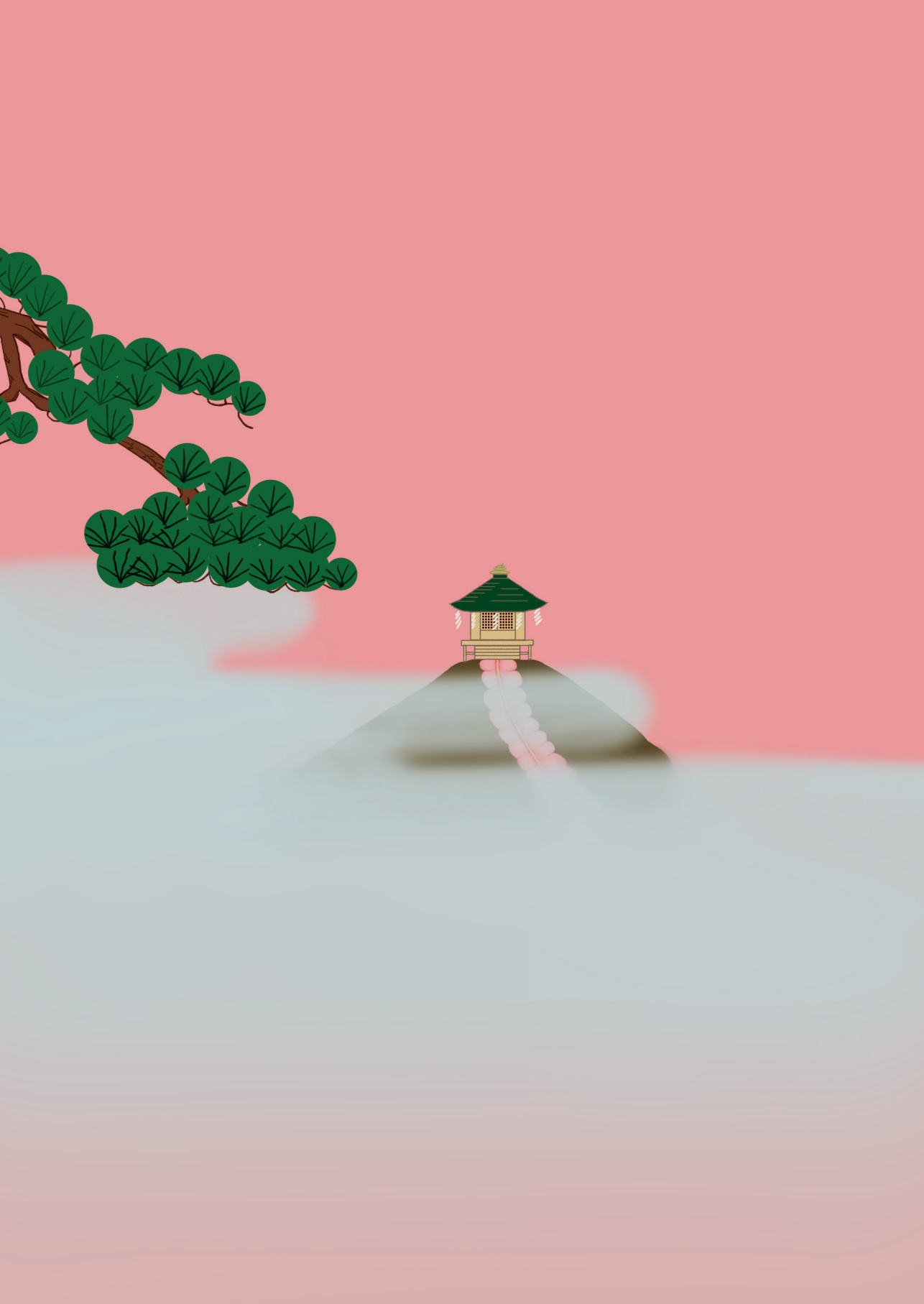
Guardians of the gut: harnessing bioinformatics to study the gut microbiome and faecal microbiota transplantation in intestinal disorders

Nooij, S.



Chapter 5

Dynamics of Gut Microbiota after Fecal Microbiota Transplantation in Ulcerative Colitis: Success Linked to Control of *Prevotellaceae*

Susanne Pinto[1], Dominika Šajbenová[1], Elisa Benincà[2], Sam Nooij[3], Elisabeth M. Terveer[3], Josbert J. Keller[4], Andrea E. van der Meulen – de Jong[4], Johannes A. Bogaards[5]*, and Ewout Steyerberg[1]*

[1] Leiden University Medical Center, Biomedical Data Sciences, Leiden, The Netherlands
[2] National Institute for Public Health and the Environment (RIVM), Centre for Infectious Disease Control, Bilthoven, The Netherlands
[3] Leiden University Medical Center, Center for Infectious Diseases (LU-CID), Leiden, The Netherlands
[4] Leiden University Medical Center, Gastroenterology and Hepatology, Leiden, The Netherlands
[5] Amsterdam UMC location Vrije Universiteit Amsterdam, Epidemiology and Data Science, Amsterdam, Netherlands; Institute for Infection and Immunity (AII), Amsterdam, The Netherlands
* These authors contributed equally

Journal of Crohn's and Colitis, *jjae137, DOI: 10.1093/ecco-jcc/jjae137, 3 September 2024*

Abstract

Background

Fecal microbiota transplantation (FMT) is an experimental treatment for ulcerative colitis (UC). We aimed to study microbial families associated with FMT treatment success.

Methods

We analyzed stools from 24 UC patients treated with four FMTs weekly after randomization for pretreatment during three weeks with budesonide (n = 12) or placebo (n = 12). Stool samples were collected nine times pre-, during, and post-FMT. Clinical and endoscopic response was assessed 14 weeks after initiation of the study using the full Mayo score. Early withdrawal due to worsening of UC symptoms was classified as non-response.

Results

Nine patients (38%) reached remission at week 14, and 15 patients had a partial response or non-response at or before week 14. With a Dirichlet Multinomial Mixture model we identified five distinct clusters based on the microbiota composition of 180 longitudinally collected patient samples and 27 donor samples. A Prevotellaceae-dominant cluster was associated with poor response to FMT treatment. Conversely, the families Ruminococcaceae and Lachnospiraceae were associated with a successful clinical response. These associations were already visible at the start of the treatment for a subgroup of patients and were retained in repeated measures analyses of family-specific abundance over time. Responders were also characterized by a significantly lower Simpson dominance compared to non-responders.

Conclusions

The success of FMT treatment of UC patients appears to be associated with specific gut microbiota families, such as control of Prevotellaceae. Monitoring the dynamics of these microbial families could potentially be used to inform treatment success early during FMT.

Clinical trial registration number

This research project was reviewed and approved by the Medical Ethical Committee in the LUMC, with reference number NL 65976.098.18. The study was registered in the Netherlands Trial Register, with reference number NL9858.

Introduction

Ulcerative colitis (UC) is a chronic inflammatory disorder affecting the colon. Symptoms experienced by patients during disease exacerbation include bloody stools, diarrhea, and abdominal pain.[1] The etiology of UC is multifactorial, involving a complex interplay between the host immune system, gut microbiota, and genetic and environmental factors.[2-4] UC patients exhibit a reduced microbial diversity and alterations in the composition of their gut microbiota compared to healthy individuals.[4, 5] Notably, a decrease in Bacillota (formally Firmicutes), especially Clostridia (such as *Clostridium, Roseburia,* and *Faecalibacterium)*, and Verrucomicrobia, along with an overgrowth of species from the Enterobacteriaceae family (such as *Escherichia coli* or *Klebsiella spp.*), have been observed.[6-9] Studies investigating associations with common Bacteroidota in the human gut, such as the Bacteroidaceae and Prevotellaceae families, have yielded conflicting results.[6, 7, 9-13]

The current approach to treat UC focuses on attenuating the hyperactive immune response using pharmaceutical drugs, such as local immune suppression with 5-aminosalicylates (5-ASA), or systemic immune suppression with prednisolone, thiopurines, biologics, or small molecules.[6] However, many patients do not derive lasting benefits from these interventions and may even experience severe side effects.[14] Fecal microbiota transplantation (FMT) has emerged as a promising alternative treatment for microbiota-associated disorders, particularly in the treatment of recurrent *Clostridioides difficile* infection.[15-17] FMT involves transferring fecal matter from a healthy donor to a patient with the aim of modulating the microbiota composition towards a more favorable state. The effectiveness of FMT in UC is limited, with a lower response rate observed as compared to FMT treatment of *Clostridioides difficile* infection.[18, 19] A recent meta-analysis comprising six randomized controlled trials (RCT) reported a short-term clinical response in only half of the patients with active UC following FMT administration.[19] The specific host factors influencing successful FMT response in UC are still unclear, and the donor characteristics that influence patient response to clinical success after FMT remain uncertain.[20, 21] In this study, we explore differences in gut microbiota dynamics between patients with clinical remission and non-responders following FMT treatment. Recently, our group performed an RCT assessing the effects of pretreatment with budesonide on FMT outcome and engraftment of donor bacteria.[22] Budesonide reduces inflammation, and therefor might promote greater engraftment of the healthy donor microbiota, as inflammation can disrupt microbiota homeostasis. Interestingly, the primary analysis showed no association between pretreatment or overall engraftment with clinical response, but there was a significant donor-dependent effect on engraftment.[22] Although the study was not powered to detect differences with regard to clinical endpoints, here we aimed to further identify longitudinal associations between the microbiota compositions and clinical response to FMT treatment.

Materials and methods

The study population

For the current study we used the stool samples collected of the 24 UC patients included in our previously described FMT trial (Supplementary file 1).[22] Patients were randomly assigned to be pretreated daily for three weeks with oral budesonide (9 mg) or with a placebo, and for treatment with FMT suspensions from donor D07 or D08 (block randomization). Inclusion criteria were: at least 18 years old and a confirmed diagnosis of mild to moderate UC (i.e., a full Mayo score (including partial Mayo score and endoscopic sub score of 1 or 2) ranging from 4 to 9). Exclusion criteria were among others proctitis, antibiotic use or surgery within the last 6 weeks, or received other treatments within 12 weeks prior to study entry.

The following clinical and demographic information was collected for each patient in the study (Supplementary file 1): sex, age at baseline (years), donor ID (D07 or D08), pretreatment (placebo or budesonide) and clinical outcome at week 14. Patients that did not complete the study because of progressive symptoms/disease were considered treatment failures and classified as non-responders. At week 14, nine patients were in clinical and endoscopic remission (hereafter called responders), 14 patients were non-responder, and one patient was a partial responder. We included this last patient in the non-responder group.

Clinical and laboratory procedures

Patients received a weekly FMT for four times (at end of week 3, 4, 5, and 6) from the Netherlands Donor Feces Bank (NDFB), either from donor D07 or donor D08 following standard protocols for donor screening, sample collection, sample preparation, sample storage and FMT infusion.[23] The samples used for the different FMTs came from different donations. Before every FMT the patients fasted for at least six hours. A bowel lavage with two liters of macrogol solution (Kleanprep) was performed one day before the first FMT to cleanse the intestine. No changes in diet or medication were reported during the study.

Stool samples of the patients were collected once at baseline, once after the pretreatment phase (but still before the FMT treatment), one week after every FMT (four times; designated Post-1 to Post-4) and three times as a follow-up, at 8, 10 and 14 weeks after randomization.[22] In total we collected 81 stool samples in the responder group (n = 9) and 99 stool samples in the non-responder group (n = 15). Stool samples of donors D07 and D08 were collected regularly and a total of 27 samples (n = 13 for donor D07 and n = 14 for donor D08) were used for analysis.

Microbiota composition

DNA was extracted from the collected stool samples (both from the donors and recipients) and sequenced by Diversigen (new Brighton, USA) with the Illumina NovaSeq platform (100 bp single-end reads to a median depth of 2.9m reads). Raw reads mapping to the human genome were removed using bowtie2 (version 2.4.2)[24] and the GRCh37 reference genome and reads were quality-trimmed using fastq (version 0.20.1)[25], both of which are part of an in-house workflow (https://git.lumc.nl/snooij/metagenomics-preprocessing). The mOTUs3 workflow (version 3.0.1) was used to generate taxonomic profiles.[26, 27] Unassigned, human-derived, Archaeal, and low-quality reads were removed from the data, which resulted in 93 different families (i.e., 1552 unique mOTUs). The database of mOTUs3 includes taxa based on metagenomic bins that have not formally been classified. These are listed as '*incertae sedis*' or '*i.s.*'. Due to the sparsity of the data and the relatively small number of patients, the analyses performed at taxonomic genus rank lacked the statistical power needed to provide robust and reliable results. For this reason, the data were aggregated to family level prior to the statistical analysis. All analyses were performed using R software (R version 4.2.2).

Differences in relative abundances of specific microbial families amongst responders and non-responders were tested for statistical significance in repeated measurements analyses, as described below in the 'longitudinal models' section. The average relative abundances of the same bacterial families were calculated for each donor from multiple samples, considering the donor samples were not collected at the same timepoints as the patient samples. Differences amongst donor D07 and donor D08 were tested with Pearson's chi-squared test and the p-values were corrected for multiple hypothesis testing with the Bonferroni method.

Principal component analysis

We performed Principal Component Analysis (PCA) on the Aitchison distances calculated between each pair of patient microbiota profiles. The Aitchison distance is often used in microbiota data because it takes into account the compositionality of the data.[28, 29] The Aitchison distance involved each patient sample undergoing the centered-log-ratio (clr) transformation and then obtaining the Euclidean distance between each pair of samples as implemented in the microViz R package.[30]

Dirichlet Multinomial Mixture models

We used the Dirichlet Multinomial Mixtures (DMM) clustering algorithm to identify distinct clusters of samples based on their microbial abundance profiles. DMM assumes that the microbial abundances in each sample follow one of a given number of multinomial distributions, the number of which is determined by the assumed number of clusters in the data. We used the dmn function from the DirichletMultinomial R package to cluster patient and donor samples.[31] The parameters of the different clusters are estimated by maximizing the likelihood of the observed data given the

assumed model, with a Dirichlet prior for relative abundances of the bacterial families to facilitate parameter estimation and prevent overfitting. The prior consisted of a mixture of Dirichlets with $k = 1, ..., K$ to represent the K clusters, with hyperparameters denoting cluster-specific weights and relative abundances. Next, the bacterial families in each cluster were ranked based on the posterior difference between the cluster in a multi-cluster solution versus a one-cluster model. A more detailed description of DMM models is presented elsewhere.[32] Considering that the DMM clustering algorithm uses stochastic likelihood optimization with random initial parameter values, we performed the clustering algorithm 1000 times and chose the model with the lowest Laplace value, indicating a better parsimonious fit of the model to the data.

Data were clustered according to a combination of patient and donor samples. As a sensitivity analysis, we also applied the algorithm in the following situations: patient samples only; patient samples excluding a patient who was placed in a distinct cluster compared to all other patients (patient 102); patient samples excluding patients who both had only two samples available (patients 109 and 117).

Longitudinal models of bacterial relative abundances

Mixed models were used to model the changes over time in relative abundance for each of the 15 most abundant bacterial families in the patient samples. Regarding the distribution of relative abundance, many families had a high proportion of zeros, resulting in right-skewed distributions. All abundances, except for Ruminococcaceae, were therefore transformed with an arcsine square root transformation to approximate normally distributed data. We modelled the relative abundances of the 15 selected bacterial families separately in 15 different longitudinal models with a linear mixed effect model (LMM), possibly augmented with a zero-inflation component (ZILMM). The lme4 package was used for constructing LMMs and the glmmTMB package was used for constructing ZILMMs.[33, 34] To account for the correlation of repeated observations within each patient, both random slopes and random intercepts were considered as potential models for each bacterial family. Note that the dataset was too small for specification of predictors in the zero-inflation component. To incorporate possible non-linearity in relative abundance trajectories over time into the model, a natural cubic spline (with the ns function from the splines package in R) with node at week 8 (the beginning of the follow-up phase) was considered for all models. Model preference was based on the lowest AIC and model diagnostics, judged by QQ-plot and a plot of residuals against predicted values. All choices per family are given in Supplementary file 2.

The longitudinal models further included the variables clinical outcome (non-responder vs. responder), time (possibly with a cubic spline), and an interaction with time and clinical outcome (non-responder vs. responder). The interaction determined whether there was a divergence in relative abundance of a particular family between non-responders and responders, with statistical significance assessed by Wald tests.[35]

The inclusion of the patient specific variables donor (donor D07 vs. D08), pretreatment (budesonide vs. placebo), age and sex (female vs. male) in the model was dependent upon testing their role as confounders or contribution to the model fit. This assessment involved examining whether their inclusion led to a greater than 15% change in the primary coefficients (notable influence on the model's outcome) or a significant Likelihood Ratio Test (contribution of the variable to the model); with flexibility allowed for a variable to meet one of these criteria during the evaluation process.

Simpson dominance

Simpson dominance was used to summarize microbiota diversity of each sample. We calculated this measure (the sum of the squared relative abundances) with the dominance function from the microbiome package.[36] The Simpson dominance estimates the probability that two random entities taken from a sample represent the same bacterial family within a patient's microbiota. Hence, a higher Simpson dominance means a higher concentration of species from the same family in the sample, which corresponds with a less diverse microbiota. To account for the correlation of repeated observations within each patient, the Simpson dominance was modelled with a random-intercepts LMM (with the lme function from the nlme package).[37] A log transformation was applied to the Simpson dominance measure to correct for non-normality. The regression parameter of primary interest was the relation between Simpson dominance and clinical response, either as a main effect (denoting baseline differences in diversity) or in interaction with time (denoting divergence in diversity between responders and non-responders over time). Additional parameters included the effects sex and time. Similar to the longitudinal LMM of bacterial families, time was modelled as a continuous variable with a natural cubic spline (node at week 8). The effects of pretreatment, donor, and age were negligible and therefore not included in the model. Wald tests were performed to test for statistical significance of the clinical response variables jointly in the model.

Results

Microbiota community composition of donors, responders, and non-responders

The fecal microbiota composition between the two donors was distinctly different (-ure 1 and Supplementary file 3). Donor D07 had a significantly higher relative abundance of the families Clostridiaceae, Clostridiales fam. *i.s.* (i.e., an unclassified family within the order Clostridiales), Ruminococcaceae, and Veillonellaceae compared to donor D08, while donor D08 had a significantly higher relative abundance of Bacillota fam. *i.s.* and Lachnospiraceae (Figure 1, Supplementary file 3, and Supplementary file 4).

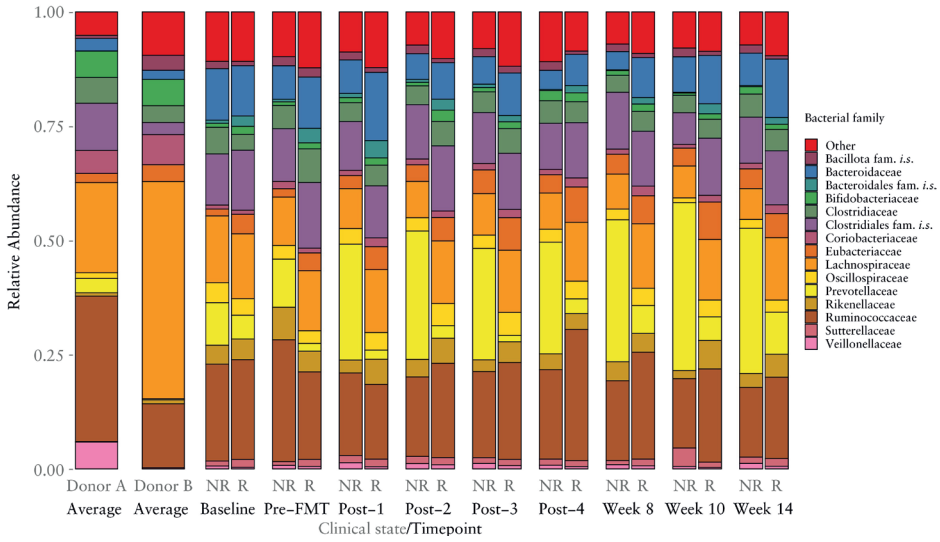


Figure 1. Average microbiota composition of the 15 most abundant bacterial families. Abundances are followed over time for the two donors, non-responders (NR) and responders (R). Here, the 'other' category includes all remaining bacterial families.

Overall, the most abundant bacterial family in the patients was Ruminococcaceae. However, from the second timepoint onwards, the relative abundance of Prevotellaceae continued to increase in the microbiota of the non-responders. Prevotellaceae overtook Ruminococcaceae as the most abundant family for non-responders at Post-1 and remained the most abundant for the remaining timepoints (Figure 1, Supplementary file 3, and Supplementary file 5). Compared to the non-responders, Lachnospiraceae and Oscillospiraceae seemed to become more abundant in the responder group over time (Figure 1, Supplementary file 3, and Supplementary file 5).

PCA results for donors and patients

The first two components in PCA analysis of patient and donor samples, based on the Aitchison distance, explained 24% of the total variation in the data (Figure 2). The samples of donor D08 clustered away from the patients' samples, driven by a difference in the relative abundance of Lachnospiraceae (Figure 2). Patients treated with an FMT from donor D08 showed a higher responder rate than those from donor D07 (Supplementary file 1). The difference in distance between non-responders and responders seemed to be explained by the relative abundance of Prevotellaceae (Figure 2). This applied particularly to the patients who received an FMT from donor D08 (Supplementary file 6). Only a few of the patient samples seemed to traverse considerable Aitchinson distance over time. Notably, the patients whose microbiota became more donor-like over time were more often non-responders (e.g. patients 110 and 111) (Supplementary file 7).

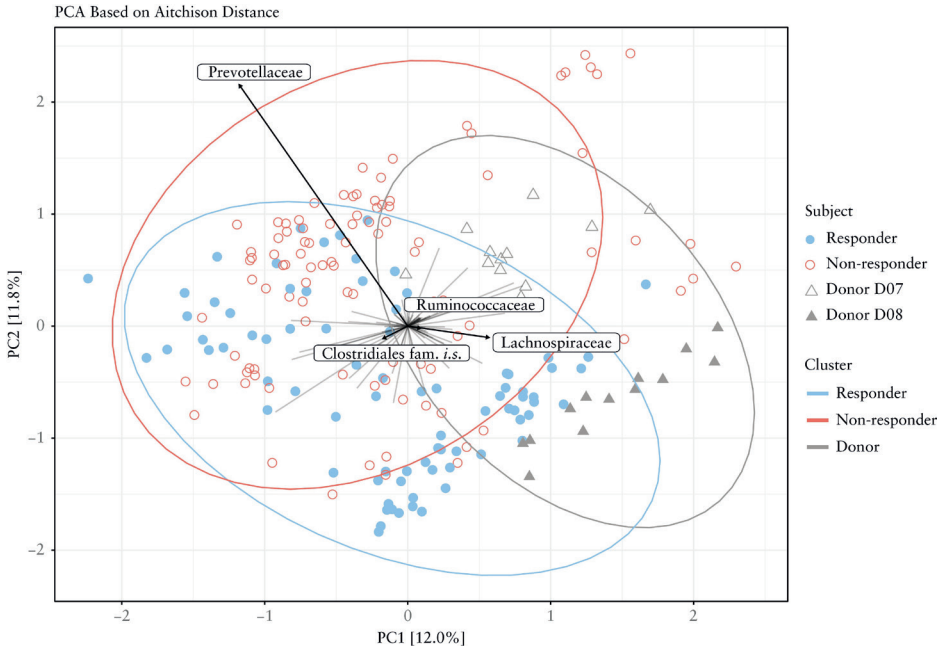


Figure 2. PCA plot with Aitchison distances in microbiota profiles for the distance between sample types. Responders are given in blue, non-responders in green and donors in grey. The PCA plots include data ellipses around the different groups and loading vectors of families to obtain an initial visualization about the extent of separation between non-responder, responder, and donor samples. The different symbols, closed circles, open circles, open triangles, and closed triangles, indicate responder, non-responder patients, donor D07, and donor D08, respectively.

Sample clustering with Dirichlet Multinomial Mixture models

Over 1000 iterations, a five-clusters model was selected as the best-fitting model (i.e., having the lowest Laplace value). Figure 3 and Supplementary file 8 show that Ruminococcaceae was present in all clusters whereas Lachnospiraceae, Bacteroidaceae and Clostridiales fam. *i.s.* were present in four of the five clusters. The relative abundances of those families in each cluster differed: clusters 1 and 4 were dominated by Ruminococcaceae and Lachnospiraceae, whereas clusters 2 and 5 were dominated by Ruminococcaceae and Clostridiales fam. *i.s.*. Prevotellaceae was the only family almost defining an entire cluster (cluster 3). Cluster 1 appeared to be associated with a successful clinical response, while cluster 3 appeared to be associated with non-response (Figure 4). For the patient samples, 56% of responder samples were classified into cluster 1, and 38% into cluster 2, whereas 42% of non-responder samples were classified into cluster 3 (Figure 4B). All donor samples, except for one, were assigned to cluster 4 (Supplementary file 9). Five non-responder patient samples were also assigned to cluster 4 (Figure 4A). This donor-dominated cluster disappeared in sensitivity analysis on patient samples only (Supplementary file 10A), resulting in the re-assignment

of the corresponding patient samples to cluster 2. Patient 102 was responsible for the existence of a separate cluster (cluster 5), with all its measurements belonging to that cluster. Removal of this patient in sensitivity analysis resulted in the deletion of that cluster, with re-assignment of the other corresponding samples to cluster 2 (Supplementary file 10B). Removal of patients with only two measurements (107 and 119) had minor impact on the results (Supplementary file 10C).

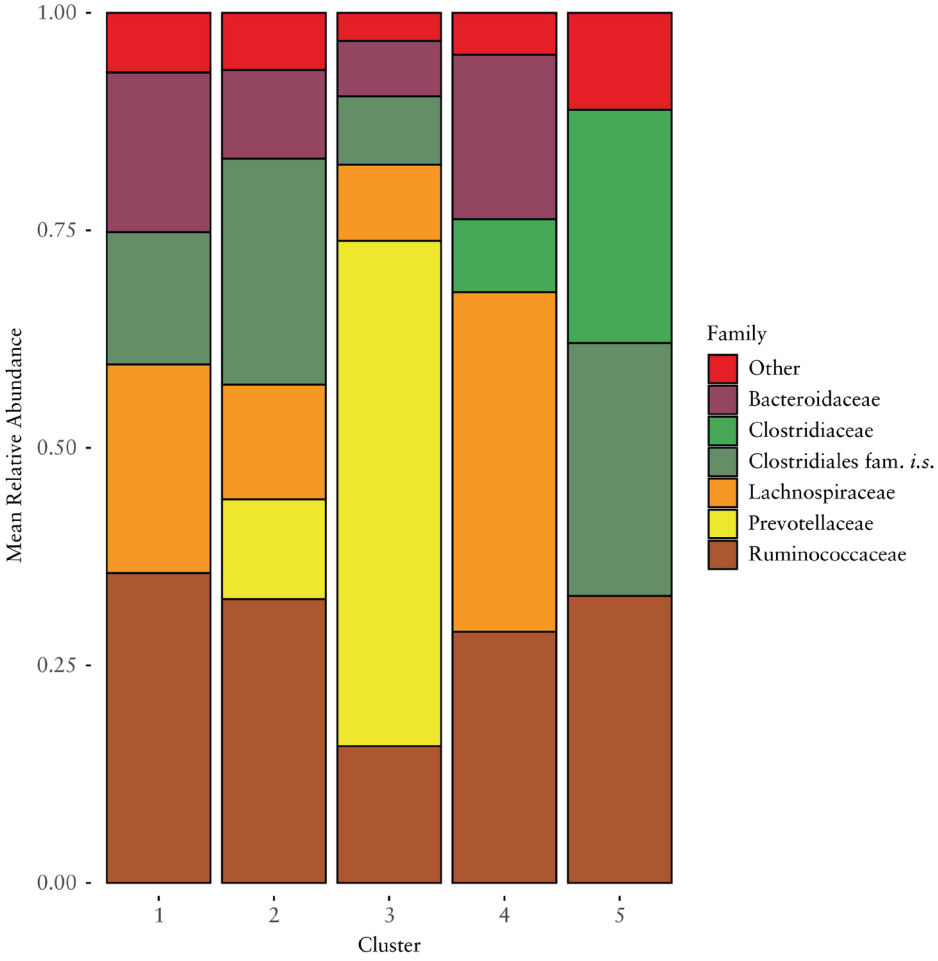


Figure 3. Mean relative abundance of bacterial families in the five clusters. Clusters are detected by the Dirichlet Multinomial Mixture model.

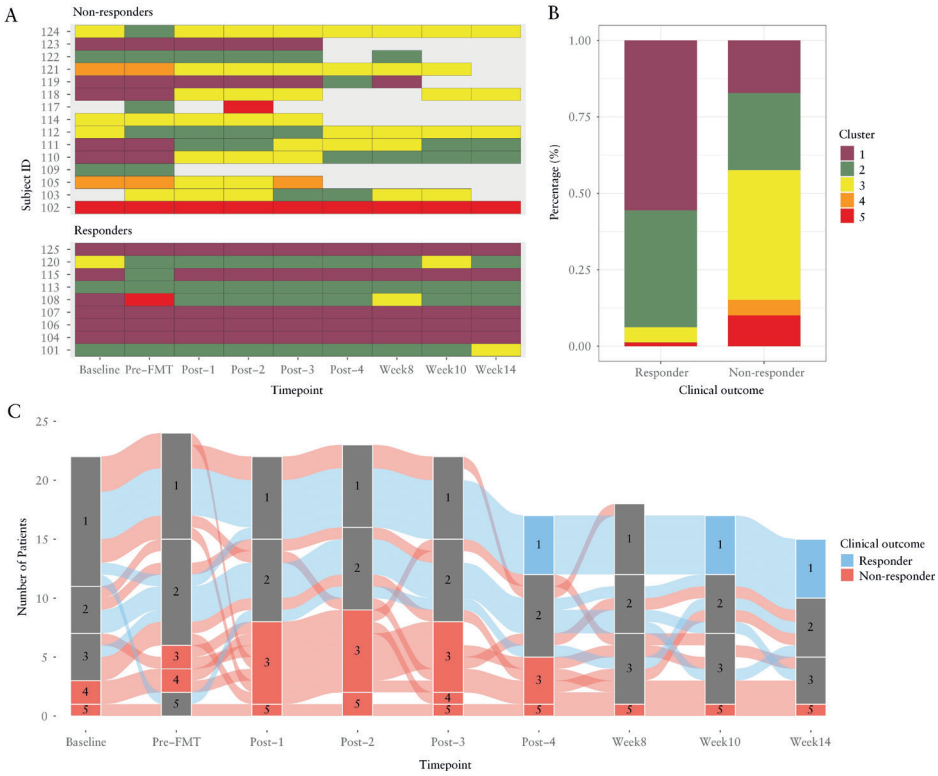


Figure 4. Clustering of donor and patient samples. A) Cluster membership over time per patient for the non-responders (upper facet) and responders (lower facet). Lack of colored bar indicates that no stool sample was collected at that timepoint. B) Percentage of each cluster for non-responders and responders. C) Alluvial plot of patients distributed over the different clusters over time. This plot displays the distribution of clusters per timepoint and whether each cluster is comprised of only one clinical group (e.g., only non-responders) for every timepoint. A grey box means that the cluster at that timepoint contains both samples from responder and non-responder patients, a red box only contains non-responder samples, and a blue box only contains responder samples.

Out of 24 patients, only nine patients (38%) remained in the same cluster for all of their provided samples (Figure 4A). An alluvial plot of patient samples showed the substantial changes in sample membership and cluster size throughout the clinical trial (Figure 4C). There was a mixture of non-responder and responder samples in cluster 1 at the beginning, with most samples at baseline being classified into cluster 1. There was then a shift towards more responder samples in cluster 1 from Pre-FMT onwards. Samples in cluster 1 were exclusively composed of responder samples at timepoints Post-4, Week 10, and Week 14. Cluster 3 was fully composed of non-responder samples after pretreatment and after every FMT treatment (Figure 4C).

Coloring samples by their cluster membership in the Aitchison distances PCA plot showed separation between clusters 1, 2, and 3, with cluster 2 being the intermediate cluster (Supplementary file 9). The Prevotellaceae vector was pointed in the direction of cluster 3, corresponding to a potential association between this cluster and non-response (Supplementary file 9), possibly driven by the donor (Figure 2 and Supplementary file 6). There appeared to be some separation between donor samples, a majority of which were in cluster 4, and patient samples. Donor D08 samples were close to cluster 1 samples. Meanwhile, donor D07 samples were positioned near cluster 2 samples (Supplementary file 9). Finally, samples from cluster 5 were grouped together tightly, likely as a result of belonging to the same patient.

Mixed models of bacterial families

Responders and non-responders showed significantly different trajectories in relative abundance over time for the families Prevotellaceae, Lachnospiraceae, Ruminococcaceae, Oscillospiraceae, and Sutterellaceae (Figure 5, Supplementary file 2, and Supplementary file 5). Prevotellaceae showed the greatest difference in trajectory between non-responders and responders over time. Note that the preferred model for Prevotellaceae had a straightforward linear trajectory and used the original time variable instead of splines. The family Prevotellaceae consisted of four named genera, of which *Prevotella* (especially *Prevotella copri*) was the most abundant (Supplementary file 11).

There were four families with a significant donor effect: Veillonellaceae, Rikenellaceae, Sutterellaceae, and Bifidobacteriaceae (Supplementary file 2). Notably, removal of the donor variable from the model for Sutterellaceae diminished the significance of the main effect related to clinical response. This observation underscores the role of the donor variable in influencing the association between Sutterellaceae and clinical response. Rikenellaceae and Bacillota fam. *i.s.* had a significant sex effect, Veillonellaceae had a significant pretreatment effect (Supplementary file 2). None of these other significant covariates altered the statistical significance of clinical response. This observation suggests that the estimated associations were not confounded by these covariates.

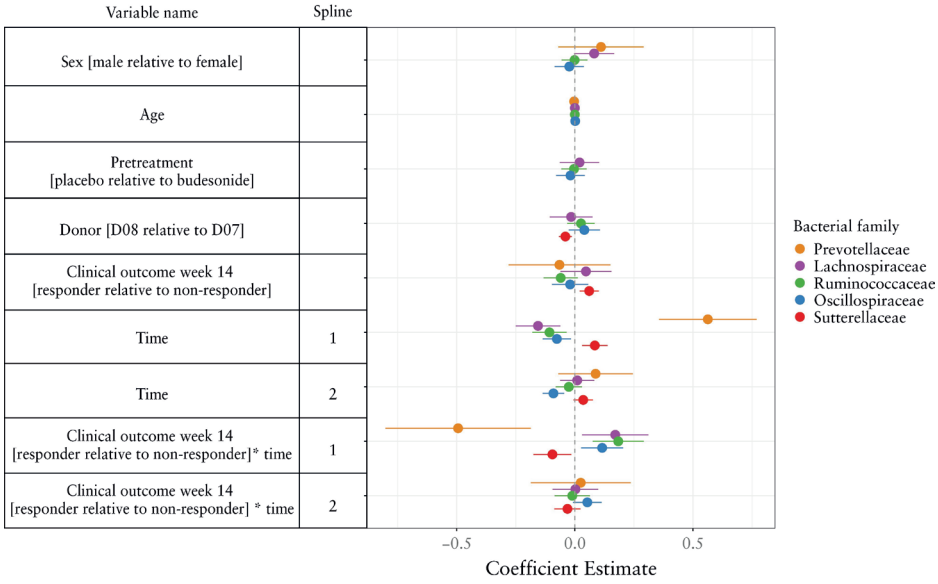


Figure 5. Results of the mixed models. Only the families among the 15 most abundant families (Prevotellaceae, Lachnospiraceae, Ruminococcaceae and Oscillospiraceae) for whom we found a significant effect in relation to clinical response with the Wald test are shown. The point estimates, 95% confidence intervals and a reference line at zero are shown. When the horizontal lines do not cross the vertical reference line, this means that the coefficients are significantly different from 0. P-values are given in Supplementary file 2.

Simpson dominance

The steadily increasing relative abundance of Prevotellaceae in non-responders found before was reflected in the Simpson dominance. Simpson dominance was higher for non-responders compared to responders, especially throughout the follow-up period (Figure 6). There was a significant difference between the Simpson dominance in responder and non-responder patients (Wald test: $p = 0.004$). Our study was too small to determine whether this difference already existed at baseline or developed over time (Supplementary file 12). The LMM random-intercept model suggested that there was also a significant sex effect (Supplementary file 12). However, sex did not alter the statistical significance of clinical response. This observation suggests that the estimated associations were not confounded by the sex of the patients.

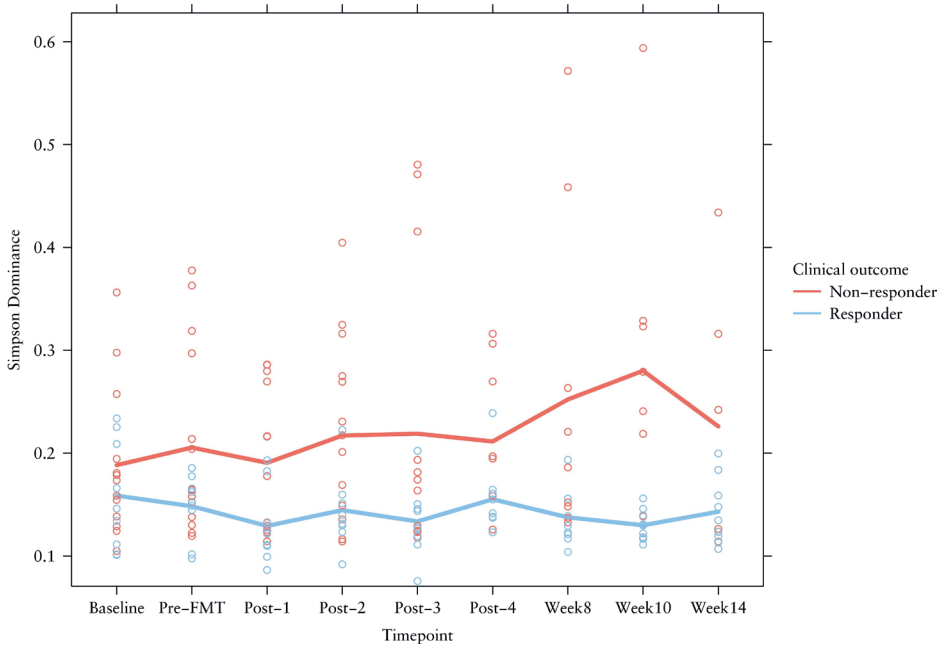


Figure 6. Change in Simpson dominance calculated for non-responders and responders. The points indicate the individual measurements of the patients. The lines are the mean Simpson dominance per group (non-responders in red and responders in blue).

Discussion

Inflammatory bowel diseases (IBD) such as ulcerative colitis (UC) have been linked to alterations in both the composition and metagenomic function of the gut microbiota.[4, 5] In this study, we employed a wide range of analytical techniques to investigate potential associations between microbiota and clinical outcomes following FMT in UC patients. A subgroup of the cohort (9 of 24 patients) reached a successful combined clinical and endoscopic remission after the FMT treatment, and our results suggest that this response may be related to certain gut microbiota families. Specifically, longitudinal models and cluster analysis of repeatedly measured compositional data indicated that the success of FMT treatment of UC patients appears to be associated with control of Prevotellaceae. Conversely, our analyses also highlighted a potentially beneficial role of Lachnospiraceae and Ruminococcaceae in FMT treatment response. Furthermore, we identified several other bacterial families, including Oscillospiraceae, Enterobacteriaceae, and Sutterellaceae, that exhibited associations with clinical remission. The clustering results indicated that differences in the gut microbiota of responders versus non-responders might already be apparent early during the treatment. If this result can be confirmed by larger studies, clinical success may be predicted from early microbiota analysis after the first FMT treatment and mitigating

actions, for example stopping, personalizing, or changing the treatment, might be envisioned.

Donor-related microbiota characteristics may potentially impact the clinical efficacy of FMT.[38] Intriguingly, we observed marked differences between the donors' and the patients' microbiota. Amongst patients who responded well to FMT, gut microbiota composition did not transition fully to resemble that of the donors at the end of follow-up. This contrasts with earlier studies that suggested that a donor-like microbiota is preferred after FMT treatment[20, 21, 38, 39], and suggest that some complementarity in microbiota compositions between donors and recipients is required for successful clinical response.[12, 21] However, additional research is needed to determine whether attaining a donor-like microbiota by the end of FMT affects the ability of patients to achieve remission. The samples of donor D08 clustered closer to cluster 1 (associated with a successful clinical response), and the samples of donor D07 were closer to cluster 3 (indicating non-response). Note that, an FMT from donor D08 resulted in relatively more treatment success in the patients than donor D07. Also, donor D08 seemed to have a more diverse microbiota than donor D07, although not statistically significant. Donor diversity has been associated with a higher clinical response before.[40] In addition, higher post-FMT diversity has been associated with remission, suggesting that the variety of introduced organisms may promote recovery.[20] It was already noted in [22] that donor D08 was the more successful donor; however, intriguingly, this was the donor with the least engraftment. This observation suggests that the persistent transfer of microbes may not be the prime reason for clinical success. Possibly, the transient exposure to an external microbial community might still induce a beneficial change in the recipient's gut environment. It is also possible that patients who received FMT from donor D08 had a more favorable starting state, while those who received FMT from donor D07 required stronger microbiota changes to move to a more favorable state. Further investigations are warranted to unravel the intricate dynamics underlying the observed outcomes.

This study provides novel evidence for a potential association between control of Prevotellaceae at moderate abundance and favorable clinical outcomes following FMT in UC patients. Additionally, the Simpson dominance measure suggests that Prevotellaceae constituted a sizable proportion of the microbiota in non-responsive FMT patients throughout the course of the clinical trial. Screening the patients (and donors) for Prevotellaceae might improve the response rate. However, previous study suggested that higher levels of *Prevotella* (a genus level within Prevotellaceae) may confer health benefits in UC patients after treatment. For instance, studies on UC patients who underwent drug and surgical treatments, excluding FMT, demonstrated that responders had higher baseline levels of *Prevotella* compared to non-responders.[11] Notably, a previous FMT trial on IBD patients did not report any detrimental effects of increased *Prevotella* abundance, despite observing a substantial increase in this

bacterium in their patients after FMT treatment.[12] They classified *Prevotella* as a colonizing bacterium, as its abundance in patients reached a level comparable to that in the donors. Of note, in our study, responders also maintained levels of Prevotellaceae comparable to donors, but in non-responders there was a clear overgrowth. The conflicting role of *Prevotella* in human health has been attributed to the high diversity within the *Prevotella* genus. While the majority of *Prevotella* species are commonly found in healthy individuals, certain strains may be implicated in disease pathogenesis.[41, 42] For instance, *Prevotella intestinalis* has been shown to induce intestinal inflammation upon colonization in mice.[13] *Prevotella melaninogenica* and *Prevotella oralis* have been characterized as tipping elements.[43] This means that *Prevotella* stands out as a bimodal group, with either a high or low abundance state, and can be a pivotal driver in the context of microbial ecosystem stability. This finding was reiterated in a recent investigation into the involvement of gut microbiota families with Crohn's disease activity, where we found that associations with Prevotellaceae were amongst the most heterogeneous across individual patients.[44]

In contrast to Prevotellaceae, other bacterial families have shown associations with positive clinical outcomes. Specifically, the families Lachnospiraceae, Ruminococcaceae, and Oscillospiraceae have also been found to increase following FMT in patients with UC in other studies.[45] Lachnospiraceae and Ruminococcaceae may play a role in modulating the immune response and inflammatory pathways in the colon.[45] Interestingly, contrary to previous literature, the expected increase in Clostridiaceae among responders was not observed in the present study. This discrepancy in Clostridiaceae abundance may be attributed to variations in FMT protocols employed across different clinical trials or the low number of patients in this study.[46] Additionally, in contrast to the present study, previous research has reported increased abundance of Enterobacteriaceae in UC patients who did not respond to drug and surgical interventions, with higher levels being associated with mucosal inflammation.[11] Discrepancies in Enterobacteriaceae abundance may stem from differences in the types of UC treatments employed, for example when FMT was not involved as a treatment modality.[11] In the context of FMT, a study involving IBD patients who underwent FMT revealed the presence of a dysbiotic *Bacteroides* cluster, as well as an Enterobacteriaceae cluster. Donors were subjected to cluster analysis and categorized into *Prevotella* or *Bacteroides* clusters. Interestingly, the clinical outcome of FMT varied depending on the cluster of both the patients and their respective donors.[12]

The longitudinal study design of our trial, with protocolized data collection across all stages of FMT, enabled a uniquely fine-grained view of gut microbiota dynamics after FMT in UC patients. Our study allowed us to assess changes on an almost weekly basis. RCTs with a strong longitudinal component often involve a smaller number of patients with more frequent repeated measurements, as compared to RCTs that focus on clinical outcomes. For example, in a recent clinical trial 42 patients provided

one stool sample for microbiota analysis before FMT, followed by one sample after FMT.[12] Another clinical trial included 12 patients who submitted stool samples weekly throughout their 12-week FMT treatment and at the 18-week follow-up.[47] A limitation of our study is that the results of statistical analyses should be interpreted with caution due to multiple testing in a small number of patients. Yet, most associations found in cluster analysis were retained in repeated measures analyses where we also accounted for the correlation of repeated observations within each patient. Moreover, despite the relatively small number of patients (n = 24) and donors (n = 2), both DMM and PCA clustering utilize all 180 patient samples and 27 donor samples available, rather than considering observations per patient.

Microbiota data is compositional, high-dimensional, and often zero-inflated.[28, 48] Moreover, the intestinal microbiota exhibits complex interactions, including competition and cooperation, forming intricate networks.[49, 50] These characteristics pose challenges to analytical methods, such as mixed models, which are commonly employed to investigate temporal variation and potential differences in bacterial abundance trajectories among clinical groups. Our analysis was limited by the individual modeling of each bacterial family, neglecting the interplay and interactions between families within the microbiota network. However, results obtained by supervised models of family-specific abundance over time were in line with results obtained by unsupervised methods (PCA and DMM clustering) that use community characteristics. Cluster analysis has been widely employed to explore the relationship between gut microbiota and conditions such as child gut development, depression, obesity, and IBD.[12, 51-53] Conventionally, unsupervised methods are suitable for exploratory analyses.[32] If the distinct clusters that we identified are confirmed in further larger-scale longitudinal analyses, this may lead to tailored diagnosis and treatment approaches based on specific cluster characteristics.[54] In our study, this would for example mean that the FMT treatment is stopped or changed to another donor when patients are found to be in the Prevotellaceae-dominated cluster during the treatment. While clustering techniques provide valuable insights, it is important to recognize that they depend on various choices by the modeler, including cutoffs and priors, which may lead to different clustering results.

Our study is admittedly rather exploratory in nature, but consistently revealed indications of a potential association between controlled abundances of Prevotellaceae with successful clinical and endoscopic remission following FMT treatment in UC patients. Moreover, we also highlighted a potential beneficial role of Lachnospiraceae and Ruminococcaceae. This provides a basis for new hypotheses regarding the role of gut microbiota in UC. Therapeutic interventions may be refined in the future, with early prediction of clinical outcomes and more personalized FMT treatments.

List of abbreviations

UC	Ulcerative colitis
FMT	Fecal microbiota transplantation
mOTU	Metagenomic-based Operational Taxonomic Unit
RCT	Randomized Controlled Trial
i.s.	*incertae sedis*
PCA	Principal Component Analysis
DMM	Dirichlet Multinomial Mixtures
AIC	Akaike Information Criterion
LMM	Linear Mixed effect Model
ZILMM	Zero-Inflated Linear Mixed effect Model

Declarations

Acknowledgements
We would like to acknowledge dr. E. van Lingen for her contribution to the data collection process.

Ethics approval and consent to participate
This research project was reviewed and approved by the Medical Ethical Committee in the LUMC, with reference number NL 65976.098.18. The study was registered in the Netherlands Trial Register, with reference number NL9858.

Consent for publication
Not applicable

Availability of data and material
R code is available via GitHub (https://github.com/susannepinto/FECBUD_microbiome.git) and the in-house preprocessing workflow is available via (https://git.lumc.nl/snooij/metagenomics-preprocessing/-/releases/v1.0). We have uploaded the metagenomic sequences to NCBI with: SRA Bioproject PRJNA1071720

Supplementary documents are available online on the publisher's website:

https://academic.oup.com/ecco-jcc/advance-article/doi/10.1093/ecco-jcc/jjae137/7748263?searchresult=1#supplementary-data

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by the Dutch Organization for Scientific Research (NWO) through the program Complexity in Health and Nutrition (NWO grant 645.001.002; www.nwo.nl/onderzoeksprogrammas/complexiteit), with co-funding by the National Institute for Public Health and the Environment (RIVM) of the Netherlands. The clinical trial was supported by a research grant from Vedanta Bioscience (Boston, US).

Authors' contributions

The authors confirm contribution to the paper as follows: study conception and design: SP, DS, EB, JAB, ES; funding acquisition: JAB, EB; data collection: SN, EMT, JJK, AEvdM-dJ; analysis and interpretation of results: SP, DS, EB, JAB, ES; draft manuscript preparation: SP, DS, EB, JAB, ES. All authors reviewed the results and approved the final version of the manuscript.

References

1. Seyedian SS, Nokhostin F, Malamir MD. A review of the diagnosis, prevention, and treatment methods of inflammatory bowel disease. J Med Life. 2019;12(2):113-122.
2. Khor B, Gardet A, Xavier RJ. Genetics and pathogenesis of inflammatory bowel disease. Nature. 2011;474(7351):307-317.
3. Molodecky NA, Kaplan GG. Environmental risk factors for inflammatory bowel disease. Gastroenterol Hepatol (N Y). 2010;6(5):339-346.
4. Halfvarson J, Brislawn CJ, Lamendella R, Vazquez-Baeza Y, Walters WA, Bramer LM, et al. Dynamics of the human gut microbiome in inflammatory bowel disease. Nat Microbiol. 2017;2:17004.
5. Alam MT, Amos GCA, Murphy ARJ, Murch S, Wellington EMH, Arasaradnam RP. Microbial imbalance in inflammatory bowel disease patients at different taxonomic levels. Gut Pathog. 2020;12(1).
6. Schierova D, Brezina J, Mrazek J, Fliegerova KO, Kvasnova S, Bajer L, et al. Gut Microbiome Changes in Patients with Active Left-Sided Ulcerative Colitis after Fecal Microbiome Transplantation and Topical 5-aminosalicylic Acid Therapy. Cells. 2020;9(10).
7. Zhang SL, Wang SN, Miao CY. Influence of Microbiota on Intestinal Immune System in Ulcerative Colitis and Its Intervention. Front Immunol. 2017;8:1674.
8. Rinninella E, Raoul P, Cintoni M, Franceschi F, Miggiano GAD, Gasbarrini A, et al. What is the Healthy Gut Microbiota Composition? A Changing Ecosystem across Age, Environment, Diet, and Diseases. Microorganisms. 2019;7(1).
9. Kumari R, Ahuja V, Paul J. Fluctuations in butyrate-producing bacteria in ulcerative colitis patients of North India. World J Gastroenterol. 2013;19(22):3404-3414.
10. Sha S, Xu B, Wang X, Zhang Y, Wang H, Kong X, et al. The biodiversity and composition of the dominant fecal microbiota in patients with inflammatory bowel disease. Diagn Microbiol Infect Dis. 2013;75(3):245-251.
11. Nishihara Y, Ogino H, Tanaka M, Ihara E, Fukaura K, Nishioka K, et al. Mucosa-associated gut microbiota reflects clinical course of ulcerative colitis. Sci Rep. 2021;11(1):13743.
12. He R, Li P, Wang J, Cui B, Zhang F, Zhao F. The interplay of gut microbiota between donors and recipients determines the efficacy of fecal microbiota transplantation. Gut Microbes. 2022;14(1):2100197.
13. Iljazovic A, Roy U, Galvez EJC, Lesker TR, Zhao B, Gronow A, et al. Perturbation of the gut microbiome by *Prevotella* spp. enhances host susceptibility to mucosal inflammation. Mucosal Immunol. 2021;14(1):113-124.
14. Maconi G, Camatta D, Cannatelli R, Ferretti F, Carvalhas Gabrielli A, Ardizzone S. Budesonide MMX in the Treatment of Ulcerative Colitis: Current Perspectives on Efficacy and Safety. Ther Clin Risk Manag. 2021;17:285-292.
15. Ianiro G, Puncochar M, Karcher N, Porcari S, Armanini F, Asnicar F, et al. Variability of strain engraftment and predictability of microbiome composition after fecal microbiota transplantation across different diseases. Nat Med. 2022;28(9):1913-1923.

16. van Nood E, Vrieze A, Nieuwdorp M, Fuentes S, Zoetendal EG, de Vos WM, et al. Duodenal infusion of donor feces for recurrent *Clostridium difficile*. N Engl J Med. 2013;368(5):407-415.

17. Debast SB, Bauer MP, Kuijper EJ, European Society of Clinical M, Infectious D. European Society of Clinical Microbiology and Infectious Diseases: update of the treatment guidance document for *Clostridium difficile* infection. Clin Microbiol Infect. 2014;20 Suppl 2:1-26.

18. El-Salhy M, Hatlebakk JG, Gilja OH, Brathen Kristoffersen A, Hausken T. Efficacy of faecal microbiota transplantation for patients with irritable bowel syndrome in a randomised, double-blind, placebo-controlled study. Gut. 2020;69(5):859-867.

19. El Hage Chehade N, Ghoneim S, Shah S, Chahine A, Mourad FH, Francis FF, et al. Efficacy of Fecal Microbiota Transplantation in the Treatment of Active Ulcerative Colitis: A Systematic Review and Meta-Analysis of Double-Blind Randomized Controlled Trials. Inflamm Bowel Dis. 2023;29(5):808-817.

20. Sommer F, Anderson JM, Bharti R, Raes J, Rosenstiel P. The resilience of the intestinal microbiota influences health and disease. Nat Rev Microbiol. 2017;15(10):630-638.

21. Schmidt TSB, Li SS, Maistrenko OM, Akanni W, Coelho LP, Dolai S, et al. Drivers and determinants of strain dynamics following fecal microbiota transplantation. Nat Med. 2022;28(9):1902-1912.

22. van Lingen E, Nooij S, Terveer E, Crossette E, Prince A, Bhattarai S, et al. Fecal Microbiota Transplantation engraftment after budesonide or placebo in patients with active ulcerative colitis using pre-selected donors: a randomized pilot study. J Crohns Colitis. 2024;jjae043.

23. Terveer EM, Vendrik KE, Ooijevaar RE, Lingen EV, Boeije-Koppenol E, Nood EV, et al. Faecal microbiota transplantation for *Clostridioides difficile* infection: Four years' experience of the Netherlands Donor Feces Bank. United European Gastroenterol J. 2020;8(10):1236-1247.

24. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9(4):357-359.

25. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. 2018;34(17):i884-i890.

26. Milanese A, Mende DR, Paoli L, Salazar G, Ruscheweyh HJ, Cuenca M, et al. Microbial abundance, activity and population genomic profiling with mOTUs2. Nat Commun. 2019;10(1):1014.

27. Ruscheweyh HJ, Milanese A, Paoli L, Karcher N, Clayssen Q, Keller MI, et al. Cultivation-independent genomes greatly expand taxonomic-profiling capabilities of mOTUs across various environments. Microbiome. 2022;10(1):212.

28. Gloor GB, Macklaim JM, Pawlowsky-Glahn V, Egozcue JJ. Microbiome Datasets Are Compositional: And This Is Not Optional. Front Microbiol. 2017;8:2224.

29. Aitchison J, Barceló-Vidal C, Martín-Fernández JA, Pawlowsky-Glahn V. Mathematical Geology. 2000;32(3):271-275.

30. Barnett D, Arts I, Penders J. microViz: an R package for microbiome data visualization and statistics. Journal of Open Source Software. 2021;6(63).

31. Morgan M. Dirichlet-Multinomial Mixture Model Machine Learning for Microbiome Data. 2022.

32. Holmes I, Harris K, Quince C. Dirichlet multinomial mixtures: generative models for microbial metagenomics. PLoS One. 2012;7(2):e30126.

33. Brooks ME KK, van Benthem KJ, Magnusson A, Berg CW, Nielsen A, Skaug HJ, Maechler M, Bolker BM. glmmTMB Balances Speed and Flexibility Among Packages for Zero-inflated Generalized Linear Mixed Modeling. The R Journal. 2017;9(2):378-400.

34. Bates D, Mächler M, Bolker B, Walker S. Fitting Linear Mixed-Effects Models Usinglme4. Journal of Statistical Software. 2015;67(1).

35. Wald A. Tests of statistical hypotheses concerning several parameters when the number of observations is large. Transactions of the American Mathematical Society. 1943;54(3):426-482.

36. Lahti L, Shetty S. microbiome R package. 2012-2019.

37. Pinheiro J BD, R Core Team. nlme: Linear and Nonlinear Mixed Effects Models. 2022.

38. 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;149(1):102-109.e106.

39. Rossen NG, Fuentes S, van der Spek MJ, Tijssen JG, Hartman JH, Duflou A, et al. Findings From a Randomized Controlled Trial of Fecal Transplantation for Patients With Ulcerative Colitis. Gastroenterology. 2015;149(1):110-118 e114.

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

41. Larsen JM. The immune response to *Prevotella* bacteria in chronic inflammatory disease. Immunology. 2017;151(4):363-374.

42. Precup G, Vodnar DC. Gut *Prevotella* as a possible biomarker of diet and its eubiotic versus dysbiotic roles: a comprehensive literature review. Br J Nutr. 2019;122(2):131-140.

43. Lahti L, Salojärvi J, Salonen A, Scheffer M, Vos WMd. Tipping elements in the human intestinal ecosystem. Nature Communications 2014;4344.

44. Pinto S, Beninca E, Galazzo G, Jonkers D, Penders J, Bogaards JA. Heterogeneous associations of gut microbiota with Crohn's disease activity. Gut Microbes. 2024;16(1):2292239.

45. Rees NP, Shaheen W, Quince C, Tselepis C, Horniblow RD, Sharma N, et al. Systematic review of donor and recipient predictive biomarkers of response to faecal microbiota transplantation in patients with ulcerative colitis. EBioMedicine. 2022;81:104088.

46. Tian Y, Zhou Y, Huang S, Li J, Zhao K, Li X, et al. Fecal microbiota transplantation for ulcerative colitis: a prospective clinical study. BMC Gastroenterol. 2019;19(1):116.

47. Crothers JW, Chu ND, Nguyen LTT, Phillips M, Collins C, Fortner K, et al. Daily, oral FMT for long-term maintenance therapy in ulcerative colitis: results of a single-center, prospective, randomized pilot study. BMC Gastroenterol. 2021;21(1):281.

48. Kodikara S, Ellul S, Le Cao KA. Statistical challenges in longitudinal microbiome data analysis. Brief Bioinform. 2022;23(4).

49. Faust K, Sathirapongsasuti JF, Izard J, Segata N, Gevers D, Raes J, et al. Microbial co-occurrence relationships in the human microbiome. Plos Computational Biology. 2012;8:1-17.

50. Faust K, Raes J. Microbial interactions: from networks to models. Nature Reviews Microbiology volume. 2012;10:538–550.

51. Christensen L, Roager HM, Astrup A, Hjorth MF. Microbial enterotypes in personalized nutrition and obesity management. Am J Clin Nutr. 2018;108(4):645-651.

52. Lee SH, Yoon SH, Jung Y, Kim N, Min U, Chun J, et al. Emotional well-being and gut microbiome profiles by enterotype. Sci Rep. 2020;10(1):20736.

53. Roswall J, Olsson LM, Kovatcheva-Datchary P, Nilsson S, Tremaroli V, Simon MC, et al. Developmental trajectory of the healthy human gut microbiota during the first 5 years of life. Cell Host Microbe. 2021;29(5):765-776 e763.

54. Costea PI, Hildebrand F, Arumugam M, Backhed F, Blaser MJ, Bushman FD, et al. Enterotypes in the landscape of gut microbial community composition. Nat Microbiol. 2018;3(1):8-16.