BMC Microbiology

RESEARCH Open Access

Multidrug resistant hypervirulent ST307 clone from genomic surveillance of extended spectrum beta-lactamase-producing *Klebsiella pneumoniae* species complex in East Africa

Benson R. Kidenya[1*], Ivan Sserwadda[2,3,4], Stephen Kanyerezi[2,4], Inyasi Lawrence Akaro[5], Baraka Mkinze[5], Suhaila O. Hashim[3], Everlyne Isoe[3], Moses L. Joloba[2], Gerald Mboowa[2,4,7*] and Jeremiah Seni[6]

Abstract

Background Extended spectrum beta-lactamase-producing *Klebsiella pneumoniae* species complex (ESBL-KpSC) presents a significant therapeutic challenge to global public health, making it essential to assess the risks associated with recovered isolates. In this study, we utilized whole-genome sequencing (WGS) to comprehensively analyze ESBL-KpSC isolates from hospitalized orthopedic patients, their caretakers, and the hospital environments at tertiary referral hospitals in Uganda and Tanzania, East Africa.

Methods ESBL-KpSC isolates were collected between September 2019 and May 2020. Rectal swabs were obtained from patients shortly after their admission to orthopedic wards for ESBL-KpSC screening. Additional swabs were collected from caretakers, healthcare workers, and the surrounding hospital environments. Confirmed ESBL-KpSC isolates underwent DNA extraction for WGS, and the resulting sequences were analyzed to identify species, sequence type (ST), antimicrobial resistance (AMR) genes, virulence genes, and to calculate antimicrobial resistance and virulence scores.

Results We analyzed 24 ESBL-KpSC isolates, 7 (29.2%) from Uganda and 17 (70.8%) from Tanzania. Of these, 14 (58.3%) were identified as *Klebsiella pneumoniae*, 7 (29.2%) as *Klebsiella quasipneumoniae* subsp. *similipneumoniae*, and 3 (12.5%) as *Klebsiella variicola* subsp. *variicola*. The 24 ESBL-KpSC genomes were distributed across 19 sequence types (STs), indicating a high diversity of 79.2% (19/24). Among the 19 STs, two genomes were found in each of the following STs: ST17, ST307, ST2478, ST367-2LV, and ST3946-1LV, with the remaining genomes being singletons. Of the 24 ESBL-KpSC genomes, 6 (25.0%) had a virulence score greater than 0, and one isolate was identified as multidrug-resistant hypervirulent *K. pneumoniae* (MDR-hvKp). The most prevalent ESBL gene was $bla_{CTX-M-15}$, present in 95.8% (23/24) of isolates. Other common antimicrobial resistance (AMR) genes included bla_{TEM-1D} (79.2%, 19/24), *sul2* (75.0%, 18/24), *strB* (66.7%, 16/24), *qnrS1* (58.3%, 14/24), and *sul1* (58.3%, 14/24). Additionally, 14 different plasmid replicon

*Correspondence:
Benson R. Kidenya
benkidenya@gmail.com
Gerald Mboowa
gmboowa@gmail.com

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

types were identified, with one isolate carrying up to five plasmids. The most common plasmids were IncFIB(K) (87.5%, 21/24) and IncR (40.0%, 10/24).

Conclusion We report, for the first time, the presence of the MDR-hvKp ST307 clone from East Africa. One out of four ESBL-KpSC isolates was identified as a virulent strain harboring multiple AMR genes, underscoring the urgent need for routine screening of patients with prolonged hospital stays. Additionally, our findings emphasize the critical importance of robust infection prevention and control measures to mitigate the spread of AMR within hospitals in East Africa.

Keywords Clone, East africa, Extended spectrum beta-lactamase-producing *Klebsiella pneumoniae* species complex (ESBL-KpSC), Genomic surveillance, *Klebsiella pneumoniae* ST307, Multidrug-resistant hypervirulent

Introduction

The *Klebsiella pneumoniae* species complex (KpSC) is an opportunistic pathogen and has been reported as a leading cause of various multidrug-resistant (MDR) community-acquired and nosocomial infections, including pneumonia, liver abscesses, bacteremia, urinary tract infections, and wound infections [1–5]. Patients with prolonged hospital stays, such as those with exposed fractures, are at an increased risk of acquiring nosocomial infections [6–8]. The extended spectrum beta-lactamase-producing *Klebsiella pneumoniae* species complex (ESBL-KpSC) was the most frequently isolated pathogen from the blood of patients admitted to intensive care units in Tanzania [9]. The ESBL-KpSC were listed as pathogens of critical priority for the research, discovery, and development of antimicrobials by the World Health Organization (WHO) in 2017 [10]. In recent years there has been an emergence of multidrug resistant hypervirulent *K. pneumoniae* (MDR-hvKp) isolates which poses a severe therapeutic challenge to global public health [11]. Recent genomic analyses of *K. pneumoniae* have identified the emergence of international high-risk pandemic multidrug-resistant (MDR) clones, such as ST11, ST14, ST86, ST147, and ST307, in tertiary healthcare settings in Uganda [12].

Genomic analysis of clinical isolates identified as *K. pneumoniae* through biochemical tests and proteomics techniques, such as matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS), has revealed that *K. pneumoniae* is split into several related species and subspecies, collectively known as the *K. pneumoniae* species complex (KpSC) [13]. These KpSC member species exhibit 3% to 4% nucleotide divergence across core chromosomal genes, yet share a common pool of antimicrobial resistance (AMR) and virulence genes. Infections and outbreaks caused by KpSC members other than *K. pneumoniae* have been reported to contribute to 10% to 20% of the disease burden associated with KpSC [14]. Therefore, distinguishing *K. pneumoniae* isolates identified by biochemical methods into specific species and subspecies, particularly during genomic surveillance, is critical for accurate tracking and understanding of pathogenicity and resistance patterns.

Members of KpSC include *K. pneumoniae* (Kp1), *K. quasipneumoniae* subsp. *quasipneumoniae* (Kp2), *K. variicola* subsp. *variicola* (Kp3), *K. quasipneumoniae* subsp. *similipneumoniae* (Kp4), *K. variicola* subsp. *tropica* (Kp5), *K. quasivariicola* (Kp6) and *K. africana* (Kp7) [14]. All the KpSC members can be recovered from humans, animals as well as the environment, and can exhibit resistance to multiple drugs and virulent phenotypes globally [15]. The species in KpSC, particularly *K. pneumoniae*, can develop multidrug resistance to several antimicrobial classes, including carbapenems and colistin, through the accumulation of mutations in chromosomal core genes and the acquisition of various AMR genes through horizontal gene transfer [15]. The possession of mobile genetic elements has also led to the occurrence of hypervirulent *K. pneumoniae* (hvKp) that carry plasmids and integrative conjugative elements (ICEs) harboring numerous key virulence determinants, including siderophores [salmochelin (*iro*), aerobactin (*iuc*), and yersiniabactin (*ybt*), the genotoxin colibactin (*clb*)] and genes encoding for a hypermucoviscosity phenotype (*rmpA/rmpA2*) [16]. Genomic studies offer valuable insights into the dissemination dynamics of KpSC, including speciation, genotyping, serotyping, and the identification of genetic determinants linked to AMR and virulence. These studies also facilitate the assessment of virulence and AMR risk scores based on the presence of clinically relevant gene markers. However, despite these advances, knowledge regarding KpSC ecology, population structure, and pathogenicity remains relatively limited. Therefore, this study aimed to conduct a comprehensive genomic analysis of KpSC isolates from East African settings to address these gaps.

Materials and methods

Study design and settings

We utilized ESBL-KpSC isolates from two previously published surveillance studies conducted between September 2019 and May 2020 in the orthopedic units of two tertiary hospitals, one in Uganda and the other one

in Tanzania. In these studies, we screened a total of 428 participants, including index orthopedic patients, their neighboring patients during admission, non-medical caretakers, and immediate healthcare workers attending to the index patients. Additionally, we sampled 59 hospital environmental surfaces (such as walking crutches, wheelchairs, floors, door knobs, bathroom sinks, patient beds, headboards, and handrails), yielding 118 environmental samples. A total of 289 ESBL isolates were obtained, of which 56 (19.4%) were ESBL-KpSC, with 15 (26.8%) from Uganda and 41 (73.2%) from Tanzania. Detailed metadata and clinical data of the study participants and isolates have been previously published in our studies from Uganda [17] and Tanzania [18]. For this study, we randomly selected 24 ESBL-KpSC isolates for genomic analysis, including seven (29.1%) from Uganda and 17 (70.8%) from Tanzania. Whole genome sequencing (WGS) was performed to determine species, sequence type (ST), capsular polysaccharide (K) and lipopolysaccharide (LPS) O antigen serotypes/loci, plasmid replicon profiles, AMR determinants, and virulence genes. The isolates were also assessed for virulence and AMR risk.

Whole-genome sequencing

We analyzed whole-genome sequence data from 24 ESBL-KpSC isolates collected as part of the mother study titled "Understanding Transmission Dynamics and Acquisition of AMR at Referral Hospitals and Community Settings in East Africa using Conventional Microbiology and Whole-Genome Sequencing." This comprehensive genomic analysis aimed to investigate the genetic determinants associated with antimicrobial resistance and virulence in these isolates (Grant number GCA/AMR/rnd2/058) [17, 18]. The confirmed ESBL-KpSC isolates were shipped to the Earlham Institute in Norwich, United Kingdom, for whole-genome sequencing. The sequencing was conducted using the Low Input, Transposase Enabled (LITE) Illumina protocol on the Illumina NovaSeq 6000 platform (Illumina, San Diego, CA, USA), generating paired-end short reads of 2 × 250 bp and sequencing coverage depth of 20× and coverage breadth of 99.8%.

Bioinformatics analysis

The analysis of WGS data was conducted using our previously published Linux command line-based bioinformatics pipeline called "rMAP", the Rapid Microbial Analysis pipeline [19], and Kleborate software version 2.3.2 [14]. Raw sequences in the format fastq.gz were used as the input for the pipeline. All raw sequences were checked for quality in the rMAP pipeline using the embedded FastQC version 0.11.9 [20] to generate individual sample reports using MultiQC version 1.9 [21] for aggregating all the multiple reports into one report. Trimmomatic version 0.39 [22] was used to trim off adapters and to remove low-quality reads with the selected parameters minimum base length and phred score set to 200 and 32 respectively.

The trimmed reads were loaded into the Shovill version 1.0.9 [23], a short-read assembler for prokaryotic genomes for de-novo genome assembly. *K*-mer sizes 31, 55, 79, 103, and 127 were used to determine the optimum genome assembly. Pilon [24] was used for checking assembly errors, correcting ambiguous gaps, insertions, deletions and finally polishing the genomes.

Characterization of the assembled extended spectrum beta-lactamase-producing *Klebsiella pneumoniae* species complex genomes

The characterization of assembled ESBL-KpSC genomes was conducted using Kleborate version 2.3.2 [14], which was employed to identify species, determine sequence types (ST), predict serotypes or loci for K (capsule) and O antigens (LPS), and detect acquired virulence and AMR genes. The AMR and virulence scores were also computed using this software. Sequence typing was determined via multilocus sequence typing (MLST) against a set of seven KpSC housekeeping genes (*gapA*, *phoE*, *rpoB*, *infB*, *mdh*, *pgi*, and *tonB*).

For the prediction of K (capsular lipopolysaccharide) and O antigen (LPS) serotypes or loci, Kleborate utilizes *wzi* alleles and the Kaptive tool [25]. Kleborate also provided AMR and virulence scoring based on clinically relevant gene markers. Virulence scores ranged from 0 to 5, depending on the presence of three key loci: yersiniabactin (*ybt*), colibactin (*clb*), and aerobactin (*iuc*). The virulence score was assigned as follows: *yersiniabactin* alone (score = 1), *colibactin* without *aerobactin* (score = 2), *aerobactin* alone (score = 3), *aerobactin* and *yersiniabactin* without *colibactin* (score = 4), and the presence of all three (score = 5).

AMR scores ranged from 0 to 3, depending on the presence of genes associated with ESBL production (score = 1), carbapenemases (score = 2), and carbapenemases plus colistin resistance (score = 3). Kleborate also facilitated the rapid identification of strains with AMR-virulence convergence, defined as a virulence score ≥ 3 and resistance score ≥ 1. The plasmid replicon profile was determined from the assembled genomes using the PlasmidFinder database [26], analyzed through ABRicate version 1.0.1 [27].

Phylogenetic analysis of assembled extended spectrum beta-lactamase-producing *Klebsiella pneumoniae* species complex genomes

From the assembled genome contigs, Prodigal [28] was used to predict open reading frames, while Prokka

version 1.14.6 [29] facilitated rapid and efficient functional annotation of the ESBL-KpSC assembled genomes. The reads were trimmed and aligned to the indexed reference sequence in FASTA format (GenBank reference *K. pneumoniae* subsp. *pneumoniae* HS11286, Accession: NC_016845.1) using Burrows-Wheeler Aligner (BWA)-MEM version 0.7.17 [30] to generate sequence alignment map (SAM) files. Soft and hard clipped alignments were then removed from the SAM files using SAMclip version 0.4.0 [31]. Next, SAMtools version 1.9 [32] was used to sort, mark duplicates, and index the binary alignment map (BAM) files. Variants were called using Freebayes version 1.3.2 [33], which applies Bayesian models, generating variant call format (VCF) files containing single nucleotide polymorphism (SNP) data. These VCF files were subsequently filtered using BCFtools version 1.9 [34]. The filtered VCF files were annotated using snpEff version 4.5covid19 [35].

Phylogenetic inference based on maximum likelihood was performed using MAFFT version 7.471 [36], IQtree version 2.0.3 [37], vcf2phylip version 1.5 [38], and BMGE version 1.12 [39]. The rMAP pipeline aggregated all individual VCF files into a single VCF containing all samples and their SNPs, which was then transposed by vcf2phylip into a multi-alignment FASTA file. Multiple sequence alignment was conducted using the MAFFT software package, while ambiguous alignments were removed and informative sites extracted for phylogeny inference using BMGE. IQtree tested various substitution models and constructed trees using the maximum likelihood method with 1,000 bootstrap replicates. The resulting phylogenetic trees were visualized as rectangular phylograms and annotated using iTOL version 6 [40].

Plasmid reconstruction and comparative genomic analysis of extended spectrum beta-lactamase-producing *Klebsiella pneumoniae* species complex

Putative plasmid contigs were extracted from de novo assembled genomes, and plasmid replicon types were identified using the PlasmidFinder [26]. Contigs containing the most prevalent [IncFIB(K), IncR, IncY, IncFIB(Mar) and Col440I] replicons were selected for downstream comparative analysis. For each isolate, the corresponding IncFIB(K)-carrying contigs were compiled into multi-FASTA files for similarity assessment.

To evaluate sequence similarity among the selected plasmids, we employed Mash v2.3 [41], a tool that estimates pairwise genomic distances using the MinHash algorithm. Plasmid sequences were first converted into sketch files using the mash sketch command. Subsequently, pairwise comparisons were performed using *Mash dist*, which calculates the Mash distance, *p*-value, and number of shared k-mers between sequences.

Mash distance values were interpreted as follows: lower distances indicate higher nucleotide similarity, with values below 0.05 suggestive of highly similar or clonally related plasmids. Conversely, distances approaching 1.0 indicate negligible or no detectable sequence similarity, reflecting divergent or unrelated plasmid backbones. This approach enabled a rapid and scalable comparison of plasmid sequences to infer clonal relationships and potential evolutionary patterns among the isolates.

Results

Source of extended spectrum beta-lactamase-producing *Klebsiella pneumoniae* species complex isolates

A total of 24 ESBL-KpSC isolates were analyzed. Of these, 15 (62.5%) were obtained from orthopedic patients with open fractures, four (16.7%) from non-medical caretakers of orthopedic patients, four (16.7%) from the immediate hospital environment (including a sink in a common bathroom, a bed in the ward, the ward floor, and a wheelchair), and one (4.2%) isolate was collected from a healthcare worker (Table 1).

Species and lineages of extended spectrum beta-lactamase-producing *Klebsiella pneumoniae* species complex isolates

Of the 24 ESBL-KpSC isolates analyzed using Kleborate taxonomic criteria, three *Klebsiella* species were identified: *Klebsiella pneumoniae* [14 isolates (58.3%)], *Klebsiella quasipneumoniae* subsp. *similipneumoniae* [seven isolates (29.2%)], and *Klebsiella variicola* subsp. *variicola* [three isolates (12.5%)]. The 24 ESBL-KpSC genomes were grouped into 19 sequence types (STs), reflecting a high genetic diversity of 79.2% (19/24). Among the 19 STs, two genomes were represented in each of the following STs: ST17, ST307, ST2478, ST367-2LV, and ST3946-1LV, while the remaining genomes were singletons. (Table 1; Fig. 1).

Predicted capsular polysaccharide and lipopolysaccharide serotypes or loci of extended spectrum beta-lactamase-producing *Klebsiella pneumoniae* species complex isolates

Among the 24 ESBL-KpSC genomes analyzed, K (Capsular polysaccharide) loci were successfully predicted in 21 out of 24 (87.5%), while 3 genomes (12.5%) had Kaptive K loci confidence calls of 'Low' or 'None,' indicating fragmented K locus assemblies (i.e., the K locus was not fully assembled in a single contig). Of the 21 genomes with successfully predicted K loci, 16 different K serotypes/loci were detected. Three genomes (14.3%) had the KL102 locus, and two genomes (9.5%) were assigned to each of the following K serotypes: K12, K25, K28, and KL133. The remaining serotypes/loci were each represented by a single genome (Table 1).

Table 1 Demographic characteristics of 24 extended spectrum beta-lactamase-producing *Klebsiella pneumoniae* species complex isolates

SN	Country	ID	Source	Species	ST	ESBL gene	Antimicrobial resistance genes	Virulence genes (score)	Plasmid Replicon Profile	Serotype/ Locus K: O
1	Uganda	ID014	Patient	*Klebsiella pneumoniae*	297	$bla_{CTX-M-15}$	$bla_{SHV-220}$; *aac(6')-Ib-cr; aadA16; strA; strB; qnrB6; catII.2; arr-3; sul1; sul2; tet(D); dfrA27*	- (0)	Col440I; IncFIB(pKPHS1); IncR; IncY; pESA2	*K10 : O1*
2	Uganda	ID022	Patient	*Klebsiella pneumoniae*	567-1LV	$bla_{CTX-M-15}$	bla_{SHV-11}; bla_{OXA-1}; bla_{TEM-1D};*aac(3)-IIa; aac(6')-Ib-cr; aadA5; aph3-Ia; strA; strB; qnrS1;mphA; catII.2; sul1; sul2; dfrA17*	- (0)	IncFIB(K); IncFIB(Mar); IncHI1B	*K67 : O1*
3	Uganda	ID051II	Patient	*Klebsiella quasipneumoniae* subsp. *similipneumoniae*	824	$bla_{CTX-M-15}$	bla_{OXA-1}; *aac(3)-IIa; aac(6')-Ib-cr; aadA5; aph3-Ia; strB; qnrS1; mphA; catII.2; sul1; sul2; tet(A); dfrA14; dfrA17*	- (0)	IncFIB(K); IncFIB(Mar); IncHI1B	*K8 : O12*
4	Uganda	ID064	Patient	*Klebsiella pneumoniae*	188	$bla_{CTX-M-15}$	bla_{TEM-1D}; bla_{SHV-37}; *aac(6')-Ib-cr; aadA16; strA; strB; qnrS1; arr-3; sul1; sul2; dfrA27*	- (0)	Col440I; IncFIA(HI1); IncR	*KL116 : O1*
5	Uganda	ID100	Patient	*Klebsiella pneumoniae*	3586	$bla_{CTX-M-15}$	bla_{TEM-1D}; *strA; strB; qnrS1; catII.2; sul2; tet(A); dfrA14;*	- (0)	IncFIB(K); IncR	*KL102 : O2afg*
6	Uganda	ID009	Patient	*Klebsiella quasipneumoniae* subsp. *similipneumoniae*	489	bla_{SHV-18}	bla_{OXA-2}; *ant(2")-Ia; sul1*	- (0)	IncFIB(K); IncL/M(pMU407)	*K53 : -**
7	Uganda	ID096I	Patient	*Klebsiella pneumoniae*	39	$bla_{CTX-M-15}$	bla_{TEM-1D}; bla_{SHV-11}; *aac(3)-IIa; catA1; sul1; tet(A); dfrA7*	*ybt* (1)	IncFIB(K)	*K62 : O1*
8	Tanzania	ID014CT	Care taker	*Klebsiella quasipneumoniae* subsp. *similipneumoniae*	2478	$bla_{CTX-M-15}$	bla_{TEM-1D}; $bla_{OKP-B-5}$; *strA; strB; qnrS1;catII.2; sul2; dfrA14*	- (0)	IncFIB(K); IncY	*K28 : O3/ O3a*
9	Tanzania	ID025	Patient	*Klebsiella variicola* subsp. *variicola*	344-1LV	$bla_{CTX-M-15}$	bla_{TEM-1D}; *aadA; aadA5; sat-2; strA; strB; qnrS1; catII.2; dfrA17*	-	Col440I; IncFIB(K); IncY	*-* : -**
10	Tanzania	1D037	Patient	*Klebsiella pneumoniae*	607	$bla_{CTX-M-15}$	bla_{TEM-1D}; $bla_{SHV-148}$; *aac(3)-IIa; aac(6')-Ib-cr; aadA16; qnrB6;arr-3; sul1; sul2; tet(D); dfrA27*	*ybt* (1)	IncFIA(HI1); IncFIB(K); IncR	*K25 : 01*

Table 1 (continued)

SN	Country	ID	Source	Species	ST	ESBL gene	Antimicrobial resistance genes	Virulence genes (score)	Plasmid Replicon Profile	Serotype/ Locus K: O
11	Tanzania	ID06CT	Care taker	*Klebsiella pneumoniae*	17	$bla_{CTX-M-15}$	bla_{TEM-1D}; bla_{SHV-11}; *strA; strB; qnrS1; sul2; dfrA14*	*ybt* (1)	Col(IRGK); IncFIB(K); IncFIB(pKPHS1)	*K25 : O5*
12	Tanzania	ID08CT	Care taker	*Klebsiella pneumoniae*	307	$bla_{CTX-M-15}$	bla_{TEM-1D}; bla_{SHV-28}; *aac(3)-IIa; aac(6')-Ib-cr; aadA16; strA; strB; arr-3; sul1; sul2; dfrA27*	*ybt* (1)	IncFIA(HI1); IncFIB(K); IncR	*KL102 : O2afg*
13	Tanzania	ID0125	Patient	*Klebsiella pneumoniae*	20	$bla_{CTX-M-15}$	bla_{TEM-1D}; $bla_{SHV-187}$; *aac(6')-Ib-cr; aadA16; qnrS1; catII.2; arr-3; sul1; tet(A); dfrA27*	*ybt* (1)	IncFIB(K); IncFIB(Mar)	*K24 : O1*
14	Tanzania	ID135	Patient	*Klebsiella pneumoniae*	495	$bla_{CTX-M-15}$	bla_{TEM-1D}; bla_{SHV-1}; *aac(6')-Ib-cr; aadA16; qnrS1; catII.2; arr-3; sul1; tet(A); dfrA27*	- (0)	IncFIB(K); IncFIB(Mar) IncR	*KL169 : OL104*
15	Tanzania	ID187CT	Care taker	*Klebsiella pneumoniae*	17	$bla_{CTX-M-15}$	bla_{TEM-1D}; bla_{SHV-11}; *strA; strB; qnrS1; sul2; dfrA14*	- (0)	IncFIB(K)	*KL112 : O2afg*
16	Tanzania	ID188	Patient	*Klebsiella pneumoniae*	491	$bla_{CTX-M-15}$	bla_{TEM-1D}; bla_{SHV-11}; *strB; qnrS1; sul2; tet(A); dfrA14*	- (0)	Col440I; IncFIB(K); IncHI1B; IncY	*-* : -**
17	Tanzania	ID198	Patient	*Klebsiella pneumoniae*	307	$bla_{CTX-M-15}$	*AmpC1;* bla_{OXA-1}; bla_{TEM-1D}; *aac(3)-IIa; aac(6')-Ib-cr; strA; strB; qnrB1; sul1; sul2; tet(A); dfrA14*	*ybt; iuc* (4)	IncFIB(K); IncFII	*KL102 : O2afg*
18	Tanzania	ID260	Patient	*Klebsiella quasipneumoniae* subsp. *similipneumoniae*	2478	$bla_{CTX-M-15}$	bla_{TEM-1D}; $bla_{OKP-B-5}$; *strA; strB; qnrS1; catII.2; sul2; tet(A); dfrA14*	- (0)	IncFIB(K); IncY	*K28 : O3/ O3a*
19	Tanzania	ORTHO053E	Sink bathroom	*Klebsiella quasipneumoniae* subsp. *similipneumoniae*	367-2LV	$bla_{CTX-M-15}$	bla_{TEM-1D}; *strB; qnrS1; sul2; tet(A); dfrA14*	- (0)	IncFIB(K); IncR; IncY	*K12 : OL103*
20	Tanzania	ORTHO066E	Bed	*Klebsiella quasipneumoniae* subsp. *similipneumoniae*	367-2LV	$bla_{CTX-M-15}$	bla_{ACT-6}; bla_{OXA-1}; *aac(3)-IIa; aac(6')-Ib-cr; aadA; strB; fosA2; qnrB1; qnrS1; sul1; sul2; tet(A); dfrA14; dfrA15*	- (0)	IncFIB(K); IncFIB(pB171); IncFII(Yp); IncR; IncY	*K12 : OL103*

Table 1 (continued)

SN	Country	ID	Source	Species	ST	ESBL gene	Antimicrobial resistance genes	Virulence genes (score)	Plasmid Replicon Profile	Serotype/ Locus K: O
21	Tanzania	ORTHO067E	Floor	*Klebsiella pneumoniae*	290	$bla_{CTX-M-15}$	bla_{OXA-1}; bla_{TEM-1D}; bla_{SHV-1}; *aac(3)-IIa; aac(6')-Ib-cr; strA; strB; qnrB1; sul2; tet(A); tet(D); dfrA14*	- (0)	Col440I; IncFIB(pKPHS1); IncR	*K21 : O1*
22	Tanzania	ORTHO-8HWII	Health care worker	*Klebsiella variicola* subsp. *variicola*	3946-1LV	$bla_{CTX-M-15}$	bla_{OXA-1}; bla_{TEM-1D}; *aac(3)-IIa; qnrB6;catII.2; arr-3; sul1; sul2; dfrA27*	- (0)	IncFIB(K); IncFIB(Mar)	*KL133 : O2a*
23	Tanzania	ID233	Patient	*Klebsiella variicola* subsp. *variicola*	3946-1LV	$bla_{CTX-M-15}$	bla_{OXA-1}; bla_{TEM-1D}; *aac(3)-IIa; aadA16; qnrB6; catII.2; arr-3; sul1; sul2; dfrA27*	- (0)	IncFIB(K); IncFIB(Mar)	*KL133 : O2a*
24	Tanzania	ORTHO055E	Wheel chair	*Klebsiella quasipneumoniae* subsp. *similipneumoniae*	334-1LV	$bla_{CTX-M-15}$	$bla_{OKP-B-10}$; *aac(3)-IIa*	- (0)	IncFIB(K); IncFIB(Mar); IncHI1B; IncR	*K48 : -**

SN Serial number, *ID* Identification Number for the isolates, *ESBL* Extended spectrum beta-lactamase

*The locus for K (capsular polysaccharide) and/or O (lipopolysaccharide) was not successfully determined as Kaptive locus confidence call was 'Low' or 'None' and the Kaptive output indicated a fragmented locus assembly (i.e. the locus assembly was not found in one contig)

Similarly, O (Lipopolysaccharide) loci were successfully predicted in 20 out of 24 genomes (83.3%), while 4 genomes (16.7%) had Kaptive O loci confidence calls of 'Low' or 'None,' indicating fragmented O locus assemblies (i.e., the O locus was not found in one contig and/or gene deletions occurred). Among the 20 genomes with successfully predicted O loci, 8 different O serotypes/loci were identified. Seven genomes (35.0%) were O1 serotype, four genomes (20.0%) were O2afg serotype, and two genomes (10.0%) each corresponded to the O2a, O3/O3a, and OL103 loci. The remaining loci were each represented by a single genome (Table 1).

Virulence genes and scores of extended spectrum beta-lactamase-producing *Klebsiella pneumoniae* species complex isolates

Of the 24 ESBL-KpSC isolates, six (25.0%) had a virulence score greater than 0, with five genomes (20.8%) scoring 1 and one genome (4.2%) scoring 4. Virulence score of 0 was assigned to 18 (75.0%) genomes. The identified virulence genes encoded for the siderophores yersiniabactin (*ybt*) and aerobactin (*iuc*). None of the genomes contained genes coding for the genotoxin colibactin (*clb*), salmochelin (*iro*), or hypermucoviscosity-associated genes (*rmpADC* and *rmpA2*) (Table 1; Fig. 1).

Antimicrobial resistance genes and scores of extended spectrum beta-lactamase-producing *Klebsiella pneumoniae* species complex isolates

All 24 ESBL-KpSC isolates analyzed had an AMR score of 1, as none of them harbored resistance genes coding for carbapenemases or conferring resistance to colistin. The most common ESBL gene identified was $bla_{CTX-M-15}$, present in 95.8% (23/24) of the isolates, followed by bla_{SHV-18}, found in 4.2% (1/24). The number of drug resistance classes per genome ranged from two to nine, with a median of seven [IQR: 7–8]. The number of AMR genes per genome varied from two to 15, with a median of 10 [IQR: 7–12].

Apart from ESBL genes, the most frequently identified AMR genes encoded resistance to beta-lactams (bla_{TEM-1D} 79.2% [19/24]), sulfonamides (*sul2* 75.0% [18/24] and *sul1* 58.3% [14/24]), quinolones (*qnrS1* 58.3% [14/24]), and aminoglycosides (*strB* 66.7% [16/24], *strA* 50.0% [12/24], *acc(6')-Ib-cr* 45.8% [11/24], and *aac(3)-IIa* 45.8% [11/24]). Other notable resistance genes included those for tetracyclines (*tet(A)* 45.8% [11/24]), phenicols (*catII.2* 45.8% [11/24]), and trimethoprims (*dfrA14* 45.8% [11/24]) (Tables 1, 2 and 3; Fig. 1).

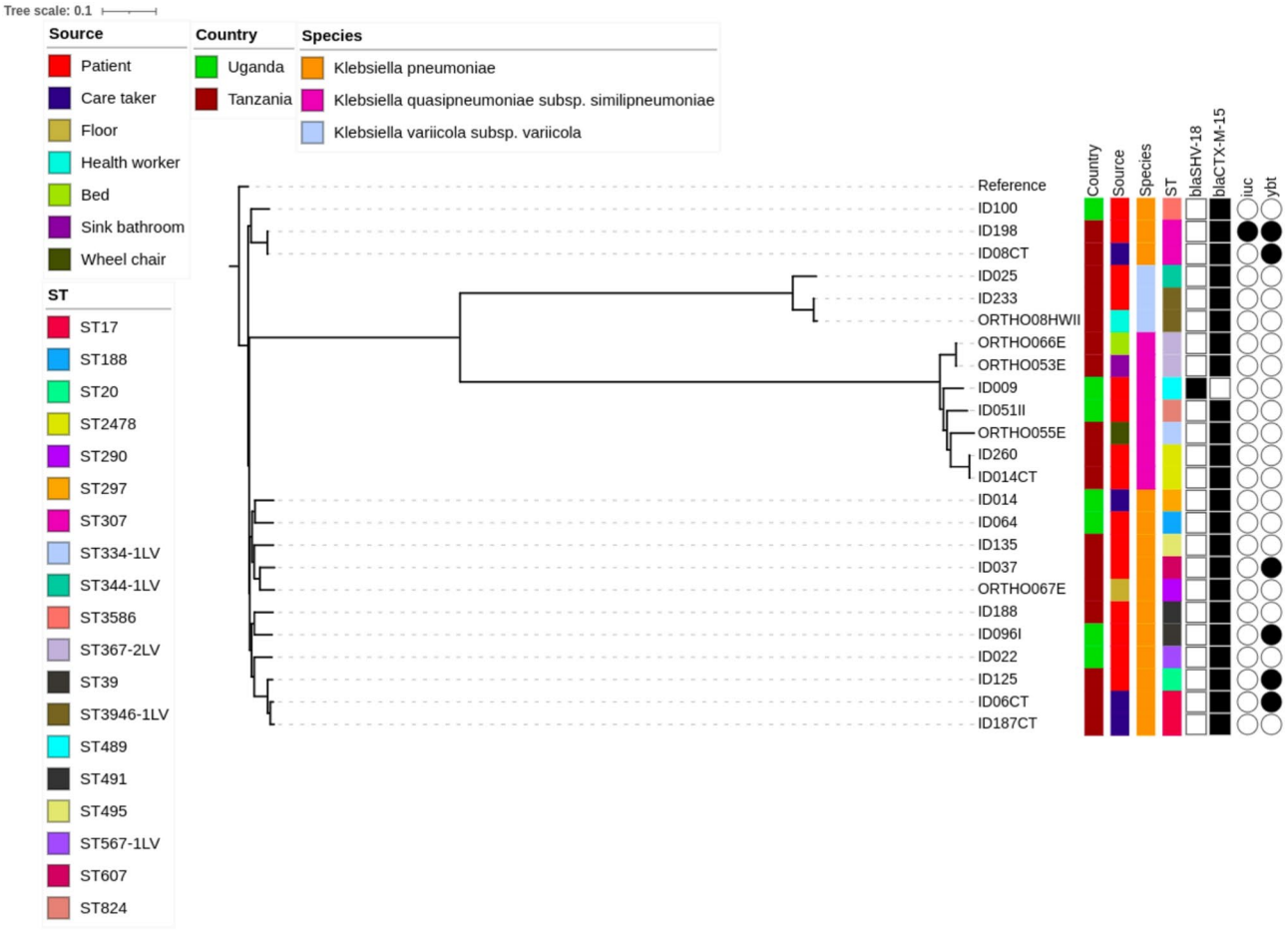

Fig. 1 A phylogenetic analysis was conducted to demonstrate the diversity of ESBL-producing *Klebsiella pneumoniae* species complex genomes from Uganda and Tanzania. The resulting phylogenetic tree highlights the significant diversity of the isolates analyzed in this study, showing multiple sequence types (STs). This wide diversity indicates a lack of a clear clonal pattern, even at the country or hospital level

Convergence of antimicrobial resistance-virulence of extended spectrum beta-lactamase-producing *Klebsiella pneumoniae* species complex isolates

Of the 24 ESBL-KpSC genomes analyzed, one ST307 isolate exhibited MDR-hypervirulence convergence, defined as genomes with a virulence score ≥ 3 and an AMR score ≥ 1 [14]. This strain, isolated from a patient in Tanzania, was identified as *K. pneumoniae* harboring the K (capsular polysaccharide) locus KL102 and O (lipopolysaccharide) serotype O2afg. Additionally, the strain contained 13 AMR genes ($bla_{CTX-M-15}$, *AmpC1*, bla_{OXA-1}, bla_{TEM-1D}, *aac(3)-IIa*, *aac(6')-Ib-cr*, *strA*, *strB*, *qnrB1*, *sul1*, *sul2*, *tet(A)*, and *dfrA14*), conferring resistance to seven classes of antimicrobials (Table 1).

Plasmid replicon profiles of extended spectrum beta-lactamase-producing *Klebsiella pneumoniae* species complex isolates

In all 24 ESBL-KpSC isolates analyzed, we used the ABRicate tool with the PlasmidFinder database and identified a total of 14 different plasmid replicon types. The isolates contained between one and five plasmids each, with the most common plasmids being *IncFIB*(K) present in 87.5% (21/24) of the isolates, followed by IncR in 41.7% (10/24), IncY in 29.2% (7/24), and IncFIB(Mar) in 29.2% (7/24) (Tables 1 and 4).

Phylogenetic relatedness of extended spectrum beta-lactamase-producing *Klebsiella pneumoniae* species complex isolates from Uganda and Tanzania

Our phylogenetic analysis of the 24 ESBL-KpSC genomes from Uganda and Tanzania revealed a heterogeneous tree, indicating a wide diversity of sequence types (STs) among the isolates. The analysis depicted limited clonal patterns, even at the country and hospital levels, further emphasizing the genetic diversity of the isolates studied (Fig. 1).

Interpretation of plasmid mash results from extended spectrum beta-lactamase-producing *Klebsiella pneumoniae* species complex isolates

Mash distance analysis of *Klebsiella pneumoniae* clinical isolates revealed distinct diversity patterns across

Table 2 Distribution of genetic variants of beta-lactamase among 24 extended spectrum beta-lactamase-producing *Klebsiella pneumoniae* species complex isolates

Beta-lactamase characteristics	Gene variant	Number (*n*)	Percent (%)
Acquired Extended spectrum beta-lactamase	$bla_{CTX-M-15}$	23	95.8
	bla_{SHV-18}	1	4.2
Acquired beta-lactamase	bla_{TEM-1D}	19	79.2
	bla_{OXA-1}	7	29.2
	bla_{OXA-2}	1	4.2
	bla_{ACT-6}	1	4.2
	bla_{AmpC1}	1	4.2
Chromosomal beta-lactamase	bla_{SHV-37}	8	33.3
	bla_{SHV-11}	5	20.8
	bla_{SHV-1}	2	8.3
	$bla_{OKP-B-5}$	2	8.3
	bla_{SHV-28}	1	4.2
	$bla_{SHV-148}$	1	4.2
	$bla_{SHV-187}$	1	4.2
	$bla_{SHV-220}$	1	4.2
	$bla_{OKP-B-10}$	1	4.2

Table 3 Distribution of antimicrobial resistance genes among 24 extended spectrum beta-lactamase-producing *Klebsiella pneumoniae* species complex isolates

Non β-lactam Antimicrobials	Antimicrobial resistance gene	Number (*n*)	Percent (%)
Quinolones	*qnrS1*	14	58.3
(Nalidixic acid, ciprofoxacin)	*qnrB6*	4	16.7
	qnrB1	3	12.5
Suphonamides	*sul2*	18	75.0
	sul1	14	58.3
Aminoglycosides	*strB*	16	66.7
(Streptomycin, gentamycin, tobramycin)	*strA*	12	50.0
	aac(6')-Ib-cr	11	45.8
	aac(3)-IIa	11	45.8
	aadA16	7	29.2
	aadA5	3	12.5
	aadA	2	8.3
	aph(3')-Ia	2	8.3
	sat-2	1	4.2
	ant(2'')-Ia	1	4.2
Fosfomycin	*fosA2*	1	4.2
Tetracyclines	*tet(A)*	11	45.8
	tet(D)	3	12.5
Phenicols (Chloramphenicols)	*catII.2*	11	45.8
	catA1	1	4.2
Trimethoprim	*dfrA14*	11	45.8
	dfrA27	8	33.3
	dfrA17	3	12.5
	dfrA15	1	4.2
	dfrA7	1	4.2
Rifampicin	*arr-3*	8	33.3
Macrolides	*mphA*	2	8.3

Table 4 Distribution of plasmid replicons among extended spectrum beta-lactamase-producing *Klebsiella pneumoniae* species complex isolates

Plasmid Replicon	Number (*n*)	Percent (%)	ST
IncFIB(K)	21	87.5	17; 17; 20; 39; 307; 307; 334-1LV; 344-1LV; 367-2LV; 367-2LV; 489; 491; 495; 567-1LV; 607; 824; 2478; 2478; 3586; 3946-1LV; 3946-1LV
IncR	10	41.7	188; 290; 297; 307; 334-1LV; 367-2LV; 367-2LV; 495; 607; 3586
IncY	7	29.2	297; 344-1LV; 367-2LV; 367-2LV; 491; 2478; 2478
IncFIB(Mar)	7	29.2	20; 334-1LV; 495; 567-1LV; 824; 3946-1LV; 3946-1LV
Col440I	5	20.8	188; 290; 297; 344-1LV; 491
IncHI1B	4	16.7	334-1LV; 491; 567-1LV; 824
IncFIA(HI1)	3	12.5	307; 188; 607
IncFIB(pKPHS1)	3	12.5	17; 290; 297
Col(IRGK)	1	4.2	17
IncFIB(pB171)	1	4.2	367-2LV
IncFII(Yp)	1	4.2	367-2LV
IncFII	1	4.2	307
IncL/M(pMU407)	1	4.2	489
pESA2	1	4.2	297

plasmid replicon types, reflecting varying degrees of conservation and evolutionary divergence. IncR plasmids exhibited moderate genetic diversity, with most pairwise Mash distances ranging between 0.072 and 0.102. This range suggests substantial sequence variability among the IncR plasmids, indicative of divergent evolutionary paths. However, select isolate pairs—such as ID014 with ORTHO053E and ORTHO066E—demonstrated lower Mash distances (0.033–0.043), suggesting higher sequence similarity and potentially shared evolutionary ancestry. These results highlight a heterogeneous IncR plasmid population, potentially harboring diverse antimicrobial resistance determinants shaped by distinct recombination and selection events.

IncY plasmids, in contrast, displayed complete sequence identity across all comparisons (Mash distance = 0, 745/745 shared k-mers), signifying the presence of a single, highly conserved IncY plasmid. This uniformity points to clonal dissemination and suggests a stable plasmid backbone that may facilitate the persistent spread of resistance or virulence genes in this clinical setting.

IncFIB(Mar) plasmids were also highly conserved. Five out of six pairwise comparisons showed zero Mash

distance, consistent with identical plasmid sequences and supporting recent clonal transmission. A single comparison (ID022 vs. ORTHO055E) yielded a Mash distance of 0.0074, reflecting near-identity and suggesting minimal variation likely due to small-scale structural changes. These findings indicate a clonally circulating IncFIB(Mar) plasmid with potential implications for the dissemination of conserved resistance loci.

IncFIB(K) plasmids revealed a bimodal pattern of diversity. The majority of comparisons—such as ID009 vs. ID051II, ID135, and ID198—produced low Mash distances (0.0045–0.0216), indicative of highly similar or nearly identical sequences, likely resulting from clonal expansion or recent horizontal gene transfer. In contrast, other comparisons involving isolates like ID014CT, ORTHO08HWII, ORTHO053E, and ORTHO066E exhibited higher Mash distances (0.085–0.094), signifying greater sequence divergence. This suggests the coexistence of a dominant, conserved IncFIB(K) lineage with more divergent plasmid variants, reflecting both clonal spread and underlying genomic plasticity.

Col440I plasmids showed a combination of conserved and slightly divergent sequences. An exact match (Mash distance = 0) between isolates ID014 and ID064 suggests recent common ancestry or clonal dissemination. Other pairwise distances ranged from 0.024 to 0.035, indicating limited sequence variation potentially driven by recombination or gene acquisition. These findings support a largely conserved Col440I plasmid population with evidence of close epidemiological links among clinical isolates.

Discussion

Conducting genomic surveillance of ESBL-KpSC is essential due to the pathogen's high genome plasticity, which plays a key role in the spread of AMR and the acquisition of virulence factors [42]. This plasticity is driven by the pathogen's strong propensity for horizontal gene transfer and the accumulation of mutations in its core genome [42], thereby promoting the emergence of high-risk clones that evolve into dominant strains and cause outbreaks. Additionally, KpSC is widely distributed and can be found in various sources, including humans, animals, hospitals, natural environments, and food [43]. Genomic surveillance of ESBL-KpSC helps us better understand the genetic diversity, prevalence of AMR, and virulence factors, allowing for an assessment of the potential pathogenic risk of circulating KpSC lineages in the region. This knowledge is vital for the prevention and control of diseases caused by KpSC.

Our study uncovered the presence of convergent multidrug-resistant hypervirulent *K. pneumoniae* (MDR-hvKp) of the ST307 lineage among human carriers. To the best of our knowledge, this is the first report of an ST307 MDR-hvKp isolate from East Africa. ST307 MDR-hvKp is recognized as a successful global clone, having been previously reported in four healthcare facilities in Germany. Multiple outbreaks of ST307 MDR-hvKp have also been documented in clinical settings, underscoring the importance of continued genomic surveillance [44, 45].

The *K. pneumoniae* ST307 isolate was first reported in the Netherlands in 2008 and then in Pakistan in 2009. Since those initial reports, erratic isolations have been observed across Europe, Asia, Africa, and the Americas [46]. Our MDR-hvKp genome did not contain the *rmpA* or *rmpA2* genes, which are typically associated with the hypermucoviscosity phenotype. It is worth noting that hvKp strains lacking *rmpA/rmpA2* genes (and the hypermucoviscosity phenotype) have been frequently identified [47, 48]. Following the initial emergence of hypervirulent *K. pneumoniae* (hvKp) strains in Asia, these pathogens have spread globally, contributing to both community- and hospital-associated infections. In our study, *K. pneumoniae* was the most frequently identified species, aligning with findings from previous research. We observed substantial genetic diversity among the ESBL-producing *K. pneumoniae* species complex (KpSC) isolates: 24 typed isolates clustered into 19 distinct sequence types (STs), representing 79.2% genetic diversity. Notably, several STs—including ST17, ST307, ST2478, ST367-2LV, and ST3946-1LV were each detected in two isolates. Among these, two ST307 isolates were recovered in Tanzania from a patient and their caretaker. The overall distribution of STs varied across countries and participant categories, with no single lineage predominating. This highlights a complex and heterogeneous landscape of *K. pneumoniae* transmission and evolution in East Africa.

The convergence of hypervirulence and AMR in hvKp strains further exacerbates the public health crisis by complicating treatment options for MDR-hvKp infections. This convergence facilitates the development of invasive and difficult-to-treat infections. From an epidemiological standpoint, the exact source and transmission route of hvKp infections remain poorly understood, but intestinal colonization appears to be a critical precursor to infection [47]. MDR *K. pneumoniae* and hypervirulent *K. pneumoniae* (hvKp) have traditionally been associated with two distinct subpopulations of *K. pneumoniae* lineages, distinguished by the presence of acquired resistance genes and specific virulence-associated loci. In recent years, however, there has been a growing number of reports on convergent MDR-hvKp strains, which simultaneously exhibit both MDR and hypervirulence [49, 50]. These MDR-hvKp isolates display high genetic diversity and possess varying AMR profiles [51]. Given this emerging threat, the development

of non-antimicrobial therapies targeting virulence factors has become an indispensable approach to combating these infections. Therefore, further research into *hvKp* is essential to improve the prevention, diagnosis, and treatment of infections caused by these strains. Moreover, hypervirulent and MDR-hv strains of *Klebsiella* species other than *K. pneumoniae*, including *K. quasipneumoniae* subsp. *similipneumoniae* and *K. variicola*, have also been reported, presenting new challenges for public health [13, 52, 53].

We found that the most common acquired ESBL gene was $bla_{\text{CTX-M-15}}$, and the most prevalent acquired beta-lactamase gene conferring resistance to beta-lactams was $bla_{\text{TEM-1D}}$. This result aligns with previous findings from a study in Malawi [54], which found $bla_{\text{CTX-M-15}}$ being most common. The Malawian study samples were blood and CSF, and carriage isolates from rectal swabs taken from adult and pediatric patients. While $bla_{\text{CTX-M-15}}$ is typically reported in *Escherichia coli*, its presence in *K. pneumoniae* can be attributed to the sharing of gut microbiota between humans and *E. coli*, as well as *K. pneumoniae*'s ability to acquire antimicrobial resistance genes from other strains, species, and genera through horizontal gene transfer [42].

The most frequently detected non-beta-lactam AMR genes included those conferring resistance to sulfonamides (*sul2* [75.0%] and *sul1* [58.3%]), aminoglycosides (*strB* [66.7%], *strA* [50.0%], *acc(6')-Ib-cr* [45.8%], and *aac(3)-IIa* [45.8%]), quinolones (*qnrS1* [58.3%]), tetracyclines (*tet(A)* [45.8%]), phenicols (*catII.2* [45.8%]), and trimethoprims (*dfrA14* [45.8%]). The high prevalence of these genes may be attributed to the extensive use of these antimicrobials as frontline treatments for infections like UTIs, bacteremia, and wound sepsis, which are common reasons for hospital visits. The increasing antimicrobial resistance (AMR) in ESBL-KpSC isolates to these essential drugs is alarming, as it significantly reduces available treatment options for such infections. This trend underscores the growing threat of pan-drug resistance within ESBL-KpSC, further complicating infection management strategies [50].

Our study highlights that one out of four ESBL-KpSC genomes contained genes encoding the siderophores yersiniabactin (*ybt*) and/or aerobactin (*iuc*), which significantly contribute to the virulence phenotype of the KpSC [14]. Iron is a crucial nutrient for both humans and pathogens, playing a key role in essential metabolic processes such as respiration, DNA synthesis, and repair [55, 56]. KpSC produces various siderophores—small molecules that sequester iron from the host's iron-chelating proteins [55, 56]. The production of multiple siderophores is a critical determinant of virulence, enhancing the pathogen's ability to cause infection. In this context, siderophore production is one of the key virulence factors in KpSC [14].

In our study, one isolate (ID198, ST307) from Tanzania harbored genes for both yersiniabactin and aerobactin biosynthesis, classifying it as a hypervirulent and MDR strain. This poses a significant threat to orthopedic patients, as these virulent ESBL-KpSC strains can lead to severe infections. Moreover, research has shown that siderophores can also act as chelators for zinc and copper, protecting *K. pneumoniae* from oxidative stress, further cementing siderophores as indispensable virulence factors [47, 57].

Our analysis revealed that two ST307 *K. pneumoniae* isolates from both the patient and caretaker in Tanzania, harbored the KL102:O2afg serotype combination. The KL102 locus has been previously associated with enhanced capsule production and evasion of host immune responses, while the O2afg (a variant of O2 associated with the *wbbY* gene) has been implicated in reduced innate immune activation [58]. This serotype profile is consistent with reports of global ST307 lineages carrying KL102 and O2afg loci, reinforcing the clone's capacity for persistence and transmission in healthcare environments. The co-occurrence of this locus with extended-spectrum β-lactamase genes underscores its contribution to the pathogenic success of this high-risk clone.

Plasmid analysis revealed that the majority of the isolates harbored IncF-type plasmids, which are well-recognized for their role in disseminating multidrug resistance determinants [59], including *blaCTX-M-15*. These plasmids are known to be highly stable, conjugative, and often carry multiple resistance genes, thus enhancing the survival and adaptability of host bacteria under antimicrobial pressure [60]. Notably, the consistent presence of these plasmids among isolates from both patients and healthcare-associated individuals in Tanzania underscores their likely contribution to the successful colonization and spread of multidrug-resistant *Klebsiella pneumoniae*.

A limitation of our study is the small sample size of ESBL-KpSC genomes analyzed from the two countries, which may limit the generalizability of our findings. The small number of isolates may not fully capture the genetic diversity and AMR profiles present in the broader population. This could affect the ability to identify fewer common strains or resistance mechanisms, which may play a significant role in the epidemiology of ESBL-KpSC infections. Future studies with larger sample sizes are necessary to validate these findings and provide a more comprehensive understanding of ESBL-KpSC distribution and evolution. Furthermore, we acknowledge that a systematic characterization of the antimicrobial resistance (AMR) profiles was not fully undertaken in this

study. This was primarily due to a lack of access to clinical prescription data availability. However, we performed phenotypic antimicrobial susceptibility test for these isolates.

These plasmid Mash distance analyses underscore the coexistence of both conserved and genetically diverse plasmid populations among clinical ESBL-KpSC isolates, highlighting the dual roles of clonal dissemination and horizontal gene transfer in shaping the local resistome and informing future surveillance strategies. Our future work will include the integration of publicly available ESBL-KpSC genomes from regional and global databases to enhance phylogenetic resolution, thereby enabling a more comprehensive understanding of transmission networks and the evolutionary origins of ESBL-producing strains in East Africa.

In conclusion, this is the first report of ST307 MDR-hvKp in East Africa. We demonstrate that orthopedic patients, their caretakers, healthcare workers, and hospital environments act as reservoirs for virulent ESBL-KpSC strains with multiple AMR genes. The study underscores the urgent need for routine screening, particularly in patients with long hospital stays or extensive antimicrobial treatment, such as orthopedic patients with open fractures. Additionally, it highlights the importance of infection prevention and control measures to reduce AMR dissemination within East Africa's hospital systems.

Abbreviations

AMR	Antimicrobial resistance
CUHAS	Catholic University of Health and Allied Sciences
ESBL	Extended spectrum beta-lactamase
KpSC	Klebsiella pneumoniae Species Complex
LPS	Lipopolysaccharide
MDR	Multidrug resistance
WGS	Whole-genome sequencing
rMAP	Rapid Microbial Analysis Pipeline
STs	Sequence types

Acknowledgements

G.M is grateful for the support of Data Science for Health Discovery and Innovation in Africa (DS-I Africa) funding Fogarty program to Makerere University Data Science Research Training to Strengthen Evidence-Based Health Innovation, Intervention and Policy (MakDARTA), award number 5U2RTW012116 through the Fogarty International Center at the U.S. National Institutes of Health (NIH). Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the supporting offices.

G.M is an EDCTP2 career development fellow supported by the PHICAMS (Pathogen detection in HIV-infected children and adolescents with non-malarial febrile illnesses using metagenomic next-generation sequencing approach in Uganda) grant which is part of the EDCTP2 program supported by the European Union (grant number TMA2020CDF-3159). The views and opinions of the authors expressed herein do not necessarily state or reflect those of EDCTP.

Authors' contributions

BRK, GM, IS, and JS made substantial contributions to the conception, designing, coordination, and execution of the study; ILA, BM, and MLJ coordinated and executed the study; IS, SK, BRK, and GM participated in the analysis and interpretation of the data; BRK, IS, GM, JS, SOH, and EI wrote the initial draft of the manuscript, which was critically revised by all authors. All the authors reviewed and approved the final manuscript.

Funding

This study was part of the research project titled "Understanding transmission dynamics and acquisition of AMR at referral hospitals and community settings in East Africa", supported through the Grand Challenges Africa program (grant number GCA/AMR/058). Grand Challenges Africa is a program of the African Academy of Sciences (AAS) implemented through the Alliance for Accelerating Excellence in Science in Africa (AESA) platform, an initiative of the AAS and the African Union Development Agency (AUDA-NEPAD). GC Africa is supported by the Bill and Melinda Gates Foundation (BMGF) and the African Academy of Sciences and partners. It was executed collaboratively between Makerere University College of Health Sciences, Kampala, Uganda, and the Catholic University of Health and Allied Sciences, Mwanza, Tanzania. The funders had no roles in the design, execution, data management, manuscript writing, or manuscript submission. This work was also in part supported by the African Association for Research and Control of Antimicrobial Resistance (AAAMR) 2020 Young Anglophone Investigator Award to G.M and the Public Health Alliance for Genomic Epidemiology (PHA4GE) grant (INV-038071) to support pathogen genomics data standards and bioinformatics interoperability awarded to I.S.

Data availability

All source code for the rMAP pipeline, installation instructions, and implementation can be accessed via GitHub (https://github.com/Gunzlvan28/rMAP). The source code is available on GitHub under the GPL3 license. Questions, bugs, or any other issues can be filed as GitHub issues. Although rMAP itself is published and distributed under a GPL3 license, some of its dependencies bundled within the rMAP volume are published under different license models.The raw read files from this study are publicly available at the Sequence Read Archive (SRA) of the National Center for Biotechnology Information (NCBI) under the study BioProject ID: PRJNA951629 at https://www.ncbi.nlm.nih.gov/bioproject/PRJNA951629.

Declarations

Ethics approval and consent to participate

This study was conducted in accordance with the Declaration of Helsinki. Ethical approvals were obtained from the; School of Biomedical Sciences-Research and Ethics Committee, College of Health Sciences, Makerere University (approval number: SBS-HDREC-650), Mulago National Referral Hospital-Research and Ethics Committee (approval number: MHREC-1702), and Uganda National Council for Science and Technology (approval number: HS411ES). It was also approved by the Joint CUHAS/BMC Research and Ethics Committee (CREC/409/2019) and the National Health Research Ethics Review Committee of the National Institute for Medical Research (NIMR/HQ/R.8a/Vol.IX/3322) in Tanzania. Informed consent was obtained from all study participants. All information obtained was coded and kept confidential. All information obtained was coded and kept confidential.

Consent for publication

Not Applicable.

Competing interests

The authors declare no competing interests.

Author details

[1]Department of Biochemistry and Molecular Biology, Weill Bugando School of Medicine, Catholic University of Health and Allied Sciences, P. O. Box 1464, Mwanza, Tanzania
[2]Department of Immunology and Molecular Biology, College of Health Sciences, School of Biomedical Sciences, Makerere University, P. O Box 7072, Kampala, Uganda
[3]Department of Biochemistry and Bioinformatics, School of Pure and Applied Sciences, Pwani University, P.O Box 195-80108, Kilifi, Kenya
[4]African Center of Excellence in Bioinformatics and Data-Intensive Sciences, the Infectious Diseases Institute, College of Health Sciences, Makerere University, P.O Box 22418, Kampala, Uganda
[5]Department of Surgery, Weill Bugando School of Medicine, Catholic University of Health and Allied Sciences, P. O. Box 1464, Mwanza, Tanzania

[6]Department of Microbiology and Immunology, Weill Bugando School of Medicine, Catholic University of Health and Allied Sciences, P. O. Box 1464, Mwanza, Tanzania
[7]Africa Centres for Disease Control and Prevention, African Union Commission, Roosevelt Street, P.O. Box 3243, Addis Ababa W21 K19, Ethiopia

Received: 19 December 2023 / Accepted: 21 October 2025
Published online: 17 November 2025

References

1. Chang D, Sharma L, Dela Cruz CS, et al. Clinical epidemiology, risk factors, and control strategies of *Klebsiella pneumoniae* infection. Front Microbiol. 2021. https://doi.org/10.3389/fmicb.2021.750662.
2. Kochan TJ, Nozick SH, Medernach RL, et al. Genomic surveillance for multidrug-resistant or hypervirulent *Klebsiella pneumoniae* among united states bloodstream isolates. BMC Infect Dis. 2022;22:603.
3. Russo TA, Marr CM. Hypervirulent *Klebsiella pneumoniae*. Clin Microbiol Rev. 2019;32:e00001–19.
4. Effah CY, Sun T, Liu S, et al. *Klebsiella pneumoniae*: an increasing threat to public health. Ann Clin Microbiol Antimicrob. 2020;19:1.
5. Dong N, Yang X, Chan EW-C, et al. Klebsiella species: Taxonomy, hypervirulence and multidrug resistance. EBioMedicine. 2022;79:103998.
6. Ait Hssain A, Vahedian-Azimi A, Ibrahim AS, et al. Incidence, risk factors and outcomes of nosocomial infection in adult patients supported by extracorporeal membrane oxygenation: a systematic review and meta-analysis. Crit Care. 2024;28:158.
7. Aiesh BM, Qashou R, Shemmessian G, et al. Nosocomial infections in the surgical intensive care unit: an observational retrospective study from a large tertiary hospital in Palestine. BMC Infect Dis. 2023;23:686.
8. Abbasi SH, Aftab RA, Chua SS. Risk factors associated with nosocomial infections among end stage renal disease patients undergoing hemodialysis: a systematic review. PLoS ONE. 2020;15:e0234376.
9. Marando R, Seni J, Mirambo MM, et al. Predictors of the extended-spectrum-beta lactamases producing Enterobacteriaceae neonatal sepsis at a tertiary hospital, Tanzania. Int J Med Microbiol. 2018;308:803–11.
10. Tacconelli E, Carrara E, Savoldi A, et al. Discovery, research, and development of new antibiotics: the WHO priority list of antibiotic-resistant bacteria and tuberculosis. Lancet Infect Dis. 2018;18:318–27.
11. Shankar C, Jacob JJ, Vasudevan K, et al. Emergence of multidrug resistant hypervirulent ST23 *Klebsiella pneumoniae*: multidrug resistant plasmid acquisition drives evolution. Front Cell Infect Microbiol. 2020;10:575289.
12. Byarugaba DK, Erima B, Wokorach G, et al. Genome analysis of *Klebsiella pneumoniae* reveals international high-risk pandemic MDR clones emerging in tertiary healthcare settings in Uganda. Pathogens. 2023;12:1334.
13. Rodríguez-Medina N, Barrios-Camacho H, Duran-Bedolla J, et al. *Klebsiella variicola*: an emerging pathogen in humans. Emerg Microbes Infect. 2019;8:973–88.
14. Lam MMC, Wick RR, Watts SC, et al. A genomic surveillance framework and genotyping tool for *Klebsiella pneumoniae* and its related species complex. Nat Commun. 2021;12:4188.
15. Morgado S, Fonseca E, Vicente AC. Genomics of *Klebsiella pneumoniae* species complex reveals the circulation of high-risk multidrug-resistant pandemic clones in human, animal, and environmental sources. Microorganisms. 2022;10:2281.
16. Choby JE, Howard-Anderson J, Weiss DS. Hypervirulent *Klebsiella pneumoniae* - clinical and molecular perspectives. J Intern Med. 2020;287:283–300.
17. Mboowa G, Sserwadda I, Bulafu D, et al. Transmission dynamics of antimicrobial resistance at a national referral hospital in Uganda. Am J Trop Med Hyg. 2021;105:498–506.
18. Seni J, Akaro IL, Mkinze B, et al. Gastrointestinal tract colonization rate of Extended-Spectrum Beta-Lactamase-Producing Gram-Negative bacteria and associated factors among orthopaedic patients in a tertiary hospital in tanzania: implications for infection prevention. Infect Drug Resist. 2021;14:1733–45.
19. Sserwadda I, Mboowa G. rMAP: the rapid microbial analysis pipeline for ESKAPE bacterial group whole-genome sequence data. Microb Genom. 2021;7:000583.
20. Babraham Bioinformatics -. FastQC A Quality Control tool for High Throughput Sequence Data, https://www.bioinformatics.babraham.ac.uk/projects/fastqc/. (accessed 29 September 2024).
21. Ewels P, Magnusson M, Lundin S, et al. Multiqc: summarize analysis results for multiple tools and samples in a single report. Bioinformatics. 2016;32:3047–8.
22. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for illumina sequence data. Bioinformatics. 2014;30:2114–20.
23. Seemann T. tseemann/shovill. 2024. https://github.com/tseemann/shovill. (Accessed 29 September 2024).
24. Walker BJ, Abeel T, Shea T, et al. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS ONE. 2014;9:e112963.
25. Lam MMC, Wick RR, Judd LM, et al. Kaptive 2.0: updated capsule and lipopolysaccharide locus typing for the *Klebsiella pneumoniae* species complex. Microb Genom. 2022;8:000800.
26. Carattoli A, Zankari E, García-Fernández A. In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. Antimicrob Agents Chemother. 2014;58:3895–903.
27. Seemann T. tseemann/abricate. 2024. https://github.com/tseemann/abricate. (Accessed 29 September 2024).
28. Hyatt D, Chen G-L, LoCascio PF, et al. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics. 2010;11:119.
29. Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinformatics. 2014;30:2068–9.
30. Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics. 2009;25:1754–60.
31. Seemann T. tseemann/samclip. 2024. https://github.com/tseemann/samclip. (Accessed 29 September 2024).
32. Li H, Handsaker B, Wysoker A, et al. The sequence alignment/map format and samtools. Bioinformatics. 2009;25:2078–9.
33. Garrison E, Marth G. Haplotype-based variant detection from short-read sequencing. Epub ahead of print 20 July 2012. https://doi.org/10.48550/arXiv.1207.3907.
34. samtools/bcftools. 2024. https://github.com/samtools/bcftools. (Accessed 29 September 2024).
35. Cingolani P, Platts A, Wang LL, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. Fly. 2012;6:80–92.
36. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013;30:772–80.
37. Minh BQ, Schmidt HA, Chernomor O, et al. IQ-tree 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol Biol Evol. 2020;37:1530–4.
38. Ortiz EM. edgardomortiz/vcf2phylip. 2024. https://github.com/edgardomortiz/vcf2phylip. (Accessed 29 September 2024).
39. Criscuolo A, Gribaldo S. BMGE (block mapping and gathering with entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. BMC Evol Biol. 2010;10:210.
40. Letunic I, Bork P. Interactive tree of life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res. 2021;49:W293–6.
41. Ondov BD, Treangen TJ, Melsted P, et al. Mash: fast genome and metagenome distance estimation using minhash. Genome Biol. 2016;17:132.
42. 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.
43. Hu Y, Anes J, Devineau S, et al. *Klebsiella pneumoniae*: prevalence, reservoirs, antimicrobial resistance, pathogenicity, and infection: a hitherto unrecognized zoonotic bacterium. Foodborne Pathog Dis. 2021;18:63–84.
44. Baek E-H, Kim S-E, Kim S, et al. Successful control of an extended-spectrum beta-lactamase-producing *Klebsiella pneumoniae* ST307 outbreak in a neonatal intensive care unit. BMC Infect Dis. 2020;20:166.
45. Kim JO, Song SA, Yoon E-J, et al. Outbreak of KPC-2-producing Enterobacteriaceae caused by clonal dissemination of *Klebsiella pneumoniae* ST307 carrying an IncX3-type plasmid harboring a truncated Tn4401a. Diagn Microbiol Infect Dis. 2017;87:343–8.
46. Schaufler K, Nowak K, Düx A, et al. Clinically relevant ESBL-producing *K. pneumoniae* ST307 and *E. coli* ST38 in an urban West African rat population. Front Microbiol. 2018. https://doi.org/10.3389/fmicb.2018.00150.
47. Zhu J, Wang T, Chen L, et al. Virulence factors in hypervirulent *Klebsiella pneumoniae*. Front Microbiol. 2021. https://doi.org/10.3389/fmicb.2021.642484.
48. Catalán-Nájera JC, Garza-Ramos U, Barrios-Camacho H. Hypervirulence and hypermucoviscosity: two different but complementary *Klebsiella* spp. phenotypes? Virulence. 2017;8:1111–23.
49. Gu D, Dong N, Zheng Z, et al. A fatal outbreak of ST11 carbapenem-resistant hypervirulent *Klebsiella pneumoniae* in a Chinese hospital: a molecular epidemiological study. Lancet Infect Dis. 2018;18:37–46.

50. Wyres KL, Nguyen TNT, Lam MMC, et al. Genomic surveillance for hypervirulence and multi-drug resistance in invasive *Klebsiella pneumoniae* from South and Southeast Asia. Genome Med. 2020;12:11.
51. Tang M, Kong X, Hao J, et al. Epidemiological characteristics and formation mechanisms of multidrug-resistant hypervirulent *Klebsiella pneumoniae*. Front Microbiol. 2020. https://doi.org/10.3389/fmicb.2020.581543.
52. Shankar C, Nabarro LEB, Muthuirulandi Sethuvel DP, et al. Draft genome of a hypervirulent *Klebsiella quasipneumoniae* subsp. *similipneumoniae* with novel sequence type ST2320 isolated from a chronic liver disease patient. J Glob Antimicrob Resist. 2017;9:30–1.
53. Yang X, Wai-Chi Chan E, Zhang R, et al. A conjugative plasmid that augments virulence in *Klebsiella pneumoniae*. Nat Microbiol. 2019;4:2039–43.
54. Musicha P, Msefula CL, Mather AE, et al. Genomic analysis of *Klebsiella pneumoniae* isolates from Malawi reveals acquisition of multiple ESBL determinants across diverse lineages. J Antimicrob Chemother. 2019;74:1223–32.
55. Riwu KHP, Effendi MH, Rantam FA, et al. A review: virulence factors of Klebsiella pneumonia as emerging infection on the food chain. Vet World. 2022;15:2172–9.
56. Cassat JE, Skaar EP. Iron in infection and immunity. Cell Host Microbe. 2013;13:509–19.
57. Behnsen J, Zhi H, Aron AT, et al. Siderophore-mediated zinc acquisition enhances enterobacterial colonization of the inflamed gut. Nat Commun. 2021;12:7016.
58. Paczosa MK, Mecsas J. *Klebsiella pneumoniae*: going on the offense with a strong defense. Microbiol Mol Biol Rev. 2016;80:629–61.
59. Yang Q-E, Sun J, Li L, et al. IncF plasmid diversity in multi-drug resistant *Escherichia coli* strains from animals in China. Front Microbiol. 2015;6:964.
60. Pitout JDD, Chen L. The significance of epidemic plasmids in the success of Multidrug-Resistant drug pandemic extraintestinal pathogenic *Escherichia coli*. Infect Dis Ther. 2023;12:1029–41.

Publisher's Note