Single Clinical Isolates from Acute Uncomplicated Urinary Tract Infections Are Representative of Dominant *In Situ* Populations

Dana Willner,[a,b] Serene Low,[a] Jason A. Steen,[a] Narelle George,[c] Graeme R. Nimmo,[c] Mark A. Schembri,[d] Philip Hugenholtz[a,d,e]

Australian Centre for Ecogenomics, School of Chemistry and Molecular Biosciences, The University of Queensland, St. Lucia, QLD, Australia[a]; Diamantina Institute, The University of Queensland, St. Lucia, QLD, Australia[b]; Pathology Queensland Central Laboratory, Herston, Brisbane, QLD, Australia[c]; Australian Infectious Diseases Research Centre, School of Chemistry and Molecular Biosciences, The University of Queensland, St. Lucia, QLD, Australia[d]; Institute for Molecular Bioscience, The University of Queensland, St. Lucia, QLD, Australia[e]

D.W. and S.L. contributed equally to this work.

ABSTRACT **Urinary tract infections (UTIs) are one of the most commonly acquired bacterial infections in humans, and uropathogenic *Escherichia coli* strains are responsible for over 80% of all cases. The standard method for identification of uropathogens in clinical laboratories is cultivation, primarily using solid growth media under aerobic conditions, coupled with morphological and biochemical tests of typically a single isolate colony. However, these methods detect only culturable microorganisms, and characterization is phenotypic in nature. Here, we explored the genotypic identity of communities in acute uncomplicated UTIs from 50 individuals by using culture-independent amplicon pyrosequencing and whole-genome and metagenomic shotgun sequencing. Genus-level characterization of the UTI communities was achieved using the 16S rRNA gene (V8 region). Overall UTI community richness was very low in comparison to other human microbiomes. We strain-typed *Escherichia*-dominated UTIs using amplicon pyrosequencing of the fimbrial adhesin gene, *fimH*. There were nine highly abundant *fimH* types, and each UTI sample was dominated by a single type. Molecular analysis of the corresponding clinical isolates revealed that in the majority of cases the isolate was representative of the dominant taxon in the community at both the genus and the strain level. Shotgun sequencing was performed on a subset of eight *E. coli* urine UTI and isolate pairs. The majority of UTI microbial metagenomic sequences mapped to isolate genomes, confirming the results obtained using phylogenetic markers. We conclude that for the majority of acute uncomplicated *E. coli*-mediated UTIs, single cultured isolates are diagnostic of the infection.**

IMPORTANCE **In clinical practice, the diagnosis and treatment of acute uncomplicated urinary tract infection (UTI) are based on analysis of a single bacterial isolate cultured from urine, and it is assumed that this isolate represents the dominant UTI pathogen. However, these methods detect only culturable bacteria, and the existence of multiple pathogens as well as strain diversity within a single infection is not examined. Here, we explored bacteria present in acute uncomplicated UTIs using culture-independent sequence-based methods. *Escherichia coli* was the most common organism identified, and analysis of *E. coli* dominant UTI samples and their paired clinical isolates revealed that in the majority of infections the cultured isolate was representative of the dominant taxon at both the genus and the strain level. Our data demonstrate that in most cases single cultured isolates are diagnostic of UTI and are consistent with the notion of bottlenecks that limit strain diversity during UTI pathogenesis.**

Received 14 January 2014 **Accepted** 23 January 2014 **Published** 25 February 2014

Citation Willner D, Low S, Steen JA, George N, Nimmo GR, Schembri MA, Hugenholtz P. 2014. Single clinical isolates from acute uncomplicated urinary tract infections are representative of dominant *in situ* populations. mBio 5(2):e01064-13. doi:10.1128/mBio.01064-13.

Editor Scott Hultgren, Washington University School of Medicine

Address correspondence to Philip Hugenholtz, p.hugenholtz@uq.edu.au, or Mark A. Schembri, m.schembri@uq.edu.au.

Urinary tract infections (UTIs) are one of the most common bacterial infectious diseases of humans, responsible for an estimated 9.6 million doctor visits in the United States each year (1). Most are acute uncomplicated infections that occur in healthy individuals with no history of urological disorders. Uropathogens may colonize the urethral opening and infect the lower urinary tract up to the bladder, resulting in urethritis and cystitis, respectively (2). Some UTIs may progress further, resulting in infection of the kidneys (pyelonephritis), and even spread into the bloodstream, leading to a systemic infection known as urosepsis (3).

Over 80% of UTIs have been attributed to *Escherichia coli* infections, while other, less common uropathogens include a variety of Gram-positive and Gram-negative bacteria such as *Staphylococcus*, *Klebsiella*, *Serratia*, *Enterococcus*, and *Proteus* species (4, 5). The standard method used for the identification of uropathogens in most clinical laboratories is microscopy followed by conventional microbiological culturing (6, 7). Culture-independent analysis using sequencing of the 16S rRNA gene has indicated that UTIs may be more polymicrobial than initially believed and has implicated organisms, such as *Actinobaculum schaalii* and *Aerococcus urinae*, which are fastidious and may be overlooked by standard techniques (8–10). Furthermore, 16S studies of culture-negative UTI samples and healthy urine samples have

demonstrated that bacterial colonization can occur despite the inability to cultivate organisms (8, 11–15).

For culture-positive UTIs, molecular analysis based on sequencing has shown that cultivation generally is able to identify the most dominant organism in infected urine at the species level (8, 10, 16). Numerous studies have compared cultured isolates across patient cross sections, but little is known about the strain-level diversity of uropathogens within individuals. The *fimH* gene has been used extensively as a phylogenetic marker for the characterization of uropathogenic *E. coli* (UPEC) isolates at the strain level (17–21). The *fimH* gene encodes the tip-located type 1 fimbrial FimH adhesin that mediates binding to α-D-mannosylated glycoproteins such as uroplakins on human bladder epithelial cells (ECs) (22–24) and is an essential virulence factor required for adhesion, invasion, intracellular bacterial community formation, and colonization of the bladder (25, 26). The sequence of *fimH* is highly conserved in virtually all *E. coli* strains sequenced to date (including nonuropathogens), with minor sequence variations often corresponding to functional differences, most likely the result of adaptation to enhance pathogenesis (19, 27–34). Compared to other established *E. coli* genotyping methods such as multilocus sequence typing (MLST) (20), typing based on *fimH* variations is a faster and more practical approach (18). A two-locus-typing scheme combining *fimH* with *fumC* (fumarase C) has also been shown to provide strong clonal discrimination power for molecular epidemiological analyses (21).

Here, we profiled the microbial communities associated with 50 uncomplicated UTIs using 16S rRNA and *fimH* amplicon pyrosequencing. These data were compared to the corresponding markers from the paired clinical strains obtained in the hospital pathology lab to determine how representative single clinical isolates are of underlying infections at both the genus (16S rRNA) and the strain (*fimH*) level. Eight paired urine UTI samples and *E. coli* isolates spanning the observed marker gene diversity were then shotgun sequenced to determine if *fimH* is sufficiently representative of strain-level diversity at the whole-genome level.

RESULTS

Patient demographics and clinical microbiology. Urine samples were obtained from 50 individuals with acute uncomplicated UTIs (see Table S1 in the supplemental material). The majority of study subjects were female (76%), and patient ages ranged from <1 year to 94 years (Table S1; median, 56 years). *E. coli* was the most commonly isolated organism (70%) followed by *Pseudomonas aeruginosa* (10%; Table S1). *E. coli* infection was associated with age ($P < 0.0001$, Mann-Whitney U test) and gender ($P = 0.03$, Fisher's exact test), occurring most often in younger females. Age was also associated with *E. coli* type, with phylogroup B2 more common in younger individuals versus D in older individuals ($P = 0.01$, Mann-Whitney U test). Males had higher white blood cell (WBC) counts on average ($P = 0.04$, Mann-Whitney U test) than did females, but no other correlations between clinical parameters and gender were observed.

Cultured isolates are representative of UTI microbial communities at the genus level. For the majority of patients, the most abundant member of the microbial community as determined by 16S pyrosequencing was concordant with the cultured isolate (Fig. 1). Over 80,000 amplicon sequences were obtained for the 50 patients, with a median of 1,359 sequences per sample (normalized to 1,000 for analysis). 16S sequences were analyzed using QIIME and CD-HIT-OTU, and microbial community profiles were compared to cultured isolates. The taxonomic identity of microbial isolates was verified by 16S Sanger sequencing (see Table S1 in the supplemental material). In all cases, the genus corresponding to the isolate(s) was observed in the profiles; however, in eight samples (16%) the cultured genus was not the most dominant. Analysis of biological replicates from these samples indicated that this was not likely due to heterogeneity within urine samples (see Fig. S1). In three of the 50 cases (6%), the most abundant genus in the community profile was anaerobic (either *Anaerococcus* or *Peptoniphilus*), while only aerobic organisms were cultured. No epithelial cells were identified microscopically in these samples; however, one of the *Anaerococcus*-dominated urine samples was noted as containing debris, which could be indicative of contamination during sample collection and a source of additional microbial diversity. Both *Anaerococcus* and *Peptoniphilus* species have also been identified in the urinary microbiome of healthy individuals (35). For a further three individuals, Gram-negative *Enterobacteriaceae* were cultured, while Gram-positive organisms were found to be most dominant by 16S sequencing (50 to 75%). In the two remaining individuals, different genera of *Enterobacteriaceae* were represented by the cultured isolate and the most abundant constituent in the community profiles.

UTI microbial communities were significantly different between age groups ($P = 0.013$, generalized PERMANOVA) (see Fig. S2A in the supplemental material). Principal component analysis using the generalized Unifrac distance indicated that communities dominated by *Enterobacteriaceae*, and more specifically *E. coli*, were more common in younger individuals. Older individuals tended to have communities with higher abundances of *Pseudomonas* and members of the phylum *Firmicutes*. Gender was a borderline significant factor influencing community composition ($P = 0.054$) (see Fig. S2B). UTI samples from females were hallmarked by microbial communities with a high relative abundance of *Enterobacteriaceae* and/or the presence of lactic acid bacteria and other bacilli. In particular, the genera *Streptococcus*, *Lactobacillus*, and *Staphylococcus* were not detected in any males in the study, even at low relative abundance (Fig. 1). These results are highly concordant with those based on isolate data alone and reflect the high correspondence between cultured isolates and microbial communities as described above.

Overall diversity in UTI microbial communities was very low (see Fig. S3 in the supplemental material). Rarefaction analysis was used to compare operational taxonomic unit (OTU) richness between samples, and overall diversity was assessed using the Shannon index at a depth of 1,000 sequences (see Table S1). The average UTI diversity by the latter metric is 0.51, which is lower than those of human skin (1 to 2.5) (36) and the gut (~6) (37). The Shannon index was inversely correlated with the relative abundance of *E. coli* (Tau = −0.54, $P < 0.0001$), and these UTIs had lower overall richness in general. *E. coli* infections also had a borderline significant association with higher microbial loads as determined by quantitative PCR (qPCR) ($P = 0.061$, Mann-Whitney U test). No significant differences in community diversity or microbial biomass associated with age or between males and females were observed. The most diverse UTI community, with a Shannon index of 3.84, was observed in an 86-year-old male (Fig. 1; indicated by asterisk). Clinical notes on this individual's sample indicated that a large amount of debris was observed

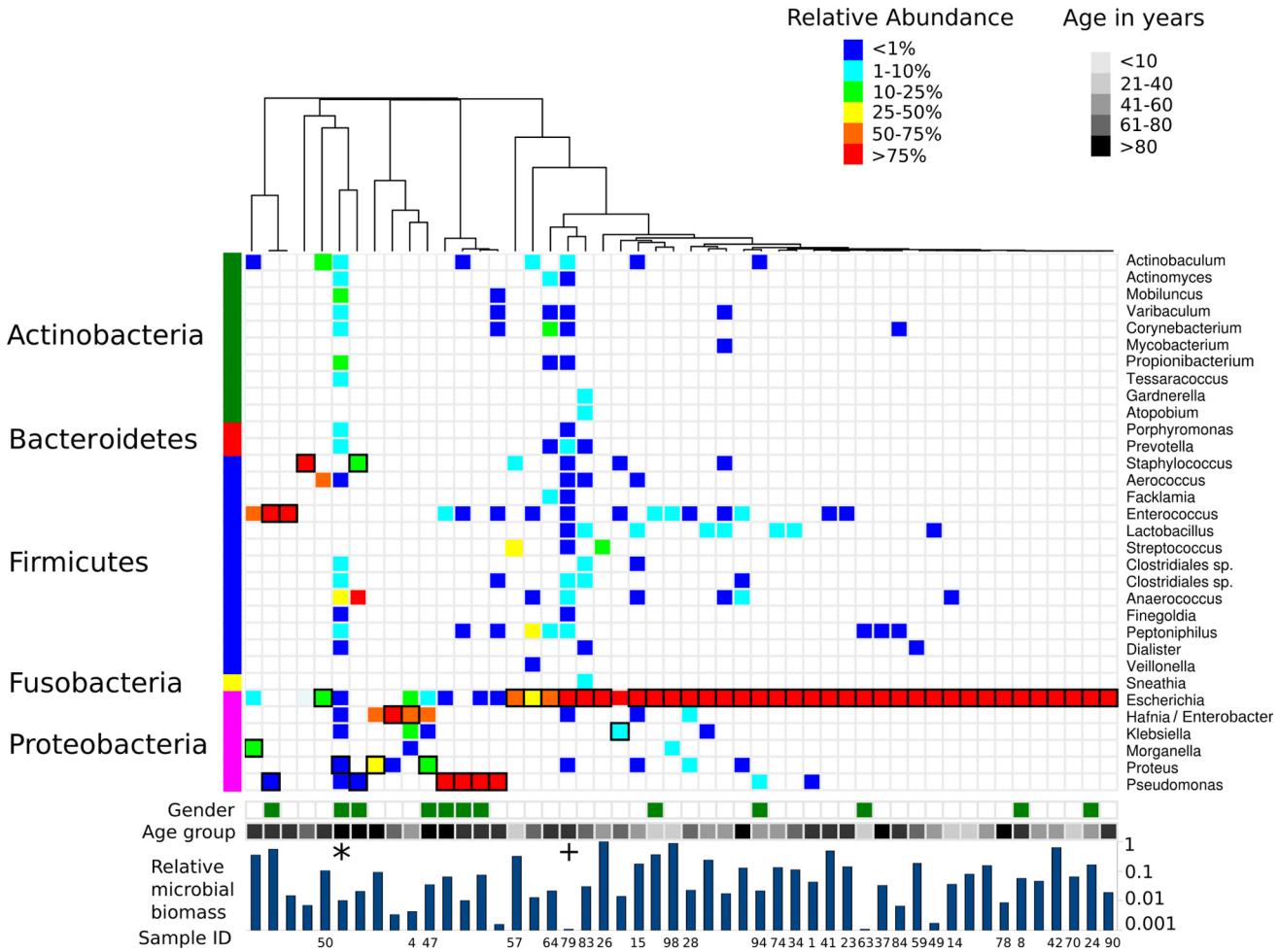

FIG 1 Microbial relative abundance and biomass in UTI samples. The heat map of genus (right)- and phylum (left)-level taxonomy for UTI microbial community profiles is based on 16S pyrosequencing with each sample normalized to 1,000 sequence reads. Olive green boxes indicate males, and darker gray shading indicates older individuals. The genus of the cultured isolate for each UTI is indicated by a black border around the appropriate box. Sample numbers are given for those UTIs which also underwent *fimH* pyrosequencing, and two samples with noted higher taxonomic richness are marked with an asterisk and a plus sign. Relative microbial biomass was estimated by 16S quantitative PCR.

in the urine. Increased diversity was also observed in a UTI community from a 71-year-old female, which was dominated by *E. coli* but also contained several members of the phyla *Actinobacteria*, *Bacteroidetes*, and *Firmicutes* (Fig. 1; indicated by plus sign). This sample had an epithelial cell count greater than 50, which is suggestive of contamination by normal microbiota of the anogenital region.

Isolates are congruent with the dominant strain in *E. coli* UTIs. *Escherichia* was the most common genus identified by both culture and community profiling in urine UTI samples. We further characterized *Escherichia*-positive samples by examining the urine UTI samples and their respective matching clinical isolates ($n = 27$) using the *fimH* gene as a strain-level marker. A 523-bp region of the *fimH* gene was targeted for amplicon pyrosequencing using primers designed to amplify all *fimH* genes belonging to the genus *Escherichia*. These primers were used to profile *Escherichia* strain-level diversity in the UTI samples, and representative sequences for all unique OTUs (clusters with 100% identity) were used to generate a phylogenetic tree with manually identified subtrees (Fig. 2). Two additional urine UTI samples from which other *Enterobacteriaceae* (*Enterobacter aerogenes* and *Proteus mirabilis*) were cultured but which had >1% relative abundances of *E. coli* in their community profiles were also included in the analysis. *fimH* sequences were generated from *E. coli* isolates using Sanger sequencing. Isolates were also classified using the Clermont method, which classifies *E. coli* into phylogroups based on a multiplex PCR targeting four discriminating loci within *E. coli* genomes (38). Additionally, the expression of functional type I fimbriae by all clinical isolates was confirmed using a yeast agglutination assay (39)

In all cases where *E. coli* was cultured, the dominant *fimH* type in the community matched the cultivated isolate (Fig. 2). A total of 33 distinct nucleotide sequences were observed in the UTI communities. Ten of these represented dominant *fimH* phylotypes across the 50 UTI samples examined (nucleotide dominant types; ND), all of which corresponded to their paired isolate sequences. The remaining 22 phylotypes were typically found in low abundance (nucleotide secondary types; NS) and were not represented by the cultured isolates. Three of the ND types were congruent with *fimH* lineages identified in ST131 subclones, including the globally distributed fluoroquinolone-resistant *fimH*30 subtype

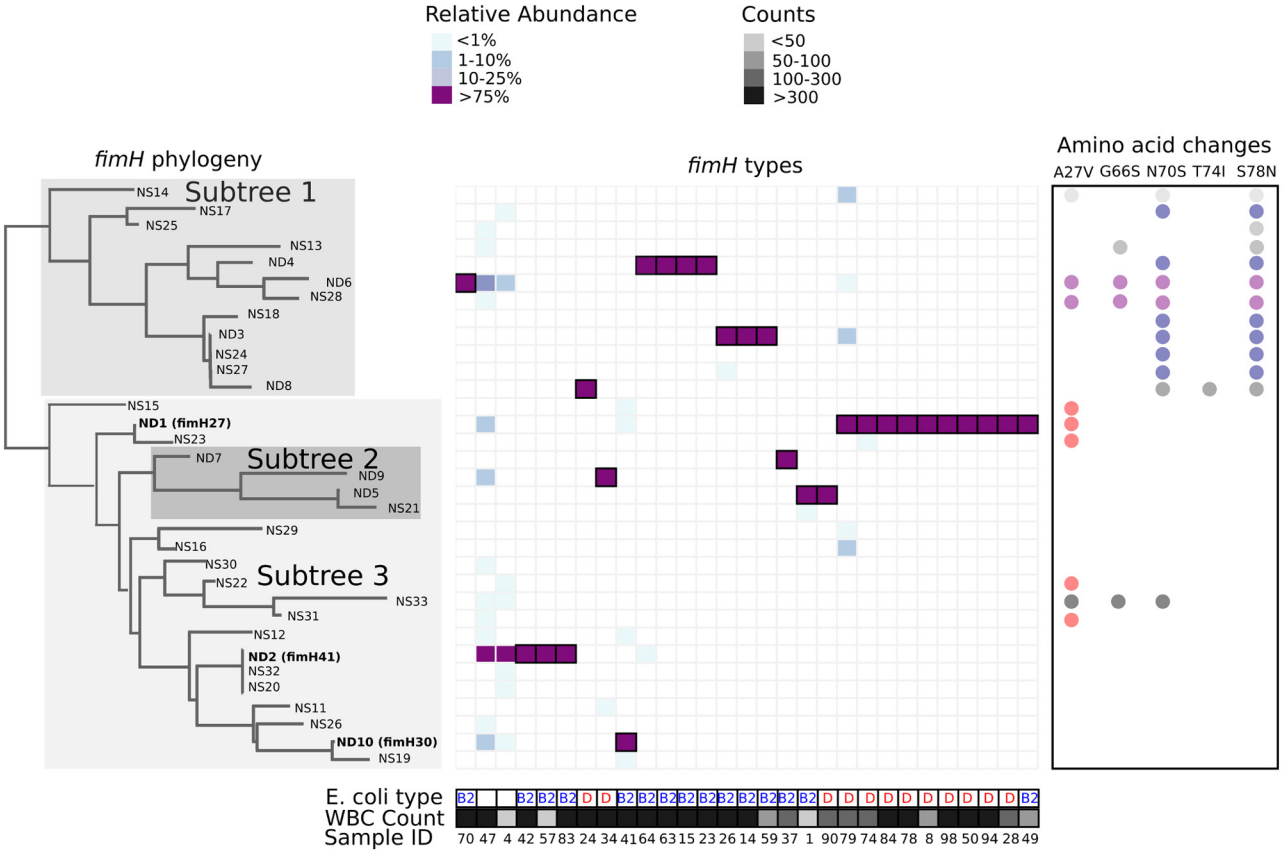

FIG 2 Heat map of strain-level taxonomy and *fimH* phylogeny for UTI microbial community profiles based on *fimH* pyrosequencing. The *fimH* type of the cultured isolate for each UTI is indicated by a black border around the appropriate box. Each node of the phylogenetic tree corresponds to a unique 100% OTU representative sequence, and gray shading indicates manually identified subtrees. Amino acid substitutions are indicated by circles for each corresponding nucleotide sequence. The phylogroup is given for samples that had *E. coli* as the cultured isolate, and darker gray squares indicate higher white blood cell counts. Sample numbers are presented below the bar for white blood cell count.

(40) (see Fig. S4 in the supplemental material). *fimH* diversity was extremely low overall, and in 20 of 29 cases (69%), only one *fimH* type could be detected in UTI samples (Fig. 2; see also Fig. S5). Biological replicates were highly concordant (Fig. S5), and rarefaction analysis showed no increase in *fimH* OTU richness even at a depth of 70,000 sequences (Fig. 2; see also Fig. S3C), indicating that these results were not due to sampling bias or insufficient sequencing. The diversity of *fimH* was notably higher in UTIs where *E. coli* was not the cultured isolate and for samples with epithelial cell counts greater than 50 (see Fig. S2B).

Translation of *fimH* nucleotide sequences resulted in 10 unique amino acid sequence types, five of which were the dominant type in at least one UTI (Fig. 2; see Table S2 in the supplemental material). Within the amino acid sequence types, five individual substitutions were observed (Fig. 2; see Table S2), all of which have been previously described (19, 27, 32). None of the amino acid types or individual substitutions were significantly associated with microbial community diversity (as measured by the Shannon index) or total microbial load. FimH allele types were labeled according to the presence of amino acid changes in dominant FimH types (AAD1 to -5) or low-abundance FimH subtypes (AAS6 to -10) (see Table S2). FimH amino acid type AAD2 was significantly more abundant in UTIs that had phylogroup D isolates ($P = 0.0002$), while AAD3 was associated with phylogroup B2 ($P = 0.0009$). AAD1 had a more widespread distribution and was the dominant FimH type in both of the UTIs not dominated by *E. coli*. AAD3 correlated significantly with an increased white blood cell (WBC) count ($P = 0.005$), while AAD1 was associated with a lower WBC count in general ($P = 0.02$). This reflected an underlying association between the amino acid substitutions N70S and S78N, both of which appear in AAD3 (see Table S2), and a higher WBC count ($P = 0.0008$ and $P = 0.001$, respectively). S78N was also correlated with patient age, occurring more often in younger individuals ($P = 0.04$).

Whole-genome analysis confirms *fimH* typing of UPEC strains. Eight *E. coli*-dominated urine UTI samples and their paired cultured isolates were selected for shotgun metagenomic and genomic sequencing, respectively. Genomic contigs were assembled *de novo* from isolate sequence data (see Table S3 in the supplemental material), and metagenomic sequences were mapped to their paired isolate genome to further evaluate the strain-level diversity of *E. coli* and dominance of the cultured isolate in the UTI (see Table S4). Metagenomic sequences were also mapped to the human genome to detect contaminating host genomic DNA. All UTI metagenomes contained sequences corresponding to human genomic DNA with relative abundances ranging from <1% to >80% (Fig. 3; see also Table S4). The two samples (UQU-8 and UQU-57) with extremely low relative

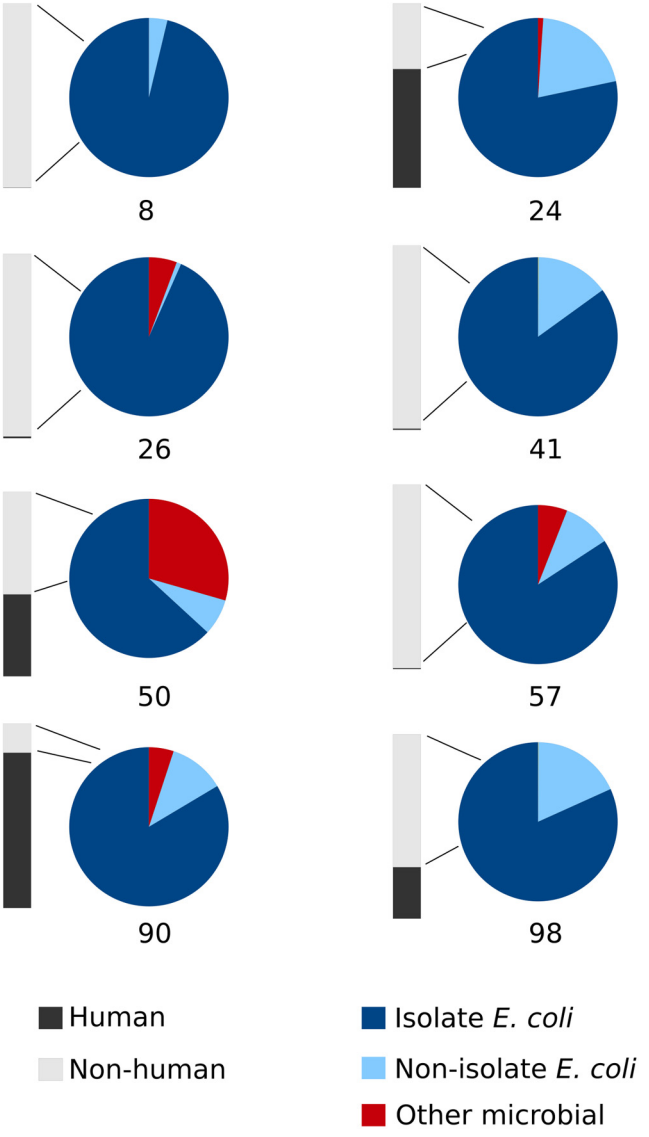

FIG 3 Taxonomic classification of metagenomic sequences. Bars show the proportion of metagenomic sequences which mapped to the human genome. Pie charts show the proportions of non-human sequences which mapped to *E. coli* isolate scaffolds and those which were assigned to non-isolate *E. coli* genomes and other microbial genomes by MEGAN.

abundances of human sequences (<1%) had white blood cell counts of ≤60, compared to >250 for all others, suggesting that WBCs may have been the major source of human genomic DNA. Sequences that did not map to isolate genomic contigs or the human genome were taxonomically classified using MEGAN as described in Materials and Methods.

For all UTI metagenomes, the majority of microbial sequences identified corresponded to *E. coli* and more specifically mapped to the corresponding cultured isolate (Fig. 3). Community profiles for UQU-8, UQU-24, UQU-41, UQU-90, and UQU-98 identified *Escherichia* species at >90% abundance (Fig. 1), and correspondingly, >95% of microbial sequences in metagenomes for these UTIs corresponded to *E. coli*. Furthermore, in these five UTIs and all other cases, 80% or more of the *E. coli* sequences were representative of the isolate genome. In addition to high relative abundances of *E. coli*, UQU-26, UQU-50, and UQU-57 had significant (>5% of total sequences) populations of other microbes, including *Streptococcus* spp. (UQU-26 and UQU-57), *Staphylococcus* spp. (UQU-57), and *A. urinae* and *Actinomycetales* spp. (UQU-50). These results are consistent with 16S profiles for these samples.

DISCUSSION

Common clinical practice for acute uncomplicated UTIs is to diagnose and treat the infection on the basis of characterization of a single cultured bacterial isolate, under the assumption that this isolate represents the dominant *in situ* population (6, 7). Recent evidence suggests that UTIs may be more polymicrobial than previously suspected and also that detectable bacterial populations are present even in healthy urine (11, 14, 15, 41). Here, we used culture-independent sequencing-based methods to investigate microbial communities in urine samples from individuals with uncomplicated UTIs at the genus and strain level to determine if single isolates are adequate to describe the most abundant populations in the disease state.

For the majority of UTIs (80%), cultured isolates were representative of dominant organisms at the genus level. Previous culture-independent analyses of UTIs in hospitalized patients, catheter-associated UTIs, and asymptomatic bacteriuria demonstrated similar concordance between cultivation and molecular analysis for culture-positive specimens (8, 10, 16, 41). Comparable levels of agreement have been reported for other high-biomass bacterial infections, such as those implicated in cystic fibrosis lung disease (42) and chronic wounds (43).

In three of 50 UTIs, anaerobic genera were more abundant than the isolate genus, while in a fourth, fastidious *Aerococcus* spp. were most dominant. The use of denaturing high-performance liquid chromatography coupled with 16S rRNA gene sequencing in infected urine samples first identified the widespread presence of fastidious bacteria, including obligate anaerobes that were previously overlooked by routine microbiological cultivation (8, 10). While some of these, such as *A. urinae* (44, 45), have been implicated directly in the etiopathogenesis of UTIs, the role of others, especially anaerobes and *Lactobacillus* spp., is largely unclear, as they have often been isolated from polymicrobial infections in conjunction with more traditional pathogens such as *E. coli* (8, 10, 46). Furthermore, community profiling using high-throughput amplicon pyrosequencing has subsequently demonstrated that these organisms are common constituents of both healthy and diseased urine microbiomes (11–15, 41), suggesting that they may not be causally linked to UTI.

The majority of UTIs were associated with the genus *Escherichia* and more specifically *E. coli*, which is consistent with previous literature (4, 5). Most of these were largely monomicrobial, with other genera present only in low or negligible (<1%) abundances. Strain typing of *E. coli* isolates from different UTI samples using *fimH* gene sequencing indicated a wide diversity of strains in the study population. All *E. coli* strains were also shown to produce functional type 1 fimbriae. In order to determine if the isolated strains were representative of dominant *E. coli* populations within individual urine UTI samples, we used amplicon pyrosequencing of the *fimH* locus, as, in general, the phylogenetic resolution of 16S amplicon pyrosequencing is limited to the genus level (47, 48). While *fimH* has been used in a number of studies to type UPEC

isolates (18, 21, 27, 30, 32, 40), to our knowledge this is the first instance where this marker has been used to profile UTI *E. coli* strain diversity *in situ*. Similar assays have also been developed for the rapid analysis of total species and/or strain-level community diversity of methanotrophs (49) and ammonia-oxidizing archaea (50) in environmental samples.

In all urine UTI samples, cultured isolates represented the dominant and, in most cases, the only detectable *E. coli* strain. Three of the dominant *fimH* types corresponded to subclones of the globally disseminated multidrug-resistant ST131 clone of *E. coli* (40). Two of these, *fimH*30 and *fimH*41, represent common and ubiquitous *fimH* lineages, present in *E. coli* from a range of hosts and geographical locations (40). For a few samples, additional strains were detected but at extremely low abundances. This lies in stark contrast to other infections such as *Pseudomonas aeruginosa* respiratory infections, which have been shown to be highly heterogeneous and poorly described by single cultured isolates (51–53). *P. aeruginosa* respiratory disease is most often chronic, occurring in patients with anatomical or physiological abnormalities that render them susceptible to repeated bouts of infection, while acute uncomplicated UTI is more likely to be due to an episode of bacterial invasion of a normally closed system. Other types of UTIs, such as catheter-associated infections, may be more heterogeneous, as they allow for sustained exposure to colonizing microbial populations from external sources (16).

E. coli has an extensive pangenome with a large reservoir of genes undergoing frequent lateral transfer (54–56), and thus, individual marker genes may not be indicative of strain-level diversity. Multilocus sequence typing (MLST) has commonly been used for classification (57); however, it has been shown to lack discriminatory power, grouping genetically, phenotypically, and ecologically distinct strains under the same sequence type (21, 40). Molecular typing with *fimH* has been proposed as a feasible alternative to MLST for large-scale epidemiological studies of *E. coli* (18, 21, 40). It has been shown that *fimH* can be used as a single marker or in conjunction with *fumC* to successfully discriminate ecologically and clinically relevant sublineages within MLST strain types (21, 40). Comparative analysis of the genomes of two ST131 isolates confirmed the validity of this approach, revealing marked divergence in a subset of shared core genes, including *fimH*, despite identical MLST loci in the two strains (58). Differences in *fimH* loci were driven by homologous recombination as well as the acquisition of point mutations, which further distinguished closely related strains with the same ancestral recombinant *fimH* alleles (58). Here, we performed shotgun sequencing on eight isolate-urine UTI pairs to corroborate the results of *fimH* amplicon pyrosequencing analysis and determine the validity of *fimH* as a strain-level marker. Overwhelmingly, the majority of metagenomic sequences of microbial origin from urine UTI samples mapped back to their corresponding isolate genomes, confirming that *fimH* is indicative of the overall genomic content of UPEC strains.

The ability of *fimH* to uniquely define UPEC strains may be linked to infection bottlenecks that have been defined in the mouse UTI model (59). In mice, FimH-mediated binding to mannosylated glycoproteins on bladder epithelial cells is critical for UPEC colonization and invasion of the uroepithelium and the initiation of cystitis (24, 60, 61). Enhanced mannose binding due to pathoadaptive mutations under high-shear-force conditions is also beneficial as it allows bacterial cells to resist clearance by micturition (30, 62, 63). Chen et al. demonstrated that *fimH* also exhibits a second class of mutations that are independent from those that enhance mannose binding and are important for the formation and proliferation of intracellular bacterial communities (IBCs) (27). IBCs in mice infected with a panel of differentially tagged UPEC strains were clonally derived from a single invasive bacterium, despite initial infection with a diverse population (64). The formation of IBCs represents a stringent bottleneck that influences the bacterial population in the lumen of the mouse bladder (64), and UPEC diversity has been shown to decrease as infections progress from the bladder to the kidneys and into the bloodstream (65). Thus, the proliferation of a single dominant *fimH* strain in acute uncomplicated human UTI may therefore be a consequence of multiple dynamics during infection. Minor subpopulations exhibiting other *fimH* types could represent either genetic drift in the founder population or the transient flux of newly introduced strains ascending from the intestinal tract or descending from the kidneys (65).

Our results indicate that identification of dominant uropathogens in UTIs via single culture representatives is diagnostic at the genus level in the majority of cases, and also at the strain level for *E. coli* infections. This implies that the common practice of prescribing antibiotic treatment based on resistance profiles of isolates is an effective treatment strategy for acute uncomplicated UTIs. Newer, more rapid, and cost-effective methods such as flow cytometry (66) and real-time PCR assays (67) that are targeted at identifying highly abundant organisms rather than total microbial communities will similarly be effective for the diagnosis and monitoring of UTIs. However, the question remains whether other microbial populations present in infected urine samples are causally linked to UTI, and further experiments exploring the mechanisms of microbial pathogenicity in the urinary tract are necessary. Additionally, while we have demonstrated the utility of *fimH* as a molecular marker for strain typing of *E. coli* and verified the lack of strain-level diversity in *E. coli*-dominated UTIs, it remains to be determined how representative single clinical isolates are for other common UTI pathogens and other types of infections such as complicated or catheter-associated UTIs.

MATERIALS AND METHODS

Ethical approval. Ethical approval for this study was obtained from the Royal Brisbane and Women's Hospital (RBWH) ethics committee (HREC/11/QRBW/107). The need for informed consent was waived by the institutional review board.

Study inclusion criteria. In order to be included in the study, individuals were required to have no history of urological disorders as well as a urine sample clinically diagnosed with UTI based on cell counts and cultures. Epithelial cell (EC), white blood cell (WBC), and red blood cell (RBC) counts were performed using phase-contrast microscopy with Kova slides (Hycor, CA) at ×10 and ×40 magnifications. Urine cultures were performed on both MacConkey and blood agar plates, with incubation at 35°C for 16 to 18 h under aerobic conditions. A clinical diagnosis of UTI required a WBC count of $>10^7$ liter^{-1}, an EC count of $<10^7$ liter^{-1}, and bacterial cultures in excess of 10^6 CFU · liter^{-1}. Urine samples with a WBC count of $>10^7$ liter^{-1} and an EC count of $>10^7$ liter^{-1} were also considered positive if culture counts were $>10^7$ CFU · liter^{-1}.

Sample collection and initial processing. Midstream urine samples were collected from patients presenting at the RBWH for microbiological analysis. All urine samples were deidentified and frozen at −20°C prior to DNA isolation for pyrosequencing. DNA was extracted from 1 ml of each urine sample using the Nucleospin tissue kit (Macherey-Nagel, Düren, Germany) according to the manufacturer's protocol for hard-to-lyse or-

ganisms. Extractions were performed in duplicate on a subset of samples for biological replication. DNA was stored at −20°C prior to further processing. Cultured bacterial isolates were also obtained for each urine sample. Patient demographic data (gender and age) were obtained and matched to all deidentified samples.

Phenotypic and molecular analysis of clinical isolates. Bacterial isolates were identified to the species level using phenotypic assays, including the Vitek2 microbial identification system (bioMérieux). A single colony representing the dominant colony type was selected from the original culture plate and used to inoculate 5 ml of sterile Luria-Bertani (LB) liquid medium. Cultures were incubated overnight at 37°C with shaking (150 rpm). Cells were centrifuged at 4,000 × *g* for 15 min, resuspended in 20% glycerol liquid LB medium, and stored in 1-ml aliquots at −80° C.

Isolates were identified at the molecular level by Sanger sequencing of the 16S rRNA gene using the primers 926F and 1392wR (see Table S5 in the supplemental material). A 50-μl PCR was run for each sample. Direct lysis multiplex PCR was carried out using a 25-μl reaction volume containing 2.5 μl 10× buffer, 10 mM (each) deoxynucleoside triphosphate (dNTP) mix, 2 mM $MgCl_2$, 15 μg bovine serum albumin (BSA), 12.5 pmol of each primer, 1 unit *Taq* polymerase (Fisher Biotec, Australia), and 0.5 μl of the culture glycerol stock as the DNA template. PCR was performed using an Applied Biosystems Veriti thermal cycler under the following conditions: cell lysis at 96°C for 10 min; initial denaturation at 95°C for 3 min; 30 cycles of 95°C (30 s), 55°C (45 s), and 72°C (1 min 30 s); and a final extension at 72°C of 10 min. Briefly, 2 μl of exonuclease and Antarctic phosphatase master mix (Affymetrix, Santa Clara, CA) was added to 20 μl of PCR product. Samples were then incubated in an Applied Biosystems Veriti thermal cycler for 15 min at 37°C and 80°C, respectively, to completely inactivate the enzymes. Purified PCR products were submitted for Sanger sequencing using the 926F primer at Macrogen Inc. (Seoul, South Korea). 16S Sanger sequences were manually trimmed and compared to the Greengenes database for taxonomic assignment using BLAST (68, 69).

Phylogroups for *E. coli* isolates were determined by a modified direct lysis triplex PCR method using primers targeting two genes (*chuA* and *yjaA*) and one anonymous DNA fragment (TspE4.C2) (see Table S5 in the supplemental material) as previously described (38). To assay for the presence of functional type 1 fimbriae, *E. coli* isolates were grown in overnight 5-ml static cultures under aerobic conditions at 37°C. Subsequently, 10 μl of each overnight *E. coli* culture was mixed with 10 μl of 5% yeast solution, respectively, and agglutination was observed as previously described (39).

Primers targeting the *fimH* gene were designed based on regions of the gene conserved across different *E. coli* strains, including an outgroup species, *Escherichia albertii*, using ARB (70). Using the *fimH* primers 72F and 563R (see Table S5 in the supplemental material), direct lysis PCR was carried out in a 25-μl reaction volume as described for 16S sequencing. PCR products were purified using the modified Exo-Sap protocol and sequenced from both the forward and reverse primers at Macrogen Inc. (Seoul, South Korea). Each pair of *fimH* Sanger sequences (forward and reverse) was aligned and assembled, and consensus calls were performed automatically according to base quality, using the Geneious software package. Ambiguous base calls were checked manually using chromatograms. Sequences were trimmed to 320 bp for comparison with *fimH* amplicon pyrosequences as described below.

Quantitative real-time PCR. Quantitative real-time PCR was performed on extracted urine DNAs using the 16S primers 803Fa and 1392wR (see Table S5 in the supplemental material) and SYBR-Green MasterMix (Life Technologies, Carlsbad, CA, USA) on the ABI 7900 platform to assess total microbial load in the UTI samples. Real time-PCR (qPCR) was performed in triplicate at two dilutions using the following cycling conditions: 10 min at 95°C and 40 cycles of 15 s at 95°C followed by 1 min at 60°C. A melt curve was produced by running a cycle of 2 min at 95°C and a last cycle of 15 s at 60°C. 16S copy number was determined based on a standard curve constructed from a dilution series of *E. coli* strain DH10B genomic DNA.

16S community profiling. Fusion primers containing 454 adaptor sequences and oligonucleotide bar codes were used to amplify the V5, V6, V7, and V8 regions of the 16S rRNA gene. The forward primer was a mixture of four variants (803Fa, 803Fb, 803Fc, and 803Fd) in a ratio of 2:1:1:1 used in conjunction with the reverse primer 1392wR (see Table S5 in the supplemental material) (71). PCR was performed as described for isolate 16S Sanger sequencing using 50-μl reaction mixtures, with a unique bar-coded reverse primer used for each sample. Amplicons were purified with the Agencourt AmPure XP system (Beckman Coulter, Danvers, MA) according to the manufacturer's instructions. The concentration of purified amplicons was quantified on a Qubit fluorometer (Invitrogen Corp., Carlsbad, CA) with the Qubit double-stranded DNA (dsDNA) high-sensitivity (HS) assay kit. Subsequently, equimolar amounts of bar-coded amplicons were pooled and sequenced on the 454 Genome Sequencer (GS)-FLX Titanium platform at the Australian Centre for Ecogenomics.

Amplicon sequences were checked for chimeras using UCHIME version 4.1 (72) and then quality filtered, trimmed to 300 bp, and assigned to their respective samples using QIIME (73). Sequence clustering at 97% identity was performed using CD-HIT-OTU-454, which detects and adjusts homopolymer errors in amplicon sequences (74). Representative centroid sequences for each cluster were compared to the Greengenes database (February 2011 release) for taxonomy assignment using BLAST (68, 69). Rarefaction analysis was performed using QIIME (73). Sequence libraries were normalized to 1,000 sequences per sample using a repeated subsampling procedure as described in reference 75 prior to further analysis. Genus-level taxonomy for the normalized OTU table was visualized using the heatmap.2 function in the R package gplots (76). The Shannon index was calculated using QIIME, and beta-diversity was assessed using the generalized Unifrac metric (77).

***fimH* community profiling.** Partial *fimH* gene sequences from urine samples were amplified using the primers fimH72F and fimH563R (see Table S5 in the supplemental material) modified to include 454 adaptor sequences and a sample-specific oligonucleotide bar code in the forward primer. PCRs and cycling conditions were identical to those used to amplify *fimH* from clinical isolates. Bar-coded amplicon preparation and sequencing were performed as described for 16S amplicons.

fimH sequences were quality filtered, trimmed to 320 bp (to maximize phylogenetic resolution), and separated by sample using QIIME. Clustering was performed at 100% similarity using CD-HIT-OUT-454, and singleton clusters were discarded. OTU representative sequences were manually checked for uncorrected homopolymer errors in Geneious using the conserved length of *fimH*. Briefly, reference sequences were aligned with the *fimH* sequence from *E. coli* K-12 (GenBank accession number NC_000913.2), and sequences not conforming to the conserved length due to deletions or insertions in homopolymer regions were manually corrected. Following correction, representative sequences were reclustered at 100% similarity using CD-HIT-EST (74). Chimeric sequences that did not align with the reference were also removed. Rarefaction curves were generated using QIIME (70). Sequence libraries were normalized to 1,000 sequences per sample prior to the creation of heat maps and calculation of the Shannon diversity index using QIIME (73). Full-length reference and partial pyrosequence amplicon *fimH* nucleotide sequences were aligned in ARB with ClustalW using default settings, and the alignment was masked using a 50% consensus filter (70). The *fimH* ARB database is available upon request. An evolutionary distance tree was constructed in ARB using the Olsen correction and neighbor-joining method. The naming system of the phylogenetic tree was adapted from reference 27, and subtrees were manually identified.

Nucleotide substitutions were described using *fimH* from *E. coli* K-12 as the reference sequence. OTU representative sequences were compared to isolate *fimH* sequences using BLAST and pairwise alignments in Geneious. Representative sequences were translated using TranSeq (78) and aligned in Geneious to the translated *E. coli* K-12 sequence and to *fimH*

ST131 clone reference sequences (40) using ARB (70). Amino acid substitutions were named according to the convention in reference 21.

Illumina genomic and metagenomic sequencing. Indexed whole-genome sequencing libraries were prepared with the Nextera DNA sample preparation kit (Illumina, San Diego, CA). Libraries were pooled such that each isolate genome was approximately 1/80 of the pool and the metagenomes were approximately 9/80 of the pool. This pool of libraries was submitted to the Queensland Centre for Medical Genomics (University of Queensland, QLD, Australia), and a single lane of 2 × 100-bp paired-end data was generated on an Illumina HiSeq2000 sequencer (see Table S1 in the supplemental material).

Read pairs were checked for overlap using SeqPrep (https://github.com/jstjohn/SeqPrep) and clipped to remove low-quality bases using Nesoni (http://bioinformatics.net.au/software.nesoni.shtml). *De novo* assemblies were generated from the isolate reads for each sample using Velvet version 1.2.10 with the help of VelvetOptimiser to determine optimum assembly parameters (79). Isolate genomes were checked for completeness using Phylosift (https://github.com/gjospin/PhyloSift) and Metachecka (https://github.com/Ecogenomics/PhylogeneticM) to identify sequences corresponding to 111 single-copy genes.

For each of the metagenome samples, human contamination was removed by mapping reads to the human genome (hg19) using BWA on default settings (80). Reads that did not map to the human genome were recovered and mapped to their respective *de novo* isolate contigs using BWA (default settings). Reads that remained unmapped against the isolate were compared to the NCBI database of complete microbial genomes using BLAST. Microbial taxonomy from best BLAST hits was summarized using MEGAN (81).

Statistical analysis. All statistical analyses were conducted in R (82). Two-way correlations between clinical, microbiological, and molecular covariates were determined using Fisher's exact test (two discrete covariates), the Mann-Whitney U test (one discrete and one continuous covariate), or Kendall's Tau (two continuous covariates). The Mann-Whitney U test was also used to compare average Unifrac distances between microbial community biological replicates. Associations between covariates and total microbial communities were determined using generalized PERMANOVA (83) with the unweighted, weighted, and generalized Unifrac distance matrices as inputs (77). Principal component analysis based on generalized Unifrac distances was performed using the function prcomp and visualized with the s.class function in the ADE4 package (84).

Nucleotide sequence accession number. All sequence data obtained from this study have been deposited in GenBank under BioProject PRJNA174753.

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at http://mbio.asm.org/lookup/suppl/doi:10.1128/mBio.01064-13/-/DCSupplemental.

Figure S1, PNG file, 0.1 MB.
Figure S2, TIF file, 0.3 MB.
Figure S3, PNG file, 0.5 MB.
Figure S4, PNG file, 0.1 MB.
Figure S5, PNG file, 0.4 MB.
Table S1, XLSX file, 0 MB.
Table S2, DOC file, 0.1 MB.
Table S3, XLSX file, 0.1 MB.
Table S4, XLSX file, 0.1 MB.
Table S5, DOCX file, 0.1 MB.

ACKNOWLEDGMENTS

We thank hospital staff for providing clinical isolates from the 50 studied UTI samples and Sue Read and Fiona May for assistance with amplicon pyrosequencing.

This work was supported by grants from strategic funding of the Australian Centre for Ecogenomics by the University of Queensland and grants from the National Health and Medical Research Council of Australia (APP1005315 and APP1033799). P.H. is supported by an Australian Research Council (A.R.C.) Discovery Outstanding Researcher Award (DP120103498). M.A.S. is supported by an ARC Future Fellowship (FT100100662). D.W. is supported by a University of Queensland Postdoctoral Research Fellowship.

REFERENCES

1. **Foxman B.** 2002. Epidemiology of urinary tract infections: incidence, morbidity, and economic costs. Am. J. Med. **113**(Suppl 1A):5S–13S. http://dx.doi.org/10.1016/S0002-9343(02)01054-9.
2. **Hooton TM.** 2012. Uncomplicated urinary tract infection. N. Engl. J. Med. **366:**1028–1037. http://dx.doi.org/10.1056/NEJMcp1104429.
3. **Lee JBL, Neild GH.** 2007. Urinary tract infection. Medicine **35:**423–428. http://dx.doi.org/10.1016/j.mpmed.2007.05.009.
4. **Linhares I, Raposo T, Rodrigues A, Almeida A.** 2013. Frequency and antimicrobial resistance patterns of bacteria implicated in community urinary tract infections: a ten-year surveillance study (2000–2009). BMC Infect. Dis. **13:**19. http://dx.doi.org/10.1186/1471-2334-13-19.
5. **Ronald A.** 2002. The etiology of urinary tract infection: traditional and emerging pathogens. Am. J. Med. **113:**14–19. http://dx.doi.org/10.1016/S0002-9343(02)01055-0.
6. **Cardoso CL, Muraro CB, Siqueira VL, Guilhermetti M.** 1998. Simplified technique for detection of significant bacteriuria by microscopic examination of urine. J. Clin. Microbiol. **36:**820–823.
7. **Pappas PG.** 1991. Laboratory in the diagnosis and management of urinary tract infections. Med. Clin. North Am. **75:**313–325.
8. **Domann E, Hong G, Imirzalioglu C, Turschner S, Kühle J, Watzel C, Hain T, Hossain H, Chakraborty T.** 2003. Culture-independent identification of pathogenic bacteria and polymicrobial infections in the genitourinary tract of renal transplant recipients. J. Clin. Microbiol. **41:**5500–5510. http://dx.doi.org/10.1128/JCM.41.12.5500-5510.2003.
9. **Gomez E, Gustafson DR, Rosenblatt JE, Patel R.** 2011. Actinobaculum bacteremia: a report of 12 cases. J. Clin. Microbiol. **49:**4311–4313. http://dx.doi.org/10.1128/JCM.00798-11.
10. **Imirzalioglu C, Hain T, Chakraborty T, Domann E.** 2008. Hidden pathogens uncovered: metagenomic analysis of urinary tract infections. Andrologia **40:**66–71. http://dx.doi.org/10.1111/j.1439-0272.2007.00830.x.
11. **Dong Q, Nelson DE, Toh E, Diao L, Gao X, Fortenberry JD, Van der Pol B.** 2011. The microbial communities in male first catch urine are highly similar to those in paired urethral swab specimens. PLoS One **6:**e19709. http://dx.doi.org/10.1371/journal.pone.0019709.
12. **Nelson DE, Van Der Pol B, Dong Q, Revanna KV, Fan B, Easwaran S, Sodergren E, Weinstock GM, Diao L, Fortenberry JD.** 2010. Characteristic male urine microbiomes associate with asymptomatic sexually transmitted infection. PLoS One **5:**e14116. http://dx.doi.org/10.1371/journal.pone.0014116.
13. **Siddiqui H, Lagesen K, Nederbragt AJ, Jeansson SL, Jakobsen KS.** 2012. Alterations of microbiota in urine from women with interstitial cystitis. BMC Microbiol. **12:**205. http://dx.doi.org/10.1186/1471-2180-12-205.
14. **Siddiqui H, Nederbragt AJ, Lagesen K, Jeansson SL, Jakobsen KS.** 2011. Assessing diversity of the female urine microbiota by high throughput sequencing of 16S rDNA amplicons. BMC Microbiol. **11:**244. http://dx.doi.org/10.1186/1471-2180-11-244.
15. **Wolfe AJ, Toh E, Shibata N, Rong R, Kenton K, Fitzgerald M, Mueller ER, Schreckenberger P, Dong Q, Nelson DE, Brubaker L.** 2012. Evidence of uncultivated bacteria in the adult female bladder. J. Clin. Microbiol. **50:**1376–1383. http://dx.doi.org/10.1128/JCM.05852-11.
16. **Xu Y, Moser C, Al-Soud WA, Sørensen S, Høiby N, Nielsen PH, Thomsen TR.** 2012. Culture-dependent and -independent investigations of microbial diversity on urinary catheters. J. Clin. Microbiol. **50:**3901–3908. http://dx.doi.org/10.1128/JCM.01237-12.
17. **Abdallah KS, Cao Y, Wei DJ.** 2011. Epidemiologic investigation of extra-intestinal pathogenic E. coli (ExPEC) based on PCR phylogenetic group and fimH single nucleotide polymorphisms (SNPs) in China. Int. J. Mol. Epidemiol. Genet. **2:**339–353.
18. **Dias RC, Moreira BM, Riley LW.** 2010. Use of fimH single-nucleotide polymorphisms for strain typing of clinical isolates of Escherichia coli for epidemiologic investigation. J. Clin. Microbiol. **48:**483–488. http://dx.doi.org/10.1128/JCM.01858-09.
19. **Sokurenko EV, Feldgarden M, Trintchina E, Weissman SJ, Avagyan S, Chattopadhyay S, Johnson JR, Dykhuizen DE.** 2004. Selection footprint in the FimH adhesin shows pathoadaptive niche differentiation in Esche-

richia coli. Mol. Biol. Evol. **21:**1373–1383. http://dx.doi.org/10.1093/molbev/msh136.
20. **Tartof SY, Solberg OD, Riley LW.** 2007. Genotypic analyses of uropathogenic Escherichia coli based on fimH single nucleotide polymorphisms (SNPs). J. Med. Microbiol. **56:**1363–1369. http://dx.doi.org/10.1099/jmm.0.47262-0.
21. **Weissman SJ, Johnson JR, Tchesnokova V, Billig M, Dykhuizen D, Riddell K, Rogers P, Qin X, Butler-Wu S, Cookson BT, Fang FC, Scholes D, Chattopadhyay S, Sokurenko E.** 2012. High-resolution two-locus clonal typing of extraintestinal pathogenic Escherichia coli. Appl. Environ. Microbiol. **78:**1353–1360. http://dx.doi.org/10.1128/AEM.06663-11.
22. **Jones CH, Pinkner JS, Roth R, Heuser J, Nicholes AV, Abraham SN, Hultgren SJ.** 1995. FimH adhesin of type 1 pili is assembled into a fibrillar tip structure in the Enterobacteriaceae. Proc. Natl. Acad. Sci. U. S. A. **92:**2081–2085. http://dx.doi.org/10.1073/pnas.92.6.2081.
23. **Krogfelt KA, Bergmans H, Klemm P.** 1990. Direct evidence that the FimH protein is the mannose-specific adhesin of Escherichia coli type 1 fimbriae. Infect. Immun. **58:**1995–1998.
24. **Wu XR, Sun TT, Medina JJ.** 1996. 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. http://dx.doi.org/10.1073/pnas.93.18.9630.
25. **Connell I, Agace W, Klemm P, Schembri M, Mărild S, Svanborg C.** 1996. Type 1 fimbrial expression enhances Escherichia coli virulence for the urinary tract. Proc. Natl. Acad. Sci. U. S. A. **93:**9827–9832. http://dx.doi.org/10.1073/pnas.93.18.9827.
26. **Mulvey MA, Lopez-Boado YS, Wilson CL, Roth R, Parks WC, Heuser J, Hultgren SJ.** 1998. Induction and evasion of host defenses by type 1-piliated uropathogenic Escherichia coli. Science **282:**1494–1497. http://dx.doi.org/10.1126/science.282.5393.1494.
27. **Chen SL, Hung CS, Pinkner JS, Walker JN, Cusumano CK, Li Z, Bouckaert J, Gordon JI, Hultgren SJ.** 2009. 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. http://dx.doi.org/10.1073/pnas.0902179106.
28. **Dreux N, Denizot J, Martinez-Medina M, Mellmann A, Billig M, Kisiela D, Chattopadhyay S, Sokurenko E, Neut C, Gower-Rousseau C, Colombel J, Bonnet R, Darfeuille-Michaud A, Barnich N.** 2013. Point mutations in FimH adhesin of Crohn's disease-associated adherent-invasive Escherichia coli enhance intestinal inflammatory response. PLoS Pathog. **9:**e1003141. http://dx.doi.org/10.1371/journal.ppat.1003141.
29. **Ronald LS, Yakovenko O, Yazvenko N, Chattopadhyay S, Aprikian P, Thomas WE, Sokurenko EV.** 2008. Adaptive mutations in the signal peptide of the type 1 fimbrial adhesin of uropathogenic Escherichia coli. Proc. Natl. Acad. Sci. U. S. A. **105:**10937–10942. http://dx.doi.org/10.1073/pnas.0803158105.
30. **Sokurenko EV, Chesnokova V, Dykhuizen DE, Ofek I, Wu XR, Krogfelt KA, Struve C, Schembri MA, Hasty DL.** 1998. Pathogenic adaptation of Escherichia coli by natural variation of the FimH adhesin. Proc. Natl. Acad. Sci. U. S. A. **95:**8922–8926. http://dx.doi.org/10.1073/pnas.95.15.8922.
31. **Sokurenko EV, Courtney HS, Ohman DE, Klemm P, Hasty DL.** 1994. FimH family of type 1 fimbrial adhesins: functional heterogeneity due to minor sequence variations among fimH genes. J. Bacteriol. **176:**748–755.
32. **Weissman SJ, Beskhlebnaya V, Chesnokova V, Chattopadhyay S, Stamm WE, Hooton TM, Sokurenko EV.** 2007. Differential stability and trade-off effects of pathoadaptive mutations in the Escherichia coli FimH adhesin. Infect. Immun. **75:**3548–3555. http://dx.doi.org/10.1128/IAI.01963-06.
33. **Hommais F, Gouriou S, Amorin C, Bui H, Rahimy MC, Picard B, Denamur E.** 2003. The FimH A27V mutation is pathoadaptive for urovirulence in Escherichia coli B2 phylogenetic group isolates. Infect. Immun. **71:**3619–3622. http://dx.doi.org/10.1128/IAI.71.6.3619-3622.2003.
34. **Schwartz DJ, Kalas V, Pinkner JS, Chen SL, Spaulding CN, Dodson KW, Hultgren SJ.** 2013. Positively selected FimH residues enhance virulence during urinary tract infection by altering FimH conformation. Proc. Natl. Acad. Sci. U. S. A. **110:**15530–15537. http://dx.doi.org/10.1073/pnas.1315203110.
35. **Lewis DA, Brown R, Williams J, White P, Jacobson SK, Marchesi JR, Drake MJ.** 2013. The human urinary microbiome; bacterial DNA in voided urine of asymptomatic adults. Front. Cell. Infect. Microbiol. **3:**41. http://dx.doi.org/10.3389/fcimb.2013.00041.
36. **Grice EA, Kong HH, Conlan S, Deming CB, Davis J, Young AC, Bouffard GG, Blakesley RW, Murray PR, Green ED, Turner ML, Segre JA, Segre JA.** 2009. Topographical and temporal diversity of the human skin microbiome. Science **324:**1190–1192. http://dx.doi.org/10.1126/science.1171700.
37. **Zhang X, Shen D, Fang Z, Jie Z, Qiu X, Zhang C, Chen Y, Ji L.** 2013. Human gut microbiota changes reveal the progression of glucose intolerance. PLoS One **8:**e71108. http://dx.doi.org/10.1371/journal.pone.0071108.
38. **Clermont O, Bonacorsi S, Bingen E.** 2000. Rapid and simple determination of the Escherichia coli phylogenetic group. Appl. Environ. Microbiol. **66:**4555–4558. http://dx.doi.org/10.1128/AEM.66.10.4555-4558.2000.
39. **Schembri MA, Sokurenko EV, Klemm P.** 2000. Functional flexibility of the FimH adhesin: insights from a random mutant library. Infect. Immun. **68:**2638–2646. http://dx.doi.org/10.1128/IAI.68.5.2638-2646.2000.
40. **Johnson JR, Tchesnokova V, Johnston B, Clabots C, Roberts PL, Billig M, Riddell K, Rogers P, Qin X, Butler-Wu S, Price LB, Aziz M, Nicolas-Chanoine MH, Debroy C, Robicsek A, Hansen G, Urban C, Platell J, Trott DJ, Zhanel G, Weissman SJ, Cookson BT, Fang FC, Limaye AP, Scholes D, Chattopadhyay S, Hooper DC, Sokurenko EV.** 2013. Abrupt emergence of a single dominant multidrug-resistant strain of Escherichia coli. J. Infect. Dis. **207:**919–928. http://dx.doi.org/10.1093/infdis/jis933.
41. **Fouts DE, Pieper R, Szpakowski S, Pohl H, Knoblach S, Suh MJ, Huang ST, Ljungberg I, Sprague BM, Lucas SK, Torralba M, Nelson KE, Groah SL.** 2012. Integrated next-generation sequencing of 16S rDNA and metaproteomics differentiate the healthy urine microbiome from asymptomatic bacteriuria in neuropathic bladder associated with spinal cord injury. J. Transl. Med. **10:**174. http://dx.doi.org/10.1186/1479-5876-10-174.
42. **Rudkjøbing VB, Thomsen TR, Alhede M, Kragh KN, Nielsen PH, Johansen UR, Givskov M, Høiby N, Bjarnsholt T.** 2011. True microbiota involved in chronic lung infection of cystic fibrosis patients found by culturing and 16S rRNA gene analysis. J. Clin. Microbiol. **49:**4352–4355.
43. **Rhoads DD, Cox SB, Rees EJ, Sun Y, Wolcott RD.** 2012. Clinical identification of bacteria in human chronic wound infections: culturing vs. 16S ribosomal DNA sequencing. BMC Infect. Dis. **12:**321. http://dx.doi.org/10.1186/1471-2334-12-321.
44. **Christensen JJ, Vibits H, Ursing J, Korner B.** 1991. Aerococcus-like organism, a newly recognized potential urinary tract pathogen. J. Clin. Microbiol. **29:**1049–1053.
45. **Zhang Q, Kwoh C, Attorri S, Clarridge JE.** 2000. Aerococcus urinae in urinary tract infections. J. Clin. Microbiol. **38:**1703–1705.
46. **Brook I.** 2004. Urinary tract and genito-urinary suppurative infections due to anaerobic bacteria. Int. J. Urol. **11:**133–141. http://dx.doi.org/10.1111/j.1442-2042.2003.00756.x.
47. **Janda JM, Abbott SL.** 2007. 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: pluses, perils, and pitfalls. J. Clin. Microbiol. **45:**2761–2764. http://dx.doi.org/10.1128/JCM.01228-07.
48. **Tindall BJ, Rosselló-Móra R, Busse HJ, Ludwig W, Kämpfer P.** 2010. Notes on the characterization of prokaryote strains for taxonomic purposes. Int. J. Syst. Evol. Microbiol. **60:**249–266. http://dx.doi.org/10.1099/ijs.0.016949-0.
49. **Kip N, Dutilh BE, Pan Y, Bodrossy L, Neveling K, Kwint MP, Jetten MS, Op den Camp HJ.** 2011. Ultra-deep pyrosequencing of pmoA amplicons confirms the prevalence of Methylomonas and Methylocystis in Sphagnum mosses from a Dutch peat bog. Environ. Microbiol. Rep. **3:**667–673. http://dx.doi.org/10.1111/j.1758-2229.2011.00260.x.
50. **Pester M, Rattei T, Flechl S, Gröngröft A, Richter A, Overmann J, Reinhold-Hurek B, Loy A, Wagner M.** 2012. amoA-based consensus phylogeny of ammonia-oxidizing archaea and deep sequencing of amoA genes from soils of four different geographic regions. Environ. Microbiol. **14:**525–539. http://dx.doi.org/10.1111/j.1462-2920.2011.02666.x.
51. **Chung JC, Becq J, Fraser L, Schulz-Trieglaff O, Bond NJ, Foweraker J, Bruce KD, Smith GP, Welch M.** 2012. Genomic variation among contemporary Pseudomonas aeruginosa isolates from chronically infected cystic fibrosis patients. J. Bacteriol. **194:**4857–4866. http://dx.doi.org/10.1128/JB.01050-12.
52. **Foweraker JE, Laughton CR, Brown DF, Bilton D.** 2005. Phenotypic variability of Pseudomonas aeruginosa in sputa from patients with acute infective exacerbation of cystic fibrosis and its impact on the validity of antimicrobial susceptibility testing. J. Antimicrob. Chemother. **55:**921–927. http://dx.doi.org/10.1093/jac/dki146.
53. **Gillham MI, Sundaram S, Laughton CR, Haworth CS, Bilton D, Fow-**

eraker JE. 2009. Variable antibiotic susceptibility in populations of Pseudomonas aeruginosa infecting patients with bronchiectasis. J. Antimicrob. Chemother. **63:**728–732. http://dx.doi.org/10.1093/jac/dkp007.
54. **Lawrence JG, Ochman H.** 1998. Molecular archaeology of the Escherichia coli genome. Proc. Natl. Acad. Sci. U. S. A. **95:**9413–9417. http://dx.doi.org/10.1073/pnas.95.16.9413.
55. **Rasko DA, Rosovitz MJ, Myers GS, Mongodin EF, Fricke WF, Gajer P, Crabtree J, Sebaihia M, Thomson NR, Chaudhuri R, Henderson IR, Sperandio V, Ravel J.** 2008. The pangenome structure of Escherichia coli: comparative genomic analysis of E. coli commensal and pathogenic isolates. J. Bacteriol. **190:**6881–6893. http://dx.doi.org/10.1128/JB.00619-08.
56. **Vieira G, Sabarly V, Bourguignon PY, Durot M, Le Fèvre F, Mornico D, Vallenet D, Bouvet O, Denamur E, Schachter V, Médigue C.** 2011. Core and panmetabolism in Escherichia coli. J. Bacteriol. **193:**1461–1472. http://dx.doi.org/10.1128/JB.01192-10.
57. **Wirth T, Falush D, Lan R, Colles F, Mensa P, Wieler LH, Karch H, Reeves PR, Maiden MCJ, Ochman H, Achtman M.** 2006. Sex and virulence in Escherichia coli: an evolutionary perspective. Mol. Microbiol. **60:**1136–1151. http://dx.doi.org/10.1111/j.1365-2958.2006.05172.x.
58. **Paul S, Linardopoulou EV, Billig M, Tchesnokova V, Price LB, Johnson JR, Chattopadhyay S, Sokurenko EV.** 2013. Role of homologous recombination in adaptive diversification of extraintestinal *Escherichia coli*. J. Bacteriol. **195:**231–242. http://dx.doi.org/10.1128/JB.01524-12.
59. **Hannan TJ, Totsika M, Mansfield KJ, Moore KH, Schembri MA, Hultgren SJ.** 2012. Host-pathogen checkpoints and population bottlenecks in persistent and intracellular uropathogenic Escherichia coli bladder infection. FEMS Microbiol. Rev. **36:**616–648. http://dx.doi.org/10.1111/j.1574-6976.2012.00339.x.
60. **Eto DS, Jones TA, Sundsbak JL, Mulvey MA.** 2007. Integrin-mediated host cell invasion by type 1-piliated uropathogenic Escherichia coli. PLoS Pathog. **3:**e100. http://dx.doi.org/10.1371/journal.ppat.0030100.
61. **Martinez JJ, Mulvey MA, Schilling JD, Pinkner JS, Hultgren SJ.** 2000. Type 1 pilus-mediated bacterial invasion of bladder epithelial cells. EMBO J. **19:**2803–2812. http://dx.doi.org/10.1093/emboj/19.12.2803.
62. **Le Trong I, Aprikian P, Kidd BA, Forero-Shelton M, Tchesnokova V, Rajagopal P, Rodriguez V, Interlandi G, Klevit R, Vogel V, Stenkamp RE, Sokurenko EV, Thomas WE.** 2010. Structural basis for mechanical force regulation of the adhesin FimH via finger trap-like beta sheet twisting. Cell **141:**645–655. http://dx.doi.org/10.1016/j.cell.2010.03.038.
63. **Yakovenko O, Sharma S, Forero M, Tchesnokova V, Aprikian P, Kidd B, Mach A, Vogel V, Sokurenko E, Thomas WE.** 2008. FimH forms catch bonds that are enhanced by mechanical force due to allosteric regulation. J. Biol. Chem. **283:**11596–11605. http://dx.doi.org/10.1074/jbc.M707815200.
64. **Schwartz DJ, Chen SL, Hultgren SJ, Seed PC.** 2011. Population dynamics and niche distribution of uropathogenic Escherichia coli during acute and chronic urinary tract infection. Infect. Immun. **79:**4250–4259. http://dx.doi.org/10.1128/IAI.05339-11.
65. **Walters MS, Lane MC, Vigil PD, Smith SN, Walk ST, Mobley HL.** 2012. Kinetics of uropathogenic Escherichia coli metapopulation movement during urinary tract infection. mBio **3**(1):e00303-11. http://dx.doi.org/10.1128/mBio.00303-11.
66. **Wada A, Kono M, Kawauchi S, Takagi Y, Morikawa T, Funakoshi K.** 2012. Rapid discrimination of Gram-positive and gram-negative bacteria in liquid samples by using NaOH-sodium dodecyl sulfate solution and flow cytometry. PLoS One **7:**e47093. http://dx.doi.org/10.1371/journal.pone.0047093.
67. **Hansen WL, van der Donk CF, Bruggeman CA, Stobberingh EE, Wolffs PF.** 2013. A real-time PCR-based semi-quantitative breakpoint to aid in molecular identification of urinary tract infections. PLoS One **8:**e61439. http://dx.doi.org/10.1371/journal.pone.0061439.
68. **Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ.** 1990. Basic local alignment search tool. J. Mol. Biol. **215:**403–410. http://dx.doi.org/10.1016/S0022-2836(05)80360-2.
69. **McDonald D, Price MN, Goodrich J, Nawrocki EP, DeSantis TZ, Probst A, Andersen GL, Knight R, Hugenholtz P.** 2012. An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. ISME J. **6:**610–618. http://dx.doi.org/10.1038/ismej.2011.139.
70. **Ludwig W, Strunk O, Westram R, Richter L, Meier H, Yadhukumar, Buchner A, Lai T, Steppi S, Jobb G, Förster W, Brettske I, Gerber S, Ginhart AW, Gross O, Grumann S, Hermann S, Jost R, König A, Liss T, Lüssmann R, May M, Nonhoff B, Reichel B, Strehlow R, Stamatakis A, Stuckmann N, Vilbig A, Lenke M, Ludwig T, Bode A, Schleifer KH.** 2004. ARB: a software environment for sequence data. Nucleic Acids Res. **32:**1363–1371. http://dx.doi.org/10.1093/nar/gkh293.
71. **Engelbrektson A, Kunin V, Wrighton KC, Zvenigorodsky N, Chen F, Ochman H, Hugenholtz P.** 2010. Experimental factors affecting PCR-based estimates of microbial species richness and evenness. ISME J. **4:**642–647. http://dx.doi.org/10.1038/ismej.2009.153.
72. **Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R.** 2011. UCHIME improves sensitivity and speed of chimera detection. Bioinformatics **27:**2194–2200. http://dx.doi.org/10.1093/bioinformatics/btr381.
73. **Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Peña AG, Goodrich JK, Gordon JI, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters WA, Widmann J, Yatsunenko T, Zaneveld J, Knight R.** 2010. QIIME allows analysis of high-throughput community sequencing data. Nat. Methods **7:**335–336. http://dx.doi.org/10.1038/nmeth.f.303.
74. **Fu L, Niu B, Zhu Z, Wu S, Li W.** 2012. CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics **28:**3150–3152. http://dx.doi.org/10.1093/bioinformatics/bts565.
75. **Willner D, Daly J, Whiley D, Grimwood K, Wainwright CE, Hugenholtz P.** 2012. Comparison of DNA extraction methods for microbial community profiling with an application to pediatric bronchoalveolar lavage samples. PLoS One **7:**e34605. http://dx.doi.org/10.1371/journal.pone.0034605.
76. **Warnes G, Bolker B, Lumley T.** gplots: various R programming tools for plotting data. R package version 2.6.0. http://cran.r-project.org/web/packages/gplots/.
77. **Chen J, Bittinger K, Charlson ES, Hoffmann C, Lewis J, Wu GD, Collman RG, Bushman FD, Li H.** 2012. Associating microbiome composition with environmental covariates using generalized UniFrac distances. Bioinformatics **28:**2106–2113. http://dx.doi.org/10.1093/bioinformatics/bts342.
78. **Rice P, Longden I, Bleasby A.** 2000. EMBOSS: the European molecular biology open software suite. Trends Genet. **16:**276–277. http://dx.doi.org/10.1016/S0168-9525(00)02024-2.
79. **Zerbino DR, Birney E.** 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. **18:**821–829. http://dx.doi.org/10.1101/gr.074492.107.
80. **Li H, Durbin R.** 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics **26:**589–595. http://dx.doi.org/10.1093/bioinformatics/btp698.
81. **Huson DH, Mitra S, Ruscheweyh HJ, Weber N, Schuster SC.** 2011. Integrative analysis of environmental sequences using MEGAN4. Genome Res. **21:**1552–1560. http://dx.doi.org/10.1101/gr.120618.111.
82. **R Development Core Team.** 2005. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
83. **Oksanen J, Blanchet F, Kindt R, Legendre P, Minchin P, O'Hara R, Simpson G, Solymos P, Stevens M, Wagner H.** 2013. vegan: community ecology package. R package version **2.0**–9. http://cran.r-project.org/web/packages/vegan.
84. **Thioulouse J, Chessel D, Dolédec S, Olivier J-M.** 1997. ADE-4: a multivariate analysis and graphical display software. Stat. Comput. **7:**75–83. http://dx.doi.org/10.1023/A:1018513530268.