Faecal carriage of ESBL-producing *Escherichia coli* in a remote region of Niger

Hervé Jacquier[1,2], Bachir Assao[3], Françoise Chau[1], Ousmane Guindo[3], Bénédicte Condamine[1], Mélanie Magnan[1], Antoine Bridier-Nahmias[1], Nathan Sayingoza-Makombe[3], Aissatou Moumouni[3], Anne-Laure Page[4], Céline Langendorf[4], Matthew E Coldiron[4], Erick Denamur[1,5,*], Victoire de Lastours[1,6,*]

1. Université Paris Cité, IAME UMR 1137, INSERM, 75018 Paris

2. Assistance Publique – Hôpitaux de Paris, Laboratoire de Bactériologie-Hygiène, Hôpitaux Universitaires Saint-Louis – Lariboisière – Fernand Widal, 75010 Paris

3. Epicentre, Médecins Sans Frontières, Maradi, Niger

4. Epicentre, Médecins Sans Frontières, 75019 Paris, France

5. Assistance Publique – Hôpitaux de Paris, Laboratoire de Génétique Moléculaire, Hôpital Universitaire Bichat, 75018 Paris

6. Assistance Publique – Hôpitaux de Paris, Service de Médecine Interne, Hôpital Universitaire Beaujon, 92110 Clichy

* These authors contributed equally to this work.

Corresponding authors: Hervé Jacquier: herve.jacquier@aphp.fr; Victoire de Lastours: victoire.de-lastours@aphp.fr

Running title: Community faecal carriage of ESBL–*E. coli* in Niger

Keywords: ESBL – *E. coli*, community, developing country, faecal carriage, chromosomal integration

Abstract

Background

Data on extended-spectrum β--lactamase-producing *Escherichia coli* (ESBL-*E. coli*) carriage in community settings, especially in developing countries, are scarce. Here, we describe the population structure and molecular characteristics of resistance of ESBL-*E. coli* faecal isolates in a rural African population.

Methods and findings

Between April and May 2017, stools of 383 healthy participants were collected from 20 villages in rural Southern Niger during a clinical trial on ciprofloxacin prophylaxis carried out during a meningococcal meningitis outbreak. Of 383 individuals, 354 (92.4%) were carriers of ESBL-*E. coli* before any ciprofloxacin intake. A subset of 90 of these ESBL-*E. coli* containing stools were selected for further analysis, from which 109 different ESBL-*E. coli* were recovered and whole genome sequenced by short-(Illumina) and long-(Nanopore) reads. Most belonged to the commensal-adapted phylogroup A (91, 83.5%), with high clonal diversity (57 distinct clones). One-quarter harboured the high pathogenicity island previously associated with a longer duration of faecal carriage. The $bla_{\text{CTX-M-15}}$ gene was the major ESBL determinant (107, 98.1%). It was chromosome-integrated in approximately half of the cases (48, 44.9%), at multiple integration sites in diverse chromosomal genetic backgrounds. When plasmid-borne, $bla_{\text{CTX-M-15}}$ was found in a large diversity of incompatibility groups. A single genetic background was found for 20 distinct plasmids, whereas very closely related plasmids were found in different genetic backgrounds in six cases, suggesting plasmid spread among strains. No geographical or social links to resistance patterns were observed.

Conclusions

Massive prevalence of community faecal carriage of CTX-M-15-producing *E. coli* was observed in a rural region of Niger without apparent antibiotic selective pressure. *E. coli* were highly diverse, well adapted commensal strains, with chromosomal integration of CTX-M-15 encoding gene in almost half of the cases. Evidence of clonal and plasmid spread suggest a risk of sustainable implementation in community faecal carriage.

Author Summary

Community faecal carriage of extended-spectrum β--lactamase-producing *Escherichia coli* (ESBL-*E. coli*) increased alarmingly over the past two decades. Nevertheless, data on ESBL-*E. coli* carriage in community settings are scarce, especially in developing countries, yet fundamental to be able to implement specific strategies to prevent their spread and determine adequate empiric antibiotic therapies. We conducted an ancillary study focusing on ESBL-*E. coli* carriage during a clinical trial on meningitis prophylaxis in rural Southern Niger. Almost all subjects were ESBL-*E. coli* carriers despite the apparent absence of recent antibiotic use. Using whole genome sequencing, we found a wide diversity of *E. coli* isolates well adapted to commensal lifestyle and bearing the $bla_{\text{CTX-M-15}}$ gene located either on the chromosome or on a variety of plasmids, some spreading between clones. ESBL-*E. coli* isolates are endemic and well implanted in this community setting. Identification of the selective pressures involved should be a priority.

Introduction

Extended-Spectrum β-Lactamase-producing (ESBL) Enterobacterales, and particularly *Escherichia coli* (ESBL-*E. coli*) have spread worldwide and become endemic, especially in low-and middle-income countries [1–5]. The diffusion of ESBL-*E. coli* is mainly the result of the epidemiological success of $bla_{CTX\text{-}M}$ genes in the 2000s, which are ESBL-encoding genes often carried by mobile genetic elements such as plasmids. *E. coli* are commensals of the gut of vertebrates, including humans [6], and constitute the main reservoir of resistance genes in Enterobacterales, leading to ESBL-*E. coli* dissemination in communities [7]. *E. coli* are also opportunistic pathogens responsible for severe infections such as bloodstream infections, urinary tract infections and in some cases diarrhea [8]. Infections with ESBL-*E. coli*present a clinical challenge due to limited therapeutic options, which is especially crucial in low-income countries [9]. Globally, ESBL-*E. coli* infections represent the biggest burden of bacterial antimicrobial resistance (AMR) morbidity and mortality [10].

Prevalence of ESBL-*E. coli* faecal carriage in humans is a good indicator of the global burden of antibacterial resistance in a given community [11]. In Africa, the few available studies showed a prevalence of ESBL-*E. coli* community faecal carriage ranging from 15% to 92% depending on the country, the population studied and the year of stool collection, which is higher than in most high income countries [2–5,12–14]. However, little is known about the molecular epidemiology and evolution of resistance mechanisms in low-income countries [13,15]. Yet, this level of understanding is essential to unravel the evolutionary mechanisms as well as the transmission mechanisms enabling the success of ESBL-*E. coli* in community settings. It is crucial to determine specific strategies to prevent further dissemination of these resistant bacteria. We previously reported that 92.4% of healthy individuals carried ESBL-*E. coli* carried in their stools as part of a trial of ciprofloxacin prophylaxis for meningococcal meningitis performed in a rural region of South Niger [14]. Faced with such a high prevalence, we sought to describe the population structure of the strains and the molecular basis of this resistance to understand the mechanisms of resistance spread.

Patients and Methods

Study design

This work was an ancillary study of a clinical trial on ciprofloxacin prophylaxis against bacterial meningitis carried out between April 22, 2017 and May 18, 2017 during a meningococcal meningitis outbreak in the Madarounfa District in rural southern Niger near the border with Nigeria [14].

Population characteristics

Briefly, a subset of 10 villages in the control-arm and 10 villages randomized to receive village-wide prophylaxis in the primary trial were randomly selected to participate in a sub-study examining faecal

carriage of ciprofloxacin-resistant and ESBL-producing *Enterobacterales*, at baseline and after interventions [14]. In each village, 20 households were randomly selected from a census list. One household member in each of the selected households was then selected using a simple lottery. Villages were assigned an identification number and GPS coordinates were collected. Villages were further described by the presence/absence of a market or a health center, as well as the distance to the nearest paved road.

Strain collection

In the primary trial, stool samples of 383 individuals were collected at day 0 (prior to ciprofloxacin administration in the intervention arm). Samples were collected at home by participants and inoculated into Cary-Blair media (Copan, Brescia, Italy) by study staff. Of them, 99 faecal samples were randomly selected and were sent to the Infection Antimicrobials Modelling and Evolution laboratory (IAME, Université de Paris) for quality control in agar deep tubes [14]. In Paris, faeces were plated on cefotaxime containing plates, and colonies of different morphologies were identified by mass spectrometry (MALDI-TOF-MS Microflex Bruker Daltonics/ BioTyper™ version 2.0). When colonies with different morphologies on the same plate were identified as *E. coli*, each isolate was studied independently. Antimicrobial susceptibility testing was performed by the disk diffusion method according to the EUCAST guidelines, and ESBL detection using the synergy test. Among the 99 samples, 90 samples from 90 subjects carrying at least one confirmed ESBL-*E. coli*, constitute the material for the analyses performed in this work.

Genomic analysis

Detailed information of genomic analysis is available in the Supplementary Material. In brief, after DNA extraction and library preparation, Illumina sequences were analysed with a previously described in-house bioinformatic pipeline [16], integrating genome assemblies, annotation and ordering, genetic distance estimation, genome typing (phylogroups, multi-locus sequence typing(MLST), serotypes and *fim*H types), plasmid replicons, contig location (plasmid or chromosome), resistome and virulome. The main virulence factors associated with extra-intestinal or intra-intestinal pathogenic *E. coli* (respectively ExPEC and InPEC) were determined, as previously described [17].

The core genome constructed with Roary v3.12 [18] was used for SNP distance matrix computation (see Supplementary Table 1) and to develop a maximum likelihood-based phylogeny using Iqtree v1.6.12 [19]. The phylogenetic tree was designed using Itol (v6.3)[20]. Identification of clonal lineages was performed using popPUNK [21]. Clones were defined for isolates with fewer than 120 SNP differences. This threshold was based on the distribution of SNPs among isolates (Supplementary Figure1) and followingprevious work using the same strategy [22]. For strains with ESBL-encoded

genes embedded in plasmids, one representative isolate was selected for Nanopore sequencing (Oxford, UK) for isolates with close genetic distances. After filtering, assembling, and polishing, Nanopore sequences were assembled with Illumina sequences. Sequences are available in the bioproject PRJEB50347 in the European Nucleotide Archive (ENA) database.

Ethics

The parent study protocol, which included this sub-study, was reviewed and approved by the National Consultative Ethics Committee of Niger (Ref: 003/2016/CCNE) and the Ethics Review Board of Médecins Sans Frontières (Ref: 1603)[14]. The trial is registered at ClinicalTrials.gov (NCT02724046). Written informed consent was obtained from individual participants.

Results

All raw data concerning antibiotic resistance and genome typing are available in Supplementary Table 2.

Population characteristics

In this ancillary study, the median age of the 90 participants was 10.5 years [IQR 4-24.8] , and the F:M sex-ratio was 1.5. Only 3 subjects (3.3%) reported antibiotic exposure or hospitalization in the three months prior to sample collection.

ESBL-*E. coli* isolation and resistance to antibiotics

A total of distinct 109 ESBL-*E. coli* were recovered from the 90 samples. Seventy-two participants carried only one ESBL-*E. coli,* 17 carried two distinct ESBL-*E. coli* and one participant carried 3 distinct ESBL-*E. coli* by colony morphology (see above).

Among the 109 ESBL-*E. coli* isolates, 58 (53.2%) were resistant to nalidixic acid, 56 (51.4%) to ciprofloxacin, 23 (21.1%) to gentamicin, 76 (69.7%) to cotrimoxazole, and 85 (78.0%) to tetracycline. All were susceptible to amikacin and imipenem.

Genomic diversity of ESBL-*E. coli*

The genomic diversity of ESBL-*E. coli* is shown in Table 1 and Figure 1. The ESBL-*E. coli* mainly belonged to phylogroup A (91/109, 83.5%), phylogroup B1 (8/109, 7.3%) and phylogroup C (6/109, 5.5%). One single isolate was found in each of phylogroups D, F, B2 and clade I.

The 109 ESBL- *E. coli* isolates were clustered in 48 clonal lineages, 82 (75.2%) belonging to clonal lineages with two or more isolates. The main clonal lineages were clonal lineage 1 (14/109, 12.8%) and clonal lineage 2 (10/109, 9.2%). ST typing using Warwick scheme highlighted 44 different STs.

The principal STs identified were ST46 (15/109, 13.8%), ST10 (10/109, 9.2%), ST6359 (7/109, 6.4%), ST617 (6/109, 5.5%), and ST410 (6/109, 5.5%). *In silico* serotyping approaches identified 42 different serotypes, including 11 serotypes with unknown O antigens and new O antigens waiting for assignment for 24/109 (22%) isolates. The main serotypes were O101:H9 (11/109, 10.1%), O101:H4 (10/109, 9.1%), O101:H10 (9/109, 8.3%), and O9:H9 (8/109, 7.3%). A total of 21 different *fimH* alleles were identified. For 34/109 (31.2%) isolates, no *fimH* gene was found. The main *fimH* alleles identified belonged to *fimH54* (15/109, 13.8%), *fimH34* (11/109, 10.1%), *fimH24* (7/109, 6.4%), and *fimH41* (6/109, 5.5%). Based on SNP differences in the core genome, 57 distinct clones were identified among the 109 isolates. These clones are in agreement with the previous typing as they belong to the same phylogroup, clonal lineage, ST, serotype and exhibit the same *fimH* allele. Of 109 isolates, 74 (67.9%) belonged to clones represented by two or more isolates. Isolates mainly belonged to clones 36 (10/109, 9.2%), 37 (5/109, 4.6%) and 26 (5/109, 4.6%). Of note, *E. coli* strains with different morphologies from the same individual were genetically distinct (Table1, Figure 1).

In summary, despite belonging almost uniquely to phylogroup A, a huge genomic diversity of clones was observed among the ESBL-*E. coli* isolates.

ESBL-encoding genes and genomic location

The ESBL phenotype was conferred by the $bla_{\text{CTX-M-15}}$ gene for 107/109 (98.1%) isolates. The $bla_{\text{CTX-M-15}}$ gene was located on the chromosome for 48/107 (44.9%) isolates and on a plasmid for 59/107 (55.1%) isolates. The remaining two isolates had the ESBL phenotype conferred by $bla_{\text{SHV-2a}}$ (1 isolate) and $bla_{\text{CTX-M-55}}$ genes (1 isolate), and were both located on a plasmid.

Isolates with a chromosomal $bla_{\text{CTXM-15}}$ gene were phylogenetically diverse, belonging to 18 clonal lineages and 23 clones. We identified 19 different chromosome integration sites (Figure 1 and Supplementary Table 3). For most isolates, a given chromosome integration site was associated with a given clonal lineage (Figure 1). For isolates belonging to clonal lineage 1, we assume that $bla_{\text{CTX-M-15}}$ gene acquisition would have followed the most parsimonious scenario: isolates with chromosomal $bla_{\text{CTX-M-15}}$ gene could have diversified after a single event of integration. This assumption is based on the fact that the insertion site was the same for all isolates with chromosomal $bla_{\text{CTX-M-15}}$ and that IS*EcpI* upstream $bla_{\text{CTX-M-15}}$ is partially deleted. For a single clonal lineage (clonal lineage 2), we observed two different integration sites (Supplementary Table 3). Altogether, these data indicate multiple integration events in diverse genetic backgrounds. For isolates with ESBL-encoding gene located on a plasmid, 10 different replicon types were identified. The plasmids with $bla_{\text{SHV-2a}}$ and $bla_{\text{CTX-M-55}}$ genes belonged to IncN and IncX1 incompatibility groups, respectively. The plasmids harbouring $bla_{\text{CTX-M-15}}$ belonged mainly to IncF (IncFIA (5/107), IncFIB (16/107), IncFIB (H89 phagemid

plasmid) (6/107), IncFIC(FII) (8/107)) and IncY (16/107) and, to a lesser extent, by IncI1 (6/107) incompatibility groups (Table 1).

To gain further insight into the relationships between clones and plasmids, we investigated the strain genomic backgrounds for plasmids with less than 1% nucleotide identity difference or ≤2 large insertions or deletions (indels). In six cases, we found closely related plasmids in different genetic backgrounds, whereas 20 plasmids were strictly associated with a single genetic background. For incFIA plasmids, we identified similar plasmids, with insertion of 26.7 kbp in the plasmid hosted by clone 8 compared to that of clone 9, suggesting that the plasmid could have drifted from the ancestor of these two clones. For IncFIB plasmids, we identified similar plasmids in clones 28 and 50, with insertion of 26.7 kbp in the plasmid hosted by clone 28, and an insertion of 8.8 kbp in the plasmid hosted by clone 50. Moreover, the plasmid hosted by clone 29 presented a 27.9 kbp insertion compared to that of clone 45. For IncI1 plasmids, plasmids hosted by clones 16 and 25 presented more than 99% identity, and the ones hosted by the clones 17, 30, and 56 presented 2 indels of 2.14 and 1.34 kbp. For IncY plasmids, we identified similar plasmids, with insertion of 17 kbp in the plasmid hosted by clone 32 compared to that of clone 37. Altogether, these data argue for plasmid transmission between clones, except for clones 8 and 9, for which the plasmid had been hosted in an ancestor isolate, before diversification. Of note, clone 36 encompasses isolates with different evolutionary scenario: $bla_{\text{CTX-M-15}}$ is hosted in two different chromosome integration sites, J (six isolates) and K (two isolates) and in an IncFIC_FII plasmid (two isolates).

Antimicrobial susceptibility and ESBL genomic location

Co-resistance to cotrimoxazole and tetracyclines was significantly lower for isolates with ESBL-encoding genes located on the chromosome (n=48) than for isolates with ESBL-encoding genes on the plasmid (n=61) (respectively, 43.8% vs 90.2 %, $P<0.05$ for cotrimoxazole and 68.8% vs 85.5%, $P<0.05$ for tetracyclines).

Virulence genes

We focused on the main virulence factors associated with extra-intestinal or intra-intestinal pathogenic *E. coli* (respectively ExPEC and InPEC)(19). The average number of ExPEC virulence genes was low (1.24), the most prevalent being *fuyA* and *irp2* (25/109, 22.9%) (Supplementary Table 2). These two genes encode the yersiniabactin siderophore and its transporter, and are known to be part of the high pathogenicity island (HPI)[23]. These ExPEC genes were unevenly distributed among the strains, as mostly found in the clonal lineage 1, 12/14 harbouring *fuyA* and *irp2* genes, and in the phylogroup C, most isolates (4/6) harbouring the adhesin encoding *pap* genes in addition to *fuyA* and *irp2* genes. One enterotoxinogenic *E. coli* (ETEC) (*eltB* positive) from clade I and two

enteropathogenic *E. coli* (EPEC) from B1 and B2 phylogroups were identified, the latter belonging to the ST2346 and being a typical EPEC (*eae* and *bfp* positive)[24].

Geographical origin of ESBL-*E. coli* carriers

The geographical distribution of the villages of the carriers of the 109 ESBL-*E. coli* according to their clonal lineage is represented in Figure 2. Most isolates belonging to the same clonal lineage were not limited to a given geographical area, but were widely spread across the different villages. Moreover, the spread of different clones was not associated with the presence of a market or a health center, nor the distance to a paved road. We found isolates presenting different genetic backgrounds harbouring the same plasmid in only one village(one isolate belonging to clonal lineage 40 and two isolates belonging to clonal lineage 8 (Figure 1).

Discussion

After finding an unexpectedly high prevalence of ESBL-*E. coli* faecal carriage in healthy subjects (>90%) in rural South Niger [14], we analysed the population structure, molecular support of resistance and mode of spread of these commensal strains in this community, using state-of-the-art sequencing techniques with both Ilumina WGS and an additional long-read sequencing by Nanopore (to specifically study strains with ESBL-encoded genes embedded in plasmids). Previous description of ESBL carriage in African settings has varied by the region, the community and the age of the subjects [2–5,12,13], ranging from 46 % in district hospital patients in South Africa in 2018 [3], 51% in hospitalized patients and 38% in healthy healthcare workers in Chad in 2019 [13], and 59% in healthy children in the community in the Central African Republic [12]. When resistance mechanisms and phylogenetic groups were described, $bla_{\text{CTX-M-15}}$ was the most commonly encountered gene encoding for resistance to β-lactams (ranging from 75% to 90% of cases), and the clonal group ST131 was predominant among phylogroup B2 *E. coli* [12,13,15,25]. To our knowledge, this is the first description of WGS and full molecular description of the ESBL-*E. coli* isolates in a low-income country setting, with original and important findings.

First, most ESBL-*E. coli* belonged to phylogroup A (83.4%), with high diversity in terms of clonal lineages and molecular typing (serotype, ST, and *fimH* typing), even though some clonal lineages were overrepresented. *E. coli* strains belonging to phylogroup A are known to be adapted to human gut commensalism [6] and have been described as predominant in low-and middle-income countries [26]. As expected, these strains carry few virulence genes, with the exception of clonal lineage 1 (ST617/6359, O101:H9/10) strains, which have accumulated many virulence genes, especially iron capture systems such as yersiniabactin (*fyuA, irp2*), borne on the HPI (Supplementary Table 2). ST410 (phylogroup C), an emerging high-risk clone [27], is also well represented, with strains harbouring the *pap* adhesin encoding genes in addition to the HPI. Both gene clusters have been associated with ESBL-*E. coli* strains with a longer duration of faecal carriage in travellers returning from developing countries [28]. On the other hand, B2 phylogroup was represented by a single isolate, a typical EPEC belonging to the EPEC9 cluster [24,29,30]. This is in contrast with what has been observed in Western countries, where the proportion of B2 strains among commensal *E. coli* has dramatically increased in the last four decades, probably under the pressure of several factors such as changes in dietary habits and improved hygiene [6,31]. Of note, no ST131-*E. coli* isolate, (also a B2) was found here, despite the clone's global epidemiological success [12,13,32]. The pronounced diversity among A strains and the low prevalence of B2 strains supports the hypothesis that the strains described here are almost individual-specific, and aligns with the expected composition of commensal *E. coli* strains from humans living in developing countries. Along the same line, a recent study performed in healthy

pregnant women in Madagascar showed a high prevalence of ESBL-*E. coli* (34%), composed mostly of isolates belonging to various phylogroup A clones (72%)[33].

Second, we found that the ESBL phenotype was mainly associated with the $bla_{CTX\text{-}M\text{-}15}$ gene (98%), which was embedded in the chromosome in approximately 50% of cases, at different integration sites. The spread of the $bla_{CTX\text{-}M\text{-}15}$ gene has been well-described in the last decades, even though other $bla_{CTX\text{-}M}$ alleles such as $bla_{CTX\text{-}M\text{-}14}$ or $bla_{CTX\text{-}M\text{-}27}$ have had some epidemiological success [2]. As with most AMR genes, the spread of $bla_{CTX\text{-}M}$ genes has been associated with plasmid dissemination, particularly the pC15-1a plasmid [34]. However, chromosomal locations have increasingly been described. Chromosome-located $bla_{CTX\text{-}M}$ genes were found in 4.8% of ESBL-*E. coli* isolates in a veterinary collection from Germany, the Netherlands, and the UK [35]. This rate increases to 8.3% and 14.6%, respectively, in clinical isolates recovered from hospitals in Zambia [36] and China [37]. In the study from Madagascar cited above [33], the $bla_{CTX\text{-}M}$ gene was located on the chromosome in 31% of strains. The highest incidence of chromosomal $bla_{CTX\text{-}M}$ gene described so far was observed in Japan, where $bla_{CTX\text{-}M\text{-}14}$ was found on the chromosome in 45.9% of cases, and in 16.2% both on chromosome and plasmid in community commensal strains [38]. The location of $bla_{CTX\text{-}M}$ genes on both chromosome and plasmid probably reflects the ability of their genetic environment to mobilize these resistance genes. In their model, Hamamoto *et al.* display high rates of $bla_{CTX\text{-}M\text{-}14}$ mobilization, and integration in various sites [39], as was observed here with our isolates. Whether a chromosomal or plasmid location confers a selective advantage for the resistance genes is a matter of debate. In a mathematical model, Lethtinen *et al.* suggested that plasmid location could be preferred under selective pressure [40], suggesting that a chromosomal location is more adapted to a lower levels of selective pressure, as would be the case in community carriage. This hypothesis is also supported by the fitness cost associated with plasmid-carried resistance genes [41] and by the conservation of vertical transfer of chromosome-located genes. These assumptions remain challenged by other authors, as compensatory mutations alleviating plasmid fitness costs and high plasmid-chromosome beneficial interactions have also been described [41]. From an epidemiological point-of-view, a recent work suggested that chromosome integration of $bla_{CTX\text{-}M\text{-}15}$ in ST131 isolates could have contributed to the clonal expansion of a rare C2 lineage [42].

Third, for isolates with plasmid-located $bla_{CTX\text{-}M\text{-}15}$, we showed high diversity in incompatibility groups, but with a high predominance of IncF plasmids, known to be well adapted to *Enterobacterales.* Interestingly, a high similarity was found for some plasmids in different genetic backgrounds, suggesting both isolate and plasmid diffusion. This observation here in a human community is in line with a previous study of temporal intestinal carriage of ESBL-*E. coli* in calf farms [22]. Additionally, we

found a similarity between IncFIb plasmid backbones and pCFSAN061762, a plasmid described in Egypt in an isolate of *E. coli* ST38 (data not shown).

Finally, we did not find any correlation between geographical localization of the subjects (the villages they live in or their usual social interactions, such as marketplaces) and the epidemiological genomic data of ESBL-*E. coli*, which suggests a large spread of the most frequent clonal lineages. These results may unveil an underestimated mix of populations in this area, or contamination by a common source of food or water that could not be captured by mapping health centers, distance to an asphalt road or marketplaces, as represented in Figure 2.

Our work has some limitations: (i) only a subset of stool samples from the overall study was analysed, potentially introducing biases, yet the random selection should have reduced this risk; (ii) only the predominant cefotaxime-resistant *E. coli* isolates present in the faeces were studied, which may not represent the diversity of *E. coli* isolates in the gut, but does represent the quantitatively predominant *E. coli* isolates and therefore the ones that constitute the biggest threat in terms of infectious risks and dissemination [43]; (iii) no precise description of the mobile genetic elements that flank the $bla_{\text{CTX-M-15}}$ gene is available, which would improve understanding of the modalities of mobilization of this gene; this will will be the subject of future works. Finally, the search of an ESBL-*E. coli* reservoir in other environments, such as the food chain (poultry/livestock), water or sediments would help understand the routes of dissemination in a “One Health” approach [11].

Despite these limitations, this epidemiologic study in a remote region of Niger is unique and shows the unprecedented dissemination of ESBL-*E. coli* in the gut of a geographically localized human population. It suggests the widespread dissemination of multiple resistant clones, possibly fostered by underestimated inadequate antibiotic treatments and multiple population exchanges that have probably lasted for years already. On a molecular point-of-view, this work highlights the diversity of pathways followed by $bla_{\text{CTX-M-15}}$. The high frequency of chromosomal $bla_{\text{CTX-M-15}}$ in isolates of *E. coli* belonging to the commensal phylogroup A is of particular concern because it could be associated with their sustainable implementation in the community. Understanding the evolutionary forces that select for such a high prevalence of ESBL-*E. coli* should be major objective.

Altogether, antibiotic resistance in commensal *E. coli* in this remote part of Africa appears to have all the characteristics of long-term adaptation and all the tools for further dissemination. This is especially worrying as commensal *E. coli* are responsible for clinical infections and antimicrobial resistant *E. coli* infections represent the biggest threat in terms of AMR globally [10]. Noteworthy, community-acquired bacteremia in children hospitalized for severe acute malnutrition was recently followed in an intensive nutritional rehabilitation center in the same region (Madarounfa district,

Niger). Approximately two-thirds of the *E. coli* and *K. pneumoniae* strains (16/26, 61.5%) were ESBL producers, reinforcing the link between the widespread carriage of ESBL-producing *E. coli* strains and the risk of severe infection from these same microorganisms [44]. Infections with these strains will be increasingly difficult to treat especially in these remote areas, with the need for broader spectrum antibiotics, which in turn may select more resistant strains such as carbapenemase producers. This work may just be exposing the tip of the iceberg of an AMR epidemiological catastrophe well on its way.

Supporting Information

S1 Fig. Distribution of core genome SNP distances

(DOCX)

Material S1. Supplementary material of genomic analysis

(DOCX)

S1 Table. SNP distance matrix

(XLSX)

S2 Table. Raw data

(XLSX)

S3 Table. Chromosome integration sites

(DOCX)

Acknowledgments

We offer sincere thanks to all the participants of the study, as well as the staff from Maradi laboratory for their excellent fieldwork, both in terms of sample collection and phenotypic microbiological analyses. We thank Médecins Sans Frontières teams for their trust and involvement in the parent study as well as this ancillary study. We thank Epicentre and Rebecca Grais, for trusting our laboratory for the quality control and genetic analyses from the beginning of this work and for the opportunity to gather an impressive collection of faecal samples from such a remote region.

Financial support

This work was partially supported by a grant from the “Fondation pour la Recherche Médicale” (Equipe FRM 2016, grant number DEQ20161136698). The parent study was funded by Médecins Sans Frontières.

Potential conflicts of interest

None

Author Contributions

Conceptualization: Matthew E. Coldiron, Anne-Laure Page, Céline Langendorf, Hervé Jacquier, Erick Denamur, Victoire de Lastours

Formal analysis: Hervé Jacquier, Françoise Chau, Bénédicte Condamine, Mélanie Magnan, Antoine Bridier-Nahmias, Anne-Laure Page, Céline Langendorf, Matthew E Coldiron, Erick Denamur, Victoire de Lastours

Investigation: Matthew E. Coldiron, Bachir Assao, Ousmane Guindo, Nathan Sayingoza-Makombe, Aissatou Moumouni, Anne-Laure Page, Céline Langendorf

Methodology: Erick Denamur, Hervé Jacquier, Bénédicte Condamine, Antoine Bridier-Nahmias

Software : Bénédicte Condamine, Antoine Bridier-Nahmias

Project administration: Matthew E. Coldiron, Victoire de Lastours, Erick Denamur

Supervision : Matthew E. Coldiron, Victoire de Lastours, Erick Denamur

Funding Acquisition : Erick Denamur

Writing ± original draft: Hervé Jacquier, Erick Denamur, Victoire de Lastours

Writing ± review & editing: Matthew E. Coldiron, Anne-Laure Page, Céline Langendorf

References

1. Karanika S, Karantanos T, Arvanitis M, Grigoras C, Mylonakis E. Fecal Colonization With Extended-spectrum Beta-lactamase–Producing Enterobacteriaceae and Risk Factors Among Healthy Individuals: A Systematic Review and Metaanalysis. Clin Infect Dis. 2016;63: 310–318. doi:10.1093/cid/ciw283

2. Bevan ER, Jones AM, Hawkey PM. Global epidemiology of CTX-M β-lactamases: temporal and geographical shifts in genotype. J Antimicrob Chemother. 2017;72: 2145–2155. doi:10.1093/jac/dkx146

3. Founou RC, Founou LL, Essack SY. Extended spectrum beta-lactamase mediated resistance in carriage and clinical gram-negative ESKAPE bacteria: a comparative study between a district and tertiary hospital in South Africa. Antimicrob Resist Infect Control. 2018;7: 134. doi:10.1186/s13756-018-0423-0

4. Letara N, Ngocho JS, Karami N, Msuya SE, Nyombi B, Kassam NA, et al. Prevalence and patient related factors associated with Extended-Spectrum Beta-Lactamase producing *Escherichia coli* and *Klebsiella pneumoniae* carriage and infection among pediatric patients in Tanzania. Sci Rep. 2021;11: 22759. doi:10.1038/s41598-021-02186-2

5. Diriba K, Awulachew E, Tekele L, Ashuro Z. Fecal Carriage Rate of Extended-Spectrum Beta-Lactamase-Producing *Escherichia coli* and *Klebsiella pneumoniae* Among Apparently Health Food Handlers in Dilla University Student Cafeteria. Infect Drug Resist. 2020;13: 3791–3800. doi:10.2147/IDR.S269425

6. Tenaillon O, Skurnik D, Picard B, Denamur E. The population genetics of commensal *Escherichia coli*. Nat Rev Microbiol. 2010;8: 207–217. doi:10.1038/nrmicro2298

7. Woerther P-L, Burdet C, Chachaty E, Andremont A. Trends in human fecal carriage of extended-spectrum β-lactamases in the community: toward the globalization of CTX-M. Clin Microbiol Rev. 2013;26: 744–758. doi:10.1128/CMR.00023-13

8. Denamur E, Clermont O, Bonacorsi S, Gordon D. The population genetics of pathogenic *Escherichia coli*. Nat Rev Microbiol. 2021;19: 37–54. doi:10.1038/s41579-020-0416-x

9. Ouedraogo AS, Jean Pierre H, Bañuls AL, Ouédraogo R, Godreuil S. Emergence and spread of antibiotic resistance in West Africa : contributing factors and threat assessment. Med Sante Trop. 2017;27: 147–154. doi:10.1684/mst.2017.0678

10. Murray CJ, Ikuta KS, Sharara F, Swetschinski L, Aguilar GR, Gray A, et al. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. The Lancet. 2022;399: 629–655. doi:10.1016/S0140-6736(21)02724-0

11. Puspandari N, Sunarno S, Febrianti T, Febriyana D, Saraswati RD, Rooslamiati I, et al. Extended spectrum beta-lactamase-producing *Escherichia coli* surveillance in the human, food chain, and environment sectors: Tricycle project (pilot) in Indonesia. One Health. 2021;13: 100331. doi:10.1016/j.onehlt.2021.100331

12. Farra A, Frank T, Tondeur L, Bata P, Gody JC, Onambele M, et al. High rate of faecal carriage of extended-spectrum β-lactamase-producing *Enterobacteriaceae* in healthy children in Bangui, Central African Republic. Clin Microbiol Infect 2016;22: 891.e1-891.e4. doi:10.1016/j.cmi.2016.07.001

13. Ouchar Mahamat O, Tidjani A, Lounnas M, Hide M, Benavides J, Somasse C, et al. Fecal carriage of extended-spectrum β-lactamase-producing Enterobacteriaceae in hospital and community settings in Chad. Antimicrob Resist Infect Control. 2019;8: 169. doi:10.1186/s13756-019-0626-z

14. Coldiron ME, Assao B, Page A-L, Hitchings MDT, Alcoba G, Ciglenecki I, et al. Single-dose oral ciprofloxacin prophylaxis as a response to a meningococcal meningitis epidemic in the African meningitis belt: A 3-arm, open-label, cluster-randomized trial. PLoS Med. 2018;15: e1002593. doi:10.1371/journal.pmed.1002593

15. Maataoui N, Langendorf C, Berthe F, Bayjanov JR, van Schaik W, Isanaka S, et al. Increased risk of acquisition and transmission of ESBL-producing Enterobacteriaceae in malnourished children exposed to amoxicillin. J Antimicrob Chemother. 2020;75: 709–717. doi:10.1093/jac/dkz487

16. Bourrel AS, Poirel L, Royer G, Darty M, Vuillemin X, Kieffer N, et al. Colistin resistance in Parisian inpatient faecal *Escherichia coli* as the result of two distinct evolutionary pathways. J Antimicrob Chemother. 2019;74: 1521–1530. doi:10.1093/jac/dkz090

17. Clermont O, Condamine B, Dion S, Gordon DM, Denamur E. The E phylogroup of *Escherichia coli* is highly diverse and mimics the whole *E. coli* species population structure. Environ Microbiol. 2021;23: 7139–7151. doi:10.1111/1462-2920.15742

18. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, et al. Roary: rapid large-scale prokaryote pan genome analysis. Bioinforma 2015;31: 3691–3693. doi:10.1093/bioinformatics/btv421

19. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, et al. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. Mol Biol Evol. 2020;37: 1530–1534. doi:10.1093/molbev/msaa015

20. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res. 2021;49: W293–W296. doi:10.1093/nar/gkab301

21. Lees JA, Harris SR, Tonkin-Hill G, Gladstone RA, Lo SW, Weiser JN, et al. Fast and flexible bacterial genomic epidemiology with PopPUNK. Genome Res. 2019;29: 304–316. doi:10.1101/gr.241455.118

22. Massot M, Châtre P, Condamine B, Métayer V, Clermont O, Madec J-Y, et al. Interplay between Bacterial Clones and Plasmids in the Spread of Antibiotic Resistance Genes in the Gut: Lessons from a Temporal Study in Veal Calves. Appl Environ Microbiol. 2021;87: e0135821. doi:10.1128/AEM.01358-21

23. Diard M, Garry L, Selva M, Mosser T, Denamur E, Matic I. Pathogenicity-associated islands in extraintestinal pathogenic *Escherichia coli* are fitness elements involved in intestinal colonization. J Bacteriol. 2010;192: 4885–4893. doi:10.1128/JB.00804-10

24. Hazen TH, Sahl JW, Fraser CM, Donnenberg MS, Scheutz F, Rasko DA. Refining the pathovar paradigm via phylogenomics of the attaching and effacing *Escherichia coli*. Proc Natl Acad Sci. 2013;110: 12810–12815. doi:10.1073/pnas.1306836110

25. Moremi N, Claus H, Vogel U, Mshana SE. Faecal carriage of CTX-M extended-spectrum beta-lactamase-producing Enterobacteriaceae among street children dwelling in Mwanza city, Tanzania. PloS One. 2017;12: e0184592. doi:10.1371/journal.pone.0184592

26. Bailey JK, Pinyon JL, Anantham S, Hall RM. Distribution of human commensal *Escherichia coli* phylogenetic groups. J Clin Microbiol. 2010;48: 3455–3456. doi:10.1128/JCM.00760-10

27. Roer L, Overballe-Petersen S, Hansen F, Schønning K, Wang M, Røder BL, et al. *Escherichia coli* Sequence Type 410 Is Causing New International High-Risk Clones. mSphere. 2018;3: e00337-18. doi:10.1128/mSphere.00337-18

28. 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. Microb Genomics. 7: 000576. doi:10.1099/mgen.0.000576

29. Ingle DJ, Levine MM, Kotloff KL, Holt KE, Robins-Browne RM. Dynamics of antimicrobial resistance in intestinal *Escherichia coli* from children in community settings in South Asia and sub-Saharan Africa. Nat Microbiol. 2018;3: 1063–1073. doi:10.1038/s41564-018-0217-4

30. Hernandes RT, Hazen TH, dos Santos LF, Richter TKS, Michalski JM, Rasko DA. Comparative genomic analysis provides insight into the phylogeny and virulence of atypical enteropathogenic *Escherichia coli* strains from Brazil. PLoS Negl Trop Dis. 2020;14: e0008373. doi:10.1371/journal.pntd.0008373

31. Massot M, Daubié A-S, Clermont O, Jauréguy F, Couffignal C, Dahbi G, et al. Phylogenetic, virulence and antibiotic resistance characteristics of commensal strain populations of *Escherichia coli* from community subjects in the Paris area in 2010 and evolution over 30 years. Microbiol. 2016;162: 642–650. doi:10.1099/mic.0.000242

32. Nicolas-Chanoine M-H, Bertrand X, Madec J-Y. *Escherichia coli* ST131, an intriguing clonal group. Clin Microbiol Rev. 2014;27: 543–574. doi:10.1128/CMR.00125-13

33. Milenkov M, Rasoanandrasana S, Rahajamanana LV, Rakotomalala RS, Razafindrakoto CA, Rafalimanana C, et al. Prevalence, Risk Factors, and Genetic Characterization of Extended-Spectrum Beta-Lactamase *Escherichia coli* Isolated From Healthy Pregnant Women in Madagascar. Front Microbiol. 2021;12: 786146. doi:10.3389/fmicb.2021.786146

34. Lavollay M, Mamlouk K, Frank T, Akpabie A, Burghoffer B, Ben Redjeb S, et al. Clonal dissemination of a CTX-M-15 beta-lactamase-producing *Escherichia coli* strain in the Paris area, Tunis, and Bangui. Antimicrob Agents Chemother. 2006;50: 2433–2438. doi:10.1128/AAC.00150-06

35. Rodríguez I, Thomas K, Van Essen A, Schink A-K, Day M, Chattaway M, et al. Chromosomal location of blaCTX-M genes in clinical isolates of *Escherichia coli* from Germany, The Netherlands and the UK. Int J Antimicrob Agents. 2014;43: 553–557. doi:10.1016/j.ijantimicag.2014.02.019

36. Shawa M, Furuta Y, Mulenga G, Mubanga M, Mulenga E, Zorigt T, et al. Novel chromosomal insertions of ISEcp1-blaCTX-M-15 and diverse antimicrobial resistance genes in Zambian clinical isolates of *Enterobacter cloacae* and *Escherichia coli*. Antimicrob Resist Infect Control. 2021;10: 79. doi:10.1186/s13756-021-00941-8

37. Zeng S, Luo J, Li X, Zhuo C, Wu A, Chen X, et al. Molecular Epidemiology and Characteristics of CTX-M-55 Extended-Spectrum β-Lactamase-Producing *Escherichia coli* From Guangzhou, China. Front Microbiol. 2021;12: 3046. doi:10.3389/fmicb.2021.730012

38. Hamamoto K, Ueda S, Toyosato T, Yamamoto Y, Hirai I. High Prevalence of Chromosomal blaCTX-M-14 in *Escherichia coli* Isolates Possessing blaCTX-M-14. Antimicrob Agents Chemother. 2016;60: 2582–2584. doi:10.1128/AAC.00108-16

39. Hamamoto K, Hirai I. Characterisation of chromosomally-located blaCTX-M and its surrounding sequence in CTX-M-type extended-spectrum β-lactamase-producing *Escherichia coli* isolates. J Glob Antimicrob Resist. 2019;17: 53–57. doi:10.1016/j.jgar.2018.11.006

40. Lehtinen S, Huisman JS, Bonhoeffer S. Evolutionary mechanisms that determine which bacterial genes are carried on plasmids. Evol Lett. 2021;5: 290–301. doi:10.1002/evl3.226

41. Rodríguez-Beltrán J, DelaFuente J, León-Sampedro R, MacLean RC, San Millán Á. Beyond horizontal gene transfer: the role of plasmids in bacterial evolution. Nat Rev Microbiol. 2021;19: 347–359. doi:10.1038/s41579-020-00497-1

42. Ludden C, Decano AG, Jamrozy D, Pickard D, Morris D, Parkhill J, et al. Genomic surveillance of *Escherichia coli* ST131 identifies local expansion and serial replacement of subclones. Microb Genomics. 6: e000352. doi:10.1099/mgen.0.000352

43. Ruppé E, Lixandru B, Cojocaru R, Büke C, Paramythiotou E, Angebault C, et al. Relative fecal abundance of extended-spectrum-β-lactamase-producing *Escherichia coli* strains and their occurrence in urinary tract infections in women. Antimicrob Agents Chemother. 2013;57: 4512–4517. doi:10.1128/AAC.00238-13

44. Andersen CT, Langendorf C, Garba S, Sayinzonga-Makombe N, Mambula C, Mouniaman I, et al. Risk of community- and hospital-acquired bacteremia and profile of antibiotic resistance in children hospitalized with severe acute malnutrition in Niger. Int J Infect Dis IJID Off Publ Int Soc Infect Dis. 2022; S1201-9712(22)00184–9. doi:10.1016/j.ijid.2022.03.047

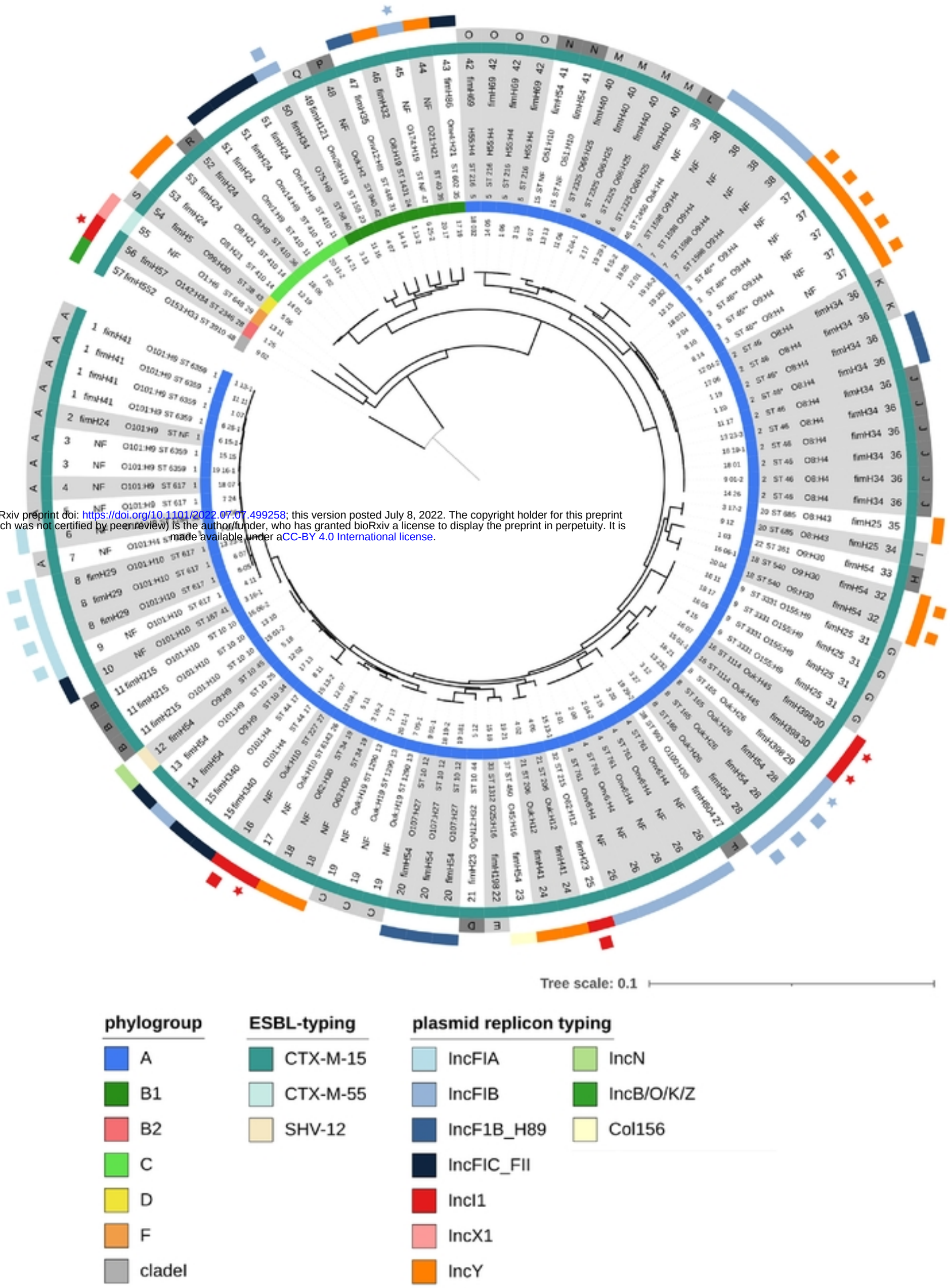
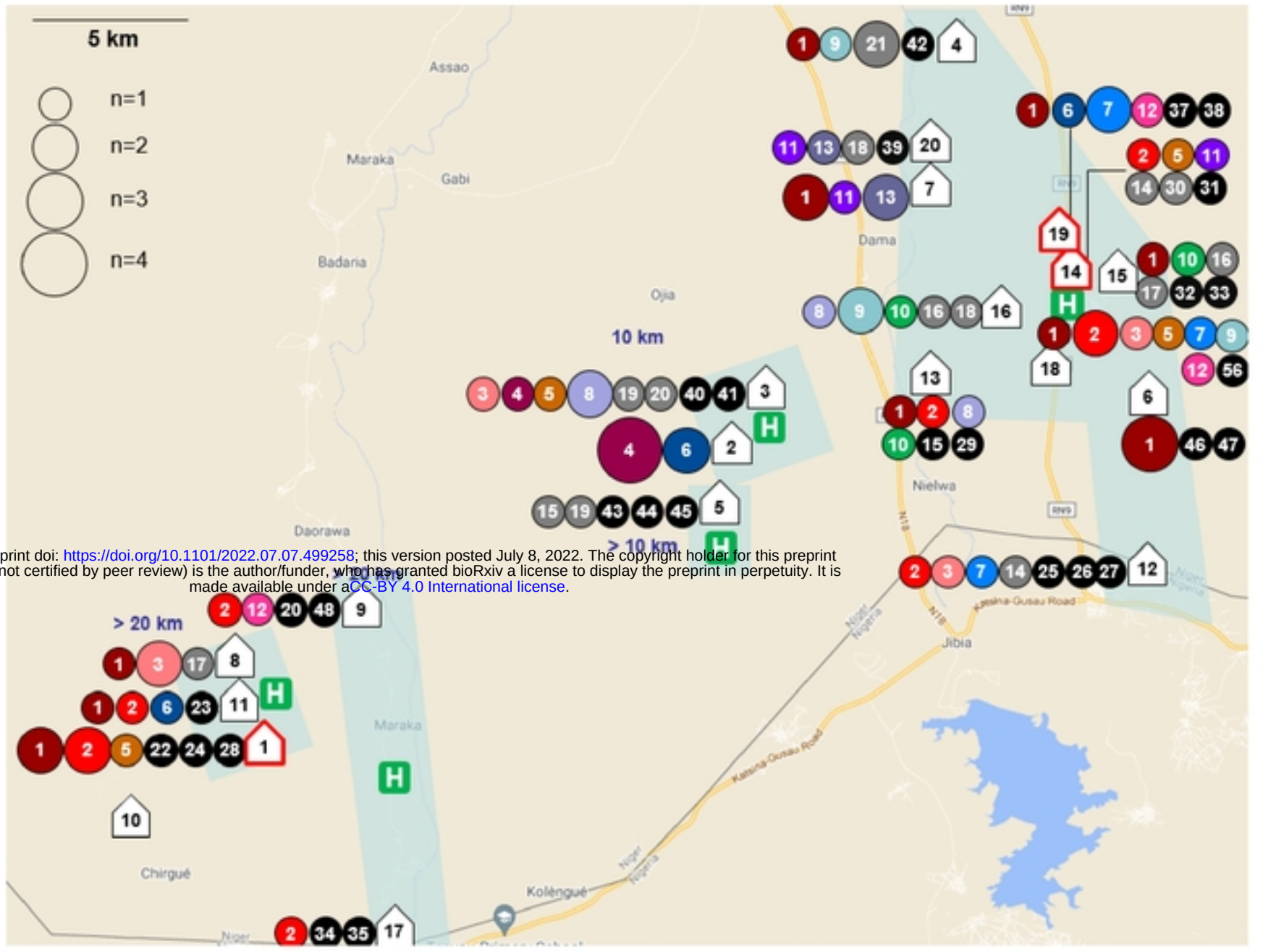


Figure 1. Core genome SNP-based phylogenetic tree of 109 ESBL-*E. coli*. The tree is rooted on *E. coli* clade I (isolate 9_02). From center to periphery, the different layers correspond to the isolate name, the phylogenetic group, the clonal lineage (defined by PopPUNK), the sequence type (according the Warwick University scheme), the serotype, the *fimH* allele, the clone number, the ESBL-typing, the chromosome integration site, and the plasmid replicon type. Squares and stars present in the last layer represent plasmids closely related (i.e. with less than 1% nucleotide differences or two indels). For the different clones, white and dark grey backgrounds are used alternatively to separate isolates. Chromosome integration sites (A to S) are described in Supplementary Table 3.

Figure1

Figure 2. Geographical location of ESBL-*E. coli* according to their clonal lineage



House icons represent the locations of the villages. No isolates were recovered from village 10. Villages with a market are bordered in red. Healthcare centres are indicated by a green box with an H. Villages belonging to the same healthcare district appear in the light blue shapes. Distance to the nearest paved road are indicated in blue. For each village, circles represent clonal lineages whose area is proportional to the number of isolates. Circles coloured with different colours correspond to clonal lineages represented by more than two clones, those in grey and black to clonal lineages represented by two or one clone, respectively.

Figure2