JCI insight

# *Escherichia coli* catheter-associated urinary tract infections are associated with distinctive virulence and biofilm gene determinants

Zongsen Zou, … , Gautam Dantas, Jeffrey P. Henderson

*JCI Insight.* 2022. https://doi.org/10.1172/jci.insight.161461.

**Research** In-Press Preview **Infectious disease** **Microbiology**

Urinary catheterization facilitates urinary tract colonization by *Escherichia coli* and increases infection risk. Here we aimed to identify strain-specific characteristics associated with the transition from colonization to infection in catheterized patients. In a single-site study population, we compared *E. coli* isolates from patients with catheter-associated asymptomatic bacteriuria (CAASB) to those with catheter-associated urinary tract infection (CAUTI). CAUTI isolates were dominated by a phylotype B2 subclade containing the multidrug resistant ST131 lineage relative to CAASB isolates, which were phylogenetically more diverse. A distinctive combination of virulence-associated genes was present in the CAUTI-associated B2 subclade. Catheter-associated biofilm formation was widespread among isolates and did not distinguish CAUTI from CAASB strains. Preincubation with CAASB strains could potently inhibit catheter colonization by multiple ST131 CAUTI isolates. Comparative genomic analysis identified a group of variable genes associated with high catheter-biofilm formation present in both CAUTI and CAASB strains. Among these, ferric citrate transport (Fec) system genes were experimentally associated with enhanced catheter biofilm formation using reporter and fecA deletion strains. Together, these results are consistent with a variable role for catheter biofilm formation in promoting CAUTI by ST131-like strains or resisting CAUTI by lower risk strains that engage in niche exclusion.



# *Escherichia coli* catheter-associated urinary tract infections are associated with distinctive virulence and biofilm gene determinants

Zongsen Zou[1,2,#], Robert F. Potter[3,4], William H. McCoy IV[1,5], John A. Wildenthal[1,2], George L. Katumba[1,2], Peter J. Mucha[6], Gautam Dantas[3,4,7,8], Jeffrey P. Henderson[1,2*]

[1]Center for Women's Infectious Diseases Research, Washington University School of Medicine, St. Louis, Missouri, USA.

[2]Department of Internal Medicine, Division of Infectious Diseases, Washington University School of Medicine, St. Louis, Missouri, USA.

[3]The Edison Family Center for Genome Sciences and Systems Biology, Washington University School of Medicine, St. Louis, Missouri, USA.

[4]Department of Pathology and Immunology, Washington University School of Medicine, St. Louis, Missouri, USA.

[5]Department of Internal Medicine, Division of Dermatology, Washington University School of Medicine, St. Louis, Missouri, USA.

[6]Department of Mathematics, Dartmouth College, Hanover, New Hampshire, USA.

[7]Department of Molecular Microbiology, Washington University School of Medicine, St. Louis, Missouri, USA.

[8]Department of Biomedical Engineering, Washington University in St. Louis, Missouri, USA.

[#]Current Address: Center for Women's Infectious Diseases Research, Department of Molecular Microbiology, Washington University School of Medicine, St. Louis, Missouri, USA.

*To whom correspondence should be addressed: Jeffrey P. Henderson: Box 8051 Washington University School of Medicine, 660 S. Euclid Ave., St. Louis, MO 63110; hendersonj@wustl.edu; Tel. (314)362-7250; Fax. (314)362-3203

**ABSTRACT**

Urinary catheterization facilitates urinary tract colonization by *Escherichia coli* and increases infection risk. Here we aimed to identify strain-specific characteristics associated with the transition from colonization to infection in catheterized patients. In a single-site study population, we compared *E. coli* isolates from patients with catheter-associated asymptomatic bacteriuria (CAASB) to those with catheter-associated urinary tract infection (CAUTI). CAUTI isolates were dominated by a phylotype B2 subclade containing the multidrug resistant ST131 lineage relative to CAASB isolates, which were phylogenetically more diverse. A distinctive combination of fvirulence-associated genes was present in the CAUTI-associated B2 subclade. Catheter-associated biofilm formation was widespread among isolates and did not distinguish CAUTI from CAASB strains. Preincubation with CAASB strains could potently inhibit catheter colonization by multiple ST131 CAUTI isolates. Comparative genomic analysis identified a group of variable genes associated with high catheter-biofilm formation present in both CAUTI and CAASB strains. Among these, ferric citrate transport (Fec) system genes were experimentally associated with enhanced catheter biofilm formation using reporter and *fecA* deletion strains. Together, these results are consistent with a variable role for catheter biofilm formation in promoting CAUTI by ST131-like strains or resisting CAUTI by lower risk strains that engage in niche exclusion.

## INTRODUCTION

Catheter-associated urinary tract infections (CAUTI) are among the most common nosocomial infections, with over one million cases annually in the United States (1-3). Accurate diagnosis and effective treatment of CAUTI and ureteral stent-associated infections, can be challenging (4). Bacteriuria alone is an insufficient criterion to establish a CAUTI diagnosis, which also requires attributable patient signs or symptoms such as suprapubic tenderness, flank pain, or fever (5, 6). For patients with catheter-associated asymptomatic bacteriuria (CAASB) who are at low risk of serious infection, antibiotics are not recommended (7). When clinical status plausibly masks symptoms or symptoms are not clearly attributable to the urinary tract, physicians must weigh the risk of progressive infection against the risks of catheter or device removal and inappropriate antibiotic therapy. In this context, professional society guidelines have long noted a need to better discriminate CAASB from CAUTI and to predict a patient's risk for progression to CAUTI (8, 9).

The pathogenic potential of any given bacterial strain is a function of both host and bacterial characteristics (10). In the urinary tract, the presence of a catheter or stent is an especially influential host characteristic, conferring a well-recognized predisposition to bacterial colonization and infection (2, 3, 11). By affecting urinary flow, providing an abiotic surface for bacterial adherence, and changing the local epithelium (12-14), these devices are associated with a distinctive pathophysiology. The ability of *E. coli* to form biofilms is generally regarded as an important virulence characteristic in catheterized patients (15, 16). Biofilms are adherent bacterial communities enmeshed in an extracellular polymeric substance (EPS) matrix that form in response to specific environmental cues, permitting a resident bacterial population to expand and persist in the urinary tract lumen. In the laboratory, a single *E. coli* strain can form qualitatively and quantitatively distinctive biofilms depending upon media composition, temperature, and flow conditions (17, 18).

*Escherichia coli* is the predominant bacterial species associated with asymptomatic bacteriuria, uncomplicated UTI, CAASB, and CAUTI (19). Unlike enteric pathotypes, there is no definitive genetic signature of a “uropathogenic” *E. coli* strain. Studies have identified *E. coli* characteristics that are common in the setting of infection but none that are definitive of a uropathogenic pathotype, consistent with the view that the uropathogenic potential of *E. coli* is multifactorial in nature (20). The virulence factors (VFs) identified to date have mostly been studied in uncomplicated UTIs, are functionally diverse, and may contribute to pathogenic potential differently in catheterized patients. Rather than achieving a strict monogenic definition for uropathogenic *E. coli,* data from uncomplicated UTIs have been most consistent with a probabilistic and combinatorial relationship between virulence determinants and disease (19, 21).

In the present study, we sought to compare *E. coli* strain characteristics between patients with CAUTI and CAASB. Intrinsically, asymptomatic isolates are more difficult to find in the clinical setting, as they must be drawn in the absence of attributable symptoms. We were able to identify CAASB isolates, along with CAUTI isolates, from a previously described observational cohort study (5, 6). Comparisons were based on whole genome sequencing analyses and quantitative biofilm phenotyping using a simulated catheter-biofilm system. Comparative genomic analyses were used in conjunction with network community analysis to identify gene combinations associated with infection and catheter biofilm formation. CAUTI strains were associated with sequence type 131, a lineage with high antibiotic resistance and distinctive virulence genes. Using a competitive catheter biofilm assay, we identified a subset of CAASB isolates capable of preventing colonization by CAUTI-associated, ST131 isolates. Multiple gene communities were associated with high catheter biofilm formation from comparative genomic analysis. Finally, we used a transcriptional reporter and a reverse bacterial genetic approach to

functionally connect the ferric citrate uptake system (Fec), which exhibited strongest relationship with catheter biofilm in genomic comparison analysis, to *E. coli* biofilm formation.

## RESULTS

### *E. coli isolates*

To compare *E. coli* strains associated with catheter-associated urinary tract infections (CAUTI) or catheter-associated asymptomatic bacteriuria (CAASB) in hospitalized patients, we identified 62 catheter-associated isolates (18.4%) from a previously described collection of 337 urinary isolates from patients at Barnes-Jewish Hospital/Washington University Medical Center between August 1st, 2009 and July 31, 2010 (5, 6). Of these 62 isolates, 12 met symptom criteria for CAUTI (concurrent fever, T > 38℃) and 16 met criteria for CAASB (lack of fever or other clinical symptoms). As an additional comparator group, 13 *E. coli* isolates corresponding to asymptomatic rectal colonization were collected by rectal swabs from healthy, adult volunteers at Barnes-Jewish Hospital/Washington University Medical Center from 2014-2015, designated as rectal colonizers (RC) (**Table 1**). In total, 41 *E. coli* isolates were collected for this study, with each isolate from a unique catheterized patient or healthy volunteer. CAUTI and CAASB subjects were of similar age and BMI but exhibited a significant sex difference ($P = 0.0093$). Bacteriuric inpatient subjects were older than non-hospitalized asymptomatic RC subjects ($P = 0.0309$), typical of inpatients in the United States (22). Moreover, *E. coli* strains isolated from urinary bacteriuria were identified with higher trimethoprim/sulfamethoxazole (TMP/SX) and quinolone resistance in clinical laboratory tests ($P < 0.01$), consistent with multidrug resistance facilitating urinary colonization in catheterized patients.

### *Phylogenomic analysis*

We characterized the genome composition of all 41 *E. coli* isolates using a whole-genome sequencing approach. Of the 15,993 genes identified in the pan-genome of these isolates, 2458 were identified in 100% of isolates, 3014 in 98% (40/41) of isolates, and over 4030 in 50% or fewer isolates. Each isolate was recognized as genetically distinct by comparing their genome differences, without clonal pairs, demonstrated by pan-genome sizes ranging from 4320 to 5983

genes (**Supplementary Table S1**) as well as their phylogenetic differences (**Fig. 1**). A maximum-likelihood tree in Fig. 1 visualizes the similarities and differences in gene content between the isolates. This unsupervised hierarchical phylogenetic clustering divides isolates into four main clades corresponding to the canonical *E. coli* phylotypes B2, F, D and a combination of A, B1, and E (23). Nearly all CAUTI isolates (11/12) belonged to phylotype B2 as is typical of extraintestinal *E. coli* (2, 3). CAASB strains were more broadly distributed among all detected phylotypes with the exception of phylotype E. RC isolates were distributed among phylotypes A, B2, and F. Of note, CAUTI strains clustered at the extreme of the phylogenetic distribution, corresponding to a subclade within phylotype B2 (designated as B2a), that disproportionately contains CAUTI isolates when compared to other B2 strains (designated B2b) (10/15 vs. 1/13, $P = 0.0021$, two-tailed Fisher's exact test). The B2a subclade consisted of 14 ST131 strains (24, 25), while the non-B2a isolates (designated as B2b) were more diverse and consisted of eight sequence types (ST 12, 73, 95, 127, 141, 144, 357, 538). The prototypical non-CAUTI model strains UTI89 and CFT073 were not associated with B2a (**Fig. 1)**. Sparse principal components analysis (sPCA) (26, 27) of B2 strain genome composition similarly distinguished B2a from B2b strains, with clear separation on the PC1 (31%) in the score plot (**Supplementary Fig. S1a&b**). Classification of these B2 subclades by logistic regression using PC1 values yielded a prediction accuracy of 1.0 (**Supplementary Fig. S1c**, SD = 0) and an AUC of 1.0 (**Supplementary Fig. S1d**, SD = 0) in five-fold cross validation. The pan-genome sizes of B2a subclade isolates range from 4578 to 5983 genes (**Supplementary Table S2**). Pairwise core-genome alignment comparison among B2a isolates identified 34 to 2784 single-nucleotide polymorphisms (**Supplementary File 1**). These results demonstrate genetic differences among B2a isolates that are inconsistent with clonal pairs. Together, these analyses identify robust, systematic differences in *E. coli* gene composition in the CAUTI-associated B2a subclade.

### *Antibiotic resistance*

The ST131 isolates that dominate B2a are a globally emergent extraintestinal pathogenic *E. coli* lineage associated with multidrug resistance, most notably to the fluoroquinolone class of antibiotics (24, 25). To determine whether increased antibiotic resistance is associated with B2a, we assessed antibiotic resistance gene (ARG) content and phenotypic resistance reported by the clinical laboratory. ARGs against aminoglycosides, beta-lactams, amphenicols, TMP/SX, macrolides/lincosamides/streptogramins (MLS), quinolones, and tetracyclines were identified from the genome assembly (**Supplementary Table S3**). Both phenotypic and genotypic fluoroquinolone resistance were more common in B2a than B2b strains (**Table 2**, $P = 0.0001$). Specific SNPs previously associated with fluoroquinolone resistance in H30 subclones of ST131 strains (*gyrA* D87N, S83L; *parC* E84V, S80I; *parE* I529L) were nearly ubiquitous in B2a isolates (**Table 2**, $P = 0.0001$). Moreover, B2a group strains exhibited higher frequencies of ARGs associated with TMP/SX, beta-lactams, and aminoglycosides (**Supplementary Table S3**, $P < 0.03$). Together, the high frequency of resistance genes, particularly those related to fluoroquinolones, is consistent with previously described ST131 isolates (24, 25).

### *Virulence factor content of CAUTI strains*

In this cohort, we hypothesized that some variable genes carried by B2a isolates enhance pathogenic potential in catheterized patients. Virulence factor genes (VFs) previously associated with pathogenic gains of function were identified from a list derived from bacterial pathogenesis literature (28). We identified 32 such VFs in our isolates (**Fig. 1**). The number of VFs per isolate, previously called the “virulence score” (27), did not distinguish ($P = 0.147$) CAUTI (9.5 ± 3.8), CAASB (8.8 ± 3.6), or RC (9.5 ± 4.3) isolates. B2 strains exhibited higher virulence scores than non-B2 strains (10.7 ± 2.7 vs. 5.8 ± 4.0, $P = 0.002$), with a non-significant trend toward a lower score in B2a than B2b (10.0 ± 2.4 vs. 11.5 ± 3.0, $P = 0.138$) (**Supplementary Fig. S2**).

We next considered that VFs may influence pathogenic potential in a non-equivalent manner that includes additive or synergistic VF combinations (28). To assess this, we used network community detection to identify co-associations between the 32 VFs. We used modularity-based community detection using the Louvain method on a weighted network of positive correlations to assess correlations between the 26 VFs that were present more than once (> 2.4%) among our isolates (21, 29). Three prominent gene communities were resolved, as visualized by the force-directed network layout (**Fig. 2a**) and its corresponding correlation matrix (**Fig. 2b**). Each community was composed of functionally diverse VFs, with iron acquisition systems and toxins prominent in communities 1 and 2, and adhesins in all three. These communities are consistent with a subgroup of VFs that additively or synergistically influence pathogenic potential, though it is also possible that these are lineage markers with no influence on human pathogenicity.

Notably, community 1 and 2 VFs were more common in phylotype B2 strains, while community 3 VFs were exclusively associated with non-B2 strains, with VFs *fyuA*, *chuA*, *ompT*, and *usp* (30-32) being nearly ubiquitous and more common in B2 strains ($P < 0.00002$). In addition, differential relationships between VFs in groups B2a and B2b were evident, with community-1 VFs *iucD, sat*, and *iha* being more common in B2a than B2b ($P < 0.008$) (33-35). Together, these results are consistent with one VF subgroup that increases the pathogenic potential of B2 strains and another VF subgroup that more specifically increases the pathogenic potential of group B2a strains in catheterized patients.

***Biofilm formation by CAUTI, CAASB, and rectal isolates***

Catheter biofilm formation is regarded as an important contributor to *E. coli* pathogenic potential. *E. coli* strains form biofilms that vary in important ways depending on media and available substrates (16, 18). To evaluate this experimentally, we compared biofilm formation between all 41 *E. coli* isolates using an *ex vivo*, continuous flow model that simulates the clinical catheter

environment in patients (**Supplementary Fig. S3**) (36). In this model, we used an artificial urine medium (AUM) (37, 38) that yielded growth kinetics (AUM vs human urine = 7.84 ± 0.19 vs 8.06 ± 0.07, $P$ = 0.3, Mann-Whitney test) and biofilm morphology (**Fig. 3a&b**) comparable to filter-sterilized human urine. Substantial inter-strain variation in catheter biofilm formation was evident between isolates, with CV retention (biofilm biomass) ranging from 0.02 to 10.60 $A_{595}/cm^2$, and adherent CFUs (sessile bacteria in biofilm matrix) from 0 to $10^{7.1}$ CFU/(mL.$cm^2$). Thirty-four of the 41 strains yielded detectable adherent CFUs. Phylotype B2 isolates exhibited significantly higher adherent CFU values (**Fig. 3d**, $P$ = 0.02) with a non-significant trend (**Fig. 3c**, $P$ = 0.22) toward higher CV retention. Neither adherent CFUs nor CV retention values significantly distinguished groups B2a and B2b (**Fig. 3f&g**, $P$ > 0.6), though it is possible the sample size was insufficient to detect a difference. Planktonic CFUs (planktonic bacteria in voided media) were not significantly different in all group-wise comparisons (**Fig. 3e&h**, $P$ > 0.1) and were not associated with adherent CFU or CV retention values, possibly reflecting bacterial persistence within loosely adherent communities and/or turbulent flow. Together, these results are consistent with widespread potential for catheter-biofilm formation among *E. coli* with greater biofilm population size in phylotype B2. Despite the CAUTI and VF gene associations, biofilm formation by B2a strains was indistinguishable from B2b strains.

***CAASB strains can inhibit CAUTI colonization***

The association between CAUTI-associated strains and the subclade B2a genotype, but not biofilm phenotype, may reflect the permissiveness of urinary catheter surfaces for bacterial colonization. In this context, we considered that biofilm-forming isolates with low pathogenic potential from CAASB subjects prevent colonization by B2a strains. A protective role for such *E. coli* strains is suggested by previous studies (39, 40). This concept has been experimentally tested in patients using *E. coli* 83972, a ST73 strain in the B2b subclade (**Fig. 1**) (41) that was isolated from a patient with persistent asymptomatic bacteriuria. *E. coli* 83972 has shown

efficacy in preventing clinical UTIs following bladder pre-colonization in patients (42, 43). To determine whether CAASB strains can also prevent colonization by CAUTI strains, we performed a two-strain competition assay using the continuous flow catheter model. In these experiments, the catheter surface was pre-colonized by a non-ST131 CAASB biofilm and subsequently challenged with a ST131 CAUTI isolate. We subsequently quantified the ST131 strains in catheter biofilm and planktonic bacterial populations using single-nucleotide polymorphism-selective qPCR (SNPs-qPCR) to distinguish them from competing non-ST131 strains.

We first assessed the ability of 11 non-ST131, non-B2a strain CAASB isolates to inhibit catheter colonization by ST131 CAUTI isolate EC20, a high-biofilm former. Total CFUs, representing both strains, were similar between different pairwise bacterial competition cultures for both catheter and planktonic populations (**Supplementary Fig. 4**). Pre-colonization with EC36 (phylotype B2) and EC25 (phylotype A) significantly inhibited both catheter-adherent (**Fig. 4a**, $P < 0.005$) and planktonic EC20 populations as assessed by SNPs-qPCR (**Fig. 4b**, $P < 0.006$). Pre-colonization with EC36 or EC25 also suppressed both catheter biofilm (**Fig. 4c**, $P < 0.0001$) and planktonic representation (**Fig. 4d**, $P < 0.0001$) of all ten CAUTI ST131 isolates. These results are consistent with the ability of a subset of CAASB strains (2/11, 18%) to markedly prevent catheter colonization and shedding by ST131 CAUTI strains with antibiotic resistance and elevated pathogenic potential (41). This raises the possibility that *E. coli* catheter biofilm formation in asymptomatic patients may play a protective role by preventing colonization with *E. coli* of greater pathogenic potential.

***Identification of catheter biofilm-associated genes***

Genomic and phenotypic results suggest that catheter biofilm formation may play a role in both promoting and preventing CAUTI, depending upon the presence of specific virulence gene

combinations. To determine whether there also exists a distinctive set of catheter biofilm-associated genes, we conducted a comparative genomic analysis of strains with high or low catheter biofilm formation. We selected isolates with high and low biofilm formation from each of the five main clades (**Fig. 1**), B2a, B2b, F, D, and A + B1, based on their CV retention (biofilm biomass) values **(Table 3)**. The criteria of "CV < 0.2" and "CV > 1" were adopted for low and high biofilm formers, which identified 13 high and 16 low biofilm isolates, respectively (**Table 3**). We next compared genome composition between these two groups using sparse partial least squares discriminant analysis (sPLSDA) (44). In the sPLSDA score plot, high and low biofilm formers were well-resolved along the PC1 axis (**Fig. 5a**). Classification between high and low biofilm formers by logistic regression using PC1 (9%, **Supplementary Fig. S5a**) values yielded a prediction accuracy of 1.0 (**Supplementary Fig. S5b**, SD = 0) and an AUC of 1.0 (**Supplementary Fig. S5c**, SD = 0) with 5-fold cross validation. Seventy-two genes with varied functional associations and significant PC1 loadings ($P < 0.05$ by two-tailed Fisher's exact test) were detected, with 46 and 26 genes associated with positive (high-biofilm) and negative (low-biofilm), respectively (**Fig. 5b, Supplementary Table S4**). Of note, the Antigen 43 gene (*flu*) (45) was among the positively associated genes, with the 44th highest PC1 loading ($P = 0.04$ by two-tailed Fisher's exact test), providing a confirmatory point of reference to a previously described *E. coli* biofilm-associated gene.

To identify co-associations between the 46 genes associated with high catheter biofilms, we performed modularity-based community detection using the Louvain method on a weighted network of positive correlations as described above for VFs (21, 29). This analysis resolved six gene communities, visualized by the force-directed network layout (**Fig. 5c**) and its corresponding correlation matrix (**Fig. 6**). Community 1 was composed of the ferric citrate transport (*fecABCDIR*) locus (46, 47) and exhibited robust co-associations and the highest betweenness centrality ranking in the weighted network. *fec* genes also exhibited the strongest

association with high catheter biofilm formation in the sPLSDA analysis ($P$ = 0.0025), suggesting a major role in this biofilm phenotype. Community 2 defined the aerobactin siderophore system locus (33) represented by the VF marker gene *iucD*. Relative to communities 1 and 2, communities 3–6 were less robust and are composed of genes associated with more divergent functions.

Together, these results are consistent with a complex polygenic contribution to catheter biofilm production that features a prominent role for iron transport systems (ferric citrate and aerobactin systems, *cir*) and autoaggregation (Antigen 43/*flu*) (**Supplementary Table S4**). The significant positive associations between multiple biofilm genes, B2, and B2a strains (**Fig. 6**) suggests that biofilm-associated genes may play a contributing role in the pathogenic potential of CAUTI isolates.

***Fec system activity is associated with increased catheter biofilm formation***

The positive association between catheter biofilm and *fec* genes raises the possibility that the Fec system plays a causative role in enhancing catheter biofilm formation (46, 47). To address this, we examined the Fec expression profile in a wild type strain and measured catheter-biofilm formation in a *fecA*-deficient mutant. The Fec system imports extracellular ferric citrate complexes to the periplasm through the outer membrane transporter FecA, after which ferric citrate activates the sigma factor FecI through the inner membrane protein FecR, followed by transcriptional activation of *fecABCDE*, which facilitates cytoplasmic iron delivery (46-48). To determine whether *fec* gene transcription is associated with *E. coli* biofilm formation, we constructed a FecI-dependent fluorescence reporter strain (EC52::*fecI-RFP*), in the *fec*+ high-biofilm rectal isolate, EC52. Shaking microplate cultures of EC52::*fecI-RFP* exhibited abrupt fluorescence activation upon stationary phase entry at 8 hours, immediately preceding

detectable biofilm formation (**Fig. 7a**). These observations temporally connect transcriptional activation of *fec* genes to *E. coli* biofilm formation.

To determine whether Fec affects catheter biofilm formation, we next sought to determine whether iron acquisition systems are active in the biofilm culture system. We assessed this by determining whether enterobactin, the conserved *E. coli* siderophore excreted under low iron conditions, is secreted by bacteria during catheter biofilm culture. Using liquid chromatography-mass spectrometry (LC-MS), we detected established MS/MS ions for enterobactin (49, 50) in voided media from catheter biofilm produced by EC52, EC52Δ*fecA*, and complemented EC52Δ*fecA* at 10 hours of growth, consistent with activation of *E. coli* iron acquisition systems (**Fig. 7b, Supplementary Fig. S6)**. To determine whether the Fec system influences catheter biofilm formation, we next compared biofilm formation by isolate EC52 to its isogenic *fecA* deletion mutant EC52Δ*fecA* in the catheter continuous flow system (47). Both CV retention (**Fig. 7c**, $P = 0.0001$) and sessile bacterial counts (**Fig. 7d**, $P = 0.0002$) were significantly lower in EC52Δ*fecA* relative to wild type EC52. Genetic complementation of EC52Δ*fecA* with a *fecA* expression plasmid (EC52Δ*fecA::fecA*) significantly reversed this biofilm formation deficit (**Fig. 7c&d**, $P < 0.006$). No significant differences in planktonic CFUs (non-biofilm growth) were observed between wild type, mutant and complemented strains (**Fig. 7e**, $P = 0.7$), consistent with indistinguishable growth curves in AUM for these three strains (**Supplementary Fig. S7**). Together, these data are consistent with activation of iron uptake systems during catheter biofilm formation and a role for the Fec system in catheter biofilm formation.

## DISCUSSION

In this study we identify a genomic lineage within the *E. coli* B2 phylotype that is associated with CAUTI in a hospitalized population. This lineage is dominated by pandemic, multidrug resistant ST131 strains (24, 25) that possess a distinctive combination of virulence factor genes. We found that experimental catheter biofilm formation (35) did not distinguish ST131 strains, and that biofilms produced by a subset of non-ST131 CAASB strains could prevent colonization by CAUTI-associated ST131 strains. In comparative metagenomic analyses, we found biofilm-associated genes to be largely distinct from those associated with CAUTI, consistent with a possible shared role for biofilm in both CAASB and CAUTI. Notably, different iron-responsive gene systems were associated with both CAUTI and biofilm formation. The ferric citrate transport system (Fec) was the most prominent catheter biofilm correlate, was transcriptionally activated early in biofilm formation, and was functionally associated with enhanced catheter biofilm formation. The overall results are consistent with a multifaceted role for *E. coli* biofilm formation in colonizing catheterized hosts, with an elevated risk for infectious progression by ST131 strains carrying a distinctive combination of virulence-associated genes. In this paradigm, catheter biofilm formation by *E. coli* may be protective in some patients and harmful in others, depending upon the presence of specific virulence function combinations.

Identification of a distinct, infection-associated *E. coli* lineage in a clinical *E. coli* bacteria cohort at this degree of resolution is unusual (20, 27). This result may reflect the study's singular focus on catheterized patients, in whom infection may arise through a relatively distinct and uniform pathophysiology in a more homogenous host population. The abundance of ST131 strains in this study may also reflect their relatively recent global proliferation (25), aided by this an ability to efficiently colonize and persist in human intestinal reservoirs (51-54), which is regarded as the source of most urinary *Enterobacterales* (55-57). While ST131 intestinal colonization exhibits no sex differences (58-60), CAUTI patients in this study were disproportionately male,

possibly reflecting the enhanced *E. coli* infection severity in males observed in an animal model of direct bladder inoculation (61, 62). It is unclear whether the association between ST131 strains and CAUTI arises from increased pathogenic potential of these strains, an association with male patients, or a combination thereof. CAUTI-associated ST131 strains compared to CAASB strains are not distinguished by their ability to form catheter biofilms in the present study, but rather by a combination of accessory genes, including virulence factors suggesting a potential role for enhanced virulence. Because blood cultures are seldom obtained from asymptomatic individuals, a sufficiently powered study to more closely distinguish these possibilities in male CAASB and CAUTI patients would likely require obtaining prospective urine cultures from asymptomatic patients.

Of note, the ability of some strains to interfere with ST131 catheter colonization raises the possibility that catheter biofilms formed by *E. coli* strains without high-risk virulence gene combinations may benefit patients. Although one such strain that has been extensively studied in this regard, *E. coli* 83972 (42, 43), was collected from an exceptional, non-catheterized patient with three years of bacteriuria, the current study suggests that protective strains are common in catheterized patients. High antibiotic resistance among ST131 strains (63) raises the possibility that antibiotic treatment preferentially eliminates protective *E. coli* strains while sparing ST131 strains, paradoxically increasing the likelihood of progression to CAUTI. This scenario further reinforces guideline recommendations to be judicious with antibiotic use (1, 4, 7). A clinical test distinguishing ST131 from non-ST131 bacteriuria could also aid treatment decisions by helping to differentiate CAUTI and CAASB, a stated area of diagnostic need (63).

The processes that predispose biofilm-bound *E. coli* to progress to CAUTI remain unclear but are of diagnostic and therapeutic interest (15). These processes are presumably complex and include biofilm efflux, host tissue adhesion, immune evasion, and nutrient acquisition (64). VFs

in this study did not appear to equip strains for infection as equally influential components with simple additive effects on pathogenic potential. Neither the number of VFs nor their general functional categories clearly distinguish B2a from B2b strains. B2a and B2b strains are, however, distinguished by the presence of specific VF combinations encoding siderophore, adhesin, and toxin systems. If VFs affect pathogenic potential in catheterized patients, this appears to occur through idiosyncratic functions of specific VFs acting within evolutionarily favored combinations, suggested by the presence of VFs in the favored network communities described here (**Fig. 2**) and in previous work (21, 29). Previously identified VFs are not the only possible contributors to pathogenic potential in ST131 strains. B2a strains in this study carry 224 unique genes that are absent in B2b strains that may also modulate pathogenic potential. Discerning the contributions of these genes would require further study.

The variable genes associated with biofilm formation are substantially different from those associated with CAUTI. In the present study, the ferric citrate uptake system (Fec) was the most prominent of these, with two other iron acquisition systems, the aerobactin siderophore system and the ferric catecholate importer Cir (65), also represented. The deficiency in catheter biofilm formation by a Fec-deficient mutant that produces enterobactin, the prototypical *E. coli* siderophore, was surprising, as previous work indicated this deficiency was only discerned in planktonic, siderophore-deficient *E. coli* mutants (66, 67). These discrepant observations may relate to important differences in iron acquisition and trafficking in the biofilm matrix. Consistent with our observation here, a recent study identified *fecA* as an *E. coli* fitness factor in a murine UTI model despite retained enterobactin function (68). Precisely how these different iron acquisition-related systems function in the context of a catheter biofilm remains unclear. It is possible that, in biofilm microenvironments, the lower metabolic cost of citrate as an iron chelator is an important feature and that host-derived urinary citrate favors bacteria that are able to use this “free” resource (46, 47). It is also possible that the Fec system mediates biofilm-

specific functions independently of its ability to mediate iron uptake. Together, investigating the Fec as well as other iron acquisition systems in UPEC provide great insights for better elucidating bacterial pathogenesis in UTI and CAUTI, aiding in the search for new therapeutic approaches.

In conclusion, we found that CAUTI in the study population was associated with *E. coli* lineage largely defined by emergent multidrug resistant ST131 strains. The pathogenic potential of these populations was associated with carriage of specific gene networks and high degree of fluoroquinolone resistance, an antibiotic class commonly used to treat UTIs. ST131 strains appeared well-adapted to cause infection in patients with urinary catheters, raising the possibility that these strains arose from co-evolution with catheterized human hosts. The gene networks associated with biofilm formation were largely distinct from the CAUTI-associated gene networks. In addition, catheter biofilm formation was widespread among *E. coli* strains, and some strains in asymptomatic bacteriuria could act to prevent the colonization by CAUTI-associated ST131 strains. These results suggest that strain-specific characteristics of urinary *E. coli* influence CAUTI pathogenesis in patients. Strain-specific testing may thus aid clinical decision making in this population. An improved understanding of how ST131 strains cause infections may suggest future therapeutic strategies for these increasingly antibiotic-resistant bacteria.

## MATERIALS AND METHODS

### *Urinary isolates*

Urinary catheter-associated *E. coli* isolates were identified from a previously described study of bacteriuric (> $5 \times 10^4$ CFU/mL) inpatients (5, 6). This study was approved by the Washington University Institutional Review Board of (WU-IRB). CAUTI was defined as fever (T > 38 °C) with contemporaneous bacteriuria and urinary catheter placement (69). Documented urinary symptoms (dysuria, lower abdominal pain, flank pain) in the absence of fever were regarded as insufficient for CAUTI diagnosis due to their poor reliability in inpatients, particularly those with urinary catheters (1, 4, 7). CAASB was defined as bacteriuria in the absence of fever and documented urinary symptoms.

### *Rectal isolates*

Rectal *E. coli* isolates were collected from healthy adult volunteers in St. Louis, MO from 2014-2015. This study was approved by the WU-IRB and all study subjects provided written informed consent. Exclusion criteria included age < 18 years old, pregnancy, current urinary tract infection, previous urogenital surgery, ongoing treatment for urogenital cancer, the use of systemic antibiotics within 30 days of the study visit, or the use of a urinary catheter within 30 days of the study visit. Each study subject used a previously published protocol (70) to procure a self-collected rectal swab (BD Eswab) and submitted it with a study survey. Swabs were processed by the clinical microbiology lab at Barnes-Jewish Hospital to identify a dominant *E. coli* isolate and assess its antibiotic susceptibilities. Fifty-seven subjects were consented, 48 submitted study materials, 41 *E. coli* isolates had matching demographic data, and 13 of those *E. coli* isolates were randomly selected for the current study.

***Human urine***

Healthy donor urine was collected from adult volunteers as approved by the WU-IRB. Participants provided written informed consent for collection of up to two specimens, at least 1 week apart, for subsequent discovery and validation analysis. Exclusion criteria included recent UTI, antibiotic therapy, pregnancy, or any urogenital diseases. Collected human urine were mixed together, filter-sterilized (0.22 µm), then stored in -80 °C until use. Before an experiment, frozen urine was taken out, thawed on ice, filter-sterilized again and then processed for usage (71).

***Bacterial strains and culture***

An isogenic mutant of *E. coli* strain EC52 was constructed as in-frame deletion using the Lambda Red recombinase method as described previously (72). Isogenic mutant complementation and fluorescence reporter construct were accomplished by ectopic expression using transformed plasmids (47). Unless otherwise specified, cultures were grown from single colonies in LB broth for 12 h at 37 °C before using in the indicated assays.

***Whole-genome sequencing***

Bacterial genomic DNA was extracted with a QIAmp BiOstic Bacteremia DNA kit (Qiagen) from ~10 colonies of overnight growth. 5 ng of DNA was used as input to create Illumina sequencing libraries using the Nextera kit (Illumina). The samples were pooled and sequenced on an Illumina NextSeq 500 High Output system to obtain 2x150 bp reads. The reads were demultiplexed by barcode and had adapter sequences removed with trimmomatic v.38 and contaminating sequenced removed with deconseq v.4.3 (73). Processed reads were assembled into draft genomes with SPAdes v3.12.0 (Bankevich). The scaffolds.fasta file from spades was annotated for protein coding sequences on all contigs > 500bp with prokka v1.12 (74). Additionally, we obtained *E. coli* genomes in the known phylogroups and annotated their

protein coding sequences. GFF files from prokka were used as input for roary to create a core-genome alignment with PRANK (75). The core-genome alignment was constructed into a maximum likelihood tree with raxML and viewed in iTOL (76). *In silico* multilocus sequence types (MLST) were identified using BLASTN to the *E. coli* MLST database (23). Previously published virulence factors (VFs) were annotated in the *E. coli* draft genomes using virulencefinder v1.5 and blastp to previously described genes (77). Antibiotic resistance genes (ARGs) in genomic assemblies were identified by BLAST comparison of protein sequences against the CARD database based on stringent cutoffs (> 95% ID and > 95% overlap with subject sequence) (78).

To examine clonality among B2a subclade isolates, Roary was repeated on the gff files of the 15 B2a isolates, to produce a core-genome alignment of 3588 genes. SNP-Sites (https://github.com/sanger-pathogens/snp-sites) was ran on the alignment file to produce a VCF which identified 3351 total SNPs within this cohort (79). Pairwise SNP distances were calculated for all genomes using vcfR (80) and custom python scripts as described in D'Souza et al (81). The genomes analyzed in this report have been deposited to NCBI WGS database under BioProject accession no. PRJNA514354.

***Genomic analysis***

The thirty-two virulence factors (VFs) were compared between phenotypic and genetic groups for identifying CAUTI-associated VFs using sparse principal component analysis (sPCA), logistic regression (LR) classification, and network analysis approaches. Biofilm-associated genes were determined by comparative genomic analyses using sparse partial least squares discriminant analysis (sPLSDA), logistic regression (LR) classification, and network analysis approaches (21, 26, 27, 30). Computational models used in these genomic analyses were configured in Python and R programming languages, mainly by using the scikit-learn module and mixOmics packages,

497 respectively, as well as the Gephi software (http://gephi.org). Because of the high sparsity of
498 genomic metadata, with 15993 genes identified in 41 genome assemblies (15993 >> 41), sparse
499 penalty was enforced in all dimensionality reduction analyses (sPCA and sPLSDA) to prevent
500 overfitting (82).
501
502 ***Network analysis***
503 Two network representations for the 26 virulence factors (VFs) and 46 high-biofilm genes,
504 connected by co-occurrences across the *E. coli* collection, were defined using statistically
505 significant positive correlations as the edge weights across the networks. Statistical significance
506 between two nodes (genes) was determined by Fisher's exact test to determine whether they
507 appeared independently, conditional on their observed marginal frequencies among the *E. coli*
508 collection. The 0.4% and 5% *P*-value thresholds (one-tailed on the right) were chosen for the
509 VFs and biofilm-positive genes networks, respectively, to ensure that the obtained gene network
510 in each case was a single connected component. An edge was defined as present between any
511 pair of positively correlated nodes that satisfied the significance threshold, with edge weight
512 equal to the positive correlation coefficient. Communities in this network were detected using
513 the Louvain method by maximizing the modularity function (21, 29). We selected the obtained 3-
514 community (Resolution = 1.0) and 6-community (Resolution = 1.25) for the network
515 visualizations of 26 VFs and 46 high-biofilm associated genes, respectively, using a force-
516 directed layout generated by the Gephi (http://gephi.org) ForceAtlas2 algorithm and the
517 corresponding correlation matrix.
518
519 ***Artificial urine medium (AUM)***
520 Artificial urine medium (AUM) (**Supplementary Table S5**) was prepared as an alternative
521 medium of human urine for characterizing biofilm formation. Iron and zinc contents of a
522 previously published AUM recipe (37, 38) was adjusted to reflect that of with human urine

specimens measured using inductively coupled plasma-mass spectrometry (ICP-MS), in this study and previous publication (83). The comparison indicated that the old AUM recipe included more iron (Old AUM vs human urine vs Sieniawska = 5 µM vs 0.86 ± 0.14 µM vs 0.21 µM) and less zinc (Old AUM vs human urine vs Sieniawska = 0 µM vs 9.60 ± 0.89 µM vs 7.0 µM) than pooled human urine. In addition, old AUM recipe without adding 5 µM $FeSO_4.7H_2O$ was also measured by ICP-MS, with the results detecting 0.74 ± 0.03 iron and 0.19 ± 0.02 zinc, suggesting that the yeast extract could provide enough iron but not zinc in AUM to mimic human urine composition. The published AUM recipe was therefore modified by removing the 5 µmol/L $FeSO_4.7H_2O$ and adding extra 7 µmol/L $ZnSO_4.7H_2O$ (**Supplementary Table S5**). All ICP-MS experiments were conducted at the Nano Research Facility (NRF), Department of Energy, Environmental and Chemical Engineering, Washington University in Saint Louis. The ICP-MS quantification was achieved using calibration curve of 1, 5, 10, 50, and 100 µg/L. Nitric acid (Fisher) was added into pooled human urine samples with a final acid concentration of 2% (71).

***Continuous flow catheter biofilm model system***

Biofilms were grown in a continuous flow catheter model, using previously published protocols with appropriate modifications (**Supplementary Fig. S3**) (36). The components used for assembling continuous flow system included the platinum-cured silicone urinary catheters (Nalgene™ 50), peristaltic pump (Watson Marlow 205U), flexible tubings (Tygon S3™), and plastic connectors (Thermo Scientific). Prior to use, all tubing, connectors and containers were autoclaved. Human urine and AUM were filter-sterilized (0.22 µm). Bacteria from single colonies were grown in LB broth under 37 °C for 12 h, washed with PBS, back-diluted 1:10 into filter-sterilized human urine or AUM, and injected into the catheter installed in the continuous flow system operating under 37 °C. *E. coli* inoculum was statically incubated for two hours to allow the bacterial attachment to catheter surface. Fresh medium was then pumped through the

548 catheter at the flow rate of 0.5 mL/min, with a 30 minute pre-flush to first wash off loosely
549 adherent bacteria. After 10 hours of continuous flow incubation, voided media and catheters
550 were collected for characterization.
551
552 Total biofilm biomass was quantified by crystal violet (CV) retention. Biofilm-bound (sessile) and
553 unbound (planktonic) bacterial counts were determined by CFU enumeration of the biofilm
554 matrix or voided media (36). Catheter (11 cm) collected from the flow system was washed with
555 PBS and cut into three pieces (3 cm), with two pieces for CV and one for CFU assays. For CV
556 staining, 3 cm catheters were stained with 0.5% CV solution for 10 min, washed with deionized
557 water, air-dried on absorbent paper overnight, and extracted with 33% acetic acid for 10 min.
558 CV extracts were diluted 20-fold and measured at 595 nm using a Spectrophotometer
559 (Beckman Coulter DU-800). To quantify sessile bacteria, 3 cm catheter was cut into fragmented
560 pieces, immersed in 3 mL PBS, sonicated (Branson 350) for 10 min, vortexed for 3 min
561 (GeneMate), and plated to quantify as CFU/(mL × $cm^2$). To quantify planktonic bacteria, voided
562 media were collected and directly plated for CFU enumeration as CFU/mL.
563
564 ***Biofilm structure characterization***
565 Biofilm grown on urinary catheter surface was processed for structure characterization using
566 transmission electron microscopy (TEM) (84). For microstructural characterization, 1 cm
567 catheter with biofilm formed on inside surface was fixed in 2% paraformaldehyde/2.5%
568 glutaraldehyde (Polysciences Inc.) in 100 mM sodium cacodylate buffer (pH 7.2) for 1 h at room
569 temperature. Samples were washed in sodium cacodylate buffer and postfixed in 1% osmium
570 tetroxide (Polysciences Inc.) for 1 h. Samples were then rinsed extensively in deionized water
571 prior to en bloc staining with 1% aqueous uranyl acetate (Ted Pella Inc.) for 1 h. Following
572 several rinses in deionized water, samples were dehydrated in a graded series of ethanol and
573 embedded in Eponate 12 resin (Ted Pella Inc.). Sections of 95 nm were cut with a Leica

Ultracut UCT ultramicrotome (Leica Microsystems Inc.), stained with uranyl acetate and lead citrate, and viewed on a JEOL 1200 EX transmission electron microscope (TEM) (JEOL USA Inc.) equipped with an AMT 8 megapixel digital camera and AMT Image Capture Engine V602 software (Advanced Microscopy Techniques).

***Bacterial interference analysis***

Competitive colonization between 11 non-ST131 CAASB and 10 ST131 CAUTI isolates (**Supplementary Table S6**) was evaluated in the continuous flow catheter-biofilm system (**Supplementary Fig. S3b**) with two consecutive flow stages (**Supplementary Fig. S8**) (85). All flow experiments were conducted aseptically under 37 °C using AUM. Urinary catheters were pre-colonized with a CAASB biofilm for 10 hours, then challenged with a CAUTI isolate for another 10 hours. At this endpoint, voided media and urinary catheters were collected to determine total CFU and the proportion of CAASB and CAUTI strains by SNPs-qPCR as described below. The proportions of CAASB and CAUTI bacteria in both planktonic and biofilm two-bacteria pellets were used to measure the affect of pre-colonized CAASB strain biofilm on CAUTI strain colonization.

***Single-Nucleotide Polymorphisms qPCR (SNPs-qPCR)***

We quantified CAASB and CAUTI strain proportions in mixed cultures using single-nucleotide polymorphisms real-time PCR (SNPs-qPCR). Housekeeping genes (25) were aligned to identify SNPs-containing segments to distinguish between two strains. Two housekeeping genes *adk* (adenylate kinase) and *gyrB* (DNA gyrase) were identified with strain-specific SNPs that could differentiate the 11 non-ST131 CAASB from the 10 ST131 CAUTI isolates (**Supplementary Fig. S9**). Three SNPs-containing portions in *adk* distinguished EC24, EC25, EC26, EC27, EC37, EC38, and EC39 from the 10 CAUTI isolates. Two SNP-containing portions in gene *gyrB* differentiated EC33, EC34, EC35, and EC36 from the 10 CAUTI isolates (**Supplementary**

**Table S7**). Primers (**Supplementary Table S8**) amplifying the SNP-containing portion in each gene were designed and validated by PCR following gel electrophoresis to confirm that SNPs-based qPCR assays distinguish between two *E. coli* isolates in mixed cultures (**Supplementary Fig. S10**).

Prior to qPCR, bacteria cultures collected from catheter biofilm system were spun down to collect two-bacteria pellets, and DNAs were extracted using Wizard Genomic DNA Purification Kit (Promega) and measured by NanoDrop 2000 Spectrophotometer (Thermo Scientific). All qPCR assays were performed on a CFX96 Real-Time System (BIO-RAD). The 20 µL PCR mixture contained 1× iTaq Universal SYBR Green Supermix (BIO-RAD), 0.2 µM of each primer, and 3 ng/µL DNA of each specimen. The standard running conditions consist of a 3 min polymerase activation and DNA denaturation at 95℃, another 10 sec DNA denaturation at 95℃, followed by 40 cycles of a 30 sec annealing at 58.5℃, ending with a melt curve with 5 sec at 65℃ first and 5 sec each at 0.5℃ increase between 65℃ and 95℃ (86), with threshold cycles (*Cq*) obtained at the end of the reactions. Calibration curves [*log(Cq) ~ log(DNA)*] for each strain's SNPs-qPCR assay were established at 0.09375, 0.1875, 0.75, 1.5, and 3 ng/µL (**Supplementary Table S7**). The acquired linear calibration curves demonstrated ability of each SNPs-qPCR assay to detect the expected proportions of bacteria in mixed cultures. Finally, in the mixed cultures of two *E. coli* isolates, threshold cycles (*Cq*) were obtained to determine the quantity of DNA (ng/µL) for each isolate, with quantification using the calibration curve.

***fec fluorescence and microplate-biofilm assays***

Fluorescence from the *fecI* red fluorescent protein (RFP, mCherry) reporter was measured to assess transcriptional activation of the ferric citrate transport (*fec*) pathway (47). The *fec* reporter plasmid was transformed into isolate EC52, a high-biofilm former, to create the reporter strain, EC52::*fecI*-RFP (**Supplementary Table S9**). A control strain, EC52::RFP, was

constructed using the same plasmid without the *fec* promoter managing RFP expression (**Supplementary Table S9**). Primers (**Supplementary Table S8**) used to construct the plasmids were designed and validated by PCR following gel electrophoresis. These strains were cultured in a Tecan Spark microplate reader at 37 °C to monitor *fec* expression levels at different stages of bacterial growth in M63/0.2% glycerol minimal medium over 24 hours. To relate *fec* expression to biofilm formation, CV determinations were performed as described above over a time series.

***Mass spectrometry***

Eluates from the catheter biofilm system were promptly centrifuged to remove bacteria and particulates (21,000g for 2 minutes), filtered (0.45 µm Millex PVDF Durapore syringe-driven filter), and stored at -80 C prior to analysis. Samples were analyzed by liquid chromatography-mass spectrometry with a Thermo Vanquish ultrahigh pressure liquid chromatograph interfaced with a Thermo ID-X Tribrid mass spectrometer with an ESI source (87). Chromatography was performed using a Ascentis-Express fused core phenyl-hexyl column (100 mm x 2 mm x 2.7 µm) with a 0.5 mL/min flow rate. The column was equilibrated in 95% A (0.1% [v/v] formic acid) and 5% B (90% acetonitrile plus 0.1% [v/v] formic acid) prior to sample injection. Percent buffer B was held at 5% until 2 minutes, then increased to 56% at 10 minutes and 98% at 12 minutes. The column was held at 98% B until 16 minutes, then returned to 5% B by 18 minutes and held until 21 minutes. Negative ion tandem mass spectra were collected for the precursor ion -668.1 m/z (cyclic enterobactin). Product ions were extracted and integrated using Thermo TraceFinder software version 5.1.

***fec deletion mutants***

f*ecA*, encoding the ferric citrate outer membrane receptor, was deleted from the high biofilm-forming strain EC52 using Lambda Red recombinase method, creating the isogenic mutant

EC52Δ*fecA* (**Supplementary Table S8**) (72). EC52Δ*fecA* was genetically complemented using a *fecA* expression plasmid, generating EC52Δ*fecA*::*fecA* (**Supplementary Table S8**) as a control. Primers (**Supplementary Table S9**) used in the *fecA* deletion and complementation were designed validated by PCR and gel electrophoresis.

***Statistical methods***

GraphPad Prism 8.0 (GraphPad software) was used to generate graphs and perform statistical analysis in this study. We used one-sample *t* test for single group comparison, Mann-Whitney test for two group comparisons, and one-way ANOVA for multigroup comparisons. Dunnett's tests was used to correct one-way ANOVA multigroup comparisons where appropriate.

## DATA AVAILABILITY

The genomes analyzed in this report have been deposited to NCBI WGS database under BioProject accession no. PRJNA514354. The computer codes for the analyses in this study are available in Github (https://github.com/QL5001/CAUTI-script). Other data that support the findings of this study are available within the paper and its Supplementary Information files.


## ACKNOWLEDGEMENTS

We thank Wandy Beatty for assistance with transmission electron microscopy (TEM). We thank the Edison Family Center for Genome Sciences and Systems Biology staff, Eric Martin, Brian Koebbe, MariaLynn Crosby, and Jessica Hoisington-López for their assistance in genome sequencing and high-throughput computing. JPH acknowledges Centers for Disease Control Prevention Epicenters Program Grant (CU54CK000162), and National Institutes of Health grants R01DK111930 and RO1DK125860. GD acknowledges National Institutes of Health grants U01AI123394 and R01AI155893. WHM acknowledges the KL2TR002346 – ICTS Institutional Career Development Program and the National Institutes of Health grants UL1TR002345 and 1K08AR076464-01. RFP acknowledges the Monsanto Excellence Fund Graduate Fellowship. The content is solely the responsibility of the authors and does not necessarily represent the official view of the CDC or NIH.


## AUTHOR CONTRIBUTIONS

ZZ and JPH conceived and designed the experiments. ZZ performed the experiments. WHM conducted rectal *E. coli* collection. RFP and GD conducted genome sequencing and alignments. JAW conducted mass spectrometry. ZZ and GLK conducted reporter construct and targeted mutagenesis. ZZ, PJM, and JPH conducted network analyses. ZZ and JPH analyzed the data. ZZ and JPH wrote the manuscript.

## COMPETING INTERESTS

The authors declare no competing interests.

**REFERENCES**


1. Trautner BW. Management of catheter-associated urinary tract infection. *Curr Opin Infect Dis.* 2010;23(1):76-82.
2. Flores-Mireles AL, Walker JN, Caparon M, and Hultgren SJ. Urinary tract infections: epidemiology, mechanisms of infection and treatment options. *Nat Rev Microbiol.* 2015;13(5):269-84.
3. Klein RD, and Hultgren SJ. Urinary tract infections: microbial pathogenesis, host-pathogen interactions and new treatment strategies. *Nat Rev Microbiol.* 2020;18(4):211-26.
4. Trautner BW, and Morgan DJ. Imprecision Medicine: Challenges in Diagnosis, Treatment, and Measuring Quality for Catheter-Associated Urinary Tract Infection. *Clin Infect Dis.* 2020;71(9):e520-e2.
5. Marschall J, Zhang L, Foxman B, Warren DK, Henderson JP, and Program CDCPE. Both host and pathogen factors predispose to Escherichia coli urinary-source bacteremia in hospitalized patients. *Clin Infect Dis.* 2012;54(12):1692-8.
6. Jonas Marschall MLP, Betsy Foxman, Lixin Zhang, David K Warrenj, Jeffrey P Henderson. Patient characteristics but not virulence factors discriminate between asymptomatic and symptomatic. *BMC Infectious Diseases.* 2013;13(213).
7. Cope M, Cevallos ME, Cadle RM, Darouiche RO, Musher DM, and Trautner BW. Inappropriate treatment of catheter-associated asymptomatic bacteriuria in a tertiary care hospital. *Clin Infect Dis.* 2009;48(9):1182-8.
8. Hooton TM, Bradley SF, Cardenas DD, Colgan R, Geerlings SE, Rice JC, et al. Diagnosis, prevention, and treatment of catheter-associated urinary tract infection in adults: 2009 International Clinical Practice Guidelines from the Infectious Diseases Society of America. *Clin Infect Dis.* 2010;50(5):625-63.

9. Nicolle LE, Gupta K, Bradley SF, Colgan R, DeMuri GP, Drekonja D, et al. Clinical Practice Guideline for the Management of Asymptomatic Bacteriuria: 2019 Update by the Infectious Diseases Society of America. *Clin Infect Dis.* 2019;68(10):1611-5.
10. Casadevall A. The Pathogenic Potential of a Microbe. *mSphere.* 2017;2(1).
11. Armbruster CE, Brauer AL, Humby MS, Shao J, and Chakraborty S. Prospective assessment of catheter-associated bacteriuria clinical presentation, epidemiology, and colonization dynamics in nursing home residents. *JCI Insight.* 2021;6(19).
12. Guiton PS, Hannan TJ, Ford B, Caparon MG, and Hultgren SJ. Enterococcus faecalis overcomes foreign body-mediated inflammation to establish urinary tract infections. *Infect Immun.* 2013;81(1):329-39.
13. Rousseau M, Goh HMS, Holec S, Albert ML, Williams RB, Ingersoll MA, et al. Bladder catheterization increases susceptibility to infection that can be prevented by prophylactic antibiotic treatment. *JCI Insight.* 2016;1(15):e88178.
14. K M Delnay WHS, H Goldman, A F Jukkola, R R Dmochowski. Bladder histological changes associated with chronic indwelling urinary catheter. *J Urol.* 1999;161(4):1106-8.
15. Trautner BW, and Darouiche RO. Role of biofilm in catheter-associated urinary tract infection. *Am J Infect Control.* 2004;32(3):177-83.
16. Donlan RM. Biofilm Formation: A Clinically Relevant Microbiological Process. *Clinical Infectious Diseases.* 2001;33(8):1387–92.
17. Donlan RM. Biofilms: microbial life on surfaces. *Emerging Infectious Diseases.* 2002;8(9):881-90.
18. Alqasim A, Emes R, Clark G, Newcombe J, La Ragione R, and McNally A. Phenotypic microarrays suggest Escherichia coli ST131 is not a metabolically distinct lineage of extra-intestinal pathogenic E. coli. *PLoS One.* 2014;9(2):e88374.

19. Marrs CF, Zhang L, and Foxman B. Escherichia coli mediated urinary tract infections: are there distinct uropathogenic E. coli (UPEC) pathotypes? *FEMS Microbiol Lett.* 2005;252(2):183-90.

20. Vejborg RM, Hancock V, Schembri MA, and Klemm P. Comparative genomics of Escherichia coli strains causing urinary tract infections. *Appl Environ Microbiol.* 2011;77(10):3268-78.

21. Parker KS, Wilson JD, Marschall J, Mucha PJ, and Henderson JP. Network Analysis Reveals Sex- and Antibiotic Resistance-Associated Antivirulence Targets in Clinical Uropathogens. *ACS Infect Dis.* 2015;1(11):523-32.

22. MacVane SH, Tuttle LO, and Nicolau DP. Demography and burden of care associated with patients readmitted for urinary tract infection. *J Microbiol Immunol Infect.* 2015;48(5):517-24.

23. Gasparrini AJ, Markley JL, Kumar H, Wang B, Fang L, Irum S, et al. Tetracycline-inactivating enzymes from environmental, human commensal, and pathogenic bacteria cause broad-spectrum tetracycline resistance. *Commun Biol.* 2020;3(1):241.

24. Kudinha T, Johnson JR, Andrew SD, Kong F, Anderson P, and Gilbert GL. Escherichia coli sequence type 131 as a prominent cause of antibiotic resistance among urinary Escherichia coli isolates from reproductive-age women. *J Clin Microbiol.* 2013;51(10):3270-6.

25. Lau SH, Reddy S, Cheesbrough J, Bolton FJ, Willshaw G, Cheasty T, et al. Major uropathogenic Escherichia coli strain isolated in the northwest of England identified by multilocus sequence typing. *J Clin Microbiol.* 2008;46(3):1076-80.

26. Ohlemacher SI, Giblin DE, d'Avignon DA, Stapleton AE, Trautner BW, and Henderson JP. Enterobacteria secrete an inhibitor of Pseudomonas virulence during clinical bacteriuria. *J Clin Invest.* 2017;127(11):4018-30.

27. Henry L Schreiber 4th MSC, Wen-Chi Chou, Michael E Hibbing, Abigail L Manson, Karen W Dodson, Thomas J Hannan, Pacita L Roberts, Ann E Stapleton, Thomas M Hooton, Jonathan Livny, Ashlee M Earl, Scott J Hultgren. Bacterial virulence phenotypes of Escherichia coli and host susceptibility determine risk for urinary tract infections. *Science Translational Medicine.* 2017;9(282):eaaf1283.

28. Johnson JR. Virulence factors in Escherichia coli urinary tract infection. *Clinical Microbiology Reviews.* 1991;4(1):80-128.

29. Robinson JI, Weir WH, Crowley JR, Hink T, Reske KA, Kwon JH, et al. Metabolomic networks connect host-microbiome processes to human Clostridioides difficile infections. *J Clin Invest.* 2019;129(9):3792-806.

30. Spurbeck RR, Dinh PC, Jr., Walk ST, Stapleton AE, Hooton TM, Nolan LK, et al. Escherichia coli isolates that carry vat, fyuA, chuA, and yfcV efficiently colonize the urinary tract. *Infect Immun.* 2012;80(12):4115-22.

31. He XL, Wang Q, Peng L, Qu YR, Puthiyakunnon S, Liu XL, et al. Role of uropathogenic Escherichia coli outer membrane protein T in pathogenesis of urinary tract infection. *Pathog Dis.* 2015;73(3).

32. Nipic D, Podlesek Z, Budic M, Crnigoj M, and Zgur-Bertok D. Escherichia coli uropathogenic-specific protein, Usp, is a bacteriocin-like genotoxin. *J Infect Dis.* 2013;208(10):1545-52.

33. N H Carbonetti PHW. A cluster of five genes specifying the aerobactin iron uptake system of plasmid ColV-K30. *Infection and Immunity.* 1984;46(1):7-12.

34. Guyer DM, Radulovic S, Jones FE, and Mobley HL. Sat, the secreted autotransporter toxin of uropathogenic Escherichia coli, is a vacuolating cytotoxin for bladder and kidney epithelial cells. *Infect Immun.* 2002;70(8):4539-46.

35. P I Tarr SSB, J C Vary Jr, S Jelacic, R L Habeeb, T R Ward, M R Baylor, T E Besser. Iha: a novel Escherichia coli O157:H7 adherence-conferring molecule encoded on a

recently acquired chromosomal island of conserved structure. *Infect Immun.* 2000;68(3):1400-7.

36. Guiton PS, Cusumano CK, Kline KA, Dodson KW, Han Z, Janetka JW, et al. Combinatorial small-molecule therapy prevents uropathogenic Escherichia coli catheter-associated urinary tract infections in mice. *Antimicrob Agents Chemother.* 2012;56(9):4738-45.

37. T Brooks CWK. A simple artificial urine for the growth of urinary pathogens. *Lett Appl Microbiol.* 1997;24(3):203-6.

38. Lehman SM, and Donlan RM. Bacteriophage-mediated control of a two-species biofilm formed by microorganisms causing catheter-associated urinary tract infections in an in vitro urinary catheter model. *Antimicrob Agents Chemother.* 2015;59(2):1127-37.

39. Darouiche RO, and Hull RA. Bacterial interference for prevention of urinary tract infection. *Clin Infect Dis.* 2012;55(10):1400-7.

40. Prasad A, Cevallos ME, Riosa S, Darouiche RO, and Trautner BW. A bacterial interference strategy for prevention of UTI in persons practicing intermittent catheterization. *Spinal Cord.* 2009;47(7):565-9.

41. Stork C, Kovacs B, Rozsai B, Putze J, Kiel M, Dorn A, et al. Characterization of Asymptomatic Bacteriuria Escherichia coli Isolates in Search of Alternative Strains for Efficient Bacterial Interference against Uropathogens. *Front Microbiol.* 2018;9:214.

42. Trautner BW, Hull RA, and Darouiche RO. Escherichia coli 83972 inhibits catheter adherence by a broad spectrum of uropathogens. *Urology.* 2003;61(5):1059-62.

43. R Hull DR, W Donovan, C Svanborg, I Wieser, C Stewart, R Darouiche. Urinary tract infection prophylaxis using Escherichia coli 83972 in spinal cord injured patients. *J Urol.* 2000;163(3):872-7.

44. Kim-Anh Lê Cao SB, Philippe Besse. Sparse PLS discriminant analysis: biologically relevant feature selection and graphical displays for multiclass problems. *BMC Bioinformatics.* 2011;12(253).

45. P N Danese LAP, S L Dove, R Kolter. The outer membrane protein, antigen 43, mediates cell-to-cell interactions within Escherichia coli biofilms. *Mol Microbiol.* 2000;37(2):424-32.

46. Yue WW, Grizot S, and Buchanan SK. Structural Evidence for Iron-free Citrate and Ferric Citrate Binding to the TonB-dependent Outer Membrane Transporter FecA. *Journal of Molecular Biology.* 2003;332(2):353-68.

47. Ehud Banin MLV, and E. Peter Greenberg. Iron and Pseudomonas aeruginosa biofilm formation. *PNAS.* 2005;102(31):11076-81.

48. Noinaj N, Guillier M, Barnard TJ, and Buchanan SK. TonB-dependent transporters: regulation, structure, and function. *Annu Rev Microbiol.* 2010;64:43-60.

49. Lv H, Hung CS, and Henderson JP. Metabolomic analysis of siderophore cheater mutants reveals metabolic costs of expression in uropathogenic Escherichia coli. *J Proteome Res.* 2014;13(3):1397-404.

50. Leslie AD, Daneshfar R, and Volmer DA. Infrared multiphoton dissociation of the siderophore enterobactin and its Fe(III) complex. Influence of Fe(III) binding on dissociation kinetics and relative energetics. *J Am Soc Mass Spectrom.* 2007;18(4):632-41.

51. Johnson JR, Clabots C, Porter SB, Bender T, Johnston BD, and Thuras P. Intestinal Persistence of Colonizing Escherichia coli Strains, Especially ST131-H30, in Relation to Bacterial and Host Factors. *J Infect Dis.* 2022;225(12):2197-207.

52. Banerjee R, and Johnson JR. A new clone sweeps clean: the enigmatic emergence of Escherichia coli sequence type 131. *Antimicrob Agents Chemother.* 2014;58(9):4997-5004.

53. Sarkar S, Hutton ML, Vagenas D, Ruter R, Schuller S, Lyras D, et al. Intestinal Colonization Traits of Pandemic Multidrug-Resistant Escherichia coli ST131. *J Infect Dis.* 2018;218(6):979-90.

54. Forde BM, Roberts LW, Phan MD, Peters KM, Fleming BA, Russell CW, et al. Population dynamics of an Escherichia coli ST131 lineage during recurrent urinary tract infection. *Nat Commun.* 2019;10(1):3643.

55. Rao K, Seekatz A, Bassis C, Sun Y, Mantlo E, and Bachman MA. Enterobacterales Infection after Intestinal Dominance in Hospitalized Patients. *mSphere.* 2020;5(4).

56. Czaja CA, Stamm WE, Stapleton AE, Roberts PL, Hawn TR, Scholes D, et al. Prospective cohort study of microbial and inflammatory events immediately preceding Escherichia coli recurrent urinary tract infection in women. *J Infect Dis.* 2009;200(4):528-36.

57. Robert Thänert KAR, Tiffany Hink, Meghan A Wallace, Bin Wang, Drew J Schwartz, Sondra Seiler, Candice Cass, C A Burnham, Erik R Dubberke, Jennie H Kwon, Gautam Dantas. Comparative Genomics of Antibiotic-Resistant Uropathogens Implicates Three Routes for Recurrence of Urinary Tract Infections. *mBio.* 2019;10(4):e01977-19.

58. Mohamed M, Clabots C, Porter SB, Bender T, Thuras P, and Johnson JR. Large Fecal Reservoir of Escherichia coli Sequence Type 131-H30 Subclone Strains That Are Shared Within Households and Resemble Clinical ST131-H30 Isolates. *J Infect Dis.* 2020;221(10):1659-68.

59. Ducarmon QR, Terveer EM, Nooij S, Bloem MN, Vendrik KEW, Caljouw MAA, et al. Microbiota-associated risk factors for asymptomatic gut colonisation with multi-drug-resistant organisms in a Dutch nursing home. *Genome Med.* 2021;13(1):54.

60. Morales Barroso I, Lopez-Cerero L, Navarro MD, Gutierrez-Gutierrez B, Pascual A, and Rodriguez-Bano J. Intestinal colonization due to Escherichia coli ST131: risk factors and prevalence. *Antimicrob Resist Infect Control.* 2018;7:135.

61. Olson PD, Hruska KA, and Hunstad DA. Androgens Enhance Male Urinary Tract Infection Severity in a New Model. *J Am Soc Nephrol.* 2016;27(6):1625-34.

62. Deltourbe L, Lacerda Mariano L, Hreha TN, Hunstad DA, and Ingersoll MA. The impact of biological sex on diseases of the urinary tract. *Mucosal Immunol.* 2022.

63. Decano AG, and Downing T. An Escherichia coli ST131 pangenome atlas reveals population structure and evolution across 4,071 isolates. *Sci Rep.* 2019;9(1):17394.

64. Ganz T. Iron and infection. *Int J Hematol.* 2018;107(1):7-15.

65. H Nikaido EYR. Cir and Fiu proteins in the outer membrane of Escherichia coli catalyze transport of monomeric catechols: study with beta-lactam antibiotics containing catechol and analogous groups. *J Bacteriol.* 1990;172(3):1361-7.

66. Wu Y, and Outten FW. IscR controls iron-dependent biofilm formation in Escherichia coli by regulating type I fimbria expression. *J Bacteriol.* 2009;191(4):1248-57.

67. Hancock V, Dahl M, and Klemm P. Abolition of biofilm formation in urinary tract Escherichia coli and Klebsiella isolates by metal interference through competition for fur. *Appl Environ Microbiol.* 2010;76(12):3836-41.

68. Arwen E Frick-Cheng AS, Sara N Smith, Ali Pirani, Evan S Snitkin, Harry L T Mobley. Ferric Citrate Uptake Is a Virulence Factor in Uropathogenic Escherichia coli. *mBio.* 2022;13(3):e0103522.

69. Livorsi DJ, and Perencevich EN. CAUTI Surveillance: Opportunity or Opportunity Cost? *Infect Control Hosp Epidemiol.* 2015;36(11):1335-6.

70. Lampinen TM, Latulippe L, van Niekerk D, Schilder AJ, Miller ML, Anema A, et al. Illustrated instructions for self-collection of anorectal swab specimens and their adequacy for cytological examination. *Sex Transm Dis.* 2006;33(6):386-8.

71. Shields-Cutler RR, Crowley JR, Hung CS, Stapleton AE, Aldrich CC, Marschall J, et al. Human Urinary Composition Controls Antibacterial Activity of Siderocalin. *J Biol Chem.* 2015;290(26):15949-60.

72. Henderson JP, Crowley JR, Pinkner JS, Walker JN, Tsukayama P, Stamm WE, et al. Quantitative metabolomics reveals an epigenetic blueprint for iron acquisition in uropathogenic Escherichia coli. *PLoS Pathog.* 2009;5(2):e1000305.

73. Bolger AM, Lohse M, and Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* 2014;30(15):2114-20.

74. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics.* 2014;30(14):2068-9.

75. Löytynoja A. Phylogeny-aware alignment with PRANK. *Methods Mol Biol.* 2014;1079:155-70.

76. Swaine L Chen MW, Jeffrey P Henderson, Thomas M Hooton, Michael E Hibbing, Scott J Hultgren, Jeffrey I Gordon. Genomic diversity and fitness of E. coli strains recovered from the intestinal and urinary tracts of women with recurrent urinary tract infection. *Sci Transl Med.* 2013;5(184):184ra60.

77. Kleinheinz KA, Joensen KG, and Larsen MV. Applying the ResFinder and VirulenceFinder web-services for easy identification of acquired antibiotic resistance and E. coli virulence genes in bacteriophage and prophage nucleotide sequences. *Bacteriophage.* 2014;4(1):e27943.

78. Sun J, Liao XP, D'Souza AW, Boolchandani M, Li SH, Cheng K, et al. Environmental remodeling of human gut microbiota and antibiotic resistome in livestock farms. *Nat Commun.* 2020;11(1):1427.

79. Page AJ, Taylor B, Delaney AJ, Soares J, Seemann T, Keane JA, et al. SNP-sites: rapid efficient extraction of SNPs from multi-FASTA alignments. *Microb Genom.* 2016;2(4):e000056.

80. Knaus BJ, and Grunwald NJ. vcfr: a package to manipulate and visualize variant call format data in R. *Mol Ecol Resour.* 2017;17(1):44-53.

81. D'Souza AW, Potter RF, Wallace M, Shupe A, Patel S, Sun X, et al. Spatiotemporal dynamics of multidrug resistant bacteria on intensive care unit surfaces. *Nat Commun.* 2019;10(1):4569.

82. Kampa K, Mehta S, Chou CA, Chaovalitwongse WA, and Grabowski TJ. Sparse optimization in feature selection: application in neuroimaging. *Journal of Global Optimization.* 2014;59(2-3):439-57.

83. Sieniawska CE, Jung LC, Olufadi R, and Walker V. Twenty-four-hour urinary trace element excretion: reference intervals and interpretive issues. *Ann Clin Biochem.* 2012;49(Pt 4):341-51.

84. Zahller J, and Stewart PS. Transmission electron microscopic study of antibiotic action on Klebsiella pneumoniae biofilm. *Antimicrob Agents Chemother.* 2002;46(8):2679-83.

85. Roos V, Ulett GC, Schembri MA, and Klemm P. The asymptomatic bacteriuria Escherichia coli strain 83972 outcompetes uropathogenic E. coli strains in human urine. *Infect Immun.* 2006;74(1):615-24.

86. Sheludchenko MS, Huygens F, and Hargreaves MH. Highly discriminatory single-nucleotide polymorphism interrogation of Escherichia coli by use of allele-specific real-time PCR and eBURST analysis. *Appl Environ Microbiol.* 2010;76(13):4337-45.

87. Behnsen J, Zhi H, Aron AT, Subramanian V, Santus W, Lee MH, et al. Siderophore-mediated zinc acquisition enhances enterobacterial colonization of the inflamed gut. *Nat Commun.* 2021;12(1):7016.

88. Johnning A, Kristiansson E, Fick J, Weijdegard B, and Larsson DG. Resistance Mutations in gyrA and parC are Common in Escherichia Communities of both Fluoroquinolone-Polluted and Uncontaminated Aquatic Environments. *Front Microbiol.* 2015;6:1355.

## FIGURES

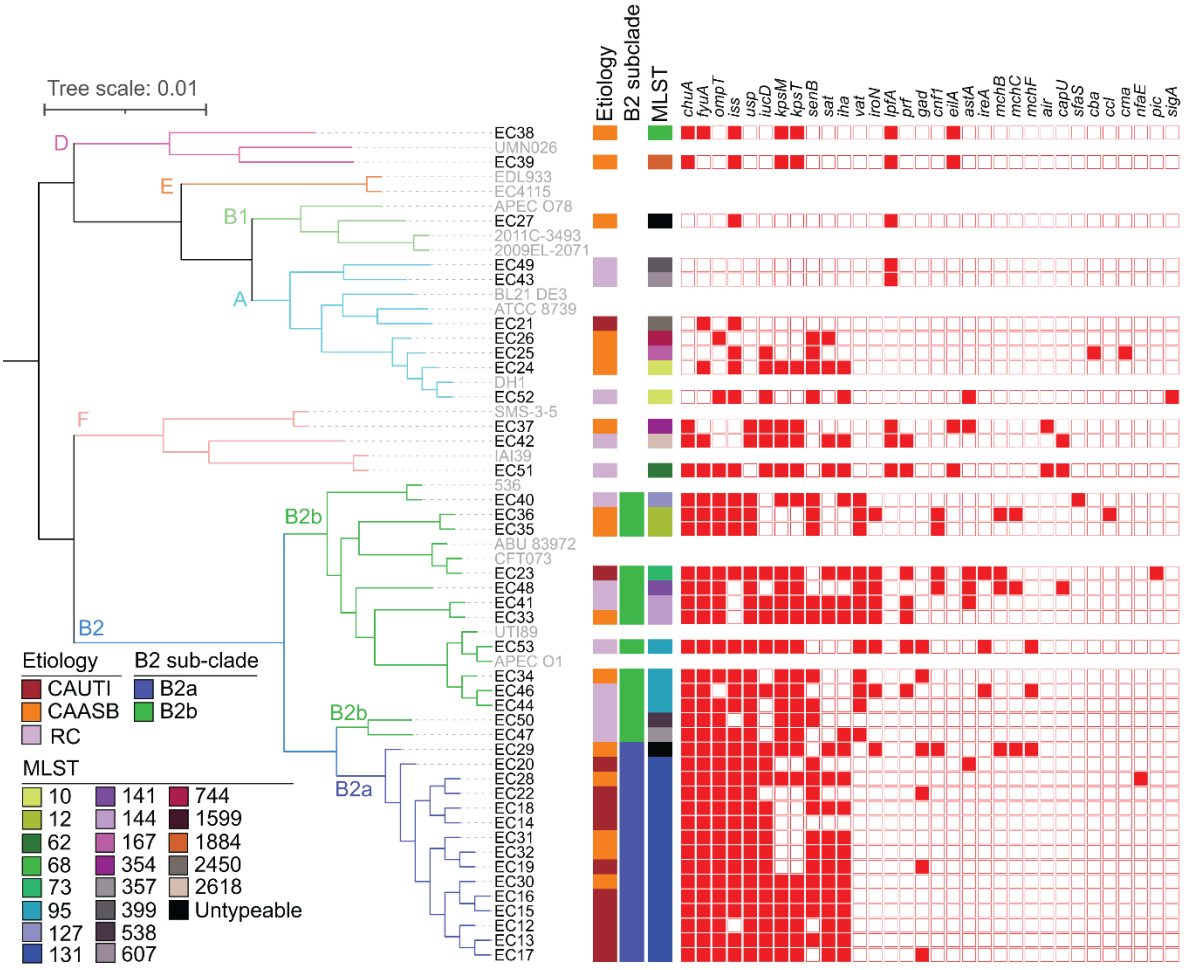


**Fig. 1 Phylogenetic distribution of 41 clinical *E. coli* isolates.** Core-genome alignment was constructed into a maximum likelihood tree with raxML and viewed in iTOL, with six phylotypes identified. *In silico* multilocus sequence types (MLST) were identified using BLASTN to the *E. coli* MLST database, with 21 STs identified. Virulence factors (VFs) were annotated in the *E. coli* draft genomes using VirulenceFinder v1.5 and blastp to previously described genes, with 32 VFs identified. The strain names of *E. coli* sequenced in this study were in black and reference *E. coli* strains were in gray.

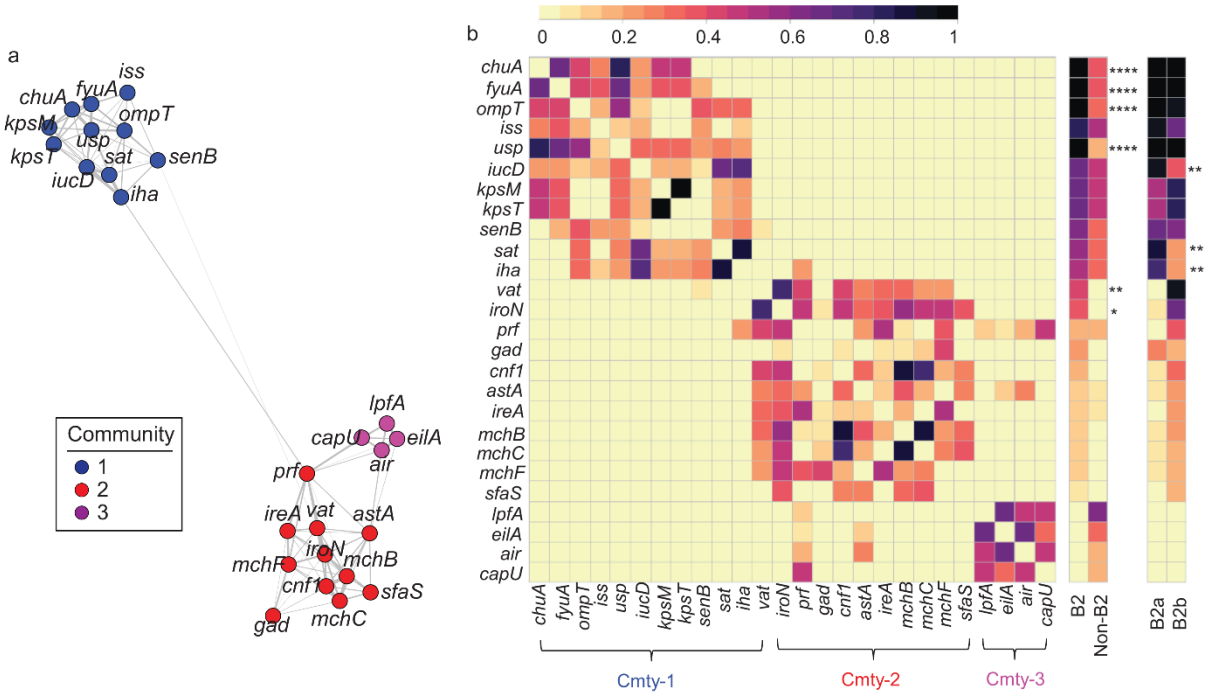


**Fig. 2 Network analysis of *E. coli* virulence factors (VFs). (a)** A force-directed network layout illustrated co-associations and three VF communities among 26 VFs. Each node represented a VF. Each connecting line (edge) represented a positive association between 2 VFs that satisfied the significance threshold (0.4% *P*-value threshold, one-tailed on the right, Fisher's exact test). Edge lengths were determined by the level of correlation between connected VFs. Nodes were colored by community assignment. **(b)** Three VF communities were discernible in the correlation matrix heatmap depicting statistically significant positive associations between 26 VFs. Presence frequency comparisons of each gene between different genetic groups, B2 vs non-B2 and B2a vs B2b, were displayed in heatmaps to the right of the correlation matrix. Cmty: community. By two-tailed Fisher's exact with $P <= 0.05$ considered statistically significant. *: $P <= 0.05$. **: $P < 0.01$. ***: $P < 0.001$. ****: $P < 0.0001$.

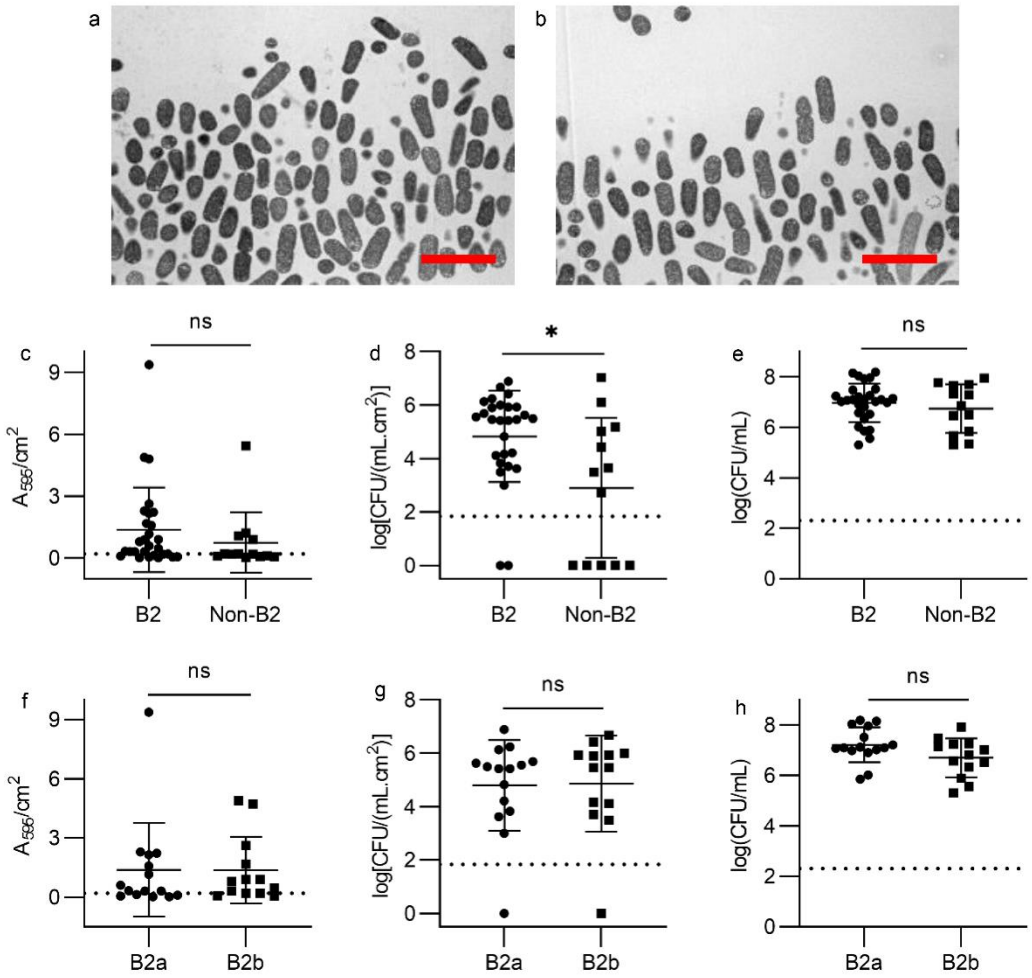


**Fig. 3 Biofilm formation by CAUTI, CAASB, and rectal isolates. (a)** Transmission electron microscopy (TEM) image of catheter-biofilm grown in human urine. Scale bar: 4 µm. **(b)** Transmission electron microscopy (TEM) image of catheter-biofilm grown in artificial urine medium (AUM). Scale bar: 4 µm. **(c)** Comparison of biofilm biomass (crystal violet retention) between B2 and non-B2 isolates. Mean with SD plotted for 28 B2 and 13 non-B2 strains. *P* = 0.22. **(d)** Comparison of catheter-adherent CFUs between B2 and non-B2 isolates. Mean with SD plotted for 28 B2 and 13 non-B2 strains. *P* = 0.02. **(e)** Comparison of planktonic CFUs in the voided media between B2 and non-B2 isolates. Mean with SD plotted for 28 B2 and 13 non-B2 strains. *P* = 0.57. **(f)** Comparison of biofilm biomass (crystal violet retention) between B2a and B2b isolates. Mean with SD plotted for 15 B2a and 13 B2b strains. *P* = 0.62. **(g)** Comparison of

catheter-adherent CFUs between B2a and B2b isolates. Mean with SD plotted for 15 B2a and 13 B2b strains. *P* = 0.73. **(h)** Comparison of planktonic CFUs in the voided media between B2a and B2b isolates. Mean with SD plotted for 15 B2a and 13 B2b strains. *P* = 0.13. By Mann-Whitney test with *P* <= 0.05 considered as statistically significant. ns: not significant. *: *P* <= 0.05.

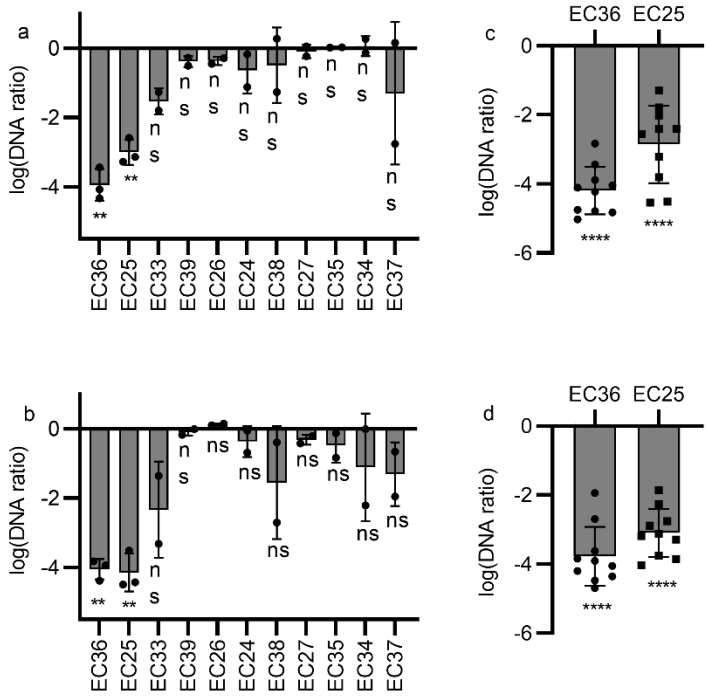


**Fig. 4 CAASB *E. coli* catheter-biofilms inhibited CAUTI colonization. (a & b)** DNA ratios of ST131 CAUTI strain EC20 in **(a)** catheter-adherent bacteria and **(b)** planktonic bacteria demonstrate its different levels of colonization when grown by itself (Control) and in competition with (Competition) 11 non-ST131 CAASB strains (EC36, 25, 33, 39, 26, 24, 38, 27, 35, 34, and 37) in the catheter colonization model. Three replicates with mean and SD plotted for EC36 and EC25, with $P < 0.005$, two replicates with mean and SD plotted for EC33, 39, 26, 24, 38, 27, 35, 34, and 37, with $P > 0.10$. **(c & d)** DNA ratios of 10 ST131 CAUTI strains (EC12, 13, 14, 15, 16, 17, 18, 19, 20, and 22) in **(c)** catheter-adherent bacteria and **(d)** planktonic bacteria demonstrate their different levels of colonization when grown by itself (Control) and in competition with (Competition) two non-ST131 CAASB strains (EC36 and EC25) in the catheter colonization model. Ten ST131 CAUTI strains with mean and SD plotted, with $P < 0.0001$. DNA ratio = (DNA-Competition)/(DNA-Control). By one-sample *t* test. $P <= 0.05$ is considered statistically significant. ns: not significant. *: $P <= 0.05$. **: $P < 0.01$. ***: $P < 0.001$. ****: $P < 0.0001$.

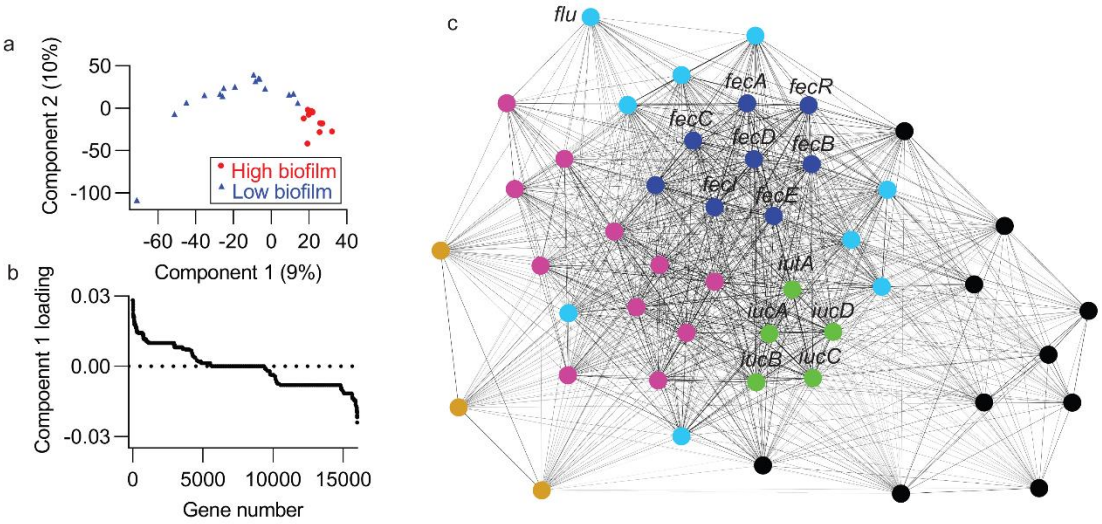


**Fig. 5 Identification of catheter biofilm-associated genes. (a)** Score plot of the first two components from sparse partial least squares discriminant analysis (sPLSDA) for displaying group-wise clustering between high and low biofilm formers. **(b)** Component 1-associated top loadings from sPLSDA identified 72 biofilm-correlated genes, including 46 positive (high-biofilm) and 26 negative (low-biofilm) genes. **(c)** A force-directed network layout illustrated co-associations and three gene communities among 46 biofilm positively associated genes. Each node represented a gene. Each connecting line (edge) represented a positive association between 2 genes that satisfied the significance threshold (5% *P*-value threshold, one-tailed on the right, Fisher's exact test). Edge lengths were determined by the level of correlation between connected genes. Nodes were colored by community assignment.

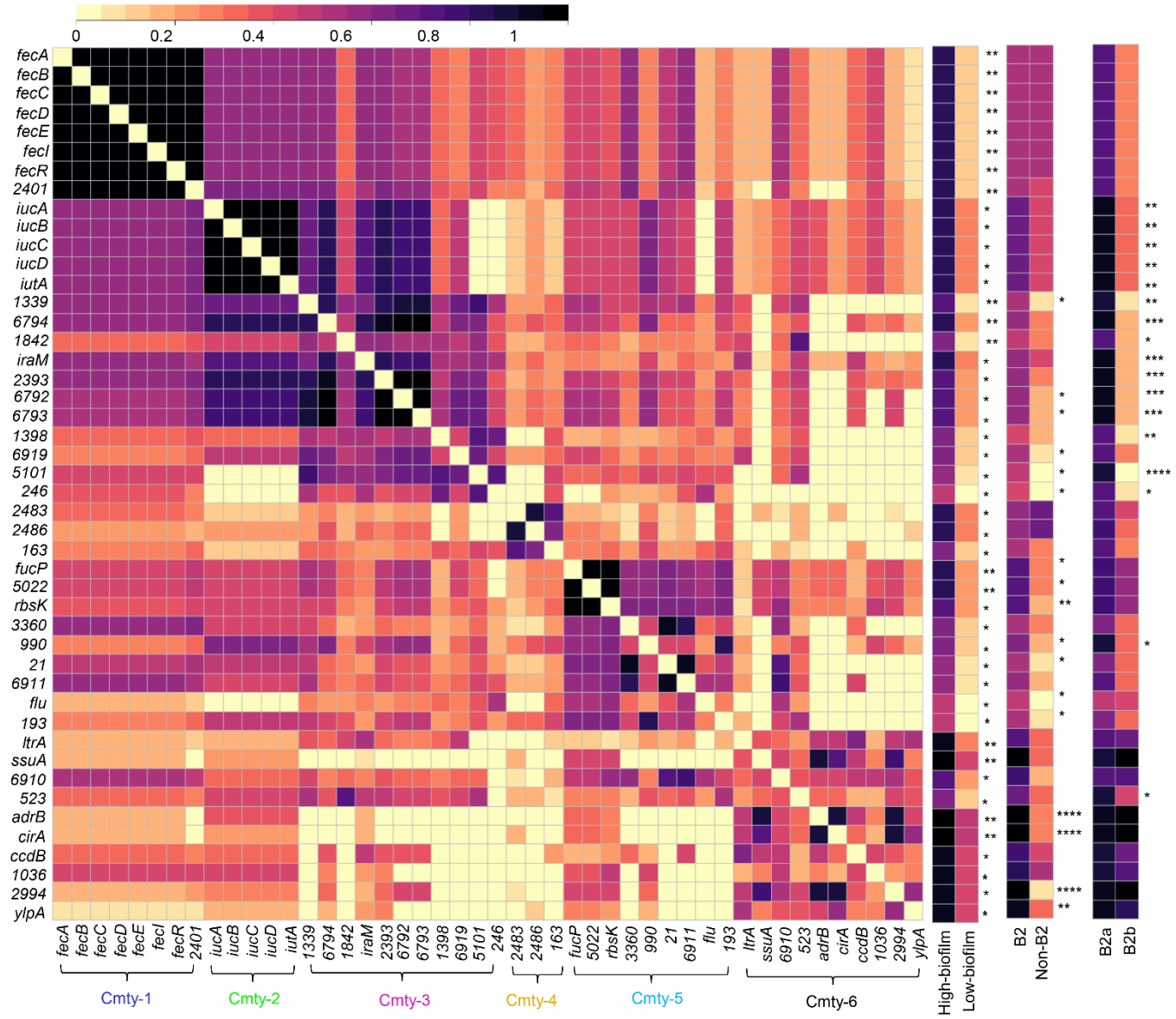


**Fig. 6 Six gene communities are discernible in the correlation matrix heatmap of catheter biofilm-associated genes.** The correlation matrix heatmap depicts statistically significant (5% *P*-value threshold, one-tailed on the right, Fisher's exact test) positive associations between 46 biofilm positively associated genes. Presence frequency comparisons of each gene between different phenotypic and genetic groups, high-biofilm vs low-biofilm, B2 vs non-B2, and B2a vs B2b, were displayed to the right of the correlation matrix. Cmty: community. By two-tailed Fisher's exact with $P <= 0.05$ considered statistically significant. *: $P <= 0.05$. **: $P < 0.01$. ***: $P < 0.001$. ****: $P < 0.0001$.

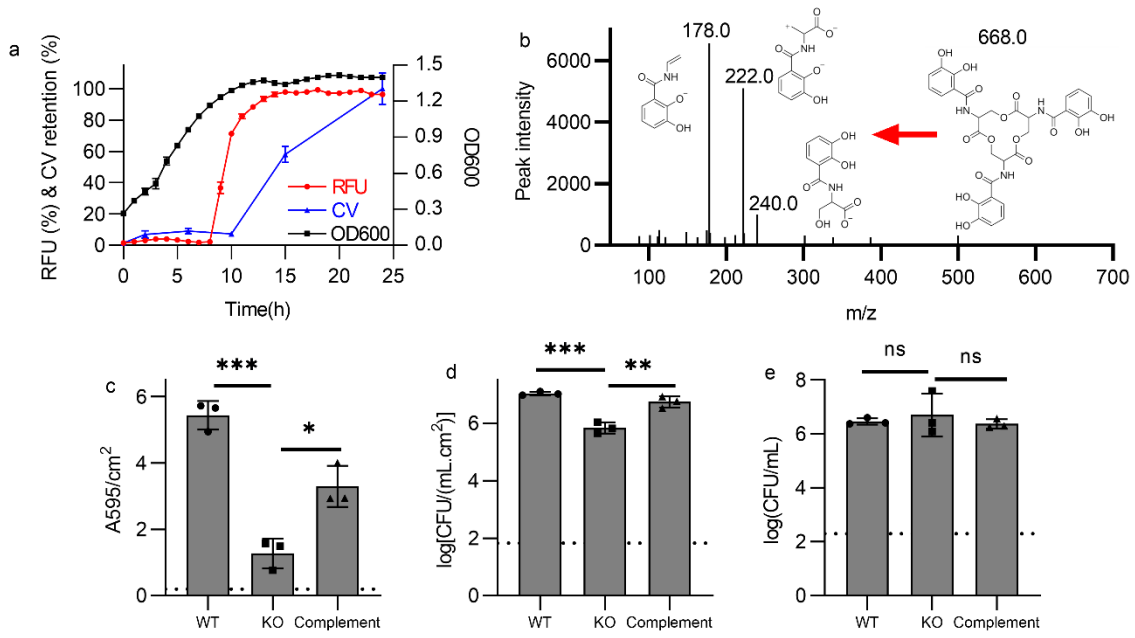


**Fig. 7 Fec expression and extent of catheter biofilm formation**. **(a)** *fec* expression (RFU), catheter-biofilm formation (CV retention, A595/cm$^2$), and bacterial growth (OD600) by the fluorescent reporter strain, EC52::*fecI*-RFP was measured in a microplate assay. **(b)** MS/MS product ion scan spectrum demonstrating presence of the *E. coli* siderophore enterobactin (m/z=-668.1) in the voided media from wild type EC52 cultured in the continuous flow catheter-biofilm assay system. Biofilm biomass (crystal violet retention, **c**), catheter-adherent CFUs (**d**), and planktonic CFUs (**e**) of EC52 (WT), EC52Δ*fecA* (KO), and EC52Δ*fecA*::*fecA* (complement) cultured in the catheter-biofilm assay system. Three replicates with mean and SD plotted. Comparisons conducted using one-way ANOVA with Dunnett's multiple comparisons test. $P <= 0.05$ is considered statistically significant. ns: not significant, *: $P <= 0.05$, **: $P < 0.01$, ***: $P < 0.001$.

## TABLES

**Table 1.** Comparison of study cohort demographics

| Variable[a] | Clinical cohort[b] | | | Two-tailed Fisher's exact test (*P*)[c] | |
|---|---|---|---|---|---|
| | CAUTI (n = 12) | CAASB (n = 16) | RC (n = 13) | (CAUTI + CAASB) vs RC | CAUTI vs CAASB |
| Sex (female) | 2 (17%) | 11 (69%) | 7 (54%) | 0.7442 | 0.0093 |
| Age in years (>= 65) | 6 (50%) | 7 (44%) | 1 (8%) | 0.0309 | 1.0000 |
| Body mass index in $kg/m^2$ (>= 25) | 8 (67%) | 6 (38%) | 8 (62%) | 0.5240 | 0.7022 |
| $TMP/SMX^R$ | 6 (50%) | 9 (56%) | 0 (0%) | 0.0011 | 1.0000 |
| $Fluoroquinolone^R$ | 11 (92%) | 9 (56%) | 0 (0%) | 0.0001 | 0.0882 |

[a] $TMP/SMX^R$, resistance to trimethoprim/sulfamethoxazole antibiotic. $Fluoroquinolone^R$, resistance to fluoroquinolone antibiotics.

[b] CAUTI: catheter-associated urinary tract infection. CAASB: catheter-associated asymptomatic bacteriuria. RC: rectal colonizer.

[c] $P <= 0.05$ is considered statistically significant.

**Table 2.** Assessment of fluoroquinolone resistance in phylotype B2 *E. coli* strains.

| Fluoroquinolone | | B2 subclade | | Two-tailed Fisher's exact test (*P*)[a] |
|---|---|---|---|---|
| | | B2a (15) | B2b (13) | |
| Clinical lab fluoroquinolone resistance result | | 14 (93%) | 0 (0%) | 0.0001 |
| Presence of fluoroquinolone resistance genes (*oqxA*, *oqxB*, *qepA1*, *gyrA*, *parC* , *parE*) | | 15 (100%) | 3 (23%) | 0.0001 |
| SNPs identified in fluoroquinolone resistance genotypes[b] | *gyrA* p.D87N | 14 (93%) | 0 (0%) | 0.0001 |
| | *gyrA* p.S83L | 14 (93%) | 1 (8%) | 0.0001 |
| | *parC* p.E84V | 15 (100%) | 0 (0%) | 0.0001 |
| | *parC* p.S80I | 15 (100%) | 0 (0%) | 0.0001 |
| | *parE* p.I529L | 15 (100%) | 0 (0%) | 0.0001 |

[a] *P* <= 0.05 is considered statistically significant.

[b] *P* <= 0.05 is considered statistically significant. The change here are amino acid substitutions. Several resistance mutations have been characterized in Escherichia coli, and the majority of these are located in the quinolone resistance-determining region (QRDR) defined as codons 67–106 in gyrA and 56–108 in parC (E. coli numbering). (88)

**Table 3.** Select low- and high-biofilm *E. coli* isolates for comparative genomic analysis to identify catheter biofilm-associated genes

| Clade | 95% confidence interval of CV | | Biofilm criteria | | Selected strains | |
|---|---|---|---|---|---|---|
| | Lower bound | Upper bound | Low | High | Low-biofilm | High-biofilm |
| B2a | 0.1 | 2.7 | CV < 0.2[a] | CV > 1.0 | EC12, EC13, EC22, EC28, EC29 | EC14, EC15, EC16, EC18, EC20, EC31 |
| B2b | 0.3 | 2.4 | | | EC34, EC44, EC48, EC50 | EC23, EC33, EC36, EC40 |
| F | -0.8 | 2.3 | | | EC37 | EC42, EC51 |
| D | -0.7 | 1.0 | | | EC39 | |
| A + B1 | -0.7 | 2.4 | | | EC24, EC26, EC27, EC43, EC49 | EC52 |

[a] Low-biofilm criteria of CV < 0.2 is chosen based on the calculation of (0.1 + 0.3)/2 = 0.2.