Genomic analysis of *Klebsiella pneumoniae* causing community-acquired respiratory deaths among Zambian infants and children using targeted RNA-probe hybridization-capture metagenomics

(short title) *Klebsiella pneumoniae* fatal community-acquired pneumonia in Zambian children

Kenneth Lindstedt[1,2,3,*], Alyse Wheelock[4,*], Mulemba Samutela[5], Wasifa Kabir[6], Mwelwa Chasaya[7], Natasha Namuziya[8], Emilia Jumbe Marsden[9], Monica Kapasa[8], Chibamba Mumba[10], Bwalya Mulenga[10], Lisa Nkole[8], Rachel Pieciak[11], Victor Mudenda[10], Chilufya Chikoti[7], Benard Ngoma[7], Charles Chimoga[7], Sarah Chirwa[7], Lilian Pemba[7], Diana Nzara[7], James Lungu[7], Leah Forman[12], Edgar Simulundu[13, 14], William MacLeod[11], Crispin Moyo[7], Somwe Wa Somwe[5,8], Kathryn E. Holt[15,16], Arnfinn Sundsfjord[1,2,3], Christopher J. Gill[11]

1 Norwegian Centre for Detection of Antimicrobial Resistance, Department of Microbiology and Infection Control, University Hospital of North Norway, Norway

2 Department of Medical Biology, Faculty of Health Sciences, UiT the Arctic University of Norway, Norway

3 Centre for New Antibacterial Strategies, UiT the Arctic University of Norway, Norway

4 Section of Preventive Medicine and Epidemiology, Chobanian & Avedisian School of Medicine, Boston University, Boston, MA, USA

5 School of Health Sciences, University of Zambia, Lusaka, Zambia

6 Aquaculture and Environment Group, UiT the Arctic University of Norway, Norway

7 Avencion Limited, Lusaka, Zambia

8 University Teaching Hospital – Children's Hospital, Lusaka, Zambia

9 Pendleton Children's Clinic, Lusaka, Zambia

10 Department of Pathology, University Teaching Hospital, Lusaka, Zambia

25 11 Department of Global Health, Boston University School of Public Health, Boston, MA, USA

26 12 Center for Health Data Science (CHDS), Boston University School of Public Health, Boston,
27 MA, USA

28 13 Macha Research Trust, Choma, Zambia

29 14 Zambia National Public Health Institute, Lusaka, Zambia

30 15 Department of Infection Biology, London School of Hygiene and Tropical Medicine, London,
31 UK

32 16 Department of Infectious Diseases, School of Translational Medicine, Monash University,
33 Melbourne, Victoria 3004, Australia

34 *Co-first authors, contributed equally

35 Corresponding author (technical questions): Kenneth Lindstedt, MBBS, PhD, Norwegian Centre
36 for Detection of Antimicrobial Resistance, Department of Microbiology and Infection Control,
37 University Hospital of North Norway, Norway. E-mail: kenneth.w.lindstedt@uit.no

38 Corresponding author (study information): Alyse Wheelock, MD, Section of Preventive Medicine
39 and Epidemiology, Chobanian & Avedisian School of Medicine, Boston University, Boston, MA,
40 USA. E-mail: alyse.wheelock@gmail.com

41

42 **Abstract**

43 *Klebsiella pneumoniae* (Kp) is a leading cause of neonatal and infant deaths in sub-Saharan
44 Africa and frequently associated with antimicrobial resistance. Previously, we identified Kp as a
45 major cause of fatal community-associated lower respiratory infections among infants and
46 children under five years in Lusaka, Zambia, using postmortem tissue sampling and pathogen
47 specific multiplex qPCR. In this follow-up study, we employed a novel culture-independent RNA-
48 probe hybridization-capture metagenomic sequencing approach, targeting Kp pan-genome core

and accessory genes, to perform in-depth genomic analysis of Kp from eleven post-mortem lung biopsy samples from seven of these children. Analysis detected Kp in all cases except one, which identified *Klebsiella quasipneumoniae* subspecies *similipneumoniae*. Core-genome multi-locus sequence typing (cgMLST) revealed six clonal groups (CG607, CG1123, CG10072, CG280, CG3648, and CG10344) belonging to five sublineages (SL607, SL17, SL280, SL37, and SL10072), with perfect concordance between paired samples from the same case. Two infants sampled the same month harbored SL607 lineages sharing 621 out of 629 cgMLST alleles, suggesting clonal spread. Kp capsule (K) loci were detected in all but one case and included potential vaccine targets KL25, KL23, and KL122. Antimicrobial resistance genes were widespread among samples, particularly encoding resistance toward aminoglycosides, β-lactams, sulphonamides, tetracyclines, and trimethoprim. Extended spectrum β-lactamases were identified in four cases, three of which were $bla_{\text{CTX-M-15}}$. The acquired Kp sideophore yersiniabactin (lineage *ybt14*) was identified in both cases associated with SL607, and the acquired siderophore aerobactin (lineage *iuc5*) was identified in one of these, suggesting possible convergence of antimicrobial resistance and hypervirulence. The detection of Kp with extensive antimicrobial resistance causing fatal community acquired pneumonia signals a deeply concerning epidemiologic shift from a largely nosocomial pathogen. This calls for urgent epidemiological investigations to better understand the burden, transmission dynamics, antimicrobial resistances, and potential vaccine targets for Kp in other community settings across sub-Saharan Africa.

Author Summary

Klebsiella pneumoniae is a major cause of infections and death among newborns and young children, particularly in low-income countries, where it is frequently resistant to antibiotics. While well-known as a hospital-associated pathogen, we previously showed *K. pneumoniae* is also a

leading cause of fatal community lung infections among infants and children in Lusaka, Zambia. In this follow-on analysis, we performed deeper genetic analysis of *K. pneumoniae* detected from the cluster of community pneumonia deaths using lung tissue samples from seven of these children. Since traditional bacterial cultures were unavailable, we instead used a novel approach that enriched and sequenced specific regions of the *K. pneumoniae* genome directly from the biopsy samples without culturing bacterial isolates. We identified five different *K. pneumoniae* genetic subtypes, known as sublineages. Two sublineages, which came from children sampled the same month, were highly similar, suggesting clonal spread. Multiple acquired antimicrobial resistance genes were detected across all sublineages. Acquired virulence factors, which may cause more aggressive infections, were also detected in two cases. We also identified capsule types previously suggested as potential vaccine targets. This study underscores the urgent need to better understand and address the emerging burden of antibiotic-resistant *K. pneumoniae* pneumonia and other invasive infections among infants and children in community settings in sub-Saharan Africa.

Introduction

Historically, *Klebsiella pneumoniae* (Kp) has not been recognized as a significant cause of community-acquired acute lower respiratory infection (ALRI) among children, as informed by landmark epidemiologic research conducted in high-resource settings (1, 2). Previous large, multi-center studies of ALRI etiologies in low- to- middle-income countries (LMICs) have systematically excluded Kp from their analyses (3, 4). However, prior research suggests that nasal carriage and subsequent ALRI from enteric Gram-negative bacilli such as Kp vary geographically, with these pathogens featuring more prominently as causes of ALRI in warmer and more humid climates (5-9). More recent data from the Drakenstein birth cohort in South Africa strongly implicate Kp as an

important cause of childhood community-acquired ALRI in certain settings (10). Moreover, Kp has emerged as a frequent pathogenic cause of childhood mortality in the Child Health and Mortality Prevention Surveillance (CHAMPS) network across sub-Saharan Africa and south Asia (11-13), including in deaths due to pneumonia (14). While the majority (86%) of deaths investigated through the CHAMPS network occurred in-hospital, of 329 deaths occurring in the community, 47 (14%) featured Kp in the causal chain, raising concern for possible community-acquired infection. We previously reported on a post-mortem study investigating 106 community deaths among children under age five in Lusaka, Zambia. Using multiplex PCR testing of lung samples collected via minimally invasive tissue sampling (MITS) (15-17), we were concerned to identify a cluster of community-acquired ALRI caused by Kp in 13 of 49 (27%) respiratory-related deaths.

Kp as a potential cause of childhood ALRI in community settings is alarming because 1) this pathogen is not well-covered by current treatment guidelines (18) and 2) global molecular epidemiology of Kp has been characterized by rising prevalence of multidrug resistant (MDR) and hypervirulent sublineages, as well as convergence of MDR and hypervirulent phenotypes (19). Molecular epidemiology data are sparser in sub-Saharan Africa than other regions of the world, but evidence indicates widespread MDR among clinical Kp isolates (13, 20-22). Genotypic and phenotypic studies have shown common to near-universal levels of extended-spectrum β-lactamase (ESBL) carriage among Kp in African settings, and molecular studies have revealed frequent detections of virulence factors, particularly the acquired siderophore yersiniabactin (23-27).

Previous in-depth genomic analysis has revealed Kp phylogeny is composed of a highly diverse population structure, consisting of hundreds of deep branching lineages (28). While acquisition of

121 resistance and virulence determinants is widespread throughout the phylogeny, the majority of
122 AMR and hypervirulence within Kp has been disproportionately driven by the expansion of a
123 distinct subset of globally disseminated lineages, termed 'global problem clones' (28). Monitoring
124 Kp genomic epidemiology is therefore of high importance to understand regional and international
125 patterns of clonal spread, identify high-risk lineages, and inform infection control and intervention
126 strategies, including vaccine development.

127

128 The finding of community deaths due to Kp ALRI in our original analysis was surprising and carried
129 important clinical and public health implications. We therefore sought to better characterize the
130 causative Kp in terms of antimicrobial resistance genes, virulence factors, lineage and capsule
131 serotype by sequencing Kp DNA detected from the preserved lung tissue biopsies. Due to the
132 unavailability of bacterial culture from these samples, we utilized culture-independent targeted
133 hybridization-capture metagenomic sequencing to achieve these aims.

134

135 **Results**

136 **Description of cases**

137 Our previous post-mortem surveillance study enrolled 121 deceased infants and children between
138 January 2020 to December 2022, with pauses in enrollment due to the COVID-19 pandemic. As
139 previously reported, a total of 106 cases had complete molecular and histopathologic data and
140 were adjudicated by the DeCoDe panel, of which 49 (46%) deaths were determined to be due to
141 respiratory causes (15). Among these, Kp pneumonia was the most common respiratory cause
142 of death, accounting for 13 (27%) of the respiratory deaths. For Kp targeted hybridization-capture
143 analysis, DNA was re-extracted from 16 paired samples of 8 selected cases with the strongest

144 molecular signal supporting Kp pneumonia. Of these, 5 samples proved insufficient for target
145 hybridization-capture due to low DNA concentration (<10 ng/μL) and poor-quality metrics
146 (260/280 ratio <1.5 or >2.2; 260/230 ratio <0.5 or >2.3). The remaining 11 samples from 7
147 deceased infants and children (4 paired samples and 3 individual samples) had sufficient quality
148 and proceeded to target metagenomic hybridization-capture sequencing (**Supp. Table 1**).

149 The 7 deceased infants and children whose samples underwent Kp probe hybridization-capture
150 analysis ranged in age from 7 days to 20 months and were enrolled in the study between February
151 2020 and December 2022. Based on enrollment criteria, none were hospitalized at the time of
152 death; the deaths all occurred at home, en route to care, or while awaiting medical evaluation at
153 a healthcare facility. All the deceased had a cough preceding their death, some had notable
154 shortness of breath, and other symptoms reported included fever, vomiting, and diarrhea (**Table**
155 **1**). From the narratives provided by family members describing the events leading up to the death,
156 some of the deceased had a noted prior interaction with a health facility, including prior
157 hospitalizations (**Supp. Table 2**). In all but one case (where the information was insufficient), the
158 DeCoDe panel determined that the death had been preventable, such as through earlier access
159 to care, earlier escalation of care, avoiding inappropriate discharge, or via infection prevention
160 measures during a prior hospitalization.

161 **Table 1**: Demographics of the deceased, case characteristics, *Klebsiella pneumoniae*
162 sublineages and clonal groups detected among 7 infants and children who died of suspected
163 *Klebsiella pneumoniae* acute lower respiratory tract infection in the community in Lusaka, Zambia.
164 Kp, *Klebsiella pneumoniae*; Kqs, *Klebsiella quasipneumoniae* subsp. *similipneumoniae*.

Case	Age	Sex	Month and year	Location of death	Cough/ severe cough	Non-respiratory	Species	Sub lineage	Clonal Group

			of death		(duration in days)	symptoms			
2337	2 weeks	Female	Dec 2022	En route to a facility	Yes (1)	Fever	Kp	SL607	CG607
4803	7 months	Male	Dec 2022	At home	Yes (2)	Fever, vomiting, diarrhea	Kp	SL607	CG607
4771	8 months	Male	Oct 2022	Waiting to be seen by a doctor	Yes (2)	Vomiting, diarrhea	Kqs	SL10072	CG10344
4796	17 months	Male	Nov 2022	At home	Yes (1)	Fever	Kp	SL17	CG1123
4201	7 days	Female	Feb 2020	En route to hospital or other health facility	Yes (2)	Fever, vomiting, diarrhea	Kp	SL17	CG10072
4220	18 weeks	Male	Jun 2020	En route to hospital or other health facility	Yes (7)	Fever	Kp	SL280	CG280
4774	20 months	Male	Oct 2022	Waiting to be seen by a doctor	Yes (1)	Fever	Kp	SL37	CG3648

Histopathology findings varied across participants as well as across the 6 biopsies obtained from each of the deceased, from interstitial expansion to histologic pneumonia. Most of the 7 individuals included in this study had Kp detected by FTD Respiratory Pathogens 33 multiplex PCR in 5 or 6 out of all 6 biopsies sent for molecular testing (**Supp. Table 3**). Most cases had

170 other pathogens detected by the multiplex PCR on one or more biopsies; based on the totality of
171 evidence (including number of detecting biopsies, qPCR cycle thresholds, host and pathogen
172 characteristics), these were considered not to be the primary driver of the infection. Reanalysis of
173 the 11 biopsy samples using the highly Kp-specific ZKIR-qPCR (29, 30) confirmed the presence
174 of Kp in all samples, ranging from 45.5 to 6586.4 genome copies/ng DNA (median 147.6) (**Supp.**
175 **Table 1**). A strong correlation was also observed between the ZKIR-qPCR quantified abundances
176 and the FTD Respiratory Pathogens 33 multiplex PCR Kp Cq values (Pearson's r = -0.94, $p <$
177 0.001, **Supp. Fig 1**), demonstrating high concordance between these two qPCR assays.

178

179 **Hybridization-capture of *K. pneumoniae* by targeted RNA-probe metagenomics**

180 We performed in-depth genomic analysis of Kp lineages from lung biopsy samples by employing
181 RNA-probe hybridization-capture metagenomics targeting Kp core and accessory genes (see
182 Methods). This targeted Kp hybridization-capture was able to overcome considerable residual
183 human DNA contamination.

184 Short-read sequencing of the 11 target-enriched samples produced an average of 30.3 x 10^6 75
185 bp paired-end reads per sample (range: 5.7 x10^6 – 51.2 x10^6). Following removal of human DNA,
186 an average of 2.4 x10^6 bacterial reads per sample remained (range: 6.4 x10^4 - 13.5 x10^6),
187 indicating considerable residual human DNA contamination in samples despite targeted Kp
188 hybridization-capture. Of these bacterial reads, an average of 57.8% were assigned to Kp (range:
189 30.1% - 89.6%), with the number of on-target Kp reads correlating strongly with the abundance
190 of Kp in samples pre-hybridization-capture (Pearson's r = 0.91, p < 0.001, **Supp. Fig 2**) (**Supp.**
191 **Table 4**). With these limitations noted, our analysis was able to detect key genomic features of
192 the Kp lineages, including sublineage (SL), clonal group (CG), capsule serotype, as well as
193 acquired AMR genes and virulence determinants (**Table 2**).

194 **Table 2.** Key genomic features of *Klebsiella pneumoniae* sequenced from post-mortem lung

195 biopsies from 7 infants and children who died of suspected *Klebsiella pneumoniae* acute lower

196 respiratory infection in Lusaka, Zambia

197

	Phylogenetic Classification			Capsule	Acquired Antimicrobial Resistance										Acquired Virulence	
Sample	**Species**	**SL**	**CG**	**K-locus**	**AGly**	**Flq**	**Phe**	**Rif**	**Sul**	**Tet**	**Tmt**	**Bla**	**Bla_ESBL**	**Bla_chr**	**Ybt**	**Iuc**
2337-V2A	Kp	SL607	CG607	KL25	AAC(3)-IIe, AAC(6')-Ib-cr4*, AAC(6')-Iz , aadA16, aadA5, APH(6)-Id	AAC(6')-Ib-cr4*, QnrB17	-	arr-3	sul2	tetD	dfrA27	TEM-239	CTX-M-15	SHV-1	ybt 14; ICEKp12	
4803-V1A	Kp	SL607	CG607	KL25	AAC(3)-IIe, AAC(6')-Ib-cr4*, aadA16, aadA5, APH(6)-Id	AAC(6')-Ib-cr4*, QnrB17	-	arr-3	sul2	tetA, tetD	dfrA14, dfrA27	TEM-239	CTX-M-15	SHV-1	ybt 14; ICEKp12	iuc 5
4803-V2B	Kp	SL607	CG607	KL55	AAC(3)-IIe, AAC(3)-IId, AAC(6')-Ib-cr4*, aadA16, aadA5, APH(6)-Id	AAC(6')-Ib-cr4*	-	-	sul2	tetA, tetB	dfrA17	TEM-239, OXA-1	CTX-M-55	-	-	iuc 5
4201-V2B	Kp	SL17	CG10098	KL128	APH(6)-Id	-	-	-	sul2	tetB	dfrA8, dfrA14	TEM-239	-	SHV-155	-	-
4201-V2C	Kp	SL17	CG10098	KL128	aadA, aadA5, APH(6)-Id	-	-	-	sul2	tetB	dfrA8, dfrA14	TEM-239	-	SHV-97	-	-
4796-V2C	Kp	SL17	CG1123	KL122	AAC(6')-Ib-cr4*, aadA16, APH(6)-Id	AAC(6')-Ib-cr4*, QnrB6	-	arr-3	sul2	tetD	dfrA27	-	TEM-116	SHV-155	-	-
4771-V2A	Kqs.	10072	CG10344	KL46	aadA, APH(6)-Id, AAC(6')-Ib7	-	-	-	sul2	tetB, tetD	dfrA14	SHV-217, TEM-239	-	OKP-B-22	-	-
4771-V2B	Kqs.	10072	CG10344	KL46	APH(6)-Id	-	-	-	sul2	-	dfrA14	-	-	-	-	-
4774-V1A	Kp	SL37	CG3648	KL23	AAC(6')-Ib-cr4*, AAC(3)-IId	AAC(6')-Ib-cr4*	-	-	-	-	-	-	-	SHV-85	-	-
4774-V1C	Kp	SL37	CG3648	-	AAC(6')-Ib-cr4*	AAC(6')-Ib-cr4*	-	-	-	-	-	-	-	-	-	-

4220-V1A	Kp	SL280	CG280	KL23	AAC(3)-IIe, AAC(6')-Ib-cr4*, APH(3")-Ib, APH(6)-Id	AAC(6')-Ib-cr4*, QnrB1	atA1, catB3	-	-	tetA, tetD	dfrA14	OXA-1, TEM-239	CTX-M-15	SHV-27	-	-

Kp = *Klebsiella pneumoniae*, Kqs = *Klebsiella quasipneumoniae* subspecies *similipneumoniae*, SL = sublineage, CG = clonal group, AGly = aminoglycoside, Flq = fluoroquinolone, Phe = phenicols, Rif = rifampicin, Sul = sulfomamides, Tet = tetracycline, Tmt = trimethoprim, Bla = β-lactamase, Bla_ESBL = extrended spectrum β-lactamase, Bla_chr = chromosomal β-lactamase, Ybt = yersiniabactin, Iuc = aerobactin. *bifunctional ARG: catalyzes acetylation of both fluoroquinolones and aminoglycosides

Determination of *K. pneumoniae* cgMLST and MLST from enriched samples.

Targeted RNA-probe metagenomics allowed for sequence typing according to current conventions. SL and CG analysis of Kp from the target enriched samples was performed using the previously described 629-gene core genome multilocus sequence typing (cgMLST) scheme and dual bar-coding system (31, 32). A median of 627 out of 629 cgMLST genes were detected per sample (IQR 625.5 - 628.5) (**Supp. Table 5**). When profiled against the Kp cgMLST profile database available through BIGSdb (https://bigsdb.pasteur.fr/klebsiella/), five SLs and six CGs were identified from the 11 enriched metagenomic samples: SL607 (CG607), SL17 (CG10098), SL17 (CG1123), SL37 (CG3648), SL10072 (CG10344), and SL280 (CG280) (**Table 1**). cgMLST profiles from enriched samples matched to cgMLST profiles within the database between 159 to 615 out of the 629 allelic positions (median 564) (**Supp. Table 5**). A strong correlation was observed between the number of matching alleles and on-target Kp reads in samples post-hybridization-capture (Spearman's rho = 0.95, $p < 0.001$) (**Supp. Fig 3**), suggesting this was a major factor determining correct allele detection. All paired samples from the same case had identical SLs and CGs identified, supporting accurate cgMLST reconstruction. Additionally, all identified SLs corresponded to *K. pneumoniae sensu stricto*, except SL10072 (case 4771), which was identified as *Klebsiella quasipneumoniae* subspecies *similipneumoniae* (Kqs). Notably, the

222 SL607 cgMLST allele profiles detected in samples 4803-V1A and 2237-V2A, which were taken
223 during the same month from different individuals, were perfectly concordant at 621 out of 629
224 detected cgMLST allelic positions, indicating a very high degree of phylogenetic relatedness
225 between the Kp strains present in these two samples (**Supp. Table 6**).

226 No cgMLST genes were detected in the negative control sample above our cut-off limit (≥80%
227 gene coverage with ≥90% sequence identity). This confirmed no significant Kp DNA
228 contamination of samples occurred during the hybridization-capture procedure, either through
229 background laboratory contamination or through cross-contamination between samples, giving
230 confidence in our cgMLST gene detection.

231 To complement our cgMLST analysis, reconstruction of sequence types (STs) using the
232 previously described seven-gene Kp MLST scheme was also performed (33). Although this
233 approach produced less consistent ST matches between sample pairs, the closest matching
234 MLST allele profile was still concordant with our cgMLST analysis in 7 out of 11 samples (**Supp.**
235 **Table 7**). Notably, 6 out of the 11 samples detected either 6 or 7 out of 7 MLST genes with 100%
236 gene coverage with 100% sequence identity. All these alleles matched their expected MLST
237 profiles positions based on our cgMLST analysis, except samples 4796-V2C and 4771-V1A which
238 had single allele mismatches. Like our cgMLST analysis, the number of detected MLST alleles
239 that matched with the expected MLST allele profile correlated strongly with the number of on-
240 target Kp reads in samples post-hybridization-capture (Spearman's rho = 0.86, $p < 0.001$, **Supp.**
241 **Fig 4**). Moreover, all samples that produced over 800,000 on-target Kp reads following
242 hybridization-capture produced an MLST profile matching that predicted by cgMLST analysis.
243 Thus, although less precise than cgMLST reconstruction, MLST analysis was able to produce
244 accurate ST profiles in the majority of samples, supporting our cgMLST analysis. The strong

245 dependence on the number of on-target Kp reads indicated MLST analysis is best suited to
246 samples that produce a higher yield of reads from the targeted hybridization-capture method.

247

248 **Capsule type (K-locus) prediction**

249 Prediction of the polysaccharide capsule synthesis locus (K-locus), a key potential target for future
250 development of Kp vaccines, was successfully performed using hybridization-capture. K-locus
251 prediction was performed by aligning reads to alleles corresponding to the K locus marker genes
252 *wzi*, *wzy*, and *wzx*, using allelic sequences from the Kaptive database (34). Using this approach,
253 K-locus marker genes were detected in all samples except 4774-V1C (**Table 2**). The K-loci
254 predicted from these marker genes were concordant between both paired samples from cases
255 4771 (SL10072, KL46) and 4201 (SL17, KL128), while unpaired samples 4220–V1A (SL280),
256 4796-V2C (SL17), and 2337-V2A (SL607) predicted K-loci KL23, KL122, and KL25, respectively.
257 A mismatch occurred between paired samples from case 4803 (SL607), with sample 4803-V1A
258 predicting KL25, in agreement with SL607 and KL25 detected in sample 2337-V2A, while 4803-
259 V2B predicted KL55. Notably, samples 4803-V2B and 4774-V1C had the two lowest number of
260 Kp-assigned reads of all samples, both less than 100,000 reads, which may suggest these were
261 below the detection threshold for accurate K-locus detection by this method.

262

263 **Acquired antimicrobial resistance genes**

264 We next analyzed samples for the presence of antimicrobial resistance genes (ARGs) known to
265 be acquired by Kp, which constrain effective options for clinical treatment. A high prevalence of
266 acquired ARGs was observed in all samples, with a median of 9 acquired ARGs from a median
267 of 5 resistance classes detected per sample (**Table 2**). A strong concordance was observed in

268 ARG profiles detected between paired samples from the same case. SL607 from samples 4803-
269 V1A, 4803-V2B, and 2337-V2A, as well as SL280 from sample 4220-V1A, were associated with
270 the highest number of acquired ARGs (n = 13, 13, 12, and 12, respectively). ARGs encoding
271 acquired resistance to aminoglycosides, trimethoprim, β-lactams, tetracyclines, and
272 sulphonamides were highly prevalent, detected in n = 11, 9, 8, 8, and 8 samples, respectively.
273 Transferable fluoroquinolone resistance determinants were also detected in seven samples;
274 *AAC(6')-Ib-cr4* plus *Qnr* in four samples and *AAC(6')-Ib-cr4* only in three samples.

275 Within the β-lactamase class, the globally disseminated extended-spectrum β-lactamase (ESBL),
276 $bla_{\text{CTX-M-15}}$, was detected in three samples; 2337-V2A and 4803-V1A associated with SL607, and
277 4220-V1A associated with SL280. Notably, the other sample from case 4803 associated with
278 SL607 (4803-V2B) had a different CTX-M allele detected, $bla_{\text{CTX-M-55}}$, possibly a misidentification
279 due to low recovered reads within this sample. Another ESBL, $bla_{\text{TEM-116}}$, was detected in sample
280 4796-V2C associated with SL17. The additional acquired broad-spectrum β-lactamases, $bla_{\text{TEM-}}$
281 $_{239}$ and $bla_{\text{OXA-1}}$, were detected in 7 and 2 samples, respectively.

282 The vast majority of Kp and Kqs lineages carry the intrinsic broad-spectrum β-lactamases bla_{SHV}
283 and bla_{OKP}, respectively, providing resistance to amoxicillin/ampicillin (35). This is of high clinical
284 importance, as amoxicillin is the empiric therapy for outpatient community acquired pneumonia in
285 Lusaka, Zambia (18). These β-lactamases were detected in all samples containing Kp (bla_{SHV}) or
286 Kqs (bla_{OKP}) except 4774-V1C, 4803-V2B, and 4771-V2B. These samples had the lowest, second
287 lowest, and fifth lowest number of on-target Kp reads recovered (median 62,020 reads), providing
288 further support for a minimum threshold of approximately 100,000 reads for reliable gene
289 detection. Thus, the Kp lineages associated with these cases of fatal community-acquired LRTIs
290 were highly resistant, with most harboring multiple ARGs affecting clinically important drug
291 classes.

292

293 **Acquired virulence determinants**

294 Analysis of known acquired virulence loci amongst the identified Kp lineages revealed the
295 presence of virulence factors in two cases. This analysis showed that Kp sublineage SL607 from
296 samples 4803-V1A and 2337-V2A were associated with the acquired siderophore yersiniabactin
297 (*ybt*), specifically lineage *ybt* 14 which is associated with the ICEKp12 mobile element (**Table 2**).
298 Additionally, nine out of eleven alleles matching this *ybt* lineage were also detected in the other
299 sample from case 4803 (4803-V2B) at ≥ 80% coverage and ≥ 99% sequence identity, highly
300 suggestive of presence within this sample. None of these alleles, however, reached our minimum
301 threshold of 50x coverage depth required for reliable detection. Similarly, samples 4774-V1A and
302 4774-V1C associated with SL37 had 8 and 6 of the 11 alleles matching the *ybt 9* lineage
303 associated with ICEKp3, detected at ≥ 80% coverage and ≥ 95% sequence identity, suggestive
304 of the presence of this locus within these samples, however these also did not meet our minimum
305 threshold of coverage depth.

306 Both samples from case 4803 associated with SL607 also had the acquired siderophore
307 aerobactin (*iuc*), specifically lineage *iuc5,* detected above the minimum identity and coverage
308 thresholds. The other major known virulence loci which can be acquired by Kp (siderophore
309 salmochelin, hypermucoidy locus *rmp*, and genotoxin colibactin) were not detected in any
310 samples.

311 Detection of possible convergence of MDR and hypervirulence in sublineage SL607 in case 4803
312 by the acquisition of *iuc5* is a potentially important finding with significant potential public health
313 implications. It was therefore important to investigate whether *iuc5* detected in this case may have
314 been due to contamination with *Escherichia coli*, which is also a known carrier of this *iuc* lineage.
315 Unique reads were assigned to *E. coli* in all 11 samples, ranging from 161 to 30896 (median:
316 13637, IQR: 3791 – 23003), suggesting either spurious hits during taxonomy profiling or low-level
317 background contamination (**Supp. Table 8**). Notably, the two samples with *iuc5* detected, 4803-
318 V1A and 4803-V2B, contained among the highest levels of *E. coli* reads detected, with 25012 and
319 22469 reads, respectively, ranking second and fourth highest overall. Thus, although detection of
320 *iuc5* in case 4803 may suggest acquisition by Kp, resulting in a higher propensity to cause
321 invasive infection, spurious detection from contaminating *E. coli* could not be ruled out.

322

323 **Plasmid content**

324 Identification of multiple MDR-associated plasmid replicon markers by hybridization-capture
325 provided insights into potential mechanisms of ARG acquisition and dissemination among Kp
326 lineages. Analysis of plasmid content associated with Kp lineages identified seven different
327 plasmid replicon types, with a median of four detected per sample (range 0 – 8) (**Supp. Table 9**).
328 Detected plasmid replicon families included IncFIA, IncFIB, IncFII, IncI, IncN, IncR, and Col, with
329 similar replicons detected between paired samples. Of these, the commonly Kp- and MDR-
330 associated replicons Col(pHAD28), IncFIB(K), IncFIA(HI1), and IncR were most prevalent,
331 detected in 9, 5, 5, and 5 samples, respectively. Other detected plasmid replicons associated with
332 high-risk Kp clones and MDR carriage included IncFII(K), IncFII(pKP91), IncFII(pHN7A8),
333 IncColRNAI, and IncN (28, 36, 37).

334

335 **Analysis of *K. pneumoniae* lineages from assembled contigs**

336 To complement our read-based analysis, reads were assembled into longer contiguous
337 sequences (contigs) and Kp lineages were analyzed for MLST, serotyping, acquired ARGs, and
338 virulence determinants. Although contig-based analysis can have reduced detection sensitivity,
339 reconstructing entire gene sequences can offer improved precision and confidence of results
340 compared to read-based analysis alone. This analysis resulted in complete MLST profiles,
341 containing perfect allele matches for all seven genes, in four samples (2337-V2A, 4803-V2B,
342 4796-V2C, and 4220-V1A) (**Supp. Table 10**). All these samples identified the same ST lineage
343 as our read-based cgMLST and MLST analysis (ST607, ST607, ST17, and ST280, respectively).
344 Of the remaining samples, all alleles with a 100% identical match to the MLST database produced
345 the expected allele for that position based on our previous read-based analysis, except for one
346 (*pgi* in sample 4771-V2B).

347 K-locus prediction from contigs was also highly concordant with our read-based analysis, with the
348 best match K-locus agreeing with the K-locus detected from reads in almost all samples. Notably,
349 the K-locus predicted in SL607 from case 4803 that had a mismatch between sample pairs from
350 read-based analysis was concordant when performed on contigs. Both samples from this case
351 predicted KL25, making it also concordant with SL607 detected in sample 2337-V2A. Additionally,
352 paired samples 4774-V1A that detected KL23 and 4774-V1C which did not detect a K-locus from
353 read-based analysis had a mismatch from contig-based analysis, predicting KL15 and KL51,
354 respectively.

355 Analysis of acquired ARGs also produced highly concordant ARG profiles to our read-based
356 analysis, with strong agreement between sample pairs. Similar to our read-based analysis, this
357 demonstrated the presence of genes associated with resistance to β-lactams, aminoglycosides,
358 trimethoprim, sulphonamides, and tetracycline was widespread among samples, as well as

359 resistance to fluoroquinolones, MLS, and chloramphenicol in the majority of samples. Within the
360 β-lactamase class, all ESBLs detected from contig analysis were concordant with our previous
361 read-based analysis, with an additional detection of $bla_{CTX\text{-}M\text{-}14}$ and $bla_{CTX\text{-}M\text{-}15}$ in sample 4774-V1A
362 which were undetected from reads.

363 Finally, analysis of acquired virulence from contigs also showed good agreement with read-based
364 detection, with samples 2337-V2A and 4803-V1A associated with SL607 matching *ybt14* at 11
365 out of 11 allelic positions, while sample 4803-V1A also matched *iuc5* at four out of five allelic
366 positions. Additionally, the acquired siderophore salmochelin (*iro*), lineage *iro5*, was also detected
367 in sample 4803-V1A with two out of four allelic matches.

368

369 **Discussion**

370 This analysis of a series of Kp pneumonia cases among children who died in the community in
371 Lusaka, Zambia, revealed several previously identified Kp genetic lineages that had been
372 observed in NICU settings and other environments. Notably, the sequencing revealed high rates
373 of AMR, global problem clones SL17 and SL37 (36, 38), multiple virulence factors, and multiple
374 capsular types. Two of the cases from the same month were both found to have SL607
375 sublineages that were highly phylogenetically similar, suggestive of clonal spread. Given that the
376 majority of the deceased had no apparent contact with the hospital setting, the most parsimonious
377 explanation for this cluster is that they represent spillover of hospital isolates that are now
378 circulating outside the facility and causing community acquired infections. This has significant
379 implications as it reinforces observations that the epidemiology of Kp has evolved and is no longer
380 exclusively a hospital acquired pathogen. Given the lack of AMR testing, and the ubiquitous use
381 of amoxicillin as first line therapy for pneumonia in community settings such as Lusaka, these
382 findings should be viewed with significant alarm.

383 RNA-probe hybridization-capture metagenomic sequencing adds to evidence generated by our
384 prior MITS study implicating Kp ALRI as the cause of deaths in these infants and children. While
385 detection of Kp does not equate to invasive infection or cause of death, the histopathologic
386 evidence and PCR data across multiple (six) samples adjudicated by an expert panel strongly
387 supported Kp as the causative pathogen in these deaths. These findings, though limited by a
388 small sample size, add to mounting evidence of a community burden of severe Kp, including as
389 a cause of ALRI, in parts of sub-Saharan Africa and other LMIC settings. Of note, this study did
390 not collect sufficient information to prove that the site of acquisition occurred in the community;
391 narrative verbal autopsies did not mention any such hospital contact in the majority of cases.

392 If reproduced within larger studies, the implications of a shift in the epidemiology of Kp towards
393 causing severe and fatal pediatric ALRI in the community should be viewed with extreme concern;
394 this is especially true in resource-limited settings, where ALRI treatment is algorithmic and
395 empiric, and almost never guided by pathogen diagnostics, let alone antimicrobial resistance
396 testing. Current WHO guidelines for childhood pneumonia treatment at health facilities
397 recommend amoxicillin for pneumonia without warning signs and ampicillin with gentamicin in
398 severe pneumonia (18). Due to the chromosomally encoded broad-spectrum β-lactamase bla_{SHV},
399 Kp is intrinsically resistant to amoxicillin/ampicillin (35). Among the 7 Kp ALRI deaths we
400 investigated, aminoglycoside resistance genes were universal. Extended-spectrum β-lactamase,
401 fluoroquinolone, and trimethoprim-sulphamethoxazole resistance determinants were also
402 prevalent in our study. This leaves the carbapenem class, which is rarely available even in tertiary
403 care settings in Lusaka, although notably, acquired resistance to this class has been detected
404 recently within Kp isolated from Zambian hospitals (39, 40).

405 Several genomic findings bear highlighting. While most of the sublineages identified between
406 children were distinct, two of the cases from the same month contained highly phylogenetically
407 similar SL607 sublineages which shared 621 out of 629 cgMLST alleles, likely representing clonal

408 spread. The biopsies associated with these related sublineages came from a neonate of 2 weeks
409 and a 7-month-old infant, whose verbal autopsy described a hospital admission the previous
410 week. Further information that could point to a common transmission source is not available. In
411 contrast, the SL17 sublineages detected from two separate cases were less closely
412 phylogenetically related, belonging to two different CGs, and thus were unlikely to represent a
413 direct epidemiologic link.

414 Regarding the specific SLs identified in this study, SL607, SL17, SL10072, SL37, and SL280,
415 each has been previously identified as known MDR-associated sublineages, and most have been
416 reported to be important causes of hospital-acquired infections and outbreaks among neonates
417 (27, 38, 41, 42). SL17 in particular is recognized as a MDR global problem clone that is prevalent
418 throughout sub-Saharan Africa and has been reported as a cause of neonatal infections both
419 within hospitals and the community (38, 43). SL37 is also a global problem clone that was
420 identified among the most prominent Kp lineages responsible for neonatal sepsis in low- and
421 middle-income countries within the BARNARDS study (23, 36). One detected sublineage,
422 SL10072, was additionally found to belong to the species *K. quasipneumoniae subsp.*
423 *similipneumoniae*, and encoded MDR. *K. quasipneumoniae* is increasingly recognized as a cause
424 of invasive infection capable of acquiring MDR, including in sub-Saharan Africa (44, 45),
425 supporting a potential need for increased surveillance of this species.

426 All SLs identified in this study, with the exception of SL10072, were also identified in a recent
427 hospital-based study of Kp isolates from Malawi (27). Among these overlapping SLs, all except
428 SL17 from case 4201 had matching K-loci with those predicted in our study and exhibited highly
429 similar profiles of acquired ARGs. These findings suggest that the lineages identified in our cohort
430 may form part of a broader circulating Kp population spanning both community and hospital
431 settings across sub-Saharan Africa. Notably however, only a single sublineage (SL280)
432 overlapped with a recent hospital-based study of Kp isolates from Zambia, which also had a

433 matching K-locus (KL23) (46). While this limited overlap could indicate differences between
434 community- and hospital-associated Kp populations within Zambia, it may also reflect temporal
435 differences in sampling between studies, with several years separating sample collection resulting
436 in shifts in the circulating lineage composition.

437 Another key finding was the detection of the acquired siderophores *ybt14* associated with SL607
438 in cases 2337 and 4803, as well as *iuc5* in case 4803. Analysis of all ST607 isolates available
439 through PathogenWatch via Kleborate (n = 39, downloaded July 2025) (47) demonstrated that
440 66.7% of these were associated with carriage of *ybt14*, consistent with our findings. Notably,
441 however, none of these isolates were associated with *iuc5* carriage. Acquisition of these
442 siderophores may suggest convergence of MDR genes and hypervirulence factors within this
443 lineage. Converged MDR-hypervirulent lineages may cause highly aggressive infections with
444 limited treatment options that many primary and secondary health settings in limited-resource
445 settings are unequipped to address. This finding therefore places the community death associated
446 with this case within a well-documented and concerning trend of high mortality related to
447 converged MDR-hypervirulent lineages of Kp, which is being observed globally, including other
448 countries across sub-Saharan Africa (27, 48).

449 Serotyping of lineages causing community-acquired pneumonia may be an important contribution
450 to choosing capsule targets for vaccine development, after further study in larger populations (49).
451 Five different K-loci were identified within this study, three of which (KL25, KL23, and KL122)
452 were present among the 30 most prevalent identified globally as potential vaccine targets in a
453 recent metanalysis of 1930 Kp neonatal blood isolates from 13 countries, with KL25 ranking as
454 the third most prevalent (49). Although our study consisted of only a small sample size, this
455 nevertheless suggests these K loci may be important targets for vaccine development within the
456 community as well as the hospital setting. There is interest in development of Kp vaccines for the
457 prevention of neonatal sepsis through maternal vaccination as well as for other high-risk groups

458 (50). Importantly, however, assuming a duration of effect of a maternal vaccination candidate of
459 three months, only 2 of the 7 infants and children would have been protected by a hypothetical
460 maternal vaccine in this small study of ALRI deaths in the community. We believe our findings
461 call for larger and more systematic studies of Kp causing severe community ALRI in infants and
462 children to inform vaccine target profiles and policy.

463 Plasmids are major vehicles through which Kp acquires and disseminates ARGs and virulence
464 determinants via horizontal gene transfer. MDR-carrying plasmids can circulate among Kp
465 lineages and even transfer to other species, facilitating ARG dissemination and contributing to
466 AMR spread and difficult-to-control outbreaks (51, 52). In this study, several clinically important
467 plasmid replicons were detected, including FIB(K), IncFII(K), Inc(R), ColRNAI, IncFII(pKP91),
468 IncFIAHI1, IncN, IncFII(pHN7A8), and Col(pHAD28). These replicons are frequently associated
469 with MDR plasmids and high-risk global Kp clones (28, 36, 37). Although detection of the same
470 plasmid replicon across multiple Kp sublineages suggests possible horizontal gene transfer
471 between these lineages, further investigation using long-read sequencing would be needed to
472 confirm this.

473 Several data points support the validity of the genomic analysis generated through this novel
474 application of RNA-probe hybridization-capture of Kp from lung tissue samples. Kp C_q values from
475 the FTD-32 multiplex PCR correlated highly to a singleplex Kp qPCR and to the on-target Kp
476 reads from RNA-probe hybridization-capture sequencing. We observed high concordance in
477 detected SLs, CGs, K-loci, ARGs, and acquired virulence determinants between paired samples
478 obtained from the same deceased individual. Previously identified SL and K-locus associations
479 were detected, including the association between SL280 and KL23, SL607 and KL25, SL17 and
480 KL122, as well as SL37 and KL23 (27, 41, 42, 46). Additionally, SLs from our study that
481 overlapped with strains reported in other recent sub-Saharan African studies in Zambia and
482 Malawi showed highly similar ARG profiles, including acquired aminoglycoside, β-lactamase,

483 fluoroquinolone, sulphonamide, tetracycline, and trimethoprim resistance genes (27, 46). Our
484 results thus demonstrate this method allows detailed analysis of genetic epidemiology including
485 STs (MLST and cgMLST), serotyping, acquired AMR and virulence of Kp lineages from tissue
486 samples. RNA-probe hybridization-capture may be useful for other settings, including post-
487 mortem surveillance studies, where genetic epidemiology investigation is warranted but cultures
488 are not available. The costs, probe-target development needs, and inability to perform phenotypic
489 testing, however, may limit its broader applicability.

490 This study was not without limitations. First, the small sample size limits the generalizability of the
491 sequencing findings. Larger, representative studies incorporating genetic analyses of Kp causing
492 severe pediatric ALRI in the community are urgently needed. Second, determining the pathogenic
493 cause in ALRI is complex and the sequencing findings in our study do not remove uncertainty
494 from the true cause of death for the infants and children in this study. To incorporate multiple
495 modalities of evidence into this determination, we relied on the collective clinical reasoning of a
496 DeCoDe panel of experts. The biopsies that underwent Kp sequencing in this study had multiple
497 pieces of evidence in support, first, of a respiratory cause of death and, second, of Kp as the
498 causative pathogen. However, some cases contained multiple conflicting pieces of evidence, for
499 example one case (**Table 1**), where the same three virulent pathogens were detected on multiple
500 biopsies, leading the DeCoDe panel to give a designation of moderate certainty for the diagnosis
501 of Kp ALRI.

502 The RNA-probe hybridization-capture technique also had limitations. High human DNA
503 contamination despite hybridization-capture resulted in low Kp reads in several samples, which
504 limited the resolution of lineage typing and detection of ARGs. While sufficient for analysis, higher
505 coverage would have given confidence to some of the lower coverage hits. In addition, the method
506 may have benefited from a second round of hybridization-capture as with other high sensitivity
507 RNA-probe hybridization-capture approaches. Moreover, while the acquired ARGs and virulence

508 determinants we detected frequently associate with Kp, especially within sub-Saharan Africa, we
509 could not definitively attribute these to Kp over other species potentially present at low
510 abundances within this study. For example, *iuc5* which was detected in case 4803, is also
511 commonly associated with *E. coli*, of which reads were detected in all samples. Although it is
512 possible these *E. coli* reads may have arisen from misclassification during taxonomy profiling,
513 due to shared homology between Kp and *E. coli*, they may also suggest background
514 contamination, possibly due to increased gut barrier permeability post-mortem. Given these
515 uncertainties, future culture-based isolation and whole genome sequencing studies would be
516 required to confirm the association of these genes with the detected Kp lineages.

517

518 **Conclusions**

519 This small study characterizing Kp genomes from Zambian children who died of suspected Kp
520 ALRI in the community warrants further study into the prevalence of severe and fatal Kp ALRI and
521 the underlying genomic epidemiology. This study investigated Kp sequenced from preserved lung
522 biopsies in 7 suspected Kp ALRI deaths identified through a post-mortem MITS study of under-5
523 community respiratory deaths in Lusaka, Zambia. ALRI remains the second most common cause
524 of under-5 mortality globally and growing evidence suggesting Kp as a contributing pathogen
525 raises concerns around current clinical management approaches. Across the samples, we
526 detected several distinct sublineages of global importance, genes encoding acquired resistance
527 to important antimicrobial classes in all cases and acquired virulence factors in two cases. Given
528 the critical implications for prevention and treatment of severe under-5 ALRI in the community,
529 additional, larger scale genomic epidemiology studies should be carried out in similar populations.

530

531 **Methods**

532 This study used a target hybridization-capture metagenomic approach to sequence Kp from
533 metagenomic DNA extracted from post-mortem lung biopsies of infants and children who died
534 outside of hospital settings in Lusaka, Zambia. The methods of the original post-mortem study
535 have been previously detailed (15). Briefly, enrollments occurred at the University Teaching
536 Hospital morgue, which serves the city of Lusaka. All next-of-kin of the deceased approached for
537 the study were offered grief counseling, and, if interested in study participation, provided informed
538 consent followed by a verbal autopsy. The body of the deceased was then brought to the autopsy
539 theater, the thorax was cleaned with ethanol and iodine, and six minimally invasive tissue samples
540 (MITS) were obtained from upper, middle, and lower lung zones from the left and right side along
541 the axillary line. Molecular testing and histopathology was performed on specimens from each of
542 the six lung zones sampled. The molecular testing performed was a 32-respiratory pathogen PCR
543 panel (FTD-33 kit, Fast-Track Diagnostics). For histopathology, lung samples were fixed in neutral
544 buffered formalin, followed by paraffin embedding, microtomy, hematoxylin and eosin staining,
545 and finally, reviewed by two experienced pathologists.

546 A panel adjudication approach modeled after the Child Health and Mortality Prevention
547 Surveillance Determination of Cause of Death (DeCoDe) Process was used to determine the
548 causes of death in this post-mortem study, based on the totality of subjective and objective data.
549 The panel convened for this study included 7 clinicians with infectious diseases and pediatrics
550 experience practicing in Lusaka. The panel met in person and reviewed case dossiers, including
551 representative photographs of the histopathology for each of the deceased, to achieve a
552 consensus determination of whether the death was due to respiratory or non-respiratory causes,
553 the cause of death, and whether the death was preventable or not. For the present sequencing
554 study, we selected 8 cases with the strongest molecular data (at least two biopsies with Kp
555 detected at qPCR cycle thresholds of less than 30) from among the 13 deceased infants and

556 children adjudicated to have died from Kp pneumonia. We attempted to perform target
557 hybridization-capture sequencing on two separate lung biopsy samples from each of these select
558 cases.

559

560 **DNA extraction, quality control, and *K. pneumoniae*-specific ZKIR-qPCR analysis**

561 DNA was extracted using the InviSorb® Spin Universal Kit (Invitek Diagnostics, Germany)
562 following the manufacturer's protocol for tissue biopsy samples in final elution volume of 60 μL
563 DNA concentration. The concentration and quality of the extracted DNA from each of the lung
564 biopsy samples was measured by Quibit 3.0 fluorometer using dsDNA Narrow Range Assay
565 (Thermo Fisher Scientific, Waltham, USA) and Nanodrop 1000 Spectrophotometer (Thermo
566 Fisher Scientific), respectively. All DNA samples which passed quality control assessment
567 underwent analysis using the highly Kp sensitive and specific ZKIR-qPCR assay as previously
568 described (29, 30). Briefly, 2.5 ng of 10 ng/μL DNA (25 ng total) was used as input for each 25 μL
569 qPCR reaction, with reaction mixture and cycling conditions as described (30). All samples were
570 assayed in triplicate using an Applied Biosystems 7500 Real-Time PCR System (Thermo Fisher
571 Scientific), and criteria used for positive Kp detection were as described previously. Kp genome
572 copy number was calculated from the standard curve using 7500 Real-Time PCR Analysis
573 Software v2.3 (Applied Biosystems, Life Technologies, Waltham, USA) as described (30).

574 Due to the very high detection sensitivity of the RNA-probe target hybridization-capture method
575 employed in this study, a blank negative control was used to account for background/cross-
576 contamination during the hybridization-capture process which consisted of DNA elution buffer
577 only. This blank sample had no detectable DNA by qubit analysis and was negative for Kp by the
578 ZKIR-qPCR.

579

580 **Library construction, targeted RNA probe hybridization-capture, and sequencing**

581 Hybridization-capture of Kp from lung biopsy metagenomic samples was performed by
582 SureSelect XT HS2 targeted RNA probe hybridisation-capture (Agilent, Santa Clara, USA). A
583 custom panel of 145,787 120bp RNA probes (total size: 2.635 Mbp) was designed targeting core
584 and acquired genes from the *Klebsiella pneumoniae* pan-genome, including cgMLST, MLST, K-
585 and O- loci, ARGs, virulence loci, and plasmid replicons (**Supp. Table 11**). Probes were designed
586 with at least 90% sequence homology to the reference sequences with 1–2x tiling density, by an
587 in-house developed algorithm by Agilent Technologies. For regions with ambiguous bases,
588 among all possible sequences, representative probes with a 90% base homology threshold were
589 selected. The algorithm reduces the number of relevant baits to ensure cost efficiency. The final
590 bait sets comprised baits (Design ID S3491653 and 61512) of 120bp RNA probes (total size:
591 2.635 Mbp).

592 Sequencing library construction and targeted RNA probe hybridization-capture of samples was
593 performed as per the SureSelect XT HS2 DNA Kit Library Preparation (+/- MBC)/Fast-Hyb Target
594 Hybridization-capture/Post-hybridization-capture Pooling Workflow protocol (Agilent; catalog nr
595 G9982A) with modifications based on the Utilization of Agilent SureSelect Target Hybridization-
596 capture for Whole Genome Sequencing of Viruses and Bacteria protocol (Agilent; catalog nr
597 PR7000-2035). Briefly, 150 - 200 ng DNA was used as input for library construction. Samples
598 with <150 ng DNA were bulked with human genomic DNA (Promega, Madison, USA). Following
599 enzymatic fragmentation, end repair and A-tailing, libraries underwent PCR amplification and
600 indexing, followed by purification on AMPure Beads. 1000 ng of DNA library were hybridized with
601 1:10 diluted probes for 90 mins according to the protocol. After hybridization, a hybridization-
602 capture step using Streptavidin beads, followed by several steps of washing were done according

603 to the protocol. A final post-hybridization-capture PCR was performed with 19 cycles for samples
604 with > 200 Kp genome copies/ng DNA and 21 cycles for samples with < 200 Kp genome copies/ng
605 DNA.

606 All target enriched samples were sequenced on the Illumina NextSeq 500 platform at 75 bp paired
607 end reads. Processing of FASTQ files and removal of residual human DNA was performed as
608 previously described (30).

609

610 **Analysis of Target-enriched samples**

611 **Taxonomy profiling**

612 Taxonomic profiling, including total bacterial read count as well as Kp read count and relative
613 abundance estimates were performed using Kraken2 v2.1.2 and Bracken v2.6.1 with
614 MiniKrakenDB_8GB v202003 (53, 54). Quantification of unique Kp and *E. coli* reads in samples
615 was performed using Centrifuge v1.0.4 using the standard p_compressed+h+v database (55).

616 **cgMLST and MLST analysis**

617 Kp lineage analysis was performed using both the previously described 7-gene multi-locus
618 sequence typing (MLST) scheme and the core genome multi-locus sequence typing (cgMLST)
619 scheme consisting of 629 core genome genes (32, 33). Both MLST and cgMLST alleles were
620 identified from sequenced reads using the aligner KMA v1.4.18 (56) against either the MLST or
621 cgMLST allelic nucleotide sequence databases downloaded from BIGsdb
622 (https://bigsdb.pasteur.fr/klebsiella/). Detected MLST/cgMLST alleles were filtered to include only

623 those with ≥80% gene coverage and ≥90% sequence identity. The allele corresponding to each

624 MLST/cgMLST gene detected above this threshold that had the highest ConClave score
625 (accumulated alignment score) was then used to construct the final Kp MLST/cgMLST profile for
626 each sample. The resulting detected MLST/cgMLST profiles were compared to the Kp
627 MLST/cgMLST allelic profile database available at BIGSdb to identify the closest matching
628 ST/cgST lineage.

629 **K-locus prediction**

630 Within our custom RNA-probe panel were probes targeting the K-locus genes *wzi*, *wzx*, and *wzy*.
631 To predict the Kp K-locus from our samples we used KMA v1.4.18 (56) to align sequence reads
632 to the K-locus gene nucleotide sequence database downloaded from Kaptive v3.0.0 (34) using
633 the command 'kaptive extract kpsc_k --ffn'. Detected alleles were filtered to include only those
634 with ≥80% gene coverage and ≥90% sequence identity. The allele for each K-locus gene with the
635 highest ConClave score was then retained. K-locus was called if at least two out of the three
636 detected K-locus genes (*wzi*, *wzx*, and *wzy*) corresponded to the same K-locus.

637 **Detection of acquired ARGs, virulence loci, and plasmid replicons**

638 Detection of ARGs, virulence loci, and plasmid replicons was performed using the read alignment
639 tool KMA v1.4.18 (56). For ARGs, reads were aligned to the CARD v3.0.8 database. For virulence
640 loci, allelic sequences were downloaded from BIGSdb (https://bigsdb.pasteur.fr/klebsiella/) and
641 ST profiles for yersinabactin (ybt), colibactin (clb), aerobactin (iuc), salmochelin (iro) and rmpADC

642 were analyzed using schemes described previously (57-59). For detection of plasmid replicon
643 sequences, the PlasmidFinder database was used. Presence of a gene was called if it was
644 detected with ≥ 90% sequence identity, ≥ 80% gene coverage, and ≥ 50x coverage depth.
645 Detection of an acquired virulence loci was called if ≥ 50% of the expected alleles for that loci
646 were detected above the minimum thresholds. Lineage was reported if ≥ 50% of the detected
647 alleles matched with an allelic profile when queried against the BIGSdb database.

648 **Contig-based analysis**

649 Paired-end reads were assembled into contigs using MegaHit v1.2.9 and --presets meta-sensitive
650 parameter. Assembled contigs were analyzed for ST, acquired virulence loci and ARGs via
651 Kleborate v3.1.2 (19) and capsule and O-antigen loci via Kaptive v3.0.0 (34) using default
652 parameters.

653 **Statistical analysis and data visualization**

654 Statistical analysis and data visualization was performed in R v4.5.0 using RStudio
655 2025.05.01+513.

656 **Computational resources**

657 Bioinformatic analysis of target enriched metagenomic samples was performed on the Norwegian
658 academic high-performance computing and storage services maintained by Sigma2 Norwegian
659 Research Infrastructure (NRIS) (http://www.sigma2.no/) (Sigma2 project number: nn9794k).

660 **Data availability**

Raw Illumina sequence reads from target enriched samples (human DNA removed) are available on ENA, Bioproject PRJEB105599.

Disclosure Statement

All authors report no conflict of interest.

Ethics statement

Ethical oversight for the ZPRIME and LCIS studies were provided by institutional review boards at Boston University and the University of Zambia. Written informed consent was obtained from the deceased's family members or representatives.

Acknowledgements

The authors thank Sylvain Brisse (Institut Pasteur, France), Kelly Wyres (Monash University, Australia), Erik Hjerde (UiT The Arctic University of Norway), and Ryan Wick (The University of Melbourne, Australia) for assistance in RNA probe target selection and design. We thank Even Østli and his team Agilent for their assistance in RNA-probe production and technical support with the hybridization-capture. We thank Ruth Paulsson and Hagar Taman (the Genome Sequencing Support Centre, Norway) for their assistance with library construction and sequencing. We also thank Tom Stanton (Monash University, Australia) for his advice when analysing Kp K-loci.

Funding sources

The ZPRIME and LCIS studies, DeCoDe process, and RNA probe design and hybridization-capture reagent kit costs were funded by the Bill and Melinda Gates Foundation, grant number OPP1163027. The salary of KL, target enrichment library construction, and sequencing costs were funded by the Trond Mohn Foundation, grant number TMF2019TMT03 and the Northern

682 Norway Regional Health Authority, grant numbers HNF1756-25 and HNF1589-21. AW received
683 salary support through National Institutes of Health 2T32HL125232-07.

684

685 **Author contributions**

686 The study was conceptualized by CG, AS, KH, AW, KL, and MS. Funding acquisition was
687 performed by CG, AS, and KL. Participant recruitment, grief counseling, and sample acquisition
688 were performed by BN, CC, SC, LP, DN, and JL. Adjudications were performed by NN, EJM, MK,
689 CM, BM, LN, VM, SWS, and coordinated by MC. Lung biopsy DNA sample preparation and
690 shipping was organized by AW and MS. RNA-probe synthesis was coordinated by KL and AW.
691 RNA-probe hybridization capture of samples was performed by WK. DNA sequencing and library
692 preparation was arranged by KL and WK. Data analysis was performed by KL with clinical
693 summaries compiled by AW. KH and AS provided expertise on *Klebsiella pneumoniae* genomics
694 and clinical aspects. AW and KL prepared the first manuscript draft. All authorships contributed
695 with reviewing and editing the manuscript and approved the final version.

696 **Abbreviations**

697 ALRI = acute lower respiratory infection

698 AMR = antimicrobial resistance

699 ARG = antimicrobial resistance gene

700 CHAMPS = Child Health and Mortality Prevention Surveillance

701 CG = clonal group

702 cgMLST = core genome multilocus sequence typing

703 cgST = core genome sequence type

704 ESBL = Extended spectrum beta-lactamase

705 Kp = *Klebsiella pneumoniae*

706 LMIC = low- to- middle-income countries

707 MDR = multidrug resistant

708 MITS = minimally invasive tissue sampling

709 MLST = multilocus sequence typing

710 SL = sublineage

711 ST = sequence type

712

713 **References**

714 1. Jain S, Williams DJ, Arnold SR, Ampofo K, Bramley AM, Reed C, et al. Community-
715 acquired pneumonia requiring hospitalization among U.S. children. N Engl J Med.
716 2015;372(9):835–45.

717 2. Madhi SA, De Wals P, Grijalva CG, Grimwood K, Grossman R, Ishiwada N, et al. The
718 burden of childhood pneumonia in the developed world: a review of the literature. Pediatr Infect
719 Dis J. 2013;32(3):e119–27.

720 3. Benet T, Sanchez Picot V, Messaoudi M, Chou M, Eap T, Wang J, et al. Microorganisms
721 Associated With Pneumonia in Children <5 Years of Age in Developing and Emerging
722 Countries: The GABRIEL Pneumonia Multicenter, Prospective, Case-Control Study. Clin Infect
723 Dis. 2017;65(4):604–12.

724 4. Pneumonia Etiology Research for Child Health Study G. Causes of severe pneumonia
725 requiring hospital admission in children without HIV infection from Africa and Asia: the PERCH
726 multi-country case-control study. Lancet. 2019;394(10200):757–79.

727 5. Wolf B, Gama A, Rey L, Fonseca W, Roord J, Fleer A, et al. Striking differences in the
728 nasopharyngeal flora of healthy Angolan, Brazilian and Dutch children less than 5 years old.
729 Ann Trop Paediatr. 1999;19(3):287–92.

730 6. Wolf B, Rey LC, Moreira LB, Milatovic D, Fleer A, Verhoef J, et al. Carriage of gram-
731 negative bacilli in young Brazilian children with community-acquired pneumonia. Int J Infect Dis.
732 2001;5(3):155–9.

733 7. Ghafoor A, Nomani NK, Ishaq Z, Zaidi SZ, Anwar F, Burney MI, et al. Diagnoses of
734 acute lower respiratory tract infections in children in Rawalpindi and Islamabad, Pakistan. Rev
735 Infect Dis. 1990;12 Suppl 8:S907–14.

736 8. Ikeogu MO, Wolf B, Mathe S. Pulmonary manifestations in HIV seropositivity and
737 malnutrition in Zimbabwe. Arch Dis Child. 1997;76(2):124–8.

738 9. El Seify MY, Fouda EM, Ibrahim HM, Fathy MM, Husseiny Ahmed AA, Khater WS, et al.
739 Microbial Etiology of Community-Acquired Pneumonia Among Infants and Children Admitted to
740 the Pediatric Hospital, Ain Shams University. Eur J Microbiol Immunol (Bp). 2016;6(3):206–14.

741 10. Zar HJ, MacGinty R, Workman L, Burd T, Smith G, Myer L, et al. Klebsiella pneumoniae
742 Lower Respiratory Tract Infection in a South African Birth Cohort: a Longitudinal Study. Int J
743 Infect Dis. 2022;121:31–8.

744 11. Taylor AW, Blau DM, Bassat Q, Onyango D, Kotloff KL, Arifeen SE, et al. Initial findings
745 from a novel population-based child mortality surveillance approach: a descriptive study. Lancet
746 Glob Health. 2020;8(7):e909–e19.

747 12. Bassat Q, Blau DM, Ogbuanu IU, Samura S, Kaluma E, Bassey IA, et al. Causes of
748 Death Among Infants and Children in the Child Health and Mortality Prevention Surveillance
749 (CHAMPS) Network. JAMA Netw Open. 2023;6(7):e2322494.

750 13. Verani JR, Blau DM, Gurley ES, Akelo V, Assefa N, Baillie V, et al. Child deaths caused
751 by Klebsiella pneumoniae in sub-Saharan Africa and south Asia: a secondary analysis of Child
752 Health and Mortality Prevention Surveillance (CHAMPS) data. Lancet Microbe. 2024;5(2):e131–
753 e41.

754 14. Mahtab S, Blau DM, Madewell ZJ, Ogbuanu I, Ojulong J, Lako S, et al. Post-mortem
755 investigation of deaths due to pneumonia in children aged 1-59 months in sub-Saharan Africa
756 and South Asia from 2016 to 2022: an observational study. Lancet Child Adolesc Health.
757 2024;8(3):201–13.

758 15. Wheelock A, Chasaya M, Namuziya N, Jumbe Marsden E, Kapasa M, Mumba C, et al.
759 Using Minimally Invasive Tissue Sampling and Determination of Cause of Death to Establish
760 Etiologies of Community Respiratory Deaths Among Zambian Infants and Children. J Pediatric
761 Infect Dis Soc. 2025;14(1).

762 16. Gill CJ, Mwananyanda L, MacLeod WB, Kwenda G, Pieciak R, Mupila Z, et al. Infant
763 deaths from respiratory syncytial virus in Lusaka, Zambia from the ZPRIME study: a 3-year,
764 systematic, post-mortem surveillance project. Lancet Glob Health. 2022;10(2):e269–e77.

765 17. Mwananyanda L, Gill CJ, MacLeod W, Kwenda G, Pieciak R, Mupila Z, et al. Covid-19
766 deaths in Africa: prospective systematic postmortem surveillance study. BMJ. 2021;372:n334.

767 18. Revised WHO Classification and Treatment of Pneumonia in Children at Health
768 Facilities: Evidence Summaries. Geneva: The World Health Organization, ISBN: 978-92-4-
769 150781-3; 2014 Available from https://www.ncbi.nlm.nih.gov/books/NBK264162/.

770 19. Lam MMC, Wick RR, Watts SC, Cerdeira LT, Wyres KL, Holt KE. A genomic
771 surveillance framework and genotyping tool for Klebsiella pneumoniae and its related species
772 complex. Nat Commun. 2021;12(1):4188.

773 20. Thomson KM, Dyer C, Liu F, Sands K, Portal E, Carvalho MJ, et al. Effects of antibiotic
774 resistance, drug target attainment, bacterial pathogenicity and virulence, and antibiotic access
775 and affordability on outcomes in neonatal sepsis: an international microbiology and drug
776 evaluation prospective substudy (BARNARDS). Lancet Infect Dis. 2021;21(12):1677–88.

777 21. Collaborators GAR. Global burden of bacterial antimicrobial resistance 1990-2021: a
778 systematic analysis with forecasts to 2050. Lancet. 2024;404(10459):1199–226.

779 22. Global antibiotic resistance surveillance report 2025: WHO Global Antimicrobial
780 Resistance and Use Surveillance System (GLASS). Geneva: World Health Organization; 2025.
781 Licence: CC BY-NC-SA 3.0 IGO. ISBN 978-92-4-011633-7.

782 23. Sands K, Carvalho MJ, Portal E, Thomson K, Dyer C, Akpulu C, et al. Characterization
783 of antimicrobial-resistant Gram-negative bacteria that cause neonatal sepsis in seven low- and
784 middle-income countries. Nat Microbiol. 2021;6(4):512–23.

785 24. Mills RO, Dadzie I, Le-Viet T, Baker DJ, Addy HPK, Akwetey SA, et al. Genomic
786 diversity and antimicrobial resistance in clinical Klebsiella pneumoniae isolates from tertiary
787 hospitals in Southern Ghana. J Antimicrob Chemother. 2024;79(7):1529–39.

788 25. Muraya A, Kyany'a C, Kiyaga S, Smith HJ, Kibet C, Martin MJ, et al. Antimicrobial
789 Resistance and Virulence Characteristics of Klebsiella pneumoniae Isolates in Kenya by Whole-
790 Genome Sequencing. Pathogens. 2022;11(5).

791 26. Amulele AV, Orindi B, Ndumba M, Tigoi C, Omuoyo D, Kahindi E, et al. Genomic
792 epidemiology of Klebsiella pneumoniae neonatal and infant sepsis in Kenyan hospitals.
793 MedRxiv. 2025.

794 27. Heinz E, Pearse O, Zuza A, Bilima S, Msefula C, Musicha P, et al. Longitudinal analysis
795 within one hospital in sub-Saharan Africa over 20 years reveals repeated replacements of
796 dominant clones of Klebsiella pneumoniae and stresses the importance to include temporal
797 patterns for vaccine design considerations. Genome Med. 2024;16(1):67.

798 28. Wyres KL, Lam MMC, Holt KE. Population genomics of Klebsiella pneumoniae. Nat Rev
799 Microbiol. 2020;18(6):344–59.

800 29. Barbier E, Rodrigues C, Depret G, Passet V, Gal L, Piveteau P, et al. The ZKIR Assay, a
801 Real-Time PCR Method for the Detection of Klebsiella pneumoniae and Closely Related
802 Species in Environmental Samples. Appl Environ Microbiol. 2020;86(7).

803 30. Lindstedt K, Buczek D, Pedersen T, Hjerde E, Raffelsberger N, Suzuki Y, et al.
804 Detection of Klebsiella pneumoniae human gut carriage: a comparison of culture, qPCR, and
805 whole metagenomic sequencing methods. Gut Microbes. 2022;14(1):2118500.

806 31. Hennart M, Guglielmini J, Bridel S, Maiden MCJ, Jolley KA, Criscuolo A, et al. A Dual
807 Barcoding Approach to Bacterial Strain Nomenclature: Genomic Taxonomy of Klebsiella
808 pneumoniae Strains. Mol Biol Evol. 2022;39(7).

809 32. Bialek-Davenet S, Criscuolo A, Ailloud F, Passet V, Jones L, Delannoy-Vieillard AS, et
810 al. Genomic definition of hypervirulent and multidrug-resistant Klebsiella pneumoniae clonal
811 groups. Emerg Infect Dis. 2014;20(11):1812–20.

812 33. Diancourt L, Passet V, Verhoef J, Grimont PA, Brisse S. Multilocus sequence typing of
813 Klebsiella pneumoniae nosocomial isolates. J Clin Microbiol. 2005;43(8):4178–82.

814 34. Stanton TD, Hetland MAK, Lohr IH, Holt KE, Wyres KL. Fast and accurate in silico
815 antigen typing with Kaptive 3. Microb Genom. 2025;11(6).

816 35. Holt KE, Wertheim H, Zadoks RN, Baker S, Whitehouse CA, Dance D, et al. Genomic
817 analysis of diversity, population structure, virulence, and antimicrobial resistance in Klebsiella

818 pneumoniae, an urgent threat to public health. Proc Natl Acad Sci U S A. 2015;112(27):E3574–
819 81.

820 36. Navon-Venezia S, Kondratyeva K, Carattoli A. Klebsiella pneumoniae: a major
821 worldwide source and shuttle for antibiotic resistance. Fems Microbiol Rev. 2017;41(3):252–75.

822 37. Spadar A, Perdigao J, Campino S, Clark TG. Large-scale genomic analysis of global
823 Klebsiella pneumoniae plasmids reveals multiple simultaneous clusters of carbapenem-resistant
824 hypervirulent strains. Genome Med. 2023;15(1):3.

825 38. Hetland MAK, Hawkey J, Bernhoff E, Bakksjo RJ, Kaspersen H, Rettedal SI, et al.
826 Within-patient and global evolutionary dynamics of Klebsiella pneumoniae ST17. Microb
827 Genom. 2023;9(5).

828 39. Chizimu JY, Kalungia AC, Mudenda S, Muhimba Z, Lukwesa C, Mufwambi W, et al.
829 Diagnostic Stewardship Trends and Antimicrobial Resistance Profiles of Bacteria Isolated in
830 Zambia: A Five-Year Retrospective Study (2020-2024). Antibiotics (Basel). 2025;14(11).

831 40. Namukonda S, Shawa M, Siame A, Mwansa J, Mulundu G. Prevalence and antibiotic
832 resistance profiles of ESKAPE pathogens in the neonatal intensive care unit of the women and
833 newborn hospital in Lusaka, Zambia. Antimicrob Resist Infect Control. 2025;14(1):96.

834 41. Peltier F, Choquet M, Decroix V, Adjide CC, Castelain S, Guiheneuf R, et al.
835 Characterization of a multidrug-resistant Klebsiella pneumoniae ST607-K25 clone responsible
836 for a nosocomial outbreak in a neonatal intensive care unit. J Med Microbiol. 2019;68(1):67–76.

837 42. Castro J, Araujo D, Oliveira H, Fernandes L, Oliveira R, Brinks E, et al. Multidrug-
838 resistant Klebsiella pneumoniae and Klebsiella variicola isolated from patients in Portuguese
839 hospitals: Genomic and phenotypic characterization. The Microbe. 2024;5:100172.

840 43. Henson SP, Boinett CJ, Ellington MJ, Kagia N, Mwarumba S, Nyongesa S, et al.
841 Molecular epidemiology of Klebsiella pneumoniae invasive infections over a decade at Kilifi
842 County Hospital in Kenya. Int J Med Microbiol. 2017;307(7):422–9.

843 44. Brinkac LM, White R, D'Souza R, Nguyen K, Obaro SK, Fouts DE. Emergence of New
844 Delhi Metallo-beta-Lactamase (NDM-5) in Klebsiella quasipneumoniae from Neonates in a
845 Nigerian Hospital. mSphere. 2019;4(2).

846 45. Mishra K, Banerjee T, Yadav G, Kumar A, Pratap A, Chaurasiya S, et al. Emergence of
847 drug-resistant Klebsiella pneumoniae phylogroups (K. quasipneumoniae and K. variicola)
848 causing human infections. Microbiol Spectr. 2025;13(9):e0019825.

849 46. Phillips LT, Bates M, Coffin SE, Foster-Nyarko E, Kapasa M, Machona S, et al.
850 Transmission dynamics of Klebsiella pneumoniae in a neonatal intensive care unit in Zambia
851 before and after an infection control bundle. PLOS Glob Public Health. 2026;6(2):e0005965.

852 47. Argimon S, David S, Underwood A, Abrudan M, Wheeler NE, Kekre M, et al. Rapid
853 Genomic Characterization and Global Surveillance of Klebsiella Using Pathogenwatch. Clin
854 Infect Dis. 2021;73(Suppl_4):S325–S35.

855 48. Lan P, Jiang Y, Zhou J, Yu Y. A global perspective on the convergence of
856 hypervirulence and carbapenem resistance in Klebsiella pneumoniae. J Glob Antimicrob Resist.
857 2021;25:26–34.

858 49. Stanton TD, Keegan SP, Abdulahi JA, Amulele AV, Bates M, Heinz E, et al. Distribution
859 of capsule and O types in Klebsiella pneumoniae causing neonatal sepsis in Africa and South
860 Asia: meta-analysis of genome-predicted serotype prevalence and potential vaccine coverage.
861 MedRxiv. 2025. https://doi.org/10.1101/2025.06.28.25330253.

862 50. Dangor Z, Benson N, Berkley JA, Bielicki J, Bijsma MW, Broad J, et al. Vaccine value
863 profile for Klebsiella pneumoniae. Vaccine. 2024;42(19 Suppl 1):S125–S41.

864 51. David S, Cohen V, Reuter S, Sheppard AE, Giani T, Parkhill J, et al. Integrated
865 chromosomal and plasmid sequence analyses reveal diverse modes of carbapenemase gene
866 spread among Klebsiella pneumoniae. Proc Natl Acad Sci U S A. 2020;117(40):25043–54.

867 52. Wyres KL, Holt KE. Klebsiella pneumoniae as a key trafficker of drug resistance genes
868 from environmental to clinically important bacteria. Curr Opin Microbiol. 2018;45:131–9.

869 53. Lu J, Breitwieser FP, Thielen P, Salzberg SL. Bracken: estimating species abundance in
870 metagenomics data. Peerj Comput Sci. 2017.

871 54. Wood DE, Lu J, Langmead B. Improved metagenomic analysis with Kraken 2. Genome
872 Biol. 2019;20(1):257.

873 55. Kim D, Song L, Breitwieser FP, Salzberg SL. Centrifuge: rapid and sensitive
874 classification of metagenomic sequences. Genome Res. 2016;26(12):1721–9.

875 56. Clausen P, Aarestrup FM, Lund O. Rapid and precise alignment of raw reads against
876 redundant databases with KMA. BMC Bioinformatics. 2018;19(1):307.

877 57. Lam MMC, Wyres KL, Judd LM, Wick RR, Jenney A, Brisse S, et al. Tracking key
878 virulence loci encoding aerobactin and salmochelin siderophore synthesis in Klebsiella
879 pneumoniae. Genome Med. 2018;10(1):77.

880 58. Lam MMC, Salisbury SM, Treat LP, Wick RR, Judd LM, Wyres KL, et al. Genomic and
881 functional analysis of rmp locus variants in Klebsiella pneumoniae. Genome Med.
882 2025;17(1):36.

883 59. Lam MMC, Wick RR, Wyres KL, Gorrie CL, Judd LM, Jenney AWJ, et al. Genetic
884 diversity, mobilisation and spread of the yersiniabactin-encoding mobile element ICEKp in
885 Klebsiella pneumoniae populations. Microb Genom. 2018;4(9).
886