Title: Pathogenic potential in catheter associated *Escherichia coli* is associated with separable biofilm and virulence gene determinants

Authors: Zongsen Zou[1,2], Robert Potter[3,4], William McCoy[1,5], George Katumba[1,2], Peter J. Mucha[6], Gautam Dantas[3,4,7,8], Jeffrey P. Henderson[1,2*]

Affiliations:

[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.

*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

Conflicts of interest

The authors have declared that no conflict of interest exits.

ABSTRACT

Urinary catheterization facilitates asymptomatic bacterial colonization of the urinary tract and increases the risk of urinary tract infection (UTI). Differentiating catheter-associated asymptomatic bacteriuria (CAASB) from catheter-associated UTI (CAUTI) can be challenging when the presence or origins of symptoms are unclear. To determine whether strain-specific *Escherichia coli* characteristics are associated with CAUTI, we compared genomic composition and experimental catheter biofilm production by urinary isolates from catheterized inpatients and rectal isolates from healthy volunteers. CAUTI isolates were associated with a distinctive phylotype B2 sub-clade dominated by the multidrug resistant, pandemic ST131 lineage. While catheter biofilm formation was widespread among *E. coli* isolates, phylotype B2 biofilms were more extensive and biofilm-associated genes were preferentially found in B2 and ST131 isolates. Distinctive suites of iron-responsive genes were associated with both the ST131 lineage and catheter biofilm formation. Catheter biofilms produced by some CAASB strains could inhibit ST131 colonization, suggesting a potential beneficial function for these strains. These results suggest that the combination of biofilm and non-biofilm-associated gene networks in urinary *E. coli* influence CAUTI risk in catheterized patients.

INTRODUCTION

Catheter-associated urinary tract infections (CAUTI) are among the most common nosocomial infections, with over one million cases yearly in the United States (1-3). Accurate diagnosis and effective treatment of CAUTI and ureteral stent-associated infections, can be challenging. 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. In the absence of UTI symptoms, the risk of serious infection is low and antibiotics are not recommended (4). Altered mental status, neurological deficits, or confounding conditions, however, can diminish the sensitivity and specificity of diagnostic criteria (5-8). When confronted with an uncertain diagnosis, treating physicians must weigh the risk of progressive infection against the risks of catheter or device removal and unnecessary antibiotic therapy. In this context, professional society guidelines have long noted a need to better discriminate catheter-associated asymptomatic bacteriuria (CAASB) from CAUTI and to predict a patient's risk for progression to CAUTI (6, 9). While some bacterial species are more commonly recovered from UTI patients, species identification alone does not distinguish infection from colonization.

Escherichia coli is the predominant bacterial species associated with both uncomplicated UTI and CAUTI (10). This species is remarkably diverse, with some strains categorized into pathotypes associated with a particular type of infection. Unlike enteric pathotypes, there is no definitive genetic signature of a "uropathogenic" *E. coli* strain. Rather, the designation is typically conferred upon a strain with compatible genetic elements collected in the context of a clinical UTI diagnosis (11). Defining a uropathogenic strain in this manner presents numerous challenges. For one, host variation in human clinical populations means that a probabilistic, not an absolute, relationship between strain characteristics and clinical outcome should be assumed (12, 13). Additionally, in a genetically diverse species like *E. coli*, it is plausible that

more than one virulence strategy has evolved, requiring careful attention to gene combinations and lineage within the species. Rather than achieving a strict genomic definition for uropathogenic *E. coli,* a more realistic goal may instead be to identify genes and phenotypes that increase virulence – termed “pathogenic potential” (14) – and to be mindful that host factors may modulate their influence.

In UTI pathophysiology, the presence of a catheter or stent is a distinctive host characteristic that confers a well-recognized predisposition to infection (15, 16). The ability of *E. coli* to form biofilms on these devices is generally regarded as a critical virulence characteristic in these patients (17-19). Biofilms are adherent bacterial communities enmeshed in an extracellular polymeric substance (EPS) matrix. A single *E. coli* strain may form biofilms that are qualitatively distinct, depending on media composition, substrate, and temperature conditions (20-22). In addition to this, clinical *E. coli* isolates are known to carry non-conserved, variable genes proposed to be virulence factors (VFs). These VFs are functionally diverse and include adhesins, toxins, iron acquisition systems, and immune evasion pathways that often appear likely to interact with the host (23, 24). How these biofilm formation and the presence of VFs influence the pathogenic potential of *E. coli* in catheterized patients is incompletely understood.

Here, we compared forty-one *E. coli* isolates collected from subjects with CAUTI, CAASB, and rectal colonization (RC) at a single location. Given the potential role for biofilm formation in catheterized patients, we combined comparative genomic analyses with a quantitative catheter-biofilm phenotyping approach to identify and test relationships between biofilm formation and lineages associated with CAUTI or CAASB. We used a reverse genetics approach to experimentally test the observed association between an iron acquisition system and enhanced catheter biofilm formation. Given the high frequency of catheter biofilm formation among CAASB

strains, we assessed the ability of these isolates to protect the host by preventing catheter colonization by CAUTI-associated strains in an experimental system.

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 337 isolate collection from patients at Barnes-Jewish Hospital/Washington University Medical Center between August 1st, 2009 and July 31, 2010 (25, 26). Of these 62 isolates, 12 met symptom criteria for CAUTI (concurrent fever, T > 38℃) and 16 met criteria for CAASB (lack of fever or reported suprapubic, flank pain). 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). 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 (27, 28). 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 all 41 *E. coli* isolates using whole-genome sequencing and genome alignment. 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 without clonal pairs based on genome differences identified by comparing their carriage profiles of 32 previously published

VFs and carriage numbers of pan-genome genes. A maximum-likelihood tree (Figure 1) based on variable gene content divided isolates into four main clades corresponding to canonical *E. coli* phylotypes B2, F, D and a combination of A, B1, and E (29, 30). Nearly all CAUTI isolates (11/12) belong to phylotype B2 as is typical of extraintestinal *E. coli* (16). 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 clade within a high level B2 subdivision composed entirely of urinary isolates that significantly separates CAUTI from non-CAUTI isolates within B2 (10/15 vs. 1/13, $P = 0.0021$ by two-tailed Fisher's exact test). This same distinctive group of strains (designated as B2a) consisting of 14 ST131 strains (31, 32) and one untypeable strain was also distinguished using sparse principal components analysis (sPCA) (Figure S1) of B2 strains gene composition (33, 34). The two B2 subclades were significantly resolved on PC1 (31%, Figure S1A) axis in the sPCA plot (Figure S1B), with logistic regression using PC1 values yielding a prediction accuracy of 1.0 (Figure S1C, SD = 0) and an AUC of 1.0 (Figure S1D, SD = 0) in five-fold cross validations. The non-B2a isolates within phylotype B2 (designated as B2b) are more diverse and consist of eight sequence types (ST 12, 73, 95, 127, 141, 144, 357, 538). These results indicate that systematic differences in *E. coli* gene composition associated with ST131 are associated with CAUTI.

Antibiotic resistance

The ST131-dominated B2a clade is a globally emergent extraintestinal pathogenic *E. coli* lineage associated with clinically significant antibiotic resistance, most notably to the fluoroquinolone class of antibiotics (31, 32). To determine whether increased antibiotic resistance is associated with B2a, we assessed phenotypic resistance reported from clinical laboratories together with carriage of known resistance genes and polymorphisms. Antibiotic resistance genes (ARGs) against aminoglycosides, beta-lactams, amphenicols, TMP/SX,

macrolides/lincosamides/streptogramins (MLS), quinolones, and tetracyclines were identified in the genomic assemblies (Table S1). Both phenotypic and genotypic fluoroquinolone resistance were significantly more common in B2a strains 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, $P = 0.0001$). B2a group strains also exhibited increased frequencies of ARGs associated with TMP/SX, beta-lactams, and aminoglycosides (Table S1, $P < 0.03$). The high frequency of resistance genes, particularly those related to fluoroquinolones, is consistent with the presence of globally disseminated ST131 lineages in group B2a.

Virulence factor content of CAUTI strains

In this cohort, we considered that gene combinations carried by B2a isolates may enhance their pathogenic potential in catheterized patients. The presence of virulence factors (VFs) associated with pathogenic gains of function are inferred by the presence of representative genes (23). We identified 32 such VFs in isolates from our cohort (Figure 1). The number of VFs per isolate, previously called the “virulence score” (35), did not distinguish ($P = 0.147$) CAUTI (9.5 ± 3.8), CAASB (8.8 ± 3.6), or RC (9.5 ± 4.3) isolates (Figure S2). The virulence score was higher in phylotype B2 isolates (10.7 ± 2.7 vs. 5.8 ± 4.0, $P = 0.002$) than in non-B2, and with a non-significant trend toward a lower score in B2a than B2b (10.0 ± 2.4 vs. 11.5 ± 3.0, $P = 0.138$).

We next considered that VFs exert non-equivalent or synergistic effects on pathogenic potential in which qualitative or combinatorial differences in VF carriage are relevant (35). For this, we used network community detection to identify co-associations between the 32 VFs. We performed modularity-based community detection using the Louvain method on a weighted network of positive correlations to assess correlations between the 26 VFs present more than

once (> 2.4%) among isolates (26, 36). Three gene communities were resolved, as visualized by the force-directed network layout (Figure 2A) and the corresponding correlation matrix (Figure 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 the evolutionary accumulation of cooperative functions but may also represent accumulation of non-selected, incidental genetic features.

VF communities 1 and 2 were more common in B2 than non-B2 isolates, while community 3 VFs were exclusively associated with non-B2 isolates. In B2 isolates, the VF genes *fyuA* (37), *chuA* (37), *ompT* (38), and *usp* (39) were both nearly ubiquitous and significantly ($P < 0.00002$) more common than in non-B2 strains. Within B2 isolates, reciprocal relationships among VFs in groups B2a and B2b were evident. In B2a isolates, the community 1 VF genes *iucD* (40)*, sat* (41), and *iha* (42) were nearly ubiquitous and significantly ($P < 0.008$) more common than in B2b. In B2b isolates, the community 2 VF genes *vat* (37)*, iroN* (43), and *prf* (35) were significantly ($P < 0.015$) more common than in B2a. Together, these results are consistent with one set of VFs that raise the pathogenic potential of B2 strains with another set of VFs raising the pathogenic potential of group B2a strains. Some of these VFs may be more consequential than others or may exert synergistic effects when present in specific combinations with other VFs.

Biofilm formation by CAUTI, CAASB, and rectal isolates

The ability of *E. coli* to form biofilms on catheters is regarded as an important contributor to pathogenic potential. *E. coli* strains can form multiple biofilm and biofilm-like structures that may vary in important ways depending on media and available substrates (20-22). Previous biofilm characterization of multidrug resistant *E. coli* ST131 strains revealed significant biofilm differences among strains and assay conditions (44). In this study, we sought to simulate the

clinical catheter environment in a controlled experimental system (45, 46). We therefore compared biofilm formation between *E. coli* isolates using an *ex vivo*, continuous flow model (Figure S3) engineered to simulate distinguishing features of the urinary catheters used in patients. Bacterial strains are inoculated into a silicone tube and urine or standardized artificial urine medium (AUM, Table S2) (47, 48) is flowed through the lumen, which is maintained at 37 °C. We found that AUM yielded growth kinetics (Figure S4, $P = 0.3$) and biofilm morphology in this system (Figure 3A&B) comparable to that observed with filter-sterilized human urine. We used this standardized quantitative phenotyping system to compare all 41 *E. coli* isolates using AUM. Biofilm formation was measured using crystal violet (CV) dye retention and catheter-adherent live bacteria as colony-forming units (CFUs). We also measured non-adherent "planktonic" CFUs in the voided media effluent from the flow system, as an analog to the clinically accessible fraction of colonizing bacteria.

We measured substantial inter-strain variation in catheter biofilm formation, with CV retention ranging from 0.02 to 10.60 A_{595}/cm^2, and adherent CFUs 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 (Figure 3D, $P = 0.02$) with a non-significant trend (Figure 3C, $P = 0.22$) toward higher CV retention. Within phylotype B2 isolates, neither adherent CFUs nor CV retention was significantly different between groups B2a and B2b (Figure 3F&G, $P > 0.6$). Planktonic CFUs were relatively uniform regardless of adherent CFU or CV retention values, with non-significant differences between groups in all comparisons (Figure 3E&H, $P > 0.1$) that may reflect bacterial persistence within loosely adherent communities or turbulent flow through the system. Together, these results are consistent with a widespread potential for catheter biofilm formation among *E. coli*, with evidence for greater extent of biofilm in phylotype B2.

Identification of catheter biofilm-associated genes

We considered the possibility that lateral transfer of genes related to catheter biofilm formation occurs as it does with VFs (49, 50). To identify these putative catheter biofilm-associated genes, we conducted a comparative genomic analysis supervised by high versus low CV retention. Within each of the four main clades, B2, F, D, and A + B1, on the phylogenetic dendrogram (Figure 1), high and low biofilm isolates were designated based on the 95% confidence intervals of their CV retention values. This analysis identified 13 high and 16 low catheter biofilm formers respectively. We next used sparse partial least squares discriminant analysis (sPLSDA) to identify contributors to the genetic differences between the two isolate groups (33, 34). In the resulting sPLSDA plot, high and low biofilm formers were well-resolved along PC1 (Figure 4A), with logistic regression analysis of PC1 (9%, FigureS5A) values yielding a prediction accuracy of 1.0 (Figure S5B, SD = 0) and an AUC of 1.0 (Figure S5C, SD = 0) with 5-fold cross-validation. Seventy-two genes with the highest PC1 loading ($P < 0.05$ by two-tailed Fisher's exact test) were associated with catheter biofilm, with 46 and 26 genes identified with positive (high-biofilm) and negative (low-biofilm) correlations, respectively (Table S3). Of note, the Antigen 43 gene (*flu*) was among the positively associated genes and exhibited the 44th highest PC1 loading ($P = 0.04$ by two-tailed Fisher's exact test), providing a point of reference to a previously described *E. coli* biofilm gene (51, 52).

To determine whether *E. coli* isolates carry the 46 biofilm positively associated genes in co-associated groups, we performed modularity-based community detection using the Louvain method on a weighted network of positive correlations as described above for VFs (26, 36). From this analysis, six gene communities were resolved, as visualized by the force-directed network layout (Figure 4C) and the corresponding correlation matrix (Figure 5). Community 1 genes exhibited the highest PC1 loading, were most highly associated with high biofilm isolates ($P = 0.0025$), and exhibited the highest betweenness centrality ranking, consistent with a

possibly central role in the catheter biofilm phenotype. Community 1 was maximally robust, defining the ferric citrate transport (*fecABCDIR*) locus (53, 54). Community 2 defined the aerobactin siderophore system (40) represented by the *iucD* VF marker gene. Communities 3–6 were less robust and consist of a variety of genes with and without known functions. The previously described biofilm gene *flu* (51, 52) was present in community 5. Multiple genes in the catheter biofilm gene communities 2-6 were also identified to be more common in B2 than non-B2, and B2a than B2b strains (Figure 5). Moreover, the *fec*-defined community 1 genes exhibited a non-significant trend (P = 0.056) trend toward higher frequency in B2a strains, where they were present in 11/15 isolates.

These results are consistent with a complex polygenic contribution to catheter biofilm production that involves iron-responsive transport systems (ferric citrate and aerobactin systems, the Cir importer) and autoaggregation (Antigen 43/Flu). The significant positive associations of multiple biofilm genes with B2 and B2a strains suggests a potential role for biofilm-associated genes in the increased pathogenic potential of CAUTI isolates, though their presence does not differentiate these strains as strongly as canonical the virulence factors as described above.

Fec system activity is associated with increased catheter biofilm formation

The strong association between biofilm formation and the Fec system, as well as its extensive genomic connectedness with other biofilm associated genes, raises the possibility that it is functionally linked to increased biofilm formation (54). To address this, we first examined its transcriptional profile and measured catheter biofilm formation in a defined mutant in which Fec function was knocked out. The *fec* system imports iron as ferric citrate complexes from the extracellular space to the periplasm through the *TonB*-dependent outer membrane transporter (FecA) (55, 56). Upon ferric citrate binding, FecA also actives a signaling cascade by interacting with the inner membrane protein FecR which activates the alternative sigma factor FecI.

Activated FecI then initiates transcription of other *fec* genes (*fecABCDE*) (53, 54). To determine whether the *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, phylotype A rectal isolate EC52, shaking microplate cultures of EC52::*fecI-RFP* exhibited a sharp fluorescence increase at 8 hours when bacteria enter the stationary phase, immediately preceding detectable biofilm formation (Figure 6A). *Fec* transcription was thus associated with *E. coli* biofilm formation.

Next, to determine whether the Fec system directly promotes catheter biofilm formation, we compared isolate EC52 with its isogenic *fecA* deletion mutant EC52Δ*fecA* (54). In the catheter biofilm culture system, both CV retention (Figure 6B, $P = 0.0001$) and sessile bacterial counts (Figure 6C, $P = 0.0002$) were significantly lower in EC52Δ*fecA* relative to wild type isolate EC52, while the genetic complementation of EC52Δ*fecA* with a *fecA* expression plasmid reversed these biofilm formation deficits (Figure 6B&C, $P < 0.006$). No reduction with regards to the planktonic bacteria counts in the voided media was observed (Figure 6D, $P = 0.7$), and Fec knock-out was not associated with differences in shaking, non-biofilm growth density (Figure S6). These data are consistent with the Fec system directly promoting catheter-associated biofilm formation.

CAASB strains can inhibit CAUTI colonization

Although there is evidence of increased catheter biofilm predisposition among CAUTI-associated lineages, competence for catheter biofilm formation is widespread among all *E. coli* isolates in this study. Differences in VF carriage, in contrast, are more prominent. This state of affairs may reflect a role for biofilm in localizing strains to the catheterized urinary tract that is not, in the absence of VF activities, sufficient to substantially increase pathogenic potential. The

relationship between biofilm and CAUTI-associated lineages may be further weakened if biofilm-forming strains of low pathogenic potential protect the human hosts from infection. This may occur when CAASB strains prevent colonization with strains with higher pathogenic potential (57-59). A prominent example of this is *E. coli* 83972, a B2 strain isolated from a patient with persistent asymptomatic bacteriuria that has shown efficacy in preventing clinical UTIs following catheter-mediated bladder colonization of patients (60, 61). To determine whether CAASB-associated strains can similarly prevent colonization by CAUTI-associated strains, we modified our experimental catheter colonization system to measure competitive colonization by multiple *E. coli* strains. We first pre-colonized catheters with each of 11 non-ST131 CAASB isolates and subsequently challenged these catheters with ST 131 CAUTI isolate EC20, a high-biofilm former as defined above. We then used single-nucleotide polymorphism-selective qPCR (SNPs-qPCR) to distinguish and quantify genomes corresponding to ST131 CAUTI isolates and non-ST131 CAASB isolates.

Upon screening the 11 CAASB isolates, we found that pre-colonization with EC36 (phylotype B2, group B2b) and EC25 (phylotype A) markedly inhibited both the catheter biofilm-adherent (Figure 7A, $P < 0.004$) and planktonic populations (Figure 7B, $P < 0.0003$) by ST131 isolate EC20 (> 2.5 log inhibition). Pre-colonization with either CAASB isolate suppressed the planktonic population of the ten CAUTI ST131 isolates (Figure 7D, $P < 0.0001$, Mann-Whitney test), with isolate EC36 also inhibiting the catheter biofilm -adherent population (Figure 7C, $P = 0.0005$). These results demonstrate that a subset (2/11, 18%) of CAASB catheter biofilms are capable of preventing subsequent catheter colonization by CAUTI-associated ST131 *E. coli* isolates. This is consistent with the ability of a subset of CAASB strains to prevent colonization by *E. coli* strains with antibiotic resistance and elevated pathogenic potential (77). These results raise the possibility that, in a substantial proportion of catheterized patients, *E. coli* catheter biofilm formation may not be exclusively pathogenic in nature.

DISCUSSION

This study identifies a lineage within the canonical extraintestinal pathogenic B2 phylotype that was associated with CAUTI in a hospitalized population. This CAUTI-associated lineage is dominated by pandemic ST131 strains associated with multiple antibiotic resistance and possessing a distinctive combination of virulence factor genes relative to the other isolates (31, 32). Using a simulated catheter biofilm culture system (45, 46), we found that most *E. coli* were competent for catheter biofilm formation, that phylotype B2 exhibited enhanced biofilm production, and that ST131 biofilms were comparable to those of other B2 strains. A comparative genomic analysis identified a diverse series of new catheter biofilm-associated genes, many of which are associated with B2 or ST131 (B2a) isolates (34, 36). The ferric citrate transport system, which was most strongly associated with catheter biofilms, was functionally linked to an enhanced biofilm phenotype. Catheter biofilm was not exclusive to CAUTI-associated lineages and pre-colonization with CAASB isolates from other lineages could diminish ST131 colonization, suggesting both a protective and pathogenic role for catheter biofilm formation among *E. coli.*

These results are consistent with both harmful and beneficial roles for *E. coli* biofilm formation on device surfaces (57, 58). It is notable that infection risk is associated with specific *E. coli* strain-level genetic differences found in pandemic ST131 strains. The combined virulence-associated characteristics, catheter biofilm potential, and antibiotic resistance in this lineage reinforces the need for a judicious therapeutic approach to catheterized patients (1, 5). Antibiotic treatment of ST131 CAUTI is likely to be hampered by both phenotypic (biofilm) and intrinsic (antibiotic resistance genes) resistance, increasing the chances of antibiotic failure or infection recurrence and increasing the importance of device removal. In contrast, non-ST131 *E. coli* colonization or bacteriuria may have salutary effects for catheterized patients (62). As a result,

fluoroquinolones and other antibiotics may paradoxically increase infection risk by eliminating antibiotic susceptible non-ST131 CAASB strains, lowering the barrier for colonization by resistant ST131 strains (31, 32). These scenarios generally support guideline recommendations to avoid antibiotic treatment of CAASB and suggest that early awareness of ST131 bacteriuria could be of clinical value. A ST131 biomarker might 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 cause CAUTI remain unclear but are of diagnostic and therapeutic interest (17-19). These processes are presumably complex and include biofilm efflux, host tissue adhesion, immune evasion, and nutrient acquisition (64, 65). 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. This is suggested by the presence of VFs in the favored network communities described here (Figure 2) and in previous work (26). 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.

E. coli biofilms are a complex phenotype arising from multiple processes. The uropathogenic *E. coli* strain UTI89, for example, is capable of forming multiple qualitatively distinct biofilm types depending upon media composition, substrate, and temperature conditions (20-22). In catheterized patients, *E. coli* biofilms are influenced by urinary composition, urinary flow, and

catheter properties, features not typically incorporated into biofilm assays (66). In the present study, some of the most prominent biofilm correlates are associated with iron uptake (Figure 5) and likely reflect the low iron conditions of natural and simulated human urine. The pivotal role identified for the Fec system, in particular, may reflect the availability of citrate, a distinctive human urinary component (53, 54). That some of the genes distinguishing high biofilm strains also distinguish B2 from non-B2 strains, and B2a from B2b strains, is consistent with a role in CAUTI pathogenesis. An additional role for catheter biofilm formation in preventing infections, however, may weaken the association between CAUTI and biofilm determinants in the comparative genomic analyses used in the present study.

Infection prevention by colonizing bacteria is a widely accepted concept in the large intestine and has also been therapeutically attempted in patients with urinary catheters. Some clinical trials of this bacterial interference approach have shown efficacy and benefits over pharmacologic treatments (57-59). Most studies of catheterized patients have focused upon the prototypical asymptomatic bacteriuria *E. coli* strain 83972 (60, 61) a B2 strain resembling B2b strains in the present study. Here, CAASB strains were diverse and widely distributed among multiple lineages. One interpretation of this diversity relative to the more restricted CAUTI lineage is that there is a low evolutionary barrier to colonization of catheterized patients that is surmounted in many lineages with an additional, higher evolutionary barrier to tissue invasion is surmounted by fewer strains. In this paradigm, the interaction between catheterized hosts and urinary bacteria favors asymptomatic colonization. Whether there is an additional evolutionary barrier to niche exclusion of CAUTI-associated strains by CAUTI strains is unclear, though the variation we observe in competitive catheter colonization (Figure 7) is consistent with this.

In conclusion, we found that CAUTI in the study population was associated with *E. coli* lineage largely defined by emergent ST131 strains. The pathogenic potential of these populations was

associated with carriage of specific gene networks that was largely distinct from networks associated with biofilm formation. While biofilm gene representation tended to be greater in CAUTI-associated lineages, biofilm formation was widespread and, in some strains, acted to prevent ST131 strain colonization. ST131 strains appear well-adapted to cause infection in patients with urinary catheters. The high degree of resistance to fluoroquinolones, an antibiotic class commonly used to treat UTIs, in CAUTI-associated B2a strains raises the intriguing possibility that these strains arose from co-evolution with catheterized human hosts. 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.

METHODS AND MATERIALS

Urinary isolate selection

Urinary catheter-associated *E. coli* isolates were identified from the study of bacteriuria (> 5×10^4 CFU/mL) in patients conducted by Marschall *et al.* (25, 26). 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 (67). Documented UTI symptoms were not required for CAUTI definition due to their poor reliability in inpatients (5). CAASB was defined as bacteriuria in the absence of either fever or attributable urinary symptoms (dysuria, lower abdominal pain, flank pain).

Rectal isolate collection

Rectal *E. coli* isolates were collected from healthy, adult volunteers in St. Louis, MO from 2014-2015. This prospective, observational study was approved by the WU-IRB, and all study subjects provided written informed consent. Exclusion criteria were age < 18 years old, pregnancy, current urinary tract infection, previous urogenital surgery (e.g., nephrectomy, neobladder formation, hysterectomy, prostate resection), 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 (68) to procure a self-collected rectal swab (BD Eswab, Catalog# 220245), stored the swab at room temperature, and then submitted it with a study survey immediately after swab collection. Swabs were processed by the clinical microbiology lab at Barnes-Jewish Hospital (St. Louis, MO) 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 collection

Healthy donor urine was collected from adult volunteers as approved by the Institutional Review Board of Washington University School of Medicine. 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 through 0.22 µm filters, 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 (69, 70).

Bacterial strains and culture conditions

Isogenic *E. coli* mutant strain was constructed as in-frame deletion in chosen isolate using the Lambda Red recombinase method as described previously. Isogenic mutant complementation and fluorescence reporter construct were accomplished by ectopic expression using transformed plasmids (36, 71). 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 the bacteremia 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, San Diego, CA). 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 (72, 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, 76). The core-genome alignment was constructed into a maximum likelihood tree with raxML and viewed in iTOL (77). *In silico* multilocus sequence types (MLST) were identified using BLASTN to the *E. coli* MLST database (29, 30). Previously published virulence factors (VFs) were annotated in the *E. coli* draft genomes using virulencefinder v1.5 and blastp to previously described genes (Kleinheinz). 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). The genomes analyzed in this report have been deposited to NCBI WGS database under BioProject accession no. PRJNA514354.

Genomic analysis

The thirty-two documented 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 (33, 34). Biofilm genes were determined by whole-genome comparative analyses using sparse partial least squares discriminant analysis (sPLSDA), logistic regression (LR) classification, and network analysis approaches. 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, respectively (36), as well as the Gephi software (http://gephi.org). Computational algorithms with sparsity regularization can perform feature selection by maximizing the original objective function under conditions of minimizing incorporated features to prevent overfitting (79). Because of the high sparsity of whole genome sequence metadata, with 15993 genes identified in 41 genome assemblies (15993 >> 41), sparse penalty therefore was enforced in all

dimensionality reduction analyses (sPCA and sPLSDA) in the present work. The network analysis approach is described in more detail in next section.

Network analysis

Two network representations for the 26 virulence factors (VFs) and 46 biofilm positively associated genes, connected by co-occurrences across the clinical *E. coli* collection, were defined using statistically significant positive correlation coefficients as the edge weights across the two groups of gene targets. Statistical significance between two nodes (genes) was determined by Fisher's exact test to determine whether they appeared independently within the *E. coli* isolates population, conditional on their observed marginal frequencies among the *E. coli* collection. The 0.4% and 5% *P*-value thresholds (one-tailed on the right) were chosen for the VFs and biofilm-positive genes networks, respectively, to ensure that the obtained gene network in each case was a single connected component. An edge was defined as present between any pair of positively correlated nodes that satisfied the significance threshold, with edge weight equal to the positive correlation coefficient. Communities in this network were detected using the Louvain method by maximizing the modularity function (26, 36). We selected the obtained 3-community (Resolution = 1.0) and 6-community (Resolution = 1.25) for the network visualizations of 26 VFs and 46 biofilm-positive genes, respectively, using a force-directed layout generated by the ForceAtlas2 algorithm in Gephi (http://gephi.org) and the corresponding correlation matrix.

Artificial urine medium (AUM)

Artificial urine medium (AUM) (Table S2) was prepared as an alternative medium of human urine for characterizing biofilm formation. The AUM in this study was based on a previously published recipe (47). Comparison between old AUM recipe (47) with human urine composition from our inductively coupled plasma-mass spectrometer (ICP-MS) test and previously

publication (48) revealed that the old recipe included more iron and less zinc than human urine (Table S4). Therefore, to better mimic human urine composition, the old AUM recipe was optimized by removing the 5 µmol/L $FeSO_4.7H_2O$ while adding extra 7 µmol/L $ZnSO_4.7H_2O$ (Table S2). 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% (69).

Continuous flow biofilm-catheter model system

Biofilms were grown in a continuous flow catheter model, using previously published protocols with appropriate modifications (Figure S3) (45, 46). The components used for assembling continuous flow system included the platinum-cured silicone urinary catheters (Thermo Scientific Nalgene™ 50 8060-0030), peristaltic pump (Watson Marlow 205U), flexible tubings (Fisher Scientific Tygon S3™), and plastic connectors (Thermo Scientific). Prior to use, all tubing, connectors and containers were autoclaved. Human urine and AUM were filter-sterilized through 0.22 µm filters. 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. Subsequently, fresh medium was pumped through the catheter at the flow rate of 0.5 mL/min, with thirty-minute pre-flush allowed to wash off the loosely adherent bacteria first. Next, at 10 hours of continuous flow incubation, the voided media and catheters were collected for subsequent characterization.

Biofilm was quantified by crystal violet (CV) retention measuring the total biofilm biomass, and CFU assays detecting the sessile bacteria in the biofilm matrix and shedding planktonic bacteria in the void media (45). Catheter (11 cm) collected from the flow system was washed with PBS and cut into three pieces (3 cm), with two pieces for CV staining and the other one for CFU enumeration. For CV staining, 3 cm catheters were stained in 0.5% CV solution for 10 min, followed by three times deionized water washing to remove the excess dye. The washed catheters were first dried by capillary action on adsorbent paper, and then allowed to air dry at room temperature for overnight. Next, the stained catheters were filled with 33% acetic acid to dissolve bound dye for 10 min. Finally, the acetic acid solutions with dissolved crystal violet were taken out, diluted by 20 times, and then measured at 595 nm using a UV/Vis Spectrophotometer (Beckman Coulter DU-800). The biofilm biomass was quantified as A_{595}/cm^2 by the CV staining method. For sessile bacteria quantification, 3 cm catheter was cut into fragmented pieces, immersed into 3 mL PBS, sonicated (Branson 350) for 10 min, vortexed for 3 min (GeneMate), and proceed for CFU plating to quantify the adhered bacteria detached from the catheter surface. The amount of biofilm bacteria was quantified as CFU/(mL × cm^2). For planktonic bacteria quantification, voided media were collected and plated for CFU enumeration, quantified as CFU/mL.

Characterization of biofilm structure

Biofilm grown on urinary catheter surface was processed for structure characterization using transmission electron microscopy (TEM) (80). For microstructural characterization, 1 cm catheter with biofilm formed on inside surface was fixed in 2% paraformaldehyde/2.5% glutaraldehyde (Polysciences Inc., Warrington, PA) in 100 mM sodium cacodylate buffer (pH 7.2) for 1 h at room temperature. Samples were washed in sodium cacodylate buffer and postfixed in 1% osmium tetroxide (Polysciences Inc.) for 1 h. Samples were then rinsed extensively in deionized water prior to en bloc staining with 1% aqueous uranyl acetate (Ted Pella Inc., Redding, CA) for 1

h. Following several rinses in dH20, samples were dehydrated in a graded series of ethanol and embedded in Eponate 12 resin (Ted Pella Inc.). Sections of 95 nm were cut with a Leica Ultracut UCT ultramicrotome (Leica Microsystems Inc., Bannockburn, IL), stained with uranyl acetate and lead citrate, and viewed on a JEOL 1200 EX transmission electron microscope (TEM) (JEOL USA Inc., Peabody, MA) equipped with an AMT 8 megapixel digital camera and AMT Image Capture Engine V602 software (Advanced Microscopy Techniques, Woburn, MA)

fec fluorescence and microplate-biofilm assays

To determine the expression of ferric citrate transport (*fec*) pathway during biofilm formation, a *fec* fluorescence assay was performed using transformed plasmid (54). The promoter sequence of gene *fecI*, chosen from the gene cluster of *fecABCDIR*, was inserted in one plasmid to manage the expression of red fluorescence mCherry protein. Next, the *fec* fluorescence plasmid was transformed into isolate EC52, a high-biofilm former, for making the fluorescence reporter strain, EC52::*fecI*-RFP (Table S5). A control group comparator, EC52::RFP, was also constructed using the same plasmid but without incorporated *fec* genes managing the fluorescence protein expression (Table S5). Primers used to construct the red fluorescence reporter plasmid, *fecI*-RFP, were designed using OligoEvaluator (Sigma-Aldrich), synthesized using Integrated DNA Technologies, and validated by polymerase chain reaction (PCR) prior to use.

The fluorescence assay was performed using a microplate method to monitor the *fec* expression levels at different stages of bacterial growth in M63 minimum medium over 24 hours. To compare between *fec* expression and biofilm formation, the biofilm growth of reporter strain, EC52::*fecI*-RFP, was also characterized in a microplate test under the same conditions as the fluorescence assay. At 0, 2, 6, 10, 15, and 24 hours of shaking incubation, bacterial cultures were removed from the microplate. To quantify the biofilm formed on the inner surface of

microplate wells, the wells were washed using PBS, stained using 0.5% CV solution for 10 min, washed with deionized water three times, dried at room temperature overnight, dissolved using 33% acetic acid for 10 min. Samples were then diluted 20 folds and measured at 595 nm using a UV/Vis Spectrophotometer (Beckman Coulter DU-800). The biofilm formation was quantified as A_{595}/cm^2. All microplate assays were performed using a Tecan microplate reader under 37 °C.

fec mutant biofilm assay

Strain EC52, one high-biofilm former, was chosen for making the isogenic *fec* mutant, as another method to investigate the *fec* pathway affecting catheter-biofilm formation. Gene f*ecA*, encoding the ferric citrate outer membrane receptor, was deleted using Lambda Red recombinase method as described before, creating the isogenic mutant EC52Δ*fecA* (Table S5) (81). Mutant strain EC*52*Δ*fecA* was also genetically complemented using the transformed plasmid, generating EC52Δ*fecA*::*fecA* (Table S5) as the control-group comparison. The catheter-biofilm formation of EC52, EC*52*Δ*fecA, and EC52*Δ*fecA*::*fecA* were measured in the continuous flow system using the same protocol as described above, and compared between each other. Primers used in the *fecA* deletion and complementation were designed using OligoEvaluator (Sigma-Aldrich), ordered from Integrated DNA Technologies, and validated by following procedures including PCR and gel electrophoresis.

Bacterial interference test

Bacterial interference between 11 non-ST131 CAASB and 10 ST131 CAUTI isolates (Table S6) were tested in the continuous flow catheter-biofilm system with consecutive flow stages (Figure S3B, Figure S7). Urinary catheter was pre-colonized with a low-virulence CAASB biofilm for 10 hours. The CAASB biofilm was then challenged by one high-virulence CAUTI isolate for another 10 hours. At the end of the two consecutive flow stages (10 hours + 10 hours), the voided media

and urinary catheters were collected for subsequent quantifications. Using the same protocols as described above, both catheter-bound and planktonic bacteria were first plated for CFU enumeration. Next, bacteria cultures were spun down, with the pellets collected and proceeded to determine the percentages of CAASB and CAUTI bacteria in the mixed cultures using a real-time PCR (qPCR) method. The qPCR quantification method was established based on employing single-nucleotide polymorphisms (SNPs) for inter-bacteria distinction, with more details described in the “Single-nucleotide polymorphisms (SNPs-qPCR)” section. The amounts of CAASB and CAUTI bacteria in both planktonic and biofilm pellets were quantified based on the SNPs-qPCR and CFU results, and used to determine the efficacy of pre-colonized CAASB biofilm for defending against the colonization by CAUTI isolates. All flow experiments were conducted aseptically under 37 °C using AUM.

Single-Nucleotide Polymorphisms qPCR (SNPs-qPCR)

The method used for quantifying CAASB and CATUI in mixed bacteria pellets was called single-nucleotide polymorphisms real-time PCR (SNPs-qPCR) in the present work. First, genes containing strain-specific SNPs were identified to distinguish between the two strains in one bacterial interference test (32), and chosen to be used in following qPCR assays. Consequently, two housekeeping genes *adk* (adenylate kinase) and *gyrB* (DNA gyrase) were identified with strain-specific SNPs for differentiating between 11 non-ST131 CAASB and 10 ST131 CAUTI isolates. Three SNPs-containing regions in gene *adk* were identified to distinguish CAASB isolates EC24, EC25, EC26, EC27, EC37, EC38, and EC39 from ten CAUTI isolates. Two SNPs harboring portions in another gene, *gyrB*, could differentiate CAASB isolates EC33, EC34, EC35, and EC36 from the ten CAUTI isolates. PCR primers amplifying the SNPs-containing portion in each gene were designed using OligoEvaluator (Sigma-Aldrich), ordered from Integrated DNA Technologies, and validated by PCR following gel electrophoresis prior to qPCR tests. Prior to qPCR tests, cellular DNAs were extracted using Wizard Genomic DNA

Purification Kit (A1120, Promega), and measured by NanoDrop 2000 Spectrophotometer (Thermo Scientific). All qPCR assays were performed on a CFX96 Real-Time System (BIO-RAD). The 20 uL PCR mixture contained 1× iTaq Universal SYBR Green Supermix (BIO-RAD), 0.2 uM of each primer, and 3 ng/uL 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℃, ended with a melt curve with 5 sec at 65℃ first and 5 sec each at 0.5℃ increase between 65℃ and 95℃ (82). Threshold cycles (*Cq*) were obtained to determine the amount of DNA (ng/uL) for each isolate in the mixed bacteria culture, and the quantification was achieved using the calibration curve [*log(Cq) ~ log(DNA)*] of 0.09375, 0.1875, 0.75, 1.5, and 3 ng/uL (Table S7). The DNA percentages was determined based the standard curve and used to calculate the bacterial CFU of each isolate in the mixed cultures collected from interference tests.

Statistical methods

GraphPad Prism 8.0 (GraphPad software) was used to generate graphs and perform statistical analysis in this study. We used 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.

AUTHOR CONTRIBUTIONS

ZZ and JPH developed the concept, designed the overall study approach and constructed patient cohorts for analysis. ZZ performed the experiments. WM designed, funded, and conducted the healthy volunteer rectal *E. coli* collection. RP and GD conducted genome sequencing and alignments. ZZ and GK designed gene deletion and ectopic expression approaches. ZZ, PM, and JPH designed and performed the network analyses. ZZ and JPH analyzed the data. ZZ and JPH wrote the manuscript.

ACKNOWLEDGEMENTS

We thank Wandy Beatty for assistance with transmission electron microscopy (TEM). We also thank the Edison Family Center for Genome Sciences and Systems Biology staff, Eric Martin, Brian Koebbe, MariaLynn Crosby, and Jessica Hoisington-López for their expertise and support in in sequencing and high-throughput computing. The authors acknowledge funding from the Centers for Disease Control Prevention Epicenters Program Grant (CU54 CK 000162). JPH acknowledges National Institutes of Health grants R01DK099534 and R01DK111930. GD acknowledges National Institutes of Health grants U01AI123394 and R01AI155893. PJM acknowledges National Institutes of Health grant R01DK111930. WM acknowledges support by the Washington University School of Medicine, Department of Medicine, Mentors in Medicine (MiM) Program, a KL2TR002346 – ICTS Institutional Career Development Program, the Washington University Institute of Clinical and Translational Sciences grant UL1TR002345 from the National Center for Advancing Translational Sciences (NCATS) of the National Institutes of Health (NIH), and the National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS) of the NIH under award number 1K08AR076464-01. RFP acknowledges support from 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 and NIH.

REFERENCES

1. Trautner BW. Management of catheter-associated urinary tract infection. *Curr Opin Infect Dis.* 2010;23(1):76-82.
2. Foxman B. Epidemiology of urinary tract infections: incidence, morbidity, and economic costs. *The American Journal of Medicine.* 2002;113(1):5-13.
3. 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.
4. 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.
5. 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.
6. 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.
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. Skelton-Dudley F, Doan J, Suda K, Holmes SA, Evans C, and Trautner B. Spinal Cord Injury Creates Unique Challenges in Diagnosis and Management of Catheter-Associated Urinary Tract Infection. *Top Spinal Cord Inj Rehabil.* 2019;25(4):331-9.
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. 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.
11. Subashchandrabose S, and Mobley HLT. Virulence and Fitness Determinants of Uropathogenic *Escherichia coli*. *Microbiol Spectr.* 2015;3(4).
12. Reigstad CS, Hultgren SJ, and Gordon JI. Functional genomic studies of uropathogenic *Escherichia coli* and host urothelial cells when intracellular bacterial communities are assembled. *J Biol Chem.* 2007;282(29):21259-67.
13. Hannan TJ, Totsika M, Mansfield KJ, Moore KH, Schembri MA, and Hultgren SJ. Host-pathogen checkpoints and population bottlenecks in persistent and intracellular uropathogenic *Escherichia coli* bladder infection. *FEMS Microbiol Rev.* 2012;36(3):616-48.
14. Casadevall A. The Pathogenic Potential of a Microbe. *mSphere.* 2017;2(1).
15. Letica-Kriegel AS, Salmasian H, Vawdrey DK, Youngerman BE, Green RA, Furuya EY, et al. Identifying the risk factors for catheter-associated urinary tract infections: a large cross-sectional study of six hospitals. *BMJ Open.* 2019;9(2):e022137.

16. 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.

17. Trautner BW, and Darouiche RO. Role of biofilm in catheter-associated urinary tract infection. *Am J Infect Control.* 2004;32(3):177-83.

18. Goodman SD, Obergfell KP, Jurcisek JA, Novotny LA, Downey JS, Ayala EA, et al. Biofilms can be dispersed by focusing the immune system on a common family of bacterial nucleoid-associated proteins. *Mucosal Immunol.* 2011;4(6):625-37.

19. Soto SM. Importance of Biofilms in Urinary Tract Infections: New Therapeutic Approaches. *Advances in Biology.* 2014;2014:1-13.

20. Donlan RM, and Costerton JW. Biofilms: survival mechanisms of clinically relevant microorganisms. *Clin Microbiol Rev.* 2002;15(2):167-93.

21. Donlan RM. Biofilms: Microbial Life on Surfaces. *Emerg Infect Dis.* 2002;8(9):881–90.

22. Donlan RM. Biofilm formation: a clinically relevant microbiological process. *Clinical Infectious Diseases.* 2001;33(8):1387-92.

23. Johnson JR. Virulence factors in *Escherichia coli* urinary tract infection. *Clin Microbiol Rev.* 1991;4(1):80-128.

24. Swaine L. Chen CSH, Jerome S. Pinkner, Jennifer N. Walker, Corinne K. Cusumano, Zhaoli Li, Julie Bouckaert, Jeffrey I. Gordon, and Scott J. Hultgren. Positive selection identifies an in vivo role for FimH during urinary tract infection in addition to mannose binding. *PNAS.* 2009;106(52):22439-44.

25. Jonas Marschall MLP, Betsy Foxman, Lixin Zhang, David K Warren, Jeffrey P Henderson. Patient characteristics but not virulence factors discriminate between asymptomatic and symptomatic *E. coli* bacteriuria in the hospital. *BMC Infectious Diseases.* 2013;13.

26. 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.

27. 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.

28. Simmering JE, Tang F, Cavanaugh JE, Polgreen LA, and Polgreen PM. The Increase in Hospitalizations for Urinary Tract Infections and the Associated Costs in the United States, 1998-2011. *Open Forum Infect Dis.* 2017;4(1):ofw281.

29. 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.

30. Pablo Tsukayama MB, Sanket Patel, Erica C. Pehrsson, Molly K. Gibson, Kenneth L. Chiou, Clifford J. Jolly, Jeffrey Rogers, Jane E. Phillips-Conroy, Gautam Dantas. Characterization of Wild and Captive Baboon Gut Microbiota and Their Antibiotic Resistomes. *mSystems.* 2018;3(3):e00016-18.

31. 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.

32. 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.

33. 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.

34. 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. 2017;9(982):eaaf1283.

35. Kaper JB, Nataro JP, and Mobley HL. Pathogenic *Escherichia coli*. *Nat Rev Microbiol.* 2004;2(2):123-40.

36. 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.

37. 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.

38. 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).

39. 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.

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

41. 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.

42. 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.

43. Caza M, Lepine F, Milot S, and Dozois CM. Specific roles of the *iroBCDEN* genes in virulence of an avian pathogenic *Escherichia coli* O78 strain and in production of salmochelins. *Infect Immun.* 2008;76(8):3539-49.

44. 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.

45. 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.

46. B B Christensen CS, J B Andersen, R J Palmer Jr, A T Nielsen, M Givskov, S Molin. Molecular tools for study of biofilm physiology. *Methods Enzymol.* 1999;310:20-42.

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

48. 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.

49. Frazao N, Sousa A, Lassig M, and Gordo I. Horizontal gene transfer overrides mutation in *Escherichia coli* colonizing the mammalian gut. *Proc Natl Acad Sci U S A.* 2019;116(36):17906-15.

50. Hasegawa H, Suzuki E, and Maeda S. Horizontal Plasmid Transfer by Transformation in *Escherichia coli*: Environmental Factors and Possible Mechanisms. *Front Microbiol.* 2018;9:2365.

51. Henrik Hasman TC, and Per Klemm. Antigen-43-Mediated Autoaggregation of *Escherichia coli* Is Blocked by Fimbriation. *J Bacteriol.* 1999;181(16):4834–41.

52. 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.

53. 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.

54. Ehud Banin MLV, E Peter Greenberg. Iron and Pseudomonas aeruginosa biofilm formation. *Proc Natl Acad Sci U S A.* 2005;102(31):11076-81.

55. Celia H, Noinaj N, Zakharov SD, Bordignon E, Botos I, Santamaria M, et al. Structural insight into the role of the Ton complex in energy transduction. *Nature.* 2016;538(7623):60-5.

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

57. Sunden F, Hakansson L, Ljunggren E, and Wullt B. Bacterial interference--is deliberate colonization with *Escherichia coli* 83972 an alternative treatment for patients with recurrent urinary tract infection? *Int J Antimicrob Agents.* 2006;28 Suppl 1:S26-9.

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

59. 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.

60. 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.

61. Richard Hull DB, William Donovan, Catharina Svandorg, Inge Wieser, Colleen Steward, Rabih Darouiche. Urinary tract infection prophylaxis using *Escherichia coli* 83972 in spinal cord injured patients. *Journal of Urology.* 2000;163(3):832-77.

62. 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.

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. Abraham SN, and Miao Y. The nature of immune responses to urinary tract infections. *Nat Rev Immunol.* 2015;15(10):655-63.

66. Stickler DJ. Bacterial biofilms in patients with indwelling urinary catheters. *Nat Clin Pract Urol.* 2008;5(11):598-608.

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

68. 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.

69. 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.

70. Shields-Cutler RR, Crowley JR, Miller CD, Stapleton AE, Cui W, and Henderson JP. Human Metabolome-derived Cofactors Are Required for the Antibacterial Activity of Siderocalin in Urine. *J Biol Chem.* 2016;291(50):25901-10.

71. 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.

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

73. Schmieder R, and Edwards R. Fast identification and removal of sequence contamination from genomic and metagenomic datasets. *PLoS One.* 2011;6(3):e17288.

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

75. Löytynoja A. Phylogeny-aware alignment with PRANK. *Methods in Molecular Biology.* 2014;1079:155-70.

76. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MT, et al. Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics.* 2015;31(22):3691-3.

77. 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.

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. 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.

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

81. Kirill A. Datsenko BLW. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *PNAS.* 2000;97(12):6640-5.

82. 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.

FIGURES & 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.

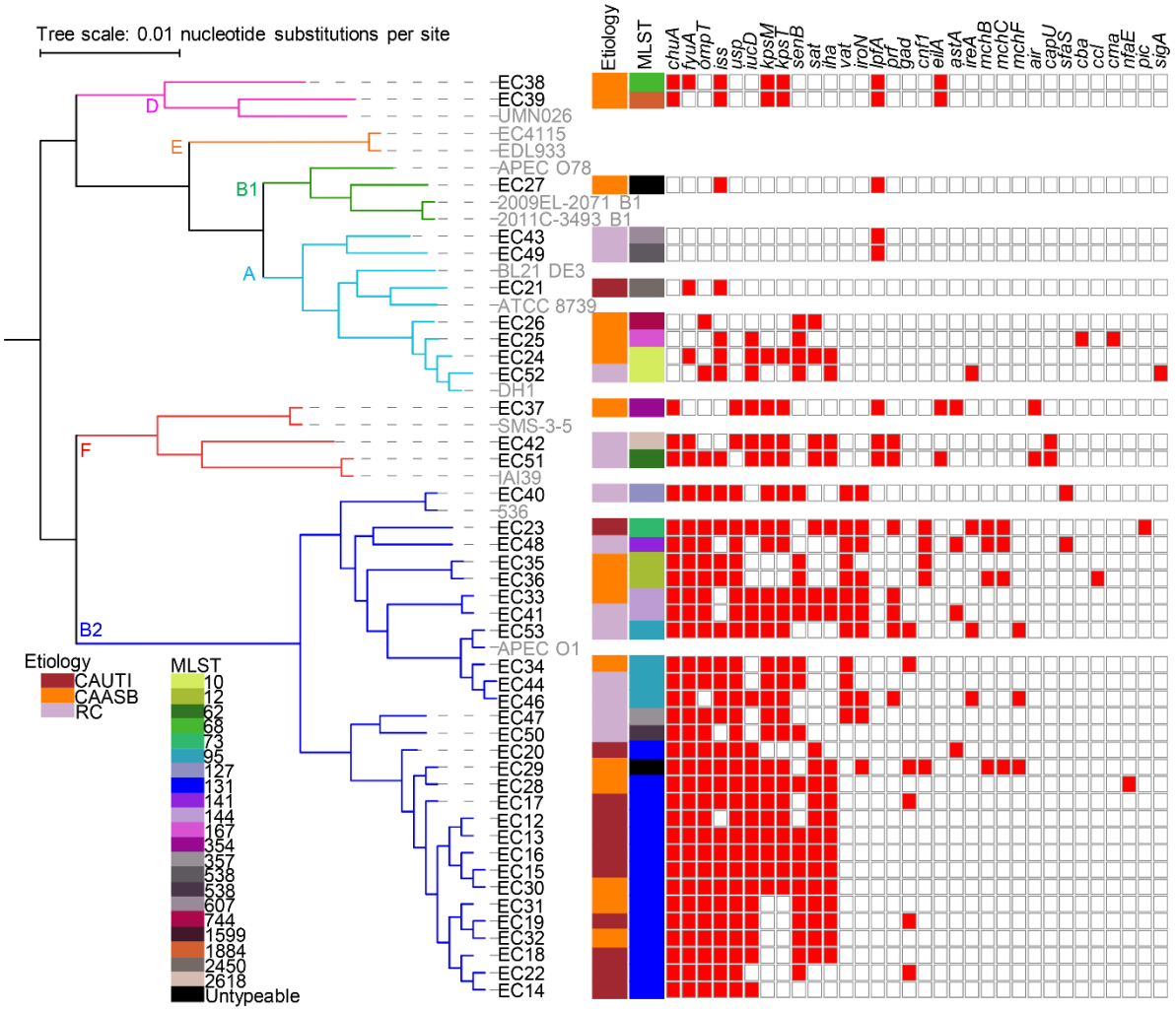

Figure 1. Phylogenetic distribution of 41 clinical *E. coli* isolates. Core-genome alignment was constructed into a maximum likelihood tree with raxML and viewed in iTOL, with six phylotypes identified. *In silico* multilocus sequence types (MLST) were identified using BLASTN to the *E. coli* MLST database, 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.

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	*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.

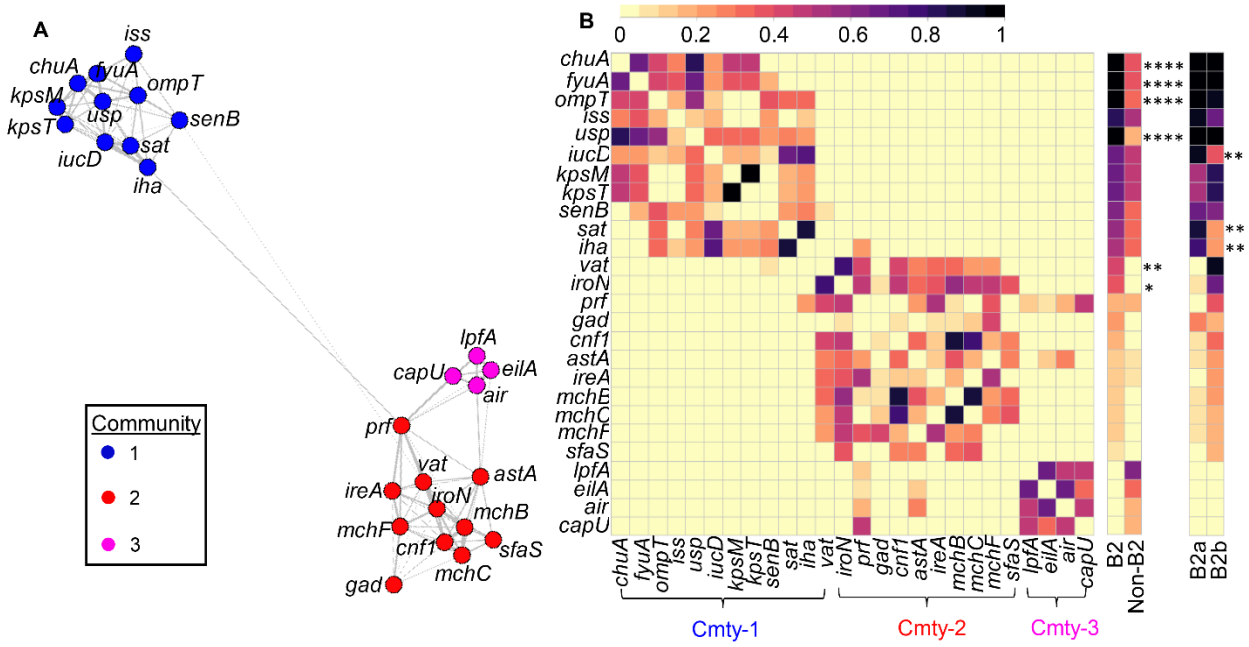

Figure 2. Network analysis of *E. coli* virulence factors (VFs). (A) A force-directed network layout illustrated co-associations and three VF communities among 26 VFs. Each node represented a VF. Each connecting line (edge) represented a positive association between 2 VFs that satisfied the significance 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. *: $P <= 0.05$. **: $P < 0.01$. ***: $P < 0.001$. ****: $P < 0.0001$.

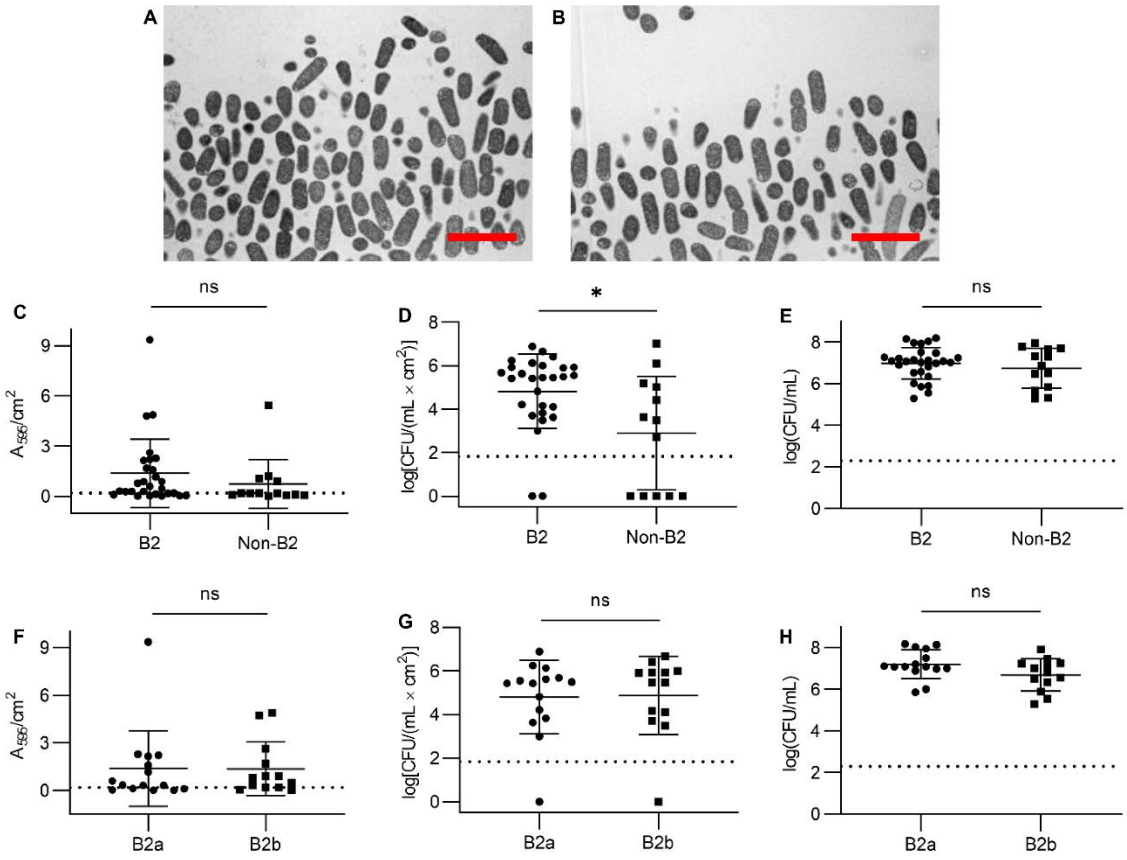

Figure 3. Biofilm formation by CAUTI, CAASB, and rectal isolates. (A) Transmission electron microscopy (TEM) image of catheter-biofilm grown in human urine. Scale bar: 4 µm. **(B)** Transmission electron microscopy (TEM) image of catheter-biofilm grown in artificial urine medium (AUM). Scale bar: 4 µm. **(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.

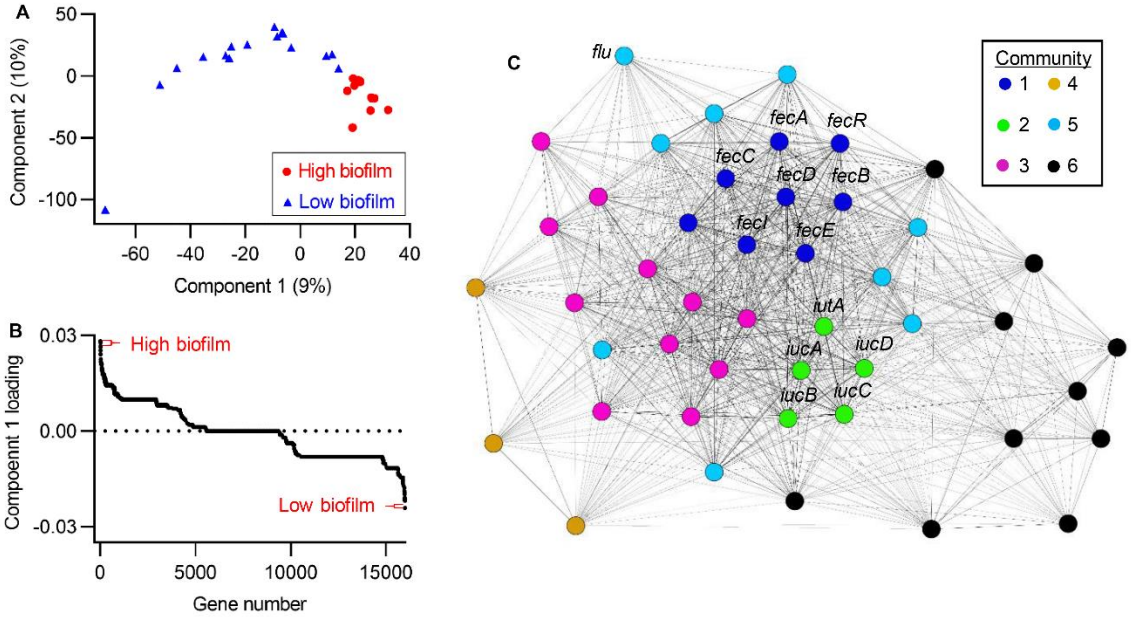

Figure 4. Identification of catheter biofilm-associated genes. (A) Score plot of the first two components from sparse partial least squares discriminant analysis (sPLSDA) for displaying group-wise clusterings between high and low biofilm formers. **(B)** Component 1-associated top loadings from sPLSDA 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.

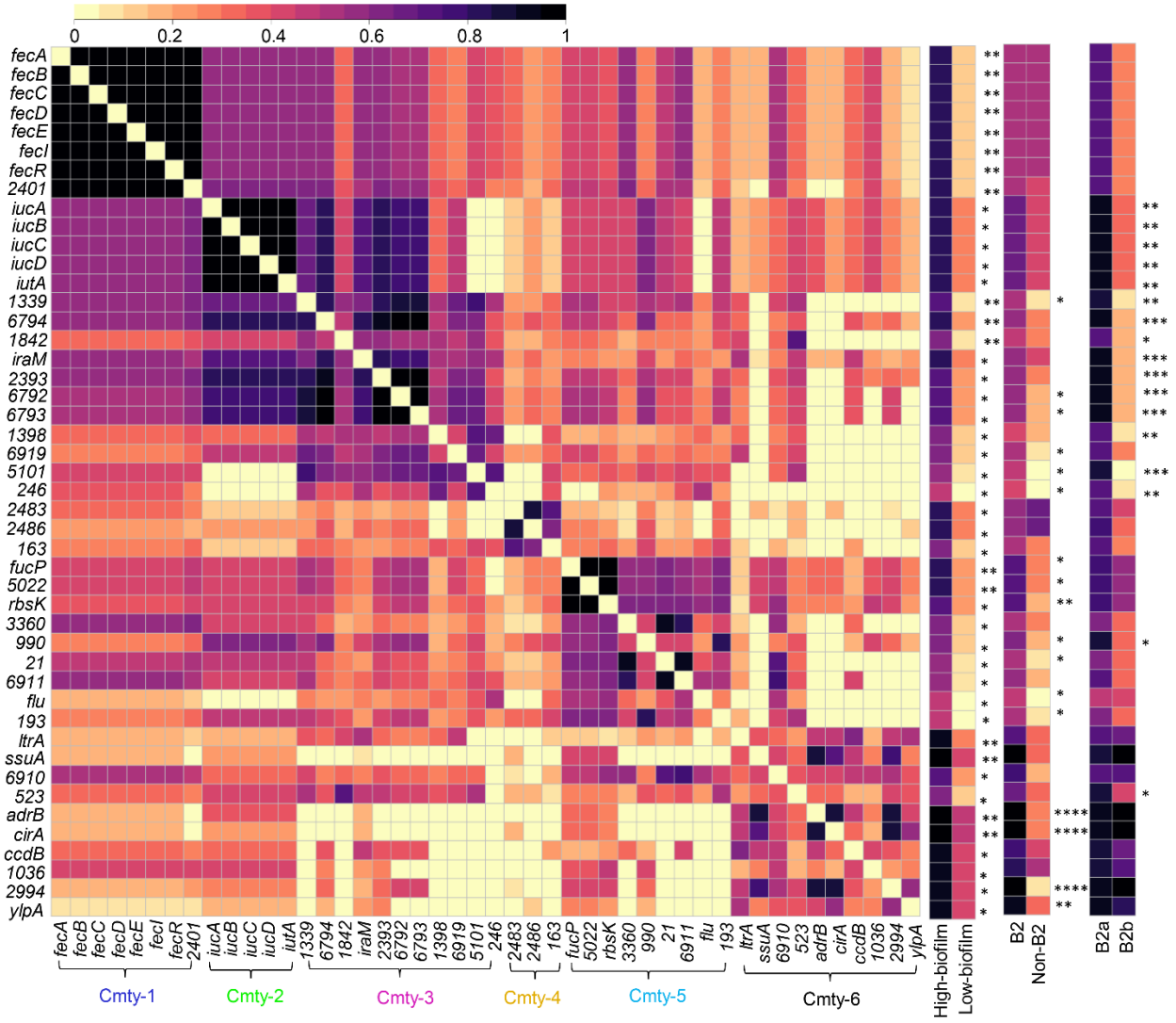

Figure 5. Correlation matrix of biofilm positively associated genes. Six gene communities are discernible in the correlation matrix heatmap depicting statistically significant (5% *P*-value threshold, one-tailed on the right, Fisher's exact test) positive associations between 46 biofilm positively associated 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. *: $P <= 0.05$. **: $P < 0.01$. ***: $P < 0.001$. ****: $P < 0.0001$.

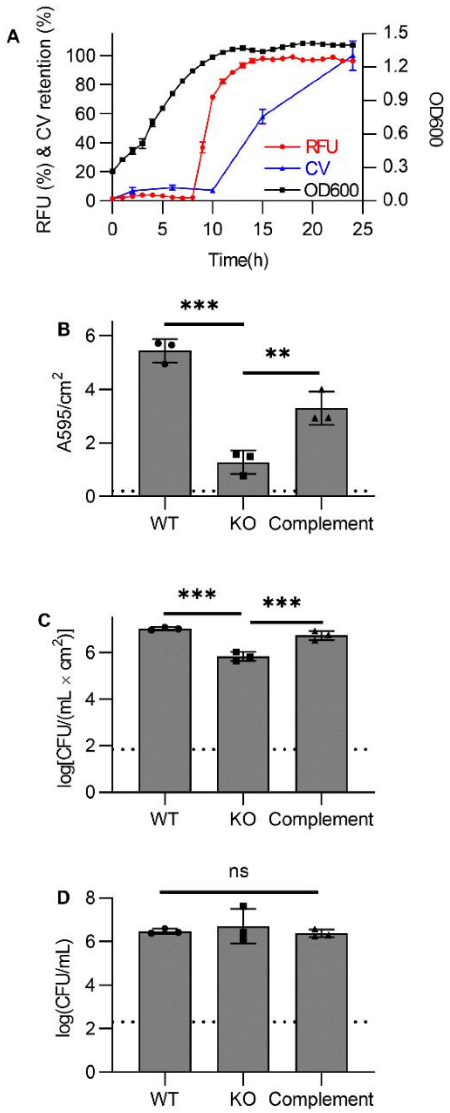

Figure 6. Fec expression is associated with increased catheter biofilm formation. **(A)** Comparison between *fec* expression (RFU), catheter-biofilm formation (CV retention, A595/cm^2), and bacterial growth (log(CFU/mL)) in the red fluorescence reporter strain, EC52::*fecI*-RFP, measured in a microplate assay. **(B)** Comparison of biofilm biomass (crystal violet retention,) between EC52 (WT), EC52Δ*fecA* (KO), and EC52Δ*fecA*::*fecA* (Complement), measured in the continuous flow model. P < 0.006, by one-way ANOVA with Dunnett's multiple comparisons test. **(C)** Comparison of catheter-adherent CFUs between EC52 (WT), EC52Δ*fecA* (KO), and EC52Δ*fecA*::*fecA* (Complement), measured in the continuous flow model. P < 0.002, by one-way ANOVA with Dunnett's multiple comparisons test. **(D)** Comparison of planktonic CFUs in the voided media between EC52 (WT), EC52Δ*fecA* (KO), and EC52Δ*fecA*::*fecA* (Complement), measured in the continuous flow model. *P* = 0.7, by one-way ANOVA multiple comparisons test. Three replicates with mean and SD plotted. *P* <= 0.05 is considered statistically significant. ns: not significant. *: *P* <= 0.05. **: *P* < 0.01. ***: *P* < 0.001.

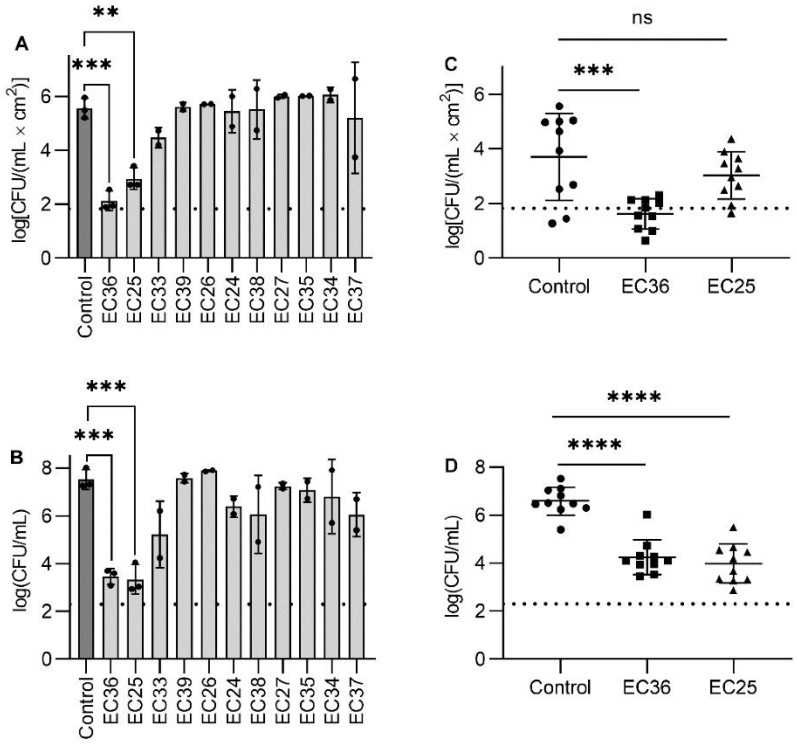

Figure 7. CAASB *E. coli* catheter-biofilms inhibited CAUTI colonization. (A) Catheter-adherent CFUs (log[CFU/(mL.cm^2)]) of strain EC20 in solely biofilm formation test (Control) and in bacterial competition tests with 11 non-ST131 CAASB strains (EC36, EC25, EC33, EC39, EC26, EC24, EC38, EC27, EC35, EC34, EC37). Three replicates with mean and SD plotted for Control, EC36, and EC25, with $P < 0.004$, Two replicates with mean and SD plotted for EC33, EC39, EC26, EC24, EC38, EC27, EC35, EC34 and EC37, with $P > 0.5$. **(B)** Planktonic CFUs (log(CFU/mL)) in the voided media of strain EC20 in solely biofilm formation test (Control) and in bacterial competition tests with 11 non-ST131 CAASB strains (EC36, EC25, EC33, EC39, EC26, EC24, EC38, EC27, EC35, EC34, EC37). Three replicates with mean and SD plotted for Control, EC36, and EC25, with $P < 0.0003$. Two replicates with mean and SD plotted for EC33, EC39, EC26, EC24, EC38, EC27, EC35, EC34 and EC37, with $P > 0.05$. **(C)** Catheter-adherent CFUs (log[CFU/(mL.cm^2)]) of 10 ST131 CAUTI strains in solely biofilm formation tests (Control) and in bacterial competition tests with two non-ST131 CAASB strains (EC36 and EC25). Test results of 10 ST131 CAUTI strains with mean and SD plotted, with P = 0.0005 and P = 0.3 for EC36 and EC35, respectively. **(D)** Planktonic CFUs (log(CFU/mL)) in the voided media of 10 ST131 CAUTI strains in solely biofilm formation tests (Control) and in bacterial competition tests with two non-ST131 CAASB strains (EC36 and EC25). Test results of 10 ST131 CAUTI strains with mean and SD plotted, with P < 0.0001. By one-way ANOVA with Dunnett's multiple comparisons test. $P <= 0.05$ is considered statistically significant. ns: not significant. *: $P <= 0.05$. **: $P < 0.01$. ***: $P < 0.001$. ****: $P < 0.0001$.