Linking Strain Engraftment in Fecal Microbiota Transplantation With Maintenance of Remission in Crohn's Disease

Lingjia Kong, Jason Lloyd-Price, Tommi Vatanen, Philippe Seksik, Laurent Beaugerie, Tabassome Simon, Hera Vlamakis, Harry Sokol, Ramnik Xavier

Version of Record: https://www.sciencedirect.com/science/article/pii/S0016508520351167
Manuscript_27d97f8dc24553a2f48185070860715c

Linking strain engraftment in fecal microbiota transplantation with maintenance of remission in Crohn's disease

Authors
Lingjia Kong[1,10], Jason Lloyd-Price[1,†], Tommi Vatanen[1,2], Philippe Seksik[3,4,5,6], Laurent Beaugerie[4,6], Tabassome Simon[7,8], Hera Vlamakis[1], Harry Sokol[3,4,5,6,*], Ramnik J. Xavier[1,9,10,*]

Affiliations
[1]Broad Institute of MIT and Harvard, Cambridge, MA, USA

[2]Liggins Institute, University of Auckland, Auckland, New Zealand

[3]Centre de Recherche Saint-Antoine, CRSA, AP-HP, Hôpital Saint Antoine, Service de Gastroenterologie, Sorbonne Université, Inserm, 75012, Paris, France.

[4]Department of Gastroenterology, Saint Antoine Hospital, Assistance Publique-Hopitaux de Paris (APHP), 184 rue du Faubourg Saint-Antoine, 75571, Paris, CEDEX 12, France.

[5]French Group of Fecal Transplantation (GFTF), Paris, France.

[6]Paris Center for Microbiome Medicine (FHU PaCeMM), Paris, France.

[7]Clinical Research Platform (URC-CRC-CRB), AP-HP Saint-Antoine Hospital, Paris, France.

[8]Department of Clinical Pharmacology, APHP, Saint Antoine Hospital, Paris, France.

[9]Center for Microbiome Informatics and Therapeutics, Massachusetts Institute of Technology, Cambridge, MA, USA

[10]Center for Computational and Integrative Biology and Department of Molecular Biology, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA

*Co-corresponding authors.
†Now at Google LLC, New York, NY 10011, USA

Grant support
The clinical trial was supported by a grant from Programme Hospitalier de Recherche Clinique - PHRC PHRCR-13-029 (Ministère de la Santé), Assistance Publique – Hôpitaux de Paris (CRC16), Fondation de France (fond Inkermann), and Association Francois Aupetit.
H.S. received funding from the European Research Council (ERC) under the European Union's Horizon 2020 Research and Innovation Programme (ERC-2016-StG-71577). R.J.X. received funding from Center for the Study of Inflammatory Bowel Disease (CSIBD, P30DK043351), The Crohn's & Colitis Foundation, and NIH (AT009708).

Correspondence
harry.sokol@aphp.fr; xavier@molbio.mgh.harvard.edu

Disclosures
R.J.X. is a consultant to Novartis and Nestle.
H.S. received consultant, board membership or lecture fees from Carenity, Abbvie, Astellas, Danone, Ferring, Mayoly Spindler, MSD, Novartis, Roche, Tillots, Enterome, Maat, BiomX, Biose, Novartis, and Takeda. He is also a co-founder of Exeliom bioscience. L.B. received consulting fees from Janssen, Pfizer and Takeda; lecture fees from Abbvie, Janssen, MSD, Ferring Pharmaceuticals, Mayoly-Spendler, and

Takeda; research support from Abbott, Ferring Pharmaceuticals, Hospira-Pfizer, Janssen, MSD, Takeda and Tillots. P.S. received consulting/lecture fees or grant funding from: Takeda, Abbvie, Merck-MSD, Biocodex, Janssen, Amgen, Astellas, Pfizer

Author contributions

L.K. and J.L.-P. carried out the microbiome data analysis and interpretation. T.V. performed data preprocessing and quality control. H.S., P.S., L.B and T.S. initiated the project and collected clinical samples. H.S. and R.J.X served as principal investigators. J.L.-P., L.K., T.V., H.V., and R.J.X. drafted the manuscript. All authors revised the article and approved the final version for publication.

Acknowledgements

We thank Luke Besse for project management and making data available through the SRA and the Broad Institute Genomics Platform for sample processing and sequencing data generation. We also thank Heather Kang for editorial assistance and Melanie Schirmer for helpful discussions.

Abstract

Background and aims

Crohn's Disease (CD) is a chronic gastrointestinal disease resulting from the dysfunctional interplay between genetic susceptibility, the immune system and commensal intestinal microbiota. Emerging evidence suggests that treatment by suppression of the immune response and replacement of the microbiota through Fecal Microbiota Transplantation (FMT) is a promising approach for the treatment of CD.

Methods

We obtained stool metagenomes from CD patients in remission and assessed gut microbiome composition before and after FMT at the species and strain levels. Longitudinal follow-up evaluation allowed us to identify the gain, loss, and strain replacement of specific species, and link these events to the maintenance of remission in CD.

Results

We find that FMT had a significant long-term effect on patient microbial compositions, though this was primarily driven by the engraftment of donor species which remained at low abundance. 38% of FMT-driven changes were strain replacements, emphasizing the importance of detailed profiling methods such as metagenomics. Several instances of long-term coexistence between donor and patient strains were also observed. Engraftment of some Actinobacteria, and engraftment or loss of Proteobacteria, were related to better disease outcomes in CD patients who received FMT, while transmission of Bacteroidetes was deleterious.

Conclusion

Our results suggest clades that may be beneficial to transmit/eliminate through FMT, and thus provide criteria that may help identify personalized FMT donors to more effectively maintain remission in CD patients. The framework established here creates a foundation for future studies centered around the application of FMT and defined microbial communities as a therapeutic approach for treating CD.

Keywords: Crohn's Disease (CD), Fecal Microbiota Transplantation (FMT), relapse, metagenomes

Introduction

Crohn's Disease is one type of inflammatory bowel disease (IBD), which causes chronic relapsing and remitting inflammation of the gastrointestinal tract. The incidence and prevalence of CD is increasing worldwide[1]. While infrequently fatal, CD has a profound influence on quality of life. The pathogenesis of CD is complex, involving the interplay between the immune system and a dynamic and dysbiotic microbial community[2,3]. Inflammation in CD results from a dysregulated immune response which is accompanied by community-wide changes in the commensal microbiome. This is most readily seen in a reduction in community diversity during disease activity, largely brought on by a loss of obligate anaerobes such as *Faecalibacterium prausnitzii* and *Roseburia intestinalis*. Many of these species are considered to be beneficial producers of anti-inflammatory compounds such as short-chain fatty acids, in particular butyrate. Concurrent with this loss, an increase is observed in facultative anaerobes which are opportunistic pathogens, such as *Escherichia coli* and *Haemophilus parainfluenzae*. *Akkermansia muciniphila* has also

been observed to bloom during CD activity, and is hypothesized to contribute to the degradation of the mucous layer in the gut triggering a larger immune response.

Management of CD currently revolves around treating symptoms, by calming the overactive immune system with immunosuppressants. Frontline treatments include corticosteroids, aminosalicylates, and biologics. Long-term use of these therapies come with significant side effects for patients, and can lead to complications including infections[4] and cancers[5]. Furthermore, primary non-response or secondary loss of response occurs frequently in patients treated with biologics such as anti-TNF agents and vedolizumab[6]. Given the immune-microbiota feedback loop in CD, therapeutic approaches that additionally target the microbiota, such as Fecal Microbiota Transplant (FMT), are gaining interest[7–9]. FMT has already been used successfully to treat other microbially-linked diseases, notably infections of recurrent *Clostridium difficile* in hospitals worldwide[10–12], including among IBD patients[7,8,13]. Microbiome modifications are particularly attractive since, rather than solely treating current symptoms, they have the potential to reduce the risk of relapse long after treatment[14].

The effectiveness of FMT has not been extensively studied in IBD. Most FMT studies have been done in patients with ulcerative colitis (UC), where the results have been mixed, with some studies finding that FMT is beneficial[15–17], and at least one observing no effect[18]. One recent pilot study in CD followed 17 patients for 24 weeks after FMT or sham transplantation[19], measuring microbial community changes by 16S rRNA gene profiling. This study aimed to jointly treat immune dysfunction and microbial dysbiosis by performing FMT after inducing clinical remission using corticosteroids and performing a colon cleansing with polyethylene glycol to facilitate donor microbiota engraftment. FMT was found to significantly improve clinically-relevant disease severity metrics, including C-reactive protein levels and Crohn's Disease Endoscopic Index of Severity. Greater donor microbiota engraftment was also associated with a reduction in the rate of relapse. In particular, two patients who were identified as having "failed FMT"s based on their 6-week similarity to the donor microbial composition eventually relapsed.

Here, from metagenomic data generated from the aforementioned pilot study[19], we first characterize changes in microbiome composition over the 26 week study period, and identify changes induced by FMT both at the species and strain levels which was not possible with 16S rRNA gene sequencing data. We then quantify disease activity-induced changes in patients that eventually relapsed. Finally, we link FMT-induced changes with disease outcome, and identify microbial clades that are both beneficial and antagonistic to the probability of relapse.

Results

Study overview

We performed deep shotgun metagenomics to analyze 115 samples, originally collected in Sokol et al[19], for changes in the microbiota (**Fig. 1a**). Among these are samples from 8 patients who received FMTs from 5 donors (each patient received stool from one donor, and samples from two donors were administered to more than one patient), and 9 patients who received sham transplantation. In total, 7 patients relapsed (2 from the FMT group, and 5 in sham; difference is not significant), though the two patients which relapsed in the FMT group both showed an absence of donor microbiota engraftment[19]. Samples were taxonomically

profiled using MetaPhlAn2, and principal coordinate analysis of these species abundance Bray-Curtis dissimilarities indicates the major patient differences are age and gender (**Fig. S1**), and FMT effects are not apparent in the first two axes of variation (**Fig. 1b**). Patient profiles tended to be consistent over time (**Fig. 1c, Fig. S4-5**). Unlike many western gut datasets, we found that Actinobacteria and Firmicutes were the most prevalent, while Bacteroidetes is considerably less prevalent (**Fig. S2**). *Bifidobacterium* and *Faecalibacterium* make up the most abundant genera (**Fig. 1c, Fig. S3**).

FMT induces long-term changes in patients' microbiota

We found that the diversity of species composing the communities (alpha diversity) transiently increases following FMT (**Fig. 2a**), as previously observed with 16S data[19], but observed no effect beyond 14 weeks with metagenomics data. Sorensen similarity between patient and donor communities showed a significant shift towards the donor community following FMT, which was persistent across the timeframe of the study. This shift is visible, but not as significant in Bray-Curtis similarities, indicating that the changes observed with the Sorensen similarity were driven by donor species which successfully colonize the patient, but remain at low abundance.

At the strain level (**Methods**), we found further evidence of replacement of patient strains with donor strains (**Fig. 2b**). Specifically, the dominant strain haplotype in patients frequently shifted towards the donor's haplotype after FMT (**Fig. 2c**), as is exemplified by *Bacteroides dorei* (**Fig. 2d**). Engraftment rates varied between patients, with three patients exhibiting >40% of species transmitted. These included both 1_9 and 1_17, for which 6 and 5 species engrafted, respectively (out of 12 and 11 species, respectively, with strain profiles). While both of these patients were previously identified as failed transplantations and did not show a large shift in community taxonomic composition towards their donors[19], our results show clear evidence of partial transmission by strain replacement and/or coexistence. One other patient, 1_15, showed engraftment by 8 of 19 species. We found that some species engrafted more frequently than others, including *Alistipes putredinis* and the butyrate producer *Coprococcus comes* (**Fig. 2b**). Others were more resistant such as the *Ruminococci*. For example, *R. torques*, a known IBD-related species[3], was prevalent but did not engraft in any patient. *F. prausnitzii*, considered a beneficial species[20], did not show clear evidence of engraftment in spite of its prevalence in both donors and patients (**Fig. S4-5**).

Among the clearest engraftment events (**Table S1**), we observed numerous instances where the donor strain took more than two weeks to engraft (low donor-patient and within-patient distances). These included instances of long-term coexistence of donor and patient strains, such as *Eubacterium halii* (**Fig. 2e**), where the dominant haplotype at weeks 2 and 24 in patient 1_22 matched that of the donor, while weeks 6 and 10 matched the patient's original strain. Despite its low abundance in this study, we note numerous examples of engraftment of Bacteroidetes species (**Table S1**), a phylum considered the most stable within-individuals[21]. Finally, we observed one clear engraftment event of *Dialister invisus*. Although typically considered an oral species, long-term engraftment by FMT shows that this species has successfully colonized the patient's GI tract.

We observed changes in the microbial community's functional potential after FMT (**Table S2**). These included an increase in a creatinine degradation pathway (**Fig. 2f**). Creatinine, a uremic toxin generated in muscles and primarily cleared through the kidney, is partially metabolized by the gut microbiota under

normal conditions[22]. Increased creatinine degradation by the gut bacteria may therefore contribute to lower serum creatinine.

Relapse-associated changes

We observed a decrease in community diversity at the first time point after relapse (**Fig. 3a**), with a recovery by the next time point. This is consistent with a loss/reduction of commensal microbial community members during CD disease activity, particularly due to the oxygenated environment of the inflamed colon. Consistent with this, we also observed a depletion in community potential for anaerobic energy metabolism (FDR = 0.045, **Fig. 3b**). Other functional differences included depletion of the NAD biosynthesis and tRNA charging pathways, two core metabolic functions (all results in **Table S2**).

Changes in FMT that affect relapse

We next assessed FMT-related changes that may affect the probability of relapse. From a joint analysis of species and strain profiles, we first categorized each species/patient pair based on the evidence for species gain/loss/strain replacement in FMT (**Methods**; **Fig. 4a**). We found evidence for 143 total engraftment events (88 gains and 55 replacements), with few losses (18 total, i.e. patient did not carry the species after FMT, but did before). This amounts to 20 changes per patient, on average, with a minimum of 12, implying that even the microbiomes of patients classified as "FMT failure" were profoundly impacted by the procedure (14 changes for 1_17 and 27 for 1_9; **Fig. 4b, right**). Losses were also observed in the sham group (20 total, or 2.5 losses per patient; one sham patient did not have a pre-FMT sample, so is not counted here), likely due to the bowel cleansing procedure prior to FMT. The majority of engraftment events in FMT patients were for Firmicutes species, though Bacteroidetes had the most gain events total (**Fig. 4b, left**), exemplified by *Barnesiella intestinihominis* with 3 gain events (**Fig. 4c**). This rate was observed despite the low relative abundance of Bacteroidetes species in this study, echoing the strain-level results above and implying a much greater rate of engraftment for Bacteroidetes than for Firmicutes (**Fig. 4b, left**). Strain replacements were not limited to low-abundance species, with 25 of the 55 replacements for species with >1% mean relative abundance before and after FMT (**Table S3**), primarily for Firmicutes (15 events) and Actinobacteria (8 events).

To connect engraftment in FMT with probability of remission maintenance, we scored each species based on the number of engraftment events in patients that maintained remission versus those that ultimately relapsed (**Fig. 4a**). We found that the phyla are non-randomly distributed within this ranking. In particular, engraftment of Proteobacteria (Wilcoxon test p = 0.028) and Bacteroidetes (p = 0.032) are associated with the likelihood of relapse. High-ranking Proteobacteria in this list included *Sutterella wadsworthensis*, *Haemophilus parainfluenzae*, and *Escherichia coli*, all of which included losses from patients that did not relapse, consistent with these species being antagonists [23]. However, there were three instances of patients gaining *S. wadsworthensis* in FMT. The majority of Bacteroidetes species transmitted frequently in patients who relapsed, thus these largely had a negative influence on maintenance of remission. However, the top species in this ranking included *Bacteroides massiliensis*, which was gained by two patients in FMT and did not have any abundance in any of the 5 patients that relapsed in sham (**Fig. 4a**). *Prevotella copri* is also high in the ranking. The *Prevotella* genus tends to be irregularly distributed within western populations, with low prevalence but high abundance when present, making it unusual how frequently it was transmitted here. In non-western populations with lower IBD incidence, these also tend to be more prevalent. Also,

none of the sham patients carried *P. copri*. Actinobacteria were marginally significantly non-randomly distributed ($p = 0.041$). The highest-ranking species, *Bifidobacterium longum*, is highly abundant and prevalent in this study, and was gained once and replaced with the donor strain three times in FMT patients that did not relapse. The donors for the two FMT patients that relapsed did not carry *B. longum* (**Fig. S4**). Notably, despite its prevalence and abundance in both donors and recipients in this study, we observed only a single instance of engraftment of Faecalibacterium prausnitzii. This engraftment was not associated with a beneficial outcome as this patient eventually relapsed. Therefore, this commonly-cited beneficial butyrate-producer which is heavily depleted in CD activity[24].

Performing strain-level PERMANOVA tests, we found that *Roseburia intestinalis* has the strongest association with probability of relapse (**Fig. 4d**; PERMANOVA $R^2 = 0.15$, $p = 0.042$; **Table S4**). However, this species transmitted poorly in FMT, with only one clear engraftment event, and was not significantly associated with FMT at the strain level ($p = 1$). Species with significant strain-level associations with FMT included *Alistipes putredinis* ($p < 0.001$), two *Dorea* species (*D. longicatena* $p = 0.003$ and *D. formicigenerans* $p = 0.002$), and two *Coprococci* (*C. cactus* $p = 0.013$ and *C. comes* $p < 0.001$). All of these species are depleted in IBD disease activity[3,25].

Discussion

In this study, we examined whether changes to the gut microbial community in FMT were associated with disease activity in CD. This is the first study to use metagenomics to obtain species- and strain-resolution profiling for FMT in IBD. We found that species-level community composition experienced a transient increase in diversity following FMT. Long-term compositional changes, however, were primarily driven by the acquisition of donor species which remained at low abundance, and the abundance distribution of donor species only weakly carried over to patients. 55 instances of strain replacements were also observed, accounting for 38% of engraftment events (143 total). Many of these occurred in patients which were previously classified as FMT failures, showing that broader profiling methods such as 16S miss a non-negligible fraction of the differences occurring in FMT. We further found multiple instances of long-term coexistence of donor and patient strains, up to the final follow-up 24 weeks after FMT.

Despite the additional evidence for post-FMT microbiome shifts, total replacement of the microbiota was not achieved in any patient, even with bowel cleansing performed prior to FMT. Consistent with this, FMT to treat recurrent infection by *C. difficile*[7] has a reduced impact on microbiome composition in IBD patients compared to non-IBD patients. This indicates that either the existing microbial communities of IBD patients are more resilient to change, or the host environment continues to actively select for the original community. Future studies should investigate how to increase the efficiency of engraftment in IBD specifically, as this may be necessary to obtain larger effects on disease activity. Engraftment potential is likely dependent on the unoccupied ecological niches in the patient pre-FMT, e.g. for *B. longum*[26]. The bowel cleansing prior to FMT here was intended to "reset" the microbiome and clear existing ecological niches, though this was not sufficient to obtain complete replacement. Other methods may include repeated rounds of FMT[17,27] or complementation with targeted dietary interventions[28,29].

Whole-microbiome replacement may not be necessary if we can determine a subset of microbes that reduces the probability of relapse. Given the complex interplay between the host and microbiome in IBD, it is

unlikely that single species will have such an effect, though some sub-communities of microbes may collectively confer protective benefits. We found a significant association between engraftment of three phyla, Actinobacteria, Proteobacteria, and Bacteroidetes and future disease relapse. Engraftment of Actinobacteria and engraftment or loss of Proteobacteria positively impacted disease outcome. Proteobacteria contains several clades that bloom in the inflamed gut[23], and species which were affected in FMT here included *S. wadsworthensis*, *H. parainfluenzae*, and *E. coli*. Changes to this phylum, particularly loss of species, may therefore confer some benefit. Engraftment of Bacteroidetes negatively impacted disease outcome. Despite its general low abundance in this study, we found evidence of widespread Bacteroidetes engraftment. This phylum contains species with persistent inter-individual variation, and thus constitutes an individual's personal core microbiome[21]. We observe that the two FMT patients that eventually relapsed primarily received Bacteroidetes species from their donors, suggesting that widespread replacement of Bacteroidetes species may not be beneficial. These species, therefore may provide a "donor compatibility" signature that could be used to help predict whether a patient's response to FMT will be pro- or anti-inflammatory[11]. Larger studies will be needed to confirm this, though evidence from a previous study with 38 patients indicated that overall relatedness of the original recipient community with the donor community is not predictive of relapse[7]. Still, given the large inter-personal differences in microbiomes, it may be possible to identify criteria by which patients can be carefully matched with donors in order to elicit a more precise set of changes through FMT.

In summary, our results show that subsets of the engrafted microbiome have a measurable impact on disease activity after FMT, both positive and negative. These effects appear to be patient-specific, suggesting that some donors (i.e. so-called "super-donors") can more effectively reduce disease activity[30]. The small number of study participants limits the resolution at which disease-associated features can be identified, and these results will need to be validated in larger studies. Future studies should aim to use high-resolution measurement techniques such as metagenomics, as our results show that strain replacement was pervasive in FMT, and the effects of FMT were more profound than originally measured. We further found that the post-FMT perturbation to the patient's microbiome settles into its long-term state starting from the 14-week time point. Future studies may therefore benefit from focusing their sampling before this point to more efficiently allocate samples. The framework established here creates a foundation for future studies centered around FMT as a therapeutic approach for treating CD.

Methods

Sample collection, preparation and sequencing

Stool samples were collected as described in Sokol, et al[19]. Nucleic acid was extracted using the AllPrep 96 PowerFecal DNA/RNA kit from QIAGEN (custom product #1114341). This method pairs bead-beating on a Tissuelyser II (QIAGEN) with a 96 well AllPrep protocol. Purified DNA was stored at -20°C. Illumina sequencing libraries were prepared from 2ng of input DNA using the Nextera XT DNA Library Preparation kit (Illumina) according to the manufacturer's recommended protocol. Prior to sequencing, libraries were pooled by collecting equal volumes of each library. Insert sizes and concentrations for each pooled library were determined using an Agilent Bioanalyzer DNA 1000 kit (Agilent Technologies) prior to sequencing on an Illumina NovaSeq 6000 with 151bp paired-end reads to yield ~10 million paired end reads per sample. Data was analyzed using the Broad Picard Pipeline which includes de-multiplexing and data aggregation (https://broadinstitute.github.io/picard).

Profile generation

Quality control for metagenomic shotgun sequencing data was performed with KneadData v0.7.2, with additional adapter detection and trimming at a minimum overlap of 5 bp by Trim Galore!. Three of 115 samples were excluded due to low read count. Taxonomic profiles were generated using MetaPhlAn v2.7.7[31]. Functional profiles were generated with HUMAnN2 v0.11.2[32], providing gene family level (here, 90% similarity) quantifications of microbial genes that are further stratified by contributing organisms. The gene families were further mapped to MetaCyc pathways[33].

Strain analysis

Strain SNP haplotypes were generated using StrainPhlAn[34] with preset "relaxed_parameters3" settings. Strain-level changes in FMT (**Fig. 2b**) were characterized from the dominant haplotype sequences in donors and patients. Briefly, Kimura 2-parameter distances (using the *ape* R package) were first normalized within species by dividing by the mean distance between all sample pairs for that species. We then calculated the minimum distance from each donor to their matched patient before FMT (time points: -2w, 0w) and after FMT (time points: 2w, 6w, 10w, 14w, 18w, 24w), with missing values excluded. Species are only included if at least one such measurement was possible for the species. A successful engraftment event was defined as events where: the after-FMT donor-patient distance was 0.25 lower than before FMT, the before-FMT distance was greater than 0.5, and the after-FMT distance was smaller than 0.5. We applied the same thresholds to the sham group, and found that this results in a 7% false positive rate (**Fig. S6**).

Phylogenetic trees (**Fig. 2d-e**) were generated based on the StrainPhlAn SNP haplotypes using the *phangorn* R package[35], using the Jukes and Cantor (JC69) model. Briefly, an initial tree was constructed using Unweighted Pair Group Method with Arithmetic Mean (UPGMA) hierarchical clustering. The tree was then optimized using maximum likelihood, by iterative optimization of edge lengths, base frequencies and topology. Visualizations were generated with the *ggtree* R package.

Differential abundance testing

Differential abundance testing was performed with MaAsLin2[36] v1.0.0, using arc-sin square root transformation of abundances. The model included sex, age, whether FMT had occurred (1 after week 0 in FMT patients, 0 otherwise), and relapse, with patient as the random effect. P-values are from Wald tests, and multiple hypothesis correction was performed with the Benjamini-Hochberg false discovery rate method.

Engraftment event identification and quantification

We built an integrated view of changes in FMT by categorizing each species/patient pair based on the evidence of changes in both the abundance and strain profiles. MetaPhlAn2 abundance profiles were first filtered to the set of species with prevalence >10% (abundance > 0.001), and low abundance values (<0.0001) were clamped to 0. StrainPhlAn profiles were limited to species for which a MetaPhlAn2 profile exists. Let b be the abundance of a species before FMT, let a be the abundance after FMT, let d be the donor abundance, and s_b and s_a be the normalized Kimura 2-parameter distance of the donor strain to the patient's strain before and after FMT, respectively. Filtered abundances and strain measurements were then grouped into five categories:

- **Gain:** b = 0, a > 0, d > 0

- **Loss:** b > 0, a = 0
- **Replace:** a > 0, b > 0, s_b - s_a > 0.25
- **Inconclusive:** a > 0, b > 0, s_b or s_a not measured
- **No change:**
    - b > 0, b > 0, s_b - s_a <= 0.25
    - b = 0, a = 0, s_b and s_a not measured
- **Inconsistent:**
    - b = 0, a > 0, d = 0
    - b = 0, a = 0, s_b or s_a is measured

Species were sorted according to the association between their changes in FMT with whether or not they relapsed. For this, a species score was defined as:

$$S(x) = \sum_i v_{x,i} p_i$$

Here, $v_{x,i}$ denotes a score associated with the engraftment of species x in patient i as determined above, and is 1 for a gain or loss, 0.5 for a replacement, -1 for no change, and 0 otherwise. p_i denotes the patient weight, and is 1 for non-relapsing FMT patients and -2.5 for FMT patients that eventually relapsed (weights chosen such that relapse and non-relapse have equal total weight).

PERMANOVA

PERMANOVA was performed with the *adonis* function in the R package *vegan* (version 2.5-6). Dissimilarity matrices were obtained from the Kimura 2-parameter distance between dominant haplotypes returned by StrainPhlAn for each species. When a species did not have sufficient coverage for haplotype calling in StrainPhlAn, its dissimilarity with all samples which had sufficient coverage was imputed with the 90th percentile of the strain dissimilarity matrix. Samples with the species were therefore considered to be very different from samples without the species. Variables were introduced in the model one at a time. The R^2 statistic should thus be considered the maximum variance explainable by that measure. Repeat measures were accounted for as in Lloyd-Price *et al.*, 2019[3]. Specifically, permutations of time-varying features (FMT and relapse) were performed within-patient, and permutations of patient-specific features (sex and age) were permuted between patients and samples were relabeled with the patient's permuted metadata.

Data availability

The datasets generated during and/or analyzed during the current study are available in the Sequence Read Archive (https://www.ncbi.nlm.nih.gov/sra) repository, under BioProject PRJNA625520.

Figure legends

Figure 1. FMT study design. (a) 115 samples were collected from 9 sham patients (blue), 8 FMT patients (yellow), and 5 donors (numbered 1-5, light brown) over 24 weeks. Relapse (black triangles), and samples that failed QC (grey dots) are also marked. At week 0, FMT patients received fecal transplantation and sham patients received sham transplantation [19]. **(b)** Principal coordinates analysis of Bray-Curtis dissimilarities from species-level MetaPhlAn2 profiles. Lines connect samples from the same

patient/donor. Separate samples were obtained at different times for the donor with 3 patients (7_2). **(c)** Barplot of species abundances over time in one FMT patient (1_19) and their donor (7_2_D7).

Figure 2. FMT alters the patient's microbiome. **(a)** Difference in species-level Shannon index from week 0 (top) shows a transient increase in diversity after FMT, but not in sham. Species-level Sorensen index (middle) shows successful and persistent engraftment of donor species in FMT patients, which is less apparent for Bray-Curtis similarity (bottom). *p-value<=0.05; **p-value<=0.001. **(b)** Heatmap of the minimum normalized Kimura 2-parameter distance (Methods) between donor and patient before FMT (-2w, 0w; left) and after FMT (2-24w; right). **(c)** Scatter plot of the minimum normalized Kimura 2-parameter distance between patient and their donor before and after FMT. Each point represents a patient/species pair. Points are colored by the distance between patient's week -2 to week 2 strains (grey indicates a missing measurement). The 1:1 line (dashed black) indicates no changes in FMT. Thresholds for defining the top engraftment events (**Table S1**) are in red (**Methods**). 17% (23/143) of points lie in the highlighted region. **(d)** Phylogenetic tree of *B. dorei*, one of the strongest engraftment events from (c), shows a successful engraftment from donor to patient. Color indicates subject and numbers indicate sample collection time in weeks relative to FMT. **(e)** Phylogenetic tree of *E. halii*, showing long-term coexistence of donor and original patient strains. **(f)** Community-level potential for the top three MetaCyc pathways associated with FMT (all associations in **Table S2**; the model included age, sex, relapse, FMT (**Methods**)).

Figure 3. Relapse-associated metagenomic features. (a) Shannon index over time for all patients who relapsed, synchronized to the first relapse time point, shows a decrease in diversity at the onset of disease activity. **(b)** Community-level potential for the top three MetaCyc pathways associated with relapse (all associations in **Table S2**; the model included age, sex, relapse, FMT (**Methods**)).

Figure 4. Joint analysis of strain and abundance levels reveals relapse-associated changes in FMT. (a) Engraftment events in FMT (left) and species abundances in sham (right), ordered by the number of events in patients which eventually relapsed. Phyla are non-randomly distributed in this ranking, with Firmicutes providing apparent benefits, while Bacteroidetes shows no benefit (or harmful). **(b)** Number of engraftment events per phylum in each category from (a). Despite the study's general underrepresentation of Bacteroidetes, these accounted for almost half of engraftment events. **(c)** Abundance of *Barnesiella intestinihominis* over time in all FMT patients, with abundance in the matched donor shown in red. **(d)** PERMANOVA detects strain-level associations with FMT and relapse (20 species with $p < 0.05$ shown; all results in **Table S3**).

References

1. Kaplan GG, Ng SC. Understanding and Preventing the Global Increase of Inflammatory Bowel Disease. Gastroenterology. 2017 Feb;152(2):313–21.e2.

2. Khor B, Gardet A, Xavier RJ. Genetics and pathogenesis of inflammatory bowel disease. Nature. 2011 Jun 15;474(7351):307–17.

3. Lloyd-Price J, Arze C, Ananthakrishnan AN, Schirmer M, Avila-Pacheco J, Poon TW, et al. Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases. Nature. 2019 May;569(7758):655–62.

4. Kirchgesner J, Lemaitre M, Carrat F, Zureik M, Carbonnel F, Dray-Spira R. Risk of Serious and Opportunistic Infections Associated With Treatment of Inflammatory Bowel Diseases. Gastroenterology. 2018 Aug;155(2):337–46.e10.

5. Beaugerie L, Itzkowitz SH. Cancers complicating inflammatory bowel disease. N Engl J Med. 2015 Apr 9;372(15):1441–52.

6. Ananthakrishnan AN, Luo C, Yajnik V, Khalili H, Garber JJ, Stevens BW, et al. Gut Microbiome Function Predicts Response to Anti-integrin Biologic Therapy in Inflammatory Bowel Diseases. Cell Host Microbe. 2017 May 10;21(5):603–10.e3.

7. Khanna S, Vazquez-Baeza Y, González A, Weiss S, Schmidt B, Muñiz-Pedrogo DA, et al. Changes in microbial ecology after fecal microbiota transplantation for recurrent C. difficile infection affected by underlying inflammatory bowel disease. Microbiome. 2017 May 15;5(1):55.

8. Hirten RP, Grinspan A, Fu S-C, Luo Y, Suarez-Farinas M, Rowland J, et al. Microbial Engraftment and Efficacy of Fecal Microbiota Transplant for Clostridium Difficile in Patients With and Without Inflammatory Bowel Disease. Inflamm Bowel Dis. 2019 May 4;25(6):969–79.

9. Plichta DR, Graham DB, Subramanian S, Xavier RJ. Therapeutic Opportunities in Inflammatory Bowel Disease: Mechanistic Dissection of Host-Microbiome Relationships. Cell. 2019 Aug 22;178(5):1041–56.

10. Buffie CG, Bucci V, Stein RR, McKenney PT, Ling L, Gobourne A, et al. Precision microbiome reconstitution restores bile acid mediated resistance to Clostridium difficile. Nature. 2015 Jan 8;517(7533):205–8.

11. Pamer EG. Fecal microbiota transplantation: effectiveness, complexities, and lingering concerns. Mucosal Immunol. 2014 Mar;7(2):210–4.

12. Staley C, Kaiser T, Vaughn BP, Graiziger C, Hamilton MJ, Kabage AJ, et al. Durable Long-Term Bacterial Engraftment following Encapsulated Fecal Microbiota Transplantation To Treat Clostridium difficile Infection. MBio [Internet]. 2019 Jul 23;10(4). Available from: http://dx.doi.org/10.1128/mBio.01586-19

13. Fischer M, Kao D, Kelly C, Kuchipudi A, Jafri S-M, Blumenkehl M, et al. Fecal Microbiota Transplantation is Safe and Efficacious for Recurrent or Refractory Clostridium difficile Infection in Patients with Inflammatory Bowel Disease. Inflamm Bowel Dis. 2016 Oct 1;22(10):2402–9.

14. Kim SG, Becattini S, Moody TU, Shliaha PV, Littmann ER, Seok R, et al. Microbiota-derived lantibiotic restores resistance against vancomycin-resistant Enterococcus. Nature. 2019 Aug;572(7771):665–9.

15. Moayyedi P, Surette MG, Kim PT, Libertucci J, Wolfe M, Onischi C, et al. Fecal Microbiota Transplantation Induces Remission in Patients With Active Ulcerative Colitis in a Randomized Controlled Trial. Gastroenterology. 2015 Jul;149(1):102–9.e6.

16. Paramsothy S, Kamm MA, Kaakoush NO, Walsh AJ, van den Bogaerde J, Samuel D, et al. Multidonor intensive faecal microbiota transplantation for active ulcerative colitis: a randomised placebo-controlled trial. Lancet. 2017 Mar 25;389(10075):1218–28.

17. Costello SP, Hughes PA, Waters O, Bryant RV, Vincent AD, Blatchford P, et al. Effect of Fecal

Microbiota Transplantation on 8-Week Remission in Patients With Ulcerative Colitis: A Randomized Clinical Trial. JAMA. 2019 Jan 15;321(2):156–64.

18. Rossen NG, Fuentes S, van der Spek MJ, Tijssen JG, Hartman JHA, Duflou A, et al. Findings From a Randomized Controlled Trial of Fecal Transplantation for Patients With Ulcerative Colitis. Gastroenterology. 2015 Jul;149(1):110–8.e4.

19. Sokol H, Landman C, Seksik P, Berard L, Montil M, Nion-Larmurier I, et al. Fecal microbiota transplantation to maintain remission in Crohn's disease: a pilot randomized controlled study. Microbiome. 2020 Feb 3;8(1):12.

20. Sokol H, Pigneur B, Watterlot L, Lakhdari O, Bermúdez-Humarán LG, Gratadoux J-J, et al. Faecalibacterium prausnitzii is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. Proc Natl Acad Sci U S A. 2008 Oct 28;105(43):16731–6.

21. Lloyd-Price J, Mahurkar A, Rahnavard G, Crabtree J, Orvis J, Hall AB, et al. Strains, functions and dynamics in the expanded Human Microbiome Project. Nature. 2017 Oct 5;550(7674):61–6.

22. Ramezani A, Massy ZA, Meijers B, Evenepoel P, Vanholder R, Raj DS. Role of the Gut Microbiome in Uremia: A Potential Therapeutic Target. Am J Kidney Dis. 2016 Mar;67(3):483–98.

23. Macfarlane GT, Blackett KL, Nakayama T, Steed H, Macfarlane S. The gut microbiota in inflammatory bowel disease. Curr Pharm Des. 2009;15(13):1528–36.

24. Hall AB, Yassour M, Sauk J, Garner A, Jiang X, Arthur T, et al. A novel Ruminococcus gnavus clade enriched in inflammatory bowel disease patients. Genome Med. 2017 Nov 28;9(1):103.

25. Nagao-Kitamoto H, Kamada N. Host-microbial Cross-talk in Inflammatory Bowel Disease. Immune Netw. 2017 Feb;17(1):1–12.

26. Maldonado-Gómez MX, Martínez I, Bottacini F, O'Callaghan A, Ventura M, van Sinderen D, et al. Stable Engraftment of Bifidobacterium longum AH1206 in the Human Gut Depends on Individualized Features of the Resident Microbiome. Cell Host Microbe. 2016 Oct 12;20(4):515–26.

27. Paramsothy S, Paramsothy R, Rubin DT, Kamm MA, Kaakoush NO, Mitchell HM, et al. Faecal Microbiota Transplantation for Inflammatory Bowel Disease: A Systematic Review and Meta-analysis [Internet]. Vol. 11, Journal of Crohn's and Colitis. 2017. p. 1180–99. Available from: http://dx.doi.org/10.1093/ecco-jcc/jjx063

28. Kearney SM, Gibbons SM, Erdman SE, Alm EJ. Orthogonal Dietary Niche Enables Reversible Engraftment of a Gut Bacterial Commensal. Cell Rep. 2018 Aug 14;24(7):1842–51.

29. Shepherd ES, DeLoache WC, Pruss KM, Whitaker WR, Sonnenburg JL. An exclusive metabolic niche enables strain engraftment in the gut microbiota. Nature. 2018 May;557(7705):434–8.

30. Wilson BC, Vatanen T, Cutfield WS, O'Sullivan JM. The Super-Donor Phenomenon in Fecal Microbiota Transplantation. Front Cell Infect Microbiol. 2019 Jan 21;9:2.

31. Segata N, Waldron L, Ballarini A, Narasimhan V, Jousson O, Huttenhower C. Metagenomic microbial community profiling using unique clade-specific marker genes. Nat Methods. 2012 Jun 10;9(8):811–4.

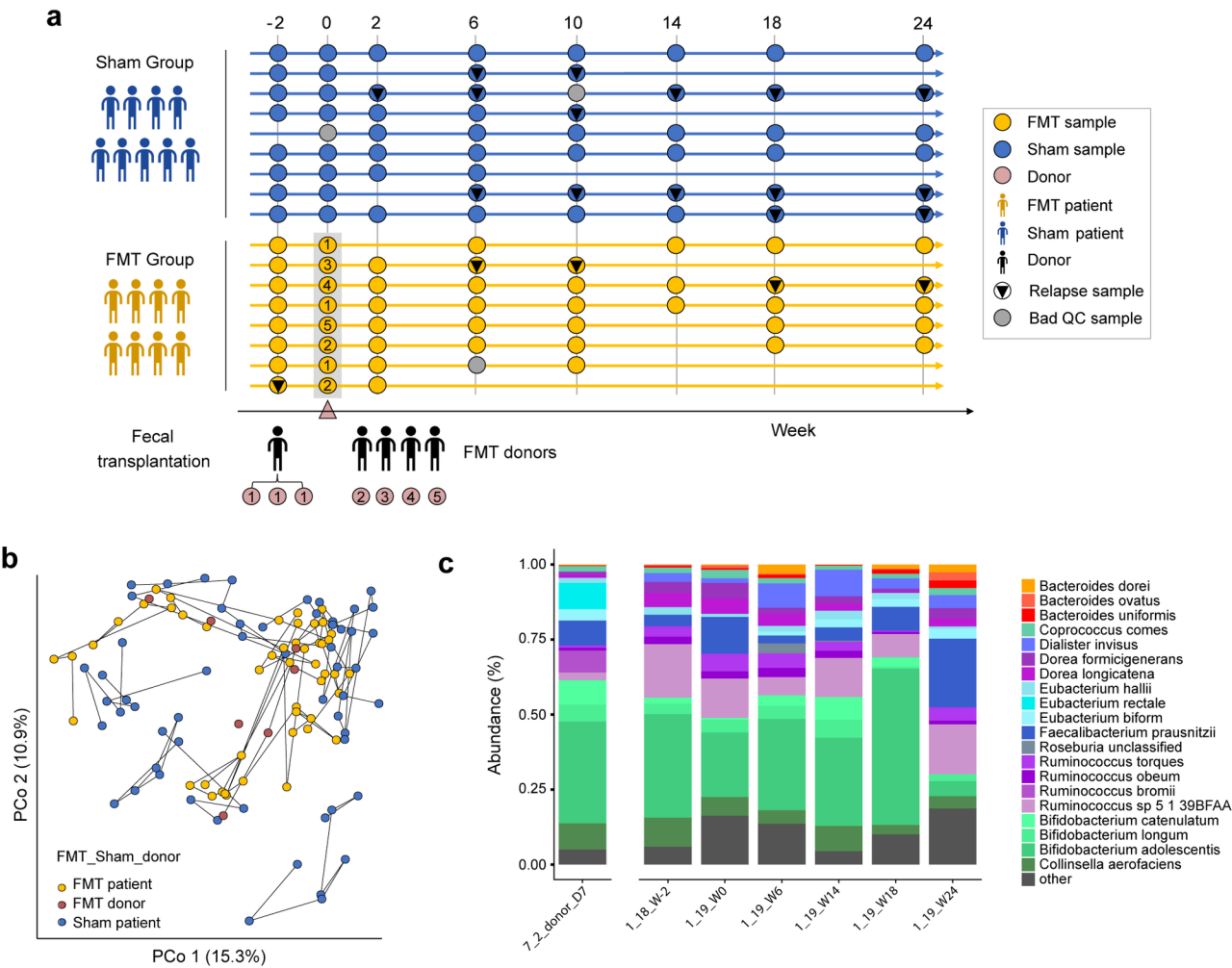
32. Franzosa EA, McIver LJ, Rahnavard G, Thompson LR, Schirmer M, Weingart G, et al. Species-level functional profiling of metagenomes and metatranscriptomes. Nat Methods. 2018 Nov;15(11):962–8.

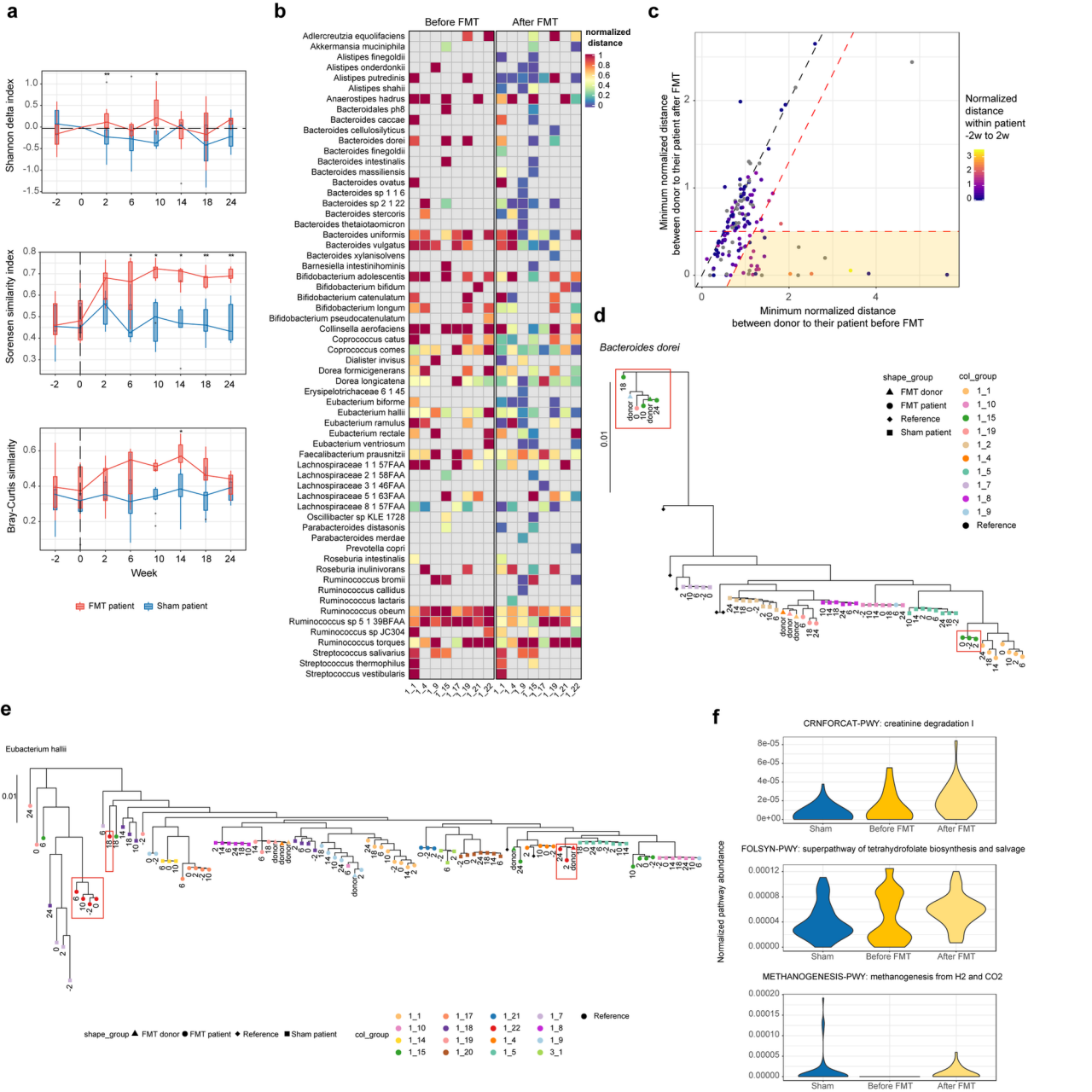
33. Caspi R, Billington R, Fulcher CA, Keseler IM, Kothari A, Krummenacker M, et al. The MetaCyc database of metabolic pathways and enzymes. Nucleic Acids Res. 2018 Jan 4;46(D1):D633–9.

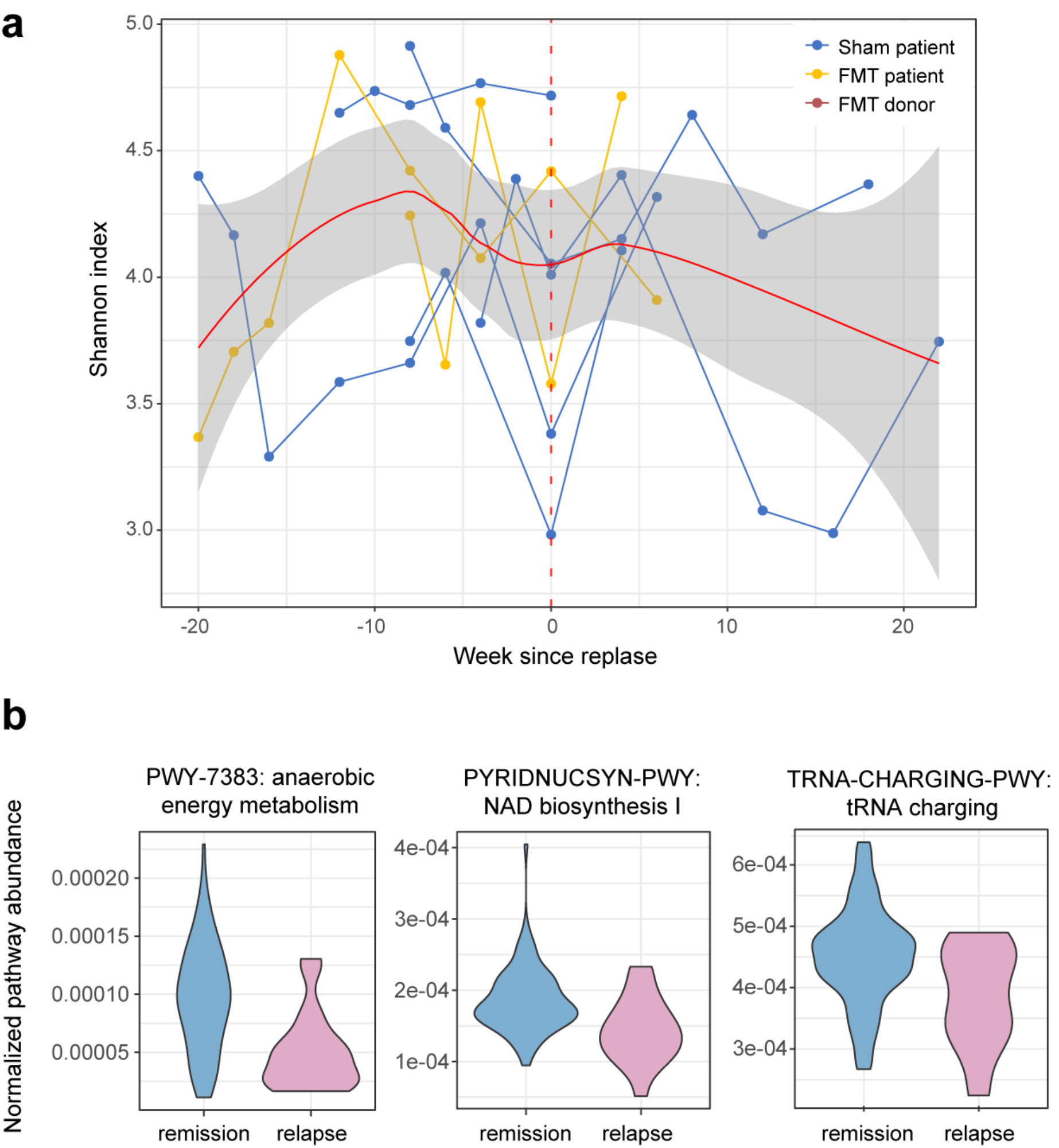
34. Truong DT, Tett A, Pasolli E, Huttenhower C, Segata N. Microbial strain-level population structure and genetic diversity from metagenomes. Genome Res. 2017 Apr;27(4):626–38.

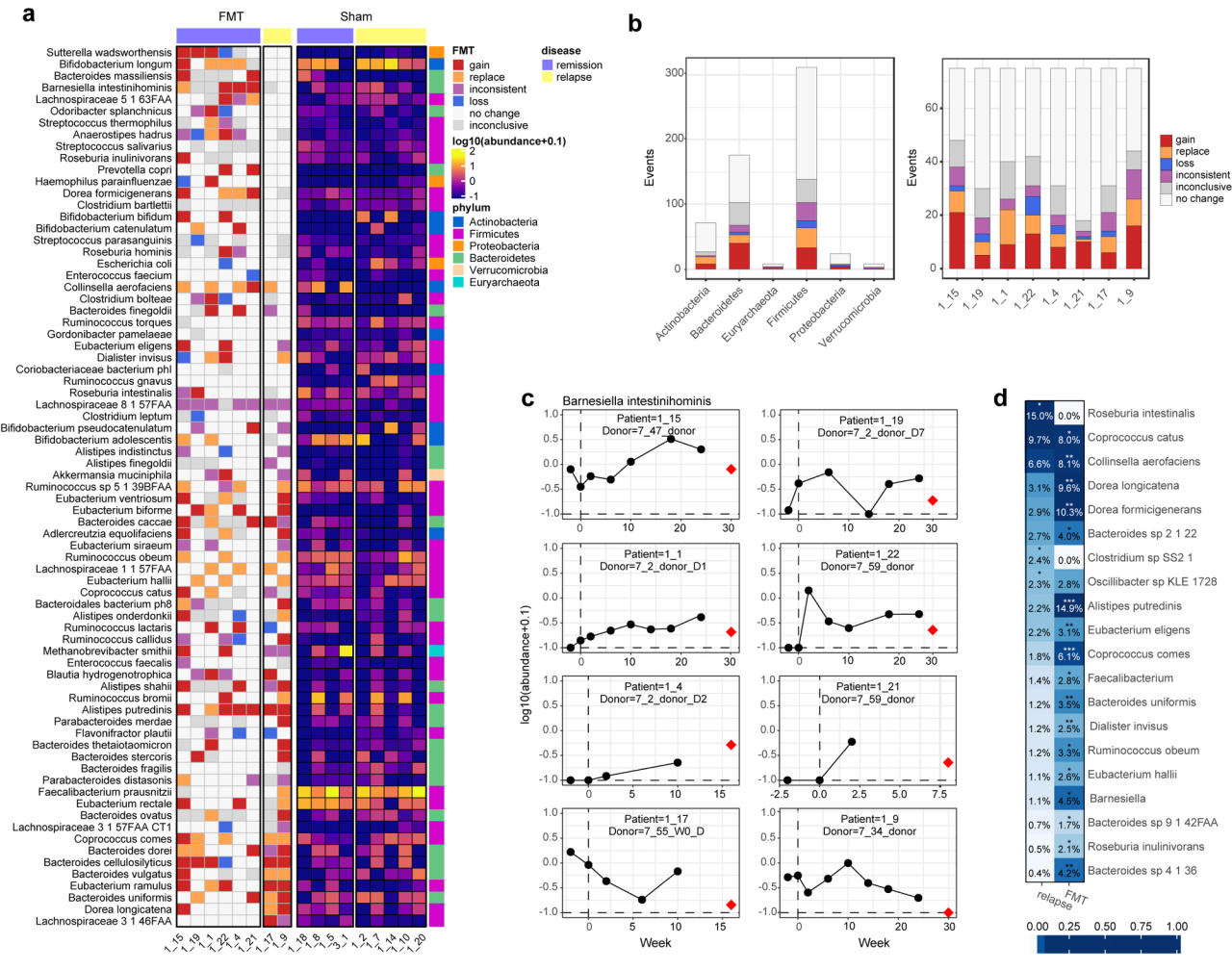
35. Schliep KP. phangorn: phylogenetic analysis in R. Bioinformatics. 2011 Feb 15;27(4):592–3.

36. Morgan XC, Tickle TL, Sokol H, Gevers D, Devaney KL, Ward DV, et al. Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment. Genome Biol. 2012 Apr 16;13(9):R79.


a
-2
0
2
6
10
14
18
24
Sham Group
FMT Group
FMT sample
Sham sample
Donor
FMT patient
Sham patient
Donor
Relapse sample
Bad QC sample
Week
Fecal transplantation
FMT donors
b
PCo 2 (10.9%)
PCo 1 (15.3%)
FMT_Sham_donor
FMT patient
FMT donor
Sham patient
c
Abundance (%)
1.00
0.75
0.50
0.25
0.00
7_2_donor_D7
1_18_W-2
1_19_W0
1_19_W6
1_19_W14
1_19_W18
1_19_W24
Bacteroides dorei
Bacteroides ovatus
Bacteroides uniformis
Coprococcus comes
Dialister invisus
Dorea formicigenerans
Dorea longicatena
Eubacterium hallii
Eubacterium rectale
Eubacterium biform
Faecalibacterium prausnitzii
Roseburia unclassified
Ruminococcus torques
Ruminococcus obeum
Ruminococcus bromii
Ruminococcus sp 5 1 39BFAA
Bifidobacterium catenulatum
Bifidobacterium longum
Bifidobacterium adolescentis
Collinsella aerofaciens
other


a
Shannon delta index
Sorensen similarity index
Bray-Curtis similarity
Week
FMT patient
Sham patient
b
Before FMT
After FMT
normalized distance
Adlercreutzia equolifaciens
Akkermansia muciniphila
Alistipes finegoldii
Alistipes onderdonkii
Alistipes putredinis
Alistipes shahii
Anaerostipes hadrus
Bacteroidales ph8
Bacteroides caccae
Bacteroides cellulosilyticus
Bacteroides dorei
Bacteroides finegoldii
Bacteroides intestinalis
Bacteroides massiliensis
Bacteroides ovatus
Bacteroides sp 1 1 6
Bacteroides sp 2 1 22
Bacteroides stercoris
Bacteroides thetaiotaomicron
Bacteroides uniformis
Bacteroides vulgatus
Bacteroides xylanisolvens
Barnesiella intestinihominis
Bifidobacterium adolescentis
Bifidobacterium bifidum
Bifidobacterium catenulatum
Bifidobacterium longum
Bifidobacterium pseudocatenulatum
Collinsella aerofaciens
Coprococcus catus
Coprococcus comes
Dialister invisus
Dorea formicigenerans
Dorea longicatena
Erysipelotrichaceae 6 1 45
Eubacterium biforme
Eubacterium hallii
Eubacterium ramulus
Eubacterium rectale
Eubacterium ventriosum
Faecalibacterium prausnitzii
Lachnospiraceae 1 1 57FAA
Lachnospiraceae 2 1 58FAA
Lachnospiraceae 3 1 46FAA
Lachnospiraceae 5 1 63FAA
Lachnospiraceae 8 1 57FAA
Oscillibacter sp KLE 1728
Parabacteroides distasonis
Parabacteroides merdae
Prevotella copri
Roseburia intestinalis
Roseburia inulinivorans
Ruminococcus bromii
Ruminococcus callidus
Ruminococcus lactaris
Ruminococcus obeum
Ruminococcus sp 5 1 39BFAA
Ruminococcus sp JC304
Ruminococcus torques
Streptococcus salivarius
Streptococcus thermophilus
Streptococcus vestibularis
c
Minimum normalized distance between donor to their patient after FMT
Minimum normalized distance between donor to their patient before FMT
Normalized distance within patient -2w to 2w
d
Bacteroides dorei
shape_group
FMT donor
FMT patient
Reference
Sham patient
col_group
e
Eubacterium hallii
shape_group FMT donor FMT patient Reference Sham patient
col_group
Reference
f
CRNFORCAT-PWY: creatinine degradation I
FOLSYN-PWY: superpathway of tetrahydrofolate biosynthesis and salvage
METHANOGENESIS-PWY: methanogenesis from H2 and CO2
Normalized pathway abundance
Sham
Before FMT
After FMT


a
Sham patient
FMT patient
FMT donor
Shannon index
5.0
4.5
4.0
3.5
3.0
-20
-10
0
10
20
Week since replase
b
PWY-7383: anaerobic energy metabolism
PYRIDNUCSYN-PWY: NAD biosynthesis I
TRNA-CHARGING-PWY: tRNA charging
Normalized pathway abundance
0.00020
0.00015
0.00010
0.00005
4e-04
3e-04
2e-04
1e-04
6e-04
5e-04
4e-04
3e-04
remission
relapse
remission
relapse
remission
relapse


a
FMT
Sham
Sutterella wadsworthensis
Bifidobacterium longum
Bacteroides massiliensis
Barnesiella intestinihominis
Lachnospiraceae 5 1 63FAA
Odoribacter splanchnicus
Streptococcus thermophilus
Anaerostipes hadrus
Streptococcus salivarius
Roseburia inulinivorans
Prevotella copri
Haemophilus parainfluenzae
Dorea formicigenerans
Clostridium bartlettii
Bifidobacterium bifidum
Bifidobacterium catenulatum
Streptococcus parasanguinis
Roseburia hominis
Escherichia coli
Enterococcus faecium
Collinsella aerofaciens
Clostridium bolteae
Bacteroides finegoldii
Ruminococcus torques
Gordonibacter pamelaeae
Eubacterium eligens
Dialister invisus
Coriobacteriaceae bacterium phl
Ruminococcus gnavus
Roseburia intestinalis
Lachnospiraceae 8 1 57FAA
Clostridium leptum
Bifidobacterium pseudocatenulatum
Bifidobacterium adolescentis
Alistipes indistinctus
Alistipes finegoldii
Akkermansia muciniphila
Ruminococcus sp 5 1 39BFAA
Eubacterium ventriosum
Eubacterium biforme
Bacteroides caccae
Adlercreutzia equolifaciens
Eubacterium siraeum
Ruminococcus obeum
Lachnospiraceae 1 1 57FAA
Eubacterium hallii
Coprococcus catus
Bacteroidales bacterium ph8
Alistipes onderdonkii
Ruminococcus lactaris
Ruminococcus callidus
Methanobrevibacter smithii
Enterococcus faecalis
Blautia hydrogenotrophica
Alistipes shahii
Ruminococcus bromii
Alistipes putredinis
Parabacteroides merdae
Flavonifractor plautii
Bacteroides thetaiotaomicron
Bacteroides stercoris
Bacteroides fragilis
Parabacteroides distasonis
Faecalibacterium prausnitzii
Eubacterium rectale
Bacteroides ovatus
Lachnospiraceae 3 1 57FAA CT1
Coprococcus comes
Bacteroides dorei
Bacteroides cellulosilyticus
Bacteroides vulgatus
Eubacterium ramulus
Bacteroides uniformis
Dorea longicatena
Lachnospiraceae 3 1 46FAA
1_15 1_19 1_1 1_22 1_4 1_21 1_17 1_9 1_18 1_8 1_5 3_1 1_2 1_7 1_14 1_10 1_20
FMT
gain
replace
inconsistent
loss
no change
inconclusive
log10(abundance+0.1)
2
1
0
-1
phylum
Actinobacteria
Firmicutes
Proteobacteria
Bacteroidetes
Verrucomicrobia
Euryarchaeota
disease
remission
relapse
b
Events
300
200
100
0
60
40
20
0
Actinobacteria
Bacteroidetes
Euryarchaeota
Firmicutes
Proteobacteria
Verrucomicrobia
1_15 1_19 1_1 1_22 1_4 1_21 1_17 1_9
gain
replace
loss
inconsistent
inconclusive
no change
c
Barnesiella intestinihominis
Patient=1_15
Donor=7_47_donor
Patient=1_19
Donor=7_2_donor_D7
Patient=1_1
Donor=7_2_donor_D1
Patient=1_22
Donor=7_59_donor
Patient=1_4
Donor=7_2_donor_D2
Patient=1_21
Donor=7_59_donor
Patient=1_17
Donor=7_55_W0_D
Patient=1_9
Donor=7_34_donor
log10(abundance+0.1)
Week
Week
d
15.0% 0.0% Roseburia intestinalis
9.7% 8.0% Coprococcus catus
6.6% 8.1% Collinsella aerofaciens
3.1% 9.6% Dorea longicatena
2.9% 10.3% Dorea formicigenerans
2.7% 4.0% Bacteroides sp 2 1 22
2.4% 0.0% Clostridium sp SS2 1
2.3% 2.8% Oscillibacter sp KLE 1728
2.2% 14.9% Alistipes putredinis
2.2% 3.1% Eubacterium eligens
1.8% 6.1% Coprococcus comes
1.4% 2.8% Faecalibacterium
1.2% 3.5% Bacteroides uniformis
1.2% 2.5% Dialister invisus
1.2% 3.3% Ruminococcus obeum
1.1% 2.6% Eubacterium hallii
1.1% 4.5% Barnesiella
0.7% 1.7% Bacteroides sp 9 1 42FAA
0.5% 2.1% Roseburia inulinivorans
0.4% 4.2% Bacteroides sp 4 1 36
relapse
FMT
0.00 0.25 0.50 0.75 1.00