Ecology meets human health: studies on human gut microbiota in health and disease

Pinto, S.

Part III

Ecological determinants of FMT treatment success

Chapter 5

Dynamics of gut microbiota after FMT

Susanne Pinto[1], Dominika Šajbenová[1], Elisa Benincà[2], Sam Nooij[3], Elisabeth M. Terveer[3, 4], Josbert J. Keller[4, 5, 6], Andrea E. van der Meulen-de Jong[5], Johannes A. Bogaards[7, 8*], Ewout W. Steyerberg[1*]

1 Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, the Netherlands
2 Centre for Infectious Disease Control, National Institute for Public Health and the Environment (RIVM), Bilthoven, the Netherlands
3 Leiden University Center for Infectious Diseases (LUCID) Research, Leiden University Medical Center, Leiden, the Netherlands
4 Netherlands donor Feces bank, Department of Medical Microbiology, Leiden University Medical Center, Leiden, the Netherlands
5 Department of Gastroenterology and Hepatology, Leiden University Medical Center, Leiden, the Netherlands
6 Department of Gastroenterology, Haaglanden Medisch Centrum, The Hague, the Netherlands
7 Department of Epidemiology and Data Science, Amsterdam UMC location Vrije Universiteit Amsterdam, Amsterdam, the Netherlands
8 Amsterdam Institute for Infection and Immunity (AI&I), Amsterdam UMC, Amsterdam, the Netherlands

* These authors contributed equally

J Crohns Colitis. 2025 Feb 4;19(2):jjae137. doi: 10.1093/ecco-jcc/jjae137.

Dynamics of gut microbiota after fecal microbiota transplantation in ulcerative colitis: success linked to control of Prevotellaceae

Abstract

Fecal microbiota transplantation (FMT) is an experimental treatment for ulcerative colitis (UC). We aimed to study microbial families associated with FMT treatment success. We analysed stools from 24 UC patients treated with four weekly FMTs 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. 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. 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.

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.[353] The etiology of UC is multifactorial, involving complex interplay between the host immune system, gut microbiota, and genetic and environmental factors.[354, 355, 374] UC patients exhibit reduced microbial diversity and alterations in the composition of their gut microbiota compared to healthy individuals.[355, 375] Notably, a decrease in Bacillota (formerly 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.[91, 196, 376, 377] Studies investigating associations with common Bacteroidota in the human gut, such as the Bacteroidaceae and Prevotellaceae families, have yielded conflicting results.[196, 200, 356, 376-379]

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.[196] However, many patients do not derive lasting benefits from these interventions and may even experience severe side effects.[380] Fecal microbiota transplantation (FMT) has emerged as a promising alternative treatment for microbiota-associated disorders, particularly in the treatment of recurrent *Clostridioides difficile* infection.[189, 190, 381] FMT involves transferring fecal matter from a healthy donor to a patient with the aim of modulating the microbiota composition towards a more favourable state. The effectiveness of FMT in UC is limited, with a lower response rate observed as compared to FMT treatment of *Clostridioides difficile* infection.[193] 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.[193] 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.[23, 199]

A small pilot study in patients with Crohn's disease suggests an additional value of FMT in maintaining remission after successful induction therapy with corticosteroids.[380, 382] Achieving or maintaining remission after FMT may be associated with engraftment of donor bacteria.[382, 383] We hypothesized that reducing inflammation promotes engraftment of the healthy donor microbiota, which in turn may result in clinical improvement in inflammatory bowel disease (IBD). To further explore the effects of corticosteroids on engraftment and clinical response, we performed a randomized study investigating the effects of three weeks budesonide pretreatment prior to FMT in patients with UC. The primary analysis showed no association between pretreatment or overall engraftment with clinical response. This may be because the anti-inflammatory potential of budesonide is limited after three weeks. However, there was a significant donor-dependent effect on engraftment, although the study was not powered to detect differences regarding clinical endpoints.[384] In the current study we aimed to further identify longitudinal associations between the microbiota composition and clinical response to FMT treatment. We explored differences in gut microbiota dynamics between patients with clinical remission and non-responders following FMT treatment.

Methods

The study population

For the current study we used the stool samples collected from 24 UC patients included in our previously described FMT trial (Appendix Table 5.1).[384] 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 included being at least 18 years old and having a confirmed diagnosis of mild to moderate UC, defined as a full Mayo score ranging from 4 to 9 (including a partial Mayo score and endoscopic sub score of 1 or 2). Exclusion criteria included, among others, proctitis, antibiotic use, 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 (Appendix Table 5.1): sex, age at baseline (years), donor ID (D07 or D08), pretreatment (placebo or budesonide), and clinical outcome at week 14.

Patients who did not complete the study because of progressive symptoms or 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-responders, 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 the end of weeks 3, 4, 5, and 6 after randomization) 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.[385] 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 by the physician who monitored the patients 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.[384] In total we collected 81 stool samples in the responder group (n = 9 patients) and 99 stool samples in the non-responder group (n = 15 patients). Stool samples of donors D07 and D08 were collected regularly, and a total of 27 samples (n = 13 samples for donor D07 and n = 14 samples 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, MN, USA) with the Illumina NovaSeq platform (100 bp single-end reads to a median depth of 2.9 million reads). Raw reads mapping to the human genome were removed using bowtie2 (version 2.4.2)[386] and the GRCh37 reference genome, and reads were quality-trimmed using fastp (version 0.20.1),[387] both part of an in-house workflow (git.lumc.nl/snooij/metagenomics-preprocessing). The mOTUs3 workflow (version 3.0.1) was used to generate taxonomic profiles.[388, 389] 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 mOTUs3 database includes taxa based on metagenomic bins that have not been formally classified, which are listed as '*incertae sedis*' (*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) and R code is available on the GitHub repository (susannepinto/FECBUD_microbiome).

Differences in relative abundances of specific microbial families among responders and non-responders were tested for statistical significance in repeated measures analyses, as described in the 'longitudinal models of bacterial relative abundances' 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 time points as the patient samples. Differences between donor D07 and donor D08 were tested with Pearson's χ^2 test and 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.[219, 390] 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.[391]

Dirichlet multinomial mixture models

We used the Dirichlet multinomial mixture (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.[392] 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.[393] 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 relative 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-effects model (LMM), possibly augmented with a zero-inflation component (ZILMM). The 'lme4' R package was used for constructing LMMs and the 'glmmTMB' R package was used for constructing ZILMMs.[394, 395] 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 the 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' R package) with a knot at week 8 (the beginning of the follow-up phase) was considered for all models.

Model preference was based on the lowest Akaike Information Criterion (AIC) and model diagnostics, judged by a QQ-plot and a plot of residuals against predicted values. All choices per family are given in Appendix Table 5.2.

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 the relative abundance of a particular family between non-responders and responders, with statistical significance assessed by Wald tests.[396] 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 fit); 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' R package.[397] 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' R package).[398] A log transformation was applied to the Simpson dominance measure to correct for non-normality. The regression parameter of primary interest was the relationship 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 of sex and time. Similar to the longitudinal LMM of bacterial families, time was modelled as a continuous variable with a natural cubic spline (knot 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 (Figure 5.1 and Appendix Figure 5.1). 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 with donor D08, whereas donor D08 had a significantly higher relative abundance of Bacillota fam. *i.s.* and Lachnospiraceae (Figure 5.1, Appendix Figure 5.1, and Appendix Table 5.3).

Overall, the most abundant bacterial family in the patients was Ruminococcaceae. However, from the second time point 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 time points (Figure 5.1, Appendix Figure 5.1, and Appendix Figure 5.2). Compared to the non-responders, Lachnospiraceae and Oscillospiraceae seemed to become more abundant in the responder group over time (Figure 5.1, Appendix Figure 5.1, and Appendix Figure 5.2).

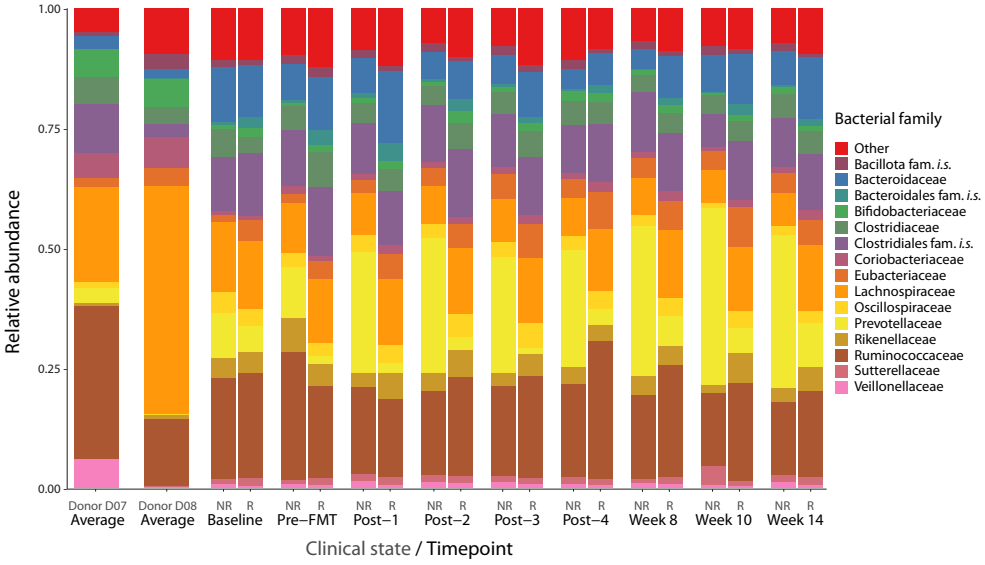


Figure 5.1. Average microbiota composition of the 15 most abundant bacterial families. Abundances were followed over time for the two donors, non-responders (NR), and responders (R). Here, the 'other' category includes all remaining bacterial families.

PCA results for donors and patients

The first two components in PCA of patient and donor samples, based on the Aitchison distance, explained 24% of the total variation in the data (Figure 5.2). The samples of donor D08 clustered away from the patients' samples, driven by a difference in the relative abundance of Lachnospiraceae (Figure 5.2). Patients treated with an FMT from donor D08 showed a higher responder rate than those from donor D07 (Appendix Table 5.1). The difference in distance between non-responders and responders seemed to be explained by the relative abundance of Prevotellaceae (Figure 5.2). This applied particularly to the patients who received an FMT from donor D08 (Appendix Figure 5.3). Only a few patient samples seemed to traverse considerable Aitchison 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) (Appendix Figure 5.4).

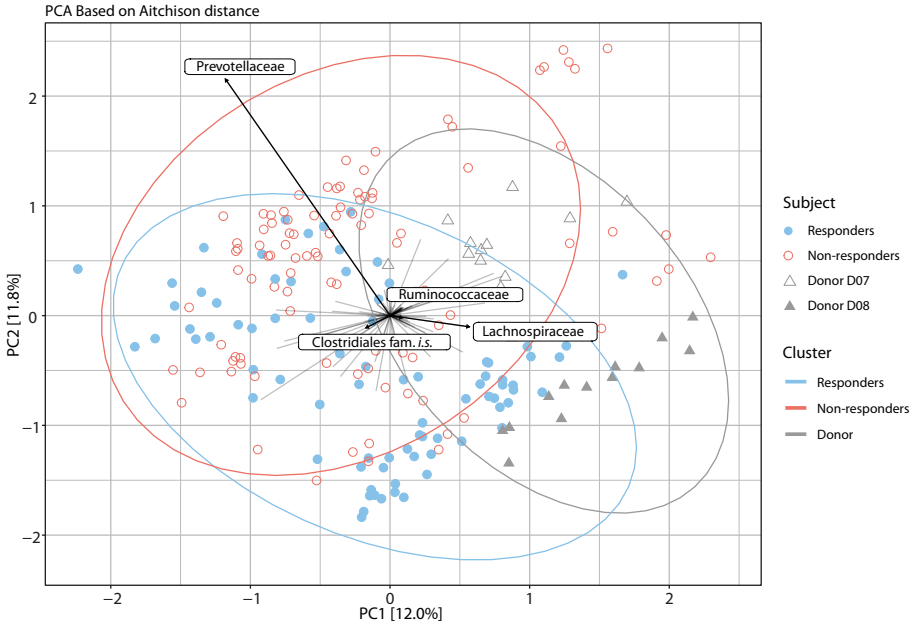


Figure 5.2. PCA plot with Aitchison distances in microbiota profiles, showing the distance between sample types. The PCA plots include data ellipses around the different groups and loading vectors of families to obtain an initial visualization about the extent of separation between non-responder, responder, and donor samples. The different symbols, closed circles, open circles, open triangles, and closed triangles, represent responders, non-responders, donor D07, and donor D08, respectively, while the different colors indicate the various groups (responders, non-responders, and donors).

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 5.3 and Appendix Figure 5.5 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 5.4). 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 5.4B). All donor samples, except for one, were assigned to cluster 4 (Appendix Figure 5.6). Five non-responder patient samples were also assigned to cluster 4 (Figure 5.4A). This donor-dominated cluster disappeared in sensitivity analysis on patient samples only (Appendix Figure 5.7A), resulting in the reassignment 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 a sensitivity analysis resulted in the disappearance of that cluster, with re-assignment of the other corresponding samples to cluster 2 (Appendix Figure 5.7B). Removing patients with only two measurements (patients 107 and 119) had a minor impact on the results (Appendix Figure 5.7C).

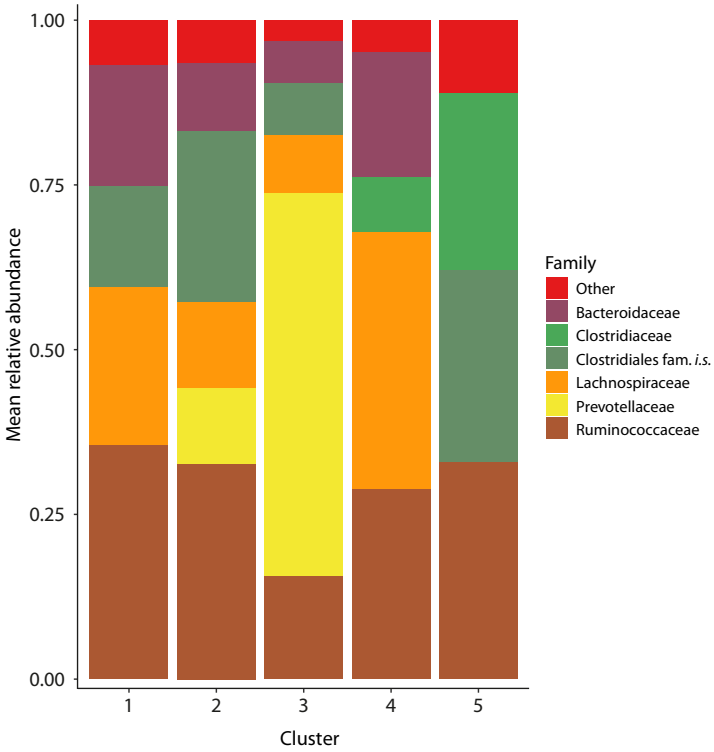


Figure 5.3. Mean relative abundance of bacterial families in the five clusters. Clusters are detected by the Dirichlet multinomial mixture model.

Out of 24 patients, nine patients (38%) remained in the same cluster for all of their provided samples (Figure 5.4A). An alluvial plot of patient samples showed the substantial changes in sample membership and cluster size throughout the clinical trial (Figure 5.4C). 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 toward more responder samples in cluster 1 from Pre-FMT onwards. Samples in cluster 1 were exclusively composed of responder samples at time points Post-4, Week 10, and Week 14. Cluster 3 was fully composed of non-responder samples after pretreatment and after every FMT treatment (Figure 5.4C).

Coloring samples by their cluster membership in the PCA plot of Aitchison distances showed separation among clusters 1, 2, and 3, with cluster 2 being the intermediate cluster (Appendix Figure 5.6). The Prevotellaceae vector was pointed in the direction of cluster 3, corresponding to a potential association between this cluster and non-response (Appendix Figure 5.6), possibly driven by the donor (Figure 5.2 and Appendix Figure 5.3). 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 (Appendix Figure 5.6). Finally, samples from cluster 5 were tightly grouped together, likely because they all originated from the same patient.

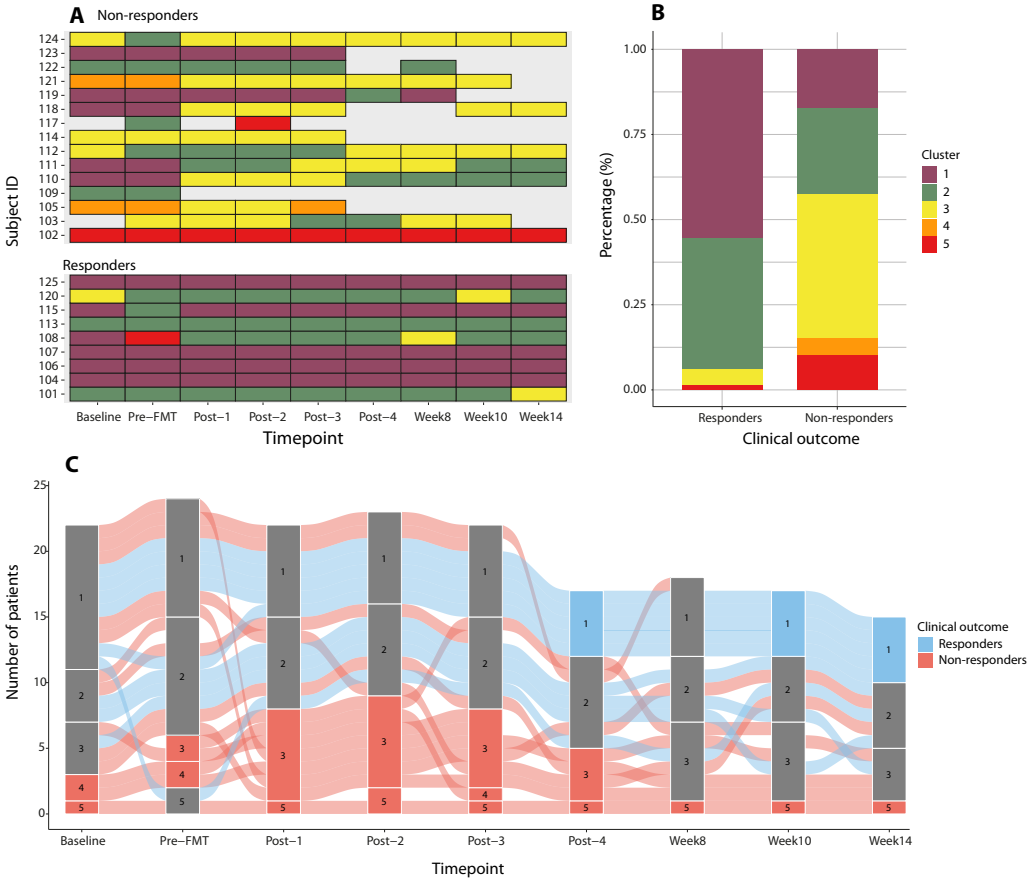


Figure 5.4. Clustering of donor and patient samples. A) Cluster membership over time per patient for the non-responders (upper facet) and responders (lower facet). Lack of colored bar indicates that no stool sample was collected at that time point. B) Percentage of each cluster for non-responders and responders. C) Alluvial plot of patients distributed over the different clusters over time. This plot displays the distribution of clusters per time point and whether each cluster is comprised of only one clinical group (e.g., only non-responders) for every time point. A grey box means that the cluster at that time point contains both samples from responder and non-responder patients, a colored box only contain responder samples or only non-responder samples.

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 5.5, Appendix Table 5.2, and Appendix Figure 5.2). Prevotellaceae showed the greatest difference in trajectory between responders and non-responders over time. Note that the preferred model for Prevotellaceae had a straightforward linear trajectory and used only 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 (Appendix Figure 5.8).

There were four families with a significant donor effect, namely Veillonellaceae, Rikenellaceae, Sutterellaceae, and Bifidobacteriaceae (Appendix Table 5.2). 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 (Appendix Table 5.2). 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.

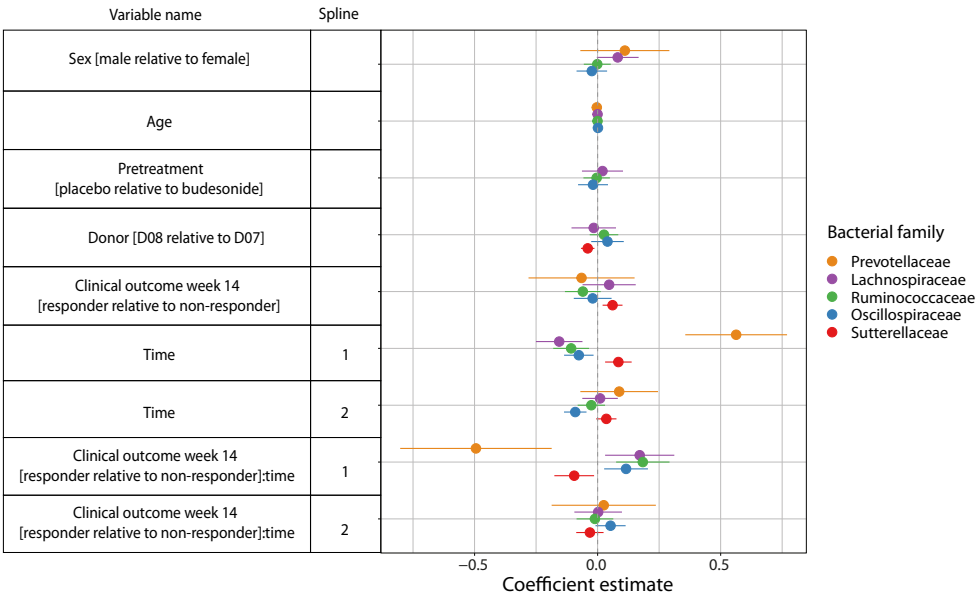


Figure 5.5. Results of the mixed models. Only the families among the 15 most abundant families (Prevotellaceae, Lachnospiraceae, Ruminococcaceae, and Oscillospiraceae) for whom we found a significant effect in relation to clinical response with the Wald test are shown. The point estimates, 95% confidence 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. All *p*-values are given in Appendix Table 5.2.

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 5.6). There was a significant difference between the Simpson dominance in responder and non-responder patients (Wald test: *p*-value = 0.004). Our study was too small to determine whether this difference already existed at baseline or developed over time (Appendix Table 5.4). The LMM random-intercept model suggested that there was also a significant sex effect (Appendix Table 5.4). 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.

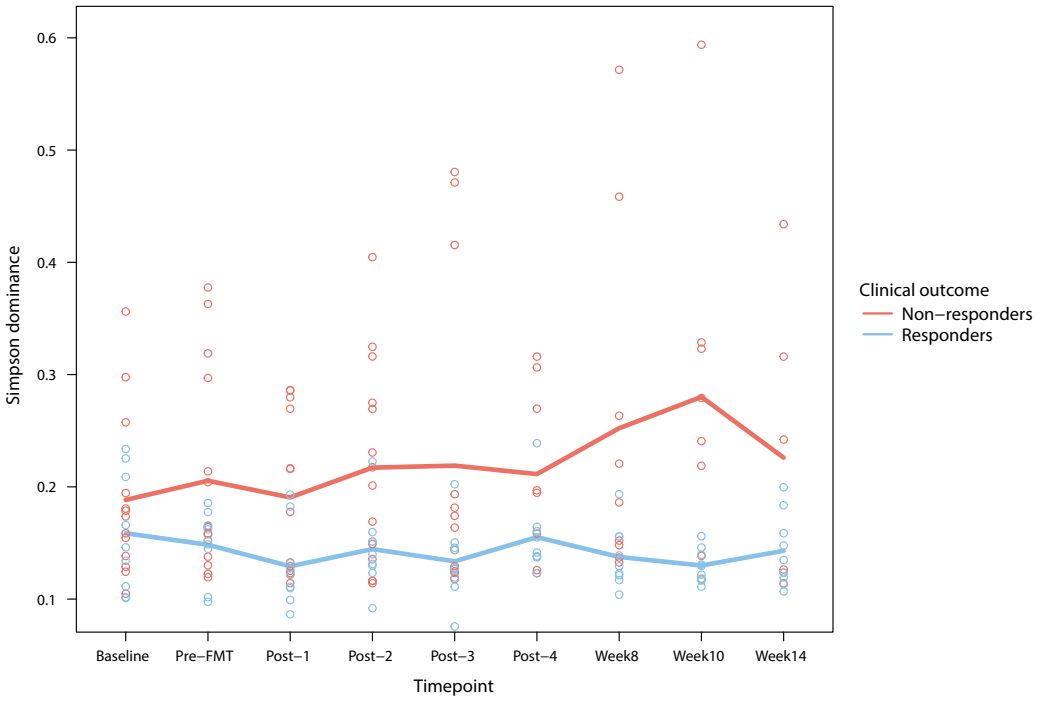


Figure 5.6. Change in Simpson dominance calculated for non-responders and responders. The points indicate the individual measurements of the patients. The lines are the mean Simpson dominance per group.

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.[355, 375] 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 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.[198] 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,[23, 198, 199, 383] and suggests that some complementarity in microbiota compositions between donors and recipients is required for a successful clinical response.[199, 200] In other words, the complementarity of the donor-patient pairing seems more important to achieve clinical remission than attaining a donor-like microbiota. 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 gut microbiome diversity has been associated with a higher clinical response before.[399] In addition, higher post-FMT diversity has been associated with remission, suggesting that the variety of introduced organisms may promote recovery.[23] It was already noted that donor D08 was the more successful donor; however, intriguingly, this was the donor with the least engraftment.[384] 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 favourable starting state, while those who received FMT from donor D07 required stronger microbiota changes to move to a more favourable 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 a moderate abundance and favourable clinical outcomes following FMT in UC patients. In addition, 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 before and during treatment, and matching donors to patients accordingly might improve the response rate. However, a 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.[378] 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.[200] 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.[400, 401] For instance, *Prevotella intestinalis* has been shown to induce intestinal inflammation upon colonization in mice.[379] *Prevotella melaninogenica* and *Prevotella oralis* have been characterized as tipping elements.[402] 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 among the most heterogeneous across individual patients (see Chapter 4).[403]

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.[195] Lachnospiraceae and Ruminococcaceae may play a role in modulating the immune response and inflammatory pathways in the colon.[195] Earlier attempts to cluster the gut microbiota of healthy and unhealthy individuals showed clusters dominated by *Bacteroides*, *Prevotella*, and *Ruminococcus*.[404-406] While our study identified clusters dominated by Prevotellaceae and Ruminococcaceae, we did not find clusters dominated by *Bacteroides* (i.e., Bacteroidaceae). This discrepancy could be due to differences in the study populations, or the specific methodologies used for microbiota analysis. 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.[407] In addition, in contrast to the present study, previous research has reported an increased abundance of Enterobacteriaceae in UC patients who did not respond to drug and surgical interventions, with higher levels being associated with mucosal inflammation.[378] 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.[378] 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.[200]

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 during and 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 measures, as compared to RCTs that focus on clinical outcomes. For example, in a recent clinical trial 42 patients provided a single stool sample for microbiota analysis before FMT, followed by another single sample after FMT.[200] Another clinical trial included 12 patients who submitted stool samples weekly throughout their 12-week FMT treatment and at the 18-week follow-up.[408] A limitation of our study is that the results of statistical analyses should be interpreted with caution due to multiple tests 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 are compositional, high-dimensional, and often zero-inflated.[217, 219] Moreover, the intestinal microbiota exhibits complex interactions, including competition and cooperation, that form intricate networks.[8, 252] 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.[200, 409-411] Conventionally, unsupervised methods are suitable for exploratory analyses.[393] 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.[412] 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.

Appendices of Chapter 5

Appendix Table 5.1 - Clinical and demographic information of responders and non-responders.

	Responders[a]	Non-responders[b]
	Number (Percentage)	**Number (Percentage)**
Patients	9 (38%)	15 (63%)
Samples	81 (45%)	99 (55%)
Missing	0	36
Sex		
% Female[c]	6 (67%)	6 (40%)
Pretreatment		
% Budesonide[c]	5 (56%)	8 (53%)
Donor		
% D07[c]	2 (22%)	10 (67%)
	Mean (SD)	**Mean (SD)**
Age	45 (17)	48 (16)

[a] Remission (i.e., response) was defined at week 14 as no symptoms (partial Mayo score of 2 with no individual sub score of > 2) and an endoscopic Mayo score 0–1.

[b] All other patients, including those with a partial response (a decrease of at least 3 points in the partial Mayo score and at least 1 point at the endoscopic Mayo score) at week 14 and patients who left the study early, were classified as non-responders.

[c] Percentages calculated separately for responders and non-responders.

Appendix Table 5.2 - Model choice and mixed models results for the 15 most abundant families.

Significant results are obtained via a χ^2 statistic (Wald test). Significant results are highlighted in bold and blue. Absence of a *p*-value means that the variable was not included in the model.

Families	Model choice	Sex (*p*-value)	Age (*p*-value)	Pretreatment (*p*-value)	Donor (*p*-value)	Clinical outcome[a] (*p*-value)
Bacillota fam. *i.s.*	LMM (random intercepts)	**0.004**	0.760	0.877	0.535	0.428
Bacteroidaceae	LMM (random intercepts)	0.243	0.377	-	0.794	0.052
Bacteroidales fam. *i.s.*	ZILMM (random intercepts)	0.182	-	-	-	0.546
Bifidobacteriaceae	ZILMM (random intercepts)	0.230	-	-	**0.023**	0.104
Clostridiaceae	ZILMM (random intercepts)	0.377	0.694	-	-	0.439
Clostridiales fam. *i.s.*	ZILMM (random slopes)	0.197	0.280	0.821	0.629	0.909
Coriobacteriaceae	ZILMM (random intercepts)	0.618	0.027	-	0.825	0.146
Eubacteriaceae	LMM (random slopes)	0.509	0.701	0.499	0.337	0.661
Lachnospiraceae	LMM (random intercepts)	0.059	0.904	0.640	0.734	**0.014**
Oscillospiraceae	LMM (random intercepts)	0.459	0.135	0.550	0.233	**0.020**
Prevotellaceae	LMM (random intercepts)	0.230	0.251	-	-	**< 0.001**
Rikenellaceae	ZILMM (random intercepts)	**< 0.001**	0.061	-	**0.038**	0.181
Ruminococcaceae[b]	LMM (random intercepts)	0.963	0.708	0.891	0.381	**0.011**
Sutterellaceae	ZILMM (random intercepts)	-	-	-	**0.004**	**0.010**
Veillonellaceae	ZILMM (random slopes)	0.589	0.503	**< 0.001**	**0.046**	0.435

[a] Wald test on multiple parameters: Responders, Responders x time point (first and second spline)

[b] No transformation

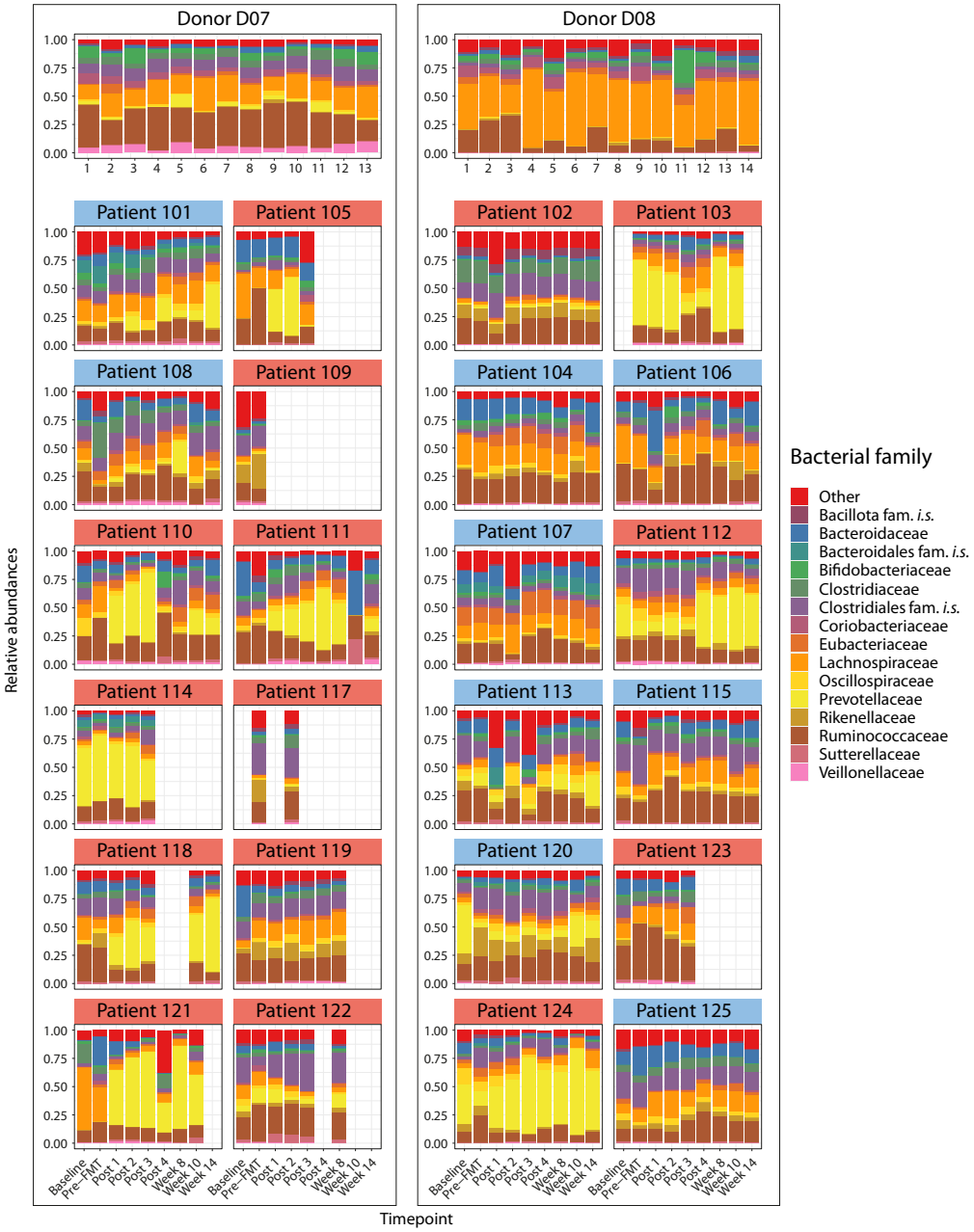
Appendix Table 5.3 - Significant differences in bacterial abundances between the two donors (for donor D07 *n* = 13 and for donor D08 *n* = 14 samples).
The results are obtained with the independence test. Significant results are highlighted in bold and blue.

Family	Mean relative abundance		*p*-value[a]
	Donor D07	Donor D08	
Bacillota fam. *i.s.*	0.0072	0.0327	**< 0.001**
Bacteroidaceae	0.0265	0.0199	0.311
Bacteroidaceae fam. *i.s.*	0.0003	0.0002	0.722
Bifidobacteriaceae	0.0575	0.0575	0.100
Clostridiaceae	0.0572	0.0365	**0.003**
Clostridiales fam. *i.s.*	0.1033	0.0265	**< 0.001**
Coriobacteriaceae	0.0501	0.0660	0.144
Eubacteriaceae	0.0204	0.0367	0.060
Lachnospiraceae	0.1971	0.4755	**< 0.001**
Oscillospiraceae	0.0126	0.0024	0.006
Prevotellaceae	0.0314	0.0000	**0.004**
Ruminococcaceae	0.3183	0.1400	**< 0.001**
Sutterellaceae	0.0018	0.0027	0.371
Veillonellaceae	0.0582	0.0000	**< 0.001**

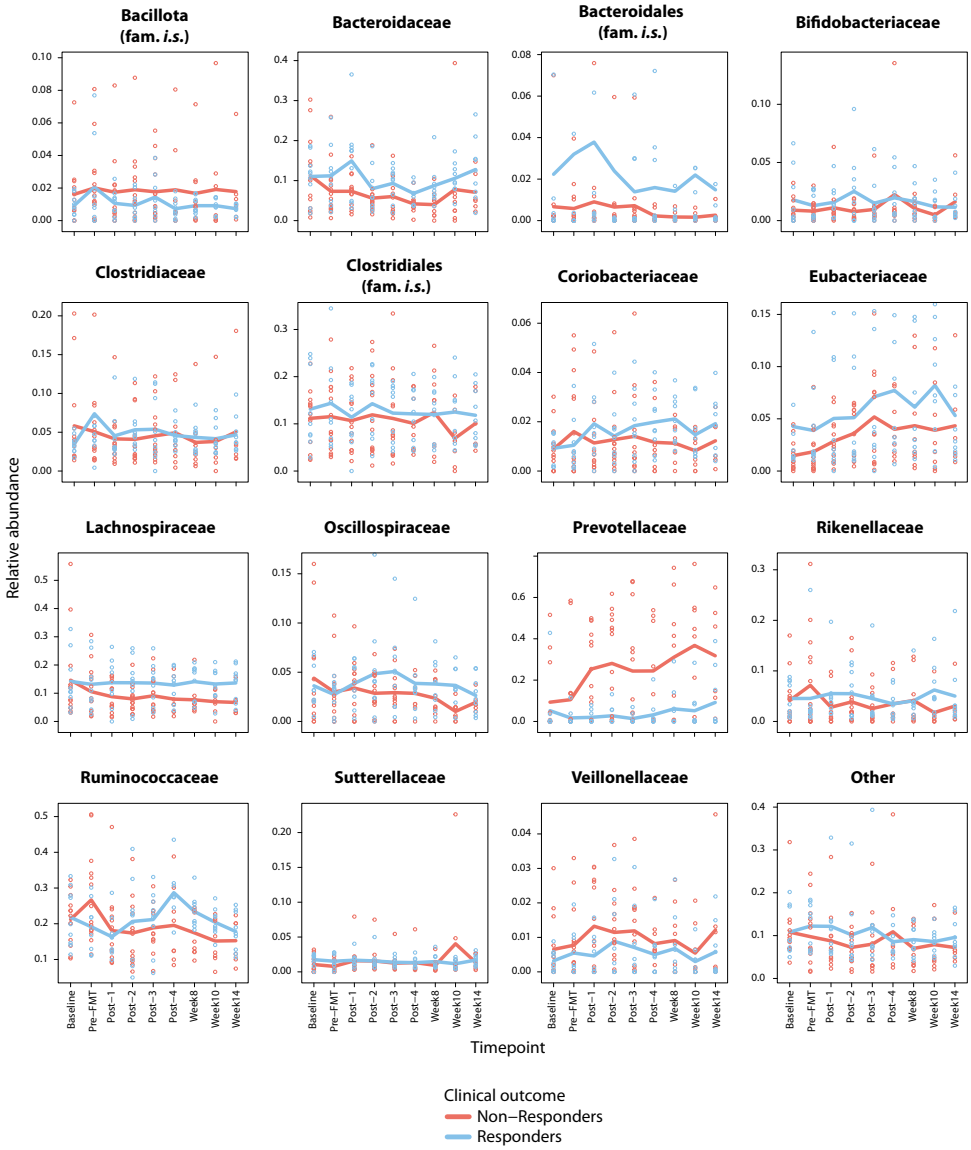
[a] After a Bonferroni correction in which the adjusted *p*-value threshold was 0.004

Appendix Table 5.4 - Regression coefficients and *p*-values of the Simpson dominance random-intercepts LMM.
Significant results are highlighted in bold and blue.

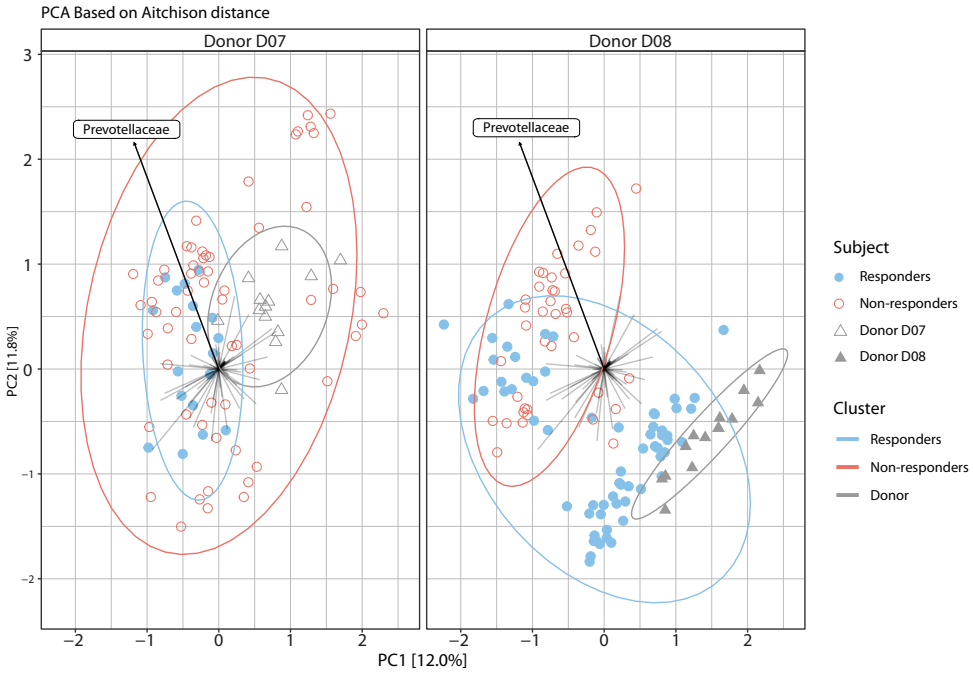
Predictors	Estimates	Standard error	*p*-value
(Intercept)	–1.88	0.1	**< 0.001**
Sex (male relative to female)	0.27	0.09	**0.01**
Clinical outcome (responder relative to non-responder)	–0.14	0.13	0.30
Time (1st spline)	0.30	0.16	0.06
Time (2nd spline)	0.23	0.12	0.06
Clinical outcome (responder relative to non-responder) * Time (1st spline)	–0.2	0.24	0.07
Clinical outcome (responder relative to non-responder) * Time (2nd spline)	–0.22	0.16	0.17



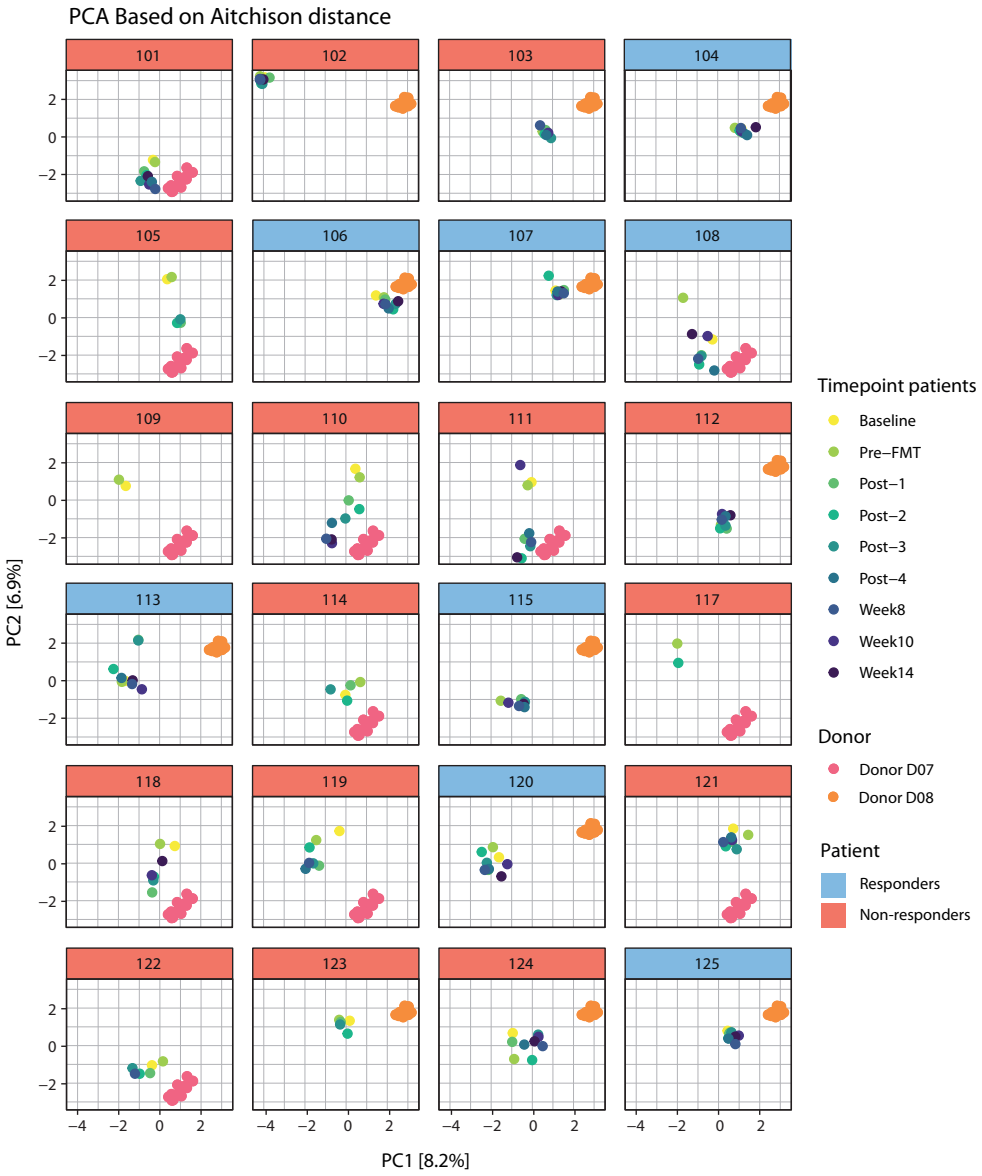
Appendix Figure 5.1 - Composition of the 15 most abundant families in the donors and the patients' microbiota over time. The 12 patients at the left-hand side of the plot (under the plot of donor D07) were treated with feces from donor D07. The 12 patients at the right-hand side of the plot (under the plot of donor D08) were treated with samples of donor D08. Patients with a blue title are responders, patients with a red title are non-responders.



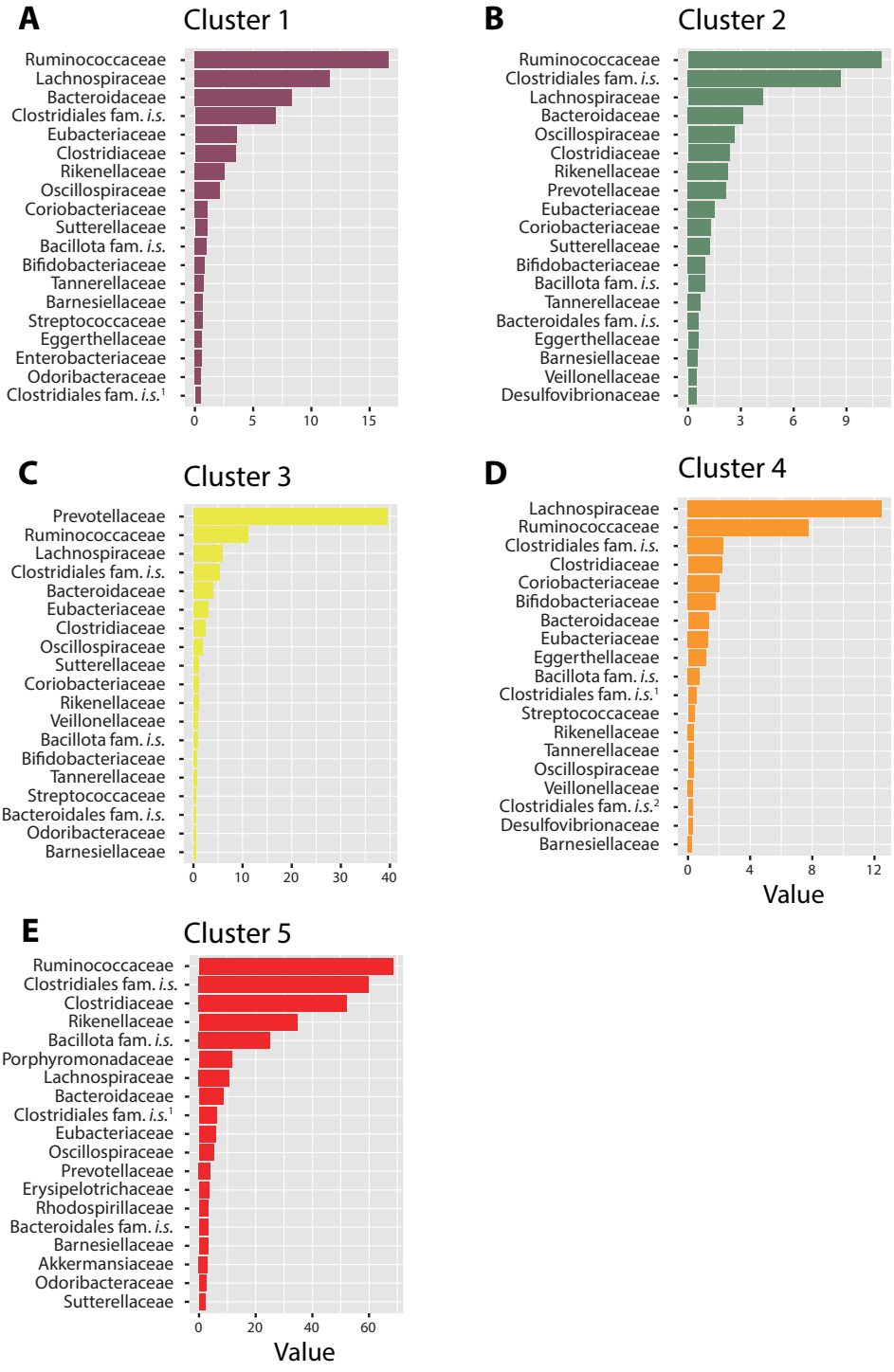
Appendix Figure 5.2 - Relative abundances over time of the 15 most abundant bacterial families. The points indicate the individual measurements of the patients. The lines are the mean relative abundances per group (responders in blue and non-responders in red).



Appendix Figure 5.3 - PCA plot with Aitchison distances in microbiota profiles differentiated per donor. The PCA plots include data ellipses around the different groups (e.g., blue for the responders, red for the non-responders, and grey for the donors) and a loading vector of Prevotellaceae to obtain an initial visualization about the extent of separation between responders, non-responders, and donor samples. The different symbols, closed circles, open circles, open triangles, and closed triangles, indicate responders, non-responders, donor D07, and donor D08, respectively.

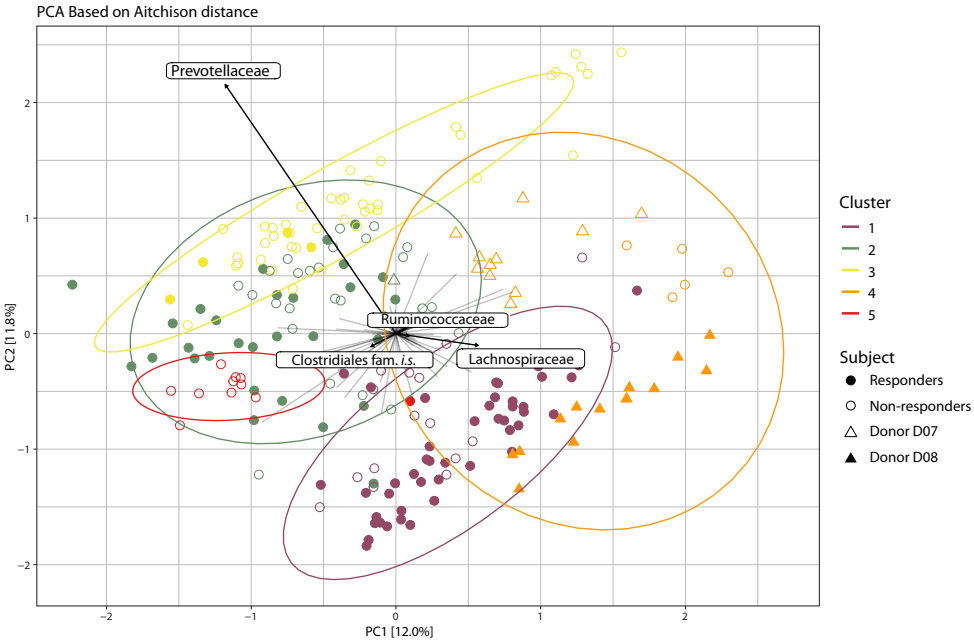


Appendix Figure 5.4 - Plot with Aitchison distances in microbiota profiles differentiated per patient and corresponding donor. Patients with a blue title are responders, patients with a red title are non-responders.

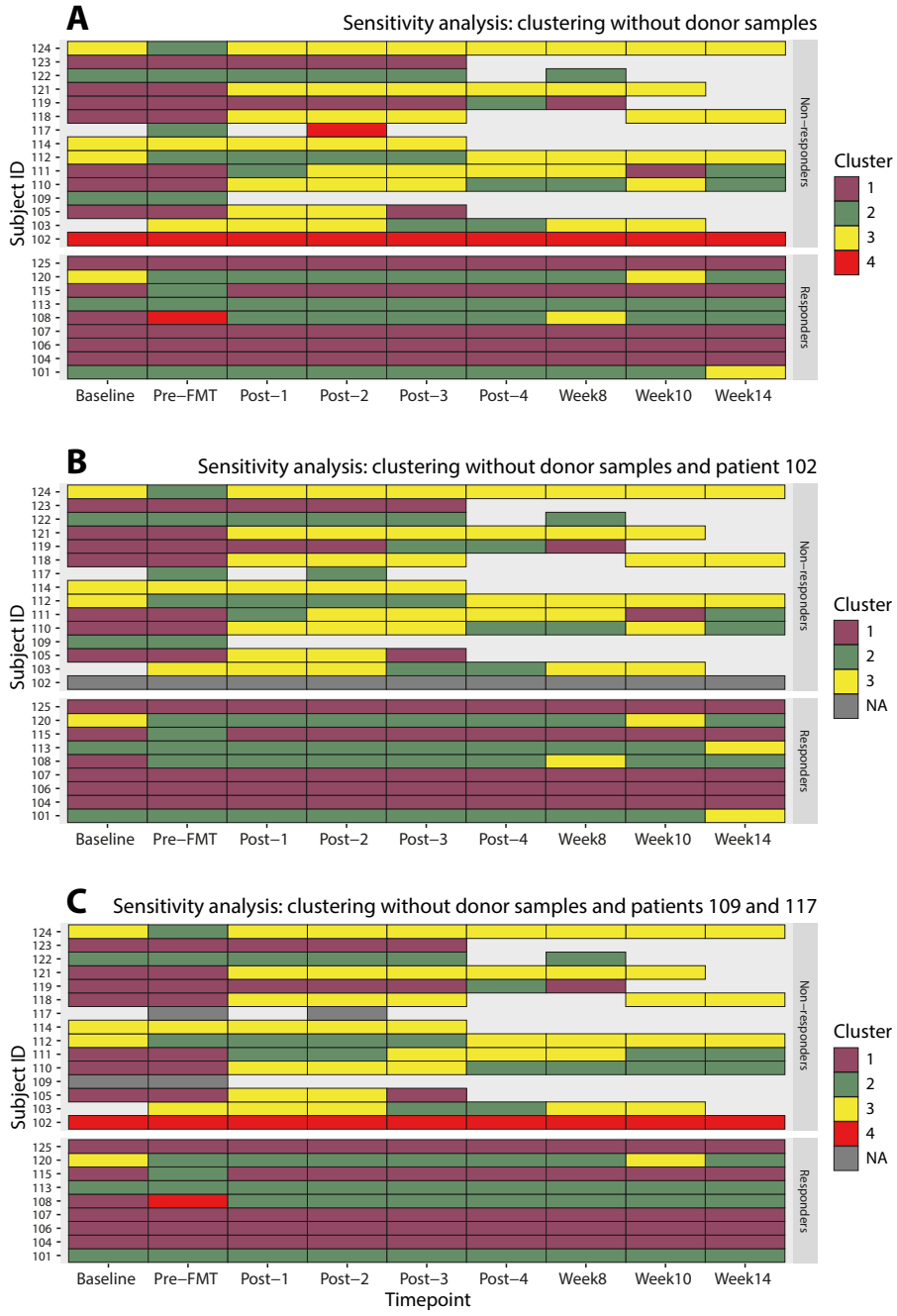


[1] Family is Lachnospiraceae or Clostridiaceae
[2] Family is Eubacteriaceae or Ruminococcaceae

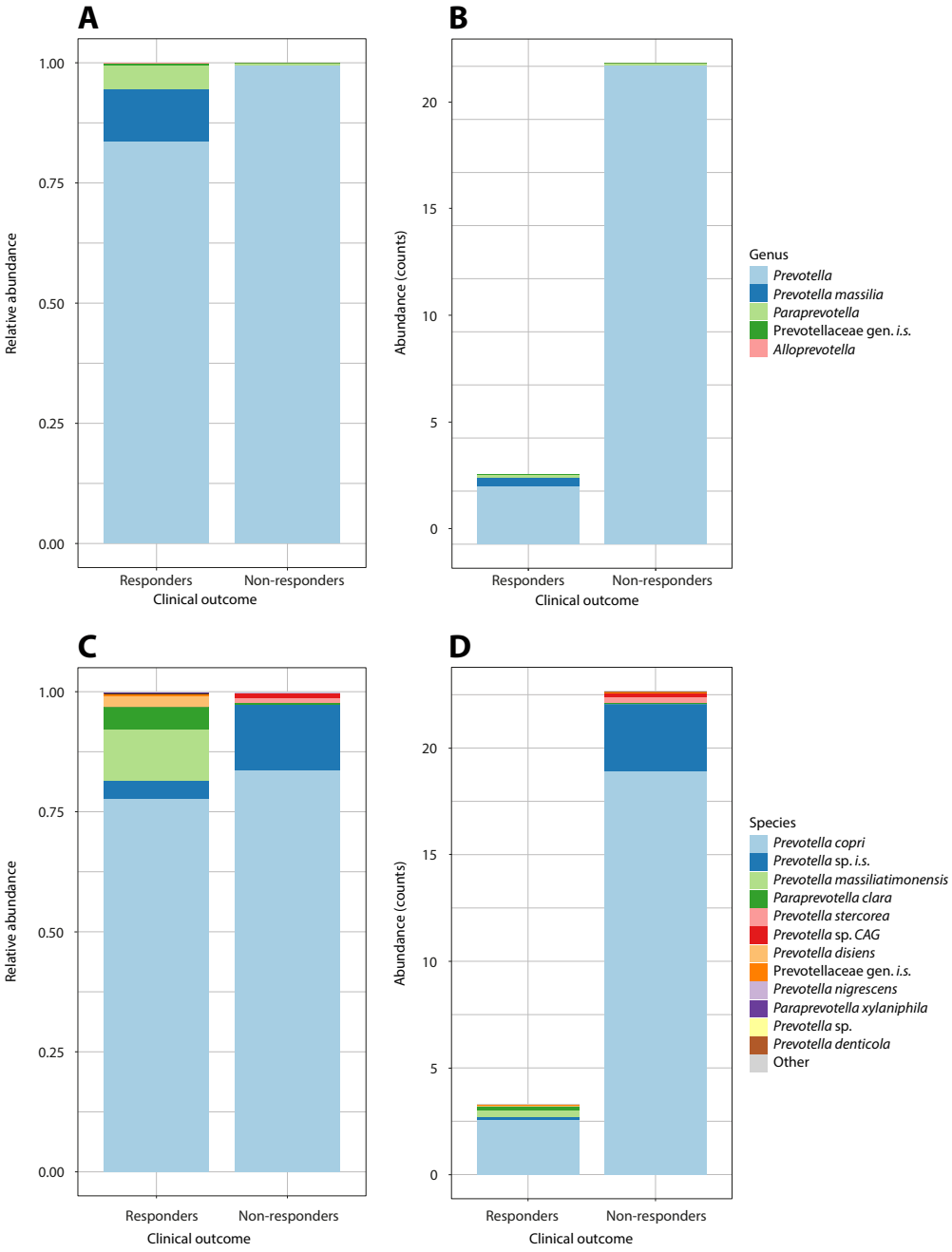
Appendix Figure 5.5 - Importance of the contribution of different families to each cluster. A) Cluster 1, B) Cluster 2, C) Cluster 3, D) Cluster 4, and E) Cluster 5.



Appendix Figure 5.6 - PCA plot with Aitchison distances in microbiota profiles for different clusters, showing the taxa that generally differ across the samples. The PCA plots include data ellipses around the different Dirichlet clusters and loading vectors of families to obtain an initial visualization about the extent of separation between patient (responders and non-responders) and donor samples. The different symbols, closed circles, open circles, open triangles, and closed triangles, indicate responders, non-responders, donor D07, and donor D08, respectively.



Appendix Figure 5.7 - Sensitivity analyses of DMM models. A) patient samples only, B) patient samples excluding patient 102 (with a distinct microbiota from all other patients), and C) patient samples excluding patients 109 and 117 (only two samples available for those patients).



Appendix Figure 5.8 - Genera (panels A and B) and species (panels C and D) within the Prevotellaceae family. Relative abundances (panels A and C) and counts (panels B and D) are given.