Whole genome sequencing and phylogenetic analysis to examine acquisition of *Escherichia coli* clones during international travel

Running title: WGS for defining travel acquisition of *E. coli* clones

Samantha dos S Tufic-Garutti, PhD[a,b]. Luís G de A Longo, PhD[a,c]. Roobinidevi Ragupathy, PhD[d]. Maliha Akram, PhD[d]. Mark C Enright, PhD[d]. Beatriz M Moreira, MD, PhD[a]. Káris M de P Rodrigues, MD, PhD [c,e #]

[a]Instituto de Microbiologia Paulo de Góes, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil

[b]Universidade Nilton Lins, Manaus, Brazil

[c]Faculdade de Medicina, Instituto de Educação Médica (IDOMED), Universidade Estácio de Sá, Rio de Janeiro, Brazil

[d]Manchester Metropolitan University, Manchester, United Kingdom

[e]Centro de Informação em Saúde para Viajantes, Faculdade de Medicina, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil

[#]To whom correspondence should be addressed. Email: karismpr@gmail.com

Funding

This study was supported by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) - Finance Code 001; Conselho Nacional de Desenvolvimento Científico e Tecnológico - INPRA- Instituto Nacional de Pesquisa em Resistência Antimicrobiana-Brazil CNPq 465718/2014-0; CNPq312205/2019-8 (Characterization of antimicrobial resistance, virulence genes and population structure of clinical and environmental strains of *Escherichia coli* obtained in the state of Rio de Janeiro). S.S.T.G. was a CNPq fellow. L.G.A.L. is a Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ) [grant number: E-26/200.228/2022 – 274327] fellow.

Abstract

International travel facilitates the acquisition and carriage of extended-spectrum beta-lactamase-producing *Escherichia coli* (ESBL-E). We describe genomes of predominant ESBL-E clones detected before and after travel among subjects departing from Rio de Janeiro, Brazil, during 2015-2021, and genomes publicly available from countries visited by travelers. WGS (Illumina NovaSeq) was performed on 70 ESBL-E isolates from 66 travelers (18 pre- and 52 post-travel). Sequence type (ST), antimicrobial resistance (AMR), virulence genes, and plasmids were determined by Center for Genomic Epidemiology tools. Phylogenetic trees were constructed with each of the most frequent ST of travelers' genomes (TG) and genomes with the same ST from Enterobase (EG). Other analyses were performed with Prokka, Roary, SNP-sites, and FastTree. Trees were visualized with iTOL. Among 70 ESBL-E, the clonal composition was quite diverse, with 41 different ST, with predominance before of CC10 and CC131, and after travel, CC10, CC131, CC69, and CC38, and three ST described for the first time: ST14408 and ST14412 (CC10), and ST4411 (CC20). Core genome phylogenetic analysis revealed 17 clusters, eight of which formed by post-travel TG and EG detected in the same country visited by traveler. We observed an increased number and diversity of AMR genes, plasmids, and virulence genes in post-travel isolates, although we only found statistical significance for IncFIB plasmid. Genome clustering supported the high-risk clone acquisition and AMR during international trips. More than half of detected clones were related to ExPEC and showed an increased number and diversity of AMR and virulence-related genes, as well as plasmids in post-travel isolates.

Keywords: Acquisition during travel, Colonization by ESBL-producing *Escherichia coli*, Whole genome sequencing

Introduction

Escherichia coli is among the most frequently detected bacteria worldwide and the most isolated pathogen from the so-called 'ESKAPE' group. These pathogens are significant contributors to life-threatening infections and are renowned for their ability to acquire and disseminate antimicrobial resistance [1,2]. International travel is critical for the spread of highly resistant strains [3]. Aminopenicillin-resistant *E. coli* was the predominant antimicrobial-resistant pathogen associated with death in 47 countries in 2019 [4]. Infections caused by extended-spectrum beta-lactamase (ESBL)-producing *E. coli* (ESBL-E) pose an increased challenge due to limited therapeutic choices. ESBL, along with AmpC and carbapenemase genes, are frequently located on plasmids, facilitating their dissemination within the gut microbiota, which serves as a major reservoir for the persistence and spread of antimicrobial resistance. 'Successful' bacterial clones also play a pivotal role in the dissemination of antimicrobial resistance. Notably, since the 2000s, the global spread of the pandemic *E. coli* sequence type 131 (ST131) high-risk clone is a significant contributor to the worldwide increase in ESBL prevalence, particularly CTX-M-15 ESBL [5,6].

International travel increases the risk of ESBL-E carriage, particularly in certain risk areas such as Southeast Asia and the Indian subcontinent [7]; isolates acquired as gut colonizers may be a major source of infection, as urinary tract infections (UTI), with *E. coli* being the primary causative agent of these infections globally [8–10], including in Brazil [11–13]. The detection of ESBL-E in meat and raw vegetables underscores the importance of additional sources contributing to the persistence and dissemination of such microorganisms within the community [14,15].

Most previous studies on ESBL-E gut colonization during travel were performed using culture-based techniques; which can underestimate colonization risk [9,16]. In a recent study with travelers departing from three US cities, by culture-based technique, 41% of travelers acquired ESBL-E strains. However, by metagenomic analysis, almost 70% of these travelers acquired new *E. coli* strains, some of which might have been new ESBL-E isolates. In addition, some of the strains regarded as new acquisitions in travelers' guts may have already been present before travel, but undetected by culture-based techniques [17]. Whole genome sequencing (WGS) could provide more accurate data, as a high-resolution surveillance tool that allows a

more thorough exploration of the fecal microbiome, with more precise tracking of antimicrobial-resistant pathogens at the individual strain level [18].

We recently reported that among international travelers departing from Brazil, the prevalence of ESBL-E gut colonization before travel was 9% in 2015-2019, increasing to 22% among returning subjects with a negative anal swab culture before departure [19]. We hypothesized that WGS could help to assess whether newly detected isolates were indeed new acquisitions; strains supposedly acquired in a certain country during travel should be more similar to local strains than to those detected in Brazil. Thus, the objective of the present study was to analyse and compare whole genome sequences of the predominant ESBL-E clones detected before and after travel of subjects participating in our project with genomes available in a public database, obtained from humans, animals, food, and the environment in countries visited by travelers.

Methods

Study setting and bacterial isolates

Study isolates were recovered during a project to determine the frequency of gut colonization with multidrug-resistant Enterobacterales acquired by travelers departing from Rio de Janeiro, Brazil. Briefly, surveillance for gut colonization was performed by self-collected anal swabs before and after travel by subjects who consented to participate and answer a questionnaire. Isolates were coded as EC, followed by the participant study number, and the letter A when detected before travel, or B when detected after return. Among 243 recruited travelers from 2015 to 2021, 66 (27%) carried ESBL-E: 14 had pre-travel colonization only, 48 had post-travel colonization only, and four had pre- and post-travel colonization, comprising 70 total isolates. Gut carriage of ESBL-E of the first 59 travelers was described previously [8]. Seven new subjects were included in the present report. Of the 52 travelers who had ESBL-E post-travel colonization, 20 (38%) returned from Sub-Saharan Africa, 11 (21%) from South America, nine (17%) from Southeast Asia, five (10%) from Indian Subcontinent, five (10%) from other regions of Asia, and one (2%) each from North Africa and Oceania.

Whole-genome sequencing, assembly, and annotation

ESBL-E isolates were sent for WGS by the Illumina NovaSeq 6000 (Illumina, USA) at MicrobesNG (United Kingdom). Adapters were removed and reads trimmed using Trimmomatic

v.0.39 with a sliding window cutoff of Q15 [20]. Reads were *de novo* assembled with Unicycler v.0.4.8 [21], assemblies quality checked with Quast v.5.2.0 [22] and annotated with Prokka v.1.14.6 [23].

Genomic features, resistome, virulome, and plasmidome

Sequence types (ST) were determined with MLST 2.0, serotype with SerotypeFinder 2.0, FimH type with FimTyper 1.0, acquired antimicrobial resistance (AMR) genes with ResFinder 4.1, virulence-related genes with VirulenceFinder 2.0, and plasmids incompatibility groups with PlasmidFinder 2.1 databases, all with a minimum threshold of 90% identity, available at https://www.genomicepidemiology.org/.

Core genome phylogenetic analyses

We constructed a phylogenetic tree for each of the most frequent ST of travelers' genomes (TG) with a group of genomes of the same ST retrieved from Enterobase (https://enterobase.warwick.ac.uk/) referred to as Enterobase genomes (EG). The tree included only complete EG obtained in the same country visited by the traveler, with information about the specimen source (human, animal, food, or environment) and year of isolation.

After the annotation, Roary (v.3.13.0) [24] was used for the alignment and inference of the isolates' core genome. Single nucleotide polymorphisms (SNP) were determined with SNP-sites (v.2.5.1) [25]. The core genome alignment of the isolates was used to build the maximum likelihood tree with the FastTree v.2.1.10 tool [26]. The phylogeny was tested with 1,000 replicates and the tree was visualized with iTOL v.5.6.1 [27]. We defined a cluster empirically by visual inspection of genome groupings in phylogenetic trees and proposed the likely travel acquisition of isolates in clusters including only EG obtained off Brazil. Clusters were named with the corresponding ST followed by "C" and a numeral.

Statistical analysis

Data were analysed with SPSS statistical suite (v.24). Categorical variables were tested using Fischer's exact test. Continuous and aggregated variables were evaluated with parametric mean comparison (unpaired t-test) or, if non-normal distribution, with a non-parametric Wilcoxon rank-sum test of median or distribution. A *p*-value <0.05 was considered statistically significant.

Data availability

This whole genome shotgun project has been deposited at DDBJ/ENA/GenBank under the accession PRJNA1071675.

Results

General characteristics and clonal composition of ESBL-producing *E. coli* isolates

Among the 70 ESBL-E, genome sizes ranged from 4,447,043 to 5,497,874 bp, the number of contigs ranged from 59 to 258, and the G+C content from 50.33 to 50.98% (Table S1). The isolates were included in 15 CC and 13 ST characterized as singletons (Table S2). Three ST were described for the first time: ST14408 and ST14412 of CC10, and ST4411 of CC20. The predominant clones in pre-travel colonization were CC10 (n=8, 44%) and CC131 (n=4, 22%), and as post-travel colonization, CC10 (n=18, 35%), CC131 (n=5, 10%), CC69 (n=4, 8%) and CC38 (n=4, 8%). Of the four travelers with pre- and post-travel ESBL-E colonization, in three the post-travel isolate belonged to an ST different from the pre-travel isolate, and in one, both were the same ST. These two isolates (EC165A and EA165B) were obtained on May 5th, 2017 and June 27th, 2017, and belonged to ST1201, O18ab:14 and *fimH53* type, carried AMR genes related to beta-lactamase production ($bla_{\text{TEM-1B}}$ and $bla_{\text{CTX-M-2}}$), resistance to aminoglycosides (*aph(3'')-Ib*, *aph(3'')-Ia*, *aph(6)-Id)*, trimethoprim (*dfrA8*), fluoroquinolones (*qnrS1*), sulphonamides (*sul2*), tetracycline (*tetA* and *tetB*); virulence genes *terC22* and *terC5*; and IncHI2 and IncN plasmids. The post-travel genome additionally carried the AMR genes *dfrA7* and *sul1*, the virulence gene *traT162*, and the IncFII plasmid. The core genomes shared only 81 SNPs.

Among these 70 isolates, except for one isolate with $bla_{\text{TEM-176}}$, all carried a $bla_{\text{CTX-M}}$ encoding gene. The most frequent variant was $bla_{\text{CTX-M-15}}$, present in four (36%) of the 11 ST10 isolates, seven (78%) of the nine ST131, two of the five ST38, and two of the four ST69. Although just isolates with phenotypic and genotypic tests indicating ESBL production had been selected for WGS, 14 isolates did not return any ESBL-producing genes. These TG were maintained in the phylogeny analyses but excluded from the assessment of AMR genes, plasmids, and virulence gene content.

Phylogenetic analysis

This analysis included TG of ST38 (n=5, shown in Fig. 1), ST69 (n=4, shown in Fig. 2), ST10 (n=11, shown in Fig. 3 and Fig. S1), and ST131 (n=9, shown in Fig. 4 and Fig. S2), in addition to EG of isolates from the same ST obtained in countries visited by travelers. The

phylogenetic trees revealed 17 clusters; in eight of these, the post-travel genomes did not include any isolates from Brazil. In five of the eight clusters, one or more EG were obtained previously to the TG in the same country visited by the traveler: ST38-C2, EG in 2013, and TG in 2017 and 2018, from Peru; ST69-C2, two EG in 2018 and 2020, and TG in 2021, from Thailand; ST10-C5, EG in 2015 and 2016, and TG in 2018, from Kenya; ST131-C1, EG in 2012-2013 and two TG in 2018, from Thailand; ST131-C3, EG in 2018 and TG in 2019, from Madagascar. In two clusters, the EG was obtained previously to the TG, but in a country different from that visited by the traveler. One of these clusters had isolates from countries sharing a border in Southeast Asia: ST69-C3, two EG in 2015, from Laos, and TG in 2018, from Thailand, Cambodia, and Vietnam. The other cluster had isolates from countries geographically apart: ST10-C2, two EG in 2009, from Tanzania, and TG in 2018 from Egypt. Finally, in one cluster, the EG was obtained in the same country visited by the traveler, but in a later year: ST69-C1, TG in 2017, and EG in 2018, from Thailand. Three of these eight clusters included a TG and an EG with the same $bla_{\text{CTX-M}}$ gene type: ST38-C2 ($bla_{\text{CTX-M-14}}$), ST131-C1 ($bla_{\text{CTX-M-27}}$) and ST131-C3 ($bla_{\text{CTX-M-15}}$); the other clusters included isolates without an ESBL encoding gene or had only genomes with different $bla_{\text{CTX-M}}$ gene types. One of these eight clusters (ST10-C5) had the group formed by EG from animal and environment, all others were from human sources.

In three other clusters the post-travel genome grouped with EG from Brazil: ST38-C3, ST10-C1, and ST10-C3. The clusters also had different $bla_{\text{CTX-M}}$ gene types or included isolates without ESBL genes. ST10-C3 had a particular feature: the cluster included a post-travel isolate with $bla_{\text{CTX-M-14}}$ obtained from the same anal swab specimen source of the ST167/CC10 isolate carrying the $bla_{\text{OXA-181}}$ carbapenemase gene, as previously described [28].

The other six clusters included pre- and post-travel genomes grouped with isolates of different countries: ST38-C1, ST10-C4, ST131-C6, ST131-C2, ST131-C4, and ST131-C5.

Pre- and post-travel acquired antimicrobial resistance genes

In the group of pre- (n=13) and post-travel (n=43) genomes returning with ESBL encoding genes considered as a group, the median number of AMR genes per isolate was eight; the number of plasmids was three and four, respectively; and of virulence genes, eight and 13, respectively, as shown in Fig. 5. These differences were not statistically significant.

The most frequently detected non-beta-lactam AMR genes were to the class of tetracyclines and aminoglycosides, both present in 77% of isolates, and sulphonamides, present

in 75% of isolates (Fig. S3). The predominant resistance genes detected among pre- and post-travel isolates, respectively, were: to tetracyclines, *tetA* (10, 77% and 18, 42%) and *tetB* (1, 8% and 16, 37%); to aminoglycosides, *aph(6)-Id* and *aph(3)-Ib* (8, 61%, and 24, 56%); and to sulphonamides, *sul2* (8, 61% and 22, 51%, shown in Fig. S4).

Pre- and post-travel plasmids and virulence genes

The predominant plasmid groups among pre- and post-travel isolates, respectively, were Col (7, 54% and 30, 70%) and IncFIB (5, 38% and 31, 72%, p=0.046) (shown in Fig. S4 and S5). Finally, the most frequent virulence gene groups were protectin, serum resistance, and dispersin (11, 85% and 38, 88%), siderophores (11, 85% and 35, 81%), adhesins (6, 46% and 28, 65%) and toxins (6, 46% and in 26, 60%) (shown in Fig. S6 and S7). The most frequent genes in the protectin, serum resistance, and dispersin group were *iss* (9, 69% and 28, 65%), *traT* (7, 54% and 29, 67%), and *ompT* (5, 38% and 22, 51%); among siderophores were *fyuA* (8, 61% and 24, 56%), *sitA* (7, 54% and 24, 56%), and *irp2* (7, 54% and 23, 53%). The most frequent genes in the adhesins group were *pap* (5, 38% and 11, 26%), *iha* (3, 23% and 12, 28%), and *hra* (2, 15% and 11, 26%), and among toxins were *sat* (3, 23% and 11, 26%), and *hlyEF* (only in post-travel isolates, 10, 23%). Except for the frequencies of IncFIB plasmid group, all differences between pre- and post-travel isolates were not statistically significant. Additional features such as *fimH* type, O:H type, and the virulence genes per isolate are presented in Fig. S6.

Discussion

To make a deeper analysis of ESBL colonization during travel, we selected for WGS all ESBL-E isolates obtained during the study period. Among the 70 isolates, in 14 (20%), the CTX-M-encoding genes were not found in the sequenced genome. We suppose that the *E. coli* isolate subcultured from our stocks and sent for sequencing lost the $bla_{\text{CTX-M}}$ gene-carrying plasmid due to transportation stress or lack of AMR selective pressure, a phenomenon previously described [29,30].

One single isolate (EC287B, of ST1421) did not carry a $bla_{\text{CTX-M}}$-type gene. This isolate carried the $bla_{\text{TEM-176}}$ and the plasmid-mediated AmpC bla_{MIR} gene, had a phenotypic test indicating ESBL production, and was resistant to cefoxitin, cefotaxime, and ceftazidime in a disk diffusion susceptibility test [31]. TEM-176 has been occasionally listed among ESBL-producing isolates; however, to date, there is no characterization of this beta-lactamase as an ESBL, and the

Beta-Lactamase DataBase (BLDB) (http://bldb.eu/BLDB.php?prot=A#TEM, accessed on January 11, 2024) has a question mark under ~~–~~functional information" for this enzyme [32].

The clonal composition of ESBL-E isolates was quite diverse, with 41 different ST, as expected for the intestinal microbiota. More than half (56%) of the STs corresponded to high-risk clones and extra-intestinal pathogenic *E. coli* (ExPEC), [33–37] frequently detected as post-travel colonization, reinforcing the potential risk of travel as a source of bacterial pathogens. The chance of acquiring high-risk clones was further evidenced in the present project among isolates of a subject returning from Sub-Saharan Africa with a carbapenemase encoding gene, as we described previously [28]; isolate 184/2aE of ST167/CC10, carrying the carbapenemase encoding gene $bla_{\text{OXA-181}}$, was isolated from the same specimen with isolate EC184B carrying $bla_{\text{CTX-M-14}}$ included in cluster ST10-C3. These results illustrate the increased diversity observed in the gut microbiota after travel, as reported by others [38,39]. Indeed, the present study collection included four travelers with pre- and post-travel ESBL-E colonization, but in only one the ST of both isolates was the same (ST1201), persisting for at least two months.

In each of the phylogenetic trees with the most frequent ST of travelers', in addition to the TG, we had one EG from at least five countries. In this way, genomes from countries other than that visited by the traveler were used as comparison groups. We propose that clusters formed with a post-travel TG and an EG obtained in the same country visited by the traveler were probably acquired in these countries. Therefore, ST38 isolates were probably acquired in Peru, ST69 in Southeast Asia, ST10 in Kenya, and ST131 in Thailand and Madagascar.

ST38 has been described in Peru in different years [40], ST69 has been described as a cause of various infections in Laos [41], and was acquired by travelers in this country [39]. The ST10 clone has been reported to colonize and cause human infection in the visited country [42]. And the ST131 was reported in animals, causing infection in hospitalized patients in Madagascar and Thailand, and the environment in Thailand [43–46].

A different scenario occurred with clusters of post-travel TG and EG from Brazil, as observed in ST38 and ST10 trees. A version of the clone without an ESBL gene may already have been present in travelers' guts before departure and may have acquired a new plasmid with AMR genes during the trip. Indeed, in a subject with the same ST before and after travel (ST1201), we observed that the post-travel isolate (EC165B) acquired a plasmid, two AMR

genes, and one virulence gene, compared with the pre-travel isolate. This same phenomenon has been described by others [47], showing the importance of studies with WGS.

Several studies have shown the presence of ESBL-E in vegetables, commercial meat, and the gut of food handlers. The findings in all these sources reinforce the circulation of isolates among humans, animals, and environmental settings, showing the importance of studying microbiology as a One Health system for a deeper understanding of microbial diseases [5,14,15].

Several local factors have been associated with a higher prevalence of ESBL-E colonization, including an increased proportion of individuals >65 years old in the community, frequent use of cephalosporins, the number of hospital facilities, poultry, and pig population densities, and the percentage of agricultural land [48–50]. Additionally, individual risk factors for those without prior healthcare exposure extend beyond local determinants, including antimicrobial use, diarrhea, and contact with local people, animals, the environment, and food [50–52].

When we investigated if pre- and post-travel isolates could differ in the number of AMR and virulence genes and plasmids, we detected an increased number of genes encoding resistance to several antimicrobial classes, but differences were not statistically significant, possibly due to our small sample size and to the bias towards being a group in which there is a tendency for predominance of resistance genes. Other studies using metagenomic analysis and without selection of isolates by AMR markers, or including other bacteria in addition to *E. coli*, have reported an increase in the abundance of AMR genes in post-travel compared to pre-travel isolates [9,53,54]. Travel is indeed an important factor in the spread of resistance determinants, mainly genes encoding ESBL, carbapenemase, plasmid-mediated quinolone resistance, and colistin resistance (*mcr* genes) [55]. Among the ESBL genes, $bla_{\text{CTX-M}}$ is quite widespread globally in several niches, including among humans, other animals, and the environment [56,57] the acquisition of high-risk clones harboring this gene is widely reported [19,50,55].

We observed increased diversity of plasmids among post-travel isolates, with a statistically significant difference for IncFIB. IncF and its replicon types are pandemic conjugative plasmids mainly detected in ExPEC, and the plasmid incompatibility group most frequently detected in travel-acquired isolates. This finding may be related to the acquisition of AMR genes, such as $bla_{\text{CTX-M}}$ and carbapenemase-encoding genes [29,38].

The virulence genes were abundant in both pre- and post-travel isolates. Despite the lack of significance, adhesins, toxins, and bacteriocins were proportionally increased in post-travel isolates. These genes are particularly related to colonization (adhesins and bacteriocins) and intestinal infection (toxins), present in intestinal pathotypes and ExPEC [38,58,59]. One post-travel isolate carried the *aggR* gene, which characterizes Enteroaggregative *E. coli* (EAEC), and two others had *eae*, of Enteropathogenic *E. coli* (EPEC), also described in travelers departing from the Netherlands [60]. Other virulence genes, such as the *fyuA*, *irp2*, *sat,* and *senB* genes, the *pap* encoding the P fimbria, detected in our study, have already been detected with greater frequency in isolates considered as persistent and in epidemic clones [38].

Limitations

The relatively small sample size and wide dispersion of sequenced isolates may have prevented the finding of statistically significant differences in prevalences of genome features. The putative loss of ESBL-encoding plasmids in some isolates may have contributed to this constraint. However, these limitations did not weaken the relevance of the results presented, which highlight the high frequency of high-risk clones and the different possibilities for the contamination and dispersion of these microorganisms around the world. The use of one single sample from each specimen could have impacted the chances of finding resistant samples but was decided after a pilot study where we found little differences among the isolated *E. coli* samples in the same specimen [8].

Conclusions

Data obtained with phylogenetic analyses of *E. coli* core genomes from travelers reinforced the hypothesis of the acquisition of high-risk clones in countries visited during international trips. More than half of the clones detected in this collection were related to ExPEC pathotypes with increased number and diversity of AMR genes, virulence-related genes, and plasmids in post-travel isolates, though with differences not statistically significant.

Funding

This work was supported by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) - Finance Code 001; Conselho Nacional de Desenvolvimento Científico e Tecnológico - INPRA- Instituto Nacional de Pesquisa em Resistência Antimicrobiana-Brazil CNPq 465718/2014-0; CNPq312205/2019-8 (Characterization of antimicrobial resistance, virulence

genes and population structure of clinical and environmental strains of *Escherichia coli* obtained in the state of Rio de Janeiro). S.S.T.G. was a CNPq fellow. L.G.A.L. is a Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ) [grant number: E-26/200.228/2022 – 274327] fellow.

CRediT authorship contribution statement

S.S.T.G., L.G.A.L., K.M.P.R., and B.M.M.: conceptualization and supervision. S.S.T.G. and L.G.A.L.: formal analysis and writing original draft. L.G.A.L. and S.S.T.G.: formal analysis and data curation. S.S.T.G. and L.G.A.L.: investigation. M.C.E., R.R. and M.A.: investigation and formal analysis. K.M.P.R., B.M.M., and M.C.E. Writing – review & editing

Conflict of Interest

The authors have declared no conflict of interest.

Acknowledgments

None

References

[1] Benkő R, Gajdács M, Matuz M, Bodó G, Lázár A, Hajdú E, et al. Prevalence and antibiotic resistance of ESKAPE pathogens isolated in the emergency Department of a tertiary care teaching hospital in Hungary: A 5-Year Retrospective Survey. Antibiotics 2020;9:624. https://doi.org/10.3390/antibiotics9090624.

[2] Tozluyur A. Fosfomycin in the treatment of New Delhi Metallo-β-Lactamase-5 (blaNDM-5)-producing *Escherichia coli* infection. Ger J Microbiol 2024;4:1–5. https://doi.org/10.51585/gjm.2024.1.0028.

[3] Frost I, Van Boeckel TP, Pires J, Craig J, Laxminarayan R. Global geographic trends in antimicrobial resistance: the role of international travel. Journal of Travel Medicine 2019:taz036. https://doi.org/10.1093/jtm/taz036.

[4] Mestrovic T, Robles Aguilar G, Swetschinski LR, Ikuta KS, Gray AP, Davis Weaver N, et al. The burden of bacterial antimicrobial resistance in the WHO European region in 2019: a cross-country systematic analysis. The Lancet Public Health 2022;7:e897–913. https://doi.org/10.1016/S2468-2667(22)00225-0.

[5] Sallem N, Hammami A, Mnif B. Trends in human intestinal carriage of ESBL- and carbapenemase-producing Enterobacterales among food handlers in Tunisia: emergence of C1-M27-ST131 subclades, *bla* OXA-48 and *bla* NDM. Journal of Antimicrobial Chemotherapy 2022;77:2142–52. https://doi.org/10.1093/jac/dkac167.

[6] Karaiskos I, Giamarellou H. Carbapenem-Sparing Strategies for ESBL Producers: When and How. Antibiotics 2020;9:61. https://doi.org/10.3390/antibiotics9020061.

[7] Rodríguez-Molina D, Berglund F, Blaak H, Flach C-F, Kemper M, Marutescu L, et al. International travel as a risk factor for carriage of extended-spectrum β-lactamase-producing *Escherichia coli* in a large sample of European individuals—The AWARE study. IJERPH 2022;19:4758. https://doi.org/10.3390/ijerph19084758.

[8] Patjas A, Martelius A, Ollgren J, Kantele A. International travel increases risk of urinary tract infection caused by extended-spectrum beta-lactamase-producing *Enterobacterales* —three-arm case-control study. Journal of Travel Medicine 2024;31:taad155. https://doi.org/10.1093/jtm/taad155.

[9] Dallman TJ, Neuert S, Fernandez Turienzo C, Berin M, Richardson E, Fuentes-Utrilla P, et al. Prevalence and persistence of antibiotic resistance determinants in the gut of travelers returning to the United Kingdom is associated with colonization by pathogenic *Escherichia coli*. Microbiol Spectr 2023;11:e05185-22. https://doi.org/10.1128/spectrum.05185-22.
[10] Woerther P-L, Andremont A, Kantele A. Travel-acquired ESBL-producing *Enterobacteriaceae*: impact of colonization at individual and community level. Journal of Travel Medicine 2017;24:S29–34. https://doi.org/10.1093/jtm/taw101.
[11] Tano ZN, Kobayashi RK, Candido EP, Dias JB, Perugini LF, Vespero EC, et al. Susceptibility to first choice antimicrobial treatment for urinary tract infections to *Escherichia coli* isolates from women urine samples in community South Brazil. The Brazilian Journal of Infectious Diseases 2022;26:102366. https://doi.org/10.1016/j.bjid.2022.102366.
[12] Bezabih YM, Sabiiti W, Alamneh E, Bezabih A, Peterson GM, Bezabhe WM, et al. The global prevalence and trend of human intestinal carriage of ESBL-producing *Escherichia coli* in the community. Journal of Antimicrobial Chemotherapy 2021;76:22–9. https://doi.org/10.1093/jac/dkaa399.
[13] Mohamed HS, Houmed Aboubaker M, Dumont Y, Didelot M-N, Michon A-L, Galal L, et al. Multidrug-resistant Enterobacterales in community-acquired urinary tract Infections in Djibouti, Republic of Djibouti. Antibiotics 2022;11:1740. https://doi.org/10.3390/antibiotics11121740.
[14] Soncini JGM, Cerdeira L, Sano E, Koga VL, Tizura AT, Tano ZN, et al. Genomic insights of high-risk clones of ESBL-producing *Escherichia coli* isolated from community infections and commercial meat in southern Brazil. Sci Rep 2022;12:9354. https://doi.org/10.1038/s41598-022-13197-y.
[15] Colosi IA, Baciu AM, Opriș RV, Peca L, Gudat T, Simon LM, et al. Prevalence of ESBL, AmpC and carbapenemase-producing Enterobacterales isolated from raw vegetables retailed in Romania. Foods 2020;9:1726. https://doi.org/10.3390/foods9121726.
[16] von Wintersdorff CJH, Penders J, Stobberingh EE, Lashof AMLO, Hoebe CJPA, Savelkoul PHM, et al. High rates of antimicrobial drug resistance gene acquisition after international travel, the Netherlands. Emerging Infectious Diseases 2014;20:649–57. https://doi.org/10.3201/eid2004.131718.
[17] Worby CJ, Sridhar S, Turbett SE, Becker MV, Kogut L, Sanchez V, et al. Gut microbiome perturbation, antibiotic resistance, and Escherichia coli strain dynamics associated with international travel: a metagenomic analysis. The Lancet Microbe 2023;4:e790–9. https://doi.org/10.1016/S2666-5247(23)00147-7.
[18] Baker KS, Jauneikaite E, Nunn JG, Midega JT, Atun R, Holt KE, et al. Evidence review and recommendations for the implementation of genomics for antimicrobial resistance surveillance: reports from an international expert group. The Lancet Microbe 2023;4:e1035–9. https://doi.org/10.1016/S2666-5247(23)00281-1.
[19] Tufic-Garutti S dos S, Almeida Ramalho JV, Guilherme de Araújo Longo L, Caramano de Oliveira G, Rocha GT, Vilar LC, et al. Acquisition of antimicrobial resistance determinants in *Enterobacterales* by international travelers from a large urban setting in Brazil. Travel Medicine and Infectious Disease 2021;41:102028. https://doi.org/10.1016/j.tmaid.2021.102028.
[20] Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 2014;30:2114–20. https://doi.org/10.1093/bioinformatics/btu170.
[21] Wick RR, Judd LM, Gorrie CL, Holt KE. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput Biol 2017;13:e1005595. https://doi.org/10.1371/journal.pcbi.1005595.
[22] Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for genome assemblies. Bioinformatics 2013;29:1072–5. https://doi.org/10.1093/bioinformatics/btt086.
[23] Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinformatics 2014;30:2068–9. https://doi.org/10.1093/bioinformatics/btu153.

[24] Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, et al. Roary: rapid large-scale prokaryote pan genome analysis. Bioinformatics 2015;31:3691–3. https://doi.org/10.1093/bioinformatics/btv421.
[25] Page AJ, Taylor B, Delaney AJ, Soares J, Seemann T, Keane JA, et al. SNP-sites: rapid efficient extraction of SNPs from multi-FASTA alignments. Microbial Genomics 2016;2. https://doi.org/10.1099/mgen.0.000056.
[26] Price MN, Dehal PS, Arkin AP. FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. PLoS ONE 2010;5:e9490. https://doi.org/10.1371/journal.pone.0009490.
[27] Letunic I, Bork P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Research 2021;49:W293–6. https://doi.org/10.1093/nar/gkab301.
[28] Tufic-Garutti S dos S, De Araújo Longo LG, Fontana H, Garutti LHG, De Carvalho Girão VB, Fuga B, et al. OXA-181 carbapenemase carried on an IncX3 plasmid in high-risk *Escherichia coli* ST167 isolated from a traveler returning from Sub-Saharan Africa to Brazil. Diagnostic Microbiology and Infectious Disease 2022;102:115570. https://doi.org/10.1016/j.diagmicrobio.2021.115570.
[29] Bevan ER, Powell MJ, Toleman MA, Thomas CM, Piddock LJV, Hawkey PM. Molecular characterization of plasmids encoding bla_{CTX-M} from faecal *Escherichia coli* in travellers returning to the UK from South Asia. Journal of Hospital Infection 2021;114:134–43. https://doi.org/10.1016/j.jhin.2021.03.030.
[30] Carroll AC, Wong A. Plasmid persistence: costs, benefits, and the plasmid paradox. Can J Microbiol 2018;64:293–304. https://doi.org/10.1139/cjm-2017-0609.
[31] Clinical and Laboratory Standards Institute C. M 100 Performance Standard Susceptibility Testing. 32nd ed. 30; 2024.
[32] Naas T, Oueslati S, Bonnin RA, Dabos ML, Zavala A, Dortet L, et al. Beta-lactamase database (BLDB) – structure and function. J Enzyme Inhib Med Chem 2017;32:917–9. https://doi.org/10.1080/14756366.2017.1344235.
[33] Fuga B, Sellera FP, Cerdeira L, Esposito F, Cardoso B, Fontana H, et al. WHO critical priority *Escherichia coli* as one health challenge for a post-pandemic scenario: genomic surveillance and analysis of current trends in Brazil. Microbiol Spectr 2022;10:e01256-21. https://doi.org/10.1128/spectrum.01256-21.
[34] Giufrè M, Errico G, Monaco M, Del Grosso M, Sabbatucci M, Pantosti A, et al. Whole genome sequencing and molecular analysis of carbapenemase-producing *Escherichia coli* from intestinal carriage in elderly inpatients. Microorganisms 2022;10:1561. https://doi.org/10.3390/microorganisms10081561.
[35] Kocsis B, Gulyás D, Szabó D. Emergence and dissemination of extraintestinal pathogenic high-risk international clones of *Escherichia coli*. Life (Basel) 2022;12:2077. https://doi.org/10.3390/life12122077.
[36] Manges AR, Geum HM, Guo A, Edens TJ, Fibke CD, Pitout JDD. Global extraintestinal pathogenic *Escherichia coli* (ExPEC) lineages. Clin Microbiol Reviews 2019;32:e00135-18, /cmr/32/3/CMR.00135-18.atom. https://doi.org/10.1128/CMR.00135-18.
[37] Riley LW. Pandemic lineages of extraintestinal pathogenic *Escherichia coli*. Clinical Microbiology and Infection 2014;20:380–90. https://doi.org/10.1111/1469-0691.12646.
[38] Armand-Lefèvre L, Rondinaud E, Desvillechabrol D, Mullaert J, Clermont O, Petitjean M, et al. Dynamics of extended-spectrum beta-lactamase-producing *Enterobacterales* colonization in long-term carriers following travel abroad. Microbial Genomics 2021;7. https://doi.org/10.1099/mgen.0.000576.
[39] Kantele A, Kuenzli E, Dunn SJ, Dance DAB, Newton PN, Davong V, et al. Dynamics of intestinal multidrug-resistant bacteria colonisation contracted by visitors to a high-endemic setting: a prospective, daily, real-time sampling study. The Lancet Microbe 2021;2:e151–8. https://doi.org/10.1016/S2666-5247(20)30224-X.

[40] Fonseca EL, Morgado SM, Caldart RV, Vicente AC. Global genomic epidemiology of *Escherichia coli* (ExPEC) ST38 lineage revealed a virulome associated with human infections. Microorganisms 2022;10:2482. https://doi.org/10.3390/microorganisms10122482.
[41] Stoesser N, Crook DW, Moore CE, Phetsouvanh R, Chansamouth V, Newton PN, et al. Characteristics of CTX-M ESBL-producing *Escherichia coli* isolates from the Lao People's Democratic Republic, 2004-09. Journal of Antimicrobial Chemotherapy 2012;67:240–2. https://doi.org/10.1093/jac/dkr434.
[42] Decano AG, Pettigrew K, Sabiiti W, Sloan DJ, Neema S, Bazira J, et al. Pan-resistome characterization of uropathogenic *Escherichia coli* and *Klebsiella pneumoniae* strains circulating in Uganda and Kenya, isolated from 2017–2018. Antibiotics (Basel) 2021;10:1547. https://doi.org/10.3390/antibiotics10121547.
[43] Bubpamala J, Khuntayaporn P, Thirapanmethee K, Montakantikul P, Santanirand P, Chomnawang MT. Phenotypic and genotypic characterizations of extended-spectrum beta-lactamase-producing *Escherichia coli* in Thailand. IDR 2018;Volume 11:2151–7. https://doi.org/10.2147/IDR.S174506.
[44] Rakotonirina HC, Garin B, Randrianirina F, Richard V, Talarmin A, Arlet G. Molecular characterization of multidrug-resistant extended-spectrum β-lactamase-producing Enterobacteriaceae isolated in Antananarivo, Madagascar. BMC Microbiol 2013;13:85. https://doi.org/10.1186/1471-2180-13-85.
[45] Springer A, Mellmann A, Fichtel C, Kappeler PM. Social structure and *Escherichia coli* sharing in a group-living wild primate, Verreaux's sifaka. BMC Ecol 2016;16:6. https://doi.org/10.1186/s12898-016-0059-y.
[46] Tansawai U, Walsh TR, Niumsup PR. Extended spectrum ß-lactamase-producing *Escherichia coli* among backyard poultry farms, farmers, and environments in Thailand. Poultry Science 2019;98:2622–31. https://doi.org/10.3382/ps/pez009.
[47] Bevan ER, McNally A, Thomas CM, Piddock LJV, Hawkey PM. Acquisition and loss of CTX-M-producing and non-producing *Escherichia coli* in the fecal microbiome of travelers to South Asia. mBio 2018;9:e02408-18. https://doi.org/10.1128/mBio.02408-18.
[48] Wuethrich I, W. Pelzer B, Khodamoradi Y, Vehreschild MJGT. The role of the human gut microbiota in colonization and infection with multidrug-resistant bacteria. Gut Microbes 2021;13:1911279. https://doi.org/10.1080/19490976.2021.1911279.
[49] Larramendy S, Gaultier A, Fournier J-P, Caillon J, Moret L, Beaudeau F. Local characteristics associated with higher prevalence of ESBL-producing *Escherichia coli* in community-acquired urinary tract infections: an observational, cross-sectional study. Journal of Antimicrobial Chemotherapy 2021;76:789–95. https://doi.org/10.1093/jac/dkaa514.
[50] Arcilla MS, Hattem JMV, Bootsma MCJ, Genderen PJJV, Goorhuis A, Grobusch MP, et al. Prevalence and risk factors for carriage of ESBL-producing *Enterobacteriaceae* in a population of Dutch travellers: A cross-sectional study. Travel Medicine and Infectious Disease 2020;33:101547. https://doi.org/10.1016/j.tmaid.2019.101547.
[51] Garcia CR, Norfolk WA, Howard AK, Glatter AL, Beaudry MS, Mallis NA, et al. Long-term gut colonization with ESBL-producing *Escherichia coli* in participants without known risk factors from the southeastern United States 2024. https://doi.org/10.1101/2024.02.03.24302254.
[52] Arcilla MS, van Hattem JM, Haverkate MR, Bootsma MCJ, van Genderen PJJ, Goorhuis A, et al. Import and spread of extended-spectrum β-lactamase-producing *Enterobacteriaceae* by international travellers (COMBAT study): a prospective, multicentre cohort study. The Lancet Infectious Diseases 2017;17:78–85. https://doi.org/10.1016/S1473-3099(16)30319-X.
[53] D'Souza AW, Boolchandani M, Patel S, Galazzo G, van Hattem JM, Arcilla MS, et al. Destination shapes antibiotic resistance gene acquisitions, abundance increases, and diversity changes in Dutch travelers. Genome Med 2021;13:79. https://doi.org/10.1186/s13073-021-00893-z.
[54] Langelier C, Graves M, Kalantar K, Caldera S, Durrant R, Fisher M, et al. Microbiome and antimicrobial resistance gene dynamics in international travelers. Emerg Infect Dis 2019;25:1380–3. https://doi.org/10.3201/eid2507.181492.

[55] Ruppé E, Andremont A, Armand-Lefèvre L. Digestive tract colonization by multidrug-resistant *Enterobacteriaceae* in travellers: An update. Travel Medicine and Infectious Disease 2018;21:28–35. https://doi.org/10.1016/j.tmaid.2017.11.007.
[56] Pehrsson EC, Tsukayama P, Patel S, Mejía-Bautista M, Sosa-Soto G, Navarrete KM, et al. Interconnected microbiomes and resistomes in low-income human habitats. Nature 2016;533:212–6. https://doi.org/10.1038/nature17672.
[57] Vila J, Sáez-López E, Johnson JR, Römling U, Dobrindt U, Cantón R, et al. *Escherichia coli* : an old friend with new tidings. FEMS Microbiology Reviews 2016;40:437–63. https://doi.org/10.1093/femsre/fuw005.
[58] Gomes TAT, Elias WP, Scaletsky ICA, Guth BEC, Rodrigues JF, Piazza RMF, et al. Diarrheagenic *Escherichia coli*. Brazilian Journal of Microbiology 2016;47:3–30. https://doi.org/10.1016/j.bjm.2016.10.015.
[59] Pakbin B, Brück WM, Rossen JWA. Virulence factors of enteric pathogenic *Escherichia coli*: A review. Int J Mol Sci 2021;22:9922. https://doi.org/10.3390/ijms22189922.
[60] van Hattem JM, Cabal A, Arcilla MS, Alvarez J, De Jong MD, Melles DC, et al. Risk of acquisition of human diarrhoeagenic *Escherichia coli* virulence genes in intercontinental travellers: A prospective, multi-centre study. Travel Medicine and Infectious Disease 2019;31:101362. https://doi.org/10.1016/j.tmaid.2018.12.005.

Figures List

Figure 1: Phylogenetic tree based on the core genome of five ST38 ESBL-producing *E. coli* isolates from the present study and 20 additional genomes from isolates of the same ST deposited in Enterobase. The Figure was generated by iTOL v5.6.1.
*Isolate with *bla*$_{\text{CTX-M}}$ confirmed by PCR but absent in genome analysis.

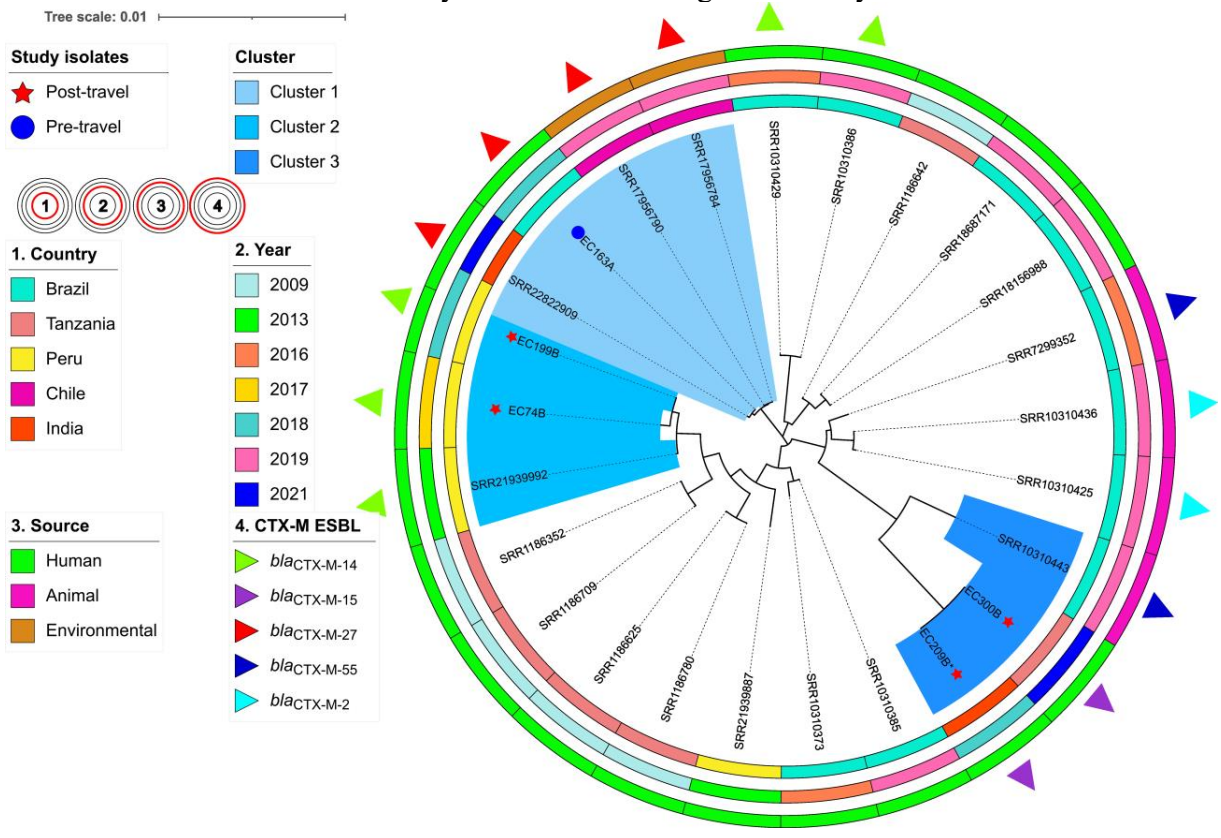

Figure 2: Phylogenetic tree based on the core genome of four ST69 ESBL-producing *E. coli* isolates from the present study and 44 genomes from the same ST deposited in Enterobase. The Figure was generated by iTOL v5.6.1.
*Isolates with *bla*$_{\text{CTX-M}}$ confirmed by PCR but absent in genome analysis.

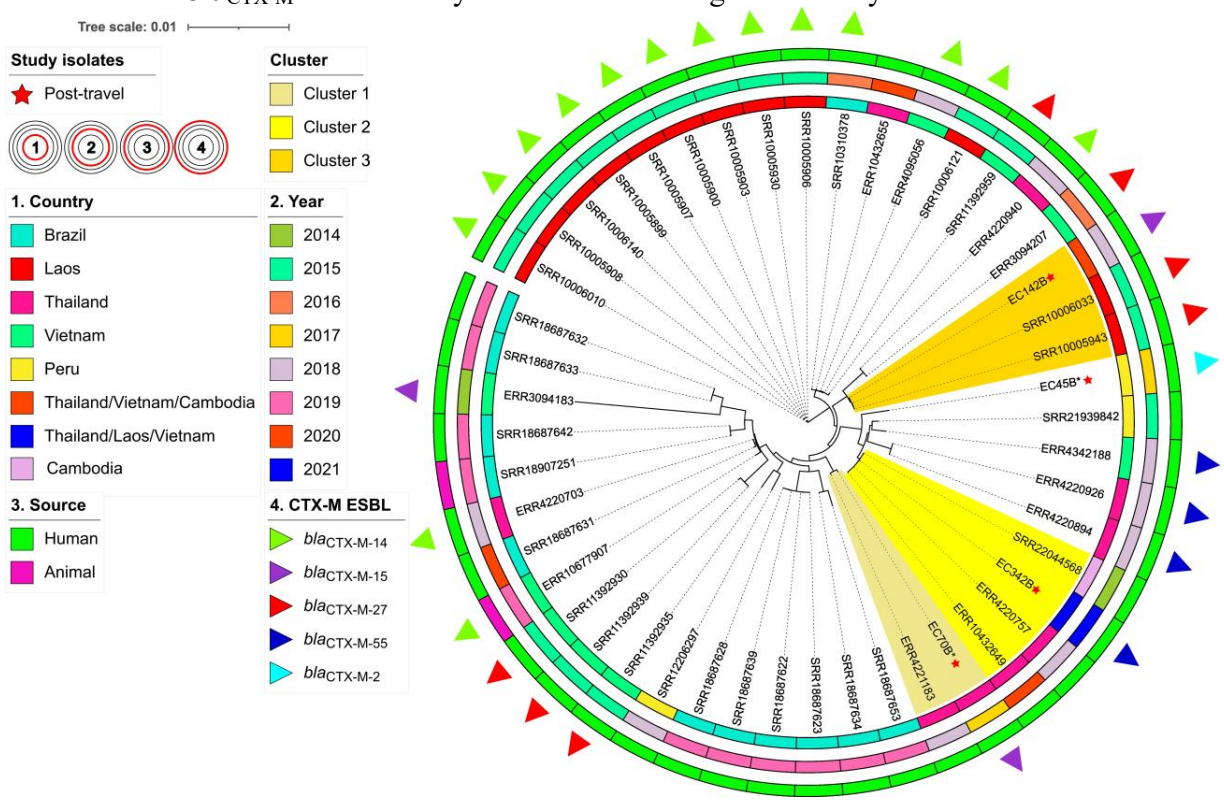

Figure 3 Clusters highlights of phylogenetic tree based on the core genome of 11 ST10 ESBL-producing *E. coli* isolates from the present study and 197 genomes from the same ST deposited in Enterobase. Clusters 1-5 are shown in squared trees (A) and (B). The circular phylogenetic tree is shown in Fig. S1. The Figure was generated by iTOL v5.6.1.

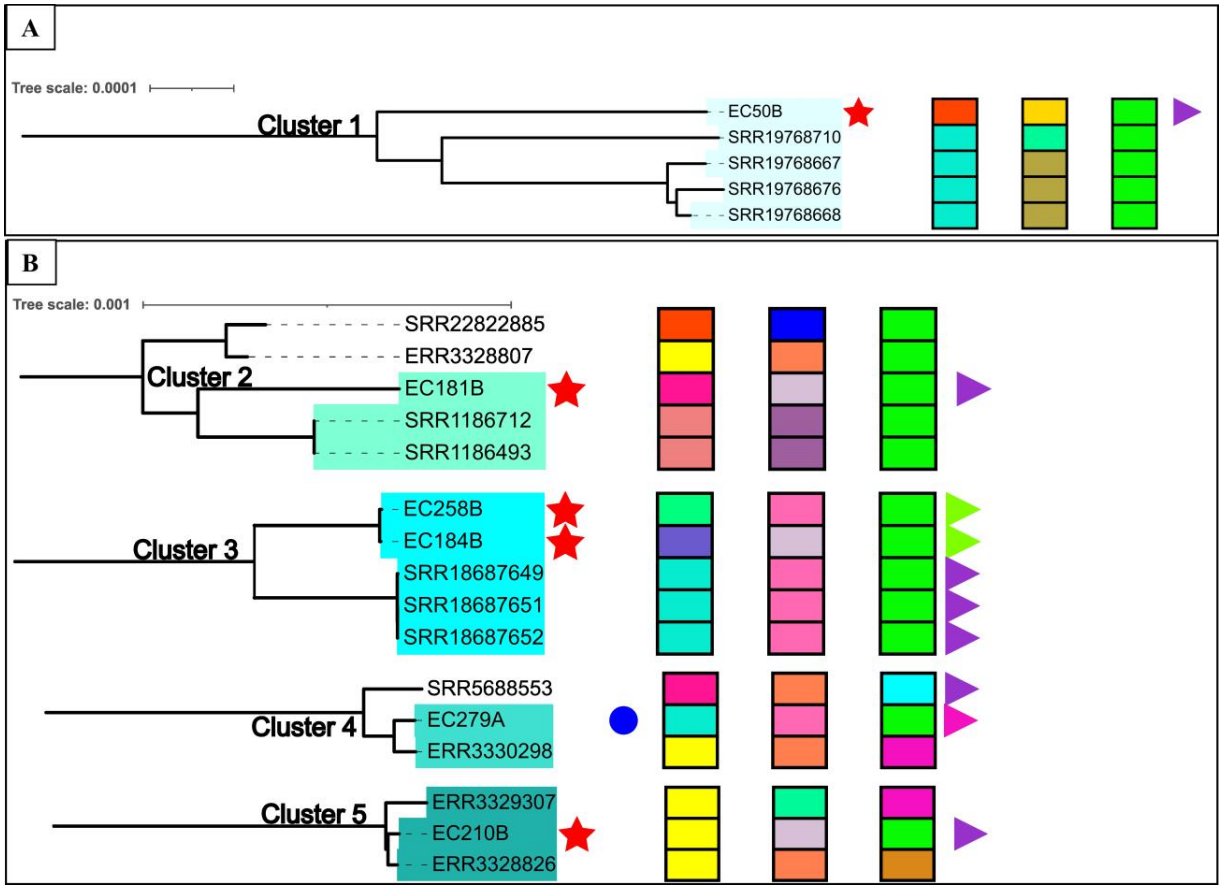

Figure 4 Clusters highlights of phylogenetic tree based on the core genome of nine ST131 ESBL-producing *E. coli* isolates from the present study and 329 genomes from the same ST deposited in Enterobase. Clusters 1-6 are shown in squared trees (B), (C) and (D). The circular phylogenetic tree is shown in Fig. S2. The Figure was generated by iTOL v5.6.1.

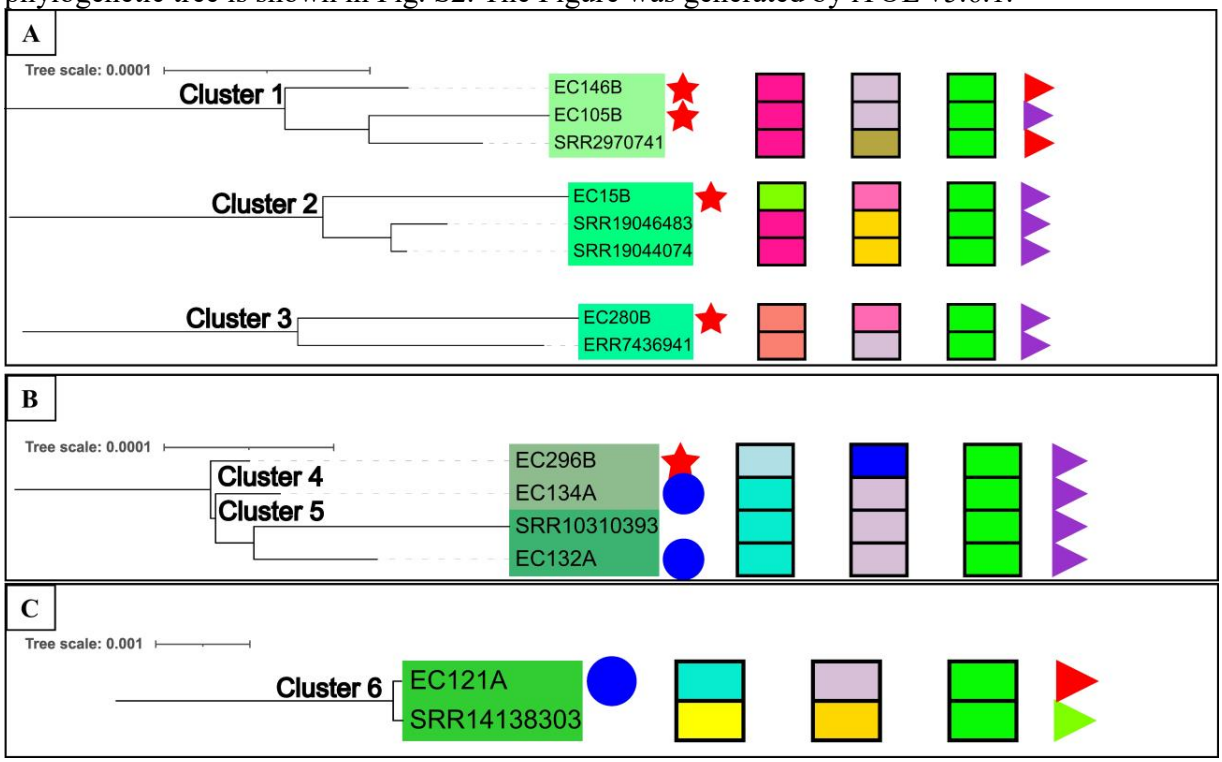

Figure 5: Number of (A) AMR genes, (B) plasmids, and (C) virulence genes in genomes of ESBL-producing *E. coli* isolates obtained pre- (n=13) and post-travel (n=43). We excluded 14 genomes of isolates with $bla_{\text{CTX-M}}$ confirmed by PCR but absent in genome analysis.