Design of the EPIGENEC study: assessing the EPIdemiology and GENetics of *Escherichia coli* in the Netherlands

Denise van Hout[1*], Tess D. Verschuuren[1], Patricia C.J. Bruijning-Verhagen[1,2], Thijs Bosch[2], E. Ascelijn Reuland[2,3,4], Ad C. Fluit[4], Anita C. Schürch[4], Rob J.L. Willems[4], Sabine C. de Greeff[2], Annemarie van 't Veen[3,4], Jan A.J.W. Kluytmans[1,5], Marc J.M. Bonten[1,2,4]

1. Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, The Netherlands
2. Center for Infectious Disease Control, National Institute for Public Health and the Environment, Bilthoven, The Netherlands
3. Saltro Diagnostic Center for Primary Care, Utrecht, The Netherlands
4. Department of Medical Microbiology, University Medical Center Utrecht, Utrecht, The Netherlands
5. Department of Infection Control, Amphia Hospital, Breda, The Netherlands

* Corresponding author: Denise van Hout, D.vanHout-3@umcutrecht.nl

ABSTRACT: Background: Infections caused by *E. coli* cause considerable disease burden and range from frequently occurring and relatively innocent urinary tract infection (UTI) to severe bloodstream infection (BSI). The incidence of infections caused by ESBL-producing *E. coli* (ESBL-PEc) is increasing, justifying surveillance and development of preventive strategies in several domains. Faecal carriage is universal and believed to be the most important reservoir for *E. coli* from which infections can originate. It is currently unknown to what extent Dutch *E. coli* carriage strains in the community reflect isolates causing disease. In this study, we will perform comparative genomics to infer the population structures of human-derived ESBL-PEc from community- and hospital-acquired infections and from community-based faecal carriage samples in the Netherlands. Furthermore, we will describe the molecular epidemiology of *E. coli* isolates causing invasive disease (BSI). **Methods**: This study uses four different microbiological data sources: 1) ESBL-PEc from patients with community-acquired UTI tested in primary care between May and November 2017, 2) ESBL-PEc from urine cultures obtained from patients hospitalized between January 2014 and December 2016, 3) *E. coli* from blood cultures obtained from patients hospitalized between January 2014 and December 2016, and 4) ESBL-PEc from faecal samples collected in a national population- prevalence study performed between January 2014 and January 2017. Clinical epidemiological data was collected from all patients and all isolates were subjected to whole genome sequencing. **Discussion**: The EPIGENEC study (EPIdemiology and GENetics of *E. coli*) will describe the molecular epidemiology of *E. coli* BSI and assess the genomic population structure of ESBL-PEc strains from community-acquired and nosocomial infections, and of ESBL-PEc reflecting community-based faecal carriage. Information from these studies may assist in optimizing surveillance strategies and determining targets and potential impact of future new preventive measures.

Keywords: *Escherichia coli*, antimicrobial resistance, infection, molecular epidemiology

© 2019 by the author(s). Distributed under a Creative Commons CC BY license.

38 **BACKGROUND**

39 *Escherichia coli* (*E. coli*) is commonly found as a gut commensal in humans. Besides its
40 commensal lifestyle *E. coli* is also an important pathogen in humans, as it can establish disease
41 in tissues other than the gastrointestinal tract. These so-called extra-intestinal pathogenic *E. coli*
42 (ExPEC) can cause a wide spectrum of diseases, from uncomplicated cystitis to bloodstream
43 infections (BSI) with 30-day mortality up to 18% [1–3]. *E. coli* is a very heterogeneous species,
44 only 20% of the genes in a typical *E. coli* genome is usually shared among all strains [4]. *E. coli* is
45 known to easily acquire antimicrobial resistance. Molecular characterization studies have shown
46 that *E. coli* strains predominantly become resistant through the exchange of mobile genetic
47 elements carrying resistance genes, such as those encoding for extended-spectrum beta-
48 lactamases (ESBL) [5]. ESBL-producing *E. coli* (ESBL-PEc) are often co-resistant to other
49 classes of antibiotics [6]. Infections caused by antibiotic-resistant *E. coli* strains occur with
50 increasing frequency, which potentially increase the total overall *E. coli* disease burden [3,7,8].
51 Furthermore, in a recent modelling study, ESBL-PEc was found to be responsible for
52 approximately a third of the estimated 33.000 antibiotic-resistance related deaths in Europe in
53 2015 [9]. The increasing availability of whole genome sequencing (WGS) has allowed a more
54 detailed insight into the genetics of *E. coli* virulence and resistance and provided further insight
55 into the distribution of acquired virulence and resistance genes in pathogenic and commensal *E.*
56 *coli* strains of different genetic backgrounds [10–12].

57 Intestinal carriage is believed to be the most important human reservoir for ESBL-PEc from
58 which infections can originate [13]. The estimated prevalence of ESBL-PEc faecal carriage in
59 Dutch community-dwelling inhabitants ranges from 5.2% in the general population [14] to 10.1% in
60 urbanized areas [15], and from 5.0% [16] to 6.1% [17] in hospitalized patients. Surveillance of the
61 molecular epidemiology of antibiotic resistance in the community reservoir is important to
62 identify trends in resistance development. Yet, such surveillance is labour-intensive and costly,

63 and, therefore, not regularly performed. It is currently unknown to what extent the molecular
64 epidemiology of these ESBL-PEc strains present in the Dutch community relates to the
65 molecular epidemiology of ESBL-PEc strains causing community-acquired and nosocomial *E.*
66 *coli* infections. In case of good correlations, urine *E. coli* isolates from primary care patients or
67 from hospitalized patients could be used for surveillance of the molecular epidemiology of
68 antibiotic-resistant *E. coli* in the community in the Netherlands.

69 Information on to what extent *E. coli* strains from different niches and patient populations in the
70 Netherlands differ genomically, is scarce. Possibly, there is also a difference in pathogenic
71 potential within invasive *E. coli* isolates, reflected for example by molecular differences at the
72 genome level in strains that have caused community-acquired BSI as compared to strains that
73 cause BSI in a population that is already vulnerable to infection. Such information is critical for
74 informing strategies around surveillance, prevention and treatment of this important pathogen.
75 In particular for *E. coli* BSI, which is characterized by high morbidity and mortality, more insight
76 in the clinical as well as molecular epidemiology in the Netherlands is needed to help identify
77 targets and potential impact of future preventive strategies such as *E. coli* vaccines, of which
78 one is currently being developed [18].

79 Here, the rationale and study design of the EPIGENEC Study (EPIdemiology and GENetics of
80 *E. coli*) is described, which aims to 1) Investigate the association between the molecular
81 epidemiology of faecal ESBL-PEc isolates obtained from carriage in the community and from
82 clinical cultures (UTI and BSI) in the Netherlands, and 2) Characterize *E. coli* causing BSI in the
83 Netherlands by evaluating clinical characteristics and mapping the genomic population structure
84 of *E. coli* causing BSI based on WGS.

85 # METHODS

86 ## Primary objectives

87 1. To compare the genomic population structure of ESBL-PEc isolates from community-
88 acquired and nosocomial infections with ESBL-PEc isolates found in community faecal samples
89 based on WGS.

90 2. To describe the clinical and molecular epidemiology of *E. coli* isolates causing BSI, and in
91 particular, to assess the association between different clinical and molecular characteristics.

92 ## Study design and population

93 This observational study consists of a prospective as well as a retrospective part. Four sources
94 of data and samples will be obtained from clinical care and the community (see Figure 1).

95 *1. Community-acquired UTI*

96 Patients with a community-acquired UTI caused by ESBL-PEc were identified prospectively by a
97 positive urine culture result at Saltro, a medical laboratory providing service to primary care
98 practices, primarily in the Utrecht (city) region. Urine samples were either inoculated in
99 enrichment broth (Isobouillon with tobramycin, vancomycin and nystatin) if specifically
100 requested on ESBL or identified by elevated MIC for cephalosporins. Screening for ESBL-
101 producing Enterobacteriaceae was performed by inoculation onto a selective screening agar,
102 the Brilliance ESBL screening agar (Oxoid, Basingstoke, United Kingdom). All broths and plated
103 were incubated overnight at 36°C. Species identification and antibiotic susceptibility testing of
104 colonies growing on the Brilliance ESBL plates were performed with respectively the
105 MALDITOF-MS (Bruker, Bremen, Germany) and the Vitek 2 system (Vitek AST, bioMérieux,
106 Marcy-l'Étoile, France). The MIC breakpoints used for interpreting the results were according to
107 the criteria of the EUCAST. Phenotypic confirmation of ESBL was performed by combination
108 disk diffusion test, as recommended by the Dutch national guideline for laboratory ESBL

109 detection [19]. All ESBL-PEc isolates from positive urine cultures between May 2017 and
110 November 2017 were stored at Saltro, at -80°C.

111 *2. Nosocomial UTI*

112 Patients with nosocomial UTI caused by ESBL-PEc were retrospectively identified from medical
113 microbiological records in two participating hospitals: 1) University Medical Center Utrecht
114 (UMCU), and 2) Amphia Hospital in Breda. The UMCU is a 1,042-bed tertiary hospital, providing
115 care to the Utrecht (province) region and serves as a regional referral center. The Amphia
116 Hospital is an 837-bed teaching hospital that provides service to a region of approximately
117 ~400,000 residents. Sample inoculation and confirmation of phenotypic ESBL production was
118 performed as described for the community UTI samples, except that CHROMagar ESBL plates
119 were used (CHROMagar, Paris, France). In both hospitals, every first ESBL-PEc isolate per
120 patient is routinely stored and frozen at -80°C by the medical microbiology department. For this
121 study, we selected all ESBL-PEc isolates from nosocomial UTIs (sample taken >2 days after
122 hospital admission) during the years 2014, 2015 and 2016.

123 *3. BSI*

124 In the same two hospitals, patients with *E. coli* BSI, both ESBL-producing and non-ESBL-
125 producing, were retrospectively identified from medical microbiological records by growth of *E.*
126 *coli* in blood cultures. In these hospitals, *E. coli* isolates from blood cultures are routinely stored
127 at -80°C. For the years 2014, 2015 and 2016, a random sample of 40 isolates per year,
128 comprising ~25% of all bacteraemic *E. coli* isolates in a year, was drawn from each hospital. In
129 addition to the random sample, all ESBL-PEc isolates from 2014-2016 were selected for WGS.
130 Consequently, this set of ESBL-PEc, together with the random sample of the bacteraemic *E.*
131 *coli* strains, comprises the total blood isolate collection for the current study (see Figure 1).

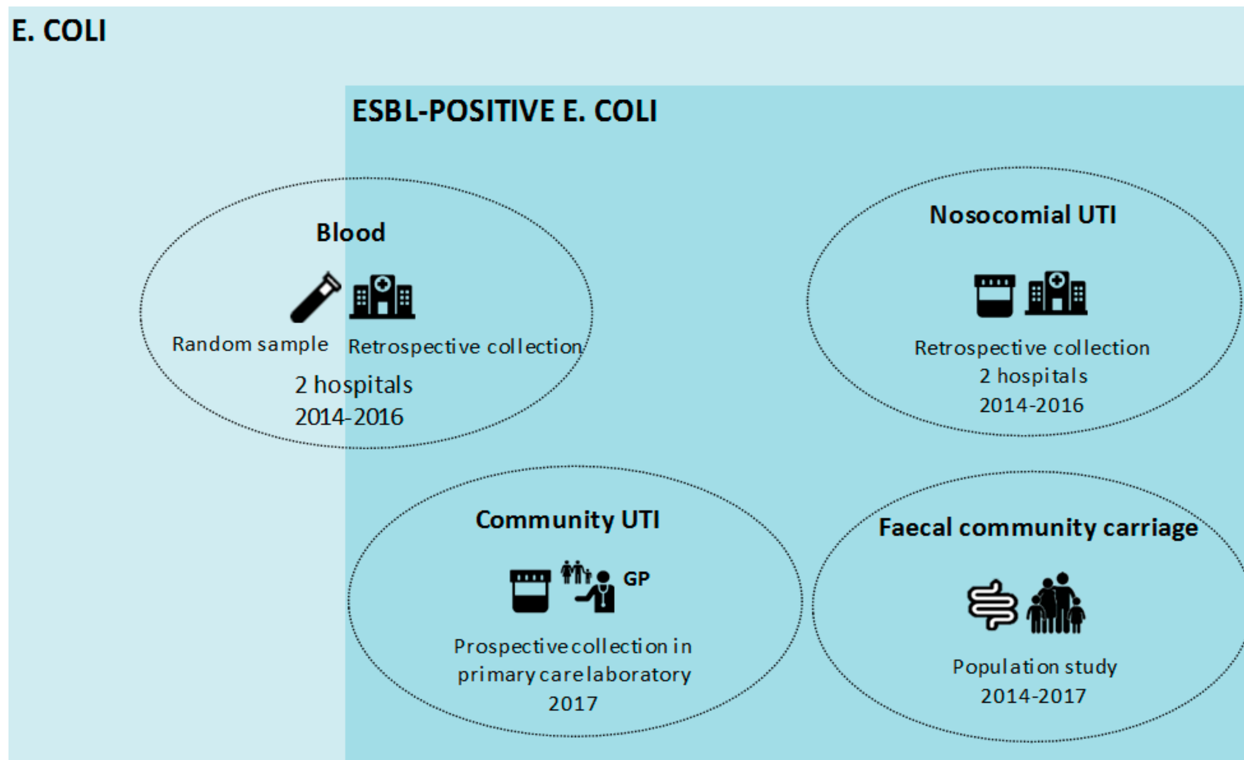
132 All ESBL-PEc isolates obtained from the clinical samples (community UTI, nosocomial UTI and
133 BSI) and the random sample of *E. coli* BSI isolates were included for further molecular analyses
134 at The Netherlands National Institute for Public Health and the Environment (RIVM). Because
135 we expected follow-up cultures to often grow the same *E. coli* isolate as the first culture, and for
136 efficiency reasons, we selected only the first available *E. coli* isolate for each patient (all ages),
137 irrespective of time between cultures.

138 *4. Community-based intestinal carriage*

139 The fourth dataset consists of ESBL-PEc isolates collected from faecal samples of a national
140 population study for ESBL-producing Enterobacteriaceae, performed between November 2014
141 and November 2016. In this cross-sectional study, every month a random sample of ~2,000
142 residents of the Netherlands was drawn from Dutch municipalities (covering the entire
143 population of the Netherlands). One person per household was invited to fill in a web-based
144 questionnaire, and upon completion of the questionnaire, the participant was asked to provide a
145 faecal sample. ESBL-producing *Escherichia coli*, *Klebsiella pneumoniae*, and the *Enterobacter*
146 *cloacae* complex were isolated using MacConkey agar with 1 mg/L cefotaxime or after
147 enrichment 2 mL of LB with 1 mg/mL cefotaxime. Up to five colonies with different morphologies
148 were selected. Species identification was performed using MALDITOF-MS (Bruker, Bremen,
149 Germany). ESBL-encoding genes were identified by PCR and isolates negative in the PCR
150 were tested for the presence of other ESBL encoding genes by the Check MDR CT-101
151 microarray (Check-points, Wageningen, the Netherlands). The genes were identified by
152 conventional sequencing. PCR-based Replicon Typing (PBRT) was performed to identify the
153 plasmid type that encoded the ESBL [20]. All ESBL-producing Enterobacteriaceae were stored at
154 -80°C in the UMCU and were subjected to WGS (see Genotyping). Further details of the study
155 design can be found elsewhere [21]. For the current study, only genetic data of the first sampled
156 faecal ESBL-PEc isolate of a patient were collected. No age restrictions were used.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 February 2019 doi:10.20944/preprints201902.0066.v1

157 **Figure 1. Data sources**



158

159 Legend Figure 1: this figure depicts the sources and time-span of the different *E. coli* collections
160 that are used in this study.
161 GP, general practitioner; UTI, urinary tract infection

162 **Epidemiological variables**

163 The following information was collected from all patients: age, sex, postal code, type of infection
164 (community UTI, nosocomial UTI, BSI), date of sample collection, and community or nosocomial
165 (i.e. sample taken >2 days after hospital admission) onset of infection. In addition, for UTIs it
166 was recorded whether the urine sample was a catheter sample. For patients with *E. coli* BSI,
167 additional information regarding presence of a urinary catheter, hospital ward (ICU versus non-
168 ICU), 30-day and 1-year mortality and the primary focus of BSI was obtained from electronic
169 medical records. Possible primary foci were: urinary tract (i.e. pyelonephritis, prostatitis), gastro-
170 intestinal (i.e. diverticulitis, bacterial translocation), hepatic-biliary (i.e. cholangitis), respiratory,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 February 2019 doi:10.20944/preprints201902.0066.v1

171 gynaecological, other (i.e. meningitis, venous catheter), and unknown. The primary focus of BSI
172 (portal of entry) was defined on the basis of clinical and/or radiologic features and the isolation
173 of *E. coli* from the presumed source of infection. If *E. coli* was not isolated from the presumed
174 primary focus (i.e. because of previous antimicrobial treatment or invasive procedure that was
175 needed to isolate *E. coli* from primary source), the presumed primary focus was based on a firm
176 clinical suspicion (given that all other possible sources of infection were excluded). In case of
177 multiple possible primary foci, consensus was reached by discussion by DH and TV.

178 **Genotyping**

179 All *E. coli* isolates that were selected for the current study were inoculated on (non-selective)
180 blood agar and species confirmation was performed by MALDITOF-MS prior to WGS, which
181 was performed at the RIVM. All *E. coli* strains, except the strains from the external dataset, were
182 subjected to WGS using the Illumina HiSeq 2500 (BaseClear, Leiden, the Netherlands). For
183 this, cell pellets were made from 1500µl of overnight culture in Brain Heart Infusion (BHI) broth
184 (Tritium Microbiologie BV, Eindhoven, the Netherlands). The pellets were washed in 500µl
185 saline and subsequently re-suspended in 200µl DNA/RNA Shield (Zymo Research, Irvine CA,
186 U.S.A.). From the pellets, nucleic acid extraction was performed by BaseClear (Leiden, the
187 Netherlands) using an in-house protocol. Library preparation and sequencing of the bacterial
188 genomes was performed using the Illumina Nextera XT kit and the HiSeq 2500 with a paired-
189 end 100 cycles protocol. The ESBL-PEc from the fourth dataset, i.e. ESBL-PEc recovered from
190 faeces of healthy humans, were sequenced using the MiSeq or NextSeq platforms (Illumina,
191 San Diego, CA). Contigs of all four datasets were assembled with SPAdes genome assembler
192 v.3.6.2. The resulting WGS data were used to determine the multi-locus sequence type (MLST),
193 virulence gene and resistance gene content using the mlst2.0
194 (https://github.com/tseemann/mlst) to scan contig files against the *E. coli* PubMLST typing
195 scheme (updated May 12th, 2018), and Abricate (https://github.com/tseemann/abricate) for

196 mass screening of contigs for antimicrobial resistance and virulence genes. Abricate comes
197 bundled with multiple resistance gene and virulence gene databases. For this study, the
198 ResFinder and VFDB databases were used. Serotypes were assigned by using the web-tool
199 SerotypeFinder 2.0 from the Center for Genomic Epidemiology at the Danish Technical
200 University, Lyngby, Denmark (http://www.genomicepidemiology.org). This tool uses presence of
201 O- and H-antigen-processing genes to predict *E. coli* serotypes [22].

202 **Planned analyses**

203 *Primary objective 1*

204 The population structure of ESBL-PEc from the clinical and faecal samples will be compared on
205 three levels. Firstly, the core genome will be assessed with MLST, a core genome phylogeny
206 based on SNP and allelic profile variation using SeqSphere, and the ESBL-PEc populations will
207 be partitioned in sequence clusters. For this, different methods are available like hierarchical
208 Bayesian Analysis of Population Structure (BAPS) or PopPUNK [23,24]. Secondly, the accessory
209 genome will be assessed by comparing acquired resistance genes in the ESBL-PEc populations
210 using Resfinder, and the plasmid composition will be predicted using the recently developed
211 mlplasmid© algorithm [25]. Lastly, a pan-genome analysis will be performed using PANINI, to
212 assess if the total gene content differs per different ESBL-PEc population [26].

213 *Primary objective 2*

214 To assess the association between epidemiological characteristics and molecular
215 characteristics of *E. coli* blood isolates, MLST, virulence and antimicrobial resistance gene
216 content will be described according to the different epidemiological subgroups. A core-genome
217 tree will be constructed with the same method as mentioned above. A virulence score will be
218 made per isolate and will be defined as the total number of virulence genes present in that
219 strain. These virulence scores will then be compared between isolates with different
220 epidemiological characteristics and between ST131 and non-ST131 isolates, respectively.

doi:10.20944/preprints201902.0066.v1

221 Serotype distribution of the bacteraemia population will be compared to current *E. coli* vaccine
222 candidates. Furthermore, a genome-wide association approach will be used to see whether any
223 epidemiological characteristics are associated with certain molecular traits.

224 Ethics

225 This study is conducted according to the principles of the Declaration of Helsinki (World Medical
226 Association, 2013) and does not fall under the scope of the Medical Research Involving Human
227 Subjects Act, the Medical Research Ethics Committee of the UMCU has therefore waived the
228 need for official approval by the Ethics Committee (IRB number 18/056). The study uses
229 pseudonymised data and informed consent is not obtained from study participants. Patients that
230 participated in the open population study (ESBLAT study, IRB number 14/219-C) have provided
231 informed consent for the use of clinical data and faecal samples in future studies such as the
232 current study. In this study, in case of age <13 years, parents provided informed consent. In
233 case of age 13-17 years, both the child and parents provided informed consent.

234 DISCUSSION

235 The EPIGENEC study aims to assess the genomic population structure of ESBL-PEc strains
236 from community and nosocomial infections and ESBL-PEc strains representing community
237 faecal carriage. It will also carefully describe the clinical epidemiology and genomic population
238 structure of *E. coli* BSI, which is important in determining the targets and impact of possible new
239 preventive measures.

240 *Strengths*

241 One of the key aspects of the current study is the combined use of epidemiological data and
242 detailed whole genome sequence data of strains from several different domains in order to
243 obtain a more complete picture of the current molecular epidemiology of (ESBL-producing) *E.*
244 *coli* in the Netherlands. Furthermore, the use of WGS techniques allows us to map the

245 population structure of *E. coli* and the association of the genomic make-up of strains with their
246 isolation source with high resolution and discriminatory power. Also, all strains were uniformly
247 assembled and analysed, reducing the risk of information bias.

248 *Limitations*

249 This study also has limitations. Guidelines for Dutch primary care physicians recommend to only
250 send in urine cultures for microbiological testing for patients with complicated UTI (i.e.
251 symptoms accompanied with fever, or in case of male patients with UTI symptoms), clinical
252 treatment failure, recurrent UTIs, or a possibly resistant infection, which implies selection of
253 patients with community UTI. However for our study, we do not consider this to cause selection
254 *bias*, since we are particularly interested in the molecular epidemiology of ESBL-PEc from urine
255 samples in the way they are currently being performed, so as according to clinical practice.
256 Also, ideally we would be able to pick up time-trends in the change in molecular epidemiology of
257 community faecal carriage of ESBL-PEc and assess whether these trends are reflected in the
258 molecular epidemiology of clinical cultures, for example from community or nosocomial UTI.
259 One could imagine using such results to assess the possible value of ESBL-PEc isolates from
260 clinical cultures as a proxy of changes in the molecular epidemiology of community faecal
261 carriage. However, considering the heterogeneity in the *E. coli* species and the limited amount
262 of years of which we have faecal samples, this will prohibit us to draw hard conclusions. We still
263 believe this comparison will provide us valuable information and will guide future research on
264 the possible use of routine clinical samples in the assessment of the molecular epidemiology of
265 ESBL-PEc.

266

267

268 # DECLARATIONS

269 ## Acknowledgements

270 We sincerely thank Kim van der Zwaluw, Judith Vlooswijk, Carlo Verhulst and Jelle Scharringa
271 for their contribution in the preparation of this study.

272 ## Conflicts of interest

273 None to declare.

274 ## Study funding

275 This study is internally funded by the Netherlands National Institute for Public Health and the
276 Environment (RIVM) and the UMCU and has not received external funding or assistance from
277 any commercial organization.

278

279 **REFERENCES**

280 1. Abernethy, J. K. *et al.* Thirty day all-cause mortality in patients with *Escherichia coli*
281 bacteraemia in England. *Clin. Microbiol. Infect.* **21,** 251.e1-8 (2015).

282 2. Fitzpatrick, J. M. *et al.* Gram-negative bacteraemia; a multi-centre prospective evaluation
283 of empiric antibiotic therapy and outcome in English acute hospitals. *Clin. Microbiol.*
284 *Infect.* **22,** 244–251 (2016).

285 3. Vihta, K. D. *et al.* Trends over time in *Escherichia coli* bloodstream infections, urinary
286 tract infections, and antibiotic susceptibilities in Oxfordshire, UK, 1998-2016: a study of
287 electronic health records. *Lancet. Infect. Dis.* **18**, 1138-1149 (2018).

288 4. Lukjancenko, O., Wassenaar, T. M. & Ussery, D. W. Comparison of 61 Sequenced
289 *Escherichia coli* Genomes. *Microb. Ecol.* **60,** 708–720 (2010).

290 5. Mathers, A. J., Peirano, G. & Pitout, J. D. D. The role of epidemic resistance plasmids
291 and international high-risk clones in the spread of multidrug-resistant Enterobacteriaceae.
292 *Clin. Microbiol. Rev.* **28,** 565–591 (2015).

293 6. European Centre for Disease Prevention and Control. Surveillance of antimicrobial
294 resistance in Europe 2016. Annual report of the European Antimicrobial Resistance
295 Surveillance Network (EARS-Net). (2016).

296 7. Schlackow, I. *et al.* Increasing incidence of *Escherichia coli* bacteraemia is driven by an
297 increase in antibiotic-resistant isolates : electronic database study in Oxfordshire. *J*
298 *Antimicrob. Chemother.* **67,** 1514–1524 (2012).

299 8. de Kraker, M. E. A. *et al.* The changing epidemiology of bacteraemias in Europe: trends
300 from the European Antimicrobial Resistance Surveillance System. *Clin. Microbiol. Infect.*
301 **19,** 860–868 (2013).

302 9. Cassini, A. *et al.* Attributable deaths and disability-adjusted life-years caused by
303 infections with antibiotic-resistant bacteria in the EU and the European Economic Area in
304 2015: a population-level modelling analysis. *Lancet Infect. Dis.* **19,** 56-66 (2019).

305 10. Köhler, C. D. & Dobrindt, U. What defines extraintestinal pathogenic *Escherichia coli*? *Int.*
306 *J. Med. Microbiol.* **301,** 642–647 (2011).

307 11. Johnson, J. R. *et al*. Epidemiological Correlates of Virulence Genotype and Phylogenetic
308 Background among *Escherichia coli* Blood Isolates from Adults with Diverse-Source
309 Bacteremia. *J. Infect. Dis.* **185**, 1439–1447 (2002).

310 12. Johnson, J. R. *et al.* Accessory Traits and Phylogenetic Background Predict *Escherichia*
311 *coli* Extraintestinal Virulence Better Than Does Ecological Source. *J. Infect. Dis.* **219,**
312 121-132 (2019).

313 13. Yamamoto, S. *et al.* Genetic evidence supporting the fecal-perineal-urethral hypothesis in
314 cystitis caused by *Escherichia coli*. *J. Urol.* **157,** 1127–1129 (1997).

315 14. van Hoek, A. H. *et al.* Molecular characteristics of extended-spectrum cephalosporin-
316 resistant Enterobacteriaceae from humans in the community. *PLoS One* **10,** e0129085
317 (2015).

318 15. Reuland, E. A. *et al.* High prevalence of ESBL-producing Enterobacteriaceae carriage in
319 Dutch community patients with gastrointestinal complaints. *Clin. Microbiol. Infect.* **19,**
320 542–549 (2013).

321 16. Willemsen, I. *et al.* Trends in Extended Spectrum Beta-Lactamase (ESBL) Producing
322 Enterobacteriaceae and ESBL Genes in a Dutch Teaching Hospital, Measured in 5
323 Yearly Point Prevalence Surveys (2010-2014). *PLoS One* **10,** e0141765 (2015).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 February 2019 doi:10.20944/preprints201902.0066.v1

324 17. Zhou, X. *et al.* Epidemiology of Extended-Spectrum beta-Lactamase-Producing *E. coli*
325 and Vancomycin-Resistant Enterococci in the Northern Dutch-German Cross-Border
326 Region. *Front. Microbiol.* **8,** 1914 (2017).

327 18. Huttner, A. *et al.* Safety, immunogenicity, and preliminary clinical efficacy of a vaccine
328 against extraintestinal pathogenic *Escherichia coli* in women with a history of recurrent
329 urinary tract infection: a randomised, single-blind, placebo-controlled phase 1b trial.
330 *Lancet. Infect. Dis.* **17,** 528–537 (2017).

331 19. Netherlands Society for Medical Microbiology. NVMM Guideline Laboratory detection of
332 highly resistant microorganisms. 24–46 (2012).

333 20. Carattoli, A. *et al.* Identification of plasmids by PCR-based replicon typing. *J. Microbiol.*
334 *Methods* **63,** 219–228 (2005).

335 21. van den Bunt, G. *et al.* Intestinal carriage of ampicillin- and vancomycin-resistant
336 Enterococcus faecium in humans, dogs and cats in the Netherlands. *J. Antimicrob.*
337 *Chemother.* **73**, 607–614 (2017).

338 22. Joensen, K. G. *et al*. Rapid and Easy In Silico Serotyping of *Escherichia coli* Isolates by
339 Use of Whole-Genome Sequencing Data. *J. Clin. Microbiol.* **53,** 2410–2426 (2015).

340 23. Cheng, L. *et al*. Hierarchical and spatially explicit clustering of DNA sequences with
341 BAPS software. *Mol. Biol. Evol.* **30,** 1224–1228 (2013).

342 24. Lees, J. *et al.* Fast and flexible bacterial genomic epidemiology with PopPunk. *BioRxiv*
343 doi: https://doi.org/10.1101/360917 (2018).

344 25. Arredondo-Alonso, S. *et al.* Mlplasmids: a User-Friendly Tool To Predict Plasmid- and
345 Chromosome-Derived Sequences for Single Species. *Microb. Genomics*

346 doi:10.1099/mgen.0.000224 (2018).

347 26. Abudahab, K. *et al.* PANINI: Pangenome Neighbour Identification for Bacterial
348 Populations. *Microb. Genomics* doi:10.1099/mgen.0.000220 (2018).

349