Responses of the Human Gut Escherichia coli Population to Pathogen and Antibiotic Disturbances

Taylor K. S. Richter,a Jane M. Michalski,a Luke Zanetti,a Sharon M. Tennant,b Wilbur H. Chen,b David A. Rasko

aDepartment of Microbiology and Immunology, Institute for Genome Sciences, University of Maryland School of Medicine, Baltimore, Maryland, USA
bCenter for Vaccine Development, Department of Medicine, University of Maryland School of Medicine, Baltimore, Maryland, USA

ABSTRACT  Studies of Escherichia coli in the human gastrointestinal tract have focused on pathogens, such as diarrhea-causing enterotoxigenic E. coli (ETEC), while overlooking the resident, nonpathogenic E. coli community. Relatively few genomes of nonpathogenic E. coli strains are available for comparative genomic analysis, and the ecology of these strains is poorly understood. This study examined the diversity and dynamics of resident human gastrointestinal E. coli communities in the face of the ecological challenges presented by pathogen (ETEC) challenge, as well as of antibiotic treatment. Whole-genome sequences obtained from E. coli isolates from before, during, and after ETEC challenge were used in phylogenomic and comparative genomic analyses to examine the diversity of the resident E. coli communities, as well as the dynamics of the challenge strain, H10407, a well-studied ETEC strain (serotype O78:H11) that produces both heat-labile and heat-stable enterotoxins. ETEC failed to become the dominant E. coli clone in two of the six challenge subjects, each of whom exhibited limited or no clinical presentation of diarrhea. The E. coli communities of the remaining four subjects became ETEC dominant during the challenge but reverted to their original, subject-specific populations following antibiotic treatment, suggesting resiliency of the resident E. coli population following major ecological disruptions. This resiliency is likely due in part to the abundance of antibiotic-resistant ST131 E. coli strains in the resident populations. This report provides valuable insights into the potential interactions of members of the gastrointestinal microbiome and its responses to challenge by an external pathogen and by antibiotic exposure.

IMPORTANCE  Research on human-associated E. coli tends to focus on pathogens, such as enterotoxigenic E. coli (ETEC) strains, which are a leading cause of diarrhea in developing countries. However, the severity of disease caused by these pathogens is thought to be influenced by the microbiome. The nonpathogenic E. coli community that resides in the human gastrointestinal tract may play a role in pathogen colonization and disease severity and may become a reservoir for virulence and antibiotic resistance genes. Our study used whole-genome sequencing of E. coli before, during, and after challenge with an archetype ETEC isolate, H10407, and antibiotic treatment to explore the diversity and resiliency of the resident E. coli population in response to the ecological disturbances caused by pathogen invasion and antibiotic treatment.

KEYWORDS  Escherichia coli, human challenge, resident E. coli, variation

Escherichia coli isolates are perhaps best known as laboratory workhorses and are also known to be the etiological agents of several gastrointestinal and extraintestinal infections (1, 2). However, this species of bacteria is prevalent in the healthy
human gastrointestinal tract, reaching up to $10^9$ CFU per gram in fecal matter (3). While *E. coli* is thought to play a critical role in the prevention of pathogen colonization (3–5), relatively little is understood about the genomic diversity and ecology of the non-pathogenic isolates that are native to the gastrointestinal tract (3, 6, 7).

With a suite of diseases, including gastrointestinal and urinary tract infections and invasive disease (e.g., meningitis), resulting from *E. coli* infections, the interest in the pathogenicity of this organism is warranted (1). Enterotoxigenic *E. coli* (ETEC), one of six generally accepted *E. coli* diarrheal pathotypes (1, 2), is the causative agent in nearly 80% of cases of traveler’s diarrhea, as well as the leading cause of diarrhea in developing countries, particularly in children under the age of 5 years (8, 9). ETEC comprises *E. coli* isolates that possess plasmid-encoded enterotoxins, including heat-labile (LT) and/or heat-stable (ST) enterotoxins (1). These enterotoxins increase the amount of cAMP in intestinal epithelial cells, resulting in increased secretion of chloride ions and water from the cell and into the gut lumen, resulting in the observed diarrhea (1, 10). The majority of ETEC strains also encode an identifiable set of host-specific colonization factors (CFs) that enable attachment to the host intestinal cells (1, 11).

Relatively little is known about the genome content of the *E. coli* bacteria that reside in the human gastrointestinal tract (3, 6, 7). Whole-genome sequencing has been primarily focused on pathogenic or laboratory-derived strains, ignoring much of the potential diversity of the resident, nonpathogenic *E. coli* isolates (3). Previous studies suggested that the strains of *E. coli* in the human intestine are diverse (12), having multiple distinct genotypes identified per subject in both nonpathogenic (13–16) and pathogenic (17, 18) *E. coli* strains. The majority of the studies of resident *E. coli* have been completed using non-whole-genome assays such as multilocus sequence typing (MLST), multilocus enzyme electrophoresis, and/or serotyping (3). Additionally, most work on nonpathogenic *E. coli* has focused on isolates from single time points, leaving much to be learned about *E. coli* genomic diversity within and between human hosts over time. However, these methods do not examine samples at the complete-genome level in longitudinal samples and thus overlook the detailed dynamics of the members of the resident, nonpathogenic *E. coli* community, especially those within the healthy gut community. This study sought to address a number of these gaps in our understanding of the resident, nonpathogenic *E. coli* community. We were interested in using whole-genome sequencing to determine the impact of ecological stressors on an *E. coli* community’s genomic diversity, in particular, those stresses imparted by a closely related pathogen and by an antibiotic to which *E. coli* is generally sensitive.

A recent ETEC challenge study at the University of Maryland Center for Vaccine Development provided a unique opportunity to investigate the human gastrointestinal *E. coli* community before, during, and after ETEC challenge and antibiotic treatment. In addition to increasing the available collection of resident *E. coli* genomes, the genome sequences of multiple *E. coli* isolates, collected longitudinally, provide insights into the diversity, dynamics, and resiliency of the members of the resident *E. coli* community in the human gastrointestinal tract. Dogma suggests that during diarrheal infection with ETEC, the pathogen becomes the dominant clone(s) in the gastrointestinal tract (19, 20). Treatment of the host infected with the pathogen, often with antibiotics, is then thought to further disrupt the native gastrointestinal bacteria by reducing the number of susceptible bacteria and encouraging the restructuring of the community (19, 20). However, little is known about the members of the resident *E. coli* community prior to challenge or about their recovery from antibiotic treatment. Furthermore, variations in the prevalence of virulence and antibiotic resistance genes in this important species are examined in the context of the observed genome variation. Overall, this report serves as a useful starting point for understanding the role of resident, nonpathogenic *E. coli* in resisting and recovering from incoming pathogens such as ETEC during episodes of traveler’s diarrhea.
**RESULTS**

Clinical results show differential subject responses to challenge. Details of the challenge study and outcomes have been published previously by McArthur et al. (21). Briefly, there was an observed diarrheal attack rate of 83% (5 of 6) among those receiving the *E. coli* H10407 challenge. As shown in Table 1, two subjects (008 and 009) had severe diarrhea consisting of cumulative loose stools of ≥3 liters, one subject (015) had moderate diarrhea with cumulative loose stool of >1 liter but <3 liters, two subjects (001 and 006) had mild diarrhea with stools of >200 ml but <1 liter, and one subject (016) had no symptoms of diarrhea. These classifications of the subjects are used in the remainder of this paper. Additionally, two subjects, 004 and 019, did not receive the challenge strain and had no symptoms of diarrhea. This observed variation in clinical outcome led to questions regarding the role of the resident, nonpathogenic *E. coli* community in protection against or promotion of ETEC diarrheal diseases.

Whole-genome sequencing of *E. coli* isolates. *E. coli* isolates were obtained from extensive plating of the stool and were PCR screened for the presence of the *cfaB* gene, a marker for the *E. coli* H10407 isolate (the challenge pathogen), which encodes colonization factor antigen I subunit B and is required for virulence (22, 23). The *cfaB* gene was not identified in the isolates from the resident community. Ten colonies per subject time point had genomic DNA extracted, and the extracted DNA was sequenced on the Illumina platform. Where possible, both *cfaB*-positive (*cfaB*+) (labeled with “E” in Table S2) and *cfaB*-negative *E. coli* colonies were selected from each sample to provide insight into the pathogen and resident *E. coli* populations. The *cfaB*-positive isolates were found in subjects only from day 0 to day 4 and comprised 9.9% of the genomes obtained from subject 001, 58.3% of those obtained from subject 006, 80% of those obtained from subject 008, 100% of those obtained from subject 009, 71% of those obtained from subject 015, and 8.8% of those obtained from subject 016 (Table S2). As predicted, the subjects with moderate and severe diarrhea showed greater proportions of *cfaB*+ isolates than the subjects with no or mild diarrhea (see Fig. S1 in the supplemental material).

The assembled resident genomes had an average of 143 contigs (range, 65 to 422), an average size of 5.17 Mb (range, 4.56 to 6.32Mb), and an average of 50.73% G+C content (range, 50.32% to 50.96%). The assembled H10407-like genomes had an average of 255 contigs (range, 201 to 344), an average size of 5.30 Mb (range, 5.08 to 6.12 Mb), and an average of 50.67% G+C content (range, 50.49% to 50.78%). Details for individual genome assemblies are presented in Table S2.

**The members of the *E. coli* community show a subject-specific response to pathogen challenge and antibiotic treatment.** We used phylogenomic analyses to explore alterations in the dominant *E. coli* community throughout the course of the challenge study. The inferred phylogenomic trees represent the *E. coli* genomes from

---

**TABLE 1** The diarrheal output volume for each subject during each day of the challenge

| Subject ID | Diarrheal output vol (ml) at day postchallenge | Cumulative | Disease severity |
|------------|-----------------------------------------------|------------|-----------------|
| 016        | 0 0 0 0 0 0 0 0 0 | 0          | None            |
| 006        | 0 0 49 99 269 91 89 0 597 | 597        | Mild            |
| 001        | 0 0 146 109 186 219 0 0 660 | 660        | Mild            |
| 015        | 0 0 913 393 60 0 0 1,366 | 1,366      | Moderate        |
| 008        | 0 1,102 1,141 1,106 258 65 0 0 3,672 | 3,672      | Severe          |
| 009        | 0 567 2,043 1,066 159 88 0 0 3,923 | 3,923      | Severe          |

*a* The subjects are listed in order of increasing disease severity as follows: no diarrhea (None), mild diarrhea (defined as 2 or more loose stools of ≥200 ml within 48 h or a single loose stool of ≥300 ml), moderate diarrhea (cumulative loose stool of ≥1 liter), and severe diarrhea (cumulative loose stool of ≥3 liters). ID, identifier.
each subject and 32 reference genomes that represent archetype isolates from each of the E. coli pathotypes and Shigella species (Fig. 1; see also Table S3). These relationships confirm that the prechallenge E. coli isolates were not closely related to H10407 and that isolates closely related to H10407 were acquired during the challenge period. The resolution of the challenge with antibiotics led to a number of different patterns of genomic relatedness.

(i) Prechallenge. Prechallenge, resident communities were not expected to contain ETEC strains, as all volunteers were screened for previous exposure and did not demonstrate any signs of diarrheal illness. Overall, the genomes from the prechallenge isolates were all within phylogroup B2, supporting previous studies that suggested that urinary tract isolates, also from the B2 phylogroup, are present in the gastrointestinal tract of humans which can then be selected for in the urinary tract (24) (Fig. 1, yellow isolates; see also Table S4).

(ii) Challenge. During the challenge period, phylogenomics analyses identified H10407-like genomes among isolates from each of the subjects. The E. coli communities in most (4/6) subjects (subjects 006, 008, 009, and 015) became dominated by ETEC isolates.

---

**FIG 1** Phylogenetic analysis of isolates by subject. The whole-genome sequences of the isolates from each subject were compared with 32 previously sequenced E. coli and Shigella genomes (listed in Table S3) using a single nucleotide polymorphism (SNP)-based approach as previously described (57, 58). SNPs were detected relative to the completed genome sequence of laboratory isolate E. coli UTI89 using the *in silico* Genotyper (ISG) tool (58). A range of 171,581 to 175,765 conserved SNP sites which were present in all of the genomes analyzed were concatenated into a representative sequence for each genome. A maximum-likelihood phylogeny with 100 bootstrap replicates was inferred using RAxML v.7.2.8 (60). Phylogenetic trees of all isolates in each subject are listed in order of increasing disease severity. Isolates collected prechallenge are shown in yellow, isolates collected during challenge in red, isolates collected postchallenge in blue, reference strains in gray, and the challenge strain in black with an arrow.
H10407-like isolates during the diarrheal phase of the challenge (isolates labeled in red in Fig. 1). This was anticipated, as historical challenge studies performed with *E. coli* H10407 have resulted in individuals that excrete large volumes of diarrhea with significant CFU of the challenge pathogen (25). The inferred phylogeny confirmed that the *E. coli* communities in subjects 001 and 016 had only a limited number of H10407-like genomes and never became ETEC H10407 dominant, with the majority of the isolates belonging to the B2 phylogroup, similarly to the resident isolates in the prechallenge community. Among these subjects whose *E. coli* communities failed to reach ETEC H10407 dominance, subject 016 had no observable diarrhea and subject 001 had mild diarrhea (less than 1-liter total volume).

**iii) Treatment.** Following the administration of antibiotic treatment, we were interested in (i) the speed with which nonpathogenic *E. coli* strains regained dominance and (ii) the degree to which the resulting community reflected the previous resident community, given that resident *E. coli* strains should also be sensitive to ciprofloxacin. All subjects were cleared of ETEC-like isolates within 17 h of the initiation of antibiotic treatment (Fig. 1; see also Table S1), with a return to a resident community reflective of the prechallenge condition. This rapid restoration of the prechallenge, resident *E. coli* strains was unexpected in that time frame. The genomes from the posttreatment isolates were found to group with the same phylogenomic and pathotype clades as those identified among the prechallenge samples (Fig. 1), with a few exceptions. Subjects 008 and 009 both had isolates with genomes from the day 28 sample that are in phylogroup A along with enteroinvasive *E. coli* (EIEC) reference isolates (26) (Fig. 1, blue). While these strains were phylogenetically similar to EIEC reference strains, the genomes lacked homologs of typical EIEC and the closely related *Shigella* virulence factors (1) (Table S6). Interestingly, subject 001 had a single isolate genome that grouped most closely with *Shigella dysenteriae* 197 (phylogroup B1) from the day 21 time point. As before, the virulence profile differed from that of *Shigella dysenteriae* 197 (Table S6). Subjects at these time points demonstrated no overt clinical symptoms.

Only two subjects (006 and 008) had samples from both the prechallenge and posttreatment periods that became ETEC dominated during the challenge period, making it difficult to make direct comparisons between the *E. coli* populations present before and after challenge and treatment (Fig. 1B and E; see also Table S1). The clade of non-ETEC genomes observed in these subjects (Fig. 1B and E) showed phylogenomic similarities of the prechallenge and postantibiotic treatment isolate genomes, suggesting that the *E. coli* members of the community had returned to the original population state following antibiotic treatment in these subjects.

*E. coli* community relatedness across subjects. To investigate the relatedness of isolates across subjects at each period of the challenge study, we used phylogenomic analyses that compared isolate genomes from all subjects within a time point (Fig. 2). The phylogenomic relationships of the isolates from all prechallenge samples revealed that the *E. coli* isolates formed distinct subject-specific lineages (Fig. 2A). Despite the fact that the resident *E. coli* strains were phylogenetically closely related across subjects, each subject was found to host subject-specific strains and communities of *E. coli*.

During the challenge period (days 0 to 4), the isolates from ETEC-dominant subjects lost their subject-specific clustering, with all H10407-like isolates forming a single, indistinguishable clade regardless of the subject of origin (Fig. 2B). These challenge strain genomes cannot be phylogenetically differentiated, suggesting that the challenge isolate genomes remained conserved throughout the challenge, regardless of host. Those subjects whose *E. coli* communities did not become ETEC dominated (subjects 001 and 016; Fig. 2B) maintained their individual *E. coli* phylogroup B2 communities throughout the challenge period.

The *E. coli* isolate genomes largely returned to their pattern of subject-specific clades within phylogroup B2 following treatment with ciprofloxacin (days 4 to 8), with some notable exceptions (Fig. 2C). Of particular interest are the samples from subjects 008 and 009 collected on day 6, which were phylogenomically similar, in contrast to the
individual-subject-based groupings seen elsewhere. This occurred again on day 28, the day when both subjects were dominated by phylogroup A strains that were similar but not identical (outlined in Table S8). The posttreatment E. coli isolates from subject 009 were cultivable for only days 6 and 28 postchallenge, so no long-term pattern can be determined (Table S1).

Diversity among resident E. coli strains and relevance to disease severity. Given the observed variations in disease severity, we were interested in the relationship between diversity in the E. coli resident community and disease severity. On the basis of the diversity-stability hypothesis and the notion that the ETEC pathogen would be less likely to establish a niche in a more diverse community (27, 28), we anticipated an inverse correlation between disease severity and the diversity of the E. coli isolates identified. Although the phylogenomic analyses suggest that the resident E. coli isolates within a subject were closely related, the level of diversity within these communities was further explored using in silico multilocus sequence typing (MLST), serotyping, and gene content comparisons. As summarized in Table 2 (and detailed in Table S4), E. coli diversity was subject specific and was not fully captured by any single in silico analysis method. The most common MLST sequence type was ST131 (with serotype O25:H4), which dominated the population in all subjects except subject 015 (dominated by ST2015) and subjects 004 and 019, neither of whom completed the challenge portion of the study (Table S4).

### TABLE 2 Summary of the diversity measurements of the resident isolates from each subject, with the subjects ordered for increasing disease severity

| Parameter                  | Diversity value(s) for resident E. coli from subject: |
|----------------------------|-------------------------------------------------------|
|                            | 016  | 006  | 001  | 015  | 008  | 009  |
| No. of phylotypes          | 1    | 1    | 2    | 1    | 2    | 2    |
| No. of MLST sequence types | 4    | 5    | 6    | 3    | 5    | 2    |
| Serotype                   | 1    | 1    | 3    | 1    | 3    | 4    |
| % unique genes compared to reference | 21–30 | 34–48 | 35–46 | 21–26 | 35–48 | 24–27 |
The large-scale BLAST score ratio (LS-BSR) assay was used to analyze the gene content variation among the resident E. coli isolates as a measure of diversity among closely related strains of the same species. Resident E. coli strains from all subjects demonstrated variability in gene content differences relative to the E. coli UTI89 isolate, with the range in variability indicating genetic diversity within the community (Table 2). The isolates from subjects 001, 006, and 008 demonstrated a greater degree of genomic diversity (Table 2).

While these data demonstrate that the levels of diversity of the resident E. coli community differed across host subjects, there was no apparent correlation between the detected E. coli diversity and susceptibility to the H10407 pathogen or the observed disease state.

**Impacts of challenge and treatment on virulence and resistance gene profiles.**

After observing that the prechallenge E. coli community was restored following treatment, we were interested in the lasting impacts of challenge and treatment on the virulence and resistance potential of the genomes of the resident community. Results from previous studies performed with the challenge strain suggest that H10704 derived from a resident E. coli strain that had acquired plasmid-bound virulence genes, which suggests that the virulence plasmids can be exchanged with the resident E. coli (29). Furthermore, antibiotic use is known to contribute to the acquisition of antibiotic resistance among exposed bacteria (30, 31). We queried the data for any lingering effects of challenge and treatment by examining the canonical virulence and resistance gene profiles of the E. coli communities at each stage of the challenge study.

Results of comparisons of the gene profiles from prechallenge and posttreatment isolates mirror what was observed in the phylogenetic analyses (Fig. 3 and 4). None of the resident E. coli isolates contained canonical ETEC virulence factors, including heat-labile toxin (LT), heat-stable toxin (ST), or colonization factors (CFs), either before or after E. coli H10407 challenge, indicating that the subjects were not colonized with an ETEC isolate prior to the challenge and that transfer of the ETEC virulence factors to the resident E. coli strains did not appear to occur during the challenge period in any of the subjects. This ETEC virulence gene presence was less pronounced in subjects 001 and 016, for whom ETEC H10407 never became a dominant member of the E. coli population.

Examination of a panel of common antimicrobial resistance genes in the collection of isolates in the study showed that the resident E. coli genomes contained at least three genes or mutations that are known to result in resistance to fluoroquinolones (such as ciprofloxacin), even in the isolates collected prior to challenge and following treatment (Fig. 4). In a functional examination, ciprofloxacin resistance of the resident E. coli isolates demonstrated that pre- and postchallenge isolates from all subjects, where available, could grow on 30 μg/ml of ciprofloxacin. This resistance to fluoroquinolones, most likely due to a mutation in the gyrA and parC genes, is common among E. coli ST131 isolates (32–35). As ST131 isolates have garnered much attention for harboring drug resistance and virulence genes, we further explored the ST131 isolates from the subjects in this study. The isolates were all ST131 subclone H30R1 clade C (based on fimH typing), containing the TEM and SHV-type beta-lactamases but lacking the CTX-M extended-spectrum beta-lactamases (32).

**DISCUSSION**

This study examined the E. coli community of the human gastrointestinal tract, using genomics to explore the diversity and dynamics of this prevalent member of the gastrointestinal microbiome, particularly in response to the introduction of a closely related pathogen and treatment with antibiotics. Overall, the majority of the members of each E. coli population appeared to be subject specific, with genomes from any single subject at any time point being more closely related to those from the same subject than to those from other subjects at that or any other time point.

During the challenge study, most subjects became ETEC H10407 dominant, with the ETEC isolates overwhelming the resident, nonpathogenic E. coli strains, leading to the
loss of subject-specific communities (Fig. 2). Within 6 to 17 h after the first antibiotic dose, the H10407-like organisms were no longer identifiable by culture or genomics, leaving the previous resident *E. coli* population to recover its prechallenge population. On the basis of the genomic comparisons, it appears as though the members of the resident *E. coli* population survive or tolerate the introduction of a known pathogen at high doses in addition to subsequent antibiotic treatment and are able to reestablish dominance within the community. In each case where it was capable of being interrogated, the resident *E. coli* community reestablished dominance, and additional isolates were identified in two subjects (008 and 009) only at later time points (highlighted in Fig. 2).

Perhaps a key to the resiliency of the resident *E. coli* population is its resistance to the antibiotic ciprofloxacin. Genome analyses suggest that the resident *E. coli* genome contains genes and mutations known to confer resistance to several classes of antibiotics, including the fluoroquinolones, such as ciprofloxacin (Fig. 4). The predicted ciprofloxacin resistance was confirmed by growth of prechallenge resident *E. coli* isolates in the presence of ciprofloxacin. This resistance to the ciprofloxacin treatment was able to allow the resident population to recover following the removal of the antibiotic-sensitive *E. coli* H10407 challenge strain. The resident populations in most subjects (001, 006, 008, 009, and 016) were dominated by the isolates of ST131.

**FIG 3** Virulence gene presence and absence. The proportions of isolate genomes containing virulence factors in each subject and time period are presented as a heat map. Gene sequences of known *E. coli* virulence genes were queried against the isolate genomes using LS-BSR. Genes with a significant sequence match (BSR > 0.8) were deemed present, and the total number of present virulence genes is presented as a proportion of the total number of isolates in that time period (in parentheses). Genes involved in ETEC toxin (LT and ST) production are highlighted in bold. Other listed genes are involved in ETEC adhesion or virulence in other *E. coli* pathotypes, with details in Table S5. BSR assay results from individual isolates for each investigated virulence factor gene are presented in Table S6.
Members of this sequence type are well known for their resistance to antibiotics, particularly fluoroquinolones and beta-lactams, and are common members of resident gut communities throughout the developing world (33, 34, 36, 37). While we are unable to say for certain why the subjects in this study were colonized by ST131, studies have demonstrated high rates of colonization among people who have traveled to countries of endemicity or have stayed in long-term health care facilities and that CTX-M-001 (10)
006 (10)
008 (4)
009 (4)
015 (10)
016 (16)
001 (92)
006 (40)
008 (65)
009 (60)
015 (69)
016 (68)
001 (74)
006 (65)
008 (24)
016 (68)

Pre-challenge Challenge Post-challenge

Aminoglycoside
APH (3’)-Ib
APH (6)-Id
AadA5
KdpE

LRA-18-like
SHV-75-like
TEM-1
TEM-171

Diaminopyrimidine
DfrA17

Fluoroquinolone
ParC*
Mfd
GyrA*

Fosfomycin
GlpT*
UhpT*

Macrolide
MphA
Mrx

Peptide
BacA

Phenicol
CatI

Polymixin
ArnA
PmrC
PmrE
PmrF

Sulfonamide
LeuO
Sul1
Sul2

Proportion of gene presence

FIG 4 Proportions of antimicrobial resistance genes associated with isolates in each phase of study. The proportions of isolate genomes containing predicted antibiotic resistance in each subject and time period are presented as a heat map. Protein sequences known to confer *E. coli* antibiotic resistance were queried against the translated isolate genomes using CARD-RGI. The total number of resistant protein sequences present is presented as a proportion of the total number of isolates in that time period (in parentheses), with multiple copies in some isolates. The resistance sequences are listed by the category of antibiotic to which they provide resistance. Those sequences marked with an asterisk (*) provided antibiotic resistance due to a sequence mutant or variation (detailed in Table S7).
containing ST131 clones have spread rapidly in the United States, where it is now a dominant lineage (35, 38–41). Furthermore, ST131 is an efficient and prolonged colonizer, known to outcompete other resident strains and predicted to persist in the gastrointestinal tract for years (40, 42, 43), which may also help explain the readiness with which the resident *E. coli* population recovered following challenge and treatment.

Upon observing the variability in clinical responses to a controlled pathogen input, we were interested in exploring the role of the resident *E. coli* community in preventing or enhancing the disease severity of an introduced, virulent *E. coli* strain. While earlier studies on diarrheagenic bacteria, such as the Global Enteric Multicenter Study (GEMS) (8, 9, 25), described the detection of pathogens in the absence of clinical symptoms, they lacked the opportunity to control for a common, known, and quantified pathogenic input such as was included in the design of this challenge study. While we are unable to make sweeping generalizations on the basis of the limited number of subjects, we can conclude that human disease severity is the result of the presence of a bacterium encoding virulence genes in addition to host factors (genetics, immune system status, nutrition status, etc.) and/or the microbiome. Additional studies examining the immune status of these subjects (21) and the gastrointestinal microbiome from this cohort (Richter et al., unpublished results) integrated with this genome-based study will provide a systems view of the host-pathogen interactions. Here we specifically explored the role of the resident *E. coli* community in ETEC disease severity, with the idea that ETEC would be less likely to find a necessary niche among the members of a more diverse resident *E. coli* community such as has been described in animal models (44–46).

Differential susceptibilities to ETEC colonization were observed, in that ETEC H10407 did not become a dominant community member in some subjects (001 and 016). Unsurprisingly, diarrhea was either mild or nonexistent in each of those cases (Fig. 1; see also Table 1). On the opposite end of the spectrum, the subjects with the most severe clinical presentation as determined on the basis of diarrheal output (subjects 008 and 009) demonstrated the greatest degree of instability in their *E. coli* populations, where the dominant *E. coli* strains shifted between phylogroups B2 and A three times in the course of the study (Fig. 1). While the examined *E. coli* populations differed between subjects in their degrees of genetic diversity, the diversity of these populations was not correlated with the observed stability and determined neither susceptibility to ETEC colonization nor disease severity (Table 2).

The current study afforded a unique opportunity to longitudinally study both pathogenic and resident *E. coli* strains in the human gastrointestinal tract over the duration of the challenge study. In addition to providing new genomes for the study of the diversity of nonpathogenic *E. coli* strains, these data demonstrate the resiliency of the *E. coli* community in response to extreme ecological disturbances, namely, pathogen and antibiotic introduction. This report serves as a useful starting point for understanding the role of *E. coli* within the larger bacterial community of the human gastrointestinal tract in comparison to examining only the pathogen in isolation. Understanding these interactions between pathogen and resident nonpathogen will allow us to potentially exploit those nonpathogens as part of a therapy for resisting incoming pathogens such as ETEC during traveler’s diarrhea.

**MATERIALS AND METHODS**

**Challenge conditions and stool sample collection.** Fecal samples were obtained from adult volunteers participating in an enterotoxigenic Escherichia coli (ETEC) challenge study performed by the Center for Vaccine Development at the University of Maryland School of Medicine in Baltimore, MD (21). The study was approved by the Institutional Review Board of University of Maryland, Baltimore (UMB). Written informed consent was obtained from healthy adult volunteers 18 to 49 years of age, who were screened for the absence of chronic medical conditions and immunodeficiencies. Participants were excluded if they had received antibiotics in the 2 weeks prior to the study or if they had had previous exposure to ETEC or *Vibrio cholerae*.

After overnight fasting, subjects ingested 10⁸ CFU of challenge strain *E. coli* H10407 (29) and were observed in an inpatient research isolation ward, where they were closely monitored for signs of diarrheal illness. As has been previously described (47), each stool was graded as follows: grade 1, firm;
grade 2, soft; grade 3, thick liquid; grade 4, opaque watery; grade 5, rice water. All stools of grade 3 were considered loose, and the volume was measured (see Table S1 in the supplemental material). Any individual who developed loose stool was offered an oral rehydration salts (ORS) solution (Jianas Brothers, Kansas City, MO) or intravenous lactated Ringer’s solution at a volume 1.5 times the loose stool volume. Ciprofloxacin was administered (500 mg twice daily for 5 days) to any subject who exceeded 3 liters of cumulative loose stool output or on day 4 postchallenge, whichever occurred first. Individuals were discharged when they were asymptomatic, completed a course of ciprofloxacin therapy, and demonstrated 3 sequential stool cultures (separated by 12 h) that were negative for \textit{E. coli}.

The sample collection and challenge timeline is shown in Table S1. Subjects also provided fecal samples following discharge from the challenge facility on days 14, 21, and 28 postchallenge. Six subjects (designated 001, 006, 008, 009, 015, and 016) participated in the full challenge study, and an additional two (004 and 019) provided only prechallenge samples. A flow chart of the sample preparation and processing is shown in Fig. 5.

\textbf{E. coli colony isolation and pathogen screen.} To isolate \textit{E. coli} from the stool, samples were streaked for single colonies onto MacConkey or eosin methylene blue (EMB) media. Single colony isolates positive for lactose utilization were picked from each stool sample, inoculated into a 96-well microtiter plate containing Miller’s LB broth (Research Products International), supplemented with 20% glycerol (final), and stored at –80°C as stock isolates.

Three biochemical tests, Voges-Proskauer (VP) (48), lysine decarboxylase (49), and indole (50), were used to confirm the evidence of the presence of enteric bacteria and to differentiate \textit{E. coli} from other \textit{Enterobacteriaceae} (51). The presumptive \textit{E. coli} isolates were grown (i) in methyl red-VP (MR-VP) medium (Difco) and tested for acetoin production using Voges-Proskauer reagent (BioMérieux) according to the manufacturer’s instructions, (ii) on decarboxylase medium base (Difco) supplemented with L-lysine.
(Sigma) to test for l-lysine utilization, and (iii) in L-broth to test for tryptophan utilization using Remel indole reagent (Thermo Scientific) according to the manufacturer’s instructions. Lactose-positive isolates on MacConkey plates that were VP negative, lysine decarboxylase positive, and indole positive were identified as E. coli (Fig. 5).

The isolates were further tested with a PCR assay for the colonization factor antigen gene B (cfaB) gene (52, 53) (using forward primer 5’-GCTTATTCTCCGCATCAAAAAC and reverse primer 5’-TTACACC CGATGCAGAATATC) from enterotoxigenic E. coli H10407 (29) to distinguish the input pathogen from the resident E. coli strains, which should lack cfaB.

**DNA extraction, sequencing, and assembly.** Bacterial genomic DNA was purified from 10 presumptive E. coli isolates from each stool sample using a GenElute bacterial genomic DNA kit (Sigma) according to the instructions of the manufacturer and subjected to whole-genome sequencing. Where possible, both cfaB-positive isolates and cfaB-negative isolates were analyzed from each stool sample. DNA was sequenced on the Illumina HiSeq platform at the Genome Resource Center at the University of Maryland School of Medicine, Institute for Genome Sciences (http://www.igs.umaryland.edu/resources/grc/). The resulting 100-bp reads were assembled as previously described (54, 55) using the Maryland Super-Read Celera Assembler (MasuLCA version 2.3.2) (56). Contigs of fewer than 200 bp were excluded from assemblies. Genomes containing greater than 500 contigs (25 genomes total) were excluded from further analysis. The assembly details and corresponding GenBank accession numbers are provided in Table S2.

**Phylogenomic analyses.** Phylogenomic trees were inferred using assembled genomes from this study in addition to 32 E. coli and Shigella reference genomes from GenBank (Table S3) (57). Single nucleotide polymorphisms (SNPs) in all genomes were detected relative to the completed genome sequence of ETEC isolate E. coli H10407 using the in silico Genotypyer (ISG) v.0.12.2 tool (58), which uses MUMmer v.3.22 (59) for SNP detection. These SNP sites were concatenated and used for phylogenetic analysis as previously described (17). A maximum-likelihood phylogeny was generated with RAxML v.7.2.8 (60) using the GTR model of nucleotide substitution with the gamma model of rate heterogeneity and 100 bootstrap replicates and was visualized using FigTree v.1.4.2 (http://tree.bio.ed.ac.uk/software/figtree/).

**Serotype identification.** In silico serotype identification was performed on the assembled genomes using the online Center for Genomic Epidemiology SerotypeFinder 1.1 tool (https://cge.cbs.dtu.dk/services/SerotypeFinder/) (61, 62).

**Multilocus sequence typing (MLST).** In silico MLST was performed on the assembled genomes using the Achtman E. coli MLST scheme (63). Gene sequences were identified in the isolate genomes using BLASTn, and MLST profiles were determined by querying the PubMLST database (http://pubmlst.org).

**Typing of the fimH gene in the ST131 isolates.** In silico fimH typing was performed on the assembled genomes from ST131 isolates using the online Center for Genomic Epidemiology FimTyper 1.0 tool (https://cge.cbs.dtu.dk/services/FimTyper/) (64).

**Variation in gene distributions as a measurement of diversity of resident E. coli strains.** The gene contents across all genomes were identified and compared using the large-scale BLAST score ratio (LS-BSR) as previously described (65). Genes with a BSR value of ≥0.80 are considered to be highly conserved and present in the isolate examined. Those genes that were conserved in all genomes were removed from further analyses. The predicted protein function of each gene cluster was determined using an ergatis-based (66) in-house annotation pipeline (67). The data presented as heat maps were generated in MeV (Multi-Experiment Viewer) (68).

Results from LS-BSR assays of the resident E. coli isolates (523 of the 820 total genome isolates, as determined by phylogenomics and cfaB-negative PCR) were compared to those from reference strain E. coli UTI89 (a reference strain closely related to the resident isolates) to determine the differences between genomes with respect to the number of genes present or absent. These difference values were divided by the total number of the genes present in the query genome and multiplied by 100 to give the percentage of deviation between the query genome and the reference genome. The variation in these percentages within genomes from a single subject represents the relative level of genetic diversity from within each subject.

**Virulence factor and antibiotic resistance gene identification.** The list of common E. coli virulence factor genes used for interrogation of the study genomes is shown in Table S4. Antibiotic resistance genes were identified in the isolate genomes using the Comprehensive Antibiotic Resistance Database (version 1.1.8; http://arpcard.mcmaster.ca) Resistance Gene Identifier (version 3.2.0) (CARD-RGI) and strict and perfect cutoff values. An average number of genes per time period was calculated by dividing the total number of genes identified in all of the genomes of isolates within a time period but from a single subject by the total number of isolates from that time period. These averages are presented as a heat map generated in MeV (Multi-Experiment Viewer) (68).

**Confirmation of ciprofloxacin resistance.** Five isolates from each subject were randomly chosen from day −1 (prechallenge) and day 6 (postchallenge). Following overnight growth on LB agar at 37°C, the colonies were transferred to LB agar containing ciprofloxacin (30 µg/µl) and allowed to grow overnight at 37°C. Positive growth on ciprofloxacin was indicated by the presence of at least 10 colonies following overnight growth.

**Accession number(s).** GenBank accession numbers for the sequences determined in this work are provided in Table S2.
SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at https://doi.org/10.1128/mSystems.00047-18.

FIG S1, PDF file, 0.3 MB.
FIG S2, PDF file, 8.4 MB.
TABLE S1, PDF file, 0.1 MB.
TABLE S2, PDF file, 0.2 MB.
TABLE S3, PDF file, 0.04 MB.
TABLE S4, PDF file, 0.2 MB.
TABLE S5, PDF file, 0.1 MB.
TABLE S6, PDF file, 0.3 MB.
TABLE S7, PDF file, 0.04 MB.
TABLE S8, PDF file, 0.1 MB.

ACKNOWLEDGMENTS

We thank Lisa Chrisley, Melissa Billington, Xiaolin Wang, Joanna Becker, Robin Barnes, Jennifer Courneya, Naya Komninou, Nancy Greenberg, Takiyah Crawford, Sofie Livio, Sunil Sen, Nurulla Sayed, Alyson Kwon, and Brenda Dorsey for helping to conduct the challenge study at the Center for Vaccine Development at the University of Maryland School of Medicine.

This project was funded in part by federal funds from the National Institute of Allergy and Infectious Diseases, National Institutes of Health, Department of Health and Human Services, under grant U19 AI10820 and by the Department of Defense, Defense Advanced Research Projects (DARPA; W911NF-13-0346).

REFERENCES

1. Kaper JB, Nataro JP, Mobley HL. 2004. Pathogenic Escherichia coli. Nat Rev Microbiol 2:123–140. https://doi.org/10.1038/nrmicro818.
2. Croxon MA, Law RJ, Scholz R, Keeney KM, Wlodarska M, Finlay BB. 2013. Recent advances in understanding enteric pathogenic Escherichia coli. Clin Microbiol Rev 26:822–880. https://doi.org/10.1128/CMR.00022-13.
3. Tenaillon O, Skurnik D, Picard B, Denamur E. 2010. The population genetics of commensal Escherichia coli. Nat Rev Microbiol 8:207–217. https://doi.org/10.1038/nrmicro2298.
4. Hudault S, Guignot J, Servin AL. 2001. Intestinal commensal Escherichia coli strains colonising the gastrointestinal tract protect germfree mice against Salmonella typhimurium infection. Gut 49:47–55. https://doi.org/10.1136/gut.49.1.47.
5. Schambert-Garnier LP, Phillips RL, Jacobs JL, Diez-Gonzalez F. 2004. Reduction of Escherichia coli O157:H7 populations in cattle by addition of colicin E7-producing E. coli to feed. Appl Environ Microbiol 70:6053–6060. https://doi.org/10.1128/AEM.70.6053-6060.2004.
6. Escobar-Páramo P, Le Menac’h A, Le Gall T, Amorin C, Gouriou S, Picard B, Skurnik D, Denamur E. 2006. Identification of forces shaping the commensal Escherichia coli genetic structure by comparing animal and human isolates. Environ Microbiol 8:1975–1984. https://doi.org/10.1111/j.1462-2920.2006.0077x.
7. Lescat M, Clermont O, Woerther PL, Glodt J, Dion S, Skurnik D, Djossou F, Dupont C, Perroz G, Picard B, Catzeflis F, Andremont A, Denamur E. 2013. Commensal Escherichia coli strains in Guiana reveal a high genetic diversity with host-dependent population structure. Environ Microbiol Rep 5:49–57. https://doi.org/10.1111/j.1758-2229.2012.00374.x.
8. Kotloff KL, Blackwelder WC, Nasrin D, Nataro JP, Farag TH, van Eijk A, Adgbola RA, Alonso PL, Breiman RF, Faruque AS, Saha D, Sob SO, Sur D, Zaidi AK, Skurnik D, Panchalingam S, Clemens JD, Cohen D, Glass RI, Mintz ED, Sommerfelt H, Levine MM. 2012. The Global Enteric Multicenter Study (GEMS): a prospective, case-control study. Lancet 382:209–222. https://doi.org/10.1016/S0140-6736(13)60844-2.
9. Kotloff KL, Nataro JP, Breiman RF, Gilmour M, Nataro JP, Rasko DA. 2015. Defining the phylogenetic diversity of enterotoxigenic Escherichia coli. mBio 6:e00501-15.
10. Sears GL, Kaper JB. 1996. Enteric bacterial toxins: mechanisms of action and linkage to intestinal secretion. Microbiol Rev 60:167–215.
11. Wolf MK. 1997. Occurrence, distribution, and associations of O and H serogroups, colonization factor antigens, and toxins of enterotoxigenic Escherichia coli. Clin Microbiol Rev 10:569–584.
12. Caugant DA, Levin BR, Selander RK. 1981. Genetic diversity and temporal variation in the E. coli population of a human host. Genetics 98:467–490.
13. Gordon DM, Bauer S, Johnson JR. 2002. The genetic structure of Escherichia coli populations in primary and secondary habitats. Microbiology 148:1513–1522. https://doi.org/10.1099/mic.0.021287-148-1513.
14. Gordon DM, O’Brien CL, Pavli P. 2015. Escherichia coli diversity in the lower intestinal tract of humans. Environ Microbiol Rep 7:642–648. https://doi.org/10.1111/1758-2229.12300.
15. Sears HJ, Brownlee I. 1952. Further observations on the persistence of individual strains of Escherichia coli in the intestinal tract of man. J Bacteriol 63:47–57.
16. Sears HJ, Brownlee I, Uchiyama JK. 1950. Persistence of individual strains of Escherichia coli in the intestinal tract of man. J Bacteriol 59:293–301.
17. Sahil JW, Morris CR, Emberger J, Fraser CM, Ochieng JB, Juma J, Fields B, Breiman RF, Gilmour M, Nataro JP, Rasko DA. 2015. Defining the phylogenomics of Shigella species: a pathway to diagnostics. J Clin Microbiol 53:951–960. https://doi.org/10.1128/JCM.03527-14.
18. Sahil JW, Sistrunk JR, Fraser CM, Hine E, Baby N, Begum Y, Luo Q, Sheikh A, Qadri F, Fleckenstein JM, Rasko DA. 2015. Examination of the enterotoxigenic Escherichia coli population structure during human infection. mBio 6:e00501-15. https://doi.org/10.1128/mBio.00501-15.
19. Gonzales-Siles L, Sjöling Å. 2016. The different ecological niches of enterotoxigenic Escherichia coli. Environ Microbiol 18:741–751. https://doi.org/10.1111/1462-2920.13106.
20. David LA, Weil A, Ryan ET, Calderwood SB, Harris JB, Chowdhury F, Begum Y, Qadri F, LaRocque RC, Turnbaugh PJ. 2015. Gut microbial
succession follows acute secretory diarrhea in humans. MBio 6:e00381-15. https://doi.org/10.1128/mBio.00381-15.

21. McArthur MA, Chen WH, Magder L, Levine MM, Szein MB. 2017. Impact of CD4 + T cell responses on clinical outcome following oral administration of wild-type enterotoxic Escherichia coli in humans. PLoS Negl Trop Dis 11:e0005291. https://doi.org/10.1371/journal.pntd.0005291.

22. Gaaster W, Svennerholm AM. 1996. Colonization factors of human enterotoxic Escherichia coli (ETEC). Trends Microbiol 4:444–452. https://doi.org/10.1016/0966-842X(96)01068-5.

23. Evans DG, Silver RP, Evans DJ, Jr., Chase DG, Gorbach SL. 1975. Plasmid-controlled colonization factor associated with virulence in Escherichia coli enterotoxic for humans. Infect Immun 12:656–667.

24. Chen SL, Hung CS, Xu J, Reigstad CS, Magrini V, Sabo A, Blasiar D, Bieri T, Meyer RR, Özersky P, Armstrong JT, Frunston RS, Laterille JP, Speith J, Hooton TM, Mardis ER, Hultgren SJ, Gordon JL. 2006. 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. https://doi.org/10.1073/pnas.0600938103.

25. Porter CK, Riddle MS, Trible DR, Louis Bougeois A, McKenzie R, Isidean SD, Sebeny P, Savarino SJ. 2011. A systematic review of experimental infections with enterotoxigenic Escherichia coli (ETEC). Vaccine 29:5869–5885. https://doi.org/10.1016/j.vaccine.2011.05.021.

26. Campanile-Francino R, Francino ML, Marconcini PM, Martorana P, Leuzzi A, Grossi M, Casalino M, Niccolleti M, Micheli G, Colonna B, Prossega G. 2014. Molecular and functional profiling of the polyamine content in enteroinvasive E. coli: looking into the gap between commensal E. coli and harmful Shigella. PLoS One 9:e106589. https://doi.org/10.1371/journal.pone.0106589.

27. McCann KS. 2000. The diversity-stability debate. Nature 405:228–233. https://doi.org/10.1038/35012234.

28. Shade A, Peter H, Allison SD, Baho DL, Berga M, Bürgmann H, Huber DH, Langenheder S, Lennon JT, Martiny JB, Matulich KL, Schmid TM, Handelsman J. 2012. Fundamentals of microbial community resistance and resilience. Front Microbiol 3:417. https://doi.org/10.3389/fmicb.2012.00417.

29. Crossman LC, Chaudhuri RR, Beatson SA, Wells TJ, Desvaux M, Cunningham AF, Smith SG, Henderson IR. 2010. A commensal gone bad: complete genome of a commensal strain H10407. J Bacteriol 192:5822–5831. https://doi.org/10.1128/JB.00710-Hy.

30. Davies J, Davies D. 2010. Origins and evolution of antibiotic resistance. Microbiol Mol Biol Rev 74:417–433. https://doi.org/10.1128/MMBR.00016-10.

31. Francino MP. 2015. Antibiotics and the human gut microbiome: dysbioses and accumulation of resistances. Front Microbiol 6:1543. https://doi.org/10.3389/fmicb.2015.01543.

32. Pitout JD, DeVinney R. 2017. Escherichia coli ST131: a multidrug-resistant urinary tract infection abilities. PLoS One 7:e46547. https://doi.org/10.1371/journal.pone.0046547.

33. Sarkar S, Hutton ML, Vagenas D, Ruter R, Schüller S, Lysas D, Schembri MA, Johnson JR. 2018. Intestinal colonization traits of Pandemic multidrug-resistant Escherichia coli ST131. J Infect Dis https://doi.org/10.1093/infdis/jiy031.

34. Meador JP, Caldwell ME, Cohen PS, Conway T. 2014. Escherichia coli pathotypes occupy distinct niches in the mouse intestine. Infect Immun 82:1931–1938. https://doi.org/10.1128/IAI.01435-13.

35. Conway T, Cohen PS. 2015. Commensal and pathogenic Escherichia coli metabolism in the gut. Microbiol Spectr 3. https://doi.org/10.1128/microbiolspec.MBP-0006-2014.

36. Fabich AJ, Jones SA, Chowdhury FZ, Cernosek A, Anderson A, Smalley DL, Mclaurage JW, Hightower GA, Smith JT, Autieri SM, Leatham MP, Lins JJ, Allen RL, Laux DC, Cohen PS, Conway T. 2008. Comparison of carbon nutrition for pathogenic and commensal Escherichia coli strains in the mouse intestine. Infect Immun 76:1143–1152. https://doi.org/10.1128/IAI.00816-07.

37. Sarkk D, Tacket CO, Cohen MB, Sack RR, Losonsky GA, Simkio J, Nato JP, Edelman R, Levine MM, Giannastra RR, Schif G, Lang D. 1998. Validation of a volunteer model of cholera with frozen bacteria as the challenge. Infect Immun 66:1968–1972.

38. West FD. 1909. Notes on the Voges and proskauer reaction. Am J Public Health 19:227–230.

39. West FD. 1955. Simplified tests for some amino acid decarboxylases and for the arginine dihydrolase system. Acta Pathol Microbiol Scand 36:158–172. https://doi.org/10.1111/j.1699-0463.1955.tb04583.x.

40. Lowrance BL, Reich P, Traub WH. 1969. Evaluation of two spot-indole reagents. Appl Microbiol 17:923–924.

41. Bergéy DH, Holt JG. 1994. Bergey’s manual of determinative bacteriology, 9th ed. Lippincott Williams & Wilkins, Baltimore, MD.

42. Perez-Casal J, Swartley JS, Scott JR. 1990. Gene encoding the major subunit of C51 pili of human enterotoxigenic Escherichia coli. Infect Immun 58:3594–3600.

43. Jordi BJ, Willshaw GA, van der Zeijst BA, Gaastra W. 1992. The complete nucleotide sequence of region 1 of the CFA/I fimbrial operon of human enterotoxigenic Escherichia coli. DNA Seq 3:158–172. https://doi.org/10.1016/0169-4442(92)80039-9.

44. Allen RL, Laux DC, Cohen PS, Conway T. 2008. Comparison of carbon nutrition for pathogenic and commensal Escherichia coli strains in the mouse intestine. Infect Immun 76:1143–1152. https://doi.org/10.1128/IAI.00816-07.

45. Richter et al. 2018. Voluntary. Antimicrobial Agents Chemother 54:3564–3568. https://doi.org/10.1128/AAC.00230-15.

46. Peirano G, Laupland KB, Gregson DB, Pitout JD. 2011. Colonization of returning travelers with CTX-M-producing Escherichia coli. J Travel Med 18:299–303. https://doi.org/10.1111/j.1708-8305.2010.00548.x.

47. Overdevest I, Haverkate M, Veennemans J, Hendriks Y, Verhulst C, Mulders A, Cuppie W, Bootsma M, Johnson J, Kluytmans J. 2016. Prolonged colonisation with Escherichia coli O25b:ST131 versus other extended-spectrum beta-lactamase-producing E. coli in a long-term care facility with high endemic level of rectal colonisation, the Netherlands, 2013 to 2014. Euro Surveill 21:30376. https://doi.org/10.2807/1560-7917.ES.2014.21.42.30376.

48. Johnson JR, Porter S, Thuras P, Castanheira M. 25 July 2017. Epidemic emergence in the United States of Escherichia coli sequence type 131-H30 (ST131-H30) 2000 to 2009. Antimicrobial Agents Chemother https://doi.org/10.1128/AAC.00732-17.

49. Vimont S, Boyd A, Bleibtreu A, Bens M, Goujon JM, Garry L, Clermont O, Denamur E, Arlet G, Vandewalle A. 2012. The CTX-M-15-producing Escherichia coli clone O25b: H4-ST131 has high intestine colonization and urinary tract infection abilities. PLoS One 7:e46547. https://doi.org/10.1371/journal.pone.0046547.

50. Bergey DH, Holt JG. 1994. Bergey’s manual of determinative bacteriology, 9th ed. Lippincott Williams & Wilkins, Baltimore, MD.

51. Bergey DH, Holt JG. 1994. Bergey’s manual of determinative bacteriology, 9th ed. Lippincott Williams & Wilkins, Baltimore, MD.
58. Sahl JW, Beckstrom-Sternberg SM, Babic-Sternberg JS, Gillece JD, Hepp CM, Auerbach RK, Tembe W, Wagner DM, Keim PS, Pearson T. 2015. The in silico Genotyper (ISG): an open-source pipeline to rapidly identify and annotate nucleotide variants for comparative genomics applications. bioRxiv https://www.biorxiv.org/content/early/2015/02/20/015578.

59. Delcher AL, Salzberg SL, Phillippy AM. 2003. Using MUMmer to identify similar regions in large sequence sets. Curr Protoc Bioinformatics Chapter 10:Unit 10.3. https://doi.org/10.1002/0471250953.bi1003s00.

60. Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22:2688–2690. https://doi.org/10.1093/bioinformatics/btl446.

61. Joensen KG, Tetzschner AM, Iguchi A, Aarestrup FM, Scheutz F. 2015. Rapid and easy in silico serotyping of Escherichia coli isolates by use of whole-genome sequencing data. J Clin Microbiol 53:2410–2426. https://doi.org/10.1128/JCM.00008-15.

62. Kuzevski A, Tauschek M, Inouye M, Stinear T, Levine MM, Robins-Browne RM, Holt KE. 2016. In silico serotyping of E. coli from short read data identifies limited novel O-loci but extensive diversity of O:H serotype combinations within and between pathogenic lineages. Microb Genom 2:e000064. https://doi.org/10.1099/mgen.0.000064.

63. Wirth T, Falush D, Lan R, Colles F, Mensa P, Wieler LH, Karch H, Reeves PR, Maiden MC, Ochman H, Achtman M. 2006. Sex and virulence in Escherichia coli: an evolutionary perspective. Mol Microbiol 60:1136–1151. https://doi.org/10.1111/j.1365-2958.2006.05172.x.

64. Roer L, Tchesnokova V, Allesøe R, Muradova M, Chattopadhyay S, Ahrenfeldt J, Thomsen MCF, Lund O, Hansen F, Hammerum AM, Sokurenko E, Hasman H. 2017. Development of a web tool for Escherichia coli subtyping based on fimH alleles. J Clin Microbiol 55:2538–2543. https://doi.org/10.1128/JCM.00737-17.

65. Sahl JW, Caporaso JG, Rasko DA, Keim P. 2014. The large-scale blast score ratio (LS-BSR) pipeline: a method to rapidly compare genetic content between bacterial genomes. PeerJ 2:e332. https://doi.org/10.7717/peerj.332.

66. Orvis J, Crabtree J, Galens K, Guissman A, Inman JM, Lee E, Nampally S, Riley D, Sundaram JP, Felix V, Whitty B, Mahurkar A, Wortman J, White O, Angiuoli SV. 2010. Ergatis: a web interface and scalable software system for bioinformatics workflows. Bioinformatics 26:1488–1492. https://doi.org/10.1093/bioinformatics/btq167.

67. Galens K, Orvis J, Daugherty S, Creasy HH, Angiuoli S, White O, Wortman J, Mahurkar A, Giglio MG. 2011. The IGS standard operating procedure for automated prokaryotic annotation. Stand Genomic Sci 4:244–251. https://doi.org/10.4056/sgs.1223234.

68. Howe EA, Sinha R, Schlauch D, Quackenbush J. 2011. RNA-Seq analysis in MeV. Bioinformatics 27:3209–3210. https://doi.org/10.1093/bioinformatics/btr490.