Genomic surveillance of antimicrobial-resistant *Escherichia coli* in fecal sludge and sewage in Uganda

Ryota Gomi[a, *], Yasufumi Matsumura[b], Masaki Yamamoto[b], Mai Tanaka[a], Allan John Komakech[c], Tomonari Matsuda[a], Hidenori Harada[d, *]

[a]Department of Environmental Engineering, Graduate School of Engineering, Kyoto University, Katsura, Nishikyo-ku, 615-8540, Kyoto, Japan

[b]Department of Clinical Laboratory Medicine, Graduate School of Medicine, Kyoto University, 54 Shogoin-kawahara-cho, Sakyo-ku, 606-8507, Kyoto, Japan

[c]Department of Agricultural and Biosystems Engineering, College of Agricultural and Environmental Sciences, Makerere University, P.O. Box 7062, Kampala, Uganda

[d]Graduate School of Asian and African Area Studies, Kyoto University, Yoshida-Shimoadachi, Sakyo-ku, 606-8501, Kyoto, Japan

*Corresponding authors

E-mail address: gomi.ryota.34v@kyoto-u.jp; Tel: +81-75-383-3354

E-mail address: harada.hidenori.8v@kyoto-u.ac.jp; Tel: +81-75-753-7830

18 **ABSTRACT**

19 The global increase of antimicrobial resistance (AMR) is a major public health concern. An
20 effective AMR surveillance tool is needed to track the emergence and spread of AMR.
21 Wastewater surveillance has been proposed as a resource-efficient tool for monitoring AMR
22 carriage in the community. Here, we performed genomic surveillance of antimicrobial-resistant
23 *Escherichia coli* obtained from fecal sludge and sewage in Uganda to gain insights into *E. coli*
24 epidemiology and AMR burden in the underlying population. Selective media containing
25 different antibiotic combinations (cefotaxime, ciprofloxacin, cefotaxime + ciprofloxacin +
26 gentamicin) were used to obtain antimicrobial-resistant *E. coli* from fecal sludge and sewage.
27 Short-read sequencing was performed for the obtained isolates, and a subset of isolates
28 (selected from predominant sequence types (STs)) was also subjected to long-read sequencing.
29 Genomic analysis of the obtained *E. coli* isolates (n = 181) revealed the prevalence of clonal
30 complex 10, including ST167 (n = 43), ST10 (n = 28), ST1284 (n = 17), and ST617 (n = 4), in
31 both fecal sludge and sewage, irrespective of antibiotics used for selection. We also detected
32 global high-risk clones ST1193 (n = 10) and ST131 (n = 2 clade A, n = 3 subclade C1-M27,
33 and n = 1 subclade C2). Diverse AMR determinants, including extended-spectrum β-lactamase
34 genes (mostly $bla_{\text{CTX-M-15}}$) and mutations in *gyrA* and *parC*, were identified. Analysis of the
35 completed genomes revealed that diverse IncF plasmids and chromosomal integration were the
36 major contributors to the spread of AMR genes in the predominant STs. This study showed

that a combination of sewage surveillance (or fecal sludge surveillance) and whole-genome sequencing can be a powerful tool for monitoring AMR carriage in the underlying population.

Keywords: *Escherichia coli*, Antimicrobial resistance, One Health, Sewage surveillance, Whole-genome sequencing

43 **1. Introduction**

44 The increase of antimicrobial resistance (AMR) is a global health concern. AMR can lead to
45 adverse consequences such as increased mortality rates, prolonged hospital admissions, and
46 higher medical costs (Majumder et al., 2020). Information on AMR collected through
47 surveillance is needed to guide treatment, monitor the effectiveness of interventions, and
48 develop polices and regulations to counter AMR (WHO, 2015). However, surveillance of AMR
49 in resource-limited settings is challenging, especially when the surveillance targets populations.

50 Sewage contains pooled urine and feces from a large population, and one sewage sample can
51 represent bacteria from multiple people in the community. For this reason, sewage surveillance
52 (or wastewater-based epidemiology) has been proposed as a resource-efficient tool for
53 monitoring AMR carriage in the underlying population (Aarestrup and Woolhouse, 2020;
54 Larsson et al., 2022). Sewage surveillance is attractive in many aspects; for example, (i) sewage
55 provides a sample from a large and mostly healthy population, which otherwise would not be
56 easily monitored, and (ii) it does not collect data on individuals and thus reduces ethical risks
57 (Hendriksen et al., 2019). Although it is a promising approach, sewage surveillance of AMR
58 cannot be directly applied in areas with no connection to centralized (sewered) systems. A
59 previous study reported at least 1.8 billion people in low- and middle-income countries use on-
60 site sanitation systems like pit latrines (Berendes et al., 2017). In such areas, fecal sludge from
61 on-site systems has been proposed to be used as pooled community samples for surveillance

62 purposes (Capone et al., 2020; Chigwechokha et al., 2022). Sewage surveillance (or fecal
63 sludge surveillance) of AMR can be either culture-dependent or culture-independent. Culture-
64 independent methods (e.g., short-read metagenomics) have some advantages over culture-
65 dependent methods, such as the ability to detect thousands of AMR genes in a single sample.
66 However, it provides limited information on the host or the genetic context of AMR genes
67 (Hendriksen et al., 2019). Culture-dependent methods, on the other hand, can link AMR genes
68 to specific host bacteria, provide information on phylogenetic characteristics on the host, and
69 may even provide information on the location (chromosome or plasmid) and genetic context
70 of AMR genes.

71 *Escherichia coli* is a commensal bacterium in humans and animals and is also known to cause
72 serious infections (Denamur et al., 2021). Although *E. coli* is intrinsically susceptible to almost
73 all clinically relevant antibiotics, it can accumulate AMR genes via, e.g., horizontal gene
74 transfer (Poirel et al., 2018). *E. coli* has been suggested as an indicator for AMR surveillance
75 in the environment (Anjum et al., 2021), and previous studies have shown that *E. coli* is a good
76 target organism for sewage surveillance of AMR (Huijbers et al., 2020; Hutinel et al., 2019;
77 Kwak et al., 2015; Raven et al., 2019). WHO also selected extended-spectrum β-lactamase
78 (ESBL)-producing *E. coli* as an indicator for "One Health" AMR surveillance (including
79 sewage surveillance) in the Tricycle protocol (WHO, 2021). The protocol recommends usage
80 of whole-genome sequencing (if possible) to determine the sequence type (ST), virulence

81 genotype, plasmid type, carriage of acquired AMR genes, and phylogenetic characteristics of
82 the obtained isolates, which contribute to the understanding of the epidemiology of pathogens.

83 Uganda is a country experiencing an extremely high prevalence of AMR (Nabadda et al., 2021).
84 Previous studies have reported high prevalence of multidrug resistance (MDR) among *E. coli*
85 (Katongole et al., 2020; Odongo et al., 2020). However, most studies were based on phenotypic
86 characterization or detection of a limited set of AMR genes/virulence genes, and thus their
87 genomic characteristics, including the clonal composition and resistome, are largely unknown.
88 In the present study, we isolated antimicrobial-resistant *E. coli* from fecal sludge and sewage
89 in Uganda and performed genomic analysis on the obtained isolates, aiming to gain insights
90 into *E. coli* epidemiology and AMR burden in the underlying population.

91

92 **2. Materials and methods**

93 **2.1. Sample collection and isolation of *E. coli***

94 Samples were collected at a fecal sludge treatment plant (FSTP) and a wastewater treatment
95 plant (WWTP) in Kampala, Uganda, in October–November, 2018. We obtained samples from
96 both sources to cover a wider population in this region. For fecal sludge sampling, three grab
97 samples were collected at different times on each of five days from effluent of the primary
98 sedimentation tank in the FSTP; in total 15 samples were collected (these samples are hereafter

99 referred to simply as fecal sludge samples). For sewage sampling, up to three grab samples
100 were collected at different times over six days from influent of the WWTP, totaling 16 samples.
101 The samples were collected and stored in sterile sampling bottles and transported on ice to the
102 laboratory. The samples were then cultured, within 12 hours after collection, using a specific
103 enzyme substrate medium (CHROMagar ECC, Kanto Chemical, Tokyo, Japan) containing
104 four combinations of three antibiotics: (a) no antibiotics, (b) cefotaxime (CTX, 1 mg/L), (c)
105 ciprofloxacin (CIP, 1 mg/L), (d) CTX (1 mg/L) + CIP (1 mg/L) + gentamicin (GM, 4 mg/L).
106 The samples were serially diluted with PBS for colony counting. Up to three colonies with an
107 *E. coli* profile (blue colony) were picked from each of plate (b), (c), and (d) (i.e., we obtained
108 up to nine isolates from a single sample), and each was transferred into a screw tube containing
109 LB agar. The tubes were incubated overnight at 37 °C, stored at 4 °C, and then transported to
110 the laboratory in Japan. *E. coli* in each screw tube was again incubated on a plate with
111 CHROMAgar ECC containing the same antibiotic(s) as at the time of isolation. A colony with
112 an *E. coli* profile was picked and spread again on a plate with CHROMAgar ECC, and then a
113 single colony was picked from the plate, cultured in LB broth, and stored at –85 °C in 35%
114 glycerol.

115 **2.2. Genome sequencing and assembly**

116 A total of 192 presumptive antimicrobial-resistant *E. coli* isolates was subjected to genome
117 sequencing as below. DNA was extracted from each isolate using a QIAamp DNA Mini Kit

118 (Qiagen, Hilden, Germany) for Illumina sequencing. Illumina libraries were prepared using a
119 Nextera XT DNA Library Preparation Kit (Illumina, San Diego, CA) and sequenced on the
120 NextSeq550 platform, generating paired-end reads of 151 bp each. For some isolates, read
121 depth was not sufficient, so we sequenced the library again on the MiniSeq platform to obtain
122 additional paired-end reads of the same length. Illumina reads were processed using fastp
123 (v0.23.2) (Chen et al., 2018) and assembled using Unicycler (v0.4.8) with the --no_correct
124 option (Wick et al., 2017). The quality of each genome assembly was ensured by using the
125 quality control (QC) criteria of (i) contig number: ≤ 800; (ii) genome size: 3.7 Mbp to 6.4 Mbp;
126 and (iii) N50: >20 kb, as recommended by EnteroBase (Zhou et al., 2020).

127 For selected isolates (n = 26), we also performed Oxford Nanopore Technologies (ONT) long-
128 read sequencing to generate complete genomes. DNA was extracted from each isolate using a
129 DNeasy PowerSoil Pro Kit (Qiagen). The long-read sequencing library was prepared using the
130 SQK-LSK109 kit (Oxford Nanopore Technologies, Oxford, UK) and sequenced on the
131 MinION with a FLO-MIN106 flow cell. ONT reads were filtered using Filtlong (v0.2.1,
132 https://github.com/rrwick/Filtlong) to remove reads with low quality or short length. Hybrid
133 assembly of Illumina short reads and ONT long reads was performed by employing a long-
134 read-first approach. Briefly, filtered ONT reads were assembled using Flye (v2.9.1-b1780)
135 (Kolmogorov et al., 2019), and then the assemblies were polished with medaka (v1.7.1,
136 https://github.com/nanoporetech/medaka) using filtered ONT reads and Polypolish (v0.5.0)

137 using processed Illumina reads (Wick and Holt, 2022). Hybrid assembly was also performed
138 using a short-read-first approach, i.e., Unicycler (v0.5.0) hybrid assembly followed by
139 polishing with Polypolish (v0.5.0) using processed Illumina reads. If the long-read-first
140 approach generated a complete genome for an isolate, we chose the polished Flye assembly as
141 the final assembly. If the long-read-first approach did not complete the genome but the short-
142 read-first approach did, we chose the polished Unicycler assembly as the final assembly.

143 **2.3. Genomic analysis**

144 *E. coli* status was confirmed by performing an average nucleotide identity (ANI) analysis using
145 each genome sequenced in the present study and the genome sequence of the *E. coli* type strain
146 ATCC 11775 (= DSM 30083 = U5/41) (Wadley et al., 2019). ANI analysis was performed
147 using FastANI (v1.33) with the threshold of 95% (Jain et al., 2018). *E. coli* phylogroups were
148 determined using a combination of EzClermont (v0.7.0) and ClermonTyping (Beghain et al.,
149 2018; Waters et al., 2020). Multilocus sequence typing (MLST) was performed using MLST
150 (v2.19.0) (https://github.com/tseemann/mlst). *fimH* types, which are used for high-resolution
151 subtyping of MLST-based *E. coli* clonal groups, were determined using FimTyper 1.0 (Roer et
152 al., 2017). AMR genes were detected using a combination of ResFinder 4.1 (Bortolaia et al.,
153 2020) and Abricate (v1.0.1) (https://github.com/tseemann/abricate) with the ncbi database.
154 Genetic contexts of AMR genes were analyzed using the RAST server (Aziz et al., 2008), the
155 BLAST algorithm (https://blast.ncbi.nlm.nih.gov/Blast.cgi), and ISfinder (Siguier et al., 2006).

156 Mutations in quinolone resistance-determining regions (QRDRs) in *gyrA* (GyrA codons 83 and
157 87) and *parC* (ParC codons 80 and 84) were analyzed as described previously (Aoike et al.,
158 2013). Virulence genes were detected using the VirulenceFinder database (Joensen et al., 2014;
159 Malberg Tetzschner et al., 2020). *E. coli* pathotypes, which include extraintestinal pathogenic
160 *E. coli* (ExPEC) and several categories of intestinal pathogenic *E. coli* (InPEC), were defined
161 based on the presence of specific virulence genes as described previously (Gomi et al., 2015).
162 K1 and K5 capsular types of ST1193 isolates were determined based on blastn analysis using
163 M57382 (K1) and X53819 (K5) as references (Johnson and Stell, 2000). Plasmid replicons
164 were detected using PlasmidFinder 2.1, and plasmids were typed using pMLST 2.0 (Carattoli
165 et al., 2014).

166 A whole-genome SNP-based tree was built for isolates belonging to phylogroup A using
167 kSNP3 based on SNPs occurring in at least 90% of the genomes (Gardner et al., 2015). To
168 confirm the topology of the tree, we also built a tree for the phylogroup A genomes using
169 another software, parsnp (v1.7.4) (Treangen et al., 2014). For isolates belonging to clonal
170 complex 10 (CC10) within phylogroup A, we built a tree together with global CC10 genomes
171 (n = 4,744) obtained from the NCBI Reference Sequence (RefSeq) database. For isolates
172 belonging to ST1193 and ST131, we separately built a recombination-free phylogenetic tree
173 for each ST. To place our isolates in broader phylogenetic contexts, we also included isolates
174 analyzed in other studies which defined clades/subclades within each of ST1193 and ST131

175 (Johnson et al., 2019; Matsumura et al., 2017b). Detailed methods used for tree construction of
176 the CC10, ST1193, and ST131 genomes are provided in the Supplementary Materials and
177 methods.

178 **2.4. Accession numbers**

179 Genome assemblies were deposited in GenBank under BioProject PRJNA956323 (also see
180 **Table S1** for the accession number of each genome and **Table S2** for the accession number of
181 each replicon for the completed genomes).

182

183 **3. Results and discussion**

184 **3.1. Detection and isolation of antimicrobial-resistant *E. coli***

185 During the study period, we collected 15 fecal sludge samples and 16 sewage samples and
186 measured the concentrations of total *E. coli* and antimicrobial-resistant *E. coli* (**Table S3**). The
187 mean proportions of antimicrobial-resistant *E. coli* in the fecal sludge samples were 30.0% for
188 CTX, 26.9% for CIP, and 10.7% for CTX + CIP + GM. The mean proportions of antimicrobial-
189 resistant *E. coli* in the sewage samples were 39.2% for CTX, 21.0% for CIP, and 5.1% for CTX
190 + CIP + GM. These proportions were higher than those reported previously for *E. coli* in
191 wastewater influent in some developed countries; for example, the proportions of
192 antimicrobial-resistant *E. coli* were 0%–3.2% for CTX and 3.2%–7.1% for CIP in Sweden

193 (Hutinel et al., 2019) and 1.6%–4.4% for CTX (or ceftazidime) and 2.4%–13.4% for CIP in
194 ten European countries (Huijbers et al., 2020). On the other hand, the proportions were
195 comparable to those reported in some other low- and middle-income countries (LMICs); for
196 example, the resistance proportions were 38% for CTX and 35% for CIP in South Africa
197 (Gumede et al., 2021) and 24% for CIP in Vietnam (Le et al., 2023), potentially highlighting
198 the extremely high prevalence of AMR in Uganda and some LMICs (Nabadda et al., 2021;
199 WHO, 2022).

200 In total, we obtained 192 presumptive antimicrobial-resistant *E. coli* isolates from the samples
201 and performed short-read sequencing on these isolates. After discarding contigs shorter than
202 100 bp, 187 genomes met the assembly QC criteria. Species identification of the 187 genomes
203 by the ANI analysis revealed that 181 belonged to *E. coli*, while 6 belonged to non-*E. coli*
204 species such as *Citrobacter freundii*. These 181 *E. coli* genomes, comprising 91 from fecal
205 sludge samples (31 from CTX-containing media, 31 from CIP-containing media, and 29 from
206 CTX + CIP + GM-containing media) and 90 from sewage samples (32 from CTX-containing
207 media, 26 from CIP-containing media, and 32 from CTX + CIP + GM-containing media), were
208 further analyzed in the present study.

209 **3.2. Clonal composition**

210 In total, 36 STs from six different phylogroups were detected among the 181 isolates (**Table**
211 **S1**). Phylogroup A was the most prevalent (n = 127), followed by B1 (n = 17), B2 (n = 17), C

212 (n = 10), D (n = 8), and F (n = 2). ST167 (n = 43), ST10 (n = 28), ST46 (n = 19), and ST1284
213 (n = 17), all of which belong to phylogroup A, were prevalent in both fecal sludge isolates and
214 sewage isolates (**Figure 1a**). In total, 144 isolates (79.6%) were from the top ten STs, while 37
215 isolates (20.4%) were from minor STs (i.e., STs that are not in the top ten and detected among
216 three or fewer isolates) (**Figures 1b** and **1c**). The top ten STs were generally associated with
217 two or three types of media and both fecal sludge and sewage, with a few exceptions (e.g.,
218 ST131 was exclusively found among fecal sludge isolates). Lineages associated with
219 extraintestinal infections were also among the top ten STs and detected in fecal sludge isolates
220 (n = 9 ST1193, n = 5 ST410, n = 6 ST131, n = 2 ST69) and sewage isolates (n = 1 ST1193, n
221 = 4 ST410, n = 2 ST69) (Denamur et al., 2021). Thirty-four isolates (18.8%), including those
222 from the above-mentioned STs associated with extraintestinal infections, were classified as
223 ExPEC based on the virulence gene contents (**Figure 1**). On the other hand, none of the isolates
224 were classified as InPEC. ExPEC strains are facultative pathogens and part of the intestinal
225 microflora of a fraction of the healthy population (Kohler and Dobrindt, 2011), which might
226 explain the relatively high prevalence of ExPEC among our *E. coli* collection.
227 We next compared the clonal composition of *E. coli* in this study and those reported previously
228 for clinical *E. coli* in Uganda. However, the clonal composition of *E. coli* in clinical settings is
229 largely unknown in Uganda (a search of human-derived *E. coli* genomes from Uganda in
230 EnteroBase returned no hits as of May 2023), and we could find only one study reporting the

231 ST composition of clinical *E. coli* isolates in Uganda. That study sequenced 24 Ugandan *E.*
232 *coli* isolates causing urinary tract infections (most were MDR) and reported the prevalence of
233 ST10 (n = 4, 15.4%) as well as the occurrence of ST410 (n = 2, 8.3%), ST131 (n = 1, 4.2%),
234 ST167 (n = 1, 4.2%), ST617 (n = 1, 4.2%), and ST1193 (n = 1, 4.2%) (Decano et al., 2021).
235 These STs were also among the top ten STs in the present study, indicating that sewage
236 surveillance may serve as a possible resource-efficient complement to clinical surveillance,
237 though further studies are needed to confirm this.
238 Previous studies have suggested that fecal sludge and sewage can be used for community
239 surveillance, but fecal sludge includes waste from fewer individuals and hence may be less
240 divergent than sewage (Capone et al., 2020). However, the number of unique STs was larger
241 for fecal sludge isolates than sewage isolates (26 unique STs vs 23 unique STs) (**Table S1**).
242 Moreover, Simpson's diversity index was higher for fecal sludge isolates than sewage isolates
243 when isolates from the same type of media were compared (**Figure 1a**). This might be because
244 of the fecal sludge sampling strategy employed in the present study: we obtained samples from
245 effluent of the primary sedimentation tank in a FSTP, which probably represents fecal sludge
246 from a large number of individuals. This might have contributed to the relatively high diversity
247 of *E. coli* isolates from fecal sludge samples, suggesting sampling from FSTPs or otherwise
248 designated fecal sludge dumping sites (rather than individual pit latrines) may better capture

249 the population-scale information in fecal sludge surveillance, when such places exist in the

250 region.

251 **3.3. Antimicrobial resistance determinants**

252 Six different ESBL genes, namely $bla_{\text{CTX-M-3}}$, $bla_{\text{CTX-M-14}}$, $bla_{\text{CTX-M-15}}$, $bla_{\text{CTX-M-27}}$, $bla_{\text{CTX-M-55}}$, and

253 $bla_{\text{TEM-169}}$, were detected among the isolates, with $bla_{\text{CTX-M-15}}$ being the predominant type

254 (**Figure 2a**). Based on the short-read assemblies, there were 21 isolates with two different

255 ESBL genes (n = 18 $bla_{\text{CTX-M-14}}$ + $bla_{\text{CTX-M-15}}$, n = 1 $bla_{\text{CTX-M-15}}$ + $bla_{\text{CTX-M-27}}$, n = 1 $bla_{\text{CTX-M-3}}$ +

256 $bla_{\text{CTX-M-27}}$, and n = 1 $bla_{\text{TEM-169}}$ + $bla_{\text{CTX-M-15}}$) (note that isolates carrying multiple copies of the

257 same ESBL gene were not considered here because those genes can be collapsed into one

258 contig in short-read assemblies). As expected, all but one isolate obtained from CTX-

259 containing media and CTX + CIP + GM-containing media carried at least one ESBL gene.

260 Importantly, many isolates obtained from CIP-containing media carried ESBL genes (41.9%

261 for fecal sludge isolates and 42.3% for sewage isolates), indicating the convergence of ESBL

262 genes and fluoroquinolone resistance.

263 All but four isolates obtained from CIP-containing media and CTX + CIP + GM-containing

264 media carried three or four mutations in QRDRs (**Figure 2b**). In all isolates with three

265 mutations in QRDRs, two were in codons 83 and 87 in GyrA, and the remaining mutation was

266 in ParC codon 80 or 84. These are congruent with previous studies reporting that two mutations

267 in the QRDR of *gyrA* and one or two mutations in the QRDR of *parC* are the main mechanism

268 of ciprofloxacin resistance in *E. coli* (Matsumura et al., 2017a; van der Putten et al., 2019).
269 Importantly, most isolates obtained from CTX-containing media also carried three or four
270 mutations in QRDRs (87.1% for fecal sludge isolates and 75.0% for sewage isolates), again
271 highlighting the convergence of ESBL genes and fluoroquinolone resistance determinants.

272 Gentamicin resistance in isolates obtained from CTX + CIP + GM-containing media could be
273 mostly explained by the presence of *aac(3)-II* genes (**Table S1**). Other resistance genes not
274 associated with antibiotics used in the selective media were also prevalent in our isolates; for
275 example, *sul* genes were detected in 175 isolates (96.7%) and *tet* genes were detected in 161
276 isolates (89.0%). However, genes conferring resistance to last-resort antibiotics, such as
277 carbapenems and colistin, were not detected in any isolates (**Table S1**). We did not use last-
278 resort antibiotics to select *E. coli* isolates in the present study, but we note that large-scale
279 genomic surveillance of isolates resistant to last-resort antibiotics should also be performed in
280 LMICs in future studies.

281 **3.4. Genomic characteristics of phylogroup A isolates**

282 Whole genome-based phylogenetic analysis was performed using kSNP3 to gain insights into
283 clonal relationships of 127 isolates belonging to phylogroup A, which was the dominant
284 phylogroup in the present study (**Figure 3**). The tree was largely separated according to STs,
285 though we observed separate clusters for ST10, ST46, and ST167. These separate clusters were

286 also confirmed in a tree generated using different software, parsnp (**Figure S1**). The ST10 and
287 ST167 clusters were characterized by distinct *fimH* types, namely *fimH54*, *fimH435*, and
288 *fimH789* for ST10, and *fimH5*, *fimH29*, and *fimH54* for ST167. ST46 isolates from both
289 clusters did not carry *fimH*.

290 According to EnteroBase (Zhou et al., 2020), ST167, ST10, ST1284, and ST617 are all
291 classified into CC10. Classifying all isolates from a monophyletic clade containing these four
292 STs (the bottom clade in **Figure 3**) into CC10 identified that 99 isolates (54.7% of isolates in
293 the present study) belong to this lineage, indicating CC10 poses a significant AMR burden in
294 the underlying population. CC10 has been globally reported as a resident of the intestinal tract
295 of humans and other animals, and carriage of AMR genes (including ESBL genes) has also
296 been widely reported in this lineage (Reid et al., 2019). To determine the relationships between
297 CC10 isolates in this study and global CC10 isolates, we built a Mash-based tree using our
298 CC10 genomes and RefSeq CC10 genomes (**Figure S2** and **Table S4**). We found that our
299 CC10 genomes formed clusters with global CC10 genomes (e.g., ST617-ST1284 cluster and
300 ST167 cluster in **Figure S2**). On the other hand, some of the top 15 STs detected in global
301 CC10 genomes (e.g., ST48 and ST746) were not detected in our CC10 genomes. We also found
302 that our CC10 genomes were not evenly distributed in the tree, while African RefSeq CC10
303 genomes were scattered throughout the tree. These findings indicate that only a subset of global
304 clones within CC10 were prevalent in our study setting.

305 Although there were some small clusters specific to fecal sludge isolates (e.g., ST10 with
306 *fimH789*) and sewage isolates (e.g., a cluster containing six sewage isolates within ST46), fecal
307 sludge isolates and sewage isolates were generally intermingled in the tree (**Figure 3**). This
308 trend was also seen for media types; isolates selected with different antibiotics were generally
309 intermingled in the tree. Most phylogroup A isolates (except six isolates from minor STs
310 obtained using CTX-containing media) carried three or four mutations within QRDRs. As
311 mentioned above, ESBL genes were also prevalent even in isolates obtained using CIP-
312 containing media (n = 21/40, 52.5%). Notably, all ST167 isolates with *fimH*29 (n = 18) carried
313 both $bla_{\text{CTX-M-14}}$ and $bla_{\text{CTX-M-15}}$. Analysis of the completed genome of KFS-C10 from this clade
314 identified that $bla_{\text{CTX-M-14}}$ was integrated into the chromosome while $bla_{\text{CTX-M-15}}$ was carried on
315 an IncF[F31:F36:A4:B1] plasmid (note that this plasmid carried two different FII alleles, F31
316 and F36) (**Table S2**). There were seven isolates with ExPEC status (n = 4 ST46, n = 1 ST10, n
317 = 1 ST93, n = 1 ST1286). All seven isolates were obtained from CIP-containing media and
318 lacked ESBL genes, but they carried 3 to 12 non-ESBL AMR genes in addition to QRDR
319 mutations.

320 **3.5. Genomic characteristics of ST1193 and ST131 isolates**

321 We investigated the detailed genetic characteristics of isolates belonging to ST1193 (n = 10)
322 and ST131 (n = 6), most of which were obtained from fecal sludge, because these STs are

323 recognized as established/emerging global high-risk clones and the only MDR clones that are
324 dominant among unselected ExPEC populations (Pitout et al., 2022).

325 ST1193 is derived from the ST14 clonal complex, carries *fimH64*, and harbors *gyrA* (S83L and
326 D87N) and *parC* (S80I) mutations (Pitout et al., 2022). Previous studies reported that ST1193
327 isolates cluster together within the phylogenetic tree according to three major IncF plasmid
328 types, namely F-:A1:B1, F-:A1:B10, and F-:A1:B20, though the clusters are not strictly
329 monophyletic (Johnson et al., 2019; Wyrsch et al., 2022). These clusters were named as green
330 (F-:A1:B1), red (F-:A1:B10), and pink (F-:A1:B20) (Johnson et al., 2019). The phylogenetic
331 tree constructed in our study was congruent with these previous studies (i.e., largely separated
332 according to IncF types), and our isolates fell into an F-:A1:B10 cluster (n = 7) and an
333 F-:A1:B20 cluster (n = 3) (**Figure 4a**). Three isolates (KFS-B24, KFS-B25, KFS-C19) and two
334 isolates (KFS-D07, KSE-D22) with the F-:A1:B10 plasmid type clustered together in the tree.
335 Three isolates (KFS-D21, KFS-D24, KFS-D26) with the F-:A1:B20 plasmid type also
336 clustered together. ST1193 isolates with F-:A1:B:20 plasmids generally contain the K5
337 capsular type, and those with F-:A1:B1 or F-:A1:B10 plasmids generally contain the K1
338 capsular type (Johnson et al., 2019). However, all of our ST1193 isolates, including those with
339 F-:A1:B:20 plasmids, contained the K1 capsular type. A variety of AMR genes associated with
340 different drug classes were detected in our ST1193 isolates (**Figure 4c**). Notably, among six
341 isolates with complete genomes, three (KFS-B24, KFS-D13, KSE-D22) carried chromosomal

342 AMR genes, including $bla_{CTX\text{-}M\text{-}15}$, within different genetic contexts (discussed in detail below
343 and in **Figure S3**). To determine if global ST1193 genomes carry chromosomal AMR genes,
344 we checked the location of AMR genes in all complete ST1193 genomes deposited in GenBank
345 (as of October 2023). Among 46 genomes, excluding our own genomes, eight (17%) carried
346 chromosomal AMR genes and six (13%) carried chromosomal $bla_{CTX\text{-}M}$ (**Table S5**). We also
347 found that the genetic contexts of chromosomal $bla_{CTX\text{-}M}$ in ST1193 genomes were diverse and
348 the contexts detected in two of our ST1193 genomes (KFS-B24 and KFS-D13) were not
349 described previously, indicating multiple events contributed to integration of $bla_{CTX\text{-}M}$ into the
350 ST1193 chromosome. The transposition unit detected in the chromosome of KSE-D22, which
351 contains $bla_{CTX\text{-}M\text{-}15}$ and four other AMR genes, was found to be inserted into the same position
352 in the chromosome of a Canadian ST1193 *E. coli* genome (GCA_032296485.1), potentially
353 suggesting global circulation of this strain. These observations indicate chromosomal
354 integration of AMR genes, as well as plasmid carriage, can contribute to the MDR phenotype
355 of ST1193. All ST1193 isolates in this study were classified as ExPEC based on the virulence
356 gene profiles, which potentially indicates these isolates can cause extraintestinal infections.

357 ST131 is known to be divided into clades A, B, and C (containing subclades C0, C1, and C2)
358 (Pitout and Finn, 2020). Within subclade C1 there is a clone known as C1-M27, which is
359 responsible for the spread of ESBL-producing *E. coli* in some countries (Matsumura et al.,
360 2017b; Merino et al., 2018). The ST131 isolates in the present study belonged to subclade C1-

361 M27 (n = 3), clade A (n = 2), and subclade C2 (n = 1) (**Figure 4b**). Three subclade C1-M27
362 isolates clustered together within the tree, while two clade A isolates were distantly related. As
363 reported previously (Matsumura et al., 2017b), the C1-M27 isolates carried *fimH*30, LNIV-
364 type QRDRs, and $bla_{\text{CTX-M-27}}$ (**Figure 4c**). The two clade A isolates were typed as *fimH41* as
365 previously reported (Matsumura et al., 2017b), but carried $bla_{\text{CTX-M-15}}$ and mutations in both
366 *gyrA* and *parC* QRDRs. Although clade A *E. coli* is generally sensitive to antibiotics (Pitout
367 and Finn, 2020), clade A isolates with $bla_{\text{CTX-M}}$ and *gyrA*/*parC* mutations have been
368 sporadically reported (Stoesser et al., 2016). The subclade C2 isolate carried *fimH*30, LNIV-
369 type QRDRs, and $bla_{\text{CTX-M-15}}$, which are the expected traits of the subclade C2 *E. coli* (Pitout
370 and Finn, 2020). A variety of AMR genes were detected in our ST131 isolates, and all the
371 isolates met the ExPEC criteria (**Figure 4c**).

372 Previous studies also identified the high-risk clones ST1193 and ST131 in sewage (Gomi et al.,
373 2017; Raven et al., 2019), indicating sewage surveillance can be an effective tool to monitor
374 the occurrence of these high-risk pathogens among populations. Genomic analysis, as
375 implemented in the present study, resolved fine-scale phylogenies of these high-risk clones at
376 the clade/subclade level (**Figure 4**), showing a combination of sewage (fecal sludge)
377 surveillance and whole-genome sequencing is even more effective and informative.

378 **3.6. Locations and genetic contexts of AMR genes in isolates from major clones**

379 In total, 26 genomes were selected from the top five STs (n = 8 ST167, n = 4 ST10, n = 3 ST46,
380 n = 2 ST1284, and n = 6 ST1193) and ST131 (n = 3) and subjected to long-read sequencing to
381 elucidate the locations and genetic contexts of the AMR genes in these major STs (care was
382 taken not to select multiple closely-related isolates). Hybrid assembly using short and long
383 reads completed 22 of 26 genomes (n = 6 ST167, n = 3 ST10, n = 3 ST46, n = 2 ST1284, n =
384 6 ST1193, and n = 2 ST131). The remaining four genomes could not be completed, probably
385 due to the presence of long repeats in the genomes.

386 The completed genomes carried one to eight plasmids (median four plasmids). Among the 89
387 plasmids in the completed genomes, 38 (42.7%) carried at least one AMR gene and thus were
388 defined as AMR plasmids. IncF AMR plasmids were the most prevalent (n = 25), followed by
389 Col(pHAD28) (n = 6), IncB/O/K/Z (n = 4), IncX1 (n = 1), IncX2 (n = 1), and IncHI2-IncHI2A
390 (n = 1) (**Table 1**). We note that the completed genomes represent only a subset of our isolates,
391 and the actual AMR plasmid diversity in the whole population may be different. Nonetheless,
392 the prevalence of AMR IncF plasmids is noteworthy, and 11 (79%) of 14 plasmidic $bla_{\text{CTX-M}}$
393 genes were located on IncF plasmids. IncF plasmids are frequently detected among *E. coli* from
394 humans and can stably be maintained in commensal *E. coli* within the human gastrointestinal
395 tract without antibiotic selection pressure (Bevan et al., 2017), which might explain this high
396 prevalence. IncF replicon sequence typing identified 14 different allele combinations, revealing
397 diverse IncF plasmids contributed to the spread of AMR genes within these major STs.

398 Although AMR plasmids of other replicon types were not common, two IncB/O/K/Z plasmids
399 and one IncHI2-IncHI2A plasmid contained a bla_{CTX-M} gene. We found plasmids from other
400 continents that were highly similar to 13 (34%) of 38 AMR plasmids in our study under the
401 stringent criteria applied (i.e., >99% query coverage and >99% identity) (**Table 1**, **Table S2**).
402 This indicates circulation of global AMR plasmids (or clones with such plasmids) in our study
403 setting. On the other hand, close blastn hits (defined here as hits with >80% query coverage)
404 were not found for some AMR plasmids (CP124985, CP124987, and CP124993), indicating
405 the occurrence of unique/rare AMR plasmids in this region.

406 Among the 22 completed genomes, AMR genes were detected on the chromosome in eight
407 isolates (n = 3 ST167, n = 3 ST1193, n = 2 ST46). Notably, all eight isolates carried
408 chromosomal bla_{CTX-M} (n = 7 $bla_{CTX-M-15}$ and n = 1 $bla_{CTX-M-14}$), and three isolates carried
409 additional chromosomal AMR genes (**Table S2**). Detailed analyses of genetic contexts of
410 chromosomal bla_{CTX-M} genes and adjacent AMR genes are provided in **Figure S3**. We note that
411 recent studies have also shown the frequent occurrence (more than 50%) of chromosomal
412 bla_{CTX-M} integration among *E. coli* isolates using both short- and long-read sequencing (Gomi
413 et al., 2022; Milenkov et al., 2023). Chromosomal integration of AMR genes might have been
414 previously underestimated due to the inability of PCR or short-read sequencing to determine
415 the location of AMR genes. Long-read sequencing will become more common in the near

416 future, which may uncover more examples of chromosomal integration of AMR genes and
417 underline the importance of this mechanism in the stabilization and maintenance of AMR genes.

418 **3.7. Study limitations**

419 This study has some limitations. First, the clonal composition of antimicrobial-resistant *E. coli*
420 observed in the present study might not directly reflect that of the gut flora of the underlying
421 population due to the noise caused by e.g., differential survival of *E. coli* strains in sewer
422 pipes/fecal sludge (Huijbers et al., 2020). Moreover, contamination from non-human animal
423 sources in the samples cannot be excluded. Dominance of CC10 in our collection, which is
424 prevalent not only in human feces but also in animal feces (Reid et al., 2019; Reid et al., 2017;
425 Zingali et al., 2020), also implies this possibility. However, we think contamination from non-
426 human animals is limited because the contribution from animal feces is lower than that from
427 human feces in this region (Komakech et al., 2014), and *E. coli* concentrations are generally
428 much higher in human feces than in domestic animals (Tenaillon et al., 2010). We also note
429 that these non-human sources should also be considered in terms of the One Health concept.
430 Inclusion of a sufficient number of fecal *E. coli* isolates from the matched human population
431 in genomic analysis will further confirm the usefulness of fecal sludge/sewage in AMR
432 surveillance. Second, we performed sampling over a relatively short period. Longitudinal
433 sampling with timeframes over one year, for example, will provide information on the change
434 of clonal composition. Third, long-read sequencing was performed only on the selected isolates

435 from major STs. Future studies employing larger-scale long-read sequencing of isolates from
436 various clones will elucidate AMR vectors also in minor STs. Nevertheless, our analyses
437 revealed detailed genomic characteristics of antimicrobial-resistant *E. coli* in fecal sludge and
438 sewage in Uganda, providing valuable insights into AMR burden in this region.

439

440 **4. Conclusions**

441 Here, we applied whole-genome sequencing to characterize antimicrobial-resistant *E. coli*
442 obtained from fecal sludge and sewage in Uganda. Our study revealed that CC10, including
443 ST167, ST10, ST1284, and ST617, were prevalent among both fecal sludge and sewage
444 isolates, indicating this lineage significantly contributed to AMR burden in the underlying
445 population. Global high-risk clones ST1193 and ST131 were also detected, adding potential
446 health risks to the population. We also revealed that many of the AMR genes were carried by
447 IncF plasmids or integrated into the chromosome by incorporating data obtained from long-
448 read sequencing. We admit that performing whole-genome sequencing in routine surveillance
449 is currently not realistic. However, as sequencing costs will continue to decrease, we believe
450 that a combination of fecal sludge/sewage surveillance and whole-genome sequencing will be
451 a powerful tool for monitoring AMR carriage in the underlying population.

452

Acknowledgements

This work was supported by the Japan Society for the Promotion of Science KAKENHI (grant numbers JP18K18881 and JP19H02274). Computations were partially performed on the NIG supercomputer at ROIS National Institute of Genetics. We acknowledge the NGS core facility of the Genome Information Research Center at the Research Institute for Microbial Diseases of Osaka University for the support in DNA sequencing.

References

Aarestrup, F.M. and Woolhouse, M.E.J. 2020. Using sewage for surveillance of antimicrobial resistance. Science 367(6478), 630-632.

Anjum, M.F., Schmitt, H., Borjesson, S., Berendonk, T.U. and network, W. 2021. The potential of using E. coli as an indicator for the surveillance of antimicrobial resistance (AMR) in the environment. Curr Opin Microbiol 64, 152-158.

Aoike, N., Saga, T., Sakata, R., Yoshizumi, A., Kimura, S., Iwata, M., Yoshizawa, S., Sugasawa, Y., Ishii, Y., Yamaguchi, K. and Tateda, K. 2013. Molecular characterization of extraintestinal *Escherichia coli* isolates in Japan: relationship between sequence types and mutation patterns of quinolone resistance-determining regions analyzed by pyrosequencing. J Clin Microbiol 51(6), 1692-1698.

Aziz, R.K., Bartels, D., Best, A.A., DeJongh, M., Disz, T., Edwards, R.A., Formsma, K., Gerdes, S., Glass, E.M., Kubal, M., Meyer, F., Olsen, G.J., Olson, R., Osterman, A.L., Overbeek, R.A., McNeil, L.K., Paarmann, D., Paczian, T., Parrello, B., Pusch, G.D., Reich, C., Stevens, R., Vassieva, O., Vonstein, V., Wilke, A. and Zagnitko, O. 2008. The RAST Server: rapid annotations using subsystems technology. BMC Genomics 9, 75.

Beghain, J., Bridier-Nahmias, A., Le Nagard, H., Denamur, E. and Clermont, O. 2018. ClermonTyping: an easy-to-use and accurate in silico method for *Escherichia* genus strain phylotyping. Microb Genom 4(7).

Berendes, D.M., Sumner, T.A. and Brown, J.M. 2017. Safely Managed Sanitation for All Means Fecal Sludge Management for At Least 1.8 Billion People in Low and Middle Income Countries. Environ Sci Technol 51(5), 3074-3083.

483 Bevan, E.R., Jones, A.M. and Hawkey, P.M. 2017. Global epidemiology of CTX-M beta-
484 lactamases: temporal and geographical shifts in genotype. J Antimicrob Chemother
485 72(8), 2145-2155.
486 Bortolaia, V., Kaas, R.S., Ruppe, E., Roberts, M.C., Schwarz, S., Cattoir, V., Philippon, A.,
487 Allesoe, R.L., Rebelo, A.R., Florensa, A.F., Fagelhauer, L., Chakraborty, T., Neumann,
488 B., Werner, G., Bender, J.K., Stingl, K., Nguyen, M., Coppens, J., Xavier, B.B.,
489 Malhotra-Kumar, S., Westh, H., Pinholt, M., Anjum, M.F., Duggett, N.A., Kempf, I.,
490 Nykasenoja, S., Olkkola, S., Wieczorek, K., Amaro, A., Clemente, L., Mossong, J.,
491 Losch, S., Ragimbeau, C., Lund, O. and Aarestrup, F.M. 2020. ResFinder 4.0 for
492 predictions of phenotypes from genotypes. J Antimicrob Chemother 75(12), 3491-3500.
493 Capone, D., Berendes, D., Cumming, O., Knee, J., Nala, R., Risk, B.B., Stauber, C., Zhu, K.
494 and Brown, J. 2020. Analysis of Fecal Sludges Reveals Common Enteric Pathogens
495 in Urban Maputo, Mozambique. Environ Sci Tech Let 7(12), 889-895.
496 Carattoli, A., Zankari, E., Garcia-Fernandez, A., Voldby Larsen, M., Lund, O., Villa, L., Moller
497 Aarestrup, F. and Hasman, H. 2014. *In silico* detection and typing of plasmids using
498 PlasmidFinder and plasmid multilocus sequence typing. Antimicrob Agents Chemother
499 58(7), 3895-3903.
500 Chen, S., Zhou, Y., Chen, Y. and Gu, J. 2018. fastp: an ultra-fast all-in-one FASTQ
501 preprocessor. Bioinformatics 34(17), i884-i890.
502 Chigwechokha, P., Street, R. and Holm, R.H. 2022. Advancing the Use of Fecal Sludge for
503 Timelier and Better-Quality Epidemiological Data in Low- and Middle-Income
504 Countries for Pandemic Prevention. Environ Sci Technol.
505 Decano, A.G., Pettigrew, K., Sabiiti, W., Sloan, D.J., Neema, S., Bazira, J., Kiiru, J., Onyango,
506 H., Asiimwe, B. and Holden, M.T.G. 2021. Pan-Resistome Characterization of
507 Uropathogenic *Escherichia coli* and *Klebsiella pneumoniae* Strains Circulating in
508 Uganda and Kenya, Isolated from 2017-2018. Antibiotics (Basel) 10(12).
509 Denamur, E., Clermont, O., Bonacorsi, S. and Gordon, D. 2021. The population genetics
510 of pathogenic *Escherichia coli*. Nat Rev Microbiol 19(1), 37-54.
511 Gardner, S.N., Slezak, T. and Hall, B.G. 2015. kSNP3.0: SNP detection and phylogenetic
512 analysis of genomes without genome alignment or reference genome. Bioinformatics
513 31(17), 2877-2878.
514 Gomi, R., Matsuda, T., Fujimori, Y., Harada, H., Matsui, Y. and Yoneda, M. 2015.
515 Characterization of Pathogenic *Escherichia coli* in River Water by Simultaneous
516 Detection and Sequencing of 14 Virulence Genes. Environ Sci Technol 49(11), 6800-
517 6807.
518 Gomi, R., Matsuda, T., Matsumura, Y., Yamamoto, M., Tanaka, M., Ichiyama, S. and Yoneda,
519 M. 2017. Occurrence of Clinically Important Lineages, Including the Sequence

520 Type 131 C1-M27 Subclone, among Extended-Spectrum-beta-Lactamase-Producing
521 *Escherichia coli* in Wastewater. Antimicrob Agents Chemother 61(9).
522 Gomi, R., Yamamoto, M., Tanaka, M. and Matsumura, Y. 2022. Chromosomal integration
523 of bla_{CTX-M} genes in diverse *Escherichia coli* isolates recovered from river water in
524 Japan. Curr Res Microb Sci 3, 100144.
525 Gumede, S.N., Abia, A.L.K., Amoako, D.G. and Essack, S.Y. 2021. Analysis of
526 Wastewater Reveals the Spread of Diverse Extended-Spectrum beta-Lactamase-
527 Producing *E. coli* Strains in uMgungundlovu District, South Africa. Antibiotics (Basel)
528 10(7).
529 Hendriksen, R.S., Munk, P., Njage, P., van Bunnik, B., McNally, L., Lukjancenko, O., Roder,
530 T., Nieuwenhuijse, D., Pedersen, S.K., Kjeldgaard, J., Kaas, R.S., Clausen, P., Vogt,
531 J.K., Leekitcharoenphon, P., van de Schans, M.G.M., Zuidema, T., de Roda Husman,
532 A.M., Rasmussen, S., Petersen, B., Global Sewage Surveillance project, c., Amid, C.,
533 Cochrane, G., Sicheritz-Ponten, T., Schmitt, H., Alvarez, J.R.M., Aidara-Kane, A.,
534 Pamp, S.J., Lund, O., Hald, T., Woolhouse, M., Koopmans, M.P., Vigre, H., Petersen,
535 T.N. and Aarestrup, F.M. 2019. Global monitoring of antimicrobial resistance
536 based on metagenomics analyses of urban sewage. Nat Commun 10(1), 1124.
537 Huijbers, P.M.C., Larsson, D.G.J. and Flach, C.F. 2020. Surveillance of antibiotic resistant
538 *Escherichia coli* in human populations through urban wastewater in ten European
539 countries. Environ Pollut 261, 114200.
540 Hutinel, M., Huijbers, P.M.C., Fick, J., Ahren, C., Larsson, D.G.J. and Flach, C.F. 2019.
541 Population-level surveillance of antibiotic resistance in *Escherichia coli* through
542 sewage analysis. Euro Surveill 24(37).
543 Jain, C., Rodriguez, R.L., Phillippy, A.M., Konstantinidis, K.T. and Aluru, S. 2018. High
544 throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries.
545 Nat Commun 9(1), 5114.
546 Joensen, K.G., Scheutz, F., Lund, O., Hasman, H., Kaas, R.S., Nielsen, E.M. and Aarestrup,
547 F.M. 2014. Real-time whole-genome sequencing for routine typing, surveillance,
548 and outbreak detection of verotoxigenic *Escherichia coli*. J Clin Microbiol 52(5), 1501-
549 1510.
550 Johnson, J.R. and Stell, A.L. 2000. Extended virulence genotypes of *Escherichia coli*
551 strains from patients with urosepsis in relation to phylogeny and host compromise. J
552 Infect Dis 181(1), 261-272.
553 Johnson, T.J., Elnekave, E., Miller, E.A., Munoz-Aguayo, J., Flores Figueroa, C., Johnston, B.,
554 Nielson, D.W., Logue, C.M. and Johnson, J.R. 2019. Phylogenomic Analysis of
555 Extraintestinal Pathogenic *Escherichia coli* Sequence Type 1193, an Emerging
556 Multidrug-Resistant Clonal Group. Antimicrob Agents Chemother 63(1).

557 Katongole, P., Nalubega, F., Florence, N.C., Asiimwe, B. and Andia, I. 2020. Biofilm
558 formation, antimicrobial susceptibility and virulence genes of Uropathogenic
559 *Escherichia coli* isolated from clinical isolates in Uganda. BMC Infect Dis 20(1), 453.
560 Kohler, C.D. and Dobrindt, U. 2011. What defines extraintestinal pathogenic *Escherichia*
561 *coli*? Int J Med Microbiol 301(8), 642-647.
562 Kolmogorov, M., Yuan, J., Lin, Y. and Pevzner, P.A. 2019. Assembly of long, error-prone
563 reads using repeat graphs. Nat Biotechnol 37(5), 540-546.
564 Komakech, A.J., Banadda, N.E., Gebresenbet, G. and Vinneras, B. 2014. Maps of animal
565 urban agriculture in Kampala City. Agron Sustain Dev 34(2), 493–500.
566 Kwak, Y.K., Colque, P., Byfors, S., Giske, C.G., Mollby, R. and Kuhn, I. 2015.
567 Surveillance of antimicrobial resistance among *Escherichia coli* in wastewater in
568 Stockholm during 1 year: does it reflect the resistance trends in the society? Int J
569 Antimicrob Agents 45(1), 25-32.
570 Larsson, D.G.J., Flach, C.F. and Laxminarayan, R. 2022. Sewage surveillance of antibiotic
571 resistance holds both opportunities and challenges. Nat Rev Microbiol, 1-2.
572 Le, T.H., Tong, D.K., Pham, T.T.V., Hoang, T.L., Tran, T.K. and Ng, C. 2023. Occurrence
573 and removal of antibiotic resistant *Escherichia coli* and antibiotic resistance genes in
574 an urban wastewater treatment plant. Int J Environ Sci Technol.
575 Majumder, M.A.A., Rahman, S., Cohall, D., Bharatha, A., Singh, K., Haque, M. and Gittens-
576 St Hilaire, M. 2020. Antimicrobial Stewardship: Fighting Antimicrobial Resistance
577 and Protecting Global Public Health. Infect Drug Resist 13, 4713-4738.
578 Malberg Tetzschner, A.M., Johnson, J.R., Johnston, B.D., Lund, O. and Scheutz, F. 2020.
579 *In Silico* Genotyping of *Escherichia coli* Isolates for Extraintestinal Virulence Genes
580 by Use of Whole-Genome Sequencing Data. J Clin Microbiol 58(10).
581 Matsumura, Y., Noguchi, T., Tanaka, M., Kanahashi, T., Yamamoto, M., Nagao, M., Takakura,
582 S., Ichiyama, S. and th, J.B. 2017a. Population structure of Japanese extraintestinal
583 pathogenic *Escherichia coli* and its relationship with antimicrobial resistance. J
584 Antimicrob Chemother 72(4), 1040-1049.
585 Matsumura, Y., Pitout, J.D.D., Peirano, G., DeVinney, R., Noguchi, T., Yamamoto, M., Gomi,
586 R., Matsuda, T., Nakano, S., Nagao, M., Tanaka, M. and Ichiyama, S. 2017b. Rapid
587 Identification of Different *Escherichia coli* Sequence Type 131 Clades. Antimicrob
588 Agents Chemother 61(8).
589 Merino, I., Hernandez-Garcia, M., Turrientes, M.C., Perez-Viso, B., Lopez-Fresnena, N., Diaz-
590 Agero, C., Maechler, F., Fankhauser-Rodriguez, C., Kola, A., Schrenzel, J., Harbarth,
591 S., Bonten, M., Gastmeier, P., Canton, R., Ruiz-Garbajosa, P. and Group, R.G.S.
592 2018. Emergence of ESBL-producing *Escherichia coli* ST131-C1-M27 clade
593 colonizing patients in Europe. J Antimicrob Chemother 73(11), 2973-2980.

594 Milenkov, M., Proux, C., Rasolofoarison, T.L., Rakotomalala, F.A., Rasoanandrasana, S.,
595 Rahajamanana, L.V., Rafalimanana, C., Ravaoarisaina, Z., Ramahatafandry, I.T.,
596 Westeel, E., Petitjean, M., Marin, J., Mullaert, J., Han, L., Clermont, O., Raskine, L.,
597 Endtz, H., Andremont, A., Denamur, E., Komurian-Pradel, F., Samison, L.H. and
598 Armand-Lefevre, L. 2023. Tricycle surveillance in Antananarivo, Madagascar:
599 circulation of both extended-spectrum beta-lactamase producing *Escherichia coli*
600 strains and plasmids among humans, chickens and the environment. medRXiv.
601 Nabadda, S., Kakooza, F., Kiggundu, R., Walwema, R., Bazira, J., Mayito, J., Mugerwa, I.,
602 Sekamatte, M., Kambugu, A., Lamorde, M., Kajumbula, H. and Mwebasa, H. 2021.
603 Implementation of the World Health Organization Global Antimicrobial Resistance
604 Surveillance System in Uganda, 2015-2020: Mixed-Methods Study Using National
605 Surveillance Data. JMIR Public Health Surveill 7(10), e29954.
606 Odongo, I., Ssemambo, R. and Kungu, J.M. 2020. Prevalence of *Escherichia Coli* and Its
607 Antimicrobial Susceptibility Profiles among Patients with UTI at Mulago Hospital,
608 Kampala, Uganda. Interdiscip Perspect Infect Dis 2020, 8042540.
609 Pitout, J.D.D. and Finn, T.J. 2020. The evolutionary puzzle of *Escherichia coli* ST131.
610 Infect Genet Evol 81, 104265.
611 Pitout, J.D.D., Peirano, G., Chen, L., DeVinney, R. and Matsumura, Y. 2022. *Escherichia*
612 *coli* ST1193: Following in the Footsteps of *E. coli* ST131. Antimicrob Agents
613 Chemother 66(7), e0051122.
614 Poirel, L., Madec, J.Y., Lupo, A., Schink, A.K., Kieffer, N., Nordmann, P. and Schwarz, S.
615 2018. Antimicrobial Resistance in *Escherichia coli*. Microbiol Spectr 6(4).
616 Raven, K.E., Ludden, C., Gouliouris, T., Blane, B., Naydenova, P., Brown, N.M., Parkhill, J.
617 and Peacock, S.J. 2019. Genomic surveillance of *Escherichia coli* in municipal
618 wastewater treatment plants as an indicator of clinically relevant pathogens and their
619 resistance genes. Microb Genom 5(5).
620 Reid, C.J., DeMaere, M.Z. and Djordjevic, S.P. 2019. Australian porcine clonal complex
621 10 (CC10) *Escherichia coli* belong to multiple sublineages of a highly diverse global
622 CC10 phylogeny. Microb Genom 5(3).
623 Reid, C.J., Wyrsch, E.R., Roy Chowdhury, P., Zingali, T., Liu, M., Darling, A.E., Chapman,
624 T.A. and Djordjevic, S.P. 2017. Porcine commensal *Escherichia coli*: a reservoir
625 for class 1 integrons associated with IS*26*. Microb Genom 3(12).
626 Roer, L., Tchesnokova, V., Allesoe, R., Muradova, M., Chattopadhyay, S., Ahrenfeldt, J.,
627 Thomsen, M.C.F., Lund, O., Hansen, F., Hammerum, A.M., Sokurenko, E. and Hasman,
628 H. 2017. Development of a Web Tool for *Escherichia coli* Subtyping Based on
629 *fimH* Alleles. J Clin Microbiol 55(8), 2538-2543.

630 Siguier, P., Perochon, J., Lestrade, L., Mahillon, J. and Chandler, M. 2006. ISfinder: the
631 reference centre for bacterial insertion sequences. Nucleic Acids Res 34(Database
632 issue), D32-36.
633 Stoesser, N., Sheppard, A.E., Pankhurst, L., De Maio, N., Moore, C.E., Sebra, R., Turner, P.,
634 Anson, L.W., Kasarskis, A., Batty, E.M., Kos, V., Wilson, D.J., Phetsouvanh, R.,
635 Wyllie, D., Sokurenko, E., Manges, A.R., Johnson, T.J., Price, L.B., Peto, T.E., Johnson,
636 J.R., Didelot, X., Walker, A.S., Crook, D.W. and Modernizing Medical Microbiology
637 Informatics, G. 2016. Evolutionary History of the Global Emergence of the
638 *Escherichia coli* Epidemic Clone ST131. mBio 7(2), e02162.
639 Tenaillon, O., Skurnik, D., Picard, B. and Denamur, E. 2010. The population genetics of
640 commensal *Escherichia coli*. Nat Rev Microbiol 8(3), 207-217.
641 Treangen, T.J., Ondov, B.D., Koren, S. and Phillippy, A.M. 2014. The Harvest suite for
642 rapid core-genome alignment and visualization of thousands of intraspecific microbial
643 genomes. Genome Biol 15(11), 524.
644 van der Putten, B.C.L., Remondini, D., Pasquini, G., Janes, V.A., Matamoros, S. and Schultsz,
645 C. 2019. Quantifying the contribution of four resistance mechanisms to
646 ciprofloxacin MIC in *Escherichia coli*: a systematic review. J Antimicrob Chemother
647 74(2), 298-310.
648 Wadley, T.D., Jenjaroenpun, P., Wongsurawat, T., Ussery, D.W. and Nookaew, I. 2019.
649 Complete Genome and Plasmid Sequences of *Escherichia coli* Type Strain ATCC
650 11775. Microbiol Resour Announc 8(9).
651 Waters, N.R., Abram, F., Brennan, F., Holmes, A. and Pritchard, L. 2020. Easy
652 phylotyping of *Escherichia coli* via the EzClermont web app and command-line tool.
653 Access Microbiol 2(9), acmi000143.
654 WHO 2015. Global action plan on antimicrobial resistance.
655 WHO 2021. WHO integrated global surveillance on ESBL-producing *E. coli* using a "One
656 Health" approach: implementation and opportunities.
657 WHO 2022. Global antimicrobial resistance and use surveillance system (GLASS) report:
658 2022.
659 Wick, R.R. and Holt, K.E. 2022. Polypolish: Short-read polishing of long-read bacterial
660 genome assemblies. PLoS Comput Biol 18(1), e1009802.
661 Wick, R.R., Judd, L.M., Gorrie, C.L. and Holt, K.E. 2017. Unicycler: Resolving bacterial
662 genome assemblies from short and long sequencing reads. PLoS Comput Biol 13(6),
663 e1005595.
664 Wyrsch, E.R., Bushell, R.N., Marenda, M.S., Browning, G.F. and Djordjevic, S.P. 2022.
665 Global Phylogeny and F Virulence Plasmid Carriage in Pandemic *Escherichia coli*
666 ST1193. Microbiol Spectr 10(6), e0255422.

667 Zhou, Z., Alikhan, N.F., Mohamed, K., Fan, Y., Agama Study, G. and Achtman, M. 2020.
668 The EnteroBase user's guide, with case studies on *Salmonella* transmissions, *Yersinia*
669 *pestis* phylogeny, and *Escherichia* core genomic diversity. Genome Res 30(1), 138-152.
670 Zingali, T., Reid, C.J., Chapman, T.A., Gaio, D., Liu, M., Darling, A.E. and Djordjevic, S.P.
671 2020. Whole Genome Sequencing Analysis of Porcine Faecal Commensal
672 *Escherichia coli* Carrying Class 1 Integrons from Sows and Their Offspring.
673 Microorganisms 8(6).

674

675

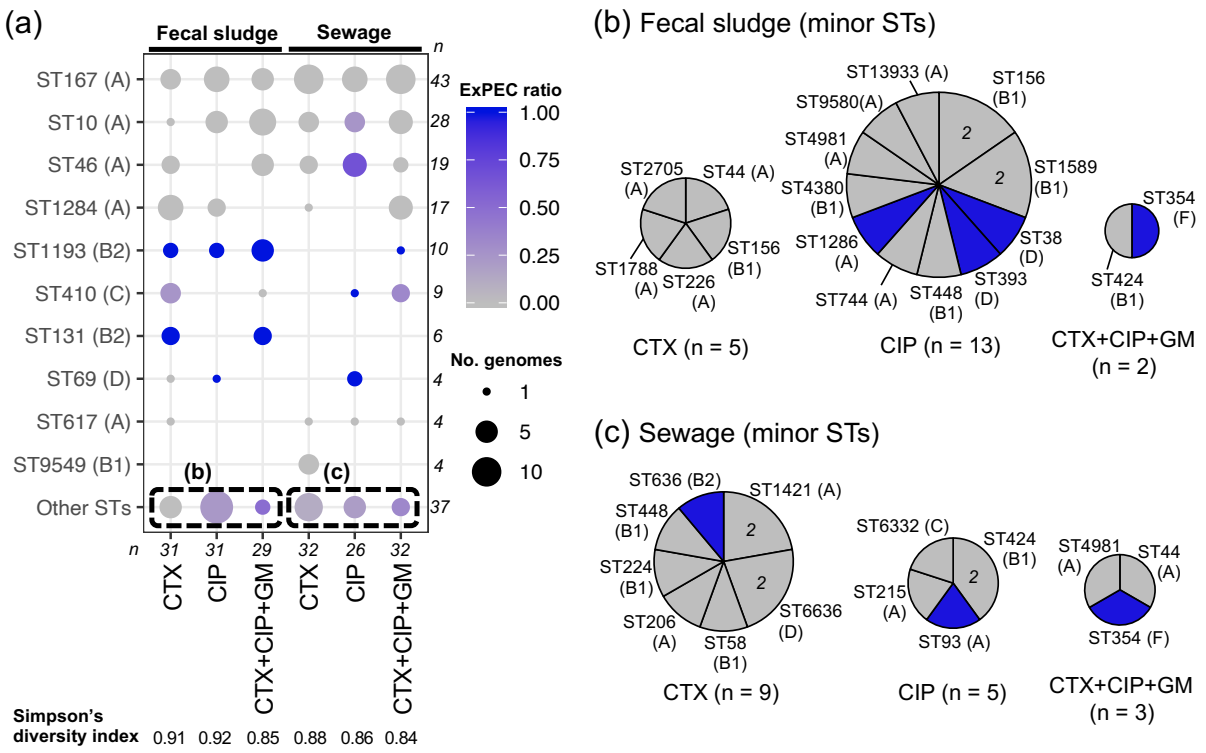


Figure 1. (a) Clonal composition of 181 *E. coli* isolates. STs are shown separately for each media type (CTX, CIP, and CTX + CIP + GM) and each sample type (fecal sludge and sewage). The top ten STs are shown in descending order. The remaining STs, all of which were with three or fewer isolates, are combined into "other STs". Phylogroup corresponding to each ST is shown in parentheses. Circle size indicates the number of genomes, and intensity of circle shading indicates the proportion of genomes classified as ExPEC (as per inset legends). Clonal composition of minor STs (i.e., STs classified as "other STs" in (a)) from (b) fecal sludge and (c) sewage. Phylogroup corresponding to each ST is shown in parentheses, and STs classified as ExPEC are colored blue.

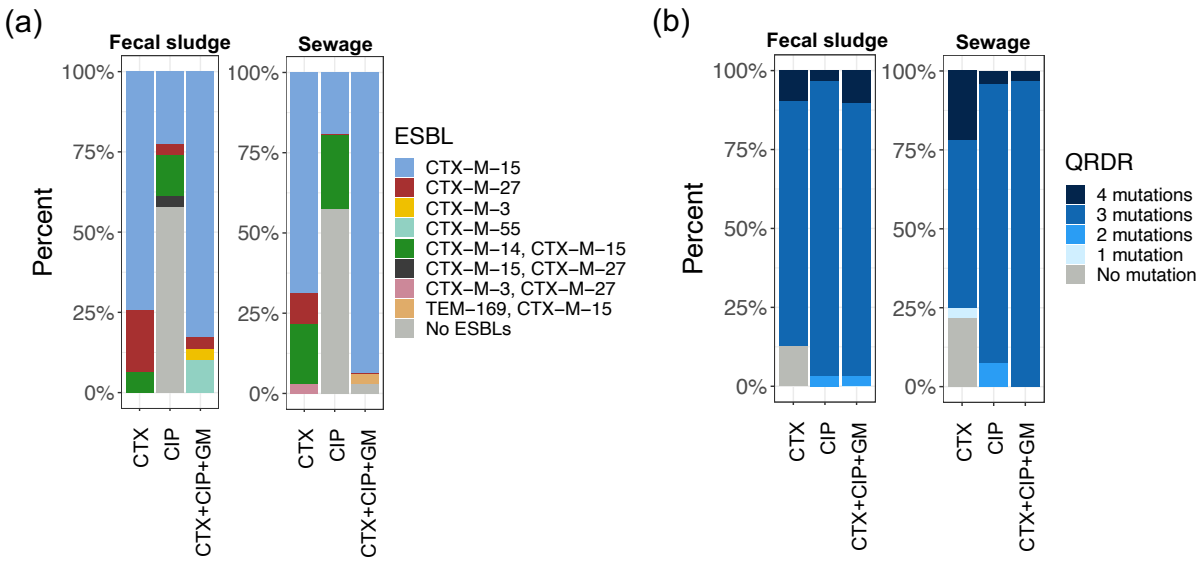
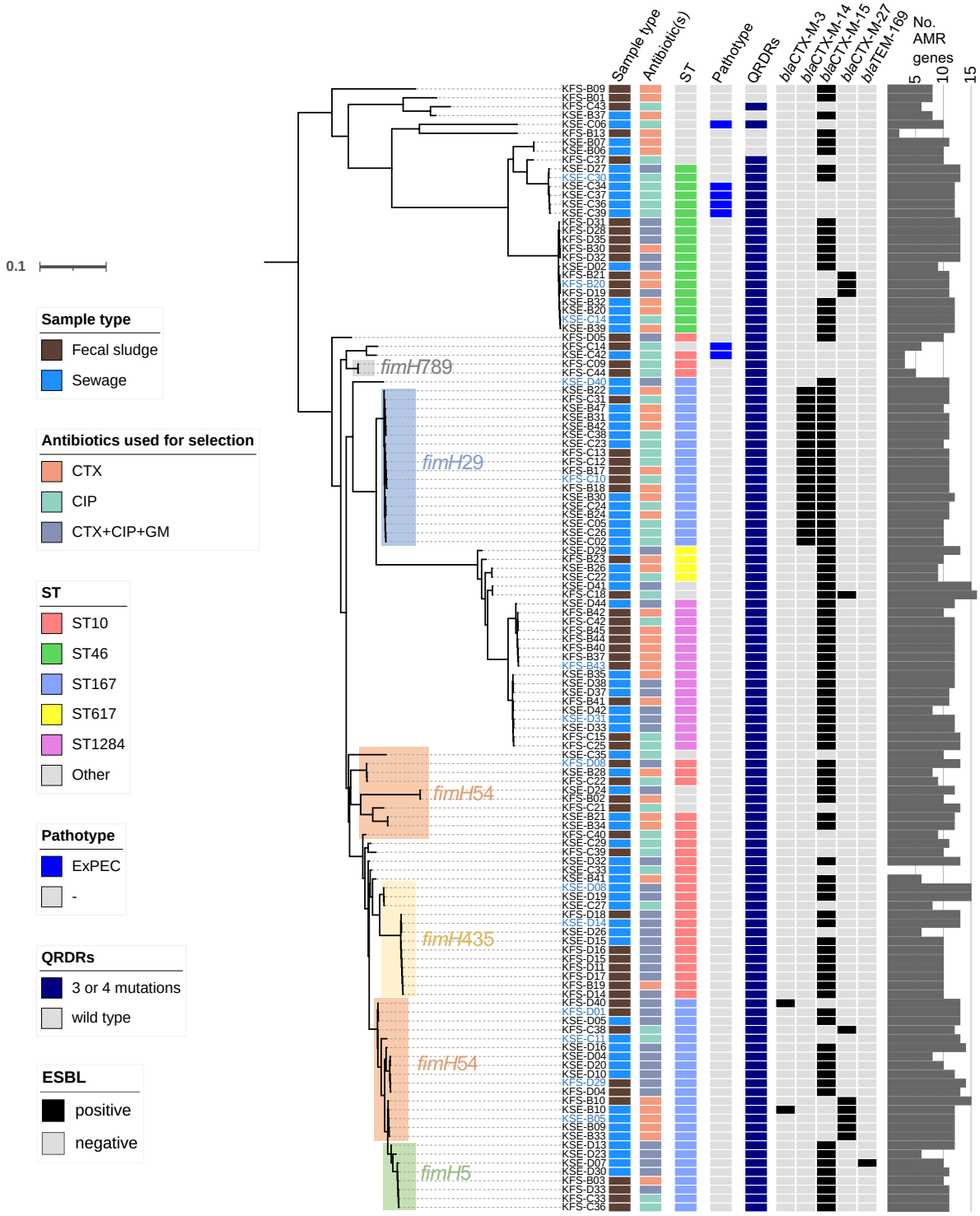


Figure 2. (a) ESBL genes and (b) QRDR mutations in *E. coli* isolates. AMR determinants are shown separately for each media type (CTX, CIP, and CTX + CIP + GM) and each sample type (fecal sludge and sewage). This figure is based on the analysis of short-read assembly data and isolates with multiple copies of the same ESBL gene can be reported to carry a single copy. For QRDRs, mutations in GyrA codons 83 and 87 and ParC codons 80 and 84 were counted.



695

696 **Figure 3**. kSNP3 whole-genome SNP-based phylogenetic tree of 127 phylogroup A isolates.

697 The tree was constructed using 54,632 SNP-loci, which occurred in at least 90% of the genomes.

698 A k-mer length of 19 was used for SNP identification. The presence of traits is indicated by

699 colored boxes (as per inset legends). For the ST column, major STs (i.e., ST10, ST46, ST167,

700 ST617, and ST1284) are indicated by the corresponding color boxes, while other STs are

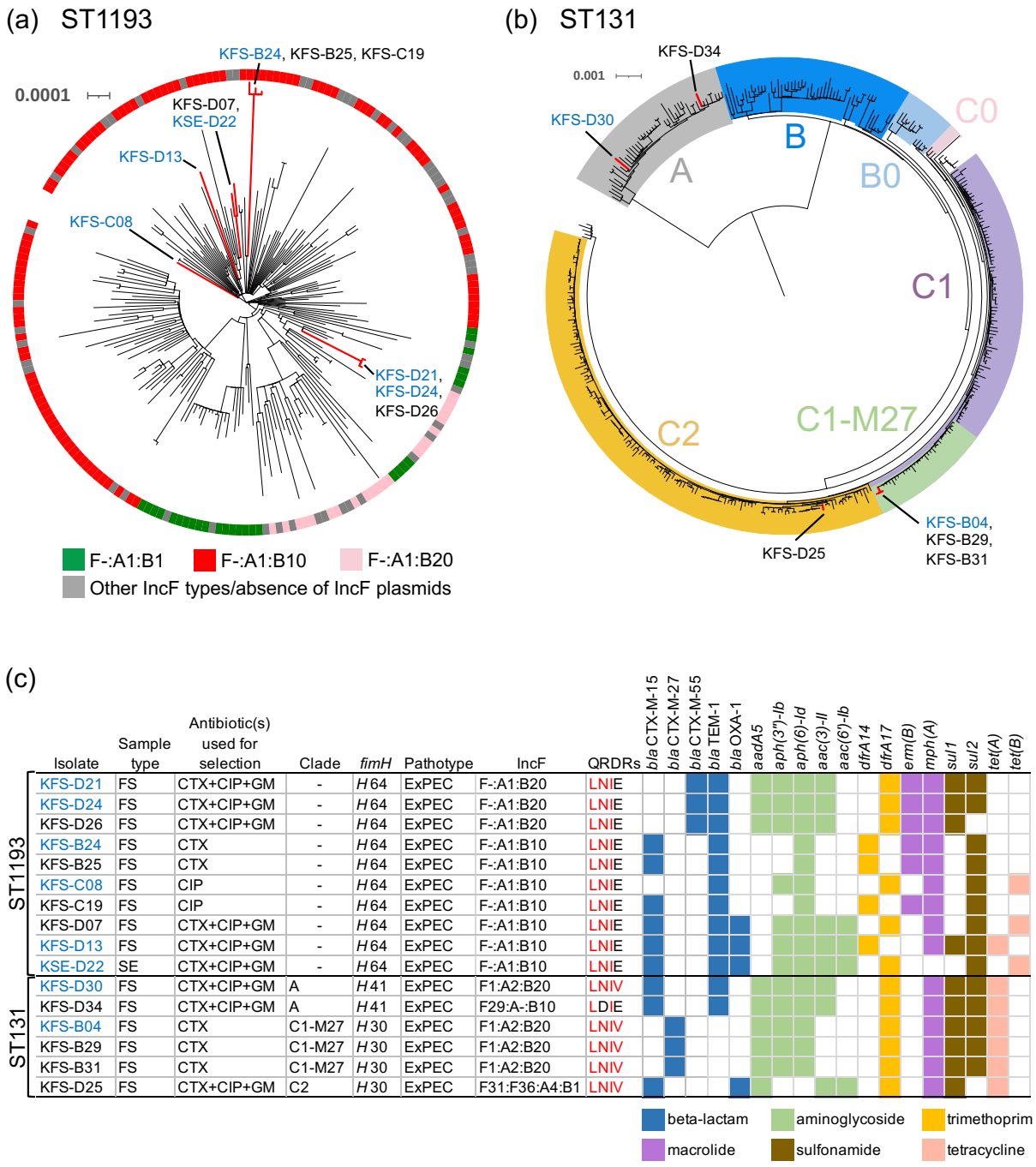
701 represented by gray boxes. ST10 and ST167 clusters are highlighted in the tree according to

702 the *fimH* alleles, except for singletons. The names of completed genomes are colored blue.

703

704

705



706

707 **Figure 4**. Recombination-free phylogenetic trees for (a) ST1193 and (b) ST131. For ST1193,

708 we used the Assembly Dereplicator tool (v0.3.0, https://github.com/rrwick/Assembly-

709 Dereplicator) and subsampled global isolates to remove near-identical redundant isolates and

710 to reduce computational burden. The ST1193 tree was built using a 3,899 bp alignment

711 generated from our isolates (n = 10) and global isolates (n = 198). The ST131 tree was built
712 using a 10,712 bp alignment generated from our isolates (n = 6) and global isolates (n = 379).
713 Branches leading to our isolates are colored in red. The trees were midpoint rooted. Accession
714 numbers of global ST1193 genomes and ST131 genomes are available in **Table S6** and **Table**
715 **S7**. (c) Key genetic features of ST1193 and ST131 isolates sequenced in the present study. FS,
716 fecal sludge; SE, sewage. We did not assign clades to ST1193 isolates because the previously
717 defined ST1193 clades are not monophyletic as discussed in the main text. For QRDRs, GyrA
718 codons 83 and 87 and ParC codons 80 and 84 that are different from the wild type (SDSE) are
719 colored in red. Presence of AMR genes is indicated by colored boxes (colored by drug class).
720 All four isolates with *aac(6')-Ib* carried the *aac(6')-Ib-cr* variant, which is associated with
721 decreased susceptibility to aminoglycosides and to ciprofloxacin and norfloxacin. KFS-D13
722 carried *aadA1*, *dfrA5*, *ere(A)*, *qnrB1*, and *catA1* in addition to AMR genes shown in (c).
723 Genomes with blue names in (a)–(c) were completed.

724

725

726 **Table 1**. Basic characteristics of AMR plasmids in major STs.

Inc type (n)	Plasmid ST (n)	Chromosomal ST (n)	Size	ESBL gene (n)	No. of other AMR genes	Identical or closely related plasmids in other continents[a] (n)	Accession number(s)
Inc F[b] (25)	F-:A1:B10 (4)	ST1193 (4)	88,138 bp–108,071 bp	-	3–7	Yes (3), No (1)	CP124974, CP125028, CP125054, CP125066
	F31:F36:A4:B1 (4)	ST167 (2)	168,379 bp, 173,762 bp	*bla*$_{CTX-M-15}$ (2)	6–8	No (2)	CP124971, CP125044
		ST10 (1)	174,470 bp	*bla*$_{CTX-M-15}$ (1)	8	No (1)	CP125035
		ST1284 (1)	174,534 bp	*bla*$_{CTX-M-15}$ (1)	8	No (1)	CP124977
	F48:A1:B49 (4)	ST167 (4)	81,682 bp–131,123 bp	*bla*$_{CTX-M-27}$ (1)	8–11	No (4)	CP124996, CP125002, CP125007, CP125040
	F-:A1:B20 (2)	ST1193 (2)	90,924 bp	-	7	No (2)	CP125014, CP125021
	F1:A2:B20 (2)	ST131 (clade A) (1)	146,809 bp	-	10	Yes (1)	CP125005
		ST131 (subclade C1-M27) (1)	104,558 bp	*bla*$_{CTX-M-27}$ (1)	8	Yes (1)	CP125077

	F-:A-:B- (1)	ST131 (clade A) (1)	110,850 bp	*bla*$_{CTX-M-15}$ (1)	0	Yes (1)	CP125006
	F-:A-:B1 (1)	ST46 (1)	97,539 bp	*bla*$_{CTX-M-27}$ (1)	10	No (1)	CP125071
	F2:A-:B- (1)	ST46 (1)	73,681 bp	-	11	No (1)	CP124987
	F22:F36:A1:B49 (1)	ST46 (1)	131,446 bp	-	11	No (1)	CP124991
	F36:A1:B1 (1)	ST1284 (1)	122,643 bp	*bla*$_{CTX-M-15}$ (1)	9	No (1)	CP125060
	F36:A1:B20 (1)	ST10 (1)	109,736 bp	*bla*$_{CTX-M-15}$ (1)	13	No (1)	CP124993
	F40:A-:B- (1)	ST10 (1)	70,669 bp	-	5	No (1)	CP125034
	F48:A1:B1 (1)	ST10 (1)	103,762 bp	*bla*$_{CTX-M-15}$ (1)	14	No (1)	CP124985
	F77:A-:B- (1)	ST167 (1)	75,825 bp	-	7	No (1)	CP124997
Col (pHAD28) (6)	-	ST167 (3)	6,200 bp–6,222 bp	-	3	Yes (3)	CP124972, CP125041, CP125046
		ST1193 (2)	6,200 bp	-	3	Yes (2)	CP125017, CP125023
		ST1284 (1)	6,200 bp	-	3	Yes (1)	CP125063
IncB/O/K/Z (4)	-	ST1193 (3)	95,063 bp–97,447 bp	*bla*$_{CTX-M-55}$ (2)	0–3	No (3)	CP125015, CP125022, CP125065
		ST1284 (1)	95,267 bp	-	3	No (1)	CP124978

IncX1 (1)	-	ST167 (1)	37,322 bp	-	5	No (1)	CP125008
IncX2 (1)	-	ST1284 (1)	34,566 bp	-	2	No (1)	CP125061
IncHI2-IncHI2A (1)	ST1 (1)	ST1193 (1)	327,916 bp	$bla_{CTX-M-15}$ (1)	12	Yes (1)	CP125030

727 [a]blastn searches against the nucleotide collection (nr/nt) of the NCBI were performed to
728 identify identical/closely related plasmids (>99% query coverage and >99% identity) in May
729 2023. Also see **Table S2** for detailed information on blast hits.

730 [b]Various types of IncF replicons were detected but were represented by IncF. Replicon
731 sequences of small plasmids (Col and IncQ) were detected in nine IncF plasmids but are not
732 listed here. Please see **Table S2** for details.

733