

1 **Full title**

2 Plasmid characterization in bacterial isolates of public health relevance in a tertiary healthcare
3 facility in Kilimanjaro, Tanzania

4 **Short title**

5 Plasmids characterization and public health relevance

6 **Authors**

7 Lameck Pashet Sengeruan ^{1, 2, *} (LP), Marco van Zwetselaar ³ (MvZ), Happiness Kumburu ³ (HK),
8 Frank M. Aarestrup ⁴ (FM), Katharina Kreppel ¹ (KK), Elingarami Sauli ¹ (ES), Tolbert Sonda ³ (TS)

9 **Affiliations**

10 ¹ School of Life Science and Bioengineering, The Nelson Mandela African Institution of Science
11 and Technology, Arusha, Tanzania.

12 ² Training and Capacity Building Department, Ifakara Health Institute, Bagamoyo, Tanzania.

13 ³ Genomics and Bioinformatics unit, Kilimanjaro Clinical Research Institute, Moshi, Tanzania

14 ⁴ DTU-Food, Technical University of Denmark

15 * Corresponding author

16 E-mail: lpashet@ihi.or.tz/ lameckpashet@gmail.com (LP)

17

18

Abstract

Plasmids are infectious double stranded DNA molecules that are found within bacteria. Horizontal gene transfer promotes successful spread of different types of plasmids within or among bacteria species, making their detection an important task for guiding clinical treatment. We used whole genome sequenced data to determine the prevalence of plasmid replicons in clinical bacterial isolates, the presence of resistance and virulence genes in plasmids, and the relationship between resistance and virulence genes within each plasmid. All bacterial sequences were de novo assembled using Unicycler before extraction of plasmids. Assembly graphs were submitted to Gplas+plasflow for plasmid prediction. The predicted plasmid components were validated using PlasmidFinder.

A total of 159 (56.2%) out of 283 bacterial isolates were found to carry plasmids, with *E. coli*, *K. pneumoniae* and *S. aureus* being the most prevalent plasmid carriers. A total of 27 (87.1%) combined plasmids were found to carry both resistance and virulence genes compared to 4 (12.9%) single plasmids. No statistically significant correlation was found between the number of antimicrobial resistance and virulence genes in plasmids ($r = -0.25$, $p > 0.05$). Our findings show a relatively high proportion of plasmid-carrying isolates suggesting selection pressure due to antibiotic use in the hospital. Co-occurrence of antibiotic resistance and virulence genes in clinical isolates is a public health relevant problem needing attention.

Keywords: whole genome sequencing; plasmids; antibiotic resistance; virulence; public health, Tanzania

40

41 **Introduction**

42 Plasmids are circular, double-stranded DNA molecules that occur naturally in bacterial cells [1],
 43 whose genes often provide evolutionary advantages for bacteria, such as antimicrobial resistance
 44 and/or virulence [2,3]. Plasmids are important vehicles in disseminating and acquiring antibiotic
 45 resistance and virulence, and can thus constitute a major burden on human health [4]. Recent
 46 studies have suggested that the prevalence of antimicrobial resistance (AMR) is higher in Low- and
 47 Middle-income Countries (LMICs) compared to European countries and the United States [5,6].
 48 There is however, limited knowledge regarding the dissemination of antibiotic resistance genes
 49 (ARGs) and virulence among clinical isolates in Sub-Saharan Africa (SSA). This study was
 50 conducted to determine the proportion of bacterial isolates carrying plasmids, to identify plasmids
 51 that mediate resistance and virulence genes, and to investigate the relationship between
 52 antimicrobial resistance genes and virulence genes within plasmids using whole genome sequence
 53 data from bacterial isolates among inpatients at Kilimanjaro Christian Medical Centre (KCMC) in
 54 Tanzania.

55 **Materials and methods**

56 **Study setting, Whole-genome sequencing and library preparation**

57 Kilimanjaro Christian Medical Centre (KCMC) is one of Tanzania's five zonal referral hospitals,
 58 located in Moshi, northern Tanzania. KCMC has a bed capacity of 650 and serves a catchment area
 59 of about 15 million people. It serves around 500 outpatients daily, from different parts of Tanzania

[7]. The whole genome sequence data that was analyzed originated from a prospective cross-sectional study that was conducted at KCMC between August 2013 and August 2015. In this study, a total of 56 stool, 122 sputum, 126 blood and 286 wound swabs (wound/pus) clinical samples, with patients' clinical and socio-demographic characteristics, were collected from 575 patients admitted to KCMC hospital [8,9]. A written informed consent was obtained from each participant and from parents or guardians of children before enrolled into the study.

Collected specimens were taken to the microbiology unit at Kilimanjaro Clinical Research Institute (KCRI) for culture and identification of bacterial isolates. Out of 590 specimens collected, 249 were culture positive, resulting in 377 isolates [8]. All bacterial isolates were sequenced in the KCRI genomics lab, and all sequences were archived on the KCRI compute cluster. In brief, whole genome sequencing (WGS) was performed for genomic DNA that was extracted from cultures of bacterial isolates using the Easy-DNA Extraction Kit (Invitrogen®). Short-read WGS was performed using the Illumina MiSeq platform (Illumina Inc.). Libraries for Illumina sequencing were constructed using the Illumina Nextera XT kit (Illumina Ltd., San Diego, CA, USA) according to the manufacturer's recommendations. The libraries were sequenced on Illumina MiSeq platform using the 2 x 250bp paired-end protocol, previously reported by Kumburu et al [8] & Sonda et al [9]. For the purpose of this study, a total of 283 bacterial whole genomes isolates with associated metadata were retrieved for analysis. Additional ethical approval was obtained from the Ifakara Health Institute Research Ethics Committee (IHI/IRB/No: 14-2021) for plasmid characterization.

82 **Bioinformatics Analyses**

83 **Quality Control and Trimming of Illumina Sequences**

84 The following steps were followed: (i) All bacterial raw reads were submitted to in-house bacterial
 85 analysis pipeline (BAP), available at <https://github.com/zwets/kcri-cge-bap>. Assembly was
 86 performed using SKESA 2.4.0 [10]. (ii) All resulting assemblies were then processed in batch by
 87 the Genome Taxonomy Database Toolkit (GTDB-Tk) 0.3.2 [11] for detailed taxonomic
 88 assignment. (iii) Metrics produced by the BAP and GTDB-Tk were then used to assess the quality
 89 of each assembly. Assessment was based on read counts, coverage depth, assembly structure
 90 (contig count, N1, N50, L50), deviation of assembly length from reference, GTDB alignment
 91 fraction, and GTDB Multi-Locus Sequence Alignment (MSA) coverage. A six-point scale was used
 92 for assembly quality rating: 0 (Unusable), 1 (Mix), 2 (Bad), 3 (Usable), 4 (Good), and 5 (Excellent).
 93 (iv) Finally, categories 0 to 2 were excluded, while categories 3 through 5 were used for subsequent
 94 analysis. Every assembly in these categories was for a single isolate that had (nearly) complete
 95 genome coverage, at sufficient sequencing depth.

96 **Plasmid extraction and validation**

97 Raw reads assembly was repeated with Unicycler 0.4.7 [12] for its ability as a “SPAdes optimiser”
 98 to produce long and, in the ideal case, circular contigs. Assembly graphs (GFA) were submitted to
 99 Gplas+plasflow for plasmid prediction. Gplas 0.6.1 [13] + Plasflow 1.1 [14] take into account the
 100 connected components in the assembly graph when predicting plasmids. The components predicted
 101 to be plasmids were extracted from the assemblies and submitted to PlasmidFinder version 1.3 [15]
 102 for validation.

103 **Identification of Plasmid-Mediated Antibiotic Resistance Genes (ARGs) and** 104 **Virulence genes**

105 To identify antibiotic resistance and virulence genes carried in plasmids, the assembled putative
106 plasmid sequences for each isolate were submitted to Resfinder 4.0 [16] and VirulenceFinder
107 1.4[17] respectively. In both Resfinder and VirulenceFinder, 90% identity and 60% coverage
108 settings to call a gene were selected.

109 **Statistical analysis**

110 Stata 14 (College Station, TX, 77845, USA) was used for descriptive statistics and determination of
111 the relationship between antimicrobial resistance and virulence genes in plasmids.

112 **Results**

113 **Study population**

114 In total, 128 patients whose whole genome bacterial isolates were analyzed were included in this
115 study (Table 1). One-hundred twenty eight patients were plasmid positive isolates. The mean age in
116 years (SD) was 46.2 (18.0). Male patients were 77 (60.2%), females were 47 (36.7%) and 4 (3.1%)
117 missed gender identification. A total of 62 (48.4%) patients were admitted to surgical ward, 9
118 (7.1%) surgical ICU, 52 (40.6%) medical ward, 4 (3.1%) medical ICU ward, and 1 (0.8%) missed
119 ward admission identification. Eighty-seven (67.9%) specimens were swabs, 19 (14.8%) were
120 stool, 13 (10.2%) were sputum, 8 (6.3%) were blood and 1(0.8%) specimen missed identification.
121 A total of 28 (21.9%) patients were diabetic, 6 (4.7%) were cancer patients, 6 (4.7%) were suffering
122 from TB, 2 (1.5%) were HIV positive and 86 (67.2%) were others: Of those 61 (47.7%) had no

underlying conditions and 25 (19.5%) had other underlying conditions. Of the wound swabs, twenty-four (18.8%) were from patients with diabetic wounds, 11 (8.6%) were from burn wounds, 10 (7.8%) were from post-surgical wounds, 6 (4.6%) were from motor traffic accidents wounds and 77 (60.2%) were others: Of those 35 (27.3%) other wounds and 42 (32.8%) had no wounds.

Seventy-eight (60.9%) patients had no history of hospitalization, 42 (32.8%) had hospitalization history and 8 (6.3%) missed identification. A total of 81 (63.3%) patients were transferred from another hospital, 44 (34.4%) were not and 3 (2.3%) missed identification. Among all participants, the median number of days stayed in the hospital before the survey was 8 days (Table 1).

142 **Table 1. Demographic and clinical characteristics of patients that samples were taken from**

Patient characteristics	Missing ^a	Total (%)
Number of patients		128 (100)
Mean age in years (SD)		46.2 (18.0)
Gender	4 (3.1)	
Female		47 (36.7)
Male		77 (60.2)
Ward of admission	1 (0.8)	
Surgical		62 (48.4)
Surgical ICU		9 (7.1)
Medical		52 (40.6)
Medical ICU		4 (3.1)
Specimen collected	1 (0.8)	
Blood		8 (6.3)
Sputum		13 (10.2)
Stool		19 (14.8)
Swab		87 (67.9)
Underlying conditions		
Cancer		6 (4.7)
Diabetes		28 (21.9)
HIV		2 (1.5)
TB		6 (4.7)
Others		86 (67.2)
Type of wound		
Burn wound		11 (8.6)
Diabetic wound		24 (18.8)
Motor traffic wound		6 (4.6)
Post-surgical wound		10 (7.8)
Others		77 (60.2)
History of hospitalization	8 (6.3)	
No		78 (60.9)
Yes		42 (32.8)
Patient hospital transfer	3 (2.3)	
No		44 (34.4)
Yes		81 (63.3)
Patient ward transfer	5 (3.9)	
No		110 (85.9)
Yes		13 (10.2)
Median time in days stayed in the hospital before survey (IQR)	8 (6.3)	8 (4-11.5)

143 SD, standard deviation; ICU, intensive care unit; IQR, interquartile range; Missing ^a, were the missing values
144 in each variable.

145

146

147

148

149

150

151

152

153

154

155

156

157

158

159

160

161

162

Proportion of bacterial species carrying plasmids

A total of 283 whole genome bacterial sequences were analyzed. One hundred fifty-nine (56.2%) bacterial isolates were detected to carry plasmids. Out of 159 plasmids, 94 non-repetitive plasmids were predicted. Of 94 plasmids, 48 (51.1%) were single plasmids and 46 (48.9%) were combined plasmids (two or more recombined plasmids). *K. pneumoniae* isolates were the most carriers of combined plasmids (17, 28.3%), followed by *S. aureus* (15, 25.0%) and *E.coli* (15, 25.0%). *E.coli* isolates were the most single plasmid carriers (23, 23.2%), followed by *S. aureus* (15, 15.2%) and *P. mirabilis* (14,14.4%) (Table 2).

163

164 **Table 2. Proportion of bacterial isolates carrying plasmids**

			Plasmids	
	Isolates		Single	Combined ^b
Species	N	%	N	N
<i>Enterobacter asburiae</i>	1	0.6	1	0
<i>Enterobacter cloacae</i>	3	1.9	1	2
<i>Enterobacter hormaechei</i>	10	6.3	6	4
<i>Enterobacter kobei</i>	1	0.6	0	1
<i>Enterobacter roggenkampii</i>	1	0.6	1	0
<i>Enterobacter soli</i>	1	0.6	1	0
<i>Enterobacter sp. n18-03635</i>	1	0.6	1	0
<i>Enterococcus faecalis</i>	7	4.4	5	2
<i>Enterococcus faecium</i>	3	1.9	2	1
<i>Enterococcus gallinarum</i>	1	0.6	1	0
<i>Escherichia coli</i>	38	23.9	23	15
<i>Klebsiella Michiganensis</i>	2	1.3	1	1
<i>Klebsiella oxytoca</i>	2	1.3	1	1
<i>Klebsiella pneumoniae</i>	25	15.7	8	17
<i>Klebsiella variicola</i>	2	1.3	2	0
<i>Micrococcus sp. Kbs0714</i>	1	0.6	1	0
<i>Morganella morganii</i>	4	2.5	4	0
<i>Proteus columbae</i>	1	0.6	1	0
<i>Proteus mirabilis</i>	14	8.8	14	0
<i>Proteus penneri</i>	1	0.6	1	0
<i>Proteus vulgaris</i>	1	0.6	1	0
<i>Pseudomonas aeruginosa</i>	2	1.3	2	0
<i>Shewanella algae</i>	1	0.6	1	0
<i>Staphylococcus aureus</i>	30	18.9	15	15
<i>Staphylococcus capitis</i>	1	0.6	1	0
<i>Staphylococcus epidermidis</i>	1	0.6	1	0
<i>Staphylococcus haemolyticus</i>	3	1.9	3	0
<i>Staphylococcus hominis</i>	1	0.6	0	1

165 N, total number of isolates/plasmids; combined^b, two or more plasmids merged or fused together.

166

167

168

169 **Plasmids concurrently mediating resistance and virulence genes**

170 A total of 31 plasmids were identified to carry both resistance and virulence genes, of which 27
171 (87.1%) were combined plasmids and 4 (12.9%) were single plasmids (Tables 3 and 4). All four
172 single plasmids were carried by *E. coli*. Resistance gene *Sul1* was found the most common across
173 three single plasmids IncFII, IncQ1 and IncFII(pRSB107). Virulence genes *iucC* and *iutA* were also
174 seen the most common across three single plasmids such as IncQ1, IncFII(pRSB107) and IncFIA
175 (Table 3).

176 Among the 27 combined plasmids, 12 (44.4%) were carried by *E.coli* isolates, 10 (37.1%) by *K.*
177 *pneumoniae* isolates, 2 (7.4%) by *E. hormaechei* isolates, 1 (3.7%) by *E. cloacae* isolate, 1(3.7%)
178 by *K. oxytoca* isolate and 1 (3.7%) by *K. michiganensis* isolate. Virulence gene *traT* was seen in 19
179 (70.4%) of the 27 combined plasmids, followed by *terC* which was identified in 7 (25.9%)
180 plasmids. Resistance genes in combined plasmids, *sul2* was observed in 17 (62.9%) plasmids,
181 followed by *blaTEM-1B* in 15 (55.6%) plasmids, followed by *blaCTX-M-15* in 14 (51.9%)
182 plasmids and *blaOXA-1* in 13 (48.1%) plasmids (Table 4).

183

184 **Table 3. Single plasmids mediating both resistance and virulence genes**

Single plasmids	Resistance genes	Virulence genes
IncFII	<i>aac(3)-IIa,aadA5,blaCTX-M-15,dfrA17,qacE,sul1,tet(B)</i>	<i>traT</i>
IncQ1	<i>aph(3'')-Ib,aph(6)-Id,blaTEM-1B,dfrA7,qacE,sul1,sul2</i>	<i>cea,focCsfaE,focG,focI,iha,ireA,iucC,iutA,mchB,mchC,mchF,mcmA,papA_F48,papC,sat</i>
IncFIA	<i>sitABCD,tet(A)</i>	<i>iucC,iutA,sitA</i>
IncFII(pRSB107)	<i>dfrA5,qacE,sul1,sul2</i>	<i>capU,iroN,iss,iucC,iutA,mchB,mchC,mchF,mcmA,vat</i>

185

186

187

188

189

190

191

192

193

194

195

196

197

198

199

200

Table 4. Combined plasmids mediating both resistance and virulence genes

Combined plasmid	Resistance genes	Virulence genes
IncFIB(K),IncFII(pKP91),IncR	<i>ARR-2,aac(3)-IIa,aac(6')-Ib-cr,aadA1,aph(3'')-Ib,aph(6)-Id,blaCTX-M-15,blaOXA-1,catB3,cmlA1,dfrA14,ere(A),qacE,sul1,sul2,tet(A)</i>	<i>traT</i>
IncFIB(K)(pCAV1099-114),IncHI2,IncHI2A,IncX3	<i>aac(6')-Ib-cr,aadA5,aph(3'')-Ib,aph(6)-Id,blaCTX-M-15,blaOXA-1,blaTEM-1B,catB3,dfrA17,qacE,qnrB1,sul1,sul2,tet(A),tet(B)</i>	<i>terC</i>
Col156,IncFIA,IncFIB(AP001918),IncFII(pRSB107)	<i>aac(3)-IIa,aac(6')-Ib-cr,aadA5,aph(3'')-Ib,aph(6)-Id,blaCTX-M-15,blaOXA-1,catB3,dfrA17,mph(A),qacE,sul1,sul2,tet(A)</i>	<i>hra,ihai,iucC,iutA,sat,senB,traT</i>
IncFIB(K),IncFII(K),IncQ1,IncR	<i>ARR-2,aac(3)-IIa,aadA1,aph(3'')-Ib,aph(3')-Ia,aph(6)-Id,blaCTX-M-15,blaTEM-1B,cmlA1,ere(A),qacE,qnrB1,sul1,sul2</i>	<i>traT</i>
IncFIA(HI1),IncFIB(K),IncFII(Yp),IncHI2,IncHI2A,IncN3,pKP1433	<i>aac(3)-IIa,aac(6')-Ib-cr,aph(3'')-Ib,aph(6)-Id,blaCTX-M-15,blaOXA-1,blaTEM-1B,catB3,qacE,qnrB1,sul1,sul2,tet(A)</i>	<i>terC</i>
IncFIB(K),IncFII(K)	<i>aac(3)-IIa,aac(6')-Ib-cr,aph(3'')-Ib,aph(6)-Id,blaCTX-M-15,blaOXA-1,blaTEM-1B,catB3,dfrA14,qnrB1,sul2,tet(A)</i>	<i>traT</i>
IncHI2,IncHI2A	<i>aac(3)-IIa,aac(6')-Ib-cr,aph(3'')-Ib,aph(6)-Id,blaCTX-M-15,blaOXA-1,blaTEM-1B,catB3,dfrA14,qnrB1,sul2,tet(A)</i>	<i>terC</i>
Col156,IncFIA,IncFIB(AP001918),IncFII	<i>aac(3)-IIa,aac(6')-Ib-cr,aadA5,blaCTX-M-15,blaOXA-1,catB3,dfrA17,mph(A),qacE,sitABCD,sul1,tet(A)</i>	<i>capU,fyuA,irp2,iucC,iutA,senB,sitA,traT</i>
IncFIB(pECLA),IncFII(pECLA),IncHI2,IncHI2A	<i>aac(3)-IIa,aac(6')-Ib-cr,aph(3'')-Ib,aph(6)-Id,blaCTX-M-15,blaOXA-1,catB3,dfrA14,qnrB1,sul2,tet(A)</i>	<i>terC</i>
Col156,IncFIA,IncFIB(AP001918)	<i>aac(6')-Ib-cr,aadA5,blaCTX-M-15,blaOXA-1,catB3,dfrA17,mph(A),qacE,sitABCD,sul1,tet(A)</i>	<i>iucC,iutA,senB,sitA,traT</i>
IncFIA,IncFIB(AP001918)	<i>aac(3)-IIa,aac(6')-Ib-cr,aadA5,blaCTX-M-15,blaOXA-1,catB3,dfrA17,qacE,sitABCD,sul1</i>	<i>fyuA,irp2,iucC,iutA,sitA,traT</i>
IncFIA,IncFII	<i>aac(6')-Ib-cr,aadA5,blaCTX-M-15,blaOXA-1,catB3,dfrA17,mph(A),qacE,sul1,tet(A)</i>	<i>afaA,afaC,afaD,ihai,iucC,iutA,nfaE,papA_F43,sat,traT</i>
IncFIA,IncFIB(AP001918),IncFII(pAM A1167-NDM-5)	<i>aac(3)-IIa,aadA2,blaTEM-1B,catA1,dfrA12,mph(A),qacE,qepA4,sul1</i>	<i>traT</i>

209 **Table 4. Continued**

IncFIA(HI1),IncFIB(K),IncFII(pKP91),IncR	<i>aac(6')-Ib-cr,aph(3'')-Ib,aph(6)-Id,blaCTX-M-15,blaOXA-1,blaTEM-1B,catB3,dfrA14,sul2</i>	<i>traT</i>
IncFIA,IncFIB(AP001918),IncFII	<i>aac(6')-Ib-cr,aadA5,blaOXA-1,catB3,dfrA17,qacE,sul1,tet(B)</i>	<i>traT</i>
IncFIB(K),IncFII(K),IncQ1	<i>aph(3'')-Ib,aph(3')-Ia,aph(6)-Id,blaTEM-1B,dfrA14,mph(A),sul2</i>	<i>traT</i>
Col156,IncFIB(AP001918),IncFII	<i>aph(3'')-Ib,aph(6)-Id,blaTEM-1B,catA1,dfrA7,sul2,tet(D)</i>	<i>afaA,afaB,afaC,afaD,afaE,hra,iha,iss,iucC,iutA,papA_F43,sat,senB,traT</i>
Col156,IncFIA,IncFIB(AP001918),IncQ1	<i>aph(3'')-Ib,aph(6)-Id,blaTEM-1B,dfrA17,sul2,tet(B)</i>	<i>iha,iucC,iutA,papA_F43,sat,senB</i>
IncFIB(K)(pCAV1099-114),IncHI1B(pNDM-MAR)	<i>aph(3'')-Ib,aph(6)-Id,dfrA15,qacE,sul1,sul2</i>	<i>terC</i>
IncFIB(K),IncFII(K),IncR	<i>aac(3)-IId,blaCTX-M-15,blaTEM-1B,dfrA30,sul2</i>	<i>traT</i>
IncFIB(K)(pCAV1099-114),IncHI1B(pNDM-MAR),IncR	<i>blaTEM-1B,dfrA5,qacE,sul1,tet(D)</i>	<i>fyuA,irp2,traT</i>
IncFII(K),IncR	<i>aac(3)-IId,blaCTX-M-15,blaTEM-1B,dfrA30,sul2</i>	<i>traT</i>
IncFIB(AP001918),IncFII,IncQ1	<i>aph(3'')-Ib,aph(6)-Id,blaTEM-1B,dfrA5,sul2</i>	<i>cia,cvaC,etsC,hlyF,ireA,iroN,iss,iucC,iutA,mchF,ompT,papA_F11,papC,traT</i>
IncFIB(K)(pCAV1099-114),IncY	<i>sul2,tet(D)</i>	<i>terC</i>
IncFIB(AP001918),IncFII	<i>dfrA5</i>	<i>cia,cvaC,etsC,hlyF,ireA,iroN,iss,mchF,ompT,papA_F11,papC,traT</i>
IncFIB(pHCM2),IncHI2,IncHI2A	<i>blaTEM-1B</i>	<i>terC</i>
IncFIB(pB171),IncFII(pCoo)	<i>mdf(A)</i>	<i>eae,espA,espF,nleB,nleC,perA,tir,traT</i>

210 **Correlation between antimicrobial resistance and virulence genes**

211 We explored the relationship between the number of antibiotic resistance genes and virulence genes
212 in 27 combined plasmids using Pearson correlation. There was an inconclusive negative
213 relationship between antibiotic resistance and virulence genes existence in plasmids ($r = -0.25$, $p >$
214 0.05).

215 **Discussion**

216 In the present study a high proportion of clinical bacterial isolates from inpatients at KCMC
217 hospital was found to carry plasmids. The present findings are in concordance with previous
218 studies elsewhere [18]. The observed high carriage of plasmids by the analyzed isolates might
219 plausibly be a reflection of resistance selection pressure due to high antibiotic exposure in hospital
220 settings [19].

221 *E. coli* isolates were the most prevalent carriers of single plasmids followed by *S. aureus* and *P.*
222 *mirabilis*. On other hand, *K. pneumonia* were the most prevalent carriers of combined plasmids,
223 followed by *S. aureus* and *E. coli*. The present study findings are in line with a study results in a
224 tertiary care hospital in south India [18]. Possible explanation could be that the mentioned bacterial
225 species have great medical relevance and thus are relatively highly isolated in hospital settings
226 compared to other species [20,21]. However, the present study findings show a larger proportion of
227 *P. mirabilis* carrying plasmids than the study in south India. This difference might be due to the fact
228 that majority of the present study isolates were from wound specimens in which *P. mirabilis* were
229 identified [22,23].

230 This study identified bacterial species with low plasmid prevalence including *Enterobacter sp. n18-*
 231 *03635*, *Enterobacter kobei*, *Klebsiella variicola* and *Klebsiella oxytoca*. The study findings are
 232 consistent with other studies conducted in Canada, Greece and Mexico [24–26]. Interestingly, the
 233 study observed other low plasmid prevalence species that were reported elsewhere in soil samples,
 234 fish flesh samples [27–29] and pigeon flesh specimens [30] such as *Micrococcus sp. Kbs0714*,
 235 *Enterobacter soli* and *Proteus columbae*. The observed species with low plasmid prevalence might
 236 be due to rarity and in most cases misidentification [31,32]. However, reports of bacterial species
 237 with low plasmid indicates the possible emerging and transmission of bacterial pathogens in
 238 humans both in community and hospital settings [33].

239 Contrary to previous studies reporting IncF plasmid group in *E.coli* to carry resistance and
 240 virulence genes often [34] the present study, however shows IncQ1 carried the highest number of
 241 both resistance and virulence genes in *E.coli*, but the finding is in line with study conducted in
 242 Brazil [35,36]. This is possibly due to the fact that IncQ1 plasmids have high-level mobility,
 243 stability, replication at high copy number and transferred in wide range of bacterial species through
 244 conjugative plasmids [37–40].

245 In this study it was also identified that there are different combined plasmids ranging from two to
 246 seven plasmids. This is probably an indicative of bacterial evolution to adapt and thrive in hospitals
 247 where they are excessively exposed to antimicrobials, antiseptics and disinfectants [41–43]. A
 248 similar distribution of some combined plasmids in other regions carrying similar or different
 249 antibiotic resistance and virulence genes was noted in the present study. This suggests resistant
 250 bacteria arising in one geographical area can spread countrywide/worldwide either by direct
 251 exposure or through the food chain or climate change and the environment [6].

There was no significant relationship found in the present study between numbers of antibiotic resistance and virulence genes in plasmids ($r = -0.25$, $p > 0.05$), indicating acquisition of antibiotic resistance can induce the loss of virulence factors. Previous studies support this study finding [44], but does not agree with a study by Dionisio [45]. This discordance might be due to the fact that in other studies the relationship between resistance and virulence genes was determine at species level and were from gut and environmental samples [34].

Limitations

We acknowledge there are a number of limitations in the present study that warrant careful interpretation. Bioinformatics analysis was performed on Illumina short reads, which limited the ability to assemble completed plasmid genomes, and consequently the ability to ‘tease out’ individual plasmids from assembled contigs. Assembly graphs were classified by Gplas+PlasFlow for plasmid prediction. As for any machine learning-based approach or indeed any method based on inference from similarity with known sequences, including tools such as PlasmidFinder, the predictive ability of the model is strongly dependent on the data in its reference database or training set. A bias toward plasmids in well-studied organisms is therefore likely.

Conclusion

There is a high proportion of isolates carrying resistance and virulence plasmids. This shows a significant concern of AMR development and spread in Tanzanian health settings and other LMIC settings. With limited resources and health service capacities, the increasing AMR trends are expected to high impact on bacterial-associated mortalities and morbidities.

272 **Acknowledgements**

273 The parent study whose data we analyzed in this study was funded by DANIDA with grant number
274 DFC No. 12-007DTU. We would like to express our sincere gratitude to study participants whose
275 specimens yielded the bacterial isolates analyzed.

276 **Authors' contributions**

277 LP, MvZ, HK, FMA and TS conceived the initial idea. LP, TS, KK and ES developed the study
278 protocol. TS, MvZ and LP analyzed the data. LP wrote the first draft of the manuscript. TS, MvZ,
279 KK, ES, HK and FM reviewed and edited drafts of the manuscript. LP, TS, MvZ, KK and ES, HK
280 and FMA wrote the final Manuscript. All authors read and approved the final manuscript.

281 **Funding**

282 This study was supported by the Ifakara Health Institute (Training and Capacity Building
283 Department) as part of the Master scholarship to Lameck Pashet Sengeruan. The funder and the
284 data providers had no role in the design, analysis, interpretation of the results, and preparation of
285 manuscript or decision to publish.

286 **Availability of data**

287 All relevant data are within the manuscript and its Supporting Information files.

288 **Consent for publication**

289 Not applicable

290 **References**

- 291 1. Dionisio F, Zilhão R, Alves J. Plasmid Interactions between plasmids and other mobile genetic
292 elements affect their transmission and persistence ☆. 2019;102: 29–36.
- 293 2. Smillie C, Garcilla MP, Francia MV, Rocha EPC, Cruz F De. Mobility of Plasmids †. 2010;74:
294 434–452. doi:10.1128/MMBR.00020-10
- 295 3. Brown SP, Cornforth DM, Mideo N. Evolution of virulence in opportunistic pathogens□:
296 generalism , plasticity , and control §. Trends Microbiol. 2014;20: 336–342.
297 doi:10.1016/j.tim.2012.04.005
- 298 4. Foley SL, Kaldhøne PR, Ricke SC, Han J. Incompatibility Group I1 (IncI1) Plasmids: Their
299 Genetics, Biology, and Public Health Relevance. 2021;1.
- 300 5. Neill et al. TACKLING DRUG-RESISTANT INFECTIONS GLOBALLY□: FINAL REPORT
301 AND RECOMMENDATIONS THE REVIEW ON. 2016.
- 302 6. Iskandar K, Molinier L, Hallit S, Sartelli M, Catena F, Coccolini F, et al. Drivers of Antibiotic
303 Resistance Transmission in Low- and Middle-Income Countries from a “ One Health ”
304 Perspective — A Review. 2020.
- 305 7. Sonda et al. Whole genome sequencing reveals high clonal diversity of Escherichia coli isolated

from patients in a tertiary care hospital in. 2018; 1–12.

8. Kumburu HH, Sonda T, Mmbaga BT, Alifrangis M, Lund O, Kibiki G, et al. Patterns of infections, aetiological agents and antimicrobial resistance at a tertiary care hospital in northern Tanzania. *Trop Med Int Heal*. 2017;22: 454–464. doi:10.1111/tmi.12836

9. Sonda T, Kumburu H, van Zwetselaar M, Alifrangis M, Mmbaga BT, Lund O, et al. Prevalence and risk factors for CTX-M gram-negative bacteria in hospitalized patients at a tertiary care hospital in Kilimanjaro, Tanzania. *Eur J Clin Microbiol Infect Dis*. 2018;37: 897–906. doi:10.1007/s10096-018-3196-8

10. Souvorov A, Agarwala R, Lipman DJ. Open Access SKESA: strategic k-mer extension for scrupulous assemblies. 2018; 1–13.

11. Chaumeil P, Mussig AJ, Parks DH, Hugenholtz P. Genome analysis GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. 2020;36: 1925–1927. doi:10.1093/bioinformatics/btz848

12. Wick RR, Judd LM, Gorrie CL, Holt KE. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. 2017; 1–22. doi:10.1371/journal.pcbi.1005595

13. Arredondo-alonso S, Bootsma M, Rogers MRC, Corander J, Willems RJL, Schu AC. Genome analysis gplas: a comprehensive tool for plasmid analysis using short-read graphs. 2020;36: 3874–3876. doi:10.1093/bioinformatics/btaa233

14. Krawczyk PS, Lipinski L, Dziembowski A. PlasFlow: predicting plasmid sequences in metagenomic data using genome signatures. 2018; 1–14. doi:10.1093/nar/gkx1321

- 326 15. Carattoli A, Zankari E, García-fernández A, Larsen V, Lund O, Villa L, et al. In Silico
327 Detection and Typing of Plasmids using PlasmidFinder and Plasmid Multilocus Sequence
328 Typing. 2014;58: 3895–3903. doi:10.1128/AAC.02412-14
- 329 16. Bortolaia V, Kaas RS, Ruppe E, Roberts MC, Schwarz S, Cattoir V, et al. ResFinder 4.0 for
330 predictions of phenotypes from genotypes. J Antimicrob Chemother. 2020;75: 3491–3500.
331 doi:10.1093/jac/dkaa345
- 332 17. Joensen KG, Scheutz F, Lund O, Hasman H, Kaas RS, Nielsen EM, et al. Real-time whole-
333 genome sequencing for routine typing, surveillance, and outbreak detection of verotoxigenic
334 Escherichia coli. J Clin Microbiol. 2014;52: 1501–1510. doi:10.1128/JCM.03617-13
- 335 18. Devi Y, Mathur P, Pragasam AK. Plasmid profiles among some ESKAPE pathogens in a
336 tertiary care centre in south India. 2019. doi:10.4103/ijmr.IJMR
- 337 19. Sonda T, Kumburu H, Zwetselaar M Van, Alifrangis M, Lund O, Kibiki G, et al. Meta-analysis
338 of proportion estimates of producing Enterobacteriaceae in East Africa hospitals. Antimicrob
339 Resist Infect Control. 2016. doi:10.1186/s13756-016-0117-4
- 340 20. Malachowa N, Deleo FR. Mobile genetic elements of Staphylococcus aureus. 2010; 3057–3071.
341 doi:10.1007/s00018-010-0389-4
- 342 21. Ramirez MS, Iriarte A, Reyes-Lamothe R, Sherratt DJ, Tolmasky ME. Small Klebsiella
343 pneumoniae Plasmids: Neglected Contributors to Antibiotic Resistance. Front Microbiol.
344 2019;10: 1–14. doi:10.3389/fmicb.2019.02182
- 345 22. Mordi RM, Momoh MI. Incidence of Proteus species in wound infections and their sensitivity

pattern in the University of Benin Teaching Hospital. 2009;8: 725–730.

23. Endimiani A, Luzzaro F, Brigante G, Perilli M, Lombardi G, Amicosante G, et al. *Proteus mirabilis* Bloodstream Infections: Risk Factors and Treatment Outcome Related to the Expression of. 2005;49: 2598–2605. doi:10.1128/AAC.49.7.2598

24. Tsakris A, Poulou A, Markou F, Pitiriga V, Piperaki ET, Kristo I, et al. Dissemination of clinical isolates of *Klebsiella oxytoca* harboring CMY-31, VIM-1, and a new OXY-2-type variant in the community. *Antimicrob Agents Chemother*. 2011;55: 3164–3168. doi:10.1128/AAC.00102-11

25. Boyd DA, Lefebvre B, Mataseje LF, Gagnon S, Roger M, Savard P, et al. *Enterobacter* sp. N18-03635 harbouring blaFRI-6 class A carbapenemase, Canada. *J Antimicrob Chemother*. 2020;75: 486–488. doi:10.1093/jac/dkz438

26. Rodríguez-Medina N, Barrios-Camacho H, Duran-Bedolla J, Garza-Ramos U. *Klebsiella variicola*: an emerging pathogen in humans. *Emerg Microbes Infect*. 2019;8: 973–988. doi:10.1080/22221751.2019.1634981

27. Manter DK, Hunter WJ, Vivanco JM. *Enterobacter soli* sp. nov.: A lignin-degrading γ -Proteobacteria isolated from soil. *Curr Microbiol*. 2011;62: 1044–1049. doi:10.1007/s00284-010-9809-9

28. Bhassu S, brahim Eid II, Annie Tan GY. First report of *Enterobacter soli* associated with gastro-hepato infection of farmed fish, *Channa striata* in Malaysia. *Gastroenterol Hepatol Open Access*. 2019;10: 61–65. doi:10.15406/ghoa.2019.10.00357

- 366 29. Kuo et al. Whole-Genome Sequence of the Soil Bacterium *Micrococcus* sp. KBS0714. 2017;5:
367 23–24. Available: <http://arxiv.org/abs/1303.2223>
- 368 30. Dai H, Wang Y, Fang Y, Xiao T, Huang Z, Kan B, et al. *Proteus columbae* sp. nov., isolated
369 from a pigeon in Ma'an Shan, China. *Int J Syst Evol Microbiol*. 2018;68: 552–557.
370 doi:10.1099/ijsem.0.002541
- 371 31. Potter et al. Population Structure, Antibiotic Resistance, and Uropathogenicity of. 2018;9: 1–17.
- 372 32. Jousset et al. False-Positive Carbapenem-Hydrolyzing Confirmatory Tests Due to ACT-28, a
373 Chromosomally Encoded AmpC with Weak Carbapenemase Activity from *Enterobacter kobei*
374 Agnè. 2019;63: 1–11.
- 375 33. Rodríguez-medina N, Barrios-camacho H, Duran- J, Garza-ramos U, Barrios-camacho H,
376 Duran-bedolla J. *Klebsiella variicola*: an emerging pathogen in humans. 2019.
377 doi:10.1080/22221751.2019.1634981
- 378 34. Beceiro A, Tomás M, Bou G. Antimicrobial resistance and virulence: A successful or
379 deleterious association in the bacterial world? *Clin Microbiol Rev*. 2013;26: 185–230.
380 doi:10.1128/CMR.00059-12
- 381 35. Cassiano I, Santos DO, Mattos R, Dutra M, Paula A, Carvalho-assef DA. *Journal of Global*
382 *Antimicrobial Resistance* Draft genome sequence of KPC-2-producing *Pseudomonas*
383 *aeruginosa* recovered from a bloodstream infection sample in Brazil. *Integr Med Res*. 2018;15:
384 99–100. doi:10.1016/j.jgar.2018.08.021
- 385 36. Cerdeira LT, Lam MMC, Wyres KL, Ribas RM, Morais MM, Holt KE. Small IncQ1 and Col-

Like Plasmids Harboring blaKPC-2 and Non-Tn4401 Elements (NTEKPC-IIId) in High-Risk Lineages of *Klebsiella pneumoniae* CG258 Louise. 2019; 1–4.

37. Mcmillan EA, Jackson CR, Frye JG, Thomas CM, Mcmillan EA. Transferable Plasmids of *Salmonella enterica* Associated With Antibiotic Resistance Genes. 2020;11. doi:10.3389/fmicb.2020.562181

38. Erchia AMD, Pazzani C. A novel group of IncQ1 plasmids conferring multidrug resistance Oliva. *Plasmid*. 2016. doi:10.1016/j.plasmid.2016.11.005

39. Rawlings DE, Tietze E. Comparative Biology of IncQ and IncQ-Like Plasmids. 2001;65: 481–496. doi:10.1128/MMBR.65.4.481

40. Lima GJ De, Maria A, Scavuzzi L, Maria E, Beltrão B. Short Communication Identification of plasmid IncQ1 and NTE KPC -IIId harboring bla KPC-2 in isolates from *Klebsiella pneumoniae* infections in patients from Recife-PE , Brazil. 2020; 1–5. doi:10.1590/0037-8682-0526-2019

41. Xie M, Li R, Liu Z, Wai E, Chan C, Chen S. Recombination of plasmids in a carbapenem-resistant NDM-5-producing clinical *Escherichia coli* isolate. 2018; 1230–1234. doi:10.1093/jac/dkx540

42. Donnell GMC. Antiseptics and Disinfectants□: Activity , Action , and Resistance. 1999;12: 147–179.

43. Smalla K, Jechalke S, Top EM. HHS Public Access. 2015;3: 1–21. doi:10.1128/microbiolspec.PLAS-0038-2014.Plasmid

44. Pan Y, Zeng J, Li L, Yang J, Tang Z, Xiong W, et al. crossm Coexistence of Antibiotic

Resistance Genes and Virulence Factors Deciphered by Large-Scale Complete Genome
Analysis. 2020.

45. Dionisio F. Antibiotic Resistance Gene Diversity and Virulence Gene Microbiomes. 2019.