

RESEARCH

Open Access



Exploring the resistome, virulome, and mobilome of multidrug-resistant *Klebsiella pneumoniae* isolates: deciphering the molecular basis of carbapenem resistance

Sidra Rahmat Ullah¹, Sidra Irum¹, Iqra Mahnoor¹, Humaira Ismatullah², Mariam Mumtaz¹, Saadia Andleeb^{1*}, Abdur Rahman¹ and Muhsin Jamal³

Abstract

Background *Klebsiella pneumoniae*, a notorious pathogen for causing nosocomial infections has become a major cause of neonatal septicemia, leading to high morbidity and mortality worldwide. This opportunistic bacterium has become highly resistant to antibiotics due to the widespread acquisition of genes encoding a variety of enzymes such as extended-spectrum beta-lactamases (ESBLs) and carbapenemases. We collected *Klebsiella pneumoniae* isolates from a local tertiary care hospital from February 2019–February 2021. To gain molecular insight into the resistome, virulome, and genetic environment of significant genes of multidrug-resistant *K. pneumoniae* isolates, we performed the short-read whole-genome sequencing of 10 *K. pneumoniae* isolates recovered from adult patients, neonates, and hospital tap water samples.

Results The draft genomes of the isolates varied in size, ranging from 5.48 to 5.96 Mbp suggesting the genome plasticity of this pathogen. Various genes conferring resistance to different classes of antibiotics e.g., aminoglycosides, quinolones, sulfonamides, tetracycline, and trimethoprim were identified in all sequenced isolates. The highest resistance was observed towards carbapenems, which has been putatively linked to the presence of both class B and class D carbapenemases, *bla_{NDM}*, and *bla_{OXA}*, respectively. Moreover, the biocide resistance gene *qacEdelta1* was found in 6/10 of the sequenced strains. The sequenced isolates exhibited a broad range of sequence types and capsular types. The significant antibiotic resistance genes (ARGs) were bracketed by a variety of mobile genetic elements (MGEs). Various spontaneous mutations in genes other than the acquired antibiotic-resistance genes were observed, which play an indirect role in making these bugs resistant to antibiotics. Loss or deficiency of outer membrane porins, combined with ESBL production, played a significant role in carbapenem resistance in our sequenced isolates. Phylogenetic analysis revealed that the study isolates exhibited evolutionary relationships with strains from China, India, and the USA suggesting a shared evolutionary history and potential dissemination of similar genes amongst the isolates of different origins.

Conclusions This study provides valuable insight into the presence of multiple mechanisms of carbapenem resistance in *K. pneumoniae* strains including the acquisition of multiple antibiotic-resistance genes through mobile genetic elements. Identification of rich mobilome yielded insightful information regarding the crucial role of insertion

*Correspondence:

Saadia Andleeb

saadiamarwat@yahoo.com; saadia.andleeb@asab.nust.edu.pk

Full list of author information is available at the end of the article



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

sequences, transposons, and integrons in shaping the genome of bacteria for the transmission of various resistance-associated genes. Multi-drug resistant isolates that had the fewest resistance genes exhibited a significant number of mutations. *K. pneumoniae* isolate from water source displayed comparable antibiotic resistance determinants to clinical isolates and the highest number of virulence-associated genes suggesting the possible interplay of ARGs amongst bacteria from different sources.

Keywords *Klebsiella pneumoniae*, Whole genome sequencing, Comparative genomics, Carbapenem resistance, Resistome, Mobilome, Virulome, Pakistan, ARGs

Introduction

Klebsiella pneumoniae (*K. pneumoniae*) is an important member of the *Enterobacteriaceae* family and a natural inhabitant of the gastrointestinal tract (GIT) microflora of humans and animals. It has evolved as an “antimicrobial resistant opportunist” accountable for one-third of all Gram-negative infections [1–3]. This bacteria lies in the list of ESKAPE organism amongst “critical priority pathogens” by the World Health Organization (WHO) [4]. *K. pneumoniae* has risen as a significant pathogen in Neonatal Intensive Care Units (NICUs) in developing countries having a mortality rate of 18–68% [5, 6]. The NICU-admitted and premature newborns face a higher susceptibility to risk due to their underdeveloped immune system, unsettled microflora, and slightly enhanced permeability of the GIT mucosa [7]. Over the past decade, a significant increase has been observed in sepsis and bacteremia in the neonatal population due to multi-drug resistant bacteria, globally [8]. With the ongoing pandemic, the situation has become increasingly perilous. Recent reports indicate that a considerable proportion of COVID-19 patients, with a mortality rate of 56.7%, are experiencing secondary bacterial infections. Among these infections, the predominant pathogen isolated from all affected individuals is MDR *K. pneumoniae* [9]. Despite the appropriate use of antimicrobial therapy, fatality rates have risen to 50% making *K. pneumoniae* the leading cause of mortality and morbidity by causing nosocomial infections [10]. The spread of antibiotic resistance genes (ARGs) in bacterial populations, is facilitated by horizontal gene transfer (HGT) aided by mobile genetic elements (MGEs). These MGEs encompass plasmids, transposons, and integrons. These elements serve as a carrier for ARGs and enable their transfer amongst different Gram-negative and Gram-positive bacterial species [11].

The rise of *K. pneumoniae* strains producing extended-spectrum β -lactamases (ESBLs), *Klebsiella pneumoniae* carbapenemases (KPCs), and Metallo β -lactamases (MBL) has become a pressing concern. The two primary non-specific porins produced by *K. pneumoniae* are OmpK35 and OmpK36, which facilitate the passive diffusion of small hydrophilic molecules and

clinically important antibiotics like β -lactams and fluoroquinolones. Mutations in these outer membrane porins are linked with carbapenemase enzymes to enhance carbapenem resistance in this pathogen. Factors such as point mutations or insertional interruptions in the coding sequences or the promoter region can modify the expression of porins in clinical isolates of *K. pneumoniae*. Strains lacking both OmpK35 and OmpK36 have demonstrated elevated levels of antibiotic resistance [12, 13]. These resistant strains pose a substantial challenge in the clinical management of infections, as they are increasingly becoming untreatable [14]. According to the CDC Antimicrobial Resistance (AMR) action plan, tracking antimicrobial-resistant organisms is one of the four core actions proposed to limit the emergence and spread of multi-drug-resistant (MDR) pathogens [15]. To combat the frequent threat of outbreaks due to multi-drug resistant bacteria in immunocompromised patients in hospital settings, whole-genome sequencing (WGS) has become an important tool. With *K. pneumoniae* being established as an urgent and major health problem globally the WGS has become the most reliable tool for investigation and keeping track of bacterial pathogens [16]. Whole genome analysis of a pathogen can provide insights into its related lineages, pathogenicity, virulence, and antimicrobial resistance mechanisms that can revolutionize outbreak analysis and management [17].

We aim to perform a comprehensive genomics analysis to check the diversity of virulence-associated genes, antibiotic resistance genes (particularly involved in conferring resistance to last-resort antibiotics), and mobile genetic elements responsible for the dissemination of significant genes involved in drug resistance amongst *K. pneumoniae* strains recovered from different sources.

Results

AST Profile & Biofilm Assay

Antimicrobial susceptibility testing was carried out following Clinical and Laboratory Standards Institute (CLSI) 2018 guidelines. Based on the acquired antibiotic resistance profiles, 19% (41/214) strains were classified as susceptible, other 35% (75/214) strains as

multi-drug-resistant (MDR), and the remaining 46% (99/214) strains as extensively drug resistant (XDR). Most of the isolates exhibited resistance to beta-lactam antibiotics. Resistance to fluoroquinolones and aminoglycosides was also observed in the majority of *K. pneumoniae* strains. The isolates were found susceptible to fosfomycin and did not exhibit any resistance towards polymyxin.

Based on the AST profiles generated using multi-resistance classification guidelines [18], four XDR and six MDR *K. pneumoniae* strains isolated from adults, neonates, and hospital tap water were subjected to whole genome sequencing. The biofilm detection assay revealed all XDR and one MDR strain (kp10) as excellent biofilm-former regardless of their source (Table 1). The AST profiles of our sequenced isolates are illustrated in Table 2.

Genome characterizations

The draft genome sizes of the isolates ranged from 5.48 to 5.96 Mbp. The specific values for L50, N50, and contig numbers are shown in Supplementary Table 2. The water isolate kp10 exhibited the largest genome size, measuring 5.96 Mbp, in comparison to the genome sizes of other clinical isolates. The Whole Genome Shotgun project of the individual isolate has been deposited at DDBJ/ENA/GenBank under the accession numbers and the version described in this paper are described in Supplementary Table 1.

A comparative genomic analysis using BRIG was conducted to visually compare our sequenced isolates with the reference *K. pneumoniae* genome HS11286 [19] (Supplementary Fig. 1). The visual representation allowed a comprehensive comparison between the sequenced isolates and the reference genome, aiding in the identification of similarities, differences, and potential genomic variations.

Table 1 Source, host age, and biofilm formation ability of selected *Klebsiella pneumoniae* isolates

Strains	Year of Collection	Source	Host Age	Biofilm Former
kp5	2018	Blood	10 Days	Excellent
kp6	2018	Urine	27 years	Excellent
kp10	2018	Hospital Tap Water	NA	Excellent
kp47	2020	Blood	New Born	Excellent
kp54	2020	Blood	New Born	Excellent
kp58	2021	Blood	2 Days	Excellent
kp106	2018	Pus	37 years	Excellent
kp126	2019	Blood	5 Days	Moderate
kp127	2019	Blood	2 Days	Moderate
kp128	2019	Blood	3 Days	Moderate

MLST and capsular types

The in silico MLST analysis revealed that the isolates encompassed seven distinct sequence types. Subsequent eBURST analysis assigned the sequenced strains to their respective clonal groups (CG). The study identified globally significant sequence types and recognized problematic clonal groups, such as water isolate kp10 falling within ST15/CG15 and clinical isolate kp47 in ST70/CG70 [1, 20–22]. *K. pneumoniae* typing is based on sequence variations of capsular polysaccharides known as K antigens and lipopolysaccharides such as O antigens. (2, 27). Within our study isolates, we noted the occurrence of four distinct O capsule types and identified eight distinct K capsule types, as presented in Table 3.

Resistome of *K. pneumoniae* isolates

Numerous antibiotic-resistance genes were observed in the sequenced *K. pneumoniae* isolates (Fig. 1). Further mining of whole-genome sequences revealed diverse β-lactamase encoding genes in the isolates. The majority of the isolates harbored more than two β-lactamase encoding genes in their genomes i.e. *bla_{CTX-M-15}*, *bla_{TEM}*, and *bla_{SHV}*. The highest resistance was observed towards carbapenems, which has been putatively linked to the presence of both class B and class D carbapenemases, *bla_{NDM}*, and *bla_{OXA}*, respectively. The water isolate kp10 was the sole isolate containing the CMY-4 beta-lactamase, in addition to other resistance genes. The second most frequent group of resistance genes belonged to the *arr-2* and *arr-3* gene families against aminoglycosides. The plasmid-mediated quinolone resistance (PMQR) genes *oqx* and *qnr* were also detected amongst which *qnrB1*, *qnrB6*, and *qnrS1* were more prominent. The resistance observed towards trimethoprim-sulfamethoxazole antibiotics appears to be potentially linked to the presence of *sul1*, *sul2*, and *dfrA* genes. In addition, *tetA* and *tetD* genes were observed in the genomes of three of our isolates which have been reported to confer resistance to tetracycline. In all isolates, the biocide-resistance gene *qacEΔ1* was commonly observed. Interestingly, kp47 and kp58 exhibited phenotypic resistance to various antibiotics despite possessing the fewest number of resistance genes. On the contrary, no genetic determinant was found against colistin in any of the isolates. The correlation of genotypic and phenotypic data of major antibiotics is provided in Table 4.

Virulome of *K. pneumoniae* isolates

Several genes associated with virulence factors, including adherence, capsule formation, efflux pumps, siderophores, type VI secretion system (T6SS-III), serum

Table 2 AST profile of the selected *Klebsiella pneumoniae* isolates

Strain	β lactams												Combinations	Fosfomycin	Polymyxin						
	Penicillin						Cephalosporins									Carbapenems					
	Aminoglycosides		Fluoroquinolones		Carbapenems		Aminoglycosides		Fluoroquinolones		Carbapenems					Aminoglycosides		Fluoroquinolones		Carbapenems	
AMP	FEP	CAZ	CTX	CRO	MEM	IMP	AK	CN	TOB	CIP	LEV	NOR	OFX	TZP	SXT	FOS	CT				
kp5	R	R	R	R	R	R	R	S	R	R	S	R	R	R	R	R	R	XDR			
kp6	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	S	XDR			
kp10	R	R	R	R	R	S	S	I	I	R	R	R	R	I	R	S	S	MDR			
kp47	R	R	R	R	R	R	R	R	R	R	R	R	R	I	R	S	S	MDR			
kp54	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	S	S	XDR			
kp58	R	R	R	R	R	R	S	S	R	R	R	R	R	R	R	S	S	MDR			
kp106	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	S	S	XDR			
kp126	R	R	R	R	R	R	S	S	S	I	I	S	R	R	R	S	S	MDR			
kp127	R	R	R	R	R	R	S	S	S	S	I	I	R	R	R	S	S	MDR			
kp128	R	R	R	R	R	R	R	R	S	I	I	R	R	R	R	S	S	MDR			

AMP Ampicillin, FEP Cefepime, CAZ Ceftazidime, CTX Cefotaxime, CRO Ceftriaxone, MEM Meropenem, IMP Imipenem, AK Amikacin, CN Gentamicin, TOB Tobramycin, CIP Ciprofloxacin, LEV Levofloxacin, NOR Norfloxacin, OFX Ofloxacin, TZP Cefotolozane/Tazobactam, SXT Trimethoprim/Sulfamethoxazole, FOS Fosfomycin, CT Colistin. Sensitive (Probably susceptible to ordinary dosage therapy), I: Intermediate (Likely to respond to high dosage therapy), R: Resistant (Unlikely to respond to high dosage therapy)

Table 3 MLST and Capsular typing of the selected *Klebsiella pneumoniae* isolates

Strains	MLST	K Antigen	O Antigen	Clonal Group (CG)
kp5	ST54	KL14	O3b	CG54
kp6	ST54	KL14	O3b	CG54
kp10	ST15	KL24	O1v1	CG15
kp47	ST70	KL22	O1/O2v2	CG70
kp54	ST29	KL149	O3b	CG29
kp58	ST2703	KL149	O3b	CG586
kp106	ST231	KL51	O1v2	CG231
kp126	ST22	KL9	O2v2	CG22
kp127	ST22	KL9	O2v2	CG22
kp128	ST22	KL9	O2v2	CG22

resistance, and immune evasion, were identified in all isolates of our study. Nonetheless, the non-clinical isolate (kp10) along with clinical isolate (kp106), were found to carry genes associated with yersiniabactin (*irp1*, *irp2*, *ybtA*, *ybtE*, *ybtP*, *ybtQ*, *ybtS*, *ybtT*, *ybtU*, and *ybtX*). Remarkably, these isolates demonstrated the highest abundance of virulence genes, as illustrated in Fig. 2.

All sequenced isolates were found to possess genes associated with various virulence factors, including *wcaG*, *wza*, *sugE*, and *treC* for capsular polysaccharide (CPS), *pgaA*, *pgaB*, *pgaC*, and *bcsA* for adhesion, and *luxS* for quorum sensing. Additionally, genes related to pilli (*ecpA*), fimbriae type I (*fimABCDEFGHIK*), and type III (*mrk ABCDFHIJ*), crucial for biofilm formation, were identified in all analyzed isolates. Multiple variations were observed in the *treC* and *wza* genes, crucial for biofilm formation, as well as in genes related to the secretion system across all isolates. Notably, significant variations were identified in genes associated with the type six

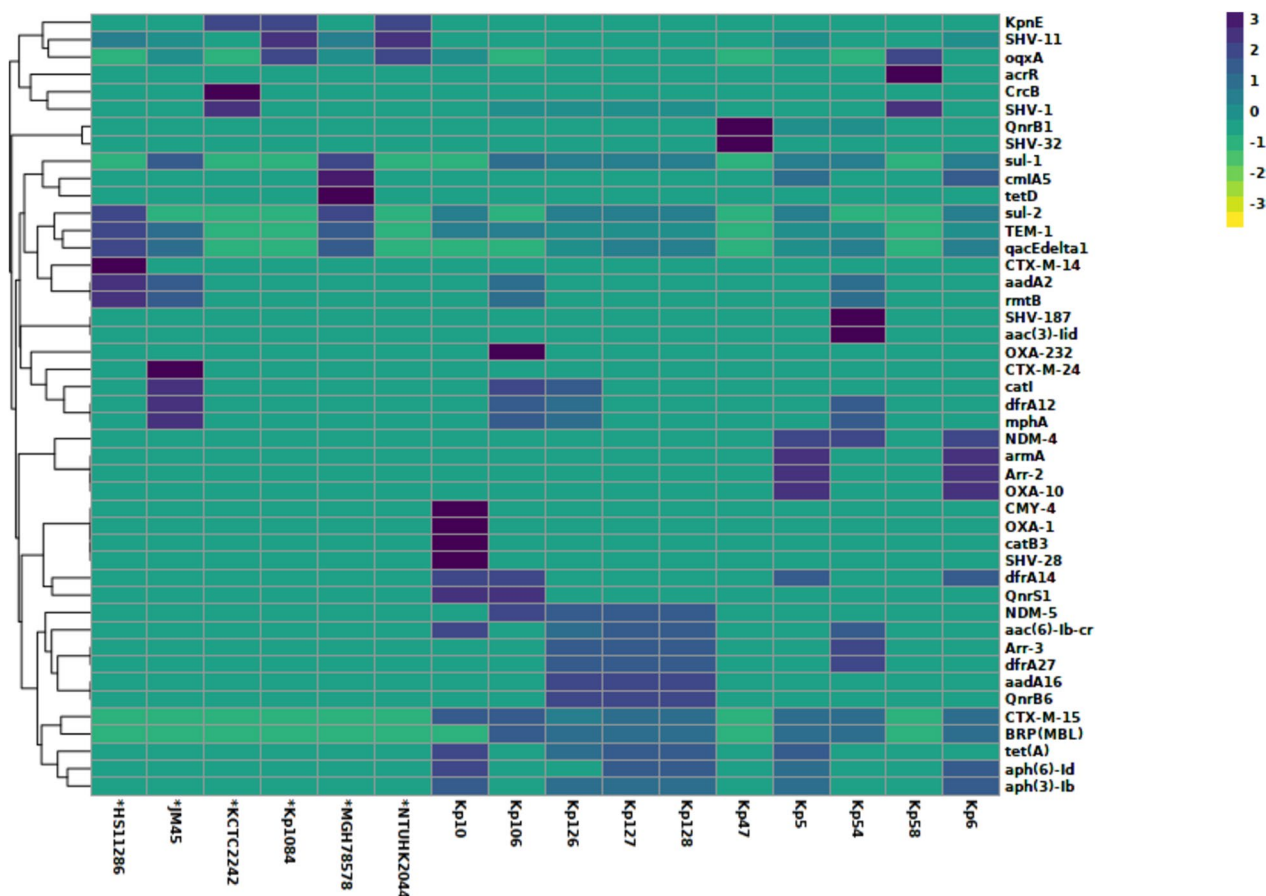


Fig. 1 The heatmap displays the relative copy numbers of the ARGs. AMR determinants in *K. pneumoniae* genomes were clustered based on Euclidean distance and hierarchical ward algorithm. Sequenced isolates along with the reference genomes (Kp1084, HS11286, KCTC2242, MGH78578, NTUH2044 and JM45) are arranged on the X axis, while the resistance genes are plotted on the Y-axis. Resistance genes with a darker blue color have a relatively low copy number, while resistance genes with a lighter blue-yellow color have a relatively high copy number in comparison to others

Table 4 Correlation between genotypic and phenotypic resistance profiles of sequenced *K. pneumoniae* isolates

Strain	β Lactams		Carbapenem		Aminoglycosides		Fluoroquinolones		Miscellaneous	
	Genotype	Phenotype	Genotype	Phenotype	Genotype	Phenotype	Genotype	Phenotype	Genotype	Phenotype
kp5	SHV-11, CTX-M-15, TEM-1,	Resistant	NDM-4, OXA-10	Resistant	armA, aph(6)-Id, aph(3'')-Ib	Resistant to all except CN	QnrB1, oqxA	Resistant to all except LEV	sul1, sul2, dfrA12, tetA	Resistant
kp6	SHV-11, TEM-1,	Resistant	NDM-4, OXA-10	Resistant	armA	Resistant	oxqA	Resistant	sul1, sul2,	Resistant
kp10	SHV-28, TEM-1, CTX-M-15,	Resistant	OXA-1, CMY-4	Resistant to MEM Susceptible to IMP	aph(3'')-Ib, aph(6)-Id, aac(6')-Ib-cr	Intermediate Susceptible	QnrS1, aac(6')-Ib-cr, oxqA	Resistant	sul-2, dfrA14, tetA	Resistant
kp47	SHV-32	Resistant	Nil	Resistant	Nil	Resistant	QnrB1	Resistant	KpnF	Resistant
kp54	SHV-187, CTX-M-15, TEM 1,	Resistant	NDM 1	Resistant	aac(3)-Iid, aadA2, rmtB	Resistant	aac(6')-Ib-cr	Resistant	sul 1, dfrA12, dfrA27	Resistant
kp58	SHV-1	Resistant	Nil	Resistant	Nil	Susceptible to AK, CN	oxqA	Resistant	Nil	Resistant
kp106	SHV-1, TEM-1, CTX-M-15	Resistant	OXA-232, NDM-5,	Resistant	rmtB, aadA2	Resistant to TOB	QnrS1	Resistant	sul1, dfrA12	Resistant
kp126	SHV-1, CTX-M-15, TEM-1,	Resistant	NDM-5	Resistant	Nil	Susceptible to AK, CN Intermediate to TOB	QnrB6	Intermediate	dfrA12, sul 1, sul 2	Resistant
kp127	SHV-1, CTX-M-15, TEM-1,	Resistant	NDM-5	Resistant	aac(6')-Ib-cr, aph(3'')-Ib, aph(6)-Id	Susceptible	QnrB6, aac(6')-Ib-cr	Intermediate	sul-1, sul-2	Resistant
kp128	SHV-1, CTX-M-15, TEM-1,	Resistant	NDM-5	Resistant	aac(6')-Ib-cr, aph(3'')-Ib, aph(6)-Id	Resistant	QnrB6, aac(6')-Ib-cr	Resistant	sul-1, sul-2	Resistant

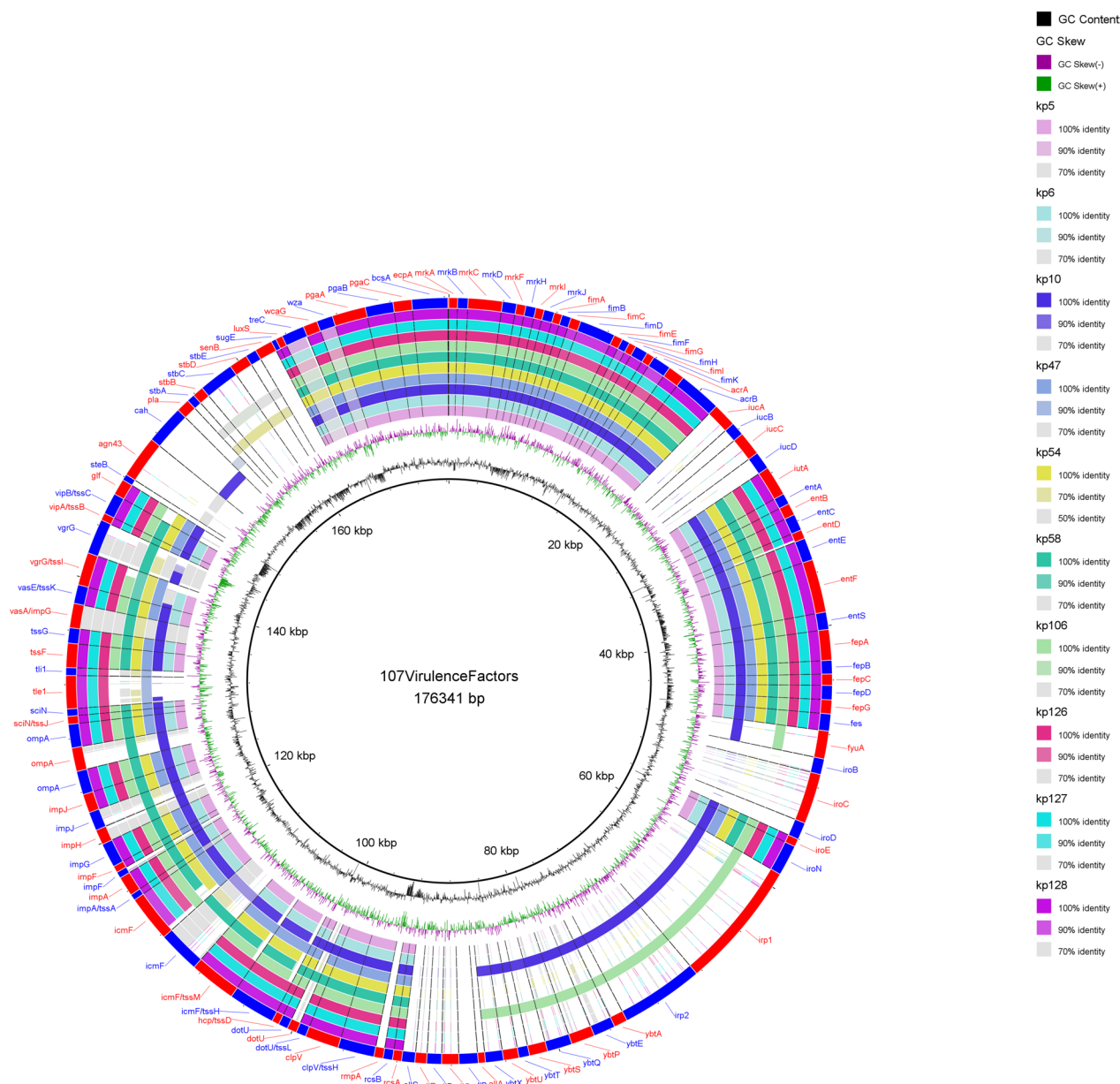


Fig. 2 A circular diagram illustrating the whole genomes (chromosomes and plasmids) of this study. The draft genomes were aligned with 107 virulence genes from reference strains curated from VFDB, to assess their presence and distribution (Kp1084, HS11286, KCTC2242, MGH78578, NTUH2044 and JM45). Each genome is visually represented by a ring in the circular diagram, where different colors were used to indicate the percentage identity of the virulence gene in the studied genomes compared to the reference sequence. The color coding provides a visual representation of the degree of similarity or variation in the virulence genes across the genomes. The isolates kp10 and kp106 harbored the highest number of virulence factors. The image was generated using BRIG (<http://brig.sourceforge.net>)

secretion system (T6SS). Virulence genes for allantoin uptake were absent in all the sequenced genomes.

Mobilome of *K. pneumoniae* isolates

The analysis of mobile genetic elements (MGEs) revealed a consistent presence of nearly identical transposons and insertion elements in all isolates. The water isolate

kp10 exhibited extensive mobilome in close proximity of the ARGs. The chromosomal encoding of all *bla_{SHV}* variants was notable, with insertion into the lac operon, forming a *glpR-bla_{SHV}-lacY-lacZ* structure. The examination MGEs revealed the presence of nearly identical transposons and insertion elements in all isolates. Furthermore, the coexistence of *bla_{CTX-M-15}* and *bla_{TEM}*,

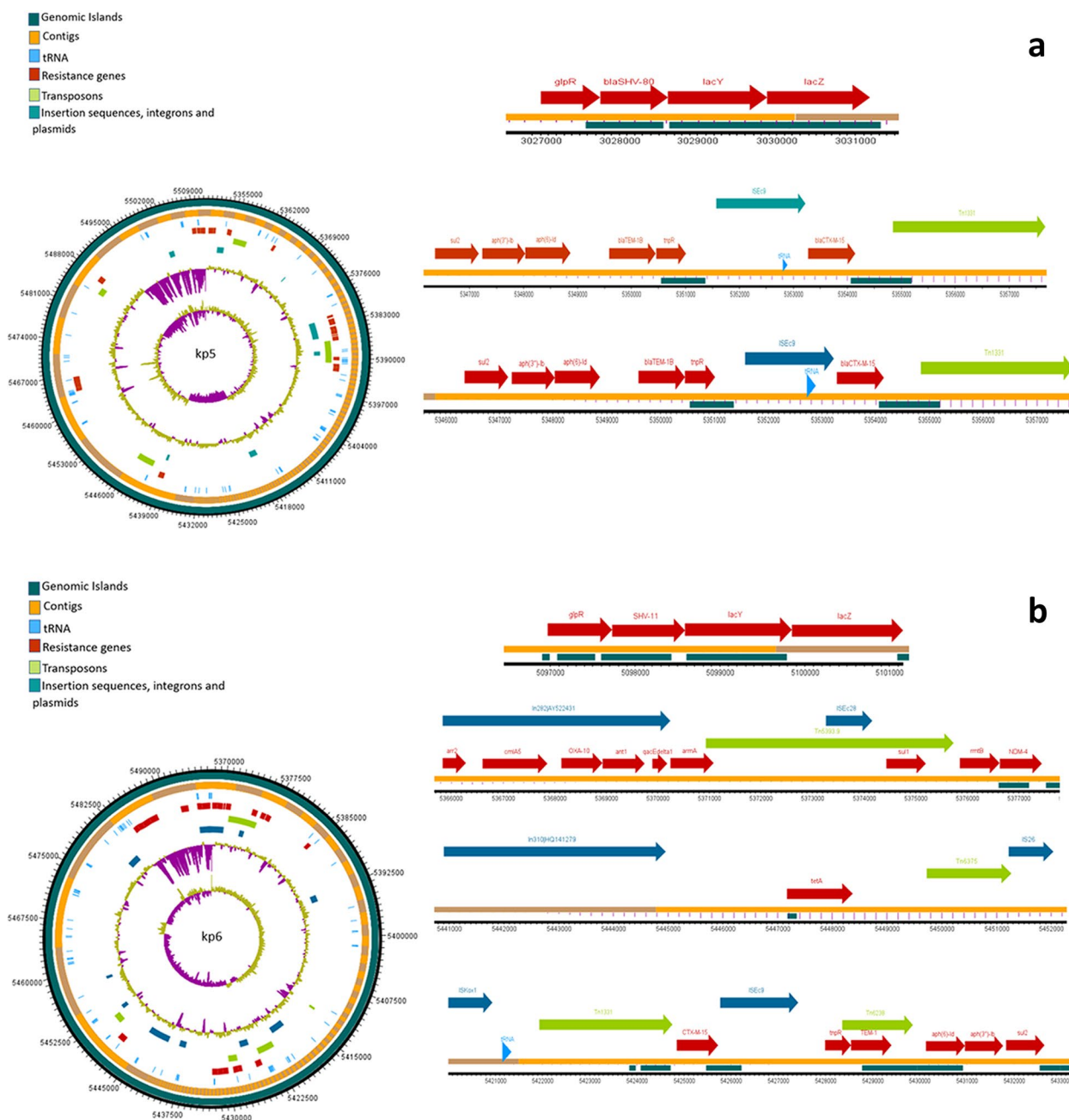


Fig. 3 Graphical representation of the genetic environment of the resistance genes and mobile genetic elements of kp5 (a) and kp6 (b). The circular visualization of the whole genome sequence represents all the resistance genes and the MGEs. The horizontal representation shows the genetic environment of the resistance genes with specific importance to beta-lactamase genes and MGEs. The genetic environment of chromosomally encoded *bla_{SHV}* was found consistent in all the genomes. The *bla_{CTX-M-15}* and *bla_{TEM}* genes were mostly found on Tn3 transposons and ISEc9

flanked by transposases/resolvase or tRNA, was observed in all isolates. Predominant mobile genetic elements, ISEc9 and Tn1331, were closely associated with both *bla_{CTX-M-15}* and *bla_{TEM}* genes. Notably, isolate kp6 exhibited additional MGEs such as Tn5393 and Tn6328. The *bla_{OXA}* variants appeared to be present as gene cassettes

carried by class I integrons. In 5/10 isolates, *bla_{NDM-1}* and *bla_{NDM-5}* genes were frequently found adjacent to the bleomycin resistance gene *ble*, while 2/10 isolates demonstrated a co-occurrence of the *rmtB* gene (encoding 16S rRNA methylase) and *bla_{NDM-4}*. Figures 3, 4, 5, 6 and 7

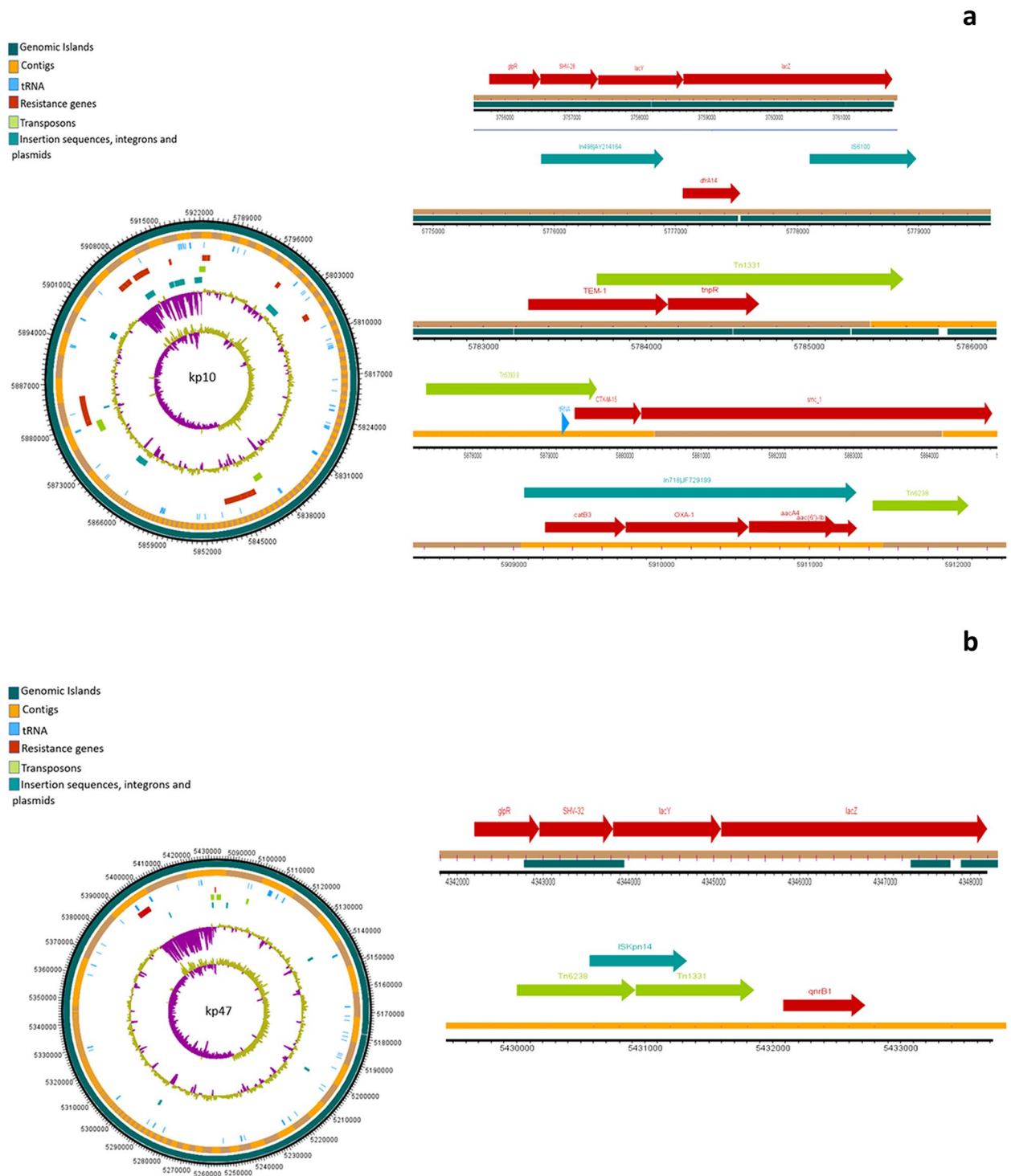


Fig. 4 Graphical representation of the genetic environment of the resistance genes and mobile genetic elements of kp10 (a) and kp47. The circular visualization of the whole genome sequence represents all the resistance genes and the MGEs. The horizontal representation shows the genetic environment of the resistance genes with specific importance to beta-lactamase genes and MGEs. The genetic environment of chromosomally encoded *bla_{SHV}* was found consistent in all the genomes. The *bla_{CTM-15}* and *bla_{TEM}* genes were mostly found on Tn3 transposons and ISEc9

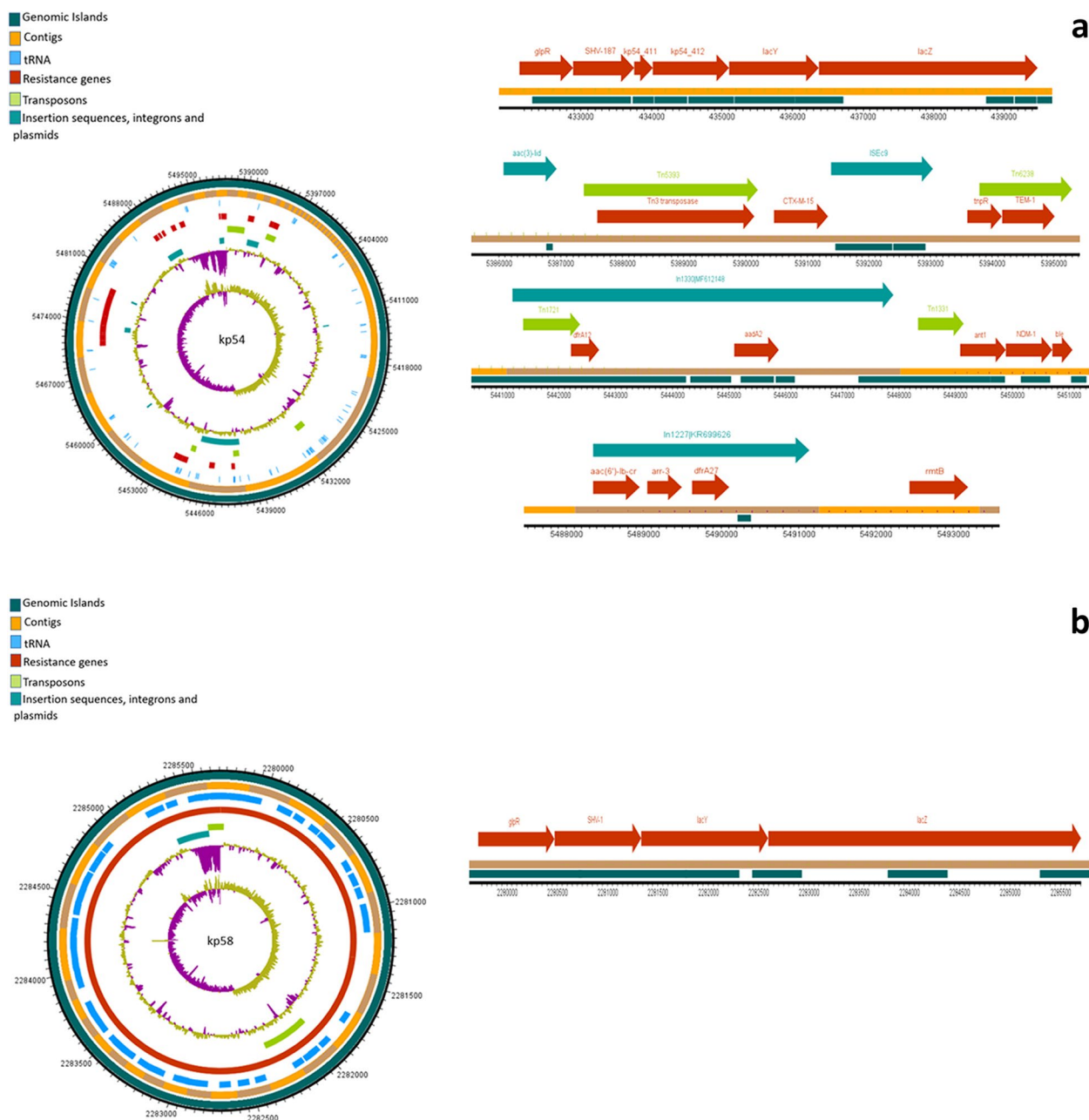


Fig. 5 Graphical representation of the genetic environment of the resistance genes and mobile genetic elements of kp54 (a) and kp58. The circular visualization of the whole genome sequence represents all the resistance genes and the MGEs. The horizontal representation shows the genetic environment of the resistance genes with specific importance to beta-lactamase genes and MGEs. The genetic environment of chromosomally encoded *bla_{SHV}* was found consistent in all the genomes. The *bla_{CTM-15}* and *bla_{TEM}* genes were mostly found on Tn3 transposons and ISEc9

illustrate the genetic environment of the resistance genes and MGEs in the sequenced isolates.

PlasmidFinder identified at least one plasmid replicon in all of the sequenced isolates from our study. Notably, isolates kp54 and kp106 were found to harbor three distinct plasmid replicons. Among the identified plasmid replicons, various incompatibility (Inc) types

were observed, such as IncH, IncF, and IncR. IncC plasmid replicon was exclusively identified in kp10. Specifically, isolates kp54, kp58, and kp106 were found to contain Col plasmids, while isolate kp47 had only the repB plasmid replicon. In addition to plasmid replicons, class I integrons were present in all the isolates. The most frequent gene cassettes observed within the

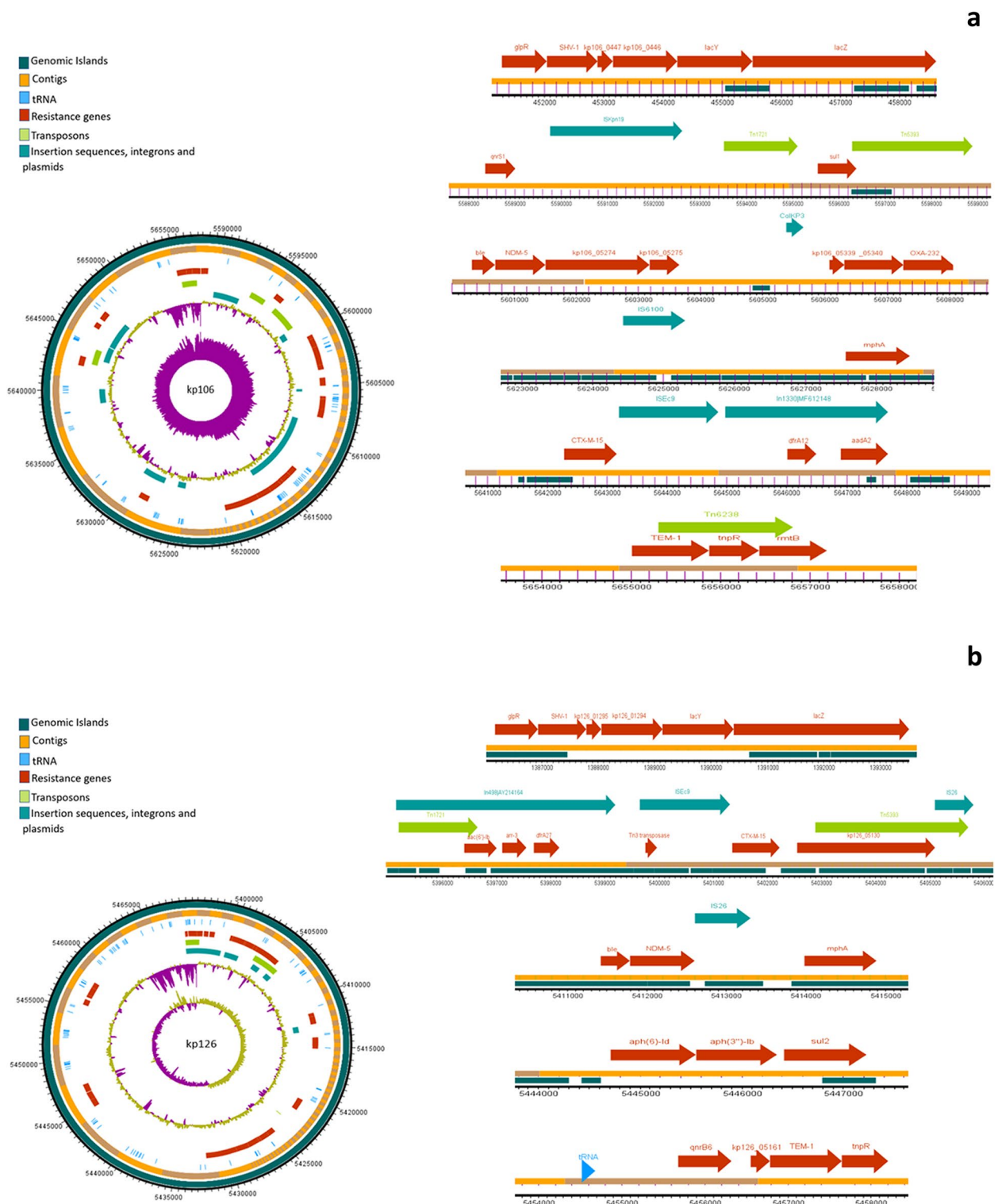


Fig. 6 Graphical representation of the genetic environment of the resistance genes and mobile genetic elements of *kp106* (a) and *kp126*. The circular visualization of the whole genome sequence represents all the resistance genes and the MGEs. The horizontal representation shows the genetic environment of the resistance genes with specific importance to beta-lactamase genes and MGEs. The genetic environment of chromosomally encoded *bla_{SHV}* was found consistent in all the genomes. The *bla_{CTX-M-15}* and *bla_{TEM}* genes were mostly found on Tn3 transposons and ISEc9

Table 5 Functional & structural consequences of SNPs

Protein	Mutation	SIFT		MUpro		
		Score	Prediction	▲G	Confidence Score	Prediction
Ompk36	Val178Pro	0.62	Tolerated	-1.295	-1	Decreases Stability
	Thr258Ser	0.04	Affect Protein Function	-1.048	0.97	Decreases Stability
	Leu307Ile	0.55	Tolerated	-0.389	-1	Decreases Stability
	Ile315Leu	1.00	Tolerated	-0.488	-1	Decreases Stability
	Asp344Glu	0.43	Tolerated	-0.809	-0.5	Decreases Stability
	His349Arg	0.56	Tolerated	-0.879	0.82	Decreases Stability

class I integrons were *dfpA* and *arr*, indicating the presence of genes associated with resistance to trimethoprim (*dfpA*) and rifampicin (*arr*), respectively. Among all the *K. pneumoniae* isolates in our study, the most prevalent type of trimethoprim resistance gene was *dfpA* type, which was found in 8/10 (80%) isolates. Further analysis of the sequenced isolates revealed that five of them contained integrons carrying the *dfpA-27* gene, three isolates had integrons with *dfpA-14* trimethoprim resistance gene.

MGE Finder and IS finder revealed the presence of multiple insertion sequence (IS) elements in all of the sequenced isolates in our study (Supplementary Table 3). Among these IS elements, ISEc9 was the most commonly identified IS element, and it was consistently found in the genetic context surrounding the beta-lactamase genes. Moreover, several intact (score > 90), questionable (score 70–90), and incomplete prophage sequences (score < 70) (Supplementary Table 1) were identified in all genomes by employing the PHASTER (Phage Search Tool Enhanced Release) tool. Except genomes kp5 and kp6, all the other genomes analyzed in our study contained at least one intact prophage sequence. Among the identified prophage sequences, the Klebsi_ST16_OXA48phi5.4_NC_049450 prophage sequence was observed in isolates kp58, kp126, kp127, and kp128. Additionally, isolates kp47 and kp106 harbored the Entero_mEp235_NC_019708 prophage sequence, which is approximately 46.3 kilobases (kb) in size. Furthermore, the Entero_mEp237_NC_019704 prophage sequence, measuring approximately 57.3kb, was identified in isolates kp126, kp127, and kp128.

Mutations in antibiotic resistance associated genes

To assess the prevalence and consequences of non-synonymous mutations in genes indirectly associated with antibiotic resistance, we performed Single Nucleotide Polymorphism (SNP) calling on the entire genome sequences using Snippy (Supplementary Table 5). The highest number of variations in genes encoding antibiotic efflux, antibiotic inactivation, antibiotic target alteration, outer membrane proteins, and lipid modification were observed

in the XDR strains kp128 (n=52). The water isolate kp10 also presented a significant number of non-synonymous mutations (n=43) in various genes similar to other clinical isolates. All the isolates harbored genetic determinants with variations in the Type 6 Secretion System (T6SSI-III) which is an important virulence factor [23]. Remarkably the isolates kp47 and kp58 that had the lowest resistance gene frequency were among the genomes that had the significant number of mutations. Most of the mutations were observed in antibiotic efflux pump genes i.e. *mdtC*, *mdtM*, *efpC* with the highest number of snps (n=63) in the *mdtO* gene. Several variations were found in genes encoding for significant outer membrane proteins OmpK35, Ompk36, and *tolC*. Amongst the observed variation Ile140Thr, Thr258Ser, and p.Gly3fs were the main mutations in Ompk35. Several types of mutations were found in the genes encoding the Ompk36 porin protein, including insertions, deletions, and missense mutations. Furthermore, all the isolates were observed to possess at least one mutation in the *pmrB* gene, *PmrAB* (two-component regulatory system), which is known to be one of the primary mechanisms for the development of resistance towards polymyxin. Furthermore, multiple mutations were identified in genes linked to fluoroquinolone resistance, with Ile83Ser being the predominant mutation in the *gyrA* gene and Ile80Ser in the *parC* gene. A heatmap was created to visually represent all non-synonymous mutations identified in the selected genes (Fig. 8).

Mutations in porins OmpK35 and OmpK36 and their association with Carbapenem resistance

Mutations in porins, specifically OmpK35 and OmpK36, have been linked to greater resistance to carbapenemase [12, 13]. The presence of frameshift variant c.5dupG p.Gly3fs in OmpK35 was found in all our sequenced isolates. This leads to change in the amino acid sequence, causing the replacement of glycine (Gly) at position 3 with a different amino acid, and the subsequent formation of a truncated protein. Several mutations that alter the amino acid sequence of OmpK36 were detected, and one of these mutations, Thr258Ser, was found to have an impact on the stability of protein structure.

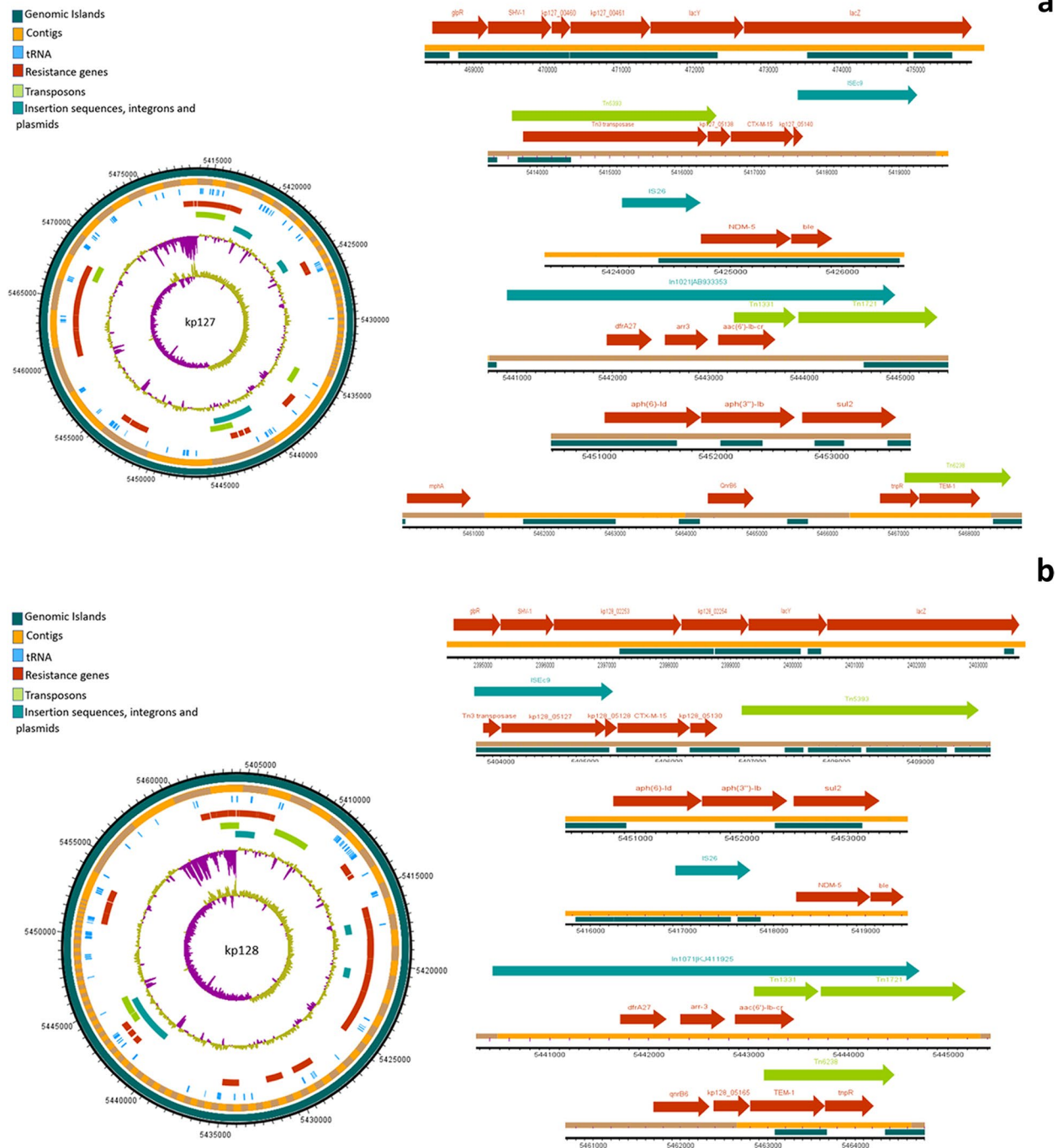


Fig. 7 Graphical representation of the genetic environment of the resistance genes and mobile genetic elements of kp127 (a) and kp128. The circular visualization of the whole genome sequence represents all the resistance genes and the MGEs. The horizontal representation shows the genetic environment of the resistance genes with specific importance to beta-lactamase genes and MGEs. The genetic environment of chromosomally encoded *bla_{SHV}* was found consistent in all the genomes. The *bla_{CTM-15}* and *bla_{TEM}* genes were mostly found on Tn3 transposons and ISEc9

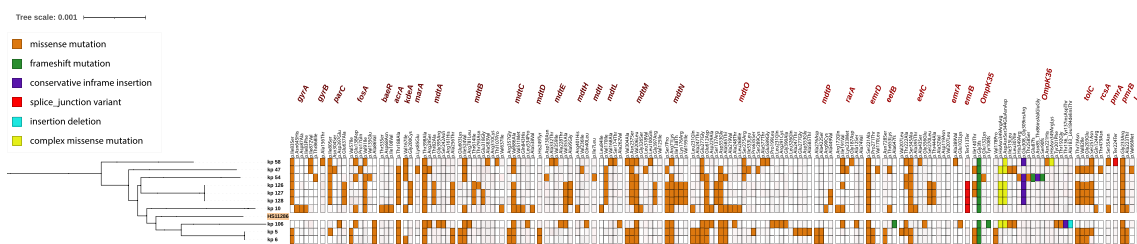


Fig. 8 Heatmap showing the Single Nucleotide Polymorphism (SNP) between the isolates obtained in this study. Different colors correspond to missense, insertion, frameshift, and complex variations. The maximum number of variations were observed in RND efflux pump gene *mdtO*. The isolates kp47 and kp58 that had the highest number of mutations

Prediction of functional & structural consequences of SNPs in OmpK36

Using the query sequence of Ompk36 proteins, SIFT conducted a sequence search with default parameters, and the median score selected was 3.00. The mutation at position serine 258 in Ompk36 was predicted to decrease protein stability when mutated. The prediction results from SIFT and MUpro are summarized in a Table 5.

Additionally, HOPE was used to explain the effect of mutations on protein function and structure. Ompk36 is a trimeric protein made up of 16-stranded β barrels that form three separate pores. The protein has long loops connecting the β strands on the outer side and short turns on the periplasmic side. Loop 3 plays a crucial role in narrowing the pore and is folded into the barrel. Mutations in these loops are considered critical for the protein’s proper functioning [24]. A mutation in OmpK36 at position 258, where threonine is replaced by serine, affects the hydrogen bond formation with Glutamic Acid at position 282. Additionally, the mutant residue is smaller than the wild-type residue, resulting in an incorrect position to form the same hydrogen bond. Consequently, external interactions may be lost due to the mutation.

Through the 3D structure analysis of the protein, it was observed that the wild-type residue is located within a β-strand, which is its preferred secondary structure. However, the mutant residue is not located in the same secondary structure, leading to a slight destabilization of the local conformation (Supplementary Fig. 2). Furthermore, this mutation creates a void space in the protein’s core, as shown in the Fig. 9.

Phylogenomic analysis

The phylogenomic analysis demonstrated substantial phylogenetic diversity among the sequenced *K. pneumoniae* isolates, as depicted in Fig. 10. Employing a core genome SNP-based approach, the analysis categorized global representative strains into five phylogroups. The sequenced genomes from our study were found to cluster

within four of these five phylogroups. In the clustering process, it was noted that the neonatal strain kp5 and adult strains kp6 and kp106 formed Group II, alongside strains from China and the USA. The water isolate kp10 was clustered in Group III, along with strains from China and India. Additionally, isolates kp54, kp126, kp127, and kp128 clustered in Group IV, aligning with isolates of Indian origin. Notably, two isolates, kp47 and kp58, were positioned in the largest phylogroup, Group V, alongside global reference strains (Kp1084 and NTUH-K2044). The neonatal strain kp47 clustered with global reference strains KCTC2242 and kp1084, originating from Korea and China, respectively. On a global scale, the sequence types (STs) clustered together with either identical or closely related STs from China, USA, and India, suggesting a shared genetic lineage among geographically diverse strains.

Discussion

The global prevalence of MDR/XDR *K. pneumoniae* has been consistently on the rise. In Pakistan, particularly, there has been a notable surge in the incidence of MDR/XDR *K.*

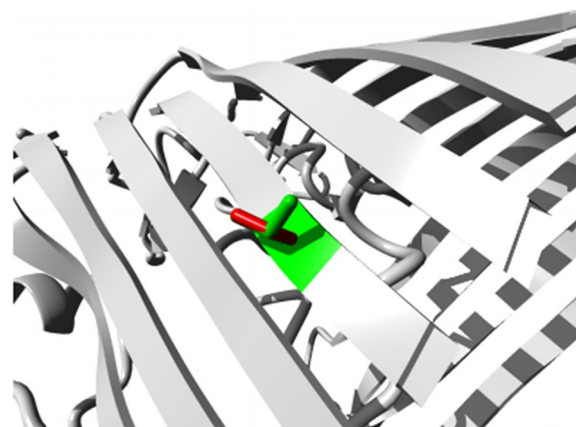


Fig. 9 Close-up of the mutation in the OmpK36 protein. The protein is colored grey, the side chains of both the wild-type and the mutant residue are shown and colored green and red respectively

pneumoniae, estimated at approximately 40%. This underscores the increasing concern regarding antibiotic-resistant *K. pneumoniae* infections in our country. Remarkably, the emergence of carbapenem-resistant *K. pneumoniae* (CRKP) has witnessed a significant increase, constituting 84.16% of reported cases in Pakistan [25]. During the investigation of antibiotic resistance, it was observed that the genotypic profiles of the sequenced isolates were fully consistent with their phenotypic profiles, indicating a strong association between the presence of ARGs and the observed resistance patterns. However, two strains, namely kp47 and kp58, exhibited a disparity between their genotypic and phenotypic profiles, suggesting the presence of additional mechanisms contributing to their antibiotic resistance beyond the detected resistance genes. This bacterial species has been known to be intrinsically resistant to penicillin because

of the core gene *bla_{SHV}* present in the chromosome [22]. All of our study isolates were resistant to ampicillin due to the presence of allelic variants of the *bla_{SHV}* gene in the core genome. Among these variants, *bla_{SHV-1}* along with *bla_{SHV-11}* were more prominent SHV variants [26]. The variant *bla_{SHV-187}* detected in one of the isolates has been recently documented in MDR *Klebsiella* strains in another study [27–29]. Furthermore, the occurrence of *bla_{CTX-M-15}*, *bla_{TEM-1}*, and *bla_{OXA}* genes, also attributed resistance to the beta-lactam class of antibiotics. Additionally, allelic variants of class B and D carbapenemase *bla_{NDM}* gene NDM-1, 4, and 5 and *bla_{OXA}* gene OXA 1, 10, and 232, respectively, were also detected, which have been recently reported from Pakistan isolates [30–32]. Among the plasmid-mediated quinolone resistance (PMQR) genes detected in our isolates, the most dominant gene was *oqxA*, present in 70% of the

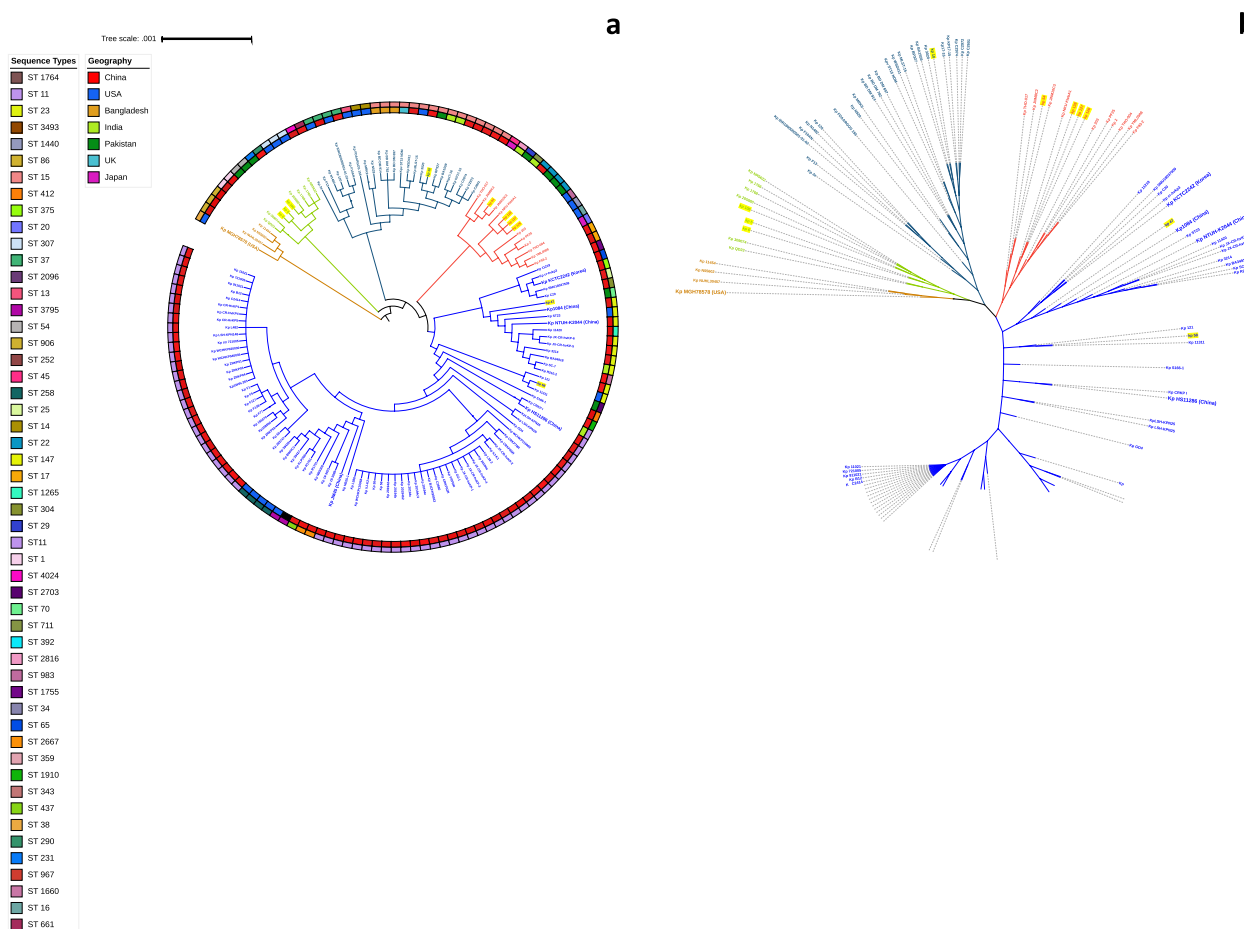


Fig. 10 **A** Phylogenetic analysis demonstrating the diversity of *Klebsiella pneumoniae* isolates from clinical and water samples. Midpoint-rooted circular representation of the phylogenetic tree, with branch colors representing the same clades. Outer ring is annotated with the sequence types (STs) while the inner ring represents the country from where the isolates were reported. The scale bar in the phylogenetic tree represents the number of substitutions per site, indicating the genetic distance between different strains. The study strains are highlighted in yellow to distinguish from the reference strains. The reference strains, which are widely recognized and extensively studied, are indicated in bold font. **B** Un-rooted approximate-maximum likelihood phylogenetic tree constructed based on a core-genome SNPs based alignment of the sequenced isolates from this study and reference genomes. The tree was refined and annotated using iTOL (<https://itol.embl.de/>)

isolates, which is consistent with the findings reported by Haeli et al. [25]. This was followed by the *qnrB* gene, present in 50% of the isolates, and the *aac(6')-Ib-cr* gene, present in 30% of the isolates. The least frequent PMQR gene observed in our study was *qnrS*, present in 20% of the isolates, which differs from previous reports [33]. Mutations in *gyrA* and *parC* have been recently reported in quinolone-resistant isolates [34, 35] that were also observed in our study, here the most frequent mutation was lle83Ser (70%) in *gyrA* and lle80Ser (60%) in *parC* gene. In our study, the most common aminoglycoside resistance genes were *aac(6')-Ib-cr*, *aph(3'')-Ib*, *aph(6)-Ia*, *aac(3)-IId* while *armA*, *aad2*, and *rmtB* were detected in two isolates, contributing High Level Aminoglycoside Resistance (HLAR) as reported by Zhang et al. [33]. Resistance genes like *LptD* and *msrE*, involved in LPS transport in ABC Transporter efflux system conferring resistance to carbapenems, rifamycin, and peptide antibiotics, have been reported elsewhere in isolates from the adult population [36, 37] were also identified in our neonatal strains. In our study samples, the biocide resistance gene *qacEdelta1*, commonly found in Gram-negative bacteria, was prevalent. This gene is associated with conferring resistance to aminoglycosides, trimethoprim, and sulfonamides through a multidrug efflux mechanism. Its presence suggests the potential for multidrug resistance development in the studied isolates, further highlighting the importance of understanding and monitoring the spread of biocide resistance genes in bacterial populations [38].

Occurrence of the *bla*_{CTX-M-15}, *bla*_{TEM}, *bla*_{OXA}, and *bla*_{SHV} ESBL genes in similar genetic contexts in our isolates and previous reports [11] indicates that antibiotic resistance is facilitated by either intrinsic or acquired genes found on the chromosome or mobile genetic elements (MGEs) [39]. Previous studies have highlighted the significant role of certain plasmids, such as IncF plasmids, and insertion sequence ISEc9, in enabling the widespread dissemination of genes like *bla*_{CTX-M-15}, *aac(6')-Ib-cr*, *bla*_{OXA-10} and *bla*_{TEM} through vertical gene transfer processes [22]. The study isolates kp54 and kp106, harbored globally prevailing IncF plasmids replicons that have been previously reported in *Enterobacteriaceae* often carrying virulence genes, and their role in HGT has already been confirmed [1, 11]. The IncC plasmid replicons that were identified in water isolate kp10 are particularly important as they harbor genes conferring resistance to antibiotic classes like chloramphenicol, sulfonamide, aminoglycoside, trimethoprim and AmpC β lactamase [1]. The globally distributed plasmid IncH1B(pNDM) was identified in two isolates, namely kp5, and kp6, which have previously been reported in Morocco, the United States, and the United Kingdom [1]. The strain kp47 exhibited a plasmid repB replicon that falls under the IncFIB family. This replicon is recognized for dynamic nature and plays a crucial role in facilitating the spread of antimicrobial resistance

within *Klebsiella* populations [40]. The presence of a wide variety of plasmid replicons carrying genetic elements associated with both resistance and virulence factors in our isolated strains is particularly noteworthy.

The identified integrons in our study exhibited a wide variety of gene cassettes, with particular emphasis on *aadA* and *dfrA* gene cassettes. These specific gene cassettes have been previously reported to play a significant part in the global spread of these mobile genetic elements [11]. In our study, it was observed that all the isolated strains contained Class I integrons, which have been extensively researched and are recognized for their significant contribution to the global spread of resistance genes. This is primarily due to their association with various mobile elements, further enhancing their ability to disseminate ARGs widely [41]. In two of our study isolates, we identified Integron 310, which has been previously reported to capture *arr-2* and *cmIA5* in *Acinetobacter baumannii* strains in South and East Asia as well as Portugal [42, 43]. The presence of Class I integrons in natural environments that have undergone evolutionary recombination events is well-documented, further emphasizing their importance in the acquisition and spread of antibiotic-resistance genes [42, 43]. The presence of two integrons In498 and In718 carrying *dfrA12* and *catB* genes, respectively, in water isolate kp10 also supports the role and spread of integrons in pathogenic bacteria. All the isolates also had composite and unit transposons particularly Tn5393 along with a wide range of insertion sequences (IS). The presence of ISEc9 in the vicinity of beta-lactamase genes suggests its potential involvement in the spread or mobilization of these resistance genes within the bacterial population. The transposons identified in our study were classified as members of the Tn3 family of replicative transposons. This group of transposons is widely distributed and known to play a significant role in the dissemination of antibiotic resistance [44]. The high diversity of transposons and insertion sequences can be attributed to the plasticity of genomes [11].

Based upon the diverse polysaccharide components with various K and O antigens, *K. pneumoniae* has been divided into 79 capsular types amongst which eight types have been described as markers for hypervirulent strains [45]. In our study, we observed a diverse range of sequence and capsular types among the isolates that belonged to different clonal groups consistent with previous findings. The capsular type K54 is well-known for its association with hypervirulence in *Klebsiella pneumoniae* strains. Specifically, it has been linked to sequence type ST29, which is recognized as a representative hypervirulent strain of *K. pneumoniae* belonging to CG29 [26, 46]. These findings highlight the diverse genetic characteristics and global distribution of different capsular types and sequence types in *K. pneumoniae*, underscoring the significance of clonal groups and their association with antimicrobial resistance and virulence

traits. The major virulence factors implicated in the pathogenicity and propensity to cause *Klebsiella pneumoniae* infections are capsule, lipopolysaccharides, fimbriae, and siderophores [3]. *RcsAB*, a two-component regulatory system that is involved in high capsule productivity [23], was identified in all the isolates. One of our study isolates had capsular type K54 one of the biomarkers for hvkp strains that had the least number of resistance genes but was found to be phenotypically more resistant and virulent. This can be explained as the hyper capsule serving as a physical barrier thus limiting the DNA uptake and consequently the HGT that's why they harbor few resistance genes but are more challenging therapeutically [23]. Our isolates were found to contain Outer membrane protein A (OmpA) and outer membrane porins OmpK35 and OmpK36, known for their contribution to virulence [23]. The virulence genes for core and acquired siderophores were identified as described previously [22, 26, 47, 48]. The genes for type I fimbriae (*fimA-fimK*) and type III fimbriae (*mrkB-mrkI*) and pili (*ecpA*) have established roles in biofilm formation and UTI [47, 49] and were present in all the isolates. In addition to fimbriae and pili, we have detected genetic elements like adhesion genes *pgaABC* and *bcsA*, *luxS* in quorum sensing, *treC* and *sugE* for capsule production, the combined effect of which have been reported crucial to produce an excellent biofilm [50, 51]. Several snps were observed in *treC* and *wza* genes in all the isolates as previously reported [52]. In order to further investigate the impact of these variants among the sequenced isolates, additional knock-out studies are necessary. The highest number of variations were observed in *mdtO* gene, involved in antibiotic efflux. Previous reports have shown that this gene facilitates antibiotic resistance and deletion of *mdtO* increases the sensitivity of the bacteria towards antibiotics [53]. The variant calling showed that Asp553Gly (90%) was the most common mutation followed by Glu612Asp (80%) which was observed in *mdtO* gene. Resistance to carbapenems in *K. pneumoniae* relies on several crucial factors, including modifications to the major outer membrane porins, OmpK35 and OmpK36. Additionally, the presence of extended-spectrum β -lactamases (ESBLs) or plasmid-encoded carbapenemases plays a significant role in conferring resistance to carbapenems in this bacterium [54]. A common finding in many studies consistent with our findings is the identification of non-functional porins resulting from frameshift mutations leading to truncated OmpK35 resulting in a shortened protein than the wild type [54–56]. The structural modifications observed in OmpK36 are caused by the insertion of amino acids into a specific region of the porin, known as loop 3 (L3). This results in a reduction in the diameter of the luminal space and restricted diffusion of substrates, such as antibiotics. The L3 insertions, are frequently observed in clinical isolates of *Klebsiella pneumoniae*. They have been identified

in approximately 12.3% (192 out of 1557) of isolates in a diverse collection of publicly available genomes [13]. Due to the high prevalence of L3 insertions in carbapenemase-producing strains of *K. pneumoniae*, a study has been initiated to examine the specific implications of these mutations on the efficacy of newly developed or recently approved drugs that target resistant strains of *K. pneumoniae* [13]. There is currently limited information available from the Indian subcontinent about the frequency of mutations in porins, OmpK35 and OmpK36, as well as the prevalence of porin deletions in *K. pneumoniae* [57]. However, given the high prevalence of XDR isolates in Pakistan, it is essential to thoroughly investigate all possible resistance mechanisms. Moreover, understanding the significance of these mutations and their effects is essential for developing strategies to combat bacterial infections.

Conclusion

In this study, we present comprehensive data from the whole genome sequences of *K. pneumoniae* strains isolated from neonatal, adults, and tap water from a tertiary care hospital in Pakistan. We observed significant genomic diversity amongst our sequenced isolates. The water isolate exhibited similar resistance determinants and the highest number of virulence genes, a finding of particular concern. Our results indicates high burden of antibiotic resistance and virulence associated genes among *K. pneumoniae* bacterial isolates, which are putatively the main culprit to poor outcome of antibiotic treatment leading to high mortality rate. The presence of various MGEs in close association with the ARGs indicates their role in HGT between diverse organisms. We also found that spontaneous mutations in porin related genes specific for the active uptake of carbapenem antibiotic e.g. OmpK35 and OmpK36 lead to the loss of porin protein function, thus causing increased bacterial resistance toward carbapenems. The phylogenomic analysis revealed the close association of our isolates with strains from China, India and USA, indicating the global spread of MDR and XDR sequence types. This study shows an urgent need for epidemiological and molecular studies to develop a better understanding of antibiotic resistance dynamics and for guiding potential treatment of multi-drug resistant *K. pneumoniae* infections.

Materials and methods

Microbiology of bacterial isolates

We collected 215 *Klebsiella pneumoniae* isolates from a local tertiary care hospital during February 2019–February 2021 in Pakistan and classified these bacterial strains on the basis of their phenotypic antimicrobial resistance profiles. The isolates were collected from diverse age groups with the maximum number of isolates ($n=95$) collected from neonates and infants accounting for 45% of the collection.

Among 215 isolates, 64 were collected from blood, 58 from pus, 41 from urine, 12 from non-bronchoscopic lung lavage (NBL), 9 each from sputum & tips, 6 from fluids, 4 from swabs, 2 from tissue and 8 from hospital tap water.

AST profiling & biofilm assay

The Kirby-Bauer disk diffusion method on Mueller-Hinton agar (Hardy Diagnostics) was used for antibiotic susceptibility testing (AST) following the Clinical and Laboratory Standards Institute (CLSI 2018) M100-S28 guidelines [58]. The isolates were tested against various classes of antibiotics including penicillin, aminoglycoside, fluoroquinolone, cephalosporin, carbapenem, phenicols and fosfomycin.

Detection of biofilm formation was performed by methods previously described [59]. Overnight cultures in fresh LB broth were inoculated in each well of a sterile 96-well polystyrene plate. After static incubation at 37°C, the bacteria were stained for 20 mins with 0.5% crystal violet dye. After discarding the supernatant and washing with deionized water was performed thrice thus removing unattached cells. Later on, 95% ethanol was used for the elution of the dye, and the optical density was determined at OD_{550nm}. NTUH-K2044 strain, notorious as biofilm former, was selected as a positive control [59]. The biofilm assay was conducted in triplicates for each isolate.

Selection of isolates for whole genome sequencing

Out of our bacterial cohort, a subset of 10 isolates was selected for whole genome sequencing (WGS) based on antimicrobial resistance profile and source by keeping the selection criteria consistent with the previous reports [36, 60–62]. The randomized selection of samples was performed from each subset of samples on the basis of host age (i.e. neonatal samples $n=7$, Adults $n=2$), antimicrobial susceptibility testing (AST) profile, and the source of the sample. The selected *K. pneumoniae* strains belonged to diverse sample types e.g. blood ($n=7$), pus ($n=1$), urine ($n=1$), and hospital water ($n=1$). Only meropenem resistant sample ($n=1$) was selected from tap water source for sequencing. The selection of the water isolate was undertaken with the primary objectives of elucidating the reservoirs harboring resistance genes, comprehending the dynamics of pathogen transmission, and conducting an in-depth characterization of genetic diversity. The strains were investigated due to their phenotypic resistance to major classes of antibiotics and variable biofilm formatting ability.

Genomic DNA preparation, sequencing, and assembly

The samples were first inoculated onto blood agar and incubated overnight. Subsequently, morphologically distinct colonies were streaked on MacConkey agar and

incubated at 37°C. The genomic DNA of the strains was isolated by using GeneJet Genomic DNA Purification Kit (Thermo Scientific). The concentration of the isolated genomic DNA was measured using a Nanodrop2000 UV-Vis Spectrophotometer (Thermo Scientific). Agarose gel electrophoresis with 1% (w/v) agarose gel was performed to evaluate DNA integrity. Nextera XT Library Prep Kit (Illumina, San Diego, USA) was used for preparing genomic DNA libraries [63]. Hamilton Microlab STAR automated liquid handling system was used for DNA quantification and library preparation. Kapa Biosystems Library Quantification Kit for Illumina on a Roche light cycler 96 qPCR machine was used for the quantification of the pooled libraries. Sequencing of the libraries was done on the Illumina HiSeq using a 250bp paired-end reads protocol. During the downstream processing of raw reads, Trimmomatic v0.30 was used to trim the adapter sequences from the reads [64]. De-novo assembly was performed using SPAdes v3.7 [65] and annotation of contigs was done by Prokka v1.11 [66].

In-silico analysis of bacterial genomes

The assembled sequences were annotated using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) (https://www.ncbi.nlm.nih.gov/genome/annotation_prok/). Comprehensive Antibiotic Resistance Database (CARD) and Resfinder were used for the prediction of antibiotic resistance genes [67, 68]. MLST 2.0 (<https://cge.cbs.dtu.dk/services/MLST/>) and PHYLOViZ 2.0 [69] were used to identify the sequence types and clonal groups (CG) of the isolates, respectively. BacAnt and Integron Finder were used to find out the transposons, integrons, and gene cassettes present within the genomes [67, 68]. Insertion sequences were identified using MGE and IS Finder [70, 71]. For the identification of plasmid replicons, PlasmidFinder was used [72]. Capsular typing was performed using Kaptive Web whereas Virulome was identified using VFDB (Virulence Factors Database) [73]. PHASTER (Phage Search Tool), an online server was used for the detection of intact prophage sequences [74].

SNP calling

In addition to acquired antibiotic resistance genes, we investigated genes that play a role in antibiotic resistance, including efflux pumps, genes responsible for antibiotic inactivation, target alteration, outer membrane proteins, lipid A protein modification, and reduction in membrane permeability. Substitutions and indels (insertions and deletions) were identified by performing Single Nucleotide Polymorphism (SNP) calling using snippy 4.6.0. On the basis of literature review 43 genes were selected

involved in antibiotic efflux for the detection of the variants and their putative role in contributing towards antibiotic resistance. High-quality non-synonymous SNPs were examined manually and were used for interpretation.

Prediction of functional & structural consequences of SNPs in OmpK36

To check to functional consequences of snps in porin OmpK36, various computational tools were utilized to examine how non-synonymous SNPs (nsSNPs) affect the structure and function, and their impact on antibiotic resistance. Sorting Intolerant from Tolerant (SIFT) a program was employed to predict the phenotypic impacts of nsSNPs [75]. Similarly, the functional impacts of mutations were predicted using SNAP which employs a neural network (NN) as a machine learning technique. SNAP uses evolutionary information from multiple sequence alignments to differentiate affected variants from neutral ones. If the annotation is available, the tool also calculates structural features such as secondary structure and solvent accessibility. SNAP achieved an accuracy of 82%, as determined by cross-validation of over 100,000 experimentally annotated variants, using sustained two-state accuracy (effect/neutral) [76].

The impact of a mutation on protein 3D structure was investigated through the use of Have Our Protein Explained (HOPE), a web-based application that leverages structural information of a specific protein to predict the structural effects of a mutation [77]. To evaluate the influence of nsSNPs on protein structure stability, the web-based server MUpPro was utilized <https://mupro.proteomics.ics.uci.edu/> [78]. MUpPro, a computational tool, predicts the effect of a single-site amino acid mutation on protein stability. It achieves an accuracy rate of 84% as determined by 20-fold cross-validation. This set of machine learning programs includes Support Vector Machines (SVM) and Neural Networks (NN) [79].

Construction of homology-based models of mutant proteins

Homology modeling was used to predict mutant models of each SNP, by mapping the SNPs on the 3-dimensional structure of a protein. The selection of a template molecule was based on the results of the Basic Local Alignment Search Tool (BLAST) and the best *E*-value [80]. The 6RCP molecule (crystal structure of the ompk36 clinical isolate ST258 from *K. pneumoniae*) was chosen as the template One hundred models were generated for each mutation using Modeller V10.4 [81]. The Discrete Optimized Protein Energy (DOPE) score is a statistical potential used in protein structure modeling to assess the

quality of a given model. The DOPE score is based on statistical distributions of atom-atom interactions observed in a large set of known protein structures [82]. The top-performing models were selected based on their DOPE scores. To assess the effect of mutations on the protein's three-dimensional structure, the mutant models were individually compared to the wild-type model, and the root-mean-square deviation (RMSD) was calculated. This allowed for quantifying the differences in spatial arrangement between the mutant and wild-type structures.

Phylogenomic analysis

Phylogenetic analysis of the sequenced genomes was conducted alongside complete whole genomes of antibiotic-resistant *Klebsiella pneumoniae* strains obtained from different geographical locations using PATRIC (www.patric.org) and NCBI (<https://ncbi.nlm.nih.gov/>) available till June 2021, to obtain a current epidemiological relatedness and evolutionary analysis.

The pangenome of the sequences was analyzed using Bacterial Pangenome Analysis Pipeline (BPGA). The core-genome alignment file output generated from BPGA was used for the construction of maximum likelihood (ML) tree with FastTree v2.1.10. For visualization and annotation of phylogenetic trees Interactive tree of life (iTOL) was used [83].

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-024-10139-y>.

Supplementary Material 1.

Acknowledgements

We would like to express our gratitude to Mahnoor Majid for her invaluable assistance.

Authors' contributions

SR, SI and SA designed the study. SR, SI and IM performed the analyses. SR and SI drafted the manuscript. HI executed the structural effects of mutations in porin proteins. IM and MM helped in figure making. SA supervised the study. SI, AR and MJ revised the manuscript critically for intellectual input and corrections. All authors read and approved the final manuscript.

Funding

This study was supported by Atta ur Rahman School of Applied Biosciences (ASAB), NUST.

Availability of data and materials

All data generated or analyzed during this study are included in this article [and its supplementary information files]. The whole genome sequences used for phylogenomic analysis are publicly available and retrieved from NCBI (<https://www.ncbi.nlm.nih.gov/genome/?term=klebsiella+pneumoniae%2B>). The Whole Genome Shotgun project of the individual isolate has been deposited at DDBJ/ENA/GenBank under the accession numbers and the version described in this paper is as follow.

Isolate	Bioproject	Web link
kp5	PRJNA743648	https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA743648
kp6	PRJNA744871	https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA744871
kp10	PRJNA826469	https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA826469
kp47	PRJNA814905	https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA814905
kp54	PRJNA766411	https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA766411
kp58	PRJNA817217	https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA817217
kp106	PRJNA827029	https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA827029
kp126	PRJNA832736	https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA832736
kp127	PRJNA832741	https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA832741
kp128	PRJNA833061	https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA833061

Declarations

Ethics approval and consent to participate

The study was approved by the Ethics Committee of the Atta ur Rahman School of Applied Biosciences (ASAB, NUST) (approval number IRB 133). The requirement for patient informed consent was waived by the Ethics Committee of the Atta ur Rahman School of Applied Biosciences because only bacterial isolates recovered from routine diagnostic laboratory tests were assessed. Clinical isolates and patient records/information were anonymous and de-identified prior to analysis. All experiments were performed in accordance with the Declaration of Helsinki.

Consent for publication

"Not applicable".

Competing interests

The authors declare no competing interests.

Author details

¹Atta-ur-Rahman School of Applied Biosciences (ASAB), National University of Sciences and Technology, Islamabad, Pakistan. ²Research Centre for Modeling & Simulation (RCMS), National University of Sciences and Technology, Islamabad, Pakistan. ³Department of Microbiology, Abdul Wali Khan University, Mardan, Pakistan.

Received: 4 June 2023 Accepted: 19 February 2024

Published online: 25 April 2024

References

- Navon-Venezia S, Kondratyeva K, Carattoli A. Klebsiella pneumoniae: a major worldwide source and shuttle for antibiotic resistance. *FEMS Microbiol Rev.* 2017;41:252–75.
- Evrard B, Balestrino D, Dosgilbert A, Bouya-Gachancard JJJ, Charbonnel N, Forestier C, et al. Roles of capsule and lipopolysaccharide O antigen in interactions of human monocyte-derived dendritic cells and Klebsiella pneumoniae. *Infect Immun.* 2010;78:210–9.
- Effah CY, Sun T, Liu S, Wu Y. Klebsiella pneumoniae: an increasing threat to public health. *Ann Clin Microbiol Antimicrob.* 2020;19:1.
- Shrivastava SR, Shrivastava PS, Ramasamy J. World health organization releases global priority list of antibiotic-resistant bacteria to guide research, discovery, and development of new antibiotics. *J Med Soc.* 2018;32(1):76–7.
- Malaki M. Klebsiella pneumoniae bacteremia presenting on a neonatal intensive care unit during the first week of life. *J Pediatr Infect Dis.* 2015;1:12–5.
- Saleem AF, Qamar FN, Shahzad H, Qadir M, Zaidi AKM. Trends in antibiotic susceptibility and incidence of late-onset Klebsiella pneumoniae neonatal sepsis over a six-year period in a neonatal intensive care unit in Karachi, Pakistan. *Int J Infect Dis.* 2013;17:e961–5.
- Paczosa MK, Meccas J. Klebsiella pneumoniae: going on the offense with a strong defense. *Microbiol Mol Biol Rev.* 2016;80:629–61.
- Ballot DE, Bandini R, Nana T, Bosman N, Thomas T, Davies VA, et al. A review of -multidrug-resistant Enterobacteriaceae in a neonatal unit in Johannesburg, South Africa. *BMC Pediatr.* 2019;19:320.
- Bharatham N, Bhowmik P, Aoki M, Okada U, Sharma S, Yamashita E, et al. Structure and function relationship of OqxB efflux pump from *Klebsiella pneumoniae*. *Nat Commun.* 2021;12(12):1–12.
- Yang Y, Higgins CH, Rehman I, Galvao KN, Brito IL, Bicalho ML, et al. Genomic diversity, virulence, and antimicrobial resistance of Klebsiella pneumoniae strains from cows and humans. *Appl Environ Microbiol.* 2019;85.
- Mbelle NM, Feldman C, Sekyere JO, Maningi NE, Modipane L, Essack SY. Pathogenomics and evolutionary epidemiology of multi-drug resistant clinical Klebsiella pneumoniae isolated from Pretoria, South Africa. *Sci Rep.* 2020;10:1–17.
- Lumbreras-Iglesias P, Rodicio MR, Valledor P, Suárez-Zarracina T, Fernández J. High-level carbapenem resistance among oxa-48-producing klebsiella pneumoniae with functional ompk36 alterations: maintenance of ceftazidime/avibactam susceptibility. *Antibiotics.* 2021;10:1174.
- David S, Wong JLC, Sanchez-Garrido J, Kwong HS, Low WW, Morecchiato F, et al. Widespread emergence of OmpK36 loop 3 insertions among multidrug-resistant clones of Klebsiella pneumoniae. *PLoS Pathog.* 2022;18(7):e1010334. <https://doi.org/10.1371/journal.ppat.1010334>.
- Norsigian CJ, Attia H, Szubin R, Yassin AS, Palsson BØ, Aziz RK, et al. Comparative genome-scale metabolic modeling of Metallo-Beta-lactamase-producing multidrug-resistant Klebsiella pneumoniae clinical isolates. *Front Cell Infect Microbiol.* 2019;9.
- Holt KE, Wertheim H, Zadoks RN, Baker S, Whitehouse CA, Dance D, et al. Genomic analysis of diversity, population structure, virulence, and antimicrobial resistance in Klebsiella pneumoniae, an urgent threat to public health. *Proc Natl Acad Sci USA.* 2015;112:E3574–81.
- Wyres KL, Holt KE. Klebsiella pneumoniae population genomics and antimicrobial-resistant clones. *Trends Microbiol.* 2016;24:944–56.
- Quainoo S, Coolen JPM, van Hijum SAFT, Huynen MA, Melchers WJG, van Schaik W, et al. Whole-genome sequencing of bacterial pathogens: the future of nosocomial outbreak analysis. *Clin Microbiol Rev.* 2017;30:1015–63.
- Magiorakos AP, Srinivasan A, Carey RB, Carmeli Y, Falagas ME, Giske CG, et al. Multidrug-resistant, extensively drug-resistant and pandrug-resistant bacteria: an international expert proposal for interim standard definitions for acquired resistance. *Clin Microbiol Infect.* 2012;18:268–81.
- Alikhan NF, Petty NK, Ben Zakour NL, Beatson SA. BLAST ring image generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics.* 2011;12:1–10.
- Andrade LN, Novais A, Stegani LMM, Ferreira JC, Rodrigues C, Darini AL, et al. Virulence genes, capsular and plasmid types of multidrug-resistant CTX-M(–2, –8, –15) and KPC-2-producing Klebsiella pneumoniae isolates from four major hospitals in Brazil. *Diagn Microbiol Infect Dis.* 2018;91:164–8.

21. Horváth M, Kovács T, Koderivalappil S, Ábrahám H, Rákhely G, Schneider G. Identification of a newly isolated lytic bacteriophage against K24 capsular type, carbapenem resistant *Klebsiella pneumoniae* isolates. *Sci Rep*. 2020;10:1–11.
22. Wyres KL, Lam MMC, Holt KE. Population genomics of *Klebsiella pneumoniae*. *Nat Rev Microbiol*. 2020;18:344–59.
23. Zhu J, Wang T, Chen L, Du H. Virulence factors in Hypervirulent *Klebsiella pneumoniae*. *Front Microbiol*. 2021;12 April:1–14.
24. Vergalli J, Bodrenko IV, Masi M, Moynié L, Acosta-Gutiérrez S, Naismith JH, et al. Porins and small-molecule translocation across the outer membrane of gram-negative bacteria. *Nat Rev Microbiol*. 2020;18:164–76.
25. Asri NAM, Ahmad S, Mohamad R, Hanafi NM, Zaidi NFM, Irekeola AA, et al. Global Prevalence of Nosocomial Multidrug-Resistant *Klebsiella pneumoniae*: A Systematic Review and Meta-Analysis. *Antibiot*. 2021;10:1508.
26. Lee CR, Lee JH, Park KS, Jeon JH, Kim YB, Cha CJ, et al. Antimicrobial resistance of hypervirulent *Klebsiella pneumoniae*: epidemiology, hypervirulence-associated determinants, and resistance mechanisms. *Front Cell Infect Microbiol*. 2017;7.
27. Tian X, Wang Q, Perlaza-Jiménez L, Zheng X, Zhao Y, Dhanasekaran V, et al. First description of antimicrobial resistance in carbapenem-susceptible *Klebsiella pneumoniae* after imipenem treatment, driven by outer membrane remodeling. *BMC Microbiol*. 2020;20:1–11.
28. Perdigão J, Caneiras C, Elias R, Modesto A, Spadar A, Phelan J, et al. Genomic epidemiology of carbapenemase producing *klebsiella pneumoniae* strains at a northern portuguese hospital enables the detection of a misidentified *klebsiella varicola* kpc-3 producing strain. *Microorganisms*. 2020;8:1–18.
29. Legese MH, Asrat D, Aseffa A, Hasan B, Mihret A, Swedberg G. Molecular Epidemiology of Extended-Spectrum Beta-Lactamase and AmpC Producing Enterobacteriaceae among Sepsis Patients in Ethiopia: A Prospective Multicenter Study. *Antibiot*. 2022;11:131.
30. Imtiaz W, Dasti JJ, Andrews SC. Draft genome sequence of a carbapenemase-producing (NDM-1) and multidrug-resistant, hypervirulent *Klebsiella pneumoniae* ST11 isolate from Pakistan, with a non-hypermucoviscous phenotype associated with *rmpA2* mutation. *J Glob Antimicrob Resist*. 2021;25:359–62.
31. Qamar MU, Walsh TR, Toleman MA, Tyrrell JM, Saleem S, Aboklaish A, et al. Dissemination of genetically diverse NDM-1, –5, –7 producing-Gram-negative pathogens isolated from pediatric patients in Pakistan. *Future Microbiol*. 2019;14:691–704. <https://doi.org/10.2217/fmb-2019-0012>.
32. Gondal AJ, Saleem S, Jahan S, Choudhry N, Yasmin N. Novel Carbapenem-resistant *Klebsiella pneumoniae* ST147 Coharboring blaNDM-1, blaOXA-48 and extended-Spectrum β -lactamases from Pakistan. *Infect Drug Resist*. 2020;13:2105.
33. Mohammed Kareem S, S Al-kadmy IM, Kazaal SS, Mohammed Ali AN, Najj Aziz S, Makharita RR, et al. Detection of *gyrA* and *parC* mutations and prevalence of plasmid-mediated quinolone resistance genes in *Klebsiella pneumoniae*. *Taylor Fr* 2021;14:555–563.
34. Hudson CM, Bent ZW, Meagher RJ, Williams KP. Resistance determinants and Mobile genetic elements of an NDM-1-encoding *Klebsiella pneumoniae* strain. *PLoS One*. 2014;9:e99209.
35. Kareem SM, Al-Kadmy IMS, Kazaal SS, Ali ANM, Aziz SN, Makharita RR, et al. Detection of *gyrA* and *parC* mutations and prevalence of plasmid-mediated quinolone resistance genes in *Klebsiella pneumoniae*. *Infect Drug Resist*. 2021;14:555.
36. Surleac M, Barbu IC, Paraschiv S, Popa LI, Gheorghe I, Marutescu L, et al. Whole genome sequencing snapshot of multi-drug resistant *Klebsiella pneumoniae* strains from hospitals and receiving wastewater treatment plants in southern Romania. *PLoS One*. 2020;15:e0228079.
37. Altayb HN, Elbadawi HS, Baothman O, Kazmi I, Alzahrani FA, Nadeem MS, et al. Genomic Analysis of Multidrug-Resistant Hypervirulent (Hypermucoviscous) *Klebsiella pneumoniae* Strain Lacking the Hypermucoviscous Regulators (*rmpA/rmpA2*). *Antibiot*. 2022;11:596.
38. Fraise AP. Biocide abuse and antimicrobial resistance—a cause for concern? *J Antimicrob Chemother*. 2002;49:11–2.
39. Founou RC, Founou LL, Allam M, Ismail A, Essack SY. Whole genome sequencing of extended Spectrum β -lactamase (ESBL)-producing *Klebsiella pneumoniae* isolated from hospitalized patients in KwaZulu-Natal, South Africa. *Sci Rep*. 2019;9:1–11.
40. Nazir A, Zhao Y, Li M, Manzoor R, Tahir RA, Zhang X, et al. Structural genomics of repA, repB1-carrying IncFIB family pA1705-qnrS, P911021-tetA, and P1642-tetA, multidrug-resistant plasmids from *Klebsiella pneumoniae*. *Infect Drug Resist*. 2020;13:1889.
41. Ghaly TM, Chow L, Asher AJ, Waldron LS, Gillings MR. Evolution of class 1 integrons: mobilization and dispersal via food-borne bacteria. *PLoS One*. 2017;12:e0179169.
42. Revathi G, Siu LK, Lu PL, Huang LY. First report of NDM-1-producing *Acinetobacter baumannii* in East Africa. *Int J Infect Dis*. 2013;17.
43. Stalder T, Barraud O, Casellas M, Dagot C, Ploy MC. Integron involvement in environmental spread of antibiotic resistance. *Front Microbiol*. 2012;3 APR:1–14.
44. Nicolas E, Lambin M, Dandoy D, Galloy C, Nguyen N, Oger CA, et al. The Tn 3-family of replicative transposons. *Microbiol Spectr*. 2015;3.
45. Zhao L, Xia X, Yuan T, Zhu J, Shen Z, Li M. Molecular Epidemiology of Antimicrobial Resistance, Virulence and Capsular Serotypes of Carbapenemase-Carrying *Klebsiella pneumoniae* in China. *Antibiot*. 2022;11:1100.
46. Turton JF, Payne Z, Micah K, Turton JA. Capsular type K54, clonal group 29 and virulence plasmids: an analysis of K54 and non-K54 closely related isolates of *Klebsiella pneumoniae*. *Epidemiol Infect*. 2018;146:1813–23.
47. Li B, Zhao Y, Liu C, Chen Z, Zhou D. Molecular pathogenesis of *Klebsiella pneumoniae*. *Future Microbiol*. 2014;9:1071–81. <https://doi.org/10.2217/fmb1448>.
48. Khaertynov KS, Anokhin VA, Rizvanov AA, Davidyuk YN, Semyenova DR, Lubin SA, et al. Virulence factors and antibiotic resistance of *Klebsiella pneumoniae* strains isolated from neonates with sepsis. *Front Med*. 2018;5 AUG:1–9.
49. Schroll C, Barken KB, Krogfelt KA, Struve C. Role of type 1 and type 3 fimbriae in *Klebsiella pneumoniae* biofilm formation. *BMC Microbiol*. 2010;10.
50. Urooj M, Shoukat M, Imran M, Ansar M, Faryal R. Emergence of hypermucoviscous colistin-resistant high-risk convergent *Klebsiella pneumoniae* ST-2096 clone from Pakistan. *Future Microbiol*. 2022;17:989–1000.
51. Shadkam S, Goli HR, Mirzaei B, Gholami M, Ahanjan M. Correlation between antimicrobial resistance and biofilm formation capability among *Klebsiella pneumoniae* strains isolated from hospitalized patients in Iran. *Ann Clin Microbiol Antimicrob*. 2021;20:1–7.
52. Devanga Ragupathi NK, Muthuirulandi Sethuvel DP, Triplicane Dwarakanathan H, Murugan D, Umashankar Y, Monk PN, et al. The influence of biofilms on Carbapenem susceptibility and patient outcome in device Associated K. *Pneumoniae* infections: insights into phenotype vs genome-wide analysis and correlation. *Front Microbiol*. 2020;11:3220.
53. Miryala SK, Ramaiah S. Exploring the multi-drug resistance in *Escherichia coli* O157:H7 by gene interaction network: a systems biology approach. *Genomics*. 2019;111:958–65.
54. Hamzaoui Z, Ocampo-Sosa A, Fernandez Martinez M, Landolsi S, Ferjani S, Maamar E, et al. Role of association of *OmpK35* and *OmpK36* alteration and blaESBL and/or blaAmpC genes in conferring carbapenem resistance among non-carbapenemase-producing *Klebsiella pneumoniae*. *Int J Antimicrob Agents*. 2018;52(6):898–905. <https://doi.org/10.1016/j.ijantimicag.2018.03.020>.
55. Humphries RM, Hemarajata P. Resistance to ceftazidime-avibactam in *Klebsiella pneumoniae* due to porin mutations and the increased expression of KPC-3. *Antimicrob Agents Chemother*. 2017;61(6):10–1128.
56. Balabanian G, et al. Effect of porins and bla KPC expression on activity of imipenem with relebactam in *Klebsiella pneumoniae*: can antibiotic combinations overcome resistance? *Microb Drug Resist*. 2018;24(7):877–81.
57. Shankar C, Kumar S, Venkatesan M, Veeraraghavan B. Emergence of ST147 *Klebsiella pneumoniae* carrying blaNDM-7 on IncA/C2 with *ompK35* and *ompK36* mutations in India. *J Infect Public Health*. 2019;12:741–3.
58. Weinstein MP, Lewis JS. The clinical and laboratory standards institute subcommittee on antimicrobial susceptibility testing: background, organization, functions, and processes. *J Clin Microbiol*. 2020;58.
59. Zheng J, Lin Z, Chen C, Chen Z, Lin F, Wu Y, et al. Biofilm Formation in *Klebsiella pneumoniae* Bacteremia Strains Was Found to be Associated with CC23 and the Presence of *wcaG*. *Front Cell Infect Microbiol*. 2018;8 FEB:21.

60. Fergestad ME, Touzain F, De Vlieghe S, De Visscher A, Thiry D, Ngassam Tchamba C, et al. Whole genome sequencing of staphylococci isolated from bovine Milk samples. *Front Microbiol.* 2021;12:3976.
61. Dixit A, Freschi L, Vargas R, Calderon R, Sacchetti J, Drobniewski F, et al. Whole genome sequencing identifies bacterial factors affecting transmission of multidrug-resistant tuberculosis in a high-prevalence setting. *Sci Rep.* 2019;9.
62. Luo K, Tang J, Qu Y, Yang X, Zhang L, Chen Z, et al. Nosocomial infection by *Klebsiella pneumoniae* among neonates: a molecular epidemiological study. *J Hosp Infect.* 2021;108:174–80.
63. Baym M, Kryazhimskiy S, Lieberman TD, Chung H, Desai MM, Kishony RK. Inexpensive multiplexed library preparation for Megabase-sized genomes. *PLoS One.* 2015;10:e0128036.
64. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* 2014;30:2114–20.
65. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol.* 2012;19:455–77.
66. Prokka ST. Rapid prokaryotic genome annotation. *Bioinformatics.* 2014;30:2068–9.
67. McArthur AG, Waglechner N, Nizam F, Yan A, Azad MA, Baylay AJ, et al. The comprehensive antibiotic resistance database. *Antimicrob Agents Chemother.* 2013;57:3348–57.
68. Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O, et al. Identification of acquired antimicrobial resistance genes. *J Antimicrob Chemother.* 2012;67:2640–4.
69. Pagès JM, James CE, Winterhalter M. The porin and the permeating antibiotic: a selective diffusion barrier in Gram-negative bacteria. *Nat Rev Microbiol.* 2008;6:893–903.
70. Siguier P, Perochon J, Lestrade L, Mahillon J, Chandler M. ISfinder: the reference Centre for bacterial insertion sequences. *Nucleic Acids Res.* 2006;34 Database issue.
71. Hua X, Liang Q, Deng M, He J, Wang M, Hong W, et al. BacAnt: a combination annotation server for bacterial DNA sequences to identify antibiotic resistance genes, Integrons, and transposable elements. *Front Microbiol.* 2021;12.
72. Carattoli A, Zankari E, García-Fernández A, Larsen MV, Lund O, Villa L, et al. In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. *Antimicrob Agents Chemother.* 2014;58:3895–903.
73. Chen L, Zheng D, Liu B, Yang J, Jin Q. VFDB 2016: hierarchical and refined dataset for big data analysis—10 years on. *Nucleic Acids Res.* 2016;44:D694–7.
74. Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, et al. PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res.* 2016;44:W16–21.
75. Ng PC, Henikoff S. SIFT: predicting amino acid changes that affect protein function. *Nucleic Acids Res.* 2003;31:3812–4.
76. Bromberg Y, Rost B. SNAP: predict effect of non-synonymous polymorphisms on function. *Nucleic Acids Res.* 2007;35:3823.
77. Venselaar H, te Beek TAH, Kuipers RKP, Hekkelman ML, Vriend G. Protein structure analysis of mutations causing inheritable diseases. An e-science approach with life scientist friendly interfaces. *BMC Bioinform.* 2010;11.
78. Prediction of Protein Stability Changes upon Mutations. <https://mupro.proteomics.ics.uci.edu/>. Accessed 20 Dec 2023.
79. Cheng J, Randall A, Baldi P. Prediction of protein stability changes for single-site mutations using support vector machines. *Proteins.* 2006;62:1125–32.
80. Johnson M, Zaretskaya I, Raytselis Y, Merezuk Y, McGinnis S, Madden TL. NCBI BLAST: a better web interface. *Nucleic Acids Res.* 2008;36 Web Server issue.
81. Šali A, Potterton L, Yuan F, van Vlijmen H, Karplus M. Evaluation of comparative protein modeling by MODELLER. *Proteins Struct Funct Bioinforma.* 1995;23:318–26.
82. Maximiano MR, Rezende SB, Rios TB, Leite ML, Vilas Boas LCP, da Cunha NB, et al. Screening for cysteine-stabilized scaffolds for developing proteolytic-resistant AMPs. *Methods Enzymol.* 2022;663:67–98.
83. Letunic I, Bork P. Interactive tree of life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics.* 2007;23:127–8.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.