Molecular epidemiology of extended-spectrum beta-lactamase-producing extra-intestinal pathogenic Escherichia coli strains over a 2-year period (2017-2019) from Zimbabwe

This study was designed to characterize extended-spectrum beta-lactamase (ESBL)–producing extra-intestinal pathogenic *Escherichia coli* (*E.coli*) (ExPEC) associated with urinary tract infections in nine different geographic regions of Zimbabwe over a 2-year period (2017–2019). A total of 48 ESBL-positive isolates from urine specimen were selected for whole-genome sequencing from 1246 *Escherichia coli* isolates biobanked at the National Microbiology Reference laboratory using phenotypic susceptibility testing results from the National *Escherichia coli* Surveillance Programme to provide representation of different geographical regions and year of isolation. The majority of ESBL *E. coli* isolates produced cefotaximase-Munich (CTX-M)-15, CTX-M-27, and CTX-M-14. In this study, sequence types (ST) 131 and ST410 were the most predominant antimicrobial-resistant clones and responsible for the increase in ESBL–producing *E. coli* strains since 2017. Novel ST131 complex strains were recorded during the period 2017 to 2018, thus showing the establishment and evolution of this antimicrobial-resistant ESBL clone in Zimbabwe posing an important public health threat. Incompatibility group F plasmids were predominant among ST131 and ST410 isolates with the following replicons recorded most frequently: F1:A2:B20 (9/19, 47%), F2:A1: B (5/19, 26%), and F1:A1:B49 (8/13, 62%). The results indicate the need for continuous tracking of different ESBL ExPEC clones on a global scale, while targeting specific STs (e.g. ST131 and ST410) through control programs will substantially decrease the spread of ESBLs among ExPEC.

Introduction

The production of extended-spectrum beta-lactamases (ESBLs) by clinical isolates of extra-intestinal pathogenic *Escherichia coli* (ExPEC) is a serious global therapeutic threat [1]. These ESBLs can reduce the efficacy of the extended-spectrum cephalosporins, except for cephamycin and carbapenems [2]. ESBL production is associated with the presence of bla_{TEM}, bla_{SHV}, and $bla_{\text{CTX-M}}$ [3]. CTX-M producers are increasingly detected and replacing TEM and SHV producers in European and African countries [4]. A number of important *E. coli* clones have been detected among the ESBL producers, these include sequence type (ST) 131, ST405, ST38, ST648, ST410, and ST1193 [1].

Among these, the most studied phylogenetic lineage of *E. coli* in terms of antibiotic resistance is *E. coli* ST131, phylogenetic group B2 serotype O25:H4. This lineage harbours a wide range of core sets of virulence genes and various plasmid-mediated resistance genes. ST131 is involved in the global spread of the ESBL phenotype linked to the production of CTX-M-15 and CTX-M-27, which is the main mechanism of resistance to beta-lactams [1]. ST131's virulence combined with its carriage of transferable elements encoding multidrug resistance is likely responsible for the pandemic success of ST131 strains [5, 6].

The ST131 population structure is divided into three clades, namely, A, B, and C. The C clade comprises two subclades, C1 and C2, which are defined by the presence of a specific fimbrial adhesin allele (*fimH30*) corresponding to the *H30R* and *H30Rx* clades [7]. The C1 subclade contains mutations in the chromosomal genes *gyrA* and *parC* that confer fluoroquinolone resistance, while the C2 subclade contains the same *gyrA* and *parC* mutations but is also associated with $bla_{\text{CTX-M-15}}$. Reports of the occurrence of ST131 subgroup C1-M-27 associated with $bla_{\text{CTX-M-27}}$ in some parts of the world especially in Europe [8], Asia, and parts of Africa [9] emerged.

ESBL–producing *E. coli* associated with nosocomial- and community-acquired infections have been reported in most regions around the world [10]. These bacteria are resistant to most of the antimicrobials used for the treatment of urinary tract infections (UTIs), such as ciprofloxacin, trimethoprim-sulfamethoxazole, and most of the cephalosporins [11]. A few studies have been carried out in sub-Saharan Africa (South Africa, Nigeria, Tanzania, and Democratic Republic of Congo) to investigate the molecular epidemiology and characteristics of the ExPEC strains [12,13,14,15], while in Zimbabwe, this remained unexplored. Thus, the main aim of this study was to characterize the ESBL ExPEC associated with UTIs in different geographic regions of Zimbabwe using whole-genome sequencing (WGS).

Materials and methods

Selection of bacterial isolates for genomic evaluation

A total of 48 ESBL-positive isolates from urine specimens were selected for WGS from 1246 *E. coli* isolates biobanked at the National Microbiology Reference laboratory using phenotypic susceptibility testing results from the National *Escherichia coli* Surveillance Programme to provide representation of different geographical regions and years of isolation (2017 (14); 2018 (22), and 2019 (12)) (Supplementary File 1*Epidemiological features of the 48 ESBL isolates*). Demographic data associated with 48 ESBL-positive isolates were analyzed. The geographical regions included Bulawayo, Chitungwiza, Marondera, and Harare. The selection was also based on community-acquired UTIs, which was defined as an infection of the urinary tract that occurs in the community or within 48 h of hospital admission and was incubating during time of hospital admission [16]. No information of previous use of antibiotics was collected.

Study population phenotypic antibiotic resistance profiling

All the selected isolates were sub-cultured on MacConkey or eosin methylene blue (EMB) (Mast Group, Merseyside, UK), incubated in Memmert ICH110 (Germany) at 37 °C for 18 to 24 h and then stored in 20% glycerol broth at 80 °C. Antimicrobial susceptibility testing was determined by the Kirby Bauer disc diffusion method using the Clinical Laboratory Standards Institute (CLSI) guidelines [17]. The antimicrobial drugs tested included ampicillin, trimethoprim-sulfamethoxazole, ciprofloxacin, ceftriaxone, tetracycline, ceftazidime, nalidixic acid, cefepime, and ertapenem; results were interpreted as described by the CLSI [17]. The presence of ESBLs was confirmed according to the CLSI criteria for ESBL screening and confirmation [17]. *E. coli* ATCC 25922 was used as quality control strain. Additional data included on collection of each isolate were the year, location of isolation, travel associated, cities, age, and gender (Supplementary File 1.

Genomic DNA isolation and whole-genome sequencing

Genomic DNA (gDNA) of *E. coli* was purified using the Wizard® Genomic DNA Purification Kit (Promega, Madison, WI, USA) according to the manufacturer's instructions [18] and stored at – 20 °C. Library preparation was performed using the Nextera XT DNA Library Preparation Kit (Illumina, San Diego, CA, USA) and sequenced on a MiSeq benchtop sequencer (Illumina, San Diego, CA, USA) at Quadram Institute Biosciences in the UK. The data was uploaded to BaseSpace (http://www.basespace.illumina.com) and then converted to FASTQ files.

Genomic sequence analysis

The sequences were analyzed on the Cloud Infrastructure for Microbial Bioinformatics [19]. Paired-end short-read sequences were concatenated, then quality-checked using FastQC v0.11.7. De novo assembly was performed with SPAdes 3.11 [20], and quality was assessed using QUAST 4.5 [21]. Snippy v4.3.2 (https;//github.com/tseemann/snippy) was used to generate a core SNP alignment using default parameters. The complete genome sequence of *E. coli* strain K12 sub-strain MG1655 was used as reference genome (NCBI accession: NC_000913.3). After the core-genome alignment, a reconstruction of maximum likelihood phylogeny with 1000 bootstrap replicates with RAxML v8.2.4 based on a general time-reversible nucleotide substitution model was used [22, 23]. The phylogenetic tree was rooted using *Escherichia fergusonii (E. fergusonii)* as an outgroup (NCBI accession: GCA_000026225.1). The phylogenetic tree was visualized in Figtree v1.4.4 (https://github.com/rambaut/figtree/) [24] and annotated in RStudio v3.5.2 and Adobe illustrator v 23.0.3 (Adobe Inc. San Jose, CA). Recombination was detected and masked using Gubbins [25] before the phylogenetic reconstruction. The pairwise SNP distances were computed between genomes from the core-genome alignment using snp-dists v0.6 (https://github.com/tseemann/snp-dists).

Comparative genomics analysis

The assembled draft genomes were used to define the presence of genes and their alleles. The following databases or typing schemes were used to determine (i) STs (multi-locus sequence typing (MLST) method according to Achtman scheme, https://guthub.com/tseemann/mlst) [26]; (ii) phylogenetic groups (ClermonTyper v1.0.0) [27]; (iii) resistance genes (ARIBA database) [28]; (iv) virulence factors (virulence factor database, VFDB) [29]; (v) serotypes (serotypeFinder O:H typing database) [30]; (vi) plasmids (PlasmidFinder) [26]; and (vii) sequence types for plasmids (plasmidMLST) [26] (Supplementary File 1). Novel STs were assigned by Enterobase [26].

A phylogenetic tree was constructed using recombination-free core genomes including ST131 genomes from Africa retrieved from Enterobase, representative clade strains from Matsumura et al. [9] and this study's ST131 sequence data to demonstrate the phylogenetic relatedness of isolates. The isolates EC958 *fim*H30 and KUN2145 *fim*H22 were used as reference and an outgroup, respectively. A similar comparative analysis was done for this study's ST410 isolates in comparison to ST410 genomes from Africa submitted to Enterobase, Roer et al. [31], and this study's sequence data. The reference genome used for the phylogenetic tree construction was YD786 (GenBank Accession Number: NZ_CP013112.1), while *fim*H53 isolates were used as outgroups for the comparative analysis of ST410 isolates. Only ST131 and ST410 genomic data with relevant metadata (year of collection, country, source type, etc.) and availability of raw reads on Enterobase were included. Mauve was used to visualize similarities of this study's ST131 C1-M27 genomic environments against a reference KUN 5781 for the presence of a phage integrase region annotated as M27PP1 and M27PP2.

Ethical approvals

Ethical clearance was obtained from the Faculty of Health Sciences Research Ethics Committee, University of Pretoria (Ethics Reference Number: 782/2018) and the Medical Research Council of Zimbabwe, Approval Number: MRCZ/A/2394.

Results

Baseline characteristics of sequenced isolates

A total of 48 ESBL-positive isolates from three provinces (Harare, Bulawayo, and Mashonaland East) were analyzed. Among these, 27 (55%) were linked to UTIs in female patients and 21 (45%) to male patients. Thirty (62%) patients were aged 21 to 60 years. Thirty-eight (80%) of the isolates were recovered from Harare with isolation years (2017 (11); 2018 (16); 2019 (11)), while five isolates (10%) originated from Chitungwiza (2017 (2); 2018 (3)), three isolates (6%) were collected from Bulawayo (2018 (2); 2019 (1)), and two (4%) were from Marondera (2017 (1); 2018 (1)). These ESBL-producing ExPEC isolates originating from the outpatient clinics of hospitals in Harare, Chitungwiza, Marondera, and Bulawayo displayed increased resistance to all antimicrobials included in the study.

Whole-genome sequence analysis

The phylogenetic tree in Fig. 1 was constructed using representative reference strains of phylogenetic groups A, B1, B2, D, and E and an outgroup *E. fergusoni*. The isolates in this study were compared in terms of the virulence factors detected and those isolates belonging to the same phylogenetic group as the representative reference strains clustered together as shown in Fig. 1.

Genomic assembly, quality control, and phylogenetic tree

Forty-eight isolates had a combined length of contigs of assembled genomes ranging from ~ 4.7 to 9.8 Mbp, with a minimum contig length required to cover 50% of the genome (N50), ranging between 84 and 340 kbp. The SNP matrix output tables with sequenced isolates aligned and compared to reference genomes were used to construct a phylogenetic tree as shown in Fig. 1.

Beta-lactamases, plasmid-mediated quinolone resistance determinants, and phylogenetic groups

The 48 ESBL *E. coli* isolates harboured diverse ESBL genes; $bla_{\text{CTX-M-15}}$ (34), $bla_{\text{CTX-M-27}}$ (11), $bla_{\text{CTX-M-14}}$ (1), $bla_{\text{CTX-M-3}}$ (1), and $bla_{\text{CTX-M-82}}$ (1). Other ESBL genes were also observed, including $bla_{\text{TEM-1}}$, $bla_{\text{OXA-1}}$, and $bla_{\text{OXA-10}}$ (Table 1). Table 2 illustrates the different beta-lactamases that were detected in different geographic regions and years. Other antimicrobial resistance (AMR) genes included the *aac (6′)-lb-cr* gene (27) encoding aminoglycoside acetyltransferase, the *qnrS* gene (4) encoding a plasmid-mediated quinolone resistance, and the *qepA* (1) encoding a quinolone pump. Twenty isolates belonged to phylogenetic group B2; 14 to phylogenetic group C; four to phylogenetic group B1; three to phylogenetic groups A, D, and F each; and lastly one to phylogenetic group G.

Multi-locus sequence typing

MLST identified five major ST clonal complexes (CCs), which included the following: ST131CC as ST131 (17), ST11380 (1), and ST11387 (1); ST23CC as ST410 (13) and ST6332 (1); ST10CC as ST617 (2) and ST218 (1); ST405CC as ST405 (1) and ST11615 (1); ST648CC as ST648 (1); ST354CC as ST354 (1); ST155CC as ST155 (1). Most of the isolates (43/48) belonged to one of these defined CCs (Table 1). Five isolates belonged to an undefined clonal complex, referred to as "unknown CC", which included ST2448 (1), ST117 (1), ST224 (1), ST678 (1), ST636 (1). The clinical presentation, year of collection, antimicrobial susceptibility profile, plasmid-mediated quinolone-resistant (PMQR) determinants, ESBL types, and phylogenetic groups of the above defined and undefined CCs are shown in Tables 1, 2, and Supplementary File 1.

Plasmid replicon types

The most predominant plasmids were Col156, ColBS512, and ColMG828, which were distributed among the ST131 and ST410 isolates as shown in Table 3. All ST131 and ST410 isolates characterized in this study harboured incompatibility group F1A (IncF1A) plasmids. Different plasmid types belonging to IncF were distributed across all strains. The IncF plasmids with F1:A2:B20 (9/19) and F2:A1: B (5/19) replicons were predominant among ST131 isolates, while the F1:A1:B49 replicon (8/14) predominated among the ST410 strains. The F36:A4: B1 replicon was detected among isolates from ST23, ST131, and ST10 CCs as shown in Supplementary File 1.

coli ST131 core genome SNP-based phylogenetic tree in an African and global context

The mapping and alignment of 19 ST131 study isolates to the reference genome EC958 produced a core genome of 4,700,950 bp. The fluoroquinolone-resistant isolates with *gyr*A and *par*C mutations formed the C/H30R cluster that comprised the C2/H30Rx and C1/H30R clades. The C2/H30Rx clade included isolates with $bla_{\text{CTX-M-15}}$ (n = 43) and isolates with both $bla_{\text{CTX-M-14}}$ and $bla_{\text{CTX-M-15}}$ (n = 1), and $bla_{\text{CTX-M-15}}$ and $bla_{\text{CTX-M-27}}$ (n = 1) and isolates without $bla_{\text{CTX-M}}$ (n = 2). The C1/H30R clade included isolates with $bla_{\text{CTX-M-27}}$ (n = 17) and $bla_{\text{CTX-M-14}}$ (n = 8) and isolates without ESBLs (n = 1) as shown in Fig. 2**.** Most Zimbabwean isolates belonged to the C2/H30Rx clade (10/19), with between 10 and 15 SNP differences between isolates from neighbouring African countries, such as Tanzania, DRC, Nigeria, Sudan, Niger, and Ethiopia as shown in Supplementary File 2. Within the C1/H30R clade, 15 of the 17 CTX-M-27-producing isolates clustered into a distinct group, the C1-M27 clade. *E. coli* ST131 C1-M27 comprised isolates from Zimbabwe (n = 9, 2017–2019); Canada (n = 1, 2008); USA (n = 2, 2013–2014); Japan (n = 2 2007, 2010); and Australia (n = 1, 2009). The novel STs assigned by Enterobase in the ST131 CC

clustered in the C1-M27 clade, and these were closely related to Canadian strains (Fig. 2).

The C1-M27 clade-specific region

The genome analysis of C1-M27 clade isolates identified an 11,894-bp region that is named M27PP1, which is a prophage integrase specific to all isolates from the C1-M27 clade and to some non-ST131 from this study that had a CTX-M27 resistance gene. BLAST and Mauve software were used to align the isolates to a reference with these regions clearly annotated (KUN5781). Three isolates from C1-M27 clade (i.e. 58EC, 60EC, 92EC) and two non-ST131 isolates (i.e. 72EC, 69EC) from this study had M27PP1 alone. Six *E. coli* ST131 C1-M27 isolates (i.e. 38EC, 45EC, 51EC, 53EC, 91EC, 67EC) aligned against the reference KUN5781 as shown in Fig. 3 had an additional insertion region named the M27PP2 situated upstream of M27PP1.

E. coli ST410 in Africa and global context

The phylogenetic reconstruction of 127 ST410 international isolates, 55 African isolates retrieved from Enterobase, and 17 isolates from Zimbabwe in this study revealed two distinct clonal lineages of ST410: lineage A with *fim*H53 (A/H53) and lineage B with *fim*H24 (B/H24) (Fig. 4). The B/H24 lineage was further divided into three sublineages: B2/H24R with the introduction of fluoroquinolone resistance by mutations in the *gyr*A and *par*C, B3/H24Rx with the introduction of $bla_{\text{CTX-M-15}}$, and B4/H24RxC with the introduction of $bla_{\text{OXA-181}}$. Zimbabwean isolates were all part of the B3/H24Rx clade.

Discussion

Emerging and established high-risk clones of ESBL-producing ExPEC are factors that negatively impact on human health management globally. This study focused on elucidating the molecular characteristics of ESBL–producing *E. coli* isolates associated with UTIs over a 2-year period (2017–2019) in Zimbabwean communities (Harare, Marondera, Bulawayo, and Chitungwiza). The Zimbabwean uropathogenic *E. coli* isolates harboured a high rate (28%) of ESBL production; similarly, high rates have been reported in studies from Tanzania (24%) [32], Algeria (31%) [33], and Rwanda (38%) [34]. In Zimbabwe's hospital setting, beta-lactam antibiotics, such as ceftriaxone, are frequently used as first-line treatment for bacterial infections, which create a selective pressure for the pathogens to evolve and adapt [35, 36].

The high diversity of uropathogenic ESBL–producing *E. coli* isolates in this study suggests the circulation of diverse community-acquired UTI clones within Zimbabwe. Most isolates belonged to major CCs, such as ST131, ST23 (ST410), ST10, ST405, ST648, ST69, ST354, and ST155. The pandemic clone, ST131, was the most detected clone in this study and is known to harbour $bla_{\text{CTX-M-15}}$, $bla_{\text{CTX-M-27}}$, and fluoroquinolone resistance. The presence of this

clone has been reported in small sample size studies from South Africa [37], Nigeria [38], and Tunisia [39]. This clone is known for its high virulence leading to infections, such as invasive bloodstream, urinary tract, and intra-abdominal infections [40].

A comparative analysis on the phylogenetic relatedness of this pandemic clone in relation to those reported from other African countries revealed a possible constant influx of ST131 strains as these isolates are mixed with other African and European sublineages in the phylogeny (Fig. 2). A local national transmission cluster of ST131 *bla*$_{\text{CTX-M-27}}$ was observed. This clade named C1-M27 within C1/H30R in ST131 has been previously reported as responsible for epidemics of ESBL-producing ExPEC in Japan [9] and has so far disseminated across different continents including Africa. The isolates clustering in this clade originated from Harare and Marondera. The proximity of these two cities could have contributed to the spread of C1-M27, although information on travel could not be retrieved upon collection of samples. Local transmission could be the largest contributor to the spread of infections with ESBL–producing *E. coli* ST131 in Zimbabwe given the small number of SNPs between isolates in the C1-M27 clade. Travel history becomes an important characteristic to consider in future studies as a better method to link these transmission dynamics.

The Zimbabwean isolates are defined by the presence of either the M27PP1 unique region or by both the M27PP1 and M27PP2 insertion regions. The *E. coli* ST131 C1-M27 isolates harboured the unique prophage-like region (M27PP1) within its chromosome, while in ST131 C2/H30Rx, it was not identified (Fig. 3). The direct flanking repeat sequences surrounding M27PP1 suggest that this region was introduced into *E. coli* ST131 C1/H30R with *bla*$_{\text{CTX-M-27}}$ by a recombination event that was then followed by the clonal expansion of the C1-M27 clade [8, 9]. Therefore, the screening for the presence of M27PP1 genetic environment is important to check for recombination since other studies have noticed that some ST131 isolates might have acquired *bla*$_{\text{CTX-M-27}}$ independently from the C1-M27 clade [8, 9].

UTIs are often preceded by colonization of the gut [41]. In a recent study by Wilmore et al. [42] on the carriage of ESBL–producing *Enterobacteriaceae* in HIV-infected children in Zimbabwe, it was observed that out of 175 collected stools, 24 isolates were ESBL-positive and nine isolates belonged to ST131 producing either CTX-M-15 or CTX-M-27. Infection as a result of colonization by these resistant strains may complicate treatment. The presence of such high rates of ESBL–producing commensal bacteria is a reflection of the high usage of these antibiotics in the public sector in Zimbabwe and its contribution to the creation of selective pressure for pathogens to evolve and spread.

Novel strains of the ST131 complex (ST11380 and ST11387) and ST405 complex (ST11615) were detected over a period of 1-year (2017–2018) harbouring *bla*$_{\text{CTX-M-15}}$, and *bla*$_{\text{CTX-M-27}}$. These novel strains originated from patients who live in one of the densely populated suburbs, where frequently poor sanitation and hygiene conditions are reported. As these resistant clones can persist in the environment, it is important to improve sanitation, water quality, and patient care through education and awareness campaigns in communities as a way of controlling the spread of such resistant clones [43, 44]. From our global analysis, other environmental-associated isolates were included, which showed the existence of different clades including C1-M27 among animal and environment; this shows how such isolates might act as reservoirs for the introduction of such clades in humans.

Enterobacterales are known for having relatively open pan-genomes that can rapidly adapt to changing selection pressures (including antibiotic usage) as observed in the ST131 and ST410 strains in this study. To the best of the authors' knowledge, this is the first report of ST410 ESBL ExPEC from Zimbabwe harbouring $bla_{CTX\text{-}M\text{-}15}$ along with other antimicrobial resistance (AMR) genes. Recent studies have indicated the *E. coli* ST410 as another successful pandemic ExPEC lineage [31]. Our study results corroborate this theory; however, as national surveillance programmes monitoring only local epidemiology, global surveillance programmes are required to follow the dissemination of pandemic clones [45]. A comparative analysis of the ST131 and ST410 isolates revealed the presence/predominance of the self-transmissible IncF plasmids in ST410, which like in ST131 allows the bacteria to capture additional virulence genes and resistance determinants [5]. Similar studies on poultry, companion animals, freshwater fisheries, and swine from Tunisia and South America have described the presence of IncF plasmids in ST410 as a contributing factor to the spread of ESBL [46, 47]. Due to the lack of ecological barriers and trading of food items between nations and continents, feacal carriage of these resistant clones may contribute greatly to the spread of AMR [48].

IncF-type plasmids have a narrow host range (limited to *Enterobacteriaceae*) and contribute to bacterial fitness via antibiotic resistance and virulence determinants [49]. These plasmids have been associated with the rapid emergence and global spread of $bla_{CTX\text{-}M\text{-}15}$, as well as genes encoding resistance to aminoglycosides and fluoroquinolones (e.g. *aac(6')-Ib-cr*, *qnr*, *armA*, *rmtB*) [49]. There was an association between $bla_{CTX\text{-}M\text{-}15}$ and $bla_{OXA\text{-}1}$, as well as the *aac(3)-IIa* and *aac(6')-Ib-cr* in clade C2 ST131 carrying IncF plasmids, the majority of which came from patients in Harare.

Previous work from North America suggested that the *H*30-R/C1 clade of ST131 most commonly carries IncF-type F1:A2:B20 plasmids and the *H*30-Rx/C2 clade is associated with IncF-type F2:A1:B-plasmids [50]. In this study, plasmid types were associated with different sublineages of ST131. For example, IncF-type F36:A4:B1 plasmids were most frequently seen in clade C2, whereas IncF-type F2:A1:B-plasmids were mostly seen in clade C1. A similar observation in ST410 was noted with F1:A1:B49 frequently being observed in B3/H24Rx.

The ST10 CC with $bla_{CTX\text{-}M\text{-}15}$ together with other broad-spectrum beta-lactamases, such as $bla_{OXA\text{-}1}$ and $bla_{TEM\text{-}1}$, has been described in clinical isolates from Nigeria [32] and Egypt [51], as well as in poultry from Gambia [52], vegetables from South Africa [39], and enteroaggregative *E. coli* (EAEC) from Nigeria [53]. EAEC is an established diarrhoeagenic pathotype, transmitted either via consumption of meat products and vegetables or through contact with animals, which has been suggested as a potential source of ESBL bacteria causing diarrhoea and UTIs at the same time. EAEC with ExPEC markers belonging to ST10 has been found to be potential agents of UTIs. Recombination events are important in the evolution of pathogenic EAEC-ST10 UTI linked clones [54].

Several of the other clones identified in this study have been detected elsewhere. ST405 and ST648 are emerging clones associated with carbapenemases, specifically New Delhi metallo-beta-lactamases [55]. ST405 *E. coli* isolates producing CTX-M-3, CTX-M-14, or CTX-M-15 have been described for using self-transmissible IncF plasmids in acquiring different resistance genes within its clonal lineage [56]. ST354 with CTX-M-15 has been described in humans [57] and stray dogs [58], while to the best of the authors' knowledge, this is the first

report on the presence of *bla*$_{CTX-M-27}$ and the insertion prophage integrase M27PP1 in this clone. ST155 has been identified in the current study to harbour *bla*$_{CTX-M-82}$. This was also observed in a Gambian study reporting the diversity of *E. coli* isolates from backyard chickens and guinea fowl; 32% (22/68) of the isolates harboured *bla*$_{CTX-M-82}$ [52]. ST354 and ST155 might be strains, which can be exchanged between animals and humans.

This study offers a better understanding of the epidemiology of Zimbabwean ESBL-producing ST131 and ST410 and there close relationship to internationally disseminating ST131 and ST410 strains. It is evident that the Zimbabwe ST131 and ST410 strains are part of the international lineages and that several introductions combined with national transmission have formed the current *E. coli* population. The clonal nature of the ST131 and ST410 lineage, with highly conserved plasmids in some sublineages, complicates estimation of local circulation and transmission, and highlights the importance of the space time epidemiological link events in the genomic era. However, very close genetic relationships (10–15 SNPs) could indicate a direct transmission even if the epidemiological link is unknown.

Improved strategies for the control of these clones will impact positively on public health management. A number of antimicrobial resistance control strategies have been put in place by different countries. These strategies include (i) improvement of awareness and understanding of antimicrobial resistance through effective communication, education, and training; (ii) strengthening the knowledge and evidence base through surveillance and research; (iii) reduction of the incidence of infection through effective sanitation, hygiene, and infection prevention measures; (iv) optimization of the use of antimicrobial medicines in human and animal health; and (v) the development of the economic case for sustainable investment that takes account of the needs of all countries, and increased investment in new medicines, diagnostic tools, vaccines, and other interventions [59].

The One Health approach is an important initiative for all countries especially in the sub-Saharan countries where information is scarce. The significance of this study from Zimbabwe was to define the scope of the resistance problem. A limitation of this study was the small sample size of only 48 ESBL *E. coli* isolates with most isolates selected from Harare; however, results from this study contributed to the baseline molecular information for isolates/clones currently and previously linked to resistance in Zimbabwe. Short-read plasmid profile analysis was explored, which does not give a full description of the plasmids as compared to long read sequencing, although this technique provides some indication on the different plasmids within each isolate. Plasmid conjugation experiments could not be performed due to funding and time constraints; therefore, future studies should focus on such experiments to get an in-depth understanding of the characteristics of the ESBL-encoding plasmids. As far as we are aware, this is the first study to be done in Zimbabwe to provide baseline data on virulence, antimicrobial resistance, and detection of specific lineages of ExPEC circulating in our communities using WGS.

Conclusion

Our study has shown a high level of *E. coli* diversity in terms of STs, antimicrobial resistance genes, serotypes, and virulence genes, which underlines the necessity for concerted efforts for continuous surveillance of the ESBL-producing ExPEC clones. Targeting specific STs (e.g.

ST131, ST410, and ST405) through control programs will substantially decrease the spread of ESBLs among ExPEC.

Availability of data and material

The raw data supporting the conclusions of this manuscript will be made available by the authors, without undue reservation, to any qualified researcher. All short reads and assemblies associated with this study are available at NCBI under BioProject: PRJNA721804; individual BioSamples are listed in Supplementary File 3.