INFECTIOUS DISEASE

Bacterial virulence phenotypes of *Escherichia coli* and host susceptibility determine risk for urinary tract infections

Henry L. Schreiber IV,[1,2] Matt S. Conover,[1] Wen-Chi Chou,[2] Michael E. Hibbing,[1] Abigail L. Manson,[2] Karen W. Dodson,[1] Thomas J. Hannan,[3] Pacita L. Roberts,[4] Ann E. Stapleton,[4] Thomas M. Hooton,[5] Jonathan Livny,[2]* Ashlee M. Earl,[2]* Scott J. Hultgren[1,6]*

Urinary tract infections (UTIs) are caused by uropathogenic *Escherichia coli* (UPEC) strains. In contrast to many enteric *E. coli* pathogroups, no genetic signature has been identified for UPEC strains. We conducted a high-resolution comparative genomic study using *E. coli* isolates collected from the urine of women suffering from frequent recurrent UTIs. These isolates were genetically diverse and varied in their urovirulence, that is, their ability to infect the bladder in a mouse model of cystitis. We found no set of genes, including previously defined putative urovirulence factors (PUFs), that were predictive of urovirulence. In addition, in some patients, the *E. coli* strain causing a recurrent UTI had fewer PUFs than the supplanted strain. In competitive experimental infections in mice, the supplanting strain was more efficient at colonizing the mouse bladder than the supplanted strain. Despite the lack of a clear genomic signature for urovirulence, comparative transcriptomic and phenotypic analyses revealed that the expression of key conserved functions during culture, such as motility and metabolism, could be used to predict subsequent colonization of the mouse bladder. Together, our findings suggest that UTI risk and outcome may be determined by complex interactions between host susceptibility and the urovirulence potential of diverse bacterial strains.

INTRODUCTION

Urinary tract infections (UTIs) are a health and financial burden, particularly affecting sexually active, premenopausal women (*1*). About 50% of women will experience a UTI in their lifetime, with 10.5 million cases of UTIs occurring in 2007 in the United States alone (*2*). In placebo-controlled studies of UTI outcome, most women maintain bladder infections for weeks after an acute cystitis episode if not treated with antibiotics, whereas other women clear the infection without antibiotic treatment (*3*, *4*). These two outcomes of infection, either self-limiting or chronic cystitis, have been modeled in mice (*5–7*). Effective treatment of UTIs is becoming more challenging because antibiotic resistance rates are increasing (*8*). In addition, although antibiotic therapy effectively resolves most UTIs, about 20 to 30% of women will experience a recurrent UTI (rUTI) within 6 months of initial infection (*9*, *10*). This highlights the need to better understand the pathogenic mechanisms that facilitate acute bacterial colonization of the urinary tract and determine UTI outcome.

The most common cause of community-acquired UTIs is uropathogenic *Escherichia coli* (UPEC) (*2*). *E. coli* strains are diverse and have been characterized into pathotypes based on their ability to cause particular diseases (*11*, *12*). Pathotypes include intestinal pathogens such as Shiga toxin–producing *E. coli* and extraintestinal pathogenic *E. coli* strains such as UPEC. *E. coli* are also subdivided into phylogenetic clades (for example, A, B1, B2, D, and E) based on their genetic similarity (*13*, *14*). In the United States and Europe, most UPEC strains are from the B2 clade, although members of clades A, B1, and D cause 25 to

50% of cystitis (*15–19*). In East Asia, clade D strains predominate in community-acquired UTI followed by B2 strains (*20*, *21*). All commonly studied UPEC isolates, including UTI89, CFT073, NU14, and 536 (*22–25*), are B2 members, whereas non-B2 strains have not been extensively characterized. Thus, it is unknown whether non-B2 UPEC strains use the same virulence mechanisms as B2 UPEC strains.

Study of prototypical B2 UPEC isolates, such as the model UPEC strain UTI89 originally isolated from a woman with cystitis (*26*), has shown that the chaperone usher pathway (CUP) type 1 pilus is critical for bladder colonization in diverse mouse models (*27*). The *fim* operon, which encodes the proteins for assembly of type 1 pili, is part of the core *E. coli* genome, found in nearly every *E. coli* strain (*28*). Further, in research of the model strain UTI89, it has been shown that type 1 pili expression is tightly coregulated with numerous factors including flagella and S and P CUP pili (*29*). Type 1 pili are tipped by the adhesin FimH, which binds specifically to mannosylated uroplakins lining the superficial bladder epithelium (*30*) and mannosylated proteins expressed on underlying layers of epithelial cells (*31*), thus mediating critical steps in bladder colonization and UTI progression. Subsequent to FimH attachment, *E. coli* can invade the superficial umbrella bladder cells (*32*) and escape into the cytoplasm where they rapidly replicate to form clonal, biofilm-like intracellular bacterial communities (IBCs) of $\sim 10^4$ colony-forming units (CFU) while protected from the host immune response (*33*, *34*). Upon IBC maturation, UTI89 bacteria transform into long filaments and burst out of the superficial umbrella cells of the bladder into the bladder lumen where they invade other epithelial cells and form new IBCs (*34*). Most clinical UPEC isolates that have been tested are able to form IBCs in mice, and each of six different inbred mouse strains tested support IBC formation (*35*). After the discovery of IBCs in mouse UTI models, IBCs and filamentous bacteria were documented in human urine collected from clinical UTI cases (*36*, *37*).

Despite decades of research using model UPEC strains, universal bacterial features that enable *E. coli* uropathogenesis remain

[1]Department of Molecular Microbiology, Washington University, St. Louis, MO 63110, USA. [2]Broad Institute of Massachusetts Institute of Technology and Harvard, Cambridge, MA 02142, USA. [3]Department of Pathology and Immunology, Washington University, St. Louis, MO 63110, USA. [4]Division of Allergy and Infectious Diseases, Department of Medicine, University of Washington, Seattle, WA 98195, USA. [5]Division of Infectious Diseases, Department of Medicine, University of Miami Miller School of Medicine, Miami, FL 33136, USA. [6]Center for Women's Infectious Disease Research, Washington University, St. Louis, MO 63110, USA.
*Corresponding author. Email: livny@broadinstitute.org (J.L.); aearl@broadinstitute.org (A.M.E.); hultgren@wusm.wustl.edu (S.J.H.)

largely undefined. Carriage of putative urovirulence factors (PUFs) is thought to enhance *E. coli* uropathogenicity and is used to measure and categorize clinical UPEC strains isolated from different patient populations (16, 20, 38, 39). PUFs were identified through their enrichment in UTI isolates when compared to non–UTI-associated *E. coli* (38–41). Although numerous sets of PUFs have been used to assess urovirulence (16, 20, 38–41) and strains with more PUFs are considered uropathogenic than strains with fewer PUFs, a direct role in pathogenicity remains to be defined for most PUFs.

Here, we examined a set of 43 human urine-associated *E. coli* (UAEC) strains isolated at symptomatic and asymptomatic time points from the urine of 14 women with frequent rUTIs. We tested whether UAEC isolates shared common genes that classified strains with increased virulence from those with lesser virulence in defined experimental mouse models. Surprisingly, we observed no correlation between gene content, including PUFs, and the ability to cause cystitis in C3H/HeN mouse models. However, we found that environmentally responsive phenotypes and specific transcriptional responses to in vitro conditions could discriminate whether strains would be more or less efficient at bladder colonization when tested in the C3H/HeN acute cystitis mouse model. We also determined that some UAEC were robust colonizers in multiple mouse models, whereas others were only efficient at colonizing the bladder in some models.

RESULTS

Human UAEC strains are diverse and caused both same-strain and different-strain rUTI

A total of 43 *E. coli* isolates were collected from 14 women with frequent rUTI (fig. S1 and tables S1 and S2). Because all of these strains were isolated from the urine of women but not necessarily collected during diagnosed UTIs, we refer to these strains as UAEC rather than UPEC. Whole-genome sequences of all 43 UAEC were determined, assembled, annotated, and combined with genomic data from 46 previously sequenced representative *E. coli* from various clades and pathotypes to evaluate aspects of their relatedness and gene content (table S3). We found that 60 to 75% of each strain's genome consisted of core genes, defined as genes present in all strains being analyzed, whereas the variable genome comprised the remaining 25 to 40% (fig. S2A). We constructed a maximum-likelihood phylogenetic tree using the 2746 single-copy genes present in the core genome of all the *E. coli* isolates (Fig. 1). Consistent with previous studies of UPEC (15–19), more than two-thirds (67.4%, 29 of 43) of our UAEC strains fell within the B2 clade. The remaining UAEC fell within three other *E. coli* clades (four in clade A, eight in clade B1, and two in clade D).

In some patients, UAEC isolated at consecutive time points were tightly grouped within the phylogenetic tree, indicating same-strain rUTI. In contrast, UAEC collected at consecutive time points in other patients were not closely related, which is consistent with different-strain rUTI. As has been carried out in previous studies (42), cluster analysis of pairwise single-nucleotide polymorphism (SNP) distances between all UAEC isolates revealed 13 different clonal groups comprised of two to four isolates each (fig. S2B and table S2). Eight UAEC did not group with any other isolate, and each was assigned to its own core clonal group. On the basis of clonal group membership, we identified 12 same-strain recurrences and 6 different-strain recurrences from a total of 18 rUTI events (table S2). Eleven of 12 same-strain rUTIs were caused by B2 strains, which was a statistically significant enrichment relative to the phylogenetic distribution of the UAEC tested in our
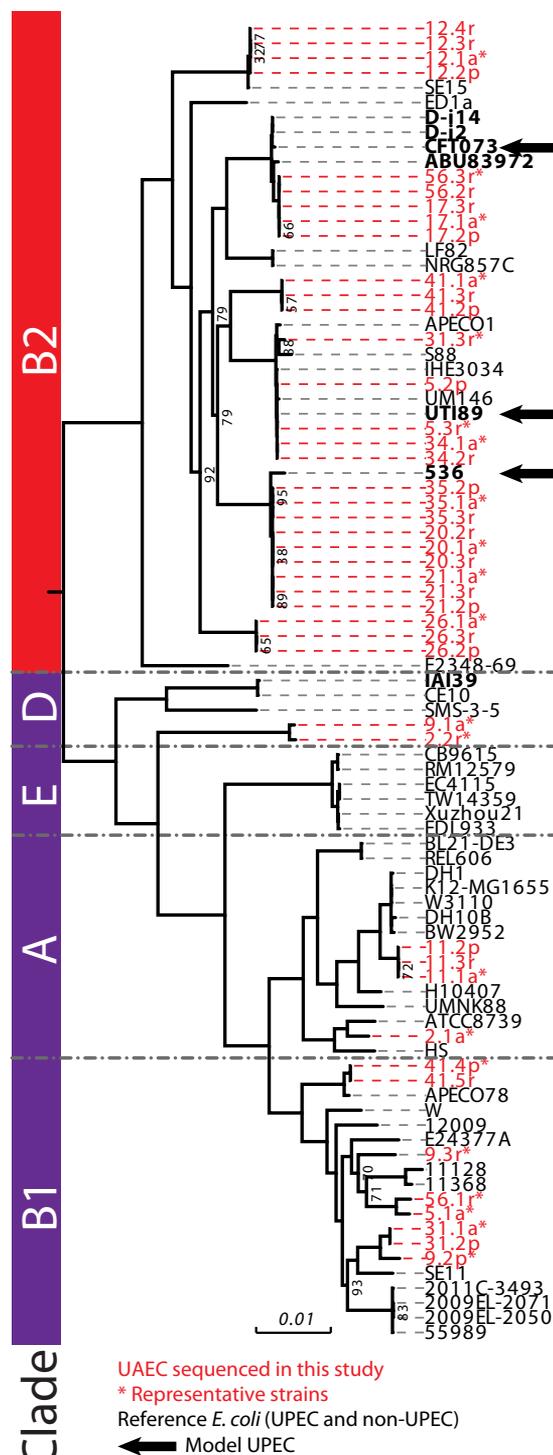

Fig. 1. Phylogenetic distribution of UAEC strains from rUTI patients. The phylogenetic relatedness of the UAEC strains (*n* = 43; taxon labels in red) was contextualized within the broader phylogeny of reference *E. coli* strains (*n* = 46; taxon labels in black) by comparing the single-copy core genes of the strains using the RAxML algorithm. Reference *E. coli* strains that were associated with urinary disease (for example, cystitis, pyelonephritis, or asymptomatic bacteriuria) are set in boldface. Asterisks indicate UAEC strains chosen as representative isolates for their clonal clusters (fig. S2B). Bootstrap supports are indicated at internal nodes, and bootstrap values >95 have been removed. UAEC strains were found in four out of five *E. coli* clades (indicated by red and purple bars on the left). Black arrows indicate model UPEC strains commonly used in UTI research.

Schreiber *et al.*, *Sci. Transl. Med.* **9**, eaaf1283 (2017)   22 March 2017

2 of 13

cohort ($P < 0.05$, hypergeometric test). For further analyses, we then selected a representative UAEC strain from each of the clonal groups, including core clonal groups with a single member, resulting in a representative collection of 21 UAEC isolates that encompassed the genetic, phylogenetic, and phenotypic diversity of our strain set, including 11 clade B2, 2 clade A, 6 clade B1, and 2 clade D strains (table S2).

Carriage of PUFs correlates with B2 clade membership in human urine and nonurine isolates of *E. coli*

As carried out previously (38–40), we calculated a PUF score for each representative UAEC and non-UAEC control strain using an alignment-based analysis of 31 previously defined PUFs (table S4) (20, 38–40). Carriage of PUFs varied considerably among strains, ranging from 2 to 24 PUFs per UAEC genome (median, 13; the model B2 UPEC strain, UTI89, encodes 17 PUFs) to 2 to 19 PUFs per non-UAEC genome (median, 5) (Fig. 2). Of the 31 PUFs, 28 were found in at least one UAEC genome. We found that B2 *E. coli* strains carried more PUFs than did non-B2 *E. coli* strains whether they were associated with urinary disease in humans ($P < 0.001$, Mann-Whitney *U* test) (Fig. 2A). Further, using unsupervised hierarchical clustering of PUF carriage, we found that strains generally clustered on the basis of B2 clade membership, rather than their association with urinary disease (Fig. 2B), indicating that PUF carriage was biased by phylogeny and was not specific to uropathogens.

Bladder colonization in C3H/HeN mice by UAEC is variable and not restricted to B2 strains

Although each of our UAEC strains was recovered from the urine of a patient and thus able to colonize a human bladder, differences in patient genetics, behavior, and infection history can influence susceptibility to UTI (9, 43–45). To control for host genetics and environmental factors that could confound comparisons of colonization efficiency among the UAEC strains, we characterized the ability of each of 21 representative strains to colonize the genetically homogenous C3H/HeN mouse model of cystitis. Each UAEC strain and a model UPEC strain, UTI89, were inoculated into the bladder of at least five C3H/HeN mice. Bacterial burdens, measured in CFU, were determined in both the bladder and kidneys at 24 hours postinfection (hpi) (Fig. 3A and fig. S3). We observed a range of bladder bacterial burdens (200 to 10^8 CFU per bladder) at 24 hpi with three colonizer groups emerging among the tested UAEC strains. Twelve of 21 strains were "robust colonizers" where UAEC titers were greater than 10^4 CFU in every infected mouse bladder, 4 of 21 strains were "deficient colonizers" where UAEC titers were less than 10^4 in every infected bladder, and 5 of 21 strains were "variable colonizers" where UAEC titers ranged both above and below 10^4 CFU per bladder of infected mice (Fig. 3A). Kidney colonization generally correlated with bladder colonization (fig. S3). Notably, both B2 and non-B2 strains were among the robust colonizers, capable of colonizing the mouse bladder at rates similar to UTI89, indicating that efficient bladder colonization of C3H/HeN mice was achievable by strains outside of the B2 clade. Whereas C3H/HeN bladder colonization varied among UAEC at 24 hpi, the titers of these isolates were markedly higher than those of gut-associated *E. coli* strains collected from adult humans who did not have urinary tract disease. These strains included model *E. coli* strains and strains from the feces of healthy adults (46, 47) and were comparable to our UAEC collection in the distribution of B2 versus non-B2 strains (table S5). We found that only one of nine gut-associated
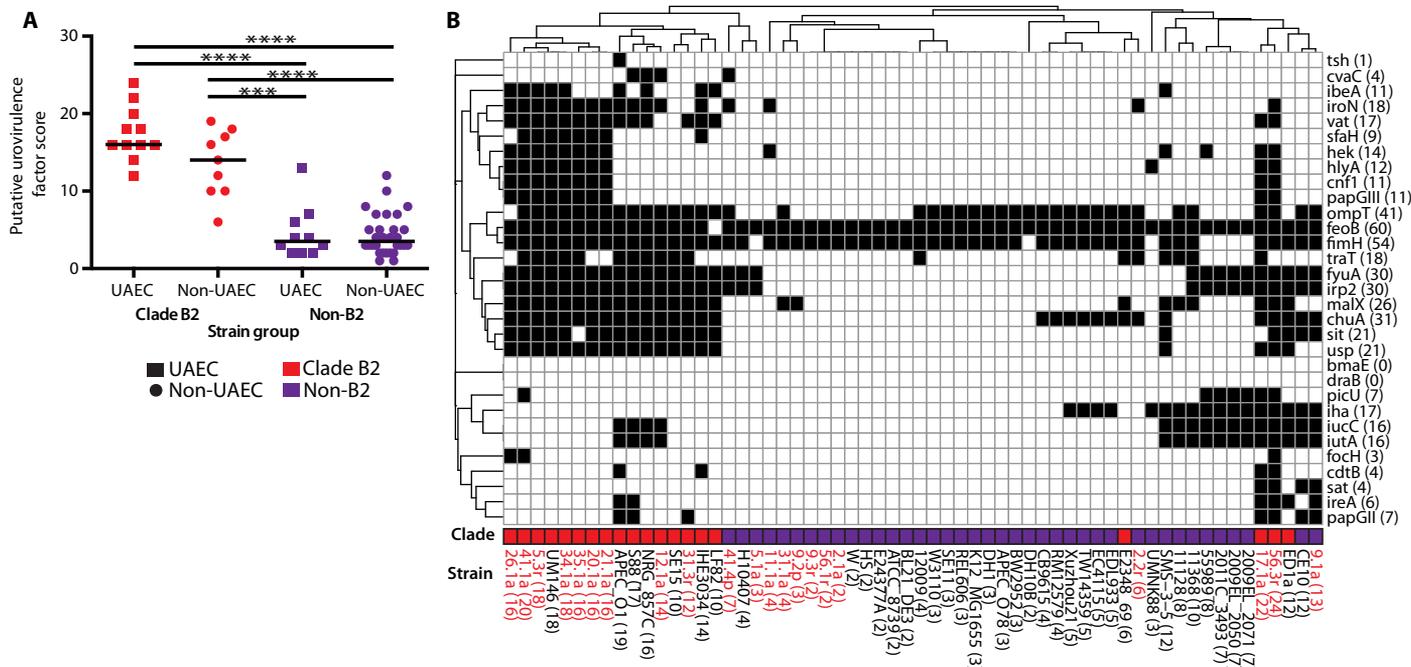

Fig. 2. Carriage of PUFs is enriched in both UAEC and non-UAEC strains from the B2 clade. (**A**) Comparisons between indicated groups were performed using the Mann-Whitney *U* test. Significant differences were calculated for each group, and significant results are indicated as follows: ***$P < 0.001$; ****$P < 0.0001$. (**B**) Clinical UAEC strains and reference *E. coli* strains that were not associated with urinary disease (non-UAEC) (x axis) were examined for the presence of 31 PUF sequences (y axis) using the Basic Local Alignment Search Tool (BLAST) and alignment-based searches. The binary presence of PUFs (indicated by black squares) was tallied for each strain and for each PUF (indicated in parentheses on both axes). Two-dimensional hierarchical clustering identified clusters of PUFs that tended to co-occur in UAEC strains (dendrogram along the y axis) and showed that PUF carriage was associated with phylogeny (dendrogram along the x axis; phylogeny indicated in the column labeled Clade).
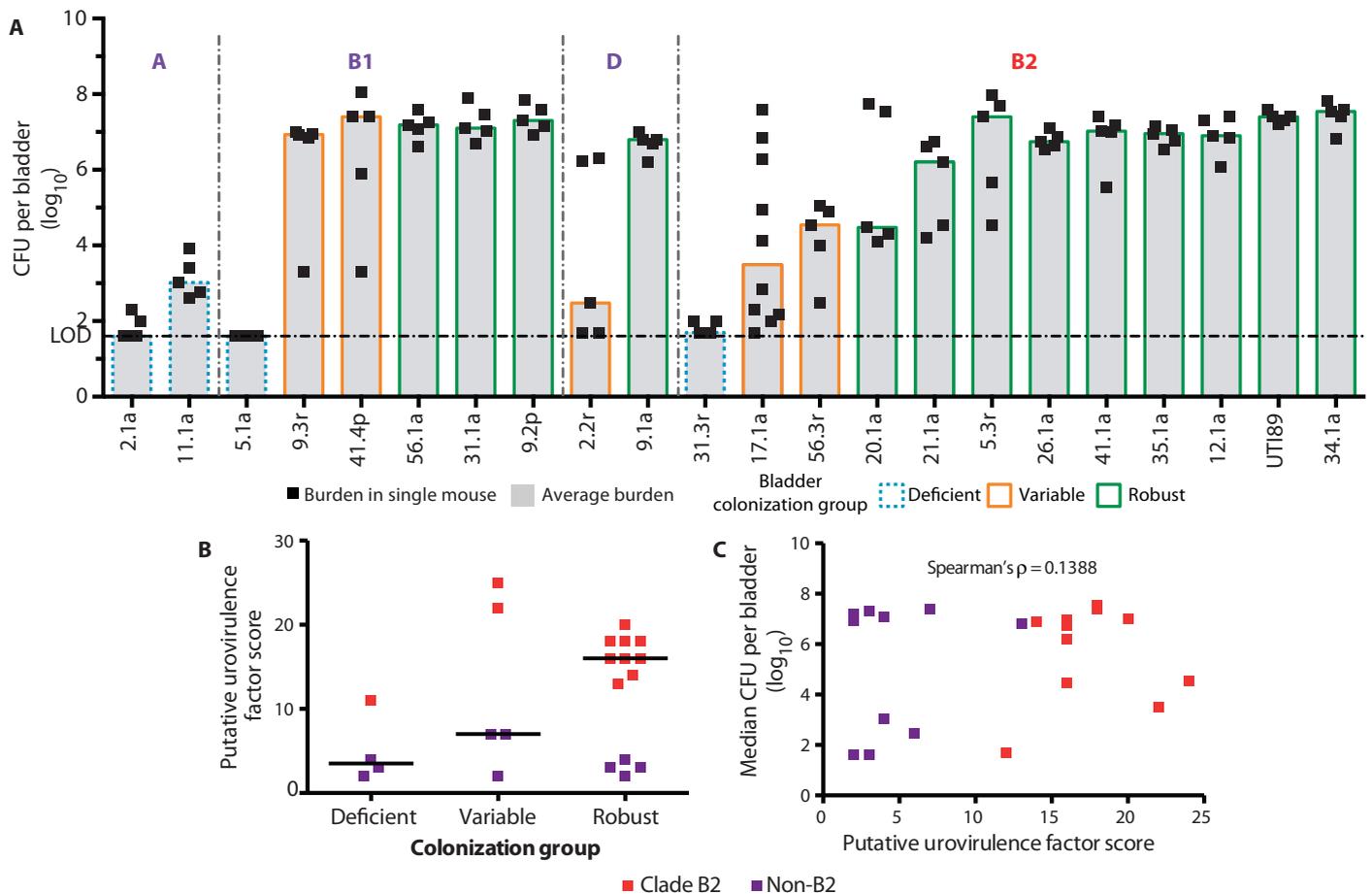
Schreiber et al., *Sci. Transl. Med.* **9**, eaaf1283 (2017)   22 March 2017

3 of 13

Fig. 3. Increased PUF carriage does not correlate with increased colonization efficiency. (**A**) The colonization efficiencies of 21 representative UAEC strains and the model strain UTI89 from indicated clades were tested in C3H/HeN mice. Bacteria were enumerated from individual harvested bladders at 24 hpi (black boxes). Each UAEC strain was categorized as "deficient" ($n = 4$; blue stippled outline) at $<10^4$ CFU per bladder, "variable" ($n = 5$; orange outline) above and below 10^4 CFU per bladder, or "robust" ($n = 12$; green outline) at $>10^4$ CFU per bladder. The black broken horizontal line represents the limit of detection (LOD) of bacteria. Data presented represent the median (gray bar) of mouse inoculations for each strain. (**B**) No enrichment of PUF carriage, as measured by PUF scores, was found in comparisons of robust, variable, or deficient colonizer strains. Comparisons were performed with Mann-Whitney U test; horizontal solid bars indicate median values. (**C**) Using Spearman's rank correlation (ρ statistic indicated at the top), there was no significant correlation between carriage of PUFs and bladder burden with UAEC strains at 24 hpi. Squares indicate the median bladder colonization of B2 (red) or non-B2 (purple) UAEC strains.

E. coli, a clade A strain, was capable of robustly colonizing mouse bladders at 24 hpi, whereas five of nine strains were deficient and the remaining four strains were variable (fig. S4A). Thus, a significantly higher proportion of UAEC strains robustly colonized the C3H/HeN mouse model than did gut-associated *E. coli* at 24 hpi ($P < 0.05$, Fisher's exact test) (table S5 and fig. S4).

Given that early bladder colonization is enhanced by the formation of IBCs characterized using model B2 UPEC strains (*33, 34, 48, 49*), we sought to determine whether diverse B2 and non-B2 UAEC strains from our collection followed the same pathogenic cascade as model B2 strains. As for the UPEC strain UTI89 (*49*), we found that five of six phylogenetically diverse UAEC were competent for IBC formation (fig. S5 and table S6) in C3H/HeN mice at 6 hpi, suggesting that both B2 and non-B2 UAEC strains exhibited pathogenic mechanisms similar to that of the model B2 UPEC strain. Notably, the only deficient colonizer strain tested among the group clade A strain 11.1a was also the strain that did not form IBCs at this single time point, supporting previous reports that IBC formation is linked to bladder colonization (*49*).

Carriage of PUFs is not required for bladder colonization in C3H/HeN mice but may provide a competitive advantage in chronic cystitis

Although UAEC strains with both high and low PUF scores caused rUTI among the women in our cohort, we sought to determine whether PUF content could explain the observed differences in bladder colonization in the C3H/HeN mouse model (Fig. 3A). We observed no significant differences in the carriage of PUFs when comparing UAEC strains that were robust (range, 2 to 20; median, 16), variable (range, 2 to 24; median, 7), or deficient (range, 4 to 12; median, 3.5) colonizers of the C3H/HeN mice ($P > 0.05$, Kruskal-Wallis test and Mann-Whitney U tests) (Fig. 3B). Further, we found no significant correlation between PUF carriage and bladder bacterial burden in C3H/HeN mice ($P > 0.05$, Spearman's rank correlation) (Fig. 3C).

In mice, chronic cystitis is defined as the development of persistent high-titer ($>10^4$ CFU/ml) bacteriuria, high-titer bladder bacterial burdens, and chronic inflammation with persistent lymphoid aggregates at sacrifice >4 weeks postinfection. Chronic cystitis and bladder

inflammation result in large-scale remodeling of the bladder mucosa, which can last for weeks or months after antibiotic treatment to clear the infection (6, 7). This remodeling is evident in changes to the bladder itself and in host susceptibility to recurrent infection, exemplified by (i) an altered transcriptional profile, (ii) an altered urothelial membrane proteome, (iii) defects in superficial cell maturation, (iv) changes to bacterial occupation of different habitats, and (v) differential host responses upon subsequent bacterial exposure (6, 7). To determine the effects of PUF carriage on the ability of UAEC to cause chronic cystitis, we examined three non-B2 strains [41.4p (clade B1, PUF score = 7), 9.1a (clade D, PUF score = 13), and 9.2p (clade B1, PUF score = 3)] and three B2 UAEC strains [20.1a (PUF score = 16), 35.1a (PUF score = 16), and 41.1a (PUF score = 20)] for their ability to cause persistent, high-titer bacteriuria ($>10^4$ CFU/ml of urine), high-titer bladder colonization ($>10^4$ CFU per bladder), and chronic inflammation at 28 days postinfection (dpi) in juvenile C3H/HeN mice. We found that all of the tested UAEC strains were capable of causing chronic cystitis at varying frequencies, including the non-B2 UAEC strains (range, 20 to 90%) (Fig. 4). These rates were comparable to the rates of chronic cystitis caused by B2 UAEC strains and UTI89 (range, 10 to 60%) (5). Together, our results indicate that several non-B2 strains carrying few PUFs were capable of causing both acute and chronic UTI in C3H/HeN mice in addition to causing clinical infection in women. Whereas the carriage of many PUFs was not essential for non-B2 UAEC strains to cause chronic cystitis in mice, B2 UAEC strains that carried many PUF genes were enriched in isolates causing same-strain rUTI (11 of 12 same-strain UTIs were caused by B2 UAEC). We found that specific deletion from UTI89 of a large pathogenicity-associated island, PAI II$_{UTI89}$, carrying 124 genes including four PUFs (papgIII, hek, cnf1, and hlyA) did not result in a competitive defect against the wild-type strain during acute time points (1 to 3 dpi) in urine. However, the mutant strain was outcompeted in the urine during chronic time points (≥7 dpi) (fig. S6A) and in the bladder and kidney tissue at sacrifice 28 dpi (fig. S6B).

Later time point rUTI isolates outcompete those collected at enrollment regardless of PUF carriage

Previous research has shown that different-strain rUTIs can occur when the recurrent strain outcompetes the initial UTI strain in the gastrointestinal tract and bladder (50). In two patients in our cohort,
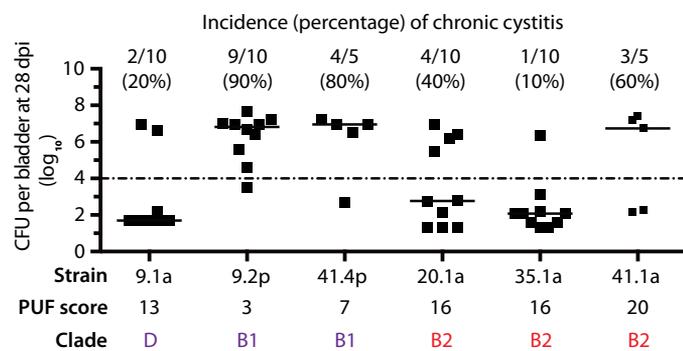
patients 9 and 41, a UAEC strain with a higher PUF score (strain 9.1a, PUF score = 13; strain 41.1a, PUF score = 20) was supplanted by a UAEC strain with a lower PUF score (strain 9.2p, PUF score = 3; strain 41.4p, PUF score = 7). We hypothesized that the supplanting strains would be fitter in colonizing the murine bladder than the initial UTI strains collected from the same patient. In two separate experiments, we inoculated equal numbers of the initial and supplanting strains from patients 9 and 41 into the bladders of C3H/HeN mice, respectively, and the relative bacterial titers were determined from urine at 24 hpi using strain-specific antibiotic resistance markers (Fig. 5). In both experiments, we found that the supplanting strains outcompeted the initial UTI strains at 24 hpi despite the fact that strains 41.4p and 9.2p carried considerably fewer PUF genes. Together, these data suggest that bladder colonization was not limited to B2 UAEC strains or strains with high carriage of PUFs.

Variations in UAEC gene expression under defined conditions predict the outcome of acute cystitis in mice

It is possible that UAEC strains share some genomic feature that is not currently identified as a PUF. Thus, we also performed an in-depth, unbiased comparative genomic analysis between robust and deficient colonizer strains of UAEC to find any possible cryptic PUF genes or genomic signature that would clearly delineate these two groups. Our analysis revealed that no orthologous genes were exclusive to either robust or deficient colonizers and that there was no significant enrichment of specific functional domains or enzymatic pathways in either colonizer group after correcting for variance between strains. Using a relaxed definition of enrichment (that is, present in >80% of one group and <20% of the other group), we found that two orthogroups were enriched in the deficient colonizer UAEC relative to robust colonizer strains, namely, the brnTA genes encoding a type 2 toxin-antitoxin (table S7) (51). The brnTA genes, whose function in pathogenesis
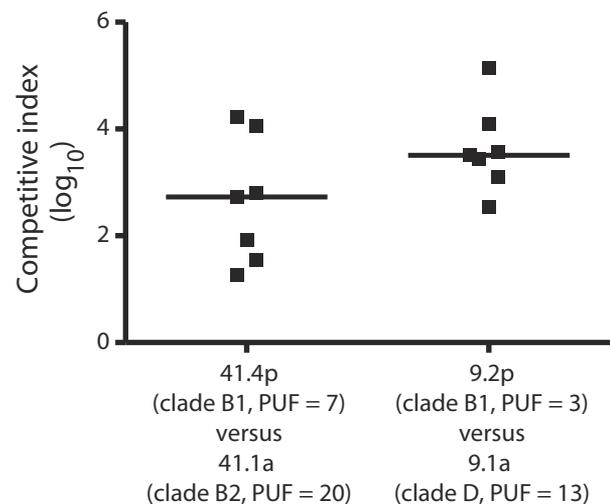

Fig. 4. **Both B2 and non-B2 UAEC strains cause chronic cystitis in mice.** Incidence of chronic cystitis, defined as persistent high-titer bacteriuria ($>10^4$ CFU/ml of urine), high-titer bladder colonization ($>10^4$ CFU per bladder), and chronic inflammation at 28 dpi, was measured for a subset of B2 and non-B2 UAEC strains infecting C3H/HeN mice. Both B2 and non-B2 UAEC strains could cause chronic cystitis in mice. PUF scores and incidence of chronic cystitis are indicated for each strain. Horizontal bars indicate median values.

Fig. 5. **PUF carriage does not increase competitive advantage during co-infection.** UAEC strains isolated from two patients (9 and 41) with different strain-induced rUTIs were differentially marked with antibiotic resistance markers. C3H/HeN mice were then coinfected with the supplanting strain (9.2p or 41.4p) at an equal dose with the enrollment strain from the same patient (9.1a or 41.1a, respectively). A competitive index was calculated from the ratio of the supplanting strains over their enrollment strains in the urine from each mouse at 24 hpi. Clades and PUF scores are indicated for each strain. Horizontal bars indicate median values.

Schreiber et al., Sci. Transl. Med. 9, eaaf1283 (2017)   22 March 2017

5 of 13

and regulation in *E. coli* are unknown, were identified in all four deficient colonizer strains; however, they were not specific to deficient colonizers because they were also found in 2 of the 12 robust colonizers (12.1a and 26.1a). Further, the number of hits identified here fell below our threshold for type 1 errors, which was determined by extensive permutations of the enrichment analysis (see Supplementary Materials and Methods). Thus, we found that the enrichment of these two genes in deficient colonizers was likely due to random chance.

Whereas differences in gene carriage could not clearly discriminate between these two groups of UAEC strains, we hypothesized that differences in how core bacterial functions are regulated, such as pili production, might be able to explain the variation that we observed in mouse bladder colonization and UTI outcome. We found variability in pili function among our UAEC strains with mannose-sensitive hemagglutination titers, denoting the expression of type 1 pili, ranging from 0 to 2^{10} (fig. S7A) despite the presence of an intact *fim* operon in 20 of the 21 UAEC strains, as determined using custom BLAST and alignment-based searches (see Supplementary Materials and Methods). Notably, the closely related B2 strains 20.1a, 21.1a, and 35.1a had high mannose-resistant hemagglutination titers ($>2^4$), denoting expression of a non–type 1 pilus, under culture conditions that typically induce mannose-sensitive hemagglutination phenotypes in model B2 strains such as UTI89, including S pili (fig. S7B) (*52, 53*). We found that robust colonizer strains had significantly higher hemagglutination titers (median, $2^{7.25}$) than did deficient colonizers (median, $2^{2.25}$; $P < 0.01$, Mann-Whitney *U* test) or variable colonizers (median, 2^6; $P < 0.01$, Mann-Whitney *U* test) (fig. S7C) and that hemagglutination titer was significantly correlated with bladder bacterial burden in C3H/HeN mice at 24 hpi ($P < 0.001$, Spearman's rank correlation) (fig. S7D). Together, these data suggest that the regulatory networks that drive the expression of CUP pili genes differ between UAEC strains and that hemagglutination titer was correlated with the ability to cause acute cystitis in C3H/HeN mice.

To systematically compare core gene expression between phylogenetically and phenotypically diverse UAEC strains, we generated RNA sequencing (RNA-Seq) data from triplicate cultures of a subset of 11 strains after growth in static LB (bladder infection inoculum). These strains included (i) representative UAEC from four phylogenetic clades; (ii) robust, variable, and deficient colonizers; (iii) strains with disparate hemagglutination titers; and (iv) the model B2 strain UTI89 (table S8). Pairwise comparisons of transcript abundance for the 10 UAEC strains relative to the UTI89 strain revealed 753 genes (of 3516 core genes or ~21.4%) that were differentially expressed in at least 1 of the 10 UAEC strains under

these culture conditions ($P_{adj} < 0.05$; fold change, >3) (table S9) with B2 strains having fewer differentially expressed genes relative to UTI89 than non-B2 strains (fig. S8A). However, the expression of a large number of core genes differed among the clinical UAEC strains, including both the B2 and non-B2 strains, although they were grown under identical culture conditions. Differentially expressed genes were scattered throughout the genome, and many were clustered in operons encoding flagella, pilus, and chemotaxis machinery (fig. S8B). Notably, under this culture condition, the average expression of genes in the *fim* operon encoding type 1 pili was lower in deficient colonizers than robust colonizers and corresponded with the results of our hemagglutination assay. We also found that most PUFs present in each strain were expressed under these conditions, although at different levels when comparing across strains (table S10).

To explore whether transcriptional profiles under inoculum growth conditions were able to predict bladder colonization efficiency, we performed principal components analysis (PCA) on the complete set of core genome transcriptional profiles relative to UTI89 (Fig. 6A). We found a distinct pattern that separated the profiles of deficient and robust colonizers along the principal component 2 (PC2) axis, which explained 19.2% of the variance between the UAEC strains. Variable colonizers grouped with robust colonizers and away from deficient colonizers in this analysis. Further, PCA showed a separation between
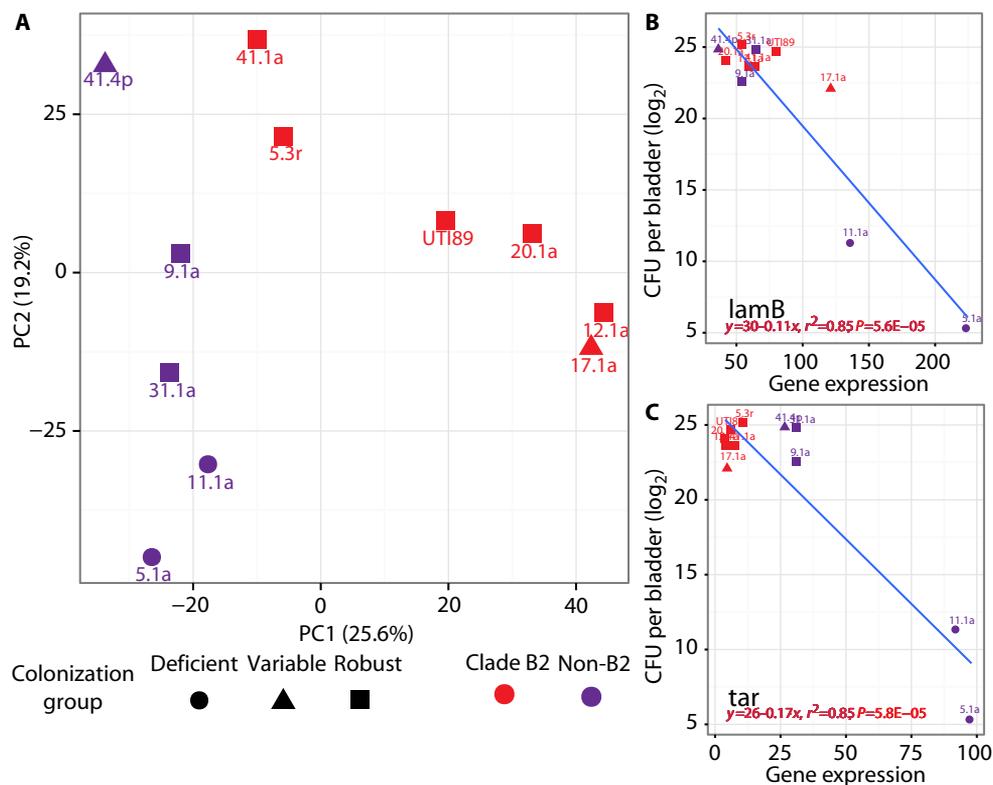

Fig. 6. Differential expression of core genes by UAEC strains distinguishes robust from deficient colonizers. (A) PCA was used to cluster UAEC strains based on their expression of 3340 core genes under type 1 pili–inducing culture conditions used to culture bacteria before inoculation into mouse bladders. B2 strains separated from non-B2 strains along PC1, whereas robust colonizers separated from deficient colonizers along PC2. Linear regression identified the expression of 42 core genes under type 1 pili–inducing conditions, as correlated with bladder bacterial burden at 24 hpi in C3H/HeN mice after correction for false discovery ($P < 0.1$). These 42 genes included pathway genes mediating nutrient uptake, such as *lamB* (**B**), and chemotaxis, such as *tar* (**C**), shown here as representative examples. Data presented represent averages derived from three independent experiments that were corrected for variance in read depth between samples and gene length.

Schreiber *et al.*, *Sci. Transl. Med.* **9**, eaaf1283 (2017) 22 March 2017

6 of 13

UAEC strains along PC1 (25.6% of variance) that correlated with clade membership (that is, B2 versus non-B2), suggesting that differences in gene expression associated with phylogeny are distinct from those that separate deficient and robust colonizers. Using a linear regression analysis, we systematically evaluated how expression of each gene (using normalized read counts per gene) correlated with bacterial burden at 24 hpi in the mouse bladder. After correcting for false discovery, we identified a total of 42 core genes whose expression in static LB correlated with subsequent bladder colonization efficiency in C3H/HeN mice at 24 hpi (fig. S9 and table S11). These included genes encoding sugar-transport proteins, such as the LamB maltoporin (Fig. 6B), and other genes associated with maltose transport (54, 55), as well as the Tar protein (Fig. 6C), a chemotactic sensor of aspartate and maltose (56, 57). Most linear regression curves had negative beta coefficients, indicating that deficient colonizers had higher expression of these genes than robust colonizers. We grouped these 42 core genes into functional pathways (58, 59), which showed a total of six functions that were enriched, including motility, chemotaxis, and carbohydrate transport (table S12). These results indicate that, under culture conditions used for growth of bacteria before infecting mice, core bacterial behaviors, such as swimming motility and nutrient transport, were regulated differently between robust and deficient colonizers in UAEC strains. Despite differences in how genes mediating motility and chemotaxis were expressed, the motility of the clinical UAEC strains did not correlate with bladder colonization in C3H/HeN mice at 24 hpi (fig. S10), suggesting that transcriptional control of motility, rather than the in vitro phenotype itself, was a predictor of subsequent colonization efficiency.

UAEC strains vary in their ability to colonize the bladders of mice of different genetic backgrounds

Our data suggest that dynamic features (for example, transcriptional control of core *E. coli* genes) rather than static characteristics (for example, carriage of PUFs) could predict the relative efficiency of UAEC strain colonization of mouse bladder. Thus, differential regulation of transcriptional networks during in vitro growth in the colonization inoculum may have served to prime UAEC bladder colonization. To determine whether this "priming" transcriptional response was universal in predicting colonization among different mouse strains, we inoculated a subset of our UAEC strains into C57BL/6 mice, which was a different host environment with a different immune response to infection. As with the C3H/HeN mouse model, the UAEC strains showed a variable ability to colonize the bladders of C57BL/6 mice, with bladder bacterial burdens ranging from undetectable to >10^7 CFU per bladder at 24 hpi (Fig. 7). Overall, we observed less variation in bladder colonization for each UAEC strain in C57BL/6 mice than seen in C3H/HeN mice; for example, the highly variable strains in C3H/HeN colonization experiments, 2.2r and 17.1a, showed a more narrow distribution of bladder burden at 24 hpi in the C57BL/6 mice. Whereas

most UAEC strains colonized the C57BL/6 mouse bladders at rates similar to the C3H/HeN bladders at 24 hpi, two strains, 31.1a (clade B1, PUF score = 4) and 41.1a (clade B2, PUF score = 20), showed significantly lower bladder bacterial burdens in C57BL/6 mice (*P* < 0.01, Mann-Whitney *U* test) (Fig. 7). Notably, the reduction in colonization ability did not appear to be based on genetic background or phylogeny of the bacterial strains because other UAEC strains carrying similar numbers of PUFs and strains from the same clades as 31.1a and 41.1a colonized C3H/HeN and C57BL/6 mouse bladders at similar levels.

Quiescent intracellular reservoirs are small collections of dormant UPEC cells that remain in LAMP-1 (lysosomal-associated membrane protein 1)–positive vesicles in the underlying uroepithelium and represent another mode of bladder colonization within a different host environment. Quiescent intracellular reservoirs persist in the bladder even after apparent clearance of bacteria from the kidneys and urine and may be able to seed same-strain rUTI (60). In addition to acute cystitis, C57BL/6 mice have been used to model the formation of quiescent intracellular reservoirs. In general, quiescent intracellular reservoirs can be detected through plating bladder homogenates 14 dpi after the clearance of bacteriuria and kidney infection, as has been carried out previously (55). We tested four robust colonizer UAEC strains, including 9.1a (clade D, PUF score = 13), 9.2p (clade B1, PUF score = 3), 35.1a (clade B2, PUF score = 16), and 41.1a (clade B2, PUF score = 20), and the model B2 UTI89 strain (PUF score = 17) and found that all five strains resulted in detectable bladder CFU at 14 dpi (fig. S11) despite sterile urine, indicating the existence of quiescent intracellular reservoirs. However, the non-B2 strain 9.1a produced lower bladder bacterial burdens at this time point than did other tested strains, including the non-B2 strain 9.2p, which was isolated from the same patient. Strain 9.1a also produced fewer IBCs than did other strains tested (table S6). Thus, diverse UAEC strains could colonize multiple host backgrounds. However, the relative efficiency of mouse bladder colonization varied between different UAEC strains in the same host background, and individual UAEC strains had different relative abilities to infect the bladders of mice with different genetic backgrounds.
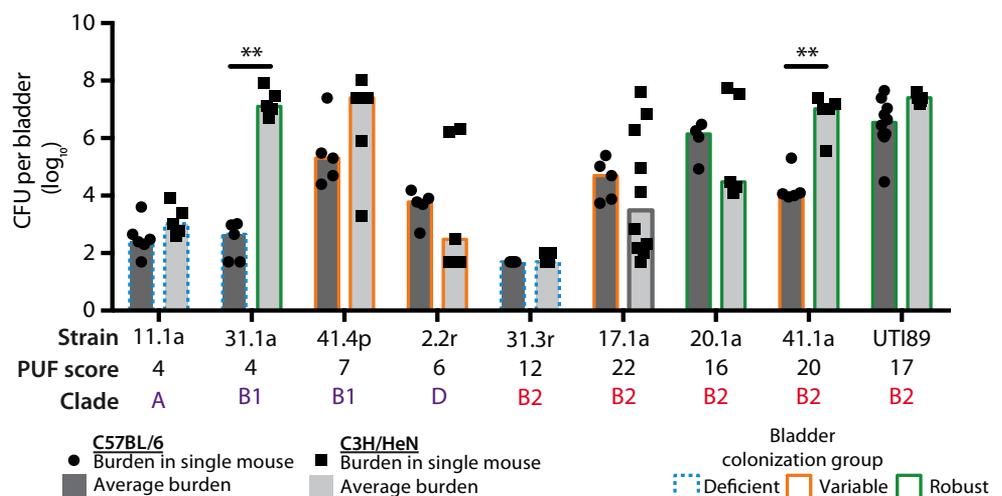

Fig. 7. Colonization efficiency by UAEC strains varies between mouse models. Select UAEC strains were inoculated into C3H/HeN mice (black squares and light gray bars) and C57BL/6 mice (black circles and dark gray bars). Bacteria were enumerated from individual harvested bladders at 24 hpi (black squares and circles). Median values for each infection are given (gray bars). Each UAEC infection of the two mouse strains was categorized separately as "deficient" (blue stippled outline), "variable" (orange outline), or "robust" (green outline). Bladder colonization between C3H/HeN and C57BL/6 mice was significantly different for strains 31.1a and 41.1a by Mann-Whitney *U* test, **P* < 0.01.

UAEC strains that robustly colonized mice correlated with markers of increased UTI severity in patients

An exuberant inflammatory response to an initial UTI is associated with severity of subsequent infections in mice (5) and susceptibility to recurrent infections in women (6). This inflammatory response is cyclooxygenase-2–dependent and is associated with increased neutrophil infiltration across the urothelium during infection, which results in increased numbers of white blood cells in the urine (5, 6). In addition to neutrophils, the host protects itself from UTI through secretion of lipocalin-2, which is induced upon infection of the bladder and helps to sequester iron away from the infecting organism (61). Levels of lipocalin-2 in the urine correlate with an increase in bacteria in the urine (bacteriuria) of women (61). To determine whether UAEC strains having increased virulence in the mouse model caused more severe UTI in women, we examined white blood cells and lipocalin-2 in our cohort of patients infected with robust, variable, or deficient colonizer UAEC strains. We found that strains defined as robust and variable colonizer in the C3H/HeN mouse model were associated with higher numbers of white blood cells in the urine than deficient colonizers during UTI episodes (table S13). Further, we found that robust colonizers were associated with higher lipocalin-2 in the urine than either variable or deficient colonizers (table S13).

DISCUSSION

Although virulence in other *E. coli* pathotypes has been tightly linked to carriage of specific genes, UPEC are distinct in their lack of a signature set of genes that universally distinguishes them from non-UPEC strains. This likely reflects the broad definition of UPEC—any *E. coli* strain that is recovered from the urine of a symptomatic UTI patient—a classification that fails to account for differences in host susceptibility or the possibility that multiple evolutionary and mechanistic paths can lead to urovirulence. To better define the determinants of urovirulence, we focused our study on a collection of 43 diverse UAEC strains isolated from a cohort of 14 women with frequent rUTI. Here, we have integrated genomic and transcriptomic approaches with in vitro and in vivo characterization of isolate phenotypes, including their colonization and pathogenic potential in defined experimental mouse models of cystitis. Our findings expand our understanding of the genomic diversity of UPEC and show that differential transcriptional regulation of core genes contributes to pathogenic potential and that the relative virulence of different *E. coli* strains varies depending on host background. We conclude that the range of clinical UTI outcomes reflects a range of complex host-pathogen interactions driven, in part, by variation in how uropathogens regulate core functions, affecting their pathogenic potential in hosts with different genetics and health status.

Phylogenomic analysis of this collection of UAEC strains revealed patterns of phylogenetic diversity previously reported in other clinical studies (15–20), including an enrichment of clade B2 strains (67% of total) and a mixture of same-strain and different-strain rUTIs, with B2 strains dominating same-strain rUTIs (11 of 12 same-strain rUTI isolates fell within the B2 clade) (20). We observed variation in the genomic content of these isolates, up to 40% in pairwise comparisons between strains, similar to the level of genomic diversity observed in *E. coli* overall (13, 14). Genomic diversity among *E. coli* strains is thought to be key to the species' ability to thrive in a multitude of environments (14). Thus, it is possible that the genetic diversity observed in our cohort of UAEC strains, where two strains may share

only 60% of the genes in their genome, reflects the sum of selection pressures across multiple habitats including but not limited to reservoirs in the gut, intracellular and extracellular locales in the bladder, and diverse host backgrounds. In addition to other genes, we found that UAEC strains were variable in their carriage of PUF genes, which have previously been identified as being enriched in UTI isolates relative to other *E. coli* strains (20, 38–40). In contrast to this, we found that, if phylogeny was taken into account, there was no enrichment of PUFs in UAEC strains relative to non-UAEC strains. Thus, the enrichment of PUFs in urinary isolates of *E. coli* reported previously may be due to a combination of the enrichment of clade B2 in urinary isolates and the phylogenetic bias in PUF carriage in clade B2 strains.

The molecular mechanisms of pathogenicity for many PUFs remain poorly defined and may have functions outside of bladder pathogenicity that still enhance the overall virulence of UPEC strains, such as increasing successful migration from the gut reservoir to the bladder or enabling persistence in host niches outside of the bladder, such as the vagina. One limitation of our study is that it is solely focused on bladder virulence, and thus, functions outside of the bladder were not tested. Further, given the fact that many of these factors are colocalized on pathogenicity-associated islands common to the B2 clade and absent from other *E. coli* (62), it is possible that the enrichment of PUFs results from a type of "genetic hitchhiking" due to some selection for B2 strains, such as their transcriptional profiles or persistence in gut reservoirs, rather than selection for the PUF genes themselves. Strikingly, we found that PUF carriage did not correlate with urovirulence in C3H/HeN mice. Non-B2 UAEC that carried few PUFs were as proficient as B2 UAEC carrying many PUFs in causing cystitis in mice. We also observed that, in two patients, a UAEC strain with a higher number of PUFs was supplanted by a UAEC strain with a lower PUF score. In addition, we found in a mouse model that the supplanting strains outcompeted the initial UTI strains in co-infection experiments. However, we found that same-strain rUTI isolates were much more likely to be from clade B2 than from other clades, which suggests that B2-specific genomic features, such as carriage of PUFs, other genes, and/or transcriptional regulatory machinery, may provide an advantage for persisting within hosts between symptomatic episodes of UTI. Further, whereas PUF carriage was neither necessary nor sufficient for bladder colonization at acute time points, we did find that deletion of a large pathogenicity island from UTI89, which carried 124 genes including four PUFs, resulted in a loss of fitness during competition in chronic cystitis. Although this result does not directly implicate these four PUFs in bladder colonization, it does show that carriage of at least some of the genes in this island may provide competitive advantages during different aspects of disease. All of our studied isolates came from the urine of women either during or between episodes of rUTI, and thus, all of them, by definition, were uropathogens that are capable of colonizing at least one human bladder. However, because host genetics, environment, and behavior are important for UTI susceptibility, it is likely that many UAEC strains isolated from a single patient may be unable to cause infection in all or even most human bladders. To control for these interhost variations, we assessed the colonization potential of UAEC and non-UAEC strains in genetically defined animal models under identical environmental conditions. Although mouse bladder colonization by UAEC strains was markedly higher than that by gut-associated *E. coli*, we observed that bladder colonization varied among UAEC isolates under these conditions. However, even after comparing the gene content of robust and deficient colonizer UAEC isolates, we found no significant differences in specific

gene content or overall functional potential that could account for the variability of these strains in colonizing mouse bladders. Another important limitation of our study is our focus on gene carriage among strains. Because our comparative analysis focused solely on gene content, as determined by orthology, we would not have been able to identify more subtle genomic differences that could influence uropathogenicity, including variation in promoter sequences or in coding sequences that might affect gene expression or protein structure and function.

Although we found no clear genomic signatures that could be identified to discriminate between robust and deficient colonizers, we found that differences in the expression of core genes shared among all UAEC strains were predictive of urovirulence in C3H/HeN mouse bladders. These included genes involved in maltose transport, chemotaxis, and flagellar assembly that mediate core bacterial behaviors such as nutrient utilization and motility. The variance in expression of core functions contrasts with studies showing that shared genes are expressed similarly in different *E. coli* strains (63), possibly reflecting different definitions of the "core genome/transcriptome" between the studies. Our definition was based on comparisons of genomic sequences, whereas previous reports have identified genes that were "commonly transcribed" to identify a core transcriptome. This variance in transcriptional regulation of core bacterial behaviors has a direct impact on the ability of UAEC strains to colonize host bladders. This is exemplified by the differential regulation of the conserved *fim* operon, which encodes type 1 pili. Although all but one of our UAEC strains carried an intact *fim* operon, the strains varied considerably in their hemagglutination titers, a measure of piliation. Further, hemagglutination titers correlated well with colonization of the C3H/HeN mouse bladder. We posit that that the regulatory networks that control type 1 pili expression within the different UAEC strains respond differently to environmental cues. For example, a set of highly related strains (from patients 20, 21, and 35), all carrying the *fim* gene cluster, expressed S pili instead of type 1 pili when grown under standard type 1 pilus–inducing conditions (24). These strains were robust colonizers of C3H/HeN mice, indicating that their increased expression of S pili under inoculum conditions did not prevent their urovirulence. In the UTI89 strain, expression of S and type 1 pili is inversely controlled so that blocking expression of one pilus type induces the expression of the other (28). Further, differences in type 1 pili regulation are seen in the ways that UAEC regulate their motility and pili expression relative to model B2 UPEC strains. For example, increased expression of type 1 pili is inversely correlated with expression of flagellar genes in UTI89 (64). Here, we found that the non-B2 strains exhibiting high expression of type 1 pili also had increased motility, indicating that the networks mediating coordinated regulation of flagella and type 1 pilus expression may differ between B2 and non-B2 UAEC strains. Together, our findings implicate the divergence of transcriptional and regulatory networks as a key driver of UPEC pathogenesis and underscore the importance of examining transcriptional regulation of genes in addition to their patterns of carriage to more fully understand the factors underlying bacterial virulence and host-pathogen interactions. Moreover, our findings suggest that transcriptional responses to the host environment likely also diverge among UAEC strains and that, to gain richer insights into the determinants of UTI risk and outcome, we need a better understanding of how gene expression differences during infection influence bacterial physiology and interactions with the host.

The progression and outcome of a UTI are determined not only by the virulence potential of the infecting bacteria but also by myriad factors in the host environment. Mouse models of UTI recapitulate the histological markers of human UTI. Multiple mouse models have been developed, each reflecting a portion of the diversity of UTI pathology seen in the clinic [reviewed in (65, 66)]. Consistent with their ability to cause UTIs in humans, many of our UAEC strains were able to elicit key hallmarks of human pathogenesis in mouse models of UTI, including acute cystitis, the formation of IBCs, the development of chronic cystitis, and persistence in quiescent intracellular reservoirs. We found that the ability of certain UAEC strains to cause acute cystitis in the C3H/HeN mice correlated with markers of increased UTI severity in patients. Specifically, UAEC strains that exhibited an increased ability to cause acute cystitis in the C3H/HeN mouse model were associated with an increase in leukocytes and lipocalin-2 in the urine of the patients from which the strains were derived. Lipocalin-2 is known to be associated with increased activation of the immune response in human UTIs (61). In addition, transmigration of neutrophils across the bladder epithelium, which results in an increase in leukocytes in the urine, has been associated with UTI severity and recurrence (6).

Our study suggests that each infection by a single UAEC strain in a single mouse background captures only a small part of the complex landscape of factors that govern UTI progression and outcome. Specifically, we found that, although most UAEC strains had similar success colonizing the bladder of C57BL/6 and C3H/HeN mice, two phylogenetically and genetically diverse strains were able to colonize the C3H/HeN bladder much better than the C57BL/6 bladder, indicating that some feature of the C57BL/6 bladder environment presents a specific barrier to colonization that could be overcome by some strains but not others. These findings suggest that the barriers to infection vary among mouse backgrounds as do the capacities of different UAEC strains to overcome these barriers. Thus, the outcome of each UTI is determined by the compatibility of bacterial virulence and host susceptibility factors involved in that infection. Gaining a more comprehensive understanding of the factors that determine UTI risk and outcome will require studies of urovirulence using more combinations of UAEC isolates and mouse backgrounds that better represent the diversity of bacterial urovirulence potentials and host susceptibilities.

MATERIALS AND METHODS
Study design
This study was conducted to identify conserved bacterial features that enabled bladder colonization and virulence by *E. coli* and to determine whether these features enabled virulence in all host backgrounds. In a previous study that was approved by the Human Subjects Review Committee at the University of Washington, 104 women aged 18 to 49 years with a self-reported history of UTI and a current diagnosis of acute cystitis were enrolled in the analysis cohort and self-collected midstream urine samples daily for 90 days, as described previously (67). Women were recruited from University of Washington Health Centers in Seattle, WA. Exclusion criteria for this cohort included known anatomic or functional abnormalities of the urinary tract, chronic illness, pregnancy, or development of acute pyelonephritis (67). From this cohort, a total of 29 women experienced a rUTI within the 90-day study window and provided urine samples containing UAEC. We examined strains isolated from the first 14 of these 29 women, resulting in a total of 43 UAEC isolates collected throughout the study (see fig. S1, table S2, and Supplementary Materials and Methods). Fourteen isolates were collected

at enrollment, 18 isolates were collected during rUTI, and 11 isolates were from daily urine samples collected at home by the patients in the days preceding the diagnosis of UTI (Table 1). A single *E. coli* isolate was collected from each urine sample using selective plating techniques (67), and all isolates were stored at −80°C. When possible, white blood cells and lipocalin-2 were measured in the urine samples, as described previously (67). *E. coli* rUTI was defined on the basis of (i) the presence of *E. coli* in the urine at ≥10^2 CFU/ml and (ii) medical diagnosis of a UTI based on symptoms of acute cystitis, such as foul-smelling urine, increased frequency of urination, and/or urgency to urinate.

Additional strains used in this project
The model strains MG1655, Sakai, and Nissle 1917 were derived from pure cultures and stored in glycerol at −80°C. As described in a previous study, gut-associated *E. coli* were isolated from human feces collected from healthy adults (46, 47). Briefly, single colonies were identified as *E. coli* with isolation plating on differential media, and glycerol stocks were stored at −80°C.

Genome data set construction
We constructed draft genomes of the 43 UAEC strains by sequencing extracted DNA on the Illumina HiSeq 2000 platform and assembling the reads into ordered contigs using Velvet v1.2 (68) and Mauve v2.3.1 (see table S14 and Supplementary Materials and Methods) (69). To contextualize the UAEC genomes, we constructed a database containing the 43 draft UAEC genomes and 46 closed *E. coli* genomes from the National Center for Biotechnology Information (table S3). All genomes were reannotated using the Broad Institute's prokaryote annotation pipeline (70), and genes were clustered into orthogroups using the reciprocal best BLAST hits method (71).

Phylogenetic and gene content analysis
Phylogenetic trees were generated using RAxML (72) with bootstrapping on a concatenated alignment of 2746 single-copy core orthogroups across 89 organisms aligned using MUSCLE (73). SNPs were calculated in this single-copy core genome and used to identify clusters of highly related strains (see Supplementary Materials and Methods). Targeted analysis of PUF genes (table S4) and the *fim* operon was performed using a combination of BLAST-based interrogation and confirmed by mapping DNA reads to a database of the PUF and *fim* genes (see Supplementary Materials and Methods). We compared global gene content differences between deficient colonizing UAEC and robust colonizers (with UTI89) using a stringent set of criteria and more relaxed definition of enrichment (see Supplementary Materials and Methods). In addition to gene content, we also investigated differences in pathways or functional domains between groups of strains using Pfam (74) and KEGG (Kyoto Encyclopedia of Genes and Genomes) (75). To provide context for our analysis and estimate the chance of "false positives" in our results, we calculated the rate type 1 errors arising from random chance, as described previously (see Supplementary Materials and Methods) (76).

Mouse infections
Mouse infections were carried out using previously established models (5, 77, 78). In preparation for inoculation into mice, bacterial strains were grown statically in LB at 37°C overnight and then subcultured 1:1000 into 10 ml of fresh LB for a second static outgrowth at 37°C for 18 to 24 hours to induce expression of type 1 pili (52). Bacteria from these cultures were harvested and resuspended in 1× phosphate-buffered saline (PBS) for inoculation. Mice used in these experiments were purchased from Charles River Laboratories and Harlan Laboratories. Urine collections and harvested bladders and kidneys were serially plated on LB agar plates with or without appropriate antibiotics for enumeration of bacterial burdens. Competitive indexes were calculated on the basis of the ratios of marked strains in each mouse replicate.

Detection of IBCs and persistent bladder reservoirs
The formation of IBCs was measured in 7- to 8-week-old female C3H/HeN mice that were inoculated with the indicated strain as described earlier (see "Mouse infections" above). IBCs were visualized at 6 hpi with α–*E. coli* antibodies, and bladder cells were counterstained with wheat germ agglutinin. Persistent bladder reservoirs were measured in C57BL/6 mice that were inoculated with the indicated strain. Mice were sacrificed after 2 weeks for enumeration of bacterial burdens in the bladders and kidneys. Mice that developed chronic infections, as determined by urine titers collected during the 2-week infection, or mice that had visible kidney abscesses were removed from this analysis.

Phenotypic analyses
For all phenotypic analyses, clinical UAEC were first grown under type 1 pili–inducing conditions (see "Mouse infections" above). Hemagglutination assays were performed with guinea pig erythrocytes in PBS with or without 2% methyl-α-D-mannopyranoside incubated at 4°C overnight. Swimming motility assays were performed in 0.25% agar LB, and the diameter of the spread of the bacteria was measured after 6-hour incubation at 37°C.

Generation of RNA-Seq data
After growth under type 1 pili–inducing conditions (see "Mouse infections" above), RNA was extracted from select UAEC strains grown in triplicate, and Illumina complementary DNA (cDNA) libraries were constructed following the RNAtag-seq protocol (see Supplementary Materials and Methods) (79). Barcoded cDNA was then sequenced on the Illumina HiSeq 2000 platform. Triplicate RNA-Seq libraries were demultiplexed and mapped to their cognate strain using BWA (Burrows-Wheeler Aligner) (80). Targeting a set of 3516 "core transcriptome genes" found in UTI89 and 9 of 10 UAEC strains tested, we calculated differences in expression using DESeq (81). We normalized for read abundance and gene length, averaged expression levels across the triplicate samples, and then removed genes whose average expression level across the 11 UAEC fell into the lowest quartile of genes. The remaining genes were then used in PCA and linear regression with appropriate controls for multiple hypothesis testing.

Table 1. Summary of 43 UAEC isolates from women with rUTIs.

Isolation time point	Number of isolates
Enrollment*	14
rUTI*	18
Before recurrence[†]	11
Total	43

*Diagnosed UTI.   [†]No UTI diagnosis.

Statistical analysis

Colonization groups were defined on the basis of the median value of mouse bladder colonization by UAEC strains at 24 hpi. Repeated nonparametric Mann-Whitney U tests were used to identify statistically significant differences in PUF gene carriage, hemagglutination assay titers, and swimming motility between colonization groups. Further, repeated Mann-Whitney U tests were used to identify differences in PUF gene carriage between clades, colonization efficiency of different inbred mouse strains by UAEC isolates, and persistence of quiescent intracellular reservoirs by UAEC strains in C57BL/6 mice. Correlations between median bladder burden in C3H/HeN mice and PUF gene carriage, hemagglutination assay titer, or motility in UAEC strains were determined using Spearman's rank correlation test in separate analyses. UAEC strains were clustered into related groups using unsupervised hierarchical clustering of SNP distances defined through pairwise alignments of the core genome. Further, unsupervised clustering was used to identify UAEC strains that shared similar carriage of PUF genes. A hypergeometric distribution test was used to compare enrichment of clade B2 strains causing same-strain rUTI relative to the distribution of B2 strains with our collection of clinical isolates. Differences in expression of core genes were measured using the DESeq algorithm. PCA was used to visualize similarities in core gene expression, and linear regression analysis was used to correlate these differences in gene expression to median bladder burden at 24 hpi in C3H/HeN mice. Type 1 error modeling was established with repeated permutation analysis with 1000 replicates and counts of gene presence/absence and is described in detail in the Supplementary Materials. Where P values could be assigned, a threshold of significance was established at $P < 0.05$ after multiple hypothesis testing correction ($P_{adj} < 0.05$), except for the linear regression analysis of gene expression to bladder burden, which used a significance threshold of $P_{adj} < 0.1$.

SUPPLEMENTARY MATERIALS

www.sciencetranslationalmedicine.org/cgi/content/full/9/382/eaaf1283/DC1

Materials and Methods

Fig. S1. Sample collection and patient time line.
Fig. S2. Measurement of UAEC gene carriage and nucleotide diversity in core genome.
Fig. S3. Kidney colonization by B2 and non-B2 UAEC strains.
Fig. S4. Gut-associated *E. coli* are poor colonizers in the C3H/HeN mouse model of UTI.
Fig. S5. IBC formation by B2 and non-B2 UAEC strains.
Fig. S6. Carriage of PAI II$_{UTI89}$ enhances competitive fitness in chronic UTI in C3H/HeN mice.
Fig. S7. Hemagglutination is correlated to colonization efficiency in both B2 and non-B2 UAEC strains.
Fig. S8. Differentially expressed genes in UAEC relative to UTI89.
Fig. S9. Expression of core genes under type 1 pili–inducing culture conditions correlates with bladder colonization in UAEC.
Fig. S10. Motility does not correlate with colonization efficiency in either B2 or non-B2 UAEC strains.
Fig. S11. Both B2 and non-B2 UAEC strains are capable of forming persistent bladder reservoirs in C57BL/6 mice.
Table S1. Clinical information of enrolled patients.
Table S2. Characteristics of UAEC isolates.
Table S3. Reference *E. coli* strains.
Table S4. List of PUFs.
Table S5. Gut-associated *E. coli* strain characteristics.
Table S6. IBC formation in select UAEC strains at 6 hpi in C3H/HeN mice.
Table S7. Presence of *brnAT* genes in robust and deficient colonizer UAEC strains.
Table S8. Reads mapped to core genome of UAEC strains.
Table S9. DESeq analysis of UAEC transcription of core genome.
Table S10. Expression of PUF genes in select UAEC strains.
Table S11. Linear correlation of core gene expression with bladder colonization.
Table S12. Functions enriched in genes associated with bladder colonization.
Table S13. UAEC strain clinical data.
Table S14. UAEC strain sequencing and genome assembly results.
References (82–86)

REFERENCES AND NOTES

1. B. Foxman, Epidemiology of urinary tract infections: Incidence, morbidity, and economic costs. *Am. J. Med.* **113** (suppl. 1A), 5–13 (2002).
2. B. Foxman, Urinary tract infection syndromes: Occurrence, recurrence, bacteriology, risk factors, and disease burden. *Infect. Dis. Clin. North Am.* **28**, 1–13 (2014).
3. C. E. Mabeck, Treatment of uncomplicated urinary tract infection in non-pregnant women. *Postgrad. Med. J.* **48**, 69–75 (1972).
4. S. A. Ferry, S. E. Holm, H. Stenlund, R. Lundholm, T. J. Monsen, The natural course of uncomplicated lower urinary tract infection in women illustrated by a randomized placebo controlled study. *Scand. J. Infect. Dis.* **36**, 296–301 (2004).
5. T. J. Hannan, I. U. Mysorekar, C. S. Hung, M. L. Isaacson-Schmid, S. J. Hultgren, Early severe inflammatory responses to uropathogenic *E. coli* predispose to chronic and recurrent urinary tract infection. *PLOS Pathog.* **6**, e1001042 (2010).
6. T. J. Hannan, P. L. Roberts, T. E. Riehl, S. van der Post, J. M. Binkley, D. J. Schwartz, H. Miyoshi, M. Mack, R. A. Schwendener, T. M. Hooton, T. S. Stappenbeck, G. C. Hansson, W. F. Stenson, M. Colonna, A. E. Stapleton, S. J. Hultgren, Inhibition of cyclooxygenase-2 prevents chronic and recurrent cystitis. *EBioMedicine* **1**, 46–57 (2014).
7. V. P. O'Brien, T. J. Hannan, L. Yu, J. Livny, E. D. O. Roberson, D. J. Schwartz, S. Souza, C. L. Mendelsohn, M. Colonna, A. L. Lewis, S. J. Hultgren, A mucosal imprint left by prior *Escherichia coli* bladder infection sensitizes to recurrent disease. *Nat. Microbiol.* **2**, 16196 (2016).
8. K. Gupta, N. Bhadelia, Management of urinary tract infections from multidrug-resistant organisms. *Infect. Dis. Clin. North Am.* **28**, 49–59 (2014).
9. B. Foxman, B. Gillespie, J. Koopman, L. Zhang, K. Palin, P. Tallman, J. V. Marsh, S. Spear, J. D. Sobel, M. J. Marty, C. F. Marrs, Risk factors for second urinary tract infection among college women. *Am. J. Epidemiol.* **151**, 1194–1205 (2000).
10. K. Gupta, T. M. Hooton, K. G. Naber, B. Wullt, R. Colgan, L. G. Miller, G. J. Moran, L. E. Nicolle, R. Raz, A. J. Schaeffer, D. E. Soper, International clinical practice guidelines for the treatment of acute uncomplicated cystitis and pyelonephritis in women: A 2010 update by the Infectious Diseases Society of America and the European Society for Microbiology and Infectious Diseases. *Clin. Infect. Dis.* **52**, e103–e120 (2011).
11. T. A. Russo, J. R. Johnson, Proposal for a new inclusive designation for extraintestinal pathogenic isolates of *Escherichia coli*: ExPEC. *J. Infect. Dis.* **181**, 1753–1754 (2000).
12. J. B. Kaper, J. P. Nataro, H. L. T. Mobley, Pathogenic *Escherichia coli. Nat. Rev. Microbiol.* **2**, 123–140 (2004).
13. D. A. Rasko, M. J. Rosovitz, G. S. A. Myers, E. F. Mongodin, W. F. Fricke, P. Gajer, J. Crabtree, M. Sebaihia, N. R. Thomson, R. Chaudhuri, I. R. Henderson, V. Sperandio, J. Ravel, The pangenome structure of *Escherichia coli*: Comparative genomic analysis of *E. coli* commensal and pathogenic isolates. *J. Bacteriol.* **190**, 6881–6893 (2008).
14. M. Touchon, C. Hoede, O. Tenaillon, V. Barbe, S. Baeriswyl, P. Bidet, E. Bingen, S. Bonacorsi, C. Bouchier, O. Bouvet, A. Calteau, H. Chiapello, O. Clermont, S. Cruveiller, A. Danchin, M. Diard, C. Dossat, M. E. Karoui, E. Frapy, L. Garry, J. M. Ghigo, A. M. Gilles, J. Johnson, C. Le Bouguénec, M. Lescat, S. Mangenot, V. Martinez-Jéhanne, I. Matic, X. Nassif, S. Oztas, M. A. Petit, C. Pichon, Z. Rouy, C. S. Ruf, D. Schneider, J. Tourret, B. Vacherie, D. Vallenet, C. Médigue, E. P. C. Rocha, E. Denamur, Organised genome dynamics in the *Escherichia coli* species results in highly diverse adaptive paths. *PLOS Genet.* **5**, e1000344 (2009).
15. M. Rijavec, M. S. Erjavec, J. A. Avguštin, R. Reissbrodt, A. Fruth, V. Križan-Hergouth, D. Žgur-Bertok, High prevalence of multidrug resistance and random distribution of mobile genetic elements among uropathogenic *Escherichia coli* (UPEC) of the four major phylogenetic groups. *Curr. Microbiol.* **53**, 158–162 (2006).
16. M. S. Erjavec, M. Rijavec, V. Križan-Hergouth, Chloramphenicol- and tetracycline-resistant uropathogenic *Escherichia coli* (UPEC) exhibit reduced virulence potential. *Int. J. Antimicrob. Agents* **30**, 436–442 (2007).
17. G. Piatti, A. Mannini, M. Balistreri, A. M. Schito, Virulence factors in urinary *Escherichia coli* strains: Phylogenetic background and quinolone and fluoroquinolone resistance. *J. Clin. Microbiol.* **46**, 480–487 (2008).
18. L. Skjøt-Rasmussen, A. M. Hammerum, L. Jakobsen, C. H. Lester, P. Larsen, N. Frimodt-Moller, Persisting clones of *Escherichia coli* isolates from recurrent urinary tract infection in men and women. *J. Med. Microbiol.* **60**, 550–554 (2011).
19. K. Ejrnæs, M. Stegger, A. Reisner, S. Ferry, T. Monsen, S. E. Holm, B. Lundgren, N. Frimodt-Møller, Characteristics of *Escherichia coli* causing persistence or relapse of urinary tract infections: Phylogenetic groups, virulence factors and biofilm formation. *Virulence* **2**, 528–537 (2011).
20. Y. Luo, Y. Ma, Q. Zhao, L. Wang, L. Guo, Y. Ye, Y. Zhang, J. Yang, Similarity and divergence of phylogenies, antimicrobial susceptibilities, and virulence factor profiles of *Escherichia*

coli isolates causing recurrent urinary tract infections that persist or result from reinfection. *J. Clin. Microbiol.* **50**, 4002–4007 (2012).

21. Y. Wang, S. Zhao, L. Han, X. Guo, M. Chen, Y. Ni, Y. Zhang, Z. Cui, P. He, Drug resistance and virulence of uropathogenic Escherichia coli from Shanghai, China. *J. Antibiot.* **67**, 799–805 (2014).

22. S. L. Chen, C.-S. Hung, J. Xu, C. S. Reigstad, V. Magrini, A. Sabo, D. Blasiar, T. Bieri, R. R. Meyer, P. Ozersky, J. R. Armstrong, R. S. Fulton, J. P. Latreille, J. Spieth, T. M. Hooton, E. R. Mardis, S. J. Hultgren, J. I. Gordon, Identification of genes subject to positive selection in uropathogenic strains of *Escherichia coli*: A comparative genomics approach. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 5977–5982 (2006).

23. R. A. Welch, V. Burland, G. Plunkett III, P. Redford, P. Roesch, D. Rasko, E. L. Buckles, S.-R. Liou, A. Boutin, J. Hackett, D. Stroud, G. F. Mayhew, D. J. Rose, S. Zhou, D. C. Schwartz, N. T. Perna, H. L. T. Mobley, M. S. Donnenberg, F. R. Blattner, Extensive mosaic structure revealed by the complete genome sequence of uropathogenic *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 17020–17024 (2002).

24. S. J. Hultgren, W. R. Schwan, A. J. Schaeffer, J. L. Duncan, Regulation of production of type 1 pili among urinary tract isolates of *Escherichia coli*. *Infect. Immun.* **54**, 613–620 (1986).

25. E. Brzuszkiewicz, H. Brüggemann, H. Liesegang, M. Emmerth, T. Ölschläger, G. Nagy, K. Albermann, C. Wagner, C. Buchrieser, L. Emődy, G. Gottschalk, J. Hacker, U. Dobrindt, How to become a uropathogen: Comparative genomic analysis of extraintestinal pathogenic *Escherichia coli* strains. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 12879–12884 (2006).

26. M. A. Mulvey, J. D. Schilling, S. J. Hultgren, Establishment of a persistent *Escherichia coli* reservoir during the acute phase of a bladder infection. *Infect. Immun.* **69**, 4572–4579 (2001).

27. D. J. Schwartz, V. Kalas, J. S. Pinkner, S. L. Chen, C. N. Spaulding, K. W. Dodson, S. J. Hultgren, Positively selected FimH residues enhance virulence during urinary tract infection by altering FimH conformation. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 15530–15537 (2013).

28. D. J. Wurpel, S. A. Beatson, M. Totsika, N. K. Petty, M. A. Schembri, Chaperone-usher fimbriae of *Escherichia coli*. *PLOS ONE* **8**, e52835 (2013).

29. S. E. Greene, J. S. Pinkner, E. Chorell, K. W. Dodson, C. L. Shaffer, M. S. Conover, J. Livny, M. Hadjifrangiskou, F. Almqvist, S. J. Hultgren, Pilicide ec240 disrupts virulence circuits in uropathogenic *Escherichia coli*. *mBio* **5**, e02038 (2014).

30. X.-R. Wu, T.-T. Sun, J. J. Medina, In vitro binding of type 1-fimbriated *Escherichia coli* to uroplakins Ia and Ib: Relation to urinary tract infections. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 9630–9635 (1996).

31. D. S. Eto, T. A. Jones, J. L. Sundsbak, M. A. Mulvey, Integrin-mediated host cell invasion by type 1–piliated uropathogenic *Escherichia coli*. *PLOS Pathog.* **3**, e100 (2007).

32. J. J. Martinez, M. A. Mulvey, J. D. Schilling, J. S. Pinkner, S. J. Hultgren, Type 1 pilus-mediated bacterial invasion of bladder epithelial cells. *EMBO J.* **19**, 2803–2812 (2000).

33. G. G. Anderson, J. J. Palermo, J. D. Schilling, R. Roth, J. Heuser, S. J. Hultgren, Intracellular bacterial biofilm-like pods in urinary tract infections. *Science* **301**, 105–107 (2003).

34. S. S. Justice, C. Hung, J. A. Theriot, D. A. Fletcher, G. G. Anderson, M. J. Footer, S. J. Hultgren, Differentiation and developmental pathways of uropathogenic *Escherichia coli* in urinary tract pathogenesis. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 1333–1338 (2004).

35. C. K. Garofalo, T. M. Hooton, S. M. Martin, W. E. Stamm, J. J. Palermo, J. I. Gordon, S. J. Hultgren, *Escherichia coli* from urine of female patients with urinary tract infections is competent for intracellular bacterial community formation. *Infect. Immun.* **75**, 52–60 (2007).

36. D. A. Rosen, T. M. Hooton, W. E. Stamm, P. A. Humphrey, S. J. Hultgren, Detection of intracellular bacterial communities in human urinary tract infection. *PLOS Med.* **4**, e329 (2007).

37. L. Robino, P. Scavone, L. Araujo, G. Algorta, P. Zunino, R. Vignoli, Detection of intracellular bacterial communities in a child with *Escherichia coli* recurrent urinary tract infections. *Pathog. Dis.* **68**, 78–81 (2013).

38. J. R. Johnson, T. T. O'Bryan, P. Delavari, M. Kuskowski, A. Stapleton, U. Carlino, T. A. Russo, Clonal relationships and extended virulence genotypes among *Escherichia coli* isolates from women with a first or recurrent episode of cystitis. *J. Infect. Dis.* **183**, 1508–1517 (2001).

39. J. R. Johnson, S. Porter, B. Johnston, M. A. Kuskowski, R. R. Spurbeck, H. L. T. Mobley, D. A. Williamson, Host characteristics and bacterial traits predict experimental virulence for *Escherichia coli* bloodstream isolates from patients with urosepsis. *Open Forum Infect. Dis.* **2**, ofv083 (2015).

40. J. R. Johnson, A. L. Stell, Extended virulence genotypes of *Escherichia coli* strains from patients with urosepsis in relation to phylogeny and host compromise. *J. Infect. Dis.* **181**, 261–272 (2000).

41. J. R. Johnson, P. Delavari, M. Kuskowski, A. L. Stell, Phylogenetic distribution of extraintestinal virulence-associated traits in *Escherichia coli*. *J. Infect. Dis.* **183**, 78–88 (2001).

42. J. W. Sahl, J. R. Sistrunk, C. M. Fraser, E. Hine, N. Baby, Y. Begum, Q. Luo, A. Sheikh, F. Qadri, J. M. Fleckenstein, D. A. Rasko, Examination of the enterotoxigenic *Escherichia coli* population structure during human infection. *mBio* **6**, e00501 (2015).

43. A. Stapleton, T. M. Hooton, C. Fennell, P. L. Roberts, W. E. Stamm, Effect of secretor status on vaginal and rectal colonization with fimbriated Escherichia coli in women with and without recurrent urinary tract infection. *J. Infect. Dis.* **171**, 717–720 (1995).

44. T. R. Hawn, D. Scholes, S. S. Li, H. Wang, Y. Yang, P. L. Roberts, A. E. Stapleton, M. Janer, A. Aderem, W. E. Stamm, L. P. Zhao, T. M. Hooton, Toll-like receptor polymorphisms and susceptibility to urinary tract infections in adult women. *PLOS ONE* **4**, e5990 (2009).

45. D. Scholes, T. M. Hooton, P. L. Roberts, A. E. Stapleton, K. Gupta, W. E. Stamm, Risk factors for recurrent urinary tract infection in young women. *J. Infect. Dis.* **182**, 1177–1182 (2000).

46. D. M. Gordon, S. E. Stern, P. J. Collignon, Influence of the age and sex of human hosts on the distribution of *Escherichia coli* ECOR groups and virulence traits. *Microbiology* **151**, 15–23 (2005).

47. E. Snyder, D. M. Gordon, D. M. Stoebel, *Escherichia coli* lacking RpoS are rare in natural populations of non-pathogens. *G3* **2**, 1341–1344 (2012).

48. M. A. Mulvey, Y. S. Lopez-Boado, C. L. Wilson, R. Roth, W. C. Parks, J. Heuser, S. J. Hultgren, Induction and evasion of host defenses by type 1-piliated uropathogenic *Escherichia coli*. *Science* **282**, 1494–1497 (1998).

49. D. J. Schwartz, S. L. Chen, S. J. Hultgren, P. C. Seed, Population dynamics and niche distribution of uropathogenic *Escherichia coli* during acute and chronic urinary tract infection. *Infect. Immun.* **79**, 4250–4259 (2011).

50. S. L. Chen, M. Wu, J. P. Henderson, T. M. Hooton, M. E. Hibbing, S. J. Hultgren, J. 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.* **5**, 184ra60 (2013).

51. B. E. Heaton, J. Herrou, A. E. Blackwell, V. H. Wysocki, S. Crosson, Molecular structure and \function of the novel BrnT/BrnA toxin-antitoxin system of *Brucella abortus*. *J. Biol. Chem.* **287**, 12098–12110 (2012).

52. S. L. Chen, C. S. Hung, J. S. Pinkner, J. N. Walker, C. K. Cusumano, Z. Li, J. Bouckaert, J. I. Gordon, S. J. Hultgren, Positive selection identifies an in vivo role for FimH during urinary tract infection in addition to mannose binding. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 22439–22444 (2009).

53. J. Parkkinen, G. N. Rogers, T. Korhonen, W. Dahr, J. Finne, Identification of the O-linked sialyloligosaccharides of glycophorin A as the erythrocyte receptors for S-fimbriated *Escherichia coli*. *Infect. Immun.* **54**, 37–42 (1986).

54. S. Szmelcman, M. Hofnung, Maltose transport in *Escherichia coli* K-12: Involvement of the bacteriophage lambda receptor. *J. Bacteriol.* **124**, 112–118 (1975).

55. Q. K. Beg, A. Vazquez, J. Ernst, M. A. de Menezes, Z. Bar-Joseph, A.-L. Barabási, Z. N. Oltvai, Intracellular crowding defines the mode and sequence of substrate uptake by *Escherichia coli* and constrains its metabolic activity. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 12663–12668 (2007).

56. M. S. Springer, M. F. Goy, J. Adler, Sensory transduction in *Escherichia coli*: Two complementary pathways of information processing that involve methylated proteins. *Proc. Natl. Acad. Sci. U.S.A.* **74**, 3312–3316 (1977).

57. S. L. Mowbray, D. E. Koshland Jr., Additive and independent responses in a single receptor: Aspartate and maltose stimuli on the tar protein. *Cell* **50**, 171–180 (1987).

58. D. W. Huang, B. T. Sherman, R. A. Lempicki, Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* **37**, 1–13 (2009).

59. D. W. Huang, B. T. Sherman, R. A. Lempicki, Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57 (2009).

60. I. U. Mysorekar, S. J. Hultgren, Mechanisms of uropathogenic *Escherichia coli* persistence and eradication from the urinary tract. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 14170–14175 (2006).

61. M. Steigedal, A. Marstad, M. Haug, J. K. Dåmås, R. K. Strong, P. L. Roberts, S. D. Himpsl, A. Stapleton, T. M. Hooton, H. L. T. Mobley, T. R. Hawn, T. H. Flo, Lipocalin 2 imparts selective pressure on bacterial growth in the bladder and is elevated in women with urinary tract infection. *J. Immunol.* **193**, 6081–6089 (2014).

62. U. Dobrindt, M. G. Chowdary, G. Krumbholz, H. Hacker, Genome dynamics and its impact on evolution of *Escherichia coli*. *Med. Microbiol. Immunol.* **199**, 145–154 (2010).

63. P. Bielecki, U. Muthukumarasamy, A. Eckweiler, A. Bielecka, S. Pohl, A. Schanz, U. Niemeyer, T. Oumeraci, N. von Neuhoff, J. M. Ghigo, S. Häussler, In vivo mRNA profiling of uropathogenic *Escherichia coli* from diverse phylogroups reveals common and group-specific gene expression profiles. *mBio* **5**, e01075–14 (2014).

64. M. C. Lane, A. N. Simms, H. L. T. Mobley, Complex interplay between type 1 fimbrial expression and flagellum-mediated motility of uropathogenic *Escherichia coli*. *J. Bacteriol.* **189**, 5523–5533 (2007).

65. A. J. Carey, C. K. Tan, D. S. Ipe, M. J. Sullivan, A. W. Cripps, M. A. Schembri, G. C. Ulett, Urinary tract infection of mice to model human disease: Practicalities, implications and limitations. *Crit. Rev. Microbiol.* **42**, 780–799 (2016).

Schreiber *et al.*, *Sci. Transl. Med.* **9**, eaaf1283 (2017)   22 March 2017

12 of 13

66. A. E. Barber, J. P. Norton, T. J. Wiles, M. A. Mulvey, Strengths and limitations of model systems for the study of urinary tract infections and related pathologies. *Microbiol. Mol. Biol. Rev.* **80**, 351–367 (2016).

67. C. A. Czaja, W. E. Stamm, A. E. Stapleton, P. L. Roberts, T. R. Hawn, D. Scholes, M. Samadpour, S. J. Hultgren, T. M. Hooton, Prospective cohort study of microbial and inflammatory events immediately preceding *Escherichia coli* recurrent urinary tract infection in women. *J. Infect. Dis.* **200**, 528–536 (2009).

68. D. R. Zerbino, E. Birney, Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res.* **18**, 821–829 (2008).

69. A. I. Rissman, B. Mau, B. S. Biehl, A. E. Darling, J. D. Glasner, N. T. Perna, Reordering contigs of draft genomes using the Mauve aligner. *Bioinformatics* **25**, 2071–2073 (2009).

70. F. Lebreton, W. van Schaik, A. M. McGuire, P. Godfrey, A. Griggs, V. Mazumdar, J. Corander, L. Cheng, S. Saif, S. Young, Q. Zeng, J. Wortman, B. Birren, R. J. L. Willems, A. M. Earl, M. S. Gilmore, Emergence of epidemic multidrug-resistant *Enterococcus faecium* from animal and commensal strains. *mBio* **4**, e00534–e13 (2013).

71. L. Salichos, A. Rokas, Evaluating ortholog prediction algorithms in a yeast model clade. *PLOS ONE* **6**, e18755 (2011).

72. A. Stamatakis, T. Ludwig, H. Meier, RAxML-III: A fast program for maximum likelihood-based inference of large phylogenetic trees. *Bioinformatics* **21**, 456–463 (2005).

73. R. C. Edgar, MUSCLE: A multiple sequence alignment method with reduced time and space complexity. *BMC Bioinf.* **5**, 113 (2004).

74. R. D. Finn, J. Tate, J. Mistry, P. C. Coggill, S. J. Sammut, H.-R. Hotz, G. Ceric, K. Forslund, S. R. Eddy, E. L. L. Sonnhammer, A. Bateman, The Pfam protein families database. *Nucleic Acids Res.* **36**, D281–D288 (2008).

75. H. Ogata, S. Goto, K. Sato, W. Fujibuchi, H. Bono, M. Kanehisa, KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **27**, 29–34 (1999).

76. E. Sefik, N. Geva-Zatorsky, S. Oh, L. Konnikova, D. Zemmour, A. M. McGuire, D. Burzyn, A. Ortiz-Lopez, M. Lobera, J. Yang, S. Ghosh, A. Earl, S. B. Snapper, R. Jupp, D. Kasper, D. Mathis, C. Benoist, Individual intestinal symbionts induce a distinct population of RORγ⁺ regulatory T cells. *Science* **349**, 993–997 (2015).

77. C.-S. Hung, K. W. Dodson, S. J. Hultgren, A murine model of urinary tract infection. *Nat. Protoc.* **4**, 1230–1243 (2009).

78. M. S. Conover, A. L. Flores-Mireles, M. E. Hibbing, K. Dodson, S. J. Hultgren, Establishment and characterization of UTI and CAUTI in a mouse model. *J. Vis. Exp.* **100**, e52892 (2015).

79. A. A. Shishkin, G. Giannoukos, A. Kucukural, D. Ciulla, M. Busby, C. Surka, J. Chen, R. P. Bhattacharyya, R. F. Rudy, M. M. Patel, N. Novod, D. T. Hung, A. Gnirke, M. Garber, M. Guttman, J. Livny, Simultaneous generation of many RNA-seq libraries in a single reaction. *Nat. Methods* **12**, 323–325 (2015).

80. H. Li, R. Durbin, Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics* **26**, 589–595 (2010).

81. S. Anders, W. Huber, Differential expression analysis for sequence count data. *Genome Biol.* **11**, R106 (2010).

82. T. J. Hannan, I. U. Mysorekar, S. L. Chen, J. N. Walker, J. M. Jones, J. S. Pinkner, S. J. Hultgren, P. C. Seed, *LeuX* tRNA-dependent and -independent mechanisms of *Escherichia coli* pathogenesis in acute cystitis. *Mol. Microbiol.* **67**, 116–128 (2008).

83. K. J. Wright, P. C. Seed, S. J. Hultgren, Uropathogenic *Escherichia coli* flagella aid in efficient urinary tract colonization. *Infect. Immun.* **73**, 7657–7668 (2005).

84. M. C. Enright, B. G. Spratt, Multilocus sequence typing. *Trends Microbiol.* **7**, 482–487 (1999).

85. M. Kearse, R. Moir, A. Wilson, S. Stones-Havas, M. Cheung, S. Sturrock, S. Buxton, A. Cooper, S. Markowitz, C. Duran, T. Thierer, B. Ashton, P. Meintjes, A. Drummond, Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**, 1647–1649 (2012).

86. K. A. Datsenko, B. L. Wanner, One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 6640–6645 (2000).

Citation: H. L. Schreiber IV, M. S. Conover, W.-C. Chou, M. E. Hibbing, A. L. Manson, K. W. Dodson, T. J. Hannan, P. L. Roberts, A. E. Stapleton, T. M. Hooton, J. Livny, A. M. Earl, S. J. Hultgren, Bacterial virulence phenotypes of *Escherichia coli* and host susceptibility determine risk for urinary tract infections. *Sci. Transl. Med.* **9**, eaaf1283 (2017).

Science Translational Medicine

Bacterial virulence phenotypes of *Escherichia coli* and host susceptibility determine risk for urinary tract infections

Henry L. Schreiber IV, Matt S. Conover, 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 and Scott J. Hultgren

Examining the many facets of *Escherichia coli*

Urinary tract infections (UTIs) are most commonly caused by uropathogenic *Escherichia coli* (UPEC). In a new study, Schreiber *et al.* undertook an interdisciplinary analysis of an extensive panel of UTI-associated *E. coli* strains. They discovered strain-dependent variation in the colonization of two mouse models of UTI. The expression of conserved bacterial behaviors, such as metabolism and motility, diverged markedly among these strains and was a better predictor of bladder colonization than carriage of any set of genes, including previously identified putative urovirulence factors. These findings suggest that UTI risk is determined by the pairing of variable host susceptibilities with bacterial virulence phenotypes governed by both gene content and expression.

ARTICLE TOOLS	http://stm.sciencemag.org/content/9/382/eaaf1283
SUPPLEMENTARY MATERIALS	http://stm.sciencemag.org/content/suppl/2017/03/20/9.382.eaaf1283.DC1
RELATED CONTENT	http://stm.sciencemag.org/content/scitransmed/9/403/eaal2717.full http://stm.sciencemag.org/content/scitransmed/9/410/eaal3693.full http://stm.sciencemag.org/content/scitransmed/10/445/eaap8423.full http://stm.sciencemag.org/content/scitransmed/10/445/eaao2586.full http://stke.sciencemag.org/content/sigtrans/11/547/eaat5750.full http://stm.sciencemag.org/content/scitransmed/11/487/eaat8418.full http://stke.sciencemag.org/content/sigtrans/12/592/eaax3938.full http://science.sciencemag.org/content/sci/369/6506/1005.full http://science.sciencemag.org/content/sci/369/6506/917.full
REFERENCES	This article cites 86 articles, 32 of which you can access for free http://stm.sciencemag.org/content/9/382/eaaf1283#BIBL
PERMISSIONS	http://www.sciencemag.org/help/reprints-and-permissions