



HAL
open science

HlyF, an underestimated virulence factor of uropathogenic *Escherichia coli*

Camille Chagneau, Delphine Payros, Audrey Goman, Cécile Goursat, Laure David, Miki Okuno, Pierre-Jean Bordignon, Carine Séguy, Clémence Massip, Priscilla Branchu, et al.

► **To cite this version:**

Camille Chagneau, Delphine Payros, Audrey Goman, Cécile Goursat, Laure David, et al.. HlyF, an underestimated virulence factor of uropathogenic *Escherichia coli*. *Clinical Microbiology and Infection*, 2023, 29 (11), pp.1449.e1-1449.e9. 10.1016/j.cmi.2023.07.024 . hal-04330628

HAL Id: hal-04330628

<https://hal.inrae.fr/hal-04330628>

Submitted on 8 Dec 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.

Copyright

1 **Original article: HlyF, an underestimated virulence factor of**
2 **uropathogenic *Escherichia coli***

3

4 **Running title: Spreading of ColV plasmids carrying *hlyF* in**
5 **uropathogenic *E. coli***

6 Camille V. Chagneau^{1,2}, Delphine Payros¹, Audrey Goman¹, Cécile
7 Goursat¹, Laure David¹, Miki Okuno³, Pierre-Jean Bordignon¹, Carine
8 Séguy¹, Clémence Massip^{1,2}, Priscilla Branchu¹, Yoshitoshi Ogura³ Jean-
9 Philippe Nougayrède¹, Marc Marends⁴ and Eric Oswald^{1,2}

10 ¹*Digestive Health Research Institute (IRSD), INSERM, Université de Toulouse, INRAE,*
11 *ENVT, UPS, Toulouse, France;*

12 ²*CHU Toulouse, Hôpital Purpan, Service de Bactériologie-Hygiène, Toulouse, France;*

13 ³*Division of Microbiology, Department of Infectious Medicine, Kurume University*
14 *School of Medicine, Kurume, Japan;*

15 ⁴*Faculty of Science, University of Melbourne, Melbourne, Australia*

16 *corresponding author: eric.oswald@inserm.fr

17 Pr Eric Oswald

18 Institut de recherche en santé digestive

19 CHU Purpan – Place Baylac

20 CS 60039

21 31024 TOULOUSE CEDEX 3

22 France

23

24 **HlyF, an underestimated virulence factor of uropathogenic *Escherichia***
25 ***coli***

26

27 **Running title: Spreading of ColV plasmids carrying *hlyF* in**
28 **uropathogenic *E. coli***

29 Urinary tract infections (UTIs) are predominantly caused by uropathogenic
30 *Escherichia coli* (UPEC). By analysing a representative collection of UPEC
31 strains from community-acquired infections, we showed that 20 % of these
32 strains had the ability to produce the protein HlyF. These *hlyF*⁺ UPEC strains
33 were the most virulent, mostly responsible for pyelonephritis, often with
34 bloodstream infections. Using a mouse model of UTI, we showed that HlyF was
35 associated with the ability of UPEC to develop a urosepsis, with the presence of
36 bacteria in the spleen and an exacerbated inflammatory response. In contrast to
37 archetypical UPEC strains, *hlyF*⁺ UPEC strains are not restricted to
38 phylogroup B2 and harbor a specific repertoire of virulence factors reflecting the
39 fact that HlyF is encoded by conjugative ColV-like plasmids. These plasmids
40 also carry antimicrobial resistance genes, which may facilitate their selection and
41 spreading amongst people receiving antimicrobial therapy. Overall, our data
42 suggest that HlyF is a virulence factor in UPEC and spreading of ColV-like
43 plasmids encoding *hlyF* warrants further investigation.

44 **Keywords:** uropathogenic *E. coli*, urinary tract infections, HlyF, ColV plasmids

45 **Introduction**

46 Urinary tract infections (UTIs) are one of the most common infections worldwide [1].
47 UTIs are associated with a decrease in the quality of life of patients and a significant
48 clinical and economic burden [2]. In both community and hospital settings, UTIs pose a
49 threat to public health. They are the most common outpatient infections and at least half
50 of adult women will have one UTI or more in their lifetime [3]. Uropathogenic *E. coli*
51 (UPEC) are responsible for more than three quarters of community-acquired UTI, and
52 about half of nosocomial infections [4]. The majority of these infections are benign, but

53 their management can be complicated by frequent recurrences and the emergence of
54 antibiotic resistance, leading to therapeutic impasses. In more severe cases,
55 complications such as kidney damage in young children or the onset of sepsis may arise.
56 A large proportion of sepsis originates from the urinary tract (accounting for 20-30%,
57 *i.e.* 2 to 9 million cases per year) and urosepsis may progress to septic shock with
58 significant mortality and morbidity [5]. The mortality from urosepsis is estimated to be
59 more than 1.5 million deaths per year, making it a major public health threat.

60 The pathogenicity of UPEC involves a variety of factors [6] such as specific adhesins,
61 toxins or iron uptake systems [7,8]. The *hlyF* gene encodes a protein that was
62 previously thought to be an haemolysin (haemolysin F) [9]. Recent work has shown that
63 HlyF is in fact a cytoplasmic enzyme that increases the formation of outer membrane
64 vesicles (OMVs) allowing the release of the *bona fide* haemolysin E (ClyA) responsible
65 for the previously observed haemolytic phenotype [10]. HlyF-induced OMVs not only
66 act as cargos for toxins, but also have the ability to block autophagic flux in eucaryotic
67 cells and to exacerbate the activation of the inflammasome through the non-canonical
68 pathway [10,11].

69 So far the *hlyF* gene was shown to be associated with the virulence of avian pathogenic
70 *E. coli* (APEC) and neonatal meningitis-causing *E. coli* (NMEC) [10,12–14]. In this
71 study, we observed that *hlyF*⁺ UPEC were isolated from the most severe cases of
72 human UTI. The ColV plasmids carrying *hlyF* are widely spread among UPEC strains
73 and encode several virulence factors as well as antimicrobial resistances which can
74 favour their dissemination. In a mouse model of UTI, we have shown that HlyF
75 promoted pyelonephritis and consecutive bloodstream infection.

76

77 **Materials and methods**

78 ***Collection of clinical strains***

79 We collected 225 *E. coli* strains from prescribed urine cultures of 223 patients attending
80 the Adult Emergency Department of Toulouse University Hospital, France, between
81 July and October 2017, corresponding to community-acquired UTIs as previously
82 described [15]. Patients at risk of misdiagnosis due to age (> 75 years) or comorbidities,
83 and patients with urinary catheters were excluded. In accordance with French
84 regulations on the analysis of observational databases, no specific informed consent was
85 required for the collection of clinical *E. coli*. Data were analysed anonymously.

86

87 ***Bacterial strains***

88 UPEC strain ECC166 was isolated from a 23-year-old woman without comorbidities
89 suffering from pyelonephritis. Whole genome sequence of ECC166 indicates it is of
90 serotype O1:H7, phylogroup B2 and sequence type ST95. In addition to the ability to
91 produce HlyF, ECC166 exhibits a wide array of virulence factors such as multiple iron
92 acquisition systems (locus *iro*, *iucABCD*, *fyuA*, *sitABCD*), Vat toxin and PapGII
93 adhesin. We constructed a deletion mutant of *hlyF* in UPEC strain ECC166 as
94 previously described [16]. The $\Delta hlyF$ mutant was constructed using primers PB3-mut-
95 hlyF-F (5'-
96 TAAGATAATTTATTTTTATAATGATCACATGAAAACAAAAGAGGTTAGATgtg
97 taggctggagctgcttcg-3') and PB4-mut-hlyF-R (5'-
98 TTTATATATTATGAGTGCAACACCAACAATAATTCTGATTATGATAAATAcata
99 tgaatatcctccttagt-3'). This mutant was complemented with plasmids expressing a wild-
100 type form of HlyF (referred to as HlyF) or a mutant in the catalytic domain (referred to
101 as SDM) [10].

102

103 ***Mouse model of UTI***

104 *Ethic statement*

105 All the experimental procedures were carried in accordance with the European
106 directives for the care and Use of animals for Research Purposes and were validated by
107 the local ethics committee from CREFRE US006 and by the national ethics committee
108 (Regional Centre for the Functional Exploration and Experimental Ressources) (number
109 21-U1220-EO/MT-128).

110 *Mouse model*

111 Female, 6-8 weeks old, C3H/HeN mice (Janvier Labs, Le Genest Saint Isle, France)
112 were infected twice transurethrally as previously described [17]. Briefly, bacterial
113 strains were cultivated statically in LB and resuspended to an inoculum of 2.10^7 CFU
114 in 50 μ L of PBS 1X. For bacterial enumeration, bladder, kidney and spleen were
115 harvested, homogenised in FastPrep Lysing Matrix D with 800 μ L PBS 1X and serial
116 dilutions were plated onto solid LB agar with adequate antibiotic.

117 *Body weight and clinical scoring*

118 Body weight was assessed before and at the end point of the experiment. The severity of
119 the clinical signs was evaluated blindly by scoring (body temperature, coat condition,
120 mobility of the animals and signs of pain such as grimace) (Table S1). A weight loss of
121 more than 15 % associated with a clinical score of more than 11 led to stop the
122 experiment and to humanely sacrifice the animal.

123 *Cytokines quantification*

124 Tissue proteins were extracted with a solution of RIPA (0.5% deoxycholic acid, 0.1%
125 sodium dodecyl sulfate, 1% Igepal in Tris-buffered saline 1X ; pH = 7.4) added with a
126 protease inhibitor cocktail (Roche diagnostic, France Ref 11697498001). Clear lysates
127 of spleen were processed for ELISA using commercial kits (Duoset R&D Systems,

128 Lille, France) for Interleukin-1 β (IL-1 β) and Interleukin-6 (IL-6) according to the
129 manufacturer's instructions. Data are expressed as picograms of cytokines per milligram
130 of tissue protein.

131

132 *Sequencing data, sequence alignments and phylogenetic analyses*

133 *Illumina sequencing*

134 The Illumina NextSeq500 Mid Output platform (Integragen, Evry, France) was used to
135 generate 2 x 150 bp paired-end reads for whole genome sequencing of the UPEC strains
136 as already described [17].

137 *Nanopore sequencing*

138 Genomic DNAs were purified from 600 μ L of overnight bacterial culture using the
139 Wizard HWM genomic DNA prep kit (Promega) following the manufacturer's
140 instructions. Nanopore sequencing libraries were prepared with the rapid barcoding kit
141 SQK-RBK004 (Oxford Nanopore) and loaded on a MinION MK-Ib device fitted with a
142 R9.4 flowcell. Live rapid basecalling was performed with MinKnow 20.13.3 using the
143 guppy version 4.2.2 (Oxford Nanopore). The resulting fastq reads were processed with
144 the program guppy_barcode (Oxford Nanopore) for quality filtering >7 ,
145 demultiplexing, and barcode removal.

146 *Genome analysis*

147 Genome *de novo* assembly and analysis (search for virulence and antimicrobial
148 resistance genes, phylogroup, plasmid typing...) were performed with the BioNumerics
149 7.6 software (Applied Maths), Enterobase and Center for Genomic Epidemiology
150 (<http://enterobase.warwick.ac.uk/> ; <http://genomicepidemiology.org/>). For SNP-based
151 phylogenetic trees, core genome alignments were generated after mapping raw reads to
152 the *E. coli* MG1655 genome using LASTAL after identifying dispersed repeats

153 (BLASTN) and tandem repeats (trf) but without considering recombination [18]. The
154 core genome phylogenetic tree was inferred with the Maximum-likelihood algorithm
155 (RAxML) using Enterobase for *hlyF*+ strains [18].

156 Alignment, comparisons, and detection of mutations in *hlyF* locus were performed
157 either using BioNumerics software or Clustal Omega [19] with pECOS88 as a reference
158 (CU928146.1: 130 790... 133 302) from *E. coli* S88 serotype O45:K1:H7 isolated from
159 neonatal meningitis and including other reference strains: *E. coli* SP15 of serotype
160 O18:K1:H7 isolated from neonatal meningitis (AP024132.2) and *E. coli* Combat1119 of
161 serotype O-:H28 isolated from UTI (CP021728.1). MEGA X software was used to
162 construct UPGMA tree of *hlyF* locus [20]. The evolutionary distances were computed
163 using the Maximum Composite Likelihood method and are in the units of the number of
164 base substitutions per site.

165 Multiple alignments for the *hlyF* and *traM-traX* sequences were performed by mafft
166 v7.429 [21] and then phylogenetic estimations were conducted using the maximum-
167 likelihood (ML) method implemented in RAxML-ng v.0.9.0 (--all, --bs-trees 100) [22].
168 The best-fit model for each phylogenetic analysis was selected by using ModelTest-NG
169 v.0.1.6 [23]. Co-phylogenetic analysis between the ML trees was performed using the
170 ‘cophylo’ function of the R package Phytools 0.7.80 [24]. Phylogenetic clusters were
171 determined by RAMI [25].

172 Hybrid assembly and annotation of the fully assembled plasmids was performed using
173 Unicycler and Prokka from the Galaxy interface (<https://usegalaxy.org/>). Plasmid linear
174 comparison and representation was performed using Easyfig 2.2.5 software
175 (<https://mjsull.github.io/Easyfig/>).

176 ***Statistical analyses***

177 Graphical representations of data were performed using the GraphPad Prism 8.0
178 (GraphPad Software, Inc, San Diego, CA). The experimental data are represented as
179 means \pm standard errors of the mean (SEM). For time-to-death experiment, the
180 difference between the experimental groups was evaluated by the log-rank (Mantel-
181 Cox) test. For the statistical analysis of clinical score, body weight loss and bacterial
182 load, a Student's t-test was used to differentiate HlyF expressing strains and strains not
183 expressing HlyF. For cytokine analysis, statistical significance between experimental
184 groups was determined by one-way ANOVA followed Bonferroni's comparison test. P
185 value of < 0.05 was considered significant. For epidemiological analysis on UPEC
186 strains, P values were calculated using Fischer's exact test.

187

188 **Results**

189 ***hlyF is epidemiologically associated with severe urinary tract infections in humans.***

190 We investigated the presence of *hlyF* in a representative collection of 225 sequenced
191 UPEC strains from community-acquired UTIs [15]. This collection was shown to be
192 representative of the commonly described UPEC collections, for classically described
193 virulence factors, phylogeny and for antibiotic resistance profiles commonly reported in
194 France [15,17]. We found that *hlyF* was present in 42/225 (19%) of the strains. These
195 *hlyF*⁺ UPEC were present in all phylogroups (Figure S1). The majority of strains
196 (30/42) belonged to phylogenetic group B2. However, the proportion of *hlyF*⁺ strains in
197 the B2 phylogenetic group (30/155) was not higher than that of non-B2 strains (12/70;
198 $p=0.8536$). Within the B2 phylogenetic group, the strains belonged to classically
199 described sequence types in UPEC (ST), especially ST95. Most of ST95 strains carried
200 the *hlyF* gene (25/27 ST95).

201 The isolation rate for *hlyF*⁺ UPEC strains was significantly higher in patients with
202 pyelonephritis (26%; 27/104) compared to patients with less severe UTI, *e.g.* cystitis or
203 asymptomatic bacteriuria (12.4%; 15/121; *p*=0.0104). We also observed that *hlyF*⁺
204 UPEC-infected patients had a trend towards higher prevalence of concomitant
205 bloodstream infection (30%; 7/21; *p*=0.08).

206

207 ***HlyF is a virulence factor of UPEC in a mouse model of UTI.***

208 To test whether HlyF plays a role during a UTI, we used a well-established mouse
209 model of UTI based on transurethral injection of bacteria [26]. We selected an UPEC
210 strain isolated from pyelonephritis: ECC166, of sequence type ST95 and serotype
211 O1:K1:H7, which is representative of the majority of *hlyF*⁺ strains in our UPEC
212 collection. Both ECC166 wild-type and *hlyF* mutant strains induced infection in mice.
213 However, infection was more severe with the wild-type compared to the Δ *hlyF* isogenic
214 mutant, with increased clinical signs (Table S1) and body weight loss (Figure S2), and
215 ultimately higher lethality (40% versus less than 10% of mortality for the wild-type and
216 mutant strain respectively) (Figure 1). We confirmed that this difference in virulence
217 was due to the presence of the *hlyF* gene by complementing the mutant with a plasmid
218 expressing HlyF (Figure S3). Mice infected with the wild-type strain had more bacteria
219 in the kidneys than mice infected with the isogenic mutant although no difference was
220 observed in the bladder (Figure 1C). In addition, mice infected with the *hlyF*⁺
221 complemented strain had a higher bacterial load in the spleen, confirming bloodstream
222 infection from the urinary tract (Figure 1C). An enhanced inflammatory response with
223 higher levels of interleukin IL-1 β and IL-6 in the spleen was observed in these mice
224 (Figure 1D). These results were confirmed in the *hlyF* mutant complemented with HlyF
225 (Figure S3C).

226 Collectively, these results indicate that the production of HlyF in UPEC strain is
227 responsible for an increased severity of UTI in a mouse model, promotes bloodstream
228 infection and induces a higher systemic inflammatory response.

229

230 **HlyF+ UPEC strains have a specific virulence signature.**

231 Using WGS of the strains collection, we compared the virulome of *hlyF+* and of *hlyF-*
232 UPEC strains (Table 1). Although evasion of the host response during UTI is often
233 mediated by toxin production [27], we observed that the vast majority of *hlyF+* strains
234 did not carry the genes for toxins classically associated with UPEC such as the alpha-
235 haemolysin (HlyA), the Cytotoxic Necrotizing Factor 1 (CNF1), or the autotransporter
236 protease Sat. By contrast, there was an association between *hlyF* and the adhesin
237 PapGII, although this chromosomal allele involved in renal colonisation has been
238 previously found almost exclusively in ST95 strains [28]. There was also a strong
239 association with virulence genes coding for colicins, the increased serum survival type 1
240 plasmid variant and iron and metals scavengers and transporter systems such as
241 aerobactin, salmochelins and the transporters encoded by *sitABCD* and *etsABC* [29,30]
242 (Table 1). The genes encoding this arsenal of virulence factors are known to be carried
243 by plasmids of the pColV family [29,30], which also harbour *hlyF*, raising the
244 possibility of a common genetic origin.

245 **Table 1. Association between HlyF and classical UPEC virulence factors.** Among
246 the strains carrying each given virulence factor, the number of *hlyF+* and *hlyF-* strains is
247 indicated, with the percentage among the *hlyF+* or *hlyF-* strains in brackets. Fischer's
248 exact test.

	gene	number of positive strains (% total)	<i>hlyF+</i> (%)	<i>hlyF-</i> (%)	p(Fischer)
--	------	--------------------------------------	------------------	------------------	------------

Toxins					
haemolysin a	<i>hlyA</i>	82 (36)	3 (7)	79 (43)	<0.0001
serine protease autotransporter	<i>sat</i>	92 (41)	0 (0)	92 (50)	<0.0001
vacuolating autotransporter toxin	<i>vat</i>	133 (59)	30 (71)	103 (56)	0.08
cytotoxic distending toxin	<i>cdt</i>	14 (6)	4 (10)	10 (5)	0.3027
cytotoxic necrotizing factor 1	<i>cnf1</i>	70 (31)	3 (7)	67 (37)	<0.0001
colibactin	<i>pks</i>	96 (43)	9 (21)	87 (48)	0.0019
increased serum survival type 1 plasmid variant	<i>iss_{plasmid}</i>	40 (18)	40 (95)	0	<0.0001
Iron transport					
salmochelins	<i>iroN</i>	117 (52)	41 (98)	76 (42)	<0.0001
aerobactin	<i>iutA</i>	138 (61)	36 (86)	102 (56)	0.0002
yersiniabactin	<i>fuyA</i>	210 (93)	39 (93)	171 (93)	1
ABC transporter	<i>sitABCD</i>	202 (90)	42 (100)	160 (87)	0.0098
putative ABC transporter	<i>etsABC</i>	37 (16)	37 (88)	0	<0.0001
Adhesins					
PapgII	<i>papgII</i>	87 (39)	25 (60)	62 (34)	0.0027
PapgIII	<i>papgIII</i>	41 (18)	4 (10)	37 (20)	0.1239

F17-like fimbriae	<i>uclD</i>	70 (31)	3 (7)	67 (37)	<0.0001
S/F1C fimbriae	<i>sfa/foc</i>	117 (52)	5 (12)	66 (36)	0.0017
Bacteriocins					
colicin V	<i>cvaA</i>	35 (16)	35 (83)	0	<0.0001
colicin Ia	<i>cia</i>	35 (16)	26 (62)	9 (5)	<0.0001
microcin M	<i>mcmA</i>	69 (31)	5 (12)	64 (35)	0.0028
microcin H47	<i>mchB</i>	65 (29)	5 (12)	60 (33)	0.0076
colicin B	<i>cbi</i>	11 (5)	7 (17)	4 (2)	0.0009
colicin M	<i>cmi</i>	12 (5)	8 (19)	4 (2)	0.0002

249

250 ***hlyF* is carried by pColV conjugative plasmids.**

251 Assembly and annotation performed after both Nanopore and Illumina sequencing of
252 the UPEC strain ECC166 yielded a chromosome of 4,945,664 bp and two plasmids of
253 115,445 bp and 1,552 bp, respectively. *hlyF* was carried by the largest plasmid
254 belonging to pColV family (Figure S4). As previously described, *hlyF* was present in an
255 operon found on pColV family plasmids, here referred to as *hlyF* locus, containing the
256 *hlyF* gene and a putative *mig-14* ortholog gene (*mig-14-like*) which encodes an
257 antimicrobial resistance factor [31–33] (Figure 2). This locus was systematically
258 associated with *ompTp* gene, which encodes for plasmid variant of OmpT, a protease
259 with an activity against host antimicrobial peptides [10,30,32,34]. In all UPEC strains,
260 *hlyF* locus was highly conserved with pairwise similarities of at least 98% between two
261 loci (Figure 2, 3 and 4). The *hlyF* gene itself showed even more similarity amongst the
262 strains with the presence of only 11 SNP sites within the 42 strains, 6 of which
263 generated silent mutations (Table S2). The predicted catalytic domain (YTHSK) and
264 NAD binding domain (GATGFLG) of HlyF [10] were conserved in all strains.

265 The largest plasmid of ECC166 also contained an arsenal of additional virulence factors
266 characteristic of pColV family plasmids : colicin *colV*, aerobactin operon (*iuc/iut*),
267 salmochelins locus *iro*, *sitABCD* metal transport system, and increased serum survival
268 type 1 plasmid variant (*iss*) [29,30]. This plasmid also encoded a large conjugation
269 system (*tra* genes). We confirmed its ability of transfer by conjugation, by using as a
270 donor strain the $\Delta hlyF$ mutant in which a kanamycin resistance cassette was inserted in
271 *hlyF* gene and strain *E. coli* J53 as recipient [35].

272

273 ***hlyF* is associated with two main lineages of conjugative plasmids in UPEC.**

274 The vast majority of *hlyF*⁺ UPEC strains showed the association of *hlyF*, *ompTp*, *iutA*,
275 *iroN*, *sitABCD*, *etsABC*, *iss* and *cvaA* genes which is characteristic of the conserved part
276 of the pColV plasmids family, together with *cia*, which encodes colicin Ia [29,30]
277 (Figure 2). We investigated whether the same plasmid was responsible for the spread of
278 *hlyF* in the UPEC population and more broadly in the *E. coli* population. Based on the
279 sequence of the *hlyF* locus, we identified two main groups of *hlyF*⁺ UPEC plasmids:
280 one with the *hlyF* locus identical to the one of *E. coli* SP15 and one identical to the one
281 of *E. coli* ECOS88 (Figure 2). ECOS88 and SP15 are two archetypal *E. coli* strains
282 isolated from neonatal meningitis [36,37]. To analyse the relationship and compare the
283 evolution of the *hlyF* locus with the structure of the plasmid carrying it, we performed a
284 co-phylogeny analysis between the *hlyF* locus and the *traMtraX* locus, which includes
285 the region encoding the conjugation system that constitutes the backbone of the
286 plasmid. Interestingly, the two main groups of strains showed conserved loci and
287 plasmid structures that have evolved in parallel (Figure S5). The gene *hlyF* is thus
288 associated with two main lineages of plasmids in UPEC.

289

290 ***hlyF*-associated pColV plasmids are platforms for dissemination of virulence**
291 **factors and antimicrobial resistance genes.**

292 In addition to the plasmids belonging to the pSP15 and pS88 lineages, more variable
293 *hlyF* locus determinants were present in some plasmids, along with a less conserved
294 *traMtraX* locus, associated with a diverse repertoire of pColV determinants (Figure 3 &
295 S5), suggesting more variable plasmids. To better characterize this heterogeneity, we
296 sequenced representative strains using the Nanopore technology and compared these
297 plasmids with previously sequenced plasmids related to pColV (pSP15, pECOS88,
298 pCombat11I9-2).

299 The structure of all the sequenced plasmids was conserved (Figure 3 and 4). However,
300 many inversion sequences were present in less conserved regions, suggesting frequent
301 IS-mediated recombinations (Figure 4). In addition, some plasmids were missing certain
302 regions or new regions were present. These new insertions affected either (1) accessory
303 virulence factors that are less frequently associated with pColV plasmids (*eit*, *tsh*), or
304 (2) antimicrobial resistance genes carried by transposons or integrons. These results
305 indicated that plasmids carrying *hlyF* in UPECs had a conserved scaffold, but that they
306 were also platforms for the dissemination of various virulence and antimicrobial
307 resistance genes. In summary, *hlyF* was carried by two main lineage of conjugative
308 plasmids that have evolved in parallel and could promote the spread of virulence factors
309 and antibiotic resistance genes in UPEC, a feature that may have an impact on UTI
310 epidemiology and treatment options.

311

312 **Discussion**

313 *Virulence plasmids and UPEC*

314 The role of plasmids in the pathophysiology of UTI has anecdotally been studied [39].
315 Acquisition of the pColV plasmid may increase growth in urine and colonisation of the
316 murine kidney [40]. The pColV family of plasmids harbour multiple virulence genes
317 encoding iron uptake and transport functions, resistance to host response or bacteriocins
318 that may be involved in the digestive and urinary colonisation stages of UTI. For
319 example, both Mig-14-like and OmpTp which are highly conserved and consistently
320 associated in *hlyF*+ UPEC strains, are involved in resistance to host antimicrobial
321 peptides [32,33]. In particular, OmpTp is a protease that has been shown to be
322 particularly active against antimicrobial peptides produced in the urinary tract [34].
323 However, independently of the other virulence factors encoded by ColV, we show in a
324 mouse model that HlyF alone can control the severity of UTI and contribute to the
325 occurrence of bloodstream infection.

326

327 *HlyF: an emerging virulence factor associated with severe UTI*

328 The gene encoding HlyF has been previously described in APEC and NMEC strains
329 [12–14] but has never been studied in UPEC, probably because *hlyF* is absent from the
330 genome of archetypal UPEC strains such as *E. coli* CFT073, 536 or UTI89. We have
331 observed in our representative UPEC collection that almost 20% of strains isolated from
332 community-acquired infections carry this virulence factor. This frequency is confirmed
333 by data from the literature based on the epidemiology of genes classically associated
334 with *hlyF* (*ompTp*, *cvaC*...) [28,34,41–43]. *hlyF*+ strains seem especially frequent in
335 UPEC belonging to ST95. Interestingly, in larger collections, some subgroups of ST95
336 strains (C and D) are much more likely to carry pColV plasmids and are becoming more
337 prevalent in patients with UTI, suggesting an on-going dissemination [44]. These strains
338 are mostly devoid of classical UPEC virulence genes such as *hlyA* or *cnfI* [28], which

339 may suggest either alternative virulence mechanisms specific to HlyF or pColV
340 determinants, or that the accumulation of the virulence genes may be too detrimental for
341 the bacterial host and would result in strains being unable to colonise the urinary tract.
342 We found that UPEC strains producing HlyF are epidemiologically associated with
343 more severe UTI in humans. Interestingly, we observed a trend towards a higher
344 frequency of bloodstream infections, in accordance with larger studies [28]. We also
345 confirmed HlyF role *in vivo* in mouse model of UTI.

346

347 *A role of OMVs in urosepsis*

348 The mode of action of HlyF in the development of more severe UTIs remains elusive.
349 HlyF induces the formation of OMVs [10]. OMVs can play the role of cargos. For
350 example, it has recently been shown that UPEC OMVs can contain and transport the
351 enzyme AroB, which is responsible for the synthesis of aromatic amino acids, thereby
352 increasing the motility of recipient bacteria [45]. They can also increase the secretion of
353 toxins such as CNF1 or CDT [10,46]. However, most of the *hlyF*⁺ UPEC strains,
354 including ECC166 used in this study in the mouse UTI model, do not produce such
355 toxins. It is therefore more likely that the intrinsic properties of OMVs could explain the
356 pathogenicity of the strain producing HlyF. We have previously shown that OMVs
357 produced by *hlyF*⁺ *E. coli* possess exacerbated pro-inflammatory properties that could
358 play a role in the pathophysiology of infections and their severity [11].

359

360 *HlyF and pColV plasmid as virulence and antimicrobial resistance genes platforms: 361 beyond UTI*

362 HlyF displays wide dissemination in UPEC, suggesting that *hlyF*, and/or other
363 determinants harboured by pColV plasmids, could give an advantage in a wide range of

364 genetic backgrounds of *E. coli* for inducing UTI. However, we notice that major
365 sequence types (ST) of *hlyF*⁺ strains also include NMEC and APEC strains which share
366 an important genetic proximity. This is the case for ST95, one of the predominant
367 UPEC ST in our collection and a highly represented ST within the *hlyF*⁺ UPEC strains
368 [36,47,48]. However, other ST such as ST58, ST117 or ST131-*H22* are often strongly
369 associated with avian infections which emergence and pathogenicity seem to be partly
370 linked to the acquisition of a pColV plasmid [49–51].
371 *hlyF* and pColV plasmids are much more widely distributed than just in APEC, NMEC
372 and UPEC and may contribute to the acquisition of virulence factors in other pathogens
373 and to emergence of “high-risk” pathogenic clones [52]. Recently, a new emerging
374 hybrid clone and lineage of *E. coli* has been described: an entero-haemorrhagic (EHEC)
375 of serotype O80:H2 possessing typical EHEC virulence factors, which has acquired a
376 pColV-like plasmid carrying *hlyF* and virulence factors usually associated with extra-
377 intestinal pathogenic *E. coli* ExPECs [53,54]. Acquisition of this plasmid seems to give
378 the EHEC host an atypical pathogenicity with a greater propensity to generate
379 bloodstream infections, a known characteristic of ExPEC. Beside carrying the genes
380 coding for various virulence factors, these plasmids are becoming concerning vectors of
381 antimicrobial resistance genes, as seen for EHEC O80:H2 [55]. Other examples include
382 the emerging ExPEC ST58 pColV⁺ sublineage and adherent-invasive *E. coli* strain
383 NRG857c associated to Crohn disease, as well as an epidemic clone of *Salmonella*
384 *enterica* serovar Kentucky: in both cases, the bacterial hosts harbour pColV-related
385 plasmids that carry antimicrobial resistance determinants [49,54,56,57]. While these
386 plasmids are all derived from the same conserved structure, their additional genes
387 coding for antimicrobial resistance mechanisms may favour their selection and
388 dissemination.

389

390 In conclusion, the results of our study and the sequencing of increasing numbers of *E.*
391 *coli* isolates reflect a paradigm shift in our understanding of pColV plasmids: from
392 being confined to APEC and NMEC, to their dissemination within new STs or even
393 new *E. coli* pathotypes. Given the ageing of the population and the increase in co-
394 morbidities and complex medical procedures such as the use of transplants or
395 immunosuppressive drugs, the frequency and severity of UTIs is likely to increase [58].
396 Moreover, antimicrobial resistance is set to become one of the leading causes of
397 mortality in the coming decades. It is therefore essential to monitor in clinical isolates
398 the presence and evolution of plasmids that promote the dissemination of virulence
399 factors such as HlyF, favour the onset of bloodstream infections, and carry antibiotic
400 resistance genes.

401

402 **Acknowledgments**

403 We thank the staff of the Tri GenoToul imaging facility, Toulouse.

404

405 **Funding**

406 This work was supported by the French National Agency for Research under grant
407 (UTI-TOUL ANR-17-CE35-0010) and the French National Institute for Health and
408 Medical Research under grant ("poste d'accueil INSERM 2018").

409

410 **Declaration of interest statement**

411 All the authors have declared that no competing interests exist.

412

413 **Data availability statement**

414 Sequencing data (Illumina and Nanopore for some strains) are available in the NCBI
415 Database, Bioproject number PRJNA615384.

416

417 **Authors contribution statement**

418 Conceptualization: CVC, JPN, EO

419 Data curation: CVC, CM

420 Formal analysis: CVC, DP

421 Funding acquisition: CVC, JPN, EO

422 Investigation: CVC, DP, AG, CG, LD, MO, PJB, CS, CM, PB, JPN, MM

423 Methodology: YO, DP, LD, PB, JPN, MM

424 Project administration: EO

425 Resources: AG, PJB, CM, PB, MM

426 Visualization: CVC, DP, MO

427 Writing –original draft: CVC, DP

428 Writing –review & editing: CVC, JPN, MM, EO

429

430 **References**

- 431 [1] Flores-Mireles AL, Walker JN, Caparon M, Hultgren SJ. Urinary tract infections:
432 epidemiology, mechanisms of infection and treatment options. *Nat Rev Microbiol*
433 2015;13:269–84. <https://doi.org/10.1038/nrmicro3432>.
- 434 [2] Öztürk R, Murt A. Epidemiology of urological infections: a global burden. *World*
435 *J Urol* 2020;38:2669–79. <https://doi.org/10.1007/s00345-019-03071-4>.
- 436 [3] Wagenlehner F, Tandogdu Z, Bartoletti R, Cai T, Cek M, Kulchavenya E, et al.
437 The Global Prevalence of Infections in Urology Study: A Long-Term, Worldwide
438 Surveillance Study on Urological Infections. *Pathog Basel Switz* 2016;5:10.
439 <https://doi.org/10.3390/pathogens5010010>.
- 440 [4] Foxman B. The epidemiology of urinary tract infection. *Nat Rev Urol* 2010;7:653–
441 60. <https://doi.org/10.1038/nrurol.2010.190>.
- 442 [5] Bonkat G, Cai T, Veeratterapillay R, Bruyère F, Bartoletti R, Pilatz A, et al.
443 Management of Urosepsis in 2018. *Eur Urol Focus* 2019;5:5–9.
444 <https://doi.org/10.1016/j.euf.2018.11.003>.
- 445 [6] Schreiber HL, Conover MS, Chou W-C, Hibbing ME, Manson AL, Dodson KW,
446 et al. Bacterial virulence phenotypes of *Escherichia coli* and host susceptibility

- 447 determines risk for urinary tract infections. *Sci Transl Med* 2017;9:eaaf1283.
448 <https://doi.org/10.1126/scitranslmed.aaf1283>.
- 449 [7] Klein RD, Hultgren SJ. Urinary tract infections: microbial pathogenesis, host-
450 pathogen interactions and new treatment strategies. *Nat Rev Microbiol* 2020.
451 <https://doi.org/10.1038/s41579-020-0324-0>.
- 452 [8] Subashchandrabose S, Mobley HLT. Virulence and Fitness Determinants of
453 Uropathogenic *Escherichia coli*. *Microbiol Spectr* 2015;3.
454 <https://doi.org/10.1128/microbiolspec.UTI-0015-2012>.
- 455 [9] Morales C, Lee MD, Hofacre C, Maurer JJ. Detection of a Novel Virulence Gene
456 and a *Salmonella* Virulence Homologue Among *Escherichia coli* Isolated from
457 Broiler Chickens. *Foodborne Pathog Dis* 2004;1:160–5.
458 <https://doi.org/10.1089/fpd.2004.1.160>.
- 459 [10] Murase K, Martin P, Porcheron G, Houle S, Helloin E, Pénary M, et al. HlyF
460 Produced by Extraintestinal Pathogenic *Escherichia coli* Is a Virulence Factor That
461 Regulates Outer Membrane Vesicle Biogenesis. *J Infect Dis* 2016;213:856–65.
462 <https://doi.org/10.1093/infdis/jiv506>.
- 463 [11] David L, Taieb F, Pénary M, Bordignon P-J, Planès R, Bagayoko S, et al. Outer
464 membrane vesicles produced by pathogenic strains of *Escherichia coli* block
465 autophagic flux and exacerbate inflammasome activation. *Autophagy* 2022;0:1–
466 13. <https://doi.org/10.1080/15548627.2022.2054040>.
- 467 [12] Kaczmarek A, Budzyńska A, Gospodarek E. Prevalence of genes encoding
468 virulence factors among *Escherichia coli* with K1 antigen and non-K1 *E. coli*
469 strains. *J Med Microbiol* 2012;61:1360–5. [https://doi.org/10.1099/jmm.0.044263-](https://doi.org/10.1099/jmm.0.044263-0)
470 0.
- 471 [13] Johnson TJ, Wannemuehler Y, Doetkott C, Johnson SJ, Rosenberger SC, Nolan
472 LK. Identification of Minimal Predictors of Avian Pathogenic *Escherichia coli*
473 Virulence for Use as a Rapid Diagnostic Tool. *J Clin Microbiol* 2008;46:3987–96.
474 <https://doi.org/10.1128/JCM.00816-08>.
- 475 [14] Moulin-Schouleur M, Répérant M, Laurent S, Brée A, Mignon-Grasteau S,
476 Germon P, et al. Extraintestinal pathogenic *Escherichia coli* strains of avian and
477 human origin: link between phylogenetic relationships and common virulence
478 patterns. *J Clin Microbiol* 2007;45:3366–76. [https://doi.org/10.1128/JCM.00037-](https://doi.org/10.1128/JCM.00037-07)
479 07.
- 480 [15] Massip C, Chagneau CV, Boury M, Oswald E. The synergistic triad between
481 microcin, colibactin, and salmochelin gene clusters in uropathogenic *Escherichia*
482 *coli*. *Microbes Infect* 2020;22:144–7. <https://doi.org/10.1016/j.micinf.2020.01.001>.
- 483 [16] Datsenko KA, Wanner BL. One-step inactivation of chromosomal genes in
484 *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci* 2000;97:6640–5.
- 485 [17] Chagneau CV, Massip C, Bossuet-Greif N, Fremez C, Motta J-P, Shima A, et al.
486 Uropathogenic *E. coli* induces DNA damage in the bladder. *PLOS Pathog*
487 2021;17:e1009310. <https://doi.org/10.1371/journal.ppat.1009310>.
- 488 [18] Zhou Z, Alikhan N-F, Mohamed K, Fan Y, Agama Study Group, Achtman M. The
489 Enterobase user's guide, with case studies on *Salmonella* transmissions, *Yersinia*
490 *pestis* phylogeny, and *Escherichia* core genomic diversity. *Genome Res*
491 2020;30:138–52. <https://doi.org/10.1101/gr.251678.119>.
- 492 [19] Madeira F, Park Y mi, Lee J, Buso N, Gur T, Madhusoodanan N, et al. The
493 EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res*
494 2019;47:W636–41. <https://doi.org/10.1093/nar/gkz268>.

- 495 [20] Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular
496 Evolutionary Genetics Analysis across Computing Platforms. *Mol Biol Evol*
497 2018;35:1547–9. <https://doi.org/10.1093/molbev/msy096>.
- 498 [21] Katoh K, Toh H. Recent developments in the MAFFT multiple sequence
499 alignment program. *Brief Bioinform* 2008;9:286–98.
500 <https://doi.org/10.1093/bib/bbn013>.
- 501 [22] Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A. RAxML-NG: a fast,
502 scalable and user-friendly tool for maximum likelihood phylogenetic inference.
503 *Bioinforma Oxf Engl* 2019;35:4453–5.
504 <https://doi.org/10.1093/bioinformatics/btz305>.
- 505 [23] Darriba D, Posada D, Kozlov AM, Stamatakis A, Morel B, Flouri T. ModelTest-
506 NG: A New and Scalable Tool for the Selection of DNA and Protein Evolutionary
507 Models. *Mol Biol Evol* 2020;37:291–4. <https://doi.org/10.1093/molbev/msz189>.
- 508 [24] Revell LJ. phytools: an R package for phylogenetic comparative biology (and
509 other things). *Methods Ecol Evol* 2012;3:217–23. <https://doi.org/10.1111/j.2041-210X.2011.00169.x>.
- 511 [25] Pommier T, Canbäck B, Lundberg P, Hagström Å, Tunlid A. RAMI: a tool for
512 identification and characterization of phylogenetic clusters in microbial
513 communities. *Bioinformatics* 2009;25:736–42.
514 <https://doi.org/10.1093/bioinformatics/btp051>.
- 515 [26] Hung C-S, Dodson KW, Hultgren SJ. A murine model of urinary tract infection.
516 *Nat Protoc* 2009;4:1230–43. <https://doi.org/10.1038/nprot.2009.116>.
- 517 [27] Welch RA. Uropathogenic Escherichia coli-Associated Exotoxins. *Microbiol*
518 *Spectr* 2016;4. <https://doi.org/10.1128/microbiolspec.UTI-0011-2012>.
- 519 [28] Biggel M, Xavier BB, Johnson JR, Nielsen KL, Frimodt-Møller N, Matheeussen
520 V, et al. Horizontally acquired papGII-containing pathogenicity islands underlie
521 the emergence of invasive uropathogenic Escherichia coli lineages. *Nat Commun*
522 2020;11:5968. <https://doi.org/10.1038/s41467-020-19714-9>.
- 523 [29] Johnson TJ, Siek KE, Johnson SJ, Nolan LK. DNA sequence of a ColV plasmid
524 and prevalence of selected plasmid-encoded virulence genes among avian
525 Escherichia coli strains. *J Bacteriol* 2006;188:745–58.
526 <https://doi.org/10.1128/JB.188.2.745-758.2006>.
- 527 [30] Johnson TJ, Nolan LK. Pathogenomics of the Virulence Plasmids of Escherichia
528 coli. *Microbiol Mol Biol Rev* 2009;73:750–74.
529 <https://doi.org/10.1128/MMBR.00015-09>.
- 530 [31] Brodsky IE, Ernst RK, Miller SI, Falkow S. mig-14 Is a Salmonella Gene That
531 Plays a Role in Bacterial Resistance to Antimicrobial Peptides. *J Bacteriol*
532 2002;184:3203–13. <https://doi.org/10.1128/JB.184.12.3203-3213.2002>.
- 533 [32] Zhuge X, Sun Y, Xue F, Tang F, Ren J, Li D, et al. A Novel PhoP/PhoQ
534 Regulation Pathway Modulates the Survival of Extraintestinal Pathogenic
535 Escherichia coli in Macrophages. *Front Immunol* 2018;9:788.
536 <https://doi.org/10.3389/fimmu.2018.00788>.
- 537 [33] McPhee JB, Small CL, Reid-Yu SA, Brannon JR, Le Moual H, Coombes BK.
538 Host Defense Peptide Resistance Contributes to Colonization and Maximal
539 Intestinal Pathology by Crohn’s Disease-Associated Adherent-Invasive
540 Escherichia coli. *Infect Immun* 2014;82:3383–93.
541 <https://doi.org/10.1128/IAI.01888-14>.
- 542 [34] Desloges I, Taylor JA, Leclerc J-M, Brannon JR, Portt A, Spencer JD, et al.
543 Identification and characterization of OmpT-like proteases in uropathogenic

- 544 Escherichia coli clinical isolates. *MicrobiologyOpen* 2019;8:e915.
545 <https://doi.org/10.1002/mbo3.915>.
- 546 [35] Matsumura Y, Peirano G, Pitout JDD. Complete Genome Sequence of Escherichia
547 coli J53, an Azide-Resistant Laboratory Strain Used for Conjugation Experiments.
548 *Genome Announc* 2018;6:e00433-18. <https://doi.org/10.1128/genomeA.00433-18>.
- 549 [36] Peigne C, Bidet P, Mahjoub-Messai F, Plainvert C, Barbe V, Médigue C, et al. The
550 Plasmid of Escherichia coli Strain S88 (O45:K1:H7) That Causes Neonatal
551 Meningitis Is Closely Related to Avian Pathogenic E. coli Plasmids and Is
552 Associated with High-Level Bacteremia in a Neonatal Rat Meningitis Model.
553 *Infect Immun* 2009;77:2272–84. <https://doi.org/10.1128/IAI.01333-08>.
- 554 [37] Auvray F, Perrat A, Arimizu Y, Chagneau CV, Bossuet-Greif N, Massip C, et al.
555 Insights into the acquisition of the pks island and production of colibactin in the
556 Escherichia coli population. *Microb Genomics* 2021;7:000579.
557 <https://doi.org/10.1099/mgen.0.000579>.
- 558 [38] Alikhan N-F, Petty NK, Zakour NLB, Beatson SA. BLAST Ring Image Generator
559 (BRIG): simple prokaryote genome comparisons. *BMC Genomics* 2011;12:1–10.
560 <https://doi.org/10.1186/1471-2164-12-402>.
- 561 [39] Cusumano CK, Hung CS, Chen SL, Hultgren SJ. Virulence Plasmid Harbored by
562 Uropathogenic Escherichia coli Functions in Acute Stages of Pathogenesis. *Infect*
563 *Immun* 2010;78:1457–67. <https://doi.org/10.1128/IAI.01260-09>.
- 564 [40] Skyberg JA, Johnson TJ, Johnson JR, Clabots C, Logue CM, Nolan LK.
565 Acquisition of Avian Pathogenic Escherichia coli Plasmids by a Commensal E.
566 coli Isolate Enhances Its Abilities To Kill Chicken Embryos, Grow in Human
567 Urine, and Colonize the Murine Kidney. *Infect Immun* 2006;74:6287–92.
568 <https://doi.org/10.1128/IAI.00363-06>.
- 569 [41] Ewers C, Li G, Wilking H, Kiessling S, Alt K, Antão E-M, et al. Avian
570 pathogenic, uropathogenic, and newborn meningitis-causing Escherichia coli: how
571 closely related are they? *Int J Med Microbiol IJMM* 2007;297:163–76.
572 <https://doi.org/10.1016/j.ijmm.2007.01.003>.
- 573 [42] Johnson JR, Kuskowski MA, Gajewski A, Soto S, Horcajada JP, de Anta MTJ, et
574 al. Extended Virulence Genotypes and Phylogenetic Background of Escherichia
575 coli Isolates from Patients with Cystitis, Pyelonephritis, or Prostatitis. *J Infect Dis*
576 2005;191:46–50. <https://doi.org/10.1086/426450>.
- 577 [43] Owrangi B, Masters N, Kuballa A, O’Dea C, Vollmerhausen TL, Katouli M.
578 Invasion and translocation of uropathogenic Escherichia coli isolated from
579 urosepsis and patients with community-acquired urinary tract infection. *Eur J Clin*
580 *Microbiol Infect Dis* 2018;37:833–9. <https://doi.org/10.1007/s10096-017-3176-4>.
- 581 [44] Gordon DM, Geyik S, Clermont O, O’Brien CL, Huang S, Abayasekara C, et al.
582 Fine-Scale Structure Analysis Shows Epidemic Patterns of Clonal Complex 95, a
583 Cosmopolitan Escherichia coli Lineage Responsible for Extraintestinal Infection.
584 *MSphere* 2017;2:e00168-17. <https://doi.org/10.1128/mSphere.00168-17>.
- 585 [45] Liu L, Law COK, Nie Q, Pham HQ, Ma H, Zhang L, et al. Comparative analysis
586 of outer membrane vesicles from uropathogenic Escherichia coli reveal the role of
587 aromatic amino acids synthesis proteins in motility. *Int J Med Microbiol*
588 2023;313:151573. <https://doi.org/10.1016/j.ijmm.2023.151573>.
- 589 [46] Davis JM, Carvalho HM, Rasmussen SB, O’Brien AD. Cytotoxic necrotizing
590 factor type 1 delivered by outer membrane vesicles of uropathogenic Escherichia
591 coli attenuates polymorphonuclear leukocyte antimicrobial activity and
592 chemotaxis. *Infect Immun* 2006;74:4401–8. <https://doi.org/10.1128/IAI.00637-06>.

- 593 [47] Denamur E, Clermont O, Bonacorsi S, Gordon D. The population genetics of
594 pathogenic *Escherichia coli*. *Nat Rev Microbiol* 2021;19:37–54.
595 <https://doi.org/10.1038/s41579-020-0416-x>.
- 596 [48] Jørgensen SL, Stegger M, Kudirkiene E, Lilje B, Poulsen LL, Ronco T, et al.
597 Diversity and Population Overlap between Avian and Human *Escherichia coli*
598 Belonging to Sequence Type 95. *MSphere* 2019;4:e00333-18.
599 <https://doi.org/10.1128/mSphere.00333-18>.
- 600 [49] Reid CJ, Cummins ML, Börjesson S, Brouwer MSM, Hasman H, Hammerum
601 AM, et al. A role for ColV plasmids in the evolution of pathogenic *Escherichia*
602 *coli* ST58. *Nat Commun* 2022;13:683. [https://doi.org/10.1038/s41467-022-28342-](https://doi.org/10.1038/s41467-022-28342-4)
603 4.
- 604 [50] Xia F, Jiang M, Wen Z, Wang Z, Wang M, Xu Y, et al. Complete genomic
605 analysis of ST117 lineage extraintestinal pathogenic *Escherichia coli* (ExPEC) to
606 reveal multiple genetic determinants to drive its global transmission: ST117 *E. coli*
607 as an emerging multidrug-resistant foodborne ExPEC with zoonotic potential.
608 *Transbound Emerg Dis* 2022;69:3256–73. <https://doi.org/10.1111/tbed.14678>.
- 609 [51] Liu CM, Stegger M, Aziz M, Johnson TJ, Waits K, Nordstrom L, et al. *Escherichia*
610 *coli* ST131-*H* 22 as a Foodborne Uropathogen. *MBio* 2018;9:e00470-18.
611 <https://doi.org/10.1128/mBio.00470-18>.
- 612 [52] Johnson TJ. Role of Plasmids in the Ecology and Evolution of “High-Risk”
613 Extraintestinal Pathogenic *Escherichia coli* Clones. *EcoSal Plus* 2021;9:eESP-
614 0013-2020. <https://doi.org/10.1128/ecosalplus.ESP-0013-2020>.
- 615 [53] Soysal N, Mariani-Kurkdjian P, Smail Y, Liguori S, Gouali M, Loukiadis E, et al.
616 Enterohemorrhagic *Escherichia coli* Hybrid Pathotype O80:H2 as a New
617 Therapeutic Challenge. *Emerg Infect Dis* 2016;22:1604–12.
618 <https://doi.org/10.3201/eid2209.160304>.
- 619 [54] Cointe A, Birgy A, Mariani-Kurkdjian P, Liguori S, Courroux C, Blanco J, et al.
620 Emerging Multidrug-Resistant Hybrid Pathotype Shiga Toxin-Producing
621 *Escherichia coli* O80 and Related Strains of Clonal Complex 165, Europe. *Emerg*
622 *Infect Dis* 2018;24:2262–9. <https://doi.org/10.3201/eid2412.180272>.
- 623 [55] Moran RA, Hall RM. Evolution of Regions Containing Antibiotic Resistance
624 Genes in FII-2-FIB-1 ColV-Colla Virulence Plasmids. *Microb Drug Resist*
625 2018;24:411–21. <https://doi.org/10.1089/mdr.2017.0177>.
- 626 [56] Nash JH, Villegas A, Kropinski AM, Aguilar-Valenzuela R, Konczyk P,
627 Mascarenhas M, et al. Genome sequence of adherent-invasive *Escherichia coli* and
628 comparative genomic analysis with other *E. coli* pathotypes. *BMC Genomics*
629 2010;11:667. <https://doi.org/10.1186/1471-2164-11-667>.
- 630 [57] Fricke WF, McDermott PF, Mammel MK, Zhao S, Johnson TJ, Rasko DA, et al.
631 Antimicrobial Resistance-Confering Plasmids with Similarity to Virulence
632 Plasmids from Avian Pathogenic *Escherichia coli* Strains in *Salmonella enterica*
633 Serovar Kentucky Isolates from Poultry. *Appl Environ Microbiol* 2009;75:5963–
634 71. <https://doi.org/10.1128/AEM.00786-09>.
- 635 [58] Yang X, Chen H, Zheng Y, Qu S, Wang H, Yi F. Disease burden and long-term
636 trends of urinary tract infections: A worldwide report. *Front Public Health*
637 2022;10:888205. <https://doi.org/10.3389/fpubh.2022.888205>.
- 638

640 **Figure legends**

641

642 **Figure 1. HlyF increases pathogenicity during urinary tract infection in a mouse**
643 **model.**

644 C3H/HeN mice were infected transurethrally with wild-type UPEC strain ECC166
645 (WT) or *hlyF* isogenic mutant ($\Delta hlyF$).

646 **A.** Time to humane euthanasia (upon a body weight loss and/or clinical score reaching a
647 predefined threshold) was monitored to build the survival curve (panel A). The results
648 are pooled from three independent experiments, the total number of animals is shown.
649 The difference between the experimental groups was evaluated by the log-rank (Mantel-
650 Cox) test: * $p < 0.05$.

651 **B.** Clinical score according to the Table S1 at 20h \pm 2h post-inoculation. Mean values \pm
652 SEM are shown, each circle represents a mouse. The presented results are pooled from
653 two independent experiments. Student's t-test: * $p < 0.05$.

654 **C.** Bacterial load in bladder, kidney and spleen (CFU/g organ) at end point. Mean
655 values \pm SEM are shown. The presented results are pooled from two independent
656 experiments. Student's t-test: * $p < 0.05$; ** $p < 0.01$

657 **D.** IL-1 β and IL-6 in spleen (pg/mg of protein) at endpoint. Mean values \pm SEM are
658 shown. The presented results are pooled from two independent experiments. Student's t-
659 test: * $p < 0.05$; *** $p < 0.001$.

660

661 **Figure 2. *hlyF* locus is conserved and associated with ColV plasmids determinants.**

662 The *hlyF* locus from pECOS88 used for comparison is shown together with the *ompTp*
663 gene. Phylogenetic tree of the *hlyF* locus from UPEC reads and three fully sequenced
664 and available genomes of plasmids from *E. coli* S88, SP15 and Combat11I9 was
665 constructed using iTol software (<https://itol.embl.de/>). The phylogroup is indicated by a
666 coloured square (column 1). The presence of classical ColV plasmid determinants is
667 indicated by a full circle (constant determinants according to Johnson et al. [29]) and a
668 full triangle (column 2). The empty version corresponds to the presence of a truncated
669 form. The size of the circle is proportional to the number of antibiotic resistance genes
670 found in the ResFinder search (column 3). pMLST plasmid typing is indicated column
671 4.

672

673 **Figure 3. *hlyF* is carried by various mosaic plasmids combining virulence factors**
674 **and antibiotic resistance genes.** pMLST plasmid typing is indicated under the plasmid
675 name. Genes of interest (virulence, plasmid replication, antimicrobial resistance genes
676 (ARG)) are highlighted and coloured. Lines follow genes of interest throughout the
677 figure to illustrate gene presence and overall gene conservation.

678

679 **Figure 4. Plasmids carrying *hlyF* have a conserved scaffold but show variability.**
680 pECOS88 sequence is compared to both reference plasmids and Nanopore-sequenced
681 plasmids from the UPEC collection using the Blast Ring Image Generator software
682 [38]. Long tick marks on the outer and inner circumference of the ring indicate 500
683 kilobase pair increments and short tick marks indicate 100 kilobase pair increments. The
684 outer black circle corresponds to ECOS88 plasmid annotation with insertion sequences
685 highlighted in orange.

686 **Supplementary Figure S1. *hlyF* is widely disseminated in UPEC strains.** A
687 phylogenetic tree based on whole genome analysis of *hlyF*⁺ strains was constructed
688 with *E. coli* MG1655 (*hlyF*⁻) as reference. Size of the circle is proportional to the
689 number of *hlyF*⁺ strains of each ST, indicated in square brackets. Phylogroups are
690 circled.

691

692 **Supplementary Figure S2.** Body weight gain/loss was compared before and at the end
693 point of the experiment. Mean values \pm SEM are shown, each circle represents a mouse.
694 The presented results are pooled from two independent experiments. Student's t-test: *
695 $p < 0.05$.

696

697 **Supplementary Figure S3. Complementation restores the pro-pathogenic effect of**
698 **HlyF during UTI.**

699 C3H/HeN mice were infected trans-urethrally with ECC166 Δ *hlyF* transformed with a
700 plasmid carrying a wild-type *hlyF* gene (HlyF) or the *hlyF* gene with mutation in the
701 catalytic domain (SDM).

702 **A.** Time to humane euthanasia (upon a body weight loss and/or clinical score reaching a
703 predefined threshold) was monitored to build the survival curve. The difference
704 between the experimental groups was evaluated by the log-rank (Mantel-Cox) test.

705 **B.** Clinical score according to Table S1 at 20h \pm 2h post-inoculation.

706 **C.** Bacterial load in spleen (CFU/g organ) at end point.

707 **D.** Inflammatory cytokines in spleen (pg/mg proteins) at endpoint.

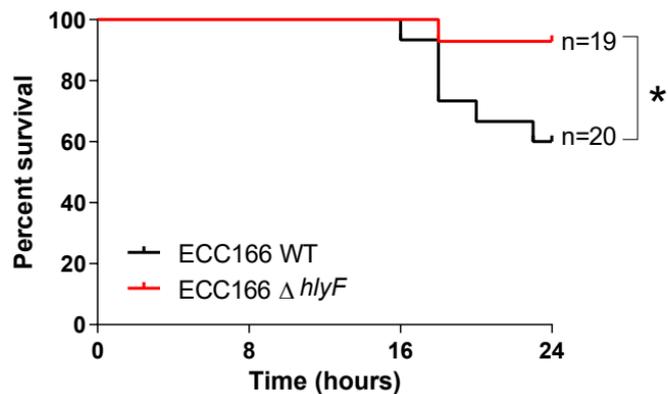
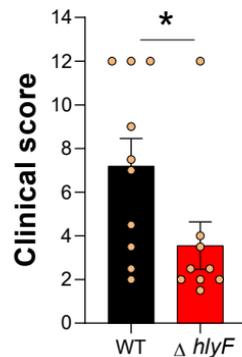
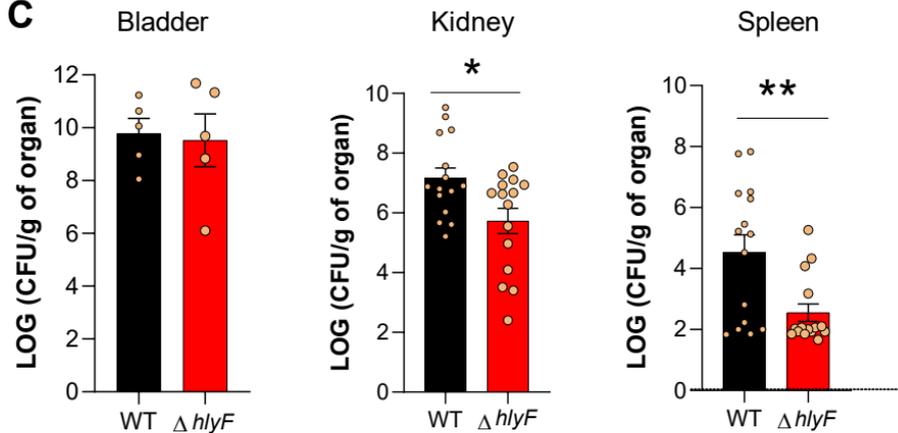
708 For B, C, D, mean values \pm SEM are shown. Each circle represents a mouse. Student's
709 t-test: * $p < 0.05$; ** $p < 0.01$.

710

711 **Supplementary Figure S4. Global structure of ECC166 pColIV conjugative plasmid**
712 **carrying *hlyF*.** Genes of interest (virulence, plasmid replication, antibiotic resistance,
713 conjugation system) are marked and coloured.

714

715 **Supplementary Figure S5. Co-phylogenetic relationship between *hlyF* and *traM-***
716 ***traX* locus.** The links between *hlyF* (left) and *traM-traX* (right) are indicated by grey
717 lines. Strain names are color-coded according to their phylogenetic clusters determined
718 by using RAMI.

A**B****C****D**