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.

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

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

Camille V. Chagneau[1,2], Delphine Payros[1], Audrey Goman[1], Cécile Goursat[1], Laure David[1], Miki Okuno[3], Pierre-Jean Bordignon[1], Carine Séguy[1], Clémence Massip[1,2], Priscilla Branchu[1], Yoshitoshi Ogura[3] Jean-Philippe Nougayrède[1], Marc Marenda[4] and Eric Oswald[1,2]

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

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

[3]*Division of Microbiology, Department of Infectious Medicine, Kurume University School of Medicine, Kurume, Japan;*

[4]*Faculty of Science, University of Melbourne, Melbourne, Australia*

*corresponding author: eric.oswald@inserm.fr

Pr Eric Oswald

Institut de recherche en santé digestive

CHU Purpan – Place Baylac

CS 60039

31024 TOULOUSE CEDEX 3

France

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

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

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

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

Introduction

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

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

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

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

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 Δ*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

Mouse model of UTI

Ethic statement

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

Mouse model

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

Body weight and clinical scoring

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

Cytokines quantification

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

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

Sequencing data, sequence alignments and phylogenetic analyses

Illumina sequencing

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

Nanopore sequencing

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

Genome analysis

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

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

Alignment, comparisons, and detection of mutations in *hlyF* locus were performed either using BioNumerics software or Clustal Omega [19] with pECOS88 as a reference (CU928146.1: 130 790… 133 302) from *E. coli* S88 serotype O45:K1:H7 isolated from neonatal meningitis and including other reference strains: *E. coli* SP15 of serotype O18:K1:H7 isolated from neonatal meningitis (AP024132.2) and *E. coli* Combat11I9 of serotype O-:H28 isolated from UTI (CP021728.1). MEGA X software was used to construct UPGMA tree of *hlyF* locus [20]. The evolutionary distances were computed using the Maximum Composite Likelihood method and are in the units of the number of base substitutions per site.

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

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

Statistical analyses

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

Results

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

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

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

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

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

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

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

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

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

	gene	**number of positive strains (% total)**	***hlyF*+ (%)**	***hlyF*- (%)**	**p(Fischer)**

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

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

hlyF *is carried by pColV conjugative plasmids.*

Assembly and annotation performed after both Nanopore and Illumina sequencing of the UPEC strain ECC166 yielded a chromosome of 4.945.664 bp and two plasmids of 115.445 bp and 1.552 bp, respectively. *hlyF* was carried by the largest plasmid belonging to pColV family (Figure S4). As previously described, *hlyF* was present in an operon found on pColV family plasmids, here referred to as *hlyF* locus, containing the *hlyF* gene and a putative *mig-14* ortholog gene (*mig-14-like*) which encodes an antimicrobial resistance factor [31–33] (Figure 2). This locus was systematically associated with *ompTp* gene, which encodes for plasmid variant of OmpT, a protease with an activity against host antimicrobial peptides [10,30,32,34]. In all UPEC strains, *hlyF* locus was highly conserved with pairwise similarities of at least 98% between two loci (Figure 2, 3 and 4). The *hlyF* gene itself showed even more similarity amongst the strains with the presence of only 11 SNP sites within the 42 strains, 6 of which generated silent mutations (Table S2). The predicted catalytic domain (YTHSK) and NAD binding domain (GATGFLG) of HlyF [10] were conserved in all strains.

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

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

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

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*

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

HlyF: an emerging virulence factor associated with severe UTI

The gene encoding HlyF has been previously described in APEC and NMEC strains [12–14] but has never been studied in UPEC, probably because *hlyF* is absent from the genome of archetypal UPEC strains such as *E. coli* CFT073, 536 or UTI89. We have observed in our representative UPEC collection that almost 20% of strains isolated from community-acquired infections carry this virulence factor. This frequency is confirmed by data from the literature based on the epidemiology of genes classically associated with *hlyF* (*ompTp*, *cvaC*...) [28,34,41–43]. *hlyF*+ strains seem especially frequent in UPEC belonging to ST95. Interestingly, in larger collections, some subgroups of ST95 strains (C and D) are much more likely to carry pColV plasmids and are becoming more prevalent in patients with UTI, suggesting an on-going dissemination [44]. These strains are mostly devoid of classical UPEC virulence genes such as *hlyA* or *cnf1* [28], which

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

A role of OMVs in urosepsis

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

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

HlyF displays wide dissemination in UPEC, suggesting that *hlyF*, and/or other 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 pathovars
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**

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

Authors contribution statement

Conceptualization: CVC, JPN, EO

Data curation: CVC, CM

Formal analysis: CVC, DP

Funding acquisition: CVC, JPN, EO

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

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

Project administration: EO

Resources: AG, PJB, CM, PB, MM

Visualization: CVC, DP, MO

Writing –original draft: CVC, DP

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

References

[1] Flores-Mireles AL, Walker JN, Caparon M, Hultgren SJ. Urinary tract infections: epidemiology, mechanisms of infection and treatment options. Nat Rev Microbiol 2015;13:269–84. https://doi.org/10.1038/nrmicro3432.

[2] Öztürk R, Murt A. Epidemiology of urological infections: a global burden. World J Urol 2020;38:2669–79. https://doi.org/10.1007/s00345-019-03071-4.

[3] Wagenlehner F, Tandogdu Z, Bartoletti R, Cai T, Cek M, Kulchavenya E, et al. The Global Prevalence of Infections in Urology Study: A Long-Term, Worldwide Surveillance Study on Urological Infections. Pathog Basel Switz 2016;5:10. https://doi.org/10.3390/pathogens5010010.

[4] Foxman B. The epidemiology of urinary tract infection. Nat Rev Urol 2010;7:653–60. https://doi.org/10.1038/nrurol.2010.190.

[5] Bonkat G, Cai T, Veeratterapillay R, Bruyère F, Bartoletti R, Pilatz A, et al. Management of Urosepsis in 2018. Eur Urol Focus 2019;5:5–9. https://doi.org/10.1016/j.euf.2018.11.003.

[6] Schreiber HL, Conover MS, Chou W-C, Hibbing ME, Manson AL, Dodson KW, 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-
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-
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.

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

[47] Denamur E, Clermont O, Bonacorsi S, Gordon D. The population genetics of pathogenic Escherichia coli. Nat Rev Microbiol 2021;19:37–54. https://doi.org/10.1038/s41579-020-0416-x.

[48] Jørgensen SL, Stegger M, Kudirkiene E, Lilje B, Poulsen LL, Ronco T, et al. Diversity and Population Overlap between Avian and Human Escherichia coli Belonging to Sequence Type 95. MSphere 2019;4:e00333-18. https://doi.org/10.1128/mSphere.00333-18.

[49] Reid CJ, Cummins ML, Börjesson S, Brouwer MSM, Hasman H, Hammerum AM, et al. A role for ColV plasmids in the evolution of pathogenic Escherichia coli ST58. Nat Commun 2022;13:683. https://doi.org/10.1038/s41467-022-28342-4.

[50] Xia F, Jiang M, Wen Z, Wang Z, Wang M, Xu Y, et al. Complete genomic analysis of ST117 lineage extraintestinal pathogenic Escherichia coli (ExPEC) to reveal multiple genetic determinants to drive its global transmission: ST117 E. coli as an emerging multidrug-resistant foodborne ExPEC with zoonotic potential. Transbound Emerg Dis 2022;69:3256–73. https://doi.org/10.1111/tbed.14678.

[51] Liu CM, Stegger M, Aziz M, Johnson TJ, Waits K, Nordstrom L, et al. Escherichia coli ST131- *H* 22 as a Foodborne Uropathogen. MBio 2018;9:e00470-18. https://doi.org/10.1128/mBio.00470-18.

[52] Johnson TJ. Role of Plasmids in the Ecology and Evolution of "High-Risk" Extraintestinal Pathogenic Escherichia coli Clones. EcoSal Plus 2021;9:eESP-0013-2020. https://doi.org/10.1128/ecosalplus.ESP-0013-2020.

[53] Soysal N, Mariani-Kurkdjian P, Smail Y, Liguori S, Gouali M, Loukiadis E, et al. Enterohemorrhagic Escherichia coli Hybrid Pathotype O80:H2 as a New Therapeutic Challenge. Emerg Infect Dis 2016;22:1604–12. https://doi.org/10.3201/eid2209.160304.

[54] Cointe A, Birgy A, Mariani-Kurkdjian P, Liguori S, Courroux C, Blanco J, et al. Emerging Multidrug-Resistant Hybrid Pathotype Shiga Toxin-Producing Escherichia coli O80 and Related Strains of Clonal Complex 165, Europe. Emerg Infect Dis 2018;24:2262–9. https://doi.org/10.3201/eid2412.180272.

[55] Moran RA, Hall RM. Evolution of Regions Containing Antibiotic Resistance Genes in FII-2-FIB-1 ColV-Colla Virulence Plasmids. Microb Drug Resist 2018;24:411–21. https://doi.org/10.1089/mdr.2017.0177.

[56] Nash JH, Villegas A, Kropinski AM, Aguilar-Valenzuela R, Konczy P, Mascarenhas M, et al. Genome sequence of adherent-invasive Escherichia coli and comparative genomic analysis with other E. coli pathotypes. BMC Genomics 2010;11:667. https://doi.org/10.1186/1471-2164-11-667.

[57] Fricke WF, McDermott PF, Mammel MK, Zhao S, Johnson TJ, Rasko DA, et al. Antimicrobial Resistance-Conferring Plasmids with Similarity to Virulence Plasmids from Avian Pathogenic Escherichia coli Strains in Salmonella enterica Serovar Kentucky Isolates from Poultry. Appl Environ Microbiol 2009;75:5963–71. https://doi.org/10.1128/AEM.00786-09.

[58] Yang X, Chen H, Zheng Y, Qu S, Wang H, Yi F. Disease burden and long-term trends of urinary tract infections: A worldwide report. Front Public Health 2022;10:888205. https://doi.org/10.3389/fpubh.2022.888205.

Figure legends

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

C3H/HeN mice were infected transurethrally with wild-type UPEC strain ECC166 (WT) or *hlyF* isogenic mutant *(*Δ*hlyF*).

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

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

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

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

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

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

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.

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

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

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

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

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

B. Clinical score according to Table S1 at 20h+/-2h post-inoculation.

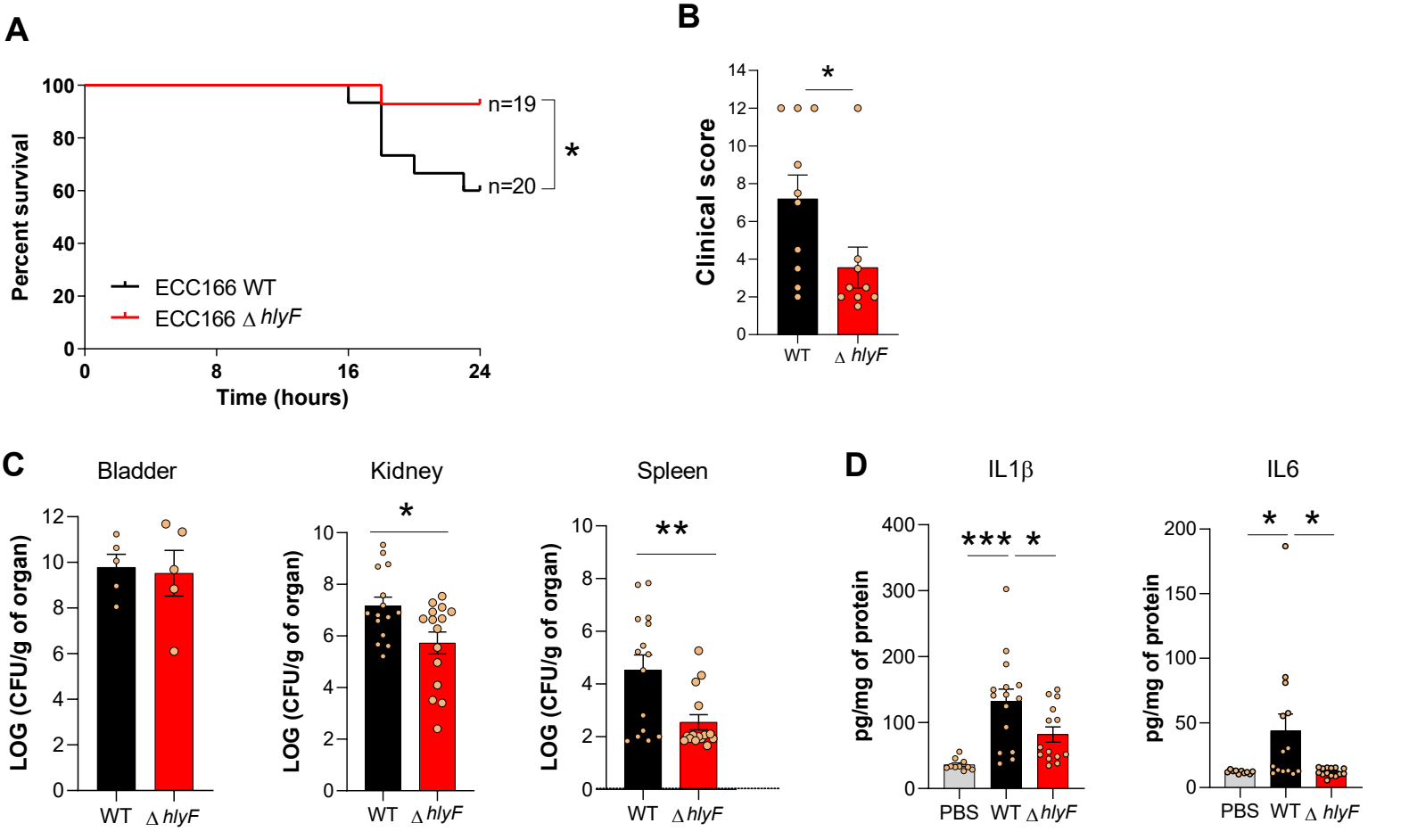
C. Bacterial load in spleen (CFU/g organ) at end point.

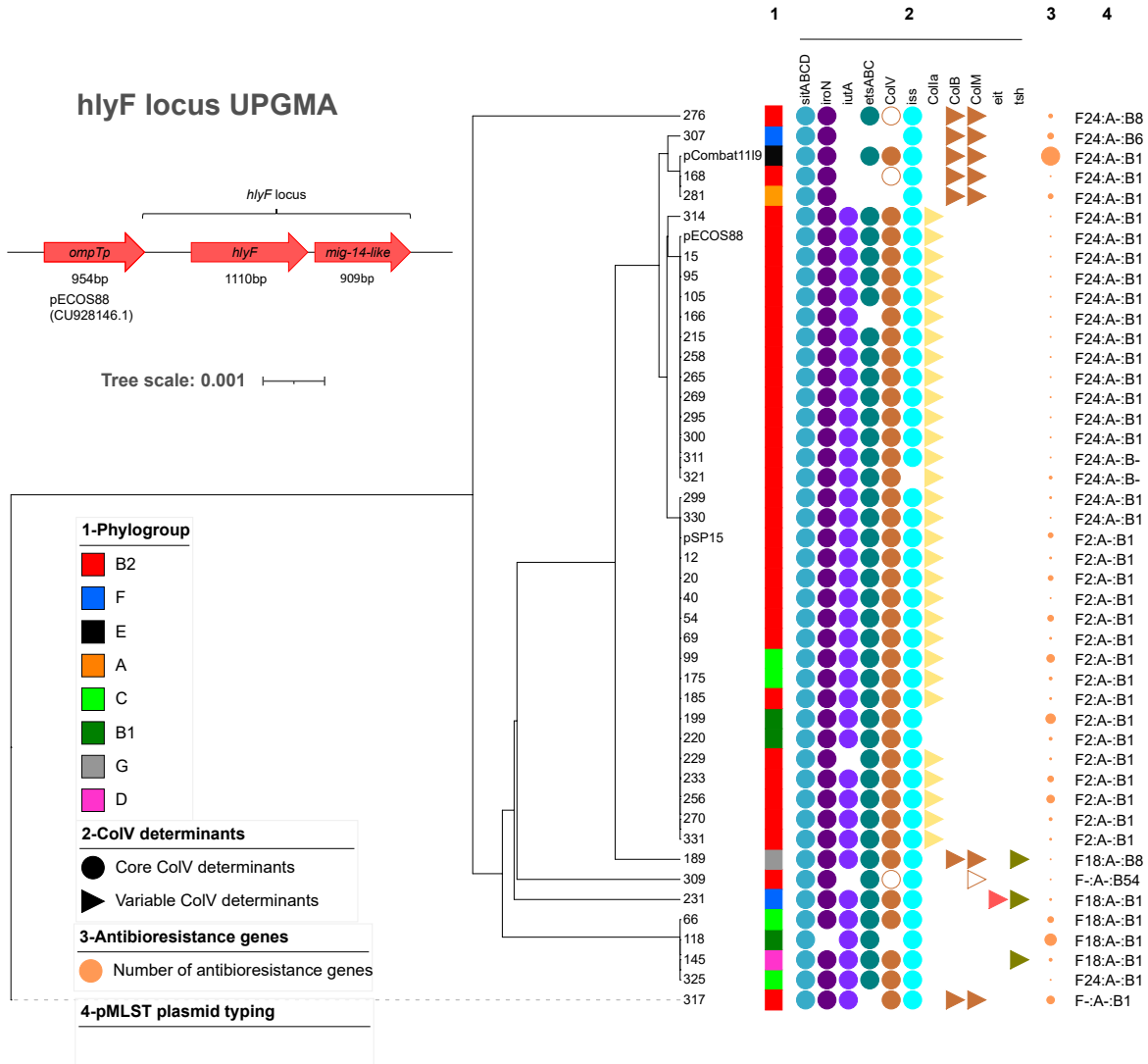
D. Inflammatory cytokines in spleen (pg/mg proteins) at endpoint.

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

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

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


A
Percent survival
Time (hours)
n=19
n=20
*
ECC166 WT
ECC166 Δ hlyF
B
Clinical score
*
WT
Δ hlyF
C
Bladder
Kidney
Spleen
LOG (CFU/g of organ)
*
**
WT
Δ hlyF
D
IL1β
IL6
pg/mg of protein
***
PBS
WT
Δ hlyF


hlyF locus UPGMA
hlyF locus
ompTp
hlyF
mig-14-like
954bp
1110bp
909bp
pECOS88
(CU928146.1)
Tree scale: 0.001
1
2
3
4
sitABCD
iroN
iutA
etsABC
ColV
iss
Colla
ColB
ColM
eit
tsh
276
307
pCombat11I9
168
281
314
pECOS88
15
95
105
166
215
258
265
269
295
300
311
321
299
330
pSP15
12
20
40
54
69
99
175
185
199
220
229
233
256
270
331
189
309
231
66
118
145
325
317
F24:A-:B8
F24:A-:B6
F24:A-:B1
F24:A-:B1
F24:A-:B1
F24:A-:B1
F24:A-:B1
F24:A-:B1
F24:A-:B1
F24:A-:B1
F24:A-:B1
F24:A-:B1
F24:A-:B1
F24:A-:B1
F24:A-:B1
F24:A-:B1
F24:A-:B-
F24:A-:B-
F24:A-:B1
F24:A-:B1
F2:A-:B1
F2:A-:B1
F2:A-:B1
F2:A-:B1
F2:A-:B1
F2:A-:B1
F2:A-:B1
F2:A-:B1
F2:A-:B1
F2:A-:B1
F2:A-:B1
F2:A-:B1
F2:A-:B1
F2:A-:B1
F2:A-:B1
F18:A-:B8
F-:A-:B54
F18:A-:B1
F18:A-:B1
F18:A-:B1
F18:A-:B1
F24:A-:B1
F-:A-:B1
1-Phylogroup
B2
F
E
A
C
B1
G
D
2-ColV determinants
Core ColV determinants
Variable ColV determinants
3-Antibioresistance genes
Number of antibioresistance genes
4-pMLST plasmid typing