Optimised methods for the targeted surveillance of extended-spectrum beta-lactamase producing *Escherichia coli* in human stool

Sarah Gallichan[1], Sally Forrest[1], Esther Picton-Barlow[1], Claudia McKeown[1], Maria Moore[1], Eva Heinz[1,2,] Nicholas A Feasey[1, 4], Joseph M Lewis[1,3], Fabrice E Graf[1, *]

[1]Department of Clinical Sciences, Liverpool School of Tropical Medicine, Liverpool, UK

[2]Department of Vector Biology, Liverpool School of Tropical Medicine, Liverpool, UK

[3]Department of Clinical Infection, Microbiology & Immunology, University of Liverpool, Liverpool, UK

[4]The School of Medicine, University of St Andrews, St Andrews, UK

Keywords

Molecular epidemiology; Antimicrobial resistance; Transmission; Surveillance; Microbiological methods; DNA extraction

***Correspondence:**

Fabrice Graf

Liverpool School of Tropical Medicine

Pembroke Place

Liverpool, UK

Abstract

Understanding transmission pathways of important opportunistic, drug resistant pathogens, such as extended-spectrum beta-lactamase (ESBL) producing *Escherichia coli,* is essential to implementing targeted prevention strategies to interrupt transmission and reduce the number of infections. To link transmission of ESBL-producing *E. coli* (ESBL-EC) between two sources, single nucleotide resolution of *E. coli* strains as well as *E. coli* diversity within and between samples is required. However, the microbiological methods to best track these pathogens are unclear. Here we compared different steps in the microbiological workflow to determine the impact different pre-enrichment broths, pre-enrichment incubation times, selection in pre-enrichment, selective plating, and DNA extraction methods had on recovering ESBL-EC from human stool samples, with the aim to acquire high quality DNA for sequencing and genomic epidemiology. We demonstrate that using a 4-hour pre-enrichment in Buffered Peptone Water, plating on cefotaxime supplemented MacConkey agar and extracting DNA using Lucigen MasterPure DNA Purification kit improves the recovery of ESBL-EC from human stool and produced high-quality DNA for whole genome sequencing. We conclude that our optimised workflow can be applied for single nucleotide variant analysis of an ESBL-EC from stool.

Introduction

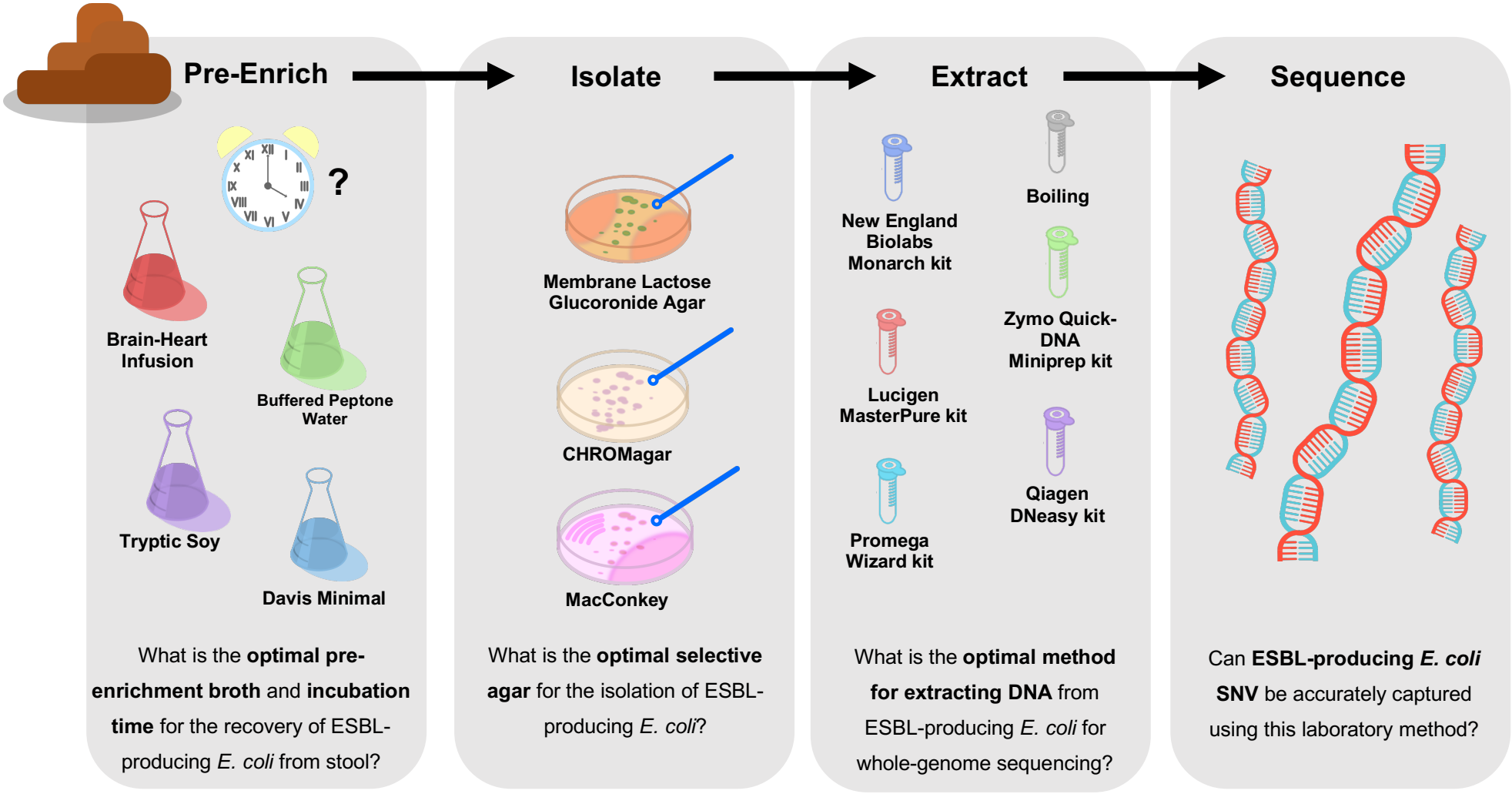
The spread of antimicrobial resistance (AMR) in bacterial pathogens is a global threat to public health. Of particular concern are infections caused by extended-spectrum beta-lactamase (ESBL) producing *Escherichia coli*, which is on the World Health Organisations (WHO) priority AMR pathogen list[1]. ESBL-producing *E. coli* (ESBL-EC) are resistant to aminopenicillins and 3rd-generation cephalosporin (3GC) antibiotics, which are commonly used to treat infections caused by Gram-negative bacteria. ESBL-EC infections can result in a higher mortality and longer hospital stays and may require access to reserve antibiotics such as carbapenems[2–4].

Infection prevention is preferable to cure and will concurrently reduce antibiotic use. To prevent the acquisition of enteric pathogens in care facilities, it is critical to understand their transmission pathways. This is challenging for near ubiquitous organisms such as *E. coli*, which are frequently part of a healthy gut flora and can also be opportunistic pathogens[5]. Typically, human faecal samples have been shown to contain one to three different *E. coli* genotypes at any one time[6] although as many as nine or more different *E. coli* genotypes per stool sample have been described[7]. The number of multidrug resistant *E. coli* strains detected has also been shown to increase with host age[8]. Therefore, differentiating intraspecies diversity of different *E. coli* strains requires comparisons at single nucleotide variation (SNV) resolution[9]. To track transmission events of ESBL-EC in care settings, microbiological methods need to capture the within-host diversity of ESBL-EC within and between different samples accurately and cost-effectively, even when present at low levels. Consensus on which microbiological methods are optimal and a comprehensive comparison of the impact of different microbiological methods are currently lacking.

Currently, the most common healthcare diagnostic method of ESBL-EC isolation from human stool is direct plating on cefotaxime-supplemented MacConkey or chromogenic agar and confirming the production of ESBL using the double disk diffusion method[10,11]. This is often followed by molecular investigation of ESBL resistance genes via multiplex PCR[12,13]. While these methods are effective in identifying the presence of an ESBL-EC to inform treatment options, they lack the resolution to infer transmission because they do not allow us to clearly distinguish between closely related *E. coli*[14]. The distinction between two different clades within the same ST of *E. coli* can be as little as 70 nucleotides, therefore cannot be accurately distinguished using PCR methods[15].

Surveillance studies focused on describing transmission pathways pre-enrich samples with low microbial load, i.e., from rectal swabs, to increase detection of present ESBL-EC[16,17]. Combined with whole-genome sequencing (WGS) this allows for accurate clustering of isolates at SNV resolution. However, there is little consensus between studies, or data on controlled comparisons, on the impact of pre-enrichment broth, pre-enrichment duration, selective agar, and method of DNA extraction has on sample diversity. Here, we describe an optimized approach for targeted surveillance of ESBL-EC from human stool samples at SNV resolution to inform transmission modelling with the potential to capture within-host diversity.

Methods


Pre-Enrich
?
Brain-Heart Infusion
Buffered Peptone Water
Tryptic Soy
Davis Minimal
What is the optimal pre-enrichment broth and incubation time for the recovery of ESBL-producing E. coli from stool?
Isolate
Membrane Lactose Glucoronide Agar
CHROMagar
MacConkey
What is the optimal selective agar for the isolation of ESBL-producing E. coli?
Extract
New England Biolabs Monarch kit
Boiling
Zymo Quick-DNA Miniprep kit
Lucigen MasterPure kit
Qiagen DNeasy kit
Promega Wizard kit
What is the optimal method for extracting DNA from ESBL-producing E. coli for whole-genome sequencing?
Sequence
Can ESBL-producing E. coli SNV be accurately captured using this laboratory method?

112 ***Figure 1*** *Graphical summary of the main research questions and comparisons preformed for each step of the microbiological*

113 *processing of ESBL-EC from stool*

114 **ESBL-EC growth dynamics in different pre-enrichment broths**

115 Pre-enrichment broths (Brain-Heart Infusion, Tryptic Soy, Buffered Peptone Water and
116 Davis Minimal) commonly used in the recovery of ESBL-EC from clinical sources were
117 selected for comparison in this study. These pre-enrichment broths were prepared
118 according to the manufacturer's recommendations **(Supplementary Table 1)**. We
119 used three reference strains encoding ESBLs: NCTC 13341 ($bla_{CTX-M-15}$); NCTC 13476
120 (bla_{IMP}); NCTC 13846 ($bla_{CTX-M-27}$), with the latter also encoding plasmid-derived
121 colistin resistance **(Supplementary Table 2)**. We further included a clinical isolate,
122 CAB17W, encoding $bla_{CTX-M-15}$ **(Supplementary Table 2)**. The growth of single ESBL-
123 EC isolates in different pre-enrichment broths was determined to establish growth
124 dynamics. Growth of the reference ESBL-EC (NCTC 13441) in four pre-enrichment
125 broths (BPW, DM, BHI, and TS) was monitored every 10 minutes through regular
126 optical density readings at 600 nm wavelength in the CLARIOstar plate reader (BMG
127 Labtech, Ottenberg, Germany) for 24 h. A population logistic model was fitted to the
128 resulting data with the Growthcurver package[18]. The recovery of *E. coli* from different
129 pre-enrichment broths with varying incubation times was tested. *E. coli* (NCTC 13441)
130 was inoculated (approximately 2 x 10^7 CFU/mL) into 5 mL of four pre-enrichment
131 broths (BPW, DM, BHI, and TS) and grown shaking at 37 °C and 220 rpm. After 2, 4,
132 6 and 8 hours, 1:10 serial dilutions of each pre-enrichment culture were performed
133 and dilutions 10^{-4} to 10^{-12} were spotted onto LB agar and incubated at 37°C overnight.
134 The following day, the number of colony forming units (CFUs) recovered from each
135 broth was calculated.

136 Stool was collected from three healthy volunteers, defined as individuals not using
137 antibiotics or experiencing any gastrointestinal problems. The stool was resuspended
138 in a stool diluent solution (1:5 w/v) prepared according to the maltodextrin-trehalose

139 method with the addition of 10% glycerol [19]. The stool slurry was mixed using a
140 magnetic stirrer for 15 minutes and allowed to sediment for 5 minutes. Supernatants
141 from each of the three stool slurries were mixed in equal proportions and stored in
142 cryotubes at - 80 °C. Before use in experiments, the stool slurry was rapidly thawed in
143 a water bath at 37°C for 10 minutes and vortexed thoroughly. To determine the optimal
144 pre-enrichment broth and incubation time for the recovery of ESBL-EC from stool, a
145 spiked stool model was developed using 1 mL of the pooled stool slurry and a 1 μL
146 spike of the NCTC 13441 or CAB17W (approximately 1 x 10^5 CFU/mL). The ESBL-
147 non-producing *E. coli* strain (NCTC 12241) was used as a negative control to
148 determine if the antibiotic supplement sufficiently inhibited growth. The spiked stool
149 models were pre-enriched using TS and BPW with and without the addition of
150 selection (1 μg/ mL cefotaxime) and incubated for 4 and 18 hours, shaking at 37 °C
151 and 220 rpm. Following incubation, 1:10 serial dilutions of each pre-enrichment culture
152 were made and dilutions (10^{-4} to 10^{-7} for the 4-hour incubation and 10^{-5} to 10^{-8} for the
153 18-hour incubation) were plated onto LB agar and CHROMagar in 10 μL spots and
154 incubated at 37 °C overnight. The number of CFUs recovered from each broth was
155 calculated.
156

157 **Recovery of ESBL-EC using different selective agars**

158 Selective agar (MacConkey, Membrane Lactose Glucoronide Agar (MLGA) and
159 CHROMagar) commonly used in the recovery of ESBL-EC from environmental, clinical
160 and food sources were selected for comparison in this study. These selective agars
161 were prepared according to the manufacturer's recommendations **(Supplementary**
162 **Table 1)**. MacConkey and MLGA agars were supplemented to a final concentration of
163 1 µg/mL cefotaxime (Melford, Suffolk, UK). To assess the recovery of ESBL-EC from

164 stool using selective agar, the stool spike model described above was used. The
165 spiked stool was pre-enriched in BPW for 4 hours, shaking at 37°C and 220 rpm.
166 Following incubation, 1:10 serial dilutions of each spiked stool sample were made, and
167 10^{-4} to 10^{-7} dilutions were plated onto MacConkey agar, CHROMagar and MLGA
168 plates in 10 μL spots and incubated at 37 °C overnight. The number of CFUs
169 recovered from each selective agar was calculated.
170

171 **Limit of ESBL-EC detection with pre-enrichment compared with direct plating**

172 A series of 100 μL of CAB17W ranging from 1 CFU/mL to 1 x 10^8 CFU/mL was spiked
173 into 100 μL of stool slurry. The spiked stools were serially diluted (1:10) and 10 μL of
174 each dilution was spotted onto cefotaxime supplemented MacConkey agar and
175 incubated at 37 °C overnight. A 96 deep well plate was then inoculated with 10 μL of
176 the remaining stool spikes in 1 mL of BPW, this was done in triplicate for each spike
177 input concentration. The deep well plate was then sealed with a breathable sealing
178 membrane and placed in an Innova 42R 19 mm Orbit shaking incubator set at 37 °C
179 and 220 rpm. After 4 hours the plate was removed from the incubator, each well was
180 again diluted in series 1:10 and spots of 10μL were plated onto cefotaxime
181 supplemented MacConkey and incubated at 37°C overnight. The number of CFUs
182 recovered were then calculated.
183

184 **DNA extraction kit comparison**

185 A single colony of each of four *E. coli* strains (**Supplementary Table 2:** NCTC 13441,
186 NCTC 13476, NCTC 13846, CAB17W) were inoculated in 5 mL of LB with and grown
187 overnight. Then, cultures were diluted to an OD600 of 2 (~ 2 x 10^9 CFU/mL). The cells
188 were washed three times in phosphate-buffered saline (PBS). The resulting pellets

189 were resuspended in 1 mL of PBS and split into three tubes with 250 µL and used as
190 the input for each extraction method. To assess the yield, quality, efficiency, and cost
191 of five commercially available kits **(Supplementary Table 1)** and a simple boiling
192 method[20], we extracted DNA from each of the four *E. coli* isolates in technical
193 triplicates (n=12 for each extraction method), with three experimental replicates (n=36
194 extractions per method) on different days. The DNA yield was then measured using
195 the Qubit 4 Fluorometer (Agilent Technologies, California, USA) and the Agilent
196 TapeStation System 4150 (Agilent Technologies, California, USA) was used to
197 determine the DNA integrity. The DNA quality was assessed by measuring the
198 A260/230 and A260/280 absorbance ratios using the NanoPhotometer (Implen,
199 California, United States). Scores **(Supplementary Table 3)** were allocated according
200 to the criteria in **Supplementary Table 4**. Each DNA extraction method was then
201 ranked according to the score given, from highest (1) to lowest (6) preforming in each
202 category: DNA yield, DNA quality, cost, hands-on time, ease of protocol and total
203 protocol time.
204

205 **Single nucleotide variant analysis**

206 To determine whether our protocol would provide sufficient resolution for the detection
207 of single nucleotide variants in a stool sample, we processed a spiked stool sample
208 and an ESBL-EC positive rectal swab. For the spiked stool sample, we inoculated 10
209 µL of a 100-fold diluted overnight CAB17W culture (~ 8 x 10^7 CFU/mL) into 1 mL of
210 the pooled stool slurry, vortexed the mixture thoroughly and added 100 µL to 5 mL of
211 BPW in a 15 mL Falcon tube. The ESBL-EC positive rectal swab was obtained with
212 permission from a participant in an observational cohort study of faecal ESBL carriage
213 in hospital patients and care home residents across Liverpool, UK. The rectal swab

214 was placed directly in 5 mL of BPW in a 15 mL Falcon tube. Falcon tubes were
215 incubated for 4 hours at 37 °C, shaking at 220 rpm. After 4 hours, 100 μL of the culture
216 was spread on cefotaxime supplemented (1 μg/mL) MacConkey agar and incubated
217 at 37 °C overnight. The following day, we picked 7 single colonies, restreaked each
218 colony on another cefotaxime supplemented MacConkey agar and incubated at 37 °C
219 overnight. DNA was then extracted from each of the purified colonies using the
220 MasterPure Complete DNA and RNA Purification Kit (Lucigen, Wisconsin, USA)
221 according to the manufacturer's recommendations. The DNA concentrations were
222 then measured using the Qubit 4 Fluorometer (Agilent Technologies, California, USA)
223 and the DNA sent for whole-genome sequencing using the NovaSeq platform (Illumina
224 Inc, California, USA) at Azenta Life Sciences (Frankfurt, Germany). Sequence data
225 (reads) were obtained as fastq files. Sequencing adaptors were removed from the
226 reads using Trimmomatic version 0.39[21] and the quality of the trimmed reads was
227 assessed using the quality metrics provided by FastQC version 0.11.9[22]. ARIBA[23] was
228 used to call the AMR genes using the CARD database[24] and to determine ESBL-EC
229 multilocus sequence type as defined by the seven-gene Achtman scheme[25] hosted at
230 pubMLST (https://pubmlst.org/). Single nucleotide variants (SNVs) were called
231 between the relevant reference genome (ST167: GCF_025398915.1; ST131:
232 GCF_004358405.1; ST1193: GCF_003344465.1) and the single pick sequences
233 using snippy version 4.3.6 (https://github.com/tseemann/snippy). We then generated
234 a core genome alignment of all the single pick sequences using snippy-core (with
235 default settings). All non-nucleotide characters and potentially recombinant sequence
236 regions were removed from the multisequence alignment using the snippy-
237 clean_full_aln and Gubbins version 2.3.4[26] respectively. We then extracted the SNVs
238 from the cleaned multisequence alignment using snp-sites (using the -c option) and

239 created a pairwise SNV distance comparison between each single pick sequence
240 using snp-dists version 0.8.2 (https://github.com/tseemann/snp-dists).
241
242 **Statistical analysis**
243 Measurements were summarised as mean and standard deviation or 95% confidence
244 interval, calculated as one standard deviation above and below the mean value. Box-
245 and-whisker plots, where presented, show medians as a horizontal line, 25th and 75th
246 centiles as boxes, and at most 1.5 times interquartile range as whiskers. Analysis of
247 variance (ANOVA) tests were used to determine the significance of any observed
248 differences between groups. $P < 0.05$ was considered statistically significant and,
249 where appropriate, P values were adjusted for multiple comparisons using the
250 Bonferroni correction. A post-hoc Tukey test was used for pairwise sample
251 comparisons when ANOVA tests were significant. All statistical analysis was
252 performed using R version 4.1.3[27]. All R code is accessible from our Github repository
253 (https://github.com/SarahGallichan/ESBL-EC_methods.git)
254
255 **Results**
256 **ESBL-EC recovery using different pre-enrichment conditions and selective**
257 **agars**
258 We assessed the growth dynamics of two ESBL-EC strains (NCTC 13441 and
259 CAB17W) in different pre-enrichment broths to determine the optimal incubation time
260 for recovery of bacteria in log growth phase **(Supplementary Figure 1).** To summarise
261 the dynamics of the resulting growth curves, the data was fitted to a population logistic
262 model **(Supplementary Table 5)**. ESBL-EC reached midpoint of log-phase after
263 approximately 4 hours in BHI, TS, and DM; and after approximately 7 hours in BPW.

264 Therefore, we chose incubation times of between 2 and 8 hours to quantify ESBL-EC
265 recovery from different broths and assessed bacterial count in CFU from each broth
266 after 4 incubation time points (2, 4, 6 and 8 hours). Across all 4 time points, BHI
267 resulted in the highest and DM the lowest *E. coli* recovery **(Supplementary Figure 2)**.
268 We performed pair-wise broth comparisons at each time point. There was no
269 significant difference between *E. coli* growth in BHI, BPW and TS across all 4 time
270 points **(Supplementary Table 6)**. However, *E. coli* counts in DM was significantly
271 lower than BHI and BPW after 2 hours; BHI, BPW and TS after 4 hours; and BHI and
272 TS after 8 hours of incubation. The largest increase in amount of *E. coli* CFU recovered
273 between the two time points for each broth was from 2 to 4 hours of incubation
274 **(Supplementary Table 6)**, which is within the log growth phase of *E. coli* in each broth.
275 Therefore, we selected 4 hours of incubation to investigate the growth of ESBL-EC
276 from spiked stool.
277
278 We selected TS (a high nutrient broth with multiple amino acid sources) and BPW (a
279 low nutrient broth with one amino acid source) for the spike experiments, adding
280 defined amounts of ESBL-EC (NCTC 13441 and CAB17W) to stool samples. Donated
281 stool was confirmed to have no growth of cefotaxime (a 3GC) resistant bacteria by
282 plating on 1µg/mL cefotaxime supplemented MacConkey. We assessed the recovery
283 of these spiked strains, testing whether the addition of selection with cefotaxime and
284 the length of incubation time influenced the concentration of spiked ESBL-EC strains
285 we were able to recover **(Figure 2)**. After 4 hours of incubation, there was no
286 significant difference in CAB17W growth (i.e. CFU concentration) between TS and
287 BPW and whether antibiotic was supplemented or not. However, at 4 hours the growth
288 of NCTC 13441 in the TS without antibiotic supplement was 2.75 CFU/mL (CI: 2.28 –

289 3.33, p = 0.0004) more than in BPW without supplement, while there was no significant
290 difference between the growth in either broth at 4 hours when supplemented with
291 cefotaxime.
292
293 At 18 hours the effects of the addition of cefotaxime and broth on growth were different
294 for each strain. Antibiotic supplement significantly increased the recovery of NCTC
295 13441 after 18 hours compared to no supplementation in TS, by 100.78 CFU/mL (CI:
296 39.81 – 251.19, p = 0.01). There was no significant difference in growth of NCTC
297 13441 at 18 hours between TS and BPW. However, recovery of CAB17W was
298 significantly higher in BPW after 18 hours of incubation compared to TS, by a
299 difference of 68 CFU/mL (CI: 51.11 – 91.26, p = 0.0001) with antibiotic supplement
300 and 41.16 CFU/mL (CI: 33.15 – 51.10, p < 0.0001) without an antibiotic supplement.
301 While an 18-hour pre-enrichment in antibiotic supplemented broth best recovered the
302 highest CFU/mL for both NCTC 13441 and CAB17W, there were inconsistencies
303 between broths depending on the strain. Therefore, we concluded that a 4-hour pre-
304 enrichment in BPW was better for a more consistent increase in both ESBL-EC strains
305 without requiring an antibiotic supplement.
306 Next, we then compared ESBL-EC recovery on three different ESBL-EC selective
307 agars (CHROMagar, MLGA and MacConkey) from our stool spike models. ESBL-EC
308 was recovered from all agars, with no significant differences in CFU/mL
309 **(Supplementary Table 8)**. However, there was a clear difference in the cost per
310 sample of each selective agar, with MacConkey being the most cost-effective and
311 CHROMagar being the most expensive **(Supplementary Table 1)**.
312

313 The limit of ESBL-EC detection from CAB17W spiked stool was then determined for
314 4-hour pre-enrichment in BPW and compared with directly plating on cefotaxime
315 supplemented MacConkey. Overall, the mean difference in ESBL-EC recovery
316 between 4-hour pre-enrichment in BPW and direct plating was 47 CFU/mL (CI: 9.65 –
317 233.08, $p < 0.001$) **(Supplementary Figure 3)**. The limit of detection using direct
318 plating was 1.91 CFU/mL and with a 4-hour pre-enrichment, the limit of detection
319 improved to 0.05 CFU/mL.
320
321 **Large variability in yield and quality of DNA using different extraction methods**
322 DNA from the same input concentration of four different *E. coli* strains
323 **(Supplementary Table 2)** was extracted using five commercially available kits **(Table**
324 **3)** and a simple boiling method. The DNA yield was measured using the Qubit
325 fluorometer **(Figure 3)**. The boiling method resulted in the highest mean DNA yield
326 overall for the Qubit (16.66 ng/mL, 95% CI: 16.16 – 17.16) while the DNeasy kit
327 resulted in the lowest (3.28 ng/mL; 95% CI: 4.28 – 2.28) **(Figure 3)**.

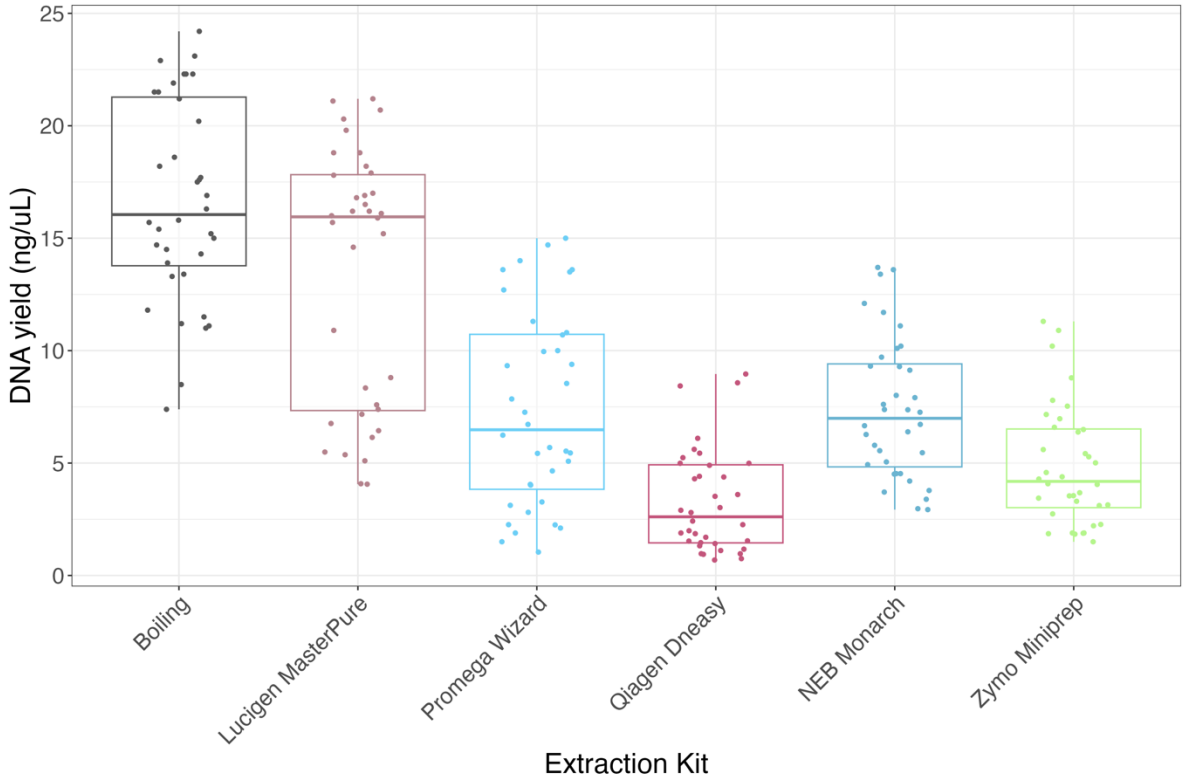


Figure 3 *DNA yield from the six extraction methods measured using the Qubit. Technical (each sample extracted in triplicates) and experimental replicates (each extraction method performed three separate times) are plotted for each extraction kit (n = 36). The outliers are shown as points outside of the boxplot whiskers.*

328

329 The DNA quality was assessed using the A260/280 absorbance ratio. Downstream
330 applications requiring high quality DNA, such as short-read and especially long-read
331 sequencing platforms, require that the A260/A280 absorbance ratio is in the range
332 between 1.8 to 2.1. The Promega Wizard kit and the DNeasy kit had the largest
333 proportion of technical and experimental replicates (67%) within this range while the
334 Zymo Miniprep kit had the lowest proportion (11%) **(Figure 4A).** Additionally, the DNA
335 integrity was assessed with the DNA Integrity Number (DIN), which is a numerical
336 assessment (1 to 10) of the DNA integrity determined by the TapeStation system
337 algorithm. A DIN score of 1 indicated very degraded DNA and DIN 10 indicated highly
338 intact DNA. The Lucigen MasterPure and the Zymo miniprep kits had the highest DIN
339 followed closely by the Promega Wizard **(Table 1).**

340

341 ***Table 1*** *Mean of measurements (technical and experimental replicates averaged*
342 *across three independent experiments) for each DNA extraction method, including*
343 *the 95% confidence intervals, for the tested E. coli and proportion (percentage of*
344 *technical replicates) within defined absorbance ratio range.*

345

	Boiling	DNeasy	Lucigen MasterPure	Promega Wizard	NEB Monarch	Zymo Miniprep
DNA yield (ng/ µL)	16.66 ± 4.66	3.28 ± 1.00	13.37 ± 0.58	7.37 ± 1.62	7.35 ± 0.97	4.85 ± 0.48
Absorbance ratio (A260/A280)	1.74 ± 0.06	1.99 ± 0.15	2.1 ± 0.03	2.06 ± 0.13	1.74 ± 0.15	1.66 ± 0.17
Proportion within range (1.8 – 2.1)	14%	67%	50%	67%	42%	11%
Absorbance ratio (A260/A230)	0.70 ± 0.04	3.00 ± 0.73	2.89 ± 0.60	4.44 ± 1.27	1.2 ± 0.63	1.74 ± 0.41
Proportion within range (2.0 – 3.0)	0%	50%	59%	0%	8%	0%
DNA Integrity Number	1.97 ± 1.56	5.40 ± 0.50	7.40 ± 0.35	7.18 ± 0.45	6.08 ± 0.45	7.40 ± 0.25

346

347 The purity of extracted DNA was assessed using the 260/230 absorbance ratio. DNA
348 with a 260/230 absorbance ratio of 2.0 – 3.0 are considered pure and mostly free of
349 contaminants (such as phenols) [28]. Lucigen Masterpure had the largest proportion of
350 replicates (59%) within the suggested A260/A230 absorbance range, while the simple
351 boiling method, the Zymo miniprep kit and the Promega Wizard kit failed to yield DNA
352 within this purity range **(Figure 4B).**

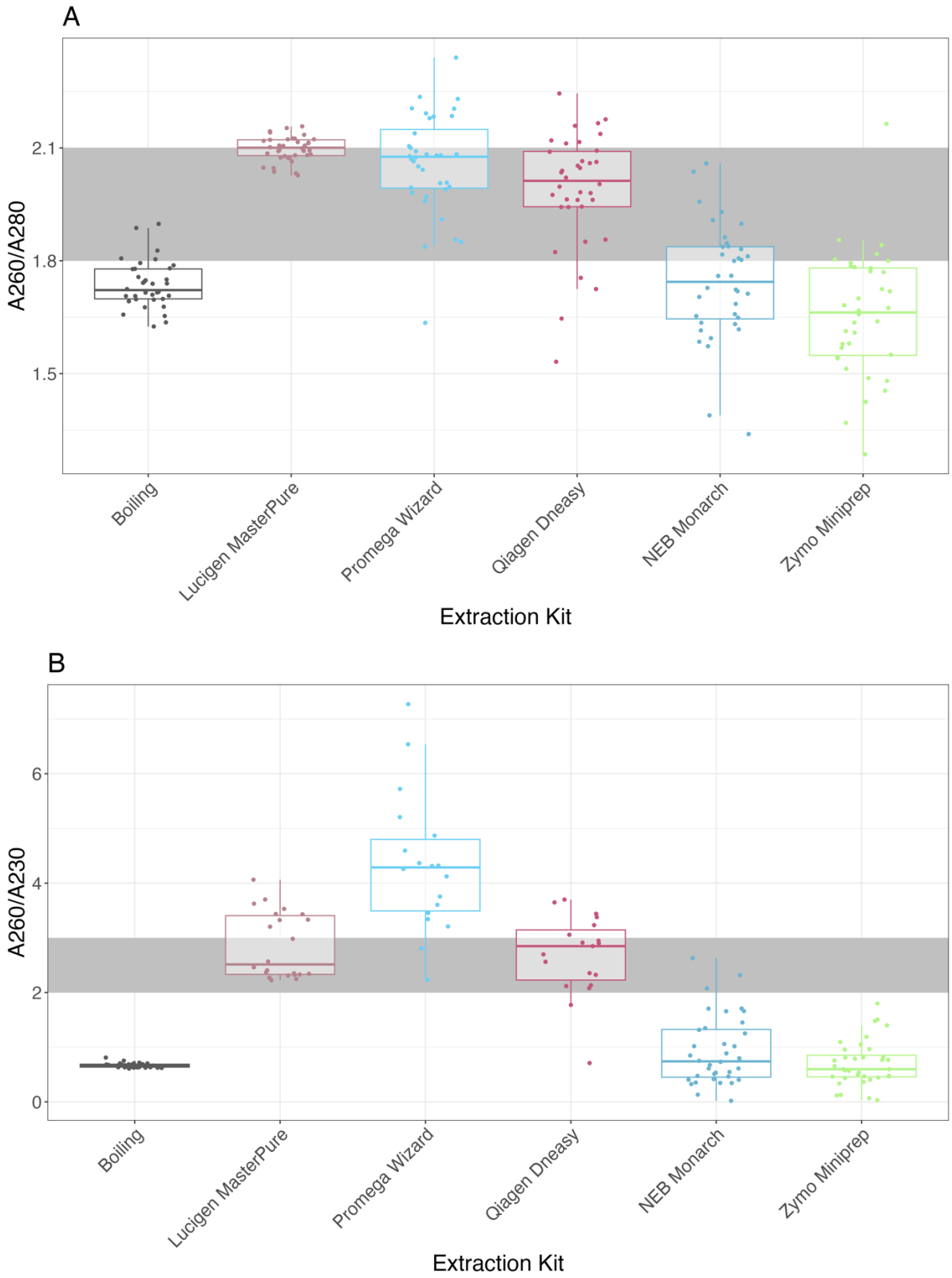


Figure 4 *Absorbance ratios* ***(A)*** *A260/280 and* ***(B)*** *A260/230 indicating the presence of contamination and quality of DNA, respectively, of the extracted DNA extracted using six extraction methods. Technical (each sample extracted in*

triplicates) and experimental replicates (each extraction method preformed three separate times) are plotted for each extraction kit (n = 36). The outliers are shown as points outside of the boxplot whiskers.

353 Using all the measurement data, cost, ease of protocol (technical difficulty and
354 equipment requirements), and time aspects (hands-on and total-protocol time) from
355 the comparison of the five commercially available DNA extraction kits, a ranking matrix
356 **(Table 2)** was produced based on the scores for each extraction method
357 **(Supplementary Table 3)**. Simple boiling scored highly for yield, cost, and time but
358 low for the purity scores (A260/230, A260/280 and DIN). Commercial kits performed
359 better in this regard, with the Promega HMW and Qiagen DNeasy ranking highest for
360 the A260/280 and the MasterPure ranking highest for the A260/230 absorbance ratio
361 and the DIN score. All other rankings were variable between kits. The Lucigen
362 MasterPure Complete DNA and RNA Purification Kit performed best for the extraction
363 of high yield and quality DNA for sequencing to investigate single nucleotide variants
364 of ESBL-EC in stool.

365
366
367 ***Table 2*** *Ranking of different DNA extraction methods, highest (1) to lowest (6).*
368

DNA extraction kit	Cost	Time		Mean DNA yield (ng/ µL)	DNA quality	Difficulty
		Hands-on	Total			
Qiagen DNeasy	5	2	3	2	5	3
Lucigen Masterpure	2	4	5	1	1	6
Promega HMW	5	4	5	2	1	5
NEB Monarch	3	1	2	2	4	2
Zymo Miniprep	4	2	3	2	1	3
Boiling	1	1	1	1	6	1

369

370 **Validation of workflow**

371 To investigate whether our workflow of a 4-hour pre-enrichment in BPW, plating on
372 cefotaxime supplemented MacConkey and extraction using the MasterPure kit was
373 sufficient for investigation of SNVs in closely related ESBL-EC strains, we performed
374 both a control experiment where we processed an ESBL-negative stool spiked with a
375 single clinical ESBL-EC strain (CAB17W) and assessed an ESBL-EC positive patient
376 rectal swab using our workflow. Seven colonies from each of the cefotaxime
377 supplemented MacConkey agar plates were then picked and sequenced. SNV
378 analysis of the spiked stool confirmed that there were no SNVs between any of the
379 single colony pick genomes as expected for a clonal spike **(Supplementary Table**
380 **10)**. Analysis of the ESBL-EC positive rectal swab revealed that the patient had two
381 different ESBL-EC sequence types ST131 that had $bla_{CTX\text{-}M\text{-}27}$ and ST1193 that had
382 $bla_{CTX\text{-}M\text{-}15}$. SNV analysis of each ESBL-EC sequence type showed that there were 5
383 SNV differences between the ST131 single picks **(Supplementary Table 11)** and 1
384 SNV difference between one of the ST1193 picks and the other four single picks
385 **(Supplementary Table 12)**. We therefore concluded that our workflow was
386 appropriate for SNV analysis of ESBL-EC from stool.

387

388 **Discussion**

389 Although *E. coli* is the dominating species of Enterobacteriaceae in the gut of healthy
390 human adults, *E. coli* only constitute around 0.5% - 5% of the gut microbial community
391 [29]. This makes it difficult for methods such as shotgun metagenomics to disentangle
392 the diversity of pathogenic and non-pathogenic *E. coli* within a highly complex human
393 gut microbiome without using costly very deep sequencing [6,30]. On the other end of the
394 spectrum, the analysis of single colony picks from stool samples directly plated on

395 antibiotic supplemented agar, as is practice in many routine diagnostic laboratories, is
396 insufficient to capture the intraspecies diversity of ESBL-EC, which can be substantial,
397 as each colony would need to be analysed which would be extremely time-, cost- and
398 labour-intensive[31,32]. Both methods are thus not feasible to implement as routine tool
399 to monitor transmission of ESBL-EC in healthcare settings. Therefore, we investigated
400 the impact of different microbiological approaches to optimise the recovery of ESBL-
401 EC from human stool. We compared different methods at each stage of laboratory
402 processing; pre-enrichment, isolation on selective agar, and DNA extraction, and were
403 able to optimise a rapid and cost-effective approach for the recovery of ESBL-EC for
404 targeted surveillance of ESBL-EC colonisation with the potential for sub-species
405 resolution.
406
407 Numerous transmission studies have highlighted the challenge of capturing ESBL-EC
408 variants present in small numbers in stool[33–35]. *E. coli* are only a small part of the
409 diversity present within complex matrices and clinical samples (e.g., stool), increasing
410 the likelihood of a test resulting in false negatives when stool is directly plated on
411 supplemented agar[36,37]. To improve ESBL-EC detection for targeted surveillance,
412 previous studies used pre-enrichment broth before plating on supplemented agar and
413 were successful in recovering multiple ESBL-EC genotypes from human stool[17].
414 Indeed, we also showed that pre-enrichment is an effective tool to increase sensitivity
415 for ESBL-EC detection from stool, thus improving the amount of ESBL-EC recovered
416 on supplemented agar.
417
418 The standard practice of pre-enriching a sample for 18 to 24 hours is not conducive to
419 rapid diagnosis and patient intervention[38,39]. A further concern with longer pre-

420 enrichment is the increased risk of creating competition with other lineages or species
421 and skewing the relative abundance. Further, the potential of movement of mobile
422 genetic elements (e.g. plasmids) can bias the recovery of the original strain[40]. We
423 found that growth of ESBL-EC was variable and strain-dependent after an 18-hour
424 pre-enrichment with and without the addition of an antibiotic supplement, suggesting
425 that there is indeed potential for competition bias with longer incubation times.
426 Therefore, we investigated a short pre-enrichment and found that a 4-hour pre-
427 enrichment yielded sufficient growth for reliable downstream identification of ESBL-
428 EC. Hence, we concluded that 4-hour pre-enrichment was not only preferable for
429 consistent ESBL-EC recovery but also did not necessitate the addition of an antibiotic
430 supplement and reduced the time and cost.
431
432 There are a variety of recommended supplemented agars available for ESBL-EC
433 detection. Cefotaxime supplemented MacConkey agar is commonly used for
434 surveillance studies and clinical microbiology laboratories since it is readily available
435 and the cheapest of the three agars compared here[10,11,41]. However, chromogenic
436 agar (CHROMagar ESBL[42]) is an increasingly popular alternative to MacConkey since
437 it is effective at distinguishing between *E. coli* and other ESBL-producing
438 Enterobacteriaceae using colour[38,43,44]. In addition to these two agars, we also tested
439 ESBL-EC recovery on cefotaxime supplemented Membrane Lactose Glucuronide
440 Agar (MLGA), which is commonly used for detecting faecal contamination in water
441 sources[45]. We found that all three agars were effective in recovering from ESBL-EC
442 from pre-enriched stool, allowing for flexibility with selective agar choice; we hence
443 selected cefotaxime supplemented MacConkey as the most cost-effective in our
444 setting.

445

446 We tested a wide variety of DNA extraction methods including commercial DNA
447 purification kits and the boiling method as the simplest and cheapest method of DNA
448 extraction from *E. coli*[46]. Since transmission studies require high resolution typing to
449 clearly distinguish between two *E. coli* strains, high quality DNA is required.
450 Commercially available kits provide standardised reagents, consumables and
451 validated DNA extraction and purification methods. Depending on the extraction
452 protocol, specialised equipment e.g. refrigerated centrifuges, skilled personnel, and
453 an increased amount of time and cost are needed to extract DNA[47]. Taking this into
454 consideration, along with the cost-per-sample, quality, and yield of DNA we compared
455 six DNA extraction methods. While the boiling method was the simplest, cheapest,
456 least time-consuming DNA extraction method with the highest yield and thus arguably
457 of high value for highly time-sensitive applications or in cost-restricted setting, we show
458 that the Lucigen MasterPure produced comparable DNA yield with much higher quality
459 scores that were consistently reproducible.

460

461 Finally, we validated our optimised ESBL-EC recovery method using a spiked stool
462 sample and a patient rectal swab. For the spiked stool sample, all sequenced 7
463 colonies were identical and clonality was confirmed at single nucleotide resolution.
464 With the patient rectal swab, we demonstrated the value of multiple colony picks
465 through the recovery of two different ESBL-EC sequence types and the analysis of
466 variation within each sequence type at single nucleotide resolution. The analysis of
467 both samples at suggests that our method is suitable for studies aiming to infer
468 transmission using SNV. While the methods we show here have focused on the
469 recovery of ESBL-EC from stool, we suggest that these methods are also applicable

470 or easily adaptable for the recovery other drug resistant Enterobacteriaceae. ESBL-
471 producing *Klebsiella* and *Enterobacter* species occupy a similar niche to ESBL-EC and
472 are all recovered from stool through direct plating on supplemented agar[10]. With the
473 increasing prevalence of drug resistant infections worldwide, it has become crucial to
474 understand transmission pathways to put targeted preventative measures in place to
475 interrupt transmission.
476
477 Our study demonstrates that a 4-hour pre-enrichment in either BPW or TS is effective
478 for the recovery of ESBL-EC from stool without the addition of a third-generation
479 cephalosporin and cefotaxime supplemented MacConkey, chromogenic agar and
480 MLGA are all viable options for the recovery of pre-enriched ESBL-EC. The quality
481 and quantity of DNA using different extraction kits is variable and not all yield high-
482 quality DNA required for whole genome sequencing. The MasterPure Complete DNA
483 and RNA Purification Kit produced a DNA yield that was comparable to the simple
484 boiling method, but with higher quality scores that were consistently reproducible.
485 Altogether this optimised method was sufficient for SNV analysis of an ESBL-EC
486 spiked stool sample.
487
488 **Ethics**
489 The healthy volunteers gave confidential consent for the use and storage of stool
490 samples (LSTM Research Tissue Bank RTB/2022/007). The observational cohort
491 study of hospital patients and care home residents in facilities in Liverpool was
492 approved by the National Research Ethics Service Greater Manchester South ethics
493 committee (ref: 22/NW/0343). Written informed consent was obtained from
494 participants or consultees, as appropriate.

495

496 **Supplementary materials and data availability**

497 Supplementary Tables 1 – 12 and Supplementary Figures 1 - 3 are accessible in the
498 supplementary material of this manuscript. The R scripts used to generate the figures,
499 supplementary figures and data analyses in this manuscript are available from the
500 GitHub repository https://github.com/SarahGallichan/ESBL-EC_methods.git.
501 Protocols are available on Protocols.io
502 (dx.doi.org/10.17504/protocols.io.kxygxyk3dl8j/v). Reads from sequenced isolates in
503 this study are accessible in the Sequence Read Archive (SRA) using BioProject ID:
504 PRJNA1095376.

505

506 **Funding**

507 This work was supported by iiCON (infection innovation consortium) via UK Research
508 and Innovation (107136) and Unilever (MA-2021-00523N).

509

510 **Author contributions**

511 Conceptualisation and method development was done by S.G., S.F., M.M., E.H.,
512 J.M.L., N.A.F., and F.E.G. Investigation was undertaken by S.G., S.F., E.P.-B, and
513 C.M. Data analysis was done by S.G., J.M.L. and F.E.G. The original draft was
514 prepared by S.G., E.H., J.M.L., N.A.F., and F.E.G and then reviewed and edited by all
515 authors. Supervision was provided by E.H., J.M.L., N.A.F., and F.E.G.

516

517 **Acknowledgements**

518 We would like to thank Ross Gray for the illustrations used in Figure 1 of this
519 manuscript, Finlay Hitchman for his assistance with the pre-enrichment broth

520 comparison work, and all the clinical staff and study participants that facilitated this
521 study.

522 **References**

523 1. WHO publishes list of bacteria for which new antibiotics are urgently needed.
524 https://www.who.int/news/item/27-02-2017-who-publishes-list-of-bacteria-for-
525 which-new-antibiotics-are-urgently-needed.

526 2. Schwaber, M. J. & Carmeli, Y. Mortality and delay in effective therapy associated
527 with extended-spectrum beta-lactamase production in Enterobacteriaceae
528 bacteraemia: a systematic review and meta-analysis. *J Antimicrob Chemother*
529 **60**, 913–920 (2007).

530 3. Melzer, M. & Petersen, I. Mortality following bacteraemic infection caused by
531 extended spectrum beta-lactamase (ESBL) producing E. coli compared to non-
532 ESBL producing E. coli. *Journal of Infection* **55**, 254–259 (2007).

533 4. Blom, A., Ahl, J., Månsson, F., Resman, F. & Tham, J. The prevalence of ESBL-
534 producing Enterobacteriaceae in a nursing home setting compared with elderly
535 living at home: a cross-sectional comparison. (2016) doi:10.1186/s12879-016-
536 1430-5.

537 5. Bettelheim, K. A. Escherichia coli in the normal flora of humans and animals. in
538 *Escherichia coli: mechanisms of virulence* (ed. Sussman, M.) 85–109
539 (Cambridge University Press, Cambridge, 1997).

540 6. Foster-Nyarko, E. & Pallen, M. J. The microbial ecology of Escherichia coli in
541 the vertebrate gut. *FEMS Microbiol Rev* **46**, (2022).

542 7. Alm, E. A. , W. S. T. , and G. D. M. The niche of Escherichia coli. in *In Population*
543 *Genetics of Bacteria: A Tribute to Thomas S. Whittam* (eds. Walk, S. T. & Feng,
544 P. C. H.) 68–89 (ASM Press, Washington, DC, USA, 2011).

545 8. Vollmerhausen, T. L. *et al.* Population structure and uropathogenic virulence
546 associated genes of faecal Escherichia coli from healthy young and elderly
547 adults. *J Med Microbiol* **60**, 574–581 (2011).
548 9. Balloux, F. *et al.* From Theory to Practice: Translating Whole-Genome
549 Sequencing (WGS) into the Clinic. *Trends Microbiol* **26**, 1035–1048 (2018).
550 10. Public Health England. *UK Standards for Microbiology Investigations Detection*
551 *of Enterobacteriaceae Producing Extended Spectrum β-Lactamases-*
552 *Standards-for-Microbiology-Investigations-Smi-Quality-and-Consistency-in-*
553 *Clinical-Laboratories*. (2016).
554 11. World Health Organization & Advisory Group on Integrated Surveillance on
555 AMR. *WHO Integrated Global Surveillance on ESBL-Producing E. Coli Using a*
556 *"One Health" Approach: Implementation and Opportunities*.
557 https://www.who.int/publications/i/item/9789240021402 (2021).
558 12. Bradford, P. A. Extended-spectrum beta-lactamases in the 21st century:
559 characterization, epidemiology, and detection of this important resistance threat.
560 *Clin Microbiol Rev* **14**, 933–951 (2001).
561 13. Naas, T., Cotellon, G., Ergani, A. & Nordmann, P. Real-time PCR for detection
562 of blaOXA-48 genes from stools. *J Antimicrob Chemother* **68**, 101–104 (2013).
563 14. Wu, G. *et al.* Comparative Analysis of ESBL-Positive Escherichia coli Isolates
564 from Animals and Humans from the UK, The Netherlands and Germany. *PLoS*
565 *One* **8**, e75392 (2013).
566 15. Petty, N. K. *et al.* Global dissemination of a multidrug resistant Escherichia coli
567 clone. *Proc Natl Acad Sci U S A* **111**, 5694–5699 (2014).

568 16. Toombs-Ruane, L. J. *et al.* Carriage of Extended-Spectrum-Beta-Lactamase-
569 and AmpC Beta-Lactamase-Producing Escherichia coli Strains from Humans
570 and Pets in the Same Households. *Appl Environ Microbiol* **86**, (2020).
571 17. Ludden, C. *et al.* Defining nosocomial transmission of Escherichia coli and
572 antimicrobial resistance genes: a genomic surveillance study. *Lancet Microbe*
573 **2**, e472–e480 (2021).
574 18. Sprouffske, K. & Wagner, A. Growthcurver: An R package for obtaining
575 interpretable metrics from microbial growth curves. *BMC Bioinformatics* **17**, 1–4
576 (2016).
577 19. Burz, S. D. *et al.* A Guide for Ex Vivo Handling and Storage of Stool Samples
578 Intended for Fecal Microbiota Transplantation. *Sci Rep* **9**, (2019).
579 20. Dashti, A. A., Jadaon, M. M., Abdulsamad, A. M. & Dashti, H. M. Heat Treatment
580 of Bacteria: A Simple Method of DNA Extraction for Molecular Techniques.
581 *Kuwait Medical Journal* **41**, 117–122 (2009).
582 21. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for
583 Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
584 22. Andrews, S. FastQC: A Quality Control tool for High Throughput Sequence
585 Data. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/ (2010).
586 23. Hunt, M. *et al.* ARIBA: rapid antimicrobial resistance genotyping directly from
587 sequencing reads. *Microb Genom* **3**, (2017).
588 24. Alcock, B. P. *et al.* CARD 2023: expanded curation, support for machine
589 learning, and resistome prediction at the Comprehensive Antibiotic Resistance
590 Database. *Nucleic Acids Res* **51**, D690–D699 (2023).
591 25. Wirth, T. *et al.* Sex and virulence in Escherichia coli: an evolutionary perspective.
592 *Mol Microbiol* **60**, 1136–1151 (2006).

593 26. Croucher, N. J. *et al.* Rapid phylogenetic analysis of large samples of
594 recombinant bacterial whole genome sequences using Gubbins. *Nucleic Acids*
595 *Res* **43**, e15–e15 (2015).
596 27. R Core Team. _R: A Language and Environment for Statistical Computing_. *R*
597 *Foundation for Statistical Computing, Vienna, Austria. <https://www.R-*
598 *project.org/>* (2023).
599 28. DNA Purification | DNA Extraction Methods | Promega.
600 https://www.promega.co.uk/resources/guides/nucleic-acid-analysis/dna-
601 purification/.
602 29. Tenaillon, O., Skurnik, D., Picard, B. & Denamur, E. The population genetics of
603 commensal Escherichia coli. *Nat Rev Microbiol* **8**, 207–217 (2010).
604 30. Huttenhower, C. *et al.* Structure, function and diversity of the healthy human
605 microbiome. *Nature 2012 486:7402* **486**, 207–214 (2012).
606 31. Stoesser, N. *et al.* Extensive Within-Host Diversity in Fecally Carried Extended-
607 Spectrum-Beta-Lactamase-Producing Escherichia coli Isolates: Implications for
608 Transmission Analyses. *J Clin Microbiol* **53**, 2122–2131 (2015).
609 32. Pallen, M. J. *et al.* Genomic diversity of Escherichia coli from healthy children in
610 rural Gambia. doi:10.7717/peerj.10572.
611 33. Aguilar-Bultet, L. *et al.* Within-host genetic diversity of extended-spectrum beta-
612 lactamase-producing Enterobacterales in long-term colonized patients. *Nature*
613 *Communications 2023 14:1* **14**, 1–14 (2023).
614 34. Adler, A. *et al.* Transmission dynamics of ESBL-producing Escherichia coli
615 clones in rehabilitation wards at a tertiary care centre. *Clinical Microbiology and*
616 *Infection* **18**, E497–E505 (2012).

617 35. Duval, A. *et al.* Close proximity interactions support transmission of ESBL-K.
618 pneumoniae but not ESBL-E. coli in healthcare settings. *PLoS Comput Biol* **15**,
619 e1006496 (2019).
620 36. Peto, L. *et al.* Selective culture enrichment and sequencing of feces to enhance
621 detection of antimicrobial resistance genes in third-generation cephalosporin
622 resistant Enterobacteriaceae. *PLoS One* **14**, (2019).
623 37. Huttenhower, C. *et al.* Structure, function and diversity of the healthy human
624 microbiome. *Nature* **486**, 207–214 (2012).
625 38. Rondinaud, E., Ruppé, E., Matheron, S., Lucet, J. C. & Armand-Lefevre, L.
626 Screening methods for intestinal carriage of multidrug-resistant
627 Enterobacterales: interest of enrichment broth. *Diagn Microbiol Infect Dis* **97**,
628 115079 (2020).
629 39. Jazmati, T., Hamprecht, A. & Jazmati, N. Comparison of stool samples and
630 rectal swabs with and without pre-enrichment for the detection of third-
631 generation cephalosporin-resistant Enterobacterales (3GCREB).
632 doi:10.1007/s10096-021-04250-1/Published.
633 40. Conrad, C. C., Stanford, K., McAllister, T. A., Thomas, J. & Reuter, T.
634 Competition during enrichment of pathogenic Escherichia coli may result in
635 culture bias. *https://doi.org/10.1139/facets-2016-0007* **1**, 114–126 (2016).
636 41. Jacob, M. E., Keelara, S., Aidara-Kane, A., Matheu Alvarez, J. R. & Fedorka-
637 Cray, P. J. Optimizing a Screening Protocol for Potential Extended-Spectrum β-
638 Lactamase Escherichia coli on MacConkey Agar for Use in a Global
639 Surveillance Program. *J Clin Microbiol* **58**, (2020).
640 42. CHROMagar™ ESBL - Chromagar.
641 https://www.chromagar.com/en/product/chromagar-esbl/.

642 43. El-Jade, M. R. *et al.* ESBL Detection: Comparison of a Commercially Available
643 Chromogenic Test for Third Generation Cephalosporine Resistance and
644 Automated Susceptibility Testing in Enterobactericeae. *PLoS One* **11**,
645 e0160203 (2016).
646 44. Réglier-Poupet, H. *et al.* Performance of chromID ESBL, a chromogenic medium
647 for detection of Enterobacteriaceae producing extended-spectrum β-
648 lactamases. *J Med Microbiol* **57**, 310–315 (2008).
649 45. Sartory, D. P. & Howard, L. A medium detecting β-glucuronidase for the
650 simultaneous membrane filtration enumeration of Escherichia coli and coliforms
651 from drinking water. *Lett Appl Microbiol* **15**, 273–276 (1992).
652 46. Shah DH, S. S. B. T. C. D. *Molecular Detection of Food Borne Pathogens*.
653 (Taylor & Francis Group, CRC Press, USA, 2010).
654 47. Barbosa, C., Nogueira, S., Gadanho, M. & Chaves, S. DNA extraction: finding
655 the most suitable method. *Molecular Microbial Diagnostic Methods: Pathways*
656 *to Implementation for the Food and Water Industries* 135–154 (2016)
657 doi:10.1016/B978-0-12-416999-9.00007-1.
658