Fitness effects of CTX-M-15-encoding IncF plasmids on their native *Escherichia coli* ST131 *H30*Rx hosts

Jana Palkovicova[1,2], Iva Kutilova[1], Javier DelaFuente[3], Adam Valcek[1,4,5], Matej Medvecky[1,6], Ivana Jamborova[1], Ibrahim Bitar[7], Minh Duy Phan[8], Alvaro San Millan[9], Monika Dolejska[1,2,7,10*]

[1]*Central European Institute of Technology, University of Veterinary Sciences Brno, Brno, Czech Republic*

[2]*Department of Biology and Wildlife Diseases, University of Veterinary Sciences Brno, Brno, Czech Republic*

[3]*Department of Microbiology, Hospital Universitario Ramon y Cajal (IRYCIS), Madrid, Spain*

[4]*Microbial Resistance and Drug Discovery, VIB-VUB Center for Structural Biology, VIB, Flanders Institute for Biotechnology, Brussels, Belgium*

[5]*Structural Biology Brussels, Vrije Universiteit Brussel (VUB), Brussels, Belgium*

[6]*Institute of Biodiversity, Animal Health and Comparative Medicine, University of Glasgow, Glasgow, United Kingdom*

[7]*Biomedical Center, Faculty of Medicine in Pilsen, Charles University, Pilsen, Czech Republic*

[8]*Australian Infectious Diseases Research Centre, School of Chemistry and Molecular Biosciences, The University of Queensland, Brisbane, Australia*

[9]*Centro Nacional de Biotecnologia-CSIC, Madrid, Spain*

[10]*Department of Clinical Microbiology and Immunology, Institute of Laboratory Medicine, The University Hospital Brno, Czech Republic*

*Corresponding author. Tel: +420 541 562 643 ; E-mail: monika.dolejska@gmail.com

Running title: Fitness effects of IncF plasmids on *E. coli* ST131 *H30*Rx

Abstract

Objectives

The objective of this study was to investigate effects of large CTX-M-15-encoding IncF plasmids on the fitness of their native *E. coli* ST131 *H30*Rx hosts in order to understand possible plasmid-host coevolution.

Methods

We selected five *E. coli* ST131 *H30*Rx strains of diverse origin, each carrying a multireplicon IncF plasmid encoding the gene $bla_{\text{CTX-M-15}}$. The plasmid was eliminated from each isolate by displacement using an incompatible plasmid vector pMDP5_cureEC958. Whole-genome sequencing (WGS) was performed to obtain complete chromosome and plasmid sequences of wild-type isolates and to detect chromosomal mutations in plasmid-free strains. Competition assays were conducted to determine the relative fitness of plasmid-free clones compared to the corresponding wild-type isolates.

Results

We were able to successfully eliminate the IncF plasmids from all of the wild-type strains using the curing vector pMDP5_cureEC958. The chromosomes of plasmid-free clones contained zero to six point mutations. Plasmid-free strains of three isolates showed no significant difference in relative fitness compared to the corresponding plasmid-free strains. In the two remaining isolates, the plasmids produced a small but significant fitness cost.

Conclusion

We conclude that IncF plasmids produce moderate fitness effects in their *E. coli* ST131 *H30*Rx hosts. This fitness compatibility is likely to promote the maintenance of antibiotic resistance in this worrisome *E. coli* lineage.

Introduction

Extraintestinal pathogenic *E. coli* (ExPEC) represent a huge public health burden[1] as these strains are a common source of numerous diseases, from mild to life-threatening infections, such as urinary tract or blood-stream associated infections, bacteraemia and meningitis[2]. Globally emerging multi-drug resistance, especially to antimicrobials of clinical importance, such as fluoroquinolones and cephalosporins, is of special concern in this species in recent years. Due to the resistance, only a few or even no therapeutic options are left for a treatment of infections caused by these bacteria[2,3].

Resistance to clinically important antimicrobials among ExPEC strains was scarce before 2000. Since then, the number of ExPEC isolates resistant to fluoroquinolones and cephalosporins has been increasing exponentially[4]. Highly virulent and multi-drug resistant *E. coli* ST131 is one of the clinically most important ExPEC strains due to its worldwide predominance. Even though *E. coli* ST131 was mainly associated with human infections[5], recent findings brought a disturbing evidence of its dissemination among companion animals, poultry, livestock, food products, wildlife and environment, including wastewater treatment plant effluents[6,7].

A study from 2015[8] analysed the evolution of *E. coli* ST131 lineage. Based on investigations from the mid-2000s, it is apparent that previous consumption and misuse of antimicrobials is linked with the emergence of resistant pathogens[8,9]. Antimicrobials, such as fluoroquinolones and cephalosporins, were often used for the treatment of human infections which resulted in the emergence of fluoroquinolone resistant subclone *E. coli* ST131 *H30*R. Furthermore, by acquisition of an incompatibility F (IncF) plasmid carrying an ESBL gene $bla_{\text{CTX-M-15}}$, a distinct fluoroquinolone resistant extended-spectrum beta-lactamase (ESBL) producing *E. coli* ST131 subclone *H30*Rx emerged[9,10]. Compensatory mutations in combination with virulence determinants and resistance to clinically important antimicrobials allowed the subclone *H30*Rx

to outcompete other sublineages of *E. coli* ST131 and in the late 2000s became a globally disseminated and most prevalent ExPEC subclone[8,11].

IncF are complex epidemic resistance plasmids composed of more than one plasmid replicon. CTX-M-15-encoding IncF plasmids present in the subclone *H30*Rx usually harbour two plasmid replicons with plasmid multilocus sequence type (pMLST) F2:A1:B-. These plasmids typically harbour other resistance genes apart from $bla_{\text{CTX-M-15}}$, such as $bla_{\text{TEM-1}}$, $bla_{\text{OXA-1}}$, *aac(6′)-Ib-cr*, *catB4*, *aadA5*, *mph*(A), *dfrA7*, *tet*(A) and *sul1*. Additionally, the narrow-host IncF plasmids encode partitioning and addiction systems to ensure their maintenance[6,8,12]. Carriage of such large plasmids providing selective advantage for a bacterial host via additional virulence and antibiotic determinants, usually imposes a fitness cost to its host[13]. On the other hand, a previous study suggested that *E. coli* ST131 *H30*Rx is adapted to large IncF plasmids[8].

In this study, we analyse plasmid-host interactions between this intriguing *E. coli* subclone and its plasmids. We aimed to estimate the fitness impact of the large F2:A1:B- IncF plasmids, previously recognised as a foundation of *H30*Rx sublineage emergence, on its native host. Five representatives of the *E. coli* ST131 *H30*Rx were selected for elimination of IncF plasmid using the curing vector pMDP5_cureEC958. Plasmid fitness effects were subsequently calculated using competition assays between the plasmid-carrying and plasmid-free isogenic clones.

Materials and Methods

Bacterial strains

Five *E. coli* ST131 *H30*Rx strains were selected out of the collection of 169 *E. coli* ST131 isolates of diverse origin from several geographic regions[6]. Selected isolates were of human (n = 3), environmental (n = 1) and companion animal (n = 1) origin. Each strain carried a large IncF plasmid harbouring $bla_{\text{CTX-M-15}}$. Additional information on isolates obtained during our previous study[6] is presented in Table S1.

Bacterial strains were routinely grown on Luria-Bertani agar (LBA; Sigma-Aldrich, Saint Louis, USA) supplemented with cefotaxime (2 mg/L) at 37 °C overnight if not specified otherwise. Competition assays were performed in Luria-Bertani broth (LBB; Becton Dickinson, MD, USA).

Construction of curing vector

The curing vector pMDP5_cureEC958 was designed based on the pCURE2 plasmid[14]. The backbone vector pMDP5 was assembled from three vectors including pUC19 (ori and MCS, 126-1480 nt), pKD3 (chloramphenicol resistant gene, 85-980 nt) and pCURE2 (*sacB* gene, 209-2076 nt). Subsequently, a curing fragment (containing RepFIIA, RepFIA, *ccdA*, *sok*, *pemI*, and *vagC*) was synthesised and cloned into pMDP5, creating the curing vector pMDP5_cureEC958 (Figure 1). All cloning and DNA synthesis steps were performed by Epoch Life Science (Texas, USA). The sequence of pMDP5_cureEC958 was deposited to GenBank under accession number MZ723317.

Plasmid curing

Plasmid-free variants were created using a curing method described by Hale and colleagues[14] with the curing vector pMDP5_cureEC958. The plasmid curing vector was introduced into the wild-type strains by electroporation (1.8 kV, 25 μFar, 200 Ω) using Gene Pulser Xcell™ electroporation system (Bio-Rad Laboratories Inc., California, USA) as described before[15]. Cultures harbouring both, the wild-type plasmid and the designed construct, were selected on LBA supplemented chloramphenicol (30 mg/L) and the presence of $bla_{\text{CTX-M}}$ and *catA1* genes was verified by PCR[16,17]. Transformants were cultivated on LBA containing chloramphenicol (30 mg/L) in order to eliminate the IncF plasmid. After successful removal of the wild-type plasmid, verified by the same PCR, the pMDP5_cureEC958 was eliminated on a non-selective LBA plates supplemented with 5% sucrose. Three to four plasmid-free clones of each isolate were selected.

Selected plasmid-free clones were sequenced on MiSeq platform (Illumina) as described below to investigate possible single nucleotide mutations (SNPs) on chromosomes. Subsequently, one plasmid-free clone of isolates without chromosomal mutations was selected to reintroduce the corresponding wild-type IncF plasmid as a control of experiment. Plasmid DNA was extracted from wild-type strains using Genopure Plasmid Midi Kit (Roche Diagnostics GmbH, Mannheim, Germany). Plasmids were reintroduced by electroporation and their presence was confirmed by PCR assays for gene *bla*$_{\text{CTX-M}}$ and for FII and multiplex (FIA, FIB, FIC) IncF replicons[16,18].

Whole-genome sequencing

Wild-type isolates were subjected for short- and long-read sequencing. Additionally, plasmid-free strains and plasmid-free strains with reintroduced wild-type IncF plasmid were selected for short-read sequencing. Genomic DNA for short-read sequencing was extracted using NucleoSpin® Tissue kit (Macherey-Nagel, GmbH & Co. KG, Germany), library was prepared by Nextera® XT Library Preparation kit (Illumina, San Diego, CA, USA) and sequenced using 2x250 bp paired-end sequencing on MiSeq (Illumina) platform. NucleoSpin® Microbial DNA kit (Macherey-Nagel) was used for the extraction of genomic DNA aimed for long-read sequencing. Libraries were constructed using SMRTbell Express Template Prep Kit 2.0 (Pacific Biosciences, PacBio, USA) followed by single molecule real-time (SMRT) sequencing on Sequel I Platform (PacBio).

Data analysis

Raw reads acquired by Illumina sequencing were trimmed using Trimmomatic v0.39[19] to remove adaptor residues and discard low quality read regions ($Q \leq 20$). SPAdes v3.13.1[20] with the "--careful" configuration was used to obtain *de novo* assemblies. Center for Genomic Epidemiology tools (PlasmidFinder v2.1, pMLST v2.0, ResFinder v4.0) were used to verify the presence of plasmid replicons and genes intermediating antibiotic resistance

(https://cge.cbs.dtu.dk/services/). HGAP4 in SMRT Link v.6 (PacBio) was used to obtain polished long reads in fastq format. Hybrid assembly of trimmed short and long reads was performed using Unicycler v0.4.8[21] and corrected with Pilon v1.23[22] in order to reconstruct chromosome and plasmid sequences of wild-type isolates. Complete circular sequences of plasmids were manually annotated using Geneious v7.1.9 (Biomatters, Auckland, New Zealand) in compliance with annotation form of previous studies[23].

Comparative genomics

Phylogenetic relatedness of wild-type isolates was estimated. Prokka v1.14.1[24] was used to predict open reading frames of isolates assemblies and their core genome was aligned using Roary v3.12.0[25]. Subsequently, the alignment was used to generate phylogenetic tree in RAxML v8.2.11[26] under GTR+GAMMA model supported by 1,000 bootstraps. A nucleotide similarity between the isolates was estimated using the core genome alignment in snp-dists v0.6.3 (https://github.com/tseemann/snp-dists) considering the number of SNPs. The phylogenetic tree was visualized in iTOL v5.7[27].

BLAST (Basic Local Alignment Search Tool) of NCBI (National Center for Biotechnology Information, MD, USA) was used to find and download a plasmid sequence with the highest coverage and identity from GenBank. Genetic content of the IncF plasmids was compared using BLAST Ring Image Generator (BRIG) v0.95[28] and Clinker v0.0.13[29]. Presence and nomenclature of specific insertion sequences in IncF plasmids were confirmed using ISfinder database[30] and toxin-antitoxin systems were verified using Conserved Domain Database[31].

Comparison of the wild-type isolates to the corresponding plasmid-free strains and to the plasmid-free strains with reintroduced plasmids was made to verify the identity of the strains as well as the identity of the wild-type and reintroduced plasmids. Corresponding sequences were aligned using algorithm BWA-MEM v0.7.17[32] and manually checked in Geneious v7.1.9.

Single nucleotide polymorphism analysis

Corrected short reads of plasmid-free variants were mapped to the corresponding wild-type *de novo* assemblies using Bowtie2 v2.3.5[33]. Variant calling was performed by VarScan v2.4.4[34] based on the coverage of mapped reads. Minimum variant frequency was 80% and called variants were manually checked in Geneious v7.1.9. Corresponding wild-type reads were mapped and analysed as well in order to normalize obtained results.

Transferability of IncF plasmids

Wild-type *E. coli* ST131 *H30*Rx isolates were used as donors while laboratory strain *E. coli* TOP10 (Invitrogen Life Technologies, Carlsbad, CA, USA) and corresponding plasmid-free variants of studied isolates were used as recipients for the estimation of conjugation ability of IncF plasmids using filter mating assays based on a previous study[35]. Conjugations were conducted in technical triplicates and biological duplicates.

Relative fitness measurements

Competition assays were performed to compare the relative fitness of the wild-type strains and their plasmid-free clones using flow cytometry as previously described[36]. Only the plasmid-free strains without chromosomal mutations were selected for fitness experiments. A small non-mobilisable pBGC plasmid (MT702881)[37] producing green fluorescent protein (GFP) was transformed to the wild-type strains by electroporation[15]. Transformants were selected on LBA plates containing cefotaxime (2 mg/L) and chloramphenicol (30 mg/L) and subsequently verified by PCR assays for genes $bla_{\text{CTX-M}}$ and *gfp*[16,37].

Two competition assays, each consisting of six biological replicates, were performed for each isolate. Competitions were performed between GFP-tagged wild-type strains and their untagged plasmid-free variants while each included a competition between tagged and untagged wild-types for normalisation. Overnight cultures were mixed in ratio 1:1 and diluted 1:400 for the

competition. GFP expression in the wild-type strains resulting in fluorescence was induced by incubation in 0.9% sodium chloride solution containing 0.5% L-arabinose for 1.5 hours. Plasmid-free and wild-type populations were competed at 37 °C for 22 hours shaking at 225 rpm. Initial and final proportions were measured on NovoCyte (ACEA Biosciences) flow cytometer recording 50,000 events of each mixture. Relative fitness of plasmid-free clones compared to corresponding wild-types was estimated using the formula:

$$w = \frac{\ln(\frac{N_{final,GFP-}}{N_{initial,GFP-}})}{\ln(\frac{N_{final,GFP+}}{N_{initial,GFP+}})}$$

where w represents relative fitness, $N_{initial,\ GFP-}$ and $N_{final,\ GFP-}$ are initial and final values of untagged population and $N_{initial,\ GFP+}$ and $N_{final,\ GFP+}$ are proportions of GFP-marked population before and after competition. Relative fitness of plasmid-free clones was statistically processed using Student's T-test where relative fitness with p value < 0.05 was evaluated as statistically significant. Obtained data was normalized using a competition between the tagged and untagged wild-type populations in order to capture relative fitness of plasmid-free clones in comparison to the corresponding (untagged) wild-type isolates. Competitions between wild-type strains and constructed plasmid-free strains with reintroduced wild-type IncF plasmid were performed as a control.

Results

Strain and plasmid features

In order to compare fitness effects of the plasmid on its native host, *E. coli* ST131 *H30*Rx isolates carrying a single large $bla_{\text{CTX-M-15}}$ harbouring IncF plasmid were selected. The phylogenetic analysis of the five selected strains (Table S1) was based on the core-genome alignment of 4,803 genes and showed 78-440 SNPs differences (Figure S1).

All five plasmids contained two IncF replicons (RepFIA, RepFII) with pMLST formula F2:A1:B-, slightly varied in size and antibiotic resistance genes content (Figure S2). All plasmids provided multi-drug resistance profile, encoding genes for ESBL as well as for other antimicrobials and contained insertion sequences, mostly IS*26* (Table 1). The *ccdAB* and *pemIK* toxin-antitoxin systems were encoded in all plasmids within replicons RepFIA and RepFII, respectively. Each plasmid harboured two copies of the addiction system *vapBC*. All but one plasmid (pM45) harboured *hok/sok* system and plasmids of human isolates encoded *parDE* system.

All IncF plasmids in our study have proved to be non-conjugative. Genetic analysis of transfer (*tra*) regions showed diverse defects in all plasmids likely resulting in their non-functionality (Figure 2). The *tra* region of pOV24 was disrupted in two parts by the $bla_{\text{TEM-1}}$ gene, usually transposed within a composite mobile genetic element, but 3' flanking sequence IS*26* was disrupted by IS*Ecp1* element. Furthermore, the *traC* gene was truncated by another IS*26* and genes *traW* and *traU*, encoding proteins for pilus assembly and DNA transfer, were missing. The *tra* region of pM70 was disrupted by composite mobile element containing $bla_{\text{TEM-1}}$ gene flanked by IS*15DI* and IS*26* similarly as in pOV24. Moreover, part of the second half of the *tra* genes was translocated 29.5 kb from the first part of the *tra* region and truncated by IS*26* in *traN* gene. The first part of the *tra* region, including *traJ*, serving as a transcriptional activator of the *tra* region, was completely missing in plasmids pM24 and pM45. Plasmid pDog168 lacked most of this part of the region as well with an exception of the *traM* gene.

Plasmid curing

To study plasmid-associated fitness effects on their native host, all five wild-type strains were cured of the naturally occurring IncF plasmids. A curing vector pMDP5_cureEC958 was designed for the generation of plasmid-free variants, harbouring selective genes (*catA1* and

241 *sacB*), replicons RepFIA, RepFII, and antitoxins of the addiction systems encoded by wild-type
242 plasmids (Figure 1).

243 Four plasmid-free clones per each of four isolates (M24, M45, M70 and Dog168) and all three
244 grown plasmid-free clones of the isolate OV24 were selected for further analyses. Sequence
245 comparison of plasmid-free clones to the corresponding wild-type isolates discovered zero to
246 six chromosomal mutations. Mutations occurred in 52.6% of plasmid-free strains (in 10 out of
247 19). Nearly all mutations (92%, 23/25) occurred in protein coding sequences, only two of them
248 were located in intergenic regions. Additionally, most of the mutations in coding sequences
249 (82.6%, 19/23) were non-synonymous. No mutations occurred only in one of the plasmid-free
250 clones of the OV24 (1/3) and M70 (1/4) isolates, in two Dog168 (2/4), two M24 (2/4) and three
251 M45 (3/4) plasmid-free clones. Detailed list of genetic changes in plasmid-free clones is in
252 Table S2.

253 Four of all plasmid-free strains with the reintroduced wild-type plasmid (40%, 4/10) harboured
254 one mutation. Although, only one of them was non-synonymous.

255 **Fitness impact of IncF plasmids on their native host**

256 To maintain isogenic conditions in the competition experiment, only the plasmid-free clones
257 with no mutations in their chromosome were selected for the measurement of the IncF-
258 associated fitness effects on their native host. Relative fitness of plasmid-free clones was
259 estimated in comparison to the corresponding wild-type isolates considering a background
260 fitness of wild-types as 1 (Table 2). Analysis of two competition assays both consisting of six
261 biological replicates for each combination plasmid-free clone and wild-type revealed non-
262 significant fitness effects ($p > 0.05$) of IncF plasmids in three isolates (Dog168, OV24, M45).
263 Moderate increase ($p < 0.05$) of relative fitness was observed in plasmid-free strains of two
264 isolates (M24 and M70) as visualised in Figure 3, revealing a small plasmid fitness cost.

Relative fitness differences of the two M24 plasmid-free clones were statistically significant (with *p*-value 1.83 x 10^{-6} and 4.07 x 10^{-4}, respectively), however, the increase in relative fitness was moderate (1.04 and 1.02). A similar result was observed for the M70 plasmid-free clone, which showed a moderate but significant increase in relative fitness compared to the wild type clone (w = 1.03, p = 1.5 x 10^{-3}).

Discussion

Even though the *E. coli* ST131 has attracted much attention in the last few years due to its predominance in ExPEC infections, its success is still not fully elucidated. Plasmids play a key role in bacterial survival under a selection pressure by providing virulence and antibiotic resistance genes, but it is known that they often impose a fitness burden to their hosts[38–40]. However, the fitness cost of a specific plasmid can differ in various hosts[41].

In our study, we investigated the fitness effects of strictly clade-specific[12] F2:A1:B- IncF plasmids, previously recognised as a source of *H30*Rx subclone emergence[8,11], on these native hosts. Five native *E. coli* ST131 *H30*Rx hosts were cured of the large $bla_{\text{CTX-M-15}}$-harbouring IncF plasmids and then competed against their corresponding wild-type strain to calculate the plasmid fitness impact.

Plasmid curing

Plasmid curing was previously recognised as a best way to study effects of a plasmid on a bacterial population[38]. Traditionally used methods for plasmid curing involved bacterial growth in the presence of a chemical factor[42]. Although these techniques were widely used, efficiency of plasmid curing was low and promoted high accumulation of unwanted mutations[43]. Recently, the CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) curing is being used more often. The CRISPR curing is based on targeting a specific conserved sequence within plasmids and subsequent plasmid cleavage. This technique is efficient and do not generate new

mutations, however, there are still plasmids which avoid targeting[44]. As IncF plasmids are known for their complexity and many combinations of the replicon alleles among the plasmids, the pCURE method[14] with the vector pMDP5_cureEC958 was used in this study. The method is based on incompatibility of the targeted plasmid with an introduced curing vector. Even though many plasmids overcome incompatibility elimination, it is possible to design a construct with more replicons and antitoxin genes to successfully cure a strain of a plasmid. The approach proved efficient in the past[14,44] as well as in our study. On the other hand, we observed point mutation accumulation in 52.6% (10 out of 19) of the plasmid-free clones. Therefore, WGS after pCURE plasmid curing followed by genomic comparison of plasmid-free clones and their corresponding wild-type isolates is necessary to exclude the clones harbouring mutations and to obtain reliable relative fitness results.

Fitness effects of the F2:A1:B- plasmids

The fitness cost imposed by plasmids is influenced by several factors. It was observed before, that solely the size of plasmids plays no role in their fitness cost, however, maintaining the plasmid-encoded accessory genes can produce an energetic burden. The increasing number of accessory genes, such as antimicrobial resistance genes, correlates with the higher fitness cost[45].

In order to evaluate the fitness impact of the F2:A1:B- plasmids providing their hosts with multi-drug resistance, we estimated relative fitness of the plasmid-free strains in comparison to the corresponding wild-type isolates. Competitions followed by measurement using flow cytometer were used for this purpose. This method is considered much more sensitive than growth curve analysis allowing to detect even subtle differences in relative fitness[13]. We demonstrated that fitness impact of these IncF plasmids on their native hosts in non-selective conditions was moderate to negligible in human as well as in animal and environmental isolates which is in concordance with previous studies on F2:A1:B- plasmids[46–48]. A study of Shin and Ko[47] focused on effects of CTX-M-14 and CTX-M-15 IncF plasmids from human clinical

314 isolates and their impact on a laboratory *E. coli* J53 strain. Based on the growth curve analysis
315 of the transconjugants, the authors proposed that strains harbouring $bla_{\text{CTX-M}}$ on IncF plasmids
316 were as competitive as the naive host. Ranjan and his colleagues[48] studied competitiveness of
317 *E. coli* ST131 harbouring IncF plasmids and their plasmid-free variants against colicin-
318 producing *E. coli* ST10. The authors observed similar fitness ($p > 0.05$) of the wild-type *E. coli*
319 ST131 and their plasmid-free variants based on growth curves analysis. However, growth
320 curves in this experiment were conducted on selective plates which could affect the fitness of
321 the strains as some antibiotic resistance genes are genetically linked to each other and co-
322 selected. Therefore, the supplementation of plates with antibiotics could create a selection
323 pressure where carriage of plasmids would be more beneficial for the strain survival[49].
324 Mahérault and colleagues[46] provided the investigation on two human clinical $bla_{\text{CTX-M-15}}$-
325 harbouring F2:A1:B- plasmids and their impact on an *E. coli* J53-2. Even though the fitness
326 cost of an IncF plasmid occurred initially after conjugation, the authors observed that this fitness
327 cost alleviated and a transconjugant carrying the IncF plasmid proved more competitive than a
328 transconjugant harbouring an IncC plasmid.

329 Previous studies also indicated that a functional conjugation system could have a negative
330 impact on a bacterial fitness and plasmids use several different ways to supress the conjugative
331 transfer[39]. Even though the silencing of the conjugative transfer results in a decrease of
332 horizontal spread of the plasmids, the vertical spread is supported by a fitness cost reduction[50].
333 The *tra* region responsible for conjugative transfer investigated during our study was
334 incomplete, the missing genes and length of the missing sequences differed among studied
335 plasmids. The rearrangements resulted in the non-functionality of the *tra* region of each
336 plasmid. Conjugation malfunction together with plasmid addiction systems ensure the vertical
337 transmission of the IncF plasmids. Additionally, it was pointed out that the initial fitness cost
338 is reduced over time. This phenomenon was observed in long-term evolution experiments, even

339 though the fitness of the plasmid-bearing strains was lower than of those without plasmids in
340 many cases[39].

341 We demonstrated that fitness impact of these IncF plasmids on their native hosts in non-
342 selective conditions was moderate to negligible among phylogenetically unrelated isolates of
343 diverse origin. Our results, combined with previous findings, strongly suggest that *E. coli*
344 *H30*Rx and IncF plasmids form successful associations promoting the world-wide
345 dissemination of this ExPEC lineage.

346 **Acknowledgements**

347 We thank Jarka Lausova and Dana Cervinkova from University of Veterinary Sciences Brno
348 (Czech Republic) and Aida Alonso del Valle and Carmen de la Vega from Hospital
349 Universitario Ramon y Cajal (Spain) for their work in the laboratory. Furthermore, we are
350 grateful to Kristina Nesporova from University of Veterinary Sciences Brno and Jaroslav
351 Hrabak from the University Hospital Pilsen (Czech Republic) for the help with short- and long-
352 read sequencing, respectively. We also thank Chris Thomas for the inspiration and help in
353 designing the curing plasmid.

354 **Funding**

355 This work was supported by the Czech Science Foundation (18-23532S), Ministry of Health of
356 the Czech Republic (NU20J-05-00033) and by CEITEC 2020 - Central European Institute of
357 Technology (LQ1601) from the Czech Ministry of Education, Youth and Sports within the
358 National Programme for Sustainability II. Jana Palkovicova was supported by Internal Grant
359 Agency of University of Veterinary Sciences Brno (grants 214/2020/FVHE and
360 215/2020/FVHE).

361 **Transparency declarations**

362 None to declare.

References

1. World Health Organization. *Antimicrobial Resistance: Global Report on Surveillance*. Geneva, Switzerland; 2014.

2. Schwaber MJ, Navon-Venezia S, Kaye KS, Ben-Ami R, Schwartz D, Carmeli Y. Clinical and Economic Impact of Bacteremia with Extended-Spectrum-β-Lactamase-Producing *Enterobacteriaceae*. *Antimicrob Agents Chemother* 2006; **50**: 1257–62.

3. Mathers AJ, Peirano G, Pitout JDD. *Escherichia coli* ST131: The Quintessential Example of an International Multiresistant High-Risk Clone. In: *Advances in Applied Microbiology*. Vol 90. Elsevier Ltd, 2015; 109–54.

4. Pitout JDD. Extraintestinal pathogenic *Escherichia coli*: an update on antimicrobial resistance, laboratory diagnosis and treatment. *Expert Rev Anti Infect Ther* 2012; **10**: 1165–76.

5. Poolman JT, Wacker M. Extraintestinal Pathogenic *Escherichia coli*, a Common Human Pathogen: Challenges for Vaccine Development and Progress in the Field. *J Infect Dis* 2016; **213**: 6–13.

6. Jamborova I, Johnston BD, Papousek I, *et al.* Extensive Genetic Commonality among Wildlife, Wastewater, Community, and Nosocomial Isolates of *Escherichia coli* Sequence Type 131 (*H30*R1 and *H30*Rx Subclones) That Carry $bla_{\text{CTX-M-27}}$ or $bla_{\text{CTX-M-15}}$. *Antimicrob Agents Chemother* 2018; **62**: e00519-18.

7. Kidsley AK, White RT, Beatson SA, *et al.* Companion Animals Are Spillover Hosts of the Multidrug-Resistant Human Extraintestinal *Escherichia coli* Pandemic Clones ST131 and ST1193. *Front Microbiol* 2020; **11**: 1968.

8. Mathers AJ, Peirano G, Pitout JDD. The role of epidemic resistance plasmids and international high- risk clones in the spread of multidrug-resistant *Enterobacteriaceae*. *Clin Microbiol Rev* 2015; **28**: 565–91.

387 9. Ben Zakour NL, Alsheikh-Hussain AS, Ashcroft MM, *et al.* Sequential acquisition of
388 virulence and fluoroquinolone resistance has shaped the evolution of *Escherichia coli* ST131.
389 *mBio* 2016; **7**: e00347-16.

390 10. Riley LW. Pandemic lineages of extraintestinal pathogenic *Escherichia coli*. *Clin Microbiol*
391 *Infect* 2014; **20**: 380–90.

392 11. Price LB, Johnson JR, Aziz M, *et al.* The epidemic of Extended-Spectrum-β-Lactamase-
393 Producing *Escherichia coli* ST131 Is Driven by a Single Highly Pathogenic Subclone, *H30*-Rx.
394 *mBio* 2013; **4**: e00377-13.

395 12. Kondratyeva K, Salmon-Divon M, Navon-Venezia S. Meta-analysis of Pandemic
396 *Escherichia coli* ST131 Plasmidome Proves Restricted Plasmid-clade Associations. *Sci Rep*
397 2020; **10**: 36.

398 13. San Millan A. Evolution of Plasmid-Mediated Antibiotic Resistance in the Clinical Context.
399 *Trends Microbiol* 2018; **26**: 978–85.

400 14. Hale L, Lazos O, Haines AS, Thomas CM. An efficient stress-free strategy to displace stable
401 bacterial plasmids. *Biotechniques* 2010; **48**: 223–8.

402 15. Wu N, Matand K, Kebede B, Acquaah G, Williams S. Enhancing DNA
403 electrotransformation efficiency in *Escherichia coli* DH10B electrocompetent cells. *Electron J*
404 *Biotechnol* 2010; **13**.

405 16. Lewis II JS, Herrera M, Wickes B, Patterson JE, Jorgensen JH. First report of the Emergence
406 of CTX-M-type Extended-Spectrum β-Lactamases (ESBLs) as the Predominant ESBL Isolated
407 in a U.S. Health Care System. *Antimicrob Agents Chemother* 2007; **51**: 4015–21.

408 17. Faldynova M, Pravcova M, Sisak F, *et al.* Evolution of Antibiotic Resistance in *Salmonella*
409 *enterica* Serovar Typhimurium Strains Isolated in the Czech Republic between 1984 and 2002.
410 *Antimicrob Agents Chemother* 2003; **47**: 2002–5.

411 18. Carattoli A. Resistance plasmid families in *Enterobacteriaceae*. *Antimicrob Agents*
412 *Chemother* 2009; **53**: 2227–38.

413 19. Bolger AM, Lohse M, Usadel B. Trimmomatic: A flexible trimmer for Illumina sequence
414 data. *Bioinformatics* 2014; **30**: 2114–20.

415 20. Bankevich A, Nurk S, Antipov D, *et al.* SPAdes: A new genome assembly algorithm and
416 its applications to single-cell sequencing. *J Comput Biol* 2012; **19**: 455–77.

417 21. Wick RR, Judd LM, Gorrie CL, Holt KE. Unicycler: Resolving bacterial genome assemblies
418 from short and long sequencing reads. *PLoS Comput Biol* 2017; **13**: e1005595.

419 22. Walker BJ, Abeel T, Shea T, *et al.* Pilon: An integrated Tool for Comprehensive Microbial
420 Variant Detection and Genome Assembly Improvement. *PLoS One* 2014; **9**: e112963.

421 23. Papagiannitsis CC, Kutilova I, Medvecky M, Hrabak J, Dolejska M. Characterization of the
422 Complete Nucleotide Sequences of IncA/C2 Plasmids Carrying In809-Like Integrons from
423 *Enterobacteriaceae* Isolates of Wildlife Origin. *Antimicrob Agents Chemother* 2017; **61**:
424 e01093-17.

425 24. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 2014; **30**: 2068–
426 9.

427 25. Page AJ, Cummins CA, Hunt M, *et al.* Roary: rapid large-scale prokaryote pan genome
428 analysis. *Bioinformatics* 2015; **31**: 3691–3.

429 26. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large
430 phylogenies. *Bioinformatics* 2014; **30**: 1312–3.

431 27. Letunic I, Bork P. Interactive Tree of Life (iTOL) v4: Recent updates and new
432 developments. *Nucleic Acids Res* 2019; **47**: 256–9.

433 28. Alikhan NF, Petty NK, Ben Zakour NL, Beatson SA. BLAST Ring Image Generator

(BRIG): Simple prokaryote genome comparisons. *BMC Genomics* 2011; **12**: 402.

29. Gilchrist CLM, Chooi Y-H. clinker & clustermap.js: automatic generation of gene cluster comparison figures. *Bioinformatics* 2021: btab007.

30. Siguier P, Perochon J, Lestrade L, Mahillon J, Chandler M. ISfinder: the reference centre for bacterial insertion sequences. *Nucleic Acids Res* 2006; **34**: D32–6.

31. Marchler-Bauer A, Derbyshire MK, Gonzales NR, *et al.* CDD: NCBI's conserved domain database. *Nucleic Acids Res* 2015; **43**: D222–6.

32. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. 2013: 1–3.

33. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 2009; **10**: R25.

34. Koboldt DC, Chen K, Wylie T, *et al.* VarScan: variant detection in massively parallel sequencing of individual and pooled samples. *Bioinformatics* 2009; **25**: 2283–5.

35. Top E, Vanrolleghem P, Mergeay M, Verstraete W. Determination of the Mechanism of Retrotransfer by Mechanistic Mathematical Modeling. *J Bacteriol* 1992; **174**: 5953–60.

36. DelaFuente J, Rodriguez-Beltran J, San Millan A. Methods to Study Fitness and Compensatory Adaptation in Plasmid-Carrying Bacteria. In: *Horizontal Gene Transfer. Methods in Molecular Biology*. Vol 2075. New York: Humana, 2020; 371–82.

37. Alonso-del Valle A, León-Sampedro R, Rodríguez-Beltrán J, *et al.* Variability of plasmid fitness effects contributes to plasmid persistence in bacterial communities. *Nat Commun* 2021; **12**.

38. San Millan A, Maclean RC. Fitness Costs of Plasmids: a Limit to Plasmid Transmission. *Microbiol Spectr* 2017; **5**: MTBP-0016-2017.

457 39. Porse A, Schønning K, Munck C, Sommer MOA. Survival and Evolution of a Large
458 Multidrug Resistance Plasmid in New Clinical Bacterial Hosts. *Mol Biol Evol* 2016; **33**: 2860–
459 73.

460 40. San Millan A, Peña-Miller R, Toll-Riera M, *et al.* Positive selection and compensatory
461 adaptation interact to stabilize non-transmissible plasmids. *Nat Commun* 2014; **5**: 5208.

462 41. Johnson TJ, Singer RS, Isaacson RE, *et al. In Vivo* Transmission of an IncA/C Plasmid in
463 *Escherichia coli* Depends on Tetracycline Concentration, and Acquisition of the Plasmid
464 Results in a Variable Cost of Fitness. *Appl Environ Microbiol* 2015; **81**: 3561–70.

465 42. Zaman M, Pasha M, Akhter M. Plasmid Curing of *Escherichia coli* Cells with Ethidium
466 Bromide, Sodium Dodecyl Sulfate and Acridine Orange. *Bangladesh J Microbiol* 2010; **27**:
467 28–31.

468 43. Lauritsen I, Porse A, Sommer MOA, Nørholm MHH. A versatile one-step CRISPR-Cas9
469 based approach to plasmid-curing. *Microb Cell Fact* 2017; **16**: 135.

470 44. Li Y, Lin Z, Huang C, *et al.* Metabolic engineering of *Escherichia coli* using CRISPR-Cas9
471 meditated genome editing. *Metab Eng* 2015; **31**: 13–21.

472 45. Vogwill T, Maclean RC. The genetic basis of the fitness costs of antimicrobial resistance:
473 a meta-analysis approach. *Evol Appl* 2015; **8**: 284–95.

474 46. Mahérault A-C, Kemble H, Magnan M, *et al.* Advantage of the F2:A1:B- IncF Pandemic
475 Plasmid over IncC Plasmids in *In Vitro* Acquisition and Evolution of $bla_{\text{CTX-M}}$ Gene-Bearing
476 Plasmids in *Escherichia coli*. *Antimicrob Agents Chemother* 2019; **63**: e01130-19.

477 47. Shin J, Ko KS. Effect of plasmids harbouring $bla_{\text{CTX-M}}$ on the virulence and fitness of
478 *Escherichia coli* ST131 isolates. *Int J Antimicrob Agents* 2015; **46**: 214–8.

479 48. Ranjan A, Scholz J, Semmler T, *et al.* ESBL-plasmid carriage in *E. coli* enhances in vitro

480 bacterial competition fitness and serum resistance in some strains of pandemic sequence types
481 without overall fitness cost. *Gut Pathog* 2018; **10**: 24.

482 49. Andersson DI, Hughes D. Antibiotic resistance and its cost: Is it possible to reverse
483 resistance? *Nat Rev Microbiol* 2010; **8**: 260–71.

484 50. Fernandez-Lopez R, del Campo I, Revilla C, Cuevas A, de la Cruz F. Negative Feedback
485 and Transcriptional Overshooting in a Regulatory Network for Horizontal Gene Transfer. *PLoS*
486 *Genet* 2014; **10**: e1004171.

487

Figure 1 Genetic map of a plasmid vector pMDP5_cureEC958 designed for plasmid curing. It contains chloramphenicol resistance encoding gene *catA1* for selective cultivation of strains harbouring the vector and sucrose sensitivity gene *sacB* for selection of plasmid-free isolates disposed of the vector. For the purpose of plasmid curing, it harbours genes encoding antitoxins VapB, PemI, Sok and CcdA and IncF plasmid replicons RepFIA and RepFII.

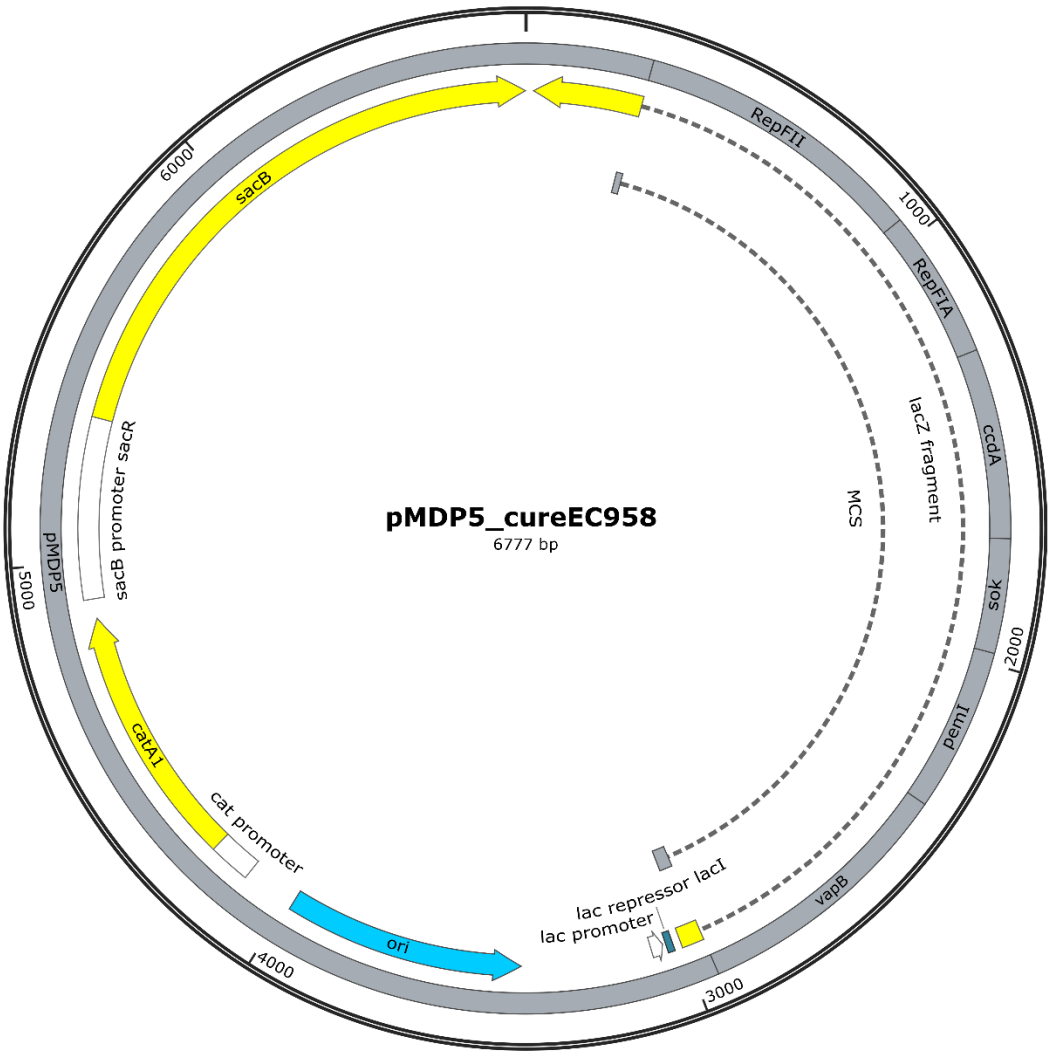

Figure 2 Transfer region of studied plasmids. The shading shows the similarity above 99.9%.

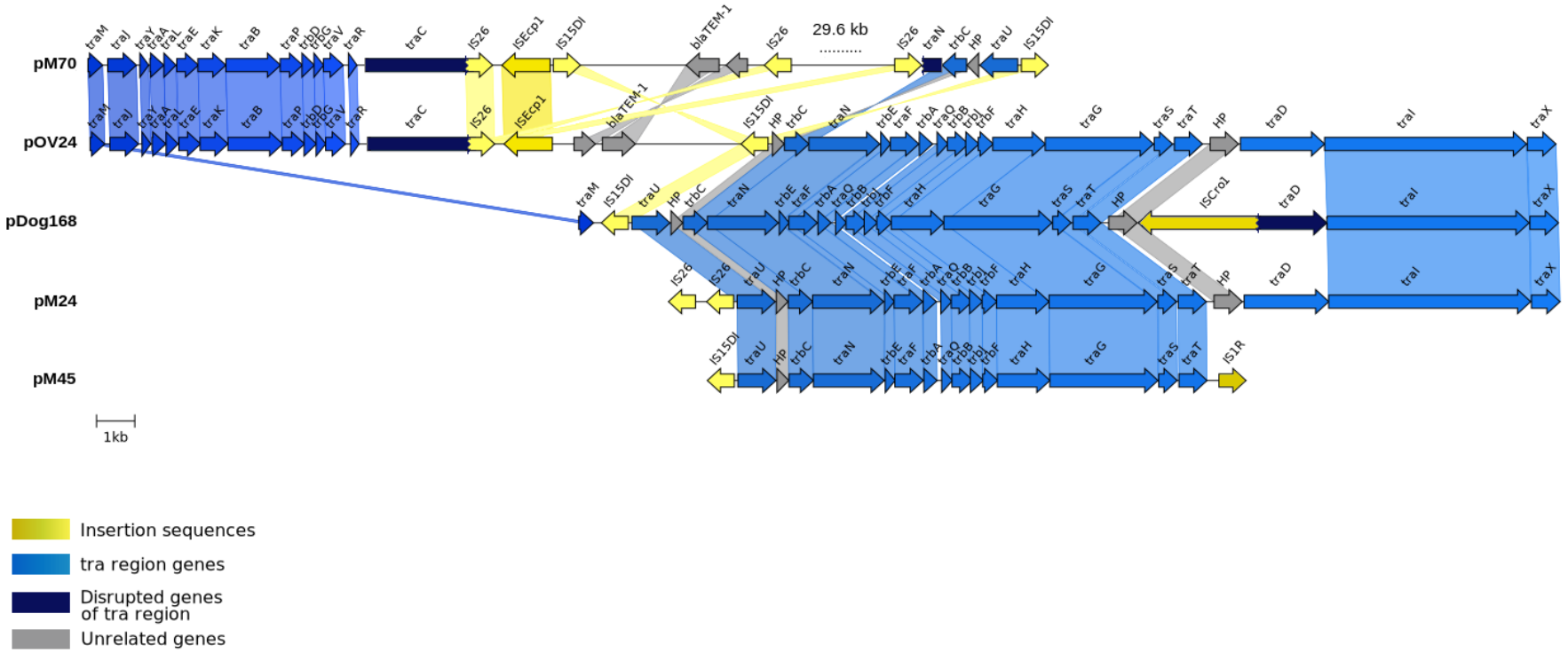

Figure 3 Relative fitness of plasmid-free clones in comparison to the corresponding wild-type isolates. Background fitness of wild-types was estimated as 1. CC stands for cured clone. Bars indicate standard deviation.

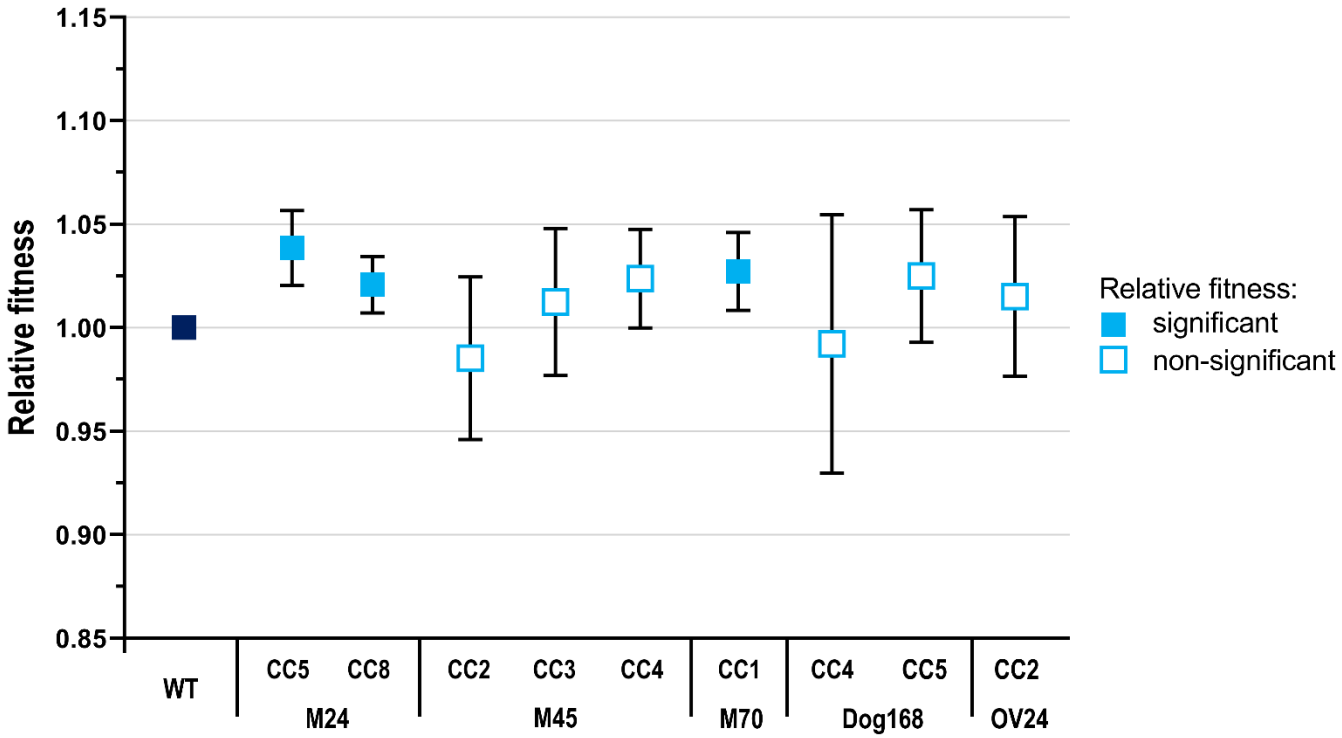

Table 1 Selected genetic characteristics of CTX-M-15-encoding IncF plasmids in our study.

Plasmid ID	pMLST[a]	Size (bp)	Toxin-antitoxins systems					ARGs[b]										GenBank accession number
			*ccd*AB	*pem*IK	*vap*CB	*par*ED	*hok/sok*	$bla_{CTX-M-15}$	bla_{TEM-1}	bla_{OXA-1}	*aac(6′)-Ib-cr*	*sul1*	*aadA5*	*mph*(A)	*dfrA17*	*tet*(A)	*catB3*	
pDog168	F2:A1:B-	131,080	■	■	■		■	■		■	■	■	■	■	■	■	*	MZ634324
pOV24	F2:A1:B-	144,582	■	■	■		■	■	■	■	■	■	■	■	■	■	*	MZ634325
pM45	F2:A1:B-	106,909	■	■	■	■		■		■	■	■	■	■	■	■	*	MZ634322
pM24	F2:A1:B-	116,543	■	■	■	■	■	■		■	■	■	■	■	■		*	MZ634326
pM70	F2:A1:B-	126,514	■	■	■	■	■	■	■			■	■	■	■	■		MZ634323

[a]pMLST – plasmid multilocus sequence type, [b]ARGs – antibiotic resistance genes

Coloured squares represent the presence of the genes, for antibiotic resistance genes with coverage 95% and identity 100%. *Gene *catB3* was disrupted by IS*26* leaving 70.3% coverage resulting in gene malfunction.

506 **Table 2** Relative fitness of the plasmid-free clones in comparison to their wild-type isolates

Isolate ID	Plasmid-free clone ID[a]	Relative fitness[b] (±SD)	*p* value
Dog168	CC4	0.992 ± 0.06	0.764
	CC5	1.025 ± 0.03	0.216
OV24	CC2	1.015 ± 0.04	0.284
M45	CC2	0.985 ± 0.04	0.386
	CC3	1.012 ± 0.03	0.444
	CC4	1.024 ± 0.03	0.108
M24	**CC5**	**1.038 ± 0.02**	**1.8×10^{-6}**
	CC8	**1.021 ± 0.01**	**4×10^{-4}**
M70	**CC1**	**1.027 + 0.02**	**1.5×10^{-3}**

507 [a] ID of constructed plasmid-free strains; CC stands for cured clone, [b] Relative fitness of plasmid-
508 free strain compared to the corresponding wild-type isolate which background fitness was
509 estimated as 1. Isolates highlighted in bold showed significant ($p < 0.05$) relative fitness
510 changes, however, the increase was moderate.