Title

Exploration of the Role Donor and Recipient Microbiome Community Structure Plays in the Efficacy of Fecal Microbiota Transplant to Treat Ulcerative Colitis.

Author

CHAVIRA, ARIES

Publication Date

2021

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA SAN DIEGO

Exploration of the Roles that Donor and Recipient Microbiome Community Structure Play in the Efficacy of Fecal Microbiota Transplant to Treat Ulcerative Colitis.

A Thesis submitted in partial satisfaction of the requirements
for the degree Master of Science

in

Biology

by

Aries Chavira

Committee in charge:

Professor Rob Knight, Chair
Professor Randolph Hampton, Co-Chair
Professor Barry Grant

2021

The Thesis of Aries Chavira is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2021

DEDICATION

I would like to dedicate my work to my mother. Her grit, compassion, and perseverance in dealing with ulcerative colitis have formed the foundation for my passion, and research focus. I hope to contribute to the literature with optimism of one day proving a better quality of life to those battling UC.

I would like to thank my PI, for providing me this incredible opportunity to conduct this research. For challenging, encouraging, and pushing me to be a better scientist.

I would like to thank my partner for remining me how capable I am, for always believing in me, and for showing me how much farther I can go. This year wouldn't nearly have been what it was without you there by my side to get through it all. I thank you endlessly.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to acknowledge Professor Rob Knight for his support as the chair of my committee. His constant efforts in challenging me have brought this work to the level it is at today.

I would like to acknowledge Yoshiki Vasquez, Justin Shaffer, Daniel McDonald, and Mehrbod Estaki, for believing in me and supporting me. They were each invaluable for formulating my ideas, challenging my understanding, and for constantly instilling their knowledge. It is because of their support that I have learned so much about microbial ecology, and bioinformatics. This work would have taken way longer if it weren't for each of them, and I am extremely grateful for all their contributions.

Material from this thesis is coauthored by Mehrbod Estaki. The thesis author was the primary author of this material.

Material from this thesis is coauthored by Justin Shaffer. The thesis author was the primary author of this material.

VITA

2019	Bachelor of Science, University of California San Diego
2021-2021	Teaching Assistant, University of California San Diego
2021	Master of Science, University of California San Diego
2016-2021	Research Assistant, University of California San Diego

FIELDS OF STUDY

Major Field: Biology

Studies in Bioinformatics
Professors Barry Grant

ABSTRACT OF THE THESIS

Exploration of the Role Donor and Recipient Microbiome Community Structure Plays in the Efficacy of Fecal Microbiota Transplant to Treat Ulcerative Colitis.

by

Aries Chavira

Master of Science in Biology

University of California San Diego, 2021

Professor Rob Knight, Chair
Professor Randolph Hampton, Co-Chair

Ulcerative colitis (UC) is a chronic autoimmune disorder of the gut affecting over one million Americans. The precise etiology is unknown; however, it commonly exhibits a loss of beneficial gut microbes that produce butyrate and short-chain fatty acids. Fecal Microbiota Transplant (FMT) is a promising therapeutic thought to restore UC gut dysbiosis. However, many questions remain regarding why FMT's work for some individuals but not for others. To decrease this gap, we conducted a quantitative meta-analysis to explore the role patient, and donor microbiomes play in the efficacy of FMT to treat UC. Twenty-eight studies met the

preliminary exclusion criteria, five of which were included in this analysis. We found that patients who experience a clinical response have a higher species richness prior to FMT than those who do not. All patients experience an increase in alpha diversity after FMT and a shift in beta diversity to their donors, irrespective of their clinical outcome. We generated a microbial log-ratio model capable of predicting a patient's response to FMT with an area under the receiver operating characteristic (AUROC) of 0.73. Differential abundance testing confirmed the increase in known butyrate producers of the genus, Blautia, Ruminococcus, and Alistipes, in donors and responder, compared to non-responders that were also predictors of clinical response in the sparse log-ratio model. These results signify that non-responders may harbor microbial communities too sparse in beneficial microbes that may not be fully restored post-FMT treatment, compared to those who experience a clinical response.

CHAPTER I

Exploration of the Role Donor and Recipient Microbiome Community Structure Plays in the Efficacy of Fecal Microbiota Transplant to Treat Ulcerative Colitis.

Introduction

Ulcerative colitis (UC) is a complex chronic inflammatory disease. It falls under the larger umbrella of Irritable Bowel Disease (IBD), which also includes Crohn's disease. To date, the precise etiology of UC is unknown, though it is thought to be multi-factorial. Studies have highlighted the influence of environmental factors, genetic predisposition, degradation of the epithelial layer along the gut, mucosal inflammation, and concurrent dysbiosis of the enteric gut microbiome (1-3). The severity of the disease is often associated with increased mucosal inflammation. Biological agents are currently the primary treatment mechanism, commonly accompanied by nonspecific immunosuppressive agents such as Prednisone, to minimize the inflammatory and immune response and disease activity (4). However, these agents impose serious risk factors for patients and are not effective for a majority of patients (5).

Advancements in sequencing technology, efficiency, and multi "omics" techniques have allowed us to quantify the constituents of the human microbiome and its functions. Humans generally harbor groups of core residential microbes that serve various metabolic processes and possess genes that are in constant communication with the host (6). The diversity of the human gut and its' respective complex functionality contribute to mucosal health and pathogenic resiliency (6). For patients with UC, dysbiosis in the enteric microbiome and subsequent loss of function is a well-established hallmark of the disease (7). However, it is not known with certainty if the shift in microbial community structure is causal or a subsequent result of the disease's pathogenesis. In either case, researchers have shown a pattern in the depletion of

beneficial gut microbes (i.e., Akkermansia muciniphila) and increases in pathogenic species (i.e., Roseburia, Suterella, Faecalibacterium, etc.) (7).

Utilizing human stool as a therapeutic can be dated back to the late 1950s where a group of surgeons provided fecal enemas to patients with pseudomembranous enterocolitis, now known as Clostridium difficile infection (CDI) (8). CDI infection is the intestinal overgrowth of Clostridium difficile, whose onset is attributed to the repeated use of antibiotics (9). It was hypothesized that a fecal enema could restore the enteric microbiome. Across many placebo-controlled randomized clinical trials published to date, Fecal Microbiota Transplant (FMT) has demonstrated remarkable remission rates in treating recurrent CDI. Its pharmacological promise has now paved the way for its implementation as a standard treatment method for CDI. FMT's potential to treat many other conditions where shifts in the human gut microbiome are critical factors of disease pathogenesis are now heavily explored.

Our understanding of the enteric microbiome disturbances in UC has made FMT a promising candidate for treatment. Regardless of its exploration dated in the late 1950s, it is only recently that many studies assessing the efficacy of FMT to treat UC have been published. Unfortunately, the studies published to date utilizing FMT to treat UC report much lower and heterogenous efficacy rates than those seen for CDI. One plausible hypothesis for this variation is the complexity of host-microbial interactions that are at play within the human gut. Not only is there heterogeneity in the microbial composition of healthy individuals, but heterogeneity in the microbial composition of patients with ulcerative colitis also exists. In the case of CDI, the gut becomes populated with one bacterium, thus any healthy donor can be capable of restoring lost function. The dysbiosis in UC is much more complex, and thus, the efficacy of one donor may be more variable.

In recent years a handful of meta-analyses have been published to evaluate the efficacy of FMT or assess its mechanism of action. Many of these studies aim their focus on the propensity of the donor microbiome to achieve remission. In some cases, there are donors capable of inducing remission for a large proportion of patients for FMT. In one randomized controlled trial (RCT), Moayyedi (2015), 78% of the patients who received donor B achieved clinical remission (10). It is plausible that without Donor B the results from their trial may have yielded a different result. However, the efficacy rates of donors for other studies that reported similar efficacy rates do not follow a similar trend of one donor contributing to a majority of those achieving clinical remission that was seen in Moayyedi (2015). In the RCT Paramsothy (2017), the efficacy rate for the donor batch which induced remission for the largest number of patients was still only 50% (11). Additionally, the rates of those who reached the primary endpoint for Paramsothy (2017) and Moayyedi (2015) were 27% and 24%, respectfully. So, while the composition of the donor's microbiome is undoubtedly an important factor, it does not fully explain the mixed rates of clinical remission from donors or donor batches observed in many other FMT studies.

An additional factor that may play a role in whether a suitable donor may induce remission is the microbial community structure of the recipients themselves. Past research has paid much attention on analyzing the enteric microbiome of donors. Still, hardly any literature exists assessing the heterogeneity in the microbial composition of patients with ulcerative colitis prior to them receiving an FMT and how that influences their respective clinical outcomes. With this in mind, we set out to conduct a quantitative meta-analysis with a holistic approach. We analyze how patient's microbiomes shift in response to FMT and look at the differences in patients' microbial communities before FMT and in relationship to their donor to understand how the donor and patient microbiome's play a role in determining whether a patient achieves

remission. Furthermore, we employ the use of novel sparse log-ratio predictive models to assess and predict the difference in the microbial composition of patients who experience clinical response to those who do not before and after FMT.

METHODS

Literature Search Strategy

With a constrained timeline, we searched the Pubmed database only, from inception to April 1st, 2020. We used the following search terms as keywords or free-text words only; FMT, or {(fecal, or faecal, or stool) and (transplant, or microbiota transplant, or transfusion, or implant, or enema, or infusion) and [bacteriotherapy, or (UC or ulcerative colitis)]. We used no other advanced search features or language limits. Additionally, the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) was performed and included (Fig. 1) (13). Of the studies identified from the search criteria, 218 were assessed for eligibility. The title and or abstract of the 218 were read and evaluated for relevance. Of the 218 identified, only 21 included titles and or abstracts indicating the assessment of FMT for UC. All 21 studies were read in full and assessed from the criteria described in the section below for inclusion in the final analysis.

Inclusion Criteria

Articles were included in this meta-analysis if they assessed the efficacy of FMT to treat ulcerative colitis. Cohort and RCTs were all included. Microbiome samples had to have been taken of the donors, and the recipients before and after FMT. To mitigate the technical effects imposed by heterogeneous sequencing platforms, the sequencing platform of the two included RCTs were used as inclusion criteria to maximize the total number of samples. Therefore, all

studies utilizing the Illumina sequencing platform were included. Additionally, minimum metadata requirements were imposed to utilize each study in the downstream analyses. These requirements included; Sample ID, de-identified patient identifiers, the clinical outcome of the patient, which donor samples were given to each patient, and the time point of each sample.

Exclusion Criteria:

Studies were excluded if microbiome samples were not taken or sequenced and if the minimum metadata requirements were not or could not be provided (Table 1, Fig. 1). One study was excluded for classifying patients with Indeterminate Colitis as UC due to low sample size. Additionally, studies were excluded if the only patients reported had a co-infection with CDI. A table is provided which lists each of the 16 studies that were excluded from this analysis and the reason for exclusion (Table 1, Fig. 1).

Data acquisition:

The representative raw sequences for each study that met the requirements described above were searched for in the European Nucleotide Archive and the Sequence Read Archive (SRA). Three studies did not have publicly available sequencing data. We contacted each of those authors directly to retrieve the raw FASTQ files. We also pulled metadata from the European Nucleotide Archive. Each study in this analysis did not have publicly available metadata that met the minimum metadata requirements described above. In the event additional metadata was required, the authors were contacted directly. Three studies were unable to either; 1) provide the additional metadata categories in a de-identified manner, 2) did not link which

donor went to which patient, or 3) did not respond to request (Table 1). The five studies included in this analysis are listed in Table 2 (Table 2).

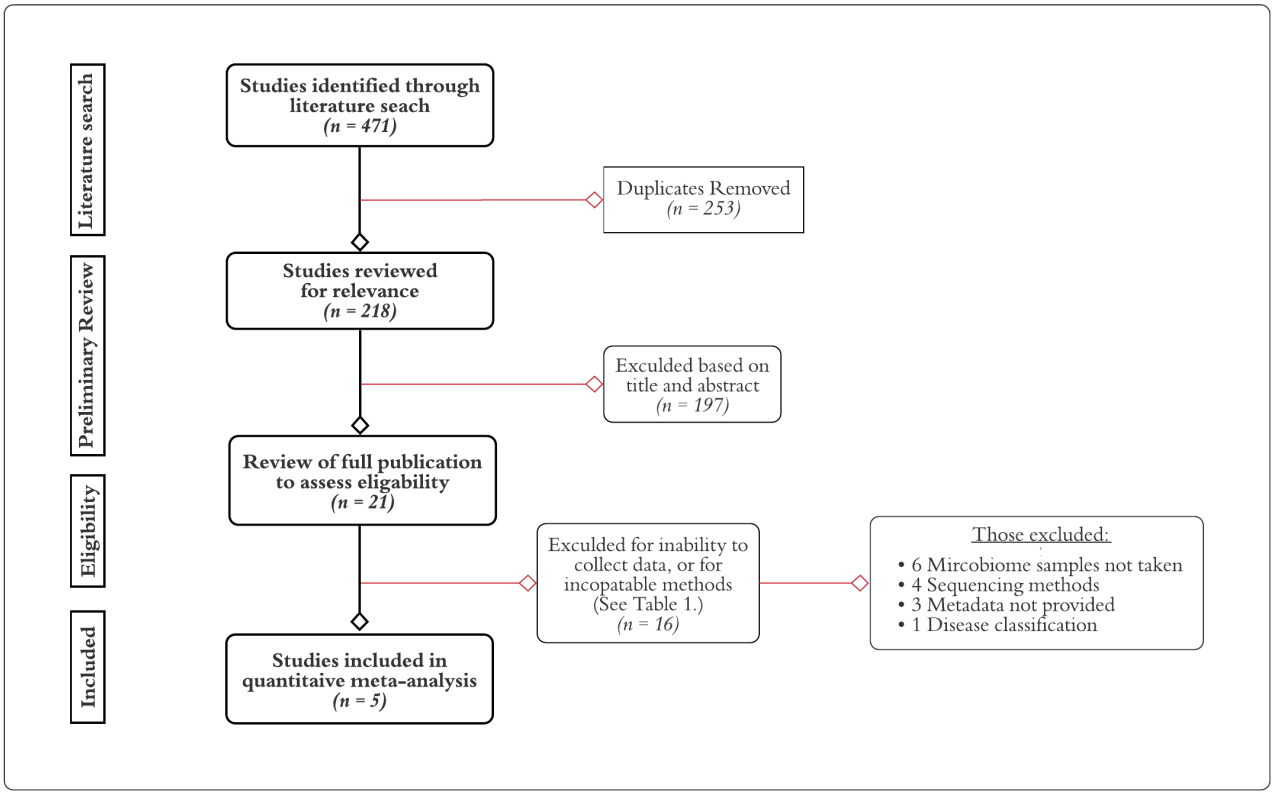


Figure 1. PRISMA flowchart for reporting search strategy. From the literature search, the title and or abstract of 218 studies were assessed for review. A total of 21 articles were read in full and assessed for eligibility. Of those reviews, 5 studies were included in this analysis.

Table 1. Excluded publications, and justification for exclusion. The table below identifies the sixteen studies excluded from this meta-analysis, along with why they were excluded.

Study *(n = 16)*	Reason for exclusion
Angelberger *et al*., 2013 (14)	16S sequencing was performed on the Roche 454 FLX instrument, likely to induce technical effects when included with the Illumina sequencing pipeline.
Kunde *et al*., 2013 (15)	Microbiome samples not taken
Costello *et al*., 2017 (16)	Unable to link which donors were given to which patients
Cui *et al*., 2015 (17)	Authors did not respond to request for metadata
Damman *et al*., 2015 (18)	Authors were unable to provide deidentified metadata
Karolewska-Bochenek *et al*., 2015 (19)	Microbiome samples not taken
Kellermayer *et al*., 2015 (20)	16s sequencing was performed via parallel pyrosequencing likely to induce technical effects when included with the Illumina sequencing pipeline.
Suskind *et al*., 2015 (21)	Microbiome samples not taken
Vermeire *et al*., 2016 (22)	16s sequencing was performed via pyrosequencing likely to induce technical effects when included with the Illumina sequencing pipeline.
Wei *et al*., 2015 (23)	Microbiome samples not taken
Ren *et al*., 2015 (24)	Microbiome samples not taken
Rossen *et al*., 2015 (25)	Microbiota profiling was performed using a phylogenetic microarray, Human Intestinal Tract Chip (HITChip), likely to induce technical effects when included with the Illumina sequencing pipeline
Goyal *et al*., 2016 (26)	Patients with Indeterminate Colitis were classified as UC patients
Laszlo *et al*., 2016 (27))	Microbiome samples not taken
Nishida *et al*., 2016 (28)	16s sequencing was performed via terminal restriction fragment length polymorphism (T-RFLP), likely to induce technical effects when included with the Illumina sequencing pipeline.
Zhang *et al*., 2016 29)	Microbiome samples not taken

Table 2. Demographic, processing information, and study parameters for the five studies included in the final meta-analysis. If other variables were reported by the original authors but are not listed below, they were excluded for the purposes of multi-study integration.

	Paramsothy 2017 (11)	**Moayyedi 2015 (10)**	**Jacob 2016**	**Kump 2015**	**Nusbaum 2018 (38)**
Patients [Total: Patients: Controls]	[81: 41 FMT: 40 controls]	[75: 38 FMT: 37 controls]	[20: 20 FMT: 0]	[27: 17 FMT: 10 controls]	[9: 9 FMT: 0]
Donors	14	6	4	14	6
FMT type	Donor batch [3–7 donors/ infusion]	Individual Donor	Donor batch [2-donor infusion]	Individual Donor	Individual Donor
Donor relationship	Unrelated	Unrelated	Unrelated	Not reported	Recipient-identified [related & unrelated]
FMT route	Colonoscopy followed by enemas	Enema	Colonoscopy [TI + right colon]	Colonoscopy [initially right colon, then left colon]	Enema
FMT Frequency	40 [5/week for 8 weeks]	6 [weekly]	Single	5 [fortnightly infusions]	20 [4/day for 5 days]
Antibiotic treatment	No	No	No	Triple therapy [not specified] for 10 days]	No
Clinical Remission	11/41 [27%]	9/38 [24%]	3/20 [15%]	4/17 [24%]	3/9 [33%]
Sequencing platform	Illumina MiSeq	Illumina MiSeq	Illumina MiSeq	Illumina MiSeq	Illumina MiSeq
16S hypervariable region	V1 - V3	V3	V4	V4	V4
Timepoints of sample collection (in weeks)	0, 4, 8	0, 6	0, 2, 4	0, 2, 4, 6, 8, 12	0, 4

Clinical outcome assessment

The three possible clinical outcomes described in this analysis include clinical remission, clinical response, and no response. Clinical remission was defined for all studies, besides Nusbaum 2018, as complete mucosal healing with a total Mayo score ≤ 1, with steroid-free remission. Clinical remission was defined as a decrease in total Mayo score ≥ 3. A patient was classified as no response if they did not achieve the primary endpoint of clinical remission nor clinical response. For Nusbaum 2018, disease activity was assessed by the Pediatric Ulcerative Colitis Activity Index (PUCAI) (38, 39). Turner 2007 proved the PUCAI to be valid and report an excellent correlation with the Mayo score, with a Pearson rho correlation coefficient of 0.95 (39). Nusbaum 2018 defined clinical remission as a decrease in PUCAI <10, in concordance with the PUCAI User Guide proposed by Tanner 2007 (38, 39).

After samples were processed in our standard pipeline, there were a total of 76 patients who were classified as no response, 31 who achieved clinical remission, and six who achieved a clinical response. There is no statistical power from the clinical response group due to the minimal sample size. Therefore, we decided to group those six clinical response patients with the clinical remission group to improve statistical power while considering the limitations of conducting a meta-analysis. We classified this combined group as clinical response.

Standardized 16s sequence pipeline:

The raw fastq files were downloaded from EBI as described in the data acquisition section above. All samples were processed using QIIME 2 (v. 2020.8) and Qiita (30, 31). Qiita is a web-based multi-omics microbiome-based platform geared for meta-analyses (31). Qiita allows the user to employ QIIME2's standardized 16S sequence pipeline easily. If needed, primers were

removed using the QIIME 2 (v. 2020.8) cutadapt trim-paired command prior to being uploaded to Qiita. Once primers were removed, all the sequences were uploaded into Qiita. In Qiita, sequences were imported as paired-end data and demultiplexed. After demultiplexing, all sequences were trimmed. Sequence trimming reduced the total number of reads. However, it increases the total number of uniquely mapped reads by removing low-quality portions due to sequencing (32). In addition, all sequences were trimmed to a length of 150 bp. After trimming, sequences were filtered and denoised using Deblur (33). The resulting Deblur filtered and trimmed sequences from all studies were merged in Qiita. Taxonomy was assigned to the representative sequences using a pre-fitted sklearn-based taxonomy classifier with the GreenGenes-trained Naive Bayes classifier (gg-13-8-99-nb-classifier.qza) provided in Qiita by QIIME 2 (qiime2 2020.11.1). The BIOM file containing the sequences for all studies was exported out of Qiita for analysis in Python 3.6.0.

Figures were generated in R using ggpubr and ggplot2 (34, 35). Additionally, ASV's were collapsed at the genus level. The genus-level collapsed table was used for all downstream analyses. ASV-based analyses are the preferred method for microbiome research, as it harbors strain-level resolution. However, as listed in Table 2, more than one hypervariable region of the 16s gene was sequenced. Collapsing ASV's at the genus level helps to minimize some of the technical effects due to differences in the hypervariable region (36). So, while it is not preferred, it is a necessary quality control step that allows us to make meaningful biological discoveries while minimizing bias. Due to this limitation, Shannon diversity, genus evenness, and genus richness were calculated utilizing the collapsed genus table. Faiths PD, and Unweighted UniFrac were calculated with ASV's because SEPP's fragment insertion is designed for sOTU phylogenetic placement, which reduces the effects of the hypervariable region (43). When

needed, samples were rarefied to a sampling depth of 3,336 reads per sample. For the Robust Atchison PCA, samples were not rarefied, as recommended (37). Finally, P values were calculated using the Wilcoxon rank-sum test. Statistical significance was defined as having a $p < 0.05$ for all analyses.

Results

Sample sizes and collection

As described in the literature search strategy section of the methods, a total of 21 studies met the inclusion criteria (Fig. 1). A total of 113 patients with mild to severe ulcerative colitis across five different studies were included in this meta-analysis, 76 no response, 31 clinical remission, and six clinical response. As described previously, the six clinical response and 31 clinical remission patients were group together as classified as clinical response in the following analyses due to their low sample number yield and subsequent limited statistical power.

Each study took microbiome samples at different time points (Table 2). Longitudinal changes in the microbiome are a significant factor when analyzing microbiome data. We thus normalized the timepoint at which microbiome samples were compared by using the most synonymous sampling day collection across all studies. Four of the five studies took microbial samples four weeks post-FMT (Table 2). Moayyedi 2015 only took samples at day zero (initial colonoscopy visit) and six weeks post FMT (10). We, therefore, defined the "After FMT" timepoint as samples taken four weeks after initial FMT and six weeks in the case for Moayyedi 2015 (10). Here, "Post FMT" and "After FMT" are used interchangeably and refer to four weeks post-FMT and six weeks post-FMT for Moayyedi 2015 (10).

Alpha Diversity

We utilized Shannon's entropy, Faith's Phylogenetic Diversity (Faith's PD), Pielou's evenness (evenness), and the observed features (richness) to assess the differences in alpha diversity. We first looked at Faith's PD, Shannon, genus richness, and evenness of patients before receiving a fecal transplant (Timepoint 0) (Fig. 2). We saw no statistical differences in the raw Faiths' PD, Shannon diversity, or species evenness values between the clinical response and no response groups before FMT (Fig. 2). There were statistical differences in the raw richness values for those who achieve clinical response to those who do not respond prior to FMT (Fig. 2). To further explore the differences in genus richness values, we stratified the values for both clinical outcome groups by study (Fig. 3). We observed that Moayyedi 2015 contained very small genus richness values that may have been driving down the separation between clinical response and no response groups. We chose to calculate the spread of genus richness values without Moayyedi 2015 (Fig. 4). Even without Moayyedi 2015, the distribution of richness values for those who go on to reach a clinical response was still higher than those who do not achieve a response and reached significance (Fig. 4).

To further analyze each study's effect on the alpha diversity measures we discussed, we stratified both the raw Shannon and Faiths' PD values for those who achieved clinical response and those who did not before FMT by study (Fig. 5). We observed no significant separation due to study influencing either metric. Additionally, the distribution of values for responders vs. non-responders for each study seems to coincide with the lack of significance between the two clinical outcomes seen in figure 2 (Fig. 2). Due to no observed significance for Faiths PD, Shannon diversity, and genus richness, we did not perform additional analyses apart from study stratification (Fig. 5).

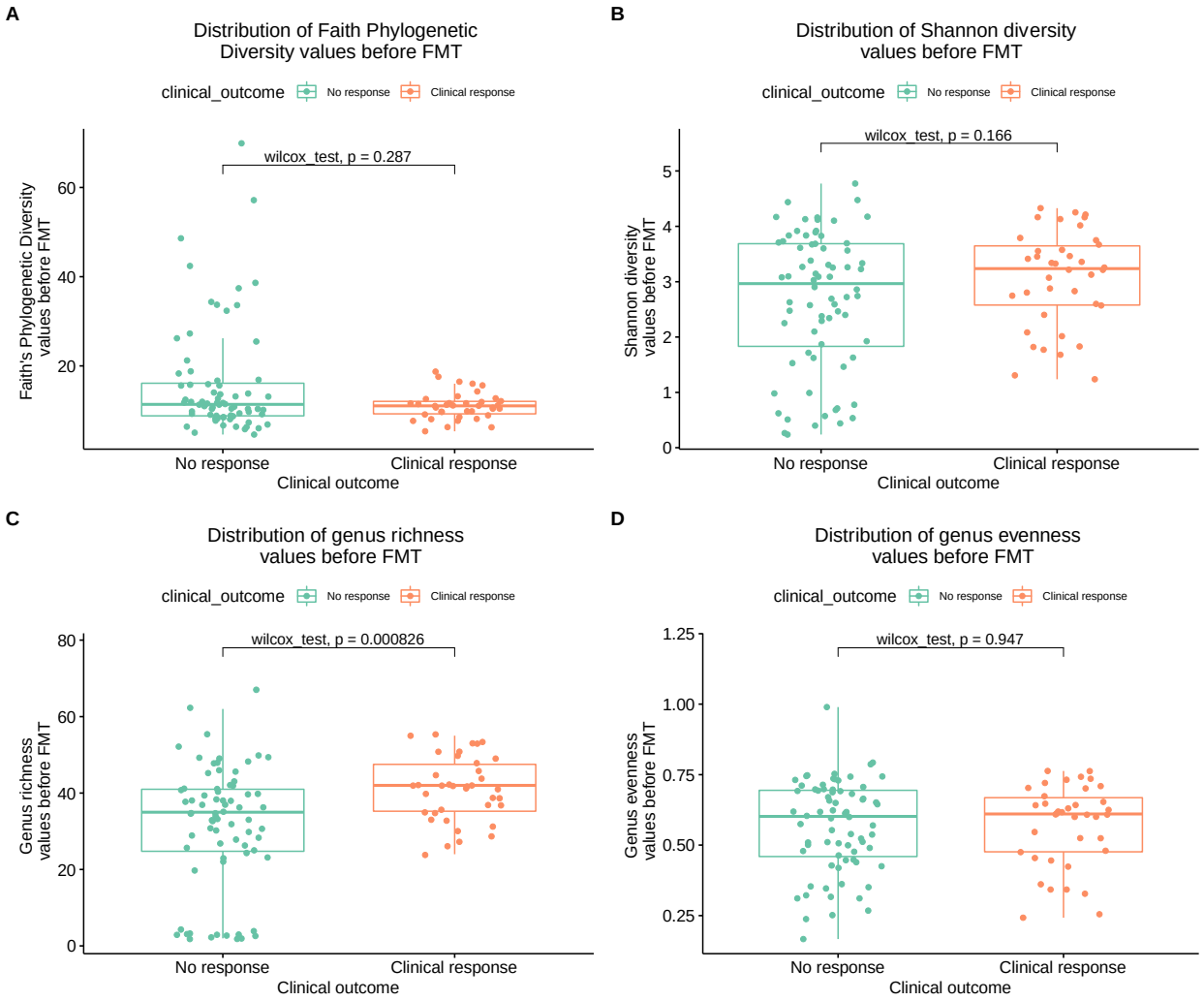


Figure 2. Distribution of Shannon, Faith's Phylogenetic Diversity, species richness, and species evenness before FMT. All panels display the respective alpha diversity measures along the y axis while stratifying patients based on their final response to FMT treatment. The distribution of species richness values seems to be higher in those who achieve who go on to achieve clinical response compared to those who do not. A Wilcoxon rank-sum test was run to determine significance at the standard $P < 0.05$ value. No other alpha diversity measure besides species richness reached significance.

The observed increase of genus richness in those who achieve clinical response highlights a potential propensity of the microbiome to regain functionality following FMT. If a patient's microbiome expresses very low species richness, it may be more challenging to restore the

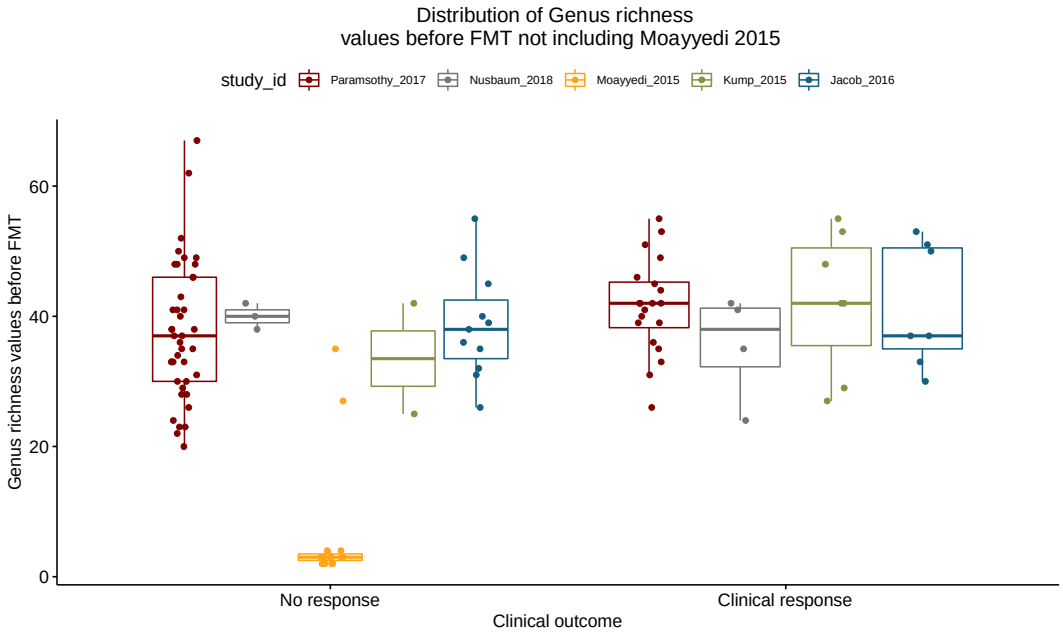


Figure 3. Distribution of genus richness values of patients before treatment, stratified by study and clinical outcome. The samples for Moayyedi 2015 can be observed to have very low genus richness when compared to the others which introduces biases when comparing genus richness for the clinical remission and no response groups when all studies are culminated.

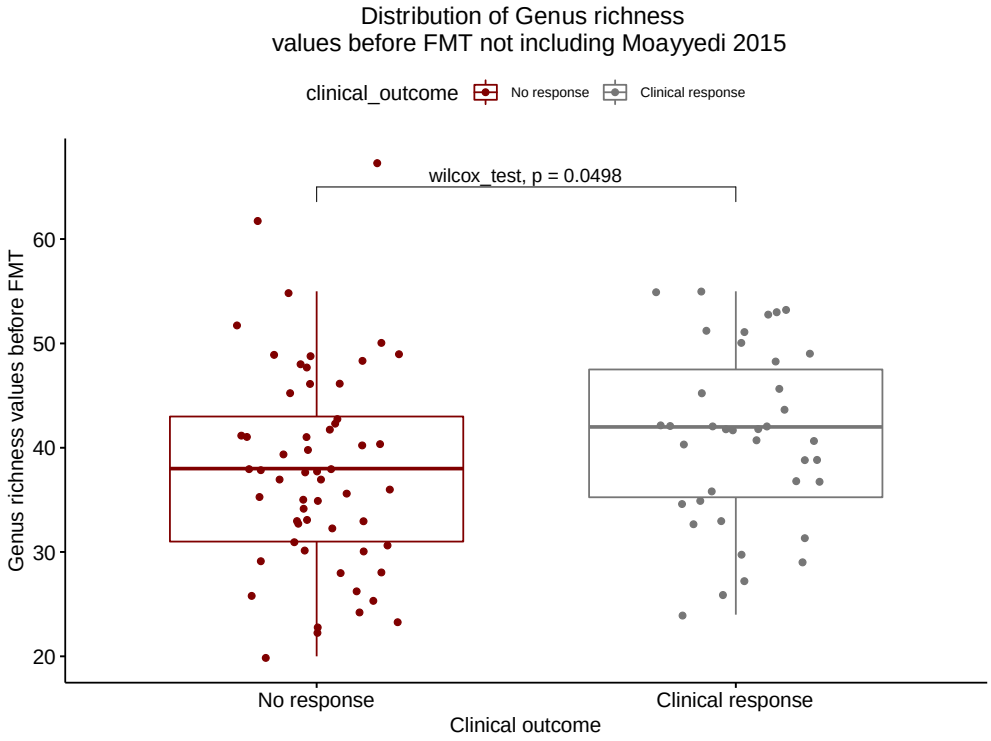


Figure 4. Distribution of genus richness values of responders and non-responders, not including Moayyedi 2015. The genus richness values for responders seems to be higher than non-responders and is statistically significant at the standard $P < 0.05$ value.

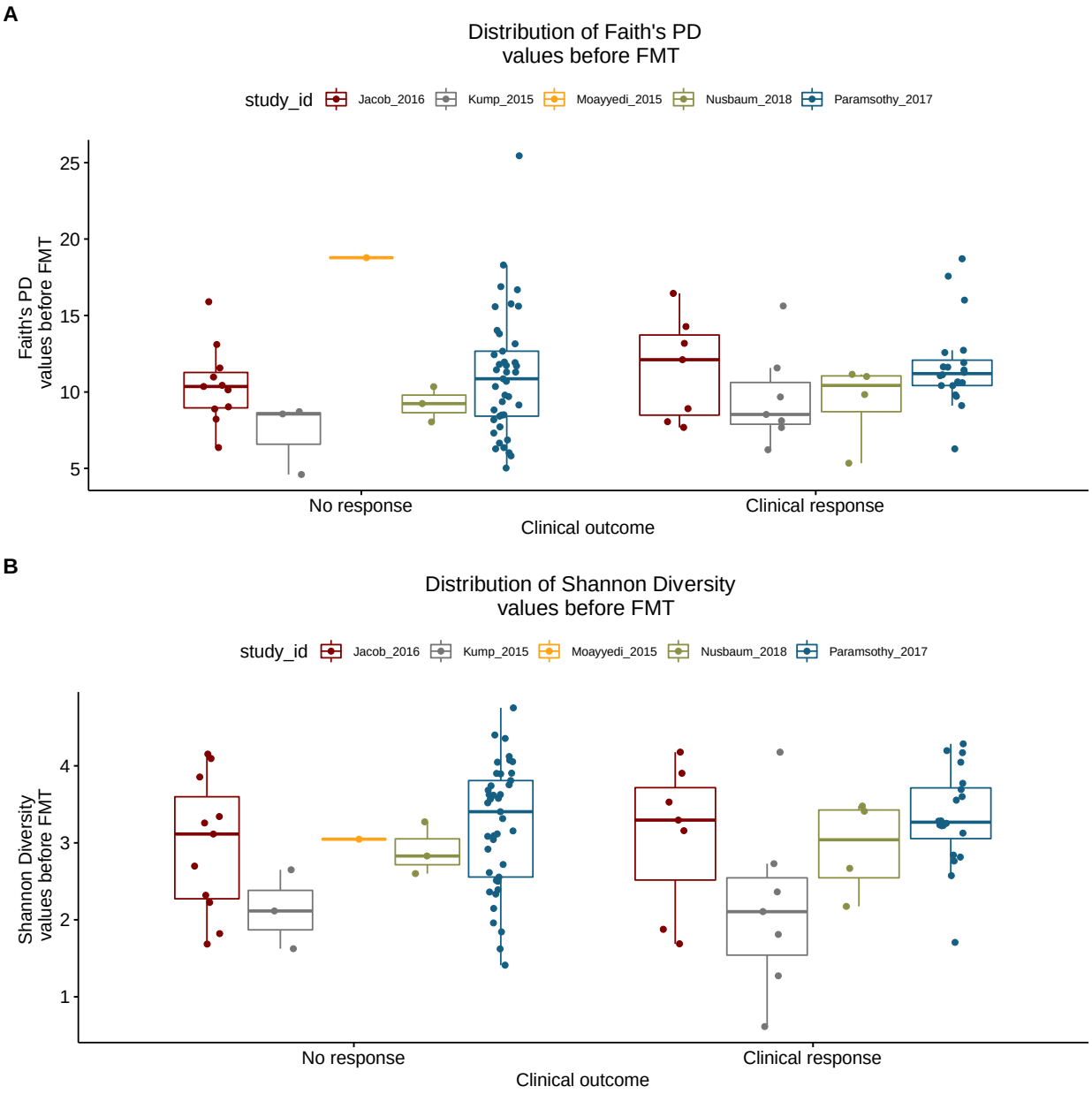


Figure 5. Distribution of Faith's PD and Shannon diversity values for responders and non-responders, stratified by study. Panel A plots the raw Faiths PD values. Panel B plots the raw distribution of Shannon diversity values. Values for responder's vs non-responders are similar for each study. No statistical tests were ran on the difference for each individual studies responders and non-responders.

diversity that was lost. Alternatively, those with high species richness may not be as clinically severe, and their microbiome may be more readily affected by FMT.

Such observations may explain why some patients achieve response, and others do not, from the same donor. However, more clinical case studies are needed to accurately assess how the severity of microbial dysbiosis in a UC patient affects the response to FMT treatment.

We then assessed whether patients who reached a clinical response experienced a more significant change in either of the four diversity after FMT than non-responders. We took the difference in alpha diversity values from after FMT to before FMT to calculate the change. We then plotted the distribution of differences for responders and non-responders for each of the four diversity measures (Fig. 6). As before, P values were calculated using the Wilcoxon rank-sum test. The change in all four alpha diversity measures between responders and non-responders was not significant. This indicated that responders do not experience a more considerable difference in alpha diversity than non-responders.

Lastly, we plotted the distribution of raw values of Faith's PD, Shannon, genus richness, and genus evenness before and after FMT, along with that of their donors, for responders vs. non-responders (Fig. 7). We see that for Faith's PD, the distribution of values before and after FMT is statistically significant. Faith's PD values after FMT are higher, and the distribution lies closer to the Faith's values of the donors for responders and non-responders. Even though they seem to be closer, the distribution of Faith's PD values after FMT is still statistically different from their donors. The distribution of Shannon's diversity values follows a similar trend as those observed for Faith's PD. The distribution of Shannon's diversity values before and after FMT for the patients that reached a clinical response and those that did not are also statistically different. The same trend is also observed for genus richness and evenness. The genus richness for responders and non-responders also increases but doesn't quite reach that of the donors. Genus evenness

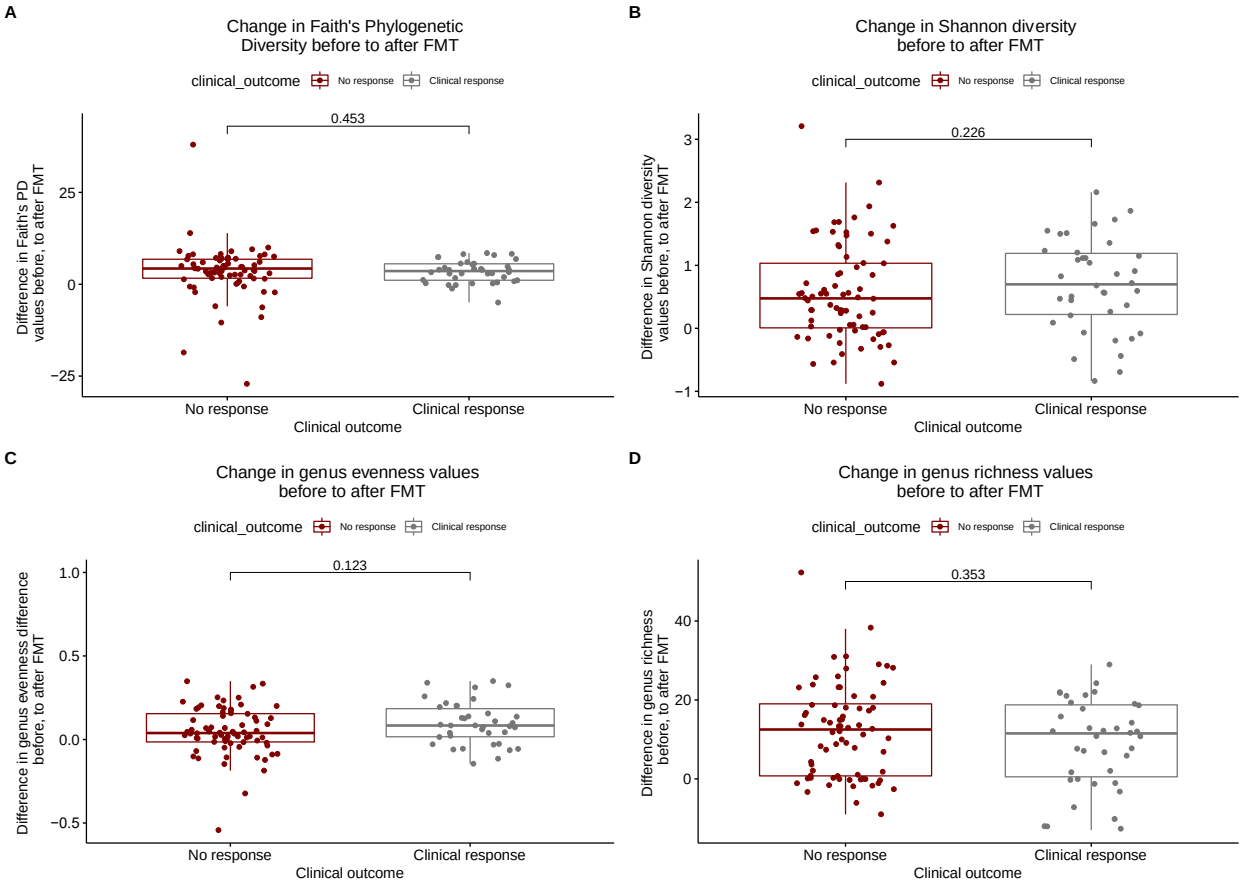


Figure 6. The delta changes of Faith's PD, Shannon diversity, genus evenness, and genus richness before to after FMT. All panels display the respective changes in alpha diversity measures before to after FMT. For all four alpha diversity measure, there seems to be no significance difference in the amount of change for responder's vs non-responders.

seems to increase for responders and non-responders to higher levels than donors (Fig. 7).

Overall, these results highlight that irrespective of clinical outcome, alpha diversity increases for all patients following FMT treatment. Patients' alpha diversity also increases but does not quite reach that of their donors. We did not stratify the changes in alpha diversity per study due to the limited statistical power of studies with minimal sample numbers.

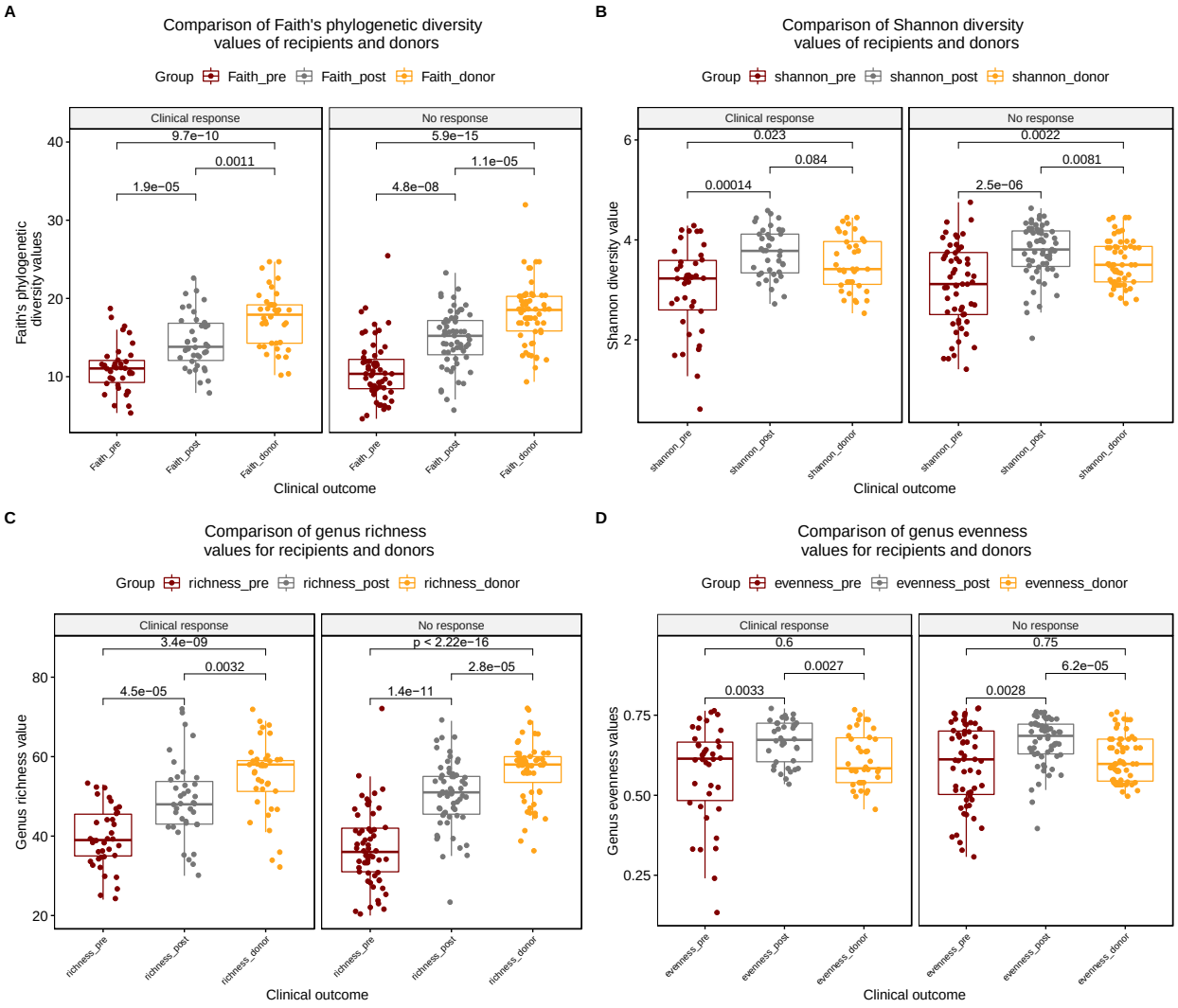


Figure 7. Distribution of Faith's PD, Shannon diversity, genus richness, and genus evenness values before and after FMT, along with their donors. Each dot in panels A, B, C, and D represents a patient or donor, with their Faiths PD, Shannon diversity, genus richness, and genus evenness values before and after FMT plotted along the y axis and stratified based on their clinical outcome. All groups are statistically significant besides genus evenness for responders and non-responders pre-FMT to their donors, utilizing the standard $P<0.05$ level (D).

Beta Diversity Comparisons

Beta diversity was assessed utilizing the unweighted UniFrac distance matrix (40). We chose to use unweighted UniFrac to evaluate the differences in --beta diversity. Unweighted UniFrac is a qualitative measure, meaning it does not consider the relative abundance of bacteria, merely the presence/absence of bacteria (40).

We first wanted to determine whether there were differences in how close a patient was to their donor before being given an FMT. We calculated the distance a patients' microbiome was to their donors before FMT from the unweighted UniFrac distance matrix and stratified them based on the patients' clinical outcome (Fig. 8a). The distribution of distances in those who achieved clinical remission to their donor or donor batch was not different from non-responders (Fig. 8a). This signifies that there is no evidence to support that the closer a patient is to their donor in terms of beta diversity, the greater likelihood they will achieve a response.

We also assessed whether those who achieve clinical response are closer in beta diversity to their donor than non-responders after FMT (Fig. 8b). We observed no statistical difference between responders and non-responders regarding their unweighted UniFrac distance to their donor after FMT (Fig. 8b).

Next, we wanted to see if a patients' beta diversity shifted to that of their donors. We plotted and compared the UniFrac distance a patient was to their donor before FMT and after FMT (Fig. 9). The pairwise Wilcoxon rank-sum test was used to determine significance at the standard $P < 0.05$ level. We computed the pairwise distance each patient was to their donor before to after FMT. For responders and non-responders, patients experienced a statistically significant drop in the distance to their donor after FMT (Fig. 9). However, the magnitude of

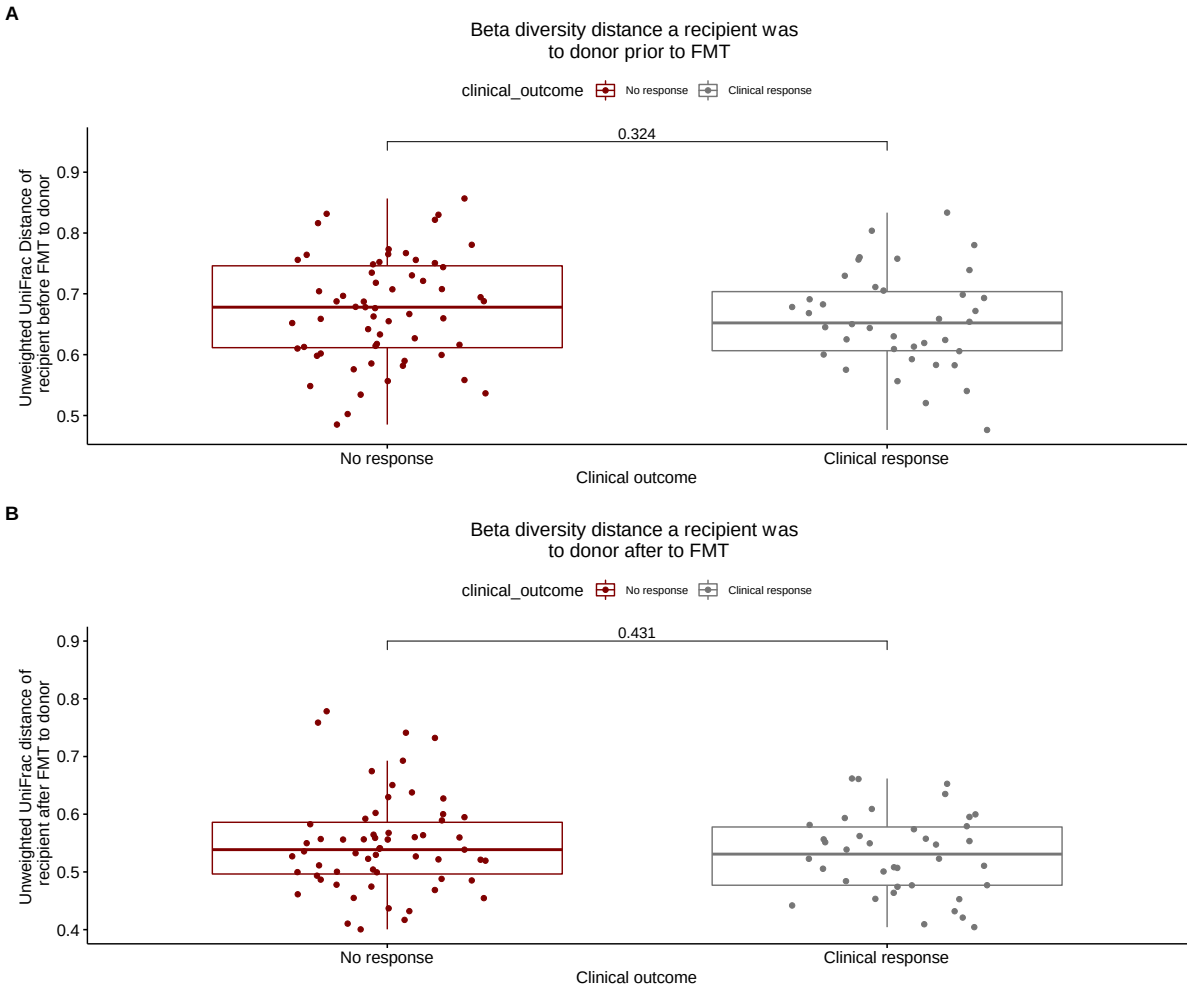


Figure 8. Beta diversity distances of recipients to their donor before FMT and after FMT stratified based on their clinical outcome. Panel A compares the UniFrac distance a patient was to their donor prior to FMT for responders and non-responders. Panel B compares the UniFrac distance a patient was to their donor after FMT for responders and non-responders. Each dot represents a patient, with their unweighted UniFrac distance to their donor plotted along the y axis. Neither comparison is statistically significant.

change is irrespective of their clinical outcome. Additionally, we would expect that those who achieve clinical response would experience a more significant decrease in the distance to their donor than the no response group. However, it seems that the difference in distance to their donor for those with no response is more significant than that of those who achieve a clinical response (Fig. 9).

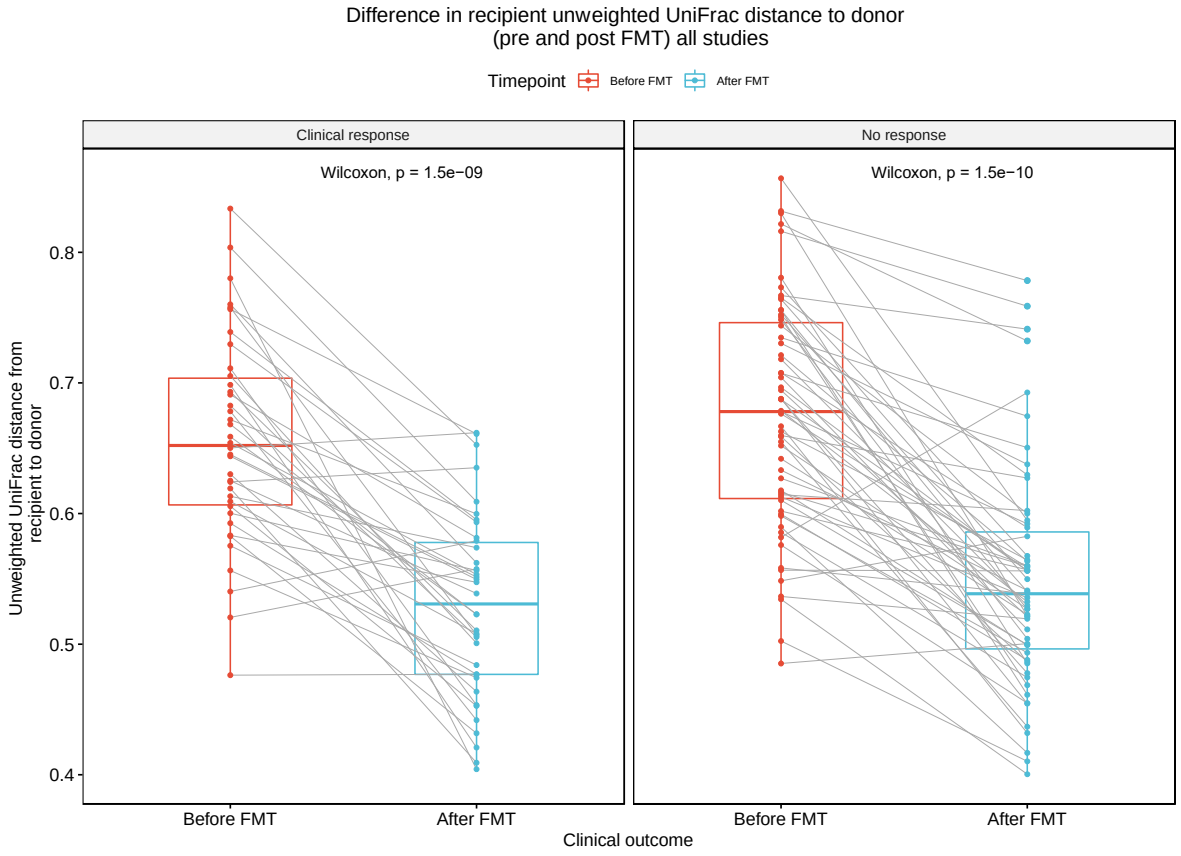


Figure 9. Beta diversity distances of recipients to their donor before and after FMT stratified based on their clinical response. Each dot represents a patient, with their unweighted UniFrac distance before and after FMT to that of their donor plotted along the y axis. The lines connect the individual patients pre timepoint, to their post FMT timepoint (4 – 6 weeks post FMT). The distance of recipients to their donor before to after FMT for responders and non-responders are statistically significant.

Differential abundance analysis

We additionally performed differential abundance testing to see if any bacteria were differentially expressed in donors and responders vs. non-responders before and after FMT. We utilized ANCOM (analysis of composition of microbes) for differential abundance testing (44). ANCOM detects differences in mean taxonomic abundance between groups. Additionally, we chose ANCOM because other differential abundance methods are unsuitable for comparing different microbial populations based on taxonomic abundances down to the ecosystem level and

are only suitable for drawing inferences of the specimen taken from the ecosystem. Additionally, ANCOM provides strict thresholds for accepting and rejecting whether a taxon is differentially expressed (44). Prior to running ANCOM, we added a pseudo count of 1, as ANCOM runs on the basis of log transformations, on the feature table collapsed to the genus level to avoid bias introduced from the hypervariable region.

We first performed ANCOM on patients before FMT. We observed enrichment of the family Enterobacteriaceae (w = 366), genus *Lactobacillus* (w = 668), and genus *Streptococcus* (w = 396) in responders vs non-responders prior to FMT. In half of the samples most abundant in Enterobacteriaceae, there was a median of 114 reads for responders, while the median value for non-responders was only 11 (Table 3). The genus *Lactobacillus* had a median of 109 in responders before FMT, while the median for non-responders was 1.0 (the pseudo count added) (Table 3). The genus *Streptococcus* had a median of 623 for responders while the median for non-responders was 1.0 (the pseudo count added) (Table 3). This data suggests that before FMT, the genus *Lactobacillus*, *Streptococcus*, and the family Enterobacteriaceae are more abundant for those who go on to achieve clinical response than those who do not.

Table 3. Differential abundance of key bacteria by percentile for both responders and non-responders prior to FMT treatment. We see that the family *Enterobacteriaceae*, and the genus *Lactobacillus* and *Streptococcus* are differentially abundant in responders than non-responders.

	Clinical response	**Clinical response**	**No response**	**No response**
Abundance by percentile:	25.0	50.0	25.0	50.0
***Enterobacteriaceae* (W = 366)**	10.75	114.0	1.0	11.0
***Lactobacillus* (W = 668)**	3.0	109.0	1.0	1.0
***Streptococcus* (W = 396)**	25.0	623.0	1.0	3.0

We performed differential abundance testing on patients after FMT as well. We show that the genus Alistipes, Bilophila, Coprococcus, Ruminococcus, Roseburia, Bifidobacterium, and Faecalibacterium were all differentially expressed in both donors and responders post-FMT, then non-responders post-FMT (Table 4). A study by Vital 2017 performed metagenomic and metatranscriptomic analyses from fifteen different data sets to classify and assess different microbe's ability to produce butyrate in the human gut (45). Their study identifies Ruminococcus, Coprococcus, Alistipes, and Faecalibacterium as major butyrate producers in the gut (45). We show here that not only are these bacteria prevalent in healthy donors, but they are also differentially expressed in responders and not prevalent in non-responders. The increase in butyrate producers in responders compared to non-responders signify that these bacteria may play a role in disease pathogenesis and contribute to mucosal health and healing.

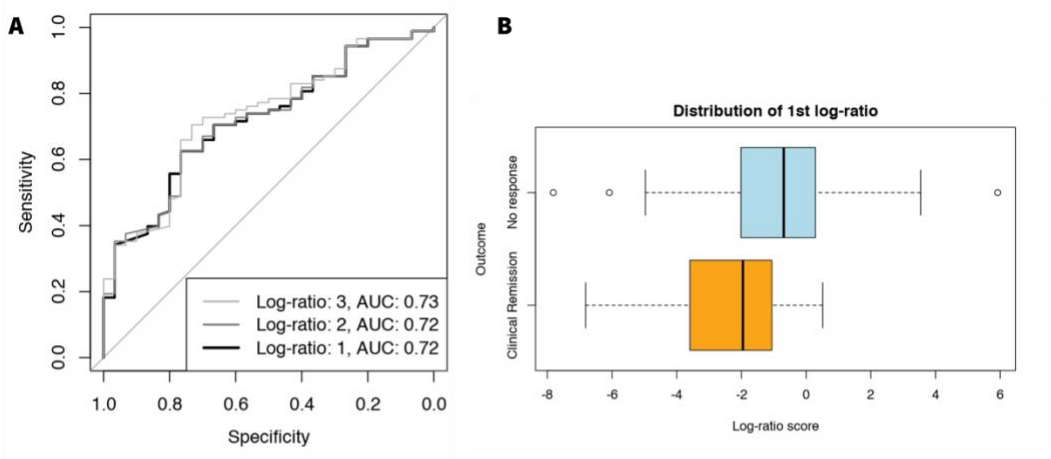
Table 4. Differential abundance of key bacteria by percentile for donors, responders, and non-responders four to six weeks post FMT treatment. We see that the family *Alistipes*, *Bilophila*, *Coprococcus*, *Ruminococcus*, *Roseburia*, *Bifidobacterium*, and *Faecalibacterium* are differentially abundant in donors and responders and not in non-responders.

	Donors	**Donors**	**Clinical response**	**Clinical response**	**No response**	**No response**
Abundance by percentile:	25.0	50.0	25.0	50.0	25.0	50.0
***Alistipes* (W = 385)**	47.75	863.0	10.0	348.0	1.0	61.5
***Bilophila* (W = 370)**	1.0	16.5	1.0	3.0	1.0	1.0
***Coprococcus* (W = 390)**	223.5	494.5	118.5	552.0	1.75	144.0
***Ruminococcus* (W = 390)**	657.00	1448.0	112.0	523.0	7.75	216.5
***Roseburia* (W = 374)**	150.75	470.0	33.5	291.0	3.0	94.5
***Bifidobacterium* (W = 379)**	1.0	11.0	1.0	7.0	1.0	2.0
***Faecalibacterium* (W = 372)**	110.0	363.5	29.5	471.0	1.0	102.0

Predictive log ratio analysis of responders and non-responders

To further assess differences in the microbial composition of responders and non-responders, we utilized a novel machine learning algorithm aimed for identifying predictive microbial log-ratio biomarkers from sequencing data, CoDaCoRe (compositional data continuous relaxations). To identify and assess predictive taxonomic log-ratios of remission in patients both before and after FMT. We first analyzed patients prior to FMT. After running CoDaCoRe, we obtained a predictive and informative log-ratio of taxa predictive of no response prior to FMT treatment with an area under the receiver operator characteristic (AUROC) = 0.72 (Fig. 10). The log ratio of the model was trained to predict and classify non-responders to responders. To maximize the specificity and sensitivity of the model, we only included patients who achieved clinical remission and no response patients and removed clinical response patients from this analysis. The log-ratio reported the classification of non-responders as having an increase in the family Lachnospiraceae, and the genus Proteobacteria and Sutterella. Additionally, there were decreases in the genus Bacteroides, Prevotella, Alistipes, Streptococcus, and the families Enterobacteriaceae and Rikenellaceae (Fig. 10). When we tested the model on a smaller testing set, we received an AUC of 0.55. This indicated very low predictability of the model for novel samples (clinical remission n = 8, no response n = 22).

In addition, we performed CoDaCoRe on patients after FMT to predict and identify bacterial log-ratios for responders vs. non-responders. We obtained a log-ratio of taxa predictive of stratifying non-responders after FMT to those achieving clinical remission with an area under the receiver operator characteristic (AUROC) = 0.79 (Fig. 11). The log ratio of the model was trained to predict and classify non-responders to responders after FMT. The log-ratio reported the classification of non-responders as having an increase in the family Lachnospiraceae, the



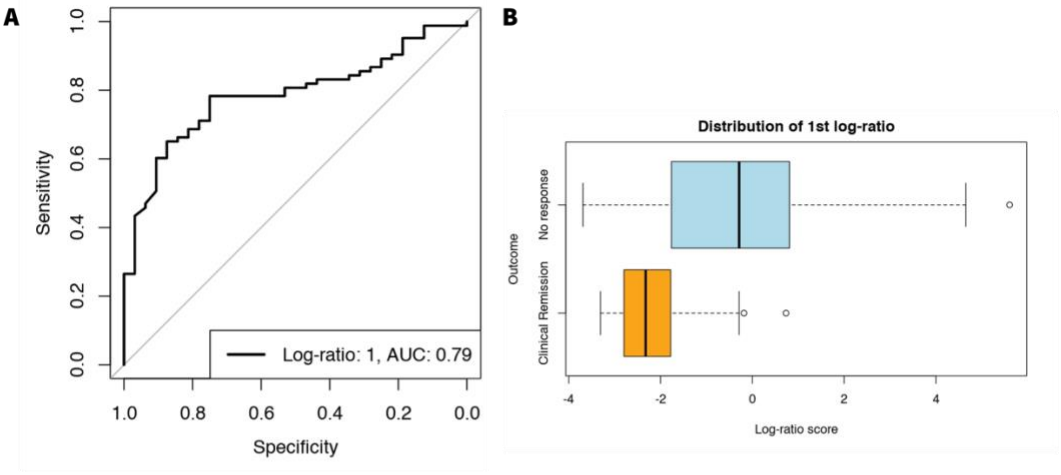
C

Log-ratio of model	
Numerator:	*o__Clostridiales;f__Lachnospiraceae* *k__Bacteria;p__Proteobacteria;* *f__Alcaligenaceae;g__Sutterella*
Denominator:	*f__Bacteroidaceae;g__Bacteroides* *f__Prevotellaceae;g__Prevotella* *o__Bacteroidales;f__Rikenellaceae;g__* *f__Rikenellaceae;g__Alistipes* *f__Streptococcaceae;g__Streptococcus* *o__Enterobacteriales;f__Enterobacteriaceae;__*
AUC: 0.718 ***Slope: 0.379***	

Figure 10. Outputs from CoDaCoRe of recipients before FMT. Panel A plots the receiver operating characteristic of the model's performance. Three log-ratios were detected, however model 1 is plotted and described in panel B and C. Panel B plots the distribution of log-ratio scores for the samples classified as clinical remission and no response, in the training set. Those classified as no response have a larger log-ratio of bacteria in the numerator, and a smaller log-ratio of bacteria in the denominator. The AUROC was 0.718, with a slope of 0.379.

phylum Proteobacteria, and the genus Fusobacterium, Sutterella, and Clostridium. Additionally, there were decreases in the genus *Bacteroides*, *Bifidobacterium*, *Parabacteroides*, *Alistipes*, *Blautia*, *Faecalibacterium*, *Roseburia*, *Phascolarctobacterium*, and the families Rikenellaceae and Erysipelotrichaceae (Fig. 10). When the model was assessed on the testing set, there was an AUC of 0.51. So, while the AUROC was relatively high (0.79), its' predictability on classifying new samples was very low. One thing to note is that the testing set contained a small sample set of clinical remission patients (n = 8) and non-responders (n = 21). Due to the nature of running machine learning models, most of the data is utilized for the training set, and little is used for the

testing set. It would be interesting to see how this set performs on other data sets developed in the future.



C

Log-ratio of model	
Numerator	f__Streptococcaceae;g__Streptococcus f__Clostridiaceae;g__Clostridium o__Clostridiales;f__Lachnospiraceae f__Fusobacteriaceae;g__Fusobacterium k__Bacteria;p__Proteobacteria f__Alcaligenaceae;g__Sutterella
Denominator	f__Bifidobacteriaceae;g__Bifidobacterium o__Bacteroidales;f__;g__ f__Bacteroidaceae;g__Bacteroides f__Porphyromonadaceae;g__Parabacteroides f__Rikenellaceae;g__ f__Rikenellaceae;g__Alistipes f__Lachnospiraceae; g__Blautia f__Lachnospiraceae; g__Coprococcus f__Lachnospiraceae; g__Roseburia f__Ruminococcaceae;g__Faecalibacterium f__Veillonellaceae;g__Phascolarctobacterium o__Erysipelotrichales;f__Erysipelotrichaceae
AUC: 0.7855798 **Slope: 0.8095308**	

Figure 11. Outputs from CoDaCoRe of recipients after FMT. Panel A plots the receiver operating characteristic of the model's performance. Panel B plots the distribution of log-ratio scores for the samples classified as clinical remission and no response, in the training set. Those classified as no response have a larger log-ratio of bacteria in the numerator, and a smaller log-ratio of bacteria in the denominator after FMT. The AUROC was 0.786, with a slope of 0.81.

The findings from running CoDaCoRe are similar to the results from the ANCOM differential abundance testing described in the previous section. We see that before given FMT, patients classified as non-responders have a decreased abundance of the genus Lactobacillus and Streptococcus (Table 3.). Streptococcus was also differentially abundant in responders before

FMT and not in non-responders (Table 3). Additionally, Streptococcus was noted in the denominator of the model developed from recipients before FMT, indicating that non-responders have a lower log-ratio of Streptococcus than clinical remission patients. The harmonious results from ANCOM and CoDaCoRe provide evidence to support that patients before FMT harbor distinct microbial taxa. Patients who achieve clinical remission may have increased Streptococcus, Enterobacteriaceae, and Lactobacillus than their clinical response counterparts.

Responders and non-responders also harbor distinct microbial signatures from one another after FMT. Differential abundance testing confirmed the increases in the genus Alistipes, Bilophila, Coprococcus, Ruminococcus, Roseburia, Bifidobacterium, and Faecalibacterium (Table 4). These taxa were also highly abundant in donors, indicating a potential increase in successful engraftment of essential beneficial bacteria contributing to mucosal healing and an increase in butyrate production as described prior. The results from CoDaCoRe also support the existence of decreased log-ratios for taxa such as Alistipes, Bilophila, Coprococcus, Faecalibacterium, and Blautia. Alistipes, Coprococcus, for patients who do not achieve remission (Fig. 11). It is also important to note that before FMT, a low log-ratio of Alistipes was also indicative of non-responders than responders.

When taken together, these results indicate an interesting duality of the importance of both the donor and the patients' microbiome before FMT. Having key taxa before FMT may allow a patient to achieve remission more readily than those who have, let's say, poor species richness. It may also be that patients with severe microbial dysbiosis and poor species richness require a much longer FMT treatment than others to successfully culminate and engraft the beneficial microbes necessary for mucosal health and healing.

Discussion

Meta-analyses

While many meta-analyses exist to date, this analysis is of the few that evaluated various aspects of both the patient and donor microbiome when assessing the efficacy of fecal microbiota transplant to treat Ulcerative Colitis. We searched the literature and contacted authors to identify and collect as much data as possible. It took eight months from the study's inception to search and collect sequence data and comprehensive metadata from authors. Many studies were excluded for reasons we should strive to mitigate in the name of reproducibility. Four studies were excluded from this analysis for their use of specific sequencing platforms. The addition of those studies would drastically increase the statistical power and subsequent biological insight that can be drawn. It was out of the scope of this analysis, but it is of interest to develop tools, if not already being explored today, to minimize technical effects due to sequencing platform. Such developments have the power to considerably increase the pool of data from which researchers can pull for future meta-analyses.

Three studies were excluded for not being able to either de-identify metadata or did not respond after multiple attempts to contact the authors. Meta-analyses are extremely valuable for their ability to generalize the observations found in individual studies. They also allow us to discover how various aspects of processing and study design impact finding. When seven studies that meet the study criteria are excluded, it severely limits the biological inferences that can be drawn. The scientific community should strive to make data open-access and properly deposit their metadata and sequences to allow researchers to synthesize the literature and draw meaningful conclusions to propel scientific exploration forwards. It might also be worthwhile to develop metadata repositories that make it easier for researchers to share their data.

Even with a portion of the data reported in the literature, we were able to culminate a decent-sized data set and tackle some of the critical questions that remain regarding the efficacy of FMT to treat ulcerative colitis. We set out first to explore aspects of the patient and donor microbiome that may play a role in inducing clinical remission. We show that Faith's PD, Shannon diversity, and genus evenness of a patient's microbiome before FMT do not differ in responders and non-responders. We highlight how patients who go on to achieve clinical response have a higher species richness than non-responders. In addition, we show that Shannon diversity, Faith's PD, genus richness, and genus evenness increase for all patients after FMT regardless of their clinical outcome. Their increase is also not different based on a patient's clinical outcome. Additionally, patient's alpha diversity does increase but does not quite match that of their donor. It is important to note here that samples after FMT were assessed four weeks after their first fecal transplant. Many patients have experienced years of mucosal inflammation, and it may take longer than four weeks to revert the dysbiosis of the enteric microbiome. It is of interest for clinicians to consider the importance of longitudinal samples and longer-term follow-ups whenever possible.

We also observed no difference in the distance a patient was in beta-diversity to their donor for responders and non-responders. There was no difference in responders' and non-responders' distance to their donor after FMT either. However, regardless of their clinical outcome, patients experience a shift in beta diversity to that of their donors.

To assess microbial composition, we performed differential abundance testing on the patients before FMT and after FMT, along with the donor microbiomes. We show that prior to FMT, those who achieve a clinical response have an increased abundance of butyrate producers such as Ruminococcus, Roseburia, Blautia, and Coprococcus (41). These findings are consistent

with results from another study that found that sustained remission of UC patients was associated with butyrate-producing bacteria (42). Additionally, those who do not achieve clinical remission have markedly lower if any differentially expressed butyrate producers. After FMT, we see that patients who achieve a clinical response have high levels of key taxa that are also differentially expressed in donors and not in non-responders.

Lastly, we developed two machine learning models to predict and identify log-ratios of taxa indicative of clinical response. We see that those who do not achieve clinical remission have lower log-ratios of Bacteroides, Prevotella, Alistipes, Streptococcus, and the families Enterobacteriaceae and Rikenellacea. Bacteria such as the genus Alistipes were also higher in responders than non-responders before FMT. Furthermore, we show that after FMT, non-responders are still identified by our model to have lower log-ratios of Alistipes and known butyrate producers such as Bilophila, Coprococcus, Faecalibacterium, and Blautia (42). Overall we show an interesting interaction that a patient's microbiome prior to FMT may indicate an increased propensity to achieve clinical remission after treatment. Like mentioned before, it is not definitive whether patients with higher species richness and increased butyrate production have a less severe case of UC and thus are more readily impacted by FMT or if it the presence of specific beneficial taxa allows for increased engraftment from the donor and increased mucosal healing.

It would be interesting to explore how the metaproteome or metabolome of the patients impacts strain engraftment from the donor. Such multi-omic analyses of the microbial ecosystem can allow us to assess if there are functional properties of the patient's microbial ecosystem that contribute to remission. Such data could help clinicians find more targeted donors or formulate

more comprehensive donor selection methods that address the state of the individual's microbiome.

In conclusion, we highlight a similar trend documented before, that all most patients experience a shift in their microbiome to that of their donor regardless of clinical outcome. We also describe the increased abundance of butyrate-producing bacteria in patients who achieve clinical remission and in donors. We make a case that both the donor and patient's microbiomes play a role in determining whether a patient will achieve remission or not and that they should be explored further. Future studies should utilize multi-omics techniques to understand better how a patients' microbiomes before FMT correlate with achieving remission. Targeted analysis of patients and their donors can lead to the development of not a one size fits all approach to FMT's but a more integrated and holistic donor selection process that is severely lacking today.

All in all, this meta-analysis highlights aspects of the patient microbiome that shift in response to fecal transplants. However, it is important to note that only five studies were included with little assessment of co-variates, so it is challenging to draw concise and definitive conclusions of how the patient and donor's enteric microbiome interact and communicate. Future analyses should also attempt to include studies of different sequencing platforms to increase the overall sample size and statistical power, and answer some of the many questions that remain regarding donor and patient matching, ideal dosage, and patient predictors for the likeliness of responding such as, length of disease, and disease severity. Additionally, future RCTs should, if possible, increase the length of follow-up and sample collection to investigate multi-study longitudinal markers of remission.

Material from this thesis is coauthored by Mehrbod Estaki. The thesis author was the primary author of this material.

Material from this thesis is coauthored by Justin Shaffer. The thesis author was the primary author of this material.

REFERENCES CITED

1. Moller, F. T., Andersen, V., Wohlfahrt, J., & Jess, T. (2015). Familial risk of inflammatory bowel disease: a population-based cohort study 1977-2011. *The American journal of gastroenterology*, *110*(4), 564–571. https://doi.org/10.1038/ajg.2015.50

2. Ananthakrishnan A. N. (2015). Epidemiology and risk factors for IBD. *Nature reviews. Gastroenterology & hepatology*, *12*(4), 205–217.https://doi.org/10.1038/nrgastro.2015.34

3. Lloyd-Price J, Arze C, Ananthakrishnan AN, Schirmer M, Avila-Pacheco J, Poon TW, Andrews E, Ajami NJ, Bonham KS, Brislawn CJ, Casero D, Courtney H, Gonzalez A, Graeber TG, Hall AB, Lake K, Landers CJ, Mallick H, Plichta DR, Prasad M, Rahnavard G, Sauk J, Shungin D, Vázquez-Baeza Y, White RA 3rd; IBDMDB Investigators, Braun J, Denson LA, Jansson JK, Knight R, Kugathasan S, McGovern DPB, Petrosino JF, Stappenbeck TS, Winter HS, Clish CB, Franzosa EA, Vlamakis H, Xavier RJ, Huttenhower C. (2019). Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases. *Nature*, *569*(7758), 655–662. https://doi.org/10.1038/s41586-019-1237-9

4. Zhou, H. Y., Guo, B., Lufumpa, E., Li, X. M., Chen, L. H., Meng, X., & Li, B. Z. (2021). Comparative of the Effectiveness and Safety of Biological Agents, Tofacitinib, and Fecal Microbiota Transplantation in Ulcerative Colitis: Systematic Review and Network Meta-Analysis. *Immunological investigations*, *50*(4), 323–337. https://doi.org/10.1080/08820139.2020.1714 650

5. Dulai, P. S., Siegel, C. A., Colombel, J. F., Sandborn, W. J. & Peyrin-Biroulet, L. (2014). Systematic review: Monotherapy with antitumour necrosis factor alpha agents versus combination therapy with an immunosuppressive for IBD. Gut 63, 1843- 1853, doi:10.1136/gutjnl-2014-307126

6. Human Microbiome Project, C. (2012). Structure, function and diversity of the healthy human microbiome. Nature 486, 207-214, doi:10.1038/nature11234

7. Sartor, R. B. & Wu, G. D. (2017). Roles for Intestinal Bacteria, Viruses, and Fungi in Pathogenesis of Inflammatory Bowel Diseases and Therapeutic Approaches. Gastroenterology 152, 327-339 e324, doi:10.1053/j.gastro.2016.10.012

8. EISEMAN, B., SILEN, W., BASCOM, G. S., & KAUVAR, A. J. (1958). Fecal enema as an adjunct in the treatment of pseudomembranous enterocolitis. *Surgery*, *44*(5), 854–859.

9. Kassam, Z., Lee, C. H., Yuan, Y., & Hunt, R. H. (2013). Fecal microbiota transplantation for Clostridium difficile infection: systematic review and meta-analysis. *The American journal of gastroenterology*, *108*(4), 500–508. https://doi.org/10.1038/ajg.2013.59

10. Moayyedi, P., Surette, M. G., Kim, P. T., Libertucci, J., Wolfe, M., Onischi, C., Armstrong, D., Marshall, J. K., Kassam, Z., Reinisch, W., & Lee, C. H. (2015). Fecal Microbiota Transplantation Induces Remission in Patients With Active Ulcerative Colitis in a Randomized Controlled Trial. *Gastroenterology*, *149*(1), 102–109.e6. https://doi.org/10.1053/j.gastro.2015.04.001

11. Paramsothy, S., Kamm, M. A., Kaakoush, N. O., Walsh, A. J., van den Bogaerde, J., Samuel, D., Leong, R., Connor, S., Ng, W., Paramsothy, R., Xuan, W., Lin, E., Mitchell, H. M., & Borody, T. J. (2017). Multidonor intensive faecal microbiota transplantation for active ulcerative colitis: a randomised placebo-controlled trial. *Lancet (London, England)*, *389*(10075), 1218–1228. https://doi.org/10.1016/S0140-6736(17)30182-4

12. Duvallet, C., Zellmer, C., Panchal, P., Budree, S., Osman, M., & Alm, E. J. (2019). Framework for rational donor selection in fecal microbiota transplant clinical trials. *PloS one*, *14*(10), e0222881. https://doi.org/10.1371/journal.pone.0222881

13. Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & PRISMA Group. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. Journal of clinical epidemiology, 62(10), 1006-1012.

14. Angelberger, S., Reinisch, W., Makristathis, A., Lichtenberger, C., Dejaco, C., Papay, P., Novacek, G., Trauner, M., Loy, A., & Berry, D. (2013). Temporal bacterial community dynamics vary among ulcerative colitis patients after fecal microbiota transplantation. *The American journal of gastroenterology*, *108*(10), 1620–1630. https://doi.org/10.1038/ajg.2013.257

15. Kunde, S., Pham, A., Bonczyk, S., Crumb, T., Duba, M., Conrad, H., Jr, Cloney, D., & Kugathasan, S. (2013). Safety, tolerability, and clinical response after fecal transplantation in children and young adults with ulcerative colitis. *Journal of pediatric gastroenterology and nutrition*, *56*(6), 597–601. https://doi.org/10.1097/MPG.0b013e318292fa0d

16. Costello, S. P., Hughes, P. A., Waters, O., Bryant, R. V., Vincent, A. D., Blatchford, P., Katsikeros, R., Makanyanga, J., Campaniello, M. A., Mavrangelos, C., Rosewarne, C. P., Bickley, C., Peters, C., Schoeman, M. N., Conlon, M. A., Roberts-Thomson, I. C., & Andrews, J. M. (2019). Effect of Fecal Microbiota Transplantation on 8-Week Remission in Patients With Ulcerative Colitis: A Randomized Clinical Trial. *JAMA*, *321*(2), 156–164. https://doi.org/10.1001/jama.2018.20046

17. Cui, B., Li, P., Xu, L., Zhao, Y., Wang, H., Peng, Z., Xu, H., Xiang, J., He, Z., Zhang, T., Nie, Y., Wu, K., Fan, D., Ji, G., & Zhang, F. (2015). Step-up fecal microbiota transplantation strategy: a pilot study for steroid-dependent ulcerative colitis. *Journal of translational medicine*, *13*, 298. https://doi.org/10.1186/s12967-015-0646-2

18. Damman, C. J., Brittnacher, M. J., Westerhoff, M., Hayden, H. S., Radey, M., Hager, K. R.,

Marquis, S. R., Miller, S. I., & Zisman, T. L. (2015). Low Level Engraftment and Improvement following a Single Colonoscopic Administration of Fecal Microbiota to Patients with Ulcerative Colitis. *PloS one*, *10*(8), e0133925. https://doi.org/10.1371/journal.pone.0133925

19. Karolewska-Bochenek, K., Grzesiowski, P., Banaszkiewicz, A., Gawronska, A., Kotowska, M., Dziekiewicz, M., Albrecht, P., Radzikowski, A., & Lazowska-Przeorek, I. (2018). A Two-Week Fecal Microbiota Transplantation Course in Pediatric Patients with Inflammatory Bowel Disease. *Advances in experimental medicine and biology*, *1047*, 81–87. https://doi.org/10.1007/5584_2017_123

20. Kellermayer, R., Nagy-Szakal, D., Harris, R. A., Luna, R. A., Pitashny, M., Schady, D., Mir, S. A., Lopez, M. E., Gilger, M. A., Belmont, J., Hollister, E. B., & Versalovic, J. (2015). Serial fecal microbiota transplantation alters mucosal gene expression in pediatric ulcerative colitis. *The American journal of gastroenterology*, *110*(4), 604–606. https://doi.org/10.1038/ajg.2015.19

21. Suskind, D. L., Singh, N., Nielson, H., & Wahbeh, G. (2015). Fecal microbial transplant via nasogastric tube for active pediatric ulcerative colitis. *Journal of pediatric gastroenterology and nutrition*, *60*(1), 27–29. https://doi.org/10.1097/MPG.0000000000000544

22. Vermeire, S., Joossens, M., Verbeke, K., Wang, J., Machiels, K., Sabino, J., Ferrante, M., Van Assche, G., Rutgeerts, P., & Raes, J. (2016). Donor Species Richness Determines Faecal Microbiota Transplantation Success in Inflammatory Bowel Disease. *Journal of Crohn's & colitis*, *10*(4), 387–394. https://doi.org/10.1093/ecco-jcc/jjv203

23. Wei, Y., Zhu, W., Gong, J., Guo, D., Gu, L., Li, N., & Li, J. (2015). Fecal Microbiota Transplantation Improves the Quality of Life in Patients with Inflammatory Bowel Disease. *Gastroenterology research and practice*, *2015*, 517597. https://doi.org/10.1155/2015/517597

24. Ren, R., Sun, G., Yang, Y., Peng, L., Zhang, X., Wang, S., Dou, Y., Zhang, X., Wang, Z., Bo, X., Liu, Q., Li, W., Fan, N., & Ma, X. (2015). *Zhonghua nei ke za zhi*, *54*(5), 411–415.

25. Rossen, N. G., Fuentes, S., van der Spek, M. J., Tijssen, J. G., Hartman, J. H., Duflou, A., Löwenberg, M., van den Brink, G. R., Mathus-Vliegen, E. M., de Vos, W. M., Zoetendal, E. G., D'Haens, G. R., & Ponsioen, C. Y. (2015). Findings From a Randomized Controlled Trial of Fecal Transplantation for Patients With Ulcerative Colitis. *Gastroenterology*, *149*(1), 110–118.e4. https://doi.org/10.1053/j.gastro.2015.03.045

26. Alka Goyal, MD, Andrew Yeh, MD, Brian R Bush, MD, Brian A Firek, MS, Leah M

Siebold, MD, Matthew Brian Rogers, PhD, Adam D Kufen, RN BS, CCRC, MBA, Michael J Morowitz, MD, FACS, Safety, Clinical Response, and Microbiome Findings Following Fecal Microbiota Transplant in Children With Inflammatory Bowel Disease, *Inflammatory Bowel Diseases*, Volume 24, Issue 2, February 2018, Pages 410–421, https://doi.org/10.1093/ibd/izx035

27. Laszlo, M., Ciobanu, L., Andreica, V., & Pascu, O. (2016). Fecal transplantation indications in ulcerative colitis. Preliminary study. *Clujul medical (1957)*, *89*(2), 224–228. https://doi.org/10.15386/cjmed-613

28. Nishida, A., Imaeda, H., Ohno, M., Inatomi, O., Bamba, S., Sugimoto, M., & Andoh, A. (2017). Efficacy and safety of single fecal microbiota transplantation for Japanese patients with mild to moderately active ulcerative colitis. *Journal of gastroenterology*, *52*(4), 476–482. https://doi.org/10.1007/s00535-016-1271-4

29. Zhang, T., Cui, B., Li, P., He, Z., Long, C., Wei, L., Peng, Z., Ji, G., & Zhang, F. (2016). Short-Term Surveillance of Cytokines and C-Reactive Protein Cannot Predict Efficacy of Fecal Microbiota Transplantation for Ulcerative Colitis. *PloS one*, *11*(6), e0158227. https://doi.org/10.1371/journal.pone.0158227

30. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, Alexander H, Alm EJ, Arumugam M, Asnicar F, Bai Y, Bisanz JE, Bittinger K, Brejnrod A, Brislawn CJ, Brown CT, Callahan BJ, Caraballo-Rodríguez AM, Chase J, Cope EK, Da Silva R, Diener C, Dorrestein PC, Douglas GM, Durall DM, Duvallet C, Edwardson CF, Ernst M, Estaki M, Fouquier J, Gauglitz JM, Gibbons SM, Gibson DL, Gonzalez A, Gorlick K, Guo J, Hillmann B, Holmes S, Holste H, Huttenhower C, Huttley GA, Janssen S, Jarmusch AK, Jiang L, Kaehler BD, Kang KB, Keefe CR, Keim P, Kelley ST, Knights D, Koester I, Kosciolek T, Kreps J, Langille MGI, Lee J, Ley R, Liu YX, Loftfield E, Lozupone C, Maher M, Marotz C, Martin BD, McDonald D, McIver LJ, Melnik AV, Metcalf JL, Morgan SC, Morton JT, Naimey AT, Navas-Molina JA, Nothias LF, Orchanian SB, Pearson T, Peoples SL, Petras D, Preuss ML, Pruesse E, Rasmussen LB, Rivers A, Robeson MS 2nd, Rosenthal P, Segata N, Shaffer M, Shiffer A, Sinha R, Song SJ, Spear JR, Swafford AD, Thompson LR, Torres PJ, Trinh P, Tripathi A, Turnbaugh PJ, Ul-Hasan S, van der Hooft JJJ, Vargas F, Vázquez-Baeza Y, Vogtmann E, von Hippel M, Walters W, Wan Y, Wang M, Warren J, Weber KC, Williamson CHD, Willis AD, Xu ZZ, Zaneveld JR, Zhang Y, Zhu Q, Knight R, Caporaso JG. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nature biotechnology*, *37*(8), 852–857. https://doi.org/10.1038/s41587-019-0209-9

31. Gonzalez A, Navas-Molina JA, Kosciolek T, McDonald D, Vázquez-Baeza Y, Ackermann G, DeReus J, Janssen S, Swafford AD, Orchanian SB, Sanders JG, Shorenstein J, Holste H, Petrus S, Robbins-Pianka A, Brislawn CJ, Wang M, Rideout JR, Bolyen E, Dillon M, Caporaso JG, Dorrestein PC, Knight R. Qiita: rapid, web-enabled microbiome meta-analysis. *Nat Methods* **15,** 796–798 (2018). https://doi.org/10.1038/s41592-018-0141-9

32. Mohsen, A., Park, J., Chen, Y. A., Kawashima, H., & Mizuguchi, K. (2019). Impact of quality trimming on the efficiency of reads joining and diversity analysis of Illumina paired-end reads in the context of QIIME1 and QIIME2 microbiome analysis frameworks. BMC bioinformatics, 20(1), 581. https://doi.org/10.1186/s12859-019-3187-5

33. Amir, A., McDonald, D., Navas-Molina, J. A., Kopylova, E., Morton, J. T., Zech Xu, Z., Kightley, E. P., Thompson, L. R., Hyde, E. R., Gonzalez, A., & Knight, R. (2017). Deblur Rapidly Resolves Single-Nucleotide Community Sequence Patterns. *mSystems*, *2*(2), e00191-16. https://doi.org/10.1128/mSystems.00191-16

34. R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.URL https://www.R-project.org/.

35. Wickham H (2016). *ggplot2: Elegant Graphics for Data Analy*sis. Springer-Verlag New York. ISBN 978-3-319-24277-4, https://ggplot2.tidyverse.org.

36. Bukin, Y. S., Galachyants, Y. P., Morozov, I. V., Bukin, S. V., Zakharenko, A. S., & Zemskaya, T. I. (2019). The effect of 16S rRNA region choice on bacterial community metabarcoding results. *Scientific data*, *6*, 190007. https://doi.org/10.1038/sdata.2019.7

37. Martino, C., Morton, J. T., Marotz, C. A., Thompson, L. R., Tripathi, A., Knight, R., & Zengler, K. (2019). A Novel Sparse Compositional Technique Reveals Microbial Perturbations. mSystems, 4(1), e00016-19. https://doi.org/10.1128/mSystems.00016-19

38. Nusbaum, D. J., Sun, F., Ren, J., Zhu, Z., Ramsy, N., Pervolarakis, N., Kunde, S., England, W., Gao, B., Fiehn, O., Michail, S., & Whiteson, K. (2018). Gut microbial and metabolomic profiles after fecal microbiota transplantation in pediatric ulcerative colitis patients. *FEMS microbiology ecology*, *94*(9), fiy133. https://doi.org/10.1093/femsec/fiy133

39. Turner, D., Otley, A. R., Mack, D., Hyams, J., de Bruijne, J., Uusoue, K., Walters, T. D., Zachos, M., Mamula, P., Beaton, D. E., Steinhart, A. H., & Griffiths, A. M. (2007). Development, validation, and evaluation of a pediatric ulcerative colitis activity index: a prospective multicenter study. *Gastroenterology*, *133*(2), 423–432. https://doi.org/10.1053/j.gastro.2007.05.029

40. Lozupone, C. A., Hamady, M., Kelley, S. T., & Knight, R. (2007). Quantitative and qualitative beta diversity measures lead to different insights into factors that structure microbial communities. *Applied and environmental microbiology*, *73*(5), 1576–1585. https://doi.org/10.1128/AEM.01996-06

41. Vital, M., Karch, A., & Pieper, D. H. (2017). Colonic Butyrate-Producing Communities in

Humans: an Overview Using Omics Data. *mSystems*, *2*(6), e00130-17. https://doi.org/10.1128/mSystems.00130-17

42. Fuentes, S., Rossen, N. G., van der Spek, M. J., Hartman, J. H., Huuskonen, L., Korpela, K., Salojärvi, J., Aalvink, S., de Vos, W. M., D'Haens, G. R., Zoetendal, E. G., & Ponsioen, C. Y. (2017). Microbial shifts and signatures of long-term remission in ulcerative colitis after faecal microbiota transplantation. *The ISME journal*, *11*(8), 1877–1889. https://doi.org/10.1038/ismej.2017.44

43. Janssen, S., McDonald, D., Gonzalez, A., Navas-Molina, J. A., Jiang, L., Xu, Z. Z., Winker, K., Kado, D. M., Orwoll, E., Manary, M., Mirarab, S., & Knight, R. (2018). Phylogenetic Placement of Exact Amplicon Sequences Improves Associations with Clinical Information. *mSystems*, *3*(3), e00021-18. https://doi.org/10.1128/mSystems.00021-18

44. Mandal, S., Van Treuren, W., White, R. A., Eggesbø, M., Knight, R., & Peddada, S. D. (2015). Analysis of composition of microbiomes: a novel method for studying microbial composition. Microbial ecology in health and disease, 26, 27663. https://doi.org/10.3402/mehd.v26.27663

45. Vital, M., Karch, A., & Pieper, D. H. (2017). Colonic Butyrate-Producing Communities in Humans: an Overview Using Omics Data. *mSystems*, *2*(6), e00130-17. https://doi.org/10.1128/mSystems.00130-17