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[$].

[$] 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 12,584 *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 inducing reduced susceptibility to diverse β-lactams spread across the species by recombination. The selection of CP-*Ec* lineages able to disseminate is more complex than the mere acquisition of carbapenemase genes.

40 **Introduction**

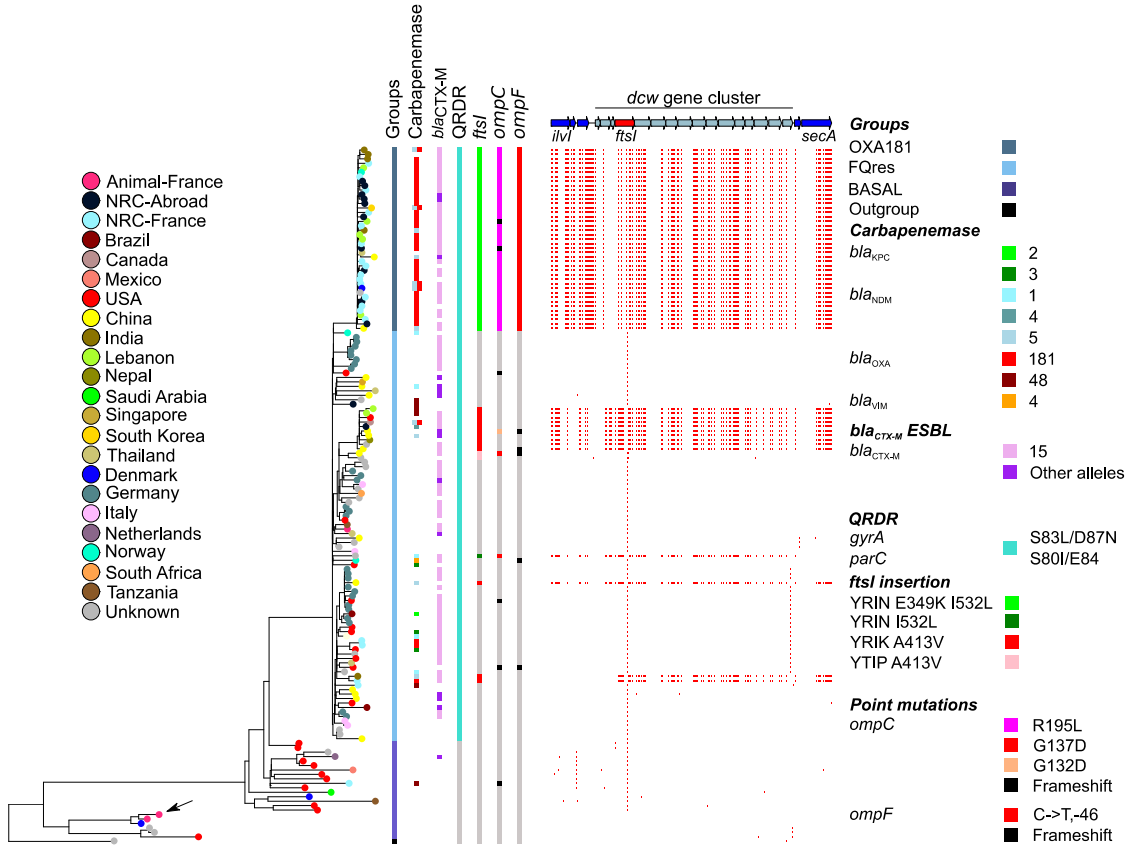
41

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

59 The epidemiology of CP-*Ec* is complex with geographical diversity in terms of
60 carbapenemase genes and of dominant lineages[4]. Most studies performed at national or
61 hospital levels point to a broad diversity of isolates as defined by multilocus sequence
62 typing (MLST), with some isolates belonging to a few dominant sequence types (STs)
63 like ST38, clonal complex (CC) 10 (ST10, ST167, ST617), ST101, ST131 and ST410 that
64 express different carbapenemases[4,8-14]. However, their prevalence varies significantly
65 worldwide. Analysis of CP-*Ec* strains isolated in 16 countries between 2008 and 2013
66 revealed that 36% belonged to the pandemic ST131, which has driven the global spread
67 of CTX-M-15 ESBL in *E. coli*[11]. Similarly, a survey of CRE strains in China showed that
68 ST131 represented 34% of the isolates and ST167, 17%[14]. But only a single ST131
69 isolate out of 140 CP-*Ec* was identified by the French National Reference Centre (Fr-
70 NRC) between 2012-2013[8]. Recently, the phylogenetic analysis of a Danish collection of
71 ST410 isolates combined to an international set of isolates revealed a globally
72 disseminated *Ec* clone carrying $bla_{\text{OXA-181}}$ on a IncX3 plasmid. This lineage was predicted

73 by a Bayesian analysis to have acquired *bla*$_{OXA-181}$ around 2003 and subsequently
74 *bla*$_{NDM-5}$ around 2014[13].
75 Despite public health implications, factors contributing to the emergence and the
76 dissemination of CP-*Ec* lineages have not been explored. Here, by using an in-depth
77 evolutionary and functional analysis of *Ec* ST410 and by extending it to the whole *E. coli*
78 species, we show that acquisition of carbapenemase genes followed different
79 evolutionary trajectories. In most STs it occurred preferentially in specific disseminated
80 lineages mutated in *ftsI* encoding penicillin-binding protein 3 (PBP3) and/or in the
81 *ompC* and *ompF* porin genes. We also show that these mutations lead to a reduced
82 susceptibility to some β-lactams including ertapenem. On the other hand, we did not
83 identify these mutations among ST131 isolates. These new data on evolution of CP-*Ec*
84 allow us to propose a model for their selection and dissemination.
85
86
87 **Results**
88
89 **Most CP-*Ec* ST410 isolates received by the French NRC belong to a single**
90 **lineage.** In order to determine the genetic bases for the dissemination of CP-*Ec* lineages,
91 we first analysed ST410 CP-*Ec* isolates, which show a high prevalence among isolates
92 collected by the Fr-NRC[8]. We sequenced the genomes of 54 CP-Ec isolates, 50 collected
93 by the Fr-NRC (including 22 from patients repatriated from 15 different countries) and
94 four from Lebanon, and three non-CP isolates of animal origin (Supplementary Table 1).
95 We reconstructed their phylogeny together with 149 *Ec* ST410 genome sequences
96 retrieved from public databases (Supplementary table 2). We filtered for redundancy in
97 this collection by removing 50 clonal isolates differing by less than seven SNPs in the
98 core genome[15] and keeping the isolate with the largest number of antibiotic resistance
99 genes (ARG). The phylogeny is in agreement with the recent analysis of CP-*Ec* ST410
100 from a Danish collection[13], with a major fluoroquinolone resistant clade (FQR-clade)
101 gathering the majority of the non-redundant (nr) isolates (133 out of 155) and of the nr-
102 isolates carrying carbapenemase genes (62 out of 63). Within the FQR-clade 77% of the
103 isolates expressed CTX-M-type ESBLs (Fig. 1). 36 out of the 40 *bla*$_{OXA-181}$ carrying
104 isolates formed a single subclade (the OXA-181 subclade) which corresponds to the
105 previously described clade B4/H24RxC[13]. The 24 CP-*Ec* isolates not belonging to the

106 OXA-181 subclade express different carbapenemases from the OXA-48, KPC, VIM and
107 NDM families.



108

109 **Fig. 1. Core genome phylogeny and genomic features of *E. coli* ST410 isolates.** ML
110 phylogeny of 155 *Ec* ST410 nr-genomes built with RAxML[16] based on the 3,419,955 bp
111 core and recombination-free alignment of 5,378 SNPs. The *Ec* ST88 isolate 789
112 (CP010315.1) was used as an outgroup. Isolates (branch tips) are colour-coded
113 according to the geographical origin as indicated in the figure key (left). Genomic
114 features are indicated as indicated in the figure key (right) from left to right: groups
115 according to the phylogeny, including the FQR clade and the OXA-181 subclade;
116 carbapenemases; CTX-M ESBL; mutations in *gyrA* and *parC* QRDR region (FQ
117 resistance); mutations in *ftsI*, *ompC,* and *ompF.* SNPs in the *dcw* cluster compared to the
118 *Ec* ST410 non-recombined strain ANSES30599 (black arrow) are indicated by small
119 vertical red bars. Upper part, genetic map of the *dcw* locus, genes are indicated by
120 arrows, *ftsI* in red. NRC stands for National Reference Centre.

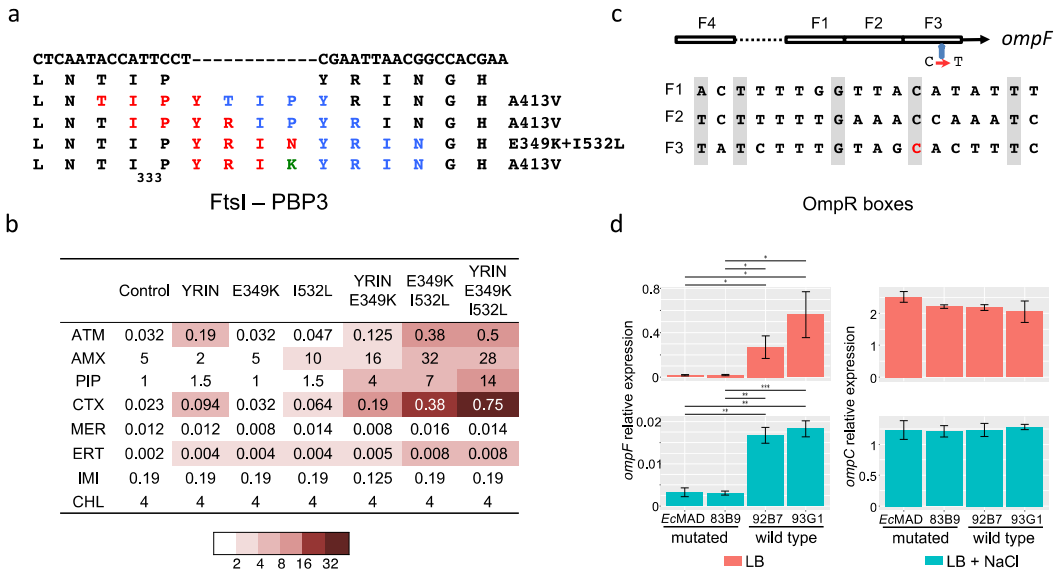
121 To accurately analyse the evolution of the OXA-181 subclade, we sequenced to
122 completion a representative isolate of this clade (*Ec*-MAD). *Ec*-MAD carries three
123 plasmids and 16 ARGs targeting seven classes of antibiotics (Supplementary Table 3).
124 Indeed, antibiotic susceptibility testing showed that it is resistant to most antibiotics
125 tested, remaining susceptible only to imipenem, meropenem, doripenem, amikacin,
126 azithromycin, chloramphenicol, tigecycline, and colistin and intermediate to mecillinam,

127 ertapenem, kanamycin and gentamicin (Supplementary Table 4). Comparison of the
128 ARG content among ST410 *Ec* isolates revealed an increase in the median number of
129 ARG between the basal isolates (n=4), the FQR-clade (n=9) and the OXA-181 subclade
130 (n=16) (Supplementary Fig. 1).

131 **Gain of specific *ftsI* alleles by recombination is a hallmark of *Ec* ST410**
132 **carbapenemase-producing strains.** Our phylogenetic analysis provided further
133 evidence of a worldwide dissemination of the OXA-181 subclade[13]. Therefore, we
134 searched for polymorphisms that, in addition to the acquisition of ARGs, have
135 contributed to the expansion of this lineage. To this end, we systematically analysed
136 mutations occurring in the branch leading to its most recent common ancestor (MRCA).
137 Besides 97 mutations in non-recombined regions, we also identified 1648 SNPs in
138 regions predicted as recombined by using Gubbins[17] (Supplementary Table 5). 92,4 %
139 occurred in a 124-kb DNA region between *yaaU* and *erpA* (Supplementary Fig. 2). In
140 contrast, this recombined region was almost identical to sequences found in four ST167
141 and eight ST617 isolates from CC 10. Strikingly, all but one of these isolates expressed a
142 carbapenemase gene. Furthermore, analysis of ST410 CP-*Ec* isolates outside the OXA-
143 181 subclade revealed four additional recombination events overlapping the 124-kb
144 recombined region identified in the OXA-181 subclade (Supplementary Fig. 2, Fig. 1).
145 These recombination events affected a subclade of ten nr-isolates of different
146 geographical origins including five CP isolates carrying different carbapenemase genes;
147 two closely related CP-*Ec* isolates, one from India ($bla_{\text{NDM-5}}$) and one from the Fr-NRC
148 ($bla_{\text{OXA-181}}$); and isolated CP-*Ec* isolates (Fig. 1). The 16.5-kb region shared by the five
149 recombined regions encompassed the *dcw* (division and cell wall) locus from *ftsI* to *secM*
150 (Fig. 1). It encodes major functions in cell wall synthesis and cell division, including *ftsI*
151 encoding PBP3, a target of diverse β-lactams[18]. Overall, 75% (47/63) of the nr CP-*Ec*
152 ST410 isolates had recombined in the *dcw* region (Fig. 1).

153 197 SNPs including 16 non-synonymous (NS) mutations differentiated the
154 common 16.5-kb region in the OXA-181 subclade from other *Ec* ST410 isolates
155 (Supplementary Table 5). Among differences, we identified insertions of four codons
156 (YRIN in two cases and YRIK in three, Fig. 2a) occurring at a same position, between
157 P333 and Y334 of FtsI. These insertions resulting from a four codons duplication (YRIN)
158 and from a subsequent mutation (YRIK) were first described in NDM-producing *E. coli*

159 isolates from different STs[19]. Additional NS SNPs were identified in the *ftsI* gene: E349K
160 and I532L in association with the YRIN insertion and A413V with the YRIK insertion.
161 The YRIK insertion in PBP3 was previously shown to confer reduced susceptibility to
162 different β-lactams including ampicillin, cefepime and aztreonam but not to
163 carbapenems[19].



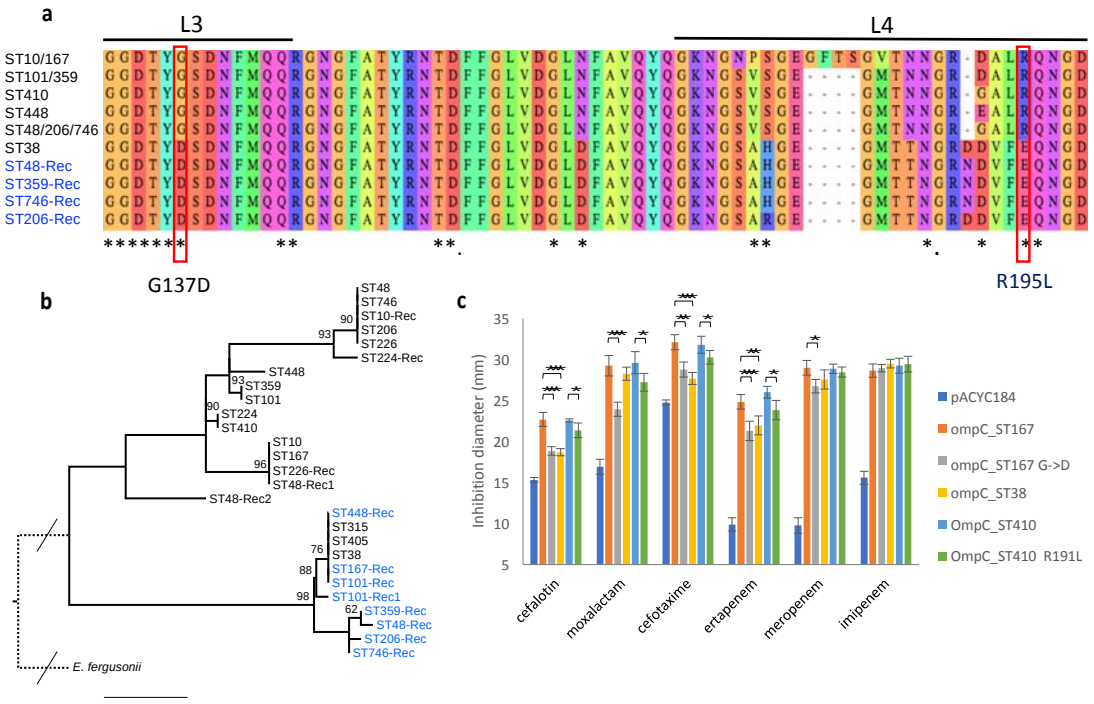
164

165 **Fig.2. Functional analyses of *ftsI* and *ompF* mutations occurring in the OXA-181 *Ec***
166 **ST410 subclade**. **a.** Mutations identified in *ftsI*. The four different insertions following
167 proline 333 resulted in the duplication of the four codons shown in red and blue. The
168 YRIK insertion derived from YRIN by a N to K AA change (in green). The first and second
169 lines represent the WT nucleotide and AA sequences respectively; on the right, AA
170 substitutions associated with each duplication. **b.** Antibiotic susceptibility testing
171 performed by Etest of MG1655 derivatives mutated in *ftsI*. Abbreviation: ATM,
172 Aztreonam; AMX, Amoxicillin; PIP, Piperacillin; CTX, Cefotaxime; MER, Meropenem; ERT,
173 Ertapenem; IMI, Imipenem; CHL, Chloramphenicol. **c.** Schematic representation of the
174 four OmpR binding sites in the *ompF* regulatory region and mutation of the conserved
175 cytosine (C=>T) in the F3 OmpR binding site in red. **d.** Expression of *ompF* and *ompC*
176 genes in two strains from the OXA-181 subclade (*Ec*MAD and 83B9, mutated) or from
177 the FQR clade (92B7 and 93G1, WT) grown in LB medium and in LB medium
178 supplemented with 0.3M NaCl. Bars represent confidence intervals; *, P < 0.05; **, P <
179 0.01; ***, P < 0.001.
180

181 **Mutations in the porin genes *ompC* and *ompF* are predicted to also have**
182 **contributed to the selection of the ST410 OXA-181 subclade.** To identify additional
183 polymorphisms that might have contributed to the dissemination of the *Ec* ST410 OXA-
184 181 subclade, we analysed the potential effect of non-synonymous mutations in the

branch leading to its MRCA by using the SIFT algorithm[20]. We identified 34 NS mutations with a predicted functional effect (nine in the recombined region) (Supplementary Table 6). Eight of these mutations affected genes from the class "transporter" including the multidrug efflux transporter components *emrD* and *emrK* and five from the class "cell envelope". These mutations might have been selected in relation to modifications in antibiotic susceptibility.

Among mutations affecting functions related to the cell envelope, one was the *ftsI* mutation I532L; another affected the porin gene *ompC* at a conserved arginine residue in the L4 loop (R195L, OmpC MG1655 numbering), one of the gateways for carbapenems (Fig. 3a)[21]. Arg_{195} is exposed, at the vestibule of the pore lumen and is conserved in OmpF[22]. Therefore, its replacement by a leucine, a non-polar AA, might affect permeation of β-lactams into the periplasm as we confirmed experimentally (see below). While we did not detect mutations in *ompF* coding sequence in the OXA-181 subclade, we identified a mutation in *ompF* regulatory region. This mutation replaces a conserved cytosine to a thymine residue in the proximal (F3) OmpR binding site. OmpR is a transcriptional activator of *ompF* and *ompC* expression and this mutation is predicted to affect *ompF* expression (Fig. 2c)[23].



202

203 **Fig 3. Mutations and recombination in the *ompC* gene. a.** Alignment of OmpC L3L4
204 region from ST in which mutations or recombination events were detected. External
205 loops L3 and L4 are indicated by lines above the sequences, and those positions
206 predicted to be exposed to the pore lumen in *E. coli* MG1655 (ST10) by asterisks[22].
207 Mutation R195L and G137D associated with the gain of carbapenemase genes are
208 highlighted by red rectangles. Numbering is according to MG1655 OmpC protein. **b.**
209 Maximum likelihood phylogenetic reconstruction of representative OmpC sequences.
210 OmpC sequences are labelled according to the ST of origin. OmpC sequences tagged with
211 "-Rec" in blue were acquired by recombination in their respective STs; independent
212 recombination events with different *ompC* alleles within a single ST are numbered.
213 Bootstrap values > 60 are indicated. **c.** Antibiotic susceptibility testing (diameters of
214 inhibition) of W3110 Δ*ompC* Δ*ompF* pOXA-232 strain complemented by different alleles
215 of the *ompC* gene cloned in the medium copy number pACYC184[24] according to the
216 figure key. The empty vector was used as control. Bars represent standard deviations; *,
217 P < 0.05; **, P < 0.01; ***, P < 0.001.

218 **Recombination at the *dcw* cluster and mutations in the porin genes *ompC***
219 **and *ompF* are frequently associated with the acquisition of a carbapenemase gene.**
220 In addition to the YRIN and YRIK insertions, two other four-AA insertions were also
221 recently reported at the same position in FtsI, YRIP and YTIP. These insertions result
222 from duplications starting two and three codons upstream the YRIN duplication
223 respectively (Fig. 3a)[25]. To determine whether the association between the acquisition
224 of a carbapenemase gene and a mutated PBP3 characterized by a four AA insertion is
225 specific to ST410 isolates or if is also observed in other *E. coli* lineages, we analysed for
226 these two features *E. coli* and *Shigella* genomes from the NCBI database. Indeed, despite
227 biases of sequenced genomes deposited in public databases, they originated from
228 different parts of the world and were analysed for different purposes and cover the
229 diversity of the species. None of the *Shigella* isolates encoded a carbapenemase gene or
230 carried an insertion in PBP3. 487 *E. coli* isolates (4.4%) encoded a carbapenemase gene
231 and 248 (2.3%) carried a four AA insertion in PBP3: 163 YRIN, 49 YRIK, 3 YTIP and 33
232 YRIP. After removing redundancy for almost identical isolates of the same origin, 80%
233 (146 out of 182) of nr-isolates mutated in *ftsI* were CP-*Ec* (Supplementary Table 7). All
234 the 123 nr-isolates showing the YRIN insertion were also mutated at position 532 (I/L)
235 and 112 at position 349 (E/K). On the other hand, all YRIK, YTIP and YRIP insertions
236 were associated with a same secondary mutation A413V, suggesting this AA change was
237 selected together with the four AA insertion either to reduce the fitness cost of the AA
238 insertion or to reduce the susceptibility to antibiotics targeting PBP3. Globally, these
239 data reveal at the species level, a strong link between these combinations of mutations
240 in PBP3 and the acquisition of a carbapenemase gene. In addition to ST410, *ftsI* was

241 mutated in the majority of the CP-*Ec* nr-isolates from ST101 (100%, N=23), ST167
242 (91%, N=49) and ST405 (81%, N=13) (Supplementary table 7). The case of the *Ec* ST167
243 isolates was particularly striking. We detected, within this single ST, after including 75
244 nr-isolates from Enterobase, 13 events of recombination leading to the replacement of
245 the endogenous PBP3 allele by an allele with the YRIN (n=11) or the YRIK (n=2)
246 insertion. Eleven recombination events affected at least one CP isolate, including a sub-
247 clade of 35 nr-isolates expressing carbapenemases from seven different types. Only five
248 out of the 54 *Ec* ST167 CP isolates did not undergo recombination at *ftsI* (Supplementary
249 Fig. 3).

250 We next reconstructed the phylogeny of the 25 STs with at least 3 CP-*Ec* isolates
251 and analysed mutations in *ompC* and *ompF* occurring during the evolution of these STs
252 (Supplementary table 7). We focused on mutations inactivating *ompF* or decreasing its
253 expression by affecting OmpR binding sites in the promoter region, as observed in the
254 ST410 OXA-181 subclade. We also looked at mutations inactivating *ompC* or predicted to
255 modify the porin permeability to β-lactams by decreasing the charge of AA located in the
256 pore lumen[21]. 117 CP-*Ec* nr-isolates (41%) out of 286 from the NCBI were mutated in
257 *ompF* compared to only 138 (8%) out of the 1659 non CP-Ec nr-isolates. This shows that
258 these *ompF* alterations are associated with the acquisition of a carbapenemase gene. In
259 89 CP-*Ec* nr-isolates (31%), OmpC was modified but in only three CP-*Ec* (1%) isolates it
260 was inactivated. In non CP-*Ec* isolates, OmpC was modified in only 44 (3%) nr-isolates
261 and inactivated or missing in 39 (2%) (supplementary table 7). Therefore, OmpC
262 modifications, but not its inactivation, are associated with the acquisition of a
263 carbapenemase gene. This might be due to the high fitness cost of OmpC loss[26]. In
264 addition to the R195L mutation in the OXA-181 ST410 subclade we identified two
265 positions in the constriction loop L3 of OmpC[22] independently mutated in different
266 isolates. The G137D replacement was identified in a ST361 lineage enriched in CP-*Ec*
267 isolates (Supplementary Fig. 4) and in four independent CP-*Ec* isolates from ST410,
268 ST448 and ST617 (Fig. 1 and Supplementary Fig. 4), and the G132D in a carbapenemase
269 resistant isolate belonging to a ST410 lineage mutated in *ftsI* and in two ST405 isolates
270 (Fig. 1 and Supplementary Fig. 5). However, the most frequent *ompC* modification
271 associated with CP-*Ec* isolates was the replacement of the original allele by alleles
272 originating from phylogroup D strains through recombination (Fig. 3b). Indeed, we

observed 16 independent recombination events, notably in the broadly distributed ST167 subcluster with a 22.7 kb recombined region from ST38 (Supplementary Fig. 6). Strikingly, OmpC proteins from phylogroup D isolates differ from other *E. coli* OmpC proteins at the two aforementioned residues G137 and R195 by negatively charged residues, D and E respectively (Fig. 3). In addition to ST38[8,9] four other STs from phylogroup D: ST354, ST405, ST457 and ST648 contained CP-*Ec* isolates (Supplementary Fig. 5). These observations point to an association between this *ompC* allele and the acquisition of a carbapenemase gene.

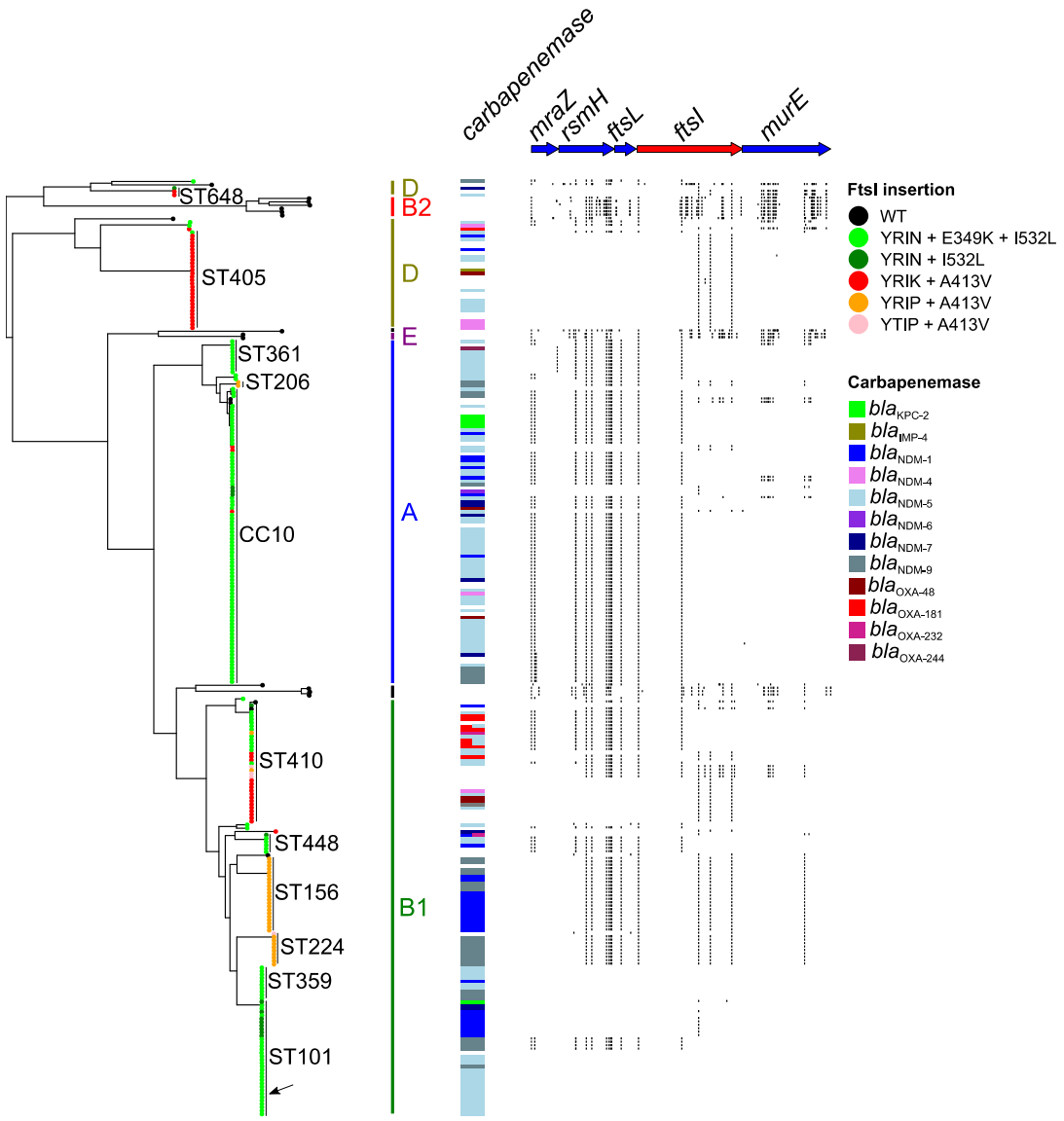


Fig. 4. Phylogenetic distribution of *E. coli* isolates mutated in *ftsI*. The genome based phylogeny of *Ec* isolates mutated in *ftsI* and of *E. coli* reference strains[27] was estimated from core non-recombinant SNP. Mutations in the *ftsI* gene are indicated at the tip of the

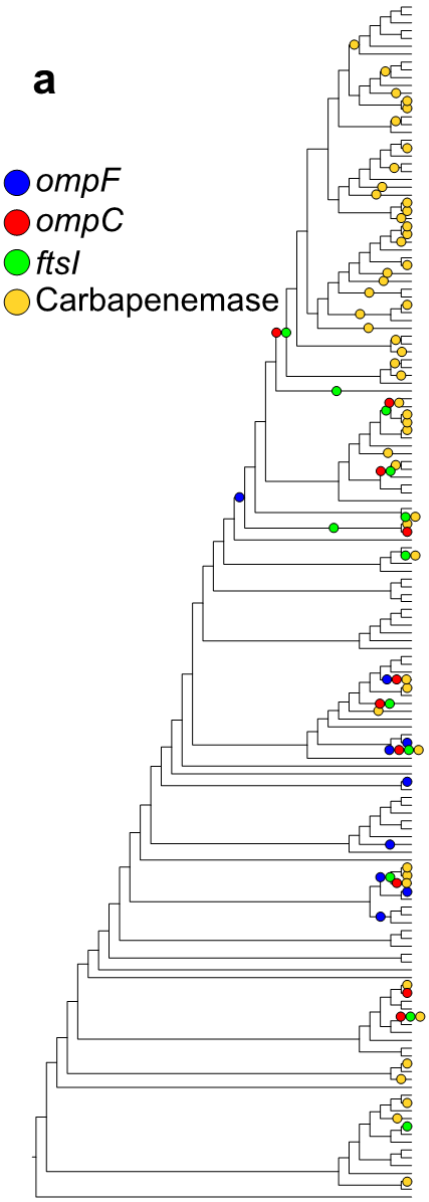
285 branch as defined in the figure keys. Major STs are indicated, as well as phylogroups, S
286 standing for *Shigella*. Carbapenemase genes are colour-coded as defined in the figure
287 keys on the right. SNPs in the *mraZ* – *murE* region compared to the *Ec* ST101 strain
288 CREC-591 (CP024821) shown by a black arrow are indicated by small vertical black
289 bars. An ST101 isolate was chosen as reference, as insertion of YRIN in *ftsI* occurred
290 among this ST. Conservation of SNP patterns between isolates showing the same
291 mutation in *ftsI* across the *E. coli* species revealed the exchange of these alleles by
292 recombination.

293 Alike for *ompC* gene, we were able to trace recombination events leading to
294 mutations in *ftsI* by analysing polymorphisms at the *dcw* locus considering genome-
295 based phylogeny (Fig. 4 and supplementary Fig. 6). The identification of mutated genes
296 with no signs of recombination (i.e. the absence of additional SNPs in *ftsI* and in the *dcw*
297 gene cluster compared to basal strains) showed that YRIN, YRIP and YTIP insertions
298 occurred first in one strain of the ST101, ST156 (Supplementary Fig. 3) and ST410
299 lineages (Fig. 1), respectively. The *ftsI* region with an YRIK insertion is almost identical
300 to the ST101 *ftsI* region with YRIN insertion and therefore likely derived from it. The
301 four *ftsI* variants were subsequently transferred to other lineages by lateral gene
302 transfer (LGT). For instance, a 29.5 kb region recombined first from ST101 to ST167 and
303 thereafter a 124-kb region from ST167 recombined into the MRCA of the ST410 OXA-
304 181 subclade (Supplementary Fig. 6). Similarly, a 65-kb region with the YRIP insertion
305 in *ftsI* from an ST156 *Ec* strain was introduced by homologous recombination in the
306 MRCA of a clade of NDM-9 expressing ST224 *Ec* isolates. The shortest recombination
307 event, detected in a ST167 CP-*Ec* strain (WCHEC16) contains only the *ftsI* gene
308 (Supplementary Fig. 3). In total, we detected 52 independent recombination events
309 involving a mutated *ftsI* allele scattered in all *E. coli* phylogroups except the B2 strains.
310 Indeed, in *Ec* ST131 from the B2 phylogroup, despite the large number of CP-*Ec* isolates
311 (n=49 nr-isolates), no isolate was mutated in *ftsI* (Supplementary Fig. 7).

312 **Acquisition of carbapenemase genes was preferentially selected in**
313 **backgrounds mutated in *ompC*, *ompF* and *ftsI*.** The observation of frequent co-
314 occurrences of mutations in the three genes and acquisition of a carbapenemase gene is
315 indicative of a genetic dependency between these events. In order to statistically test for
316 the independence of two events in the phylogeny of each ST, we applied the methods
317 (EpiCs) developed by Behdenna et al.[28]. This method takes into account the topology of
318 the tree and the node at which each event is predicted, by parsimony, to have occurred
319 (Fig. 6a). The test is based on a probabilistic framework that computes the exact

320 probability of counts of co-occurrences (two events in the same branch) and/or
321 subsequent events (one preceding the other in the tree). This statistical analysis was
322 repeated on the 20 STs containing at least four CP-*Ec* isolates after removing
323 redundancy. In each case, both models, ie: mutations occurring first or carbapenemase
324 gene being acquired first, were tested. We obtained no evidence for the model where
325 carbepenemase genes were acquired first. In contrast, in 11 STs, a significant association
326 was observed for mutations in *ftsI* and the acquisition of a carbapenemase gene, with
327 the mutation predicted to have occurred first in nine ST. Similarly, *ompC* and *ompF*
328 mutations show a significant association with carbapenemase acquisition in seven STs
329 each. For both porin genes, mutations were predicted to have occurred first in six STs. In
330 three STs, mutations in the three genes precede the acquisition of the carbapenemase
331 gene: ST167, ST101 and ST359. This result showed that carbapenemase genes within
332 these 11 ST are preferentially acquired in a genetic background with a reduced
333 susceptibility to β-lactams resulting of mutations in these three genes.

334 We did not detect such an association for ST131 (49 nr-CP-*Ec* isolates), ST10 (12
335 nr-CP-*Ec* isolates), ST648 (11 nr-CP-*Ec* isolates), ST226 (7 nr-CP-*Ec* isolates) and ST38
336 (26 nr-CP-*Ec* isolates). ST648 and ST38 belong to phylogroup D, the phylogroup that
337 served as the source for the dissemination of specific *ompC* alleles by recombination in
338 other CP-*Ec* lineages. As these alleles were present in the ancestors of the two STs and
339 did not result from mutation or recombination, they were not considered in our
340 association analysis although they probably confer a predisposition to acquire a
341 carbapenemase gene. ST10 was the most numerous ST analysed in this study with 528
342 nr-isolates and showed a low rate of CP-*Ec* of 2% (Supplementary Fig 8). Despite the
343 large number of CP-*Ec* ST131 isolates, none showed a four codons insertion in *ftsI* or an
344 AA change in OmpC predicted to decrease susceptibility. In addition, among the 29
345 ST131 isolates with an inactivated *ompF* gene, only eight carry a carbapenemase gene.
346 Furthermore CP-*Ec* were equally distributed in the four ST131 lineages A, B, C1 and C2
347 (Supplementary Fig. 7). Therefore, acquisition of carbapenemase genes in ST131
348 isolates might proceed according to a different path.



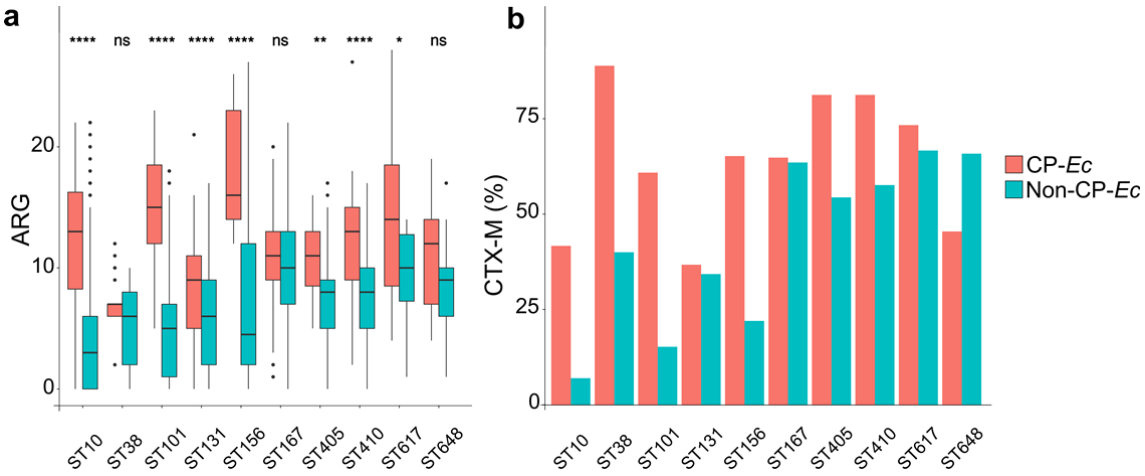
b

	ftsI → Carba.		*ompC* → Carba.		*ompF* → Carba.	
	p-value	Counts	p-value	Counts	p-value	Counts
	Genealogically ordered pairs					
ST10	1	0	1	0	1	0
ST38	1	0	1	0	1	0
ST48	1	0	1.14E-06	5	1	0
ST101	3.71E-05	5	4.20E-04	4	3.73E-06	6
ST131	1	0	1	0	0.354	1
ST156	9.45E-04	4	1	0	0.047	6
ST167	6.15E-15	37	6.50E-11	29	2.60E-10	39
ST206	1	0	1	0	1	0
ST224	1	0	1	0	1	0
ST226	1	0	0.112	1	1	0
ST354	1	0	1	0	1	0
ST359	0.042	2	0.042	2	0.042	2
ST361	0.020	2	0.020	2	1	0
ST405	3.01E-05	7	1	0	0.011	6
ST410	7.81E-06	14	0.038	6	0.038	6
ST448	0.012	2	1	0	1	0
ST457	1	0	1	0	1	0
ST617	0.006	7	1	0	1	0
ST648	0.104	1	1	0	1	0
ST746	1	0	0.128	2	0.005	2
	Inseparable pairs					
ST10	1	0	1	0	0.155	1
ST38	0.057	1	1	0	0.073	1
ST48	2.93E-04	2	0.501	1	1	0
ST101	0.079	1	0.100	1	1	0
ST131	1	0	1	0	0.056	1
ST156	0.134	2	1	0	0.201	2
ST167	0.783	4	0.373	5	0.823	2
ST206	1	0	0.179	1	1	0
ST224	0.017	1	1	0	1	0
ST226	1	0	1	0	1	0
ST354	1	0	1	0	0.028	2
ST359	0.207	1	0.248	1	0.207	1
ST361	1	0	1	0	1	0
ST405	0.007	4	0.198	1	1	0
ST410	0.266	2	0.540	2	0.578	1
ST448	0.022	2	0.001	3	1	0
ST457	0.103	1	1	0	1	0
ST617	0.360	2	0.177	2	0.501	1
ST648	0.097	1	1	0	1	0
ST746	1	0	0.065	2	0.134	1

*	**	***

Fig. 5. Test for the independence of the acquisition of carbapenemase alleles in defined genetic backgrounds. **a**. Cladogram obtained from the maximum likelihood tree estimated for the ST167 *E. coli* isolates (Supplementary Fig. 3). Four different genetic events, represented by coloured circles, have been placed by parsimony on the tree. In blue, mutations affecting the *ompF* gene (inactivation, by frameshift of premature stop codons, and mutations in the regulatory region); in red, genetic events affecting the *ompC* gene (gains of new *ompC* alleles by homologous recombination, inactivation of the gene, and non-synonymous mutations modifying the charge of AA localized in the pore lumen; in green, homologous recombination of an *ftsI* alleles with a 4 codon insertion (Fig. 2); and in yellow, acquisition of a carbapenemase gene. **b**. Testing for the independence between the acquisition of carbapenemase genes and mutations in porin genes, and/or *ftsI*. *, P < 0.05; **, P < 0.01; ***, P < 0.001.

362 **Carbapenemase genes were more frequently acquired in MDR**
363 **backgrounds.** A characteristic of the ST410 OXA-181 subclade compared to other
364 ST410 *Ec* isolates is a globally higher number of ARGs. To determine whether this
365 observation can be extended to other CP-Ec isolates, we systematically analysed for
366 their ARG content the isolates belonging to the 10 STs with more than 10 CP-Ec isolates.
367 In most STs CP-*Ec* isolates showed a significantly higher number of ARG than non-CP *Ec*
368 isolates. Only in ST38, ST167, and ST 648 the number of ARG was not significantly
369 superior in CP-*Ec* (Fig. 6). Note that both CP-*Ec* and non-CP-*Ec* ST167 isolates show a
370 high number of ARG (median=10). Similarly, we observed a higher percentage of CTX-M
371 enzymes among CP-*Ec* isolates compared to non-CP-*Ec* isolates except in ST131 and
372 ST648 (from phylogroup D).



373

374 **Fig. 6. Occurrence of ARG and *bla*$_{CTX-M}$ genes in CP-*Ec* isolates.** a. Comparison of the
375 number of ARG between CP-*Ec* and non-CP-*Ec* for the 10 STs encompassing more than
376 10 CP-Ec isolates. The horizontal lines in the boxes represent the median number of
377 ARG. The box boundaries represent the first and third quartiles of the distribution and
378 box-plot whiskers span 1.5 times the interquartile range of the distribution. Outliers are
379 denoted as black points outside whiskers. Statistical significances were tested with a
380 one-sided Wilcoxon rank-sum test. *, $P < 0.05$; **, $P < 0.01$; ****, $P < 0.0001$; ns, non-
381 significative. b Comparisons in % of the presence of *bla*$_{CTX-M}$ genes between CP-*Ec* and
382 non-CP-*Ec.*

383 **Mutations in *ftsI*, *ompC* and *ompF* associated with CP-*Ec* contribute to**
384 **decreased susceptibility to β-lactams.** Our data show that specific mutations in *ompC*,
385 *ompF* and *ftsI* have been frequently selected in lineages that thereafter acquired
386 carbapenemase genes by LGT. To further decipher the consequences of these mutations,
387 we tested experimentally their impact on the susceptibility of *E. coli* to β-lactams. We

388 first determined the contribution of the three mutations in *ftsI* (YRIN insertion, E349K
389 and I532L) identified in the OXA-181 lineage. To this aim we constructed derivatives of
390 the ST10 *s*train MG1655 with combinations of these mutations (Fig. 2b). Individually,
391 each mutation showed only a small effect on susceptibility to β-lactam targeting PBP3.
392 However, the combination of two or three mutations led to a stronger decrease in the
393 susceptibility to these antibiotics. In particular, the MG1655 PBP3 derivative with the
394 three modifications showed, in the absence of any β-lactamase, a 32, 16 and 14-fold
395 increase in the MIC to the third-generation cephalosporin cefotaxime, to the
396 monobactam aztreonam and to piperacillin respectively. This strain showed a slight
397 increase in the MIC to ertapenem (x4), which mainly targets PBP2 and to a lesser extend
398 PBB3, but no difference in the MIC to meropenem and imipenem that show low affinity
399 for PBP3[29].

400 To test the impact of mutations and recombination in *ompC* on β-lactams
401 permeability, we complemented an *E. coli* K12 strain lacking the two major porins and
402 carrying pOXA-232 (a high copy number plasmid encoding the carbapenemase bla_{OXA-}
403 $_{232}$[30]) and tested for susceptibility to β-lactams (Fig. 3c). The wild type (wt) ST167
404 (CC10) *ompC* allele, and its G137D derivative, the ST38 (phylogroup D) allele, the wt
405 ST410 allele and its R195L derivative were cloned into the medium copy number vector
406 pACYC184[24]. Complementation with the different alleles of *ompC* led to an increased
407 susceptibility to most β-lactams tested. However, we observed a differential effect of the
408 different alleles of *ompC* (Fig. 3c). In particular we observed that strains expressing the
409 R195L, the G137D and the ST38 *ompC* alleles showed reduced susceptibilities to
410 cefalotin, cefoxitin, moxalactam and ertapenem compared to the strain complemented
411 with the wt ST167 and ST410 alleles. These results confirm our prediction that the two
412 *ompC* variants and the ST38 allele associated with Cp-*Ec* isolates show lower
413 permeability towards different ß-lactams including ertapenem than their wild type
414 counterparts.

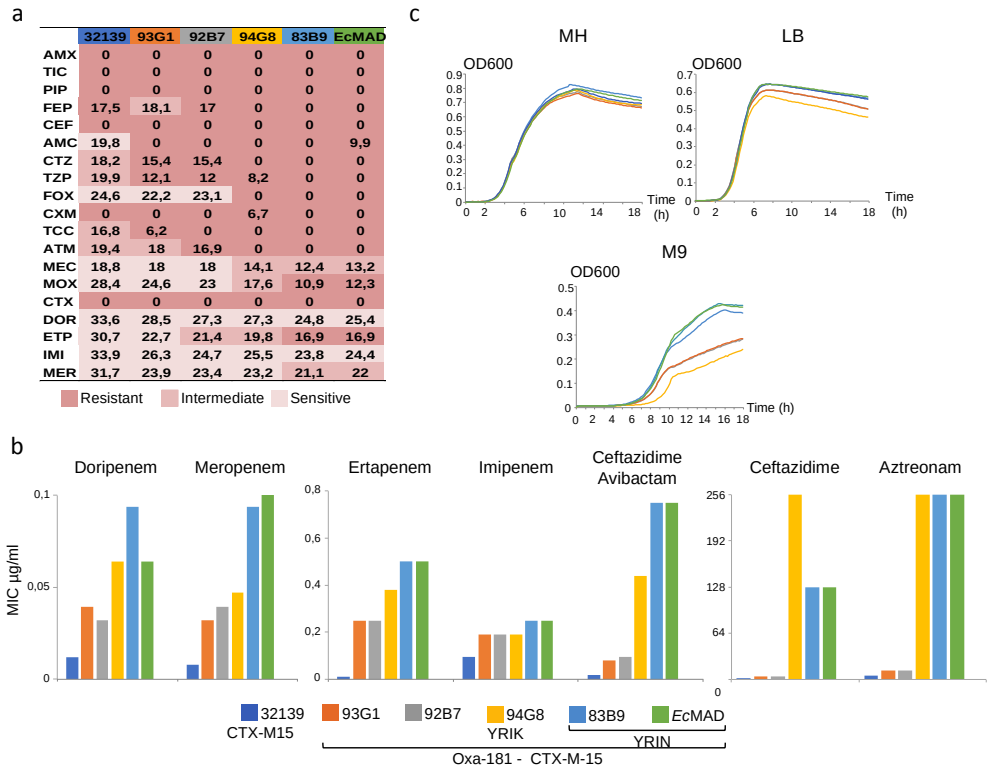
415 In *E. coli*, the OmpF porin has been shown to contribute to the permeation of β-
416 lactams into the periplasm and to susceptibility to these antibiotics[31]. To estimate the
417 impact on β-lactam susceptibility of the mutation in the *ompF* promoter region
418 identified in the ST410 OXA-181 subclade we quantified *ompF* mRNA by qRT-PCR. We
419 compared *ompF* at normal and high osmolarity (LB and LB supplemented with 1M NaCl)

420 between two isolates from the OXA-181 subclade (mutated) and two isolates from the
421 FQR clade (non-mutated). As control, we also quantified *ompC* expression. We observed
422 a 15 to 30 and 5-fold reduction in *ompF* expression in LB and LB-NaCl respectively in the
423 mutated isolates compared to wild-type, whereas *ompC* expression remained unchanged
424 (Fig. 2d). This confirmed that the regulatory mutation identified in the OXA-181
425 subclade leads to a decreased *ompF* expression in these isolates that will reduce entry of
426 β-lactams into the periplasm and antibiotic susceptibility.

427 **Isolates of the OXA-181 subclade show higher resistance without *in vitro***
428 **fitness cost as compared to other OXA-181 *Ec* isolates of the FQR clade.**
429 Independently, mutations in *ftsI* and *ompC* affected the susceptibility to β-lactams of a
430 laboratory strain. To determine the impact of mutations and ARG on β-lactam resistance
431 and fitness in ST410 isolates, we analysed the phenotype of five $bla_{\text{OXA-181}}$ $bla_{\text{CTX-M-15}}$ *Ec*
432 clinical isolates from different lineages of the ST410 FQR clade (Supplementary table 8)
433 in comparison with an isolate expressing only CTX-M-15: two isolates from the OXA-181
434 subclade mutated in *ompC*, *ompF* and *ftsI* (YRIN insertion), one isolate mutated in *ftsI*
435 (YRIK insertion) and the two other isolates without any mutation in the three genes. We
436 observed, by disk diffusion assay and Etest, a gradual decrease in the susceptibility to
437 diverse β-lactams between the four groups of isolates: CTX-M15 < CTX-M15, OXA-181 <
438 CTX-M15, OXA-181, YRIK insertion in PBP3 < OXA-181 subclade (Fig. 7a and b).
439 According to CLSI breakpoints, isolates from the OXA-181 subclade were resistant to
440 almost all β-lactams tested except doripenem and imipenem and intermediate to
441 meropenem and mecillinam. Higher resistance was partly due to differences in β-
442 lactamase gene content including $bla_{\text{OXA-181}}$ (Supplementary table 8). However,
443 mutations in *ftsI*, including the YRIK and YRIN insertions were likely responsible for the
444 decreased susceptibility of ST410 *Ec* isolates to β-lactams targeting PBP3 like
445 ceftazidime and aztreonam (Fig. 7b). The isolates from the OXA-181 subclade showed a
446 decreased susceptibility to ertapenem and meropenem likely resulting from *ompC* and
447 *ftsI* mutations for ertapenem.

448 To determine whether these mutations have an impact on fitness we compared
449 growth parameters in LB, MH and M9 media as proxy. Despite the higher resistance to
450 antibiotics, we did not detect any significant differences in the growth parameters in
451 rich medium among the six isolates tested (Fig. 7c) suggesting these mutations have no

fitness cost or their effect has been compensated by other mutations. The two isolates from the OXA-181 subclade and the non-OXA-181 isolate 32139 grew at a higher OD_{600} in minimal medium than the three other isolates (Fig. 7c). Therefore, the high level of resistance to most β-lactams and the decreased susceptibility to carbapenems of isolates of the OXA-181 subclade do not seem to be at the expense of a lower *in vitro* fitness for the isolates we have analysed.



	32139	93G1	92B7	94G8	83B9	EcMAD
AMX	0	0	0	0	0	0
TIC	0	0	0	0	0	0
PIP	0	0	0	0	0	0
FEP	17,5	18,1	17	0	0	0
CEF	0	0	0	0	0	0
AMC	19,8	0	0	0	0	9,9
CTZ	18,2	15,4	15,4	0	0	0
TZP	19,9	12,1	12	8,2	0	0
FOX	24,6	22,2	23,1	0	0	0
CXM	0	0	0	6,7	0	0
TCC	16,8	6,2	0	0	0	0
ATM	19,4	18	16,9	0	0	0
MEC	18,8	18	18	14,1	12,4	13,2
MOX	28,4	24,6	23	17,6	10,9	12,3
CTX	0	0	0	0	0	0
DOR	33,6	28,5	27,3	27,3	24,8	25,4
ETP	30,7	22,7	21,4	19,8	16,9	16,9
IMI	33,9	26,3	24,7	25,5	23,8	24,4
MER	31,7	23,9	23,4	23,2	21,1	22

Resistant Intermediate Sensitive

Fig. 7. β-lactam susceptibility profiles and fitness of *Ec* ST410 strains. **a.** β-lactam susceptibility determined by disk diffusion. Diameters are indicated in mm. Resistant, Intermediate and Sensitive according to CLSI guidelines[32] are indicted by colours as defined in the figure keys. 32139 carries $bla_{CTX\text{-}M15}$, 93G1 and 92B7 carry $bla_{CTX\text{-}M15}$ and $bla_{OXA\text{-}181}$, 94G8 carries $bla_{CTX\text{-}M15}$ and $bla_{OXA\text{-}181}$ 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 curves in rich (LB and Müller Hinton, MH) and minimal (M9) medium. Curves represent the average value of 10 experiments. Box plot representations of the area under the curve of replicates determined by using growthcurver [33] are given in Supplementary Fig. 9. 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 gene repertoires of the six strains are given in Supplementary table 8.

475

476 ## Discussion

477 During the past twenty years the increased prevalence of ESBL-producing
478 enterobacteriaceae led to an increased use of carbapenems and selection of CP-*Ec*
479 strains. The dissemination of Cp-*Ec* lineages is particularly feared. The global
480 dissemination of antibiotic resistant clones and their evolutionary trajectory result from
481 a trade-off between acquired resistance and biological cost in the absence of antibiotic[34].
482 However, how this trade-off is reached during *in vivo* evolution is largely unknown as it
483 likely depends on multiple factors, such as antibiotic usage, which differs between areas
484 of the world[35]. Also, the same clone may spread across various human and non-human
485 sectors and successively undergo highly different selective pressures. In that context
486 both the surveillance of emerging Cp-*Ec* clones in a One Health perspective and their
487 thorough genetic analysis are required to characterize their evolutionary trajectories
488 and prevent dissemination of other and possibly more virulent clones.

489 Here, we first analysed the worldwide disseminated CP-*Ec* OXA-181 ST410
490 subclade and extended this analysis to the whole *E. coli* species. We showed that
491 carbapenemase genes were preferentially acquired in lineages already mutated in three
492 genes contributing to β-lactam resistance: *ompC*, *ompF* and *ftsI*. Indeed, in 13 STs
493 representing 54% (n=234) of the non-redundant CP isolates analysed in this work, a
494 combined phylogenetic and statistical analysis revealed a significant association
495 between mutations of these genes and subsequent acquisition of a carbapenemase gene
496 (Fig. 5). Globally we found that 45%, 32% and 38% of the studied CP-*Ec* isolates were
497 mutated in *ftsI*, *ompC* and *ompF* respectively. The association was best exemplified by a
498 large clade among ST167 isolates first defined by a mutation in an OmpR box within the
499 *ompF* promoter region. Within this clade, six events of recombination led to the
500 acquisition of mutations in *ftsI* and four genetic events modified the *ompC* gene: two
501 recombination events leading to its replacement by an allele from a phylogroup D strain
502 and two homoplasic G137D mutations affecting AA of the pore lumen. Eventually,
503 multiple events of acquisition of carbapenemase genes were selected (Supplementary
504 Fig. 3). Interestingly, the encoded carbapenemases belonged to different families that
505 differ by the levels of carbapenem resistance and by the spectrum of β-lactams they
506 hydrolyze[36]. Therefore, the scenario originally detected in the ST410 OXA-181 subclade

507 expands to numerous other STs and does not seem to depend on the carbapenemase
508 family.

509 Such a situation is reminiscent of what has recently been observed in the
510 *K. pneumoniae* high-risk clones ST258, ST512 and ST11[37]. In these clones, the
511 acquisition of carbapenemase genes was frequently associated with the inactivation of
512 the porin gene *ompK35* (equivalent of *E. coli ompF*) and mutations in *ompK36*
513 (equivalent of *ompC*). In CP-*Ec*, we observed *ompC* inactivation in a few isolates,
514 probably due to a high fitness cost of this event. Instead, like in *K. pneumoniae*, the
515 selected mutations affected *ompC* permeability towards β-lactams (Fig. 3c) while
516 probably keeping the global function of the porin. This likely leads to a lower fitness cost
517 than gene inactivation, as also hypothesized in *K. pneumoniae* [37]. In contrast, both *ompF*
518 inactivation or mutations with likely a lower fitness cost such as those in its promoter
519 region were observed in CP-*Ec*. This indicates that, under *in vivo* growth conditions, the
520 OmpF porin might be more easily dispensable than OmpC. On the other hand, PBP3 is
521 essential for cell division and selection of *ftsI* mutants is expected to be highly
522 evolutionary constrained. This might explain the extremely rare *ftsI* mutations
523 previously reported in clinical lineages of *E. coli*. This contrasts with the high frequency
524 of *ftsI* mutations observed among Cp-*Ec* isolates.

525 During the evolution of CP-Ec lineages, recombination events involving *ftsI* and
526 *ompC* were pervasive throughout the *E. coli* species except among phylogroup B2. We
527 identified four combinations of mutations in *ftsI* frequently associated with CP-*Ec*
528 isolates (Fig. 2a). To our knowledge these modifications are the only mutations
529 contributing to β-lactams reported in natural *E. coli* isolates[19,25]. Nevertheless, the
530 A413V mutation found in association with three types of insertions, has recently been
531 selected during serial passages of *E. coli* in the presence of aztreonam and shown to lead
532 to a four-fold decreased susceptibility to this antibiotic [38]. This result is in agreement
533 with our prediction that this mutation might contribute to resistance to β-lactams of
534 isolates carrying YRIK, YTIP or YRIP insertions as we have shown for the two SNPs
535 associated with the YRIN duplication (Fig. 2b). The chance of mutation combinations
536 leading to a significant decrease in susceptibility is likely very low, but the selective
537 advantage strong. In agreement with this hypothesis, the phylogenetic reconstruction of
538 the recombination events showed that these combinations arose likely only once and

539 disseminated widely across the *E. coli* species by LGT, as we identified 48 events of
540 recombination (Fig. 4). Most of these recombination events were associated with at least
541 one CP-*Ec* isolates (n=46) and in 24 cases it corresponded to disseminated lineages (i.e.
542 with more than three isolates from different geographical origins). We observed a
543 similar situation for the porin OmpC, with 20 events of recombination events of *ompC*
544 alleles originating from phylogroup D isolates. The chromosomal region next to the
545 *ompC* gene has been shown as a hotspot of recombination [39]. However, we observed that
546 the acquisition of this specific allele was associated with acquisition of carbapenemase
547 genes in most cases (n=15). Recombination has been shown to play a major role in β-
548 lactam resistance in pneumococcus[40] or in *Neisseria spp*[41]. Altogether, our data show for
549 the first time that in addition to the LGT of β-lactamase genes located on MGE,
550 recombination has a significant contribution in β-lactam resistance in *E. coli*, including
551 carbapenems.

552 Further comparisons of β-lactam susceptibility and fitness among ST410 isolates
553 carrying the same $bla_{\text{OXA-181}}$-bearing plasmid and $bla_{\text{CTX-M-15}}$ gene with different patterns
554 of mutations in *ompC*, *ompF* and *ftsI* showed that in these clinical isolates and
555 particularly in the broadly disseminated OXA-181 subclade, these mutations were not
556 associated with a fitness cost (Fig. 7c). We also observed that increased resistance to β-
557 lactams could be attributable to mutations in the three genes, in agreement with our
558 experimental study of these mutations individually (Fig. 2 and 3). In particular, for the
559 OXA-181 subclade, we observed an additional decrease in the susceptibility to
560 ertapenem. Interestingly, ertapenem shows a higher biliary excretion than other
561 carbapenems and was found to have a stronger impact on the intestinal microflora[42]. *ftsI*
562 mutations were also found to be selected during *in vitro* evolution in the presence of
563 ertapenem but not meropenem[43]. However, other β-lactams like aztreonam might also
564 have contributed to the selection of these combinations of mutations[38].

565 A characteristic of the ST410 OXA-181 subclade is a higher number of ARG than
566 in other ST410 *Ec* isolates (Supplementary Fig. 1), suggesting that antibiotic pressure
567 was a major contributor to the evolution of this clone. In particular, the phylogenetic
568 analysis reveals that mutations in *ompC*, *ompF* and *ftsI* occurred in a context already
569 resistant to fluoroquinolone and expressing CTX-M-15 ESBL (Fig. 1). A systematic
570 analysis of the number of ARG across the *E. coli* species showed that a higher number of

571 ARG in CP-*Ec* compared to non-CP-Ec isolates was a common feature in all STs with
572 more than 10 CP-*Ec* isolates. A single exception was ST167 where a similarly high
573 number of ARG was observed in both CP and non-CP-*Ec* isolates analysed (Fig. 6).
574 Therefore, for these STs, acquisition of carbapenemase genes occurred more frequently
575 in an MDR background. We also observed a frequent co-occurrence of CTX-M family
576 ESBL and carbapenemase genes in most ST, but not in the ST131 lineage. In all, these
577 data suggest a long-term and step by step evolution of those lineages with episodic
578 periods of selection and dissemination. As a first step, specific mutations in *ompF, ompC*
579 or *ftsI* would have been fixed in isolates already expressing different β-lactamases
580 including ESBL enzymes, leading to low levels of resistance to carbapenems[44]. This
581 scenario is compatible with our phylogenetic analysis and the high frequency of $bla_{\text{CTX-M}}$
582 genes in isolates mutated in *ftsI* (78% of nr-isolates) and with our experimental data on
583 the effect of mutations in porin genes and in *ftsI* on β-lactam susceptibility (Fig. 2 and 3).
584 In a second step, under antibiotic pressure, the combination of these mutations and β-
585 lactamase expression might have favored the efficient conjugative transfer of plasmids
586 carrying carbapenemase genes from other CP-bacterial species by increasing the
587 proportion of donor and receptor bacteria[45]. This might have occurred in the context of
588 low levels of carbapenems or other β-lactams, such as found in the gut during parenteral
589 administration of antibiotics with biliary excretion. This model could also explain the
590 high prevalence of ST38 isolates observed both in England and in France[8,9], as 24 out of
591 27 (89%) ST38 CP-*Ec* isolates express a CTX-M class enzymes. The specific ST38 *ompC*
592 allele with reduced permeability to different β-lactams including ertapenem together
593 with CTX-M ESBL would have favored the fixation of carbapenemase genes.

594 Although CP-*Ec* were frequent among the ST131 isolates studied here, with 49
595 nr-isolates (66 in total), we did not observe any case of four AA insertions in *ftsI* or of
596 mutations affecting AA in the pore lumen of OmpC among the 402 nr-genome sequences
597 we have analysed. Furthermore, CP-*Ec* isolates were broadly distributed among the
598 different ST131 lineages and were not associated with CTX-M type ESBL since only 37%
599 expressed also β-lactamase of this class (Fig. 6b). The selection for ST131 CP-*Ec* isolates
600 might therefore follow a different path compared to others CP-*Ec*, which might be
601 related to their higher and human-specific pathogenicity[46]. ST131 CP-*Ec* isolates might
602 arise sporadically in patients following conjugation of carbapenemase gene carrying

plasmids from another CPE and subsequent selection by β-lactams, including treatments with carbapenems. These transconjugants would be more frequently detected in clinics due to their high pathogenicity. On the other hand, possibly due to their lower fitness they would not disseminate globally. Indeed, despite the high prevalence of ST131 CP-*Ec* reported in different studies[11,47,48], there was no indication of global dissemination of a specific lineage among the ST131 CP-*Ec* genome sequences we have analysed (Supplementary Fig. 7).

Material and methods

Bacterial isolates, growth conditions and antibiotic susceptibility testing. Features of the clinical *E. coli* isolates analysed in this work are listed in Supplementary table 1. 50 *Ec* ST410 isolates came from the strain collection of the Fr-NRC for Antibiotic Resistance. Four *Ec* ST410 clinical isolates came from the Microbiological collection of the Public Health Faculty of the Lebanese University (Tripoli, Lebanon) and three ST410 isolates of animal origin from the ANSES strain collection. Test for OmpC permeability to β-lactam was performed in strain W3110 deleted for *ompC* and *ompF* genes[49]. Antibiotic susceptibility was performed by the disk diffusion method following the Clinical & Laboratory Standards Institute (CLSI) guidelines[32] or by Etest (Biomérieux) following the manufacturer's recommendations. For W3110 Δ*ompC* Δ*ompF* pOXA-232 strains carrying pACYC184 derivatives, disk diffusion assays were performed on Mueller Hinton agar plates supplemented with 2mg/l chloramphenicol. Fitness was determined by growth curve analysis with an automatic spectrophotometer Tecan Infinite M200 during 24 hours in LB, Mueller Hinton or M9 media supplemented with 0.4% glucose. Growth metrics were estimated with the R package "growthcurver" [33]. The area under the curve (AUC) which include contributions of the most important growth parameters (log phase, growth rate, and the carrying capacity) was used as a growth metric.

Genome sequencing, and genome sequences retrieved from sequence libraries. *Ec* genomes were sequenced by using the Illumina HiSeq2500 platform, with 100 nucleotides (nt) single reads for the four isolates from Lebanon and 100 nt paired-end reads for the other isolates. Libraries were constructed by using the Nextera XT kit

636 (Illumina) following the manufacturer's instructions. The OXA-181-producing *Ec*-MAD
637 ST410 isolate was selected as a reference strain and sequenced to completion by using
638 the long read PacBio technology. 10 947 *E. coli* and 1 451 *Shigella* genome sequences
639 deposited in the NCBI database (19th June, 2018) were retrieved for a global analysis of
640 the specificity of CP-*Ec* (Supplementary table 2). 96 additional *Ec* ST167 isolates were
641 retrieved from Enterobase (https://enterobase.warwick.ac.uk/). Raw reads from 62 *Ec*
642 ST410 and 21 *Ec* ST38 isolates identified in Enterobase (march 2017) were retrieved
643 from the NCBI database. (Supplementary table 2). Redundancy in the genome collection
644 was removed by filtering for isolates from the same study, diverging by less than 7 SNPs.
645 We kept one randomly selected isolate. In case of differences in resistome, assuming
646 that ARG loss was more likely than ARG gain, we kept an isolate with the largest number
647 of ARG.
648 **Sequence assembly, genome annotation and mutation identification.** The PacBio
649 reads were assembled with the RS_HGAP_Assembly.3 protocol from the SMRT analysis
650 toolkit v2.3[50], and with Canu[51]. The consensus sequence was polished with Quiver[50] and
651 by mapping Illumina reads. Illumina sequenced isolates were assembled with SPAdes[52],
652 and the quality of the assemblies was assessed with Quast[53]. Contigs shorter than 500
653 bp were filtered out. All assemblies were annotated with Prokka[54]. The presence of
654 antibiotic resistance genes, and plasmid replicons was assessed with ResFinder[55] and
655 PlasmidFinder [56] respectively. The software was run in local from the script and
656 databases downloaded from the repositories of the Centre for Genomic Epidemiology
657 (https://bitbucket.org/genomicepidemiology/). Graphs of genomic regions of interest
658 were drawn with genoplotR[57]. All downloaded genomes were reannotated with
659 prokka[54]. For each ST type analysed (Warwick scheme) the pangenome was
660 characterized with Roary[58], and the AA sequences of OmpC, OmpF, GyrA, ParC, and FtsI
661 were identified from the orthologous table generated by Roary. To determine the
662 presence of different *ompC* alleles within the different STs, OmpC AA sequences were
663 clustered with cd-hit[59], with a sequence identity threshold of 0.95. AA sequences for
664 GyrA and ParC were aligned with the mafft L-INS-i approach[60], and the amino acids at
665 the QRDR positions for both genes were filtered with customized Perl scripts.
666
667 **Phylogenetic reconstruction.** A core genome alignment was obtained for all *Ec* ST410
668 isolates, and strain 789 that belongs to the ST88 as an outgroup with parsnp[61]. Non-

669 recombinant SNPs were extracted from the core genome alignment with Gubbins[17], and
670 the Maximum likelihood (ML) phylogeny was obtained with RAxML[16], using the General
671 Time-Reversible (GTR) substitution model with a gamma-distributed rate. The same
672 procedure was followed for the phylogenetic reconstruction of the different *E. coli*
673 lineages. Maximum likelihood phylogeny of OmpC protein sequences was inferred with
674 RAxML[16]. OmpC protein sequences were aligned with the mafft L-INS-i approach[60].
675 Gblocks[62] was used to "clean" the alignment and the best fit model (WAG, with a gamma
676 distribution) was estimated with protest 3[63]. The visual display of phylogenetic trees
677 was done with FigTree (http://tree.bio.ed.ac.uk/software/figtree/), and annotated trees
678 with the script plotTree.R (https://github.com/katholt/plotTree). The final version of
679 the tree was edited with Inkscape (https://inkscape.org/es/).
680
681 **Mapping, variant calling, and identification of SNPs of interest**. Sequence reads for
682 *Ec* ST410 isolates were mapped against the complete reference genome *Ec*-MAD with
683 BWA[64]. Variant calling was performed with the Genome Analysis Toolkit v 3.6.0[65]. The
684 criteria for variants were: occurrence of the alternative base in more than 90% of the
685 reads covering the position, a depth coverage of at least 10 (DP > 10), a quality by depth
686 (QD) > 2, a fisher strand bias (FS) < 60, a mapping quality (MQ) > 40, a mapping quality
687 rank sum test (MQRankSum) > -12.5, and a read position rank sum test
688 (ReadPosRankSum) > -8. Variants associated with different clades of the *Ec* ST410
689 phylogeny were extracted with vcftools[66], and finally they were annotated with snpEff[67].
690 The effect of the non-synonymous mutations was assessed with the Sorting Intolerant
691 From Tolerant (SIFT) algorithm[20]. The algorithm searches for protein homologs in the
692 refseq database using mutated proteins as query and assigns a score to each position.
693 This score is weighted by the properties of the AA changed. If this score is below a
694 threshold (0.05) the change is predicted to be functional.
695
696 **Testing the independence between mutations in *ftsI*, *ompF*, and *ompC*, and the**
697 **acquisition of carbapenemase genes.** To assess the association between the different
698 genetic events we used the method developed by Behdenna et al.[28] implemented in the
699 software EpiCs. The described events were mapped on the tree by parsimony, and the
700 probability distribution of the number of paired events occurring in the tree was
701 computed under the null model of independence. Two types of paired events are

702 described in the method: inseparable pairs, when both events occur in the same branch,
703 and genealogically ordered pairs, when the second event is found in a node more recent
704 than the first one. We have considered the following genetic events: (i) "*ompC*
705 mutations" encompassing acquisition of *ompC* alleles from phylogroup D strains by
706 recombination, *ompC* mutations changing the charge of AA in the pore lumen, and *ompC*
707 inactivation; (ii) "*ompF* mutations" including *ompF* gene inactivation and point
708 mutations in the OmpR binding sites of its promoter; (iii) "*ftsI* mutations" including the
709 four different four-codon duplications (YRIN, YRIK, TYPI, and YTIP) in *ftsI* and (iv) the
710 acquisition of a carbapenemase gene. We have focused our analysis on the
711 independence between carbapenemase gene acquisition and mutation in each of the
712 three genes, ompC, *ompF* and *ftsI*.
713
714 **Complementation of the W3110 Δ*ompC* Δ*ompF* strain.** *ompC* alleles and their
715 regulatory regions were cloned into the medium copy number vector pACYC184[24]
716 following amplification by primers *ompC_Xba*_F and *ompC_Hind*_R (Supplementary Table
717 9), digestion by *Xba*I and *Hind*III restriction enzymes and ligation into the vector
718 digested by the same enzymes. Ligation was transformed into commercial *E. coli* TOP10
719 competent cells (Invitrogen). The absence of mutation was checked by Sanger
720 sequencing. Plasmid containing the *ompC* gene from MG1655 or the G137D mutated
721 allele, the ST410 wt allele and the R195L mutated allele from *Ec*MAD and the ST38 allele
722 as well as the empty vector were introduced into competent W3110 Δ*ompC* Δ*ompF*
723 pOXA-232. Competent cells were prepared by the $CaCl_2$ method[68]. Plasmid pOXA-232[30]
724 was prepared from an ST231 *Klebsiella pneumoniae* isolate from the Bicêtre hospital
725 collection carrying this plasmid. Plasmid content in transformants was verified by
726 plasmid DNA extraction (Qiagen) and Sanger sequencing.
727
728 **Construction of *ftsI* mutant strains.** The three mutations identified in the *ftsI* gene of
729 the MAD strain were reconstructed in a MG1655 genetic background to analyse their
730 effects on antibiotic resistance. To this aim we introduced the 12 nt insertion (YRIN
731 form) and the two non-synonymous SNPs (E349K and I532L) into the *Ec* strain MGF
732 (MG1655strepR_F'tet_ΔtraD::Apra) by TM-MAGE[69]. Briefly, overnight culture of strain
733 MGF transformed by pMA7SacB was used to inoculate 5 ml LB medium supplemented
734 with tetracycline (7.5 µg/l) and carbenicillin (100 µg/l) (LB-TC) and grown at 37°C until

735 OD_{600} reached 0.6-0.7. The recombinase and Dam methylase were induced by addition
736 of L-arabinose (final concentration of 0.2% w/v), and further incubation for 10 min.
737 Cultures were then chilled for 15 min on ice and centrifuged at 7300g at 4°C. Two
738 successive washes by 50 and 10 ml of cold water were performed and the final pellet
739 was resuspended in 200 µl water. 100 µl of cells were used for electroporation with 2 µl
740 oligonucleotides Mut1ftsI or Mut2ftsI (Supplementary table 9) alone or in combination
741 at 20 µM each. The Mut1f*tsI* oligonucleotide carries both the 12 nt insertion and the
742 E349K mutation, while the Mut2*ftsI* oligonucleotide has the I532L mutation. The content
743 of the electroporation cuvette was used to inoculate 5 ml of LB-TC and submitted to
744 three additional cycles of growth-induction-preparation of electrocompetent cells and
745 electroporation. Following the last electroporation step, cells were resuspended in1 ml
746 LB and plated onto LB-TC agar plates. Mutations in isolated colonies were tested by PCR
747 using primers complementary to mutant or wild-type alleles (Supplementary table 9).
748 Mutated colonies were grown on plates containing 10 g/l tryptone, 5 g/l yeast extract,
749 15 g/l agar and 5% w/v sucrose for plasmid curing. Mutant strains were sequenced by
750 using Illumina MiSeq platform, with 150 nt paired-end reads and Nextera XT kit
751 (Illumina) for library preparation. Reads were mapped onto the MG1655 genome
752 (Genbank NC_000913.3) to confirm that mutations in *ftsI* gene have been correctly
753 introduced and to check that the rare off target mutations are not predicted to interfere
754 with the β-lactam susceptibility phenotype (Supplementary table 10).
755
756 **RNA extraction and quantitative RT-PCR.** Bacteria were grown in LB medium until
757 OD_{600} reached 0.30-0.33, an osmotic shock was applied by supplementing 10 ml of
758 culture with 0.3M, final concentration of Sodium Chloride (NaCl) or with the same
759 volume of water as control and further incubation for 20 minutes. Bacterial pellets were
760 collected and stored at -80ºC. Total RNA was extracted with the Total RNA purification
761 kit Norgen Biotek. cDNAs were synthetized from 500 ng of RNA with the Superscript II
762 reverse transcriptase (Invitrogen, Life Technologies). Primer pairs were designed for
763 the *ompC* and *ompF* genes, targeting divergent regions from these two genes, and for the
764 reference gene *recA* (Supplementary Table 9). The SYBR Green PCR kit (Applied
765 Biosystems, Life Technologies) was used to perform quantitative PCR, and the relative
766 expression of porin genes was measured by a standard curve method where the
767 regression analysis was performed from serial dilutions of a mixture of control cDNAs.

The expression value of each gene was normalized against the expression of the housekeeping gene *recA*. Each point was measured in triplicate, and three independent cultures were used for each strain in each condition.

Statistical analysis. The statistical significance of the differences in expression in RT-PCR experiments was assessed by using a two-tailed t-test. The statistical significance of the differences in the number of ARG between bacterial groups in different STs was assessed using the Wilcoxon rank sum-test implemented in R (v3.4.4). One-sided test was used for the comparison of the number of ARG, and two-sided test was used for the comparison on the area under the curve between six isolates from the ST410 FQR clade.

Data availability. Illumina reads from the 57 newly sequenced isolates and the complete genome assembly of strain Ec-MAD have been deposited in the EMBL nucleotide sequence database (http://www.ebi.ac.uk/ena) under study accession number PRJEB27293 and PRJEB27274 respectively. The accession numbers for individual isolates are listed in Supplementary Table 1. Code of software scripts could be obtained upon request.

Acknowledgements

This work was supported by grants from the French National Research Agency (ANR LabEx IBEID and ANR-10-LABX-33), from the Joint Program Initiative on Antimicrobial Resistance (ANR-14-JAMR-0002) and from One-Health EJP ARDIG. The authors thank Alexandre Almeida, Eduardo Rocha and Adriana Chiarelli for their discussion and critical reading of the manuscript, Guillaume Achaz for advices for the EpiCs analyses, Sylvie Letoffe, and Laurence Ma for their help in performing the TM-MAGE and the Illumina sequencing respectively. They also thank Jean-Michel Pages for the gift of strain W3110 Δ*ompC*, Δ*ompF*.

Authors contribution

RPN, IRC and NC performed experiments; LG, JT, JYM, MH, RB and TN collected and provided the samples for the study; RPN, retrieved genome sequences from public databases, wrote scripts and performed all the bioinformatics analyses; RPN, IRC, NC

and PG analysed the data; TN provided expertise on CP-*Ec*; RPN and PG designed the study; PG was responsible for management of the project; PG and IRC built the evolutionary model; RPN, IRC, TN and PG wrote the manuscript. All authors approved the manuscript prior to submission.

Conflict of interest

The authors declare no competing financial interests.

808 **References**

809

810 1 O'Neill, J. Tackling Drug-Resistant Infections Globally: Final Report and
811 Recommendations: Review on Antimicrobial Resistance. (2016).
812 2 Papp-Wallace, K. M., Endimiani, A., Taracila, M. A. & Bonomo, R. A. Carbapenems:
813 past, present, and future. *Antimicrob. Agents Chemother.* **55**, 4943-4960,
814 doi:10.1128/aac.00296-11 (2011).
815 3 Nordmann, P., Dortet, L. & Poirel, L. Carbapenem resistance in
816 Enterobacteriaceae: here is the storm! *Trends in molecular medicine* **18**, 263-272,
817 doi:10.1016/j.molmed.2012.03.003 (2012).
818 4 Logan, L. K. & Weinstein, R. A. The Epidemiology of Carbapenem-Resistant
819 Enterobacteriaceae: The Impact and Evolution of a Global Menace. *J. Infect. Dis.*
820 **215**, S28-s36, doi:10.1093/infdis/jiw282 (2017).
821 5 Bush, K. & Jacoby, G. A. Updated functional classification of beta-lactamases.
822 *Antimicrob. Agents Chemother.* **54**, 969-976, doi:10.1128/aac.01009-09 (2010).
823 6 Canton, R. & Coque, T. M. The CTX-M beta-lactamase pandemic. *Curr Opin*
824 *Microbiol* **9**, 466-475, doi:10.1016/j.mib.2006.08.011 (2006).
825 7 Kelly, A. M., Mathema, B. & Larson, E. L. Carbapenem-resistant
826 Enterobacteriaceae in the community: a scoping review. *Int J Antimicrob Agents*
827 **50**, 127-134, doi:10.1016/j.ijantimicag.2017.03.012 (2017).
828 8 Gauthier, L. *et al.* Diversity of carbapenemase-producing *Escherichia coli* isolates,
829 France 2012-2013. *Antimicrob. Agents Chemother.*, doi:10.1128/aac.00266-18
830 (2018).
831 9 Turton, J. F. *et al.* Clonal expansion of *Escherichia coli* ST38 carrying a
832 chromosomally integrated OXA-48 carbapenemase gene. *J. Med. Microbiol.* **65**,
833 538-546, doi:10.1099/jmm.0.000248 (2016).
834 10 Zong, Z., Fenn, S., Connor, C., Feng, Y. & McNally, A. Complete genomic
835 characterization of two *Escherichia coli* lineages responsible for a cluster of
836 carbapenem-resistant infections in a Chinese hospital. *J. Antimicrob. Chemother.*,
837 doi:10.1093/jac/dky210 (2018).
838 11 Peirano, G. *et al.* Global incidence of carbapenemase-producing *Escherichia coli*
839 ST131. *Emerg Infect Dis* **20**, 1928-1931, doi:10.3201/eid2011.141388 (2014).
840 12 Cerqueira, G. C. *et al.* Multi-institute analysis of carbapenem resistance reveals
841 remarkable diversity, unexplained mechanisms, and limited clonal outbreaks.
842 *Proc Natl Acad Sci U S A* **114**, 1135-1140, doi:10.1073/pnas.1616248114 (2017).
843 13 Roer, L. *et al. Escherichia coli* Sequence Type 410 Is Causing New International
844 High-Risk Clones. *mSphere* **3**, doi:10.1128/mSphere.00337-18 (2018).
845 14 Zhang, R. *et al.* Nationwide Surveillance of Clinical Carbapenem-resistant
846 Enterobacteriaceae (CRE) Strains in China. *EBioMedicine* **19**, 98-106,
847 doi:10.1016/j.ebiom.2017.04.032 (2017).
848 15 Moura, A. *et al.* Whole genome-based population biology and epidemiological
849 surveillance of Listeria monocytogenes. *Nature microbiology* **2**, 16185,
850 doi:10.1038/nmicrobiol.2016.185 (2016).
851 16 Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-
852 analysis of large phylogenies. *Bioinformatics* **30**, 1312-1313,
853 doi:10.1093/bioinformatics/btu033 (2014).

854 17 Croucher, N. J. *et al.* Rapid phylogenetic analysis of large samples of recombinant
855 bacterial whole genome sequences using Gubbins. *Nucleic Acids Res.* **43**, e15,
856 doi:10.1093/nar/gku1196 (2015).
857 18 Kocaoglu, O. & Carlson, E. E. Profiling of beta-lactam selectivity for penicillin-
858 binding proteins in *Escherichia coli* strain DC2. *Antimicrob. Agents Chemother.* **59**,
859 2785-2790, doi:10.1128/aac.04552-14 (2015).
860 19 Alm, R. A., Johnstone, M. R. & Lahiri, S. D. Characterization of *Escherichia coli* NDM
861 isolates with decreased susceptibility to aztreonam/avibactam: role of a novel
862 insertion in PBP3. *J. Antimicrob. Chemother.* **70**, 1420-1428,
863 doi:10.1093/jac/dku568 (2015).
864 20 Kumar, P., Henikoff, S. & Ng, P. C. Predicting the effects of coding non-
865 synonymous variants on protein function using the SIFT algorithm. *Nature*
866 *protocols* **4**, 1073-1081, doi:10.1038/nprot.2009.86 (2009).
867 21 Pages, J. M., James, C. E. & Winterhalter, M. The porin and the permeating
868 antibiotic: a selective diffusion barrier in Gram-negative bacteria. *Nat Rev*
869 *Microbiol* **6**, 893-903, doi:10.1038/nrmicro1994 (2008).
870 22 Basle, A., Rummel, G., Storici, P., Rosenbusch, J. P. & Schirmer, T. Crystal structure
871 of osmoporin OmpC from *E. coli* at 2.0 A. *J. Mol. Biol.* **362**, 933-942,
872 doi:10.1016/j.jmb.2006.08.002 (2006).
873 23 Huang, K. J. & Igo, M. M. Identification of the bases in the *ompF* regulatory region,
874 which interact with the transcription factor OmpR. *J. Mol. Biol.* **262**, 615-628,
875 doi:10.1006/jmbi.1996.0540 (1996).
876 24 Chang, A. C. & Cohen, S. N. Construction and characterization of amplifiable
877 multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. *J.*
878 *Bacteriol.* **134**, 1141-1156 (1978).
879 25 Zhang, Y., Kashikar, A., Brown, C. A., Denys, G. & Bush, K. Unusual *Escherichia coli*
880 PBP 3 Insertion Sequence Identified from a Collection of Carbapenem-Resistant
881 Enterobacteriaceae Tested In Vitro with a Combination of Ceftazidime-,
882 Ceftaroline-, or Aztreonam-Avibactam. *Antimicrob. Agents Chemother.* **61**,
883 doi:10.1128/aac.00389-17 (2017).
884 26 Ferenci, T. & Phan, K. How Porin Heterogeneity and Trade-Offs Affect the
885 Antibiotic Susceptibility of Gram-Negative Bacteria. *Genes* **6**, 1113-1124,
886 doi:10.3390/genes6041113 (2015).
887 27 Touchon, M. *et al.* Organised genome dynamics in the *Escherichia coli* species
888 results in highly diverse adaptive paths. *PLoS Genet* **5**, e1000344,
889 doi:10.1371/journal.pgen.1000344 (2009).
890 28 Behdenna, A., Pothier, J., Abby, S. S., Lambert, A. & Achaz, G. Testing for
891 Independence between Evolutionary Processes. *Syst Biol* **65**, 812-823,
892 doi:10.1093/sysbio/syw004 (2016).
893 29 Kohler, J. *et al.* In vitro activities of the potent, broad-spectrum carbapenem MK-
894 0826 (L-749,345) against broad-spectrum beta-lactamase-and extended-
895 spectrum beta-lactamase-producing *Klebsiella pneumoniae* and *Escherichia coli*
896 clinical isolates. *Antimicrob. Agents Chemother.* **43**, 1170-1176 (1999).
897 30 Potron, A. *et al.* Genetic and biochemical characterisation of OXA-232, a
898 carbapenem-hydrolysing class D beta-lactamase from Enterobacteriaceae. *Int J*
899 *Antimicrob Agents* **41**, 325-329, doi:10.1016/j.ijantimicag.2012.11.007 (2013).
900 31 Jaffe, A., Chabbert, Y. A. & Semonin, O. Role of porin proteins OmpF and OmpC in
901 the permeation of beta-lactams. *Antimicrob. Agents Chemother.* **22**, 942-948
902 (1982).

903 32 Clinical and Laboratory Standards Institute (CLSI). Performance Standards for
904 Antimicrobial Susceptibility Testing, Twenty-Seventh Informational Supplement,
905 M100-S25 edition. Clinical and Laboratory Standards Institute, Wayne, PA.
906 33 Sprouffske, K. & Wagner, A. Growthcurver: an R package for obtaining
907 interpretable metrics from microbial growth curves. *BMC Bioinformatics* **17**, 172,
908 doi:10.1186/s12859-016-1016-7 (2016).
909 34 Hughes, D. & Andersson, D. I. Evolutionary Trajectories to Antibiotic Resistance.
910 *Annu Rev Microbiol* **71**, 579-596, doi:10.1146/annurev-micro-090816-093813
911 (2017).
912 35 Collignon, P., Beggs, J. J., Walsh, T. R., Gandra, S. & Laxminarayan, R.
913 Anthropological and socioeconomic factors contributing to global antimicrobial
914 resistance: a univariate and multivariable analysis. *The Lancet. Planetary health*
915 **2**, e398-e405, doi:10.1016/s2542-5196(18)30186-4 (2018).
916 36 Queenan, A. M. & Bush, K. Carbapenemases: the versatile beta-lactamases. *Clin*
917 *Microbiol Rev* **20**, 440-458, table of contents, doi:10.1128/cmr.00001-07 (2007).
918 37 Fajardo-Lubian, A., Ben Zakour, N. L., Agyekum, A., Qi, Q. & Iredell, J. R. Host
919 adaptation and convergent evolution increases antibiotic resistance without loss
920 of virulence in a major human pathogen. *PLoS Pathog* **15**, e1007218,
921 doi:10.1371/journal.ppat.1007218 (2019).
922 38 Dean, C. R. *et al.* Mode of Action of the Monobactam LYS228 and Mechanisms
923 Decreasing In Vitro Susceptibility in *Escherichia coli* and *Klebsiella pneumoniae*.
924 *Antimicrob. Agents Chemother.* **62**, doi:10.1128/aac.01200-18 (2018).
925 39 Didelot, X., Meric, G., Falush, D. & Darling, A. E. Impact of homologous and non-
926 homologous recombination in the genomic evolution of *Escherichia coli*. *BMC*
927 *Genomics* **13**, 256, doi:10.1186/1471-2164-13-256 (2012).
928 40 Jensen, A., Valdorsson, O., Frimodt-Moller, N., Hollingshead, S. & Kilian, M.
929 Commensal streptococci serve as a reservoir for beta-lactam resistance genes in
930 *Streptococcus pneumoniae*. *Antimicrob. Agents Chemother.* **59**, 3529-3540,
931 doi:10.1128/aac.00429-15 (2015).
932 41 Zapun, A., Morlot, C. & Taha, M. K. Resistance to beta-Lactams in *Neisseria* ssp Due
933 to Chromosomally Encoded Penicillin-Binding Proteins. *Antibiotics (Basel,*
934 *Switzerland)* **5**, doi:10.3390/antibiotics5040035 (2016).
935 42 Pletz, M. W. *et al.* Ertapenem pharmacokinetics and impact on intestinal
936 microflora, in comparison to those of ceftriaxone, after multiple dosing in male
937 and female volunteers. *Antimicrob. Agents Chemother.* **48**, 3765-3772,
938 doi:10.1128/aac.48.10.3765-3772.2004 (2004).
939 43 Adler, M., Anjum, M., Andersson, D. I. & Sandegren, L. Combinations of mutations
940 in *envZ*, *ftsI*, *mrdA*, *acrB* and *acrR* can cause high-level carbapenem resistance in
941 Escherichia coli. *J. Antimicrob. Chemother.* **71**, 1188-1198,
942 doi:10.1093/jac/dkv475 (2016).
943 44 Adler, M., Anjum, M., Andersson, D. I. & Sandegren, L. Influence of acquired beta-
944 lactamases on the evolution of spontaneous carbapenem resistance in
945 Escherichia coli. *J. Antimicrob. Chemother.* **68**, 51-59, doi:10.1093/jac/dks368
946 (2013).
947 45 Stecher, B. *et al.* Gut inflammation can boost horizontal gene transfer between
948 pathogenic and commensal Enterobacteriaceae. *Proc Natl Acad Sci U S A* **109**,
949 1269-1274, doi:10.1073/pnas.1113246109 (2012).

46 Kallonen, T. *et al.* Systematic longitudinal survey of invasive *Escherichia coli* in England demonstrates a stable population structure only transiently disturbed by the emergence of ST131. *Genome Res*, doi:10.1101/gr.216606.116 (2017).

47 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).

48 Zhang, F. *et al.* Molecular epidemiology of carbapenemase-producing *Escherichia coli* and the prevalence of ST131 subclone H30 in Shanghai, China. *Eur J Clin Microbiol Infect Dis* **34**, 1263-1269, doi:10.1007/s10096-015-2356-3 (2015).

49 Dam, S., Pages, J. M. & Masi, M. Dual Regulation of the Small RNA MicC and the Quiescent Porin OmpN in Response to Antibiotic Stress in Escherichia coli. *Antibiotics (Basel, Switzerland)* **6**, doi:10.3390/antibiotics6040033 (2017).

50 Chin, C. S. *et al.* Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nature methods* **10**, 563-569, doi:10.1038/nmeth.2474 (2013).

51 Koren, S. *et al.* Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res* **27**, 722-736, doi:10.1101/gr.215087.116 (2017).

52 Bankevich, A. *et al.* SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of computational biology : a journal of computational molecular cell biology* **19**, 455-477, doi:10.1089/cmb.2012.0021 (2012).

53 Gurevich, A., Saveliev, V., Vyahhi, N. & Tesler, G. QUAST: quality assessment tool for genome assemblies. *Bioinformatics* **29**, 1072-1075, doi:10.1093/bioinformatics/btt086 (2013).

54 Seemann, T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30**, 2068-2069, doi:10.1093/bioinformatics/btu153 (2014).

55 Zankari, E. *et al.* Identification of acquired antimicrobial resistance genes. *J. Antimicrob. Chemother.* **67**, 2640-2644, doi:10.1093/jac/dks261 (2012).

56 Carattoli, A. *et al.* In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. *Antimicrob. Agents Chemother.* **58**, 3895-3903, doi:10.1128/aac.02412-14 (2014).

57 Guy, L., Kultima, J. R. & Andersson, S. G. genoPlotR: comparative gene and genome visualization in R. *Bioinformatics* **26**, 2334-2335, doi:10.1093/bioinformatics/btq413 (2010).

58 Page, A. J. *et al.* Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics* **31**, 3691-3693, doi:10.1093/bioinformatics/btv421 (2015).

59 Li, W. & Godzik, A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **22**, 1658-1659, doi:10.1093/bioinformatics/btl158 (2006).

60 Katoh, K. & Standley, D. M. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Mol Biol Evol* **30**, 772-780, doi:10.1093/molbev/mst010 (2013).

61 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).

997 62 Talavera, G. & Castresana, J. Improvement of Phylogenies after Removing
998 Divergent and Ambiguously Aligned Blocks from Protein Sequence Alignments.
999 *Systematic biology* **56**, 564-577, doi:10.1080/10635150701472164 (2007).
1000 63 Darriba, D., Taboada, G. L., Doallo, R. & Posada, D. ProtTest 3: fast selection of
1001 best-fit models of protein evolution. *Bioinformatics* **27**, 1164-1165,
1002 doi:10.1093/bioinformatics/btr088 (2011).
1003 64 Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler
1004 transform. *Bioinformatics* **25**, 1754-1760, doi:10.1093/bioinformatics/btp324
1005 (2009).
1006 65 McKenna, A. *et al.* The Genome Analysis Toolkit: a MapReduce framework for
1007 analyzing next-generation DNA sequencing data. *Genome Res* **20**, 1297-1303,
1008 doi:10.1101/gr.107524.110 (2010).
1009 66 Danecek, P. *et al.* The variant call format and VCFtools. *Bioinformatics* **27**, 2156-
1010 2158, doi:10.1093/bioinformatics/btr330 (2011).
1011 67 Cingolani, P. *et al.* A program for annotating and predicting the effects of single
1012 nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila
1013 melanogaster strain w1118; iso-2; iso-3. *Fly* **6**, 80-92, doi:10.4161/fly.19695
1014 (2012).
1015 68 Sambrook, J., Fritsch, E. F. & Maniatis, T. *Molecular Cloning, a laboratory manual,*
1016 *Second Edition*. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor NY,
1017 1989).
1018 69 Lennen, R. M. *et al.* Transient overexpression of DNA adenine methylase enables
1019 efficient and mobile genome engineering with reduced off-target effects. *Nucleic*
1020 *Acids Res.* **44**, e36, doi:10.1093/nar/gkv1090 (2016).
1021