Toward an Understanding of Changes in Diversity Associated with Fecal Microbiome Transplantation Based on 16S rRNA Gene Deep Sequencing

Dea Shahinas,a Michael Silverman,b Taylor Sittler,c Charles Chiu,d Peter Kim, Soft Emma Allen-Vercoee Scott Weese,f Andrew Wong,9 Donald E. Low,a,b,c,d and Dylan R. Pillaii

Department of Laboratory Medicine, Pathobiology, University of Toronto, Toronto, Ontario, Canada; Lakeridge Health, Oshawa, Ontario, Canada; Department of Laboratory Medicine, University of California, San Francisco, California, USA; Department of Mathematics and Statistics, University of Guelph, Guelph, Ontario, Canada; Department of Laboratory Medicine, Pathobiology, University of California, San Francisco, California, USA; Department of Molecular and Cell Biology, University of Guelph, Guelph, Ontario, Canada; Laboratory Veterinary College, University of Guelph, Guelph, Ontario, Canada; School of Computing, Queen’s University, Kingston, Ontario, Canada; Department of Microbiology, Mount Sinai Hospital, Toronto, Ontario, Canada; Ontario Agency for Health Protection and Promotion, Toronto, Ontario, Canada; and Department of Pathology and Laboratory Medicine, University of Calgary, Calgary, Alberta, Canada

ABSTRACT Fecal microbiome transplantation by low-volume enema is an effective, safe, and inexpensive alternative to antibiotic therapy for patients with chronic relapsing Clostridium difficile infection (CDI). We explored the microbial diversity of pre- and posttransplant stool specimens from CDI patients (n = 6) using deep sequencing of the 16S rRNA gene. While interindividual variability in microbiota change occurs with fecal transplantation and vancomycin exposure, in this pilot study we note that clinical cure of CDI is associated with an increase in diversity and richness. Genus- and species-level analysis may reveal a cocktail of microorganisms or products thereof that will ultimately be used as a probiotic to treat CDI.

IMPORTANCE Antibiotic-associated diarrhea (AAD) due to Clostridium difficile is a widespread phenomenon in hospitals today. Despite the use of antibiotics, up to 30% of patients are unable to clear the infection and suffer recurrent bouts of diarrheal disease. As a result, clinicians have resorted to fecal microbiome transplantation (FT). Donor stool for this type of therapy is typically obtained from a spouse or close relative and thoroughly tested for various pathogenic microorganisms prior to infusion. Anecdotal reports suggest a very high success rate of FT in patients who fail antibiotic treatment (90%). We used deep-sequencing technology to explore the human microbial diversity in patients with Clostridium difficile infection (CDI) disease after FT. Genus- and species-level analysis revealed a cocktail of microorganisms in the Bacteroidetes and Firmicutes phyla that may ultimately be used as a probiotic to treat CDI.

Antibiotic-associated diarrhea (AAD) is a widespread phenomenon in hospitals today. The most common opportunistic pathogen in the setting of AAD is the bacterium Clostridium difficile. The incidence of Clostridium difficile infection (CDI) has been increasing over the last decade, fueled in large part by a strain identified as North American pulstype (NAP) 1 or ribotype 027 (1–3). This strain drives outbreaks in institutional settings and is associated with increased morbidity and mortality in elderly patients (4, 5), with the most severe infections leading to colectomy (surgical removal of the colon) or death. In addition, up to 30% of patients are unable to clear the infection and suffer recurrent bouts of diarrheal disease.

Current guidelines from the Infectious Diseases Society of America suggest that the first-line therapy for CDI is metronidazole, which nonspecifically targets anaerobic bacteria. The second-line therapy for patients with refractory or severe CDI is vancomycin, an antibiotic with broad activity against Gram-positive bacteria. In certain instances, tapering regimens of vancomycin have been recommended (6), and newer drugs for CDI have recently been developed (7). Nevertheless, owing to significant failure rates with all conventional antibiotic regimens for recurrent CDI, clinicians have for several decades resorted to fecal bacteriotherapy or microbiome transplantation (8–14). This type of therapy represents a paradigm shift in treatment of infectious agents, with the objective of restoration of the normal microbiota rather than specific eradication of the pathogen using a conventional antibiotic. Donor stool for this type of therapy is typically obtained from a spouse or close relative and thoroughly tested for various pathogenic microorganisms prior to infusion. Anecdotal reports suggest a very high success rate of FT in patients who fail antibiotic treatment. These clinical results have been lauded in the lay press as a radically alternative approach with great success to antibiotic therapy (15–17).
It has been suggested that the CDI state is marked by reduced microbial diversity and that persistent disruption of commensal microbe-gut interactions may result in interference with the complex interactions between the host and the commensal gut microbiota of healthy individuals (11, 12, 18). We hypothesized that FT successfully restores the gut diversity absent in patients suffering from refractory CDI. In order to test this hypothesis, we used deep-sequencing technology to explore the human microbial diversity in patients with CDI disease after direct instillation of donor stool as a clinical therapeutic intervention. Even though they are very specific, traditional methods such as culture, cloning, or Sanger sequencing of full-length 16S rRNA following PCR are limited in their capacity to elaborate the microbial community present in the gut (19–23). The use of deep-sequencing technology to address this task provides superior depth and breadth of coverage, especially of as-yet-uncultured anaerobic microorganisms in the gut. A previous study has shown that pyrosequencing provided a more precise estimate of relative abundance and an improved confidence of detection in comparison to cloned full 16S rRNA sequence analysis (19).

We targeted the V5-V6 region of bacterial 16S rRNA. The V5-V6 region reads were classified to reference sequence-based operational taxonomic units (refOTUs), which were defined at 97% sequence identity to determine the represented taxa. In the past, unguided probiotic treatments with Lactobacillus casei and Lactobacillus bulgaricus administered empirically have been shown to be inferior to vancomycin and metronidazole therapy (24). Effective probiotic formulations to prevent and treat recurrent CDI will likely require a global understanding of the dynamic changes of the gut microbiome in response to stool transplantation. Here we demonstrate that dramatic shifts in microbial diversity and composition occur with fecal transplantation, resulting in restoration of bacterial diversity and richness. These results have significant implications for tailoring specific biological therapeutics for infections such as CDI in which the microbial community plays a significant, if not primary, role.

(These data were presented in part at the 2011 American Society for Microbiology General Meeting in New Orleans, LA.)

**RESULTS**

**Number of sequence reads and sequences.** Twenty-one stool samples in total were evaluated from patients undergoing FT as described previously (14). Out of the six patients in this sample set, the first FT of patients 5 and 6 failed. The second FT for patient 5 failed, while the second FT for patient 6 was successful. FT was considered successful if the patients did not have recurrent CDI for ≥4 months. From this sample set, we obtained 148,945 V5-V6 pyrosequencing reads with an average length of 210 bp. The number of reads per sample is outlined in Table 1. The sequenced reads were trimmed, filtered, and preprocessed as described in Text S1 in the supplemental material. The distance matrix was used with a 0.03-genetic-distance, furthest-neighbor threshold in mothur to designate the operational taxonomic units (OTUs) (25). For subsequent standardized comparative analyses, the OTUs were classified at the level of 97% sequence homology into 4,029 reference operational taxonomic units (refOTUs). Taxonomic classification was performed using the Ribosome Database Project (RDP) (minimum 80% bootstrap score) (26, 27). Because several OTUs had the same genus-level classification, indicating that these OTUs represent sub-OTU lineages, further classification at the species level was carried out using the metagenomics analysis server (MG-RAST) (28). From these sequence reads, a range of 72 to 154 genera was identified in the stool specimens. The average number of genera for pooled donors was 123, that for recipients pre-FT was 96, and that for recipients post-FT was 120. The two failed FTs resulted in recipients with an average number of genera of 90. Thus, clinically successful FT resulted in an average increase of 25% in genera for recipients post-FT compared to their pre-FT state (extraction of differential gene expression [EDGE] analysis, \( P < 0.05 \)) (see Table S5 in the supplemental material). The number of species that belonged to the genera with significant changes in abundance post-FT was 90. The successful donor sample sequence classification contained a median of 211 unique species, while the pre-FT and post-FT sample sequence classifications consisted of a median of 164 and 255 unique species, respectively (see Table S4).

**Statistical analyses of changes in microbial diversity after FT.** Rank abundance curves were used for refOTUs obtained from the V5-V6 pyrosequencing read data for all six patients (Fig. 1) to evaluate the distribution and abundance of taxa in rank order. Statistical assessment of variation in the abundance of refOTUs between participants or with respect to the different states was conducted with EDGE software with a false discovery rate of 7% and a \( P \) value of 0.05 (29). EDGE analysis of significant relative changes was selected here because it improves the overall performance of multiple significance tests for multivariate data (29, 30). Patients 1 to 3 demonstrated no increase in abundance from the pre-FT to post-FT state, whereas patients 4 to 6 did show an increase for the successful post-FT state. When all samples were pooled, a median of 132 V5-V6 refOTUs was found in the pre-FT state and a median of 157 V5-V6 refOTUs was found in the post-FT state. Rarefaction curves are a measure of the number of refOTUs observed as a function of sampling effort and establish the rapidity with which taxonomic richness increased with each V5-V6 tag sequence. The rarefaction curves for the six patients in the study (Fig. 2) reveal that except for patient 1, the diversity increased after successful FT. When all specimens were pooled,

<table>
<thead>
<tr>
<th>Patient no.</th>
<th>Transplant outcome(s)</th>
<th>No. of reads</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Successful</td>
<td>R1 Pre, 15,786; R1 post-S, 13,851; D1-S, 5,346</td>
</tr>
<tr>
<td>2</td>
<td>Successful</td>
<td>R2 Pre, 6,607; R2 Post-S, 7,339; D2-S, 6,977</td>
</tr>
<tr>
<td>3</td>
<td>Successful</td>
<td>R3 Pre, 12,971; R3 Post-S, 9,633; D3-S, 8,055</td>
</tr>
<tr>
<td>4</td>
<td>Successful</td>
<td>R4 Pre, 9,368; R4 Post-S, 9,767; D4-S, 3,982</td>
</tr>
<tr>
<td>5</td>
<td>First failed; second failed</td>
<td>R5 Pre, 3,975; R5 Post-F (first), 2,959; R5 Post-F (second), 3,727; D5-F (first), 3,131; D5-F (second), 3,408</td>
</tr>
<tr>
<td>6</td>
<td>First failed; second successful</td>
<td>R6 Pre, 5,017; R6 Post-F, 6,239; R6 Post-S, 4,677; D6-S, 6,130</td>
</tr>
</tbody>
</table>
V5-V6 tag abundance ranged between 1,903 and 6,319 unique reads, corresponding to a range of 72 to 154 higher taxonomic classifications (see Table S3 in the supplemental material). In an attempt to catalogue the ecological diversity of each stool specimen, observed taxon richness, the Shannon diversity index, and the Simpson diversity index were established for each patient pre- and post-FT compared to healthy donors (Fig. 3). When the medians for each category were examined, the failed donors had significantly lower richness ($P < 0.01$) and Shannon diversity ($P < 0.0001$) than did the majority of successful donors as assessed by a normal distribution nonparametric test.

We used principal coordinate analysis (PCoA) to analyze the variability in refOTUs among donor, pre-FT, and post-FT samples. The first three axes of the PCoA, plotted in Fig. 4A, were sufficient to account for 71.8% of the total variability, and notably, the variability on these axes was higher than interindividual variability between samples. The relative contribution to variability for each successive axis decreases, with the first five axes accounting for all nonrandom variability (Fig. 4B). The Pearson correlation coefficients for using 1, 2, or 3 axes were 0.87, 0.91, and 0.90, respectively. The PCoA plot also revealed a clustering of refOTUs into two groups consisting of (i) pre-FT or failed-FT samples and (ii) donor or successful-FT samples, suggesting that key genera are shared between donors and successful FTs. Clustering may be responsive to both differences in proportions and the presence or absence of sequence data.

Examination of changes in the microbiota at the phylum level following FT showed two general trends (Fig. 5). Patients 1, 3, and 5 demonstrated no change in the predominance of Proteobacteria...
post-FT, whereas patients 2, 4, and 6 revealed a replacement of Proteobacteria with Bacteroidetes post-FT. We examined quantitative changes in specific refOTUs and genera in response to FT. Normalized and classified refOTUs are shown in detail across the samples (Fig. 6A). The greatest number of refOTUs was shared between failed transplants and pretransplant samples (1,435 refOTUs) followed by the refOTUs shared between successful transplants and donor samples (1,189 refOTUs). On the other hand, the lowest number of shared refOTUs was that between successful transplant samples and the pretransplant samples (9 refOTUs). Post-FT, within Proteobacteria significant ($P < 0.05$, EDGE analysis) decreases in the abundance of Klyuyera, Raoultella, and Parasutterella species were noted. In contrast, for Bacteroidetes a significant increase ($P < 0.05$, EDGE analysis) in the abundance of Alistipes, Bacteroides, and Parabacteroides was documented post-FT (Fig. 6B). Further significant increases were noted in Firmicutes genera and families such as Lactobacillus, Subdoligranulum, Carnobacterium, Streptococcus, Enterococcus, unclassified Ruminococcaceae, and unclassified Lachnospiraceae. Further dissection of genera that increase post-FT at the species level revealed that the following species contribute significantly to this effect: Alistipes putredinis, Alistipes shahii, Bacteroides intestinalis, Bacteroides caccae, Bacteroides ovatus, Bacteroides sp. strain 2_1_22, and Clostridium lituseburense (Fig. 6C; see also Table S4 in the supplemental material). Pyrosequencing reads corresponding to Clostridium difficile were present pre-FT but not post-FT.

**Community balance changes in response to FT.** Changes in taxon abundance after successful FT were calculated based on pooled data. Donor stool samples were characterized by an equal abundance of the taxa Bacteroides, Alistipes, Bifidobacterium, and Enterobacteriaceae, whereas in the pretransplant stool samples, Bacteroides strains were outnumbered by a factor of 10 by unclassified Enterobacteriaceae, Pseudomonadaceae, and Escherichia/Shigella strains. In the post-FT samples, a marked expansion of Bacteroides, Alistipes, and Bifidobacterium was documented, while decreases in the abundance of Escherichia/Shigella and Parabacteroides were noted. The taxonomic composition of the recipients post-FT was highly dependent on the donor, with some recipients showing a resemblance to the donor's microbiome post-FT.
teroides and Faecalibacterium content and a corresponding decrease in Escherichia/Shigella were noted. The highest fold increases after successful FT (range, 16-fold to 100-fold) were observed in Bacteroides, Lactobacillus, Dorea, Roseburia, and Faecalibacterium, with the corresponding highest fold decreases (range, 4-fold to 20-fold) being observed in Enterobacteriaceae and unclassified Proteobacteria.

DISCUSSION
Refractory CDI can be severely debilitating and life-threatening. Patients often become dependent on chronic metronidazole or vancomycin therapy for symptom alleviation at great financial expense. Discontinuation of these antibiotics often leads to recurrent episodes of diarrhea from CDI. We reported previously on self-administered FT as an effective treatment modality for the cure of patients with refractory CDI (14). FT represents a paradigm shift in the approach to the treatment of CDI, aiming to restore microbial diversity rather than to eradicate the pathogen with antibiotics. While many studies have reported on the efficacy and clinical success of FT, only two have attempted to address the mechanistic basis for this success. The first used a biased approach by amplifying predetermined gut microbes using PCR (8). The second is a single FT case report where the authors used a traditional clone-and-sequence approach to describe changes in the microbiota (12). However, these experimental approaches may underestimate the relative abundance changes that take place in the gut microbiota.

Here we undertook metagenomic analysis of stool specimens from donors and recipients of FT. It has been previously shown that classification accuracy of microbial populations is largely dependent on the choice of the 16S rRNA gene region, sequencing technology, and sequence quality (19, 31). Dethlefsen et al. showed that even though the V3 region provided more unique sequences than did the V6 region, only 85% of V3 reads had exact matches in the database to facilitate classification, relative to 90% for the V6 region (19). In addition, Claesson et al. suggested that the V4 and V5 regions provide representative sequences for the gut microbiota (31). Therefore, we used the long reads of 454 sequencing to combine the V5-V6 regions in order to take advantage of these two important features that these variable regions bring to the CDI patient microbiota analysis. Amplification bias may occur with this method beyond 20 cycles of amplification, but as this is a comparative analysis of specimens, these effects should be minimal.

Statistical analyses of pooled deep-sequencing data from successful FT recipients revealed the dynamic changes in the microbial community that occur pre- and post-FT. At the level of phyla, the healthy gut microbiota consists predominantly of Bacteroidetes, whereas CDI patients in this study demonstrate an overabundance of Proteobacteria. When we measured the ecological diversity of the microbiota, the CDI disease state (refractory cases in this study) is marked by decreased richness and diversity. The failed post-FT samples highlight the importance of increased richness and diversity for a successful FT and eradication of CDI. In the CDI state, fewer taxa are present and they are present in greater abundance (Fig. 6B). Healthy post-FT recipients appear to “adopt” certain elements of the donor microbiota (Fig. 6B). Specifically, the healthy state corresponding to a successful FT is marked by an eradication of key Proteobacteria species and a corresponding restoration of key Firmicutes and Bacteroidetes species, which may well be critical to “outcompeting” or depriving C. difficile of fitness in some way. These data agree with a previous study using a clone-and-sequence strategy (32).

We note in CDI stools the relative paucity of taxa from Lachnospiraceae (formerly Clostridium cluster XIVa), which constitute the majority of butyrate-producing bacteria (19, 33, 34). Lactate breakdown into butyrate by key microorganisms plays a major role in host metabolism, nutrient digestibility, microbiota changes, and epithelial integrity as well as host defense (35–37). In the hindgut, butyrate acts directly on tissue development, repair,
Lactate accumulation has been previously associated with disease states such as gut resections and ulcerative colitis both in humans and in horses (34, 38). We speculate that the absence of lactate-utilizing bacteria such as Lachnospiraceae is one of the metabolic reasons leading to disequilibrium and perpetuation of refractory CDI (34). Consistent with this model, successful FT replenishes butyrate-producing bacteria that utilize lactate in the distal gut, along with Roseburia and Bacteroides (33, 34, 39). Roseburia, for example, is a difficult-to-culture, mutualistic, and cellulolytic bacterial genus that is important for plant polysaccharide breakdown. Roseburia spp. such as Roseburia intestinalis and Roseburia inulinivorans are able to use soluble oligosaccharides and polysaccharides to produce butyrate (52). The abundance of these species was on average increased post-FT. One might speculate that butyrate depletion and lactate buildup may be the reason why CDI results

FIG 4  (A) PCoA. PCoA represents multivariate analysis, which explores community similarity by measuring phylogenetic branch length (UniFrac) of 16S rRNA tag sequences. The dots represent the projected shadows of the samples on the respective two-dimensional planes (xy, blue; yz, red; xz, green). The x axis, y axis, and z axis represent 27.1%, 24.0%, and 20.7% of the intersample variability, respectively. R Pre, recipient pre-FT; R Post-F, recipient post-failed FT; D-S, donor of successful transplant; R Post-S, recipient post-successful FT; D-F, donor of failed transplant. (B) PCoA axes in relation to observed variability. The percentage of variability observed is depicted in relation to the first five axes. The dashed line at 3.26% represents the amount of variability expected for each axis based on random events.
in severe disease, particularly in a certain subset of patients being treated with antimicrobials and infected with hypervirulent C. difficile strains. Further studies of the metabolic consequences of changes in microbiota will be required to investigate this possibility.

One limitation of our study is that patients were being treated with vancomycin prior to FT. It has been demonstrated that vancomycin can alter the stool microbiota, and thus, decreases in diversity and richness prior to FT may be in part due to lingering effects of vancomycin as opposed to CDI per se (41). For example, it has been shown that vancomycin has activity against Bacteroides (42). Nevertheless, we observe the same paucity of Bacteroidetes and overexpansion of Proteobacteria seen by other studies of the CDI disease state in the absence of prior therapy with vancomycin (18, 43). Furthermore, by use of 16S rRNA gene deep sequencing, our study confirms and extends the depth and breadth of our understanding of the changes in microbiota observed by other groups conducting FT (11, 12). Despite the potential interindividual variability because each FT microbiome was from a unique donor, it was possible to draw predictable patterns of striking similarity in the gut microbial communities among the individuals following a successful FT using a combination of pooled specimen analysis and PCoA. Another concern is spurious PCR amplification from reagents used in the stool extraction. To address this, we noted experimentally that reagents alone can produce bacterial 16S rRNA gene sequence data corresponding to genera such as Leuconostoc and Lactobacillus, but in this comparative analysis, this should not affect interpretation of relative microbial changes. Finally, the use of technical replicates would have increased the confidence of the findings in this study.

Clear separation of pre-FT/failed-FT and donor/successful-FT groups suggests that the presence or absence of key taxa likely outweighs interindividual differences in determining whether or not an FT is successful. Even though variability in microbiota composition depends on historical exposures to microbes and antibiotics, the post-FT samples grouped quite closely with the successful donor samples, suggesting the presence of key genera required for recolonization and bacterial community resilience. Furthermore, FT replenishment of certain specialized bacteria appears to be durable for the most part and results in the restoration of key taxa following FT, as mirrored by an increase in richness and diversity consistent with other previously published reports (11). The apparent success of FT in restoring “normal” microbial diversity raises the question of whether it can be used for other clinical conditions. For example, perturbations of the distal gut microbiota and childhood antibiotic use have been previously implicated in chronic conditions such as asthma, life-threatening pseudomembranous colitis, cancer, and obesity (44, 45). It is possible that the microbial equilibrium is specifically altered by commonly used antibiotics such as the newer-generation fluoroquinolones which have increased activity against healthy Bacteroides, lower activity against Proteobacteria, and no activity against the epidemic North American pulsotype (NAP) 1 strain, which predominates currently (5, 40, 46, 47). The clinical implications of our findings for the development of a rational probiotic drug to treat CDI are significant. It is conceivable that a probiotic containing the species which increase significantly with FT may be sufficient to outcompete CDI and restore gut homeostasis. Future investigation will require a larger analysis of patients pre- and post-FT from randomized controlled trials and animal studies to confirm mechanistic hypotheses raised by our work.

MATERIALS AND METHODS
Participants, fecal transplantation, and specimens. Patient eligibility criteria, donor screening, and FT protocol were as previously described (14). Briefly, all patients (recipients) underwent a full history and physical examination. The study was approved by the Conjoint Health Research Ethics Board, Faculty of Medicine, University of Calgary (study ID, E-24404). Potential family member fecal donors were selected by the patient and were questioned for any of the following contraindications for donation: (i) history of gastrointestinal illness, including peptic ulcer disease, gastroesophageal reflux, irritable bowel syndrome, inflammatory bowel disease, or polyps; (ii) any malignancy; and (iii) antibiotic use or hospitalization within the past 3 months. Donor screening was conducted as previously described (19). FT was conducted using a home-based en-
ema protocol conducted by a family member with 250 ml of donor stool slurry as previously described (14). Patients discontinued chronic oral vancomycin suppressive therapy 24 to 48 h pre-FT. Pre-FT recipient stool specimens for analysis were obtained after discontinuation of vancomycin, on the morning of the transplantation procedure. For failed transplants, recipient posttransplantation samples were obtained on the sec-
ond day of diarrhea (≥3 watery bowel movements/day) recurrence, prior to the reintroduction of vancomycin. For successful cases, post-FT specimens were obtained 14 days postprocedure. Stool samples were collected using standard methods in sterile containers without preservatives. Samples were frozen at −20°C within 24 h of collection prior to DNA extraction.

**DNA extraction.** DNA extraction from 200 mg of each of the 21 stool specimens was performed by following the protocol of isolation of DNA from stool for pathogen detection available with the QIAamp DNA stool minikit (Qiagen, Valencia, CA).

**16S rRNA V5-V6 tag deep sequencing.** For the amplification of the 16S rRNA gene from the genomic DNA, amplicon fusion primers were designed for the multiplexing and deep-sequencing procedure. The primers consisted of a directional GS FLX titanium primer (forward and reverse), a four-base library key sequence (5’ TCGA 3’), a multiplex identifier (MID) sequence specific for each stool sample pre- and posttransplant, and the gene complementary primer comprising the V5 and V6 tag region (31, 48). A list of the sequences of these fusion primers is provided in Table S2 in the supplemental material. To amplify the 16S rRNA V5-V6 region sequences, 43 µl of Pfu Accuprim Supermix (Invitrogen, Carlsbad, CA) was mixed with 0.25 µM forward and reverse primers (final concentration) and 5 µl of the stool-extracted DNA. Thermocycling involved a 7-min degeneration step at 95°C and 35 cycles of degeneration (95°C for 30 s), annealing (57°C for 1 min), and elongation (68°C for 1 min) as well as a final extension step of 10 min at 68°C. After amplification, the products were purified using the QIAquick PCR purification columns (Qiagen, Valencia, CA) and were quantified using a NanoDrop spectrophotometer (Thermo Scientific, Waltham, MA). The amplicons were submitted to Genome Quebec for further quality control and 454 deep sequencing. The MID sequence was used for automated software identification of samples after pooling and deep sequencing. Sequence assembly and downstream analysis was conducted using the mothur package of algorithms (25).

**Reference databases and refOTU definitions.** Operational taxonomic units (OTUs) are units of taxonomy. They were set at 97% sequence identity to resolve the data to the species level. A reference sequence-based operational taxonomic unit (refOTU) was defined as containing all query reads sharing the same reference database sequence (or group of sequences) as their nearest neighbor (19, 32). The data shown represent relative abundance in relation to the total taxon present. To determine genera, the V5 and V6 tag sequences obtained were aligned to the SILVA 16S rRNA sequence database (49) with taxonomic classifications determined at the highest taxonomic level available (44, 45, 50, 51).

To determine phyla, the V5 and V6 tag sequences obtained were aligned to the RDP (43) with taxonomic classifications determined at the highest taxonomic level available (44, 45, 50, 51).

**Statistical calculations.** The comparisons of the nearest database distance to tag abundance, the rank abundance plots, the rarefaction curves, and the diversity and richness summary statistics were calculated using absolute tag counts. All these analyses were performed using mothur (25). Principal coordinate analysis of unweighted refoOTUs was also conducted using the PCoA function of the mothur pipeline (25). To determine the statistical significance of variation in the abundance of taxa post-FT, the significance analysis using the optimal discovery procedure (26) was used with a false discovery rate of 7% and a P value of 0.05 (see Table S3 in the supplemental material). Data from SILVA and RDP were used for the EDGE significance analysis. The mothur package and the EDGE optimal discovery procedure were chosen for the comparison of community memberships and structures on the basis of the thoroughness of the results previously published for this type of analysis with pooled data using the refoOTU approach (19). Newer pipelines that make use of mothur and updated databases have become available recently, and future studies will have to compare results between these algorithms in order to ensure consistent interpretation of sequence data (25). The OTUs were normalized to relative abundance by dividing read counts relative to sample size. Log2-transformed normalized abundance was used to determine the significance of relative changes pre- and post-FT at the genus level. Significantly increasing or decreasing genera were identified at the species level using MG-RAST (29).

**SUPPLEMENTAL MATERIAL**

Supplemental material for this article may be found at http://mbio.asm.orglookup/suppl/doi:10.1128/mBio.00338-12/-/DCSupplemental.

Text S1, DOCX file, 0.1 MB.
Table S1, PDF file, 0.1 MB.
Table S2, PDF file, 0.1 MB.
Table S3, PDF file, 0.1 MB.
Table S4, XLSX file, 0.1 MB.
Table S5, XLSX file, 0.1 MB.

**ACKNOWLEDGMENTS**

We thank the patients, clinical microbiology staff, and administrative staff who supported this work. We thank Patrick Schloss of the University of Michigan for reading the manuscript and providing valuable comments. We have no conflict of interest to declare. Funding for this study was provided by a grant from the OPS Innovation Fund (D.R.P.).

**REFERENCES**

14. Silverman MS, Davis I, Pillai DR. 2010. Success of self-administered...