Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Population genomics, resistome and plasmidome of invasive Klebsiella pneumoniae isolates in adults hospitalised in Johannesburg, South Africa

Denasha L. Reddy
denasha.reddy@wits-vida.org
University of the Witwatersrand
Courtney P. Olwagen
University of the Witwatersrand
Shama Khan
University of the Witwatersrand
Nicholas J. Dean
University of the Witwatersrand
Lara van der Merwe
University of the Witwatersrand
Frank O. Williams
Imperial College London
Jeannette Wadula
National Health Laboratory Service
Trusha Nana
National Health Laboratory Service
Rispah Chomba
National Health Laboratory Service
Sinenhlanhla Ndzabandzaba
National Health Laboratory Service
Lyle Murray
University of the Witwatersrand
Jacob Merika Tsitsi
University of the Witwatersrand
Jeremy Nel
University of the Witwatersrand
Arifa Parker
Stellenbosch University
Vicky Baillie

University of the Witwatersrand
Elita Jauneikaite
Imperial College London
Ziyaad Dangor
University of the Witwatersrand
Shabir A. Madhi
University of the Witwatersrand

Article

Keywords: Klebsiella pneumoniae, molecular epidemiology, antimicrobial resistance, virulence, plasmids

Posted Date: March 13th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-8967609/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Population genomics, resistome and plasmidome of invasive *Klebsiella pneumoniae* isolates in adults hospitalised in Johannesburg, South Africa

Denasha L. Reddy[1,2]*, Courtney P. Olwagen[1], Shama Khan[1], Nicholas J. Dean[1], Lara van der Merwe[1], Frank O. Williams[3], Jeannette Wadula[4,5], Trusha Nana[4,5], Rispah Chomba[4,5], Sinenhlanhla Ndzabandzaba[4,5], Lyle Murray[2,6], Jacob Merika Tsitsi[2,7], Jeremy Nel[2,8], Arifa Parker[9], Vicky Baillie[1], Elita Jauneikaite[3], Ziyaad Dangor[1], Shabir A. Madhi[1,10].

1. South African Medical Research Council Vaccines and Infectious Diseases Analytics Research Unit, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa
2. Division of Infectious Diseases, Department of Internal Medicine, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa
3. Institute of Reproductive and Developmental Biology, Department of Metabolism, Digestion and Reproduction, Imperial College London, London, United Kingdom
4. National Health Laboratory Service, Johannesburg, South Africa
5. Clinical Microbiology and Infectious Diseases, School of Pathology, University of the Witwatersrand, Johannesburg, South Africa
6. Division of Infectious Diseases, Department of Internal Medicine, Charlotte Maxeke Johannesburg Academic Hospital, Johannesburg, South Africa
7. Division of Infectious Diseases, Department of Internal Medicine, Chris Hani Baragwanath Academic Hospital, Johannesburg, South Africa
8. Division of Infectious Diseases, Department of Internal Medicine, Helen Joseph Hospital, Johannesburg, South Africa
9. Unit for Infection Prevention and Control and Department of Medicine, Faculty of Health Sciences, Stellenbosch University, Cape Town, South Africa
10. Wits Infectious Diseases and Oncology Research Institute, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

Target journal: *npj Antimicrobials and Resistance*
Article type: Article
Corresponding author: Denasha L. Reddy
Corresponding author's institution: University of the Witwatersrand
Corresponding author's email address: denasha.reddy@wits-vida.org
Corresponding author's address: Wits-VIDA, Chris Hani Baragwanath Academic Hospital, Chris Hani Road, Diepkloof, Soweto, South Africa, 1864
First author: Denasha L. Reddy
Manuscript Region of Origin: South Africa
Abstract: 148
Word count body: 5750
Number of Tables: 3
Number of Figures: 6
Number of References: 56
Keywords: *Klebsiella pneumoniae*; molecular epidemiology; antimicrobial resistance; virulence; plasmids.

Abstract

Klebsiella pneumoniae (KPn) is a leading cause of healthcare-associated infections in adults in low- and middle-income countries. We conducted laboratory-based surveillance between 2023 and 2024 in Johannesburg, South Africa, and analysed 508 KPn whole genomes isolated from blood and cerebrospinal fluid cultures of hospitalised adults. We investigated the key antimicrobial resistance (AMR) profiles in invasive KPn isolates, and plasmid-mediated AMR. Overall, 79.1% of the KPn isolates harboured genes conferring a multidrug-resistant (MDR) phenotype. The prevalence of genotypic AMR compared to phenotypic AMR was higher for aminoglycosides and carbapenems, similar for third generation cephalosporins, and lower for colistin, respectively. Three plasmid types, IncX3, IncFII and IncFIB, were most frequently associated with harbouring beta-lactamase genes bla_{CTX}, bla_{TEM}, bla_{NDM}, and bla_{OXA}. On multiple regression modelling, ST2497 was associated with increased risk of death. Overall, our findings contribute to the genomic characterisation of drug-resistant KPn invasive disease in the WHO African Region.

Klebsiella pneumoniae (KPn) has emerged as a leading cause of community-associated infections (CAIs) and healthcare-associated infections (HAIs) in adults,[1, 2] with the highest burden of disease in low- and middle-income countries (LMICs).[3] According to a predictive statistical modelling study on the global burden of bacterial antimicrobial resistance (AMR) in 2019, KPn invasive disease (KPn-ID) accounted for 50,000 deaths attributable to AMR in Africa.[3] Due to the rising global public health crisis of bacterial AMR, third-generation-cephalosporin (3GC)- and carbapenem-resistant strains of KPn have been included in the critical group of the World Health Organisation (WHO) Bacterial Priority Pathogens List since 2017,[4] underscoring the need for novel therapeutic agents and preventive strategies. Notably, carbapenem-resistant KPn (CRKp) rose from fifth in 2017 to become the top WHO priority pathogen in 2024, due to limited treatment options, high associated mortality, and rapidly increasing AMR rates across the majority of WHO regions.[4] In a recent meta-analysis, the estimated global prevalence of hospital-acquired CRKp was 28.7% (95% CI, 26.5%–30.9%), with rates exceeding 50.0% in some regions.[5] In addition to AMR, convergence of AMR with hypervirulence genes has been reported in KPn, and poses a new diagnostic and therapeutic challenge.[6]

Genomic-based surveillance of KPn, with a particular focus on accurately identifying CRKp and hypervirulent (hvKp) strains, integrated with clinical data, is needed to accurately determine the burden of KPn disease and guide targeted public health interventions aimed at reducing KPn-associated morbidity and mortality in vulnerable populations.[2, 7] Key interventions include the implementation of robust infection prevention and control (IPC) measures, development of novel antimicrobial therapies, and advancement of preventive strategies such as vaccination.[8-10] Despite the growing global threat posed by CRKp and hvKp, molecular surveillance in the WHO African region remains limited, resulting in underestimation of the disease burden and delays in timely, evidence-based responses to emerging threats.[6, 7] We previously reported on the clinical and microbiological epidemiology of KPn-ID in African adults,[11] and now we add to that the detailed genomic epidemiology of KPn infections in hospitalised adults in Johannesburg, South Africa to address the knowledge gap of key AMR and virulence profiles in invasive KPn isolates.

RESULTS

Molecular epidemiology

Of the 617 cases of KPn-ID identified during the study period, 539 (80.3%) invasive isolates underwent whole genome sequencing (WGS) (Figure 1). Of the 539 isolates that underwent WGS, four (0.7%; 4/539) were excluded due to contamination, 94.2% (508/539) were identified as *Klebsiella pneumoniae* (KPn), with the remaining isolates comprising *Klebsiella variicola* (4.5%; 24/539), *Klebsiella quasipneumoniae* (0.4%; 2/539), and *Raoultella planticola* (0.2%; 1/539) (Figure 1). Further analysis was limited to the 508 KPn genomes.

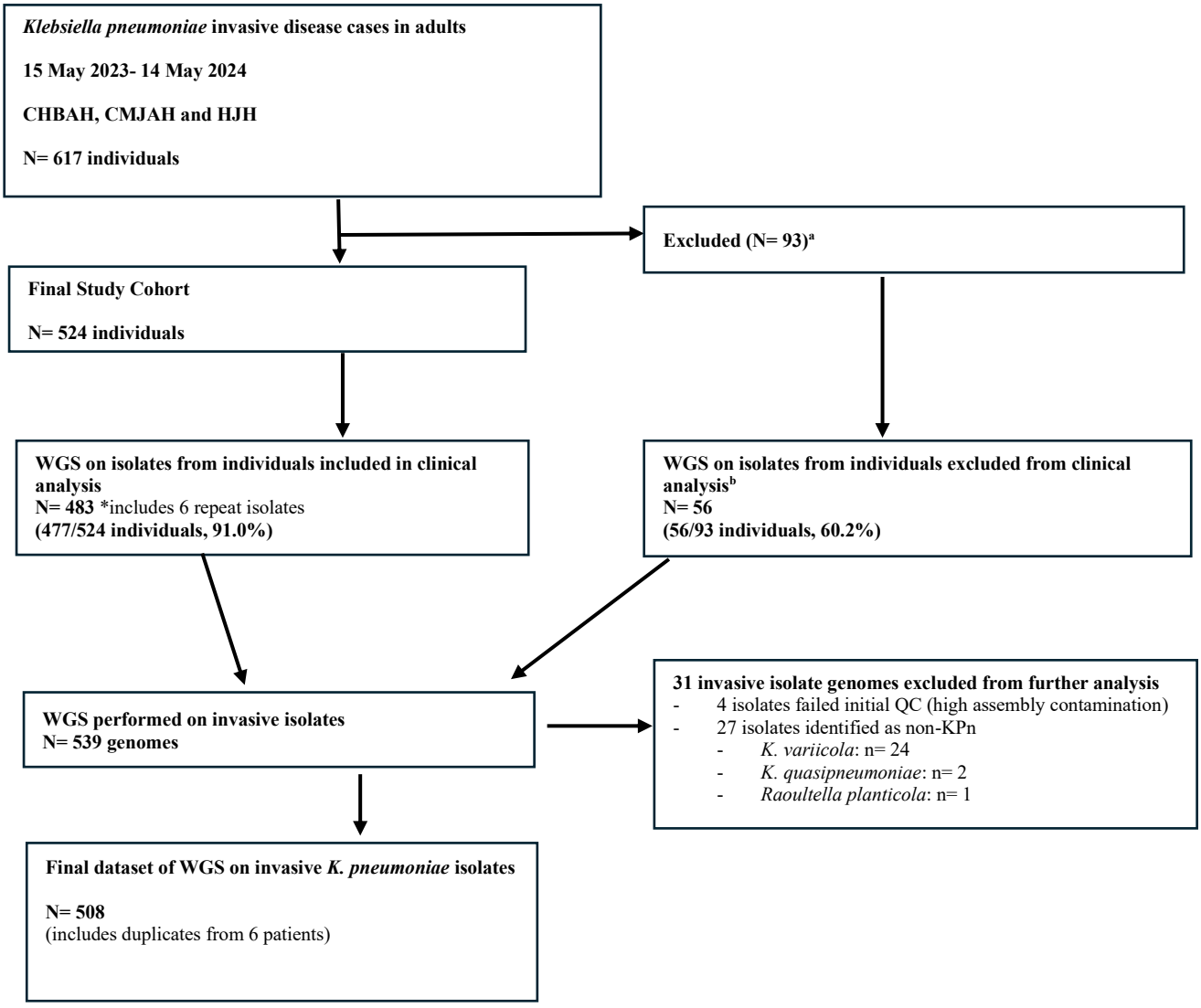


Figure 1. Consort diagram

Abbreviations: CHBAH, Chris Hani Baragwanath Academic Hospital; CMJAH, Charlotte Maxeke Johannesburg Academic Hospital; HJH, Helen Joseph Hospital; WGS, whole genome sequencing.

[a] 93 cases were excluded from clinical analysis for the following reasons: records could not be traced (n=40), declined participation in study (n=33), discharged prior to enrolment before ethical approval granted for retrospective enrolment (n=11), and missed on initial laboratory surveillance (n=9).

[b] An ethical waiver was granted for the collection of isolates, but not clinical information, from excluded individuals with KPn-ID.

Six individuals had two distinct episodes of KPn-ID each. All isolates were included in the genotypic epidemiological analysis as the second infection episodes were considered new infection events. The number of days between the first and second infection episode culture draws ranged from 27 to 169 days (median 86 days, IQR: 54- 97 days). Across the 12 genomes analysed, overall genomic distances were high (≈23,000–26,000 single nucleotide polymorphisms [SNPs]), indicating substantial genetic diversity between unrelated isolates. However, the isolates obtained from the same individual at first (I) and second (II) visits showed a range of SNP distances. KPIDA0060-I (ST2497)/ B-II (ST2497) and KPIDA0149A-I (ST307)/ B-II (ST307) differed by 6 and 8 SNPs, respectively, consistent with persistence of the same strain with limited within-host microevolution. In contrast, KPIDA0127A-I (ST9094)/ B-II (ST5785) and KPIDA0268A-I (ST985)/ B-II (ST5785) differed by 33 SNPs, indicating continuity within the same lineage with greater genetic divergence over time, while KPIDC0030A-I (ST86)/ B-II (ST86) differed by 72 SNPs, making persistence of the same strain less likely and indicating a higher inter-strain diversity within this sequence type (ST). No pairwise comparison between isolates from different individuals fell below this threshold, indicating no evidence of inter-patient transmission within the cohort. Overall, the Split Kmer Analysis (SKA) (Supplementary file S1) demonstrated that isolates from the same individual represented the same infecting strain in two of the six individuals, whereas isolates recovered from different individuals were genomically distinct (Supplementary Figure 1). The plasmid carriage in the two individuals with persistent KPn strains were similar in both infection episodes (Supplementary Figure 1).

A detailed description of the demographics of the 524 individuals enrolled in the clinical study is available.[11] Briefly, 55.2% (n=289) of KPn-ID cases were male and 85.5% (n=448) were black-African.[11] The median age of participants was 48 (IQR: 35- 61) years.[11] Eighty-four percent (442/524) of the KPn-ID cases were categorised as presumed HAIs (pHAIs), and the overall in-hospital case fatality risk (CFR) was 56.5% (296/524).[11]

Sequence types (ST), Capsular polysaccharide (K-) Loci, and Lipopolysaccharide (O-) antigens

There was a large diversity of KPn causing invasive disease in adults, with 111 different STs identified, including fifteen novel STs indicated in bold in Supplementary Table 1. Overall, ST2497 (18.1%; 92/508) predominated followed by ST39 (12.0%; 61/508), the novel ST9094 (9.4%; 48/508), ST17 (7.1%; 36/508), and ST307 (3.5%; 18/508), with the remaining STs each having a prevalence of 0.2%- 2.9% (n= 1-15).

Differences were observed in prevalence of ST2497 across the different study sites, with a prevalence of 16.0% (46/287) at Chris Hani Baragwanath Academic Hospital (CHBAH), 23.6% (45/191) at Charlotte Maxeke Johannesburg Academic Hospital (CMJAH) where it predominated, and 3.3% (1/30) at Helen Joseph Hospital (HJH) (Supplementary Table 1 and Figure 2). The dominant STs at CHBAH and HJH were ST39 (16.7%; 48/287) and ST9094 (16.7%; 5/30), respectively. The most prevalent STs overall were more commonly associated with pHAI compared with presumed community-associated infections (pCAI): ST2497 (91.3%; 84/92 vs 1.1%; 1/92), ST39 (78.7%; 48/61 vs 3.3%; 2/61), ST9094 (87.5%; 42/48 vs 6.3%; 3/48), ST17 (83.3%; 30/36 vs 2.8%; 1/36), and ST307 (83.3%; 15/18 vs 5.6%; 1/18) (Supplementary Table 1).

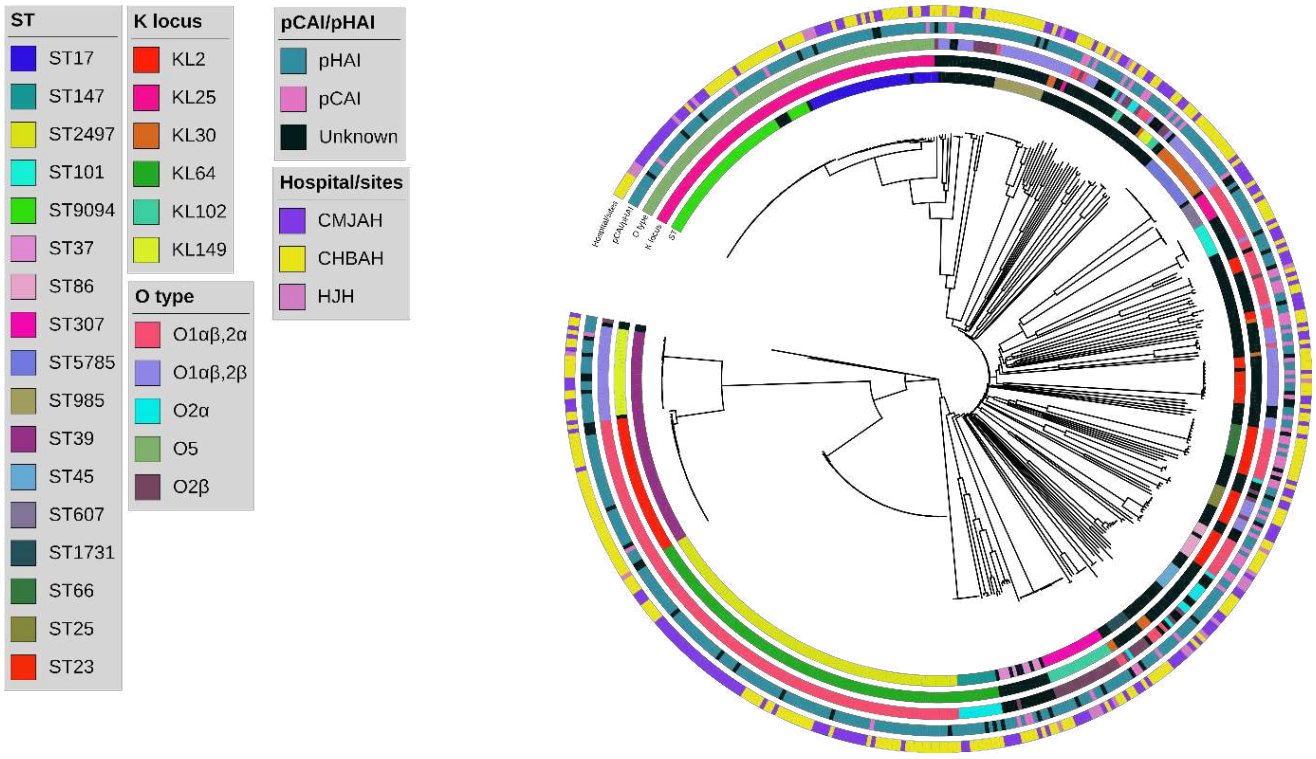
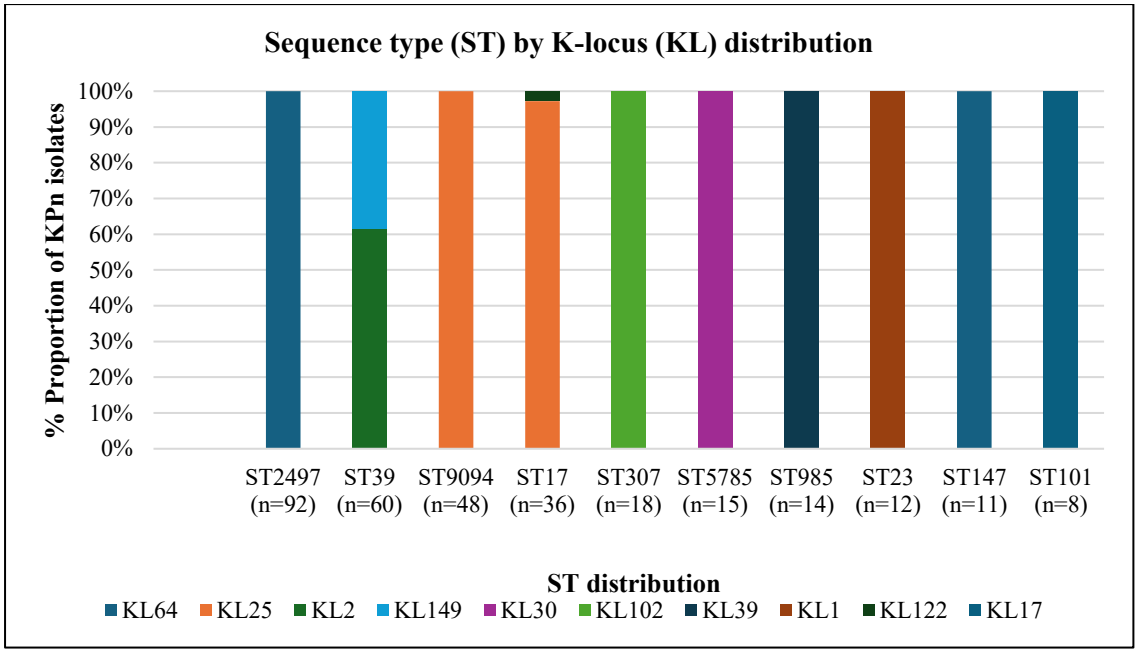


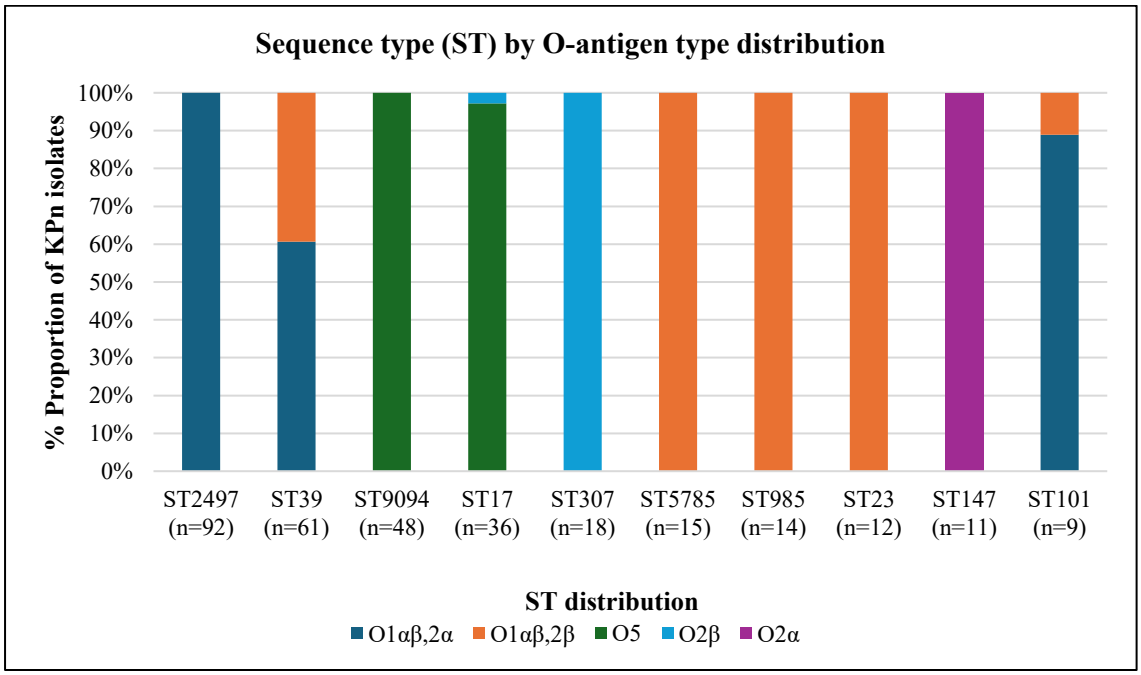
Figure 2. Unrooted phylogenetic tree of 508 *K. pneumoniae* invasive isolates in hospitalised adults in Johannesburg, South Africa, illustrating the relatedness of isolates. The metadata rings from inside out indicate (i.) Sequence types (ST) (ii.) K-locus (iii.) O-antigen type (iv.) Acquisition category- Community-associated infections (CAI) vs Hospital-associated infections (HAI) , and (v.) Hospital site of isolation, with colours corresponding to the legend.

There were significant associations between the top 10 STs, top 10 K-loci and top 5 O-antigen types (Figures 3a and 3b), with the following clonal groups (CG) predominating (Figure 3c): CG10190 (18.2%; 92/508), CG17 (ST9094) (9.4%; 48/508), CG39 (7.3%; 37/508), and CG17 (ST17) (6.9%; 35/508).

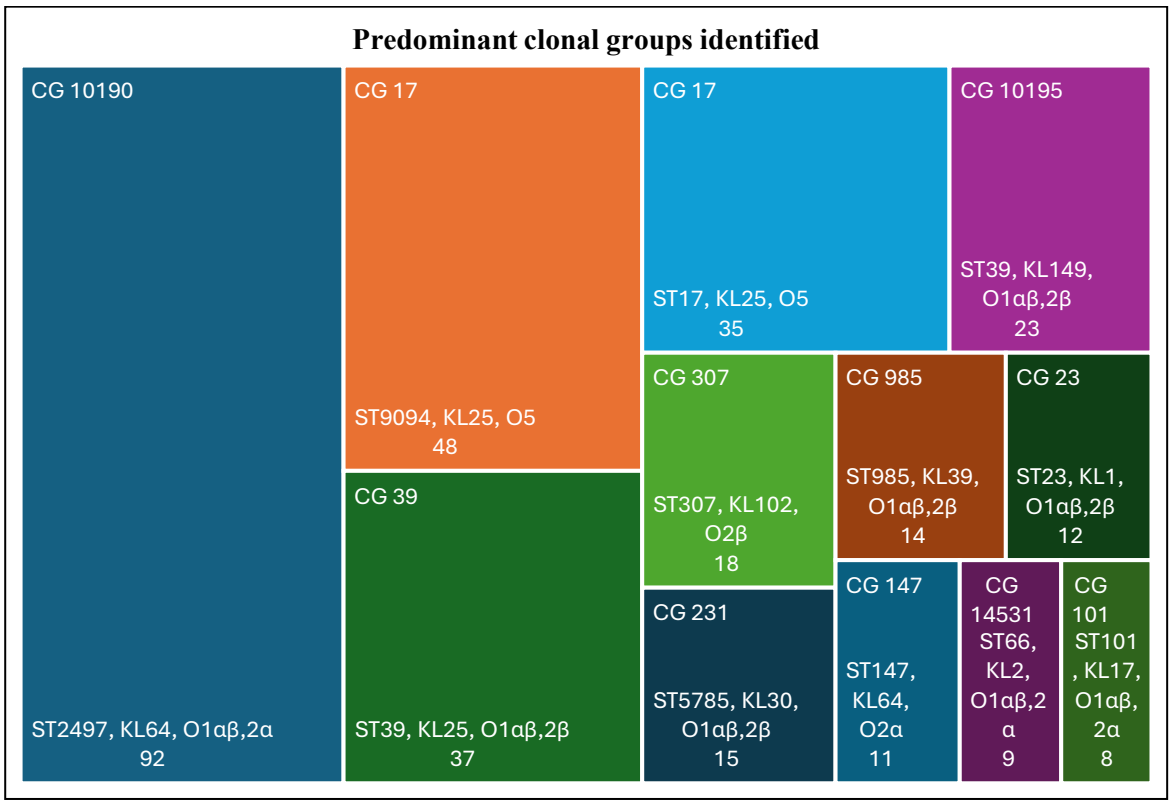
3a



3b



3c



Figures 3a-3c. 3a. Sequence type (ST) by K-locus (KL) distribution of invasive *Klebsiella pneumoniae* isolates (Top 10 STs by Top 10 KLs). Figure 3b. Sequence type (ST) by O-type distribution of invasive *Klebsiella pneumoniae* isolates (Top 10 STs by Top 5 O-types). Figure 3c. Predominant clonal groups identified

Overall, 56 different K-loci were identified (Supplementary Table 1) with KL64 (20.3%; 103/508) dominating, followed by KL25 (19.1%; 97/508), KL2 (14.8%; 75/508), and KL149 (5.1%; 26/508). The remaining 52 K-loci each had a prevalence of 0.2%- 4.9% (n= 1-25). There was heterogeneity in the predominant K-locus by site (Supplementary Table 1). At CHBAH, the predominant K-loci were KL2 (19.9%; 57/287), and KL64 (18.8%; 54/287), while KL64 (25.1%; 48/191) dominated at CMJAH, and KL25 (26.7%; 8/30) at HJH. The most prevalent K-loci were more often associated with pHAI than pCAI cases: KL64 (89.3%; 92/103 vs 1.0%; 1/103), KL25 (85.6%; 83/97 vs 4.1%; 4/97), KL2 (60.0%; 45/75 vs 25.3%; 19/75), and KL149 (73.1%; 19/26 vs 7.7%; 2/26).

Thirteen different O-antigen types were identified (Supplementary Table 1). Overall, the most common O-antigen types were O1αβ,2α (39.6%; 201/508), O1αβ,2β (20.3%; 103/508), O5 (17.5%; 89/508), and O2β (9.1%; 46/508). The remaining nine O-antigen types each had a prevalence of 0.2% to 4.7% (n= 1-24). There were no significant associations between specific O-antigen types and study sites (Supplementary Table 1).

Genomic diversity and phylogenetic analysis

The phylogenetic tree illustrates the substantial diversity of the 508 KPn isolates (Figure 2). The most predominant STs (i.e. ST2497, ST39 and, ST9094) clustered into dense clades suggestive of recent clonal spread. A third of the genomes (27.6%; 140/508) belonged to less frequent STs (Figure 2, indicated in black), signifying ongoing endemic genotypic diversity of disease-causing KPn. Mapping of pCAI versus pHAI showed that pHAI isolates were more likely to cluster together, which is suggestive of genetically related local clones circulating within hospitals. In contrast, pCAI isolates were scattered across the phylogenetic tree (Figure 2) and displayed greater diversity indicating that the diversity of KPn clones in the community is underexplored. Overall, this core-genome phylogenetic analysis revealed extensive genomic heterogeneity, reflecting a complex population structure shaped by both clonal expansion and sporadic introductions of KPn clones.

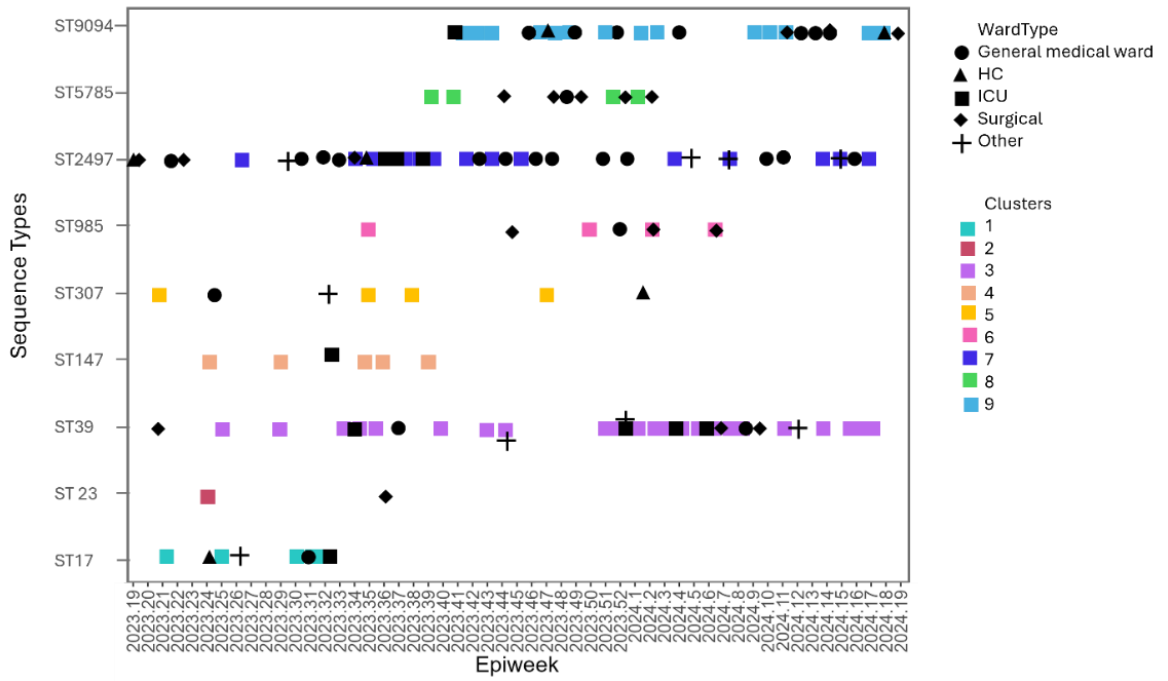
Transmission dynamics

Across the three hospitals- CHBAH, CMJAH and HJH, KPn transmission dynamics varied considerably within each hospital (Figures 4a, 4b and 4c). Analysis revealed multiple KPn isolates that differed by ≤20 SNPs within a 45-day window, suggesting potential transmission events. In CHBAH (Figure 4a), seven large, sustained, and multi-ward clusters were detected, of which three, Cluster 3 (ST39), Cluster 7 (ST2497), and Cluster 9 (ST9094), were the major clusters. These clusters persisted across several epi-weeks, and spread between ICU, surgical and medical wards, indicating both ward-specific and inter-ward transmission. Smaller clusters (Clusters 1, 2, 4, 5, and 6) were short-lived and restricted to specific wards. Many isolates fell outside cluster definitions, representing sporadic or genetically distinct introductions.

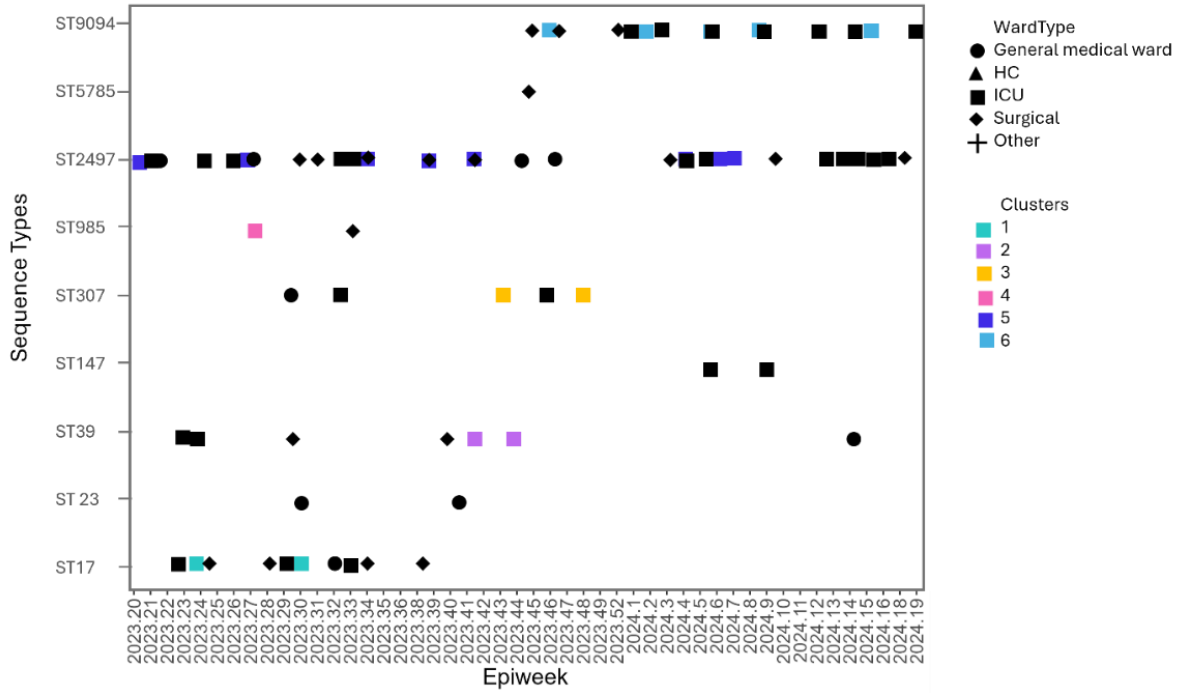
In CMJAH (Figure 4b), transmission models were less prevalent but still clinically significant. Clusters were generally smaller, temporally constrained and ward-limited with ST17, ST2497 and ST9094 being most prominently in ICU and surgical wards only. This implies frequent introductions of resistant clones into high-risk wards, followed by limited local spread rather than repeated hospital-wide transmission.

By contrast, HJH (Figure 4c) showed minimal evidence of onward transmission with only two small, short-lived clusters (ST9094 over two weeks and ST17 over 1 week), while most isolates appeared as genetically distinct singletons across different wards.

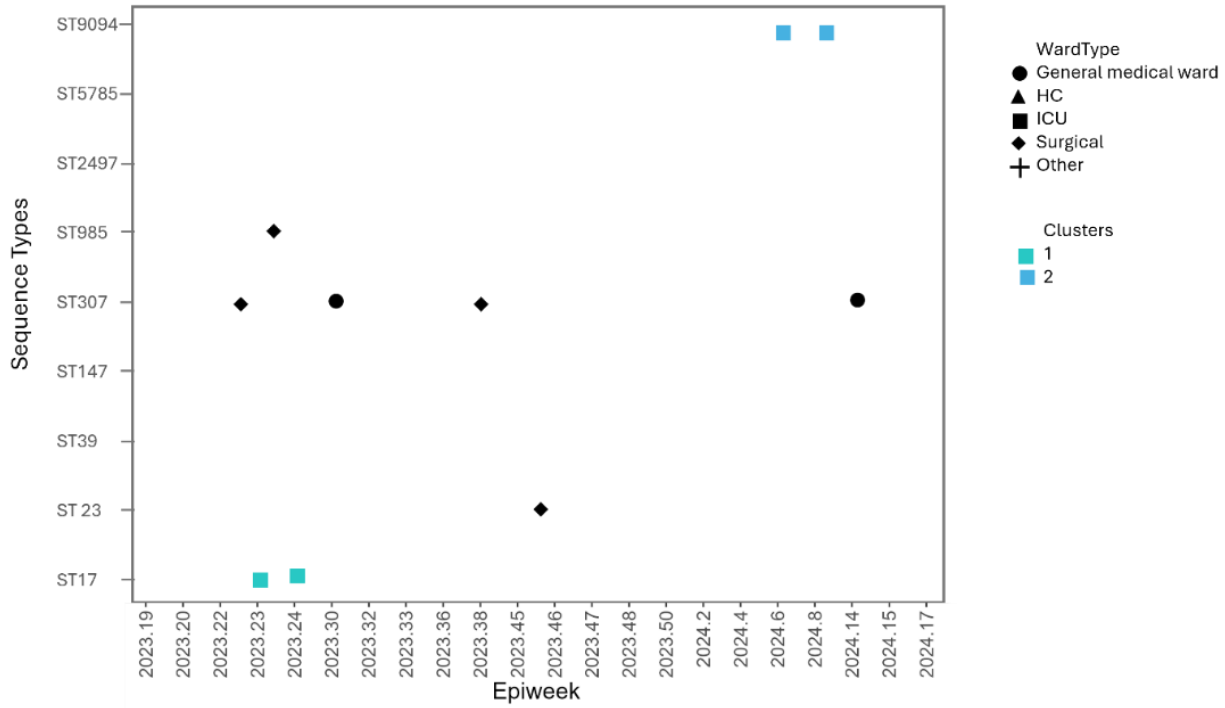
4a


ST9094
ST5785
ST2497
ST985
ST307
ST147
ST39
ST 23
ST17
Sequence Types
Epiweek
WardType
General medical ward
HC
ICU
Surgical
Other
Clusters
1
2
3
4
5
6
7
8
9

4b


ST9094
ST5785
ST2497
ST985
ST307
ST147
ST39
ST 23
ST17
Sequence Types
Epiweek
WardType
General medical ward
HC
ICU
Surgical
Other
Clusters
1
2
3
4
5
6

4c



Figures 4a-4c. Figure 4a. Transmission analysis for Chris Hani Baragwanath Academic Hospital (CHBAH). Figure 4b. Transmission analysis for Charlotte Maxeke Johannesburg Academic Hospital (CMJAH). Figure 4c. Transmission analysis for Helen Joseph Hospital (HJH).

Antimicrobial resistance

Overall, 79.1% (402/508) of the KPn isolates harboured genes conferring a multidrug-resistant (MDR) phenotype (defined as resistance to at least one agent in at least three different antimicrobial classes). The prevalence of AMR genes was: 76.6% (389/508) to aminoglycosides, 62.2% (316/508) to carbapenems, and 73.0% (371/508) to 3GC (Table 1).

Overall, the oxacillinase-type β-lactamase (*bla*$_{OXA}$) gene, including any of the allelic variants *bla*$_{OXA-181}$, *bla*$_{OXA-232}$, *bla*$_{OXA-417}$ and *bla*$_{OXA-48}$, was detected in 55.5% (282/508) of KPn isolates, with *bla*$_{OXA-181}$ predominating (37.6%; 191/508). The New Delhi metallo β-lactamase *(bla*$_{NDM}$), including any of the allelic variants *bla*$_{NDM-1}$, *bla*$_{NDM-5}$ and *bla*$_{NDM-7}$, was detected in 25.8% (131/508) of isolates. A combination of *bla*$_{NDM}$ and *bla*$_{OXA}$ (allelic variants *bla*$_{OXA-181}$ and *bla*$_{OXA-232}$) was detected in 111 isolates (21.9%) and a combination of *bla*$_{OXA-181}$ and *bla*$_{VIM}$ was detected in one isolate (0.2%). The *mcr* gene associated with colistin resistance was not detected in any of the isolates.

Table 1. Comparison of genotypic vs phenotypic antimicrobial resistance prevalence in sequenced invasive *Klebsiella pneumoniae* isolates

Phenotypic Resistance[a] (n/N with phenotypic AST) Prevalence % (95% CI)	Genotypic Resistance[b] (n/N with genotypic AST) Prevalence % (95% CI)	Concordant phenotypic and genotypic resistance, amongst isolates with phenotypic and genotypic AST results[c] % (n/N with phenotypic AST)	Concordant phenotypic and genotypic susceptibility, amongst isolates with phenotypic and genotypic AST results[c] % (n/N with phenotypic AST)	Discordant phenotypic and genotypic AST, amongst isolates with phenotypic and genotypic AST results[c] % (n/N with phenotypic AST)
Aminoglycosides Amikacin (212/506) 41.9% (37.6%- 46.3%)	**Aminoglycosides** (389/508) 76.6% (72.6%- 80.0%)	41.1% (208/506)	22.7% (115/506)	Geno AMR/ Pheno no AMR 35.4% (179/506) Geno no AMR/ Pheno AMR 0.8% (4/506)
Carbapenems Imipenem (244/505) 48.3% (43.9%- 52.7%)	**Carbapenems** (316/508) 62.2% (57.8%- 66.4%)	46.5% (235/505)	36.2% (183/505)	Geno AMR/ Pheno no AMR 15.4% (78/505) Geno no AMR/ Pheno AMR 1.8% (9/505)
Cephalosporins (3GC) Ceftriaxone (375/508) 73.8% (69.8%- 77.5%)	**Cephalosporins (3GC)** (371/508) 73.0% (68.9%- 76.8%)	70.9% (360/508)	24.0% (122/508)	Geno AMR/ Pheno no AMR 2.2% (11/508) Geno no AMR/ Pheno AMR 3.0% (15/508)
Polymyxin Colistin (22/256) 8.6% (5.5%- 12.7%)	**Colistin** (8/508) 1.6% (0.7%- 3.0%)	1.6% (4/256)	90.6% (232/256)	Geno AMR/ Pheno no AMR 0.8% (2/256) Geno no AMR/ Pheno AMR 7.0% (18/256)
Fluoroquinolones Ciprofloxacin (332/497) 66.8% (62.5%- 70.9%)	**Fluoroquinolones** (371/508) 73.0% (68.9%- 76.8%)	64.6% (321/497)	24.7% (123/497)	Geno AMR/ Pheno no AMR 8.5% (42/497) Geno no AMR/ Pheno AMR 2.2% (11/497)
Phenicols Chloramphenicol (137/268) 51.1% (45.0%- 57.0%)	**Phenicols** (233/508) 45.9% (41.5%- 50.0%)	44.0% (118/268)	46.3% (124/268)	Geno AMR/ Pheno no AMR 2.6% (7/268) Geno no AMR/ Pheno AMR 7.1% (19/268)
Sulfonamides Trimethoprim sulfamethoxazole (361/501) 72.1% (67.9%- 75.9%)	**Sulfonamides** (383/508) 75.4% (71.4%- 79.0%)	69.5% (348/501)	22.4% (112/501)	Geno AMR/ Pheno no AMR 5.6% (28/501) Geno no AMR/ Pheno AMR 2.6% (13/501)

Abbreviations: AST, antimicrobial susceptibility testing; AMR, antimicrobial resistance; CI, confidence interval; 3GC, third generation cephalosporin; Geno, genotypic; Pheno, phenotypic.

[a] Phenotypic resistance prevalence includes isolates of both complete and intermediate resistance as defined by the Clinical and Laboratory Standards Institute guidelines. For the phenotypic comparison, the antimicrobial selected for each class had the most complete phenotypic antimicrobial susceptibility profile available.

[b] Genotypic resistance profiles were based on Kleborate AMR profiles generated after upload to Pathogenwatch

[c] A comparison was made between phenotypic and genotypic resistance patterns only if phenotypic data was available for the class of antimicrobial.

Phenotypic antimicrobial susceptibility testing (AST) results are summarised in Supplementary Table 2. The top eight STs (ST2497, ST39, ST9094, ST17, ST5785, ST147, ST985 and ST307) were also most frequently associated with the beta-lactamase genes bla_{CTX}, bla_{OXA} and bla_{NDM} (Supplementary Table 3) and CRKp and MDR phenotypes (Supplementary Table 4). Similarly the top five K-loci (KL64, KL25, KL2, KL149 and KL30) and top five O-antigen types (O1αβ,2α, O1αβ,2β, O5, O2β and O2α) were most frequently associated with CRKp and MDR phenotypes (Supplementary Tables 5 and 6, respectively).

The prevalence of genotypic AMR in the invasive isolates varied by antimicrobial class (Table 1). Overall, there was a higher prevalence of genotypic AMR compared to phenotypic AMR for aminoglycosides (76.6%; 389/508 vs 41.9%; 212/506, respectively), and carbapenems (62.2%; 316/508 vs 48.3%; 244/505, respectively). In isolates that underwent phenotypic AST prior to genotypic testing, diagnostic concordance for AMR and susceptibility also varied according to antimicrobial class. Phenotypic AST and genotypic markers of AMR had a high concordance of detecting resistance in 3GC (70.9%; 360/508), where the prevalence of phenotypic resistance was 73.8% (375/508) and genotypic resistance was 73.0% (371/508), respectively. Of the 256 isolates that underwent colistin broth microdilution AST, 22 were found to be resistant (MIC ≥4 mg/L), and of those 22 isolates, only four were found to have genotypic colistin resistance determinants.

Virulence factors

There were no significant differences in prevalence of virulence factors at each study site (Supplementary Table 1). Yersiniabactin (*ybt*) was the most common virulence factor detected (70.7%; 359/508) and was most frequently associated with ST2497 (24.5%; 88/359), ST39 (17.0%; 61/359), and ST9094 (12.5%; 45/359). Genes encoding for colibactin (*clb*), aerobactin (*iuc*) and salmochelin (*iro*) were detected in 5.7% (n=29), 10.6% (n=54) and 10.6% (n=54) of isolates, respectively, and were frequently associated with ST23 and ST66 (Table 2). The hypermucoidy gene *RmpADC* was present in 9.8% (n=50) of KPn isolates, and most frequently associated with ST23 (24.0%; 12/50), ST66 (18.0%; 9/50), ST86 (18.0%; 9/50), KL2 (62.0%; 31/50) and KL1 (28.0%; 14/50). Although not clinically in keeping with hvKp, the plasmid-mediated biomarker *peg-344*, was detected on an IncFIB/ IncHI1B hybrid plasmid in one isolate (KPIDA0157) (Supplementary Figure 3). Yersiniabactin was most frequently associated with pHAI (80.8%; 290/359) compared with pCAI isolates (8.4%; 30/359). In contrast, *clb* (48.3%; 14/29 vs 31.0%; 9/29), *iuc* (44.4%; 24/54 vs 38.9%; 21/54), *iro* (42.6%;

23/54 vs 38.9%; 21/54) and *rmpADC* (50.0%; 25/50 vs 32.0%; 16/50) were more common in pCAI compared with pHAI isolates (Table 2).

Table 2. Summary of virulence factors associated with invasive *Klebsiella pneumoniae* isolates

N= 508 **n (%)**	**Yersiniabactin *ybt*** n= 359 (70.7)	**Colibactin *clb*** n= 29 (5.7)	**Aerobactin *iuc*** n= 54 (10.6)	**Salmochelin *iro*** n= 54 (10.6)	**RmpADC *rmp*** n= 50 (9.8)
Category of infection					
pCAI	30 (8.4)	14 (48.3)	24 (44.4)	23 (42.6)	25 (50.0)
pHAI	290 (80.8)	9 (31.0)	21 (38.9)	21 (38.9)	16 (32.0)
Unknown	39 (10.9)	6 (20.7)	9 (16.7)	10 (18.5)	9 (18.0)
Sequence Types (STs)[a]					
ST2497	88 (24.5)	0	6 (11.1)	6 (11.1)	0
ST39	61 (17.0)	0	0	0	0
ST23	9 (2.5)	11 (37.9)	12 (22.2)	12 (22.2)	12 (24.0)
ST9094	45 (12.5)	0	0	0	0
ST66	9 (2.5)	9 (31.0)	9 (16.6)	9 (16.6)	9 (18.0)
ST17	33 (9.2)	0	0	0	0
ST86	2 (0.6)	0	9 (16.6)	7 (13.0)	9 (18.0)
ST25	3 (0.8)	0	3 (5.6)	3 (5.6)	3 (6.0)
ST5785	11 (3.1)	0	0	0	1 (2.0)
ST985	12 (3.3)	0	0	0	0
ST147	11 (3.1)	0	0	0	0
ST65	1 (0.3)	0	3 (5.6)	3 (5.6)	3 (6.0)
ST380	2 (0.6)	1 (3.4)	2 (3.7)	2 (3.7)	2 (4.0)
K-loci (KL)[b]					
KL2	62 (17.3)	14 (48.3)	31 (57.4)	29 (53.7)	31 (62.0)
KL64	99 (27.6)	0	6 (11.1)	6 (11.1)	0
KL25	91 (25.3)	0	0	0	0
KL1	10 (2.8)	12 (41.4)	14 (25.9)	14 (25.9)	14 (28.0)
KL149	26 (7.2)	0	0	0	0
KL30	11 (3.1)	0	0	0	1 (2.0)
KL39	12 (3.3)	0	0	0	0
KL17	8 (2.2)	0	0	0	0
KL57	1 (0.3)	1 (3.4)	2 (3.7)	2 (3.7)	2 (4.0)
KL24	7 (1.9)	0	0	0	0
KL102	6 (1.7)	0	0	0	0
KL20	1 (0.3)	1 (3.4)	1 (1.9)	1 (1.9)	1 (2.0)
KL3	4 (1.1)	0	0	0	0
KL38	4 (1.1)	0	0	0	0
KL62	4 (1.1)	0	0	0	0

Abbreviations: pCAI, presumed community-associated infection; pHAI, presumed hospital-associated infection.

[a] Forty-seven different STs were associated with virulence factors. Only the top 13 are represented in the table. The remaining 34 STs include: ST101, ST45, ST607, ST9190, ST9467, ST268, ST382, ST13, ST14, ST307, ST37, ST1049, ST1063, ST152, ST9468, ST9482, ST105, ST35, ST48, ST9472, ST16, ST160, ST1621, ST1999, ST2621, ST29, ST353, ST36, ST405, ST458, ST711, ST716, ST9461, and ST9465.

[b] Twenty-eight different KLs were associated with virulence factors. Only the top 15 are represented in the table. The remaining 13 KLs include: KL5, KL14, KL54, KL110, KL127, KL15, KL151, KL16, KL19, KL22, KL46, KL52, and KL81.

Plasmidome

Of the 508 KPn genomes, 482 (94.9%) had plasmids that could be reconstructed, and 26 (5.1%) did not. Sixty-six different plasmid types were identified from a total of 2090 reconstructed plasmids identified in 482 invasive isolate genomes (range 1-11 plasmids per isolate). From the total number of 2090 reconstructed plasmids, 967 (46.3%) harboured genes encoding for AMR, primarily: $bla_{CTX\text{-}M\text{-}15}$, $bla_{NDM\text{-}1}$, bla_{OXA}, and bla_{TEM} conferring resistance to beta-lactams (67.1%; 649/967); *sul1* and *sul2* conferring resistance to sulfonamides (44.8%; 433/967); and *qnrB1* and *qnrS1* conferring resistance to quinolones (33.4%; 323/967).

Forty-three different plasmid types were associated with carriage of beta-lactamase genes, the most common being IncFIB/IncHI1B and IncX3 (Supplementary Table 7). An analysis of the AMR genomic cargo revealed 15 plasmids most frequently associated with carriage of genes encoding for beta-lactamases (bla_{CTX}, bla_{TEM}, bla_{NDM}, bla_{OXA}, and bla_{VIM}); Figure 5. While known to be present on chromosomes, bla_{SHV} was also present on the plasmids IncFIB (n=2), IncX3 (n=2) and IncFIB/IncFII (n=1). The dominant beta-lactamase gene, bla_{CTX}, was present on most of the plasmids, except IncL/M. Of the carbapenemase-encoding genes, bla_{OXA} predominated and was widely present on plasmids, notably IncX3 and IncL/M (Figure 5). In contrast, bla_{VIM} was only present on the IncL/M plasmid.

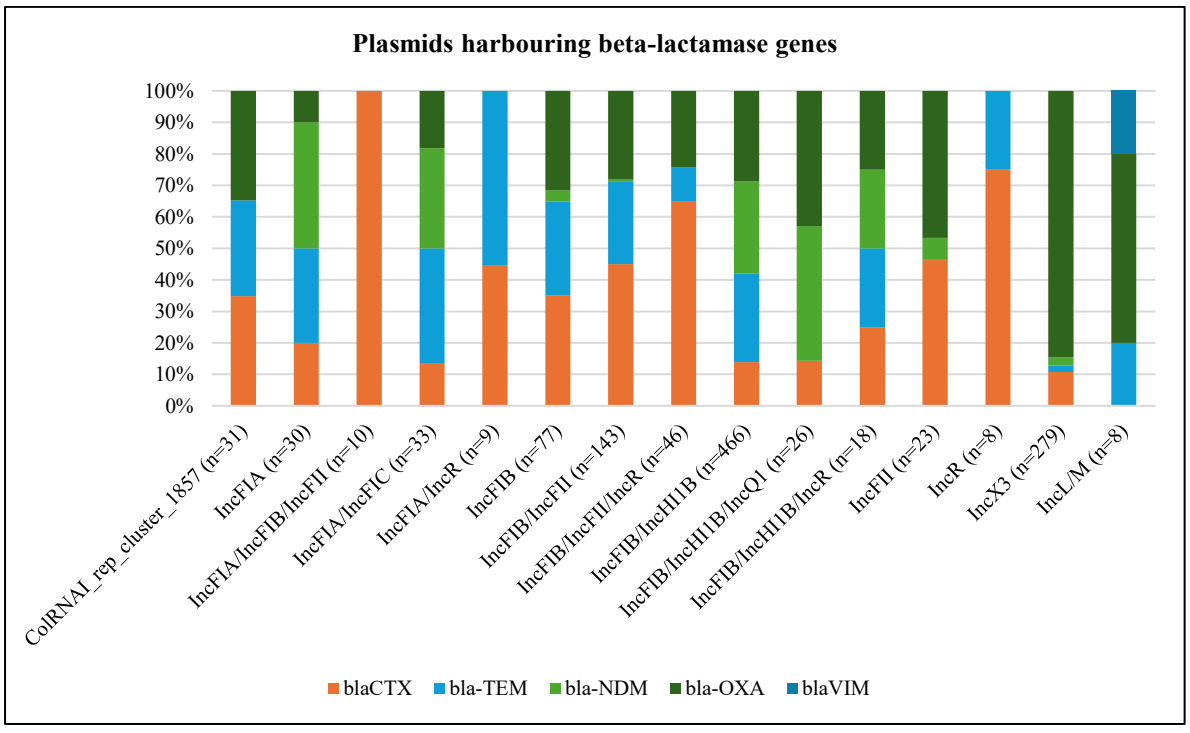


Figure 5. Plasmids harbouring beta-lactamase genes

Fifty-eight reconstructed plasmid types were associated with carriage of virulence-encoding genes (Supplementary Table 8), of which six plasmid types were most frequently associated with the virulence genes (Supplementary Figure 2). None of the identified plasmids harboured virulence-encoding genes associated with Kleborate v3 virulence scores of 1 (*ybt*), 2 (*ybt* and *clb* or *clb*), 4 (*ybt* and *iuc*), and 5 (*ybt*, *clb* and *iuc*) (Figure 6).

Isolates harbouring plasmid-mediated AMR genes (n=359) were detected from all study sites, with the greatest proportion (85.8%; 308/359) being pHAI isolates. Three plasmids frequently associated with harbouring beta-lactamase genes, IncX3, IncFII and IncFIB, were present together in 50.0% of novel ST9094 strains (24/48), and IncX3 was also present in 80.6% of ST17 strains (29/36) (Supplementary Figure 4). IncFIB/IncHI1B (harbouring both AMR and virulence genes) was present primarily in ST23 (83.3%; 10/12) and ST2497 (88.0%; 81/92) isolates (Supplementary Figure 4).

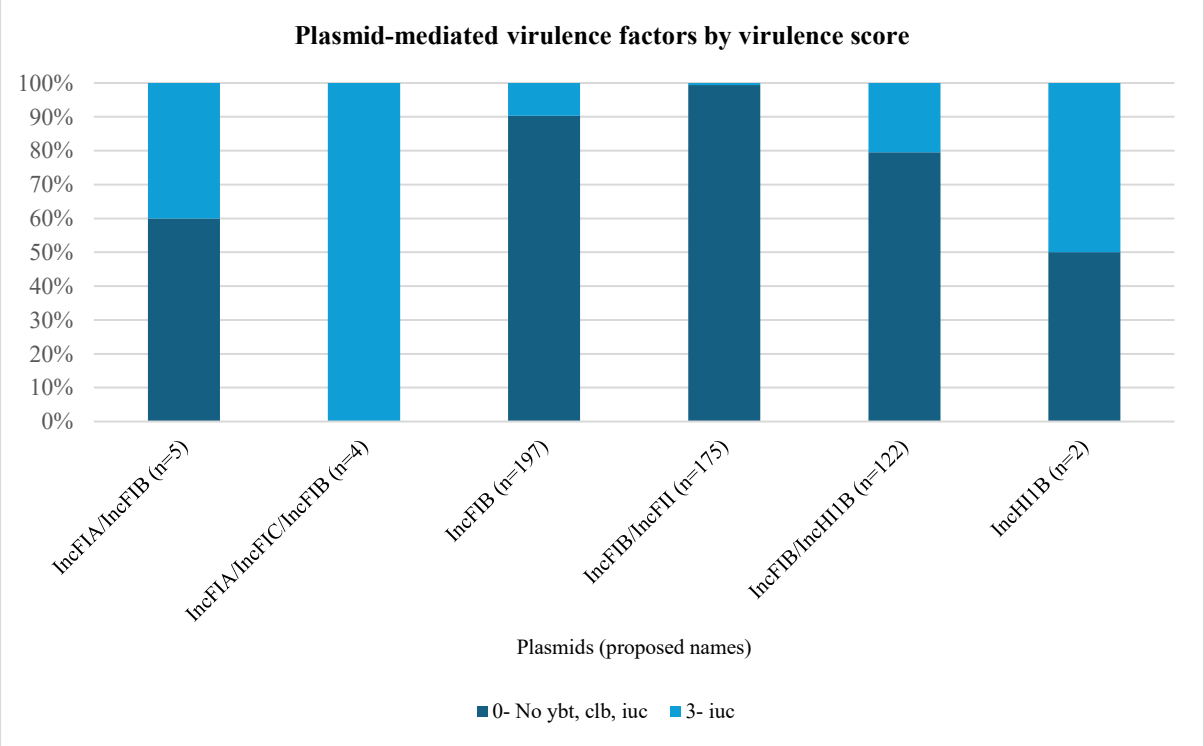


Figure 6. Plasmids harbouring virulence genes

Virulence scores were defined as per Kleborate v3: 0- negative for all of yersiniabactin, colibactin, aerobactin; 1- yersiniabactin only; 2- yersiniabactin and colibactin (or colibactin only); 3- aerobactin (without yersiniabactin or colibactin); 4- aerobactin with yersiniabactin (without colibactin); and 5- yersiniabactin, colibactin and aerobactin.

Risk factors for Death

Of the 448 individuals with KPn genomic and clinical outcome data, 60.9% (273/448) demised through to 90 days, including 250 during hospital admission and 23 within 90-days post-discharge. Logistic regression analysis showed increased risk of death for participants infected with KPn genotype ST2497 when compared with all other STs (OR: 2.91; 95% CI: 1.64- 5.17), and once the model was adjusted for patient clinical characteristics (aOR: 2.55; 95% CI: 1.39- 4.66) (Table 3). However, once the model was adjusted for K-loci, O-antigen types, AMR-encoding beta-lactamase genes and virulence-encoding genes, no significant association was found between ST2497 or any other genomic variable and increased risk of death. The patient characteristics consistently associated with increased risk of death were age over 65 years (aOR: 2.99; 95% CI: 1.48- 6.07), and having a pHAI compared to a pCAI (aOR: 2.20; 95% CI: 1.06- 4.56).

Table 3. Multiple regression modelling of association between *Klebsiella pneumoniae* genomic factors, patient characteristics and mortality risk

Factor	**ST group-unadjusted**	**Model 1:** ST group adjusted for patient characteristics	**Model 2:** Model 1 + K-loci	**Model 3:** Model 2 + O-antigen types	**Model 4:** Model 3 + AMR factors	**Model 5:** Model 4 + virulence factors
STs (other STs= reference)						
ST2497	2.91 (1.64- 5.17)***	2.55 (1.39- 4.66)*	1.91 (0.40- 9.05)	1.18 (0.14- 9.83)	1.22 (0.12- 12.29)	1.58 (0.15- 16.35)
ST39	1.39 (0.74- 2.60)	1.18 (0.60- 2.29)	0.90 (0.33- 2.43)	0.89 (0.32- 2.44)	0.80 (0.27- 2.38)	1.01 (0.28- 3.66)
ST9094	1.23 (0.64- 2.37)	1.12 (0.57- 2.22)	0.53 (0.12- 2.34)	0.91 (0.14- 5.77)	0.86 (0.13- 5.61)	0.82 (0.12- 5.40)
ST17	1.79 (0.81- 3.97)	1.65 (0.72- 3.80)	0.78 (0.17- 3.70)	1.35 (0.20- 9.09)	1.34 (0.20- 9.11)	1.31 (0.19- 9.01)
ST307	1.28 (0.44- 3.71)	1.38 (0.46- 4.17)	1.58 (0.51- 4.85)	1.47 (0.38- 5.71)	1.44 (0.36- 5.80)	1.32 (0.32- 5.33)
Sex (male= reference)						
Female	-	1.44 (0.96- 2.17)	1.45 (0.96- 2.20)	1.49 (0.98- 2.25)	1.49 (0.98- 2.27)	1.53 (1.00- 2.34)*
Age categories, years (18-35= reference)						
36-50	-	0.91 (0.54- 1.52)	0.88 (0.53- 1.48)	0.90 (0.53- 1.52)	0.89 (0.53- 1.51)	0.90 (0.53- 1.52)
51-65	-	1.60 (0.91- 2.81)	1.62 (0.92- 2.86)	1.63 (0.92- 2.89)	1.63 (0.91- 2.90)	1.64 (0.91- 2.94)
>65	-	2.71 (1.39- 5.28)*	2.84 (1.44- 5.57)*	2.82 (1.43- 5.59)*	2.78 (1.40- 5.51)*	2.99 (1.48- 6.07)*
Study site (CHBAH= reference)						
CMJAH	-	0.76 (0.50- 1.18)	0.79 (0.51- 1.22)	0.80 (0.51- 1.25)	0.81 (0.52- 1.28)	0.88 (0.55- 1.40)
HJH	-	0.57 (0.24- 1.39)	0.59 (0.24- 1.44)	0.59 (0.24- 1.46)	0.58 (0.23- 1.43)	0.62 (0.25- 1.58)
Category of infection (Community-acquired= reference)						
Healthcare-associated	-	1.99 (1.08- 3.69)*	2.09 (1.09- 4.00)*	2.08 (1.08- 3.40)*	1.99 (0.98- 4.04)	2.20 (1.06- 4.56)*
K-loci (other KL= reference group)						
KL64	-	-	1.52 (0.34- 6.89)	2.48 (0.35- 17.36)	2.42 (0.32- 18.10)	2.29 (0.30- 17.32)
KL25	-	-	2.42 (0.60- 9.79)	4.50 (0.46- 44.19)	4.41 (0.42- 45.71)	5.52 (0.52- 58.50)
KL2	-	-	1.69 (0.77- 3.74)	1.68 (0.68- 4.10)	1.69 (0.68- 4.15)	1.43 (0.41- 4.96)
KL149	-	-	1.28 (0.37- 4.51)	1.09 (0.29- 4.05)	1.18 (0.31- 4.51)	1.15 (0.26- 5.05)
KL30	-	-	1.06 (0.43- 2.62)	0.89 (0.34- 2.38)	0.89 (0.33- 2.41)	1.03 (0.37- 2.84)
O-types (other OTs= reference group)						
O1αβ,2α	-	-	-	1.13 (0.45- 2.81)	1.13 (0.45- 2.83)	1.02 (0.40- 2.61)
O1αβ,2β	-	-	-	1.42 (0.63- 3.22)	1.39 (0.61- 3.18)	1.16 (0.48- 2.78)
O5	-	-	-	0.35 (0.02- 6.72)	0.36 (0.02- 7.55)	0.29 (0.01- 6.28)
O2β	-	-	-	1.22 (0.43- 3.44)	1.23 (0.42- 3.55)	1.22 (0.42- 3.56)
O2α	-	-	-	0.70 (0.17- 2.89)	0.72 (0.17- 2.96)	0.78 (0.18- 3.32)
Resistance-encoding beta-lactamase genes (No bla= reference)						
			-	-		
bla_{CTX}	-	-	-	-	1.23 (0.52- 2.92)	1.42 (0.58- 3.45)
bla_{TEM}	-	-	-	-	0.87 (0.50- 1.52)	0.80 (0.44- 1.42)
bla_{OXA}	-	-	-	-	0.95 (0.41- 2.20)	1.06 (0.44- 2.54)
bla_{NDM}	-	-	-	-	0.99 (0.43- 2.25)	0.97 (0.42- 2.26)
Virulence-encoding genes (No VF= reference group)						
Ybt	-	-	-	-	-	0.85 (0.45- 1.61)
Clb	-	-	-	-	-	3.80 (0.79- 18.40)
Iuc	-	-	-	-	-	0.52 (0.02- 14.20)
Iro	-	-	-	-	-	0.69 (0.01- 32.97)
rmpADC	-	-	-	-	-	2.85 (0.22- 36.62)

Outcome data were available for 448 KPn-ID cases. ORs (95% CIs) are presented; *P<0.05, ***P<0.001

Abbreviations: ST, sequence type; AMR, antimicrobial resistance; CHBAH, Chris Hani Baragwanath Academic Hospital; CMJAH, Charlotte Maxeke Johannesburg Academic Hospital; HJH, Helen Joseph Hospital; KL, K-locus, OT, O-antigen type; bla, beta-lactamase; VF, virulence factor.

DISCUSSION

Our multicentre, prospective study provides a detailed analysis of the genomic epidemiology on a large cohort of KPn invasive isolates from blood and cerebrospinal fluid (CSF) clinical samples in Johannesburg, Gauteng, the province with the highest burden of CRKp in South Africa.[12] Whereas most smaller regional epidemiological studies focus on AMR-related KPn-ID, in particular CRKp[13, 14] and extended-spectrum-β-lactamase (ESBL)-producing KPn,[15] we included the clinical and molecular analysis of isolates of all susceptibilities to provide a reliable estimate of the burden of all KPn-ID in African adults. Genomic surveillance offers a unique opportunity to reduce the morbidity and mortality attributed to KPn-ID, including CRKp and ESBL KPn, by rapid identification of AMR genes and a prediction of response to therapy,[7] by mapping transmission dynamics to avert early outbreaks by using targeted IPC measures,[7, 16, 17] and by identification of hypervirulent strains.[6] The broader public health implications of KPn genomic surveillance include using serotyping to develop a KPn vaccine that has the potential to prevent KPn-ID, reduce onward transmission of the pathogen and, in turn, reduce use of antimicrobials.[8, 18]

Our study described the population genomics of KPn-ID amongst adults in Johannesburg, including predominant serotypes and the transmission dynamics noted within our study sites, three major academic hospitals. Globally, specific STs (ST11, ST258, ST437, ST147 and ST101) have been associated with MDR KPn and/or hvKp (ST23, ST65 and ST86).[1, 19] We found over thirty STs associated with the 302 CRKp isolates in our study (determined on phenotypic AST), of which ST2497 (29.1%), ST39 (17.9%), and ST9094 (14.2%) predominated. A South African study investigating the molecular epidemiology of 22 CRKp isolates from invasive and colonising samples in Pretoria, Gauteng, also found ST2497 (22.7%; 5/22) to be the predominant ST, followed by ST17 (18.2%; 4/22) and ST307 (13.6%; 3/22).[14] Another South African study of 85 CRKp isolates from Cape Town, reported ST219 to be most common (18.8%; 16/85), followed by ST307 (17.6%; 15/85), ST17 (12.9%; 11/85), ST13 (12.9%; 11/85), and ST2497 (10.6%; 9/85).[13] The collective South African findings above differ significantly from a recent review on the global epidemiology of ESKAPE pathogens which noted that the most prevalent CRKp lineage (predominantly in the United States and Europe) is clonal group 258, comprised of ST11, ST258 and ST512,[20] all of which had a prevalence less than 3.0% in our study.

We found that KL1 which is typically associated with hvKp, constituted a low percentage (<3.0%) of our isolates. KL116 (27%; 28/103) predominated amongst pHAI KPn isolates in a Ghanaian study,[15] but was not detected in our cohort. The CRKp study in Pretoria found the predominant K-locus to be KL64,[14] while KL114 predominated in Cape Town, South Africa.[13] Twenty-five different K-loci were associated with the 302 CRKp isolates in our study (determined on phenotypic AST), of which KL64 (32.4%), KL25 (25.8%), and KL2 (13.2%) predominated. The common finding of KL64 predominating amongst CRKp isolates in Pretoria and Johannesburg suggests regional homogeneity in Gauteng province.

KPn transmission dynamics varied across the three hospitals in our study, reflecting different nosocomial risks. CHBAH experienced large, sustained, multi-ward clusters dominated by high-risk STs (ST2497, ST39, ST9094) with both ward-specific and inter-ward transmission, particularly from the general medical ward. CMJAH exhibited smaller, temporally and spatially limited clusters, indicating frequent introductions into high-risk wards with restricted spread. In contrast, HJH showed minimal onward transmission, with most infections arising from sporadic, genetically distinct isolates likely originating from external sources such as the community, referral hospitals, or the patient's own microbiota, rather than from in-hospital amplification. Our analysis on the isolates from six individuals who had second episodes of KPn-ID demonstrated that isolates from the same individual represented the same infecting strain in two individuals, suggesting re-infection from the patient's own colonising KPn strain which has been reported in other studies.[21, 22]

Our study also described the resistome, plasmidome, virulome, and investigated genomic variables associated with death. While Kleborate AMR determinants are not known to directly correlate with phenotypic resistance,[23] we found it clinically informative to compare the prevalence of genomic determinants of AMR against the prevalence of phenotypic resistance in our setting. The concordance between genomic and phenotypic detection of AMR was high for 3GCs, making genomic detection reliable for surveillance, rapid diagnostics and directed treatment of clinically relevant ESBL-related AMR. In our study, phenotypic detection of resistance to aminoglycosides and carbapenems were lower than genotypic detection, and although this may represent increased sensitivity of genomic testing, clinical correlation is still warranted to avoid overtreatment and unnecessary use of antimicrobials which would further contribute to AMR and the consequences thereof. The genotypic detection of colistin resistance was lower than phenotypic methods using broth microdilution

techniques, making genomic detection of colistin resistance unreliable in our setting. The poor reliability of genotypic colistin resistance detection in KPn isolates is considered to be expected, given the complexity of the mechanisms of colistin resistance and the diversity of molecular markers involved.[24]

A recent systematic review and meta-analysis on the prevalence of carbapenemase-encoding genes in KPn clinical isolates from Africa noted an overall pooled prevalence of 34.0% (95% CI: 26.0–42.0%), with a pooled prevalence of bla_{OXA-48} of 17.0% (95% CI: 12.2–21.8%), and bla_{NDM} of 15.1% (95% CI: 9.8– 20.4%).[25] We detected a significantly higher prevalence of all carbapenemase-encoding genes (62.2%), bla_{OXA} (55.5%), and bla_{NDM} (25.8%), possibly due to inconsistent IPC practices and indiscriminate use of carbapenems, generally. The observed association of the AMR-gene harbouring plasmids IncX3, IncFII and IncFIB with ST9094, IncX3 with ST17, and IncFIB/IncHI1B with ST23 and ST2497 strains suggests clonal expansion of plasmid-associated AMR genes in these lineages.

The prevalence of hvKp, based on a combination of clinical characteristics, the string test and murine studies, is estimated to be 37.8% of KPn infections in mainland Chinese hospitals, and much lower (6.3%) in United States.[26] Whilst there is still no molecular or microbiological consensus definition of hvKp, several genomic determinants associated with hypervirulence can assist in differentiating hvKp from strains of classical KPn and genomic criteria for hvKp have subsequently been proposed.[27, 28] Russo *et al* noted the limitation of a Kleborate virulence score at predicting phenotypic hypervirulence, and proposed the detection of all five of the plasmid-mediated biomarkers *iucA*, *iroB*, *peg-344, rmpA*, and *rmpA2* as criteria for hvKp.[28] Because we enrolled KPn-ID cases of all drug susceptibilities, we were also able to report on the prevalence of KPn-associated virulence factors in our region. None of the KPn isolates in our cohort met these suggested criteria for hvKp, although we observed a prevalence of highly virulent KPn (Kleborate virulence score of 5) in 5.5% of isolates in our study.

A limitation of our study included not repeating phenotypic AST on invasive isolates that underwent WGS which may have allowed for more in-depth statistical comparisons to be performed between phenotypic and genotypic AST methods. Also, plasmids were analysed based on short read sequences only; any further and more detailed analyses should be done using long reads nanopore or PacBio technology.

In conclusion, we demonstrated significant genomic diversity amongst isolates within Johannesburg, and compared with invasive isolates in other South African provinces and sub-Saharan African countries. We found a predominance of ST2497, ST39 and ST9094 amongst our study sites, suggesting clonal expansion which was confirmed in a transmission analysis. Our findings on the resistome and plasmidome of invasive KPn isolates in South Africa add to the body of genomic surveillance data in the African region, and highlight the need for ongoing longitudinal clinical and molecular surveillance in establishing the burden of KPn-ID in South Africa, and providing baseline data to inform possible interventions against KPn-ID.

Methods

Study design and study population

We undertook a prospective laboratory-based surveillance study from May 15th, 2023 to May 14th, 2024 of blood and CSF cultures at the National Health Laboratory Service (NHLS) microbiology laboratories serving three academic hospital study sites in Johannesburg, South Africa. The study sites included Chris Hani Baragwanath Academic Hospital (CHBAH; a secondary-tertiary level care hospital), Charlotte Maxeke Johannesburg Academic Hospital (CMJAH; a tertiary level care hospital) and Helen Joseph Hospital (HJH, a secondary-level care hospital). Inclusion criteria included hospitalised adults ≥ 18 years of age with KPn-ID, defined as KPn cultured from a blood or CSF sample.

The blood and CSF cultures were obtained from hospitalised adults with presumed sepsis that were investigated at the discretion of the attending physicians. Participants suspected of having a second distinct episode of KPn-ID (defined as clinical resolution of the initial infection followed by a repeat KPn infection ≥21 days later) were not re-enrolled in the clinical surveillance. Nevertheless, the isolates related to the second infection episodes were collected and underwent genotypic testing. Each invasive disease episode was considered a separate infection and thus each isolate was analysed independently in participants with two invasive disease episodes. Presumed healthcare-associated infections (pHAIs) were defined as an invasive KPn detected ≥48 hours after admission to the hospital or if there was contact with a healthcare service in the preceding two weeks. Presumed community-associated infections (pCAIs) were defined as an invasive KPn detected on a sample collected within 48 hours of admission, and no previous contact with a healthcare service in the preceding two weeks. Clinical and laboratory information was abstracted from clinical records and entered into a password-protected electronic database (REDCap, USA). Participants were followed up until hospital discharge or in-hospital death.

A follow-up telephone interview was conducted 90 days post-discharge to assess outcomes of individuals who were discharged from the hospital.

Clinical isolates

Blood and CSF samples underwent microbiological testing and antimicrobial susceptibility testing (AST) at participating NHLS microbiology laboratories, according to standard operating procedures at each site, and has been reported previously.[11] Briefly, KPn was detected using either the VITEK 2 automated system (bioMérieux, France) or the MALDI-TOF VITEK MS (bioMérieux, France) system. Antimicrobial susceptibility testing (AST) was performed using the VITEK 2 automated system (bioMérieux, France), or by disc diffusion, gradient diffusion, or broth microdilution methods, which, along with interpretation of results, followed the 2023 and 2024 Clinical and Laboratory Standards Institute (CLSI) guidelines.[29, 30] Isolates reported with intermediate antimicrobial susceptibility were differentiated from isolates reported to have full resistance. Third-generation cephalosporin (3GC) resistance was designated when phenotypic resistance was detected to at least one of the three 3GCs (ceftriaxone, cefotaxime or ceftazidime). Carbapenem resistance was based on phenotypic resistance to at least one of the three carbapenems (ertapenem, imipenem or meropenem). Extended-spectrum-β-lactamases (ESBL) production was not specifically tested for in CRKp isolates, thus only phenotypic 3GC resistance was reported. Colistin minimum inhibitory concentration (MIC) testing in CRKp isolates was done using standard broth microdilution techniques.[29, 30] Additionally, carbapenemase detection was performed with the RESIST-5 O.K.N.V.I. lateral flow assay (Coris BioConcept, Belgium). Multidrug-resistant (MDR) was defined as KPn resistant to ≥ 3 antimicrobial categories.[31]

The available blood and CSF KPn isolates from eligible participants were collected from the NHLS laboratories and underwent confirmatory species testing at the Wits Vaccines and Infectious Diseases Analytics Unit (Wits-VIDA) laboratory. Briefly, single bacterial colonies were subcultured onto Colorex Orientation Agar plates (MediaMage, South Africa) and then incubated at 37°C for 24 hours. Following growth, the identification of KPn isolates was performed by classical microbiology techniques and confirmed by using the Analytical Profile Index 20E (API 20E, bioMérieux, France). Confirmed KPn colonies were subcultured onto 5% Columbia Sheep Blood Agar (MediaMage, South Africa) to ensure purity and adequate biomass before being stored at -80ºC in a medium containing skim milk, tryptone, glucose and 10% glycerol (STGG) for further analysis.

Whole genome sequencing

To prepare for WGS, genomic DNA was extracted directly from stored STGG, using the TANBead Maelstrom 4810 Automated DNA/RNA Purification Platform (Taiwan Advanced Nanotech Inc.) following the manufacturer's guidelines. Thirty isolates required an additional pre-lysis step to meet the minimum DNA input requirements. Briefly, 50 µL of the pure culture isolate in STGG was mixed with 200 µL of lysis buffer (Taiwan Advanced Nanotech Inc.), 10 µL of proteinase K, 10 µL of lysozyme, and 10 µL of mutanolysin (M9901, Merck KGaA, Germany), followed by a staged incubation at 37°C, 56°C, and 60°C before proceeding with the standard TANBead DNA extraction. Following DNA extraction, sequencing libraries were prepared using the Nextera XT DNA Kit (Illumina, US) and sequenced on the Illumina Nextseq 1000 and Illumina Nextseq 2000 platforms to generate short 2 x 150bp paired-end reads (Illumina, San Diego, CA) at the National Institute for Communicable Diseases (NICD), Johannesburg, South Africa. Quality control assessment of raw sequencing reads was performed at Imperial College London, with FastQC (version 0.12.1,2023)[32] and MultiQC (version 1.21, 2024)[33] files, generated for raw reads, to assess initial quality metrics. Raw reads were then trimmed, with low quality reads removed using Trimmomatic (version 0.39, 2021).[34] This used the following parameters: quality threshold set for both leading and trailing bases was 20, the minimum length for retaining informative reads was set as 36 bases, a sliding window, with a window size of 4 bases and an average quality threshold of 20. After trimming raw reads, Seqkit (version 2.8.2, 2024 https://bioinf.shenwei.me/seqkit/)[35] was used to estimate sequencing coverage by calculating the total number of bases relative to the reference genome size. Genome Assemblies of isolates were generated from trimmed reads using SPAdes (version 3.15.5, 2022)[36]. A multi-k-mer approach was used to address limitations of using a single k-mer size. Therefore k-mer sizes were set to '45,55,67,79,85,91,97,103,115,127'. Draft genomes from SPAdes were run through Quast (version 5.2.0, 2022)[37] to summarise statistics for genome assemblies. CheckM2 (version 1.0.2, 2023)[38] was used to assess the completeness and contamination of generated draft assemblies. Bioawk (version 1.0, 2015)[39] was then utilised to remove contigs less than 200bp, with Quast[37] and CheckM2[38] re-run on the filtered assembly files.

Genomic characterisation

Pathogenwatch (https://rc.pathogen.watch/en)[40] was used for species identification and KPn genomic characterisation using Kleborate v3.2.4:[23, 41] Multi-locus sequence typing (MLST) and core genome MLST (cgMLST) predictions were performed using Klebsiella Pasteur MLST[42] and BIGSdb-Pasteur,[43] respectively. Capsule K- and O- serotype predictions were made with Kaptive v3.[44] Antimicrobial resistance genes identified

with Kleborate[23] compared to the curated version (v3.2.4) of the Comprehensive Antibiotic Resistance Database program including ResFinder 9[45]. Virulence genes were identified by Kleborate[23] using the BLASTn for Key loci. Plasmid reconstruction was done using MOB-suite version 3.1.8.[46] Prior to plasmid reconstruction, bacterial species from raw reads were verified using Kraken2 (version 2.1.3, 2023)[47] and Bracken (version 2.9, 2023).[48] Multi-locus sequence typing (MLST) (v 2.23.0, 2022) was used for the genotyping of isolates. These MLST schemes are hosted on PubMLST[49] and BIGSdb-Pasteur[43] and examines and compares 7 house-keeping genes (adk, fumC, gyrB, mdh,icd, purA, recA and misn) to characterise bacterial strains and identify their unique allelic profiles. Identification of AMR genes carried on the plasmid genome was performed out using ABRicate (v1.0.0, 2020)[50] using the built in NCBI database.[51] Kleborate (version 2.3.2, 2023)[23] was used to analyse *Klebsiella* genomes, providing rapid identification of virulence factors. The tool Proksee[52] was used to visualise the hybrid plasmid IncFIB/ IncHI1B hybrid plasmid in one isolate (KPIDA0157), and Bakta[53] was used to annotate the plasmid genome.

Genomic relatedness between isolates was assessed using Split Kmer Analysis (SKA v1).[54] Draft assemblies were processed with ska align to generate split-kmer files (.skf) for each genome. Pairwise genomic distances were computed with ska distance, which reports SNP-per-site distances derived from shared split kmers. Because SKA v1 outputs proportional genomic distances rather than absolute SNP counts, values were converted to approximate SNP numbers by multiplying each distance proportion by the estimated KPn genome length (5.5 Mb). These SNP counts were used to infer strain relatedness and potential transmission events, using a threshold of ≤20 SNPs as indicative of high relatedness.

For evolutionary analysis, paired-end reads of the KPn isolates were assembled *de novo* using SPAdes.[36] A reference-free phylogeny was constructed directly from the assemblies using Split Kmer Analysis (SKA v2.0)[54] to generate a Newick-formatted tree, avoiding potential bias from using a single reference genome. Phylogenetic relationships were further assessed with wgMLST and cgMLST. The resulting phylogenetic tree was visualised and annotated in iTOL v12 (https://itol.embl.de).[55] SNP2Cluster[56] was used for transmission cluster detection based on core SNP and K-means clustering (https://github.com/stanikae/SNP2Cluster/tree/v0.5.4) using a similar methodology to a recent transmission analysis study.[17] We generated a SNP distance matrix, with default parameters of 20 SNP cut-off and 14 days collection date intervals, MLST profiles, and epidemiological data which included a single time point (date of culture) and the location of the individual at the time of culture as the movement of individuals between wards was not tracked. Closely related isolates were initially grouped via enhanced K-means clustering, with the optimal number of clusters (K) determined by silhouette scores and 500

bootstrap iterations. SNP-Epi clusters were determined by combining sequence type and temporal data resulting in transmission clusters.

Statistical Analysis

Categorical variables were summarised using frequencies in percentages. We compared the prevalence of genotypic detection of AMR with phenotypic AMR data for the following antimicrobial classes: aminoglycosides, carbapenems, 3GCs, colistin, fluoroquinolones, phenicols, and sulfonamides. The prevalences were summarised using frequencies in percentages.

Multiple logistic regression models were used to investigate the association of genomic variables with overall mortality (in-hospital and up to 90 days post discharge) in a group of 448 individuals with outcome data. As mortality is a rare outcome, multiple mortality models were used to reduce small sample bias. The KPn sequence type was included as a categorical model in all models, as well as patient clinical characteristics (sex, age category, study site, pCAI and pHAI), which were considered potential confounders for death. The following groups of genomic variables were sequentially added to each logistic regression model: the top five K-loci, the top five O-antigen types, resistance-encoding beta-lactamase genes bla_{CTX}, bla_{TEM}, bla_{OXA}, and bla_{NDM}, and virulence-encoding genes yersiniabactin (*ybt)*, colibactin (*clb)*, aerobactin (*iuc)*, salmochelin (*iro)* and *rmpADC*. In cases with two distinct episodes of KPn (i.e. two isolates per participant), the isolate associated with the first infection episode was not included in the regression analysis; the second isolate was included as it was considered temporally related to outcome. Adjusted odd ratios and the associated 95% CIs were estimated using multivariable logistic regression analysis. Statistical analysis was performed using Stata/SE version 18.5 (StataCorp). A p-value <0.05 was considered statistically significant.

Ethical approval

The study was approved by the Human Research Ethics Committee (HREC), University of the Witwatersrand (HREC reference number: M220960), and registered on the South African National Health Research Database (reference number: GP202209026). Written informed consent was obtained from individuals enrolled during their hospital admission. If individuals were too ill to provide consent at the time of enrolment, proxy consent was obtained from their next-of-kin. A waiver for consent was granted by the HREC for cases where an individual was discharged or died before consent could be obtained, to enable retrospective enrolment of the

cases for whom medical records could be retrieved. An ethical waiver was granted for the collection of isolates, but not clinical information, from individuals with KPn-ID who were excluded from the clinical study.

Data availability

All sequencing reads generated in this study have been deposited in the NCBI genome database under the BioProject number PRJNA1369306 and PRJNA1219545. The individual assembly accession numbers are included with the Supplementary materials.

Supplementary Data

Supplementary materials are available online. Consisting of data provided by the authors to benefit the reader, the posted materials are not copyedited and are the sole responsibility of the authors, so questions or comments should be addressed to the corresponding author.

Notes

Author contributions. D.L.R, Z.D and S.A.M conceptualised and designed the study. D.L.R collected and analysed the clinical and microbiological data, analysed the molecular data, performed statistical analyses and wrote the first draft of the manuscript. N.J.D, L.V.D.M, V.B and C.P.O facilitated the molecular testing (DNA extractions and whole genome sequencing) of isolates. S.K performed the bioinformatics analysis. E.J and F.O.W. performed genomic quality control, plasmid reconstruction, characterisation and analysis. D.L.R, S.A.M, E.J, C.P.O, S.K, N.J.D, L.V.D.M, Z.D, F.O.W, J.W, T.N, R.C, S.N, L.M, J.M.T, J.N, A.P, and V.B contributed to the final manuscript.

Acknowledgments. The authors thank the staff at the NHLS laboratories and the laboratory team at Vaccines and Infectious Diseases Analytics Research Unit (Wits-VIDA). We also thank the Institut Pasteur teams for the curation and maintenance of BIGSdb-Pasteur databases at http://bigsdb.pasteur.fr. E.J acknowledges funding from Genesis Research Trust. E.J and F.O.W acknowledge computational resources and support provided by the Imperial College Research Computing Services.

Financial Support. This study was funded by Wits-VIDA, Johannesburg, South Africa. D.L.R received additional funding for the study from the South African Medical Research Council (SAMRC) through its Division of Research Capacity Development under the Clinician Researcher Development Programme from funding received from the Public Health Enhancement Fund / South African National Department of Health 2023. The content hereof is the sole responsibility of the authors and does not necessarily represent the official views of the SAMRC. D.L.R. received stipend support from the Professor Bongani Mayosi Netcare Clinical Scholarship.

Competing Interests. All authors report no competing interests.

REFERENCES

1. Gonzalez-Ferrer S, Penaloza HF, Budnick JA, Bain WG, Nordstrom HR, Lee JS, et al. Finding Order in the Chaos: Outstanding Questions in *Klebsiella pneumoniae* Pathogenesis. Infect Immun. 2021.
2. Wyres KL, Lam MMC, Holt. K. Population genomics of *Klebsiella pneumoniae*. Nat Rev Microbiol. 2020;18:344-59.
3. Sartorius B, Gray AP, Weaver ND, Aguilar GR, Swetschinski LR, Ikuta KS, et al. The burden of bacterial antimicrobial resistance in the WHO African region in 2019: a cross-country systematic analysis. Lancet Glob Health. 2024;12.
4. WHO Bacterial Priority Pathogens List, 2024: World Health Organization; 2024 [Available from: https://www.who.int/publications/i/item/9789240093461 [Accessed 2 March 2025].
5. Lin X, Li C, Zhang S, Yang X, Jiang. M. The Global and Regional Prevalence of Hospital-Acquired Carbapenem-Resistant *Klebsiella pneumoniae* Infection: A Systematic Review and Meta-analysis. Open Forum Infectious Diseases. 2023.
6. World Health Organization. Disease Outbreak News; Antimicrobial Resistance, Hypervirulent *Klebsiella pneumoniae* - Global situation 31 July 2024. [Available from: https://www.who.int/emergencies/disease-outbreak-news/item/2024-DON527. [Accessed on 3 October 2024].
7. Aanensen DM, Okeke IN, Donado-Godoy P, Egyir B, Lingegowda. R, Sia. S. The genomic surveillance gap: averting the antimicrobial resistance pandemic requires global equity and action. The Lancet Infectious Diseases. 2025.
8. Dangor Z., Benson N, Berkley JA, Bielicki J, Bijsma MW, Broad J, et al. Vaccine value profile for *Klebsiella pneumoniae*. Vaccine. 2024.
9. Estimating the impact of vaccines in reducing antimicrobial resistance and antibiotic use: technical report. Geneva: World Health Organisation; 2024.
10. Wantuch PL, Rosen D. *Klebsiella pneumoniae*: adaptive immune landscapes and vaccine horizons. Trends in Immunology. 2023;44(10).
11. Reddy DL, Dangor Z, Murray L, Tsitsi M, Nel J, Nana T, et al. Clinical and microbiological epidemiology of *Klebsiella pneumoniae* invasive disease in hospitalized adults in Johannesburg, South Africa: a multicentre observational study. JAC Antimicrob Resist. 2026;8(1).

12. Lowe M, Shuping L, Perovic O. Carbapenem-resistant Enterobacterales in patients with bacteraemia at tertiary academic hospitals in South Africa, 2019 - 2020: An update. S Afr Med J. 2022;112(8):542-52.
13. Marais G, Moodley C, Claassen-Weitz S, Patel F, Prentice E, Tootla H, et al. Carbapenem-resistant *Klebsiella pneumoniae* among hospitalized patients in Cape Town, South Africa: molecular epidemiology and characterization. JAC Antimicrob Resist. 2024.
14. Salvador-Oke KT, Pitout JDD, Peirano G, Strydom KA, Kingsburgh C, Ehlers MM, et al. Molecular epidemiology of carbapenemase-producing *Klebsiella pneumoniae* in Gauteng South Africa. Nature Scientific Reports. 2024;14.
15. Mills RO, Dadzie I, Le-Viet T, Baker DJ, Addy HPK, Akwetey SA, et al. Genomic diversity and antimicrobial resistance in clinical *Klebsiella pneumoniae* isolates from tertiary hospitals in Southern Ghana. J Antimicrob Chemother. 2024;79:1529-39.
16. Odih EE, Abdulahi JA, Amulele AV, Bates M, Heinz E, Hu W, et al. Contribution of nosocomial transmission to *Klebsiella pneumoniae* neonatal sepsis in Africa and South Asia: a meta-analysis of infection clusters inferred from pathogen genomics and temporal data. Preprint. 2025.
17. Shuping L, Ismail H, Mashau R, Kwenda S, Holt KE, Magobo RE, et al. Enhanced detection of neonatal invasive infection clusters in South Africa using epidemiological and genomic surveillance data. Preprint. 2025.
18. Brink A, Kgasi A, Musyoki A, Kagina B, Feldman C, Reddy D, et al. Vaccines in the fight against antimicrobial resistance – perspectives from South Africa. S Afr Med J 2024.
19. Arcari G, Carattoli. A. Global spread and evolutionary convergence of multidrug-resistant and hypervirulent *Klebsiella pneumoniae* high-risk clones. Pathogens and Global Health. 2023;117(4):328-41.
20. Miller WR, CA A. ESKAPE pathogens: antimicrobial resistance, epidemiology, clinical impact and therapeutics. Nature Reviews Microbiology 2024;22:598-616.
21. Gorrie CL, Mirceta M, Wick RR, Edwards DJ, Thomson NR, Strugnell RA, et al. Gastrointestinal Carriage Is a Major Reservoir of *Klebsiella pneumoniae* Infection in Intensive Care Patients. Clinical Infectious Diseases. 2017;65(2):208-15.
22. Gorrie CL, Mirceta M, Wick RR, Judd LM, Lam MMC, Gomi R, et al. Genomic dissection of *Klebsiella pneumoniae* infections in hospital patients reveals insights into an opportunistic pathogen. Nature Communications. 2022.
23. Lam MMC, Wick RR, Watts SC, Cerdeira LT, Wyres KL, Holt K. A genomic surveillance framework and genotyping tool for *Klebsiella pneumoniae* and its related species complex. Nature Communications. 2021;12.
24. Elias R, Duarte A, Perdigao J. A Molecular Perspective on Colistin and *Klebsiella pneumoniae*: Mode of Action, Resistance Genetics, and Phenotypic Susceptibility. Diagnostics 2021;11.
25. Sisay A, Kumie G, Gashaw Y, Nigatie M, Gebray HM, Reta. M. Prevalence of genes encoding carbapenem-resistance in *Klebsiella pneumoniae* recovered from clinical samples in Africa: systematic review and meta-analysis. BMC Infectious Diseases. 2025;25(556).
26. Lan P, Jiang Y, ZhouJ, Yu. Y. A global perspective on the convergence of hypervirulence and carbapenem resistance in *Klebsiella pneumoniae*. Journal of Global Antimicrobial Resistance. 2021;25:26-34.
27. Stanton T, Wyres K. What defines hypervirulent Klebsiella pneumoniae? The Lancet. 2024;108.
28. Russo TA, Lebreton F, PT. M. A Step Forward in Hypervirulent *Klebsiella pneumoniae* Diagnostics. Emerging Infectious Diseases. 2025;31(1).
29. CLSI. Performance *Standards for Antimicrobial Susceptibility Testing- Thirty-Third Edition: M100*. 2023.

30. CLSI. Performance *Standards for Antimicrobial Susceptibility Testing- Thirty-Fourth Edition: M100*. 2024.
31. Magiorakos AP, Srinivasan A, Carey RB, Carmeli Y, Falagas ME, Giske CG, et al. Multidrug-resistant, extensively drug-resistant and pandrug-resistant bacteria: an international expert proposal for interim standard definitions for acquired resistance. Clin Microbiol Infect. 2012;18:268-81.
32. Andrews S. FastQC. A quality control tool for high throughput sequence data. 2023 [30 January 2026]. Available from: https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.
33. Ewels P, Magnusson M, Lundin S, Kaller. M. MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics. 2016;32(19):3047-8.
34. Bolger AM, Lohse M, Usadel. B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30(15):2114-20.
35. Shen W, Sipos B, Zhao. L. SeqKit2: A Swiss Army Knife for Sequence and Alignment Processing. iMeta. 2024.
36. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. Journal of Computational Biology. 2012.
37. Gurevich A, Saveliev V, Vyahhi N, Tesler. G. QUAST: quality assessment tool for genome assemblies. Bioinformatics. 2013;29(8):1072-5.
38. CM2 Reference Database Uniref100/KO (Version 3) [Data set]. [Internet]. Zenodo. 2025.
39. BWK awk modified for biological data [Available from: https://github.com/lh3/bioawk?tab=readme-ov-file.
40. Argimon S, David S, Underwood A, Abrudan M, Wheeler NE, Kekre M, et al. Rapid Genomic Characterization and Global Surveillance of *Klebsiella* Using Pathogenwatch. Clinical Infectious Diseases. 2021(73).
41. Whitfield C, Kelley SD, Stanton TD, Wyres KL, Clarke BR, Forrester TJB, et al. O-antigen polysaccharides in *Klebsiella pneumoniae*: structures and molecular basis for antigenic diversity. Microbiology and Molecular Biology Reviews. 2025;89(2).
42. Insitut Pasteur BIGSdb-Pasteur Klebsiella pneumoniae species complex [Internet]. 2026 [cited 30 January 2026]. Available from: https://bigsdb.pasteur.fr/klebsiella/.
43. Insitut Pasteur BIGSdb-Pasteur [Internet]. 2026 [cited 30 January 2026]. Available from: https://bigsdb.pasteur.fr/.
44. Lam MMC, Wick RR, Judd LM, Holt KE, Wyres. K. Kaptive 2.0: updated capsule and lipopolysaccharide locus typing for the *Klebsiella pneumoniae* species complex. Microbial Genomics. 2022;8.
45. Florensa AF, Kaas RS, Clausen PTLC, Aytan-Aktug D, Aarestrup. F. ResFinder – an open online resource for identification of antimicrobial resistance genes in next-generation sequencing data and prediction of phenotypes from genotypes. Microbial Genomics. 2022;8.
46. Robertson J, Nash. J. MOB-suite: software tools for clustering, reconstruction and typing of plasmids from draft assemblies. . Microb Genom 2018;4: e000206.
47. Wood DE, Lu J, Langmead. B. Improved metagenomic analysis with Kraken 2 (2019). Genome Biology. 2019.
48. Lu J, Breitwieser FP, Thielen P, Salzberg. S. Bracken: estimating species abundance in metagenomics data. PeerJ Computer Science. 2017.
49. PubMLST. Public databases for molecular typing and microbial genome diversity [Internet]. 2026 [cited 30 January 2026]. Available from: https://pubmlst.org/.
50. ABRicate [Internet]. 2026 [cited 30 January 2026]. Available from: https://github.com/tseemann/abricate.
51. Feldgarden M, Brover V, Haft DH, Prasad AB, Slotta DJ, Tolstoy I, et al. Validating the AMRFinder Tool and Resistance Gene Database by Using Antimicrobial Resistance Genotype

Phenotype Correlations in a Collection of Isolates. Antimicrobial Agents and Chemotherapy. 2019;63(11).
52. Grant JR, Enns E, Marinier E, Mandal A, Herman EK, Chen C, et al. Proksee: in-depth characterization and visualization of bacterial genomes. Nucleic Acids Research. 2023.
53. Schwengers O, Jelonek L, Dieckmann MA, Beyvers S, Blom J, Goesmann A. Bakta: rapid and standardized annotation of bacterial genomes via alignment-free sequence identification. Microb Genom. 2021;7(11).
54. Harris S. Preprint. SKA: Split Kmer Analysis Toolkit for Bacterial Genomic Epidemiology. 2018.
55. Letunic I, Bork. P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res. 2021.
56. Kwenda S, Shuping L, Mashau R, Ismail H, Govender. N. SNP2Cluster: A core SNP and K-means clustering-based tool for enhanced transmission cluster detection in outbreak scenarios (v0.5.3). Klebsiella Epidemiology and Biology Symposium 2024 (KLEBS2024), Institut Pasteur, Paris, France. Zenodo [Internet]. 2024. Available from: https://zenodo.org/records/14060296.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- 3.KPIDWGScasesSupplementarymaterialnpjAntimicrobialsandResistance.pdf
- 4.SupplementaryfileS1.xlsx
- 5.Supplementary.BioSampleObjectsAccessionnumbersCasesKPID508.pdf