

Title: Pan-resistome characterization of uropathogenic *Escherichia coli* and *Klebsiella pneumoniae* strains circulating in Uganda and Kenya isolated from 2017-2018

Authors: Arun Gonzales Decano¹, Kerry Pettigrew¹, Wilber Sabiiti¹, Derek Sloan¹,
Stella Neema², Joel Bazira³, John Kiiru⁴, Hellen Atieno Onyango⁵, Benon Asiimwe⁶,
Matthew T. G. Holden¹

¹ School of Medicine, University of St Andrews, St Andrews, UK.

² Department of Sociology and Anthropology, Makerere University, Kampala, Uganda.

³ Faculty of Medicine, Mbarara University of Science and Technology, Mbarara, Uganda.

⁴ Centre of Microbiology Research, Kenya Medical Research Institute, Kenya.

⁵ Jomo Kenyatta University of science Technology, Kenya

⁶ Department of Medical Microbiology, Makerere University College of Health Sciences, Kampala, Uganda.

Abstract

Background: Urinary tract infection (UTI) develops after a pathogen adhered to the inner lining of the urinary tract. Cases of UTIs are predominantly caused by several Gram-negative bacteria and account for high morbidity in the clinical and community settings. Of greater concern are the strains carrying antimicrobial resistance (AMR)-conferring genes. The gravity of UTI is also determined by a spectrum of other virulence factors. This study represents a pilot project to investigate the burden of AMR among uropathogens in East Africa.

Methods: We examined bacterial samples isolated in 2017-2018 from in- and out-patients in Kenya (KY) and Uganda (UG) that presented with clinical symptoms of UTI. We reconstructed the evolutionary history of the strains, investigated their population structure and performed comparative analysis their pangenome contents.

Results: We found 55 *Escherichia coli* and 19 *Klebsiella pneumoniae* strains confirmed uropathogenic following screening for the prevalence of UTI virulence genes including *fimH*, *iutA*, *feoA/B/C*, *mrkD* and *foc*. We identified 18 different sequence types in *E. coli* population while all *K. pneumoniae* strains belong to ST11. The most prevalent *E. coli* sequence types were ST131 (26%), ST335/1193 (10%) and ST10 (6%). Diverse plasmid types were observed in both collections such as Incompatibility (IncF/IncH/IncQ1/IncX4) and Col groups. Pangenome analysis of each set revealed a total of 2,862 and 3,464 genes comprised the core genome of *E. coli* and *K. pneumoniae* population, respectively. Among these are AMR determinants including fluoroquinolone resistance-conferring genes *aac(3)-Ib-cr* and other significant genes: *aad*, *tet*, *sul1*, *sul2*, and *cat*, which are associated with aminoglycoside, tetracycline, sulfonamide and chloramphenicol resistance. Accessory genomes of both species collection were detected several β -lactamase genes, *bla_{CTX-M}*, *bla_{TEM}* and *bla_{OXA}* or *bla_{NDM}*. Overall, 93% are multi-drug resistant in the *E. coli* collection while 100% of the *K. pneumoniae* strains contained genes that are associated with resistance to 3 or more antibiotic classes.

Conclusions: Our findings illustrate the abundant resistome and virulome repertoire in uropathogenic *E. coli* and *K. pneumoniae*, which are mainly disseminated via clonal and horizontal transfer, circulating in the East African region. We further demonstrate here that routine genomic surveillance is necessary for high-resolution bacterial epidemiology of these important AMR pathogens.

Introduction

Antimicrobial resistance (AMR) has raised alarms as a global health threat. AMR is often fuelled by misuse and abuse of antibiotics including self-medication (Michael et al., 2014; Rather et al. 2017) and unrestricted access to antimicrobial drugs (Nathan and Cars 2014, French 2010, Read and Woods 2014), and is further accelerated by industrialization, poor waste disposal and hygiene levels. AMR pathogens are frequently detected in food, clinical and environmental settings in East Africa and despite facing broad challenges, significant efforts have recently been put in place to curb AMR in East African countries. For instance, Kenya (KY) has adapted the National Action Plan that incorporates One Health measures to prevent AMR and is highly supported by multiple governmental policies (NAPCAR 2017). Similarly, an extensive evaluation of the AMR situation in Uganda (UG) was assessed by the Uganda National Academy of Sciences (UNAS) supported by the Global Antibiotic Resistance Partnership (GARP)-Uganda (UNAS 2015). High prevalence of multi-drug resistant bacteria particularly extended-spectrum beta-lactamase (ESBL)-producing strains is significantly recorded in both countries.

Urinary tract infection (UTI) develops after a pathogen's adherence to the inner lining of the urinary tract. UTIs occur among patients of all age groups and account for high morbidity in the clinical and community settings (Flores-Mireles 2015). Following binding within the urinary tract, uropathogens either cause asymptomatic or commensal connection, or severe disease. About 1% of the population have asymptomatic bacteriuria (ABU), wherein a pathogen ($\geq 10^5$ cfu mL⁻¹) inhabits the tract without eliciting mucosal host response (Kunin et al. 1962; Lindberg et al. 1975). Infections in the lower urinary tract region e.g. cystitis, are recognized by symptoms such as dysuria. Successful virulent strains can induce pyelonephritis where rapid immune response is mobilised via cytokine secretion and influx of immune cells. UTIs are either uncomplicated or complicated. Uncomplicated UTI cases are usually observed in patients who are otherwise healthy, while complicated UTIs are diagnosed in compromised patients e.g. if have anatomical or functional anomalies in their urinary tract or are under long-term catheterization (Najar et al. 2009). Treatment of these complicated UTI cases is often confounded by AMR uropathogens usually caused by Gram-negative bacteria (Levison and Kaye 2013). Uncomplicated UTIs are frequently caused by uropathogenic *Escherichia coli* (UPEC) while complicated cases might be caused by several pathogens such as *Proteus mirabilis*, *Providencia stuartii*, *Morganella morganii*, *Klebsiella pneumoniae* and *Pseudomonas aeruginosa* (Flores-Mireles et al. 2015). Recurrent UTI cases are also common, particularly when urinary tract anomalies linger, or treatment failed to kill resistant bacteria (Nicolle LE 2000), leading to more severe type of infections. Due to the lack of active investigation of UTI cases in East Africa, particularly in the community, access to accurate data can be challenging.

An increasing number of studies have employed whole genome sequencing (WGS) and analyses for disease surveillance in both hospital and community settings (Croucher et al. 2015; Goswami et al. 2018). The high resolution genotyping that WGS provides can be used to investigate and describe the population structure, and evolutionary history of the isolates, as well as tracing their spread. Outbreaks have been robustly detected and described using high-throughput methodologies designed for bacterial pathogens (Halachev et al. 2014; Lee et al. 2018; Ludden et al. 2020). Comprehensive AMR gene databases and prediction tools are also available that help assess AMR gene content in whole genomes with high accuracy (Feldgarden et al. 2019). Here, we used WGS to investigate the prevalence of AMR-conferring genes in *E. coli* and *K. pneumoniae* isolated from urine samples taken from patients in rural areas of KY and UG that presented UTI-like symptoms. We further explored their phylogenetic relationships with other currently circulating African and global strains. This study represents a pilot project of the HATUA consortium. HATUA stands for Holistic Approach to Unravel Antibiotic Resistance in East Africa and the team is comprised of researchers from different disciplines that aim to tackle the main drivers of AMR among uropathogens in East Africa.

Methods

Study design and Patient recruitment

A total of N=150 bacterial isolates were obtained from patients in KY (n=91) and UG (n=59) presenting UTI-like symptoms, as part of a larger study. Ethical Review Board of University of St Andrews ethical approval, Approval code MD14548 and KY (KEMRI/SERU/P00112/3865) approved verbal consent taken from all the patients. Important patient data such as name, age, gender, location was recorded, and unique identification number were assigned to each patient.

Library preparation and whole genome sequencing

Bacterial genomic DNA for the isolates were extracted using the QIAextractor (Qiagen, Valencia, CA, USA) according to the manufacturer's instructions. Library preparation was conducted according to the Illumina protocol and sequenced (96-plex) on an Illumina MiSeq X platform (Illumina, San Diego, CA, USA) using 250 bp paired-end reads.

Read library quality control, mapping and de novo genome assembly

Illumina MiSeq read libraries were rid of sequencing adapters and ambiguous bases using Fastp (Chen et al. 2018). Sets that passed the quality filtering were *de novo* assembled using Unicycler v4.6 (Wick et al. 2017) pipeline in normal mode to merge contigs.

The read libraries were mapped to reference sequences using SMALT v7.6 (<http://www.sanger.ac.uk/resources/software/SMALT/>) and the resulting SAM files were converted to BAM format, sorted and PCR duplicates removed using SAMtools v1.19 (Li et al. 2009). Strain TOP52_1721_U1 (Wiley et al. 2019) was used the reference genome for the *K. pneumoniae* samples while the strain EC958 (Forde et al. 2014) was employed as the reference sequence for the *E. coli* population.

Species identification, sero- and sequence typing, genome annotation and screening for UTI virulence/AMR genes

Prediction of bacterial species was done by uploading the assemblies on PathogenWatch website (<https://pathogen.watch>), which runs Speciator (<https://gitlab.com/cgps/mash-speciator>) for its species assignment. Speciator employs Mash (Ondov et al. 2016) to identify the most identical strain ($\geq 90\%$ identity) in a reference collection of complete genomes found in the NCBI RefSeq database (<https://www.ncbi.nlm.nih.gov/refseq/>). The strains are then grouped according to their species designation and were screened for UTI pathogen determinants. Multi-locus sequence typing was performed by running SRST2 v.0.2.0 (Inouye et al. 2014) based on the Achtman scheme (Achtman et al. 2012) for *E. coli* and Pasteur (Diancourt et al. 2010) for *K. pneumoniae* isolates. Antigenic (O polysaccharide and H flagellin) profiles of *E. coli* samples were identified by employing Serotypefinder v.2.0 (<https://cge.cbs.dtu.dk/services/SerotypeFinder/>) at 85% ID threshold and 60% minimum length.

Genome composition of the draft assemblies was assessed using Prokka v.1.10 (Seemann 2014). AMR genes were identified by aligning the genome sequences to the 2,158 gene homolog subset of the Comprehensive Antibiotic Resistance Database (CARD) v. 3.0.8 (<https://card.mcmaster.ca/>). Plasmid and replicon typing was done by comparing the genomes against the PlasmidFinder database v. (Carattoli et al. 2014) at 99% identity threshold.

Bacterial sample collection and Antimicrobial Susceptibility Testing

To determine concordance between the AMR gene content and sample phenotype, antibiotic testing and phenotypic detections of ESBL were performed by disc diffusion methods on a subset of n=16 isolates from

KY; the tests are done according to CLSI, 2016 guidelines. Isolates were examined for the insusceptibility to 9 different classes of antibiotics including Penicillin (ampicillin (AMP)), Penicillin + β -lactamase inhibitors (ampicillin-clavulanic acid (AMC)), Chloramphenicol (Chloramphenicol (CHL)), Sulfonamide (Trimethoprim-sulfamethoxazole (SXT)) and Quinolones (nalidixic acid (NA)) and Fluoroquinolone (Ciprofloxacin (CIP)). Resistance to ESBL Cephalosporins was also assessed by testing the strains with Ceftriaxone (CRO), Ceftazidime (CAZ), cefotaxime (CTX) and Cefepime (FEP) (Supplementary Table 1).

Phylogenetic reconstruction

Phylogenetic relationships and sequence variations between the samples were determined by constructing phylogenetic trees based on their chromosomal SNPs; MGEs were further excluded using an internal script. Non-recombinant SNPs were determined using ClonalframeML v. 1.12 (Didelot and Wilson 2015) and were used to create a maximum-likelihood midpoint-rooted phylogeny using RAxML v8.0.19 (Stamatakis 2014) using a General Time Reversible + gamma (GTR+G) substitution model with 100 bootstraps. Phylogenies were visualized using iTOL (<https://itol.embl.de/>) and FigTree v1.4.3 (<http://tree.bio.ed.ac.uk/software/figtree/>).

Pangenome analyses

The resulting annotation files from Prokka v.1.10 (Seemann 2014) were used as the basis for generating a pangenome for each species set. This step was done by running Roary v3.11.2 (Page et al. 2015) with a 100% BLAST v2.6.0 identity threshold using the MAFFT v7.3 setting (Kato and Standley 2013). Pangenome outputs were also used to assess the accessory genome composition of each bacterial population and as basis for constructing core genome phylogenies.

Results

Patient and bacterial strain profiles

A collection of N=100 strains, n=81 *E. coli* and n=19 *K. pneumoniae* were isolated from urine of different patients admitted or visiting rural hospitals in Kenya and from clinics in the countryside of Uganda.

*Genomic and pangenomic characterization confirmed the virulence factors present in uropathogenic *E. coli* and *K. pneumoniae**

A subset (n=55) from the total n=81 *E. coli* and all n=19 *K. pneumoniae* were confirmed uropathogenic following a thorough characterization of their pangenome contents. One thousand one hundred forty-four (1,144) and 3,464 core genes were found across the strains in *E. coli* and *K. pneumoniae* populations,

respectively. These include known UTI virulence markers that are responsible for urinary tract (mucosal) surface binding (type 1 fimbrial adhesin-coding *fimH*) and colonization (*mrkD*; *K. pneumoniae* only), iron (Fe(2+)) transport (*feoA/B/C*), enterobactin synthase production (*entB*), formate transport (*focA*), cell division (*zapA*), succinate-acetate/proton symport (*satP*), anaerobic sulfatase-maturation (*chuR*; found in 100% and 95% of *E. coli* and *K. pneumoniae*, respectively). Other important virulence genes were also found, albeit not conserved among all the isolates: *iutA* (ferric aerobactin receptor: 44.6% in *E. coli*, 10% in *K. pneumoniae*), *papA* (fimbrial major pilin protein: *E. coli* only (41%)), *papD* (import of P pilus subunits into the periplasm: 44.6% in *E. coli*, 10% in *K. pneumoniae*), *hlyE* (hemolysin E: 80.4% in *E. coli*, 10% in *K. pneumoniae*), *fyuA* (pesticin receptor: 73.2% in *E. coli*, 15% in *K. pneumoniae*), *kpsT* (polysialic acid transport ATP-binding protein: *E. coli* only (26.8%)) and *pic* (serine protease pic autotransporter: *E. coli* only (5.4%); Table 1).

Gene	Protein Product	Strain count and % in <i>E. coli</i> collection	Strain count and % in <i>K. pneumoniae</i> collection
<i>fimH</i>	Type 1 fimbrin D-mannose specific adhesin	56/56 (100)	20/20 (100)
<i>feoA/B/C</i>	Fe(2+) transport protein A/B/C	56/56 (100)	20/20 (100)
<i>entB</i>	Enterobactin synthase component B	56/56 (100)	20/20 (100)
<i>focA</i>	Formate transporter	56/56 (100)	20/20 (100)
<i>zapA</i>	Cell division protein	56/56 (100)	20/20 (100)
<i>satP</i>	Succinate-acetate/proton symporter	56/56 (100)	20/20 (100)
<i>chuR</i>	Anaerobic sulfatase-maturing enzyme	56/56 (100)	19/20 (95)
<i>mrkD</i>	Type 3 fimbrial adhesin	0	20/20 (100)
<i>hlyE</i>	Hemolysin E	45/56 (80.4)	2/20 (10)
<i>fyuA</i>	Pesticin receptor	41/56 (73.2)	3/20 (15)
<i>iutA</i>	Ferric aerobactin receptor	25/56 (44.6)	2/20 (10)
<i>papD</i>	Import of P pilus subunits into the periplasm	26/56 (46.4)	2/20 (10)
<i>papA</i>	Fimbrial major pilin protein	23/56 (41.1)	0
<i>kpsT</i>	Polysialic acid transport ATP-binding protein	15/56 (26.8)	0
<i>pic</i>	Serine protease pic autotransporter	3/56 (5.4)	0

Table 1. UTI virulence marker genes present in the pangenome of N=55 *E. coli* and N=19 *K. pneumoniae* isolates. Proportion of the samples containing the gene are shown in count of strains with gene over the total strains and % values.

Prevalence of AMR genes in *E. coli* and *K. pneumoniae* uropathogens from KY and UG

All n=55 *E. coli* and n=19 *K. pneumoniae* isolates harboured type 1 fimbrin. Among the UPEC, *fimH30* was the most common allele, followed by *fimH41*; n=4/55 samples had type *fimH22* and n=2/55 singleton were found with *fimH22*.

We further detected multiple AMR-conferring elements in the genomes of the two species collections. Alignment of the sequences against CARD v.3.0.7 with 98-100% identity revealed that the n=55 *E. coli* (n=31 from KY, n=24 from UG) were detected with the ciprofloxacin-conferring gene, *marA*. Majority (n=47/55) were also aminoglycoside resistant and harbours either *aadA* or *aac(6')-Ib/(3')-Ib* alleles or both. Only n=11/55 were not detected with resistant genes for ESBL cephalosporins. Of the n=44/55 that produce

ESBLs, n=10/44 had *bla*_{CTX-M} (allele type 15 or 88), n=24/44 had *bla*_{TEM} (type 30/2/220) and n=9/44 had *bla*_{OXA-1/140} and n=2/44 (both from UG) had all 3 ESBL genes. Sulfonamide resistance was widely observed, n=39 had either *sul1* only, *sul2* only or *sul3* only, or both *sul1* and *sul2*. Tetracycline resistance gene, *tet(A)* was present in n=34 of 55. Twenty-four (N=24/55) contained macrolide resistance-conferring *mphA*, n=9/55 (n=6 from KY, n=3 from UG) harboured *catB3* and were chloramphenicol resistant. Ninety-five percent (95%, n=52/55) had at least one gene that codes for efflux pump proteins with n=4/55 having *yojI-pmrF-emrR-bacA-acrS/B/E-msbA-avgA-kdpE-mdtP-eptA* cassette and n=1/55 containing a mixture of *yojI*, *pmrF*, *emrR*, *bacA*, *acrS/B/E*, *msbA*, *avgA*, *kdpE*, *mdtP*, *eptA*, *emtK*, *cpxA* (Table 2a). Two KY isolates (71 and 72) were found to have the fluoroquinolone resistance-conferring gene *aac(3')-Ib-cr* while its variant *aac(6')-Ib-cr* was present in UG isolates BN19, BN38 and BN44.

All n=19 *K. pneumoniae* isolates were resistant to aminoglycosides and had either *aac(6')-30/aac(6')-Ib/7/10-aadA9* (n=10/19) or *aac(6')-Ib7-aadA9* (n=1/19) combination. N=18/19 are potentially sulfonamide insusceptible and contained either *sul1* only (n=2/19), *sul2* only (n=15/19) or both (n=1). The β -lactamase *bla*_{LEN} gene is present in all but the BN7 strain, with *bla*_{LEN-4/6} (n=15 from KY) or *bla*_{LEN-3/4/5/6} alleles. BN7 was also the only susceptible isolate against ESBL cephalosporins. The rest are ESBL producers: n=15/19 had *bla*_{SHV-28}, *bla*_{CTX-M-15}, *bla*_{OXA-1/140} and *bla*_{NDM-1}, n=1/19 were observed with the *bla*_{SHV-28-CTX-M-15-OXA-1/140} combination and n=1/19 had *bla*_{SHV-28}, *bla*_{CTX-M-15/88} only. All n=3 strains from UG had *tet(B)* and *dfrA* (17 or 27 allele type); these strains also contained efflux pump-expressing genes: *yojI*, *pmrF*, *emrR*, *bacA*, *acrB*, *msbA*, *avgA*, *kdpE*, *mdtP*, *eptA*, *emtK* and *cpxA*. N=2/19 (BN14 and BN16 from UG) had ciprofloxacin-resistance gene, *marA*. Only the strain 90 from KY was not resistant to phenicols, while the rest were detected with the *cat* gene, specifically, *catB3* (Table 2b). Overall, 80% of our *E. coli* uropathogens had ESBL genes (n=15 strains from UG and n=29 from KY) and 93% of these UPEC are MDR, while all the *K. pneumoniae* isolates are MDR and only n=1 out of the total n=19 (95%) are ESBL.

Population structure of KY and UG Uropathogens

The UPEC collection was polyclonal. Eighteen (18) different sequence types were identified in the UPEC population (Achtman scheme). The most prevalent MLST sequence types were ST131 (n = 17/55, 31%), ST335 and ST1193 (n=6/55, 11%) and ST10 (n= 4/55, 7%). These sequence types were usually associated with UTI cases (Nicolas-Chanoine et al. 2014; Afset et al. 2008; Yamaji et al. 2018); the globally disseminated ESBL-ST131 stood out to be the most dominant ST. Other clones were also observed: n=3 ST73, n=2 each from ST155, ST410, ST6161 and ST162, and singletons from ST44, ST48, ST165, ST167, ST212, ST448, ST617, ST648 and ST2163; n=2 strains from UG (BN2 and BN48) were unclassified (Figure 1a, Table 2a). *E. coli* isolates from UG belong to 15 STs and were thus more diverse compared to those collected from KY, which belong to only 6 STs (Figure 1a, Table 2a). This difference in diversity is consistent with the number

of serotypes found in UG relative to those from KY: Ugandan strains belong to 20 different O:H antigen combinations while the KYn ones were found to have 9 O:H types.

Fifteen out of nineteen (n=15/19) *K. pneumoniae* isolates from KY belong to ST11 (Pasteur scheme); n=3/19 UG had no defined sequence types (BN14: 0b8e, BN16: 67b2, BN7: 6b6f) and formed their own clade (Table 2b; Figure 1b).

Plasmid characterization

Genome assemblies of the KY and UG uropathogens were screened for the presence and type of plasmids using PlasmidFinder v. N=47/55 in the *E. coli* collection were found with at least one plasmid. IncFIA was consistently found in n=10 had both *IncFIA* and *IncFII*, n=9 contained *IncFIA*, *IncFII*, *Col156* types, n=1 was detected with *IncFIA*, *IncFII* and *IncY* only or *IncI* only and *IncFII-IncFIA-IncX4* plasmid combinations.

All the samples from the *K. pneumoniae* collection were found with at least one plasmid type. *IncFII-IncFIB-IncR* is the most common combination and is found among n=15 isolates, while n=2/19 was found with *IncR*, *IncFII*, *IncFIA*, *Col* and *IncX4*. Notably, the strain 90 from KY had the *bla_{CTX-M-1}* gene-carrying plasmid *IncN* and BN7 from UG had the *bla_{NDM}*-associated *IncR*.

Discussion

We assessed the prevalence of AMR characteristics among uropathogenic *E. coli* and *K. pneumoniae* circulating in East African region using WGS. We recruited out-patients that presented UTI-like symptoms from rural areas in KY and UG, which represents a limitation of our sample collection. The lack of point-of-care diagnostic tool such as the use of dipstick test also contributed to some difficulties in our screening. This is evidenced by a high level of contaminants that comprised of strains that do not contain UTI determinants. Nevertheless, our *in silico* predictions using whole genome analysis revealed alarming rates of ESBL-producing and MDR strains in both our UPEC and *K. pneumoniae* collections, which reiterates the great necessity for effective interventions to curb their spread.

Our results firmly indicate a high diversity among *E. coli* uropathogens, which was more evident in samples taken from UG rather than KY. Strains that belong to the same clonal group had <200 core SNPs from each other. This rich genetic diversity is consistent with those observed in other isolates collected from rural or semi-rural communities of low/middle-income countries (Salinas et al. 2019; Kimata et al. 2005; Montealegre et al. 2020). The widely-disseminated UTI-causing clones ST131, ST335 and ST10 were common among our *E. coli* strains and dominated our Kenyan collection. This is unsurprising as these ST's are reported to be circulating globally (Ben Zakour 2016; Sanjar et al. 2015; Day et al. 2019). What's remarkable is the detection

of emerging clones such as ST1193 and ST617 that were unusually associated with UTI (Valenza et al. 2019; Mhaya et al. 2019) albeit observed in hospital settings. UPEC strains from UG are even more alarming as they represent higher number of unusual or novel UTI clones i.e. ST155, ST448 and ST162 with potentially higher virulence levels (Castellanos et al. 2017; Choudhury 2018; Zhang et al. 2019) compared to those globally-known STs.

*bla*_{CTX-M} genes were present in 40% of our UPEC collection and in all but one *K. pneumoniae* strain (95%). Notably, the *bla*_{CTX-M-15} gene that confers resistance to last-resort antibiotics was found in high levels in both countries. This gene was detected with other ESBL determinants, *bla*_{TEM} and *bla*_{OXA-1} in *E. coli* and with *bla*_{NDM} in *K. pneumoniae*, concordant with those in uropathogens found from the Middle East (Alqasim et al. 2018) and Asia (Zhang et al. 2015), among others. Consistent with previous findings in other African regions, *tet* genes in this study were also detected alongside ESBL genes *bla*_{CTX-M-15}, *bla*_{OXA-1} and *bla*_{TEM} in n=30/55 *E. coli* and with *bla*_{LEN-3/4/5/6} among n=3/19 *K. pneumoniae* (Mbelle et al. 2019; Adefisoye et al. 2016) which stipulates their co-selection and co-transmission in KY and UG. The presence of these genes in the identified plasmid-associated contigs suggest that the mode of transfer may have been plasmid-mediated.

Taken together, we underline in this pilot study the high frequency of AMR determinants associated with resistance to common antibiotic classes among *E. coli* and *Klebsiella pneumoniae* in East Africa, with specific focus on MDR and ESBL-producing strains from KY and UG. We further demonstrate that routine genomic surveillance is necessary for high-resolution investigation of bacterial epidemiology especially in less represented regions. Our findings have significant implications on improving interventions that aim to address the strong presence of AMR pathogens that cause UTI particularly in low/middle-income countries.

Figures and Tables

Table 2. Genomic characteristics of uropathogenic *E. coli* (a) and *K. pneumoniae* strains (b) isolated in this study for AMR-associated genes and plasmid replicon types. NF means not found; Yes means the strain is either ESBL-producing or MDR and No means the sample is either non-ESBL or non-MDR. (View or download from Figshare, DOI: https://figshare.com/articles/Table_2a_and_2b/11965455).

Figure 1. Maximum likelihood phylogenies of core genomes of *E. coli* (a) and *K. pneumoniae* (b) uropathogens isolated from KY and UG; reference genomes are in red and font. The mid-point rooted phylograms was constructed using 2,862 and 3,464 core genes from *E. coli* and *K. pneumoniae* populations, respectively and visualized with iTol. The coloured strips adjacent to the *E. coli* phylogeny represent (from left to right) the sequence type (ST), country of origin, type of *fimH* allele and the ESBL status of each strain. The coloured rings around the *K. pneumoniae* phylogenetic tree indicate the ST, country of origin and the ESBL status of each isolate. The scale bar indicates the number of SNP distance from the reference.

Supplementary Materials

Supplementary Table 1. Metadata of *E. coli* and *K. pneumoniae* strains isolated from urine samples including Antibiotic Sensitivity Test results of n=16 isolates.

Acknowledgments

We thank Prof. Andy Lynch from the Mathematical Institute at the University of St. Andrews for his helpful feedback on the manuscript.

References

- Achtman M, Wain J, Weill FX, et al. Multilocus sequence typing as a replacement for serotyping in *Salmonella enterica*. *PLoS Pathog.* 2012;8(6):e1002776. doi:10.1371/journal.ppat.1002776.
- Adefisoye MA, Okoh AI. Identification and antimicrobial resistance prevalence of pathogenic *Escherichia coli* strains from treated wastewater effluents in Eastern Cape, South Africa. *Microbiologyopen.* 2016;5(1):143–151. doi:10.1002/mbo3.319.
- Afset JE, Anderssen E, Bruant G, Harel J, Wieler L, Bergh K. Phylogenetic backgrounds and virulence profiles of atypical enteropathogenic *Escherichia coli* strains from a case-control study using multilocus sequence typing and DNA microarray analysis. *J Clin Microbiol.* 2008 Jul;46(7):2280-90. doi: 10.1128/JCM.01752-07. Epub 2008 May 7. PMID: 18463209; PMCID: PMC2446914.
- Alqasim A, Abu Jaffal A, Alyousef AA. Prevalence of Multidrug Resistance and Extended-Spectrum β -Lactamase Carriage of Clinical Uropathogenic *Escherichia coli* Isolates in Riyadh, Saudi Arabia. *Int J Microbiol.* 2018 Sep 16;2018:3026851. doi: 10.1155/2018/3026851. PMID: 30305814; PMCID: PMC6165594.
- Ben Zakour NL, Alsheikh-Hussain AS, Ashcroft MM, et al. Sequential Acquisition of Virulence and Fluoroquinolone Resistance Has Shaped the Evolution of *Escherichia coli* ST131 [published correction appears in *MBio.* 2016 Jun 28;7(3):null]. *mBio.* 2016;7(2):e00347-16. Published 2016 Apr 26. doi:10.1128/mBio.00347-16.
- Carattoli A, Zankari E, García-Fernández A, et al. In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. *Antimicrob Agents Chemother.* 2014;58(7):3895–3903. doi:10.1128/AAC.02412-14.
- Castellanos LR, Donado-Godoy P, León M, Clavijo V, Arevalo A, Bernal JF, Timmerman AJ, Mevius DJ, Wagenaar JA, Hordijk J. High Heterogeneity of *Escherichia coli* Sequence Types Harboring ESBL/AmpC Genes on IncII Plasmids in the Colombian Poultry Chain. *PLoS One.* 2017 Jan 26;12(1):e0170777. doi: 10.1371/journal.pone.0170777. PMID: 28125687; PMCID: PMC5268450.
- Choudhury NA, Paul D, Chakravarty A, Bhattacharjee A, Dhar Chanda D. IncX3 plasmid mediated occurrence of bla_{NDM-4} within *Escherichia coli* ST448 from India. *J Infect Public Health.* 2018;11(1):111–114. doi:10.1016/j.jiph.2017.06.008.

Clinical and Laboratory Standards Institute. 2016. Reference method for broth dilution antifungal susceptibility testing of filamentous fungi; approved standard, 30th Ed. CLSI document C24-Ed4. Clinical and Laboratory Standards Institute, Wayne, PA

Croucher, N. J. & Didelot, X. The application of genomics to tracing bacterial pathogen transmission. *Current Opinion in Microbiology* 23, 62–67, <https://doi.org/10.1016/j.mib.2014.11.004> (2015).

Day MJ, Hopkins KL, Wareham DW, et al. Extended-spectrum β -lactamase-producing *Escherichia coli* in human-derived and foodchain-derived samples from England, Wales, and Scotland: an epidemiological surveillance and typing study. *The Lancet. Infectious Diseases*. 2019 Dec;19(12):1325-1335. DOI: 10.1016/s1473-3099(19)30273-7.

Didelot X, Wilson DJ. ClonalFrameML: efficient inference of recombination in whole bacterial genomes. *PLoS Comput Biol*. 2015 Feb 12;11(2):e1004041. doi: 10.1371/journal.pcbi.1004041. PMID: 25675341; PMCID: PMC4326465.

Flores-Mireles AL, Walker JN, Caparon M, Hultgren SJ. Urinary tract infections: epidemiology, mechanisms of infection and treatment options. *Nat Rev Microbiol*. 2015 May;13(5):269-84. doi: 10.1038/nrmicro3432. Epub 2015 Apr 8. PMID: 25853778; PMCID: PMC4457377.

Forde BM, Ben Zakour NL, Stanton-Cook M, et al. The complete genome sequence of *Escherichia coli* EC958: a high quality reference sequence for the globally disseminated multidrug resistant *E. coli* O25b:H4-ST131 clone. *PLoS One*. 2014;9(8):e104400. Published 2014 Aug 15. doi:10.1371/journal.pone.0104400.

Goswami, C. et al. Genetic analysis of invasive *Escherichia coli* in Scotland reveals determinants of healthcare-associated versus community-acquired infections. *Microb Genom* 4, e000190, <https://doi.org/10.1099/mgen.0.000190> (2018).

Halachev MR, et al. Genomic epidemiology of a protracted hospital outbreak caused by multidrug-resistant *Acinetobacter baumannii* in Birmingham, England. *Genome Med* 61070.2014; 10.1186/s13073-014-0070-x.

French GL. The continuing crisis in antibiotic resistance. *Int J Antimicrob Agents*. 2010;36 Suppl 3:S3–S7. doi:10.1016/S0924-8579(10)70003-0.

Hull RA, Gill RE, Hsu P, Minshew BH, Falkow S. 1981. Construction and expression of recombinant plasmids encoding type 1 or D-mannose-resistant pili from a urinary tract infection *Escherichia coli* isolate. *Infection and Immunity* 33:933-938.

Inouye M, Dashnow H, Raven LA, et al. SRST2: Rapid genomic surveillance for public health and hospital microbiology labs. *Genome Med.* 2014;6(11):90. Published 2014 Nov 20. doi:10.1186/s13073-014-0090-6.

Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol.* 2013;30(4):772–780. doi:10.1093/molbev/mst010.

Kimata K, Shima T, Shimizu M, et al. Rapid categorization of pathogenic *Escherichia coli* by multiplex PCR. *Microbiol Immunol.* 2005;49(6):485–492. doi:10.1111/j.1348-0421.2005.tb03752.x

Lee RS, Seemann T, Heffernan H, et al. Genomic epidemiology and antimicrobial resistance of *Neisseria gonorrhoeae* in New Zealand. *J Antimicrob Chemother.* 2018;73(2):353–364. doi:10.1093/jac/dkx405.

Levison ME, Kaye D. Treatment of complicated urinary tract infections with an emphasis on drug-resistant gram-negative uropathogens. *Curr Infect Dis Rep.* 2013;15(2):109–115. doi:10.1007/s11908-013-0315-7.

Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics.* 2009;25(16):2078–2079. doi:10.1093/bioinformatics/btp352.

Ludden, C. et al. Genomic surveillance of *Escherichia coli* ST131 identifies local expansion and serial replacement of subclones. *Biorxiv*, <https://doi.org/10.1101/814731> (2019).

Mbelle NM, Osei Sekyere J, Amoako DG, et al. Genomic analysis of a multidrug-resistant clinical *Providencia rettgeri* (PR002) strain with the novel integron In1483 and an A/C plasmid replicon. *Ann N Y Acad Sci.* 2020;1462(1):92–103. doi:10.1111/nyas.14237.

Mhaya A, Trabelsi R, Begu D, et al. Emergence of B2-ST131-C2 and A-ST617 *Escherichia coli* clones producing both CTX-M-15- and CTX-M-27 and ST147 NDM-1 positive *Klebsiella pneumoniae* in the Tunisian community. *bioRxiv*; 2019. DOI: 10.1101/713461.

Michael CA, Dominey-Howes D, Labbate M. The antimicrobial resistance crisis: causes, consequences, and management. *Front Public Health.* 2014;2:145. Published 2014 Sep 16. doi:10.3389/fpubh.2014.00145.

Najar MS, Saldanha CL, Banday KA. Approach to urinary tract infections. *Indian J Nephrol.* 2009 Oct;19(4):129–39. doi: 10.4103/0971-4065.59333. PMID: 20535247; PMCID: PMC2875701.

Nathan C, Cars O. Antibiotic resistance--problems, progress, and prospects. *N Engl J Med.* 2014;371(19):1761–1763. doi:10.1056/NEJMp1408040.

Nicolas-Chanoine MH, Bertrand X, Madec JY. *Escherichia coli* ST131, an intriguing clonal group. *Clin Microbiol Rev.* 2014;27(3):543–574. doi:10.1128/CMR.00125-13.

Nicolle LE. Pivmecillinam in the treatment of urinary tract infections. *J Antimicrob Chemother.* 2000;46 Suppl A:35–39.

Ondov BD, Treangen TJ, Melsted P, et al. Mash: fast genome and metagenome distance estimation using MinHash. *Genome Biol.* 2016;17(1):132. Published 2016 Jun 20. doi:10.1186/s13059-016-0997-x.

Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MT, Fookes M, Falush D, Keane JA, Parkhill J. Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics.* 2015 Nov 15;31(22):3691-3. doi: 10.1093/bioinformatics/btv421. Epub 2015 Jul 20. PMID: 26198102; PMCID: PMC4817141.

Rather IA, Kim BC, Bajpai VK, Park YH. Self-medication and antibiotic resistance: Crisis, current challenges, and prevention. *Saudi J Biol Sci.* 2017 May;24(4):808-812. doi: 10.1016/j.sjbs.2017.01.004. Epub 2017 Jan 9. PMID: 28490950; PMCID: PMC5415144.

Read AF, Woods RJ. Antibiotic resistance management. *Evol Med Public Health.* 2014;2014(1):147. Published 2014 Oct 28. doi:10.1093/emph/eou024.

Salinas L, Cárdenas P, Johnson TJ, Vasco K, Graham J, Trueba G. Diverse Commensal *Escherichia coli* Clones and Plasmids Disseminate Antimicrobial Resistance Genes in Domestic Animals and Children in a Semirural Community in Ecuador. *mSphere.* 2019;4(3):e00316-19. Published 2019 May 22. doi:10.1128/mSphere.00316-19.

Sanjar F, Rusconi B, Hazen TH, Koenig SS, Mammel MK, Feng PC, Rasko DA, Eppinger M. Characterization of the pathogenome and phylogenomic classification of enteropathogenic *Escherichia coli* of the O157:non-H7 serotypes. *Pathog Dis.* 2015 Jul;73(5):ftv033. doi: 10.1093/femspd/ftv033. Epub 2015 May 10. PMID: 25962987; PMCID: PMC4849350.

Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics.* 2014;30(14):2068–2069. doi:10.1093/bioinformatics/btu153.

Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics.* 2014;30(9):1312–1313. doi:10.1093/bioinformatics/btu033.

Valenza G, Werner M, Eisenberger D, et al. First report of the new emerging global clone ST1193 among clinical isolates of extended-spectrum β -lactamase (ESBL)-producing *Escherichia coli* from Germany. *J Glob Antimicrob Resist*. 2019;17:305–308. doi:10.1016/j.jgar.2019.01.014.

Wick RR, Judd LM, Gorrie CL, Holt KE. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput Biol*. 2017;13(6):e1005595. Published 2017 Jun 8. doi:10.1371/journal.pcbi.1005595.

Wylie KM, Wylie TN, Minx PJ, Rosen DA. Whole-Genome Sequencing of *Klebsiella pneumoniae* Isolates to Track Strain Progression in a Single Patient With Recurrent Urinary Tract Infection. *Front Cell Infect Microbiol*. 2019 Feb 8;9:14. doi: 10.3389/fcimb.2019.00014. PMID: 30800643; PMCID: PMC6375827.

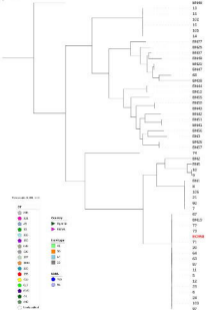
Yamaji R, Rubin J, Thys E, Friedman CR, Riley LW. Persistent Pandemic Lineages of Uropathogenic *Escherichia coli* in a College Community from 1999 to 2017. *J Clin Microbiol*. 2018;56(4):e01834-17. Published 2018 Mar 26. doi:10.1128/JCM.01834-17.

Zhao R, Shi J, Shen Y, Li Y, Han Q, Zhang X, Gu G, Xu J. Phylogenetic Distribution of Virulence Genes Among ESBL-producing Uropathogenic *Escherichia coli* Isolated from Long-term Hospitalized Patients. *J Clin Diagn Res*. 2015 Jul;9(7):DC01-4. doi: 10.7860/JCDR/2015/13234.6157. Epub 2015 Jul 1. PMID: 26393125; PMCID: PMC4572956.

Zhang P, Wang J, Wang X, Bai X, Ma J, Dang R, Xiong Y, Fanning S, Bai L, Yang Z. Characterization of Five *Escherichia coli* Isolates Co-expressing ESBL and MCR-1 Resistance Mechanisms From Different Origins in China. *Front Microbiol*. 2019 Aug 29;10:1994. doi: 10.3389/fmicb.2019.01994. PMID: 31555232; PMCID: PMC6727855.

Figure 1

a



b

