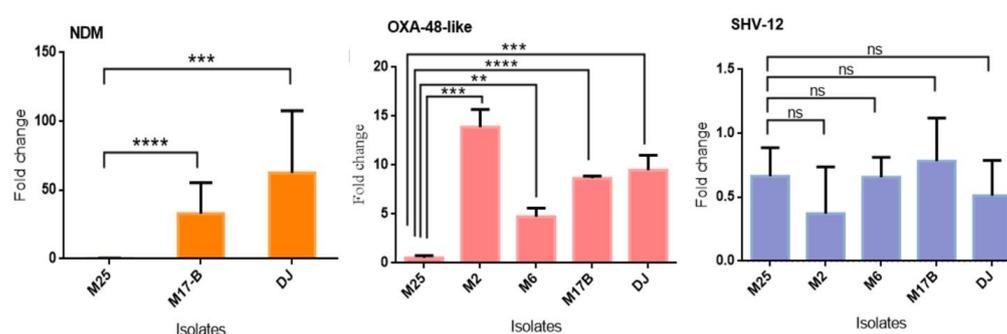
Table 4.16 Carbapenemase genes detected using multiplex PCR

Isolates	<i>blaAIM</i> set	<i>blaKPC</i> set	<i>blaIMP</i> set
M2	<i>blaGIM</i> , <i>blaSIM</i> , <i>blaAIM</i>	<i>bla<sub>OXA-48</sub></i>	<i>blaVIM</i>
M6	<i>blaGIM</i>	<i>bla<sub>OXA-48</sub></i>	<i>blaVIM</i>
M17B	<i>blaGIM</i>	<i>bla<sub>OXA-48</sub></i>	<i>blaVIM</i> , <i>blaIMP</i>
DJ	<i>blaGIM</i> , <i>blaSIM</i> , <i>blaAIM</i>	<i>bla<sub>NDM</sub></i> , <i>bla<sub>OXA-48</sub></i>	<i>blaVIM</i> , <i>blaIMP</i>
M25	-	-	-

### 4.3.5 Quantitative expression of beta-lactamase genes using qRT-PCR

qRT-PCR was performed to study the expression of *bla*<sub>NDM</sub>, *bla*<sub>OXA-48-like</sub> (important carbapenemase-producing genes) and *bla*<sub>SHV</sub> (ESBL-producing gene) genes in XDR and PDR isolates selected above.

The relative fold change of all three genes for each sample was normalized against a housekeeping gene, *rpoB* by using the comparative Ct method. Relative expression of *bla*<sub>NDM</sub> gene was higher in DJ than M17B. The relative fold change for expression of *bla*<sub>NDM</sub> was 62.54 and 32.97 for DJ and M17B, respectively; M25 showed no expression of *bla*<sub>NDM</sub>. Relative expression of *bla*<sub>OXA-48-like</sub> was 13.87 (highest), 9.48, 8.65 and 4.74-fold-change in M2, DJ, M-17B and M-6, respectively, whereas no expression of *bla*<sub>OXA-48-like</sub> was found in M-25. In case of *bla*<sub>SHV-1</sub> negligible relative expression was found in all 5 isolates with fold change of ~0.5. Statistically significant difference in the relative expression of *bla*<sub>NDM</sub> and *bla*<sub>OXA-48-like</sub> (carbapenemase genes) was observed between PDR/XDR (DJ, M-2, M-6) isolates and susceptible isolate (M-25) (relative p value is shown in figure 4.10). No significant difference was observed between the isolates in the relative expression of ESBL gene, *bla*<sub>SHV</sub>. M25 showed no expression of *bla*<sub>SHV</sub>.



**Figure 4.10 Relative-expression of  $\beta$ -lactamase genes.** Relative expression of carbapenemase producing genes (A) *bla*<sub>NDM</sub>, (B) *bla*<sub>OXA-48-like</sub>, and ESBL producing gene (C) *bla*<sub>SHV</sub> using qRT-PCR is shown in the figure. The y-axis shows the relative fold change of *bla*<sub>NDM</sub>, *bla*<sub>OXA-48-like</sub> and *bla*<sub>SHV</sub> to the *rpoB* respective isolates (x-axis). M25 (susceptible isolate) used as reference. Statistical analysis was (internal control)

in performed by the unpaired t-test using Prism-8 GraphPad. \* $p < 0.05$ ; \*\* $p < 0.01$ , \*\*\* $p < 0.001$ , \*\*\*\* $p < 0.0001$ , no significance  $p > 0.05$ . Difference not significant statistically is not shown in the figure.

#### 4.3.6 Genomic analysis of PDR, XDR, MDR and susceptible isolates using WGS by Illumina NGS

To investigate the genomic features of PDR, XDR, MDR and susceptible isolates and understand the genotypes responsible for the difference in phenotypes, WGS was performed for 8 selected isolates belonging to each resistance category; namely, M2, M6, M17B (XDR), DJ (PDR), M27, M20, M3 (MDR) and M25 (susceptible).

##### 4.3.6.1 Quality analysis and MLST

Table 4.17 shows the quality analysis performed for Illumina whole genome sequences of  $n=8$  isolates (as mentioned above) including number of contigs present in each genome, genome size (bp) and GC content (%). Number of total contigs sequenced in respective genome, size (in bp) and GC content (%) of the genomes are shown in the table. The size of the genomes sequences was between 5000000 bp to 5500000 bp. GC content of all the genomes was approximately 57%, which indicates that the genomes belonged to *K. pneumoniae sensu stricto* (*i.e.*, phylogroup Kp1). The PDR isolate, DJ belonged to ST147; M2 and M6 belonged to ST231 and M17B belonged to ST14. Susceptible isolate, M25 belonged to ST1087. M3, M20 and M27 (MDR) belonged to ST16, ST2943 and ST1715, respectively (table 4.18).

Table 4.17 Quality analysis of the sequences.

Isolate ID	Contigs	Size of genome sequenced (bp)	GC (%)
DJ	103	5567922	57.08
M-17B	117	5736445	56.55
M-2	89	5399150	57.06

M-6	127	5514447	57.31
M-27	73	5582316	57.26
M-20	55	5194719	57
M-3	74	5317738	57.18
M-25	88	5307008	57.02

Table 4.18 MLST type of 8 isolates based on allelic combinations at 7 loci.

Isolate	gapA	infB	mdh	pgi	phoE	rpoB	tonB	ST
DJ	3	4	6	1	7	4	38	<b>147</b>
M17B	1	6	1	1	1	1	1	<b>14</b>
M2	2	6	1	3	26	1	77	<b>231</b>
M6	2	6	1	3	26	1	77	<b>231</b>
M27	2	2	2	3	8	25	15	<b>1715</b>
M20	2	52	2	1	12	1	34	<b>2943</b>
M3	2	1	2	1	4	4	4	<b>16</b>
M25	2	1	1	1	3	3	38	<b>1087</b>

Comparative analysis of phylogeny and genotype of PDR (DJ), XDR (M2, M6 and M17B), MDR (M3, M20 and M27) and susceptible (M25) isolates is shown in figure 4.11A. The branch length of the phylogenetic tree suggests that M6 (0.385 substitution/site) and M2 (0.384 substitution/site) were the most recently diverged genomes among all 8 genomes, followed by DJ (0.379 substitution/site) and M25 (Branch length 0.378 substitution/site). M20 (Branch length 305 substitution/site) was the least diverged among these 8 genomes. Figure 4.11B shows number of resistance genes and plasmids present in PDR, XDR, MDR and susceptible isolates. K-locus (KL) type of DJ, M2, M6, M17B and M27 was KL-64, KL-51, KL-51, KL-2 and KL-2, respectively. KL-type of M25, M20, and M3 was not predicted.

#### 4.3.6.2 Detection of antibiotic resistance genes

Significant difference in number of resistance genes and plasmids were observed between XDR and Susceptible isolates (figure 4.11, table 4.19a and 4.19b). Only one ESBL producing gene, SHV-42 was found to be present in M-25 (susceptible) and no plasmid was found. Whereas 12 to 15 resistance genes were detected in XDR isolates along with presence of 4 to 9 various plasmids. In DJ (PDR), 8 resistance genes and 4 plasmids were found (figure 4.11B). Importantly, in DJ and M17B Coexistence of two carbapenemase genes was observed; DJ was harbouring *bla<sub>NDM-5</sub>* and *bla<sub>OXA-181</sub>* while M17B carried *bla<sub>NDM-1</sub>* and *bla<sub>OXA-232</sub>*, this explains increased resistance to carbapenems and other beta-lactams. M2 and M6 carried only one carbapenemase gene, *bla<sub>OXA232</sub>*. M3, M20, M27 (MDR) and M25 (susceptible) isolates did not carry any carbapenemase producing genes. DJ also showed coexistence of *bla<sub>NDM</sub>* and *rmtF*, which could be responsible for increased resistance to carbapenems and aminoglycosides. Coexistence of *bla<sub>NDM</sub>* and *qnrB*, responsible for resistance to carbapenems and quinolones was found in M17B. Other carbapenemase genes such as *bla<sub>KPC</sub>*, *bla<sub>VIM</sub>* and *bla<sub>IMP</sub>* were not found in any isolate. In case of ESBL producing genes, *bla<sub>CTX-M</sub>* was found in DJ, M2, M6 and M27. *bla<sub>SHV-42</sub>* and *bla<sub>SHV-27</sub>* were present in M25 and M-20, respectively. M20 and M25 did not carry any other resistant determinants. Other beta-lactamase genes such as *bla<sub>TEM-1</sub>* and *bla<sub>SHV-11</sub>* or 28 were found in all 4 XDR and PDR isolates; *bla<sub>OXA-1</sub>* was present in M17B. DHA-1, responsible for resistance to third generation cephalosporins was only present in M6. Interestingly, *mgrB* was found to be disrupted by IS5 element, insertional inactivation of *mgrB* is responsible for colistin resistance in the PDR isolate, DJ. Such genotype was not observed in any other isolate. *mcr* was not found in any isolate.

Regarding resistance to quinolones and fluoroquinolones, mutation in *parC-80I* was found in all 4 XDR and PDR isolates. *gyrA-83I* mutation was present in all DJ (PDR), M2, M6 (XDR) isolates except M17B (XDR). M17B exhibited mutation in *gyrA-87Y* and *gyrA-87G*; *gyrA-87G* was also found in M6. Mutation in *parC-84K*, *gyrA-83F* and *gyrA-87N* were only found in M3 (MDR) (along with *qnrB*); However, no resistance genes to other antimicrobial categories were found in M3. Apart from these, M3, M6 and M17B were

harbouring *qnrB*, while M2 carried *qnrS1*; *qepA* was found only in M6. Isolates DJ, M2, M6, M17B carried one or more quinolone resistance determinant regions (QRDR) mutations, whereas M20, M27 (MDR) and M20 (susceptible) did not carry genes/mutations responsible for quinolone/ fluoroquinolone resistance. Besides QRDR mutations, aminoglycosides resistance genes such as *rmt* was also found in DJ, M2 and M6, while *strA* and *strB* were only found in DJ. *aadA* was found in all 4 PDR and XDR isolates. Co-occurrence of *aadA*, *sul1*, *dfrA* and *catA/B* was observed in all 4 XDR and PDR isolates, whereas *sul2* was only present in DJ. Further, M17B carried three variants of *dfrA* (*dfrA1*, *dfrA12*, *dfrA14*). Tetracycline resistant genes *tetA* and *tetB* were found in M3 and M6, respectively. Macrolide resistance gene *arr2* was found in DJ, M2 and M6; *mphA* was found in M17B (*bla<sub>NDM</sub>* is present), M2 and M6. Collectively, *bla<sub>OXA-232</sub>*, *bla<sub>TEM-1</sub>*, *bla<sub>SHV-11/28</sub>*, *gyrA*, *parC* alterations, *aadA*, *sul1*, *dfrA*, *catA/B* and *mphA* were commonly present in all XDR and PDR isolates, which were responsible for their resistant phenotype across multiple antibiotics. Along with these, the PDR isolate, DJ also carried *bla<sub>NDM</sub>*, *bla<sub>CTX-M</sub>*, *mgrB* mutation, *sul2*, *strA*, *strB* and *rmt* (figure 4.11A). No resistance genes/determinants except *bla<sub>SHV-27/42</sub>* was present in M25 (susceptible).

#### 4.3.6.3 Detection of mobilome

Isolates were found to carry different combinations of ybt;ICEKp. Plasmid content was observed to be associated with different ybt;ICEKp variants. M-20 and M-25 with ybt15;ICEKp5 did not found to carry any plasmid. Moreover, M2 and M6 (ST231 isolates), both are ybt14;ICEKp5 variants shared similar plasmid content. Both isolates carried IncFIB(pQil), IncFII(K), Col4401 and ColKp3 plasmids. M6 carried highest number of plasmids (n=9). Moreover, IncFIA, IncFIB(AP001918), IncFII(pAMA1167-NDM-5), IncN and ColBS512 were only present in M6. M17B with ybt10;ICEKp4 was found to carry IncHI1B, IncFII, IncFIB(K), IncFIB(Mar) (virulence plasmid) and ColKP3. M17B is the only isolate carried a virulence plasmid IncFIA(Mar) and yersiniabactin along with MDR genes *bla<sub>NDM</sub>*, *bla<sub>OXA232</sub>* and *bla<sub>CTX-M</sub>*. DJ with ybt10;ICEKp4 carried IncFII, IncFII(pKPX1), IncR and ColpVC. M3 did not possess yersiniabactin and was found to carry FIA(pBK30683) and Col440II.

M27 with unknown *ybt*, *ICEKp* was harbouring IncFIB(K), IncFII(K) and IncI2 (figure 4.11 and 4.19b).

#### 4.3.6.4 Detection of virulence genes

Regarding virulence, all the isolates except M3 carried yersiniabactin synthesis genes and were associated with respective *ybt*, *ICEKp* variants. None of the isolates were found to carry aerobactin genes (*iucABC*), colibactin (*clb* cluster) and salmochelin (*iroBCDN*) siderophore producing genes. Genes for ferric uptake system (*kfuABC*) was found to be present in M2, M6, M17B and M27. *rmpA* and/or *rmpA2* were also not found in any isolate. *mrkABCD* cluster responsible for Type-3 fimbriae expression was present in all 8 isolates (figure 4.11 and 4.19b)

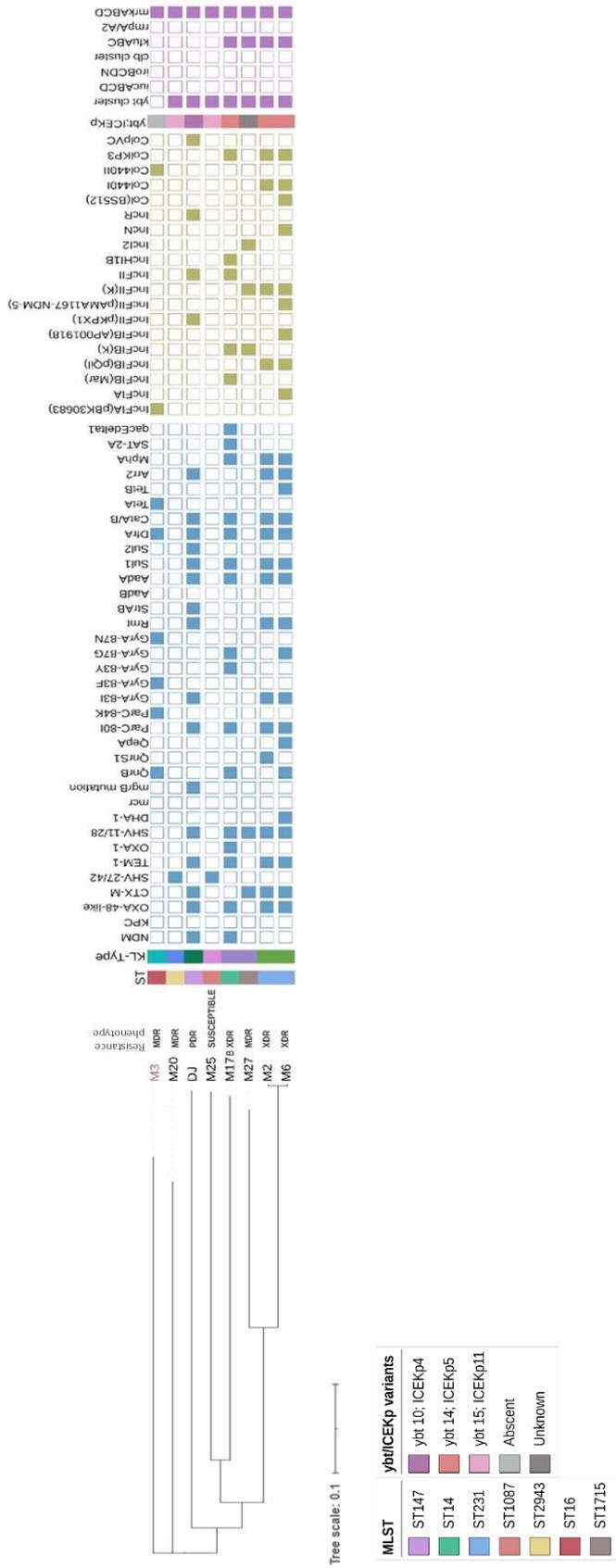
Table 4.19a Phenotypic and genomic analysis of the eight *Kp* isolates.

Strain	ST	Wzi	KL-locus	Antibiotic susceptibility testing	Carbapenemases	ESBLs/pAmpCs	Other $\beta$ -lactamases	Intrinsic $\beta$ -lactamases	Aminoglycosides
<b>DJ</b>	ST147	64	64	AMC, TCC, PIP, PIP-TAZ, CXM, CTX, CAZ, FEP, ERT, AZT, AMK, GEN, TOB, NAL, OFL, CLO, SXT, TMP, COL	NDM-5, OXA-181	CTX-M-15	TEM-1B	SHV-11	RmtF, RmtB, StrAB, AadA2
<b>17B</b>	ST14	2	2	AMC, TCC, PIP, PIP-TAZ, CXM, CTX, CAZ, FEP, ERT, AZT, AMK, GEN, TOB, NAL, OFL, CLO, SXT, TMP	NDM-1, OXA-232	-	TEM-1B, OXA-1	SHV-28	AadA2
<b>2</b>	ST231	104	51	AMC, TCC, PIP, PIP-TAZ, CXM, CTX, CAZ, FEP, ERT, AZT, AMK, GEN, TOB, NAL, OFL, CLO, SXT, TMP	OXA-232	CTX-M-15	TEM-1B	SHV-1	RmtF, AadA2
<b>6</b>	ST231	104	51	AMC, TCC, PIP, PIP-TAZ, CXM, CTX, CAZ, FEP, ERT, AZT, AMK, GEN, TOB, NAL, OFL, CLO, SXT, TMP	OXA-232	CTX-M-15, DHA-1	TEM-1B	SHV-11	RmtF, AadA2
<b>27</b>	ST1715	2	2	AMC, TIC, TCC, PIP, CXM, CTX, CAZ, FEP, AZT, TOB	-	CTX-M-15	-	SHV-1	-
<b>20</b>	ST2943	39	NP	AMC, TIC, TCC, CXM, CTX, CAZ, AZT, GEN, TOB, NAL, OFL	-	SHV-27	-	-	-
<b>3</b>	ST16	50	NP	AMC, TIC, TCC, PIP, CXM, CAZ, FEP, AZT, GEN, NAL, OFL	-	-	-	SHV-1	-
<b>25</b>	ST1087	364	NP	AMC, TIC, TCC, NAL	-	-	-	SHV-42	-

NP, not predicted; AMC- Ampicillin-sulbactam; PIP- Piperacillin; PIP-TAZ- Piperacillin-tazobactam; TCC- Ticarcillin-clavulanic acid; CTX- Cefotaxime; CXM- Cefuroxime; FEP- Cefepime; CAZ- Ceftazidime; AZT- Aztreonam; ERT- Ertapenem; AMK- Amikacin; GEN- Gentamycin; TOB- Tobramycin; NAL- Nalidixic acid; OFL- Ofloxacin; CLO- Clotrimazol; SXT- Trimethoprim-sulfamethoxazole; TMP- Trimethoprim; COL- Colistin

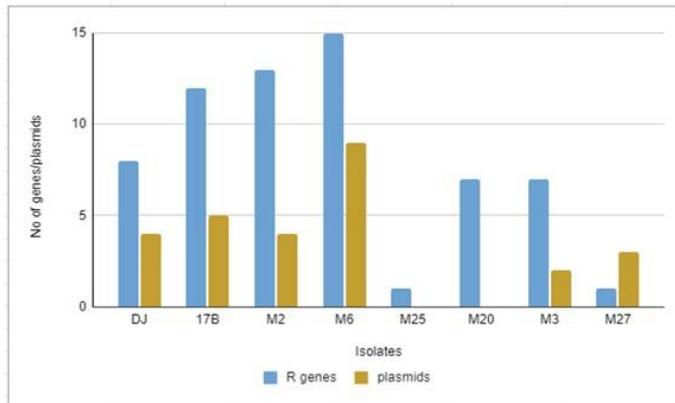
Table 4.19b Phenotypic and genomic analysis of the eight *Kp* isolates.

Strain	Fluoroquinolones	Colistin	Sulfonamides	Trimethoprim	Other resistance genes	Plasmids	Yersiniabactin	Other virulence genes
<b>DJ</b>	<i>parC</i> -80I, <i>gyrA</i> -83I	$\Delta mgrB$	<i>sul1</i> , <i>sul2</i>	<i>dfrA12</i>	<i>catA2</i> , <i>Arr2</i>	IncFII, IncFII(pKPX1), IncR, ColpVC	ybt10;ICEKp4	<i>mrk</i> ABCFHIJ
<b>17B</b>	<i>parC</i> -80I, <i>gyrA</i> -83Y, <i>gyrA</i> -87G, <i>qnrB1</i>	-	<i>sul1</i>	<i>dfrA12</i> , <i>dfrA14</i> , <i>dfrA1</i>	<i>catB4</i>	IncHII1B + IncFIB(Mar), IncFII, IncFIB(K), ColKP3	ybt14;ICEKp5	<i>mrk</i> ABCFHIJ, <i>kfu</i> ABC
<b>2</b>	<i>parC</i> -80I, <i>gyrA</i> -83I, <i>qnrS1</i>	-	<i>sul1</i>	<i>dfrA12</i>	<i>catA1</i> , <i>Arr2</i>	IncFIB(pQil), IncFII(K), ColKP3, Col440I	ybt14;ICEKp5	<i>mrk</i> ABCFHIJ, <i>kfu</i> ABC
<b>6</b>	<i>qnrB4</i>	-	<i>sul1</i>	<i>dfrA12</i>	<i>catA1</i> , <i>Arr2</i> , <i>tetB</i>	IncFIB(pQil), IncFII(K), IncFIB(AP001918), IncFIA, IncFII(pAMA1167), IncN, ColKP3, Col440I, Col(BS512)	ybt14;ICEKp5	<i>mrk</i> ABCFHIJ, <i>kfu</i> ABC
<b>27</b>	-	-	-	-	-	IncFIB(K), IncFII(K), IncI2	ybt unknown	<i>mrk</i> ABCFHIJ, <i>kfu</i> ABC
<b>20</b>	-	-	-	-	-	-	ybt15;ICEKp11	<i>mrk</i> ABCFHIJ
<b>3</b>	<i>parC</i> -84K, <i>gyrA</i> -83F, <i>gyrA</i> -87N, <i>qnrB1</i>	-	-	<i>dfrA14</i>	<i>tetA</i>	IncFIA(pBK30683), Col440II	-	<i>mrk</i> ABCFHIJ
<b>25</b>	-	-	-	-	-	-	ybt15;ICEKp11	<i>mrk</i> ABCFHIJ

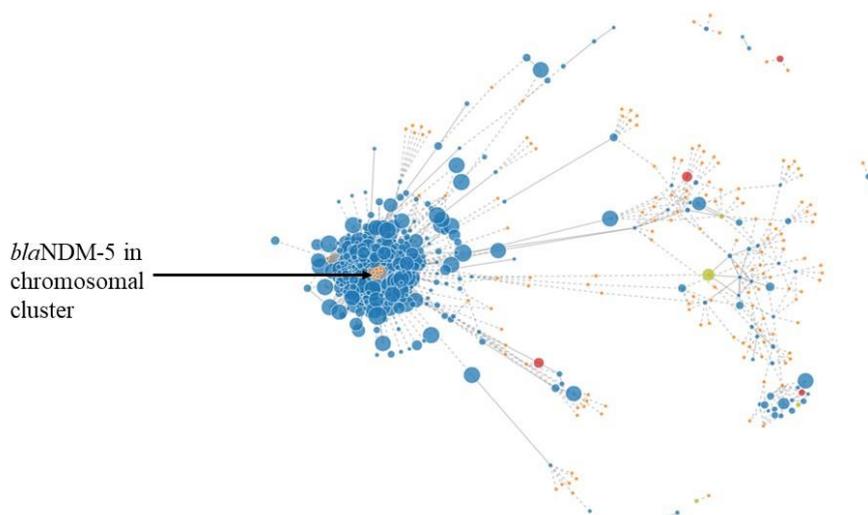


A

B



**Figure 4.11 Comparative analysis of genomic features of PDR, XDR, MDR and susceptible isolates.** (A) Analysis of phylogenetic and genomic features of  $n=8$  whole genome sequences. Branch length of the tree indicates the divergence (substitution/site) of the genomes. Resistance phenotype and genotype of the isolates along with virulence genes are shown in the figure. Empty boxes only with the outline indicates absence, and color-filled boxes indicates the presence of the respective genes (names are shown on the top). Legends for the strip data-type are also included in the figure. The was used to generate the phylogenetic tree and iTOL v5 was used to exhibit the genotypes. Maximum-likelihood phylogenetic tree was generated using IQ-TREE v1.6.11. The branch length indicates the divergence of the genomes. (B) Number of resistance genes and plasmids present in the genomes of XDR MDR PDR and susceptible isolates.



**Figure 4.12 Location of *bla*<sub>NDM-5</sub> using PLACNET.** Location of *bla*<sub>NDM-5</sub> was analysed using PLACNET v3 based on Illumina sequence. The big sized cluster in blue color consists of chromosomes present in the genome. The small sized clusters in orange color consist of plasmids present in the genome. Each tiny circles inside the

clusters represent genes belonging to that cluster. The black arrow shows the location of *bla<sub>NDM-5</sub>* gene inside the chromosomal cluster.

Interestingly, in this isolate, *bla<sub>NDM-5</sub>* carbapenemase-producing gene was found to be present on the chromosome instead of the plasmid. Upon analysis using Illumina sequence, *bla<sub>NDM-5</sub>* was detected in the chromosomal cluster (blue colored, shown by arrow) and not in the plasmid clusters (figure 4.12). The location of *bla<sub>NDM-5</sub>* in DJ was further confirmed using long read oxford nanopore NGS platform.

#### 4.3.7 High-resolution genomic analysis of pan-drug resistant strain DJ using Illumina and Nanopore NGS

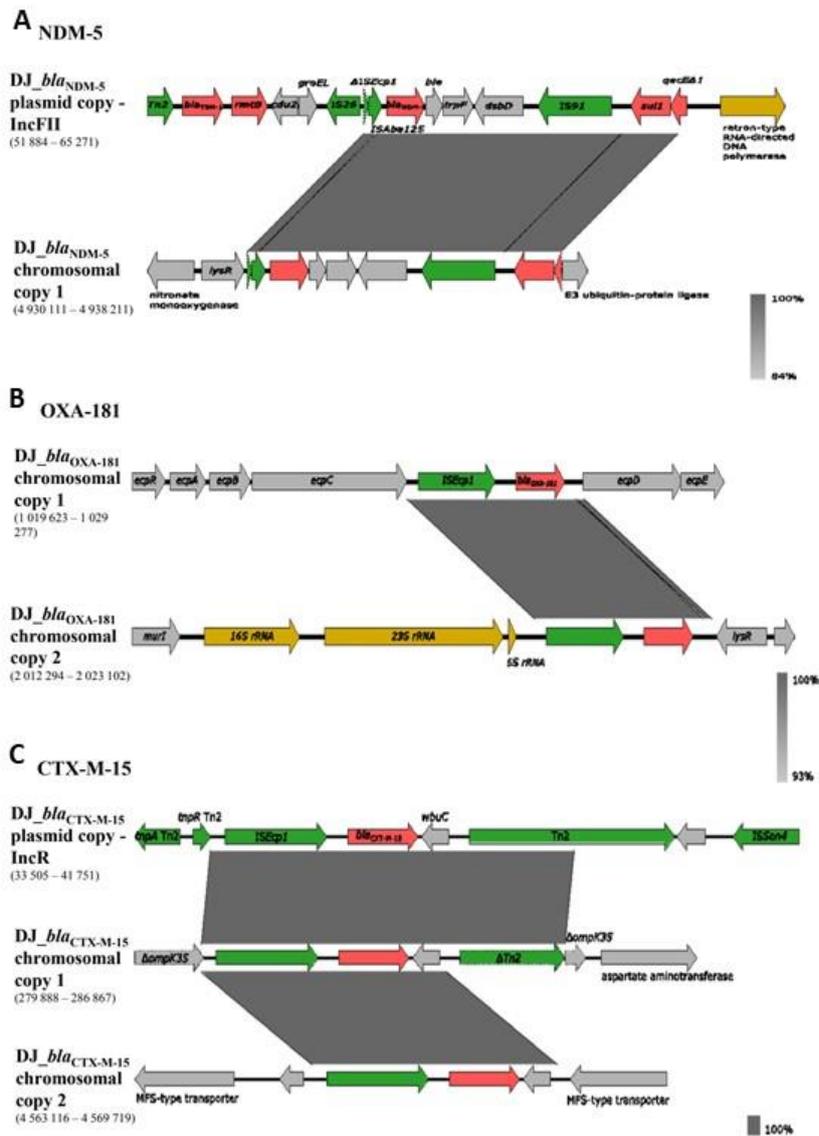
Apart from Illumina, oxford nanopore sequencing was also performed for isolate DJ to confirm the location of *bla<sub>NDM-5</sub>* and further investigate the genomic features of this PDR strain. The genome assembly (5.7 Mbp; 56.9% GC) comprised a total of 9 contigs, 7 of which were circularized. Strain DJ possessed four carbapenemase genes, corresponding to two copies of each *bla<sub>NDM-5</sub>* and *bla<sub>OXA-181</sub>*. In addition, there were three copies of *bla<sub>CTX-M-15</sub>*.

One or two copies of *bla<sub>NDM-5</sub>*, *bla<sub>OXA-181</sub>*, and *bla<sub>CTX-M-15</sub>* were integrated either in the chromosome (one or two copies of each) or in plasmids; namely one copy of *bla<sub>CTX-M-15</sub>* in the 57 kb IncR plasmid, and *bla<sub>NDM-5</sub>* in the non-circularized 84 kb-IncFII plasmid contig (figure 4.13A and C). Figure 4.13 shows these beta-lactamase genes clusters present on chromosome and plasmid of DJ. One copy of *bla<sub>NDM-5</sub>* was present on IncFII plasmid. IncFII plasmid carried *bla<sub>TEM-1</sub>* (upstream of *bla<sub>NDM-5</sub>*) and *sul-1* (downstream of *bla<sub>NDM-5</sub>*) along with *bla<sub>NDM-5</sub>*. Importantly, *bla<sub>NDM-5</sub>* was found to be located adjacent to truncated ( $\Delta$ ) *ISEcp1/IS26* between *ISAbal25* and Bleomycin resistance genes. The genetic arrangement of *bla<sub>NDM-5</sub>* present on chromosome was the same as the *bla<sub>NDM-5</sub>* present on IncFII plasmid. *bla<sub>NDM-5</sub>* was located between *ISAbal25* and Bleomycin resistance genes;  $\Delta$ *ISEcp1* and *IS26* were also present adjacent to *bla<sub>NDM-5</sub>* on the chromosome. However, *bla<sub>TEM-1</sub>* and Tn2 were not found on the chromosomal copy (figure 4.13A).

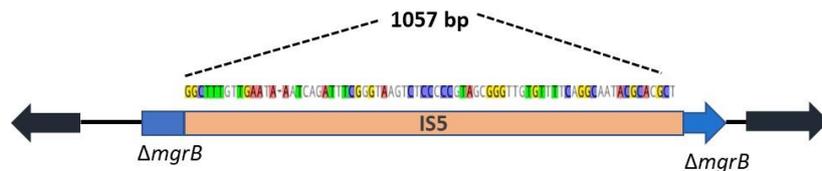
In case of *bla*<sub>OXA-181</sub>, we found two copies of *bla*<sub>OXA-181</sub> in DJ, both copies were present on chromosome instead of plasmid. Both the copies of *bla*<sub>OXA-181</sub> were present adjacent to *ISEcp1* on chromosome (figure 4.13B). Three copies of *bla*<sub>OXA-181</sub> were found in DJ; one copy was present on IncR plasmid and two copies were present on chromosome. In case of the copy on IncR plasmid, *bla*<sub>CTX-M</sub> was present between *ISEcp1* and *Tn2*; *wbuC* was present adjacent to *bla*<sub>CTX-M-15</sub> (between *bla*<sub>CTX-M-15</sub> and *Tn2*). In one of the chromosomal copies of *bla*<sub>CTX-M-15</sub>, *ISEcp1*-*bla*<sub>CTX-M-15</sub> (along with *wbuC* and partial sequence of *Tn2* ( $\Delta$ *Tn2*)) was observed to be inserted into *OmpK35* gene present on the chromosome, led to insertional inactivation of *OmpK35* (chromosomal copy-1; figure 4.13B). In another copy of *bla*<sub>CTX-M-15</sub> present on the chromosome (chromosomal copy-2; figure 4.13B), *ISEcp1* and *bla*<sub>CTX-M-15</sub> were found to be inserted in MFS-type transporter gene on chromosome along with *wbuC*. However, *Tn2* was not found to be present along with this copy. Apart from these, an insertion of 6 nucleotides (nt) in *ompk36* was also observed; resulting in a protein of 369 amino acid (aa) instead of 367 aa.

Mutation in *mgrB* was confirmed as *mgrB* was found to be disrupted by an IS5 transposase of 1057 bp length (figure 4.14). Strain DJ also carried *rmtF*, *rmtB*, *strA*, *strB* and *aadA2* genes, responsible for resistance to aminoglycosides, including amikacin. Mutations in Quinolone resistance determining regions (QRDRs) of both *gyrA-83I* and *parC-80I* were found. A premature stop codon caused by a A580T substitution was found in *ramR* (a negative regulator of RamA).

Regarding the virulence gene content of strain DJ, yersiniabactin gene cluster (*ybt10*), located on an ICE*Kp4* mobile genetic element was detected. The largely *Kp*-conserved *mrk* and *fim* gene clusters responsible for type-1 and type-III fimbriae, respectively were found. However, no gene associated with hypervirulence (*rmpA*, aerobactin, salmochelin) was carried by DJ.



**Figure 4.13 Schematic representation of *bla* gene clusters found in DJ using nanopore NGS.** (A) Copy of *bla*<sub>NDM-5</sub> present on IncFII plasmid located between ISAbal25 and Bleomycin resistance genes; ISecp1/IS26 present upstream of *bla*<sub>NDM-5</sub>. Chromosomal copy 1 of *bla*<sub>NDM-5</sub> was present on the chromosome due to insertion of ISecp1/IS26 into the chromosome. (B) Two the copies of *bla*<sub>OXA-181</sub> were present adjacent to ISecp1 on chromosome. (C) Three copies of *bla*<sub>CTX-M-15</sub> were found in DJ; one copy was present on IncR plasmid between ISecp1 and Tn2. In case of two chromosomal copies, ISecp1-*bla*<sub>CTX-M-15</sub> were inserted into OmpK35 and MFS-type transporter genes, respectively. EasyFig v2.2.5 was used to generate the figure.



**Figure 4.14 Disruption of *mgrB*.** Schematic representation of disruption of *mgrB*. 1057 bp insertion of IS5 transposase in *mgrB* of DJ.

#### 4.3.8 Analysis of global population of *Kp* sublineage CG147

In order to compare strain DJ with CG147 genomes worldwide and to investigate global epidemiological emergence as well as evolution of sublineage CG147, we performed comparative genomics using the 217 publicly available genomes and **genome of our isolate, DJ**. Figure 4.15 iTOL shows the comparative genomic analysis of 217 publicly available genome of CG147 along with strain DJ (shown with black arrow). Figure 4.16 shows the prevalence of different ST-KL clade at various geographical locations. The isolates were collected between 2002 and 2018. Majority of the isolates were from human samples [90%, 196/218 (196 out of 218)] and hospital environment (6%, 12/218). Most isolates were collected in Europe (40%), Southeast Asia (17%), South Asia (11%) and North America (10%).

The phylogenetic analysis (figure 4.15) revealed the existence of three main clades, each corresponding to a single MLST sequence type: ST147, ST273 and ST392; the two latter are single-locus variants of ST147, differing at locus *tonB*. The phylogenetic tree indicates that ST147 is the most recently evolved, followed by ST392. S273 was evolved long before ST147 and ST392. ST147 is the most represented (79%, 172/218 genomes) and widespread ST. ST392 has also been detected worldwide, but more extensively in Europe (29%) and America (38%). Four main branches characterized by distinct capsular loci were observed in CG147 genomes: KL64, KL10, KL27 and KL74. In case of ST147 genomes, KL64 (74%; 128/172) and KL10 (20%, 34/172) were found to be dominant. The ST147-KL64 clade comprised genomes were mainly from Europe (54%) and South Asia (15%). ST392 was found to carry only KL27 capsular gene cluster. In ST273, KL74 was the predominant capsular gene

cluster (50%), although in 42% of the genomes the *cps* cluster was not detected (figure 4.15; figure 4.16).

#### 4.3.8.1 Acquired resistance genes and their evolutionary dynamics within CG147

Numerous Acquired resistance genes were found among CG147 genomes ranging from 2 to 23 genes (Average: 13 genes). Regarding beta-lactam resistance, 83% of CG147 isolates carried a *bla*<sub>CTX-M</sub>. 63% of CG147 genomes carried at least one carbapenemase gene; only 37% did not harbour any known carbapenemase gene. Figure 4.17 shows the global distribution of the carbapenemase genes carried CG147 genomes. It showed that *bla*<sub>NDM</sub> was predominant in Southeast Asia and North America, whereas the combination of *bla*<sub>NDM</sub> and *bla*<sub>OXA-48-like</sub> was more frequently observed in South and East Asia. *bla*<sub>OXA-48</sub>, *bla*<sub>KPC</sub> and *bla*<sub>VIM</sub> were found to be dominated in Europe. Interestingly, few associations between resistance genes and CG147 sublineage were observed. ST147-KL10 (*wzi*420, O3a) was almost exclusively found in Asia (85%), mainly in Southeast and South Asia as well as Eurasia (figure 4.16); it was associated with *bla*<sub>CTX-M</sub>, *bla*<sub>OXA-48</sub> or *bla*<sub>OXA-232</sub>, and more occasionally with *bla*<sub>NDM-1/-9/-4</sub> and *bla*<sub>KPC-2</sub>. Importantly, all CG147 genomes presented QRDR alterations in *gyrA* (83I or 83Y and 87A) and *parC*-80I (figure 4.15). QRDR mutations observed among ST147-KL10 (*wzi*420, O3a) was in *gyrA*, 83Y instead of 83I (resulting of the nucleotide change TAC instead of ATC) and additionally 87A. Although less frequently associated with carbapenemases (35%), an association between *bla*<sub>CTX-M</sub> and *qnrB1* was identified among ST273.

Two types of mutations in OmpK35 were observed in ST147-KL64 genomes with *ybt*10;*ICEKp4* (subclade-1) and *ybt*16/*ICEKp12* (subclade-2). (1) n=18 (including strain DJ) ST147-KL64 genomes with *ybt*10;*ICEKp4* showed disruption of OmpK35 by *ISEcp1-bla*<sub>CTX-M</sub> (figure 4.15). (2) n=29 ST147-KL64 genomes with *ybt*16;*ICEKp12* exhibited insertion of 2 nucleotide (nt) resulting in a premature stop codon in OmpK35. Both the subclades were also harboring mutation in OmpK36 by insertion of 6 nt resulting in a protein of 369 amino acid (aa) instead of 367 aa. Interestingly, these two subclades also

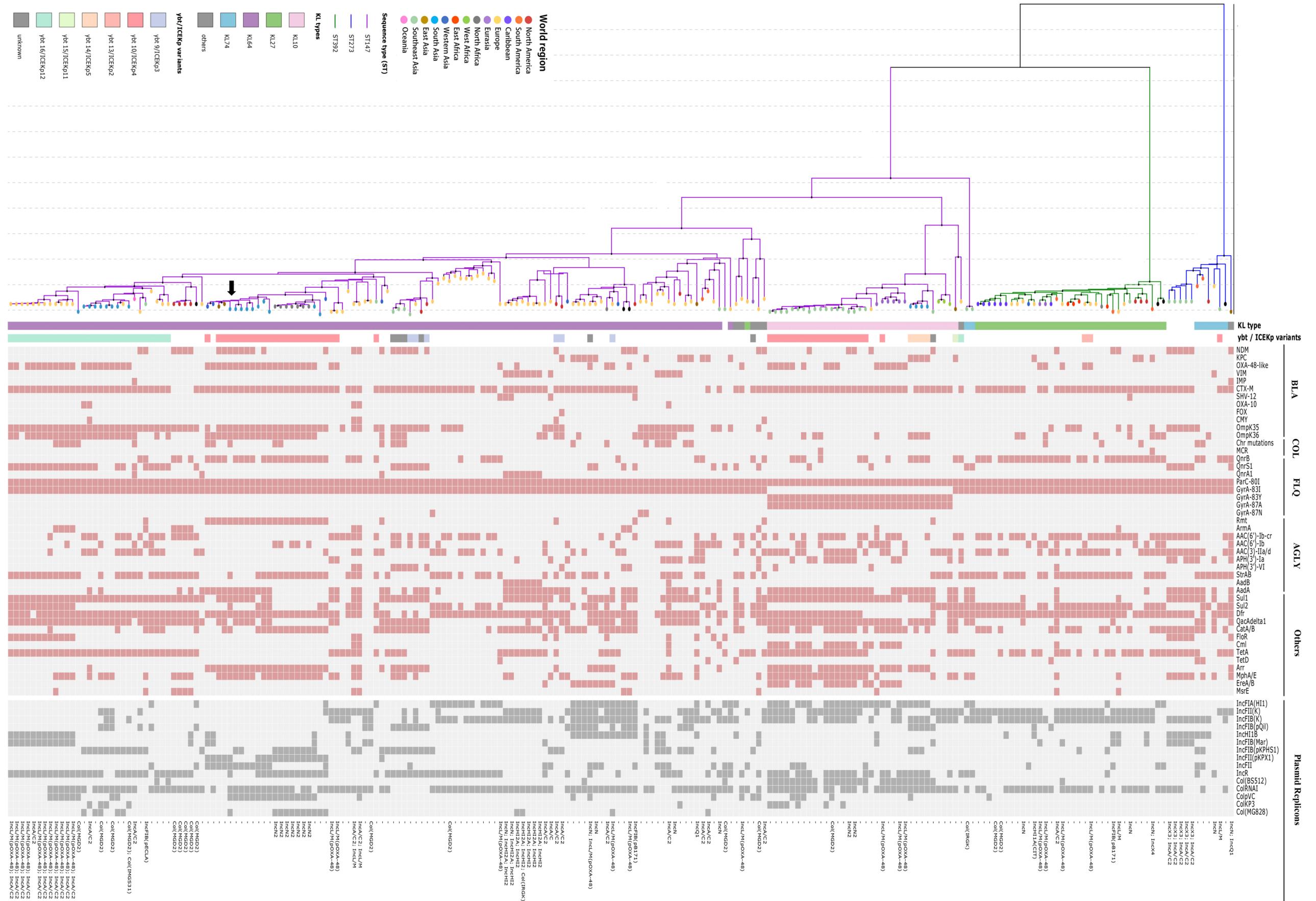
revealed differences in their plasmid replicon content. ST147-KL64-ybt16;ICEKp12 (subclade-2) (n=29) were enriched in IncHIB/IncFIB(Mar) (41%) and IncR (88%), whereas ST147-KL64-ybt10/ICEKp4 (subclade-1) (n=18) were enriched in IncFII (pKPX1) (78%), IncFII (56%) and IncR (61%). These results show rapid plasmid dynamics/turnover at small time scales within ST147 genomes. The remaining ST147-KL64 (n=81) were enriched in IncFIB<sub>K</sub> (67%), IncFII<sub>K</sub> (52%) (common pKPN-3-derived plasmids found in *K. pneumoniae* harbouring *pco* and *sil* clusters) (Rodrigues et al., 2014; Navon-Venezia et al., 2017) and IncFIA(HI1) (40%).

ybt;ICEKp was rarely detected amongst ST392 and ST273 genomes, this virulence factor was observed in 53% ST147 genomes. Despite of a high diversity of ICE observed among ST147 isolates, ybt16;ICEKp12 and ybt10;ICEKp4 were two predominant variants found in ST147 genomes.

#### 4.3.8.2 Association of resistance genes with virulence and other genetic elements

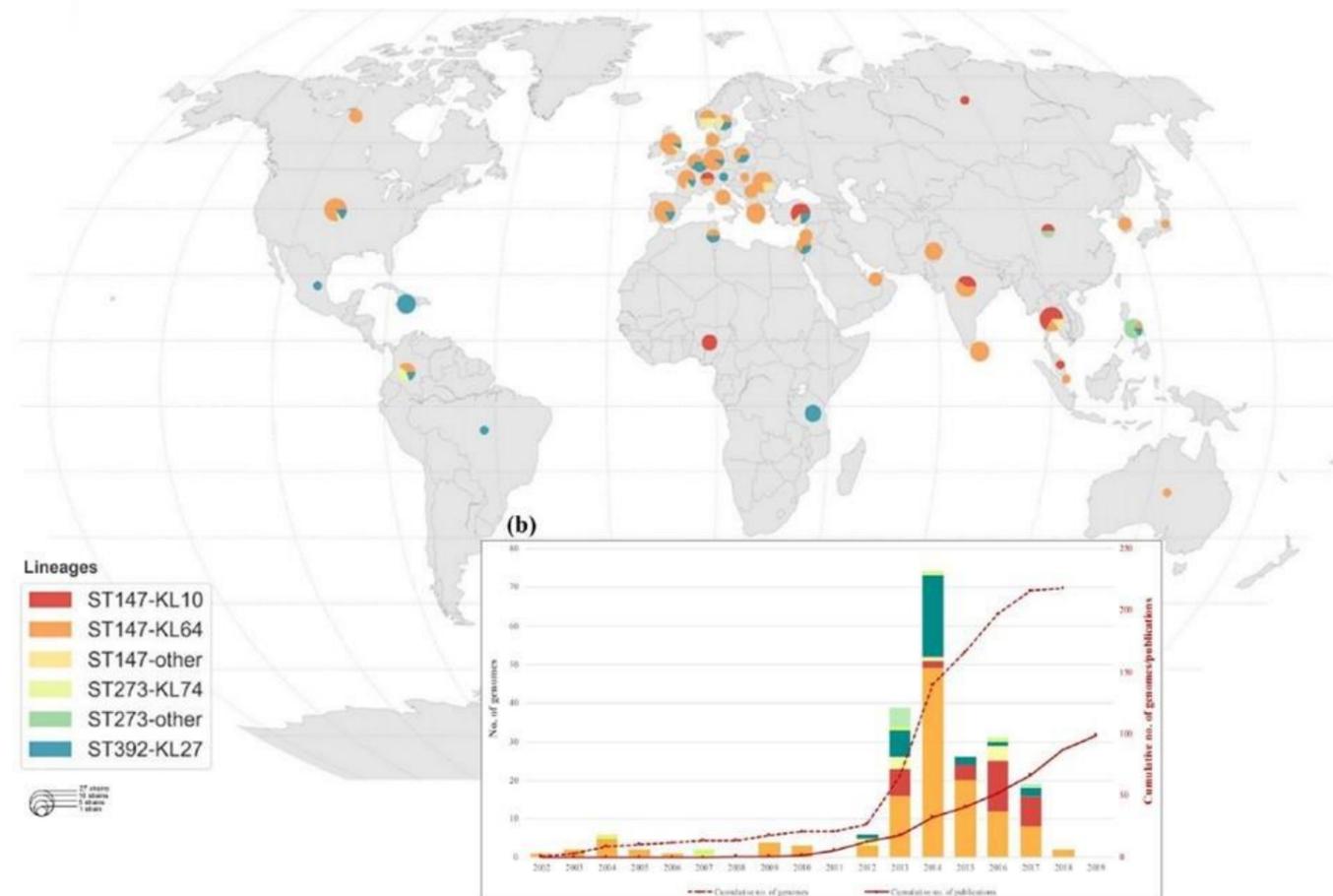
Regarding the convergence between MDR and hypervirulent genotypes (presence of *rmpA* and/or *rmpA2* and aerobactin), it was observed in two genomes, B-8658 (ST147-KL10, 2014) and KpvST147L (ST147-KL14, 2016), from Russia and United Kingdom, respectively.

Some networks of acquired resistance genes and plasmids were detected within CG147. For example, Col(BS512), ColKp3, IncFIA (HI1), *cml*, *ereA/B* and *gyrA* 83Y/87A co-occurred frequently and were commonly found within the ST147-KL10 lineage. A negative association between IncR and IncFII<sub>K</sub>/IncFIB<sub>K</sub> was detected (figure 4.15), resulting in differences in the content of genes conferring tolerance to copper and silver.

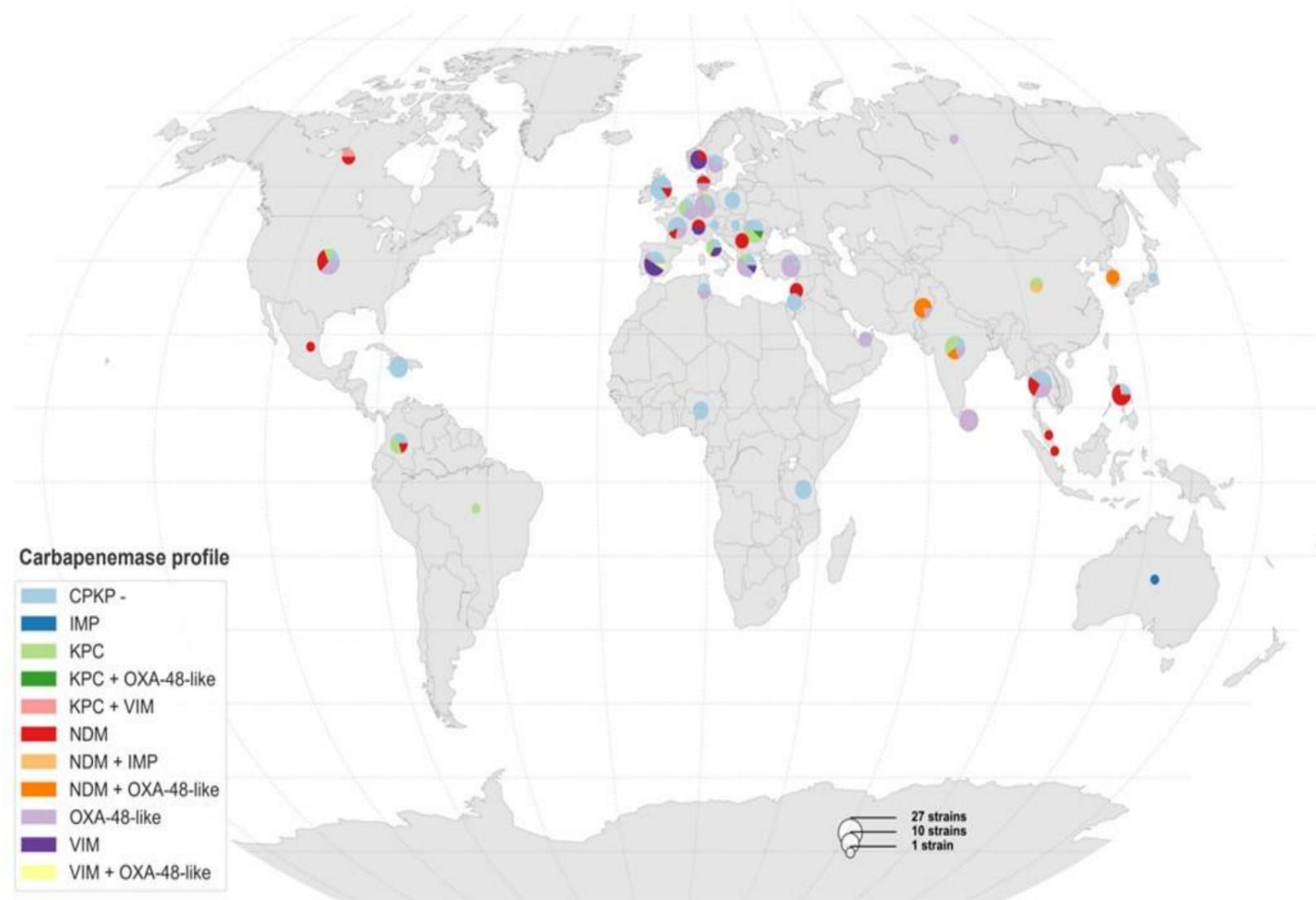


**Figure 4.15 comparative genomic analysis of CG147.** Phylogenetic tree (left side) along with presence of resistance genes and plasmids in n=245 *Kp* genomes is shown in the figure. Maximum-likelihood phylogenetic tree was generated using IQ-TREE v1.6.11. Branch length of the tree indicates the divergence (substitution/site) of the genomes. Color of the branches and clades belong to —

respective ST; Different colors in the strip represent specific KL-type and ybt/ICEKp variant (as per the legends). Circles at the end of each leaves represent geographical origin of that genome. Name of resistance genes and plasmids are mentioned on the top. More than one replicons present in the genomes are mentioned in text form (to the right end). Empty boxes indicates absence and color filled boxes indicates presence of respective genomic gene/plasmid. iTOL v3 was used to exhibit the genetic profile of the genomes.



**Figure 4.16 Prevalence of ST-KL clade and number of genome publications.** (A) prevalence of different ST-KL clade at various geographical locations. Each color belong to respective ST-KL clade and the diameter of the pie-chart represent number of strains (as per the legends). (B) Number of genome publications from 2002 to 2019. Dotted line in the graph shows cumulative number of genomes collected and deposited worldwide and continuous line shows number of publications of the genomes.



**Figure 4.17 Geographical locations and associated carbapenemase profile.** The pie chart represents the frequency of each carbapenemase in respective country. CPKP: carbapenemase-producing *Klebsiella pneumoniae*. The diameter of the pie-chart represents number of strains (as per the legends).

### 4.3.9 Global analysis of ST231 *Kp* genomes

To understand the genomic features of *Kp* ST231 lineage and compare our isolates M2 and M6 with the ST231 genomes worldwide, we performed comparative genomic analysis using n=95 publicly available genomes of ST231 lineage, collected between 2010 to 2018. Figure 4.18 shows the comparative genomic analysis of total n=97 ST231 genomes including M2 and M6.

Phylogenetic analysis revealed that genomes of our two isolates M2 and M6 are very closely related to each other genetically. Both the genomes were relatively closer to VBP1492 (2018) and PM7109 (2017) genomes from India among all n=97 ST231 genomes. Four genomes deposited in 2017 from India namely, CMC\_VB5505, C15, C13 and BA20567 have been diverged most recently. 89.7% (87/97), 7.3% (7/97) and 1.05% (1/97) isolates were from human, hospital environment and animal origin, respectively. Most of the genomes 72.2% (70/97) genomes were collected during 2016 and 2017, and shared a similar genetic profile. Unlike ST147, the diversity in KL-type in ST231 was not very high; KL-51 is the most prevalent K-locus type found in ST231 (97.9%; 95/97). Only two isolates from Vietnam (VD412) and Australia (MSB1\_8A), respectively were found to have KL-64. Regarding the geographical spread of the isolates, a large number of isolates were mainly collected from India (43.3%, 42/97) and Thailand (26.8%, 26/97); 7.2% (7/97) genomes were from Pakistan. Hence, Asia seems to be the continent with highest prevalence of ST231 *Kp* isolates (77.3%, 75/97), followed by Europe with 11.3% (11/97) isolates.

Regarding carbapenemase genes acquired by ST231 genomes, 79.4% (77/97) of the genomes carried at least one of the known carbapenemase genes. Only 18.5% (18/97) of genomes did not carry any known carbapenemase genes. Importantly, *bla*<sub>OXA-48-like</sub> found to be highly prevalent in ST231 as 77.3% (75/97) genomes were harbouring *bla*<sub>OXA-48-like</sub>, mainly *bla*<sub>OXA-232</sub>. *bla*<sub>OXA-48-like</sub> was observed in an Indian isolate in 2014 for the first time; however, it was not found in isolates from Nigeria, Greece and Portugal collected in the same year. *bla*<sub>NDM</sub> was found to be present in 3.1% (3/97) of the isolates; two isolates from USA (from human blood) and one isolate from India (from urine sample) in 2017. *bla*<sub>NDM</sub> was not seen in ST231 sublineage after 2017. *bla*<sub>KPC</sub> was found only in one isolate from Portugal (ybt9; ICEKp3) in 2014 and not seen in any genome after 2014. *bla*<sub>VIM</sub>, *bla*<sub>IMP</sub> were not found in any isolates. Coexistence of carbapenemase genes such as *bla*<sub>NDM</sub>, *bla*<sub>KPC</sub>

and *bla*<sub>OXA-48-like</sub> gene was not found in any genome. In case of ESBLs, 51.5% (50/97) genomes were found to carry *bla*<sub>CTX-M-15</sub>. 14.4% (14/97) genomes carried *bla*<sub>SHV-12</sub>; these genomes were mainly from India (12/14 genomes) (figure 4.18) DHA-1, responsible for resistance to third-generation cephalosporins was found to be present only in our isolate M6 among all ST231 genomes.

Figure 4.19 shows the geographical distribution of carbapenemase and ESBL genes. *bla*<sub>OXA-48-like</sub> genes were observed to be more prevalent in South and Southeast Asia (India, Pakistan and Thailand), whereas *bla*<sub>CTX-M</sub> disseminated globally with its presence in all 5 continents such as Asia, South America, North America, Africa, Australia and Europe. Genomes from Brazil, Tanzania, Nigeria and Australia carried only *bla*<sub>CTX-M-15</sub>. North America was observed to be highly diverse in terms of carbapenemase and ESBL genes, whereas Africa and South America were found to have *bla*<sub>CTX-M</sub> only (figure 4.19A and B). Highest number of genomes (n=31) carrying *bla*<sub>OXA-48-like</sub> and *bla*<sub>CTX-M-15</sub> (n=26 genomes) were deposited from India; *bla*<sub>SHV</sub> was observed mainly in Indian genomes (n=12). *bla*<sub>NDM</sub> was found to be present only in genomes from USA (n=2) and India (n=1) (figure 4.18 and 4.19). Carriage of *bla*<sub>NDM</sub> by ST231 genomes is lesser compared to ST147 population, where *bla*<sub>NDM</sub> was observed to be carried by genomes from 4 continents except Africa (figure 4.18, figure 4.15).

Regarding other acquired resistance genes, average 13 resistance genes were found to be present in ST231 genomes ranging from 6 to 15 resistance genes in some genomes. Interestingly, mutations in *parC80I* and *gyrA83I* were found in all genomes of ST231 sublineage along with the presence of *qnrS1* and *qnrB* in 50.5% (49/97) and 10.3% (10/97) of the genomes, respectively. These QRDR mutations indicate that it is an ancestral characteristic of this sublineage. Apart from these, mutation in *GyrA-87G* was found in 6.2% (6/97) genomes from Asian countries (India, Pakistan and Thailand); *GyrA-87N* in two genomes from Switzerland and one from India; *GyrA 87Y* in one genome from each Australia, Pakistan and Vietnam. Concerning colistin resistance, only one isolate from Thailand in 2015 was found to carry *mcr* gene; however, it did not carry any carbapenemase genes. *mcr* did not found in ST231 lineage after 2015. 55.7% (54/97) isolates carried *rmt* gene, responsible for aminoglycoside resistance. *tetD* responsible for tetracycline resistance only found in genomes from Portugal, whereas, *tetB* was present in six genomes including M6. *aadA*, *dfp* and *sul1* genes seem to be linked as their coexistence was observed in 78.3% (76/97) of the genomes. *sul1* and *sul2* were co-occurred only in

6.2% (6/97) genomes. Macrolide resistance genes, *mphA* and *ermB* found to be present in 71.2% (70/97) and 69.1% (67/97) isolates, respectively. Coexistence of *mphA* and *ermA/B* found in (57/99) of genomes.

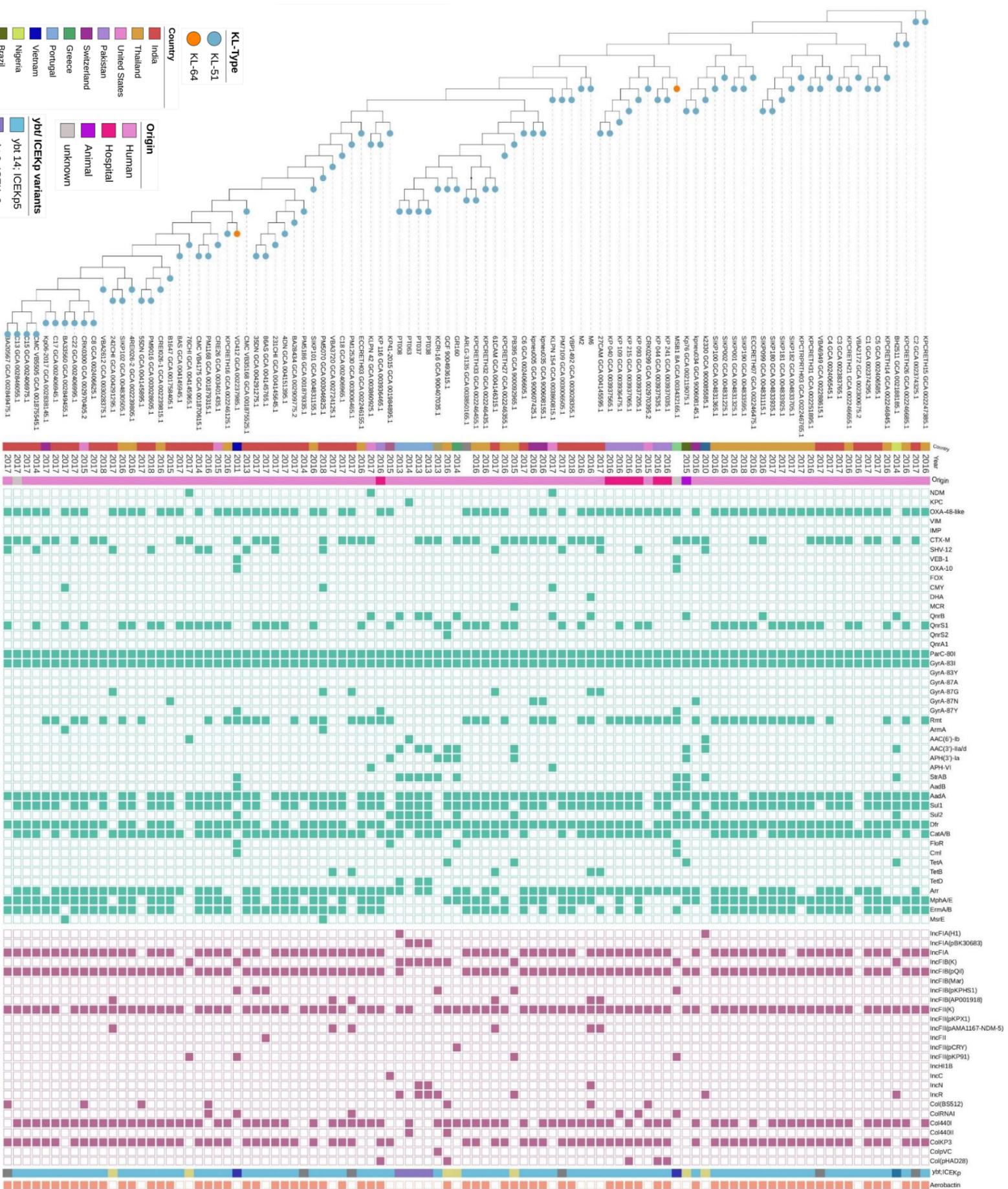
Besides these, coexistence of few other resistance genes was also observed during the analysis. For an instance, 34% (33/97) genomes carried both *bla*<sub>CTX-M</sub> and *bla*<sub>OXA-48-like</sub> genes. Coexistence of *bla*<sub>OXA-48-like</sub> and *rmtF* was found in 51.5% (50/97) of the ST231 genomes, which could be responsible for increased resistance to carbapenem and aminoglycosides. Two genomes from the USA showed coexistence of *bla*<sub>NDM</sub>, *qnrB* and *rmt*, this coexistence is reported to be responsible for increased resistance to carbapenems, quinolones and aminoglycosides.

Regarding other virulence genes, aerobactin genes(*iucABC*) were found to be present in 73.2% (71/97) and yersiniabactin (*ybt*) genes were present in 92.8% (90/97) genomes. *rmpA* and *rmpA2* were not found in any genome of ST231 lineage.

The comparative analysis of resistome, plasmidome and virulence revealed some important networks of plasmids and resistance genes. A negative association between QRDR mutations (*gyrA83I* and *parC80I*) and Col(BS12) plasmid was observed. In 69.1% (67/97) genomes co-occurrence of Col*KP3* and *bla*<sub>OXA-48-like</sub> was observed. Among these, 57.7% (56/97) of *bla*<sub>OXA-232</sub> and Col*Kp3* carrying genomes were also harbouring aerobactin genes (*iucABC*); all belong to *ybt14*;ICE*Kp5* variant. These genomes harbouring carbapenemase and aerobactin producing genes uncover the convergence of hypervirulence and MDR in ST231 lineage. These convergent genomes were mainly from India (21.6%, 21/97) and Thailand (24.7%, 24/97).

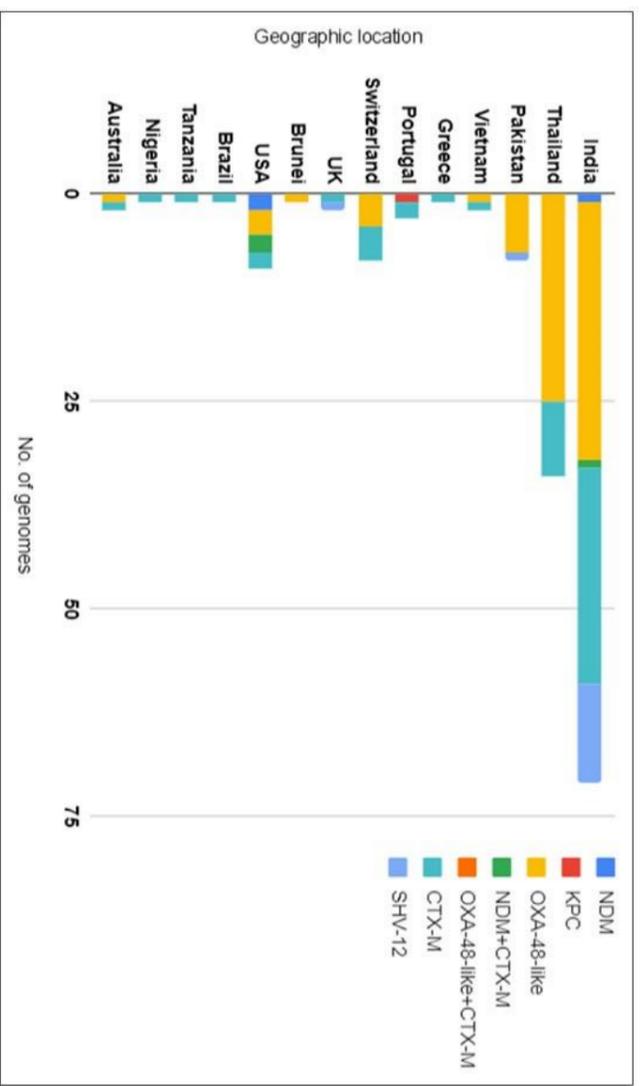
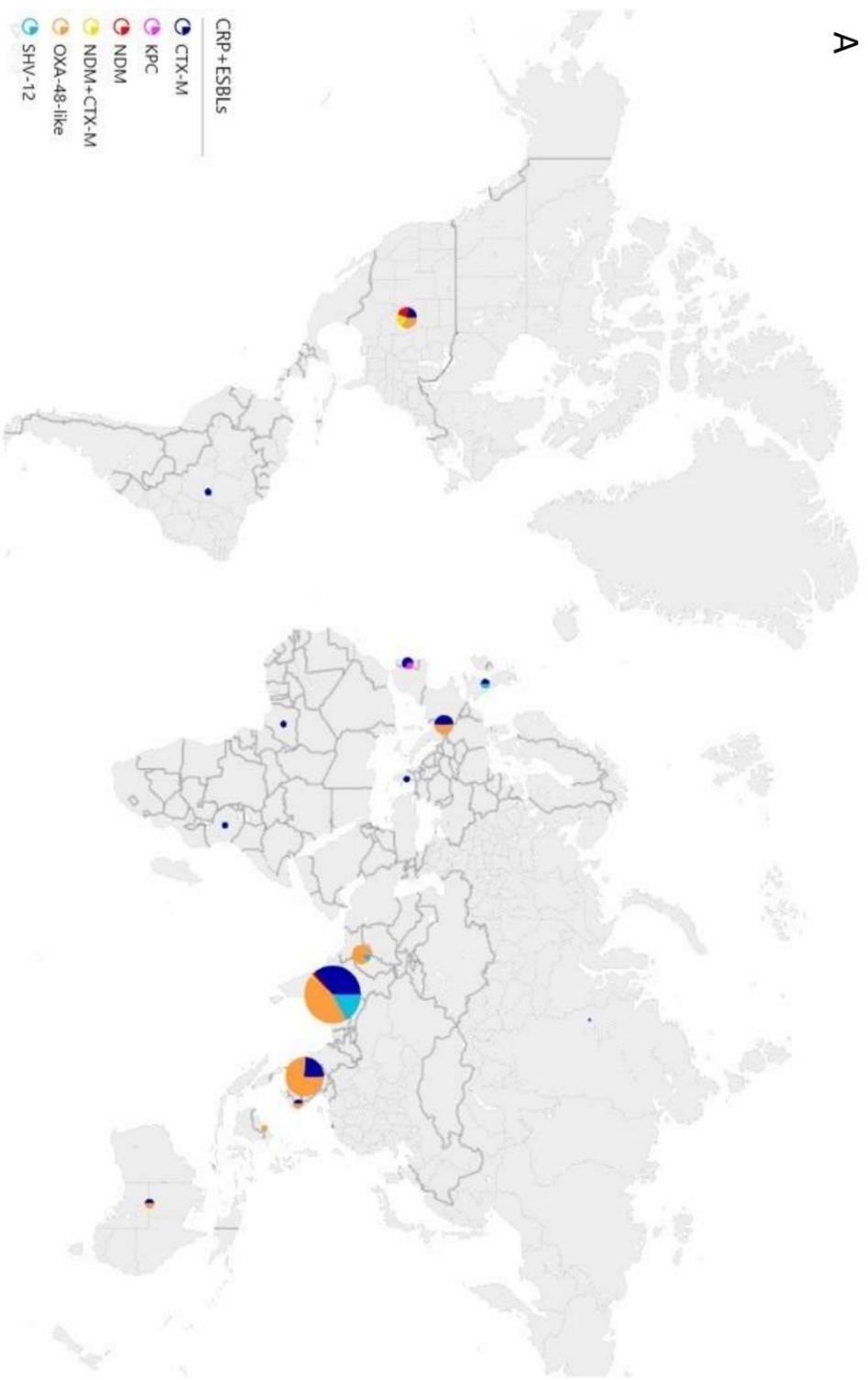
Analysis of the plasmidome also exhibited coexistence of IncFIA, IncFIB (pQil) and IncFII(K) in 76.2% (74/97) of the genomes. A negative association between IncR and IncFII(K)/IncFIB(K) was observed. Coexistence of IncFIA and Col*Kp3* was observed in 93.8% (91/97) of genomes. Interestingly, different combinations of plasmids were observed to be present in different *ybt*;ICE*Kp* variants. *ybt14*;ICE*Kp5* was the most prevalent (79.4%; 77/97) in ST231 lineage. Moreover, all the convergent genomes with both, aerobactin (*iucABCD*) (hypervirulence) and *bla*<sub>OXA-48-like</sub> or *bla*<sub>NDM</sub> (MDR) genes possessed *ybt14*;ICE*Kp5*. The genomes carrying Col*KP3* or MDR plasmid IncFII(PMA1167-NDM-5) also possessed *ybt14*;ICE*Kp5* variant; all genomes containing *ybt14*;ICE*Kp5* carried a specific set of plasmids namely, IncFIA, IncFIB(pQil), IncFII(K)

and Col440I. Further, co-occurrence of FIB(K), FIB(pKPHS1), IncFII(pKp91) and ColRNAI was only found genomes with ybt15;ICEKp11 (n=2, 2/2). ybt10;ICEKp4 was found to be present only in one isolate from Nigeria, which carried different combination of plasmids including IncFII(K), IncFIB(K), IncR and Col440I. IncN was present in only two (out of 4) ybt9;ICEKp3 genomes from Portugal with absence of Col plasmid.



**Figure 4.18. Comparative genomic analysis of ST231 lineage.** Genomic profiles including phylogeny (left side), resistance genes, plasmids and virulence genes of n=97 Kp genomes belong to ST231 are shown in the figure. Maximum-likelihood phylogenetic tree was generated using IQ-TREE v1.6.11. Branch length of the tree indicates the divergence (substitution/site) of the genomes. Different colors in the strip represent respective country, origin source of the isolate, and ybh/ICEKp variant (as per the legends). Circles at the end of each leaves represent KL-type of that genome. Name of resistance genes and plasmids are mentioned on the top. Blue arrows show our genomes of M2 and M6. Empty boxes indicate absence and color filled boxes indicates presence of respective genomic gene/plasmid. The scale bar shown indicates branch length of 0.01 substitution/site. iTOL v3 was used to generate this figure.

A



**Figure 4.19 Prevalence of carbapenemase and ESBL genes associated with ST231 lineage combined with geographical locations.** (A) World map showing the frequency of carbapenemase and ESBL genes in respective countries. (B) Prevalence of genomes harboring various carbapenemase and ESBL genes in different countries.

## 4.4 Discussion

Study of antibiotic resistance in uro-pathogenic *Klebsiella* spp., is crucial because of high prevalence of UTIs in both nosocomial and community settings of developing countries like India (Potron et al., 2013). Moreover, *Kp* (member of *Enterobacteriaceae* family) has been declared as an “urgent threat to public health” (WHO, 2014; CDC, 2013) due to the increasing antibiotic resistance (ABR), including last-resort antibiotics such as carbapenems, colistin and tigecycline. Numerous studies have been reported on antibiotic resistance of *Kp*, however, there is no report on comparison of PDR, XDR, MDR and susceptible *Kp* strains using both phenotypic and genotypic analysis. Here, we showed the comparative analysis of PDR, XDR, MDR and susceptible *Kp* strains to study the factors responsible for a resistance profile of a strain. Further, we have also performed the global analysis of publicly available genomes belong to clonal group (CG) 147 (n=218) and ST231 lineage (n=97) worldwide. PDR strains like isolate DJ, are extremely dangerous and difficult to treat as these strains are resistant to all available antibiotics including last resort drug like colistin and tigecycline; this is a burning issue in clinical settings which leaves clinicians with no available treatment options for *Kp* infections. Hence, we performed the high-resolution analysis of our PDR isolate DJ using both long read (Oxford nanopore) and short read (Illumina) sequencing. Continuous surveillance of XDR strains is also important as these strains remains susceptible only to one or two antibiotics; are on the verge of evolution to become PDR.

Regarding resistance rate, Effah et al., recorded the highest resistance rate of against Cefotaxime (79.2%) followed by aztreonam (73.3%) and cefepime (72.6%); the lowest resistance rate of 2.9% against colistin (Effah et al., 2020). Recently, high-level of resistance (92%) reported against cefotaxime (Karimi et al., 2021). Our results support these findings, as we also found highest resistance in case of extended-spectrum cephalosporins (except penicillins) and low resistance rate against colistin (polymyxin) as well as tigecycline (glycycycline) (figure 4.4). Lower rate of colistin resistance is because of their restricted use in human medicine between 1980s to 2000 (Falagas et al., 2005). Hence, carbapenems were extensively used as the last resort-drug in past years, which led to increased carbapenem resistance in *Kp* isolates. Resistance rate against carbapenems and ESBLs has been worryingly increased in India in past few years. According to report by Praveenet et al., India has witnessed an increase in carbapenem resistance rates from 9% in

2008 to 44% in 2010 (Praveen et al., 2010). A report from tertiary hospital in India reported 24.6% resistance to ESBLs in 2007 (Shahid et al., 2008). In 2017, carbapenem-resistance as high as 44% was reported from India (Pragasam et al., 2017). In 2018 a significant rise in resistance level of ESBLs (45.1-93.1%) and carbapenems (84.1%) were reported from India; the highest resistance of 84.9% was reported against cephalosporins (Odsbu et al., 2018). A report from North India suggested 29.4% resistance rate against carbapenems (Jaggi et al., 2019). Recently, the overall prevalence of multidrug-resistance (MDR) among Indian *Kp* isolates was reported 58.0%. Further they also reported that tigecycline and colistin are the most effective drugs so far (Karki et al., 2021). However, carbapenem and colistin co-resistant XDR (Dey et al., 2020; Pragasam et al., 2017) and PDR (Rodrigues et al., 2021) isolates have also been reported recently from India. Interestingly, despite of exhibition of resistance to all antibiotics by PDR, a recent report suggested that combination of ceftazidime-avibactam could be a treatment option for infective endocarditis by PDR *Kp* (Alghoribi et al., 2021). Also, phages and phage-derived enzymes have earned increasing attention for treatment of CRKP infections in recent years (Li et al., 2021; Shi et al., 2021; Hesse et al., 2021; Haines et al., 2021).

We explored resistance phenotype of an isolate using phenotypic assays for MBL, ESBL, antimicrobial susceptibility testing and qRT-PCR, while WGS approach helped us to understand the genotype responsible for the phenotype of the isolates. The results were in coordination with each other. The susceptible isolate M-25 was negative for ESBL production (according to phenotypic assay for ESBL), because despite the presence of *bla<sub>SHV-42</sub>* in its genome (according to WGS; figure 4.11A), it was not expressed (according to qRT-PCR, figure 4.10). Further, absence of all other antibiotic resistance genes including *bla<sub>CTX-M-15</sub>* (ESBL), *bla<sub>OXA-48-like</sub>*, *bla<sub>NDM</sub>* (MBL) and lack of expression of *bla<sub>SHV-42</sub>* led to its susceptible phenotype. (figure 4.11A and B). In case of presence of *bla<sub>CTX-M-15</sub>* in DJ, M2 and M6 led to resistance to broad range of beta-lactams despite of absence of *bla<sub>SHV</sub>* in their genomes. MDR phenotype of M3 could be explained by the only presence of *qnrB* and QRDR mutations and absence of resistance genes to other antimicrobial categories. M17B exhibited resistance to wide range of beta-lactams including carbapenems because of presence of *bla<sub>OXA-1</sub>*, *bla<sub>SHV-106</sub>* and *bla<sub>TEM-1</sub>*. Coexistence of *bla<sub>NDM</sub>* and *bla<sub>OXA-181</sub>* in DJ and M17B was responsible for increased resistance to carbapenems and other beta-lactams. Co-occurrence of *aadA*, *sulI*, *dfrA* and *catA* or *catB* explains resistance to

trimethoprim/sulfamethoxazole and Phenicol in XDR and PDR isolates. Moreover, co-occurrence of *sulI*, *dfrA* and *aadA* could be because of presence of *sulI* on a class-1 integron which also carry *aadA* and *dfrA* gene cassettes (Antunes et al., 2004). Presence of numerous resistance genes (n=12 to 15) to various antimicrobial categories are responsible for XDR status of M17B, M6 and M2. In case of DJ, only n=8 (lesser in number than XDRs) resistance genes to various antimicrobial categories including carbapenems and other last-generation antibiotics were present. However, disruption of *mgrB* by IS5 transposase resulted in resistance against colistin; mutation (premature stop codon) in *RamR* leading to an overexpression of the AcrAB-TolC efflux system (Villa et al., 2014) was responsible for resistance against Tigecycline (MIC=16 µg/ml). This diverse set of resistant determinants gave rise to the PDR phenotype of DJ. Hence, not only number of resistance genes, but also, type, overall diversity and expression of resistance genes present in genome collectively give rise to a resistance phenotype of an isolate.

Multiplex PCR have been widely used for various purposes such as determination of K1-K2 and K5 serotypes (Compain et al., 2014; Turton et al., 2008), detection of carbapenemase genes (Khorvash et al., 2017; Hong et al., 2012; Poirel et al., 2011); identification of *Kp*, *K. quasipneumoniae* and *K. variicola* as a clinical routine (Fonseca et al., 2017), detection of carbapenem resistant and hypervirulent *Kp* strains (Yu et al., 2018; Chen et al., 2014). Here, we performed multiplex PCR to screen only 11 genes and it is based on the conserved sequence of the gene. Hence, no information about the variants of the resistant genes present can be obtained. Further, the reason behind detection of *bla*AIM, *bla*GIM, *bla*VIM and *bla*IMP in multiplex PCR but not in WGS could be the difference in specificity and accuracy of two methods; specificity of multiplex PCR is limited as amplicon length these genes were very small (Poirel et al., 2011), but WGS identifies the gene based on full-length sequence ( $\geq 95\%$  identity).

Presence of genes such as *ybt* cluster (yersiniabactin), *kfuABC* (Krapp et al., 2012; Lawlor et al., 2007) and *mrkABCD* (fimbriae) (Struve et al., 2008) could lead to increased survival of these isolates inside the host resulting in invasive *Kp* infections; which would be difficult to treat because of their extremely resistant phenotype. However, no genes for hypervirulence (*rmpA*, *rmpA2* and aerobactin) were detected in any MDR, XDR or PDR

isolates; no convergence between antibiotic resistance and hypervirulence was observed in the collected isolates.

Nanopore sequencing of DJ revealed the presence of multiple copies of *bla*<sub>OXA-181</sub> and *bla*<sub>CTX-M-15</sub> on chromosome of DJ. This continuous redundancy and dual-existence of *bla*<sub>OXA-181</sub> and *bla*<sub>NDM-5</sub> explains with the high-levels of resistance to all beta-lactams observed in DJ. The genetic environment of the copy of *bla*<sub>NDM-5</sub> present on IncFII plasmid in DJ (figure 4.13A) was similar as the previous reports (Naha et al., 2021; Pitart et al., 2015). In case of *bla*<sub>OXA-181</sub>, the location of *bla*<sub>OXA-181</sub> was recently reported downstream of truncated ISEcp1 with restricted transposase activity, which ensured stabilization of *bla*<sub>OXA-181</sub> on the plasmid (Shankar et al., 2021). But interestingly, in our study, chromosomal insertion of *bla*<sub>OXA-181</sub> in strain DJ was mediated by functional (non-truncated) ISEcp1 (figure 4.13B), which could further mobilize and disseminate *bla*<sub>OXA-181</sub> within the population of ST147 lineage. Integration of *bla*<sub>CTX-M-15</sub> into the chromosome was occurred by ISEcp1 mediated insertional inactivation of OmpK35 (in chromosomal copy 1; (figure 4.13C), which mediated disruption/loss of OmpK35. Loss of OmpK35 and OmpK36 in carbapenemase producing strain such as DJ act as a secondary mechanism for increased resistance to carbapenems (Shi et al., 2013; Karampatakis et al., 2016). Occurrence of *bla*<sub>OXA-181</sub> and *bla*<sub>NDM-5</sub> on separate plasmid/chromosomal locations indicate two independent events of gene acquisition by the strain (Naha et al., 2021). This suggests that PDR strains of high-risk clones like DJ spread not only by clonal dissemination, but are also associated with global dispersion of *bla*<sub>OXA-181</sub> and *bla*<sub>NDM-5</sub>. This indicates higher genetic plasticity of PDR strains compared to others.

4.13A). To best of my knowledge, this is the first report on chromosomal integration of *bla*<sub>NDM-5</sub> in high risk clone ST147. The incorporation of carbapenemase and ESBL genes in the chromosome of isolates from the high-risk clone ST147 indicates a possibility of maintenance, survival and vertical dissemination of these genes within ST147 lineage, which demonstrates a novel antimicrobial resistance threat. Disruption of *mgrB* by IS5 transposase explains the observed colistin resistance phenotype (MIC=4 µg/ml). *mgrB* mutation observed in DJ was reported previously among ST258, ST512 and ST101 (Shankar et al., 2019; Poirel et al., 2015; Cannatelli et al., 2014). Apart from IS5, disruption of *mgrB* by IS1-like transposase has also been reported (Azam et al., 2021). This is extremely worrisome because the last-resort drugs such as polymyxins (especially colistin) have been re-introduced for the treatment of carbapenem-resistant *Kp* infections (Poirel et al., 2017); now the evolution of *Kp* into PDR strains like DJ left no treatment option.

Location specific epidemiological study of high risk clones such as ST147 and ST231 are crucial in surveillance because *Kp* infections and resistance phenotypes vary from country to country along with some intercontinental similarity (Ling et al., 2015). This is due to socio-demographic factors such as difference in clinical settings, source of domestic water, companion animals (cats, dogs), livestock (chicken, cattle), malnutrition, and the use and misuse of antibiotics (Effah et al; 2020). It is reported that south and southeast Asia are hub form AMR *Kp* (Wyre et al., 2020); exposure to health care facility and history of previous overseas hospitalization is one of the notifiable risk factors associated with *Kp* infections (Espenhain et al., 2018) but on the contrarily, Ling et al., reported that persons with no history of overseas travel and overseas hospitalization are also at risk of *K. pneumoniae* colonization and infections (Ling et al., 2015). In our study, we observed that South and Southeast Asian countries (India, Thailand, Pakistan, etc.) are hotspots for AMR genes, and the genomes were highly diverse in terms of carbapenemase and ESBL genes. Hence, to recognize and control the growing public health threat of AMR *K. pneumoniae* worldwide, scrutinizing clinically prevalent lineages is imperative.

CG147 is one of the most prevalent clonal group found in clinical settings of India; also isolates belonging to CG147 reported worldwide. The analysis of CG147 genomes reported globally along with our PDR isolate DJ, helped us to understand the big picture of evolution and epidemiology of this lineage, that led to emergence of this high-risk clone. The records

suggested that between 2008 to 2010, the isolates belonging to ST147 only carried *bla*<sub>CTX-M-15</sub>; reported first from Hungary and then it disseminated across this country since 2005 (Szilágyi, 2010; Damjanova et al., 2008). But between 2011-2014 the number of publications increased reporting carriage and expansion of different carbapenemases in CG147 lineage; this was consistent with higher number of genomes isolated between these years (figure 4.16B). CG147 was found to be associated with particular carbapenemases in specific regions, such as it is associated with VIM and *KPC* in Greece and Italy (Giakkoupi et al., 2011; Lascols et al., 2013; Peirano et al., 2014) with *KPC* in France and Portugal (Liapis et al., 2014; Rodrigues et al., 2016) with NDM and VIM in Spain (Peirano et al., 2014, Pérez-Vázquez et al., 2019) OXA-48-like and NDM in India, Pakistan, Turkey and Tunisia (Giske et al 2012; Lascols et al., 2013; Pitout et al., 2019) and NDM with China (Wang et al., 2014). Moreover, the literature indicates that globalization has played a role in the spread of this high-risk CG clone in countries like Canada, United States and Nordic countries, which has been strongly related to travels to endemic countries such as India (medical tourism) and Greece (Cantón et al., 2012, Giske et al 2012, Peirano et al., 2014; Rojas et al., 2014). Recent reports of isolates belonging to CG147 describe the XDR and PDR phenotypes (Nahid et al., 2017; Sonnevend et al., 2017; Avgoulea et al., 2018, Simner et al., 2018, Falcone et al., 2020) The ability of this clonal group to spread and carriage of diverse set of carbapenemase genes is of particular concern.

The presence of QRDR mutations (*gyrA* and *parC* mutations) in all genomes of CG147 and ST231 (including our isolate DJ, M2 and M6) probably suggests the consequences of the extensive use of fluoroquinolones (mainly ciprofloxacin) and extended-spectrum cephalosporins. Since DNA gyrase (encoded by *gyrA*, *gyrB*) and topoisomerase (encoded by *parC* and *parE*) are the primary target fluoroquinolones, (Woodford and Ellington, 2007); lineage of clinical strains such as CG147 have evolved with mutation in *gyrA* and *parC*, and became resistant to these drug classes over the years. This suggests a role of (fluoro)quinolones selective pressure in the emergence and dissemination of this clone.

Diversity of *ybt*:*ICEKp* variants found in CG147 genomes added to the novel diversity compared to the previous study (Lam et al., 2018). Same kind of *OmpK35/36* mutations (harboring *ybt10/ICEKp4*) were observed in n=18 genomes of CG147 (including isolate DJ); found to be predominantly circulating in Asia. Such mutations were reported in one genome from North America, in a clinical case imported from India (Simner et al., 2018).

Regarding the important question of convergence, convergence of MDR and hypervirulence (Hv) genotypes due to acquisition of virulence plasmids was recently observed (Turton et al., 2018; Pajand et al., 2020). This suggests that this MDR/Hv plasmid continues to circulate within this clonal group; adding a new risk to epidemiology of such high-risk clones. No detection of *cps* cluster in 42% of the genomes could be due to sequencing artifact or loss and further investigation is needed.

Although ST231 is a well-spread MLST across 13 countries of 5 continents; ST231 is significantly associated with Southeast Asia (Thailand, Vietnam) and South Asia (India, Pakistan) (Wyres et al., 2020). ST231 is the most prevalent sequence type found in clinical isolates circulating in the India (Shankar et al., 2019); our results of ST231 analysis are in corroboration with these findings. We found KL-51 was predominant in genomes belonging to ST231 as reported previously (Wyres et al., 2020); however, we also found two genomes with KL-64. Presence of IncH1B was reported to be associated with ST231 (Giske et al., 2021); however, we did not find such association as none of the genome of ST231 carried IncH1B plasmid.

Majority of ST231 genomes were belonging to Asian countries; genomes deposited from United Kingdom and Portugal in 2010 and 2013, respectively; might be due to sampling bias and early reporting rather than origination of ST231 in Europe. Moreover, there is a limitation of the information available in the public databases because countries that report more are not the ones with high prevalence and diversity of variants; and south Asian countries like India are hotspots, but whole genome sequencing approach has not been incorporated yet in the regular surveillance.

Recent shift from *bla*<sub>NDM</sub> to *bla*<sub>OXA-48-like</sub> in Indian clinical isolates was reported (Veeraraghavan et al., 2017). Presence of *bla*<sub>OXA-48-like</sub> in majority of Indian ST231 genomes observed in our study supports this finding. Diverse mobile genetic elements (MGEs) such as ISX4, IS1, IS3, ISK<sub>p</sub>n1, ISK<sub>p</sub>n26, ISK<sub>p</sub>n25 as well as transposons such as Tn1999, Tn1999.2 associated with *bla*<sub>OXA-48-like</sub> (Cuzon et al., 2011; Beyrouthy et al., 2014) contributed to their spread (Shankar et al., 2019). Records suggest that OXA-48-like carbapenemase was initially geographically confined to India (Lascols et al., 2011), North African countries (Barguigua et al., 2012) and Turkey (Alp et al., 2013); it was then spread to Switzerland (Mancini et al., 2018) and United states (Lascols et al., 2013) in 2013.

OXA-48-like carbapenemase were reported to be present nationwide in India including community settings (Potron et al., 2013; Poirel et al., 2012). Recently, Naha et al., also reported the high prevalence of *bla*<sub>OXA-181</sub> carried by ST231 genomes in India (Naha et al., 2021). Abundance of *bla*<sub>OXA-48-like</sub> (OXA-181 and OXA-232) genes was also observed in Indian genomes during comparative analysis of ST231, confirmed the high frequency of *bla*<sub>OXA-48-like</sub> circulating in India. Moreover, presence of *bla*<sub>OXA-48</sub> in majority of the ST231 genomes (77.3%) also suggests the role of overuse or misuse of carbapenems in evolution of carbapenemase producing strains belong to clinically prevalent lineages such as ST147 and ST231. Unlike ST147, *bla*<sub>NDM</sub> was not found to be more prevalent in ST231 as only two genomes from USA found to carry *bla*<sub>NDM</sub>. Coexistence of *bla*<sub>OXA-48</sub> and *bla*<sub>CTX-M</sub> genes observed in 34% of ST231 genomes is worrisome as the majority of ST231 isolates could be resistant to carbapenems and extended-spectrum Beta-lactams. We also found that the convergent genomes carrying *iuc* (aerobactin) and *bla*<sub>OXA-232</sub> mainly belonged to South and Southeast Asia as reported previously by Wyres et al (Wyres et al., 2020).

Comparison of ST147 and ST231 genomes exhibited that, the prevalence of *bla*<sub>NDM</sub> was high in ST147 genomes compared to very low prevalence of *bla*<sub>NDM</sub> in ST231. *bla*<sub>OXA-48-like</sub> genes are highly prevalent in ST231 than ST147 genomes. Further, aminoglycoside resistance genes such as *aac*-(6')-Ib, *aac*-(3')-II were abundantly present in ST147 population, but were observed to be absent in ST231. All genomes of both ST147 and ST231 carried *gyrA* and *parC* mutations. Hence, different set of antibiotic resistance determinants are associated with different sequence types (ST). Moreover, combinations of antibiotic genes present in each ST are also associated with geographical location.