Fecal microbiota signatures are predicitve of response to therapy among Ustekinumab-treated Crohn's Disease patients.

Running title: microbiota of Ustekinumab-treated Crohn's Disease patients.

Matthew K. Doherty[2], Tao Ding[2α], Charlie Koumpouras[2], Shannon Telesco[1], Calixte Monast[1], and Patrick D. Schloss[2†]

† To whom correspondence should be addressed: pschloss@umich.edu

1. Janssen Pharmaceutical Companies of Johnson & Johnson, Spring House, PA, USA

2. Department of Microbiology and Immunology, University of Michigan, Ann Arbor, MI, USA

α Currently at Department of Biology, New York University, New York, NY, USA.

Abstract

The fecal microbiota is a rich source of biomarkers that have previously been shown to be predictive of numerous disease states. Less well studied is whether these biomarkers can be predictive of response to therapy. Here we sought to predict the therapuetic response of Crohn's disease (CD) patients enrolled in a double-blinded, placebo-controlled, Phase 2b clinical trial to test the efficacy of Ustekinumab (UST). CD is a global health concern characterized by patches of ulceration and inflammation along the gastrointestinal tract and alterations to the microbial community structure. Using stool samples collected over the course of 22 weeks, we characterized the composition of these patients' fecal bacterial communities by sequencing the V4 region of the 16S rRNA gene. We were able to differentiate patients in remission from those with active disease 6 post treatment induction by using RF models trained on the composition of their baseline microbiome and baseline clinical metadata (AUC $= 0.844$, specificity $= 0.831$, sensitivity $= 0.774$). Top predictive OTUs that were ubiquitous among patients included *Faecalibacterium* and *Escherichia/Shigella*. Among patients in remission 6 weeks post induction, the median baseline inverse Simpson index was 1.7 times higher than treated patients with active disease at week 6. Their baseline community structures were similarly different. Two OTUs, *Faecalibacterium* and *Bacteroides*, were significantly more abundant at baseline in patients who were in remission 6 weeks post induction. In treated patients we could follow to week twenty-two, the α-diversity of UST treated clinical responders increased over time, in contrast to nonresponsive patients. The fecal microbiota at baseline was also associated with markers for disease severity, such as Crohn's Disease Activity Index (CDAI), stool frequency, CRP, fecal lactoferrin, and fecal calprotectin. The observed baseline differences in fecal microbiota and changes due to therapeutic response support using the microbiota as a biomarker for the establishment and maintenance CD remission.

Importance: Finding biomarkers that give clinicians the ability to predict response to CD treatment at diagnosis will increase the likelihood of faster induction and maintenance of remission. The fecal microbiota could be a useful non-invasive biomarker for directing or monitoring the treatment of CD patients. OTUs associated with remission post induction induction, especially *Faecalibacterium*, could be biomarkers for successful UST treatment of TNF-α refractory CD patients.

Keywords: Crohn's Disease, IBD, fecal microbiome, microbiota, biologics, prediction, biomark-

ers, remission, Faecalibacterium, Ustekinumab, Stelara, machine learning, random forest

Introduction

The microbiome has been correlated with a variety of diseases and has shown promise as a predictive tool for disease outcome for gingivitis (1), cardiovascular disease (2), *Clostridium difficile* infection (**???**, 3), and colorectal cancer (4, 5). Additionally, the microbiome has been shown to affect the efficacy of various therapies, including vaginal microbicides (6), cardiac drugs (7), and cancer treatements (8, 9), and therefore could be used to predict therapeutic response. In relation to inflammatory bowel disease (IBD), previous studies have shown that the bacterial gut microbiota correlates with disease severity in new-onset, pediatric Crohn's disease (CD) patients (10, 11). Additionally, recent studies have shown promise for the gut microbiota as it relates to IBD and therapeutic response (12, 13). It remains to be determined, however, whether the composition of the fecal gut microbiota can predict and monitor response to CD therapy. Considering the involvement of the immune system and previous evidence for involvement of the microbiome, it is likely that response to CD therapy can be predicted using microbiome data.

CD is a global health concern causing large economic and healthcare utilization impacts on society (14, 15). CD is characterized by patches of ulceration and inflammation along the entire gastrointestinal tract, though mostly the ileum and colon. Currently, individuals with CD are treated based on disease location and risk of complications using escalating immunosuppressive treatment, and/or surgery, with the goal of achieving and sustaining remission (16, 17). Faster induction of remission following diagnosis reduces the risk of irreversible intestinal damage and disability (17–19). Ideally, clinicians would be able to determine personalized treatment options for CD patients at diagnosis that would result in faster achievement of remission (20). Therefore, recent research has been focused on identifying noninvasive, prognostic biomarkers to monitor CD severity and predict therapeutic response (21–23).

The precise etiology of CD remains unknown, but host genetics, environmental exposure, and the gut microbiome appear to be involved (14, 24). Individuals with CD have reduced microbial diversity in their guts, compared to healthy individuals, with a lower relative abundance of *Firmicutes* and an increased relative abundance of *Enterobacteraciae* and *Bacteroides*, at the phylum level (10, 25–28). Additionally, genome-wide association studies of individuals with CD identified several susceptibility loci, including loci involved in the IL-23 signaling pathway, which could impact the gut microbiota composition and

function (16, 25, 29–32). If the fecal microbiota can be used to monitor disease severity and predict response to specific treatment modalities, then clinicians could use it as a noninvasive tool for prescribing therapies that result in faster remission (33).

The FDA recently approved Ustekinumab (UST), a monoclonal antibody directed against the shared p40 subunit of IL-12 and IL-23, for the treatment of CD (17, 34–36). Given the potential impact of IL-23 on the microbiota (29–32), we hypothesized that response to UST could be predicted or influenced by differences in patients' gut microbiota and that UST treatment may alter the fecal microbiota. The effects of biologic treatment of IBD on the microbiota are not yet well described, but are hypothesized to be indirect, as these drugs act on host factors (16). We analyzed the fecal microbiomes of patients who participated in a double-blinded, placebo-controlled Phase II clinical trial that demonstrated the safety and efficacy of UST for treating CD (34). We tested whether clinical responders had a microbiota that was distinct from non-responders and if the fecal microbiota changed in patients treated with UST using 16S rRNA gene sequence data from these patients' stool samples. We also quantified the association between the fecal microbiota and disease severity. Our study demonstrates that these associations are useful in predicting and monitoring UST treatment outcome and suggest the fecal microbiota may be a broadly useful source of biomarkers for predicting response to treatment.

Results

Fecal microbiota based prediction of treatment response

We characterized the fecal microbiota in a subset of anti-TNF-α refractory CD patients, with moderate to severe CD, who took part in the double-blinded, CERTIFI clinical trial that demonstrated the efficacy of UST in treating CD (34). Demographic and baseline disease characteristics of this subset are summarized in Table 1. Patients were randomly assigned to a treatment group in the induction phase of the study and at week 8 patients were re-randomized into maintenance therapy groups based on their induction response (Figure 1A). In our study response is defined as a decrease in a patient's initial Crohn's Disease Activity Index (CDAI) greater than 30%. The CDAI is the standard instrument for evaluating clinical symptoms and disease activity in CD (37, 38). The CDAI weights patient reported stool frequency, abdominal pain, and general well being over a week, in combination with weight change, hematocrit,

opiate usage for diarrhea, and the presence of abdominal masses or other complications to determine the disease severity score (37, 38). The international definitions for the levels of CD activity include relrespission (CDAI < 150), mild to moderate (CDAI 150<220), moderate to severe (CDAI 220<450), and severe (CDAI < 450) (37). Patients provided stool samples at baseline (screening) and at 4, 6, and 22 weeks post induction for analysis using 16S rRNA gene sequencing (Figure 1B).

We hypothesized that the baseline fecal microbiota could predict therapuetic response (CDAI deacrease >30%) 6 weeks post induction. To test this hypothesis, we constructed random-forest (RF) models to classify responders from non-responders 6 weeks post induction based on the relative abundance of fecal microbiota community members at baseline, clinical metadata at baseline, and the combination of microbiota and clinical data (5, 39). Clinical data included components of the CDAI, biomarkers for inflammation, and patient metadata described further in the methods section. We ran these models on 232 baseline stool samples from patients induced with UST. Clinical data alone resulted in an AUC of 0.693 (specificity = 0.76, sensitivity = 0.598) (Figure 2A). Using only microbiota data, the model predicted response with an AUC of 0.737 (specificity = 0.807, sensitivity = 0.585). When combining clinical metadata with the microbiome, the model predicted response with an AUC of 0.745 (specificity = 0.727, sensitivity = 0.744). These models were not significantly different in their ability to predict response. Optimal predictors were determined based on their mean decrease in accuracy (MDA) in the ability of the model to classify response (Figure 2B).

Prediction of remission following treatment

We also investigated whether the baseline fecal microbiota could predict therapuetic remission (CDAI < 150) 6 weeks post induction. To test this hypothesis, we again used RF models to classify patients in remission from those with active CD 6 weeks post induction. Clinical data alone resulted in an AUC of 0.616 (specificity = 0.801, sensitivity = 0.452) (Figure 2C). Using only fecal microbiota data the model had an AUC of 0.838 (specificity = 0.766, sensitivity = 0.806). Finally, when combining clinical metadata with the microbiota we achieved an AUC of 0.844 (specificity = 0.831, sensitivity = 0.774) for remission at week six. Prediction with clinical metadata alone did not perform as well as models using the baseline fecal microbiome ($p = 0.0011$) or the combined model ($p = 0.00087$). However, there was no significant difference between the baseline fecal microbiota model and the combined model **(p=)**. Also, our baseline fecal microbiota model was significatnly better able to classify remission compared to

response (p = 0.043), whereas this was not true for the combined model (p = 0.055).

The majority of OTUs identified as optimal classifiers in our model for remission were low in abundance across our cohort (Figure 2D). However, two OTUs appeared to be differentially abundant for patients in remission at week six. The relative abundance of *Escherichia/Shigella* (OTU00001) appeared lower in remitters (median = 1.07 IQR = 0.033-3.7) compared to patients with active CD (median = 4.13, IQR = 0.667-15.4). Also, the relative abundance of *Faecalibacterium* (OTU00007) was not only higher in remitters (median = 7.43, IQR = 1.43-11.9) than patients with active CD (median = 0.167, IQR = 0-5.1), it was present prior to the start of treatment in every patient who was in remission at week six post induction.

Comparison of clinical responders and non-responders

As our random forest models identified OTUs abundant across our cohort that were important in classifying response and remission, we further investigated differences in the baseline microbiota that could serve as potential biomarkers for successful UST treatment. We compared the baseline microbiomes of all 306 patients who provided a baseline sample based on treatment group and treatment outcome 6 weeks post induction. Patients in remission 6 weeks post induction with UST had significantly higher diversity based on the inverse Simpson index than patients with active CD (respective median values = 11.6 (IQR = 4.66-13.9), 6.95 (IQR = 4.4-11.8), p = 0.020). No other treatment or response groups were significantly different. β-diversity was significantly different for response and remission in UST treated patients 6 weeks post induction (response p = 0.012, remission p = 0.017). No phyla were significantly different by treatment and response (Supplemental Figure 1) and no OTUs were significantly different among patients receiving placebo for induction, regardless of response and remission status. Two OTUs were significantly more abundant in patients in remission 6 weeks post induction compared to patients with active CD; *Bacteroides* (OTU19) (p = 0.022) and *Faecalibacterium* (OTU7) (p = 0.0026) (Figure 3).

Variation in the baseline microbiota is associated with variation in clinical data

Based on the associations we identified between baseline microbial diversity and response, we hypothesized that there were associations between the microbiota and clinical variables at baseline that could support the use of the microbiota as a non-invasive biomarker for disease activity (33). To test this

hypothesis, we compared the baseline microbiota with clinical data at baseline for all 306 samples provided at baseline (Supplemental Table 1). We observed small, but significant correlations for lower α-diversity correlating with higher CDAI (ρ = -0.161, p = 0.014), higher frequency of loose stools per week (ρ = -0.193, p = 0.003), and longer disease duration (ρ = -0.225, p = 0.001). Corticosteroid use was associated with higher α-diversity (p = 0.001). No significant association was observed between α-diversity and CRP, fecal calprotectin, or fecal lactoferrin. However, the β-diversity was significantly different based on CRP (p = 0.033), fecal calprotectin (p = 0.006), and fecal lactoferrin (p = 0.004). The β-diversity was also significantly different based on weekly loose stool frequency (p= 0.024), age (p = 0.033), the tissue affected (p = 0.004), corticosteroid use (p =0.01), and disease duration (p = 0.004). No significant differences in the microbiota were observed for BMI, weight, or sex.

The diversity of the microbiota changes in UST responders

We tested whether treatment with UST alters the microbiota by performing a Friedman test comparing α-diversity at each time point within each treatment group based on response 22 weeks post induction. We included 48 patients induced and maintained with UST (20 responders, 28 non-responders) and 14 patients induced and maintained with placebo (10 responders, 4 non-responders), who provided samples at every time point (Figure 1). We saw no significant difference in the inverse Simpson index over time in patients who did not respond 22 weeks post induction, regardless of treatment, and in patients who received placebo (Figure 4). However, the median inverse Simpson index of responders 22 weeks post UST induction significantly changed over time (p = 0.005) having increased from baseline (median = 6.65, IQR = 4.61 - 9.19) to 4 weeks post UST induction(median = 11.3, IQR = 6.59 - 16.0), decreased from 4 to 6 weeks post induction (median = 8.42, IQR = 4.68 - 16.5), and was significantly higher than baseline ($p < 0.05$) at 22 weeks post induction (median = 11.4, IQR = 5.62 - 15.7).

The microbiota post induction can distinguish between treatment outcomes

Having demonstrated the microbiome changes in patients who responsd to UST treatment, we hypothesized that the microbiota could be used to monitor response to UST therapy by classifying patients based on disease activity (33). We again contrstructed a random forest classification model to distinguish between patients by UST treatment outcome based on their fecal microbiota (5, 39). The study design resulted in only 75 week twenty-two stool samples from patients induced and maintained with

UST, so we focused our analysis on the 220 week 6 stool samples from patients induced with UST. We were again better able to distinguish patients in remission from patients with active CD compared to responders from non-responders ($p = 0.0019$; Figure 5A). Our model using week 6 stool samples for response 6 weeks post induction could classify patients who responded from non-responders with an AUC of 0.708 (sensitivity = 0.769, specificity = 0.606). For remission 6 weeks post UST induction, the model had an AUC of 0.866 (sensitivity = 0.833, specificity = 0.832) when classifying patients in remission from patients with active CD. Important classifiers again included *Fecalibacterium* (OTU7) (Figure 5B).

Discussion

With this study we sought to determine whether the microbiota can be used to identify patients who will respond to UST therapy and to gain a more detailed understanding of if and how UST treatment affects the microbiota. We demonstrated that the microbiota could be useful in predicting remission due to UST therapy, compared to clinical metadata alone, in our unique patient cohort. We also showed found the fecal microbiota to be useful in uncovering associations between the microbiota and aspects of CD severity metrics and treatment outcomes. Finally, we found that the microbiota of treated responders changed over time. These results helped us to gain a better understanding of the interaction between the human gut microbiota and CD pathogenesis in adult patients refractory to anti-TNF-α therapies with moderate to severe CD.

Porgnositc model generation paragraph

The presented prognostic model is useful for biomarker discovery and hypothesis generation about the biology of CD as it relates to the microbiome. Similar models could be further developed into a clinically useful prognostic tool. *Faecalibacterium* was the most frequently occurring OTU in our models. It is associated with health, comprising up to 5% of the relative abundance in healthy individuals, and has been shown to be rare in CD patients (25, 27, 40, 41). Each patient in remission six weeks post UST induction had measurable *Faecalibacterium* present at baseline. This supports the hypothesis that *Faecalibacterium* impacts CD pathogenesis. *Escherichia/Shigella* also occurred frequently in our models. This OTU is associated with inflammation and has been shown to be associated with CD pathogenesis

(41). Many other taxa observed in our analysis had low abundance or were absent in the majority of patients. However, in many cases these taxa are related and may serve similar ecologic and metabolic roles in the gut environment. We hypothesize that these microbes may have genes that perform similar metabolic functions. These functions could be revealed by performing metagenomics on stool samples in future studies, especially in patients who achieve remission.

We were better able to classify remission status compared to response status. We hypothesize that this is due to the relative nature of the response criteria compared to the threshold used to determine remission status. We defined response as a decrease in a patient's baseline CDAI of 30% or more, while remission was defined as a CDAI below 150. The original study used a decrease in CDAI of 100 points for their measure of response, but we felt using the relative percent change better represented a meaningful difference in disase activity and patient quality of life (34). Additionally, the field appears to be moving away from CDAI and towards more objectively quantifiable biomarkers for inflammation as wells as endoscopic verification of mucosal healing. (18).

We observed several associations between the microbiota and clinical variables that could impact how CD is monitored and treated in the future. Serum CRP, fecal calprotectin, and fecal lactoferrin are used as biomarkers to measure inflammation and CD severity. We found that the microbial community structure is different among patients based on these markers. This supports the hypothesis that the fecal microbiota could function as a biomarker for measuring disease activity in patients, especially in concert with established inflammatory biomarkers (33, 42, 43). We also found that higher CDAI was associated with lower microbial diversity. This is consistent with other studies on the microbiota in individuals with CD compared to healthy individuals and studies looking at active disease compared to remission (10, 33, 44). However, these differences may have been driven by the CDAI subscore of weekly stool frequency (Supplementary Table 1), which is consistent with previous studies (45). We also observed differences in the microbial community structure based on disease localization, which is consistent with a study by Naftali et al (40). Our study also found that corticosteroid use impacts the composition of the human fecal microbiota, which is consistent with observations in mouse models (46). As corticosteroid use appears to impact diversity, corticosteroid therapy may be useful when trying to positively impact microbial diversity during biologic therapy and thereby increase the possibility of response to CD therapies. We also observed that longer disease duration is associated with a reduction

in fecal microbial diversity. This decreased diversity may be due to the long duration of inflammatory conditions in the gut. This observation and the increased likelihood of remission and mucosal healing in individuals treated with biologics earlier in the course of their disease is an argument for earlier biologic intervention (47–49). Hypothetically, earlier biologic intervention would occur before chronic inflammation resultled in reduced microbial diveristy. A more diverse microbiota may then promote remission and reduce the likelihood of relapse. However, the cost of biologics for patients is hindrance to early biologic intervention. Using aptamers in place of monoclonal antibodies may alleviate this expense (50).

We observed that the α-diversity of clinical responders increased over time, in contrast to nonresponsive patients. This observation could be due to lower inflammation and changes in disease activity corresponding to improved health in patients who responded to UST. We also addressed whether response to therapy can be predicted with the microbiota by developing a random-forest model that used relative microbial abundance data and/or clinical metadata for input. Again, we were better able to predict remission/non-remission than response/non-response. These findings are again consistent with other studies suggesting the microbiota could be useful as a biomarker in detecting remission versus active disease (33).

The positive and negative associations between the microbiota and CD allow us to hypothesize on ways to alter the microbiota in order to increase the likelihood therapeutic response. Prior to the initiation of therapy, patients could have their fecal microbiome analyzed. The community data could then be used to direct the modification of a patient's microbiome prior to or during treatment with the goal of improving the outcome of UST treatment. Additionally, further research into the microbiota as a prognostic biomarker could eventually allow for the screening of patients with stool samples at diagnosis to better inform other treatment decisions. If the fecal microbiota can be used as a prognostic tool to non-invasively predict response to specific treatment modalities or inform treatment, then more personalized treatment could result in faster achievement of remission, thereby increasing patients' quality of life and reducing economic and healthcare impacts.

Methods

Study Design and Sample Collection

Janssen Research and Development conducted a placebo-controlled, phase II clinical study of approximately 500 patients to assess the safety and efficacy of UST for treating anti-TNF-α refractory, moderate to severe CD patients (34). Both patients and clinicians were blinded to their induction and maintenance treatment groups. Participants provided a stool sample prior to the initiation of the study and were then divided into 4 groups of 125 individuals receiving placebo or 1, 3, or 6 mg/kg doses of UST by IV. Additional stool samples were provided at week four. At week six an additional stool sample was collected, patients were scored for their response to UST based on CD Activity Index (CDAI), and then divided into groups receiving either subcutaneous injection of UST or placebo at weeks 8 and 16 as maintenance therapy. Finally, at 22 weeks patients provided an additional stool sample and were then scored using CDAI for their response to therapy. Response was defined as a decrease in a patient's initial CDAI of 30% or more. This value was determined by using the approximate percent change in CDAI from mild-moderate CD (220) to remission (150). Remission was defined as a CDAI below the threshold of 150. Frozen fecal samples were shipped to the University of Michigan and stored at -80°C prior to DNA extraction.

DNA extraction and 16S rRNA gene sequencing

Microbial genomic DNA was extracted using the PowerSoil-htp 96 Well Soil DNA Isolation Kit (MoBio Laboratories) and an EPMotion 5075 pipetting system, as previously described (4, 5). The V4 region of the 16S rRNA gene from each sample was amplified and sequenced using the Illumina MiSeq™ platform as described elsewhere (43). Sequences were curated as described previously using the mothur software package (v.1.34.4) (51, 52). Briefly, we reduced sequencing and PCR errors (53), aligned the resulting sequences to the SILVA 16S rRNA sequence database (54), and removed any chimeric sequences using UCHIME (55). Sequences were clustered into operational taxonomic units (OTU) at a 97% similarity cutoff using the average neighbor algorithm (56). All sequences were classified using a naive Bayesian classifier trained against the RDP training set (version 14) and OTUs were assigned a classification based on which taxonomy had the majority consensus of sequences within a given OTU (57). All fastq

files and the MIMARKS spreadsheet with de-identified clinical metadata are available at **TBD**.

Following sequence curation using the mothur software package (51), we obtained a median of 13,732 sequences per sample (IQR = 7,863-21,978). Parallel sequencing of a mock community had an error rate of 0.017%. To limit effects of uneven sampling, we rarefied the dataset to 3,000 sequences per sample. Samples from patients that completed the clinical trial and had complete clinical metadata were included in our analysis. Of these samples, 306 were provided prior to treatment as well as 258 provided at week four, 289 at week six, and 205 at week twenty-two post-treatment, for a total of 1,058 samples.

Gut microbiota biomarker discovery and statistical analysis

R v.3.3.2 (2016-10-31) and mothur were used for our data analysis (58). To assess α-diversity, the inverse Simpson index was calculated for each sample in the dataset. Spearman correlation tests were performed to compare the inverse Simpson index and continuous clinical data. Wilcoxon rank sum tests were performed for pairwise comparisons and Kruskal-Wallis rank sum tests for comparisons with more than two groups (59, 60). To measure β-diversity, the distance between samples was calculated using the thetaYC metric, which takes into account the types of bacteria and their abundance to calculate the differences between the communities (61). These distance matrices were assessed for overlap between sets of communities using the non-parametric analysis of molecular variance (AMOVA) and homogeneity of variance (HOMOVA) tests in mothur (62), as well as the adonis function in the R package vegan (v.2.4.2) (63). Change in α-diversity over time based on week twenty-two response was assessed using a Friedman test on patients who provided a sample at each time point (64). The Friedman test is a function in the R package stats (v.3.3.2). Multiple comparisons following a Friedman test were performed using the friedmanmc function in the package pgirmess (v.1.6.5) (65). Change in beta-diversity over time by treatment group and response was assessed using the adonis function in vegan stratified by patient. We used the relative abundance of each OTU, inverse Simpson index, age, sex, current medications, BMI, disease duration, disease location, fecal calprotectin, fecal lactoferrin, C-reactive protein, bowel stricture, and CDAI subscores as input into our RF models constructed with the AUCRF R package (v.1.1) (66), in order to identify phylotypes or clinical variables that distinguish between various treatment and response groups, as well as to predict or determine response outcome

(67). Optimal predictors were determined based on their mean decrease in accuracy (MDA) of the model to classify patients. Differentially abundant OTUs and phyla were selected through comparison of clinical groups using Kruskal-Wallis and Wilcox tests, where appropriate, to identify OTUs/phyla where there was a p-value less than 0.05 following a Benjamini-Hochberg correction for multiple comparisons (68). Other R packages used in our analysis included ggplot2 v.2.2.1 (69), dplyr v.0.5.0 (70), pROC v.1.9.1 (71), knitr v.1.15.1 (72–74), gridExtra v.2.2.1 (75), devtools v.1.12.0 (76), knitcitations v.1.0.7 (77), scales v.0.4.1 (78), tidyr v.0.6.1 (79), Hmisc v.4.0.2 (80), and cowplot v.0.7.0 (81).

We compared α-diversity at baseline to clinical variables using the inverse Simpson index with Spearman correlation, Wilcoxon rank sum, or Kruskal-Wallis rank sum, tests where appropriate. We compared β-diversity with a PERMANOVA using the adonis function in the vegan R package. Following multiple comparison correction, w

Tables

Table 1: Summary of clinical metadata of chort at baseline

Clinical Variable	Treated	Placebo	Total
	n = 232	n = 74	n = 306
Age (years)	38 ± 13	40 ± 14	39 ± 13
Sex (% Male)	36.6	43.2	38.2
Race (% Caucasian)	91.8	93.2	92.2
Corticosteroid Use (%)	40.1	52.7	43.1
BMI (kg/m^2)	26 ± 6.7	25 ± 4.9	25 ± 6.3
Disease Duration (years)	12 ± 8.4	13 ± 10	12 ± 8.8
CDAI	330 ± 62	310 ± 69	320 ± 64
Bowel Stricture (%)	12.5	10.8	12.1
Tissue Involvement (%) Colon/Ileocolic/Ileal	28.9/51.3/19.8	24.3/39.2/36.5	27.8/48.4/23.9

Supplemental Table 1: Diversity differences based on clinical metadata of chort at baseline

Clinical Variable	Correlation	Alpha-Diversity (p-value)	Beta-Diversity (p-value)
CDAI	ρ = -0.2	0.014	0.324
Loose Stool Frequency (per week)	ρ = -0.2	0.003	0.024
C-Reactive Protein (mg/L serum)	ρ = 0.06	0.394	0.033
Fecal Calprotectin (µg/g)	ρ = 0.08	0.254	0.006
Fecal Lactoferrin (µg/g)	ρ = 0.1	0.070	0.004
BMI	ρ = 0.07	0.299	0.277
Weight (kg)	ρ = 0.07	0.299	0.112
Age (years)	ρ = -0.05	0.472	0.033
Sex (F/M)	-	0.539	0.277
Corticosteroid Use (Y/N)	-	0.001	0.010
Disease Duration (years)	ρ = -0.2	0.001	0.004
Tissue Involvement	-	0.190	0.004

Figures

Figure 1: Experimental design as adapted from Sandborne et al 2012. (A) Diagram of experimental design. Participants were divided into 4 groups of 125 individuals receiving placebo or 1, 3, or 6 mg/kg doses of UST by IV. At week 8, patients were divided into groups receiving either subcutaneous injection of UST or placebo at weeks 8 and 16 as maintenance therapy, based on response at week six. Finally, at 22 weeks patients were scored using CDAI for their response to therapy. (B) Stool sampling, treatment, and response evalution timeline.

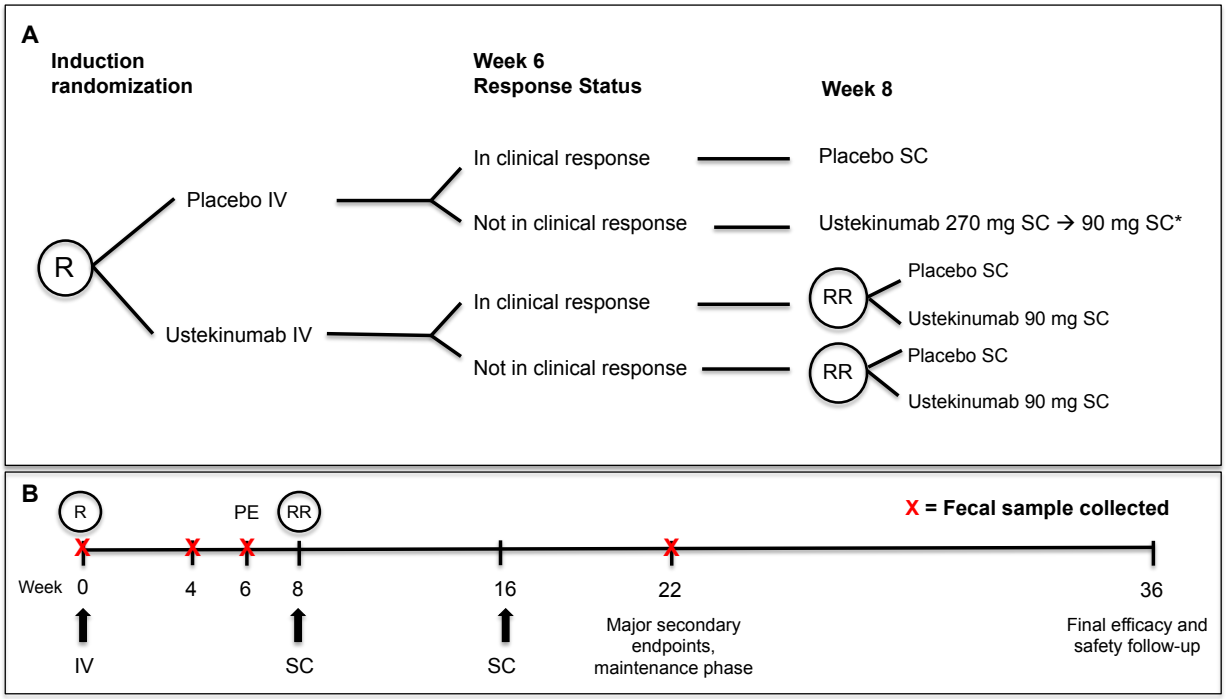


Figure 2: Change in alpha diversity over time by induction treatment and week twenty-two response status. The α-diversity of 48 patients induced and maintained with UST and 14 patients induced and maintained with placebo was assess at each time point. Friedman test were perfomed within each teatment and responder group. * indicates week twenty-two is significantly different from baseline ($p < 0.05$).

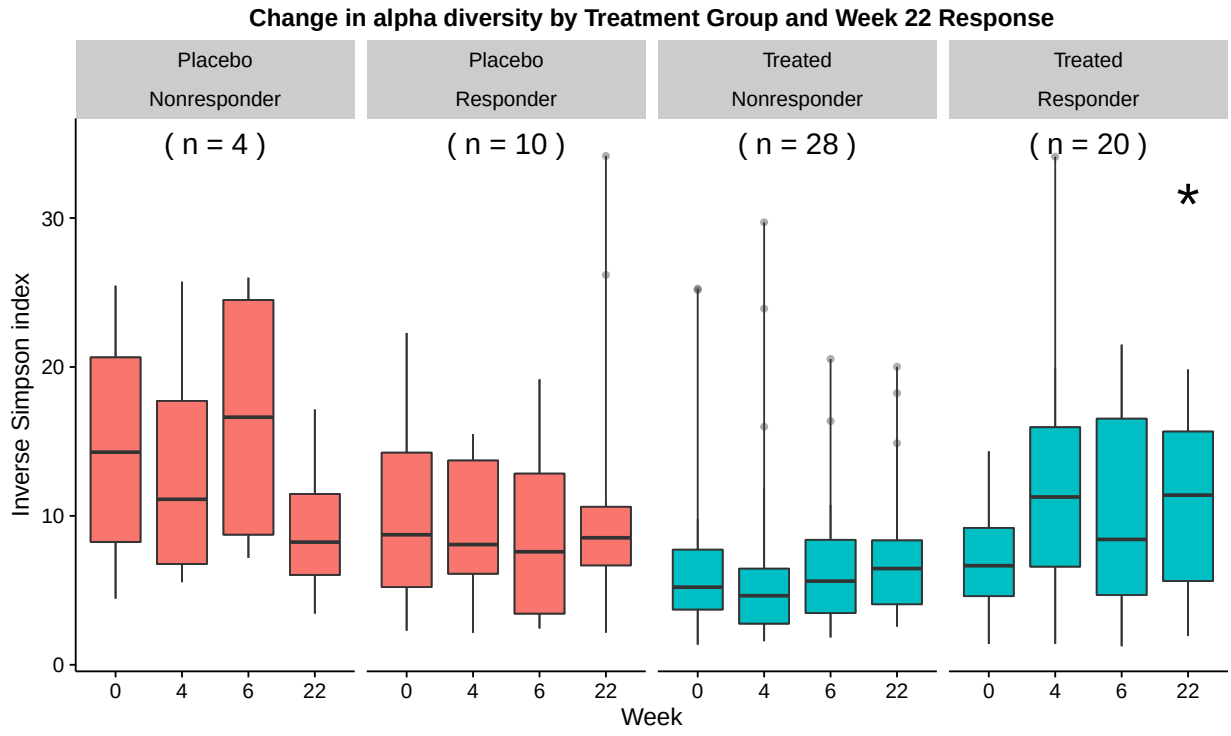
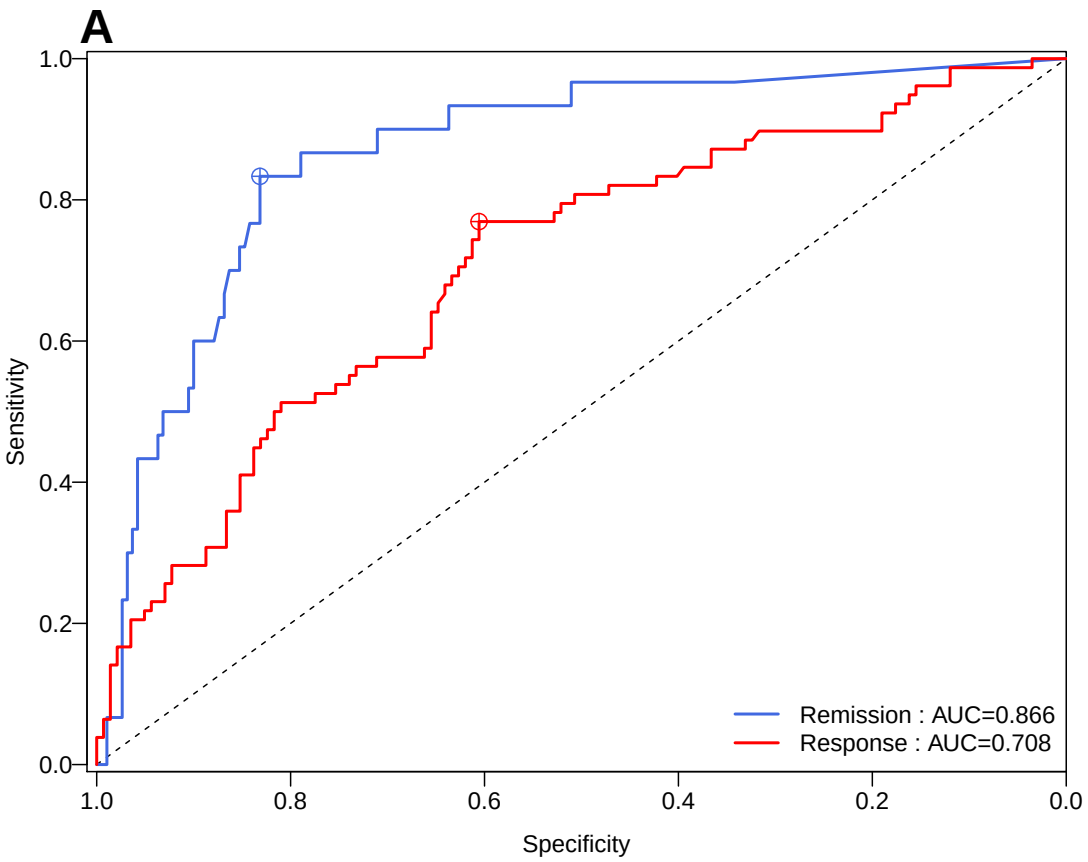
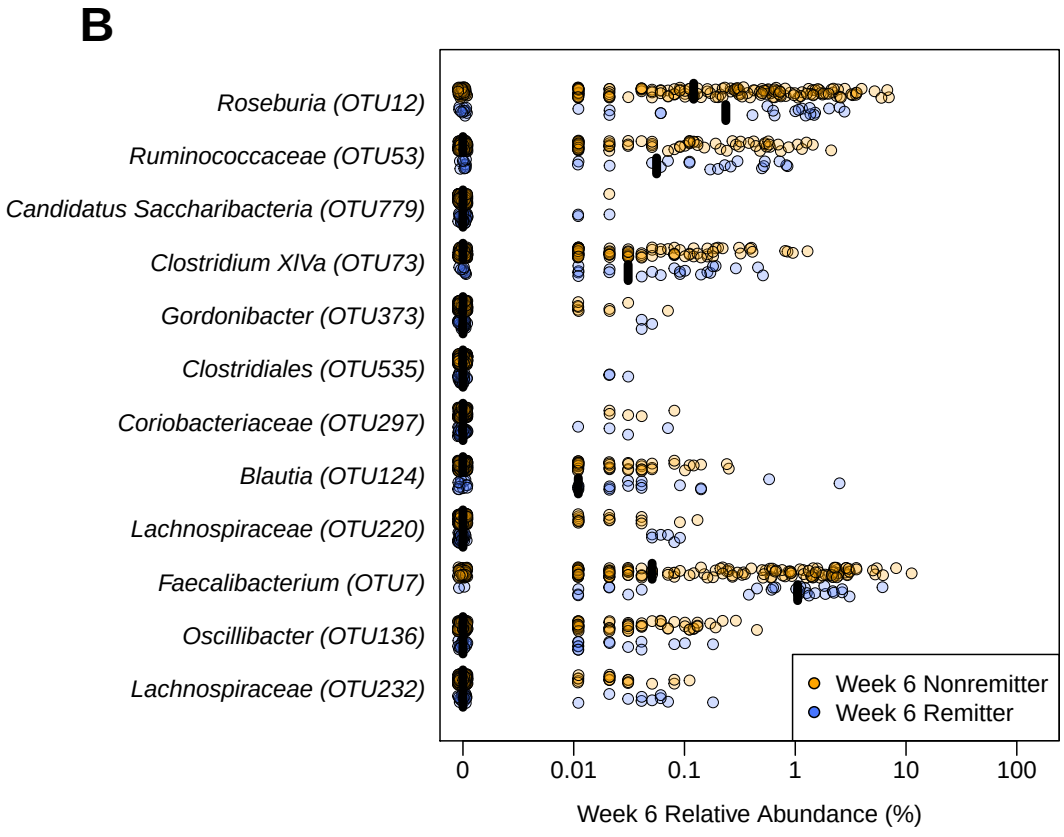


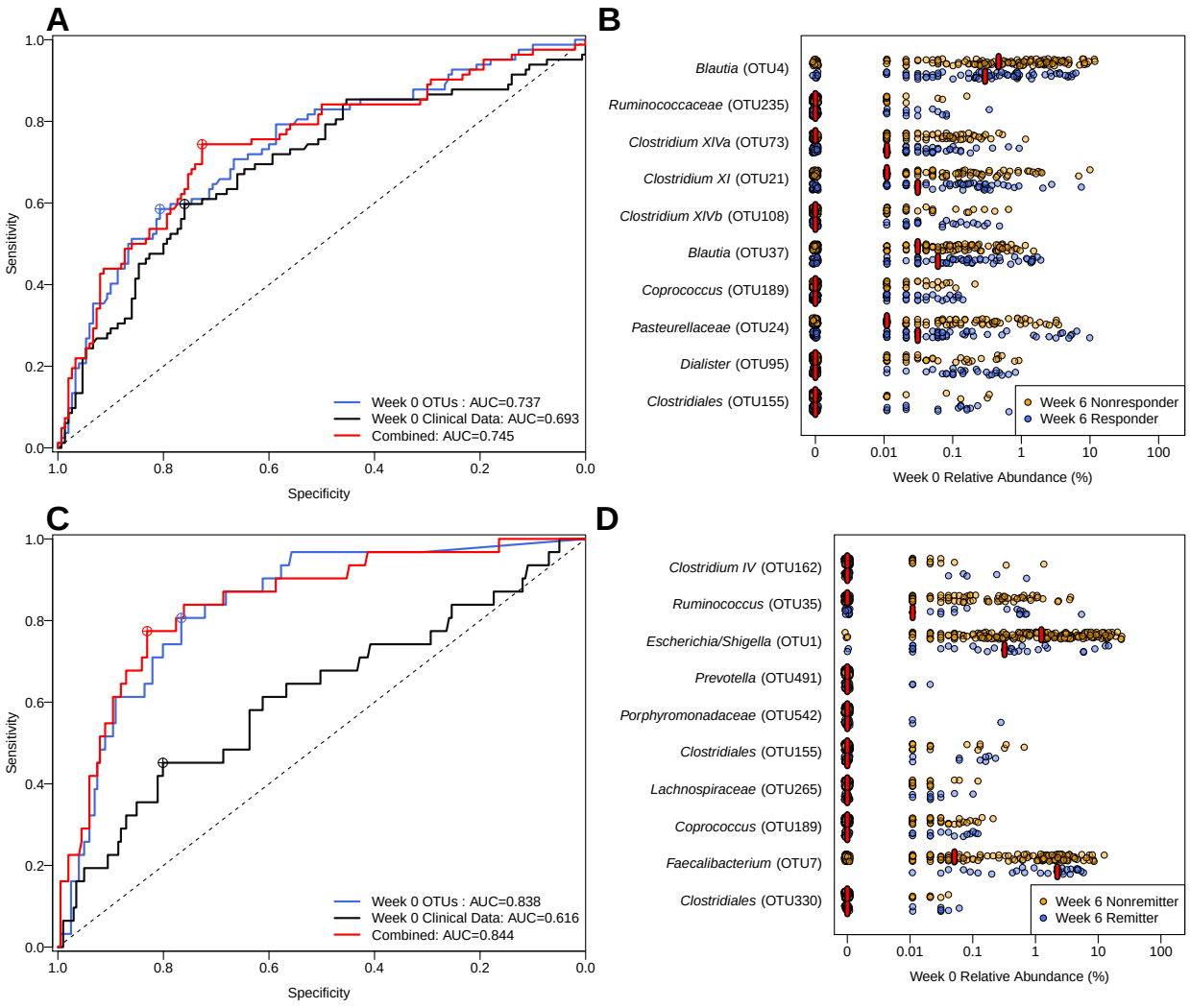
Figure 3: Classification of week six response or remission status using week six stool samples from patients treated with UST (A) ROCs for week six outcome based on the week six microbiome. (B) Top predictive taxa from week six stool for remission status at week six, based on MDA.


A
1.0
0.8
0.6
0.4
0.2
0.0
Sensitivity
1.0
0.8
0.6
0.4
0.2
0.0
Specificity
Remission : AUC=0.866
Response : AUC=0.708


B
Roseburia (OTU12)
Ruminococcaceae (OTU53)
Candidatus Saccharibacteria (OTU779)
Clostridium XlVa (OTU73)
Gordonibacter (OTU373)
Clostridiales (OTU535)
Coriobacteriaceae (OTU297)
Blautia (OTU124)
Lachnospiraceae (OTU220)
Faecalibacterium (OTU7)
Oscillibacter (OTU136)
Lachnospiraceae (OTU232)
Week 6 Nonremitter
Week 6 Remitter
0
0.01
0.1
1
10
100
Week 6 Relative Abundance (%)

350

Figure 4: Prediction of week six disease status in patients treated with UST, using baseline samples ROCs for (A) response and (C) remission using microbiota data, clinical metadata, and a combined model. Top predictive taxa for the microbiota model based on MDA for (B) response and (D) remission.



Supplemental Figure 1: Phyla from baseline stool samples in patients treated with UST by week six outcome Compared the relative abundance of each phylum in UST teated patients based on (A) response and (B) remission status using a Wilcoxon rank sum test and to identify phyla where there was a p-value less than 0.05 following a Benjamini-Hochberg correction for multiple comparisons.

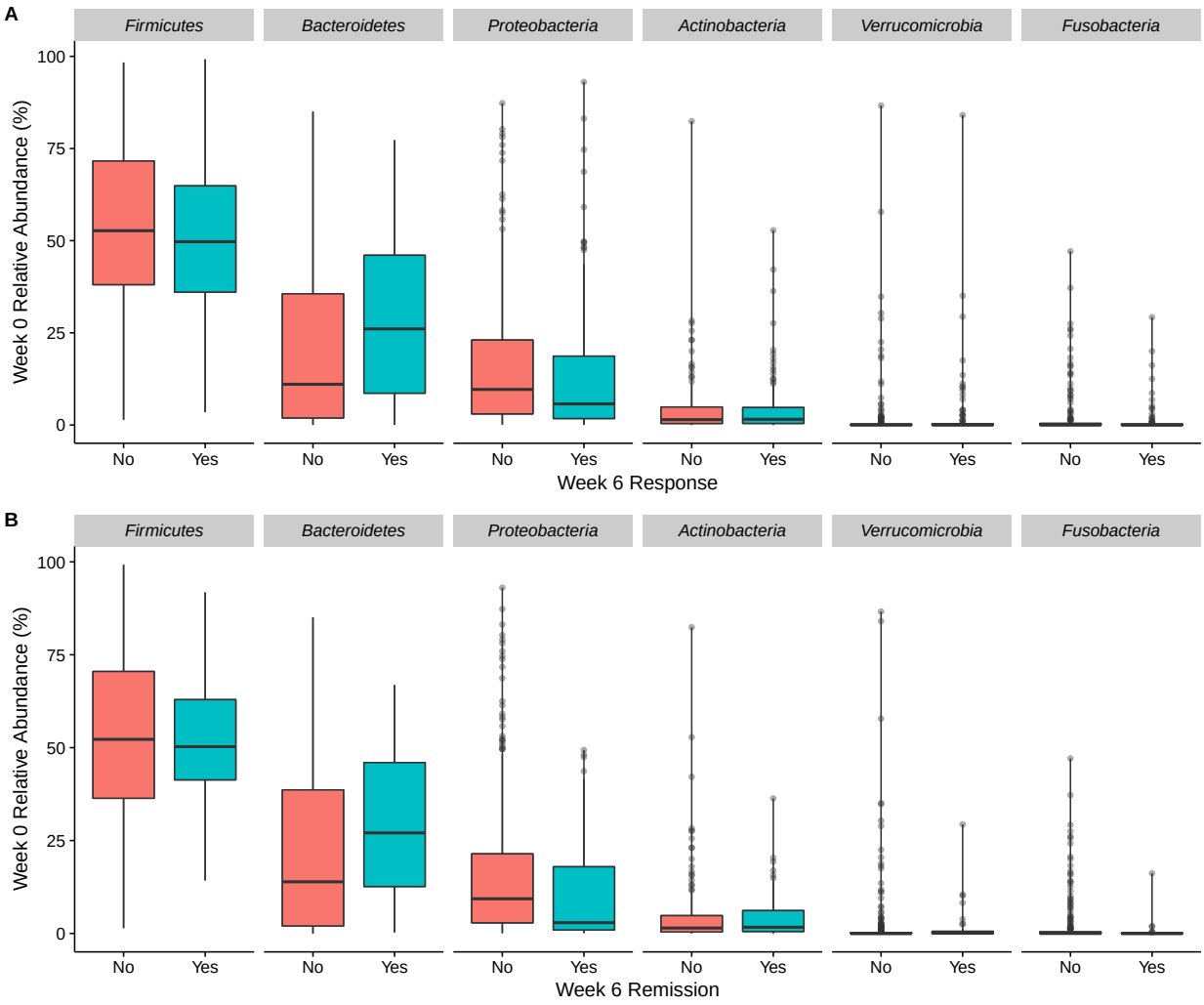
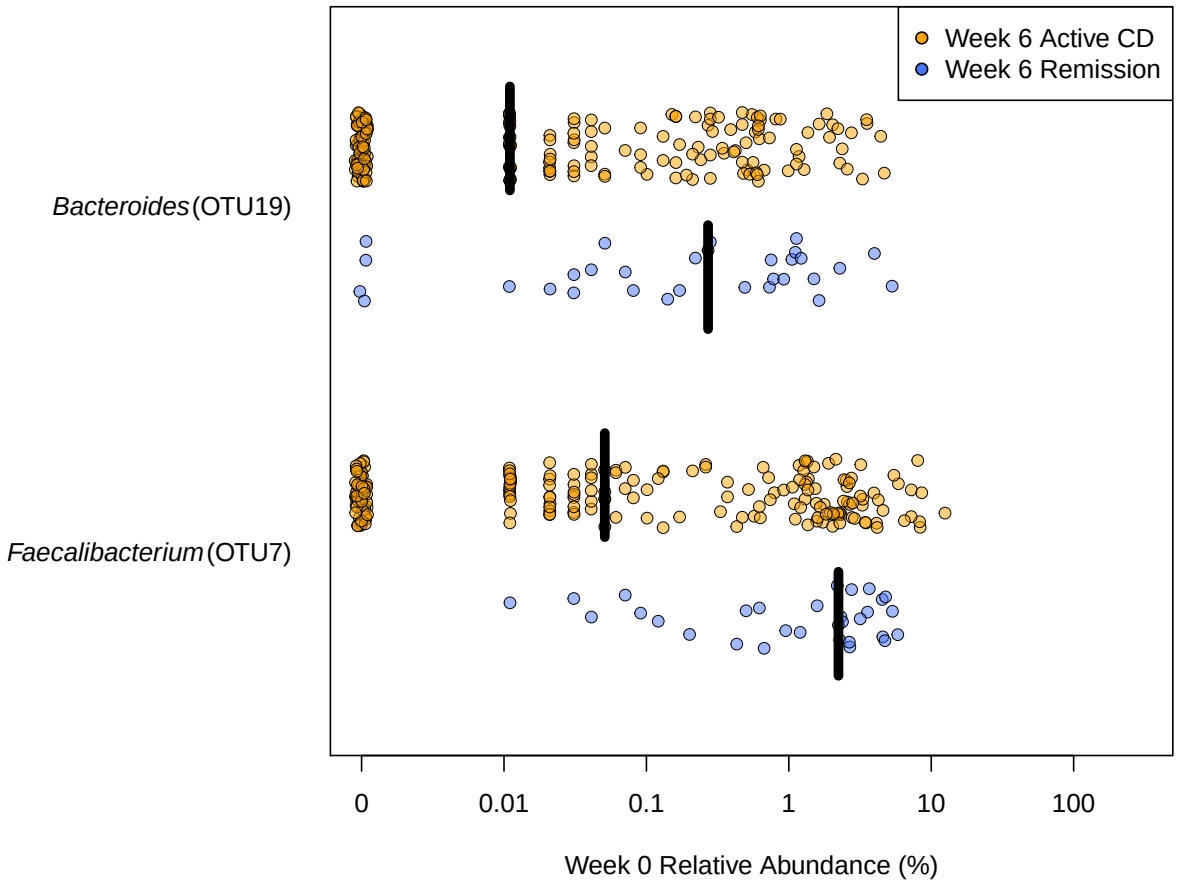


Figure 5: Differential taxa in baseline stool samples from patients treated with UST, based on week six remission status Compared the relative abundance of each OTU in UST teated patients based on (A) response and (B) remission status using a Wilcoxon rank sum test to identify OTUs where there was a p-value less than 0.05 following a Benjamini-Hochberg correction for multiple comparisons.



References

1. Huang S, Li R, Zeng X, He T, Zhao H, Chang A, Bo C, Chen J, Yang F, Knight R, Liu J, Davis C, Xu J. 2014. Predictive modeling of gingivitis severity and susceptibility via oral microbiota. ISME J 8:1768–80.

2. Wang Y, Ames NP, Tun HM, Tosh SM, Jones PJ, Khafipour E. 2016. High molecular weight barley -glucan alters gut microbiota toward reduced cardiovascular disease risk. Front Microbiol 7.

3. Seekatz AM, Rao K, Santhosh K, Young VB. 2016. Dynamics of the fecal microbiome in patients with recurrent and nonrecurrent clostridium difficile infection. Genome Med 8.

4. Zackular JP, Rogers MA, Ruffin MT th, Schloss PD. 2014. The human gut microbiome as a screening tool for colorectal cancer. Cancer Prev Res (Phila) 7:1112–21.

5. Baxter NT, Ruffin MT th, Rogers MA, Schloss PD. 2016. Microbiota-based model improves the sensitivity of fecal immunochemical test for detecting colonic lesions. Genome Med 8:37.

6. Klatt NR, Cheu R, Birse K, Zevin AS, Perner M, Noel-Romas L, Grobler A, Westmacott G, Xie IY, Butler J, Mansoor L, McKinnon LR, Passmore JS, Abdool Karim Q, Abdool Karim SS, Burgener AD. 2017. Vaginal bacteria modify hiv tenofovir microbicide efficacy in african women. Science 356:938–945.

7. Haiser HJ, Gootenberg DB, Chatman K, Sirasani G, Balskus EP, Turnbaugh PJ. 2013. Predicting and manipulating cardiac drug inactivation by the human gut bacterium eggerthella lenta. Science 341:295–8.

8. Sivan A, Corrales L, Hubert N, Williams JB, Aquino-Michaels K, Earley ZM, Benyamin FW, Lei YM, Jabri B, Alegre ML, Chang EB, Gajewski TF. 2015. Commensal bifidobacterium promotes antitumor immunity and facilitates anti-pd-l1 efficacy. Science 350:1084–9.

9. Vetizou M, Pitt JM, Daillere R, Lepage P, Waldschmitt N, Flament C, Rusakiewicz S, Routy B, Roberti MP, Duong CP, Poirier-Colame V, Roux A, Becharef S, Formenti S, Golden E, Cording S, Eberl G, Schlitzer A, Ginhoux F, Mani S, Yamazaki T, Jacquelot N, Enot DP, Berard M, Nigou J, Opolon P, Eggermont A, Woerther PL, Chachaty E, Chaput N, Robert C, Mateus C, Kroemer G, Raoult D, Boneca IG, Carbonnel F, Chamaillard M, Zitvogel L. 2015. Anticancer immunotherapy by ctla-4 blockade relies

392 on the gut microbiota. Science 350:1079–84.

393 10. Gevers D, Kugathasan S, Denson LA, Vazquez-Baeza Y, Van Treuren W, Ren B, Schwager E,
394 Knights D, Song SJ, Yassour M, Morgan XC, Kostic AD, Luo C, Gonzalez A, McDonald D, Haberman
395 Y, Walters T, Baker S, Rosh J, Stephens M, Heyman M, Markowitz J, Baldassano R, Griffiths A,
396 Sylvester F, Mack D, Kim S, Crandall W, Hyams J, Huttenhower C, Knight R, Xavier RJ. 2014. The
397 treatment-naive microbiome in new-onset crohn's disease. Cell Host Microbe 15:382–92.

398 11. Wang F, Kaplan JL, Gold BD, Bhasin MK, Ward NL, Kellermayer R, Kirschner BS, Heyman MB,
399 Dowd SE, Cox SB, Dogan H, Steven B, Ferry GD, Cohen SA, Baldassano RN, Moran CJ, Garnett
400 EA, Drake L, Otu HH, Mirny LA, Libermann TA, Winter HS, Korolev KS. 2016. Detecting microbial
401 dysbiosis associated with pediatric crohn disease despite the high variability of the gut microbiota. Cell
402 Rep.

403 12. Ananthakrishnan AN, Luo C, Yajnik V, Khalili H, Garber JJ, Stevens BW, Cleland T, Xavier RJ.
404 2017. Gut microbiome function predicts response to anti-integrin biologic therapy in inflammatory bowel
405 diseases. Cell Host Microbe 21:603–610.e3.

406 13. Shaw KA, Bertha M, Hofmekler T, Chopra P, Vatanen T, Srivatsa A, Prince J, Kumar A, Sauer C,
407 Zwick ME, Satten GA, Kostic AD, Mulle JG, Xavier RJ, Kugathasan S. 2016. Dysbiosis, inflammation,
408 and response to treatment: A longitudinal study of pediatric subjects with newly diagnosed inflammatory
409 bowel disease. Genome Med 8:75.

410 14. Ananthakrishnan AN. 2015. Epidemiology and risk factors for IBD. Nat Rev Gastroenterol Hepatol
411 12:205–217.

412 15. Floyd DN, Langham S, Severac HC, Levesque BG. 2015. The economic and quality-of-life burden
413 of crohn's disease in europe and the united states, 2000 to 2013: A systematic review. Dig Dis Sci
414 60:299–312.

415 16. Randall CW, Vizuete JA, Martinez N, Alvarez JJ, Garapati KV, Malakouti M, Taboada CM. 2015.
416 From historical perspectives to modern therapy: A review of current and future biological treatments
417 for crohn's disease. Therap Adv Gastroenterol 8:143–59.

418 17. Wils P, Bouhnik Y, Michetti P, Flourie B, Brixi H, Bourrier A, Allez M, Duclos B, Grimaud JC,

Buisson A, Amiot A, Fumery M, Roblin X, Peyrin-Biroulet L, Filippi J, Bouguen G, Abitbol V, Coffin B, Simon M, Laharie D, Pariente B. 2015. Subcutaneous ustekinumab provides clinical benefit for two-thirds of patients with crohn's disease refractory to anti-tumor necrosis factor agents. Clin Gastroenterol Hepatol.

18. Colombel JF, Reinisch W, Mantzaris GJ, Kornbluth A, Rutgeerts P, Tang KL, Oortwijn A, Bevelander GS, Cornillie FJ, Sandborn WJ. 2015. Randomised clinical trial: Deep remission in biologic and immunomodulator naive patients with crohn's disease - a SONIC post hoc analysis. Aliment Pharmacol Ther 41:734–46.

19. Baert F, Moortgat L, Van Assche G, Caenepeel P, Vergauwe P, De Vos M, Stokkers P, Hommes D, Rutgeerts P, Vermeire S, D'Haens G. 2010. Mucosal healing predicts sustained clinical remission in patients with early-stage crohn's disease. Gastroenterology 138:463–8; quiz e10–1.

20. Lichtenstein GR. 2010. Emerging prognostic markers to determine crohn's disease natural history and improve management strategies: A review of recent literature. Gastroenterol Hepatol (N Y) 6:99–107.

21. Chang S, Malter L, Hudesman D. 2015. Disease monitoring in inflammatory bowel disease. World J Gastroenterol 21:11246–59.

22. Boon GJ, Day AS, Mulder CJ, Gearry RB. 2015. Are faecal markers good indicators of mucosal healing in inflammatory bowel disease? World J Gastroenterol 21:11469–80.

23. Falvey JD, Hoskin T, Meijer B, Ashcroft A, Walmsley R, Day AS, Gearry RB. 2015. Disease activity assessment in ibd: Clinical indices and biomarkers fail to predict endoscopic remission. Inflamm Bowel Dis 21:824–31.

24. Sartor RB. 2006. Mechanisms of disease: Pathogenesis of crohn's disease and ulcerative colitis. Nat Clin Pract Gastroenterol Hepatol 3:390–407.

25. Wright EK, Kamm MA, Teo SM, Inouye M, Wagner J, Kirkwood CD. 2015. Recent advances in characterizing the gastrointestinal microbiome in crohn's disease: A systematic review. Inflamm Bowel Dis 21:1219–28.

26. Manichanh C, Rigottier-Gois L, Bonnaud E, Gloux K, Pelletier E, Frangeul L, Nalin R, Jarrin C,

446 Chardon P, Marteau P, Roca J, Dore J. 2006. Reduced diversity of faecal microbiota in crohn's disease
447 revealed by a metagenomic approach. Gut 55:205–11.

448 27. Hansen R, Russell RK, Reiff C, Louis P, McIntosh F, Berry SH, Mukhopadhya I, Bisset WM,
449 Barclay AR, Bishop J, Flynn DM, McGrogan P, Loganathan S, Mahdi G, Flint HJ, El-Omar EM, Hold
450 GL. 2012. Microbiota of de-novo pediatric IBD: Increased faecalibacterium prausnitzii and reduced
451 bacterial diversity in crohn's but not in ulcerative colitis. Am J Gastroenterol 107:1913–22.

452 28. Haberman Y, Tickle TL, Dexheimer PJ, Kim MO, Tang D, Karns R, Baldassano RN, Noe JD,
453 Rosh J, Markowitz J, Heyman MB, Griffiths AM, Crandall WV, Mack DR, Baker SS, Huttenhower C,
454 Keljo DJ, Hyams JS, Kugathasan S, Walters TD, Aronow B, Xavier RJ, Gevers D, Denson LA. 2014.
455 Pediatric crohn disease patients exhibit specific ileal transcriptome and microbiome signature. J Clin
456 Invest 124:3617–33.

457 29. Riol-Blanco L, Lazarevic V, Awasthi A, Mitsdoerffer M, Wilson BS, Croxford A, Waisman A, Kuchroo
458 VK, Glimcher LH, Oukka M. 2010. IL-23 receptor regulates unconventional il-17-producing t cells that
459 control infection1. J Immunol 184:1710–20.

460 30. Round JL, Mazmanian SK. 2009. The gut microbiome shapes intestinal immune responses during
461 health and disease. Nat Rev Immunol 9:313–23.

462 31. Eken A, Singh AK, Oukka M. 2014. INTERLEUKIN 23 in crohn'S disease. Inflamm Bowel Dis
463 20:587–95.

464 32. Shih VFS, Cox J, Kljavin NM, Dengler HS, Reichelt M, Kumar P, Rangell L, Kolls JK, Diehl
465 L, Ouyang W, Ghilardi N. 2014. Homeostatic il-23 receptor signaling limits th17 response through
466 il-22–mediated containment of commensal microbiota. Proc Natl Acad Sci U S A 111:13942–7.

467 33. Tedjo DI, Smolinska A, Savelkoul PH, Masclee AA, Schooten FJ van, Pierik MJ, Penders J, Jonkers
468 DMAE. 2016. The fecal microbiota as a biomarker for disease activity in crohn's disease. Scientific
469 Reports, Published online: 13 October 2016; doi:101038/srep35216.

470 34. Sandborn WJ, Gasink C, Gao LL, Blank MA, Johanns J, Guzzo C, Sands BE, Hanauer SB, Targan
471 S, Rutgeerts P, Ghosh S, Villiers WJ de, Panaccione R, Greenberg G, Schreiber S, Lichtiger S, Feagan
472 BG. 2012. Ustekinumab induction and maintenance therapy in refractory crohn's disease. N Engl J

Med 367:1519–28.

35. Sandborn WJ, Feagan BG, Fedorak RN, Scherl E, Fleisher MR, Katz S, Johanns J, Blank M, Rutgeerts P. 2008. A randomized trial of ustekinumab, a human interleukin-12/23 monoclonal antibody, in patients with moderate-to-severe crohn's disease. Gastroenterology 135:1130–41.

36. Kopylov U, Afif W, Cohen A, Bitton A, Wild G, Bessissow T, Wyse J, Al-Taweel T, Szilagyi A, Seidman E. 2014. Subcutaneous ustekinumab for the treatment of anti-TNF resistant crohn's disease–the McGill experience. J Crohns Colitis 8:1516–22.

37. Peyrin-Biroulet L, Panes J, Sandborn WJ, Vermeire S, Danese S, Feagan BG, Colombel JF, Hanauer SB, Rycroft B. 2016. Defining disease severity in inflammatory bowel diseases: Current and future directions. Clin Gastroenterol Hepatol 14:348–354.e17.

38. Best WR, Becktel JM, Singleton JW, Kern J F. 1976. Development of a crohn's disease activity index. national cooperative crohn's disease study. Gastroenterology 70:439–44.

39. Calle ML, Urrea V, Boulesteix A-L, Malats N. 2011. AUC-RF: A new strategy for genomic profiling with random forest. Human Heredity 72:121–132.

40. Naftali T, Reshef L, Kovacs A, Porat R, Amir I, Konikoff FM, Gophna U. 2016. Distinct microbiotas are associated with ileum-restricted and colon-involving crohn's disease. Inflamm Bowel Dis 22:293–302.

41. Sartor RB, Wu GD. 2016. Roles for intestinal bacteria, viruses, and fungi in pathogenesis of inflammatory bowel diseases and therapeutic approaches. Gastroenterology.

42. Boon GJ, Day AS, Mulder CJ, Gearry RB. 2015. Are faecal markers good indicators of mucosal healing in inflammatory bowel disease? World J Gastroenterol 21:11469–80.

43. Chang S, Malter L, Hudesman D. 2015. Disease monitoring in inflammatory bowel disease. World J Gastroenterol 21:11246–59.

44. Papa E, Docktor M, Smillie C, Weber S, Preheim SP, Gevers D, Giannoukos G, Ciulla D, Tabbaa D, Ingram J, Schauer DB, Ward DV, Korzenik JR, Xavier RJ, Bousvaros A, Alm EJ. 2012. Non-invasive mapping of the gastrointestinal microbiota identifies children with inflammatory bowel disease. PLoS

One 7:e39242.

45. Vandeputte D, Falony G, Vieira-Silva S, Tito RY, Joossens M, Raes J. 2016. Original article: Stool consistency is strongly associated with gut microbiota richness and composition, enterotypes and bacterial growth rates. Gut 65:57–62.

46. Huang EY, Inoue T, Leone VA, Dalal S, Touw K, Wang Y, Musch MW, Theriault B, Higuchi K, Donovan S, Gilbert J, Chang EB. 2015. Using corticosteroids to reshape the gut microbiome: Implications for inflammatory bowel diseases. Inflamm Bowel Dis 21:963–72.

47. Monteleone G, Neurath MF, Ardizzone S, Di Sabatino A, Fantini MC, Castiglione F, Scribano ML, Armuzzi A, Caprioli F, Sturniolo GC, Rogai F, Vecchi M, Atreya R, Bossa F, Onali S, Fichera M, Corazza GR, Biancone L, Savarino V, Pica R, Orlando A, Pallone F. 2015. Mongersen, an oral SMAD7 antisense oligonucleotide, and crohn's disease. N Engl J Med 372:1104–13.

48. Monteleone G, Di Sabatino A, Ardizzone S, Pallone F, Usiskin K, Zhan X, Rossiter G, Neurath MF. 2016. Impact of patient characteristics on the clinical efficacy of mongersen (GED-0301), an oral smad7 antisense oligonucleotide, in active crohn's disease. Aliment Pharmacol Ther 43:717–24.

49. Ardizzone S, Bevivino G, Monteleone G. 2016. Mongersen, an oral smad7 antisense oligonucleotide, in patients with active crohn's disease. Therap Adv Gastroenterol 9:527–32.

50. Orava EW, Jarvik N, Shek YL, Sidhu SS, Gariepy J. 2013. A short DNA aptamer that recognizes TNFalpha and blocks its activity in vitro. ACS Chem Biol 8:170–8.

51. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Van Horn DJ, Weber CF. 2009. Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl Environ Microbiol 75:7537–41.

52. Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD. 2013. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the miseq illumina sequencing platform. Appl Environ Microbiol 79:5112–20.

53. Schloss PD, Gevers D, Westcott SL. 2011. Reducing the effects of PCR amplification and sequencing

524 artifacts on 16S rRNA-based studies. PLoS One 6:e27310.

525 54. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO. 2013. The
526 silva ribosomal rna gene database project: Improved data processing and web-based tools. Nucleic Acids
527 Res 41:D590–6.

528 55. Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R. 2011. UCHIME improves sensitivity and
529 speed of chimera detection. Bioinformatics 27:2194–200.

530 56. Schloss PD, Westcott SL. 2011. Assessing and improving methods used in operational taxonomic
531 unit-based approaches for 16S rRNA gene sequence analysis. Appl Environ Microbiol 77:3219–26.

532 57. Wang Q, Garrity GM, Tiedje JM, Cole JR. 2007. Naive bayesian classifier for rapid assignment of
533 rRNA sequences into the new bacterial taxonomy. Appl Environ Microbiol 73:5261–7.

534 58. R Core Team. 2016. R: A language and environment for statistical computing. R Foundation for
535 Statistical Computing, Vienna, Austria.

536 59. Sokal RR, Rohlf FJ. 1995. Biometry: The principles and practice of statistics in biological research,
537 3rd ed. Freeman, New York.

538 60. Magurran AE. 2004. Measuring biological diversity. Blackwell Pub., Malden, Ma.

539 61. Yue JC, Clayton MK. 2005. A similarity measure based on species proportions. Communications in
540 Statistics-Theory and Methods 34:2123–2131.

541 62. Schloss PD. 2008. Evaluating different approaches that test whether microbial communities have
542 the same structure. ISME J 2:265–75.

543 63. Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara
544 RB, Simpson GL, Solymos P, Stevens MHH, Szoecs E, Wagner H. 2016. Vegan: Community ecology
545 package. r package version 2.4-1.

546 64. Friedman M. 1937. The use of ranks to avoid the assumption of normality implicit in the analysis
547 of variance. Journal of the American Statistical Association 32:675–701.

548 65. Giraudoux P. 2016. Pgirmess: Data analysis in ecology.

549 66. Urrea V, Calle M. 2012. AUCRF: Variable selection with random forest and the area under the

curve.

67. Breiman L. 2001. Random forests. Machine Learning 45:5–32.

68. Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: A practical and powerful approach to multiple testing. Journal of the Royal Statistical Society Series B (Methodological) 57:289–300.

69. Wickham H. 2009. Ggplot2: Elegant graphics for data analysis. Springer-Verlag New York.

70. Wickham H, Francois R. 2016. Dplyr: A grammar of data manipulation.

71. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J-C, Müller M. 2011. PROC: An open-source package for r and s+ to analyze and compare roc curves. BMC Bioinformatics 12:77.

72. Xie Y. 2014. Knitr: A comprehensive tool for reproducible research in R. *In* Stodden, V, Leisch, F, Peng, RD (eds.), Implementing reproducible computational research. Chapman; Hall/CRC.

73. Xie Y. 2015. Dynamic documents with R and knitr, 2nd ed. Chapman; Hall/CRC, Boca Raton, Florida.

74. Xie Y. 2017. Knitr: A general-purpose package for dynamic report generation in r.

75. Auguie B. 2016. GridExtra: Miscellaneous functions for "grid" graphics.

76. Wickham H, Chang W. 2016. Devtools: Tools to make developing r packages easier.

77. Boettiger C. 2015. Knitcitations: Citations for 'knitr' markdown files.

78. Wickham H. 2016. Scales: Scale functions for visualization.

79. Wickham H. 2017. Tidyr: Easily tidy data with 'spread()' and 'gather()' functions.

80. Harrell Jr FE, Charles Dupont, others. 2016. Hmisc: Harrell miscellaneous.

81. Wilke CO. 2016. Cowplot: Streamlined plot theme and plot annotations for 'ggplot2'.