Whole genome sequencing reveals high clonal diversity of *Escherichia coli* isolated from patients in a tertiary care hospital in Moshi, Tanzania

Tolbert Sonda¹², Happiness Kumburu¹², Marco van Zwetselaar¹, Michael Alifrangis³, Blandina T. Mmbaga¹², Frank M. Aarestrup⁴, Gibson Kibiki²⁶ and Ole Lund⁵*

**Abstract**

**Background:** Limited information regarding the clonality of circulating *E. coli* strains in tertiary care hospitals in low and middle-income countries is available. The purpose of this study was to determine the serotypes, antimicrobial resistance and virulence genes. Further, we carried out a phylogenetic tree reconstruction to determine relatedness of *E. coli* isolated from patients in a tertiary care hospital in Tanzania.

**Methods:** *E. coli* isolates from inpatients admitted at Kilimanjaro Christian Medical Centre between August 2013 and August 2015 were fully genome-sequenced at KCMC hospital. Sequence analysis was done for identification of resistance genes, Multi-Locus Sequence Typing, serotyping, and virulence genes. Phylogeny reconstruction using CSI Phylogeny was done to ascertain *E. coli* relatedness. Stata 13 (College Station, Texas 77845 USA) was used to determine Cohen’s kappa coefficient of agreement between the phenotypically tested and whole genome sequence predicted antimicrobial resistance.

**Results:** Out of 38 *E. coli* isolates, 21 different sequence types (ST) were observed. Eight (21.1%) isolates belonged to ST131; of which 7 (87.5%) were serotype O25:H4. Ten (18.4%) isolates belonged to ST10 clonal complex; of these, four (40.0%) were ST617 with serotype O89:H10. Twenty-eight (73.7%) isolates carried genes encoding beta-lactam resistance enzymes. On average, agreement across all drugs tested was 83.9%. Trimethoprim/sulphamethoxazole (co-trimoxazole) showed moderate agreement: 45.8%, kappa =15% and \( p = 0.08 \). Amoxicillin-clavulanate showed strongest agreement: 87.5%, kappa = 74% and \( p = 0.0001 \). Twenty-two (57.9%) isolates carried virulence factors for host cells adherence and 25 (65.7%) for factors that promote *E. coli* immune evasion by increasing survival in serum. The phylogeny analysis showed that ST131 clustering close together whereas ST10 clonal complex had a very clear segregation of the ST617 and a mix of the rest STs.

**Conclusion:** There is a high diversity of *E. coli* isolated from patients admitted to a tertiary care hospital in Tanzania. This underscores the necessity to routinely screen all bacterial isolates of clinical importance in tertiary health care facilities. WGS use for laboratory-based surveillance can be an effective early warning system for emerging pathogens and resistance mechanisms in LMICs.

**Keywords:** *E. coli*, Multi-locus sequence typing, Serotyping, And virulence, Whole genome sequencing, Tanzania

* Correspondence: lund@cbs.dtu.dk
³ Centre for Biological Sequence Analysis, Technical University of Denmark, Copenhagen, Denmark

© The Author(s). 2018 Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated.
Background

*Escherichia coli* are Gram-negative bacterial commensals that naturally inhabit the human gastrointestinal tract (GIT). Through horizontal transfer and other mechanisms, commensal *E. coli* regularly acquire virulence, pathogenicity and multi-drug resistance properties from pathogenic *E. coli*. Consequently, *E. coli* is an important causative agent for a range of nosocomial and opportunistic infections including neonatal meningitis, diarrhoea, septicaemia, urinary tract and wound infections [1–5]. The global emergence and spread of multidrug resistant (MDR) *E. coli* in both community and as nosocomial infections, as well as in animals, warrants public health concerns [6–9]. Furthermore, virulent *E. coli* strains share resistance, virulence and pathogenic factors with avirulent or less virulent strains, enabling them to cause overlapping pathogenesis beyond their classical capacities [10]. Several *E. coli* outbreaks leading to serious health, social and economic impacts have been reported in high income countries (HICs) including the Netherlands [11], the UK [12], Norway, and Georgia [10].

Africa is highly burdened by diarrhoea and urinary tract infections (UTIs) that are *E. coli* related. For instance, in rural Kenya it was reported that 64.5% of UTI or bacteriuria was *E. coli* related [13]. In north-western Tanzania, 41.2 and 70% of UTI or bacteriuria was *E. coli* related among under-fives and febrile children, respectively [14, 15]. Similarly, in north-eastern part of Tanzania, *E. coli* accounted for 56.1% of all UTI cases among children [5] and *E. coli* accounted for 79 and 75% of UTI in non-malaria febrile children and adults, respectively [16]. In Dar es Salaam, two previous reports showed that, *E. coli* accounted for 51.1% of UTI in a survey done in 2010 [17] whereas as high as 64.0% of hospital- and community-acquired UTI was accounted for by *E. coli* in 2004 [18].

Finally, we recently showed that in Moshi, north-eastern Tanzania, *E. coli* was one of the most common bacterial pathogen isolates from a range of clinical manifestations [1]. Advanced molecular diagnostics such as whole genome sequencing (WGS) have revealed the emergence of a fatal diarrhoea-causing *E. coli* strain that combines virulence factors (VFs) from two *E. coli* strains [4]. The VFs are important properties that do enable an infectious agent to effectively and efficiently establish itself on or within its host by enhancing its potential to cause harm or disease. Some of the factors include those for host cell adherence, immune evasion, toxins and protease production for disrupting host cell pathways and scavenging minerals like iron in order to increase their survival. Certain *E. coli* strains have VFs that been linked to a serious outbreaks in China in 1999, where it caused 177 deaths [19], and in Germany in 2011, where it claimed 54 lives [20].

Molecular typing studies to evaluate *E. coli* related infections in low and middle-income countries (LMICs) are rare and little information is available regarding the specific subtypes causing infections in Tanzania or whether the high prevalence is due to sporadic infectious or clonal transmission between patients. The purpose of this study was to determine the serotypes, antimicrobial resistance and virulence genes. Further, we carried out a phylogenetic tree reconstruction to determine relatedness of *E. coli* isolated from patients in a tertiary care hospital in Tanzania using the WGS-based diagnostic platform installed at KCMC hospital in Moshi, Tanzania.

Methods

Study design, participants and specimen collection

A hospital based prospective cross-sectional study was conducted at KCMC between August 2013 and August 2015. Part of the study's methods has been described in details by Kumburu et al. [1]. KCMC is located in Moshi municipality, north-eastern Tanzania and serves as a zonal referral hospital for a catchment area of around 15 million people. The hospital has a bed capacity of 650 with approximately 500 outpatients seeking medical services daily. This study was granted ethical approval by the KCMC Research Ethics Committee and the National Institute for Medical Research. A written informed consent was obtained from each participant or from parents or guardians of children before enrolment into the study. The study involved patients admitted in medical, surgical and paediatrics wards who were suspected to have bacterial infection. Specimens collected for bacterial culture were sputum, wound or pus swab, stool and blood. Bacteria culture, isolation, and identification were performed according to in-house standard operating procedures and the Clinical and Laboratory Standards Institute (CLSI) guidelines as described by Kumburu et al. [1]. Over a 2-year period, 590 samples were collected without apriori knowledge of the infecting agent. A total of 377 bacterial strains were isolated, and whole genome sequenced. A number of isolates from this collection were randomly selected for antimicrobial susceptibility testing. A total of 38 *E. coli* collected sequentially were included in this study; among which 24 *E. coli* isolates had antimicrobial susceptibility results.

Genomic DNA isolation, whole genome sequencing, and analysis

For all *E. coli* isolates genomic DNA (gDNA) was purified and its concentration determined using the Easy-DNA extraction kit (Invitrogen®) and the Qubit dsDNA Assay Kit (Invitrogen®), respectively. The gDNA library preparation was performed following Nextera® XT DNA Sample Preparation Guide [21]. In brief, each gDNA was tagmented (tagged and fragmented) by the Nextera® XT transposome.
The transposome simultaneously fragments the input DNA and adds adapter sequences to the fragment ends. Then followed a limited-cycle PCR amplification whereby indexes required for cluster formation were added to each DNA piece. Then each gDNA library was normalized to ensure equal representation during sequencing. Equal volumes of the normalized library were combined, diluted in hybridization buffer, and heat denatured prior to sequencing on the Illumina MiSeq platform (Illumina Inc.). The sequencing output was analysed using the standard WGS pipeline at KCRI, which is based on local implementations of the bioinformatics services available at https://cge.cbs.dtu.dk/services/. Quality control of the reads was performed using FastQC 0.11.4 [22].

De novo assembly was performed with SPAdes 3.11.1 [23], and quality assessed using QUAST 4.5 [24]. For this article’s purpose the analyses included: resistance genes identification using ResFinder 2.1 [25], Multi-Locus Sequence Typing (MLST) determination using MLST 1.8 [26], serotyping using SeroTypeFinder 1.1 [27], and virulence genes determination using VirulenceFinder 1.4 [28]. Phylogeny reconstruction was done using CSI Phylodynamics [29] (with reference strain EC958, NZ_HG941718.1). The 38 assembled E. coli genomes of the present study have been submitted to the European Nucleotide Archive with project accession number PRJEB23541. The phylogenetic analyses included 6 more E. coli genomes previously isolated from animals in Mwanza, north-western Tanzania, by Seni et al. [30]. The raw sequence data of the E. coli of animal origin were downloaded from the European Nucleotide Archive (ENA) under the project number PRJEB12335. Stata 13 (College Station, Texas 77845 USA) was used to determine Cohen’s kappa coefficient of agreement between the phenotypically tested and whole genome sequence predicted antimicrobial resistance.

**Results**

**Specimens and isolates**

A total of 38 non-duplicate E. coli were isolated, of which 18 (47.4%) were from wound (or pus) swabs, 13 (34.2%) from diarrhoeal stool, 4 (10.5%) from sputum and 3 (7.9%) from blood. Twenty-three (60.5%) were isolated from patients in medical wards, 9 (23.7%) from patients in surgical wards, and 6 (15.8%) from patients in intensive care unit (ICU). Out of these six E. coli isolates, four were from surgical ICU and 2 from medical ICU. Thirteen (34.2%) E. coli were isolated in 2013, 16 (42.1%) in 2014 and nine (23.7%) in 2015 (Table 1).

**Multi-locus sequence typing and serotyping**

Out of 38 E. coli isolates, 21 different STs were observed. Eight (21.1%) E. coli belonged to ST131; 3 isolated in 2013, 2 in 2014 and 3 in 2015 (Table 1). Two of the E. coli ST131 were isolated from patients in surgical wards, and six from patients in medical wards, including one from medical ICU. Of these eight, seven (87.5%) E. coli had serotype O25:H4 and one had serotype O15:H1. Out of 38 isolates, 10 (18.4%) belonged to ST10 clonal complex; of these, four (40.0%) were ST617 with serotype O89:H10, of which 3 were isolated in 2014 and 1 in 2015 and three (30.0%) were ST10, of which two had serotype O89:H10 and one had serotype O89:H9 (Table 2). These ST10 were all isolated in 2014 from medical wards. We noted that, out of the three ST410, one had serotype O8:H21, and the others had unknown O-group but belonged to H9.

**Virulence factors**

Table 2 as well show the distribution of virulence factors (VFs) among the sequenced E. coli isolates. A total of 22 (57.9%) E. coli isolates carried VFs for host cells adherence. These included 12 (31.6%) for iha, 7 (18.4%) for lpfA and 6 (15.8%) for each of aafC and nfaE. The prevalence of eilA was 3 (7.9%) and 2 (5.3%) for each of eae and bfpA. A total of 9 (23.7%) E. coli isolates had VFs that promote toxin production. The distribution of VFs that promote toxin was 7 (18.4%) for senB and 2 (5.3%) for each of astA and pic. The VFs that promote E. coli protease production were detected in 18 (47.4%) E. coli isolates. The prevalent protease VFs were sat and vat with 13 (34.2%) and 5 (13.2%), respectively. A total of 25 (65.7%) E. coli isolates had iss, a factor that promotes E. coli immune evasion by increasing serum survival.

**Acquired antimicrobial resistance**

A total of 28 (73.7%) E. coli carried genes encoding beta-lactam resistance enzymes (Table 3). The blaCTX-M-15 was harboured by 17 (44.7%) isolates, blaOXA-1 by 19 (50%) and blatem-1B by 12 (31.6%). The overall proportion of dfrA genes encoding trimethoprim resistance enzymes was 28 (73.7%). The dfrA17 was the most abundant gene with a proportion of 16 (42.1%) followed by 6 (15%) for dfrA14 genes, and 2 (5.3%) for each of dfrA5, dfrA7 and dfrA12.

A total of 19 (50%) E. coli carried mhp(A), a gene encoding macrolide resistance enzymes including towards erythromycin and azithromycin. The prevalence of genes encoding sulphonamides resistance enzyme for sul1 and sul2 were 19 (47.5%) and 24 (60.5%), respectively. The proportion of genes encoding tetracycline resistance enzymes for tet(A) and tet(B) were 16 (42.1%) and 9 (23.7%) respectively. The proportions of genes encoding chloramphenicol resistance enzymes were 19 (50%) for catB3 and 3 (7.9%) for catA1. A total of 22 (57.9%) isolates carried both strA and strB genes against aminoglycoside. The gene aadA5 was detected in 15 (39.5%), for each aadA2 and aadA1 in 2 (5.3%) and aadB in 1 (2.6%) E. coli isolates. Others included aac(3)-Iia in 9 (23.7%) and aac(3)-IId in 2 (5.3%) of the E. coli isolates.
Table 1 Origin and drug susceptibility results of 38 clinical *E. coli* isolates

| Isolate | Ward | Specimen | Year | AMC | AM | CZ | CAZ | CRO | C | CIP | GM | NA | SXT |
|---------|------|----------|------|-----|----|----|-----|-----|---|----|----|----|-----|
| 1        | MW   | Stool    | 2013 | S   | R  | S  | S   | S   | S | S  | S  | S  | R   |
| 5        | MW   | Stool    | 2013 | S   | S  | S  | S   | S   | S | S  | S  | S  | S   |
| 6        | MW   | Stool    | 2013 | S   | S  | S  | S   | S   | S | S  | S  | S  | S   |
| 10       | SW   | Swab     | 2013 | R   | R  | S  | S   | S   | R | S  | R  | S  | R   |
| 21       | MW   | Stool    | 2013 | R   | S  | S  | S   | S   | S | S  | S  | S  | S   |
| 22       | MW   | Stool    | 2013 | N   | N  | N  | N   | N   | N | N  | N  | N  | N   |
| 30       | SW   | Swab     | 2013 | S   | S  | S  | S   | S   | S | S  | S  | S  | S   |
| 44       | SICU | Swab     | 2013 | S   | R  | R  | R   | R   | S | R  | R  | R  | R   |
| 70       | MW   | Sputum   | 2013 | R   | R  | R  | R   | R   | S | R  | R  | R  | R   |
| 73       | MW   | Stool    | 2013 | R   | R  | R  | R   | R   | S | R  | S  | R  | R   |
| 82       | MW   | Stool    | 2013 | R   | R  | R  | S   | S   | R | R  | S  | R  | R   |
| 97       | MW   | Stool    | 2014 | N   | N  | N  | N   | N   | N | N  | N  | N  | N   |
| 115      | SW   | Swab     | 2014 | S   | R  | R  | R   | R   | S | S  | S  | S  | R   |
| 199      | MW   | Blood    | 2014 | R   | R  | R  | R   | R   | R | R  | R  | R  | R   |
| 203      | MW   | Stool    | 2014 | R   | R  | S  | S   | S   | R | S  | R  | S  | R   |
| 210      | MW   | Blood    | 2014 | R   | R  | R  | R   | R   | S | R  | R  | R  | R   |
| 244      | MW   | Sputum   | 2014 | R   | R  | R  | R   | R   | S | R  | R  | R  | R   |
| 245      | MW   | Swab     | 2014 | R   | R  | R  | R   | R   | S | R  | R  | R  | R   |
| 247      | MW   | Blood    | 2014 | R   | R  | R  | S   | S   | R | S  | R  | S  | R   |
| 298      | MW   | Stool    | 2014 | N   | N  | N  | N   | N   | N | N  | N  | N  | N   |
| 365      | SICU | Swab     | 2014 | R   | R  | R  | R   | R   | S | S  | R  | S  | R   |
| 393      | MICU | Stool    | 2015 | R   | R  | R  | R   | R   | S | R  | R  | R  | R   |
| 521      | MW   | Sputum   | 2014 | S   | S  | S  | S   | S   | S | S  | S  | S  | R   |
| 538      | SW   | Swab     | 2015 | N   | N  | N  | N   | N   | N | N  | N  | N  | N   |
| 554      | MW   | Stool    | 2015 | N   | N  | N  | N   | N   | N | N  | N  | N  | N   |
| 587      | SW   | Swab     | 2015 | N   | N  | N  | N   | N   | N | N  | N  | N  | N   |
| 603      | MICU | Swab     | 2015 | N   | N  | N  | N   | N   | N | N  | N  | N  | N   |
| 118A     | SW   | Swab     | 2014 | S   | R  | S  | S   | S   | S | S  | S  | S  | S   |
| 119EC    | MW   | Stool    | 2014 | S   | S  | R  | R   | R   | R | S  | S  | S  | S   |
| 163A     | SW   | Swab     | 2014 | S   | S  | S  | S   | S   | S | S  | S  | S  | S   |
| 237C     | SW   | Swab     | 2014 | N   | N  | N  | N   | N   | N | N  | N  | N  | N   |
| 340B     | MW   | Sputum   | 2014 | S   | S  | S  | S   | S   | S | S  | S  | S  | S   |
| 431D     | SICU | Swab     | 2015 | N   | N  | N  | N   | N   | N | N  | N  | N  | N   |
| 567B     | SW   | Swab     | 2015 | N   | N  | N  | N   | N   | N | N  | N  | N  | N   |
| 598A     | MW   | Swab     | 2015 | N   | N  | N  | N   | N   | N | N  | N  | N  | N   |
| 598B     | MW   | Swab     | 2015 | N   | N  | N  | N   | N   | N | N  | N  | N  | N   |
| 71E       | MW   | Swab     | 2013 | R   | R  | R  | R   | R   | S | R  | R  | R  | R   |
| 77E       | SICU | Swab     | 2013 | R   | R  | R  | R   | R   | S | R  | S  | R  | S   |

*a* Year of collection  
*b* MW Medical ward, SW Surgical ward, MICU Medical ICU, SICU Surgical ICU  
*c* Wound or pus swab, diarrhoea or stool  
*d* S Susceptible, R Resistant, N Not tested, AMC Amoxicillin-Clavulanate, AM Ampicillin, CZ Cefazolin, CAZ Ceftazidime, CRO Ceftriaxone, C Chloramphenicol, CIP Ciprofloxacin, GM Gentamycin, NA Nalidixic acid, SXT Trimethoprim Sulphamethoxazole
| Isolate | ST<sup>a</sup> | Serotype | Virulence Factors | Adherence | Toxin | Pro tease | Evasion<sup>b</sup> | Type III<sup>c</sup> |
|---------|---------------|----------|------------------|-----------|-------|----------|-----------------|-----------------|
| 70      | ST-131        | O25:H4   | iha,nfaE         | sat       |       |          | i ss            |                 |
| 73      | ST-131        | O25:H4   | iha,nfaE         | sat       |       |          | i ss            |                 |
| 199     | ST-131        | O25:H4   | iha              | astA,senB | sat   |          | i ss            | air             |
| 210     | ST-131        | O25:H4   | iha,nfaE         | senB      | sat   |          | i ss            |                 |
| 587     | ST-131        | O25:H4   | iha              |           |       |          | i ss            |                 |
| 603     | ST-131        | O15:H1   | iha,nfaE         | lpfA,eilA | senB |          | i ss            |                 |
| 567B    | ST-131        | O25:H4   | iha              |           |       |          | i ss            | capU            |
| 71E     | ST-131        | O25:H4   | iha,nfaE         | senB      | sat   |          | i ss            |                 |
| 244     | ST-617<sup>f</sup> | O89:H10  |                 |           |       |          | i ss            |                 |
| 245     | ST-617<sup>f</sup> | O89:H10  |                 |           |       |          | i ss            |                 |
| 538     | ST-617<sup>f</sup> | O89:H10  | astA,senB        |           |       |          | i ss            | capU            |
| 237C    | ST-617<sup>f</sup> | O89:H10  |                 |           |       |          | i ss            |                 |
| 393     | ST-405        | O102:H6  | eilA             |           |       |          | i ss            |                 |
| 10      | ST-410        | O7?:H9   | lpfA             |           |       |          | i ss            |                 |
| 22      | ST-410        | O7?:H9   | lpfA             |           |       |          | i ss            |                 |
| 598B    | ST-410        | O89:H21  | lpfA             |           |       |          | i ss            |                 |
| 97      | ST-10<sup>f</sup> | O89:H10  |                 |           |       |          | i ss            |                 |
| 203     | ST-10<sup>f</sup> | O89:H10  |                 |           |       |          | i ss            |                 |
| 340B    | ST-10<sup>f</sup> | O89:H9   |                 |           |       |          | i ss            |                 |
| 44      | ST-167<sup>f</sup> | O89:H21  | senB             |           |       |          | i ss            | capU            |
| 247     | ST-167<sup>f</sup> | O89:H9   | senB             |           |       |          | i ss            | capU            |
| 5<sup>a</sup> | ST-226        | O40H19   | bfpA,eae         | espA,espF,espJ | nleB,nleC |                 |                 |                 |
| 6<sup>a</sup> | ST-226        | O40H19   | bfpA,eae         | espA,espF,espJ | nleB,nleC |                 |                 |                 |
| 118A<sup>a</sup> | ST-73         | O6:H1    | iha              | pic       | sat,vat |          | i ss            | mchB,mchC,mchF,mcmA |
| 21      | ST-942        | O39H28   | lpfA             |           |       |          | i ss            |                 |
| 554<sup>d</sup> | ST-95         | O2(50):H4 |                 |           |       |          | i ss            | mchF            |
| 431D    | ST-44<sup>f</sup> | O89:H4   |                 |           |       |          | i ss            |                 |
| 119EC   | ST-4959       | O154:H4  |                 |           |       |          | i ss            |                 |
| 521<sup>b</sup> | ST-504       | O166:H7  | iha              | vat       | i ss    |          | mchB,mchC,mchF,mcmA, sigA, capU | |
| 4       | ST-2332       | O128:H45 | cfaC, lngA       | eaaA      |         |          |                 |                 |
| 30      | ST-355        | O150:H5  | pic               | vat       | i ss    |          |                 |                 |
| 115     | ST-361        | O9H30    |                 |           |       |          | i ss            |                 |
| 163A    | ST-372        | O83:H1   |                 | vat       | i ss    |          |                 |                 |
| 298     | ST-38         | O86:H18  | iha,nfaE         | eilA      | senB   | sat       | i ss            | air             |
| 365     | ST-224        | O8H23    | lpfA             |           |       |          | i ss            |                 |
| 82      | ST-156        | O61:H34  | lpfA             |           |       |          | i ss            |                 |
| 598A    | ST-1193       | O75:H5   | iha              | senB      | sat,vat |          | i ss            |                 |
| 77E     | ST-1284       | O89:H21  |                 |           |       |          | i ss            |                 |

<sup>a</sup>Has virulence factor tir
<sup>b</sup>Has virulence factor iroN
<sup>c</sup>Has virulence factors ireA,iroN
<sup>d</sup>Has virulence factor ireA
<sup>e</sup>Sequence Type (ST)
<sup>f</sup>ST-10 clonal complex
<sup>g</sup>Immune evasion
<sup>h</sup>Type III translocated protein
### Table 3: Acquired antimicrobial resistance genes of 38 clinical E. coli isolates

| Isolate | AMG | BL | FQA | Macrolide | Phenicol | Quinolone | Sulphonamide | Tetracycline | Trimethoprim |
|---------|-----|----|-----|-----------|----------|-----------|--------------|--------------|--------------|
| 4       | strA strB | blaTEM-1B | | | | | sul2 | tet(A) | dfra14 |
| 5       | | | | | | | | | |
| 6       | | | | | | | | | |
| 10      | strA strB | blaOXA-1 | oac(6')lb-cr | | catB3 | | sul2 | tet(A) | dfra14 |
| 21      | strA strB | oac(6')lb-cr | | catB3 | | sul2 | tet(A) | dfra14 |
| 22      | | | | | | | | | |
| 30      | | | | | | | | | |
| 44      | oac(3)-Iid aadA2 | blaCTX-M-15 | blaTEM-1B | mph(A) | catA1 | QnrS1 qepA QnrD | sul1 | tet(A) | dfra12 |
| 70      | aadA5 | blaCTX-M-15 | blaOXA-1 | oac(6')lb-cr | mph(A) | catB3 | sul1 | tet(A) | dfra17 |
| 73      | aadA5 | blaCTX-M-15 | blaOXA-1 | oac(6')lb-cr | mph(A) | catB3 | sul1 | tet(A) | dfra17 |
| 82      | strA strB aadA2 | blaTEM-1B | | mph(A) | | sul2 sul1 | tet(B) | tet(A) | dfra12 |
| 97      | | | | | | | | | |
| 115     | aadA1 aac(3)-Ila strB strA | blaCTX-M-15 | blaTEM-1B | blaOXA-1 | oac(6')lb-cr | | catB3 catA1 | QnrB1 | sul2 | dfra14 |
| 163A    | | | | | | | | | |
| 199     | aadA1 strB strA aadA5 aadB | blaCTX-M-15 | blaTEM-1B | blaOXA-1 | oac(6')lb-cr | | catB3 cmIA1 | sul1 sul2 | tet(A) | dfra17 |
| 203     | strB strA | blaTEM-1C | blaOXA-1 | oac(6')lb-cr | | | catB3 | sul2 | tet(A) | dfra14 |
| 210     | aadA5 oac(3)-Ila strB strA | blaCTX-M-15 | blaOXA-1 | oac(6')lb-cr | mph(A) | | catB3 | sul1 sul2 | tet(B) | dfra17 |
| 237C    | aac(3)-Ila aadA5 strA strB | blaCTX-M-15 | blaOXA-1 | oac(6')lb-cr | mph(A) | | catB3 | sul1 sul2 | tet(B) | dfra17 |
| 244     | aac(3)-Ila aadA5 strA strB | blaCTX-M-15 | blaOXA-1 | oac(6')lb-cr | mph(A) | | catB3 | sul1 sul2 | tet(B) | dfra17 |
| 245     | strA aac(3)-Ila strB aadA5 | blaCTX-M-15 | blaOXA-1 | oac(6')lb-cr | mph(A) | | catB3 | sul1 sul2 | tet(B) | dfra17 |
| 247     | strA strB aadA5 | blaOXA-1 | blaCTX-M-15 | oac(6')lb-cr | mph(A) | | catB3 | sul1 sul2 | tet(B) | dfra17 |
| 298     | strB strA | blaTEM-1B | | | | | catA1 | sul2 | tet(C) | dfra7 |
| 365     | aadA5 | blaCTX-M-15 | blaOXA-1 | oac(6')lb-cr | mph(A) | | catB3 | sul1 | tet(A) | dfra17 |
| 393     | oac(3)-Ila aadA5 | blaCTX-M-15 | blaOXA-1 | oac(6')lb-cr | mph(A) | | catB3 | sul1 | tet(B) | dfra17 |
| 431D    | erm(B) strA strB | | | | | | | | |
| 521     | | | | | | | | | |
| 538     | aac(3)-Ila aadA5 strA strB | blaCTX-M-15 | blaOXA-1 | oac(6')lb-cr | mph(A) | | catB3 | sul1 sul2 | tet(A) | dfra17 |
| 554     | strA strB | blaTEM-1B | | | | | | | | dfra5 |
| 567B    | aadA5 | | | | | | | | |
| 587     | aadA5 | | | | | | | | |
| 603     | strB aadA5 aac(3)-Ild strA | blaTEM-1B | | | | | | | | |
| 118A    | strB strA | | | | | | | | |

Senda et al. Antimicrobial Resistance and Infection Control (2018) 7:72
| Isolate | AMG<sup>a</sup> | BL<sup>b</sup> | FQA<sup>c</sup> | Macrolide | Phenicol | Quinolone | Sulphonamide | Tetracycline | Trimethoprim |
|---------|-----------------|---------------|----------------|------------|----------|-----------|--------------|--------------|--------------|
| 119EC   | strB strA        | bl<sub>aTEM-1B</sub> | mph(A)         |            |          |           | sul2         | tet(A)       | dfrA14       |
| 598A    | strB strA        | bl<sub>aTEM-1B</sub> bl<sub>CTX-M-15</sub> bl<sub>OXA-1</sub> | aac(6')ib-cr | mph(A) |          | catB3     | sul2         | tet(B)       | dfrA17       |
| 598B    | aac(3)-Ila strB strA | bl<sub>aTEM-1B</sub> |            |            |          |           | sul2         | tet(A)       | dfrA14       |
| 71E     | aadA5 aac(3)-Ila strB strA | bl<sub>CTX-M-15</sub> bl<sub>OXA-1</sub> | aac(6')ib-cr | mph(A) |          | catB3     | sul1 sul2    | tet(A)       | dfrA17       |
| 77E     | strA strB        | bl<sub>CTX-M-15</sub> bl<sub>OXA-1</sub> | aac(6')ib-cr | mph(A) |          | catB3     | sul2         | tet(B)       |              |

<sup>a</sup> Aminoglycoside  
<sup>b</sup> Beta-Lactam  
<sup>c</sup> Fluoroquinolones and aminoglycoside
Phenotype and sequence-based antimicrobial resistance comparison

Agreement between phenotype and whole genome sequence-based antimicrobial resistance was done for 24 out of 38 E. coli isolates (Table 4). On average, agreement across all drugs tested was 83.9%. Overall, the phenotypically determined resistance was higher than sequence-based resistance. However, all but trimethoprim sulpha or co-trimoxazole showed strong agreement (81–100%) between phenotype and sequence-based resistance results. Trimethoprim/sulphamethoxazole (co-trimoxazole) showed moderate agreement: 45.8%, kappa = 15% and \( p = 0.08 \). Sequence-based analysis predicted resistance in 4 (16.7%) isolates, whereas phenotypic testing revealed 17 (70.8%) isolates to be resistant. Amoxicillin-clavulanate showed strongest agreement: 87.5%, kappa = 74% and \( p = 0.0001 \). Sequence-based analysis predicted resistance to amoxicillin-clavulanate in 14 (58.3%) isolates, whereas 15 (62.5%) isolates were found to be resistant phenotypically.

Phylogeny and genome comparison

The observed minimum and maximum SNPs difference between one isolate and another in a pairwise genome comparison for ST131 were 234 and 10,425 respectively (Additional file 1: Table S1). The tree topology for ST131 showed one isolate segregating very distinctly from the rest (Fig. 1). For ST10 clonal complex the observed minimum and maximum SNPs difference between one isolate and another were 157 and 35,103 respectively (Additional file 2: Table S2). Looking at the tree for ST10 clonal complex (Fig. 2), limited pattern of the E. coli isolates was observed. In this tree, with the exception of ST617 in the middle clade of the ST10 clonal complex tree, all other STs making this complex occurred across the trees with little apparent segregation. Furthermore, the resistance and virulence genes were spread almost universally across both phylogenetic trees.

Discussion

The present study revealed a high diversity of E. coli strains circulating in KCMC hospital settings as measured by Multi-Locus Sequence Typing and Serotyping. However, strains belonging to O25:H4-ST131 and O89:H10-ST617 (ST10 clonal complex) were found to predominate. These findings are similar to the findings from other continents describing the spread and predominance of these endemic clones in health facilities [31–34]. The observed clonal diversity in this hospital may suggest sporadic introductions of diverse strains into the hospital from the community. To explore whether or not similar STs were clonally related, SNP difference between isolates and phylogenetic analysis suggested the existence of multiple clones of E. coli in these settings. A similar clonal diversity was observed when STs from the present study were compared to similar STs of E. coli isolates from companion and other domesticated animals in Mwanza, north western Tanzania.

Overall, levels of antimicrobial resistance in E. coli isolates were observed to be high. Trimethoprim/sulphamethoxazole (co-trimoxazole) resistance was observed to be lower (70%) than the one (93%) found by Seni et al. [30], but in the present study on average 54–60% resistance to amoxicillin-clavulanate, ampicillin, ceftazidime, ceftriaxone, and ciprofloxacin was higher than that reported by Seni et al. [30]. Strong agreement (81–100%) between sequence and phenotype-based resistance to all drugs tested was observed. However, trimethoprim/sulphamethoxazole revealed moderate agreement (45.8%) between the two methods. The phenotypically determined resistance to trimethoprim/sulphamethoxazole was higher than sequence-based resistance. The plausible explanations for the observed difference could be that our analysis used only known resistance genes and did not include point mutations. Also, resistance in gram negative bacteria including E. coli is multifactorial and not all genes involved in resistance mechanisms have been uncovered.

### Table 4 Agreement between phenotypically tested and whole genome sequence predicted antimicrobial resistance

| Antibiotic name          | DST a | WGS b | Agreement | Kappa | \( p \) value |
|--------------------------|-------|-------|-----------|-------|--------------|
| Amoxicillin-Clavulanate  | 15 (62.5%) | 14 (58.3%) | 0.875 | 0.74 | 0.0001 |
| Ampicillin               | 16 (66.7%) | 14 (58.3%) | 0.9167 | 0.82 | 0.00 |
| Ceftazidime              | 13 (54.2%) | 13 (54.2%) | 0.9167 | 0.83 | 0.00 |
| Ceftriaxone              | 15 (62.5%) | 13 (54.2%) | 0.9167 | 0.84 | 0.00 |
| Chloramphenicol          | 2 (8.3%) | 2 (8.3%) | 0.8333 | –0.09 | 0.672 |
| Ciprofloxacin            | 15 (62.5%) | 14 (58.3%) | 0.875 | 0.74 | 0.0001 |
| Gentamycin               | 7 (29.2%) | 7 (29.2%) | 0.9167 | 0.79 | 0.00 |
| Trimethoprim Sulpha      | 17 (70.8%) | 4 (16.7%) | 0.4583 | 0.15 | 0.0799 |

a phenotype-based resistance
b whole genome sequence-based resistance

Sonda et al. Antimicrobial Resistance and Infection Control (2018) 7:72 Page 8 of 12
Following analysis of resistance genes, over 70% of the E. coli strains carried genes encoding beta-lactamases, with bla_{OXA-1} being predominant (50%) followed by bla_{CTX-M-15} (44.7%) and bla_{TEM-1B} (31.6%). The prevalence of bla_{CTX-M-15} found in the present study contrasts with Manyahi et al. [18] who found bla_{CTX-M-15} as the most prevalent gene (90.6%) in a tertiary hospital in Dar es Salaam. Nonetheless, similar to the present study, O25:H4-ST131 was reported to be a major cause of MDR E. coli infections [35]. The proportion (70%) of E. coli strains carrying genes encoding beta-lactamases, correlates in this study with a high (50%) presence of aac(6\')ib-cr encoding ciprofloxacin resistance enzymes. Although aac(6\')ib-cr encodes low level ciprofloxacin resistance by itself (as well as aminoglycoside resistance), and usually requires additional mutations (e.g. in chromosomal gyrA or parC) to confer high level resistance, it is a threat to ciprofloxacin which in our settings is one of the most prescribed drugs. Other studies in the US [36, 37], Brazil [38] and Korea [34] have documented similar findings to the present study of the co-carriage of aac(6\')ib-cr, ESBL genes and the existence of ciprofloxacin resistance in E. coli ST131. Further, an agreement between the presence or absence of aac(6\')ib-cr and phenotypic resistance results in particular ciprofloxacin was noted by Madoshi et al. [39] who characterised E. coli isolates from healthy cattle and cattle attendants in Morogoro, Tanzania [39]. Among other fluoroquinolones, ciprofloxacin is one of the most prescribed antibiotics in Tanzania. Plausibly this explains the observed linkage between ciprofloxacin resistance and carriage of aac(6\')ib-cr and CTX-M ESBL [31]. The linkage between quinolone resistance and sub-lineages of ST131 has been recently explored by Zakour et al. [40] whereby this work suggested that quinolone use is associated with the acquisition of virulence and fluoroquinolone resistance determinants and expansion of clade C2/H30-Rx of quinolone-resistant ST131. Also a gene (mph(A)) conferring resistance to macrolides including erythromycin and azithromycin was detected in a high proportion. The presence of the mph(A) in E. coli isolates of the current study was 50%, and thus, higher than the 13% that was found in E. coli from 5 countries from 4 continents by Nguyen et al. [41].
High exposures to erythromycin and azithromycin could be one possible reason leading to emergence of resistance to macrolides [42]. Furthermore, the present study noted a high proportion (73.7%) of dfrA genes encoding trimethoprim resistance enzymes similar to findings by Madoshi et al. [39] in Morogoro, Tanzania. Nonetheless, the present study proportion of dfrA genes was relatively higher than that (43%) found in E. coli isolates from healthy college students in Ghana and Nigeria in 2005 and 2009 [43]. The observed proportion difference could be explained by the fact that resistance in hospital-based studies is relatively higher than in community-based ones. Another reason could be that the present study was hospital-based whereas the West African study population were healthy individuals.

The present study also characterised VFs in all E. coli isolates. The VFs that facilitate E. coli and host cells or E. coli and E. coli adherence including long polar fimbriae (lpfA), adhesin (iha) and intimin (eae) were predominant among the circulating E. coli O25:H4-ST131 strains. These VFs have been identified as suggestive of virulent serotypes and may be used as reliable markers for the identification of pathogenic E. coli [44]. Another prevalent group of VFs this study noted were those responsible for E. coli immune evasion by increasing serum survival (iss). Although the predominance of iss in the present study was from E. coli isolated from hospitalised patients, in Mwanza north-western Tanzania, Msahana et al. [45] noted a similar pattern of iss in E. coli that were community-acquired. In this study we observed no clear correlation or pattern of VFs with the clinical findings or patient's outcomes. However, colonisation events rather than infections were common as wound or pus swab and stool specimens constituted the majority of strains studied. The observed existence of multiple VFs further underlines the exceptional ability of E. coli in colonising or causing infections to a wide range of hosts and niches.

Potential clinical implications for the obtained results are that caution should be taken when interpreting and utilising microbiology results especially when E. coli is isolated in LMICs as most often E. coli is regarded of low threat. Contrary to that notion, these findings highlight that E. coli should not be regarded non-pathogenic until pathogenic and antimicrobial resistance determinants have been truly confirmed absent. Additionally, the WGS-based findings hint on the existence of nosocomial transmissions in the hospital thus prompting the formulation of pragmatic antimicrobial stewardships and infection prevention and control initiatives.

We understand and acknowledge the limitations of the present study. First, the analysis was performed on a
small number of E. coli isolates, which may limit generalisation of the findings. With such a small number of the isolates analysed, it is important to point out that another limitation that this work was likely to suffer from was the lack of deeper statistical analysis to correlate the isolates resistance and virulence findings with patients characteristics (age, gender, ward, room, specimen). Third, the existence of genes encoding different resistance and virulence factors is only indicative of the genes present in the isolates. On the other hand, RNA sequencing which is an avenue for future studies could have given a strong evidence of expression levels of genes encoding different resistance and virulence factors.

Conclusion
The observed high levels of E. coli diversity in terms of their antimicrobial resistance genes, serotypes and virulence genes underlines the necessity for concerted efforts to routinely screen all bacterial isolates of clinical importance especially in tertiary health care facilities. WGS use for laboratory-based surveillance can be an effective early warning system for emerging pathogens and resistance mechanisms in LMICs. The information generated in this study will not only provide updates on levels of virulence and antimicrobial resistance at hospital level but will also be used as a basis in formulating pragmatic antimicrobial stewardships and infection prevention and control initiatives.

Additional files

Additional file 1: Table S1. A pairwise genome comparison matrix of SNP differences for E. coli ST131. (XLS 34 kb)

Additional file 2: Table S2. A pairwise genome comparison matrix of SNP differences for E. coli ST10 clonal complex. (XLS 29 kb)

Acknowledgements
We thank the management of Kilimanjaro Christian Medical Centre and all patients who consented to participate in this study.

Funding
This study was supported by DANIDA through Danida Fellowship Centre award number DFC No. 12-007/DTU.

Availability of data and materials
Data are available on request to the authors.

Authors’ contributions
TS conceived the initial idea; FA, OL, MA, BM and GK refined the idea. TS and MA performed laboratory analyses. TS and MZ analysed data and prepared manuscript draft. All authors read, revised and approved the final manuscript.

Ethics approval and consent to participate
This study was granted ethical approval by the KCRC Research Ethics Committee and the National Institute for Medical Research with approval numbers 893 and NMIR/HQ/RBa/Vol.IX/2080 respectively. A written informed consent was obtained from each participant or from parents or guardians of children before enrolment into the study.

Competing interests
The authors declare that they have no competing interests.

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details
1Kilimanjaro Clinical Research Institute, Kilimanjaro Christian Medical Centre, Moshi, Tanzania. 2Kilimanjaro Christian Medical University College, Moshi, Tanzania. 3Centre for Medical Parasitology, Department of Immunology and Microbiology, University of Copenhagen and Department of Infectious Diseases, Copenhagen University Hospital, Copenhagen, Denmark. 4DTU-Food, Technical University of Denmark, Copenhagen, Denmark. 5Centre for Biological Sequence Analysis, Technical University of Denmark, Copenhagen, Denmark. 6East African Health Research Commission, Bujumbura, Burundi.

Received: 12 December 2017 Accepted: 22 May 2018
Published online: 08 June 2018

References
1. Kumburu HH, Sonda T, Mmbaga BT, Alifrangis M, Lund O, Kibiki G, Aarestrup FM. Patterns of infections, aetiological agents and antimicrobial resistance at a tertiary care hospital in northern Tanzania. Trop Med Int Health. 2017;22:454–64.
2. Moisssenet D, Salaue B, Clermont O, Birgen E, Aletf G, Denamur E, Mérenes A, Mitanchez D, Vu-Thien H. Meningitis caused by Escherichia coli producing TEM-52 extended-spectrum β-lactamase within an extensive outbreak in a neonatal ward: epidemiological investigation and characterization of the strain. J Clin Microbiol. 2010;48:2459–63.
3. Aftet JE, Bevanger L, Romundstad P, Bergh K. Association of atypical enteropathogenic Escherichia coli (EPEC) with prolonged diarrhoea. J Med Microbiol. 2004;53:1137–42.
4. Olesen B, Scheutz F, Andersen RL, Menard M, Boisen N, Johnston B, Hansen DS, Krogfelt KA, Nataro JP, Johnson JR. Enterogaugregative Escherichia coli O178:H10, the cause of an outbreak of urinary tract infection. J Clin Microbiol. 2012;50:3703–11.
5. Mahende C, Njasala B, Lusingu J, Butichi A, Lushino P, Lemnge M, Premji Z. Prevalence of Escherichia coli virulence genes in patients with diarrhea and a subpopulation of healthy volunteers in Madrid, Spain. Front Microbiol. 2016;7:1–6.
6. Shahriani M, Dehkhordi FS, Montaz H. Characterization of Escherichia coli virulence genes, pathotypes and antibiotic resistance properties in diarrheic calves in Iran. Bol Res. 2014;47:28.
7. Imuta N, Ooka T, Seto K, Kawahara R, Koriyama T, Kojyo T, Iguchi A, Tokuda K, Ogasawara Y, Takahashi Y, Ishige T, Ohashi K, Sato Y, Yano K, Torii T, Ando K, Kubota T. Emergence of ctx-m-14-producing eaec o25:h4 clones related to sequence type 131. J Clin Microbiol. 2016;54:2108–20.
8. Brittain L, Hammeri JA, Strauch E, Neef J, Dieckmann R, Kelner-Bungos Y, Martin A, Miko A, Strockbine NA, Lindstedt BA, Horn D, Monroe H, Huettel B, Muller L, Stuber K, Reinhardt R. Spread of a distinct Ssa2-encoding phage prototype among Escherichia coli O104:H4 strains from outbreaks in Germany, Norway, and Georgia. J Virol. 2012;86:3414–55.
9. Berndt H, Barberino MG, Moreira ED, Riley L, Reis JN. Distribution of strain type and antimicrobial susceptibility of Escherichia coli isolates causing meningitis in a large urban setting in Brazil. J Clin Microbiol. 2014;52:1418–22.
10. Cabal A, Garcia-Castillo M, Canton R, Gortzár C, Domínguez L, Álvarez J. Prevalence of virulence genes underlines the necessity for concerted efforts in tertiary health care facilities. WGS control initiatives.

Additional files

Additional file 1: Table S1. A pairwise genome comparison matrix of SNP differences for E. coli ST131. (XLS 34 kb)

Additional file 2: Table S2. A pairwise genome comparison matrix of SNP differences for E. coli ST10 clonal complex. (XLS 29 kb)
13. Masika WG, O’Meara WP, Holland TL, Armstrong J. Contribution of urinary tract infection to the burden of febrile illnesses in young children in rural Kenya. PLoS One. 2017;12:e0174199.

14. Msiak BF, Mishana SE, Hokororo A, Mazigo HD, Morona D. Prevalence and predictors of urinary tract infection and severe malaria among febrile children attending Makongoro health Centre in Mwanza city, north-western Tanzania. Archives of Public Health. 2012;70.

15. Ahmed M, Moreni N, Miramboro MM, Hokororo A, Mushy MF, Seni J, Kamugisha E, Mishana SE. Multi-resistant gram negative enteric bacteria causing urinary tract infection among malarious underfives admitted at a tertiary hospital, northwestern, Tanzania. Ital J Pediatr. 2015;41:1–5.

16. Hildeman W, Amos B, Mtove G, Mwana A, Cederlund K, Reyburn H. Causes of non-malarial febrile illness in outpatients in Tanzania. Trop Med Int Health. 2016;21:149–56.

17. Moyo SJ, Aboud S, Kasubi M, Lyaemuya EF, Maselle SY. Antimicrobial resistance among producers and non-producers of extended spectrum beta-lactamases in urinary isolates at a tertiary Hospital in Tanzania. BMC research notes. 2010;3:948.

18. Manyai J, Merejsa J, Tewolde MK, Nigadugulle F, Ularasa W, Blomberg B, Langeland N. Detection of CTX-M-15 beta-lactamases in Enterobacteriaceae causing hospital- and community-acquired urinary tract infections as early as 2004, in Dar es salaam, Tanzania. BMC Infect. 2017;17:382.

19. Xiong Y, Wang P, Lan R, Ye C, Wang H, Ren J, Jing H, Wang Y, Zhou Z, Bai X, Cui Z, Luo X, Zhao A, Wang Y, Zhang S, Sun H, Wang L, Xu J. A novel escherichia coli O157:H7 clone causing a major hemolytic uraemic syndrome outbreak in China. PLoS One. 2012;7:1–10.

20. Frank C, Werber D, Cramer JP, Askar M, Faber M, an der Heiden M, Bernard H, Fruth A, Parger R, Spode A, Wadl M, Zoufaly A, Jordan S, Kemper MJ, Pollin P, Müller K, Rosner B, Buchholz U, Stark K, Krause G. Epidemic profile of Shiga-toxin-producing Escherichia coli O104:H4 outbreak in Germany. N Engl J Med. 2011;365:1771–80.

21. Nextera® XT Library Prep Reference Guide. http://support.illumina.com/downloads/nextera_xt_sample_preparation_guide_15031942.html.

22. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyzhev AV, Stoytchov AV, Yaraby NH, Tesler G, Alekseyev MA, Pevzner PA. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 2012;19:455–77.

23. Gurevich A, Saveliev V, Yaraby NH, Tesler G. QUAST: quality assessment tool for genome assemblies. Bioinformatics. 2013;29:1072–5.

24. Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O, Aarestrup FM. Identification of acquired antimicrobial resistance genes. J Antimicrob Chemother. 2012;67:2640–4.

25. Larsen MV, Cosentino S, Rasmussen S, Mikkola S, Hasman H, Mikkola S, Rasmussen S, Mikkola S, Hasman H, Manling R, Jelsbak L, Schwiertz-Pontén T, Lysøe DW, Aarestrup FM, Lund O. Multilocus sequence typing of total-genome-sequenced bacteria. J Clin Microbiol. 2012;50:1355–61.

26. Joensen KG, Scheutz F, Lund O, Hasman H. Kaas RS, Nielsen EM, Aarestrup FM. Real-time whole-genome sequencing for routine typing, surveillance, and outbreak detection of verotoxigenic Escherichia coli. J Clin Microbiol. 2014;52:1501–10.

27. Kaas RS, Leekitcharoenphon P, Aarestrup FM, Lund O. Solving the problem of comparing whole bacterial genomes across different sequencing platforms. PLoS One. 2014;9:e104984.

28. Seni J, Falgenhauer L, Siemo N, Miramboro MM, Imizalilougi C, Matee M, Rweyemamu M, Chakabority T, Mishana SE. Multiple ESBL-producing Escherichia coli sequence types carrying quinolone and aminoglycoside resistance genes circulating in companion and domestic farm animals in Mwanza, Tanzania, harbor commonly occurring plasmids. Front Microbiol. 2016;7:1–8.

29. Valenza G, Nickel S, Pfeffer Y, Pietsch M, Voigtländer E, Lehner-Indrevi L, Hölter C. Prevalence and genetic diversity of extended-spectrum beta-lactamase (ESBL)-producing Escherichia coli in nursing homes in Bavaria, Germany. Vet Microbiol. 2017;200:138–41.

30. Alibru, I, Odugbemi T, Koenig W, Ghebremedhin B. Sequence Type ST131 and ST10 Complex (ST617) predominant among CTX-M-15-producing Escherichia coli isolates from Nigeria. Clin Microbiol Infect. 2012;18:49–51.

31. Oetoo J, Diestra K, Juan C, Bautista V, Novais A, Pérez-Vázquez M, Moya B, Miro-E, Coque TM, Oliver A, Canton R, Navarro F, Campos J. Extended-spectrum β-lactamase-producing Escherichia coli in Spain belong to a large variety of multilocus sequence typing types, including ST10 complex/a, ST23 complex/a and ST131/b2. Int J Antimicrob Agents. 2006;31:173–6.

32. Cho MK, Kang CJ, Kim SH, Cho SY, Ha YE, Wi YM, Chung DR, Peck KR, Song JH, Shin JS, Moon SY, Kim CK, Lee SS, Lee JA, Kim YS, Sohn KM, Rhey JY, Jung SI, Park KH, Kang SJ, Kim SW, Chang HH, Ryu SY, Kim HA, Kwon KT, Lim MH. Comparison of the microbial characteristics and virulence factors of ST31 and non-ST31 clones among extended-spectrum β-lactamase-producing Escherichia coli causing bacteremia. Diagn Microbiol Infect Dis. 2016;84:102–4.

33. Johnson JR, Johnston B, Clabots C, Kuskowski MA, Castanheira M. Escherichia coli sequence type ST131 as the major cause of serious multidrug-Resistant E. coli infections in the United States. Clin Infect Dis. 2010;51:86–94.

34. Periago G, Costello M, Pitout JDD. Molecular characteristics of extended-spectrum β-lactamase-producing Escherichia coli from the Chicago area: high prevalence of ST131 producing CTX-M-15 in community hospitals. Int J Antimicrob Agents. 2010;36:19–23.

35. Park CH, Robicsek A, Jacoby GA, Sahn D, Hooper DC. Prevalence in the United States of acrA67-Ib-cr encoding a ciprofloxacin-modifying enzyme. Antimicrob Agents Chemother. 2006;50:3953–5.

36. Periago G, Asensi MD, Pitono-Silva A, Pitout JDD. Molecular characteristics of extended-spectrum β-lactamase-producing Escherichia coli from Rio de Janeiro, Brazil. Clin Microbiol Infect. 2011;17:1039–43.

37. Madoshi BP, Kudirkene E, Matombo MMA, Muhairwa AP, Lupindo AM, Olsen JE. Characterisation of commensal Escherichia coli isolated from apparently healthy cattle and their attendants in Tanzania. PLoS One. 2016;11:1–e0168160.

38. Ben Zakour NL, Alsheikh-Husssain AS, Ashcroft MM, Khahn Nhu NT, Roberts LW, Stanton-Cook M, Schembri MA, Beaton SA. Sequential Acquisition of Virulence and Fluoroquinolone Resistance Has Shaped the evolution of Escherichia coli ST31. mBio. 2016;7:e00347-16.

39. Phuc Nguyen MC, Woerther PL, Bouvet M, Andremont A, Leclercq R, Canu A. Escherichia coli as reservoir for macrolide resistance genes. Emerg Infect Dis. 2009;15:1648-50.

40. Noguchi N, Takada K, Katayama J, Ebara A, Satatsu M. Regulation of transcription of the mphA gene for macrolide 2′-phosphotransferase I in Escherichia coli: characterisation of the regulatory gene mphR(a). J Bacteriol. 2000;182:5052–8.

41. Labar AS, Millman JS, Ruebush E, Opintan JA, Bishar RA, Aboderin AO, Nguyen WC, Larimana A, Oleke IN. Regional dissemination of a trimethoprim-resistance gene cassette via a successful transposable element. PLoS One. 2012;7:e31842.

42. Torres AG, Blanco M, Valenzuela P, Slater TM, Patel SD, Dabhi G, Lopez C, Barriga XF, Blanco JE, Gomes TAT, Vidal R, Blanco J. Genes related to long polar fimbriae of pathogenic Escherichia coli strains as reliable markers to identify virulent isolates. J Clin Microbiol. 2009;47:2442–51.

43. Mshana SE, Falgenhauer L, Miramboro MM, Mushi MF, Moreni N, Julius R, Seni J, Imizalilougi C, Matee M, Chakabority T. Predictors of blaCTX-M-15 in varieties of Escherichia coli genotypes from humans in community settings in Mwanza, Tanzania. BMC Infectious Diseases. 2016;16:1–9.

44. Sonda et al. Antimicrobial Resistance and Infection Control (2018) 7:72