

1 Long-term gut colonization with ESBL-producing *Escherichia coli* in participants without known risk
2 factors from the southeastern United States

3 Coralís Rodríguez García¹, William A. Norfolk^{2a}, Amanda K. Howard^{2,4}, Amanda L. Glatter^{2b}, Megan S.
4 Beaudry^{2c}, Nicholas A. Mallis³, Michael Welton³, Travis C. Glenn², Erin K. Lipp² and Elizabeth A.
5 Ottesen^{1#}

6 ¹ Department of Microbiology, University of Georgia, Athens, GA, USA

7 ² Department of Environmental Health Sciences, University of Georgia, Athens, GA, USA

8 ³ Department of Epidemiology and Biostatistics, University of Georgia, Athens, GA, USA

9 ⁴ Institute of Bioinformatics, University of Georgia, Athens, GA, USA

10 ^a Center for Disease Control and Prevention, Atlanta, Georgia, USA

11 ^b Department of Epidemiology of Microbial Diseases, Yale School of Public Health, New Haven, CT,
12 USA

13 ^c Daicel Arbor Biosciences, Ann Arbor, MI, USA

14 #Corresponding author.

15 Email: ottesen@uga.edu

16

17 Running Title: Gut colonization with ESBL-*E. coli*

18 **Abstract**

19 We evaluated gut carriage of extended spectrum beta lactamase producing *Enterobacteriaceae*
20 (ESBL-E) in southeastern U.S. residents without recent in-patient healthcare exposure. Study enrollment
21 was January 2021-February 2022 in Athens, Georgia, U.S. and included a diverse population of 505
22 adults plus 50 child participants (age 0-5). Based on culture-based screening of stool samples, 4.5% of
23 555 participants carried ESBL-Es. This is slightly higher than reported in studies conducted 2012-2015,
24 which found carriage rates of 2.5-3.9% in healthy U.S. residents.

25 All ESBL-E confirmed isolates (n=25) were identified as *Escherichia coli*. Isolates belonged to
26 11 sequence types, with 48% classified as ST131. Ninety six percent of ESBL-E isolates carried a *bla*_{CTX-M}
27 gene. Isolated ESBL-Es frequently carried virulence genes as well as multiple classes of antibiotic
28 resistance genes. Long-term colonization was common, with 64% of ESBL-E positive participants testing
29 positive when rescreened three months later. One participant yielded isolates belonging to two different *E.*
30 *coli* sequence types that carried *bla*_{CTX-M-1} genes on near-identical plasmids, suggesting intra-gut plasmid
31 transfer.

32 Isolation of *E. coli* on media without antibiotics revealed that ESBL-*E. coli* typically made up a
33 minor fraction of the overall gut *E. coli* population, although in some cases they were the dominant strain.
34 ESBL-E carriage was not associated with a significantly different stool microbiome composition.

35 However, some microbial taxa were differentially abundant in ESBL-E carriers. Together, these results
36 suggest that a small subpopulation of US residents are long-term, asymptomatic carriers of ESBL-Es, and
37 may serve as an important reservoir for community spread of these ESBL genes.

38 **Importance**

39 Antibiotic resistant bacteria, especially *Enterobacteriaceae* carrying ESBLs, have become an
40 increasing public health threat. Increasing numbers of community-associated infections (47% of ESBL-E
41 infections in the U.S.) in participants without healthcare exposure is particularly concerning. This study
42 found that 4.5% of a southeastern United States study population, without in-patient healthcare exposure,
43 were asymptotically colonized with ESBL-E, and 64% of ESBL-E positive participants were still
44 positive when rescreened 3 months later. This suggests that the gut microbiome of healthy participants
45 may represent an understudied community reservoir of ESBL genes and ESBL *Escherichia coli* in the
46 U.S.

47 **Keywords:** extended-spectrum beta-lactamase, antimicrobial resistance, community-acquired ESBL,
48 microbiome, ESBL-*Enterobacteriaceae*

49 **Introduction**

50 Antimicrobial resistance is a serious and growing public health threat worldwide (1, 2) that
51 contributes to increased complication rates, as well as increased treatment costs (3, 4). Extended-
52 spectrum beta-lactamase (ESBL) producing *Enterobacteriaceae* (ESBL-E) are listed as a serious threat by
53 the U.S. Center for Disease Control and Prevention (CDC), contributing to 197,400 infections and 9,100
54 deaths in 2019 (2).

55 ESBL-E were first reported in 1983 and have since spread rapidly throughout the world (5–8). ESBL
56 enzymes confer resistance to multiple antibiotics including penicillin, monobactams, and cephalosporins,
57 commonly used to treat infections caused by Gram negative bacteria (9). There are multiple classes of
58 ESBL enzymes including TEM, SHV, OXA and CTX-M (10–12). Recently, attention has been drawn to
59 the CTX-M class ESBLs as it is currently the most common ESBL found worldwide (13–16). CTX-M-15
60 is typically associated with *E. coli* sequence type 131 (ST131), a frequent causative agent of
61 extraintestinal infections and outbreaks (17, 18). The successful dissemination of this *E. coli* clone
62 contributed to the wide spread of this ESBL enzyme (19). Another factor contributing to the rapid
63 dissemination of ESBLs is their frequent association with mobile elements including plasmids (20–22).
64 Microorganisms often carry ESBL genes in plasmids from the groups IncF, IncI1, IncA/C and IncHI2,
65 facilitating horizontal gene transfer of antibiotic resistance (23). These plasmids also carry antibiotic
66 resistance genes for other classes, easily resulting in multi-drug resistant organisms (24).

67 In a meta-analysis by Bezabih et al. (25), the global intestinal carriage rate of ESBL-*E. coli* in the
68 community increased 10-fold during 2001-2020, highlighting the significance of investigating

69 community-associated ESBL-E. The global prevalence of ESBL-E fecal carriage in the community is
70 highly variable throughout the globe, with substantially higher rates (76.3%) in Tanzania and the lowest
71 (1.9%) in Australia (25, 26). The exact causes of these geographic variations are unknown, although they
72 have been associated to factors including the use of antibiotics in food animals and sanitation standards
73 (27).

74 Few studies have examined ESBL-E carriage in healthy individuals in the U.S. Doi et al. (28, 29)
75 reported a 3.9% prevalence of ESBL-*E. coli* in community-associated samples from outpatient clinics in 5
76 U.S. states collected in 2009-2010. Vento et al. (30) found that only 1 of 101 healthy U.S. military
77 personnel based in the U.S. (May-June 2011) carried an ESBL-E. Weisenberg et al. (31), reported a 2.5%
78 colonization rate for ESBL-E in 2012, among New York residents in participants at pre-travel or no
79 international travel planned. Finally, Islam et al. (32) reported an ESBL-E carriage rate of 3.5% among
80 stool samples from healthy children collected 2013-2015 in three U.S. cities. Overall, these studies
81 suggest that carriage of ESBL-E in the U.S. is low. However, the steady increase of community-
82 associated ESBL-E infections from 2013-2019 (2, 29) suggest a need for updated data on ESBL-E
83 community carriage in the U.S.

84 The aim of this work, was an examination of the prevalence and risk factors for gut colonization of
85 ESBL-E in the community among participants living in or near Athens, Georgia (GA), U.S. We also
86 evaluated the frequency of long-term (~3 month) carriage of ESBL-E among positive participants.
87 Finally, we performed an in-depth genomic analysis, including mobile elements, of ESBL-E isolated in
88 this study and a comparison of gut microbiome community composition between carriers and non-
89 carriers.

90 **Results**

91 We recruited a total of 555 participants including 505 adults and 50 children between January
92 2021 and November 2021 from the southeastern U.S. (mostly northeastern Georgia) (Fig. S1). Overall,
93 the study population reflected the demographic composition of the study area as expected based on U.S.
94 census data (Table S2), with some over-representation of participants who identified as white and non-
95 Hispanic, were 18-39 years of age, and/or reported female sex. (Table S1).

96 Carriage of ESBL *Enterobacteriaceae* (ESBL-E)

97 From 555 samples, 25 participants (4.5%) were positive for stool carriage of an ESBL-E. One of
98 the 25 ESBL-E positive participants was a child, whose parent also tested positive. Of the 25 ESBL-E
99 positive participants, 22 provided a second sample at follow-up (97-176 days later), including the parent-
100 child dyad. 14 of those 22 samples (64%) were ESBL-E positive, whereas the remaining 8 were ESBL-E
101 negative.

102 ESBL-E carriage was not identified as significantly associated with any of the demographic or
103 socioeconomic risk factors examined (Table 1, Table S1). We observed slightly, but not significantly,
104 increased incidence of ESBL-E in participants identified as biological males and Asian or Black/African
105 American race identities. The full list of occupational, lifestyle, and environmental risk factors tested is
106 available in supplemental table S1. A subset of the study population regularly experienced interaction
107 with increased risk environments; however, these exposures were not significantly associated with ESBL-
108 E carriage. ESBL-E carriage was also not associated with significant differences in self-reported
109 gastrointestinal (GI) distress, urinary tract infection, or antibiotic usage. A review of the health
110 information provided by ESBL-E positive participants did not suggest severe chronic health problems.

111 International travel has been widely associated with ESBL colonization (33). The proportion of
112 ESBL-positive participants that lived internationally in the last 5 years is higher than the ESBL-negative
113 participants (17% vs. 7%); however, it was not significantly associated with ESBL-carriage (fdr-corrected
114 p-value 0.09). Only 8 participants (all ESBL-negative) reported international travel in the previous 30
115 days prior to sample collection.

116 **Table 1: Selected demographic characteristics, risk factors, and presence of ESBL-E among all adult**
117 **participants.** *The demographic data presented on Table 1 are based on the questionnaire responses from*
118 *the first visit of all adult participants. The child demographics were not included in the tables or*
119 *statistical analyses of risk factors, as they were not considered independent of the parent participants.*

120 Characteristics of ESBL-E isolates

121 *Genome assemblies and phylogenetic relationships*

122 All confirmed ESBL-E were identified as *Escherichia coli* by MALDI-TOF. High-quality draft
123 genome sequences were obtained for the ESBL-E isolates of each 25 initial and 14 second visit samples
124 (Table S3). Genome sizes ranged from 4,981,979 bp to 5,465,567 bp.

125 Isolates belonged to phylogroups B2 (17 isolates), D (4), A (2) B1 (2) and F (1). Of the B2 isolates,
126 12 belonged to the uropathogenic group ST131, including members with *fimH41* and *fimH30* alleles (Fig.
127 S2). Other closely related groups, ST2279 (2 isolates), ST1193 (2) and ST636 (1) were also present (Fig.
128 1F). All the samples from second visits that yielded confirmed ESBL-E positive were near-clonal (3-55
129 SNPs) with the isolate of the initial visit from that participant with the sole exception of the isolates from
130 participant 497 (Fig. 1E). The initial isolate from participant 497 was assigned to ST154 in phylogroup
131 B1, whereas the isolate from this participant's second sample, 497R, was assigned to ST106 in
132 phylogroup D (Fig. 1A-F). The parent-child dyad samples were both assigned to phylogroup A but had
133 different sequence types: ST10 (parent) and ST305 (child) and 22,309 SNP differences.

134 **Figure 1: Characteristics and classification of ESBL-E isolates.** A) Core-genome single nucleotide
135 polymorphisms (SNPs) phylogenetic tree of ESBL-E isolates per participant colored by Clermont

136 phylogroup, highlighting ST131 isolates. B) Phenotypic characterization of antibiotic resistance by isolate
137 tested in Sensititre, after the confirmation as ESBL-E given their resistance to beta-lactams: Cefotaxime
138 and Ceftazidime. C) Antimicrobial resistance genes found by AMRFinderPlus in the genome, transposons
139 (T) or plasmids (1) of each isolate classified by antimicrobial class. D) Beta-lactamase genes found in
140 each isolate, chromosomally or in mobile elements. Circles labelled with a number match plasmids
141 described in Table S5 for each isolate. E) SNPs between first and second sample confirmed as ESBL-
142 positive to determine clonality of isolates. F) Sequence type (ST) classification of each isolate, colored by
143 phylogroup. Second samples collected (R) are not included in the figure as they were the same sequence
144 type (ST) and share almost the same genetic content as their original sample, with the exception of 497R.

145

146 *Antibiotic resistance profile*

147 All confirmed ESBL-E positive isolates were phenotypically resistant to Ampicillin, Cefotaxime,
148 Ceftazidime, Cefpodoxime, Ceftriaxone, Cefazolin, and Cephalothin. They presented variable resistance
149 to Cefoxitin and Cefepime as shown in Figure 1B. Of the 25 initial ESBL-E isolates, 18 had resistance to
150 Ciprofloxacin (72%) and 6 were resistant to Gentamicin (24%). ESBL-E isolates from the second visit
151 exhibited the same phenotypic antibiotic resistance profile as their original sample except for 491R,
152 which showed a decreased resistance to Cefoxitin compared to the isolate 491. Some second-visit isolates
153 (C483R and 487R) had resistance to Cefoxitin but were susceptible to other beta-lactams tested and were
154 not confirmed as ESBL-E. None of the isolates were resistant to the carbapenems tested.

155 All isolates, except for the child sample, carried at least one *bla*_{CTX-M} beta lactamase gene (Fig.
156 1D). The most predominant beta lactamase gene was *bla*_{CTX-M-15} in 12 isolates, followed by *bla*_{CTX-M-27} in 8
157 isolates, *bla*_{CTX-M-14} in 3 isolates, and *bla*_{CTX-M-1} in 2 isolates. *bla*_{CTX-M-15} was located on a plasmid in 2
158 isolates and located within a chromosomally encoded transposon in 2 isolates. *bla*_{CTX-M-15} was
159 chromosomally encoded without an obvious mobile genetic element in the remaining 8 isolates (Fig. 1D).
160 All isolates with *bla*_{CTX-M-27} carried the gene in IncF plasmids (Table 2).

161 Other beta lactamase genes identified in ESBL-E isolates included *bla*_{EC}, *bla*_{TEM-1}, and *bla*_{OXA-1}.
162 All ESBL isolates carried the *bla*_{EC} gene that confers resistance to Ampicillin, except for isolates 504
163 and 486, which carried *bla*_{EC5} (not shown). Beta lactamase genes found in lower prevalence were *bla*_{TEM-}
164 ₁ (carried by 8 isolates) and *bla*_{OXA-1} (carried by 6 isolates). Seven copies of *bla*_{TEM-1} were carried by IncF
165 plasmids, and one isolate had a copy in the chromosome. *bla*_{OXA-1} was encoded in an IncF plasmid of one
166 isolate, in transposons of two isolates and chromosomally in three isolates (Fig. 1D). Isolate 492 carried
167 three copies of *bla*_{TEM-1}, two of them in two different plasmids and one chromosomally. No currently
168 characterized ESBL beta lactamase genes were identified in the child sample, C483. It is not currently

169 clear whether this isolate carries an uncharacterized gene or may have lost its ESBL gene following initial
170 isolation on selective media.

171 In addition to these beta lactamases, all isolates carried antimicrobial resistance genes in the
172 efflux class, while 16 carried AR genes for Fosfomycin, 12 for sulfonamides, 12 for tetracycline, 10 for
173 trimethoprim, 7 for macrolides, 8 for quaternary ammonium and 5 for phenicol (Fig. 1C). Interestingly,
174 the child isolate C483 carried two operons (*pco* and *sil*) for resistance genes in metal classes like Copper
175 and Silver, both in close association with transposase genes. Nearly all isolates (22 of 26, including 497R)
176 carried quinolone resistance genes. All ST131 isolates had mutations on *gyrA* and *parE* genes and exhibit
177 Ciprofloxacin resistant phenotype, except isolate 482 (Fig. S2). For participants with a confirmed ESBL-
178 E isolate from their second visit, all but two encoded the same resistance genes on re-isolation. One of
179 these is 497R, which as discussed elsewhere, denoted a different sequence type from the initial visit
180 isolate. The other, 493R, was identified as clonal (14 SNPs) but lacked erythromycin, trimethoprim,
181 streptomycin, tetracycline, and sulfamethoxazole antibiotic resistance genes that were present in the
182 original isolate, 493. Both isolates carried a near-identical IncF plasmid, but in 493 this plasmid contained
183 an IS6 insertion element that encoded the resistance genes and is missing in the plasmid carried by 493R
184 (possibly lost in a deletion event) (Fig. S3).

185 *Virulence genes and Plasmids*

186 **Figure 2: Selected virulence genes associated with pathogenic *E. coli* pathotypes.** Virulence genes
187 were identified by VirulenceFinder using the genome assemblies of each isolate. Only selected virulence
188 genes are shown based on their associated pathotype. Colored boxes represent the genes associated with
189 each pathotype and gray circles indicate the presence of that gene in each isolate. Isolates are clustered by
190 Clermont phylogroups matching Figure 1. Circles labelled with a number match plasmids described in
191 Table S5 for each isolate.

192 Phylogroup B2 isolates carried a large variety of virulence genes associated with extraintestinal *E.*
193 *coli* (ExPEC) and had three or more genes associated with sepsis-associated *E. coli* (SEPEC) (Fig. 2).

194 However, virulence genes were not limited to phylogroup B2 isolates. The *afa* genes, encoding for
195 afimbrial adhesins, are also associated with diffusely adherent *E. coli* (DAEC) and were found in isolates
196 498 and 499 from phylogroup D and 492 from phylogroup B2. Six of 9 genes in the adhesin associated
197 locus (*aal*) that encodes for the Coli surface antigen 23 (CS23) were present in isolate 498, which also
198 carried the *eata* gene, encoding for ETEC autotransporter A. Isolate 495 carried the epsilon subtype of
199 intimin outer membrane protein gene, *eae*, associated with enteropathogenic *E. coli* (EPEC), hemorrhagic
200 *E. coli* (EHEC) and Shiga-toxin *E. coli* (STEC). Isolate 488 has 3 P fimbriae genes (A, C, F) in the same
201 transposon as *bla_{OXA-1}* in addition to other resistance genes (*aac(6')-Ib-cr5*, *ermD*, *mdtM*).

202 The most frequently identified and biologically significant plasmids identified belonged to the IncF
203 group, with 24 of the 26 unique ESBL-E isolates carrying at least one plasmid in this group. Most beta
204 lactamases genes located within plasmids were carried by IncF plasmids, except for isolates 497 and
205 497R which carried *bla_{CTX-M-1}* in an IncI1 plasmid (Table 2). The only plasmid carrying more than one
206 beta-lactamase gene was the IncF in isolate 495, which carried genes *bla_{TEM-1}*, *bla_{CTX-M-15}*, and *bla_{OXA-1}*.
207 Isolate 492 has 2 plasmids each carrying a copy of *bla_{TEM-1}*.

208 Eight isolates carried IncF plasmids that encoded *bla_{CTX-M-27}*, all of which also carried the plasmid-
209 encoded enterotoxin, *senB*. *bla_{CTX-M-27}* was encoded in an IS6 transposon array frequently associated with
210 up to 16 additional antibiotic resistance genes (Fig. S4). Two of these isolates, 486 and 504, were near-
211 clonal (52 SNPs) and carried plasmids that were 99.98% identical, with the main difference being two
212 65Kb IS6 transposon arrays encoding (among other genes) 16 resistance genes that are present in the
213 plasmid carried by isolate 504 but absent in 486 (Fig. S5).

214 Three isolates carried an IncI1 plasmid, including 497 and 497R. Isolates 497 and 497R, from the
215 same participant at different times and in different sequence groups, shared a near-identical (99.7%
216 pairwise identity) plasmid encoding *bla_{CTX-M-1}* (Fig. S6). Plasmid group IncY was identified in 2 isolates,
217 while replicons for IncI2, IncH, IncB/O/K/Z and IncN were each present in one isolate accompanied by
218 an IncF plasmid (Table 2).

219 **Table 2: Assembled plasmids carried by each isolate identified by PlasmidFinder with encoded**
220 **antibiotic genes identified by AMRFinderPlus and encoded virulence genes identified by**
221 **VirulenceFinder.** Plasmid number in each isolate matches the numbers in panels C and D of Figure 1
222 and Figure 2.

223 Carriage of antibiotic resistant *E. coli* among ESBL-E positive participants

224 We also examined the overall prevalence of AR among commensal *E. coli* in ESBL-E positive
225 participants. Up to 48 *E. coli* isolates were tested for resistance to ampicillin, ceftriaxone, ciprofloxacin,
226 tetracycline, and trimethoprim. Participants 500 and 498 yielded no *E. coli* colonies on this medium. Six
227 of 25 ESBL-E positive participants showed dominance of beta-lactam resistant strains in their commensal
228 *E. coli* community, with more than 50% of the colonies resistant to both ceftriaxone and ampicillin (Fig.
229 3). Five participants carried commensal *E. coli* resistant to ciprofloxacin. High levels (>50% colonies) of
230 tetracycline resistance were found in 10 of the 25 ESBL-E positive participants. Trimethoprim resistance
231 was less common, with only six participants showing high resistance (>50% colonies) (Fig. 3). In the
232 second samples provided, eight participants had an increase in the proportion of *E. coli* colonies resistant
233 to both ceftriaxone and ampicillin. Most of the second samples that were ESBL-E negative, had a
234 decreased proportion of commensal *E. coli* colonies that were resistant to the antibiotics tested.

235 **Figure 3: Antibiotic resistance phenotype of commensal *E. coli* isolated from ESBL-E positive**
236 **participants.** A range between 1-48 *E. coli* colonies isolated from ESBL-E positive participants were
237 patched into Mueller Hinton II agar with antibiotics at CSLI standard levels. Each circle size and color
238 intensity show the proportion of those colonies that were resistant to the antibiotic tested.

239 *Gut microbial community composition in ESBL-E positive and ESBL-E negative participants*

240 ESBL-E colonization was not associated with significantly different stool microbiome alpha
241 diversity (Fig. S7). Stool microbiome composition as evaluated by weighted Bray-Curtis distances (Fig.
242 4A) resulted in a significant PERMANOVA p-value of 0.03. However, the R^2 value was 0.003,
243 indicating that ESBL-E carriage explained only a very small proportion of community variance.
244 Participants with negative second visit samples did not exhibit significantly larger between-sample shifts
245 in stool microbiome composition than participants with continued colonization at re-sampling (Fig. 4B;
246 Fig. S8).

247 **Figure 4: Microbial community similarities based on ESBL-E carriage.** A) Bray-Curtis distances of
248 metagenomic microbiome samples from all participants, organized in a non-metric multidimensional
249 scaling (NMDS). Magenta dots represent participants that were ESBL-E negative while blue dots
250 represent participants colonized with ESBL-E. B) Comparison of samples (first, duplicate of first, second
251 and duplicate of the second sample) from the same ESBL-E positive participant, calculated by Bray-
252 Curtis distances. Each number indicates the participant ID, while circles or triangles represent the ESBL-
253 E status by sample.

254 The relative abundance of most microbial classes was similar between ESBL-E positive and
255 ESBL-E negative samples (Fig. S9). However, DESeq2 analysis identified 21 amplicon sequence variants
256 (ASVs) that were significantly enriched and 50 that were significantly depleted in samples from ESBL-E
257 positive vs. ESBL-E negative participants (Fig. 5).

258 **Figure 5: Genus assignments for ASVs identified as enriched or depleted in ESBL-E positive**
259 **participants.** ASVs identified as significantly enriched or depleted by DESeq2 analysis with local fit and
260 Wald test identifying significantly different ASVs under an $\alpha=0.001$. ASVs are colored by phylum
261 and organized by genus. Point size shows normalized mean counts and y axis shows log₂fold change in
262 ESBL-E positive participants. ASVs with a positive log₂fold change were enriched, while ASVs with a
263 negative log₂fold change were depleted in ESBL-E positive participants.

264 Discussion

265 A key goal of this work was to evaluate asymptomatic carriage of ESBL-E in U.S. residents without
266 significant healthcare exposure. We observed a carriage rate of 4.5% in our study population, which was
267 recruited from the vicinity of the city of Athens, in northeastern Georgia. This reflects a higher prevalence
268 than previous reports from the U.S. that range between 1.7% and 3.5% (26, 30, 32), although this remains

269 lower than carriage rates reported from European and African countries (34–36). One caveat is that our
270 study participants were recruited from the southeastern U.S., which has reported a higher rate (19.9%) of
271 healthcare associated ESBL-E infections than the national average of 12.2% (37).

272 Our study did not identify any demographic, socioeconomic, environmental, or health-associated risk
273 factors that were significantly associated with the carriage of ESBL-E. This contrasts with previous
274 studies that have reported antibiotic usage (38) and recent international travel (39) as correlated to ESBL-
275 E carriage. These results may have been impacted by the COVID-19 pandemic, which limited
276 international travel during the study period (January 2021-February 2022). This lack of association with
277 specific risk factors suggests that ESBL *E. coli* may be circulating, albeit at low rates, among the general
278 population in the study area.

279 A large fraction (14 of 22) of ESBL-E positive participants remained positive when re-tested at least
280 3 months after their initial visit, which is similar to previous reports in Sweden (40). That study, as well
281 as others, found that sequence type ST131 or strains classified in the B2 and D phylogroups were more
282 likely to persist (40, 41). However, our persistent isolates were distributed across many *E. coli*
283 phylotypes, suggesting that persistent colonization is not confined to these phylogroups (42).

284 Overall, the most prevalent group of ESBL *E. coli* was B2, which matches previous studies in North
285 America (43–46). In terms of sequence types, 46% of the isolates were identified as belonging to ST131,
286 a globally distributed uropathogenic clade that is widely associated with ESBL gene carriage and
287 identified as the most frequent multidrug resistant extraintestinal pathogenic *E. coli* (47). ST131
288 subclades C1, C2 and C1-M27 (as classified by *bla*_{CTX-M} and *fimH* (48)) were all present.

289 All but one isolate carried at least one gene encoding for an CTX-M type enzyme, for a total
290 prevalence of 96%, compared to the 90% reported in a previous review of community isolates from
291 different geographical regions worldwide (49). This suggests that *bla*_{CTX-M} carrying *E. coli* may be a key
292 driver of community ESBL-E spread in the southeastern U.S. In fact, an *E. coli* carrying *bla*_{CTX-M-15} has
293 been isolated from a stream sample in Athens, GA (50). Among the isolates that carried *bla*_{CTX-M} genes,
294 *bla*_{CTX-M-15} was the most predominant, followed by *bla*_{CTX-M-27} and *bla*_{CTX-M-14}. This agrees with previous
295 reports that *bla*_{CTX-M-15} is the most abundant beta lactamase gene circulating in the U.S., closely followed
296 by *bla*_{CTX-M-27} (51, 52), although a third study of urinary tract infections from gram-negative pathogens
297 isolated in hospitals of Canada and the U.S. (2010-2014) reported *bla*_{CTX-M-14} as more abundant than
298 *bla*_{CTX-M-27} (53). *bla*_{CTX-M-27} was primarily associated with ST131 C1-M27 and C1/H30R clades as
299 previously reported (54) but was also present in ST131 clones with *fimH41*. In all recovered isolates with
300 this gene, the *bla*_{CTX-M-27} was borne on an IncF-type plasmid.

301 In addition to the resistance shown to beta-lactams antibiotics, 72% of the ESBL-E isolates showed
302 resistance to Ciprofloxacin and 24% were resistant to Gentamicin. Ciprofloxacin resistance is widely

303 associated with ST131 (17); however, 7 isolates from other groups also exhibited ciprofloxacin resistance
304 suggesting a broader relationship. A previous study of extra-intestinal pathogenic ESBLs from
305 hospitalized patients in India reported a slightly lower prevalence of 65% (55). It remains unclear whether
306 this increased prevalence is because ciprofloxacin resistance is more common in gut isolates or whether
307 this is due to the population studied.

308 The most abundant plasmid type found in our isolates was IncF, which also was the plasmid type
309 carrying the most antibiotic resistance and virulence genes, as has been widely reported (56). While
310 plasmids from the incompatibility group IncII are typically associated with transfer of beta lactamase
311 genes (57), we only isolated ESBL-carrying IncII plasmids from two participants. The IncII plasmid
312 from participant 497, is of particular interest because the ESBL-*E. coli* isolates obtained from the first and
313 second fecal samples provided by this participant carried a near-identical copies of this plasmid but
314 belonged to two different lineages of *E. coli*. This suggests a recent transfer of this plasmid in the human
315 gut. Similar events have been reported elsewhere (58–61) and have been used to argue for the potential
316 role of the human gut as a key site of horizontal gene transfer of antibiotic resistance genes.

317 In agreement with previous studies (62), many of our isolates, particularly those belonging to the B2
318 phylogroup, carried virulence genes associated with extra-intestinal pathogenic *E. coli* including
319 uropathogenic (UPEC) and sepsis associated (SEPEC) strains. The *fimH* gene was the most prevalent
320 virulence factor, followed by UPEC-associated genes *chuA*, *fyuA*, *usp*, and *yfcV* and *kpsMII* (found in
321 84% of ESBL-E isolates). Other virulence genes frequently found in our study include *iutA*, *papA*, *papC*,
322 *sat*, *vat*, *cnfI* and *hlyA*, as previously shown by Iseppi et. al (63), however in contrast to their results *sfa*
323 was not found in any isolate and *afa* genes were only in 3 isolates. Remarkably, the *eae* gene, previously
324 used to classify EPEC isolates (64), was found in the isolate from participant 495, who also carried the
325 only plasmid in this study with 3 different beta-lactamase genes. The presence of virulence genes in
326 plasmids could explain the hybrid classification into different pathotypes of *E. coli* and contribute to the
327 evolution of these pathotypes by accumulating virulence factors within the commensal community (65–
328 67).

329 We also examined the frequency of antibiotic resistance among the overall population of commensal
330 *E. coli* present in stool samples from ESBL-E positive participants. More than half of ESBL-E positive
331 participants carried ampicillin- (64%) and tetracycline- (56%) resistant *E. coli*. However, presumptive
332 ESBL-E made up most commensal *E. coli* in initial samples from only 6 of 25 ESBL-positive
333 participants. This agrees with previous studies that in most cases the ESBL-producing *E. coli* is not the
334 dominant strain among gut *E. coli* (36). Interestingly, 8 of the 22 second samples showed an increase in
335 the fraction of presumptive ESBL-*E. coli* isolated which could indicate a highly dynamic population
336 structure and/or possible genetic exchange of antibiotic resistance genes among commensal *E. coli*.

337 Alpha and beta diversity analysis suggests that ESBL-E carriage was not associated with substantially
338 different gut microbiome composition or diversity, consistent with the results of other studies (68, 69).
339 However, pairwise analysis identified multiple taxa with significantly different abundance in carriers and
340 non-carriers. Similar to other studies, we observed enriched taxa belonging to *Prevotella* (69, 70) and
341 depletion of taxa belonging to *Sellimonas* (71) and *Bacteroides uniformis* (72). On the other hand, some
342 taxa that other groups found enriched in ESBL-E carriers were identified as depleted in our ESBL-E
343 communities, including *Clostridiales* (69), *Erysipelotrichaceae*, *Lactococcus*, *Bilophila* and
344 *Negativibacillus* (70). In addition, *Desulfovibrio* and *Oscillospira* genera were identified as enriched
345 among ESBL-E carriers in our study but depleted in another (68). These differences could be explained
346 by the species of ESBL-*Enterobacteriaceae* in each study, given that the microbial population differs
347 based on the ESBL-E species (71), or other factors influencing the microbiome composition as discussed
348 before (73).

349 Notable limitations of our study include that it was restricted to a specific geographic range (Athens,
350 GA and vicinity) and that the study commenced during the global COVID-19 pandemic, which caused
351 multiple large-scale changes in behavior including limitations on international travel. In addition, there
352 was a possible bias provided by the participant self-reporting of risk factors. However, overall our results
353 suggest that a subset of southeast U.S. residents are likely asymptomatic carriers of ESBL-*E. coli*.

354 **Conclusions**

355 To our knowledge this is the first genomic analysis of community associated ESBL-E carriage in the
356 southeastern U.S. including long-term colonization, as previous studies only focused on the prevalence of
357 ESBL-E or on healthcare-associated isolates. If the results from this study can be extrapolated, it suggests
358 the potential for a small but notable increase in ESBL-E carriage in the U.S. since 2015, consistent with
359 reports that the frequency of community-acquired ESBL-E infections also increased over this period (25)
360 and supporting the role community-associated isolates in the incidence of ESBL-E outbreaks. Long term
361 (>3 months) colonization was common in the study population, which underscores the potential of the
362 human gut microbiome to serve as a long-term reservoir of ESBL *Enterobacteriaceae*. Colonizing ESBL
363 *Enterobacteriaceae* were all identified as *E. coli* including strains that are unlikely to be pathogenic and
364 strains carrying virulence genes associated with extraintestinal pathogenic *E. coli*. Finally, many strains
365 carried multidrug resistance plasmids and we observed at least one participant where the same plasmid
366 was observed in isolates with different phylogenetic backgrounds, consistent with a role for the human
367 gut as a hotspot for antibiotic resistance gene exchange (74). Overall, our study suggests that the human
368 gut may represent an important but under-recognized reservoir of ESBL genes and ESBL-carrying *E. coli*
369 and highlights the relevance and importance of understanding the role of gut commensals in mediating
370 the spread of antibiotic resistance.

371 **Materials and Methods**

372 Specimen collection, metadata collection and analysis

373 Recruitment and consenting of participants were performed by the Clinical and Translational
374 Research Unit (CTRU) at the University of Georgia. Inclusion criteria included: the ability and
375 willingness to answer an online survey regarding risk factors and to physically visit the CTRU to obtain
376 and return the specimen collection kit, as well as age >18 years for adult participants. Adult participants
377 with a child age ≤ 5 years in their household were invited to enroll the child in the study. Exclusion
378 criteria included pregnancy and in-patient (overnight) hospitalization/health care in the last 12 months for
379 reasons other than uncomplicated childbirth. Prospective participants who reported systemic (oral or
380 intravenous) antibiotic use within the last 48 hours were asked to schedule specimen collection/drop-off
381 for a later date, as were participants with active COVID-19 infections. Children who did not reside with
382 participant parent for 5 days or more per week were also excluded. A signed consent was obtained from
383 each participant or parent. To protect the confidentiality of personal data, all participants were assigned a
384 unique, randomly generated identification number. All research activities involving human subject
385 research were reviewed and approved by the Institutional Review Board (IRB) at University of Georgia,
386 Athens.

387 Participants were provided a stool specimen container that was pre-labeled with their participant
388 ID (Medline Industries, Cat. No. DYND36500) with a scheduled sample return date. Participants were
389 asked to collect a stool sample using the provided collection kit and to complete an online questionnaire
390 to collect demographic information as well as possible environmental risk factors for carriage of antibiotic
391 resistant bacteria (see supplemental materials) as close in time as possible and no more than 12 hours
392 before their scheduled return appointment. After collection, they were instructed to keep the stool
393 specimen in a refrigerator or protected from heat until return to the facility. Upon return to the CTRU,
394 stool specimens were stored at 4°C until processing. Upon transfer to the study team, stool specimens
395 were subdivided for processing, typically within 1-3 hours of receipt (max. 24 hours). For culture work,
396 200 mg of stool specimens were diluted in 1 mL of 1X PBS and immediately processed as described
397 below. For DNA extraction and molecular analyses, 200 mg sub-specimens were transferred to sterile
398 cryovials and stored at -20°C until processing as described below.

399 Participant data was collected by online survey (Qualtrics) and matched to laboratory samples via
400 alphanumeric participant IDs. Participant surveys consisted of a series of questions relating to each
401 participant's demographics, lifestyle, socioeconomic status, environmental exposures, preexisting medical
402 conditions, and travel experience. All child surveys were completed by a designated parent or guardian.
403 Upon completion of sample collection, all data were downloaded, matched to laboratory results, and
404 cleaned for analysis using R statistical analysis software with 'tidyverse', 'readxl', 'magrittr', 'reshape2'

405 and ‘rcompanion’. ESBL carriage frequencies and Fisher’s test with Monte-Carlo simulations were used
406 to calculate p-values using ‘fisher.test’ function in R. P-adjusted values were calculated using the
407 ‘p.adjust’ function with the Benjamini-Hochberg procedure for False Discovery Rate (FDR) (75, 76).
408 Statistical significance was considered from p-values less than 0.05.

409 Isolation, testing and confirmation of ESBL-E

410 200 mg of each stool specimen was diluted in 1 ml of 1X PBS. 100µl of the diluted aliquot was
411 pre-enriched overnight in 5 ml tryptic soy broth (TSB) at 37°C. 10 µl of the enriched culture was then
412 spread onto CHROMagarTM ESBL plates and incubated at 37°C overnight. Well-isolated colonies from
413 CHROMagar ESBL plates were re-streaked for isolation and purification on the same medium and
414 incubated under the same conditions. Presumptive ESBL-producing colonies were re-streaked on blood
415 agar, identified by MALDI-TOF MS, and confirmed by antibiotic resistant profiling using Sensititre panel
416 ESB1F at the UGA Veterinary Diagnostics Laboratory. An isolate was considered ESBL-positive if it
417 was resistant to Cefotaxime and Ceftazidime but susceptible when clavulanic acid was added to each.
418 Other antibiotics tested with Sensititre included Ampicillin, Cefazolin, Cefepime, Cefoxitin,
419 Cefpodoxime, Ceftriaxone, Cephalothin, Ciprofloxacin, Gentamicin, Imipenem, Meropenem and
420 Piperacillin/Tazobactam. Confirmed ESBL-positive participants were requested to provide a second
421 sample at least 90 days after the initial participation. 22 of 25 positive participants provided a second
422 sample, the remaining 3 could not be contacted after multiple attempts. Second-visit samples were
423 processed under the same conditions as the first samples.

424 Whole-genome sequencing and analysis

425 Isolated colonies from ESBL-confirmed participants were used for whole-genome analysis.
426 Genomic DNA of isolates was extracted using Omega Biotek Bacterial DNA Kit. Purified DNA was
427 arrayed in 96-well plates, normalized, and run on agarose minigels for QC. Genomic libraries were
428 created using an NEBNext Ultra II FS DNA Library Prep Kit with custom primers and protocols (77) and
429 an Opentrons OT-2 robot. Libraries for all ESBL-confirmed isolates were sequenced on a NovaSeq S4
430 6000 to obtain PE150 reads. Trimming and quality filtering was performed via Trimmomatic v0.39.

431 For long-read sequencing, DNA extracts were normalized to 50 ng per 9 µl. Oxford Nanopore
432 Technologies (ONT) libraries were prepared following manufacturer’s instructions for the rapid barcode
433 96 kit and cleaned with Ampure XP beads. A total of 75 µl of library was loaded onto an ONT MinION
434 flow cell (9.3.1) for sequencing. Following sequencing, bases were called using super accuracy mode
435 with Guppy v6.1.1.

436 *De novo* assembly was performed using Unicycler v0.4.7. using the short-reads first approach and
437 default settings (78). A phylogenetic tree between ESBL-E genomes was constructed using maximum
438 likelihood and core single nucleotide polymorphisms with Parsnp, with a randomly selected reference

439 genome (AREA_490). Core genome single nucleotide polymorphisms (SNPs) between isolates from first
440 and second samples were calculated using CSI Phylogeny (79, 80). AMRFinder v3.9.8 was used to
441 identify antimicrobial resistance genes in assembled genomes using the Plus genes database (--plus) and
442 *E. coli* as the reference organism (-O) to identify point mutations (81). Mobile genetic elements were
443 identified using MobileElementFinder v1.0.3 with antimicrobial resistance genes annotated (82). Tools
444 from the Center for Genomic Epidemiology (<http://www.genomicepidemiology.org/services/>) that were
445 used to analyze the hybrid assemblies including: MLST v2.0.9 with configuration for *Escherichia coli* #1
446 and a minimum depth of 5X; SeroTypeFinder v2.0.1 with 85% threshold ID and 60% minimum length,
447 FimTyper v1.0 with 95% threshold ID, 95% threshold identity and 60% minimum coverage and
448 VirulenceFinder v2.0.3 for *E. coli* with 90% threshold ID and 60% minimum length (83–86). Assemblies
449 were classified by phylogroups using the ClermonTyping web server ([http://clermontyping.iame-](http://clermontyping.iame-research.center)
450 [research.center](http://clermontyping.iame-research.center))(87).

451 Microbial community characterization

452 Fecal *E. coli* were isolated without antibiotic selection from all stool samples using CHROMagar
453 ECC. Up to 48 well-isolated colonies (if available) were re-streaked on LB plates for purity, then arrayed
454 in 96-well plates to grow overnight with Mueller Hinton II broth media. 3 µl subsamples of the overnight
455 culture were patched in Mueller Hinton II agar plates containing antibiotics at CLSI standards: Ampicillin
456 (32 µg/ml), Ceftriaxone (32 µg/ml), Ciprofloxacin (4 µg/ml), Tetracycline (4 µg/ml) and Trimethoprim (8
457 µg/ml) to characterize the AR profile of the commensal *E. coli* community in the ESBL-E positive
458 participants.

459 For 16S rRNA gene library sequencing, DNA was isolated from 200 mg of each stool sample
460 with the Omega Biotek Stool DNA kit following the manufacturer protocol for pathogen detection
461 (Omega Biotek, Norcross, GA, USA). The optional incubation with DS buffer and Proteinase K was
462 performed to improve recovery of Gram-positive bacteria. Extracted DNA was eluted in 100µl of elution
463 buffer from the kit heated at 65°C. DNA concentrations were measured with NanoDrop. Amplification of
464 16S rRNA gene was done using primers 515F and 806R as described previously (88), followed by library
465 sequencing with Illumina MiSeq 250 x 250 bp at the Georgia Genomics Facility. Sequence analysis was
466 performed using DADA2 (version 1.18) in R (version 4.0.2) including filtering, trimming, merge and
467 taxonomy assignment with SILVA database, version 138 (89, 90). Data analyses were done in R using
468 phyloseq, vegan, dplyr, tidyverse and DESeq packages with samples rarefied to a depth of 10,000 when
469 needed (91, 92).

470 Data availability

471 Whole-genome assemblies and 16S rDNA sequences have been submitted to NCBI under the BioProject:
472 PRJNA894544.

473

474 Acknowledgements

475 This work was supported by BAA contract no. 75D30120C09496 from the Centers for Disease Control
476 and Prevention. Funders had no role in data collection and interpretation, or the decision to submit the
477 publication. Participants recruitment and sample processing was performed with the help of the CTRU at
478 University of Georgia, which is supported by the National Center for Advancing Translational Sciences
479 of the National Institutes of Health under award number: UL1TR002378. The content is solely the
480 responsibility of the authors and does not necessarily represent the official views of the National Institutes
481 of Health.

482 Author roles based on CRediT taxonomy: Coralís Rodríguez García, data curation, formal analysis,
483 investigation, methodology, visualization, writing – original draft; William A. Norfolk, data curation,
484 formal analysis, visualization; Amanda K. Howard, investigation, methodology, formal analysis; Amanda
485 L. Glatter, investigation; Megan S. Beaudry, investigation; Nicholas A. Mallis, investigation; Michael
486 Welton, conceptualization, methodology; Travis C. Glenn, conceptualization, methodology, supervision,
487 project administration; Erin K. Lipp, conceptualization, methodology, supervision, project administration;
488 Elizabeth A. Ottesen, conceptualization, funding acquisition, methodology, supervision, project
489 administration, writing – review and editing.

490

491 We also thank Andreas Handel, Jonathan Frye, and Charlene Jackson for their advice regarding
492 methodology and data analysis, as well as Marissa Howard, Kelly Petersen, Bryce Miller, Katie Dillon,
493 Kirty Muthyala, Claire Pearson for their assistance in data collection.

494

495 References

- 496 1. WHO Association Media Centre. 2015. Sixty-Eighth World Health Assembly. World Health Organization.
- 497 2. CDC. 2019. Antibiotic Resistance Threats in The United States, 2019 U.S. Department of Health and
498 Human Services.
- 499 3. Nelson RE, Slayton RB, Stevens VW, Jones MM, Khader K, Rubin MA, Jernigan JA, Samore MH. 2017.
500 Attributable mortality of healthcare-associated infections due to multidrug-resistant gram-negative bacteria
501 and methicillin-resistant *Staphylococcus aureus*. *Infect Control Hosp Epidemiol* 38:848–856.
- 502 4. Neidell MJ, Cohen B, Furuya Y, Hill J, Jeon CY, Glied S, Larson EL. 2012. Costs of healthcare-and
503 community-associated infections with antimicrobial-resistant versus antimicrobial-susceptible organisms.
504 *Clinical Infectious Diseases* 55:807–815.
- 505 5. Knothe H, Shah P, Krcmery V, Antal M, Mitsuhashi S. 1983. Transferable resistance to cefotaxime,
506 cefoxitin, cefamandole and cefuroxime in clinical isolates of *Klebsiella pneumoniae* and *Serratia*
507 *marcescens*. *Infection* 11:315–317.
- 508 6. Bush K, Bradford PA. 2020. Epidemiology of B-Lactamase-Producing Pathogens. *Clin Microbiol Rev*
509 33:1–37.
- 510 7. Rupp ME, Fey PD. 2003. Extended spectrum beta-lactamase (ESBL)-producing *Enterobacteriaceae*:
511 considerations for diagnosis, prevention and drug treatment. *Drugs* 63:353–365.
- 512 8. Whichard JM, Joyce K, Fey PD, Nelson JM, Angulo FJ, Barrett TJ. 2005. β -Lactam Resistance and
513 *Enterobacteriaceae*, United States. *Emerg Infect Dis* 11:1464–1466.
- 514 9. Rawat D, Nair D. 2010. Extended-spectrum β -lactamases in gram negative bacteria. *J Glob Infect Dis* 2:263.

- 515 10. Castanheira M, Simner PJ, Bradford PA. 2021. Extended-spectrum β -lactamases: An update on their
516 characteristics, epidemiology and detection. *JAC Antimicrob Resist* 3.
- 517 11. Doi Y, Iovleva A, Bonomo RA. 2017. The ecology of extended-spectrum β -lactamases (ESBLs) in the
518 developed world. *J Travel Med* 24:S44–S51.
- 519 12. Logan LK, Rispens JR, Medernach RL, Domitrovic TN, Hujer AM, Marshall SH, Rudin SD, Qureshi NK,
520 Zheng X, Hayden MK, Weinstein RA, Bonomo RA. 2021. A Multicentered Study of the Clinical and
521 Molecular Epidemiology of TEM- and SHV-type Extended-Spectrum Beta-Lactamase Producing
522 *Enterobacterales* Infections in Children. *Pediatric Infectious Disease Journal* 40:39–43.
- 523 13. Cantón R, González-Alba JM, Galán JC. 2012. CTX-M enzymes: Origin and diffusion. *Front Microbiol* 3.
- 524 14. Cantón R, Coque TM. 2006. The CTX-M β -lactamase pandemic. *Curr Opin Microbiol* 9:466–475.
- 525 15. Logan LK, Hujer AM, Marshall SH, Domitrovic TN, Rudin SD, Zheng X, Qureshi NK, Hayden MK,
526 Scaggs FA, Karadkhele A, Bonomo RA. 2016. Analysis of β -Lactamase Resistance Determinants in
527 Enterobacteriaceae from Chicago Children: a Multicenter Survey. *Antimicrob Agents Chemother* 60:3462–
528 3469.
- 529 16. Hu F, O’Hara JA, Rivera JI, Doi Y. 2014. Molecular Features of Community-Associated Extended-
530 Spectrum- β -Lactamase-Producing *Escherichia coli* Strains in the United States. *Antimicrob Agents*
531 *Chemother* 58:6953–6957.
- 532 17. Matsumura Y, Pitout JDD, Peirano G, Devinney R, Noguchi T, Yamamoto M, Gomi R, Matsuda T, Nakano
533 S, Nagao M, Tanaka M, Ichiyama S. 2017. Rapid Identification of Different *Escherichia coli* Sequence Type
534 131 Clades. *Antimicrob Agents Chemother* 61:1–9.
- 535 18. Salamzade R, McElheny CL, Manson AL, Earl AM, Shaikh N, Doi Y. 2023. Genomic epidemiology and
536 antibiotic susceptibility profiling of uropathogenic *Escherichia coli* among children in the United States.
537 *mSphere* 8.
- 538 19. Nicolas-Chanoine MH, Bertrand X, Madec JY. 2014. *Escherichia coli* st131, an intriguing clonal group.
539 *Clin Microbiol Rev* 27:543–574.
- 540 20. Shaikh S, Fatima J, Shakil S, Rizvi SMD, Kamal MA. 2015. Antibiotic resistance and extended spectrum
541 beta-lactamases: Types, epidemiology and treatment. *Saudi J Biol Sci* 22:90–101.
- 542 21. Johnson TJ, Wannemuehler YM, Johnson SJ, Logue CM, White DG, Doetkott C, Nolan LK. 2007. Plasmid
543 replicon typing of commensal and pathogenic *Escherichia coli* isolates. *Appl Environ Microbiol* 73:1976–
544 1983.
- 545 22. Lukac PJ, Bonomo RA, Logan LK. 2015. Extended-Spectrum β -Lactamase-Producing Enterobacteriaceae in
546 Children: Old Foe, Emerging Threat. *Clinical Infectious Diseases* <https://doi.org/10.1093/cid/civ020>.
- 547 23. Carattoli A. 2011. Plasmids in Gram negatives: Molecular typing of resistance plasmids. *International*
548 *Journal of Medical Microbiology* 301:654–658.
- 549 24. van Duin D, Paterson DL. 2020. Multidrug-Resistant Bacteria in the Community: An Update. *Infect Dis*
550 *Clin North Am* 34:709–722.
- 551 25. Bezabih YM, Bezabih A, Dion M, Batard E, Tekla S, Obole A, Dessalegn N, Enyew A, Roujeinikova A,
552 Alamneh E, Mirkazemi C, Peterson GM, Bezabhe WM. 2022. Comparison of the global prevalence and
553 trend of human intestinal carriage of ESBL-producing *Escherichia coli* between healthcare and community
554 settings: A systematic review and meta-Analysis. *JAC Antimicrob Resist* 4.
- 555 26. Bezabih YM, Sabiiti W, Alamneh E, Bezabih A, Peterson GM, Bezabhe WM, Roujeinikova A. 2021. The
556 global prevalence and trend of human intestinal carriage of ESBL-producing *Escherichia coli* in the
557 community. *Journal of Antimicrobial Chemotherapy* 76:22–29.
- 558 27. Bevan ER, Jones AM, Hawkey PM. 2017. Global epidemiology of CTX-M β -lactamases: Temporal and
559 geographical shifts in genotype. *Journal of Antimicrobial Chemotherapy* 72.
- 560 28. Doi Y, Park YS, Rivera JI, Adams-Haduch JM, Hingwe A, Sordillo EM, Lewis JS, Howard WJ, Johnson
561 LE, Polsky B, Jorgensen JH, Richter SS, Shutt KA, Paterson DL. 2013. Community-associated extended-
562 spectrum β -lactamase-producing *Escherichia coli* infection in the United States. *Clinical Infectious Diseases*
563 56:641–648.
- 564 29. Doi Y, Adams J, O’Keefe A, Qureshi Z, Ewan L, Paterson DL. 2007. Community-acquired extended
565 spectrum β -lactamase producers, United States.
- 566 30. Vento TJ, Cole DW, Mende K, Calvano TP, Rini EA, Tully CC, Zera WC, Guymon CH, Yu X, Cheatle KA,
567 Akers KS, Beckius ML, Landrum ML, Murray CK. 2013. Multidrug-resistant gram-negative bacteria
568 colonization of healthy US military personnel in the US and Afghanistan. *BMC Infect Dis* 13.

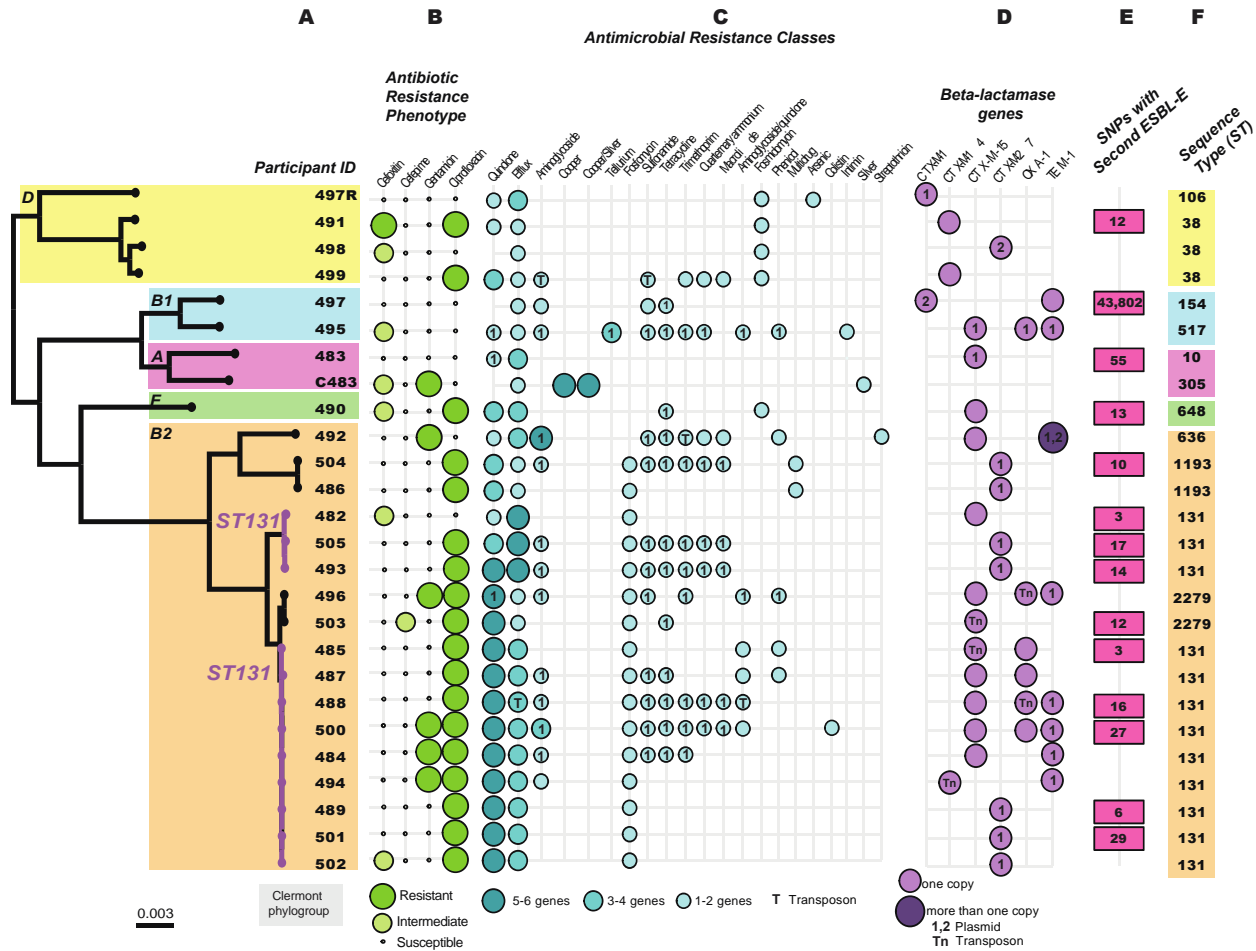
- 569 31. Weisenberg SA, Mediavilla JR, Chen L, Alexander EL, Rhee KY, Kreiswirth BN, Jenkins SG. 2012.
570 Extended Spectrum Beta-Lactamase-Producing Enterobacteriaceae in International Travelers and Non-
571 Travelers in New York City. *PLoS One* 7:2010–2013.
- 572 32. Islam S, Selvarangan R, Kanwar N, McHenry R, Chappell JD, Halasa N, Wikswo ME, Payne DC, Azimi
573 PH, McDonald LC, Gomez-Duarte OG. 2018. Intestinal Carriage of Third-Generation Cephalosporin-
574 Resistant and Extended-Spectrum β -Lactamase-Producing Enterobacteriaceae in Healthy US Children. *J*
575 *Pediatric Infect Dis Soc* 7:234–240.
- 576 33. Karanika S, Karantanos T, Arvanitis M, Grigoras C, Mylonakis E. 2016. Fecal Colonization with Extended-
577 spectrum Beta-lactamase-Producing Enterobacteriaceae and Risk Factors among Healthy Individuals: A
578 Systematic Review and Metaanalysis. *Clinical Infectious Diseases* 63:310–318.
- 579 34. Geser N, Stephan R, Korczak BM, Beutin L, Hächler H. 2012. Molecular identification of extended-
580 spectrum- β -lactamase genes from Enterobacteriaceae isolated from healthy human carriers in Switzerland.
581 *Antimicrob Agents Chemother* 56:1609–1612.
- 582 35. Mshana SE, Falgenhauer L, Mirambo MM, Mushi MF, Moremi N, Julius R, Seni J, Imirzalioglu C, Matee
583 M, Chakraborty T. 2016. Predictors of blaCTX-M-15 in varieties of *Escherichia coli* genotypes from
584 humans in community settings in Mwanza, Tanzania. *BMC Infect Dis* 16:1–9.
- 585 36. Nicolas-Chanoine MH, Gruson C, Bialek-Davenet S, Bertrand X, Thomas-Jean F, Bert F, Moyat M, Meiller
586 E, Marcon E, Danchin N, Noussair L, Moreau R, Leflon-Guibout V. 2013. 10-fold increase (2006–11) in the
587 rate of healthy subjects with extended-spectrum β -lactamase-producing *Escherichia coli* faecal carriage in a
588 parisian check-up centre. *Journal of Antimicrobial Chemotherapy* 68:562–568.
- 589 37. Castanheira M, Farrell SE, Krause KM, Jones RN, Sader HS. 2014. Contemporary diversity of β -lactamases
590 among enterobacteriaceae in the nine U.S. census regions and ceftazidime-avibactam activity tested against
591 isolates producing the most prevalent β -lactamase groups. *Antimicrob Agents Chemother* 58:833–838.
- 592 38. van den Bunt G, van Pelt W, Hidalgo L, Scharringa J, de Greeff SC, Schürch AC, Mughini-Gras L, Bonten
593 MJM, Fluit AC. 2019. Prevalence, risk factors and genetic characterisation of extended-spectrum beta-
594 lactamase and carbapenemase-producing Enterobacteriaceae (ESBL-E and CPE): A community-based
595 cross-sectional study, the Netherlands, 2014 to 2016. *Eurosurveillance* 24.
- 596 39. Mizuno Y, Miura Y, Yamaguchi T, Matsumoto T. 2016. Extended-spectrum β -lactamase-producing
597 Enterobacteriaceae colonisation in long-term overseas business travellers. *Travel Med Infect Dis* 14:561–
598 567.
- 599 40. Titelman E, Hasan CM, Iversen A, Nauc ler P, Kais M, Kalin M, Giske CG. 2014. Faecal carriage of
600 extended-spectrum β -lactamase-producing Enterobacteriaceae is common 12 months after infection and is
601 related to strain factors. *Clinical Microbiology and Infection* 20.
- 602 41. van Duijkeren E, Wienders CCH, Dierikx CM, van Hoek AHAM, Hengeveld P, Veenman C, Florijn A,
603 Lotterman A, Smit LAM, van Dissel JT, Maassen CBM, de Greeff SC. 2018. Long-term Carriage of
604 Extended-Spectrum β -Lactamase-Producing *Escherichia coli* and *Klebsiella pneumoniae* in the General
605 Population in The Netherlands. *Clin Infect Dis* 66.
- 606 42. Rumball NA, Alm EW, McLellan SL. 2023. Genetic Determinants of *Escherichia coli* Survival in Beach
607 Sand. *Appl Environ Microbiol* 89.
- 608 43. Logue CM, Doetkott C, Mangiamele P, Wannemuehler YM, Johnson TJ, Tivendale KA, Li G, Sherwood
609 JS, Nolan LK. 2012. Genotypic and phenotypic traits that distinguish neonatal meningitis-associated
610 *Escherichia coli* from fecal *E. coli* isolates of healthy human hosts. *Appl Environ Microbiol* 78.
- 611 44. Johnson JR, Owens K, Gajewski A, Kuskowski MA. 2005. Bacterial characteristics in relation to clinical
612 source of *Escherichia coli* isolates from women with acute cystitis or pyelonephritis and uninfected women.
613 *J Clin Microbiol* 43.
- 614 45. Zhang L, Foxman B, Marrs C. 2002. Both urinary and rectal *Escherichia coli* isolates are dominated by
615 strains of phylogenetic group B2. *J Clin Microbiol* 40.
- 616 46. Sannes MR, Kuskowski MA, Owens K, Gajewski A, Johnson JR. 2004. Virulence factor profiles and
617 phylogenetic background of *Escherichia coli* isolates from veterans with bacteremia and uninfected control
618 subjects. *Journal of Infectious Diseases* 190.
- 619 47. Rogers BA, Doi Y. 2013. Who Is Leading This Dance? Understanding the Spread of *Escherichia coli*
620 Sequence Type 131. *Infect Control Hosp Epidemiol* 34.
- 621 48. Petty NK, Zakour NLB, Stanton-Cook M, Skippington E, Totsika M, Forde BM, Phan MD, Moriel DG,
622 Peters KM, Davies M, Rogers BA, Dougan G, Rodriguez-Ba no J, Pascual A, Pitout JDD, Upton M,
623 Paterson DL, Walsh TR, Schembri MA, Beatson SA. 2014. Global dissemination of a multidrug resistant
624 *Escherichia coli* clone. *Proc Natl Acad Sci U S A* 111.

- 625 49. Woerther PL, Burdet C, Chachaty E, Andremont A. 2013. Trends in human fecal carriage of extended-
626 spectrum β -lactamases in the community: Toward the globalization of CTX-M. *Clin Microbiol Rev* 26:744–
627 758.
- 628 50. Cho S, Nguyen HAT, McDonald JM, Woodley TA, Hiott LM, Barrett JB, Jackson CR, Frye JG. 2019.
629 Genetic characterization of antimicrobial-resistant *Escherichia coli* isolated from a mixed-use watershed in
630 northeast Georgia, USA. *Int J Environ Res Public Health* 16:1–14.
- 631 51. Duffy N, Karlsson M, Reses HE, Campbell D, Daniels J, Stanton RA, Janelle SJ, Schutz K, Bamberg W,
632 Rebolledo PA, Bower C, Blakney R, Jacob JT, Phipps EC, Flores KG, Dumyati G, Kopin H, Tsay R, Kainer
633 MA, Muleta D, Byrd-Warner B, Grass JE, Lutgring JD, Rasheed JK, Elkins CA, Magill SS, See I. 2022.
634 Epidemiology of extended-spectrum β -lactamase-producing Enterobacterales in five US sites participating
635 in the Emerging Infections Program, 2017. *Infect Control Hosp Epidemiol* 43:1586–1594.
- 636 52. Mostafa HH, Cameron A, Taffner SM, Wang J, Malek A, Dumyati G, Hardy DJ, Pecora ND. 2020.
637 Genomic Surveillance of Ceftriaxone-Resistant *Escherichia coli* in Western New York Suggests the
638 Extended-Spectrum β -Lactamase blaCTX-M-27 Is Emerging on Distinct Plasmids in ST38. *Front Microbiol*
639 11.
- 640 53. Lob SH, Nicolle LE, Hoban DJ, Kazmierczak KM, Badal RE, Sahn DF. 2016. Susceptibility patterns and
641 ESBL rates of *Escherichia coli* from urinary tract infections in Canada and the United States, SMART
642 2010–2014. *Diagn Microbiol Infect Dis* 85:459–465.
- 643 54. Matsumura Y, Pitout JDD, Gomi R, Matsuda T, Noguchi T, Yamamoto M, Peirano G, De Vinney R,
644 Bradford PA, Motyl MR, Tanaka M, Nagao M, Takakura S, Ichiyama S. 2016. Global *Escherichia coli*
645 sequence type 131 clade with blaCTX-M-27 gene. *Emerg Infect Dis* 22:1900–1907.
- 646 55. Chakraborty A, Adhikari P, Shenoy S, Saralaya V. 2015. Clinical significance and phylogenetic background
647 of extended spectrum β -lactamase producing *Escherichia coli* isolates from extra-intestinal infections. *J*
648 *Infect Public Health* 8:248–253.
- 649 56. Carattoli A. 2009. Resistance plasmid families in Enterobacteriaceae. *Antimicrob Agents Chemother*
650 53:2227–2238.
- 651 57. Carattoli A, Villa L, Fortini D, García-Fernández A. 2021. Contemporary IncI1 plasmids involved in the
652 transmission and spread of antimicrobial resistance in Enterobacteriaceae. *Plasmid*
653 <https://doi.org/10.1016/j.plasmid.2018.12.001>.
- 654 58. Göttig S, Gruber TM, Stecher B, Wichelhaus TA, Kempf VAJ. 2015. In vivo horizontal gene transfer of the
655 carbapenemase OXA-48 during a nosocomial outbreak. *Clinical Infectious Diseases* 60.
- 656 59. Trobos M, Lester CH, Olsen JE, Frimodt-Møller N, Hammerum AM. 2009. Natural transfer of
657 sulphonamide and ampicillin resistance between *Escherichia coli* residing in the human intestine. *Journal of*
658 *Antimicrobial Chemotherapy* 63.
- 659 60. Sandegren L, Linkevicius M, Lytsy B, Melhus Å, Andersson DI. 2012. Transfer of an *Escherichia coli*
660 ST131 multiresistance cassette has created a *Klebsiella pneumoniae*-specific plasmid associated with a
661 major nosocomial outbreak. *Journal of Antimicrobial Chemotherapy* 67.
- 662 61. Hazen TH, Mettus R, McElheny CL, Bowler SL, Nagaraj S, Doi Y, Rasko DA. 2018. Diversity among
663 blaKPC-containing plasmids in *Escherichia coli* and other bacterial species isolated from the same patients.
664 *Sci Rep* 8:10291.
- 665 62. Geurtsen J, de Been M, Weerdenburg E, Zomer A, McNally A, Poolman J. 2022. Genomics and pathotypes
666 of the many faces of *Escherichia coli*. *FEMS Microbiol Rev* 46.
- 667 63. Iseppi R, Di Cerbo A, Messi P, Sabia C. 2020. Antibiotic resistance and virulence traits in vancomycin-
668 resistant enterococci (Vre) and extended-spectrum β -lactamase/ampc-producing (ESBL/ampc)
669 enterobacteriaceae from humans and pets. *Antibiotics* 9.
- 670 64. Chandra M, Cheng P, Rondeau G, Porwollik S, McClelland M. 2013. A single step multiplex PCR for
671 identification of six diarrheagenic *E. coli* pathotypes and *Salmonella*. *International Journal of Medical*
672 *Microbiology* 303:210–216.
- 673 65. Hazen TH, Michalski J, Luo Q, Shetty AC, Daugherty SC, Fleckenstein JM, Rasko DA. 2017. Comparative
674 genomics and transcriptomics of *Escherichia coli* isolates carrying virulence factors of both
675 enteropathogenic and enterotoxigenic *E. coli*. *Sci Rep* 7.
- 676 66. Yu F, Chen X, Zheng S, Han D, Wang Y, Wang R, Wang B, Chen Y. 2018. Prevalence and genetic
677 diversity of human diarrheagenic *Escherichia coli* isolates by multilocus sequence typing. *International*
678 *Journal of Infectious Diseases* 67:7–13.
- 679 67. Neuwirth C, Siebor E, Pechinot A, Duez JM, Pruneaux M, Garel F, Kazmierczak A, Labia R. 2001.
680 Evidence of in vivo transfer of a plasmid encoding the extended-spectrum β -lactamase TEM-24 and other

- 681 resistance factors among different members of the family Enterobacteriaceae. *J Clin Microbiol* 39:1985–
682 1988.
- 683 68. Gosalbes MJ, Vázquez-Castellanos JF, Angebault C, Woerther PL, Ruppé E, Ferrús ML, Latorre A,
684 Andremont A, Moya A. 2016. Carriage of enterobacteria producing extended-spectrum β -lactamases and
685 composition of the gut microbiota in an Amerindian community. *Antimicrob Agents Chemother* 60:507–
686 514.
- 687 69. Le Bastard Q, Chapelet G, Birgand G, Hillmann BM, Javaudin F, Hayatgheib N, Bourigault C, Bemer P, De
688 Decker L, Batard E, Lepelletier D, Montassier E. 2020. Gut microbiome signatures of nursing home
689 residents carrying Enterobacteria producing extended-spectrum β -lactamases. *Antimicrob Resist Infect*
690 *Control* 9.
- 691 70. Peng Y, Liang S, Poonsuk K, On H, Li SW, Pascal Maurin MM, Chan CH, Chan CL, Sin ZY, Tun HM.
692 2021. Role of gut microbiota in travel-related acquisition of extended spectrum β -lactamase-producing
693 Enterobacteriaceae. *J Travel Med* 28:1–11.
- 694 71. Prevel R, Enaud R, Orieux A, Camino A, Sioniac P, M’Zali F, Dubois V, Berger P, Boyer A, Delhaes L,
695 Gruson D. 2023. Bridging gut microbiota composition with extended-spectrum beta-lactamase
696 Enterobacteriales faecal carriage in critically ill patients (microbe cohort study). *Ann Intensive Care* 13.
- 697 72. Piewngam P, Quiñones M, Thirakittiawatthana W, Yungyuen T, Otto M, Kiratisin P. 2019. Composition of
698 the intestinal microbiota in extended-spectrum β -lactamase-producing Enterobacteriaceae carriers and non-
699 carriers in Thailand. *Int J Antimicrob Agents* 53.
- 700 73. Stoppe N de C, Silva JS, Carlos C, Sato MIZ, Saraiva AM, Ottoboni LMM, Torres TT. 2017. Worldwide
701 phylogenetic group patterns of *Escherichia coli* from commensal human and wastewater treatment plant
702 isolates. *Front Microbiol* 8.
- 703 74. Nhu NTQ, Young VB. 2023. The Relationship Between the Microbiome and Antimicrobial Resistance.
704 *Clinical Infectious Diseases* 77:S479–S486.
- 705 75. Benjamini Y, Hochberg Y. 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach
706 to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)* 57:289–300.
- 707 76. Jafari M, Ansari-Pour N. 2019. Why, when and how to adjust your P values? *Cell J* 20.
- 708 77. Glenn TC, Nilsen RA, Kieran TJ, Sanders JG, Bayona-Vásquez NJ, Finger JW, Pierson TW, Bentley KE,
709 Hoffberg SL, Louha S, Garcia-De Leon FJ, Del Rio Portilla MA, Reed KD, Anderson JL, Meece JK,
710 Aggrey SE, Rekaya R, Alabady M, Belanger M, Winker K, Faircloth BC. 2019. Adapterama I: Universal
711 stubs and primers for 384 unique dual-indexed or 147,456 combinatorially-indexed Illumina libraries (iTru
712 & iNext). *PeerJ* 2019.
- 713 78. Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: Resolving bacterial genome assemblies from
714 short and long sequencing reads. *PLoS Comput Biol* 13:1–22.
- 715 79. Treangen TJ, Ondov BD, Koren S, Phillippy AM. 2014. The harvest suite for rapid core-genome alignment
716 and visualization of thousands of intraspecific microbial genomes. *Genome Biol* 15:1–15.
- 717 80. Kaas RS, Leekitcharoenphon P, Aarestrup FM, Lund O. 2014. Solving the problem of comparing whole
718 bacterial genomes across different sequencing platforms. *PLoS One* 9:1–8.
- 719 81. Feldgarden M, Brover V, Gonzalez-Escalona N, Frye JG, Haendiges J, Haft DH, Hoffmann M, Pettengill
720 JB, Prasad AB, Tillman GE, Tyson GH, Klimke W. 2021. AMRFinderPlus and the Reference Gene Catalog
721 facilitate examination of the genomic links among antimicrobial resistance, stress response, and virulence.
722 *Sci Rep* 11.
- 723 82. Johansson MHK, Bortolaia V, Tansirichaiya S, Aarestrup FM, Roberts AP, Petersen TN. 2021. Detection of
724 mobile genetic elements associated with antibiotic resistance in *Salmonella enterica* using a newly
725 developed web tool: MobileElementFinder. *Journal of Antimicrobial Chemotherapy* 76:101–109.
- 726 83. Larsen M V., Cosentino S, Rasmussen S, Friis C, Hasman H, Marvig RL, Jelsbak L, Sicheritz-Pontén T,
727 Ussery DW, Aarestrup FM, Lund O. 2012. Multilocus sequence typing of total-genome-sequenced bacteria.
728 *J Clin Microbiol* 50:1355–1361.
- 729 84. Joensen KG, Tetzschner AMM, Iguchi A, Aarestrup FM, Scheutz F. 2015. Rapid and easy in silico
730 serotyping of *Escherichia coli* isolates by use of whole-genome sequencing data. *J Clin Microbiol* 53:2410–
731 2426.
- 732 85. Roer L, Tchesnokova V, Allesøe R, Muradova M, Chattopadhyay S, Ahrenfeldt J, Thomsen MCF, Lund O,
733 Hansen F, Hammerum AM, Sokurenko E, Hasmana H. 2017. Development of Web Tool for *Escherichia*
734 *coli* Subtyping Based on fimH Alleles. *J Clin Microbiol* 55:2538–2543.
- 735 86. Malberg Tetzschner AM, Johnson JR, Johnston BD, Lund O, Scheutz F. 2020. In Silico Genotyping of
736 *Escherichia coli* Isolates for Sequencing Data. *J Clin Microbiol* 58:e01269-20.

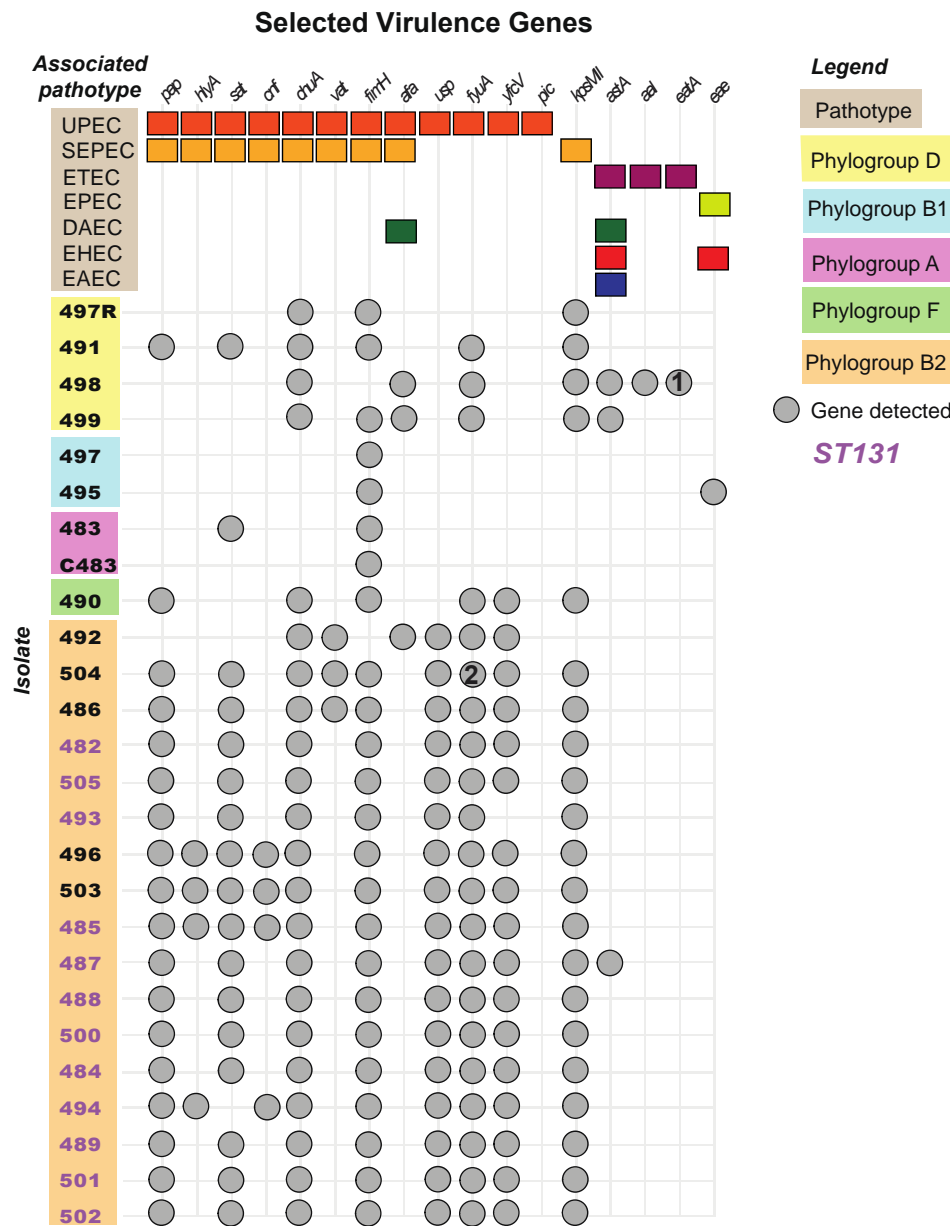
- 737 87. Beghain J, Bridier-Nahmias A, Nagard H Le, Denamur E, Clermont O. 2018. ClermonTyping: An easy-to-
738 use and accurate in silico method for Escherichia genus strain phylotyping. *Microb Genom* 4:1–8.
- 739 88. Tinker KA, Ottesen EA. 2016. The core gut microbiome of the American cockroach, *Periplaneta americana*,
740 is stable and resilient to dietary shifts. *Appl Environ Microbiol* 82:6603–6610.
- 741 89. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. 2016. DADA2: High-resolution
742 sample inference from Illumina amplicon data. *Nat Methods* 13:581–583.
- 743 90. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO. 2013. The SILVA
744 ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res*
745 41:590–596.
- 746 91. Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data
747 with DESeq2. *Genome Biol* 15:1–21.
- 748 92. McMurdie PJ, Holmes S. 2013. Phyloseq: An R Package for Reproducible Interactive Analysis and
749 Graphics of Microbiome Census Data. *PLoS One* 8.
- 750

751



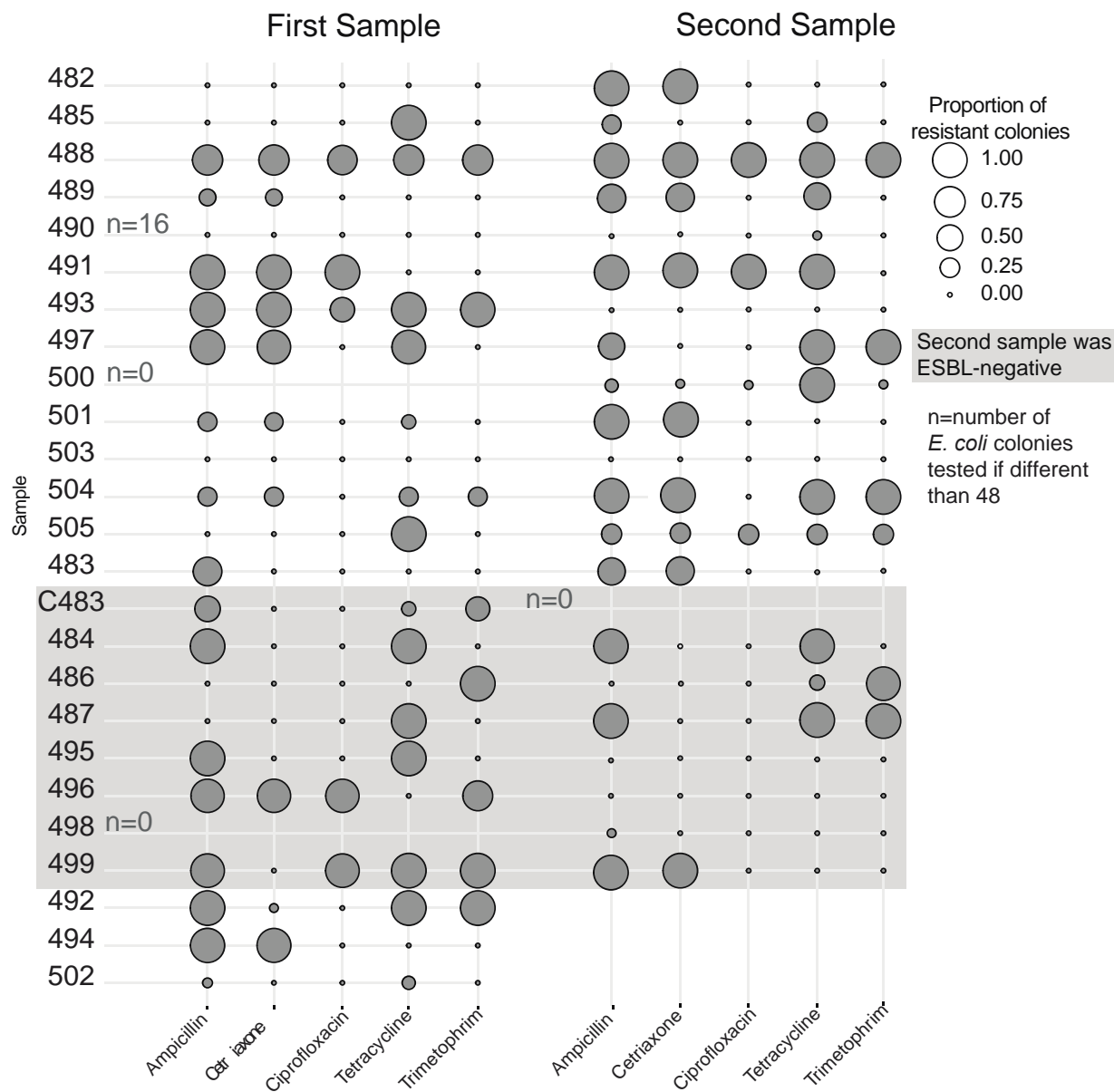
752
753

754 **Figure 1: Characteristics and classification of ESBL-E isolates.** A) Core-genome single nucleotide
755 polymorphisms (SNPs) phylogenetic tree of ESBL-E isolates per participant colored by Clermont
756 phylogroup, highlighting ST131 isolates. B) Phenotypic characterization of antibiotic resistance by isolate
757 tested in Sensititre, after the confirmation as ESBL-E given their resistance to beta-lactams: Cefotaxime
758 and Ceftazidime. C) Antimicrobial resistance genes found by AMRFinderPlus in the genome, transposons
759 (T) or plasmids (1) of each isolate classified by antimicrobial class. D) Beta-lactamase genes found in
760 each isolate, chromosomally or in mobile elements. Circles labelled with a number match plasmids
761 described in Table S5 for each isolate. E) SNPs between first and second sample confirmed as ESBL-
762 positive to determine clonality of isolates. F) Sequence type (ST) classification of each isolate, colored by
763 phylogroup. Second samples collected (R) are not included in the figure as they were the same sequence
764 type (ST) and share almost the same genetic content as their original sample, with the exception of 497R.



765
 766 **Figure 2: Selected virulence genes associated with pathogenic *E. coli* pathotypes.** Virulence genes
 767 were identified by VirulenceFinder using the genome assemblies of each isolate. Only selected virulence
 768 genes are shown based on their associated pathotype. Colored boxes represent the genes associated with
 769 each pathotype and gray circles indicate the presence of that gene in each isolate. Isolates are clustered by
 770 Clermont phylogroups matching Figure 1. Circles labelled with a number match plasmids described in
 771 Table S5 for each isolate.
 772

773



774

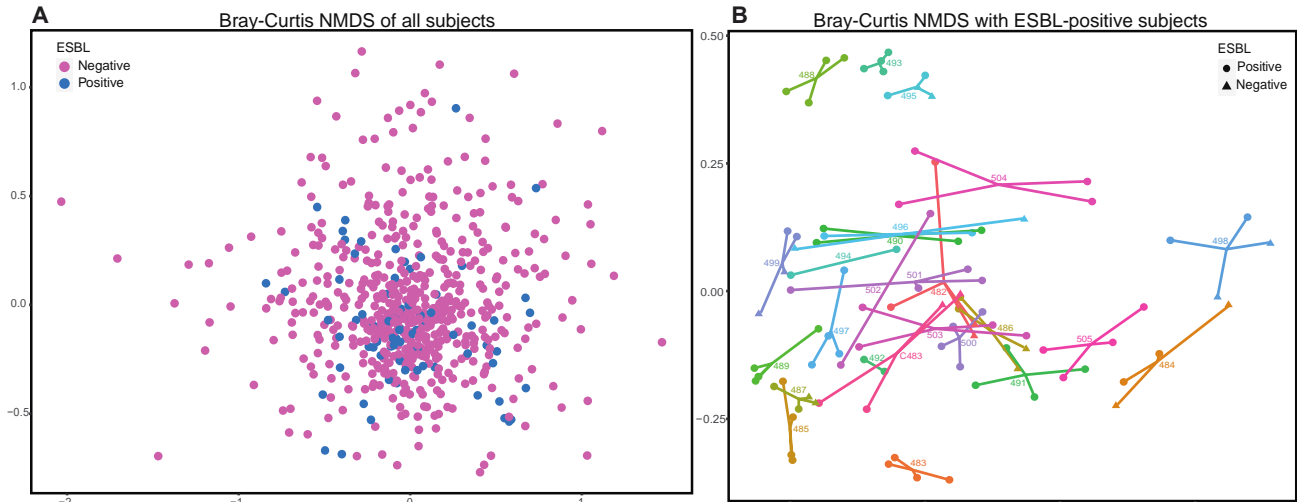
775 **Figure 3: Antibiotic resistance phenotype of commensal *E. coli* isolated from ESBL-E positive**

776 **participants.** A range between 1-48 *E. coli* colonies isolated from ESBL-E positive participants were

777 patched into Mueller Hinton II agar with antibiotics at CSLI standard levels. Each circle size and color

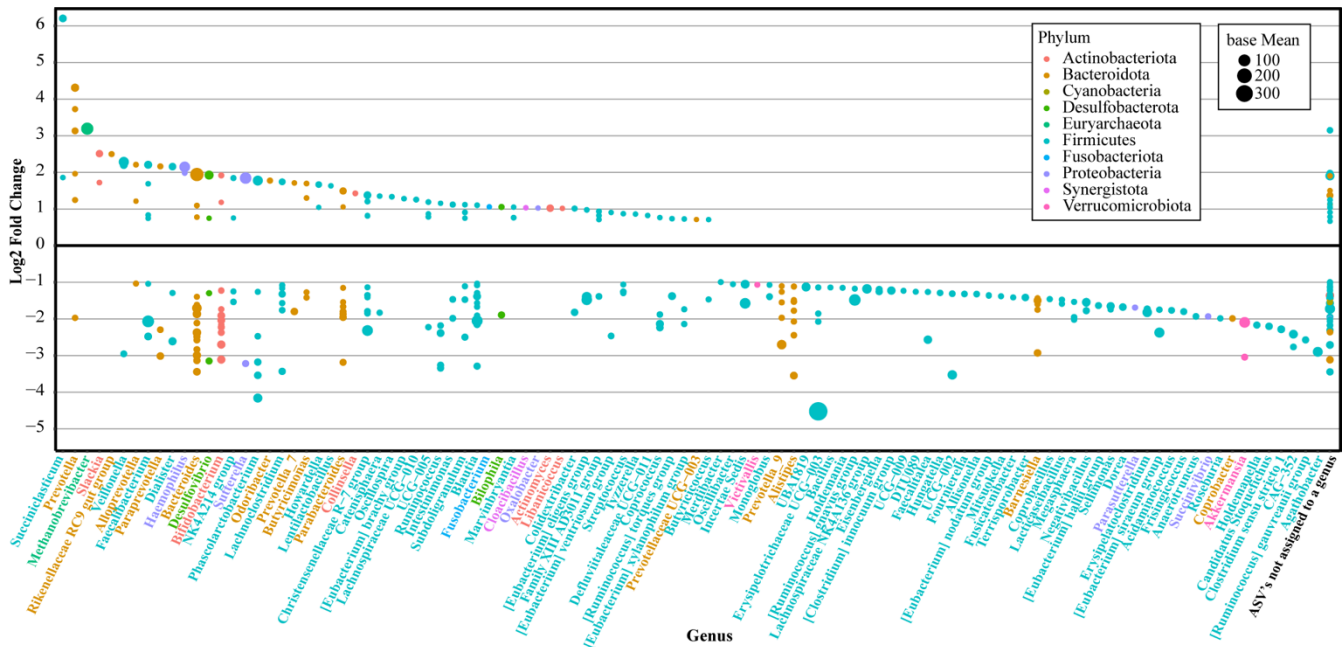
778 intensity show the proportion of those colonies that were resistant to the antibiotic tested.

779



780
781 **Figure 4: Microbial community similarities based on ESBL-E carriage.** A) Bray-Curtis distances of
782 metagenomic microbiome samples from all participants, organized in a non-metric multidimensional
783 scaling (NMDS). Magenta dots represent participants that were ESBL-E negative while blue dots
784 represent participants colonized with ESBL-E. B) Comparison of samples (first, duplicate of first, second
785 and duplicate of the second sample) from the same ESBL-E positive participant, calculated by Bray-
786 Curtis distances. Each number indicates the participant ID, while circles or triangles represent the ESBL-
787 E status by sample.
788

789



790

791

Figure 5: Genus assignments for ASVs identified as enriched or depleted in ESBL-E positive

792

participants. ASVs identified as significantly enriched or depleted by DESeq2 analysis with local fit and

793

Wald test identifying significantly different ASVs under an alpha=0.001. ASVs are colored by phylum

794

and organized by genus. Point size shows normalized mean counts and y axis shows log2fold change in

795

ESBL-E positive participants. ASVs with a positive log2fold change were enriched, while ASVs with a

796

negative log2fold change were depleted in ESBL-E positive participants.

797

798 **Table 1: Selected demographic characteristics, risk factors, and presence of ESBL-E among all adult**
 799 **participants.** The demographic data presented on Table 1 are based on the questionnaire responses from
 800 the first visit of all adult participants. The child demographics were not included in the tables or
 801 statistical analyses of risk factors, as they were not considered independent of the parent participants.

Demographic or Risk Factor	N	ESBL-E positive N=24	ESBL-E negative N=481	p-value ¹ /FDR p-adjusted ²
Biological Sex				
Female	349	13 (54%)	336 (70%)	0.12/0.79
Male	156	11 (46%)	145 (30%)	
Race				
Asian	37	4 (17%)	33 (7%)	0.17/0.79
Black or African American	59	5 (21%)	54 (11%)	
White	377	15 (63%)	362 (75%)	
Age (years)				
18-29	234	13 (54%)	221 (46%)	0.17/0.79
30-39	84	2 (8%)	82 (17%)	
40-49	53	3 (13%)	50 (10%)	
50-59	49	2 (8%)	47 (10%)	
60-69	45	1 (4%)	44 (9%)	
70-79	22	2 (8%)	20 (4%)	
80-89	5	0 (0%)	5 (1%)	
Antibiotic use in the past 1 month (oral, topical, or intravenous)				
Yes	47	2 (8%)	45 (9%)	1/1
No	454	22 (92%)	432 (90%)	
Gastrointestinal conditions or symptoms in the past 1 month				
Yes	177	8 (33%)	169 (35%)	1/1
No	327	16 (67%)	311 (65%)	
Urinary tract infection in the past 1 month (self-diagnosed or diagnosed by doctor)				
Yes	13	1 (4%)	12 (2%)	0.47/0.79
No	492	23 (96%)	469 (98%)	
International travel (past year)				
Yes	31	1 (4%)	30 (6%)	1/1
No	471	23 (96%)	448 (93%)	
Lived internationally in the last 5 years				
Yes	37	4 (17%)	33 (7%)	0.09/0.79
No	463	20 (83%)	443 (92%)	
Exposure to treated recreational water in the past month				
Yes	127	4 (17%)	123 (26%)	0.47/0.79
No	371	20 (83%)	351 (73%)	
Exposure to untreated recreational water in the past month				
Yes	107	3 (13%)	104 (22%)	0.44/0.79
No	391	21 (88%)	370 (77%)	

802 ¹p-values were calculated using 'fisher.test' function in R with Monte-Carlo simulations. ²FDR adjusted p-values
 803 were calculated using the 'p.adjust' function in R with the Benjamini-Hochberg method.

804
 805

806 **Table 2: Assembled plasmids carried by each isolate identified by PlasmidFinder with encoded**
 807 **antibiotic genes identified by AMRFinderPlus and encoded virulence genes identified by**
 808 **VirulenceFinder. Plasmid number in each isolate matches the numbers in panels C and D of Figure 1**
 809 **and Figure 2.**

<i>Isolate^a</i>	Plasmid Type	Plasmid-associated Antibiotic Resistance Genes	Virulence Genes
497R (n=4)	1. IncII 2. IncFII	1. <i>bla_{CTX-M-1}</i>	1. <i>cia</i> 2. <i>mcbA, traJ, traT</i>
491 (n=2)	-		
498 (n=3)	1. IncFII 2. IncFIB	2. <i>bla_{CTX-M-27}</i>	1. <i>aap, eatA, faeF, traT</i> 2. <i>traJ, traT</i>
499 (n=0)	-		
497 (n=2)	1. IncF 2. IncII	1. <i>tetA, tetR</i> 2. <i>bla_{CTX-M-1}</i>	1. <i>anr, ompT</i> 2. <i>cia</i>
495 (n=2)	1. IncF 2. IncF	1. <i>bla_{CTX-M-15}, bla_{OXA-1}, bla_{TEM-1}, aph(3'')-Ib, aph(6)-Id, aac(6')-Ib-cr5, catB3, floR, qacEdelta1, qnrS1, sul1, sul2, terB, terC, terD, terE, tet(A), dfrA1</i>	1. <i>terC</i> 2. <i>anr, traJ, traT</i>
483 (n=3)	1. IncFII 2. IncB/O/Z/K	1. <i>bla_{CTX-M-15}, qnrS1</i>	1. <i>traT</i> 2. <i>ireA</i>
C483 (n=0)	-		
490 (n=3)	1. IncF	1. <i>tet(B)</i>	1. <i>anr, iucC, iutA, sitA, traT</i>
492 (n=6)	1. IncF 2. IncFII	1. <i>bla_{TEM-1}, aph(3'')-Ib, aph(6)-Id, sul2, tet(B)</i> 2. <i>bla_{TEM-1}</i>	1. <i>anr, iucC, iutA, sitA, traT</i> 2. <i>mcbA, traJ, traT</i>
504 (n=3)	1. IncF 2. IncFII	1. <i>bla_{CTX-M-27}, aadA5, aph(3'')-Ib, aph(6)-Id, erm(B), mph(A), qacEdelta1, sul1, sul2, tet(A), dfrA17</i>	1. <i>senB</i> 2. <i>fyuA, traJ</i>
486 (n=3)	1. IncF	1. <i>bla_{CTX-M-27}</i>	1. <i>senB</i>
482 (n=3)	1. IncF 2. Col156		1. <i>traJ, traT</i> 2. <i>senB</i>
505 (n=1)	1. IncF	1. <i>bla_{CTX-M-27}, aadA5, aph(3'')-Ib, aph(6)-Id, mph(A), qacEdelta1, sul1, sul2, tet(A), dfrA17</i>	1. <i>anr, senB, traT</i>
493 (n=2)	1. IncF 2. IncF	1. <i>bla_{CTX-M-27}, aadA5, aph(3'')-Ib, aph(6)-Id, mph(A), qacEdelta1, sul1, sul2, tet(A), dfrA17</i>	1. <i>anr, senB, traT</i> 2. <i>traT</i>
496 (n=1)	1. IncF	1. <i>bla_{TEM-1}, aph(3'')-Ib, aph(6)-Id, qnrS1, sul2, dfrA14</i>	1. <i>anr, traJ, traT</i>
503 (n=1)	1. IncF	1. <i>tet(A)</i>	1. <i>iutA, senB</i>
485 (n=2)	1. IncF		1. <i>anr, iutA, senB, traJ, traT</i>
487 (n=2)	1. IncF	1. <i>aph(3'')-Ib, aph(6)-Id, sul2, tet(A)</i>	1. <i>anr, senB</i>
488 (n=3)	1. IncFI 2. IncFII	1. <i>bla_{TEM-1}, aadA5, aph(3'')-Ib, aph(6)-Id, mph(A), qacEdelta1, sul1, sul2, tet(A), dfrA17</i>	1. <i>traJ, traT</i> 2. <i>mcbA, traT</i>
500 (n=3)	1. IncF	1. <i>bla_{TEM-1}, aac(3)-IId, aadA5, aph(3'')-Ib, aph(6)-Id, mph(A), qacEdelta1, sul1, sul2, tet(A), dfrA17</i>	1. <i>senB, traJ, traT</i>
484 (n=3)	1. IncF 2. IncFII	1. <i>bla_{TEM-1}, aac(3)-IId, aph(3'')-Ib, aph(6)-Id, sul2, tet(B), dfrA17</i>	1. <i>anr</i> 2. <i>traT</i>
494 (n=2)	1. IncF	1. <i>bla_{TEM-1}</i>	1. <i>anr, senB, traT</i>
489 (n=2)	1. IncF	1. <i>bla_{CTX-M-27}</i>	1. <i>anr, senB, traT</i>
501 (n=3)	1. IncF 2. IncII	1. <i>bla_{CTX-M-27}</i>	1. <i>anr, senB, traT</i> 2. <i>cia</i>
502 (n=2)	1. IncF	1. <i>bla_{CTX-M-27}</i>	1. <i>anr, senB</i>

810 ^an represents total number of plasmids contigs associated with each isolate, as listed in Table S5.