Stepwise evolution and convergent recombination underlie the global dissemination of carbapenemase-producing *Escherichia coli*

[1,2]Rafael PATIÑO-NAVARRETE, [1,2]Isabelle ROSINSKI-CHUPIN, [1,2]Nicolas CABANEL, [1,3,4,5]Lauraine GAUTHIER, [1,5]Julie TAKISSIAN, [6]Jean-Yves MADEC, [7]Monzer HAMZE, [1,4,5]Remy A. BONNIN, [1,3,4,5]Thierry NAAS[$] and [1,2,*]Philippe GLASER[$].

[$] Both authors contributed equally

[1]Unité EERA, Institut Pasteur, APHP, Université Paris Saclay, [2]UMR3525, CNRS, 28 rue du Dr Roux, 75015, Paris, France; [3]EA7361 Faculty of Medicine of University Paris-Sud; [4]Department of Bacteriology-Hygiene, Bicêtre Hospital, APHP, [5]Associated French National Reference Center for Antibiotic Resistance, Le Kremlin-Bicêtre, France; [6]Université de Lyon - Agence Nationale de Sécurité Sanitaire (ANSES), Unité Antibiorésistance et Virulence Bactériennes, Lyon, France. [7]Laboratoire Microbiologie Santé et Environnement (LMSE), Ecole Doctorale des Sciences et de Technologie, Faculté de Santé Publique, Université Libanaise, Tripoli, Lebanon.

*Corresponding author:

Philippe GLASER,

Institut Pasteur, 28 Rue du Dr Roux, 75724 Paris Cedex 15,

Tel + 33 1 45 68 89 96 - E-mail: pglaser@pasteur.fr

Abstract

Carbapenem-resistant *Enterobacteriaceae* are considered by WHO as "critical" priority pathogens for which novel antibiotics are urgently needed. The dissemination of carbapenemase-producing *Escherichia coli* (CP-*Ec*) in the community is a major public health concern. However, the global molecular epidemiology of CP-*Ec* isolates, as well as the genetic bases for the emergence and global dissemination of specific lineages, remain largely unknown. Here, by combining a thorough genomic and evolutionary analysis of *Ec* ST410 isolates with a broad analysis of 12,398 *E. coli and Shigella* genomes, we showed that the fixation of carbapenemase genes depends largely on a combination of mutations in *ftsI* encoding the penicillin binding protein 3 and in the porin genes *ompC* and *ompF*. Mutated *ftsI* genes and a specific *ompC* allele spread across the species by recombination. Those mutations were in most cases selected prior to carbapenemase gene acquisition. The selection of CP-*Ec* lineages able to disseminate is more complex than the mere acquisition of carbapenemase genes and might be largely triggered by β-lactams other than carbapenems.

41 **Introduction**

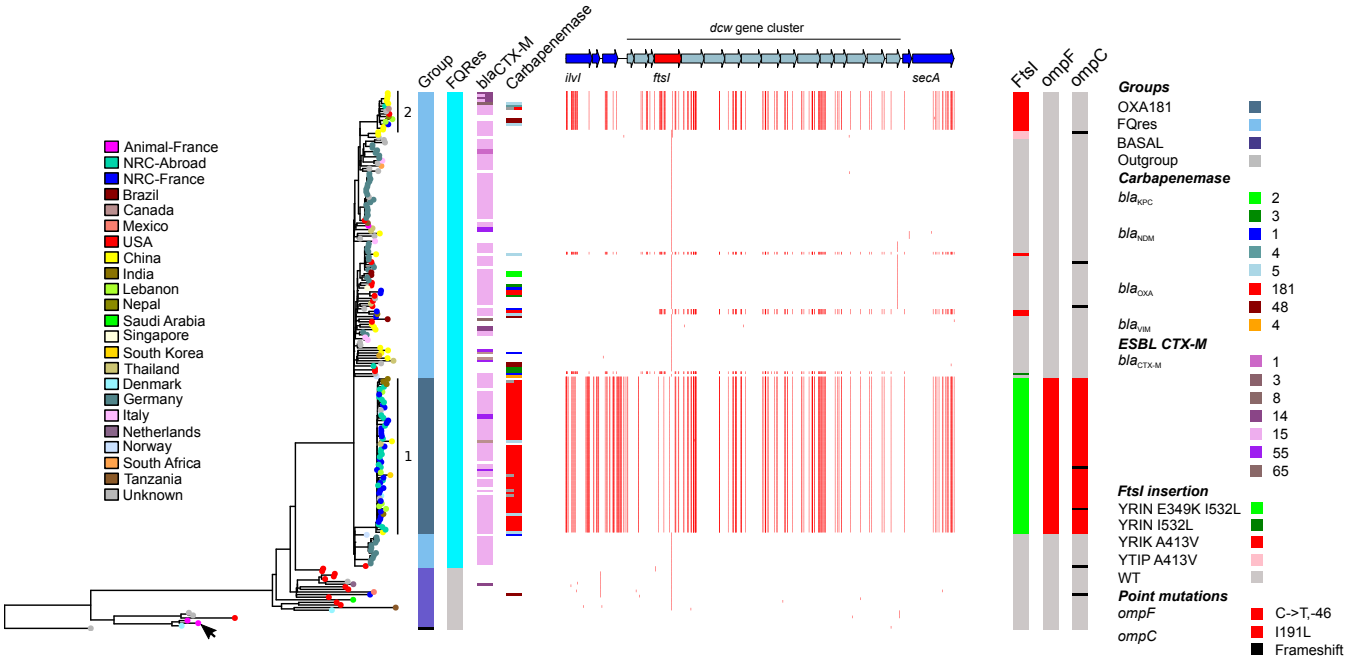
42

43 Antibiotic resistance is one of the most urgent public health concerns. The
44 increasing rates in antimicrobial resistances worldwide suggests a bleak outlook in
45 terms of morbidity, mortality and economic loss[1]. Carbapenems are the last resort
46 antibiotics used to treat infections caused by multidrug-resistant (MDR) Gram-negative
47 bacteria[2]. Dissemination of carbapenem-resistant *Enterobacteriaceae* (CRE) threatens
48 the efficacy of current treatment options. Carbapenem-resistance may result from a
49 combination of mutations leading to reduced permeability (e.g. porin deficiency) and
50 overexpression of an extended-spectrum β-lactamase (ESBL) or a cephalosporinase that
51 show a weak activity against carbapenems[3]. However, the main resistance mechanism is
52 the acquisition of a carbapenemase gene[4]. The major carbapenemases encountered in
53 *Enterobacteriaceae* belong to Ambler class A (KPC-type), class B (metallo-β-lactamases
54 IMP, VIM- and NDM-types) or class D (OXA-48-like enzymes)[5]. As these carbapenemases
55 are now frequently encountered in *Escherichia coli*, public health authorities fear that
56 carbapenemase-producing *E. coli* (CP-*Ec*) might follow the same expansion and
57 dissemination in hospitals and the community as the one observed for CTX-M-type
58 ESBL-producing *E. coli* isolates[6,7]. This is especially worrisome as these isolates are
59 usually resistant to multiple antibiotics.

60 The epidemiology of CP-*Ec* is complex with geographical diversity in terms of
61 carbapenemase genes and of dominant lineages[4]. Most studies performed at national or
62 hospital levels point to a broad diversity of isolates as defined by multilocus sequence
63 typing (MLST), with some isolates belonging to a few dominant sequence types (STs)
64 like ST38, clonal complex 10 (ST10, ST167, ST617), ST101, ST131 and ST410 that
65 express different carbapenemases[4,8-13]. However, their prevalence varies significantly
66 worldwide. Analysis of CP-*Ec* strains isolated in 16 countries between 2008 and 2013
67 revealed that 36% belonged to the pandemic ST131, which has driven the global spread
68 of CTX-M-15 ESBL in *E. coli*[11]. But only a single ST131 isolate out of 140 CP-*Ec* was
69 identified by the French National Reference Centre (Fr-NRC) between 2012-2013[8].
70 Recently, the phylogenetic analysis of a Danish collection of ST410 isolates combined to
71 an international set of isolates revealed a globally disseminated *Ec* clone carrying $bla_{\text{OXA-}}$
72 $_{181}$ on a IncX3 plasmid. This lineage was predicted by a Bayesian analysis to have
73 acquired $bla_{\text{OXA-181}}$ around 2003 and subsequently $bla_{\text{NDM—5}}$ around 2014[13].

74 Despite public health implications, factors contributing to the emergence and the
75 dissemination of CP-*Ec* lineages have not been explored. Here, by using an in-depth
76 evolutionary and functional analysis of *Ec* ST410 and by extending it to the whole *E. coli*
77 species, we show that acquisition of carbapenemase genes occurred preferentially in
78 specific disseminated lineages mutated in *ftsI* encoding penicillin-binding protein 3
79 (PBP3) and/or in the *ompC* and *ompF* porin genes. Our global analysis predicts that
80 selection by other β-lactams aside from carbapenems played a major role in fixing
81 carbapenemase genes in specific genetic backgrounds that have not yet been observed
82 among the high-risk *Ec* ST131 isolates.
83
84
85 **Results**
86
87 **Most CP-*Ec* ST410 isolates received by the French NRC belong to a single**
88 **lineage.** In order to determine the genetic bases for the dissemination of CP-*Ec* lineages,
89 we first analysed ST410 CP-*Ec* isolates, which show a high prevalence among isolates
90 collected by the Fr-NRC[8]. We sequenced the genomes of 55 CP-isolates, 51 collected by
91 the Fr-NRC (including 22 from colonized patients repatriated from 15 different
92 countries) and four from Lebanon, and three non-CP isolates of animal origin
93 (Supplementary Table 1). We reconstructed their phylogeny together with 149 *Ec*
94 ST410 genome sequences retrieved from public databases (Supplementary table 2). The
95 phylogeny is in agreement with the recent analysis of CP-*Ec* ST410 from a Danish
96 collection[13], with a major fluoroquinolone resistant clade (FQR-clade) gathering the
97 majority of the isolates (160 out of 182) and of the isolates carrying carbapenemase
98 genes (75 out of 76). Within the FQR-clade 84% of the isolates expressed CTX-M-type
99 ESBLs (Fig. 1). 55 out of the 59 $bla_{\text{OXA-181}}$ carrying isolates formed a single subclade (the
100 OXA-181 subclade) which corresponds to the previously described clade B4/H24RxC[13].
101 The 27 CP-*Ec* isolates not belonging to the OXA-181 subclade express different
102 carbapenemases from the OXA-48, KPC and NDM families. To accurately analyse the
103 evolution of the OXA-181 subclade, we sequenced to completion a representative isolate
104 of this clade (*Ec*-MAD). *Ec*-MAD carries three plasmids and expresses 16 antibiotic
105 resistance genes (ARG) targeting seven classes of antibiotics (Supplementary Table 3). It
106 is resistant to most antibiotic tested, remaining susceptible only to mecilinam,

107 imipenem, meropenem, amikacin, azithromycin, chloramphenicol, tigecycline and
108 colistin and intermediate to kanamycin, ertapenem and gentamicin (Supplementary
109 Table 4). Comparison of the antibiotic resistance gene (ARG) content among ST410 *Ec*
110 isolates revealed an increase in the median number of ARG between the basal isolates
111 (n=3), the FQR-clade (n=9) and the OXA-181 subclade (n=16) (Supplementary Fig. 1).



112 **Fig. 1. Core genome phylogeny and genomic features of *E. coli* ST410 isolates.**
113 Maximum likelihood phylogeny of 207 *Ec* ST410 genomes, built with RAxML[14] based on
114 the core and recombination-free alignment of 5,362 SNPs. The *Ec* ST88 isolate 789
115 (CP010315.1), also from CC23, was used as an outgroup. Isolates are colour-coded
116 according to the geographical origin (branch tips) as indicated in the figure key (left).
117 Genomic features are indicated on the right side of the figure as indicated in the figure
118 key (right). From left to right: groups according to the phylogeny, including the FQR
119 clade and the OXA-181 subclade; FQ resistance due to mutations in the QRDR of GyrA
120 (S83L/D87N) and ParC (S80I) as turquoise blue; CTX-M type ESBLs; Carbapenemases;
121 SNPs (small vertical red bars) compared to the *Ec* ST410 non-recombined strain
122 ANSES30599 from the basal clade (highlighted by an arrowhead) as reference;
123 mutations in *ftsI*, *ompF,* and *ompC*. For CTX-M and Carbapenemase, two colours indicate
124 the detection of two genes encoding the corresponding β-lactamases. NRC stands for
125 National Reference Centre. Upper part, genetic map of the *dcw* locus, genes are indicated
126 by arrows and the *ftsI* gene, in red. Clusters 1 and 2 as defined in Table 1 are indicated
127 by thin vertical lines next to the tree.

128 **Gain of specific *ftsI* alleles by recombination is a hallmark of Ec ST410**
129 **carbapenemase-producing strains.** As our phylogenetic analysis provided further
130 evidence of a worldwide dissemination of the OXA-181 subclade[13], we searched for
131 polymorphisms that, in addition to the acquisition of ARG, have contributed to the
132 expansion of this lineage. To this end, we systematically analysed mutations occurring in

133 the branch leading to its most recent common ancestor (MRCA). Among the 1 745 SNPs
134 identified in this branch, 92.4 % were located in a 124-kb DNA region between *yaaU* and
135 *erpA* (Supplementary Table 5). This variable region was predicted by using Gubbins[15] as
136 resulting from a recombination event (Supplementary Fig. 2). BLASTN search against
137 the NCBI database revealed that the recombined region in the OXA-181 subclade is
138 almost identical to sequences found in four ST167 and eight ST617 isolates from clonal
139 complex (CC) 10. For instance, only three SNPs over 124 kb were detected between the
140 recombined region in the OXA-181 subclade and the corresponding region in a ST167
141 NDM-7-producing *Ec* strain (SCEC76) isolated in China (Fig. 2). Strikingly, all these
142 isolates except one expressed a carbapenemase gene. Furthermore, among isolates of
143 the ST410 FQR-clade four additional recombination events were identified that
144 overlapped the 124-kb recombining region identified in the OXA-181 subclade by
145 16.5 kb (Supplementary Fig. 2). All affected CP-*Ec* isolates: OXA-181 subclade; a
146 subclade of 15 isolates of different geographical origins including six CP isolates
147 carrying different carbapenemase genes; two closely related CP-*Ec* isolates, one from
148 India ($bla_{\text{NDM-5}}$) and one from the Fr-NRC ($bla_{\text{OXA-181}}$); and two recombination events
149 affecting a single CP-*Ec* isolate each (Fig. 1). The 16.5-kb common region encompassed
150 the *dcw* (division and cell wall) locus from *ftsI* to *secM* (Fig. 2). It encodes major
151 functions in cell wall synthesis and cell division, including *ftsI* encoding PBP3, a target of
152 diverse β-lactams[16]. Overall, 79% (68/85) of the CP-*Ec* ST410 isolates had recombined
153 in the *dcw* region. Given this association between recombination at the *dcw* locus and
154 the presence of a carbapenemase gene, we hypothesized that recombination of this
155 region in the OXA-181 subclade and other CP-*Ec* ST410 isolates was positively selected
156 and confers a selective advantage in the context of MDR strains.

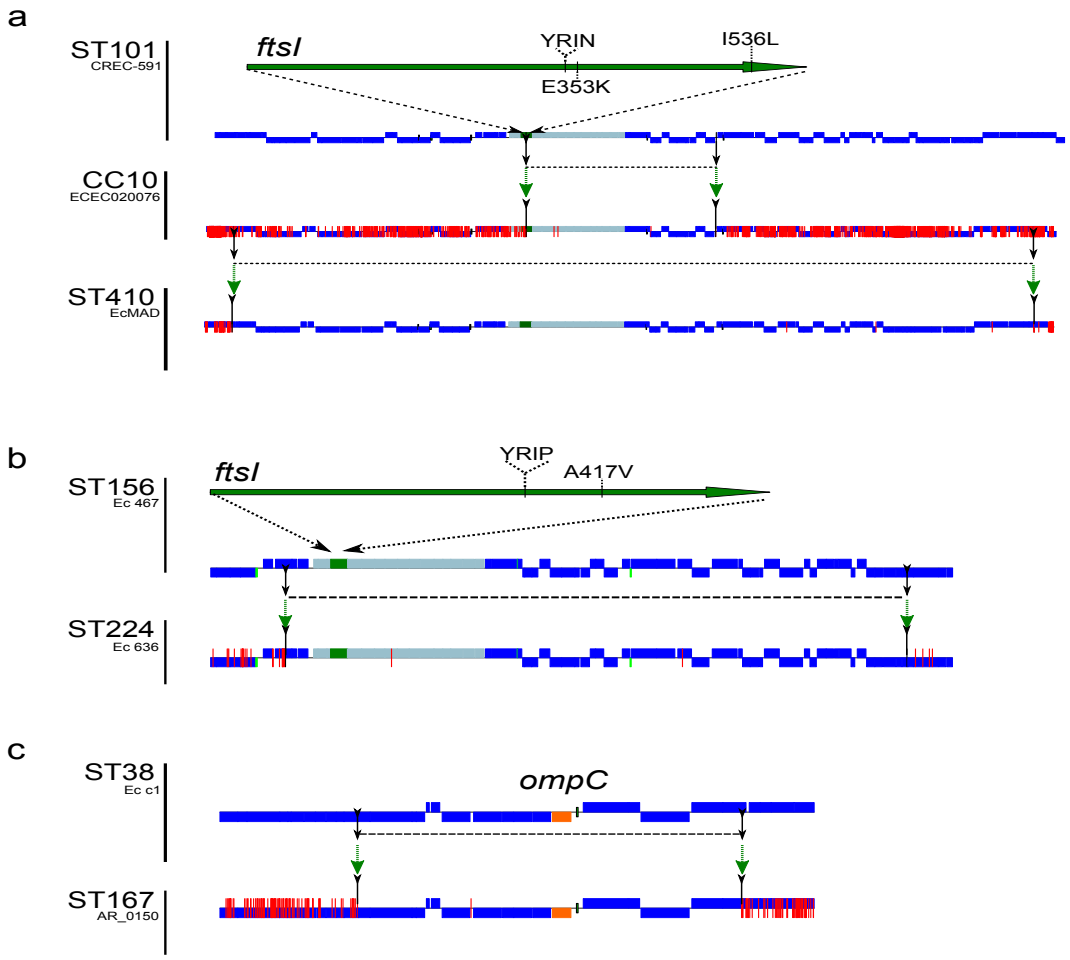
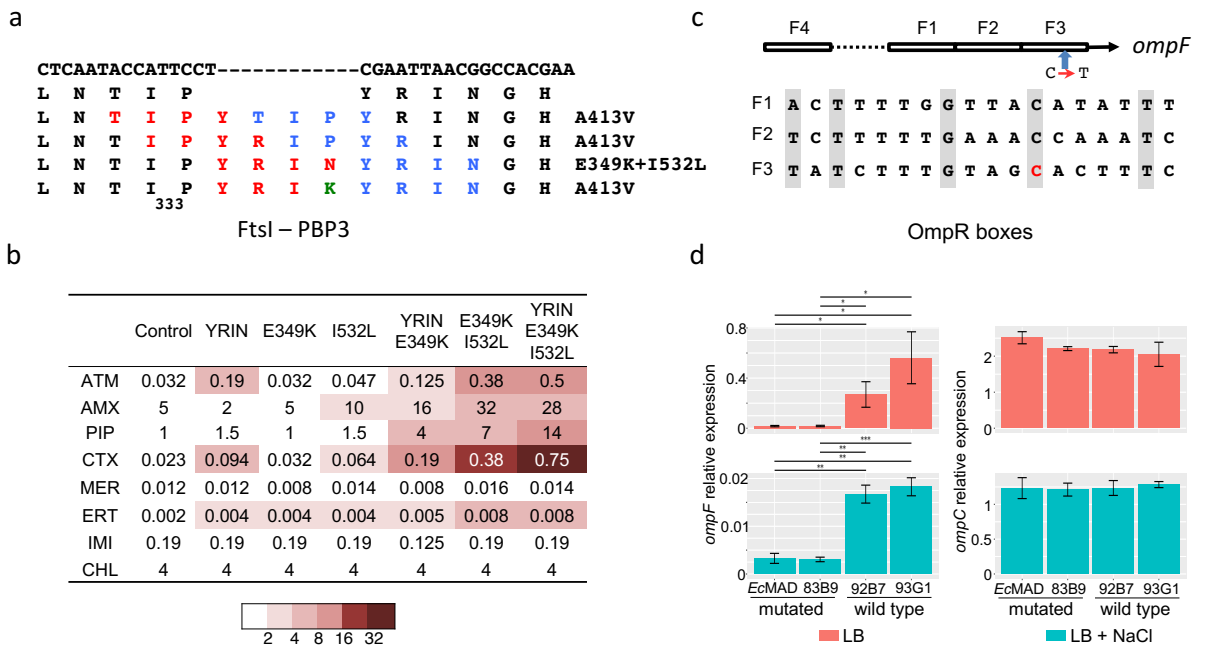


Fig. 2. Recombination events at the *dcw* and *ompC* loci. a. Recombination of *ftsI* carrying an YRIN insertion. The *ftsI* gene (in green) of an *E. coli* ST101 isolate was affected by a 12-nucleotide duplication leading to an YRIN insertion after P_{333} and by two additional mutations (E349K and I532L). A 29.5 kb chromosomal region of an ST101 isolate mutated in *ftsI* comprising most of the *dcw* gene cluster was transferred by LGT and recombination into a CC10 strain, which transferred a 124-kb fragment containing the region acquired from *E. coli* ST101 to the common ancestor of the OXA-181 ST410 subclade. **b.** Recombination of *ftsI* carrying an YRIP insertion. The *ftsI* gene of an *Ec* ST156 isolate was affected by a 12-nucleotide duplication leading to an YRIP insertion after P_{333} and a mutation (A413V). A 65-kb region was transferred by recombination into an *E. coli* ST224 strain. **c.** Recombination of *ompC* gene from an ST38 strain. A 22.7-kb region encompassing the *ompC* gene (in orange) was transferred by recombination in the MRCA from the ST167 clade of isolates expressing different carbapenemase genes. Vertical black arrows indicate the limits of the recombined regions and green arrows the direction of exchange. Small vertical red bars represent SNPs compared to the donor strains. CDSs are indicated by blocks, upper line transcribed in the rightward orientation and lower line in the opposite orientation. The *dcw* gene cluster is in grey.

197 SNPs including 16 non-synonymous (NS) mutations differentiated the common 16.5-kb region in the OXA-181 subclade from other *Ec* ST410 isolates (Supplementary Table 5). Among differences, we identified insertions of four codons (YRIN in two cases and YRIK in three, Fig. 3a) occurring at a same position, between

179 P333 and Y334 of *ftsI*. These insertions resulting from a 12-bp duplication (YRIN) and
180 from a subsequent mutation (YRIK) were first described in NDM-producing *E. coli*
181 isolates from different STs[17]. Two other insertions at the same position, YRIP and YTIP,
182 were also recently reported and result from 12-bp duplications starting two and three
183 codons upstream the YRIN duplication respectively (Fig. 3a)[18]. We identified the YTIP
184 insertion in a cluster of three ST410 *Ec* isolates (Fig. 1), two from China and one from
185 Brazil. These three isolates do not express a carbapenemase, although the two isolates
186 from China were reported as resistant to carbapenems[19]. In these three isolates, we did
187 not observe signs of recombination. This suggested that the YTIP duplication occurred
188 *de novo* in the ST410 background. Additional NS SNPs were identified in the *ftsI* gene.
189 *ftsI* with the YRIN insertion were also mutated at positions 532 (I/L) and 349 (E/K)
190 whereas isolates with the YRIK insertion were mutated at position 413 (A/V). The same
191 A413V mutation was also present in the three isolates showing the YTIP insertion. The
192 YRIK insertion in PBP3 was previously shown to confer reduced susceptibility to
193 different β-lactams including ampicillin, cefepime and aztreonam but not to
194 carbapenems[17]. Therefore, acquisition of a mutated *ftsI* gene might have contributed to
195 the fixation of the recombined *dcw* locus under selection for resistance against β-lactams
196 targeting PBP3



	Control	YRIN	E349K	I532L	YRIN E349K	E349K I532L	YRIN E349K I532L
ATM	0.032	0.19	0.032	0.047	0.125	0.38	0.5
AMX	5	2	5	10	16	32	28
PIP	1	1.5	1	1.5	4	7	14
CTX	0.023	0.094	0.032	0.064	0.19	0.38	0.75
MER	0.012	0.012	0.008	0.014	0.008	0.016	0.014
ERT	0.002	0.004	0.004	0.004	0.005	0.008	0.008
IMI	0.19	0.19	0.19	0.19	0.125	0.19	0.19
CHL	4	4	4	4	4	4	4

2 4 8 16 32

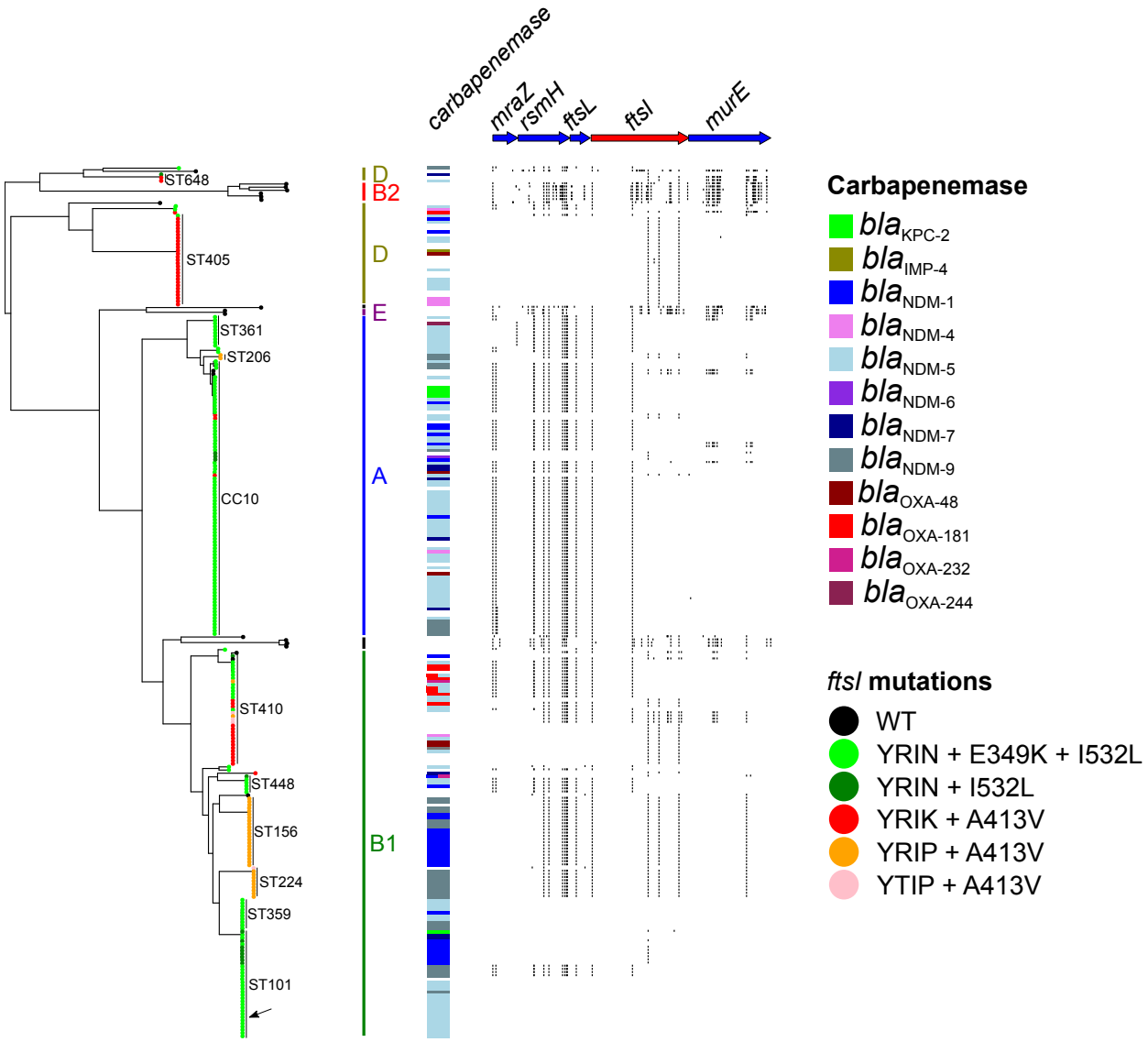
197 **Fig. 3. Functional analyses of *ftsI* and *ompF* mutations occurring in the OXA-181 *Ec***
198 **ST410 subclade. a.** Mutations identified in *ftsI*. The four different insertions following
199 proline 333 resulted in the duplication of the four codons shown in red and blue. The
200 YRIK insertion derived from YRIN by a N to K AA change (in green). The first and second

201 lines represent the wild type nucleotide and AA sequences respectively. On the right, AA
202 substitutions associated with each duplication are indicated. **b.** Antibiotic susceptibility
203 testing (AST) performed by E-test of MG1655 derivatives mutated in *ftsI* as indicated in
204 the first line. Abbreviation: ATM, Aztreonam; AMX, Amoxicillin; PIP, Piperacillin; CTX,
205 Cefotaxime; MER, Meropenem; ERT, Ertapenem; IMI, Imipenem; CHL, Chloramphenicol.
206 **c.** Schematic representation of the four OmpR binding sites in the *ompF* regulatory
207 region and mutation of the conserved cytosine (C) in the F3 OmpR binding site in red. **d.**
208 Expression of *ompF* and *ompC* genes in two strains from the OXA-181 subclade (EcMad
209 and 83B9, mutated) or from the FQR clade (92B7 and 93G1, wild type) grown in LB
210 medium and in LB medium supplemented with 1M NaCl. Bars represent the confidence
211 interval; *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.
212

213 **Recombination of the *dcw* cluster is linked to carbapenemase production.**
214 To determine how general is the association of acquisition of a carbapenemase gene and
215 recombination of the *dcw* gene cluster identified by the four AA insertion in PBP3, we
216 analysed for these two features *E. coli* and *Shigella* genomes from the NCBI database.
217 None of the *Shigella* isolates encoded a carbapenemase gene or carried an insertion in
218 PBP3. 487 *E. coli* isolates (4.4%) encoded a carbapenemase gene and 247 (2.3%) carried
219 a four AA insertion in PBP3: 163 YRIN, 48 YRIK, 3 YTIP and 33 YRIP. 82% (n=204) of
220 isolates mutated in *ftsI* were CP-*Ec* (Supplementary Table 6). All the 164 isolates
221 showing the YRIN insertion were also mutated at position 532 (I/L) and 150 also at
222 position 349 (E/K). Strikingly all YRIK, YTIP and YRIP insertions were associated with
223 the secondary mutation A413V. This A to V AA change is therefore likely selected
224 together with the four AA insertion to reduce susceptibility to antibiotics targeting PBP3
225 or to reduce the fitness cost of the AA insertion. Globally, this reveals at the species level,
226 a strong link between these combinations of mutations in PBP3 and the acquisition of a
227 carbapenemase gene.

228 We were able to trace some recombination events leading to mutations in *ftsI* by
229 analysing polymorphisms at the *dcw* locus considering genome-based phylogeny (Fig. 2
230 and 4). The identification of mutated genes with no signs of recombination showed that
231 YRIN, YRIP and YTIP insertions occurred first in one strain of the ST101, ST156 (Fig. 5)
232 and ST410 lineages (Fig. 1), respectively. *ftsI* with an YRIK insertion is almost identical
233 to ST101 *ftsI* with YRIN insertion and therefore likely derived from it. The four *ftsI*
234 variants were subsequently transferred to other lineages by lateral gene transfer (LGT).
235 For instance, a 29.5 kb region recombined first from ST101 to ST167 and second a 124-
236 kb region from ST167 was recombined into the MRCA of the ST410 OXA-181 subclade

237 (Fig. 2). Similarly, a 65-kb region with the YRIP insertion in *ftsI* from an ST156 *Ec* strain
238 was exchanged by homologous recombination in the MRCA of a clade of NDM-9
239 expressing ST224 *Ec* isolates. Analysis of the SNP distribution using the ST101 sequence
240 as reference shows that these recombination events always include the distal mutated
241 part of *ftsI*. The shortest recombination event, detected in a ST167 CP-*Ec* strain
242 (WCHEC16) contains only the *ftsI* gene (Fig. 5a). In total, we detected 52 independent
243 recombination events involving a mutated *ftsI* allele scattered in all *E. coli* phylogroups
244 except the B2 strains. Strikingly, in *Ec* ST131, despite the large number of CP-*Ec* isolates
245 (n=66), no isolate was mutated in *ftsI* (Supplementary Fig. 3).



246 **Fig. 4. Phylogenetic distribution of *E. coli* isolates mutated in *ftsI*.** The genome based
247 phylogeny of *Ec* isolates mutated in *ftsI* and of *E. coli* reference strains[20] was constructed
248 by using Parsnp[21]. Mutations in the *ftsI* gene are indicated at the tip of the branch as
249 defined in the figure keys. Major STs are indicated, as well as phylogroups, S standing for

Shigella. Carbapenemase genes are indicated as defined in the figure keys on the right. SNPs in the *mraZ* – *murE* region compared to the *Ec* ST101 strain CREC-591 (CP024821) shown by a black arrow are indicated by small vertical black bars. An ST101 isolate was chosen as reference, as insertion of YRIN in *ftsI* occurred among this ST. Conservation of SNP patterns between isolates showing the same mutation in *ftsI* across the *E. coli* species revealed the exchange of these alleles by recombination.

In addition to ST410, *ftsI* was mutated in the majority of the CP-*Ec* isolates from ST101 (100%, N=33), ST167 (89%, N=76) and ST405 (83%, N=15) (Supplementary table 6). Multiple acquisitions of carbapenemase genes across the phylogeny of *Ec* ST167 have been recently reported[10]. We detected within this single ST eleven events of recombination corresponding to YRIN (n=9) and YRIK (n=2) insertions in PBP3. Nine affected at least one CP isolate, including a sub-clade of 56 isolates expressing carbapenemases from ten different types. Only nine out of the 85 *Ec* ST167 CP isolates did not undergo recombination at *ftsI* (Fig. 5a).

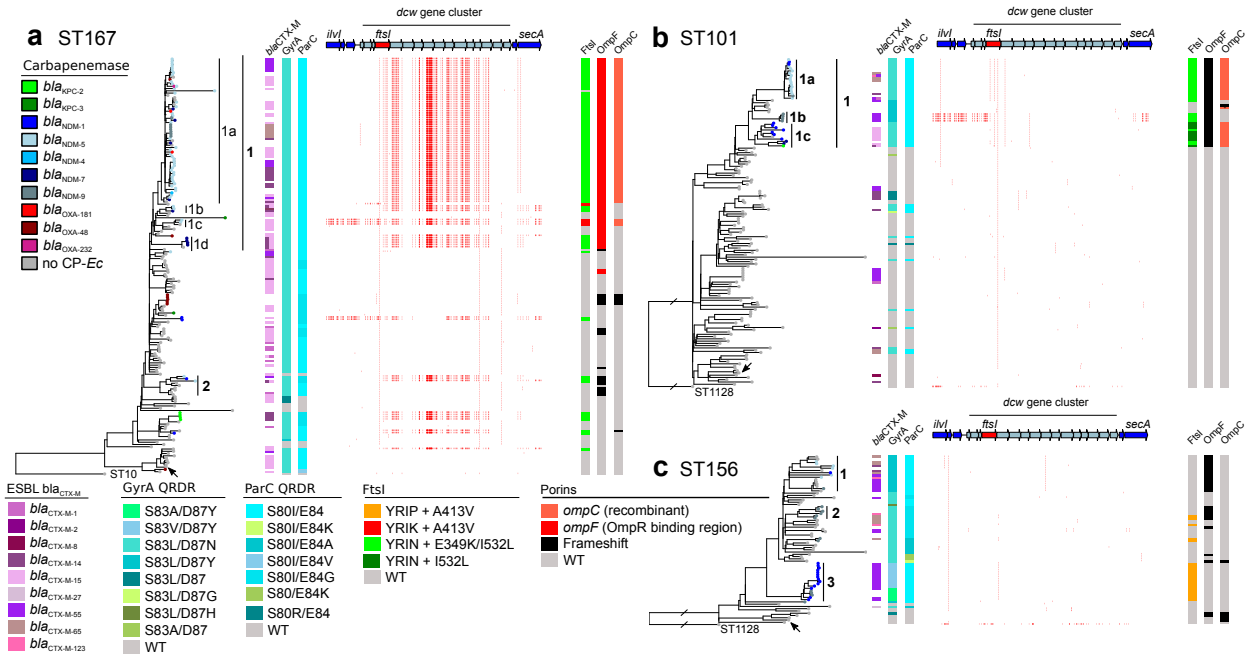


Fig. 5. Phylogeny and mutations in CP-*Ec* isolates of ST167, ST101 and ST156. Maximum likelihood phylogenies were estimated as for Fig. 1, using 5871, 2166, and 18774 non-recombinant SNPs for **a.** *Ec* ST167, rooted with the *Ec* ST10 strain MG1655 (NC_00913) **b.** *Ec* ST101, and **c.** *Ec* ST156 respectively, both rooted with the ST1128 strain IAI1 (NC_011741). 14 distantly related non CP-*Ec* ST156 isolates have been removed from the phylogenetic analysis. Branch tips indicates the presence and the type of carbapenemase according to the figure key on the left. On the right side of the tree the genetic features analysed in the present work are represented, from left to right: $bla_{\text{CTX-M}}$ ESBL, *gyrA* and *parC QRDR* mutations, the *dcw* region SNPs (represented by small vertical red bars), mutations in the *ftsI* gene, and different genetic events affecting *ompF* and *ompC* according to the figure keys at the bottom. Genes from the *dcw* locus are indicated by arrows and the *ftsI* gene in red. Black arrowheads indicate the isolate used

276 as reference for SNPs mapping in the *dcw* gene cluster. Clusters and subclusters as
277 defined in Table 1 are indicated by thin vertical lines next to the trees.
278
279 **Mutations in *ftsI* associated with the YRIN insertion contribute to a**
280 **decreased susceptibility to β-lactams targeting PBP3.** In order to determine the
281 contribution of the three mutations (YRIN insertion, E349K and I532L) in *ftsI* identified
282 in the ST410 OXA-181 subclade to the resistance to β-lactams, we constructed
283 derivatives of the ST10 *s*train MG1655 with combinations of these mutations (Fig. 3b).
284 Individually, each mutation showed a low effect on susceptibility to β-lactam targeting
285 PBP3[16]. However, the combination of two or three mutations led to a stronger decrease
286 in the susceptibility to these antibiotics. In particular, the MG1655 derivative with the
287 three mutations showed in the absence of any β-lactamase a 32, 16 and 14-fold increase
288 of the MIC to the third-generation cephalosporin cefotaxime, to the monobactam
289 aztreonam and to piperacillin respectively. This strain showed only a slight increase in
290 the MIC to ertapenem (x4), which targets mainly PBP2 but also in a lesser extend
291 PBB3[22]. Susceptibility to meropenem and imipenem that show low affinity for PBP3,
292 was not affected.
293
294 **Mutation in the porin genes *ompC* and *ompF* are predicted to also have**
295 **contributed to the selection of the ST410 OXA-181 subclade.** To identify additional
296 polymorphisms that might have contributed to the dissemination of the *Ec* ST410 OXA-
297 181 subclade, we analysed the potential effect of nucleotide substitutions in the branch
298 leading to its MRCA by using the SIFT algorithm[23]. We identified 34 NS mutations with a
299 predicted functional effect (nine in the recombined region) (Supplementary Table 7).
300 These mutations might have been selected for decreased susceptibility to antibiotics or
301 to alleviate the fitness cost of other mutations, in particular those affecting genes from
302 the class "transporter" (n=8) including the multidrug efflux transporter components
303 *emrD* and *emrK* and the class "cell envelope" (n=5) including *ftsI* mutation I532L.
304 Among other mutations affecting functions related to the cell envelope, one
305 mutation detected by SIFT affected the porin gene *ompC* at a conserved arginine residue
306 in the L4 loop (R191L), one of the gateways for carbapenems (Supplementary Fig. 4)[24].
307 Arg_{191} is exposed, at the vestibule of the pore lumen and is conserved in OmpF[25]. We
308 hypothesize that its replacement by leucine, a non-polar AA, affects permeation of β-
309 lactams into the periplasm.

310 Although *ompF* coding sequence was not mutated, we identified a mutation in the
311 regulatory region replacing a conserved cytosine to a thymine residue in the proximal
312 OmpR binding site, F3 (Fig. 3c)[26]. To quantify the impact of this mutation on *ompF*
313 expression, we compared *ompF* and *ompC* expression at low and high osmolarity (LB
314 and LB supplemented with 1M NaCl) in two isolates from the OXA-181 subclade
315 (mutated) and two isolates from the FQR clade (non-mutated). We observed a 15 to 30
316 and 5-fold reduction in *ompF* expression in LB and LB-NaCl respectively in the mutated
317 isolates compared to wild-type, whereas *ompC* expression remained unchanged (Fig.
318 3d). The strong decrease in *ompF* expression is predicted to reduce β-lactams entry into
319 the periplasm, possibly with a lower fitness cost than gene inactivation.

320 **Acquisition of carbapenemase genes occurred preferentially in genetic**
321 **backgrounds first mutated in *ompC*, *ompF* and/or *ftsI*.** In order to determine the
322 contribution of mutations in *ompC* and *ompF* in the acquisition of carbapenemase genes,
323 we analysed CP-*Ec* isolates for mutations in the two porin genes. We focused on the 18
324 STs with at least 5 CP-*Ec* isolates (Supplementary table 6). Within these STs we searched
325 for mutations in *ompC* and *ompF* coding and promoter regions compared to the
326 predicted ancestral sequence of the ST and reconstruct their occurrence in the
327 phylogeny. Out of the 489 CP-*Ec* isolates, 242 and 202 isolates were mutated in *ompF*
328 and *ompC* respectively, compared to only 138 and 119 out of the 1471 non CP-Ec
329 isolates (supplementary table 1 and 2). Therefore, likewise *ftsI* mutations, the presence
330 of a carbapenemase gene was frequently associated with mutations in *ompC* and *ompF*.

331 Analysis of the distribution of CP-*Ec* isolates within the 18 STs revealed in ten of
332 them evidence of lineage expansion, i.e. lineages containing isolates from different
333 geographical origins and at least three CP-*Ec* isolates (Table 1). Up to 24 lineages
334 enriched in CP-*Ec* isolates were detected, which encompass 338 out of the 489 CP-*Ec*
335 isolates from these 18 STs. Strikingly we identified mutations in *ompF*, *ompC* and/or *ftsI*
336 occurring in the MRCA from most clusters. Furthermore, for three of them, ST101-1,
337 ST167-1 and ST405-1, we were able to define subclusters based on successive
338 mutations in the three genes.

339 **Table 1**: Clusters of CP-*Ec* isolates

Cluster	Subcluster*	Isolates†	Dissemi-nated‡	*ftsI*§	*ompC*#	*ompF*&
ST38-1		21 (21,1)	Y		Group D	
ST38-2		13 (13,1)	Y		Group D	Frameshifted
ST48-1		18 (5,2)	Y		Recombined	
ST101-1	a	19 (18,3)	Y	YRIN - E349K-I532L	Recombined	Frameshifted
	b	4 (4,1)		YRIN - E349K-I532L		
	c	11 (10,3)		YRIN - E349K-I532L	Recombined	
	other	5 (0)				
ST156-1		16 (8,3)	Y			Frameshifted
ST156-2		6 (6,1)	N			Frameshifted
ST156-3		17 (17,2)	Y	YRIP - A413V		
ST167-1	a	64 (57,7)	Y	YRIN - E349K-I532L	Recombined	OmpR-box
	b	3 (3,2)		YRIN - E349K-I532L		
	c	3 (2,1)		YRIK - A413V	Recombined	
	d	5 (4,1)		YRIN - E349K-I532L		
	other	8 (2,2)				
ST167-2		9 (3,2)	Y			Frameshifted
ST226-1		8 (6,1)	N			Frameshifted
ST224-1		9 (9,1)	N	YRIP - A413V		
ST354-1		4 (4,1)	N			Frameshifted
ST359-1		10 (10,3)	N	YRIN - E349K-I532L	Recombined	Frameshifted
ST361-1		8 (8,2)	Y	YRIN - E349K-I532L	G138D	
ST405-1	a	6 (5,1)	Y	YRIK - A413V	Group D	Frameshifted
	b	7 (1,1)		YRIK - A413V		
	other	23 (4,3)				
ST405-2		8 (5,4)	Y	YRIK - A413V	Group D	
ST405-3		4 (3,2)	Y	YRIK - A413V	Group D	
ST410-1		60 (58,2)	Y	YRIN- E349K-I532L	R191L	OmpR-box
ST410-2		15 (6,4)	Y	YRIK - A413V		
ST448-1		5 (3,2)	Y	YRIN - E349K-I532L		
ST617-1		6 (5,2)	N	YRIN - E349K-I532L		
ST648-1		14 (14,1)	N		Group D	
ST746-1		16 (11, 1)	Y		Recombined	
ST131-1		9 (9,1)	N			
ST131-2		6 (6,1)	N			

340 *clusters and subclusters as shown in Fig. 1, *Fig*. 5 and Supplementary Fig. 3, 5, 6 and 7 and in
341 Supplementary Table 1 and 2; †number of isolates in clusters or subclusters, in parentheses,
342 number of CP-*Ec* isolates and number of carbapenemase types, ‡Considered as disseminated (Y)
343 if it includes isolates from different countries; §Mutations in *ftsI* are according to Fig. 3a, #"Group
344 D" means that the isolates are of phylogroup D and "recombined" that the *ompC* gene has been
345 replaced by an allele from the phylogroup D type. &OmpR-box corresponds to mutations in an
346 OmpR binding site.

347 In total, we identified *ompF* mutations affecting 10 clusters of isolates
348 corresponding to 212 CP-*Ec* isolates out of the 242 CP-Ec isolates mutated in this porin
349 gene. These mutations either led to *ompF* inactivation by a frameshift or altered OmpR
350 boxes in the promoter region as observed in the ST410 OXA-181 subclade. On the other
351 hand, nine genetic events in *ompC* occurred in the MRCA from clusters or subclusters

352 and affected 182 isolates out of 202. In addition to the R191L mutation in the OXA-181
353 ST410 subclade we identified in the disseminated lineage ST361-1 a mutation (G138D)
354 in the constriction loop L3 of OmpC[25] predicted to reduce its permeability to β-lactams
355 (Supplementary *Fig*. 4). The same AA change occurred independently in CP-*Ec* isolates
356 from ST448 and ST617 (Supplementary Fig. 5). However, the most frequent *ompC*
357 modification associated with CP-*Ec* clusters was its replacement by recombination by
358 alleles originating from phylogroup D strains (n=7), as observed in the broadly
359 distributed ST167-1a subcluster with a 22.7 kb recombined region from ST38 (Fig. 2c)
360 Furthermore, in addition to ST38[8,9] four other STs from phylogroup D: ST354, ST405,
361 ST457 and ST648 contained CP-*Ec* isolates (Supplementary Fig. 6). Strikingly, OmpC
362 proteins from phylogroup D isolates differ from other *E. coli* OmpC proteins at the two
363 previously mentioned G138 and R191 residues, suggesting a naturally reduced
364 permeability to ß-lactams(Supplementary Fig. 4). Mutations in *ftsI* were the most
365 frequent events as they were present in the MRCA of 19 lineages or sublineages
366 enriched in CP-*Ec* isolates and affected 235 isolates.

367 Eventually, all disseminated clusters were mutated in at least *ftsI*, *ompC* or *ompF*
368 or have a *ompC* allele from phylogroup D. Seven clusters or subclusters like the ST410
369 OXA-181 subclade were mutated in the three loci. The comparison of the occurrence in
370 the phylogeny of mutations in *ompF*, *ompC* and *ftsI*, and the acquisition of
371 carbapenemase genes revealed that mutations in most cases occurred first, contributing
372 to the emergence of lineages and sublineages, which subsequently acquired
373 carbapenemase genes. This model is best exemplified by the three lineages ST101-1,
374 ST167-1 (Fig. 5) and ST405-1 (Supplementary Fig. 6) first mutated in *ompF*, then in
375 *ompC* and/or in *ftsI* leading to nine sublineages that gained carbapenemase genes from
376 different classes. On the contrary, we did not detect clusters of isolates mutated in one of
377 these three genes among ST131 isolates. Furthermore, we did not gain evidence for
378 internationally disseminated lineages among ST131 isolates. Among the 66 ST131 CP-*Ec*
379 isolates, we identified only two clusters of isolates but without any evidence of
380 international dissemination[12,27] (Table 1, Supplementary Fig. 3).

381 **Isolates of the OXA-181 subclade show higher resistance without *in vitro* fitness**
382 **cost as compared to other isolates of the FQR clade.** To determine the impact of
383 mutations and ARG on the resistance of *Ec* isolates along the ST410 FQR clade, we
384 compared, by disk diffusion assay and E-test, β-lactam susceptibility of six isolates from
385 the *Ec* ST410 FQR-clade producing or not OXA-181 and with different mutations (Fig. 6).
386 These isolates were chosen as they all express CTX-M15 and all but one express Oxa-
387 181. We observed a gradual decrease in the susceptibility to diverse β-lactams between
388 the four groups of isolates: CTX-M15 < CTX-M15, OXA-181 < CTX-M15, OXA-181, YRIK
389 insertion in PBP3 < OXA-181 subclade (Fig. 6a). Isolates from the OXA-181 subclade
390 were resistant to almost all β-lactams tested except carbapenems. Higher resistance was
391 partly due to differences in β-lactamase gene content including $bla_{\text{OXA-181}}$
392 (Supplementary table 8). However, mutations in *ftsI*, including the YRIK and YRIN
393 insertions were likely responsible for the decreased susceptibility of ST410 Ec isolates
394 to β-lactams targeting PBP3 like ceftazidime and aztreonam (Fig. 6b). Surprisingly, the
395 isolates from the OXA-181 subclade show only a minor decrease in the susceptibility to
396 carbapenems compared to ST410 isolates $bla_{\text{OXA-181}}$ positive not mutated in *ftsI*, *ompC*
397 and *omp*F (Fig. 6b).
398 To determine whether these mutations have an impact on fitness we compared
399 growth parameters in LB, MH and M9 media as proxy. Despite the higher resistance to
400 antibiotics, we did not detect any significant differences in the growth parameters in
401 rich medium among the six isolates tested (Fig. 6c) suggesting these mutations have no
402 fitness cost or their effect has been compensated by other mutations. The two isolates
403 from the OXA-181 subclade and the non-OXA-181 isolate 32139 grew at a higher OD_{600}
404 in minimal medium than the three other isolates (Fig. 6c). Therefore, the high level of
405 resistance to most β-lactams and the decreased susceptibility to carbapenems of isolates
406 of the OXA-181 subclade do not seem to be at the expense of a lower *in vitro* fitness for
407 the isolates we have analysed.

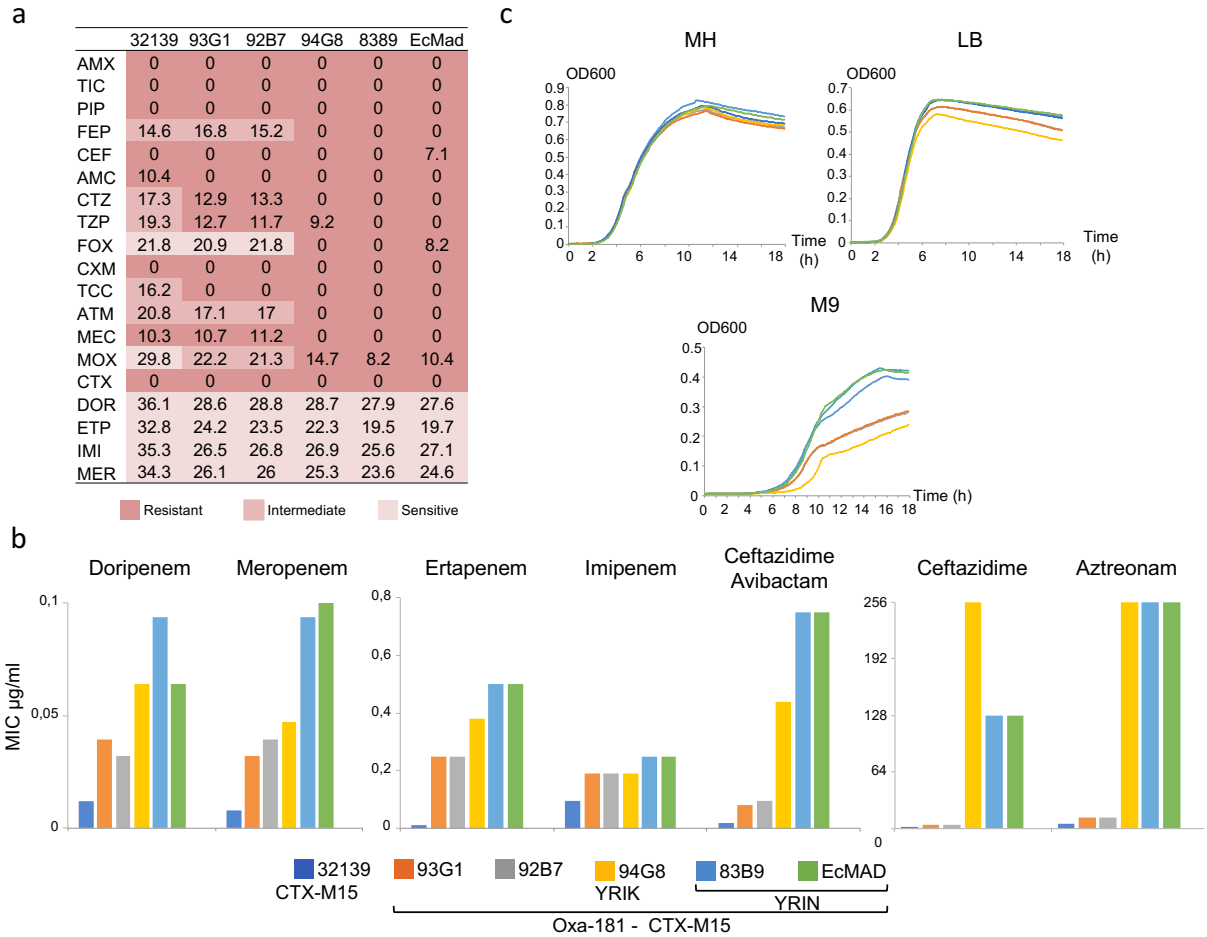


Fig. 6. β-lactams susceptibility profiles and fitness of *Ec* ST410 strains. **a.** β-lactams susceptibility determined by disk diffusion. Diameters are indicated in mm. Resistant, Intermediate and susceptible according to CLSI guidelines[28] are indicted by colours as defined in the figure keys. 32139 carries $bla_{\text{CTX-M15}}$, 93G1 and 92B7 carry $bla_{\text{CTX-M15}}$ and $bla_{\text{OXA-181}}$, 94G8 carries $bla_{\text{CTX-M15}}$ and $bla_{\text{OXA-181}}$ and a mutated *ftsI* gene (YRIK), 8389 and *Ec*Mad belong to the OXA-181 subclade. ***b.*** Minimal inhibitory concentrations (MIC) determined by E-test for selected β-lactams. **c.** Growth curved in rich (LB and Müller Hinton, MH) and minimal (M9) medium. Curves represent the average value of 10 experiments. Blox plot representations of the area under the curve of replicates determined by using growthcurver [29] are given in Supplementary Fig. 8. Abbreviations: AMX, Amoxicillin; TIC, Ticarcillin; PIP, Piperacillin; FEP, cefepime; CEF, cephalothin; AMC, Amoxicillin - clavulanic acid.; CTZ, Ceftazidime; TZP, piperacillin-tazobactam; FOX, Cefoxitine; CXM, Cefuroxime; TCC, Ticarcilline - clavulanic acid; ATM, Aztreonam; MEC, Mecillinam; MOX, Moxalactam; CTX, Cefotaxime; DOR, Doripenem; ETP, Ertapenem; IMI, Imipenem; MER, Meropenem. Antibiotic resistance genes repertoires of the six strains are given in Supplementary table 8.

424 **Discussion**

425

426 The global dissemination of antibiotic resistant clones results from a trade-off between
427 acquired resistance and biological cost[30]. Here, we show that carbapenemase genes were
428 preferentially fixed in lineages mutated in genes contributing to β-lactam resistance.
429 Mutations in *ompC* and *ompF* affect permeation into the periplasm, whereas mutations
430 in *ftsI* reduce the susceptibility to antibiotics targeting preferentially PBP3. We have
431 described 15 disseminated lineages mutated in one, two or the three loci. The recently
432 reported CP-*Ec* lineages from *Ec* ST410 B4/H24RxC[13] (OXA-181 subclade) and *Ec*
433 ST167[10] are mutated in these three genes. Most clusters include CP-*Ec* isolates
434 producing different carbapenemases indicative of multiple acquisitions in the same
435 genetic background as observed in cluster ST167-1 (Fig. 5a). However, a global
436 dissemination following the acquisition of a carbapenemase gene is evidenced in the
437 ST410 OXA-181 subclade where all isolates share the same plasmid expressing $bla_{\text{OXA-181}}$
438 [13] (Fig. 1) and in clusters ST38-1 and -2 where the $bla_{\text{OXA-48}}$ gene is inserted at conserved
439 position on the chromosome[9] (Supplementary Fig. 6). Contribution of mutations in porin
440 genes to carbapenem resistance is well documented. It is generally assumed that these
441 mutations are selected following the acquisition of carbapenemase genes under the
442 selective pressure of carbapenems. However, our analysis reveals a different scenario
443 with mutations in these genes occurring before the acquisition of carbapenemase genes.
444 In six clusters, we observed successive mutation events, like in cluster ST167-1, first
445 mutated in the *ompF* promoter region and subsequently submitted two and five times
446 independently to recombination of *ompC* and *ftsI* genes respectively (Fig. 5a).

447 We identified four combinations of mutations in *ftsI* associated with CP-*Ec*
448 isolates. Each arose only once and disseminated widely across the species by LGT
449 (Fig. 4). The probability of such combinations is very low, but the strength of the
450 selective advantage led to their multiple independent fixation in different genetic
451 backgrounds except in phylogroup B2. Whereas, the inactivation of *ompF* gene or its
452 decreased expression by mutations in OmpR binding sites was associated with the
453 dissemination of nine clusters of CP-*Ec*, inactivation was not observed for *ompC* (Table
454 1). Loss of OmpC might have a high fitness cost precluding further spreading. On the
455 other hand, the replacement by LGT of *ompC* by alleles originating from phylogroup D
456 strains is the most frequent event with a total of 15 independent events associated with

457 CP-*Ec* isolates. Overall, the contribution of the recombination of specific *ompC* and *ftsI*
458 alleles in the evolution of β-lactam resistance of CP-*Ec* isolates along with the acquisition
459 of resistance plasmids was largely underestimated.
460 Despite the high prevalence of ST131 CP-*Ec* reported in different studies[11,31,32], there was no
461 indication of global dissemination of a specific lineage among the 66 ST131 CP-*Ec* genome
462 sequences we have analysed (Supplementary Fig. 3). We did not observe replacement of *ftsI*
463 by mutated alleles or of *ompC* by alleles from the phylogroup D. ST131 isolates had a major
464 contribution to the dissemination of CTX-M family ESBL. Surprisingly, only 30% of the
465 ST131 CP-*Ec* isolates produced also a CTX-M ESBL compared to 80.5% for CP-*Ec* isolates
466 from clusters in other STs (Supplementary Table 2). Selection of ST131 CP-*Ec* isolates
467 follows therefore a different trend compared to other lineages that might not be compatible
468 with their global dissemination yet.
469 Despite the presence of the $bla_{\text{OXA-181}}$ gene and mutations in *ftsI*, *ompC* and *ompF*,
470 isolates of the *Ec* ST410 OXA-181 subclade do not show a phenotype of resistance to
471 carbapenems. In contrast, they are highly resistant to other β-lactams. This high level of
472 resistance compared to those of other isolates from the ST410 FQR clade results both
473 from the genomic mutations and from differences in β-lactamase spectrum (Fig. 6a,
474 Supplementary table 8) and seems not to be at the expense of fitness as compared to
475 isolates expressing only CTX-M15 (Fig. 6c). In particular the acquisition of mutations in
476 *ftsI* by LGT is predicted to led to a strong decrease in susceptibility to β-lactams of
477 therapeutic importance like aztreonam and cefotaxime, as shown in the ST10 MG1655
478 background (Fig. 3) with only a minor effect on ertapenem resistance. OXA-181
479 hydrolyses different β-lactams in addition to carbapenems[33]. Therefore, we propose that
480 carbapenems were not the main driving force for the acquisition of $bla_{\text{OXA-181}}$ by the *Ec*
481 ST410 subclade. This model for the selection of CP-*Ec* lineages is in line with the
482 observation that VIM-1-producing *Ec* isolates are detected in animal herds in Germany
483 despite the fact that carbapenems have not been used in this context[34] and of the significant
484 prevalence of bla_{NDM}-positive *E. coli* in Chinese poultry production [35].
485 The risk of a global spreading of CP-*Ec* strains in the hospital and in the community is
486 a major concern. These new perspectives on the selection of CP-*Ec* clones largely
487 independent of the use of carbapenems need to be taken into account in β-lactams usage
488 recommendation to prevent their dissemination. Furthermore, the follow up of mutations and
489 recombination events in *ompC*, *ompF* and *ftsI* deserves to be considered in the global

490 surveillance of β-lactam resistance in *E. coli* as lineages mutated in these genes are
491 preferential recipients of carbapenemase genes.
492
493
494 ## Material and methods
495
496 **Bacterial isolates, growth conditions and antibiotic susceptibility testing.** Features
497 of the *E. coli* isolates analysed in this work are listed in Supplementary table 1. 51 *Ec*
498 ST410 isolates came from the strain collection of the French National Reference Centre
499 (Fr-NRC) for Antibiotic Resistance. Four *Ec* ST410 clinical isolates came from the
500 Microbiological collection of the Université Libanaise (CMUL). the Public Health Faculty
501 of the Lebanese University (Tripoli, Lebanon) and three ST410 isolates of animal origin
502 from the ANSES strain collection. Antibiotic susceptibility was performed by the disk
503 diffusion method or by broth microdilution assays following the Clinical & Laboratory
504 Standards Institute (CLSI) guidelines[28] or by E-test (Biomérieux) following the
505 manufacturer's recommendations. Fitness was determined by growth curve analysis
506 with an automatic spectrophotometer Tecan Infinite M200 during 24 hours in LB,
507 Müller Hinton or M9 media supplemented with 0.4% glucose. Growth metrics were
508 estimated with the R package "growthcurver" [29]. The area under the curve (AUC) which
509 include contributions of the most important growth parameters (log phase, growth rate,
510 and the carrying capacity) was used as a growth metric.
511
512 **Genome sequencing, and genome sequences retrieved from sequence libraries.** *Ec*
513 genomes were sequenced by using the Illumina HiSeq2500 platform, with 100
514 nucleotides (nt) single reads for the four isolates from Lebanon and 100 nt paired-end
515 reads for the other isolates. Libraries were constructed by using the Nextera XT kit
516 (Illumina) following the manufacturer's instructions. The OXA-181-producing *Ec*-MAD
517 ST410 isolate was selected as a reference strain and sequenced to completion by using
518 the long read PacBio technology. 10 947 *E. coli* and 1 451 *Shigella* genome sequences
519 deposited in the NCBI database (19th June, 2018) were retrieved for a global analysis of
520 the specificity of CP-*Ec* (Supplementary table 2). 96 additional *Ec* ST167 isolates were
521 retrieved from Enterobase (https://enterobase.warwick.ac.uk/). Raw reads from 62 *Ec*

522 ST410 and 21 *Ec* ST38 isolates identified in Enterobase (march 2017) were retrieved
523 from the NCBI database. (Supplementary table 2)
524
525 **Sequence assembly, genome annotation and mutations identification.** The PacBio
526 reads were assembled with the RS_HGAP_Assembly.3 protocol from the SMRT analysis
527 toolkit v2.3[36], and with Canu[37]. The consensus sequence was polished with Quiver[36] and
528 by mapping Illumina reads. Illumina sequenced isolates were assembled with SPAdes[38],
529 and the quality of the assemblies was assessed with Quast[39]. Contigs shorter than 500
530 bp were filtered out. All assemblies were annotated with Prokka[40]. The presence of
531 antibiotic resistance genes, and plasmid replicons was assessed with ResFinder[41] and
532 PlasmidFinder [42] respectively. The software was run in local from the script and
533 databases downloaded from the repositories of the Centre for Genomic Epidemiology
534 (https://bitbucket.org/genomicepidemiology/). Graphs of genomic regions of interest
535 were drawn with genoplotR[43]. All downloaded genomes were reannotated with
536 prokka[40]. For each ST type analysed the pangenome was characterized with Roary[44],
537 and the AA sequences of OmpC, OmpF, GyrA, ParC, and FtsI were identified from the
538 orthologous table generated by Roary. To determine the presence of different *ompC*
539 alleles within the different STs, OmpC AA sequences were clustered with cd-hit[45], with a
540 sequence identity threshold of 0.95. AA sequences for GyrA and ParC were aligned with
541 the mafft L-INS-i approach[46], and the amino acids at the QRDR positions for both genes
542 were filtered with customized perl scripts.
543
544 **Phylogenetic reconstruction.** A core genome alignment was obtained for all *Ec* ST410
545 isolates, and strain 789 that belongs to the ST88 as an outgroup with parsnp[21]. Non-
546 recombinant SNPs were extracted from the core genome alignment with Gubbins[15], and
547 the Maximum likelihood (ML) phylogeny was obtained with RAxML[14], using the General
548 Time-Reversible (GTR) substitution model with a gamma-distributed rate. The same
549 procedure was followed for the phylogenetic reconstruction of the different *E. coli*
550 lineages. Maximum likelihood phylogeny of OmpC protein sequences was inferred with
551 RAxML[14]. OmpC protein sequences were aligned with the mafft L-INS-i approach[46].
552 Gblocks[47] was used to "clean" the alignment and the best fit model (WAG, with a gamma
553 distribution) was estimated with protest 3[48]. The visual display of phylogenetic trees
554 was done with FigTree (http://tree.bio.ed.ac.uk/software/figtree/), and annotated trees

555 with the script plotTree.R (https://github.com/katholt/plotTree). The final version of
556 the tree was edited with Inkscape (https://inkscape.org/es/).
557
558 **Mapping, variant calling, and identification of SNPs of interest**. Sequence reads for
559 *Ec* ST410 isolates were mapped against the complete reference genome *Ec*-MAD with
560 BWA[49]. Variant calling was performed with the Genome Analysis Toolkit v 3.6.0[50]. The
561 criteria for variants were: occurrence of the alternative base in more than 90% of the
562 reads covering the position, a depth coverage of at least 10 (DP > 10), a quality by depth
563 (QD) > 2, a fisher strand bias (FS) < 60, a mapping quality (MQ) > 40, a mapping quality
564 rank sum test (MQRankSum) > -12.5, and a read position rank sum test
565 (ReadPosRankSum) > -8. Variants associated with different clades of the *Ec* ST410
566 phylogeny were extracted with vcftools[51], and finally they were annotated with snpEff[52].
567 The effect of the non-synonymous mutations was assessed with the Sorting Intolerant
568 From Tolerant (SIFT) algorithm[23]. The algorithm searches for protein homologs in the
569 refseq database using mutated proteins as query and assigns a score to each position.
570 This score is weighted by the properties of the AA changed. If this score is below a
571 threshold (0.05) the change is predicted to be functional.
572
573 **Construction of *ftsI* mutant strains derivatives.** To analyse the potential effects on
574 antibiotic resistance of the mutations identified in the *ftsI* gene of the isolates from the
575 OXA-181 subclade, we introduced the 12 nt insertion (YRIN form) and the two non-
576 synonymous SNPs (E349K and I532L) into the *Ec* strain MGF
577 (MG1655strepR_F'tet_ΔtraD::Apra) by TM-MAGE[53]. Briefly, overnight culture of strain
578 MGF transformed by pMA7SacB was used to inoculate 5 ml LB medium supplemented
579 with tetracycline (7.5 µg/l) and carbenicillin (100 µg/l) (LB-TC) and grown at 37°C until
580 OD_{600} reached 0.6-0.7. The recombinase and Dam methylase were induced by addition
581 of L-arabinose (final concentration of 0.2% w/v), and further incubation for 10 min.
582 Cultures were then chilled for 15 min on ice and centrifuged at 7300g at 4°C. Two
583 successive washes by 50 and 10 ml of cold water were performed and the final pellet
584 was resuspended in 200 µl water. 100 µl of cells were used for electroporation with 2 µl
585 oligonucleotides Mut1ftsI or Mut2ftsI (Supplementary table 9) alone or in combination
586 at 20 µM each. The Mut1f*tsI* oligonucleotide carries both the 12 nt insertion and the
587 E349K mutation, while the Mut2*ftsI* oligonucleotide has the I532L mutation. The content

588 of the electroporation cuvette was used to inoculate 5 ml of LB-TC and submitted to
589 three additional cycles of growth-induction-preparation of electrocompetent cells and
590 electroporation. Following the last electroporation step, cells were resuspended in1 ml
591 LB and plated onto LB-TC agar plates. Mutations in isolated colonies were tested by PCR
592 using primers complementary to mutant or wild-type alleles (Supplementary table 8).
593 Mutated colonies were grown on plates containing 10 g/l tryptone, 5 g/l yeast extract,
594 15 g/l agar and 5% w/v sucrose for plasmid curing. Mutant strains were sequenced by
595 using Illumina MiSeq platform, with 150 nt paired-end reads and Nextera XT kit
596 (Illumina) for library preparation. Reads were mapped onto the MG1655 genome
597 (Genbank NC_000913.3) to confirm that mutations in *ftsI* gene have been correctly
598 introduced and to check that the rare off target mutations are not predicted to interfere
599 with the β-lactam susceptibility phenotype (Supplementary table 9).
600
601 **RNA extraction and quantitative RT-PCR.** Bacteria were grown in LB medium until
602 OD_{600} reached 0.30-0.33, an osmotic shock was applied by supplementing 10 ml of
603 culture with 0.3M, final concentration of Sodium Chloride (NaCl) or with the same
604 volume of water as control and further incubation for 20 minutes. Bacterial pellets were
605 collected and stored at -80ºC. Total RNA was extracted with the Total RNA purification
606 kit Norgen Biotek. cDNAs were synthetized from 500 ng of RNA with the Superscript II
607 reverse transcriptase (Invitrogen, Life Technologies). Primer pairs were designed for
608 the *ompC* and *ompF* genes, targeting divergent regions from these two genes, and for the
609 reference gene *recA* (Supplementary Table 8). The SYBR Green PCR kit (Applied
610 Biosystems, Life Technologies) was used to perform quantitative PCR, and the relative
611 expression of porin genes was measured by a standard curve method where the
612 regression analysis was performed from serial dilutions of a mixture of control cDNAs.
613 The expression value of each gene was normalized against the expression of the
614 housekeeping gene *recA*. Each point was measured in triplicate, and three independent
615 cultures were used for each strain in each condition.
616
617 **Statistical analysis** The statistical significance of the differences in expression in RT-
618 PCR experiments was assessed by using a two-tailed t-test. The statistical significance of
619 the differences in the number of ARG between the three Ec ST410 groups and of the area
620 under the curve between six isolates from the ST410 FQR clade was assessed using the

621 Wilcoxon rank sum-test implemented in R (v3.4.4). One-sided test was used for the
622 comparison of the number of ARG, and two-sided test was used for the comparison on
623 the area under the curve analysis."
624
625 **Data availability**. Illumina reads from the 57 newly sequenced isolates and the
626 complete genome assembly of strain Ec-MAD have been deposited in the EMBL
627 nucleotide sequence database (http://www.ebi.ac.uk/ena) under study accession
628 number PRJEB27293 and PRJEB27274 respectively. The accession numbers for
629 individual isolates are listed in Supplementary Table 1. Code of software scripts could be
630 obtained upon request.
631
632
633 **Acknowledgements**
634 This work was supported by grants from the French National Research Agency (ANR LabEx
635 IBEID and ANR-10-LABX-33), from the Joint Program Initiative on Antimicrobial
636 Resistance (ANR-14-JAMR-0002) and from One-Health EJP ARDIG. The authors thank
637 Alexandre Almeida and Adriana Chiarelli for their discussion and critical reading of the
638 manuscript and Laurence Ma for her help in performing the Illumina sequencing.
639
640
641 **Authors contribution**
642 RPN, IRC and NC performed experiments; LG, JT, JYM, MH, RB and TN collected
643 and provided the samples for the study; RPN, retrieved genome sequences from public
644 databases, wrote scripts and performed all the bioinformatics analyses; RPN, IRC, NC
645 and PG analysed the data; TN provided expertise on CP-*Ec*; RPN and PG designed the
646 study; PG was responsible for management of the project; RPN, IRC, TN and PG wrote
647 the manuscript. All authors approved the manuscript prior to submission.
648
649 **Conflict of interest**
650 The authors declare no competing financial interests.

651 **References**

652

653 1 J., O. N. Tackling Drug-Resistant Infections Globally: Final Report and
654 Recommendations: Review on Antimicrobial Resistance. (2016).
655 2 Papp-Wallace, K. M., Endimiani, A., Taracila, M. A. & Bonomo, R. A. Carbapenems:
656 past, present, and future. *Antimicrob. Agents Chemother.* **55**, 4943-4960,
657 doi:10.1128/aac.00296-11 (2011).
658 3 Nordmann, P., Dortet, L. & Poirel, L. Carbapenem resistance in
659 Enterobacteriaceae: here is the storm! *Trends in molecular medicine* **18**, 263-272,
660 doi:10.1016/j.molmed.2012.03.003 (2012).
661 4 Logan, L. K. & Weinstein, R. A. The Epidemiology of Carbapenem-Resistant
662 Enterobacteriaceae: The Impact and Evolution of a Global Menace. *J. Infect. Dis.*
663 **215**, S28-s36, doi:10.1093/infdis/jiw282 (2017).
664 5 Bush, K. & Jacoby, G. A. Updated functional classification of beta-lactamases.
665 *Antimicrob. Agents Chemother.* **54**, 969-976, doi:10.1128/aac.01009-09 (2010).
666 6 Canton, R. & Coque, T. M. The CTX-M beta-lactamase pandemic. *Curr Opin*
667 *Microbiol* **9**, 466-475, doi:10.1016/j.mib.2006.08.011 (2006).
668 7 Kelly, A. M., Mathema, B. & Larson, E. L. Carbapenem-resistant
669 Enterobacteriaceae in the community: a scoping review. *Int J Antimicrob Agents*
670 **50**, 127-134, doi:10.1016/j.ijantimicag.2017.03.012 (2017).
671 8 Gauthier, L. *et al.* Diversity of carbapenemase-producing *Escherichia coli* isolates,
672 France 2012-2013. *Antimicrob. Agents Chemother.*, doi:10.1128/aac.00266-18
673 (2018).
674 9 Turton, J. F. *et al.* Clonal expansion of *Escherichia coli* ST38 carrying a
675 chromosomally integrated OXA-48 carbapenemase gene. *J. Med. Microbiol.* **65**,
676 538-546, doi:10.1099/jmm.0.000248 (2016).
677 10 Zong, Z., Fenn, S., Connor, C., Feng, Y. & McNally, A. Complete genomic
678 characterization of two *Escherichia coli* lineages responsible for a cluster of
679 carbapenem-resistant infections in a Chinese hospital. *J. Antimicrob. Chemother.*,
680 doi:10.1093/jac/dky210 (2018).
681 11 Peirano, G. *et al.* Global incidence of carbapenemase-producing *Escherichia coli*
682 ST131. *Emerg Infect Dis* **20**, 1928-1931, doi:10.3201/eid2011.141388 (2014).
683 12 Cerqueira, G. C. *et al.* Multi-institute analysis of carbapenem resistance reveals
684 remarkable diversity, unexplained mechanisms, and limited clonal outbreaks.
685 *Proc Natl Acad Sci U S A* **114**, 1135-1140, doi:10.1073/pnas.1616248114 (2017).
686 13 Roer, L. *et al. Escherichia coli* Sequence Type 410 Is Causing New International
687 High-Risk Clones. *mSphere* **3**, doi:10.1128/mSphere.00337-18 (2018).
688 14 Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-
689 analysis of large phylogenies. *Bioinformatics* **30**, 1312-1313,
690 doi:10.1093/bioinformatics/btu033 (2014).
691 15 Croucher, N. J. *et al.* Rapid phylogenetic analysis of large samples of recombinant
692 bacterial whole genome sequences using Gubbins. *Nucleic Acids Res.* **43**, e15,
693 doi:10.1093/nar/gku1196 (2015).
694 16 Kocaoglu, O. & Carlson, E. E. Profiling of beta-lactam selectivity for penicillin-
695 binding proteins in *Escherichia coli* strain DC2. *Antimicrob. Agents Chemother.* **59**,
696 2785-2790, doi:10.1128/aac.04552-14 (2015).
697 17 Alm, R. A., Johnstone, M. R. & Lahiri, S. D. Characterization of *Escherichia coli* NDM
698 isolates with decreased susceptibility to aztreonam/avibactam: role of a novel

insertion in PBP3. *J. Antimicrob. Chemother.* **70**, 1420-1428, doi:10.1093/jac/dku568 (2015).

18 Zhang, Y., Kashikar, A., Brown, C. A., Denys, G. & Bush, K. Unusual *Escherichia coli* PBP 3 Insertion Sequence Identified from a Collection of Carbapenem-Resistant Enterobacteriaceae Tested In Vitro with a Combination of Ceftazidime-, Ceftaroline-, or Aztreonam-Avibactam. *Antimicrob. Agents Chemother.* **61**, doi:10.1128/aac.00389-17 (2017).

19 Luo, Y. *et al.* Characterization of Carbapenem-Resistant *Escherichia coli* Isolates Through the Whole-Genome Sequencing Analysis. *Microb. Drug Resist.* **24**, 175-180, doi:10.1089/mdr.2017.0079 (2018).

20 Touchon, M. *et al.* Organised genome dynamics in the *Escherichia coli* species results in highly diverse adaptive paths. *PLoS Genet* **5**, e1000344, doi:10.1371/journal.pgen.1000344 (2009).

21 Treangen, T. J., Ondov, B. D., Koren, S. & Phillippy, A. M. The Harvest suite for rapid core-genome alignment and visualization of thousands of intraspecific microbial genomes. *Genome Biol* **15**, 524, doi:10.1186/preaccept-2573980311437212 (2014).

22 Kohler, J. *et al.* In vitro activities of the potent, broad-spectrum carbapenem MK-0826 (L-749,345) against broad-spectrum beta-lactamase-and extended-spectrum beta-lactamase-producing *Klebsiella pneumoniae* and *Escherichia coli* clinical isolates. *Antimicrob. Agents Chemother.* **43**, 1170-1176 (1999).

23 Kumar, P., Henikoff, S. & Ng, P. C. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nature protocols* **4**, 1073-1081, doi:10.1038/nprot.2009.86 (2009).

24 Pages, J. M., James, C. E. & Winterhalter, M. The porin and the permeating antibiotic: a selective diffusion barrier in Gram-negative bacteria. *Nat Rev Microbiol* **6**, 893-903, doi:10.1038/nrmicro1994 (2008).

25 Basle, A., Rummel, G., Storici, P., Rosenbusch, J. P. & Schirmer, T. Crystal structure of osmoporin OmpC from *E. coli* at 2.0 A. *J. Mol. Biol.* **362**, 933-942, doi:10.1016/j.jmb.2006.08.002 (2006).

26 Huang, K. J. & Igo, M. M. Identification of the bases in the *ompF* regulatory region, which interact with the transcription factor OmpR. *J. Mol. Biol.* **262**, 615-628, doi:10.1006/jmbi.1996.0540 (1996).

27 Hazen, T. H. *et al.* Diversity among *bla*KPC-containing plasmids in *Escherichia coli* and other bacterial species isolated from the same patients. *Scientific reports* **8**, 10291, doi:10.1038/s41598-018-28085-7 (2018).

28 Clinical and Laboratory Standards Institute (CLSI). Performance Standards for Antimicrobial Susceptibility Testing, Twenty-Seventh Informational Supplement, M100-S25 edition. Clinical and Laboratory Standards Institute, Wayne, PA.

29 Sprouffske, K. & Wagner, A. Growthcurver: an R package for obtaining interpretable metrics from microbial growth curves. *BMC Bioinformatics* **17**, 172, doi:10.1186/s12859-016-1016-7 (2016).

30 Andersson, D. I. The biological cost of mutational antibiotic resistance: any practical conclusions? *Curr Opin Microbiol* **9**, 461-465, doi:10.1016/j.mib.2006.07.002 (2006).

31 Ortega, A. *et al.* Carbapenemase-producing *Escherichia coli* is becoming more prevalent in Spain mainly because of the polyclonal dissemination of OXA-48. *J. Antimicrob. Chemother.* **71**, 2131-2138, doi:10.1093/jac/dkw148 (2016).

747 32 Zhang, F. *et al.* Molecular epidemiology of carbapenemase-producing *Escherichia*
748 *coli* and the prevalence of ST131 subclone H30 in Shanghai, China. *Eur J Clin*
749 *Microbiol Infect Dis* **34**, 1263-1269, doi:10.1007/s10096-015-2356-3 (2015).
750 33 Potron, A. *et al.* Characterization of OXA-181, a carbapenem-hydrolyzing class D
751 beta-lactamase from *Klebsiella pneumoniae*. *Antimicrob. Agents Chemother.* **55**,
752 4896-4899, doi:10.1128/aac.00481-11 (2011).
753 34 Roschanski, N. *et al.* Retrospective Analysis of Bacterial Cultures Sampled in
754 German Chicken-Fattening Farms During the Years 2011-2012 Revealed
755 Additional VIM-1 Carbapenemase-Producing *Escherichia coli* and a Serologically
756 Rough *Salmonella enterica* Serovar Infantis. *Frontiers in microbiology* **9**, 538,
757 doi:10.3389/fmicb.2018.00538 (2018).
758 35 Wang, Y. *et al.* Comprehensive resistome analysis reveals the prevalence of NDM
759 and MCR-1 in Chinese poultry production. *Nature microbiology* **2**, 16260,
760 doi:10.1038/nmicrobiol.2016.260 (2017).
761 36 Chin, C. S. *et al.* Nonhybrid, finished microbial genome assemblies from long-read
762 SMRT sequencing data. *Nature methods* **10**, 563-569, doi:10.1038/nmeth.2474
763 (2013).
764 37 Koren, S. *et al.* Canu: scalable and accurate long-read assembly via adaptive k-mer
765 weighting and repeat separation. *Genome Res* **27**, 722-736,
766 doi:10.1101/gr.215087.116 (2017).
767 38 Bankevich, A. *et al.* SPAdes: a new genome assembly algorithm and its
768 applications to single-cell sequencing. *Journal of computational biology : a journal*
769 *of computational molecular cell biology* **19**, 455-477, doi:10.1089/cmb.2012.0021
770 (2012).
771 39 Gurevich, A., Saveliev, V., Vyahhi, N. & Tesler, G. QUAST: quality assessment tool
772 for genome assemblies. *Bioinformatics* **29**, 1072-1075,
773 doi:10.1093/bioinformatics/btt086 (2013).
774 40 Seemann, T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30**,
775 2068-2069, doi:10.1093/bioinformatics/btu153 (2014).
776 41 Zankari, E. *et al.* Identification of acquired antimicrobial resistance genes. *J.*
777 *Antimicrob. Chemother.* **67**, 2640-2644, doi:10.1093/jac/dks261 (2012).
778 42 Carattoli, A. *et al.* In silico detection and typing of plasmids using PlasmidFinder
779 and plasmid multilocus sequence typing. *Antimicrob. Agents Chemother.* **58**,
780 3895-3903, doi:10.1128/aac.02412-14 (2014).
781 43 Guy, L., Kultima, J. R. & Andersson, S. G. genoPlotR: comparative gene and genome
782 visualization in R. *Bioinformatics* **26**, 2334-2335,
783 doi:10.1093/bioinformatics/btq413 (2010).
784 44 Page, A. J. *et al.* Roary: rapid large-scale prokaryote pan genome analysis.
785 *Bioinformatics* **31**, 3691-3693, doi:10.1093/bioinformatics/btv421 (2015).
786 45 Li, W. & Godzik, A. Cd-hit: a fast program for clustering and comparing large sets
787 of protein or nucleotide sequences. *Bioinformatics* **22**, 1658-1659,
788 doi:10.1093/bioinformatics/btl158 (2006).
789 46 Katoh, K. & Standley, D. M. MAFFT Multiple Sequence Alignment Software
790 Version 7: Improvements in Performance and Usability. *Mol Biol Evol* **30**, 772-
791 780, doi:10.1093/molbev/mst010 (2013).
792 47 Talavera, G. & Castresana, J. Improvement of Phylogenies after Removing
793 Divergent and Ambiguously Aligned Blocks from Protein Sequence Alignments.
794 *Systematic biology* **56**, 564-577, doi:10.1080/10635150701472164 (2007).

48 Darriba, D., Taboada, G. L., Doallo, R. & Posada, D. ProtTest 3: fast selection of best-fit models of protein evolution. *Bioinformatics* **27**, 1164-1165, doi:10.1093/bioinformatics/btr088 (2011).

49 Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754-1760, doi:10.1093/bioinformatics/btp324 (2009).

50 McKenna, A. *et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* **20**, 1297-1303, doi:10.1101/gr.107524.110 (2010).

51 Danecek, P. *et al.* The variant call format and VCFtools. *Bioinformatics* **27**, 2156-2158, doi:10.1093/bioinformatics/btr330 (2011).

52 Cingolani, P. *et al.* A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly* **6**, 80-92, doi:10.4161/fly.19695 (2012).

53 Lennen, R. M. *et al.* Transient overexpression of DNA adenine methylase enables efficient and mobile genome engineering with reduced off-target effects. *Nucleic Acids Res.* **44**, e36, doi:10.1093/nar/gkv1090 (2016).