scientific reports

Check for updates

OPEN

Genomic convergence of multidrug resistance, virulence-associated loci, and phage defense systems in *Klebsiella pneumoniae* from pharmaceutical wastewater in Bangladesh

Md Firoz Ahmed[1], Md. Murshed Hasan Sarkar[3], Kaniz Mehzabin[1], Md. Ismail Hossain[4], Manotush Bhim[1], Sanjana Fatema Chowdhury[3], Showti Raheel Naser[3], Tabassum Mumtaz[2]✉ & Mohammad Omar Faruk[2]✉

Klebsiella pneumoniae **strains that combine multidrug resistance and enhanced virulence pose a growing global public health threat. Understanding the genetic basis of these high-risk lineages is critical for surveillance and mitigation. We isolated *K. pneumoniae* JU-BAEC-01 from treated effluent of antibiotic-manufacturing pharmaceutical facilities in Bangladesh and performed whole-genome sequencing with comparative genomic analyses to characterize its phylogeny, resistome, virulence-associated loci, mobile genetic elements, and predicted antiviral defense systems. JU-BAEC-01 belongs to a phylogenetically distinct lineage, serotype O3b: KL150 with resistance to nearly all clinically relevant antibiotic classes except carbapenems and colistin, mediated by an extensive acquired resistome, including tmexCD3-toprJ3 (tigecycline), armA, aac(6')-Ib-cr, qnrB4, oqxAB, blaDHA-1, blaSHV-182, and blaTEM-1B, mostly carried on conjugative IncC, IncFIB, IncHI1B, and IncR plasmids. Classical hypervirulence markers are present: complete aerobactin (iucABCD-iutA) and salmochelin (iroBCDEN) clusters, rmpA2, type 1 and type 3 fimbriae, T6SS, and pgaABCD. Notably, the strain encodes one of the most elaborate anti-phage defense arsenals reported in *Klebsiella* to date, comprising functional Type I-E, III-A, and IV-A CRISPR-Cas systems, multiple restriction-modification systems, BREX Type I, abortive infection systems (AbiE, AbiU), and additional novel defenses that coexist with phage-derived anti-CRISPR (AcrIE9) and anti-restriction (ArdA) proteins. *K. pneumoniae* JU-BAEC-01 is a "perfect storm" pathogen that combines multi-drug resistance (MDR), hypervirulence, and a multilayered, highly developed defense against bacteriophages. Together, these findings highlight the environmental emergence of a genetically distinct, multidrug-resistant *K. pneumoniae* with substantial virulence potential and complex phage–host interaction capacity, underscoring the need for genomic surveillance of pharmaceutical wastewater systems.**

Keywords *Klebsiella pneumoniae*, multi-drug resistance, Hypervirulence, Phage defense systems, CRISPR-Cas, BREX system

Antimicrobial resistance (AMR) has emerged as one of the greatest threats to global public health in the 21st century. Since the discovery of penicillin in 1928 and the subsequent "golden era" of antibiotic development, the selective pressure imposed by widespread clinical, agricultural, and veterinary use of antibiotics has driven

[1]Department of Microbiology, Jahangirnagar University, Savar, Dhaka 1342, Bangladesh. [2]Microbiology and Industrial Irradiation Division (MIID), Institute of Food and Radiation, Atomic Energy Research Establishment (AERE) Ganakbari, DEPZ, Atomic Energy Commission (BAEC), Savar, Dhaka 1342, Bangladesh. [3]Bangladesh Council of Scientific and Industrial Research (BCSIR), Dr. Qudrat-i-Khuda Road, 1205 Dhanmondi, Dhaka, Bangladesh. [4]Torrens University, 88 Wakefield St, Adelaide, SA 5000, Australia. ✉email: tabassum_baec@hotmail.com; faruk2006_du@yahoo.com

the rapid evolution and dissemination of resistant bacterial populations[1,2]. The World Health Organization now estimates that AMR could cause 10 million deaths annually by 2050 if current trends continue[3]. Of particular concern are Gram-negative pathogens exhibiting multidrug resistance or extensive drug resistance, many of which are resistant to last-resort carbapenems and colistin[4].

Among these critical-priority pathogens, *K. pneumoniae* exhibits extraordinary genomic plasticity and has been known to acquire both antibiotic resistance and hypervirulence determinants[5,6]. Traditionally, *K. pneumoniae* strains were divided into classical *K. pneumoniae* (cKP), mainly associated with nosocomial infections in immunocompromised hosts, and hypervirulent *K. pneumoniae* (hvKP), which had the capability of causing severe community-acquired infections in healthy individuals[7,8]. The hypervirulence is largely attributed to enhanced siderophore production, particularly aerobactin, hyper mucoviscosity conferred by rmpA/rmpA2, and other virulence plasmids[9]. The worrying convergence of carbapenem resistance with hypervirulence factors has generated multidrug-resistant-hvKP strains that combine near pan-drug resistance with extreme pathogenicity and high transmissibility[10,11]. Such "superbugs" have caused fatal outbreaks worldwide and represent a nightmare for modern medicine.

Horizontal gene transfer mediated by plasmids, transposons, integrative conjugative elements, and even bacteriophages facilitate the rapid spread of resistance and virulence genes in *K. pneumoniae*[12,13]. Pharmaceutical wastewater treatment plants receiving effluents from bulk-drug manufacturing facilities are extreme selective environments, with mg/L-level antibiotic concentrations, several orders of magnitude higher than in typical hospital or municipal wastewaters[14,15]. The hotspots thus created promote the emergence, persistence, and dissemination of antibiotic-resistant bacteria and antibiotic-resistant genes (ARG) into receiving water bodies and ultimately the broader environment[16,17].

Despite the increase in global awareness, pharmaceutical wastewater remains poorly regulated and under-characterized in many low- and middle-income countries. In Bangladesh, a major hub for generic drug manufacturing, high concentrations of active pharmaceutical ingredients have been detected in industrial effluents, raising serious concerns about environmental contamination and selection for high-risk clones[18]. The present study thus aimed to characterize the antibiotic residues, resistant bacterial populations, and associated resistance determinants in pharmaceutical wastewater from Bangladesh. Here, we provide a comprehensive analysis of this critical pollution source through physicochemical characterization, HPLC-based antibiotic residue analysis, culture-based isolation, antimicrobial susceptibility testing, and whole-genome sequencing. We further describe in detail a multidrug-resistant *Klebsiella pneumoniae* isolate, JU-BAEC-01, recovered from treated pharmaceutical effluent, which harbors an extensive resistome, virulence-associated determinants classically linked to hypervirulent lineages, and a diverse repertoire of predicted anti-phage defense systems. The genomic convergence of these features underscores the potential role of pharmaceutical wastewater as a reservoir for high-risk bacterial lineages and highlights the need for improved wastewater treatment and enhanced environmental surveillance.

Results

Antibiotic residues and antibiotic-resistant bacteria in pharmaceutical wastewater

Ten wastewater samples, including five influents and five effluents, were collected from antibiotic-manufacturing facilities in Dhaka and Gazipur (Fig. 1a). Influents of the wastewater were significantly polluted, characterized by mean (±SD) concentrations of COD 1248 ± 412 mg/L, BOD_5 468 ± 156 mg/L, and TDS 2156 ± 786 mg/L (Table 1). After conventional activated-sludge treatment, COD decreased to < 68 mg/L and BOD_5 to 118 ± 24 mg/L ($p < 0.01$). However, TDS declined only partly and remained above the limit in several treated effluents, as high as 2,978 mg/L. Through HPLC quantification, no ciprofloxacin or penicillin G residues above the limit of

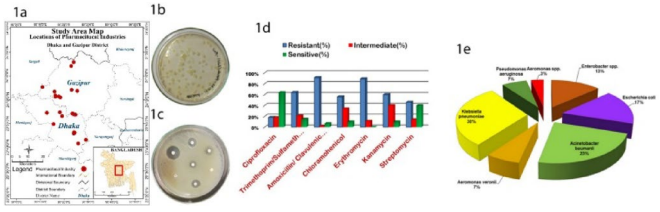

Fig. 1. Sampling sites, bacterial diversity, and antimicrobial resistance profiles of isolates recovered from pharmaceutical wastewater. **a** Geographic map generated with ArcMap 10.8 (ESRI) indicating the locations of the five pharmaceutical manufacturing facilities in Dhaka and Gazipur, Bangladesh, from which influent and effluent wastewater samples were collected. Base map data sourced from OpenStreetMap (www.openstreetmap.org). **b** Representative cultural morphology of isolated bacteria grown on non-selective Nutrient Agar (NA) after 24 h of incubation at 37 °C, demonstrating colonial diversity. **c** Phenotypic antibiotic susceptibility testing by the Kirby-Bauer disk diffusion method on Mueller-Hinton Agar (MHA). A clear zone of inhibition (ZOI) around the antibiotic disks indicates sensitivity, while confluent growth up to the disk edge indicates resistance. **d** Bar chart summarizing the antimicrobial susceptibility test results, showing the percentage of isolates classified as Sensitive, Intermediate, and Resistant (**c**) Pie chart showing the percentage distribution of identified bacterial species by an automated system such as VITEK showing *Klebsiella* spp. (30%), *Acinetobacter baumannii* (23%), *Escherichia coli* (17%), *Enterobacter* spp. (13%), *Pseudomonas aeruginosa* (7%), *Aeromonas* spp. (7%), and other genera (7%).

Parameter	Influent (mean ± SD)	Effluent (mean ± SD)	*p*-value
pH	7.1 ± 0.4	7.3 ± 0.2	n.s.
Temperature (°C)	28.4 ± 1.1	28.1 ± 0.9	n.s.
TDS (mg/L)	2,156 ± 786	1,247 ± 912	< 0.05
COD (mg/L)	1,248 ± 412	< 68	< 0.01
BOD_5 (mg/L)	468 ± 156	118 ± 24	< 0.01

Table 1. Physicochemical characteristics of influent and effluent wastewater samples collected from five pharmaceutical factories ($n = 5$ each).

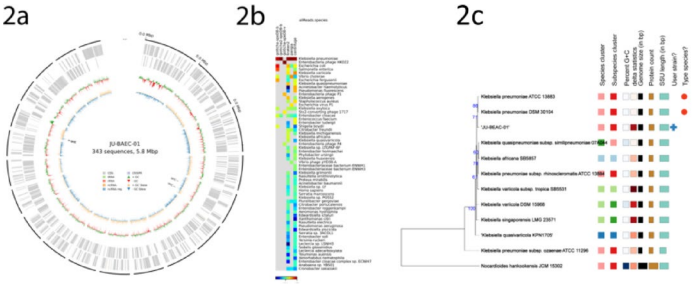

Fig. 2. Genome assembly, taxonomic composition, and phylogenetic placement of Klebsiella pneumoniae JU-BAEC-01. Initial genomic characterization confirms the species identity of the isolate. **a** Circular genome map of the draft assembly (5.77 Mbp, 343 contigs) annotated with Bakta, showing from outer to inner: COG functional classifications of CDS on forward and reverse strands, RNA genes, GC content (black), and GC skew (purple/green). **b** Heat map of taxonomic classification for JU-BAEC-01 sequencing dataset. Red color scale indicates relative abundance, and K. pneumoniae was identified as the predominant species. **c** TYGS-based phylogenetic tree indicating the evolutionary position of JU-BAEC-01, highlighted in red, within the genus Klebsiella. It clustered together with the two reference strains, namely, K. pneumoniae ATCC 13,883 and DSM 30,104.

Feature	Value
Contigs	343
GC Content	56.79%
Contig L50	51
Genome Length	5,769,218 bp
Contig N50	32,077

Table 2. Assembly Details.

quantification (0.05 µg/mL) could be detected, indicating their effective removal or degradation by the specific treatment. From the wastewater samples analyzed, 150 distinct bacterial isolates were obtained in nutrient agar (Fig. 1b). Susceptibility testing by the disk-diffusion method (Fig. 1c) showed high rates of resistance among the collection: 91.6% for amoxicillin/clavulanic acid, 89.2% for erythromycin, 64.7% for trimethoprim, 60.9% for kanamycin, 56.3% for chloramphenicol, 46.1% for streptomycin, and 19.7% for ciprofloxacin (Fig. 1d). Biochemical characterization of the obtained bacteria revealed their identity & identified bacterial species by an automated system such as VITEK: *Klebsiella* spp. (30%), *Acinetobacter baumannii* (23%), *Escherichia coli* (17%), *Enterobacter* spp. (13%), *Pseudomonas aeruginosa* (7%), *Aeromonas* spp. (7%), and other genera, which accounted for 7% (Fig. 1e).

This *K. pneumoniae* isolate was designated JU-BAEC-01 and was selected for whole-genome sequencing because it exhibited the broadest phenotypic resistance profile among the multidrug-resistant isolates recovered. Whole-genome sequencing was performed to characterize its resistome, virulence-associated loci, and potential mobile genetic elements.

Genome assembly and taxonomic identification

The genome of *K. pneumoniae* JU-BAEC-01 was assembled using the PATRIC Comprehensive Genome Analysis service and yielded a draft genome consisting of 5,769,218 bp in 343 contigs, with a G + C content of 56.79%. The assembly was visualized using Bakta (Fig. 2a). Assembly metrics are excellent, with a contig N50 of 32,077 and L50 of 51, indicating a good-quality draft genome (Table 2).

Taxonomic classification of the sequencing reads was performed using PUBMLST and EDGE bioinformatics tools, including gottcha-speDB-b, gottcha2-speDB-b, gottcha-speDB-v, kraken2, pangia, and centrifuge. Each tool uses different algorithms and databases that give a comprehensive view of the taxonomic composition of the dataset. Figure 2b shows the output in a heatmap format where a darker red cell means the species is more abundant, indicating that the JU-BAEC-01 sample contains *Klebsiella pneumoniae* as its most abundant species. This result was further confirmed by whole-genome sequencing and multi-locus sequence typing (MLST), which identified *Klebsiella pneumoniae* with 100% confidence.

The evolutionary relationship of different species of *Klebsiella*, including *Klebsiella* JU-BAEC-01, was determined using the TYGS platform. The phylogenetic tree obtained indicates that *Klebsiella* JU-BAEC-01 belongs to the *Klebsiella pneumoniae* family, a group of bacteria commonly associated with *pneumonia* (Fig. 2c). Interestingly, *Klebsiella* JU-BAEC-01 shares a close genetic relationship with *Klebsiella pneumoniae* strains ATCC 13,883 and DSM 30,104.

Identification of *Klebsiella pneumoniae* lineage

In the bacterial assembly "JU-BAEC-01," Kaptive analysis identified three main loci: KL150, KL107-D1, and O3b. Thus, the KL150 locus had a high match confidence with 98.19% coverage and 99.51% identity (Fig. 3a). On the other hand, KL107-D1 showed low confidence with 72.73% coverage and 76.74% identity therefore suggesting the presence of a partial capsular locus. Furthermore, the O3b locus presented a high match confidence with 99.43% coverage and 98.85% identity, indicating a well-preserved O-antigen biosynthesis region.

The Average Nucleotide Identity (ANI) analysis provided additional insights into the genetic relationships among the strains. *Klebsiella* JU-BAEC-01 exhibited the highest ANI similarity (99.60%) with *Klebsiella pneumoniae* HS11286, confirming its close evolutionary relationship within the *K. pneumoniae* species complex. ANI values above 95% typically indicate that genomes belong to the same species, supporting the classification of JU-BAEC-01 as a strain of *K. pneumoniae* (Fig. 3b).

To determine the genetic relatedness and phylogenetic placement of *Klebsiella pneumoniae* isolate JU-BAEC-01, a cgMLST analysis was performed using the BacWGSTdb 2.0 platform. Phylogenetic relationships were subsequently visualized using a Minimum Spanning Tree obtained by GrapeTree. The cgMLST-based phylogeny placed JU-BAEC-01 in a long, distinct branch of the MST, indicative of a highly diverged genetic background. Its closest genomic neighbor was isolate SU13_RCGF01 (Fig. 3c). However, it was genetically quite distant, with 576 allelic differences between the two cgMLST loci. This degree of genetic divergence far exceeds accepted thresholds for clonal or outbreak-related isolates (≤ 10–25 allelic differences), which confirms that JU-BAEC-01 and SU13_RCGF01 belong to different transmission networks or recent evolutionary lineages.

Genome annotation and functional insights

Whole-genome annotation of the draft genome of *Klebsiella pneumoniae* isolate JU-BAEC-01 provided an in-depth view of its structural and functional potential. The RASTtk pipeline identified 6,062 protein-coding sequences (CDSs), out of which 4,975 were functionally characterized, while 1,087 CDSs were designated as hypothetical proteins. Similarly, a considerable functional diversity as depicted by 5,771 genus-specific and 5,862 cross-genus protein families was identified (Table 3).

Functional categorization by RASTtk subsystem technology assigned 2,876 protein-coding sequences to 28 distinct functional categories (Fig. 4a), revealing a metabolic profile typical of a versatile, facultative anaerobic bacterium. The genome was significantly enriched in core metabolic processes, with the most heavily

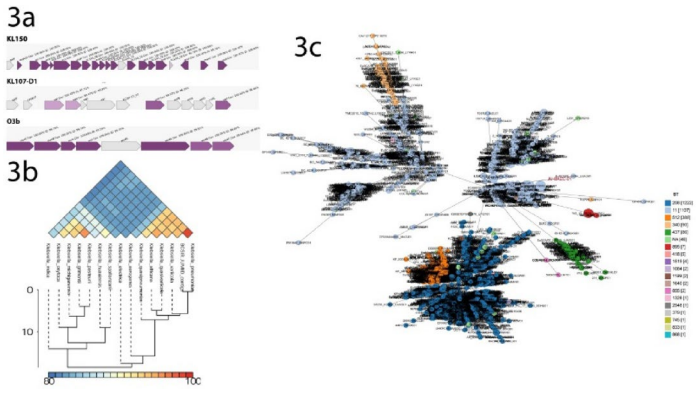

Fig. 3. Genomic characterization and phylogenetic analysis of a distinct Klebsiella pneumoniae lineage. Strain JU-BAEC-01 represents a unique genetic lineage within the species complex. **a** Kaptive analysis characterizing the capsular (K)-and O-antigen loci, identifying a high-confidence KL150 K-locus, a partial KL107-D1 locus, and a conserved O3b O-antigen locus. **b** Heatmap depicting Average Nucleotide Identity values, which indicated that JU-BAEC-01 shares 99.60% identity with K. pneumoniae HS11286, thus confirming its species-level classification, whereas it shared lower identity with K. variicola and K. africana. **c** Minimum spanning tree based on core-genome MLST, placing JU-BAEC-01 on a long, distinct branch with 576 allelic differences from its closest relative (SU13_RCGF01).

Feature	Count
CDS	6062
tRNA	49
rRNA	4
Hypothetical proteins	1087
Proteins with functional assignments	4975
Proteins with EC number assignments	1515
Proteins with GO assignments	1251
Proteins with Pathway assignments	1097
Proteins with PATRIC genus-specific family (PLfam) assignments	5771
Proteins with PATRIC cross-genus family (PGfam) assignments	5862

Table 3. Annotated Genome Features.

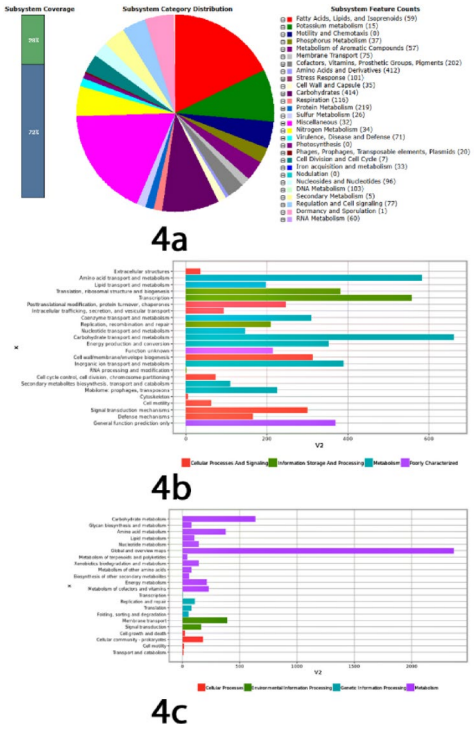

Fig. 4. Functional annotation and metabolic profiling of the Klebsiella pneumoniae JU-BAEC-01 genome. The genome is enriched with metabolic functions and pathways related to environmental adaptation. **a** Bar plot representing the distribution of protein-coding sequences across 28 functional subsystems as annotated by RASTtk. **b** Chart of the Clusters of Orthologous Groups (COG) functional classification, showing strong representation in Metabolism, Cellular Processes and Signaling, and Information Storage and Processing. **c** Histogram of KEGG pathway mapping, which shows the counts of genes in major pathway categories and enrichment in metabolic pathways.

represented subsystems including Carbohydrates (414 genes) and Amino Acids and Derivatives (412 genes), underpinning a high capacity for nutrient acquisition and biosynthesis. This extensive metabolic network was further underpinned by significant genetic investments in Protein Metabolism (219 genes), Cofactors, Vitamins, Prosthetic Groups, Pigments (202 genes), and Respiration (116 genes).

Prominent genes were associated with pathogenicity and environmental interaction. For example, the subsystems Virulence, Disease and Defense accounted for 71 genes, while 33 genes were dedicated to Iron acquisition and metabolism—a critical virulence factor. Further, 35 genes were associated with Cell Wall and Capsule biosynthesis.

COG classification likewise supported these observations, with a strong overrepresentation of metabolism and cellular processes (Fig. 4b). Main functional categories represented were Carbohydrate transport and metabolism (661 genes), Amino acid transport and metabolism (583 genes), Energy production and conversion (352 genes) and Transcription (558 genes). The genome also contained a remarkably high number of genes

related to Defense mechanisms (165) and the mobilome (226), indicating active evolutionary forces. General or unknown functions were assigned to 584 genes, pointing to a pool of untapped genetic potential.

KEGG pathway mapping further outlined the metabolic potential of JU-BAEC-01, as shown in Fig. 4c. The major category represented was metabolism, dominated by Carbohydrate metabolism (637 genes), Amino acid metabolism (382 genes), and Energy metabolism (215 genes). It had pathways for Xenobiotic degradation (145 genes) and Membrane transport (391 genes) for environmental resilience and nutrient uptake. Clinical perspectives were also posed through KEGG, highlighting genes associated with antimicrobial drug resistance (76) and Infectious diseases (37). Collectively, this integrated genomic analysis using RASTtk, COG, and KEGG shows that *K. pneumoniae* JU-BAEC-01 possesses a metabolically versatile and genetically resilient genome.

Pathogenicity, antibiotic resistance and virulence Island

Further analysis with the IslandPath-DIMOB tool revealed a diverse array of proteins, enzymes, transporters, and regulatory elements in *Klebsiella* JU-BAEC-01 isolates; most of these were directly or indirectly related to resistance mechanisms, mobile genetic elements, and other vital bacterial processes. Using bioinformatic tools such as SIGI-HMM, which uses Hidden Markov Models to detect genomic islands, phage, and resistance genes, this work has been able to highlight several important genomic features, as shown in Fig. 5a.

The isolate possesses a robust repertoire of virulence-associated genes that are classically linked to hypervirulent *K. pneumoniae*, indicating a strong genomic potential for enhanced virulence. Among the most important are its elaborated siderophore systems; it carries the aerobactin receptor gene iutA, one of the main hallmarks of hvKp, and the complete salmochelin cluster iroBCDEN, which modifies enterobactin to evade host immunity. The combination provides exceptional iron acquisition that is key to systemic infection. This isolate also carries a complete and complex Type VI Secretion System (T6SS), which is a molecular weapon for invading host cells and competing with other bacteria. The presence of entire fim (Type 1 fimbriae) and mrk (Type 3 fimbriae) operons, alongside the pgaABCD locus for poly-N-acetylglucosamine synthesis, promotes adhesion and biofilm development. Adhesion properties are further provided by the genes for the *E. coli* common pilus, ECP (ecpA-R). Immune evasion is facilitated by the presence of genes for extensive capsular (CPS) and lipopolysaccharide (LPS) synthesis, which is controlled by rcsAB, and the arn operon for resistance to cationic antimicrobial peptides (CAMPs) (Fig. 5b).

The resistome of this isolate includes an extensive range of antimicrobial resistance mechanisms that make it resistant to nearly all major classes of antibiotics. Resistance to tigecycline is mediated through the powerful RND-type efflux system TmexCD3-toprJ3. High-level, broad-spectrum aminoglycoside resistance is conferred by the 16 S rRNA methyltransferase armA. In addition, fosfomycin resistance proceeds through a two-fold mechanism that includes both the glutathione transferase gene fosA6 and a mutation in the uhpT fosfomycin transporter.

The combination of an AmpC-type cephalosporinase, DHA-1, with an extended-spectrum beta-lactamase, SHV-11, and a penicillinase, TEM-1, ensures a broad-spectrum beta-lactam resistance profile. Fluoroquinolone resistance is multi-pronged, with contributions from plasmid-mediated target protection, qnrB4; target site mutations in DNA gyrase, gyrA S83I, and topoisomerase IV, parC S80I; and efflux pumps, OqxAB. The presence of aac(6')-Ib-cr further contributes by acetylating and inactivating ciprofloxacin.

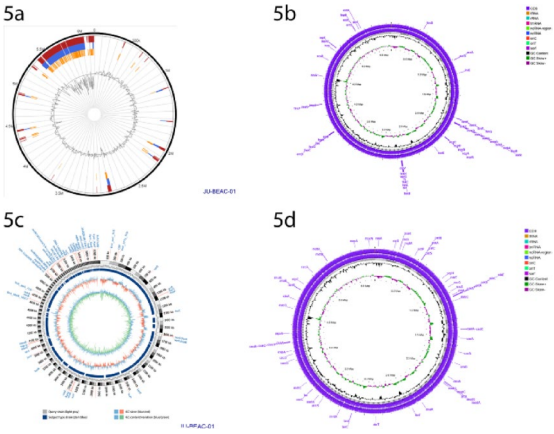

Fig. 5. Virulence potential, antibiotic resistance, and genomic islands of Klebsiella pneumoniae JU-BAEC-01. The genome harbors a diverse arsenal of virulence factors, resistance determinants, and horizontally acquired elements. **a** Circular genome map showing genomic islands as predicted by IslandPath-DIMOB and SIGI-HMM methods indicating putative horizontal gene transfer events. **b** Circular genome map showing the distribution of major virulence gene categories, including adhesion factors (adhesins), T2SS components, and genes for LPS biosynthesis. Color intensity represents the number of the respective genes. **c** Distribution of antibiotic resistance genes detected across major drug classes by CARD and ResFinder, considering efflux pumps, enzymatic inactivation, and target modification mechanisms. **d** Circular genome map showing the locations of heavy metal resistance operons for arsenic (ars), copper (cus, pco), mercury (mer), and cobalt-zinc-cadmium (czc).

Other core resistances consist of genes for macrolides [mph(A), mph(E), msr(E)], sulfonamides (sul1), trimethoprim (dfrA12), and chloramphenicol (through OqxAB efflux). Importantly, resistance mechanisms against peptide antibiotics such as polymyxins were identified: the arnT and eptB genes, which modify lipid A, and impaired permeability through OmpA and OmpK37. The combined presence of various efflux systems, including AcrAB-TolC, OqxAB, TmexCD3-toprJ3, and various regulatory mutations, such as those in marR, ensure a high degree of baseline multidrug tolerance (Fig. 5c).

Several heavy metal resistance genes were detected, indicating the ability of the isolate to survive under contaminated environments. These included: ArsA, ArsB, and ArsH, which take part in arsenic detoxification, while CzcD confers resistance to cobalt, zinc, and cadmium. Copper resistance is mediated by CusA, CusB, CusR, and PcoE, which facilitate detoxification and subsequent efflux of copper ions out of the cells. Additionally, MerD, MerE, and MerT take part in mercury detoxification, thus showing the potential of an isolate for survival in environments under toxic metal stress. In addition, copper and silver efflux systems such as CusC, CusA, and CusR further enforce the ability of the isolate to handle metal toxicity (Fig. 5d).

Bacteriophage diversity, mobile genetic elements and horizontal gene transfer

MobileOG-db classification of the *K. pneumoniae* isolate JU-BAEC-01 identified 637 protein families that could be assigned to five major functional categories of MGEs: IE, with 138 protein families; RRR, with 174 protein families; P, with 124 protein families; STD, with 81 protein families; and T, with 120 protein families. The representation of such a wide range of protein families indicates that the isolate probably carries an array of MGEs, including conjugative plasmids, transposons, and bacteriophages, which are probably contributing to MDR and virulence. Further study on specific protein families obtained from mobileOG-db will be required for understanding the mechanisms involved in these processes (Fig. 6a).

A detailed bacteriophage analysis using PHASTER revealed a diverse array of bacteriophages with varied structural and functional properties. PHASTER analysis identified six prophage regions, including a questionable 19.9 Kb region linked to PHAGE_Shigel_SfIV, and intact regions of 36 Kb, 27.1 Kb, and 24.1 Kb associated with PHAGE_Edward_GF_2, PHAGE_Klebsi_ST512_KPC3phi13.2, and PHAGE_Pseudo_phiPSA1, respectively. Additionally, two incomplete prophage regions (18.6 Kb and 11.2 Kb) were linked to PHAGE_Escher_phiV10 and PHAGE_Escher_HK639. These findings suggest that the isolate may possess a significant number of prophages that could play a role in the horizontal transfer of resistance and virulence genes, thus enhancing the pathogenic potential of the strain (Fig. 6b).

These analyses were followed by plasmid replicon detection within the isolate using the geNomad workflow revealed several plasmids carrying important resistance genes, including *qnrB4, blaDHA-1, blaTEM-1B,

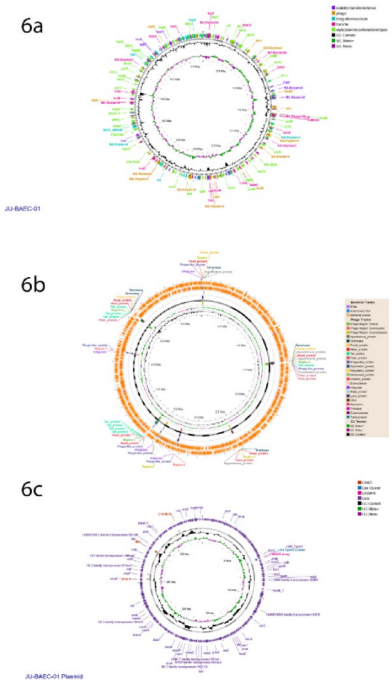

Fig. 6. Mobile genetic elements and bacteriophage diversity in Klebsiella pneumoniae JU-BAEC-01. The genome is equipped with a rich repertoire of mobile elements, including prophages and plasmids. **a**Circular genome map showing 637 mobile genetic element protein families classified by MobileOG-db, including functions for Integration/Excision, Replication/Repair, Phage, Stability/Transfer/Defense, and Conjugation. **b** Circular representation of the genome showing the location, size and completeness of six identified prophage regions. **c** Schematic of detected plasmid replicons, highlighting their associated antibiotic resistance genes and virulence factors.

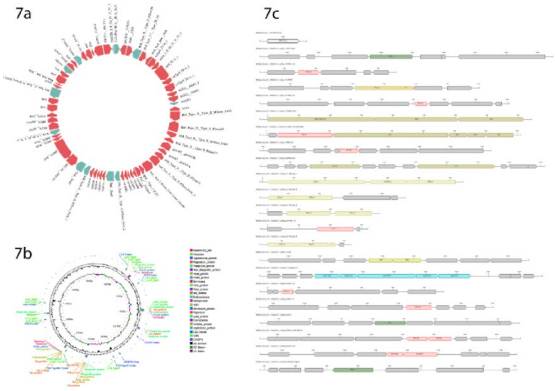

Fig. 7. Comprehensive analysis of antiviral defense systems in Klebsiella pneumoniae JU-BAEC-01. It possesses a multilayered arsenal of antiviral defense mechanisms. **a** Circular genome map from DefenseFinder, showing the locations of identified defense systems. These include the CRISPR-Cas, restriction-modification, abortive infection, and BREX systems. **b** Genome map representation of defense systems identified by PADLOC. These include a CRISPR-Cas Type IV-A system, several restriction-modification systems, and BREX modules. Colour intensity is indicative of prediction confidence. **c** Detailed visualization of CRISPR arrays and associated Cas proteins identified by CRISPRCas Finder, showing the multi-component architecture of the CRISPR-Cas systems.

blaSHV-182, OqxA, OqxB, fosA, mph(A), qacE, aac(6')-Ib-cr, mph(E), msr(E), dfrA12,* and aadA2. This is evidence of resistance to various classes of antibiotics, including quinolones, beta-lactams, and aminoglycosides. Virulence factors such as iutA, mrkA, terC, and fimH were also found in this isolate, showing enhanced pathogenicity potential (Fig. 6c).

Defense mechanisms against phage infections

The bacterial defense systems of *Klebsiella* JU-BAEC-01 were comprehensively characterized using three complementary bioinformatics tools: PADLOC, DefenseFinder, and CRISPRCasFinder 4.2.20. These analyses revealed a diverse and multi-layered defense strategy involving CRISPR-Cas systems, restriction-modification (RM) systems, abortive infection (Abi) mechanisms, and less characterized defense elements, along with phage-encoded antidefense systems, which highlight the ongoing evolutionary dynamics between bacteria and bacteriophages.

CRISPR-Cas systems were detected consistently by both DefenseFinder (Fig. 7a) and PADLOC (Fig. 7b), which identified a Class 1, Subtype IV-A CRISPR-Cas system known for targeting phage DNA. CRISPR arrays were also confirmed by CRISPRCasFinder (Fig. 7c), with a particularly strong prediction on Contig 41 (evidence level 3), suggesting a likely functional CRISPR-Cas defense. Additionally, Cas proteins necessary for viral DNA targeting were found across multiple contigs, including Type III-A on Contig 71 and Type U on Contig 41, indicating that *Klebsiella* JU-BAEC-01 harbors several functional CRISPR-Cas systems.

Both PADLOC and DefenseFinder identified multiple RM systems, including Type I and Type II, which are crucial for the protection of bacterial genomes through the cleavage of foreign DNA and methylation of specific sequences. These are among the essential components of the arsenal of *Klebsiella* JU-BAEC-01 against bacteriophages, enhancing its capacity for resistance against phage infection.

In addition to CRISPR-Cas and RM systems, the BREX (I) system was detected by both PADLOC and DefenseFinder, suggesting that *Klebsiella* JU-BAEC-01 employs this non-cleaving phage defense mechanism to inhibit viral replication. The detection of abortive infection (Abi) systems, including AbiE and AbiU, further adds to the multilayered nature of the bacterial defense, providing an additional fail-safe mechanism by inducing programmed cell death in infected cells to prevent the spread of phages.

Furthermore, both PADLOC and DefenseFinder identified several other less-studied or novel defense systems. PADLOC detected PD-T7-1, PDC-S12, and VSPR systems, along with BrxL and BrxC proteins associated with the BREX system, whereas DefenseFinder highlighted the GAPS1 system and the Mok-Hok-Sok toxin-antitoxin system, which may contribute to bacterial stress responses and phage resistance. Of particular interest was the identification by DefenseFinder (Fig. 6a) of phage-encoded antidefense systems, including Anti-RM systems (e.g., Ard proteins) and Anti-CRISPR systems (e.g., acrIE9), which allow phages to evade bacterial immune defenses. This highlights the ongoing evolutionary arms race between *Klebsiella* JU-BAEC-01 and bacteriophages, as the phages continue to evolve mechanisms to overcome bacterial defense strategies.

Methods

Sample collection and physicochemical characterization

Wastewater samples were collected from five pharmaceutical industries located in Dhaka and Gazipur, Bangladesh, between January and March 2023. Each industry was a bulk producer of β-lactam and fluoroquinolone antibiotics. Five influent (raw wastewater) and five effluent (post-treatment) samples were collected aseptically in sterile 1 L polypropylene bottles. Samples were transported at 4 °C and processed within 6 h of collection. Temperature

and pH were measured on-site using a portable meter (Hanna Instruments, USA). Total dissolved solids (TDS), chemical oxygen demand (COD), and biochemical oxygen demand (BOD_5 at 20 °C) were determined according to the standard methods[19].

Determination of antibiotic residues by high-performance liquid chromatography (HPLC)

Antibiotic residues were quantified using a Hitachi LaChrome Elite HPLC system, UV detector, and reversed-phase Nucleosil 120-5 C18 column (250 × 4.6 mm, 5 µm). The mobile phase consisted of a mixture of acetonitrile and 0.024 M orthophosphoric acid buffer (pH 3.0 ± 0.1 adjusted with triethylamine), at a ratio of 13:87 (v/v). The flow rate was 1.5 mL/min, detection wavelength 278 nm, column temperature 40 °C and injection volume 10 µL. Ciprofloxacin and penicillin G were used as external standards. Quantification was based on peak area using six-point calibration curves (0.1–100 µg/mL; $R^2 > 0.999$). LOQ was 0.05 µg/mL for both antibiotics[11,18].

Bacterial isolation, enumeration, and preservation

Serial ten-fold dilutions of wastewater were plated onto nutrient agar (Oxoid, UK), cetrimide agar (Condalab, Spain) and mFC agar (HiMedia, India) for enumeration of total heterotrophic bacteria, *Pseudomonas spp*. and fecal coliforms, respectively. Plates were incubated at 37 °C for 24–48 h (44.5 °C for mFC agar). Morphologically distinct colonies were purified and preserved at − 80 °C in tryptic soy broth containing 20% glycerol.

Antimicrobial susceptibility testing

Antibiotic susceptibility was determined by the Kirby–Bauer disk diffusion method on Mueller–Hinton agar (Oxoid, UK) in accordance with the EUCAST 2023 guidelines[20]. A total of 26 antibiotics belonging to 12 classes were tested. Isolates resistant to at least one agent in ≥ 3 antibiotic classes were classified as multidrug-resistant (MDR)[21]. MICs of selected MDR isolates were determined using the VITEK® 2 Compact automated system (bioMérieux, France) with AST-GN83 cards according to the manufacturer's recommendations.

Genomic DNA extraction and whole-genome sequencing of *K. pneumoniae* JU-BAEC-01

High-molecular-weight genomic DNA from an overnight culture of *K. pneumoniae* JU-BAEC-01 was extracted with the DNeasy Blood & Tissue Kit (Qiagen, Germany), as instructed by the manufacturer's protocol for Gram-negative bacteria. Library preparation was carried out using the Nextera XT DNA Library Preparation Kit (Illumina, USA). Paired-end sequencing (2 × 150 bp) was carried out on an Illumina MiniSeq platform, generating > 100× average coverage.

Genome assembly

High-quality reads were assembled into a draft genome using the integrated assembly pipelines of the BV-BRC version 3.56.67[22] and the EDGE bioinformatics platform[23]. These services utilize robust algorithms like SPAdes[24] to perform contig generation, scaffolding, and quality assessment.

Taxonomic and phylogenetic analysis

The isolate was confirmed as *K. pneumoniae*, an MLST scheme that defines the strain type based on seven housekeeping genes through the PubMLST[25] database. Capsular polysaccharide (K-locus) and O-antigen (O-locus) were typed with high confidence to KL150/KL107-D1 and O3b, respectively, by Kaptive v1.3.0[26]. Phylogenetic placement was evaluated by the Type (Strain) Genome Server (TYGS)[27]. Additionally, a reference genome SNP approach and the cgMLST scheme through the BacWGSTdb 2.0 database[28] were used in an effort to build phylogenetic trees, generating neighbor-joining and minimum spanning trees that help explain genetic relationships among the isolates.

Pangenome analysis

A pangenome analysis was performed with the aim of determining genetic variability. Fourteen reference genome sequences of the Klebsiella genus have been downloaded from the NCBI database. The analysis via the Integrated Prokaryotes Genome and Pan-genome Analysis tool IPGA v1.09[29] confirmed that there is a high degree of genetic relatedness between JU-BAEC-01 and other *K. pneumoniae* strains.

Genome annotation and functional analysis

Extensive genome annotation was performed to identify and characterize genetic elements. Primary annotation was carried out using the Genome Annotation Service on BV-BRC and Prokka[30] for rapid gene prediction. Functional insights were achieved by annotating genes with GO terms and KEGG pathways using the gcType database[31]. For subsystem-based annotation, RASTtk (RAST tool kit)[32] was used. The genome was visualized using Proksee and Bakta[33,34], which provided a way to map the genome in a circular fashion and to highlight features like CDSs, RNA genes, and AMR loci.

Analysis of antibiotic resistance and virulence factors

ARGs were identified using CARD[35] and ResFinder[36], and results were visualized using ProbioMinServer[37]. Resistance genes for metals and biocides were detected using the BacMet database[38]. Virulence factors and pathogenic islands were visualized using the dedicated analysis pipelines available within IslandViewer 4[54] and EDGE Bioinformatics.

Identification of mobile genetic elements and bacteriophages

MGEs were characterized with the use of Mobile Element Finder[39,41] and Plasmid Finder[40,41] for the identification of insertion sequences, transposons, and plasmid replicons. Putative plasmid sequences within assembled

scaffolds were further analyzed using geNomad[42]. Bacteriophage-derived sequences and prophage regions were identified and classified using PHASTEST[43].

Prediction of antiviral defense systems

Bacterial defense systems against viral infections, such as CRISPRCas and abortive infection systems, were identified using a suite of bioinformatics tools: DefenseFinder[44], PADLOC v2.0.0[45], and CRISPRCasFinder 4.2.20[46].

Discussion

The escalating crisis of AMR has been widely recognized as one of the most serious global public health threats of the 21st century, with current estimates of 700,000 annual deaths projected to rise to 10 million by 2050[3]. Pharmaceutical wastewater treatment plants (PWWTPs) that receive effluents from bulk-drug manufacturing facilities constitute extreme selective environments in which antibiotic concentrations can reach the mg/L range—orders of magnitude higher than in municipal or hospital wastewaters[14,47]. Such hotspots have repeatedly been shown to act as crucibles for the emergence and dissemination of multidrug-resistant (MDR) and extensively drug-resistant pathogens, as well as reservoirs of mobilizable antibiotic resistance genes (ARGs)[48,49].

Here, we report the first whole-genomic characterization of a high-risk *K. pneumoniae* isolate, JU-BAEC-01, recovered from treated pharmaceutical effluent in Bangladesh-a major hub of generic antibiotic production. Although conventional activated-sludge treatment had substantially reduced the organic load (COD and BOD_5), and ciprofloxacin and penicillin G were not detectable by HPLC in the effluents, MDR bacteria were still viable in numbers ranging from 10^3 to 10^6 CFU/mL. This suggests that current treatment processes are inadequate for removal of highly adapted resistant populations and is in agreement with findings from related PWWTPs in India and China[47,15].

Phenotypic screening of 150 isolates showed alarmingly high resistance rates to many antibiotics including 20% that were classified as MDR. Of 30 MDR isolates analyzed by VITEK® 2, the predominant species was *K. pneumoniae* at 30%, followed by *Acinetobacter baumannii* and *Escherichia coli*. Notably, all isolates remained susceptible to carbapenems and colistin, while exhibiting resistance to several non-carbapenem, non-polymyxin antibiotics, consistent with—but not demonstrative of—environmental selective pressures. Draft whole-genome sequencing of the most resistant isolate, *Klebsiella pneumoniae* JU-BAEC-01, revealed a notable convergence of antimicrobial resistance, virulence-associated loci, and predicted phage defense systems in an environmental isolate. The genome harbors an extensive acquired resistome conferring resistance to most major antibiotic classes, with the exception of carbapenems and colistin, alongside multiple virulence-associated determinants classically linked to hypervirulent *K. pneumoniae* lineages and a diverse, multi-layered repertoire of predicted anti-phage defense systems. It is important to note that the presence of hypervirulence-associated genetic determinants does not equate to a confirmed hypervirulent phenotype, as no in vitro or in vivo virulence assays were performed in this study.

In particular, the resistome of JU-BAEC-01 is alarming. It bears the newly emerged tmexCD3-toprJ3 RND efflux cluster mediating high-level tigecycline resistance that has recently spread with remarkable success globally in mobile plasmids[50]. Other relevant elements included armA mediating high-level aminoglycoside resistance, the bifunctional aminoglycoside/fluoroquinolone acetyltransferase aac(6')-Ib-cr, plasmid-mediated qnrB4 and oqxAB[43], and chromosomal quinolone resistance-conferring mutations -QRDR gyrA S83I, parC S80I-. Multiple β-lactamases -blaDHA-1, blaSHV-182, blaTEM-1B- completed a profile making the strain effectively multi-drug-resistant to first- and second-line therapies. Most of such determinants were plasmid-borne-IncC, IncFIB, IncHI1B, IncR-pointing out the role of horizontal gene transfer in assembling such a complex resistome under intense selective pressure.

Meanwhile, JU-BAEC-01 harbors multiple virulence-associated loci classically linked to hypervirulent *Klebsiella pneumoniae*, including complete aerobactin and salmochelin (iroBCDEN) siderophore systems, rmpA2, intact Type 1 and Type 3 fimbrial operons, a Type VI secretion system, and the pgaABCD biofilm operon. These determinants have been previously associated with severe community-acquired invasive infections, such as pyogenic liver abscess and metastatic dissemination, although their functional contribution was not experimentally assessed in this study[8,26].Convergence of carbapenem-susceptible yet extensively MDR profiles with hvKp determinants in an environmental isolate is decidedly alarmist, since their acquisition of carbapenemases-for example, via a single plasmid transfer event-could instantly yield untreatable CR-hvKp "superbugs" of the kind responsible for fatal nosocomial outbreaks[15,11].

A third, hitherto underappreciated dimension of JU-BAEC-01 success is the extraordinarily sophisticated arsenal of anti-phage defense systems that it possesses. Utilizing DefenseFinder, PADLOC, and CRISPRCasFinder, we have identified functional Type I-E, IIIA, and IVA CRISPR-Cas systems; multiple Type I and II restriction-modification systems; BREX Type I; abortive infection systems AbiE, AbiU; toxin-antitoxin modules; and several poorly characterized systems: PD-T7-1, VSPR, GAPS1. Cooccurrence of phage-encoded anti-defense proteins AcrIE9 and ArdA within prophage regions highlights an active evolutionary arms race. The presence of this multi-layered defense can be a plausible explanation for the ecological dominance of the strain within phage-rich environments such as wastewater and serves as a major obstacle to phage therapy-an otherwise very promising alternative against MDR infections[51].

Such extensive innate immunity should, therefore, coexist with a finely tuned evolutionary trade-off-a selective relaxation of adaptive immunity, for example, limited CRISPR activity against resistance/virulence plasmids, with retention of robust innate systems that block lytic phages but allow access to beneficial gene acquisition[52,53].

In summary, *Klebsiella pneumoniae* JU-BAEC-01 exemplifies a concerning genomic convergence of anthropogenic selection pressures, representing a genetically distinct environmental isolate that has accumulated

an extensive resistome, multiple virulence-associated loci, and a diverse repertoire of predicted phage defense systems. Although these traits are inferred from genomic data and not functionally validated, their co-occurrence in a strain recovered from treated pharmaceutical effluent highlights the potential role of industrial wastewater environments as reservoirs for the emergence and persistence of high-risk bacterial lineages. These findings underscore the need for (i) improved wastewater treatment technologies capable of reducing viable multidrug-resistant bacteria and resistance genes, (ii) stricter regulation of antibiotic emissions from pharmaceutical manufacturing, and (iii) continued development of alternative therapeutic strategies, including anti-virulence and phage-based approaches, informed by a realistic understanding of bacterial defense potential.

The present study demonstrates that even efficiently treated pharmaceutical wastewater in Bangladesh can act as a reservoir of multidrug-resistant bacteria. *Klebsiella pneumoniae* JU-BAEC-01 represents a genetically distinct environmental isolate that exhibits a notable genomic convergence of an extensive acquired resistome, multiple virulence-associated loci, and a diverse repertoire of predicted anti-phage defense systems. Although these traits are inferred from draft genome analysis and not functionally validated, their co-occurrence in an isolate recovered from treated effluent underscores the potential role of pharmaceutical wastewater environments in the emergence and persistence of high-risk bacterial lineages. These findings highlight the need for advanced wastewater treatment strategies, stricter regulation of industrial antibiotic discharges, and continuous genomic surveillance of pharmaceutical effluents to mitigate dissemination into clinical and community settings.

Data availability

Raw sequencing reads generated from this study will be available upon request.

Received: 4 December 2025; Accepted: 17 March 2026
Published online: 23 March 2026

References

1. Davies, J. & Davies, D. Origins and evolution of antibiotic resistance. *Microbiol. Mol. Biol. Rev.* **74** (3), 417–433. https://doi.org/10.1128/MMBR.00016-10 (2010).
2. MacLean, R. C. & San Millan, A. The evolution of antibiotic resistance. *Science* **365** (6458), 1082–1083. https://doi.org/10.1126/science.aax3879 (2019).
3. O'Neill, J. Tackling drug-resistant infections globally: final report and recommendations. *Rev. Antimicrob. Resist.* (2016). https://amr-review.org/sites/default/files/160518_Final%20paper_with%20cover.pdf
4. Tacconelli, E. et al. Discovery, research, and development of new antibiotics: the WHO priority list of antibiotic-resistant bacteria and tuberculosis. *Lancet Infect. Dis.* **18** (3), 318–327. https://doi.org/10.1016/S1473-3099(17)30753-3 (2018).
5. Paczosa, M. K. & Mecsas, J. Klebsiella pneumoniae: going on the offense with a strong defense. *Microbiol. Mol. Biol. Rev.* **80** (3), 629–661. https://doi.org/10.1128/MMBR.00078-15 (2016).
6. Wyres, K. L., Lam, M. M. C. & Holt, K. E. Population genomics of Klebsiella pneumoniae. *Nat. Rev. Microbiol.* **18** (6), 344–359. https://doi.org/10.1038/s41579-019-0316-1 (2020).
7. Shon, A. S., Bajwa, R. P. & Russo, T. A. Hypervirulent (hypermucoviscous) Klebsiella pneumoniae. *Virulence* **4** (2), 107–118. https://doi.org/10.4161/viru.22718 (2013).
8. Russo, T. A. & Marr, C. M. Hypervirulent Klebsiella pneumoniae. *Clin. Microbiol. Rev.* **32** (3), e00001–19. https://doi.org/10.1128/CMR.00001-19 (2019).
9. Russo, T. A. et al. Identification of biomarkers for differentiation of hypervirulent Klebsiella pneumoniae from classical K. pneumoniae. *mBio* **9** (3), e00888–e00818. https://doi.org/10.1128/mBio.00888-18 (2018).
10. Gu, D. et al. A fatal outbreak of ST11 carbapenem-resistant hypervirulent Klebsiella pneumoniae in a Chinese hospital: a molecular epidemiological study. *Lancet Infect. Dis.* **18** (1), 37–46. https://doi.org/10.1016/S1473-3099(17)30489-9 (2018).
11. Chen, L. & Kreiswirth, B. N. Emerging hypervirulent-carbapenem-resistant Klebsiella pneumoniae. *Lancet Infect. Dis.* **21** (5), 601–602. https://doi.org/10.1016/S1473-3099(21)00173-8 (2021).
12. Partridge, S. R., Kwong, S. M., Firth, N. & Jensen, S. O. Mobile genetic elements associated with antimicrobial resistance. *Clin. Microbiol. Rev.* **31** (4), e00088–e00017. https://doi.org/10.1128/CMR.00088-17 (2018).
13. Penadés, J. R., Chen, J., Quiles-Puchalt, N., Carpena, N. & Novick, R. P. Bacteriophage-mediated spread of bacterial virulence genes. *Curr. Opin. Microbiol.* **23**, 171–178. https://doi.org/10.1016/j.mib.2014.11.019 (2015).
14. Larsson, D. G. J., de Pedro, C. & Paxeus, N. Effluent from drug manufactures contains extremely high levels of pharmaceuticals. *J. Hazard. Mater.* **148** (3), 751–755. https://doi.org/10.1016/j.jhazmat.2007.07.008 (2007).
15. Guo, X. et al. Behavior of antibiotic resistance genes under extremely high-level antibiotic selection pressures. *Sci. Total Environ.* **612**, 119–128. https://doi.org/10.1016/j.scitotenv.2017.08.229 (2018).
16. Fick, J. et al. Contamination of surface, ground, and drinking water from pharmaceutical production. *Environ. Toxicol. Chem.* **28** (12), 2522–2527. https://doi.org/10.1897/09-073.1 (2009).
17. Obayiuwana, A. & Ogunjobi, A. A. Prevalence of antibiotic resistance genes in pharmaceutical wastewaters. *Water* **13** (13), 1731. https://doi.org/10.3390/w13131731 (2021).
18. Larsson, D. G. J. & Fick, J. Transparency throughout the drug production chain. *Science* **333** (6041), 405–406. https://doi.org/10.1126/science.1209031 (2011).
19. Rice, E. W., Baird, R. B., Eaton, A. D. & Clesceri, L. S. (eds) *Standard methods for the examination of water and wastewater* 23rd edn (American Public Health Association, 2017).
20. European Committee on Antimicrobial Susceptibility Testing (EUCAST). Breakpoint tables for interpretation of MICs and zone diameters, version 13.0. (2023). https://www.eucast.org/clinical_breakpoints
21. Magiorakos, A. P. 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.* **18** (3), 268–281. https://doi.org/10.1111/j.1469-0691.2011.03570.x (2012).
22. Wattam, A. R. et al. BVBRC: The Bacterial and Viral Bioinformatics Resource Center. *Nucleic Acids Res.* **49** (D1), D798–D808. https://doi.org/10.1093/nar/gkaa930 (2021).
23. Li, H. et al. EdgeBio: Integrating cloud-computing infrastructure with genomic analysis. *Bioinformatics* **33** (1), 142–143. https://doi.org/10.1093/bioinformatics/btw617 (2017).
24. Bankevich, A. et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* **19** (5), 455–477. https://doi.org/10.1089/cmb.2012.0021 (2012).
25. Jolley, K. A. & Maiden, M. C. PubMLST: a public resource for the genetic analysis of bacterial pathogens. *Methods Mol. Biol.* **693**, 1–21. https://doi.org/10.1007/978-1-60761-747-5_1 (2010).

26. Lam, M. M. C., Wick, R. R., Judd, L. M., Holt, K. E. & Wyres, K. L. Kaptive 2.0: updated capsule and lipopolysaccharide locus typing for the Klebsiella pneumoniae species complex. *Microb. Genom.* **8** (3), 000800. https://doi.org/10.1099/mgen.0.000800 (2022).
27. Meier-Kolthoff, J. P. & Göker, M. T. Y. G. S. The Type Strain Genome Server for prokaryotic taxonomy. *BMC Genom.* **20** (1), 1–8. https://doi.org/10.1186/s12864-019-6095-1 (2019).
28. Ye, F. et al. BacWGSTdb 2.0: a one-stop repository for bacterial whole-genome sequence typing and source tracking. *Nucleic Acids Res.* **49** (D1), D644–D650. https://doi.org/10.1093/nar/gkaa821 (2021).
29. Liu, D. et al. IPGA: a handy integrated prokaryotes genome and pan-genome analysis web service. *iMeta* ;**1**(4):e55. doi:https://doi.org/10.1002/imt2.55 (2022).
30. Seemann, T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30** (14), 2068–2069. https://doi.org/10.1093/bioinformatics/btu153 (2014).
31. Fan, G. et al. GCM and gcType in 2024: comprehensive resources for microbial strains and genomic data. *Nucleic Acids Res.* **53** (D1), D763–D771. https://doi.org/10.1093/nar/gkae1057 (2025).
32. Aziz, R. K. et al. The RAST Server: rapid annotations using subsystems technology. *BMC Genom.* **9**, 75. https://doi.org/10.1186/1471-2164-9-75 (2008).
33. Grant, J. R. et al. Proksee: in-depth characterization and visualization of bacterial genomes. *Nucleic Acids Res.* **51** (W1), W484–W492. https://doi.org/10.1093/nar/gkad326 (2023).
34. Beyvers, S., Jelonek, L., Goesmann, A. & Schwengers, O. Bakta Web – rapid and standardized genome annotation on scalable infrastructures. *Nucleic Acids Res.* **53** (W1), W51–W56. https://doi.org/10.1093/nar/gkaf335 (2025).
35. Jia, B. et al. CARD 2017: expansion and model-centric curation of the Comprehensive Antibiotic Resistance Database. *Nucleic Acids Res.* **45** (D1), D566–D573. https://doi.org/10.1093/nar/gkw1004 (2017).
36. Bortolaia, V. et al. ResFinder 4.0 for predictions of phenotypes from genotypes. *J. Antimicrob. Chemother.* **75** (12), 3491–3500. https://doi.org/10.1093/jac/dkaa345 (2020).
37. Liu, Y. Y., Hsu, C. Y., Yang, Y. C., Chen, C. H. & Huang, C. C. ProbioMinServer: an integrated platform for assessing the safety and functional properties of potential probiotic strains. *Bioinform Adv.* **3** (1), vbad153. https://doi.org/10.1093/bioadv/vbad153 (2023).
38. Pal, C., Bengtsson-Palme, J., Rensing, C., Kristiansson, E. & Larsson, D. G. J. BacMet: antibacterial biocide and metal resistance genes database. *Nucleic Acids Res.* **42** (Database issue), D737–D743. https://doi.org/10.1093/nar/gkt1252 (2014).
39. Brown, C. L. et al. mobileOG-db: a manually curated database of protein families mediating the life cycle of bacterial mobile genetic elements. *Appl. Environ. Microbiol.* **88** (18), e0099122. https://doi.org/10.1128/aem.00991-22 (2022).
40. Carattoli, A. et al. In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. *Antimicrob. Agents Chemother.* **58** (7), 3895–3903. https://doi.org/10.1128/AAC.02412-14 (2014).
41. Kelliher, J. M. et al. Standardized and accessible multi-omics bioinformatics workflows through the NMDC EDGE resource. *Comput. Struct. Biotechnol. J.* **23**, 3575–3583. https://doi.org/10.1016/j.csbj.2024.09.018 (2024).
42. Camargo, A. P. et al. Identification of mobile genetic elements with geNomad. *Nat. Biotechnol.* **42** (8), 1303–1312. https://doi.org/10.1038/s41587-023-01953-y (2024).
43. WishartD.S. et al. PHASTEST: Faster than PHASTER, Better than PHAST. *Nucleic Acids Res. (Web Serv. Issue)*. https://doi.org/10.1093/nar/gkad382 (2023).
44. Gautreau, G. et al. DefenseFinder: a tool for discovering bacterial defense systems. *Bioinformatics* **36** (15), 4671–4673. https://doi.org/10.1093/bioinformatics/btaa366 (2020).
45. Payne, L. J. et al. PADLOC: a web server for the identification of antiviral defence systems in microbial genomes. *Nucleic Acids Res.* **50** (W1), W541–W550. https://doi.org/10.1093/nar/gkac398 (2022).
46. Grissa, I., Vergnaud, G. & Pourcel, C. CRISPRCasFinder: an update of CRISPRdetection tool. *Bioinformatics* **23** (6), 1426–1428. https://doi.org/10.1093/bioinformatics/btm087 (2007).
47. Marathe, N. P. et al. A treatment plant receiving waste water from multiple bulk drug manufacturers is a reservoir for highly multi-drug resistant integron-bearing bacteria. *PLoS One*. **8** (10), e77310. https://doi.org/10.1371/journal.pone.0077310 (2013).
48. Pruden, A. et al. Management options for reducing the release of antibiotics and antibiotic resistance genes to the environment. *Environ. Health Perspect.* **121** (8), 878–885. https://doi.org/10.1289/ehp.1206446 (2013).
49. Bengtsson-Palme, J., Kristiansson, E. & Larsson, D. G. J. Environmental factors influencing the development and spread of antibiotic resistance. *FEMS Microbiol. Rev.* **42** (1), 68–80. https://doi.org/10.1093/femsre/fux053 (2018).
50. Li, R. et al. Global dissemination of the tmexCD-toprJ resistance pump. *Nat. Microbiol.* **8** (6), 1179–1189. https://doi.org/10.1038/s41564-023-01394-6 (2023).
51. Strathdee, S. A., Hatfull, G. F., Mutalik, V. K. & Schooley, R. T. Phage therapy: from biological mechanisms to future directions. *Cell* **186** (1), 17–31. https://doi.org/10.1016/j.cell.2022.11.001 (2023).
52. Gophna, U. & Brodt, A. The effect of CRISPR-Cas systems on horizontal gene transfer and microbial evolution. *Trends Microbiol.* **26** (6), 525–535. https://doi.org/10.1016/j.tim.2018.01.004 (2018).
53. Wheatley, R. & MacLean, R. C. CRISPR-Cas systems restrict horizontal gene transfer in clinical Staphylococcus aureus strains. *Nat. Commun.* **12**, 4562. https://doi.org/10.1038/s41467-021-24688-2 (2021).
54. Bertelli, C. et al. IslandViewer 4: expanded prediction of genomic islands for larger-scale datasets. *Nucleic Acids Res.* **45** (W1), W30–W35. https://doi.org/10.1093/nar/gkx343 (2017).

Acknowledgements

The authors would like to acknowledge all participants who took part in this study. Authors are grateful to Dr. Kamruzzaman Pramanik head of Microbiology Industrial Irradiation Division (MIID); Dr. Zahid Hasan of MIID under the Institute of Food and Radiation Biology (IFRB), Bangladesh Atomic Energy Commission (BAEC) and Department of Microbiology, Jahangirnagar University for the logistics support for this study.

Author contributions

FA. contributed to the design and execution of the research experiments, ensuring alignment with the study's objectives. **MMHS.** performed comprehensive genomic data analyses and carried out all major bioinformatics interpretations. **MIH.** led the sampling activities and conducted statistical analyses, ensuring high-quality and reliable datasets. **MB.** actively participated in data analysis and manuscript drafting, ensuring clarity, logical flow, and coherence. **KM.** provided essential research assistance and played a key role in sampling and coordination of field activities. **SFC.** conducted sample preparation and genome extraction, maintaining strict quality control throughout the process. **SRN.** managed sample preparation and sequencing workflows, ensuring accuracy and consistency in data generation. **TM.** secured funding, supervised project management, and facilitated collaboration among team members and institutions. **MOF.** developed the research plan, managed the project's data framework, and led the manuscript writing, ensuring comprehensive documentation of the study. All authors reviewed the manuscript.

Funding
Not applicable.

Competing interests
The authors declare no competing interests.

Additional information
Correspondence and requests for materials should be addressed to T.M. or M.O.F.

Reprints and permissions information is available at www.nature.com/reprints.

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