



HAL
open science

Analysis of in-patient evolution of *Escherichia coli* reveals potential links to relapse of bone and joint infections

Stanislas Thiriet-Rupert, Jérôme Josse, David Perez-Pascual, Jason Tasse, Camille Andre, Lélia Abad, David Lebeaux, Jean-Marc Ghigo, Frédéric Laurent, Christophe Beloin

► To cite this version:

Stanislas Thiriet-Rupert, Jérôme Josse, David Perez-Pascual, Jason Tasse, Camille Andre, et al.. Analysis of in-patient evolution of *Escherichia coli* reveals potential links to relapse of bone and joint infections. *Journal of Infectious Diseases*, 2024, 229 (5), pp.1546-1556. 10.1093/infdis/jiad528 . pasteur-04297206

HAL Id: pasteur-04297206

<https://pasteur.hal.science/pasteur-04297206v1>

Submitted on 21 Nov 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NoDerivatives 4.0 International License

1 **Analysis of *in-patient* evolution of *Escherichia coli* reveals potential**
2 **links to relapse of bone and joint infections**

3 Running title (40 characters): in-patient evolution of *E. coli* in BJIs

4 **Major article**

5 Stanislas THIRIET-RUPERT¹, Jérôme JOSSE², David PEREZ-PASCUAL^{1§}, Jason TASSE²,
6 Camille ANDRE², Lélia ABAD², David LEBEAUX^{1,3,4}, Jean-Marc GHIGO¹, Frédéric
7 LAURENT² and Christophe BELOIN*¹

8 ¹ *Institut Pasteur, Université de Paris-Cité, Genetics of Biofilms Laboratory F- 75015 Paris, France.*

9 ² *CIRI, Centre International de Recherche en Infectiologie, Inserm, U1111, Université Claude*
10 *Bernard Lyon 1, CNRS, UMR5308, École Normale Supérieure de Lyon, 69007, France.*

11 ³ *Département de Maladies Infectieuses et Tropicales, AP-HP, Hôpital Saint-Louis, Lariboisière, F-*
12 *75010 Paris, France.*

13 ⁴ *FHU PROTHEE*

14

15 ***Corresponding authors:** Christophe BELOIN, Institut Pasteur, Université de Paris-Cité,
16 Genetics of Biofilms Laboratory F- 75015 Paris, France, christophe.beloin@pasteur.fr (ORCID: 0000-
17 0002-0344-3443)

18

19 **Author contributions:** C.B., S.T.-R., F.L., J.J. designed the experiments. S.T.-R., J.J., D.P.-P., J.T.,
20 C.A., and L.A. performed the experiments. C.B., S.T.-R., F.L., J.J., D.L. and J.-M.G. analyzed data. S.T.-
21 R., C.B. and J.J. wrote the manuscript with significant help of D.L., F.L. and J.-M.G.

22

23 Abstract: 148 words; Main text: 3476 words

24

25 40-word-or-less summary: Here, we identify that three *E. coli* strains involved in bone and joint
26 infection (BJI) relapses adapted in-patient through various molecular mechanisms allowing to
27 outcompete other bacteria or evading immune system. Understanding in-patient adaptation may
28 help fighting relapses in BJIs.

29 Abstract

30 Bone and joint infections (BJIs) are difficult to treat and affect a growing number of patients,
31 in which relapses are observed in 10-20% of the case. These relapses, which call for prolonged
32 antibiotic treatment and increase resistance emergence risk, may originate from ill understood
33 adaptation of the pathogen to the host. Here, we investigated three pairs of *Escherichia coli*
34 strains from BJI cases and their relapses to unravel in-patient adaptation. Whole genome
35 comparison presented evidence for positive selection and phenotypic characterization showed
36 that biofilm formation remained unchanged, contrary to what is usually described in such cases.
37 Although virulence was not modified, we identified the loss of two virulence factors
38 contributing to immune system evasion in one of the studied strains. Other strategies, including
39 global growth optimization and colicin production, likely allowed the strains to outcompete
40 competitors. This work highlights the variety of strategies allowing in-patient adaptation in
41 BJIs.

42

43

44 **Keywords:** BJI, adaptation, relapse, *Escherichia coli*

45

46 Background

47 Bone and joint infections (BJIs), including prosthetic joint infections (PJI), are severe
48 and difficult to treat infections affecting a growing number of patients, and showing 10 to 20%
49 of relapse [1–6]. Although staphylococci are responsible for more than 50% of all BJI, 20% of
50 them are caused by *Enterobacteriales*, especially in the case of early and late acute infections
51 [7], whereas *Pseudomonas aeruginosa* infections are associated with particularly difficult to
52 treat BJI [6]. These data therefore support the need to study BJI caused by both Gram-positive
53 and Gram-negative bacteria and for a better understanding of the *in vivo* adaptation leading to
54 relapses.

55
56 During host colonization, pathogenic bacteria must face the endogenous microbiota or
57 other co-infecting bacteria. This competition could involve toxin production [8,9] and other
58 adaptive strategies to access resources [10] and evade the host immune system [11,12],
59 eventually leading to a decrease virulence accounting for bacterial persistence in the host.

60
61 The genetic modifications underlying these adaptations are in general quite restricted upon
62 natural host purifying selection [13,14]. In some extreme cases, mutator genotypes can evolve,
63 thus facilitating within-host adaptation [15,16]. Extensive genome reductions and
64 rearrangements [17–19] are also observed leading to the loss of non-essential genes in the host
65 environment. However, there is an increasing body of evidence that these gene losses could
66 also fuel adaptative evolution and microbial adaptation within patient is currently under intense
67 scrutiny [20,21]. However, acquiring longitudinal samples and studying pathogens adaptation
68 from the same patient in the context of infection relapses is difficult [22]. In this study, we
69 analysed three pairs of *E. coli* strains isolated from initial and recurrent BJIs in 3 patients to

70 investigate how in-patient evolution occurred in these situations and provide insights on the
71 potential mechanisms involved in infection relapse.

72

73

74 **Methods**

75

76 **Selection of clinical strains**

77 Initial/Relapse pairs of clinical *E. coli* strains used in this study were selected among the routine
78 collections of the Bacteriology Department of Hôpital de la Croix-Rousse, Hospices Civils de
79 Lyon and the Bacteriology laboratory at Cerballiance, Lyon. The first search criterion
80 corresponded to collect *E. coli* strains isolated from BJI from 2017 to 2020. Then, initial/relapse
81 pairs were formed when *E. coli* strain was isolated in two different samples at two different
82 times in the same patient (initial and relapse) and at the same site of infection. Strain
83 characteristics are available in Table 1. Antimicrobial susceptibility tests were performed by
84 using the disc diffusion method (Supplementary Table 1).

85

86 **Bacterial strains and growth conditions**

87 Bacterial strains used in this study are listed in Supplementary Table 2. Bacteria were grown in
88 Miller's Lysogeny Broth (LB) (Corning) or M63B1 minimum medium supplemented with
89 appropriate antibiotic and carbon source as specified. Liquid cultures were incubated at 37°C
90 or 30°C with 180rpm shaking. Solid cultures were done on LB with 1.5% agar supplemented
91 with the appropriate antibiotics. Chemicals and media were purchased from Sigma-Aldrich.

92

93 **Strain construction**

94 Insertion of a Zeo-GFP and Zeo-mars genetic cassettes at the permissive lambda *att* attachment
95 site as well as the deletion of *fliC* in I1 and R1 strains and the deletion of two adhesin coding
96 genes in I2 strain were done using pKOBApra plasmid [23] and lambda-red recombination.
97 Primers used for genetic constructions are listed in Supplementary Table 3.

98

99 **CFU counting on biomaterials**

100 The biofilm forming capacity was assessed on different orthopaedic biomaterials [stainless steel
101 (AISI 316L), titanium (TA6V), medical grade silicone and ultra-high molecular weight
102 polyethylene (UHMWPE)]. Uniform cylindrical pegs of orthopaedic biomaterial were
103 manufactured by a supplier of medical and surgical orthopaedics equipment (Groupe Lépine,
104 Genay, France) under the same conditions of manufacture of orthopaedic biomaterial and
105 sterilised by gamma rays. Biomaterial pegs were fixed into the lid of 24-well flat bottom cell
106 culture plates (Corning Inc., Corning, NY, USA). Bacterial biofilms were formed by immersing
107 the lid-containing pegs into 2mL of cultures at $OD_{600}=0.05$ and incubated in LB for 24 hours at
108 37°C in humid atmosphere. After 3 rinsing steps in PBS, the pegs were dropped in a Falcon
109 tube containing 5mL of PBS for biofilm disruption using sonication for 5 minutes
110 (BACTOSONIC 14.2, BANDELIN, Berlin, Germany). Then, viable cell count was determined
111 by serial dilution and plating.

112

113 **Bacterial adhesion to MG63 human osteoblastic cells**

114 The osteoblastic cell line MG63 (LGC standards, Molsheim, France) was used to test bacterial
115 adhesion. Cells were cultured in 75cm² flasks (T75, BD Falcon, Le Pont de Claix, France) at
116 37°C under 5% CO₂, in a culture medium composed of DMEM (Dulbecco's Modified Eagle
117 Medium containing D-glucose, L-glutamine, pyruvate) supplemented with 10% fetal calf serum
118 (FCS), penicillin (100µg/mL) and streptomycin (100µg/mL), all from Gibco (Paisley, UK).

119 Cells were passaged once a week and used until passage 25 at the most. The day before the
120 infection, MG63 cells were seeded at 100,000 cells per well in a 24-well plate and the different
121 bacterial strains were inoculated in LB. The day of the infection, overnight bacterial cultures
122 were centrifugated, resuspended and adjusted in DMEM + 10% FCS to reach a MOI
123 (multiplicity of infection) of 100:1 and added to cell culture. After 2 hours of infection, cells
124 were washed twice with PBS and incubated and lysed by osmotic shock in sterile water and
125 lysates were plated on tryptone soy agar plates (Biomérieux) using the Easy Spiral® automaton
126 (Interscience, Saint-Nom-la-Bretèche, France). CFU were counted after 18 hours of incubation
127 at 37°C and results were presented as CFU for 100 000 cells.

128

129 **Survival to human serum**

130 For each strain, 50µL at OD₆₀₀=0.4 was mixed in a well of a 96-well plate with 50µL of active
131 human serum. The same process was followed for a control with heat inactivated serum (HIS)
132 (30 minutes at 55°C). Plate was then incubated at 37°C under shaking (450rpm) for 1h30. Each
133 well was then serially diluted and plated on LB for CFU count and the survival of each strain
134 was expressed as the ratio between the CFUs in active serum divided by the CFUs in HIS.

135

136

137 **Zebrafish infection**

138 Overnight bacterial cultures were centrifuged, washed, and resuspended at the desired
139 concentration in PBS. Anesthetized zebrafish larvae were microinjected intravenously 52–58
140 hours post-fertilization with 1nL of each bacterial suspension at a final concentration of 2x10³
141 CFU [24]. Infected larvae were transferred into individual wells containing 1mL of Volvic
142 water (Danone, France) and 0.003% 1-phenyl 2-thiourea (used to improve optical transparency
143 in zebrafish) in 24-well plates, incubated at 28°C, and regularly observed under a

144 stereomicroscope. Protocols describing zebrafish maintenance and analysis, see the
145 corresponding supplementary method section.

146

147 **Competition assay for biofilm formation**

148 The pair of strains used in competition assay carried two different fluorescent tags and were
149 inoculated in LB medium supplemented with appropriate antibiotic and incubated overnight at
150 37°C under shaking. Cultures were then adjusted to $OD_{600} = 0.05$ and mixed in 1:1 ratio
151 (verified by FACS). Each competition mix was then inoculated in a PVC plate and incubated
152 at 37°C for 16 hours. The supernatant was removed from each well, the biofilm resuspended in
153 100µL of PBS and the proportion of each strain was assessed by FACS and the relative fitness
154 calculated as follows:

$$155 \quad \text{relative fitness} = \frac{\left(\frac{Target_{post}}{Target_{pre}}\right)}{\left(\frac{ref_{post}}{ref_{pre}}\right)}$$

156 , where $Target_{pre}$ and $Target_{post}$ are the cell concentrations of the target strain in the mixed
157 culture before and after the overnight incubation, and ref_{pre} and ref_{post} are the cell concentrations
158 of the corresponding reference strain. In order to account for growth differences due to the
159 expression of one or the other fluorescent tag (supplementary Figure 1), competitions were run
160 with both possible combinations of fluorescent tags. This protocol was used to compete R1 and
161 R3 strains against their respective ancestors. However, because GFP and RFP tags were not
162 expressed in I2 and R2 strains, we used an alternative strategy based on PCR amplification and
163 Sanger sequencing of a SNP locus allowing to distinguish both strains (see supplementary
164 method section).

165

166 **Competition assay for growth in liquid cultures**

167 The strains were incubated at 37°C in the appropriate medium and adjusted to OD₆₀₀=0.05. A
168 1:1 ratio mixed culture was verified by FACS and incubated overnight at 37°C with shaking.
169 The concentrations of target and reference strains was measured using FACS and relative
170 fitness were calculated as above.

171

172 **SNP analysis**

173 Mutations in the adapted strain were identified by using the genome assembly of its ancestral
174 strain (see supplementary methods) as a reference. breseq version 0.35.0 [25] with the
175 consensus mode was used with default parameters. As a confirmation of this analysis,
176 Freebayes version 1.3.3 [26] was also used to identify SNPs using the following stringent
177 thresholds: minimum depth of 20 reads, a maximum depth corresponding to the average depth
178 over the genome plus one standard deviation, at least 10 forward and 10 reverse reads
179 supporting the SNP, minimum variant quality of 30, minimum base quality of 20. Only variants
180 identified in both analyses were validated. The resulting identified mutations were filtered out
181 if they were too close to each other (less than 50bp). These mutations are usually the result of
182 misaligned reads, often due to repetitive regions [27]. Because these regions are not 100%
183 identical, they are disturbing the mapping process resulting in false positives. Mutations found
184 in mobile elements were kept if only one copy was present in the genome.

185

186 **Statistical analysis**

187 Figures and statistical tests were performed using either Prism 9.5.1 for Mac OS X (GraphPad
188 Software, Inc.) or Rstudio with R version 4.2.1 using the R package GGplot2 [28].

189

190 Methods related to the supplementary results are available in supplementary methods.

191

192

193 **RESULTS**

194 **1- In-patient evolved *Escherichia coli* clones showed no antibiotic resistance acquisition** 195 **but signs of natural selection.**

196

197 In order to investigate potential cases of in-patient adaptation, we obtained three pairs
198 of *E. coli* strains sampled in three patients presenting relapsing BJIs. Each pair is composed of
199 a sample at the time of the initial infection (strains hereafter denominated as I1 to I3 for cases
200 1 to 3) and a second sample corresponding to the relapse of this infection (strains hereafter
201 denominated as R1 to R3 for cases 1 to 3). The genome of each strain was sequenced and
202 analyzed to confirm that the infection and relapse strains were indeed the same and not a case
203 of re-infection (Table 1, see supplementary methods). Interestingly, antimicrobial susceptibility
204 testing (supplementary Table 1) and in silico analysis (supplementary Data 1) did not reveal
205 any acquisition of antibiotic resistance associated with the relapse events. Consequently, each
206 strain pair was examined for other phenotypic and genetic features that could have been selected
207 for to allow the infection relapses.

208

209 The first notable difference identified by genomic analysis was the number of gained
210 and lost genes (supplementary Table 4), some of which corresponding to potential plasmids
211 (supplementary Data 2, see supplementary methods). Particularly, the R2 strain that displayed
212 the longest in-patient evolution underwent an important genome reduction with the loss of 3.4%
213 of its size corresponding to 170 coding sequences involved in various functions such as
214 transport, iron acquisition, carbohydrate acquisition and utilization, defense against competitors
215 with Microcin H47 production-related genes and three adhesins (supplementary Data 2).

216

217 The number of mutation per base per year was quite high for all three relapse strains (1.11×10^{-6}
218 to 8.38×10^{-6}) as compared to what was reported in the literature (6.9×10^{-7} for *E. coli* ED1a in

219 the gut, [27]). In addition, even if no functional convergence was observed in the mutated genes,
220 the predominance of loss of function and non-synonymous mutations suggests that positive
221 selection was at play in these cases (Table 2, see supplementary methods).

222

223

224

225 **II- *In vivo* adaptation did not modify biofilm formation capacities.**

226

227 Because strain motility could impact the virulence and biofilm forming capacity, we
228 tested each strain on low-agar LB plates at 30°C and 37°C, which showed that the second and
229 third pairs of strains were non-motile (supplemental Figure 2, see supplementary methods).
230 Both I1 and R1 strains display temperature dependent motility very likely due to a flagellin
231 production restricted to 30°C (supplementary Figure 2 and 3, see Supplementary Results).

232

233 Biofilm formation was assessed on various surfaces showing that the R1 relapse strain
234 formed 1.71 more biofilms than I1 strain when grown on silicon coupons (Figure 1.A). No
235 difference was identified for the I2-R2 pair (Figure 1.B), while the I3 strain showed a 1.49
236 increase in biofilm formation on silicon (Figure 1.C). However, in all cases the tested clinical
237 strains showed a weak biofilm forming capacity, all of them being equivalently or even less
238 efficient than *E. coli* K12 strain MG1655 used as control for weak biofilm formation.
239 Interestingly, R1 strain showed a slight decrease of its adhesion capacity to osteoblasts (Figure
240 1.D). However, no evident genetic modification could be linked to this phenotype. No
241 difference was observed either in term of auto-aggregation or production of curli amyloids and
242 cellulose (supplementary Figure 4, see supplementary methods and Supplementary Results).
243 Taken together, these results show that very limited differences in biofilm formation were
244 detected within each pair suggesting that a modification of biofilm formation and biofilm-
245 related phenotypes was unlikely to be involved in *in vivo* adaptation in these three studied cases.

246

247

248 **III- Loss of adhesins induces serum resistance in R2 strain.**

249

250 To assess the strain capacity to challenge to the host immune defenses, we measured
251 the survival of all strains to exposure to active human serum. While we observed no difference
252 between I3 and R3 strains, the relapse strain R1 had a slight decrease in serum resistance
253 compared to I1 strain and the relapse strain R2 had a clear increase in serum resistance relatively
254 to its ancestral strain I2 (Figure 2.A). Although we identified no difference in the I2 vs R2
255 lipopolysaccharide profile (outer-membrane component targeted by the complement) (Figure
256 2.B), the loss of two adhesins in R2 could be linked to this phenotype (supplementary Data 2)
257 [11]. Indeed, deletion of each adhesin in I2 strain led to reduced serum resistance suggesting
258 that the loss of both adhesins could have been selected through the conferred immune system
259 evasion, contributing to the strain adaptation to the host environment (Figure 2.C).
260 Unfortunately, we were unable to assess the additive effect of these mutations since we could
261 not reconstruct the double mutant in the I2 strain despite repeated attempts.

262

263

264 **IV-Strain virulence was not targeted during in-patient selection.**

265

266 To gain insight into the potential evolution of the virulence of the studied strains, we
267 assessed their cytotoxicity on HeLa cells as well as hemolysis capacities and showed no
268 differences, except for a slightly lower hemolysis induced by R2 strain after 24h of incubation
269 in presence of defibrinated horse blood (Figure 3.A.B).

270

271 Potential differences in in vivo virulence were tested in a zebrafish (*Danio rerio*) larval model
272 [29]. While no differences were found in bacterial load for all pairs (Sup Figure S5.A.B.C),
273 zebrafish larvae exposed to R2 relapse strain had a slightly lower survival, although non-
274 significant, 24 hours post infection (hpi) than with its related ancestor I2, even though both

275 strains ultimately induced the same level of mortality 48 hpi (Figure 3.C.D). No difference was
276 identified concerning the zebrafish innate immune response level (Figure 3.E). These results
277 suggest that the R2 relapse strain induced a slightly quicker mortality, which does not seem to
278 be linked to a difference in colonization ability.

279

280

281 **V- Growth optimization and colicin production as strategies to outcompete** 282 **other bacteria.**

283

284 To assess if a difference in growth capacity could advantage the strains isolated after
285 relapses, they were pitted against their related ancestor in 1:1 ratio competition experiments in
286 biofilm and liquid conditions. First, no difference in fitness was observed in biofilm condition,
287 which further straightened the hypothesis that biofilm formation was not targeted during the in
288 vivo adaptation. In liquid, while R1 strain slightly outcompeted its ancestor I1, R2 strain had a
289 striking fitness advantage over I2 (Figure 4.A).

290 To further investigate this phenotype, we competed I2 and R2 strains in presence of
291 different sugars as well as of an iron chelator (2,2'-Bipyridyl) based on identified genetic
292 modifications and especially gene losses related to uptake and metabolism of these substrates
293 (supplementary Data 2). Surprisingly, despite these gene losses, R2 strain showed no growth
294 defect and even globally had an advantage through a capacity to better sustain growth during
295 stationary phase (supplementary Figure 6), as confirmed by competition assays (Figure 4.B).
296 Therefore, extensive gene loss seems to have resulted in a global growth optimization, which
297 could have helped the strain to outcompete other bacteria including its ancestor.

298 Another strategy to outcompete its kin and promote subsequent adaptation was
299 identified in the relapse strain R1, which acquired a colicin-producing plasmid allowing the
300 production of colicin E1. An overlay assay indeed confirmed the acquired capacity of R1
301 relapse strain to inhibit I1 strain growth (supplementary Figure 7).

302

303

304 **Discussion**

305

306 Treatment failure is a common issue of BJIs, allowing bacteria to remain and adapt to
307 the host, ultimately increasing the risks of relapse and the difficulty to cure the patient. A better
308 understanding of the mechanisms underlying pathogens adaptation to the host environment
309 would allow improving the management of relapsing BJIs. Here we investigated three cases of
310 BJIs relapse involving pathogenic *E. coli*.

311 Our genetic and phenotypic analysis revealed diverse functions arising from bacterial
312 adaptation to the host environment during BJIs. We found that adapted strains isolated from
313 relapses displayed a relatively high mutation rate, which is a known adaptive advantage
314 increasing the advantageous mutation probability that could be subsequently selected
315 [13,30,31]. However, no mutations were found in genes associated with DNA damage repair,
316 suggesting that these high mutation rates were intrinsic to these strains and could have further
317 promoted the selection of beneficial mutations, partly explaining their adaptive success.

318 Although biofilms are known to be frequently involved in infections and relapses, none of the
319 three investigated cases showed an increased biofilm formation capacity that could have played
320 a role in the relapses and all strains were even characterized as poor *in vitro* biofilm formers.

321 While a strong *in vitro* biofilm formation is often linked to the severity of clinical outcomes in
322 prosthesis-related infections [32–34], it did not seem to be a critical factor in this study.

323 The host immune system exerts a critical selection pressure on pathogenic bacteria. In I1 and
324 R1 strains, the inhibition of flagellin production at 37°C could minimize their recognition by
325 the host's immune defenses [35,36]. This could not only contribute to immune system evasion
326 but also to save energy through a down-regulation or suppression of flagellin production, as
327 shown to be the case in the context of gut infection by *E. coli* [37,38].

328 This immune system evasion strategy is even clearer in the case of the R2 relapse strain, which
329 showed increased survival to human serum through the loss of two adhesins. This suggests that
330 in these cases, evading the host's immune system may be more critical than maintaining
331 virulence as shown to happen during in vivo adaptation [39,40]. This could allow pathogens to
332 stay longer in the host as shown in the infection experiments performed in zebrafish, where no
333 difference was observed.

334 In addition, the R2 strain presented an interesting case of gene loss. Genome reduction is usually
335 considered as a strategy selected because energy consumption is reduced via the loss of genes
336 that are not used or under selective pressure [41,42]. In the case of the R2 strain, numerous gene
337 losses seem to result in a global growth capacities optimization, potentially associated to lower
338 the cost for gene production and lower amount of DNA to replicate. It is possible that some
339 nutritional resources could be directly acquired from the new host environment instead of being
340 synthesized by the cells and that the loss of genes involved in this production optimizes the cell
341 metabolism [43]. Gain of growth capacity in the context of an infection could be of paramount
342 importance to outcompete other endogenous or pathogenic bacteria and thrive in the host
343 environment.

344 The polymicrobial context of some of these BJIs is also an important parameter to stress since
345 the interactions between the different bacterial species or strains could play an important role
346 in the relapses and the mechanisms shaping bacterial adaptation to the host environment
347 [44,45]. This study may therefore have missed some adaptation that rely on bacteria-bacteria
348 direct or indirect interactions. In addition, as for any longitudinal study, there is always a
349 possibility that the strains causing the very initial infection already evolved at the time of
350 sampling when infection was clinically detected, and some of these mutations, that would be
351 undetected in our study, could potentially contribute to the success of the relapse strains.

352 Finally, the small number of strains studied, and the diversity of BJI cases, make it challenging
353 to draw definitive conclusions about typical strain profiles and mechanisms involved in in vivo
354 adaptation. However, it seems that a higher mutation rate allowing a better adaptability as well
355 as evolving immune system evasion and outcompeting other bacteria are relevant strategies
356 selected in patient. Further investigations with a larger cohort would help determine whether
357 these diverse adaptive strategies constitute a more general rule of *E. coli* pathogenic strain in
358 BJIs.

359

360

361 DATA AVAILABILITY

362 All sequencing reads as well as genome assembly for each strain were deposited in NCBI under
363 the BioProject accession number PRJNA932763. The code used in this study is available at
364 https://github.com/Sthiriet-rupert/BJIs_adaptation.

365

366 COMPETING FINANCIAL INTERESTS

367 The authors declare no competing financial interests.

368

369 ACKNOWLEDGEMENTS

370 We thank Dr. Emma Colucci-Guyon and Laurent Boucontet for critical reading of the
371 manuscript. We are grateful to the Zebrafish Projects Hub of the Institut Pasteur and especially
372 to Dr. Emma Colucci-Guyon and Laurent Boucontet who carried out the infection experiments
373 in zebrafish larvae. We are grateful to Dr. Monica Rolando for kindly providing the anti-
374 flagellin antiserum. We are grateful to Emmanuel Chanard from Cerballiance for providing
375 strains and microbiological information.

376

377 **ETICS STATEMENT**

378 Animal experiments were performed according to European Union guidelines for handling of
379 laboratory animals (http://ec.europa.eu/environment/chemicals/lab_animals/home_en.htm)
380 and were approved by the Institut Pasteur Animal Care and Use Committee and the French
381 Ministry of Research (APAFIS#31827).

382

383 **FUNDINGS**

384 This work was supported by the French National Research Agency (ANR), project EvolTolAB
385 (ANR-18-CE13-0010), by the French government's Investissement d'Avenir Program,
386 Laboratoire d'Excellence "Integrative Biology of Emerging Infectious Diseases" (grant
387 n°ANR-10-LABX-62-IBEID) and by the Fondation pour la Recherche Médicale (grant
388 DEQ20180339185). S.T.-R was supported by the French National Research Agency (ANR),
389 project EvolTolAB (ANR-18-CE13-0010).

390

391 **References**

- 392 1. Osmon DR, Barbari EF, Berendt AR, et al. Diagnosis and management of prosthetic
393 joint infection: clinical practice guidelines by the Infectious Diseases Society of America.
394 Clin Infect Dis. **2013**; 56(1):e1–e25.
- 395 2. Tande AJ, Patel R. Prosthetic joint infection. Clin Microbiol Rev. **2014**; 27(2):302–
396 345.
- 397 3. Tong SYC, Davis JS, Eichenberger E, Holland TL, Fowler VG. Staphylococcus
398 aureus infections: epidemiology, pathophysiology, clinical manifestations, and management.
399 Clin Microbiol Rev. **2015**; 28(3):603–661.
- 400 4. Laurent E, Gras G, Druon J, et al. Key features of bone and joint infections following
401 the implementation of reference centers in France. Med Mal Infect. **2018**; 48(4):256–262.
- 402 5. Ferguson J, Alexander M, Bruce S, O’Connell M, Beecroft S, McNally M. A
403 retrospective cohort study comparing clinical outcomes and healthcare resource utilisation in
404 patients undergoing surgery for osteomyelitis in England: a case for reorganising orthopaedic
405 infection services. J Bone Jt Infect. **2021**; 6(5):151–163.
- 406 6. Cerioli M, Batailler C, Conrad A, et al. Pseudomonas aeruginosa Implant-Associated
407 Bone and Joint Infections: Experience in a Regional Reference Center in France. Front Med
408 (Lausanne). **2020**; 7:513242.
- 409 7. Triffault-Fillit C, Ferry T, Laurent F, et al. Microbiologic epidemiology depending on
410 time to occurrence of prosthetic joint infection: a prospective cohort study. Clin Microbiol
411 Infect. **2019**; 25(3):353–358.

- 412 8. Joo H-S, Cheung GYC, Otto M. Antimicrobial activity of community-associated
413 methicillin-resistant *Staphylococcus aureus* is caused by phenol-soluble modulin derivatives.
414 *J Biol Chem.* **2011**; 286(11):8933–8940.
- 415 9. Majeed H, Gillor O, Kerr B, Riley MA. Competitive interactions in *Escherichia coli*
416 populations: the role of bacteriocins. *ISME J.* **2011**; 5(1):71–81.
- 417 10. Baishya J, Wakeman CA. Selective pressures during chronic infection drive microbial
418 competition and cooperation. *npj Biofilms Microbiomes.* Nature Publishing Group; **2019**;
419 5(1):1–9.
- 420 11. Klemm P, Roos V, Ulett GC, Svanborg C, Schembri MA. Molecular characterization
421 of the *Escherichia coli* asymptomatic bacteriuria strain 83972: the taming of a pathogen.
422 *Infect Immun.* **2006**; 74(1):781–785.
- 423 12. Price EP, Sarovich DS, Mayo M, et al. Within-host evolution of *Burkholderia*
424 *pseudomallei* over a twelve-year chronic carriage infection. *mBio.* **2013**; 4(4):e00388-13.
- 425 13. Lieberman TD, Flett KB, Yelin I, et al. Genetic variation of a bacterial pathogen
426 within individuals with cystic fibrosis provides a record of selective pressures. *Nat Genet.*
427 **2014**; 46(1):82–87.
- 428 14. Marvig RL, Johansen HK, Molin S, Jelsbak L. Genome analysis of a transmissible
429 lineage of *Pseudomonas aeruginosa* reveals pathoadaptive mutations and distinct evolutionary
430 paths of hypermutators. *PLoS Genet.* **2013**; 9(9):e1003741.
- 431 15. Klemm EJ, Gkrania-Klotsas E, Hadfield J, et al. Emergence of host-adapted
432 *Salmonella* Enteritidis through rapid evolution in an immunocompromised host. *Nat*
433 *Microbiol.* Nature Publishing Group; **2016**; 1(3):1–6.
- 434 16. Feliziani S, Marvig RL, Luján AM, et al. Coexistence and Within-Host Evolution of
435 Diversified Lineages of Hypermutable *Pseudomonas aeruginosa* in Long-term Cystic Fibrosis
436 Infections. *PLoS Genet.* **2014**; 10(10):e1004651.
- 437 17. Smith EE, Buckley DG, Wu Z, et al. Genetic adaptation by *Pseudomonas aeruginosa*
438 to the airways of cystic fibrosis patients. *Proc Natl Acad Sci U S A.* **2006**; 103(22):8487–
439 8492.
- 440 18. Losada L, Ronning CM, DeShazer D, et al. Continuing evolution of *Burkholderia*
441 *mallei* through genome reduction and large-scale rearrangements. *Genome Biol Evol.* **2010**;
442 2:102–116.
- 443 19. Linz B, Ivanov YV, Preston A, et al. Acquisition and loss of virulence-associated
444 factors during genome evolution and speciation in three clades of *Bordetella* species. *BMC*
445 *Genomics.* **2016**; 17(1):767.
- 446 20. Bolotin E, Hershberg R. Gene Loss Dominates As a Source of Genetic Variation
447 within Clonal Pathogenic Bacterial Species. *Genome Biol Evol.* **2015**; 7(8):2173–2187.
- 448 21. Albalat R, Cañestro C. Evolution by gene loss. *Nat Rev Genet.* **2016**; 17(7):379–391.
- 449 22. Trouillet-Assant S, Lelièvre L, Martins-Simões P, et al. Adaptive processes of
450 *Staphylococcus aureus* isolates during the progression from acute to chronic bone and joint
451 infections in patients. *Cell Microbiol.* **2016**; 18(10):1405–1414.
- 452 23. Chaveroche MK, Ghigo JM, Enfert C d'. A rapid method for efficient gene
453 replacement in the filamentous fungus *Aspergillus nidulans*. *Nucleic Acids Res.* **2000**;
454 28(22):E97.
- 455 24. Colucci-Guyon E, Tinevez J-Y, Renshaw SA, Herbomel P. Strategies of professional
456 phagocytes in vivo: unlike macrophages, neutrophils engulf only surface-associated microbes.
457 *J Cell Sci.* **2011**; 124(Pt 18):3053–3059.
- 458 25. Deatherage DE, Barrick JE. Identification of mutations in laboratory evolved microbes
459 from next-generation sequencing data using breseq. *Methods Mol Biol.* **2014**; 1151:165–188.
- 460 26. Garrison E, Marth G. Haplotype-based variant detection from short-read sequencing
461 [Internet]. arXiv; 2012 [cited 2023 Oct 16]. Available from: <http://arxiv.org/abs/1207.3907>

- 462 27. Ghalayini M, Launay A, Bridier-Nahmias A, et al. Evolution of a Dominant Natural
463 Isolate of *Escherichia coli* in the Human Gut over the Course of a Year Suggests a Neutral
464 Evolution with Reduced Effective Population Size. *Appl Environ Microbiol.* **2018**;
465 84(6):e02377-17.
- 466 28. Wickham H. *ggplot2: Elegant Graphics for Data Analysis* [Internet]. New York, NY:
467 Springer; 2009 [cited 2023 Feb 7]. Available from: [https://link.springer.com/10.1007/978-0-](https://link.springer.com/10.1007/978-0-387-98141-3)
468 [387-98141-3](https://link.springer.com/10.1007/978-0-387-98141-3)
- 469 29. Wiles TJ, Bower JM, Redd MJ, Mulvey MA. Use of Zebrafish to Probe the Divergent
470 Virulence Potentials and Toxin Requirements of Extraintestinal Pathogenic *Escherichia coli*.
471 *PLOS Pathogens*. Public Library of Science; **2009**; 5(12):e1000697.
- 472 30. Giraud A, Matic I, Tenaillon O, et al. Costs and benefits of high mutation rates:
473 adaptive evolution of bacteria in the mouse gut. *Science.* **2001**; 291(5513):2606–2608.
- 474 31. Tenaillon O, Barrick JE, Ribbeck N, et al. Tempo and mode of genome evolution in a
475 50,000-generation experiment. *Nature.* **2016**; 536(7615):165–170.
- 476 32. Svensson Malchau K, Tillander J, Zaborowska M, et al. Biofilm properties in relation
477 to treatment outcome in patients with first-time periprosthetic hip or knee joint infection. *J*
478 *Orthop Translat.* **2021**; 30:31–40.
- 479 33. Sanchez CJ, Mende K, Beckius ML, et al. Biofilm formation by clinical isolates and
480 the implications in chronic infections. *BMC Infect Dis.* **2013**; 13:47.
- 481 34. Post V, Harris LG, Morgenstern M, et al. Comparative Genomics Study of
482 *Staphylococcus epidermidis* Isolates from Orthopedic-Device-Related Infections Correlated
483 with Patient Outcome. *J Clin Microbiol.* **2017**; 55(10):3089–3103.
- 484 35. Smith KD, Andersen-Nissen E, Hayashi F, et al. Toll-like receptor 5 recognizes a
485 conserved site on flagellin required for protofilament formation and bacterial motility. *Nat*
486 *Immunol.* **2003**; 4(12):1247–1253.
- 487 36. Bambou J-C, Giraud A, Menard S, et al. In vitro and ex vivo activation of the TLR5
488 signaling pathway in intestinal epithelial cells by a commensal *Escherichia coli* strain. *J Biol*
489 *Chem.* **2004**; 279(41):42984–42992.
- 490 37. Giraud A, Arous S, De Paepe M, et al. Dissecting the genetic components of
491 adaptation of *Escherichia coli* to the mouse gut. *PLoS Genet.* **2008**; 4(1):e2.
- 492 38. Leatham MP, Stevenson SJ, Gauger EJ, et al. Mouse intestine selects nonmotile flhDC
493 mutants of *Escherichia coli* MG1655 with increased colonizing ability and better utilization of
494 carbon sources. *Infect Immun.* **2005**; 73(12):8039–8049.
- 495 39. Winstanley C, O'Brien S, Brockhurst MA. *Pseudomonas aeruginosa* Evolutionary
496 Adaptation and Diversification in Cystic Fibrosis Chronic Lung Infections. *Trends Microbiol.*
497 **2016**; 24(5):327–337.
- 498 40. Viberg LT, Sarovich DS, Kidd TJ, et al. Within-Host Evolution of *Burkholderia*
499 *pseudomallei* during Chronic Infection of Seven Australasian Cystic Fibrosis Patients. *mBio.*
500 **2017**; 8(2):e00356-17.
- 501 41. Moran NA. Microbial minimalism: genome reduction in bacterial pathogens. *Cell.*
502 **2002**; 108(5):583–586.
- 503 42. Weinert LA, Welch JJ. Why Might Bacterial Pathogens Have Small Genomes? *Trends*
504 *Ecol Evol.* **2017**; 32(12):936–947.
- 505 43. D'Souza G, Kost C. Experimental Evolution of Metabolic Dependency in Bacteria.
506 *PLoS Genet.* **2016**; 12(11):e1006364.
- 507 44. Tay WH, Chong KKL, Kline KA. Polymicrobial-Host Interactions during Infection. *J*
508 *Mol Biol.* **2016**; 428(17):3355–3371.
- 509 45. Boldock E, Surewaard BGJ, Shamarina D, et al. Human skin commensals augment
510 *Staphylococcus aureus* pathogenesis. *Nat Microbiol.* **2018**; 3(8):881–890.

511

512 **Figure legends**

513 **Figure 1.** Biofilm formation capacity for infection and relapse strains from case 1 (A), 2 (B) and 3 (C). Six different
514 surfaces were tested, *E. coli* K12 MG1655 was used as a control for weak biofilm formation and *E. coli* TG1 as a
515 control for strong biofilm formation (n=3 to 4 biological replicates). (D) Adhesion capacity on osteoblast human
516 cells (n=3 biological replicates). For each panel, statistical significance was only assessed between the infection
517 strain and its related relapse strain according to non-parametric two-tailed Mann-Whitney test. * p<0.05; **
518 p<0.01.

519

520 **Figure 2.** Serum survival capacity. A. Normalized serum survival capacity of each strain expressed as percentage
521 of the control for growth in Heat Inactivated Serum (HIS). *E. coli* strains CFT073 and MG1655 were used as
522 positive and negative controls, respectively (n=3 biological replicates). Statistical significance was only assessed
523 between the infection strain and its related relapse strain according to non-parametric two-tailed Mann-Whitney
524 test. * p<0.05. B. LPS profile of I2 and R2 strains. *E. coli* MG1655 was used as a negative control and the same
525 strain in which the O-Antigen was reconstructed was used as a positive control. C. Relative Serum survival of I2
526 ancestral strain, R2 relapse strain as well as two deletion mutants constructed in the I2 ancestral strain: one for the
527 *afa* fimbrial operon and a second for the *yadA*-like encoding gene. Results are expressed relatively to R2 relapse
528 strain (n=4 biological replicates). Statistical significance according to non-parametric two-tailed Mann-Whitney
529 test. * p<0.05; ** p<0.01.

530

531 **Figure 3.** Virulence of the strains. (A) Cytotoxicity test on HeLa cells. All six adapted and ancestral strains were
532 tested as well as MG1655 strain as a negative control. The values for each strain were expressed relatively to the
533 control without bacteria (n=3 biological replicates). No statistical significance was detected according to non-
534 parametric two-tailed Mann-Whitney test. (B) Hemolysis capacity in liquid culture of each relapse strain relatively
535 to its ancestor (n=3 biological replicates). Values higher than 1 show higher hemolysis capacity for the relapse
536 strain, while values lower than 1 show lower hemolysis capacity for the relapse strain. Statistical significance
537 according to non-parametric two-tailed Mann-Whitney test. * p<0.05. (C) Monitoring of zebrafish survival (no
538 statistical significance was detected according to log-rank test), (D) bacterial burden into the zebrafish (no
539 statistical significance was detected according to non-parametric two-tailed Mann-Whitney test) and (E) *in vivo*
540 fluorescence of macrophages and neutrophils for zebrafish infected by I2 and R2 strains as well as non-injected

541 control, MG1655 strain (negative control) and CFT073 (positive control). Each image is representative of 5
542 biological replicates. Fifteen zebrafish were used for each experiment in C and D.

543

544 **Figure 4.** Fitness competitions and growth capacities. **(A)** Fitness competitions of relapse against ancestral strain
545 for all three pairs in planktonic and biofilm conditions (n=6 biological replicates). Statistical significance
546 according to non-parametric two-tailed Mann-Whitney test (*: p-value < 0.05, **: p-value < 0.01). **(B)** Fitness
547 competitions of R2 relapse strain against I2 ancestral strain in different conditions (same as in supplementary
548 Figure S6.B-F). BIP: 2,2'-Bipyridyl (n=6 biological replicates). Statistical significance according to non-
549 parametric two-tailed Mann-Whitney test (*: p-value < 0.05).

550

551 **Supplementary Figures legends**

552 **Supplementary Figure 1.** Optimization of competition protocols. **(A)** Assessment of the fitness cost associated
553 to the expression of each fluorescence marker (GFP or mars) in each strain (n=3 biological replicates). There is a
554 slight cost to produce mars marker in I3 and R3 strains as compared to the production of GFP. Therefore, both
555 combinations of fluorescence marker were used in the competition assays for all strains. Statistical significance
556 according to two-tailed Mann-Whitney test comparing initial and relapsing strains (*: p-value < 0.05, **: p-value
557 < 0.01). **(B)** Comparison of FACS approach and Sanger based QSV analyzer approach to assess the proportion of
558 each strain in a set of mixed samples with different proportions of I1 and R1 strains tagged with GFP and mars,
559 respectively (n=3 biological replicates).

560

561 **Supplemental Figure 2.** Motility of each strain assessed on low agar plates at 30 and 37°C. A *E. coli* MG1655
562 strain deleted for *fliE* to *fliR* genes is used as a negative control and a *E. coli* MG1655 strain in which the motility
563 is increased by the insertion of an IS1 in *yeaJ/dgcJ* encoding a diguanylate cyclase resulting in an increase of c-
564 di-GMP, which in turn fuels motility was used as a positive control.

565

566 **Supplementary Figure 3.** Influence of temperature on I1 and R1 strain motility. Motility of each strain at 30°C
567 **(A)** or 37°C **(B)** when pre-cultured in liquid at 30°C or 37°C. **(C)** Motility of each strain deleted for *fliC* when pre-
568 cultured at 30°C or 37°C and incubated at 30°C for motility testing. **(D)** Production of flagellin (top panel) and
569 RNA polymerase alpha subunit used as a loading control (bottom panel) in I1 and R1 strains assessed by Western
570 blot analysis in different combinations of temperature in liquid cultures.

571

572 **Supplementary Figure 4.** **A.** Auto-aggregation kinetic of all clinical strains. In each case, the infection strain is
573 compared to the relapse strain as well as a strain of *E. coli* MG1655 in which *flu* is under the control of a
574 constitutive promoter (PCL-*flu*) used as a positive control and a strain of *E. coli* MG1655 in which *flu* is deleted
575 (Δ *flu*) used as a negative control (n=3 biological replicates). No statistical significance was identified according
576 to two-tailed Mann-Whitney test comparing initial and relapsing strains. **B.** Pictures showing the red, dry, and
577 rough (rdar) morphotypes of each strain after 24 or 48 hours of incubation at either 30°C or 37°C. The first couple
578 had a pink, dry and rough (pdar) morphotype characterizing the production of curli only. The second couple had
579 a smooth and white morphotype (saw) showing the absence of matrix component production. The third couple had
580 pink and smooth morphotype (pas) showing the production of cellulose only.

581

582 **Supplementary Figure 5.** Zebrafish *in vivo* experiments. Monitoring of zebrafish survival (left panel) and
583 bacterial burden into the zebrafish (right panel) for I1 and R1 (**A**), I2 and R2 (**B**) and I3 and R3 (**C**). MG1655
584 strain was used as a negative control and CFT073 strain as a positive control. Fifteen zebrafish were used for each
585 experiment and no significant statistical difference was identified between the infection strains and their relative
586 relapse strains according to log-rank test for the survival and two-tailed Mann-Whitney test for the bacterial burden
587 comparing initial and relapsing strains.

588

589 **Supplementary Figure 6.** Growth capacities in different conditions. (**A**) Growth curve and related parameters
590 inferred by the R package GrowthCurver (values for which the difference between both strains was significant
591 (two-tailed Mann-Whitney test, p-value < 0.05) are highlighted in bold red) for I1 and R1 strains grown in M63B1
592 minimum medium supplemented with 0.4% pyruvate (n=5 biological replicates). (**B-F**) Growth curve and related
593 parameters inferred by the R package GrowthCurver (values for which the difference between both strains was
594 significant (two-tailed Mann-Whitney test, p-value < 0.05) are highlighted in bold red) for I2 and R2 strains grown
595 in M63B1 minimum medium supplemented with 0.4% lactate (B), Maltose (C), Glucose (D) as well as in LB
596 medium (E) and LB medium supplemented with 500 μ M of 2,2'-Bipyridyl (F) (n=5 biological replicates).

597

598 **Supplemental Figure 7.** Colicin production in R1 relapse strain. **A.** Overlay of R1 strain showing no growth
599 inhibition by a drop of either I1 or R1 supernatant from a culture in presence of Mitomycin C to induce colicin
600 production. **B.** Same experiment with an overlay of I1 strain showing a growth inhibition by a drop of R1

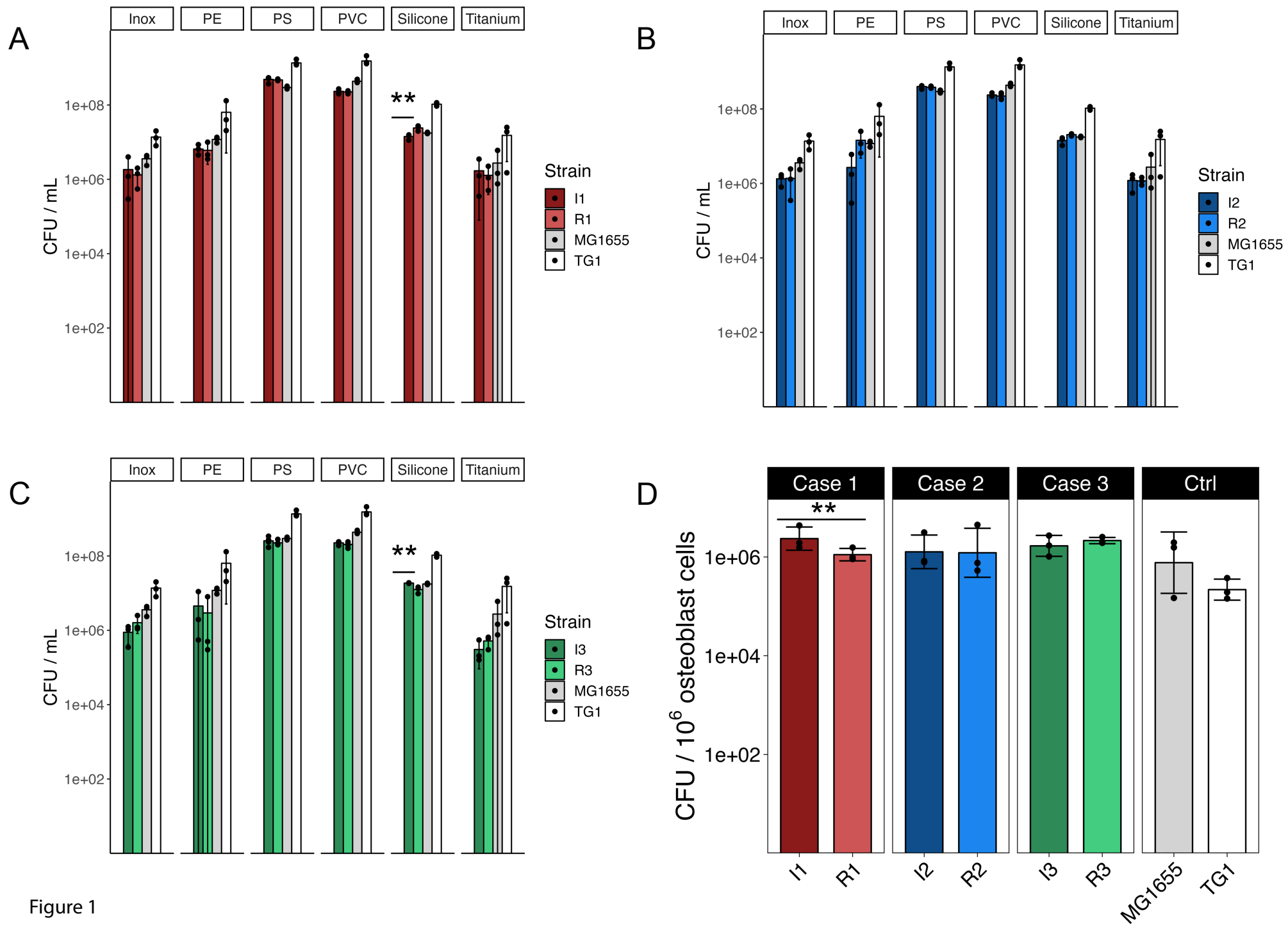
601 supernatant from a culture in presence of Mitomycin C to induce colicin production. In both cases, LB is used as
602 a negative control.
603

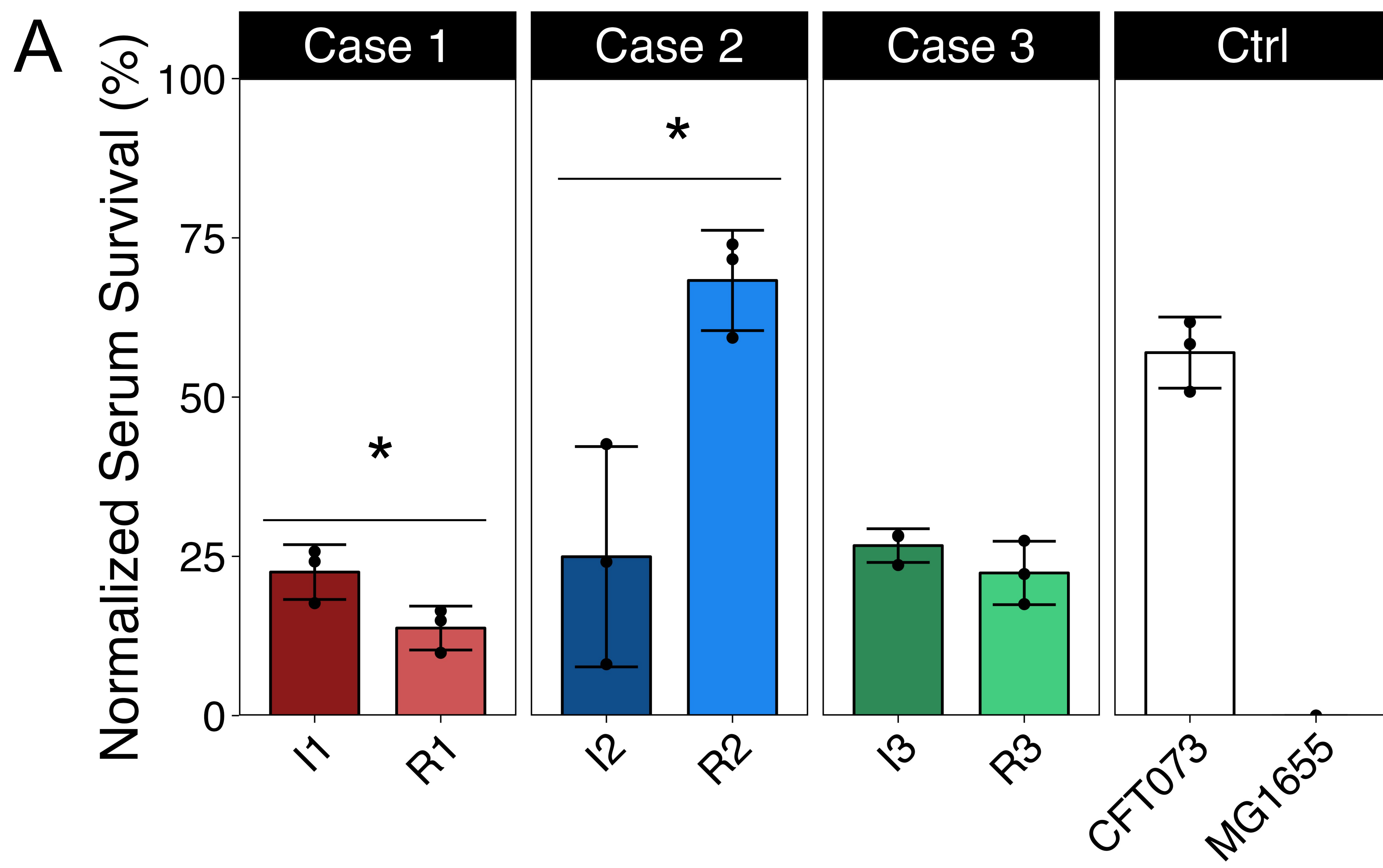
Table 1. Cases and strain information. Table summarizing the type and location of infection as well as type of sample. Each strain was typed *in silico* and, when the infection was polymicrobial, the other species involved are listed. ANI: Average Nucleotide Identity, MLST: MultiLocus Sequencing Typing.

Case N°	Strain	Infection status	Infection site	Sample type	Diagnosis	Time before relapse	Antibiotic susceptibility profile	Phylogroup	Serotype	ANI	MLST #1	MLST #2	fimH	fumC	Infection type	Other pathogens if polymicrobial
1	I1	Initial	Hand	Deep tissue by swab	Septic arthritis 4th finger (no material)	51 days	Penicillinase	B1	O76:H28	99.99%	224	730	FimH3 9	FumC4	Monomicrobial	
	R1	relapse		Deep tissue by swab			Identical								Polymicrobial	<i>Enterococcus faecium</i> , <i>Pseudomonas aeruginosa</i> , <i>Citrobacter freundii</i>
2	I2	Initial	Spine	Washing liquid	Vertebral osteomyelitis (no material)	748 days	ESBL-producing	G	O24:H4	99.99%	117	48	FimH9 7	FumC4 5	Monomicrobial	
	R2	relapse		biopsy			Identical								Polymicrobial	<i>Cutibacterium acnes</i>
3	I3	Initial	Right hip	Bone	Prosthetic joint infection	66 days	ESBL-producing	A	O89/O16 2:H9	99.99%	167	2	/	FumC1 1	Monomicrobial	
	R3	relapse		Deep tissue			Identical								Monomicrobial	

Table 2. SNP analysis in the relapse strains. SNPs, insertions and deletions identified in the evolved strains as compared to their related ancestor.

Strain	Mutation	Supporting reads (Left/Right)	Annotation	Impact	Gene	Description
R1	T→A	104 / 83	intergenic (+168/-221)	Intergenic	uidR → / → uidA	transcriptional repressor/beta-D-glucuronidase
	A→G	191 / 136	N16D (AAC→GAC)	Non-synonymous	yehT →	putative response regulator in two-component system withYehU
	T→C	273 / 260	T46A (ACA→GCA)	Non-synonymous	rppH ←	RNA pyrophosphohydrolase
	T→C	129 / 119	K577E (AAG→GAG)	Non-synonymous	ydiJ ←	putative FAD-linked oxidoreductase
R2	G→A	152 / 189	G335G (GGC→GGT)	Synonymous	COMLMOEC 00122 ←	hypothetical protein
	C→G	56 / 52	V16L (GTA→CTA)	Non-synonymous	insK 3 ←	IS150 transposase B
	C→G	62 / 82	S67T (AGC→ACC)	Non-synonymous	COMLMOEC 04582 ←	IS3 family transposase IS629
	C→T	265 / 166	A131T (GCC→ACC)	Non-synonymous	ebgR ←	transcriptional repressor
	C→T	172 / 171	A197V (GCC→GTC)	Non-synonymous	COMLMOEC 04752 →	IS66 family transposase ISEc47
	T→A	86 / 55	I263N (ATT→AAT)	Non-synonymous	btuC →	vitamin B12 ABC transporter permease
	A→G	164 / 169	intergenic (+1283/-)	Intergenic	COMLMOEC 01687 → / -	hypothetical protein/-
	A→C	66 / 84	intergenic (+352/-)	Intergenic	COMLMOEC 04947 → / -	hypothetical protein/-
R3	G→T	146 / 182	E297D (GAG→GAT)	Non-synonymous	hypT →	hypochlorite-responsive transcription factor
	G→A	73 / 103	D119N (GAT→AAT)	Non-synonymous	insF1 1 →	IS3 transposase B
	C→T	56 / 61	intergenic (+315/-76)	Intergenic	uidR → / → uidA	transcriptional repressor/beta-D-glucuronidase
	Δ1 bp	59 / 71	intergenic (-142/+229)	Intergenic	araD ← / ← araA	L-ribulose-5-phosphate 4-epimerase/L-arabinose isomerase





B

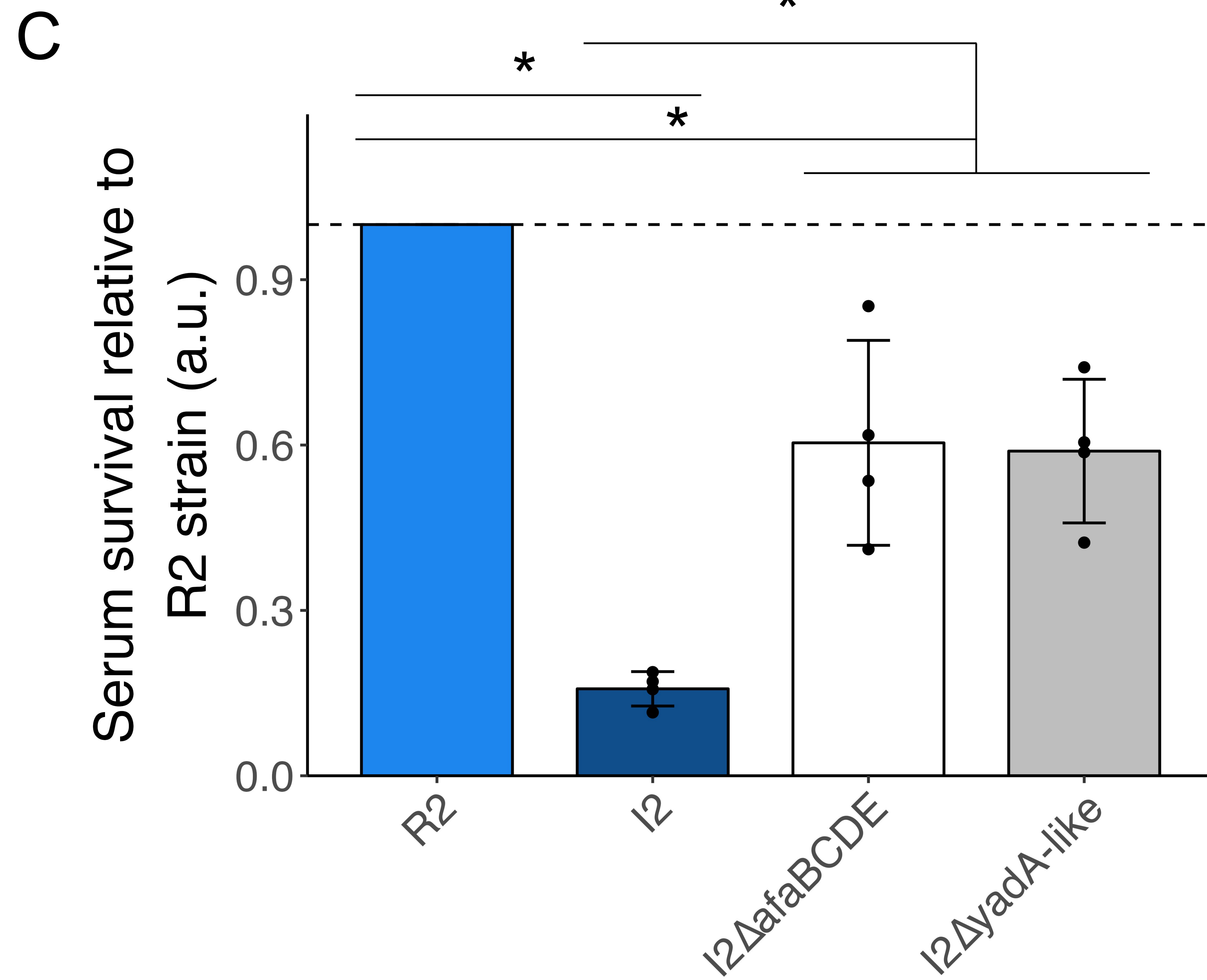


Figure 2

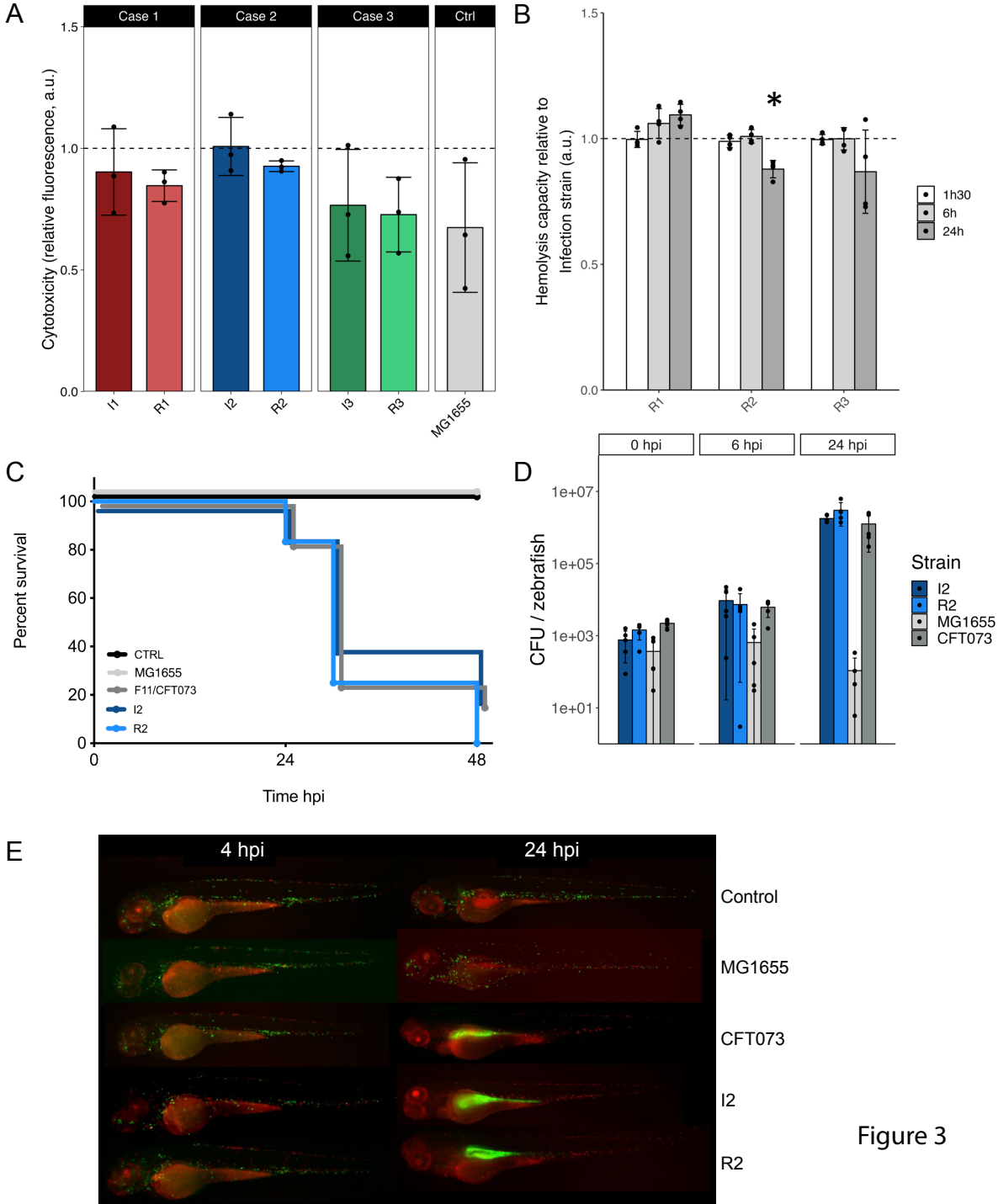


Figure 3

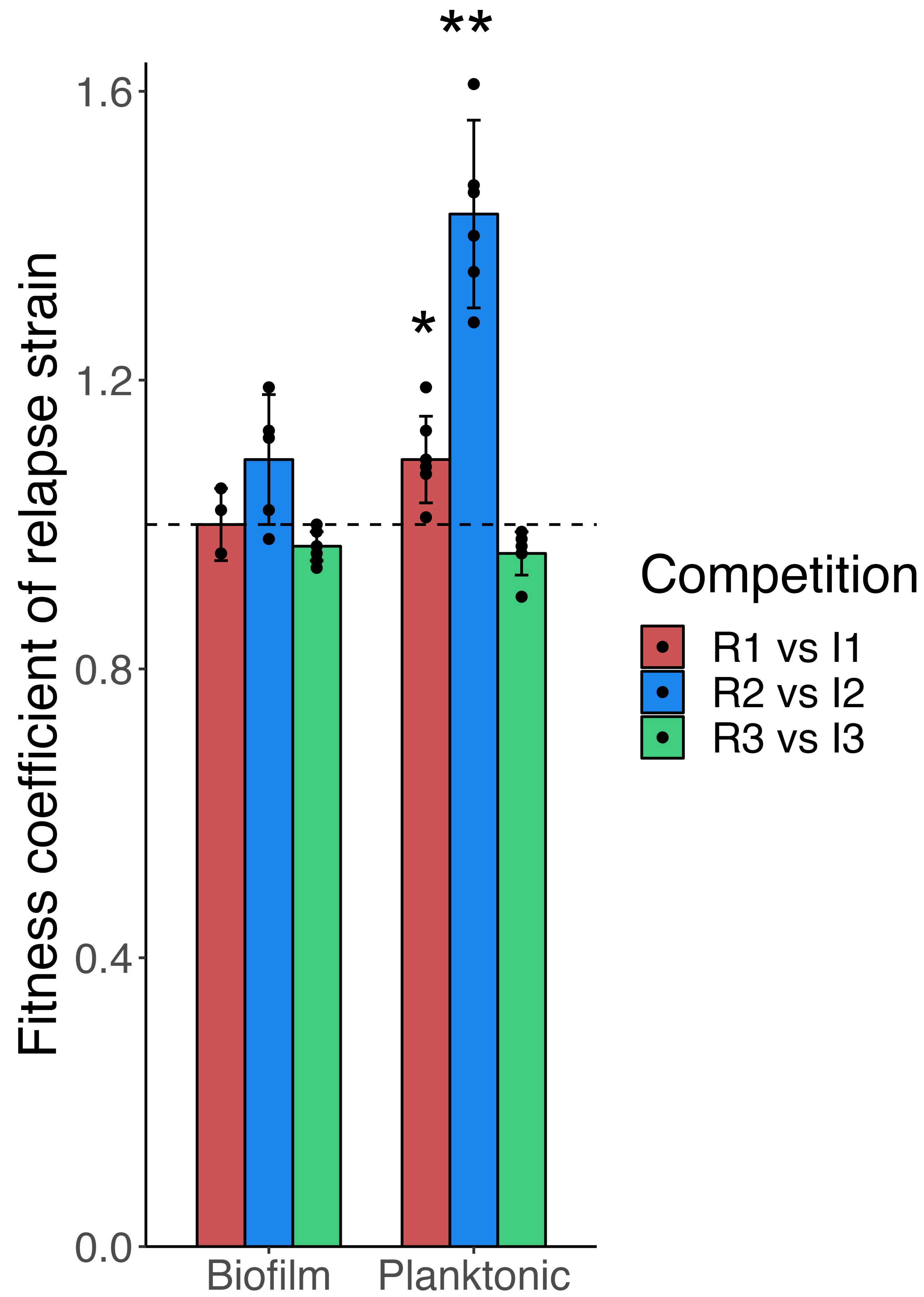
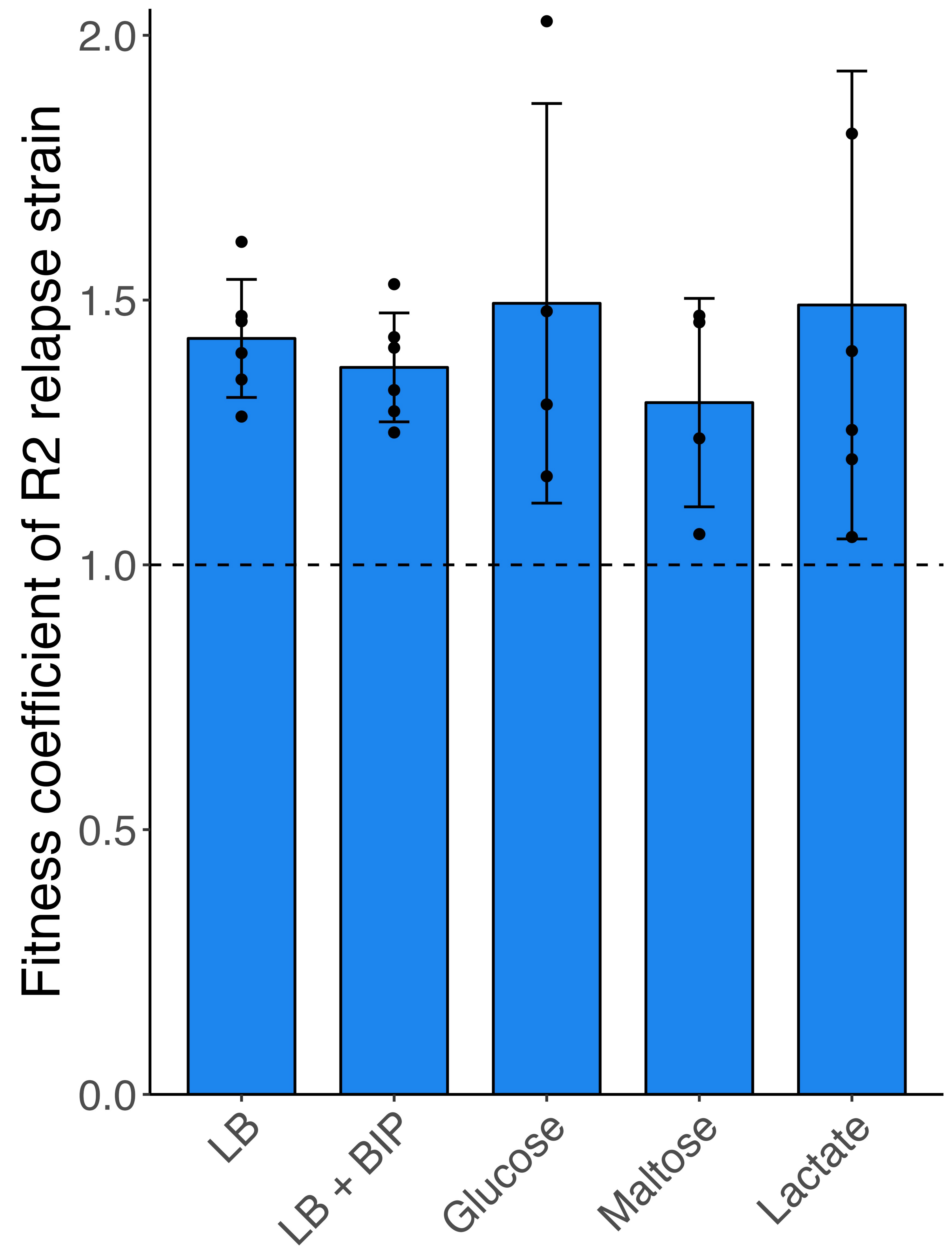
A**B**

Figure 4

Supplementary Table 1. Antibiotic susceptibility profiles. Resistant (R), sensitive (S) or intermediate (I) status of each strain to commonly used antibiotics assessed by disc diffusion assay and following EUCAST breakpoints recommendations.

Antibiotic	I1	R1	I2	R2	I3	R3
Ampicillin	R	R	R	R	R	R
Cefoxitin	S	S	S	S	R	R
Meropenem	S	S	S	S	S	S
Ticarcillin	R	R	R	R	R	R
Ticarcillin + clavulanic acid	S	S	R	R	R	R
Cefepime	S	S	R	R	R	R
Ertapenem	S	S	S	S	S	S
Mecillinam	S	S	S	S	S	S
Cefotaxime	S	S	R	R	R	R
Amoxicillin + clavulanic acid	S	S	R	R	R	R
Ceftazidime	S	S	I	I	R	R
Piperacillin	R	R	R	R	R	R
Tazocillin (Piperacillin + Tazobactam)	S	S	R	R	R	R
Aztreonam	S	S	R	R	R	R
Imipenem	S	S	S	S	S	S
Cefalexin	R	R	S	S	S	S
Tobramycin	S	S	S	S	R	R
Amikacin	S	S	S	S	S	S
Norfloxacin	R	R	S	S	R	R
Moxifloxacin	R	R	S	S	R	R
Ciprofloxacin	R	R	S	S	R	R
Co-trimoxazole (Trimethoprim + Sulfamethoxazole)	R	R	S	S	R	R
Tigecycline	S	S	S	S	S	S
Fosfomycin	S	S	S	S	S	S
Nitrofurantoin	S	S	S	S	S	S
Cefadroxil	S	S	S	S	S	S

Supplementary Table 2. List of all strains used in this study.

Strain	Genotype, phenotype	Relevant information	Source
MG1655	F- lambda- <i>ilvG rfb-50 rph-1</i>	E. coli Genetic Stock Center GCSC#6300	Laboratory collection
TG1	F ⁺ [<i>traD36 proAB + lacIq lacZDM15</i>] <i>supE hsdD5 thi</i> Δ(<i>lac-proAB</i>)	Strong biofilm former	Laboratory collection
CFT073	<i>rpoS</i> -	Human Blood Pyelonephritis (ExPEC) B2 K2	[1]
MG1655 KmFRT- <i>wbbL</i> +	F- lambda- <i>ilvG rfb-50 rph-1 wbbL</i> +, KmR	LPS O-antigen restored	[2]
MG1655 Δ <i>fliE-R</i>	F- lambda- <i>ilvG rfb-50 rph-1</i> Δ <i>fliE-R</i> ::cat, CmR	Non-motile strain	[3]
MG1655 <i>yeaJ/dgcJ</i> inter	F- lambda- <i>ilvG rfb-50 rph-1</i> with and intergenic mutation between <i>yeaJ</i> and <i>dgcJ</i>	Hyper motile strain	Laboratory collection
55989	<i>tetR</i>	rdar morphotype on congored plates	[4]
MG1655 Δ <i>csgD</i>	F- lambda- <i>ilvG rfb-50 rph-1</i> Δ <i>csgD</i> :: <i>aadA7</i> , SpecR	No cellulose production	[5]
MG1655 Δ <i>flu</i>	F- lambda- <i>ilvG rfb-50 rph-1</i> Δ <i>flu</i> :: <i>GB</i> , KmR	No auto-aggregation	[3]
MG1655 PcL <i>flu</i>	F- lambda- <i>ilvG rfb-50 rph-1</i> KmPcL- <i>flu</i> , KmR	Strong auto-aggregation	[6]
MG1655 IATT Zeo-mars	F- lambda- <i>ilvG rfb-50 rph-1</i> IATT- <i>zeo-mars</i> , ZeoR	Source of Zeo-mars cassette at the lambda site	Laboratory collection
MG1655 IATT Zeo-GFP	F- lambda- <i>ilvG rfb-50 rph-1</i> IATT- <i>zeo-gfpmut3</i> , ZeoR	Source of Zeo-GFP cassette at the lambda site	Laboratory collection
I1		WT infection strain from case 1	This study
I1 Δ <i>fliC</i>	Δ <i>fliC</i> :: <i>zeo</i> , ZeoR	WT with <i>fliC</i> gene deleted, non-motile mutant	This study
I1 IATT Zeo-GFP	IATT- <i>zeo-gfpmut3</i> , ZeoR	WT with <i>gfpmut3</i> gene inserted together with a zeocin resistance cassette at the λ att site under λ PR constitutive promoter	This study
I1 IATT Zeo-mars	IATT- <i>zeo-mars</i> , ZeoR	WT with <i>mars</i> gene inserted together with a zeocin resistance cassette at the λ att site under λ PR constitutive promoter	This study
R1		WT relapse strain from case 1	This study
R1 Δ <i>fliC</i>	Δ <i>fliC</i> :: <i>zeo</i> , ZeoR	WT with <i>fliC</i> gene deleted, non-motile mutant	This study
R1 IATT Zeo-GFP	IATT- <i>zeo-gfpmut3</i> , ZeoR	WT with <i>gfpmut3</i> gene inserted together with a zeocin resistance cassette at the λ att site under λ PR constitutive promoter	This study
R1 IATT Zeo-mars	IATT- <i>zeo-mars</i> , ZeoR	WT with <i>mars</i> gene inserted together with a zeocin resistance cassette at the λ att site under λ PR constitutive promoter	This study
I2		WT infection strain from case 2	This study
I2 Δ <i>yadA</i> -like	Δ <i>yadA-like</i> :: <i>zeo</i> , ZeoR	WT with <i>yadA</i> -like coding gene deleted	This study
I2 Δ <i>afaBCDE</i>	Δ <i>afaBCDE</i> :: <i>zeo</i> , ZeoR	WT with genes <i>afaBCDE</i> deleted	This study
R2		WT relapse strain from case 2	This study

I3		WT infection strain from case 3	This study
I3 IATT Zeo-GFP	IATT- <i>zeo-gfpmut3</i> , ZeoR	WT with <i>gfpmut3</i> gene inserted together with a zeocin resistance cassette at the λ att site under λ PR constitutive promoter	This study
I3 IATT Zeo-mars	IATT- <i>zeo-mars</i> , ZeoR	WT with <i>mars</i> gene inserted together with a zeocin resistance cassette at the λ att site under λ PR constitutive promoter	This study
R3		WT relapse strain from case 3	This study
R3 IATT Zeo-GFP	IATT- <i>zeo-gfpmut3</i> , ZeoR	WT with <i>gfpmut3</i> gene inserted together with a zeocin resistance cassette at the λ att site under λ PR constitutive promoter	This study
R3 IATT Zeo-mars	IATT- <i>zeo-mars</i> , ZeoR	WT with <i>mars</i> gene inserted together with a zeocin resistance cassette at the λ att site under λ PR constitutive promoter	This study

1. Mobley HL, Green DM, Trifillis AL, et al. Pyelonephritogenic *Escherichia coli* and killing of cultured human renal proximal tubular epithelial cells: role of hemolysin in some strains. *Infect Immun.* **1990**; 58(5):1281–1289.
2. Rendueles O, Beloin C, Latour-Lambert P, Ghigo J-M. A new biofilm-associated colicin with increased efficiency against biofilm bacteria. *ISME J.* **2014**; 8(6):1275–1288.
3. Francius G, Polyakov P, Merlin J, et al. Bacterial Surface Appendages Strongly Impact Nanomechanical and Electrokinetic Properties of *Escherichia coli* Cells Subjected to Osmotic Stress. *PLOS ONE. Public Library of Science;* **2011**; 6(5):e20066.
4. Mossoro C, Glaziou P, Yassibanda S, et al. Chronic Diarrhea, Hemorrhagic Colitis, and Hemolytic-Uremic Syndrome Associated with HEp-2 Adherent *Escherichia coli* in Adults Infected with Human Immunodeficiency Virus in Bangui, Central African Republic. *J Clin Microbiol.* **2002**; 40(8):3086–3088.
5. Da Re S, Ghigo J-M. A CsgD-independent pathway for cellulose production and biofilm formation in *Escherichia coli*. *J Bacteriol.* **2006**; 188(8):3073–3087.
6. Da Re S, Le Quéré B, Ghigo J-M, Beloin C. Tight modulation of *Escherichia coli* bacterial biofilm formation through controlled expression of adhesion factors. *Appl Environ Microbiol.* **2007**; 73(10):3391–3403.

Supplementary Table 3. List of all primers used in this study.

I1/R1 and I3/R3 IATT Zeo-mars and Zeo-GFP construction and verification	
lambda-ATT. B1.500-3	CCCTGATACTCACCAGGCATCAC
Rev K12-ZeoXFP-IATT	GTCACGCCAAAAGCCAATGC
Fw K12-ZeoXFP-IATT	GCAAGCGCCTCGATTACTGC
lambda-ATT.A1.500-5	CGATGGCGATAATATTCACC
Sanger sequencing-based fitness assessment in I2/R2	
Fw butC	CAGTCGGTTATTGCTGGCT
Rv butC	CACACTTTGAGGCGACATTG
Sanger sequencing-based fitness assessment calibration vs FACS in I1/R1	
Fw yehT	CAACATTTGATGCTGGAGATCG
Rv yehT	GCTTCATCAATTGGCTTCAGC
afaBCDE deletion in I2 and verification	
Fw afaDel size ZeoFRT	GATTCAGCACCTCAGTCAGAC
Fw Del afaB-E ZeoFRT	GGTTGCTATTAACATATTTATAAAAAGGTTTATTTGCCTTCAGGATAAAAATATGTAGGCTGGAGCTGCTTC
Rv Del afaB-E ZeoFRT	CAAAGGGAGCATATAGCCCCCTTCTTTCATCAGGTTAATTTCCAGAATATCCTCCTTAGTTCCTATTCC
Rv afaDel size ZeoFRT	GCGAGGAAGATTTCTCTTGTAGG
yadA-like deletion in I2 and verification	
Fw Size yadA Del	TGCTGCTTAAATTGAGGTCGC
Fw Del yadA ZeoFRT	CTCCGCACTACCGTTCTGGCTGGTGAAGTAATAACAGGAGAATAACAGATGTAGGCTGGAGCTGCTTC
Rv Del yadA ZeoFRT	TCCCGGGGACTCCCCCGGGACAGATATCTAAATCCTGACCTGGAATATCCTCCTTAGTTCCTATTCCG
Rev Size yadA Del	CGATTCCAAATATCTCTCGCAGG
fliC deletion in I1/I2 and verification	
Fw fliC	GGCATGATTATCCGTTTCTGC
Fw CmFRT ForFliC	TTGGCGTTGCCGTCAGTCTCAGTTAATCAGGTTACAACGAATTTAAATGGCGCGCCTTAC
Rev CmFRT ForFliC	CCCAATACGTAATCAACGACTTGCAATATAGGATAACGAATCGCCTACCTGTGACGGAAG
Rev fliC	TCCAGCGATGAAATACTTGC

Supplementary Table 4. Genome assembly characteristics and metrics. Table showing data related to sequencing data, genome assembly quality and description.

Case N°	Strain	Read number	Genome coverage average (SD)	Genome size (base pairs)	Contig number	GenBank Assembly accession	N50	L50	GC%	Complete BUSCOs	CDS	Potential plasmid gain	Potential plasmid loss	CDS gained	CDS lost	Mutation per base per year
1	I1	5,761,693	336 (170)	5,124,061	105	JARFPI000000000	193,231	9	50.81	440/440	4,740	n.a				
	R1	7,181,866	418 (200)	5,133,911	97	JARFPF000000000	193,231	10	50.81	440/440	4,754	1	0	35	19	8.38x10 ⁻⁶
2	I2	5,591,377	316 (144)	5,300,036	130	JARFPH000000000	194,756	9	50.69	440/440	4,914	n.a				
	R2	5,271,593	308 (132)	5,122,701	111	JARFPE000000000	208,913	8	50.77	440/440	4,775	0	1	31	170	1.11x10 ⁻⁶
3	I3	5,594,583	341 (138)	4,900,901	99	JARFPG000000000	137,146	11	50.79	440/440	4,511	n.a				
	R3	4,748,473	289 (118)	4,904,465	98	JARFPD000000000	137,146	11	50.79	440/440	4,518	1	0	14	9	4.51x10 ⁻⁶

1 **Analysis of *in-patient* evolution of *Escherichia coli* reveals potential**

2 **links to relapse of bone and joint infections**

3 Stanislas THIRIET-RUPERT¹, Jérôme JOSSE², David PEREZ-PASCUAL¹, Jason TASSE², Camille
4 ANDRE², Lélia ABAD², David LEBEAUX¹, Jean-Marc GHIGO¹, Frédéric LAURENT² and Christophe
5 BELOIN*¹

7 **Supplementary Results**

9 **Auto-aggregation and RDAR morphotypes on congo-red plates**

10 Once settled and in order to colonize host environment, *E. coli* has a very large arsenal
11 of adhesion factors. Some of them are known to promote biofilm formation via homotypic self-
12 interactions also called auto-aggregation [1,2], which could help facing stressful conditions
13 such as the one encountered in vivo. The production of other factors such as curli and cellulose
14 are also of critical importance [3,4] and induces a red, dry, and rough (rdar) morphotype when
15 the strain is plated on LB agar supplemented with congo red. However, none of these
16 phenotypes showed any difference when comparing the relapse strains to their related ancestors,
17 all strains even having very weak auto-aggregation capacities (supplementary Figure S1.A.B).
18 Still, inter-couple differences were identified on congo red plates: (i) the first couple had a pink,
19 dry and rough (pdar) morphotype characterizing the production of curli only; (ii) the second
20 couple had a smooth and white morphotype (saw) showing the absence of matrix component
21 production; and (iii) the third couple had pink and smooth morphotype (pas) showing the
22 production of cellulose only.

23

24

25 Temperature specific motility

26 Because strain motility could have an impact on the virulence and biofilm forming
27 capacity, we tested each strain on LB plates containing low amount of agar at 30°C and 37°C.
28 Motility is sometime more visible at lower temperature. The second and third couple strains
29 had no motility (supplemental Figure S2). However, an interesting phenotype was observed for
30 the first couple. Although both strains were motile, we observed some flares characteristic of
31 the apparition of a hypermotile clone within the population and spreading faster than the others
32 (supplementary Figure S2). Because this phenomenon was only observed at 37°C, we
33 investigated the importance of temperature by testing all combinations between the temperature
34 used to produce the liquid inoculum and the temperature the motility plates were incubated at.
35 We showed that regardless of the temperature used for the liquid culture, incubating the motility
36 plate at 30°C always showed a homogenous motile phenotype, while plates inoculated with the
37 same samples but incubated at 37°C always showed the apparition of flares (supplementary
38 Figure S3.A.B.C). This suggests that the process is reversible and a westernblot analysis using
39 anti-flagellin antibodies showed that both strains only produce flagellin at 30°C, even when
40 pre-incubated at 37°C (supplementary Figure S3.D). Therefore, it seems that some mutants
41 arise on plates incubated at 37°C in which the capacity to swim is restored.

42 Because the apparition of hyper-motile mutants suggests a high mutation rate, we used a Luria-
43 Delbruck fluctuation test to evaluate it. Resistance to Rifampicin was used in these assays and
44 strains I1 and R1 had a mutation rate of 1.58×10^{-6} and 1.39×10^{-6} mutations per generation,
45 respectively, which was by far higher than MG1655 as estimated in the same assay (3.21×10^{-8}
46 mutations per generation), which was in accordance with what was reported in the literature for
47 the same kind of experiment in *Escherichia coli* K12 MG1655 (see Figure 4 from [5]). Such a
48 high mutation rate could be the result of a loss of function in genes responsible for DNA repair

49 such as the mut genes. However, no such loss of function could be identified in the genome of
50 the clinical strains.

51

52 **Genome analysis**

53 To confirm that in each case the relapse strain was the same as the initial infection one all of
54 them were subjected to whole genome sequencing (WGS) and the genomic characteristics
55 analyzed. Each case showed that both strains were indeed the same according to their
56 phylogroup, serotype, ANI, MLST, FimH and FumC type (Table 1).

57 If the genome assemblies and completeness were good for all strain, the first notable difference
58 identified between the infection strains and their related relapse strains was the number of
59 gained and lost genes (supplementary Table S4). Some of these genes corresponded to a whole
60 contig that matched known plasmids, therefore constituting potential gain or loss of plasmids
61 (supplementary Data 2). R1 relapse strain gained a potential plasmid bearing Colicin E1
62 production genes, R2 relapse strain lost a potential plasmid containing hypothetical protein
63 coding genes and R3 relapse strain gained a potential plasmid containing a Toxin/Antitoxin
64 (T/A) system (HigA/HigB). However, PacBio sequencing technology would be required to rule
65 out what proportion of the gene gains and losses was plasmid-mediated or the consequence of
66 big chromosomal insertions or deletions.

67

68

69

70 **References**

71

- 72 1. Kjaergaard K, Schembri MA, Ramos C, Molin S, Klemm P. Antigen 43 facilitates
73 formation of multispecies biofilms. *Environ Microbiol.* **2000**; 2(6):695–702.
- 74 2. Danese PN, Pratt LA, Dove SL, Kolter R. The outer membrane protein, antigen 43,
75 mediates cell-to-cell interactions within *Escherichia coli* biofilms. *Mol Microbiol.* **2000**;
76 37(2):424–432.
- 77 3. Gualdi L, Tagliabue L, Bertagnoli S, Ieranò T, De Castro C, Landini P. Cellulose
78 modulates biofilm formation by counteracting curli-mediated colonization of solid surfaces

- 79 in *Escherichia coli*. *Microbiol Read Engl.* **2008**; 154(Pt 7):2017–2024.
- 80 4. Beloin C, Roux A, Ghigo JM. *Escherichia coli* biofilms. *Curr Top Microbiol Immunol.*
- 81 **2008**; 322:249–289.
- 82 5. Krašovec R, Richards H, Gifford DR, et al. Opposing effects of final population density
- 83 and stress on *Escherichia coli* mutation rate. *ISME J. Nature Publishing Group*; **2018**;
- 84 12(12):2981–2987.
- 85

Analysis of *in-patient* evolution of *Escherichia coli* reveals potential links to relapse of bone and joint infections

Stanislas THIRIET-RUPERT¹, Jérôme JOSSE², David PEREZ-PASCUAL¹, Jason TASSE², Camille ANDRE², Lélia ABAD², David LEBEAUX¹, Jean-Marc GHIGO¹, Frédéric LAURENT² and Christophe BELOIN*¹

Supplementary Methods

RDAR morphotypes on congo red plates

Two microliters of an overnight culture grown at 37°C in LB medium under shaking (with appropriate antibiotics when needed) was spotted onto LB plates (without NaCl) containing 0.004% Congo Red and 0.002% brilliant blue. The spotted drops were allowed to dry, and the plates were incubated for 24 and 48 hours at 30 or 37°C.

Auto-aggregation curves

Cultures were grown as described above, and 1mL aliquots were diluted to an OD₆₀₀=3 in spent medium in order to prevent growth during experiment. Once diluted, all the tubes were vortexed and an aliquot of 50µL was removed, 1cm below the top of the culture, and diluted with 50µL of LB prior to OD₆₀₀ measurement. The bacteria were then left to aggregate at RT for 6 hours, while repeating the measurement of OD₆₀₀ 1 cm below the top of the culture every hour.

Motility assay

Bacterial strains were incubated overnight in LB at the specified temperature with shaking. Then, 2µL of overnight culture was inoculated in swim plates (1% tryptone, 0.25% NaCl, 0.25%

Eiken agar) for 24 hours at 37°C or 30°C. After incubation, motility was assessed by checking the spread of inoculated bacteria into the motility agar.

Biofilm microtiter plate CV assay

Bacteria were grown overnight and diluted to an $OD_{600}=0.05$ and 100 μ L were inoculated in technical triplicates in either a polystyrene Greiner 96-well plate or a polyvinyl chloride (PVC) round bottom 96-well plates (Corning). After 24 hours at 37°C, the biofilms were washed twice by slowly pipetting PBS in each well to remove non-attached bacteria. Biofilms were then resuspended in 100 μ L of PBS by vigorous pipetting and viable cell count was determined by serial dilution and plating.

Growth capacity

An overnight culture of each strain was diluted to OD_{600} of 0.05 in appropriate medium. One hundred and fifty μ L aliquots were inoculated in a 96-well plate. The plates were then incubated in a TECAN Infinite M200 Pro spectrophotometer (Männedorf, Switzerland) for 18 hours at 37°C with shaking of 2mm amplitude. The absorbance of each culture at 600 nm was measured every 15 minutes. Growth curve of each strain was measured using 3 biological replicates.

The R package GrowthCurver [1] was used to infer growth-related parameters for all strains based on the growth curves.

Cytotoxicity on HeLa cells

The cytotoxic activity on HeLa cells of each bacterial strain was analyzed using the CellTiter-Blue® Cell Viability Assay (Promega) following manufacturer instructions. HeLa cells were cultured in complete DMEM (Life Technologies) supplemented with 10% FCS, 100 μ g/mL penicillin/streptomycin, and 10mM L-glutamine at 37C in the presence of 5% CO₂. For

infection assay, the HeLa cells were incubated in 96-well plates to reach a density of 10^5 cells. Each bacterial strain was grown overnight in LB as described above and then washed twice in PBS, to finally adjust the bacterial density to 10^7 CFU/ml. These bacterial suspensions were used to infect HeLa cells by adding 100 μ l of each to the wells (MOI of 1:100). The plates were incubated at 37C in the presence of 5% CO₂ for 2 hours. Non-infected cells and the *E. coli* K12 MG1655 strain were used as negative controls. At the end of the incubation time, the wells were washed twice with PBS to remove the bacteria. Then, 100 μ l of PBS and 20 μ l of CellTiter Blue solution were added to each well and incubated for 2 hours at 37C and 5% CO₂. The cell viability was measured by recording the fluorescence at 560/590 nm using a Tecan Infinite-M200-Pro spectrophotometer.

Hemolysis capacity

The hemolytic activity of each strain was assessed by mixing defibrinated horse blood (bioMerieux, Marcy l'Etoile, France) diluted to 10% (v/v) in PBS with a total of 10^8 bacteria in PBS. After 2, 6 and 24 hours of incubation at 37°C with shaking, the non-lysed erythrocytes were pelleted by centrifugation at 1500g for 2 minutes, and the OD₄₅₀ of 150 μ L of the cell-free supernatants was measured. The hemolytic capacity of each adapted strain was expressed relatively to its ancestral strain.

Luria-Delbrück fluctuation test

The mutation rate of I1, R1 and MG1655 strains were assessed using Luria-Delbrück fluctuation test [2]. For each strain, 60 wells of a 96-well plate were inoculated with ~100 bacteria in a volume of 200 μ L in LB before incubating the plate in a humid atmosphere for 16h at 37°C under 450 rpm shaking. After the cultures had reach saturation, ten randomly chosen wells were sampled to evaluate the number of total bacteria per well by serial dilution and plating. The

total volume of each of the remaining 50 wells was plated on individual LB plates supplemented with 2xMICs of Rifampicin to evaluate the number of resistant cells. The colony count was then used to assess the mutation rate by a maximum likelihood approach available on FALCOR web tool (<https://lianglab.brocku.ca/FALCOR/>) [3].

Zebrafish care and maintenance

Homozygous Tg(*mfap4::mCherryF*) (ump6Tg) [4] Tg(*mpx::GFP*)ⁱ¹¹⁴ [5] double transgenic fishes were raised in Institut Pasteur zebrafish facility. Eggs were obtained by natural spawning, bleached according to standard protocols, and then kept in Petri dishes containing Volvic spring water and, from 24 hpf onwards, 0.003% 1-phenyl-2-thiourea (PTU) (Sigma-Aldrich) was added to prevent pigmentation. Embryos as well as infected larvae were reared at 28°C. Larvae were anesthetized with 200 µg/mL of buffered tricaine methane sulfonate (MS-222, Sigma-Aldrich) during the injection procedure as well as during *in vivo* imaging and processing for bacterial burden evaluation.

Evaluation of the bacterial burden in infected zebrafish larvae

Infected zebrafish larvae were collected at 0, 24, and 48 hpi and lysed for the evaluation of the bacterial burden as previously described [6]. Each larva was placed in an individual 1.5 mL Eppendorf tube and anesthetized with tricaine (200 µg/mL), washed with 1 mL of sterile water, and placed in 150 µL of sterile water. Larvae were then homogenized using a pestle motor mixer (Argos). Each sample was transferred to an individual well of a 96-well plate and 10 X serial dilutions were performed. For CFU enumeration and to assess plasmid stability

throughout the infection kinetics, serial dilutions of lysates were plated on MacConkey agar plates supplemented with lactose plates and incubated overnight at 37°C.

Visualization of macrophages and neutrophils in infected zebrafish larvae

Visualization of macrophages and neutrophils on living transgenic reporter larvae was performed upon infection as we previously described [7]. Briefly, bright field, GFP, and RFP images of whole living anesthetized larvae were taken using a Leica Macrofluor Z16 APOA (zoom 16:1) macroscope equipped with a Leica PlanApo 2.0 X lens, and a Photometrics CoolSNAP HQ2 camera. Images were captured using the Metavue software version 7.5.6.0 (MDS Analytical Technologies). After capture of images, larvae were washed and transferred to a new 24-well plate filled with 1mL of fresh water per well and incubated at 28 °C.

Overlay assay colicin

I1 and R1 strains were incubated overnight at 37°C with shaking in LB supplemented with 40 µg/mL of Mitomycin C to induce colicin production. After incubation an overlay of each strain was realized by pouring 6mL of overnight culture diluted at OD₆₀₀=0.1 onto a LB agar plate and removing it after 5 minutes. Plates were allowed to dry and a 10µL drop of supernatant from each strain culture filtered using 0.22µm filter was spotted on each overlay. After an additional drying step, the plates were incubated overnight at 37°C and the production of colicin was identified by the presence of an inhibition halo when the filtered supernatant was spotted. LB was spotted as a negative control and filtered supernatant from both strains was also spotted on overlay of each of them to be sur that the growth inhibition halo was not due to the presence of Mitomycin C in the filtered supernatant.

Whole genome sequencing and analysis

Prior to genome extraction, a single colony of each strain was inoculated in LB medium for over-day culture till the OD₆₀₀ reached around 1.0 (ca. 5.0x10⁸ bacteria/mL). The bacterial cells were collected from 2mL of the culture and the genomic DNA was extracted using the Wizard Genomic DNA Purification Kit (Promega). Samples were then sent to Novogene Co., Ltd to be sequenced using Illumina NovaSeq 6000. Sequencing reads were pre-processed to remove low-quality or artefactual bases. We used fqCleaner v.0.5.0, a mini workflow implemented in Galaxy [8] (<https://galaxy.pasteur.fr/>) to process fastq files (quality trimming, duplicate and artifact filters). Draft genomes were de novo assembled using SPAdes version 3.15.0 [9] and annotated using Proka version 1.14.5 [10]. Assembly quality and completeness was assessed using BUSCO version v5.4.2 [11]. The assembled genomes were used for in silico typing and characterization of each strain using tools available at the Centre of Genomic Epidemiology (<http://www.genomicepidemiology.org/>): Resfinder [12–14], MLST [14,15], SerotypeFinder [16], CHtyper [17]. Phylogroups were identified in silico using ClermonTyping tool [18]. In order to identify contigs corresponding to potential plasmids, each contig was used as a query in a BLASTn search using blast+ version 2.2.31 [14] against the COMPASS database [19]. Only hits with 95% identity and 80% coverage of both the query and the hit sequence were kept.

Alternative strategy for competition assays of I2 vs R2 strains

Even if we managed to introduce GFP and RFP (mars) tags in I2 and R2 strains, the related proteins were likely not expressed since the fluorescence was impossible to detect. Consequently we used an alternative strategy to assess the proportion of each strain. We took advantage of the fact that both strains can be differentiated by SNPs. An aliquot of the mixed culture containing 10⁷ cells was pelleted and resuspended in 50µL of sterile distilled water and boiled for 10 minutes. Specific primers were then used to amplify a SNP locus in butC gene

and Sanger sequence the resulting DNA fragment. Then, using QSVanalyzer [20], we calculated the frequency of each mutation and used it to compute the relative fitness as above. To assess the accuracy of this strategy, we used strains I1 and R1 tagged with different fluorescent marker to compare strains proportions as assessed by FACS and QSVanalyzer and found a very good correspondance between both methods (Supplemenatry Figure S6), therefore validating our alternative approach.

Western blot analysis and LPS gel

The heat extracted proteins from 1mL at $OD_{600}=1$ culture were suspended in 1×Laemmli buffer and incubated for 5 min at 95°C. The protein extracts were run on Mini-PROTEAN TGX Stain-Free precast Gels (BioRad) in 1×TGX buffer and then transferred to nitrocellulose membrane using a Trans-Blot® Turbo Transfer System (BioRad). Blocking was performed in a 5% solution of dry milk and 0.05% Tween 1×PBS (1×PBST) overnight at 4°C with agitation. The membranes were then incubated in 1×PBST with either rabbit anti-flagellin (kindly given by Dr Monica Rolando) or mouse anti-RNAP α (BioLegend #663104) diluted 1/5,000 for 1h at room temperature with agitation. Membranes were washed in 1×PBST and then incubated with the secondary antibody (anti-rabbit or anti-mouse IgG conjugated with horse radish peroxidase at 1:10 000, Promega). The membranes were then revealed using an ECL kit (GE Healthcare) and the iBright™ CL1500 system (Thermofisher).

LPS analysis on Mini-PROTEAN TGX Stain-Free precast Gels (BioRad) were realized as in [21].

References

1. Sprouffske K, Wagner A. Growthcurver: an R package for obtaining interpretable metrics from microbial growth curves. *BMC Bioinformatics*. **2016**; 17(1):172.
2. Luria SE, Delbrück M. Mutations of Bacteria from Virus Sensitivity to Virus Resistance. *Genetics*. **1943**; 28(6):491–511.
3. Hall BM, Ma C-X, Liang P, Singh KK. Fluctuation Analysis CalculatOR: a web tool for the determination of mutation rate using Luria–Delbrück fluctuation analysis. *Bioinformatics*. **2009**; 25(12):1564–1565.
4. Phan QT, Sipka T, Gonzalez C, Levraud J-P, Lutfalla G, Nguyen-Chi M. Neutrophils use superoxide to control bacterial infection at a distance. *PLOS Pathog. Public Library of Science*; **2018**; 14(7):e1007157.
5. Renshaw SA, Loynes CA, Trushell DMI, Elworthy S, Ingham PW, Whyte MKB. A transgenic zebrafish model of neutrophilic inflammation. *Blood*. **2006**; 108(13):3976–3978.
6. Boucontet L, Passoni G, Thiry V, et al. A Model of Superinfection of Virus-Infected Zebrafish Larvae: Increased Susceptibility to Bacteria Associated With Neutrophil Death. *Front Immunol*. **2018**; 9:1084.
7. Mostowy S, Boucontet L, Mazon Moya MJ, et al. The zebrafish as a new model for the in vivo study of *Shigella flexneri* interaction with phagocytes and bacterial autophagy. *PLoS Pathog*. **2013**; 9(9):e1003588.
8. Mareuil F, Doppelt-Azeroual O, Ménager H. <p>A public Galaxy platform at Pasteur used as an execution engine for web services</p>. F1000Research [Internet]. **2017** [cited 2023 Feb 7]; 6. Available from: <https://f1000research.com/posters/6-1030>
9. Bankevich A, Nurk S, Antipov D, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol J Comput Mol Cell Biol*. **2012**; 19(5):455–477.
10. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinforma Oxf Engl*. **2014**; 30(14):2068–2069.
11. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinforma Oxf Engl*. **2015**; 31(19):3210–3212.
12. Bortolaia V, Kaas RS, Ruppe E, et al. ResFinder 4.0 for predictions of phenotypes from genotypes. *J Antimicrob Chemother*. **2020**; 75(12):3491–3500.
13. Zankari E, Allesøe R, Joensen KG, Cavaco LM, Lund O, Aarestrup FM. PointFinder: a novel web tool for WGS-based detection of antimicrobial resistance associated with chromosomal point mutations in bacterial pathogens. *J Antimicrob Chemother*. **2017**; 72(10):2764–2768.
14. Camacho C, Coulouris G, Avagyan V, et al. BLAST+: architecture and applications. *BMC Bioinformatics*. **2009**; 10:421.
15. Larsen MV, Cosentino S, Rasmussen S, et al. Multilocus sequence typing of total-genome-sequenced bacteria. *J Clin Microbiol*. **2012**; 50(4):1355–1361.
16. Joensen KG, Tetzschner AMM, Iguchi A, Aarestrup FM, Scheutz F. Rapid and Easy In Silico Serotyping of *Escherichia coli* Isolates by Use of Whole-Genome Sequencing Data. *J Clin Microbiol*. **2015**; 53(8):2410–2426.
17. L R, Tb J, F H, et al. CHTyper, a Web Tool for Subtyping of Extraintestinal Pathogenic *Escherichia coli* Based on the *fumC* and *fimH* Alleles. *J Clin Microbiol* [Internet]. *J Clin Microbiol*; **2018** [cited 2023 Feb 22]; 56(4). Available from:

<https://pubmed.ncbi.nlm.nih.gov/29436420/>

18. Beghain J, Bridier-Nahmias A, Le Nagard H, Denamur E, Clermont O. ClermonTyping: an easy-to-use and accurate in silico method for Escherichia genus strain phylotyping. *Microb Genomics*. **2018**; 4(7):e000192.
19. Douarre P-E, Mallet L, Radomski N, Felten A, Mistou M-Y. Analysis of COMPASS, a New Comprehensive Plasmid Database Revealed Prevalence of Multireplicon and Extensive Diversity of IncF Plasmids. *Front Microbiol*. **2020**; 11:483.
20. Carr IM, Robinson JI, Dimitriou R, Markham AF, Morgan AW, Bonthron DT. Inferring relative proportions of DNA variants from sequencing electropherograms. *Bioinforma Oxf Engl*. **2009**; 25(24):3244–3250.
21. Szczesny M, Beloin C, Ghigo J-M. Increased Osmolarity in Biofilm Triggers RcsB-Dependent Lipid A Palmitoylation in Escherichia coli. *mBio*. **2018**; 9(4):e01415-18.