The fecal microbiome as a tool for monitoring and predicting response outcomes in Ustekinumab-treated, anti-TNF-alpha refractory Crohn's Disease patients.

	Clinical.Variable	Summary	Species.Richness..Alpha.diversity.	Community.Structure..beta.diversity.
1	CDAI	rho = -0.2	0.014	0.324
2	Loose Stool Frequency (per week)	rho = -0.2	0.003	0.024
3	C-Reactive Protein (mg/L serum)	rho = 0.06	0.394	0.033
4	Fecal Calprotectin (µg/g)	rho = 0.08	0.254	0.006
5	Fecal Lactoferrin (µg/g)	rho = 0.1	0.07	0.004
6	BMI	rho = 0.07	0.299	0.277
7	Weight (kg)	rho = 0.07	0.299	0.112
8	Age (years)	rho = -0.05	0.472	0.033
9	Sex (F/M)	-	0.539	0.277
10	Corticosteroid Use (Y/N)	-		
11	Disease Duration (years)	rho = -0.2		
12	Tissue Involvement	-		

```
## 10                                0.001                                0.01
## 11                                0.001                               0.004
## 12                                 0.19                               0.004
```

```
## [1] 0.014
## Levels: 0.001 0.003 0.014 0.07 0.19 0.254 0.299 0.394 0.472 0.539
```

Running title: Microbiome 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

Abstract: Crohn's disease (CD) is a global health issue characterized by patches of ulceration and inflammation along the gastrointestinal tract, as well as reduced gut microbial diversity. We investigated the association between the fecal microbiome and clinical phenotypes of subjects with moderate to severe CD, refractory to anti-TNFa, treated with Ustekinumab (UST). We hypothesized that the fecal microbiome would be different between treatment response groups. Stool samples from 400 patients taking part in a double-blinded, placebo-controlled, Phase 2b clinical trial were obtained over the course of 22 weeks. The V4 region of the 16S rRNA gene was amplified and sequenced to determine the structure of the fecal bacterial communities.

The microbiome of UST treated clinical responders changed over time, in contrast to nonresponsive subjects. Using Random Forest models, the differences in the fecal microbiome following treatment could classify patients in remission from those with active disease. Additionally, the baseline microbiome and clinical metadata could effectively classify therapeutic outcome, especially for remission. Baseline fecal microbiome community structures and α-diversity were significantly different based on the outcome of UST treatment. *Faecalibacterium* and *Bacteroidies* were significantly more abundant in remitters. Additionally, the fecal microbial diversity at baseline was associated with markers for disease severity, such as Crohn's Disease Activity Index (CDAI), stool frequency, CRP, fecal lactoferrin, fecal calprotectin, corticosteroid use, disease duration, and tissue involvement.

Importance: Finding biomarkers that give clinicians the ability to predict potential treatment outcomes at diagnosis will increase the likelihood of faster induction and maintainance of remission due to more personalized treament. The fecal microbiome could be a useful biomarker for directing or monitoring the treatment of CD patients. The observed baseline differences in fecal microbiota and changes due to therapeutic response will allow further investigation into the microbes and/or the metabolic functions important in establishing and maintaining CD remission. Finally, the OTUs associated with remission following treatment induction, *Faecalibacterium* and *Bacteroidites*, could be biomarkers for successful UST treatment of TNF-α refractory CD patients.

Keywords: Crohn's Disease, fecal microbiome, biologics, prediction

Introduction

Crohn's disease (CD), an incurable inflammatory bowel disease (IBD), is a global health issue causing large economic and healthcare utilization impacts on society (1–3). 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 (4, 5). Faster induction of remission following diagnosis reduces the risk of irreversible intestinal damage and disability (5–7). Ideally, clinicians would be able to determine personalized treatment options for CD patients at diagnosis that would result in faster achievement of remission [cites]. Therefore, recent research has been focused on identifying noninvasive, prognostic biomarkers to monitor CD severity and predict therapuetic response [cites].

The precise etiology of CD remains unknown, but host genetics, environmental exposure, and the gut microbiome appear to be involved (1, 8). 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 (9–13). Additionally, genome-wide association studies of individuals with CD identified several susceptibility loci, including loci involved in the IL-23 signialing pathway, which could impact the gut microbiome structure and function (4, 9). If the fecal microbiome 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.

The microbiome has been correlated with a variety of diseases and has shown promise as a predictive tool for disease outcome for gingivitis(14), cardiovascular disease(15), *Clostridium difficile* infection (16–18), and colorectal cancer (19, 20). In relation to IBD, previous studies have shown that the gut microbiome correlates with disease severity in new-onset, pediatric CD patients (13, 21). Additionally, recent studies have shown promise for the microbiome as it relates to IBD and therapeutic response (22). It remains to be determined, however, whether the fecal microbiome can predict and monitor response to therapy in CD (9).

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 (5, 23–25). Given the potential impact of IL-23 on the microbiome [cites], we hypothesized that UST treatment may alter the fecal microbiome and that response to UST could be predicted or influenced by differences in patients' gut microbiota. We analyzed the fecal microbiomes of individuals who participated in a double-blinded, placebo-controlled Phase II clinical trial of UST in treating CD (23). Using 16S rRNA gene sequence data from these patients' stool samples, we determined associations between clinical metadata, disease severity, and the fecal microbiome. We also tested whether the microbiome changed in subjects with UST and if clinical responders had a microbiome that is distinct from non-responders. Our study demonstrates that the fecal microbiome is associated with baseline clinical metadata and that these associations are useful in predicting and monitoring treatment outcome.

Results

Characteristics of the study population

We characterized the fecal microbiota in a subset of TNF-α refractory CD patients, with moderate to severe CD, who took part in the double-blinded, CERTIFI clinical trial (23). 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). Subjects provided stool samples at screening (week 0), week 4, week 6, and week 22 post induction for analysis using 16S rRNA gene sequencing (Figure 1B).

Following sequence curation using the mothur software package (26), 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 subjects 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 4, 289 at week 6, and 205 at week 22 post-treatment, for a total of 1,058 samples.

The microbiome changes in UST responders

Given the potential impact of IL-23 on the microbiome [cites], we hypothesized that UST treatment

would alter the fecal microbiome. The effects of biologic treatment of IBD on the microbiome are not yet well described, but are hypothesized to be indirect, as these drugs act on host factors [cites]. We tested whether treatment with UST alters the microbiome by performing a Freidman test comparing α-diversity at each sample time within each treatment group based on their week 22 response status. We included 248 subjects induced and maintainted with UST (168 responders, 112 nonresponders) and 248(40 responders, 16 nonresponders) (Figure 1A), who provided samples at every timepoint (Figure 1B). As shown in Figure 4, we saw no significant difference in α-diversity over time in subjects who did not respond at week 22, regardless of treatment. However, in UST treated-week 22 responders α-diversity changed over time ($p = 0.00486$). Multiple comparisons following the freidman test showed the inverse Simspon at week 22 was significantly higher than week 0 ($p < 0.05$). No change was observed in subjects induced and maintained with placebo who responded at week 22.

The microbiome following treatment can distinguish between treatment outcomes

Having observed that the microbiome changes in subjects who responded to treatment, we hypothesized that we could use the fecal microbiome to distinguish between subjects who responsded to treatment from those who did not respond. A prior study demonstrated a link between the microbiome and disease activity, where specific microbes were associated with remission compared to active CD (27). We hypothesized that the microbiome could be used to monitor response to therapy in a similar manner. We used the AUCRF package (version 11) in R to develop a random forest classification model to distinguish between subjects by treatment outcome, based on the relative abundance of fecal microbiome community members (20, 28). The study design resulted in only 75 week 22 stool samples from subjects induced and maintained with UST, so we focused our analysis on the 220 week 6 stool samples and outcomes in subjects treated with UST. As seen in Figure 5, using stool samples from week 6 the model could classify responders from nonresponders at week 6 with an AUC of 0.708 (sensitivity =0.769, specificity = 0.606). For remission at week 6 the model had an AUC of 0.866 (sensitivity = 0.833, specificity = 0.832) using week 6 samples. We were better able to distinguish remitters from non-remitters than responders from non-responders. We hypothesize that this is due to the relative nature of the response criteria compared to the threshold used to determine remission status.

Prediction of remission based on the microbiome at screening

Having demonstrated that the microbiome following treatement could classify treatment outcomes, we hypothesized that the week 0 fecal microbiome could predict response to therapy. To test this hypothesis, we again used the AUCRF package in R to develop a random forest classification model to classify week 6 responders from non-responders, as well as week 6 remitters from non-remitters, based on the relative abundance of fecal microbiome community members at week 0, clinical metadata at week 0, and the combination of microbiome and clinical data (20, 28). We ran these models on 232 week 0 stool sampls from subjects induced with UST. The optimal models for response and remission at the primary endpoint (Week 6) are shown in *Figure 6A and C*. Using only clinical metadata, we achieved an AUC of 0.693, a specificity of 0.76, and a sensitivity of 0.598. Using only microbiome data, the model predicted response with an AUC of 0.737 with a specificity of 0.807 and a sensitivity of 0.585. When combining clinical metadata with the microbiome, the model predicted response with an AUC of 0.745, a specificity of 0.727, and a sensitivity of 0.744. With respect to Week 6 remission, using solely clinical metadata we achieved AUC of 0.616 with a specificity of 0.801 and a sensitivity of 0.452. Using only fecal microbiome data we achieved an AUC of 0.838 with a specificity of 0.766 and a sensitivity of 0.806. When combining clinical metadata with the microbiome, we achieved an AUC of 0.844 with a specificity of 0.831 and a sensitivity of 0.774. Across all weeks and responses, prediction with clinical metadata alone did not perform as well as models using the fecal microbiome at screening. Also, combining microbiome data with clinical metadata did not consistently improve prediction compared to microbiome data alone.

Additionally we found several OTUs occurred frequently across models including *Faecalibacterium*, among other taxa that were more abundant in responders/remitters (Figure 6B and D).

Comparison of clinical responders and non-responders

Given the observed differences in the fecal microbiome at baseline and week 6 in responders/remitters compared to non-responders/non-remitters, We hypothesized that there are associations between the microbiome at baseline and treatment outcome. To test this, we compared the week 0 microbiomes of subjects based on treatment group and outcome status at week 6 and week 22. Outcome status was broken into 2 catagories; response and remission. Response is a relative value defined as a decrease in a subject's initial CDAI of 30% or more, while remission is defined as a CDAI below the threshold of 150. For week 22 analysis, subjects who changed treatment for maintainence therapy were not included in our analysis. This resulted in 120 subjects induced and maintained with UST and 25 subjects induced

and maintained with placebo included in our week 22 analysis. Week 6 analysis compared to the full 306 subjects with screening samples.

With respect to α-diversity, subjects induced with UST and in remission at week 6 were significantly different from non-remitters treated with UST, having higher diversity based on inverse Simpson (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. Beta-diversity was significantly different for each outcome status (response/remission) and treatment group at each clinical endpoint (week 6 response p = 0.012, week 6 remission p = 0.017, week 22 response p = 0.012, week 22 remission p = 0.012), as seen in Table 2. No phyla were significantly different by treatment and response, however *Fusobacteria* was less frequently observed in week 6 remitters than non-remitters treated with UST (median relative abundance = 0 (IQR = -) and 0.0333 (IQR = -), respectively).

As seen in Figure 3, two taxa were significantly more abundant in UST-induced, week 6 remitters compared to non-remitters; *Bacteroides* (OTU0019) (p = 0.022) and *Faecalibacterium* (OTU0007) (p = 0.0026).

No individual taxa were significantly different among UST induced subjects at week 22, or those receiving placebo for induction, regardless of response/remission status at week 6 and 22.

The baseline microbiome associates with clinical variables

We hypothesized that there were associations between the microbiome and clinical variables at baseline related to disease severity in this unique cohort. To test this hypothesis, we compared the week 0 microbiome with clinical data at week 0 (Supplemental Table 1). We compared α-diveristy at baseline to clinical variables using the inverse Simpson index with the Spearman correlation, wilcoxon, or kurskal-wallis tests to compare groups. We compared β-diversity with a PERMANOVA using the adonis function in the vegan R package. Following multiple comparison correction, we observed small, but significant correlations for lower α-diversity correlating with higher CDAI (rho = -0.161, p = 0.014), higher frequency of loose stools per week (rho = -0.193, p = 0.003), and longer disease duration (rho = -0.225, p = 0.001), with lower diversity corresponding to longer disease. 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 =), and fecal lactoferrin (p =). The β-diversity was also significantly different based on weekly loose stool frequency (p=), age (p =), the tissue affected (p = 0.004), corticosteroid use β-diversity (p =0.01) and disease duration (p = 0.004). No significant differences in the microbiome were observed for BMI, weight, or sex.

Discussion

With this study we sought to gain a more detailed understanding of if and how biologic treatment affects the microbiome, to determine whether the microbiome can be used to identify patients who will respond to therapy, and to gain a better understanding of the interaction between the human gut microbiome and CD pathogenesis in adult patients. We found the fecal microbiome to be useful in uncovering associations between the microbiome and aspects of CD severity metrics and treatment outcomes. We also demonstrated that the microbiome of treated responders changed over time, though it is not yet possible to determine any direct effects of treatment on the microbiome. Finally, we were able to show that the microbiome could be useful in predicting response to therapy, especially clinical remission, compared to clinical metadata alone in our unique patient cohort.

We observed several associations between the microbiome and clinical variables that could play a role in how CD is monitored and treated in the future. Given that serum CRP, fecal calprotectin, and fecal lactoferrin are used as biomarkers to measure intestinal inflammation and CD severity, the observation that the microbial community structure is different among patients based on these markers supports the hypothesis that the microbiome could function as a biomarker for measuring disease activity in patients, especially in concert with these established inflammatory biomarkers (27, 29, 30). Higher CDAI was associated with lower microbial diversity. This appears to be consistent with other studies on the microbiome in individuals with CD compared to healthy individuals and studies looking at active disease compared to remission (13, 27, 31). However, these differences may have been driven by weekly stool frequency, one component of the CDAI, where higher stool frequency is also negatively associated with microbial diversity. This finding is consistent with the association between loose stools and lower diversity (32). We also observed differences in the microbial community structure based on disease localization. These results are consistent with a study by Naftali et al finding distinct microbiotas for ileal versus colonic CD using mucosal tissue (33). Our study also found that corticosteroid use

impacts the composition of the human fecal microbiome. This supports data seen in the mouse model where corticosteroid injections altered the fecal mouse microbiome (34). 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 (35–37). Hypothetically, earlier biologic intervention could 'preserve' a more diverse microbiome that promotes remission and reduces 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 reduce this cost and make earlier intervention possible. Aptamers are short strands of DNA or RNA capable of specifically binding small molecules, proteins, and whole cells. Anti-TNF aptamers have been published that could potentially be used to test this in the mouse model (38).

An important question for the microbiome and IBD is whether or not the microbiome is affected by treatment with biologics. This study attempted to answer that question by looking at the microbiome of our CD subjects across multiple time points during treatment. While we were unable to see direct effects of the drug on the fecal microbiome, we observed that the microbiome of clinical responders changed over time, in contrast to nonresponsive subjects. This was observed for responsive patients regardless of induction treatment, leading us to think we are observing the effects of change in disease activity and health, leading to lower inflammation, rather than any effects from treatment. This interpretation is consistent with studies using the microbiome to distinguish between remission and active CD (27). We did however observe a significant difference in community structure based on treatment and cannot eliminate the possibility of a direct effect on the microbiome in treated responders, however the change in community structure observe in responders treated with placebo supports the hypothesis that the change in community structure reflects a change in inflammation.

Another important question in for the importance of the microbiome in IBD is whether response to therapy can be predicted with the microbiome. We attempted to address this by developing a random-forest model that used relative microbial abundance data and/or clinical metadata for input. We

found we were better able to predict remission status compared to response status. Response may be less predictable due to the "floating target" nature of a relative decrease (>30% decrease) in CDAI compared to the hard threshold for remission (CDAI<150). We were also better able to distinguish remission/non-remission than response/non-response, using samples provided 6 weeks after treatment induction. This is consistent with other studies, again suggesting the microbiome could be useful as a biomarker in detecting remission versus active disease (27).

The presented model is useful for hypothesis generation about the biology of CD as it relates to the microbiome and could be further developed into a clinically useful theraprognostic tool. Some of the frequently occurring factors in our predictive models have already been linked to CD pathogenesis. As far as clinical biomarkers, fecal lactoferrin and fecal calprotectin occurred in the majority of models where clinical metadata was combined with the microbiome, supporting their importance as biomarkers for CD activity, especially in relation to the fecal microbiome (29, 30). *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 low in CD patients (9, 11, 33, 39). Remission was much more likely in individuals who 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 negatively impact CD pathogenesis (39). *Fusobacterium* also appeared in our predictive models and is associated with CD and CRC, something CD patients are more likely to develop than individuals without IBD (39). Many other taxa observed in our analysis had low abundance, but 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 that could be revealed by performing metagenomics on the week 0 stool samples in future studies, especially in subjects who achieved remission. These observations and the positive/negative associations of these microbes and CD also allow us to hypothesize on ways to alter the microbiome to increase the likelihood therapeutic response. Prior to the initiation of therapy, patients could get a fecal microbiome analysis. The community data could then be used to direct the patient to undergo a round of antibiotics to target and reduce the levels of Escherichia in the patient's gut. Alternatively, the microbes found to be positively associated with response could be formulated into a daily probiotic patients could take

while receiving therapy with the goal of increasing the likelihood of remission and mucosal healing. Additionally, altering the weighting or binning of important factors in the model could make prediction of response or remission more reliable. This could eventually allow for pre-screening of patients with stool samples to predict successful treatment or better direct treatment. If the fecal microbiome can be used as a theraprognostic 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 (23). 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 4. At week 6 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. 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) using an EPMotion 5075 pipetting system, as previously described (19, 20). The V4 region of the 16S rRNA gene from each sample was amplified and sequenced using the Illumina MiSeq Personal Sequencing platform as described elsewhere (30). Sequences were curated as described previously using the mothur software package (40). Briefly, we reduced sequencing and PCR errors, aligned the resulting sequences to the SILVA 16S rRNA sequence database, and removed any chimeric sequences flagged by UCHIME (41). Sequences were clustered into operational taxonomic units (OTU), as previously described (42). Briefly, OTUs were clustered at a 97% similarity cutoff and the relative abundance was calculated for OTUs in each sample. All sequences were classified using a naive Bayesian classifier trained against the RDP training set (version 11) and OTUs were assigned a classification based on which taxonomy had the majority consensus of sequences within a given OTU (43). All fastq files and the MIMARKS spreadsheet with de-identified clinical metadata are available at TBD.

Gut microbiome biomarker discovery and statistical analysis

Mothur as well as the R software package were used for our data analysis. Alpha diversity metrics (e.g. Inverse Simpson) were calculated for each sample in the dataset, and compared using non-parametric statistical tests (i.e. Kruskal-Wallace and Wilcox Test) (44, 45). Beta diversity was calculated the distance between samples using the thetaYC metric, which takes into account the types of bacteria and their abundance to calculate the differences between the communities (46). 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 as well as the adonis function in the R package vegan (47, 48). Change in alpha diversity over time was assessed using a Friedman test, whereas change in beta-diversity over time was assessed using the adonis function in vegan stratified by subject (49). Differentially abundant OTUs and phyla were selected through comparison of clinical groups using non-parametric statistical tests (i.e. Kruskal-Wallace and Wilcox Test) to identify OTUs/phyla where there is a P-value less than 0.05 following a Benjamini-Hochberg correction for multiple comparisons (50). We also used the relative abundance of each OTU across the samples and clinical metadata as input into the AUC-RF R package, in order to identify phylotypes/clinical variables that distinguish between various treatment and response groups, as well as to predict or determine response outcome (51).

, we used the friedmc function in package pgirmess (version = 165)

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

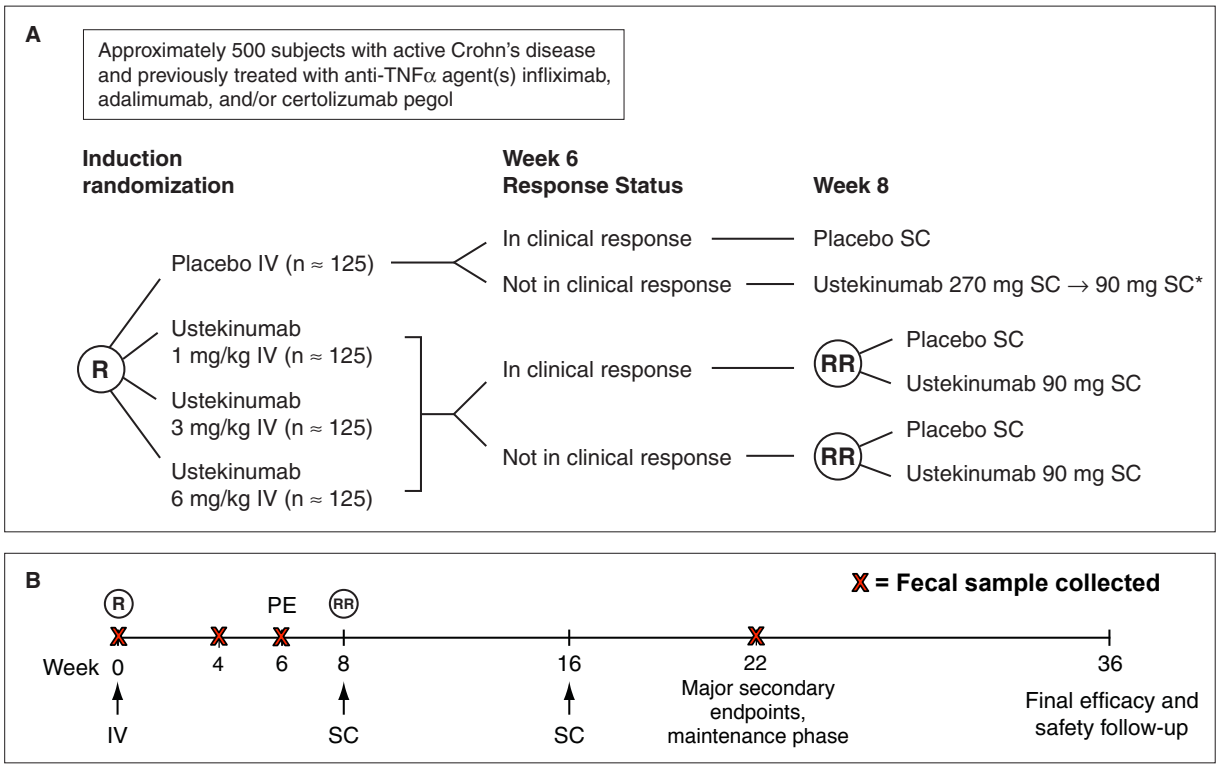
	Clinical Variable	Correlation	Alpha−Diversity (p−value)	Beta−Diversity (p−value)
1	CDAI	rho = −0.2	0.014	0.324
2	Loose Stool Frequency (per week)	rho = −0.2	0.003	0.024
3	C−Reactive Protein (mg/L serum)	rho = 0.06	0.394	0.033
4	Fecal Calprotectin (µg/g)	rho = 0.08	0.254	0.006
5	Fecal Lactoferrin (µg/g)	rho = 0.1	0.07	0.004
10	Corticosteroid Use (Y/N)	−	0.001	0.01
11	Disease Duration (years)	rho = −0.2	0.001	0.004
12	Tissue Involvement	−	0.19	0.004

Table 2: Diversity differenced bases on Response/Remission in UST treated subjects.

Clinical Variable	Alpha−Diversity (p−value)	Beta−Diversity (p−value)
Week 6 Response (No, Yes)	0.440	0.012
Week 6 Remission (No, Yes)	0.020	0.017
Week 22 Response (No, Yes)	0.900	0.012
Week 22 Remission (No, Yes)	0.440	0.012

Figures

Figure 1: Experimental design as adapted from Sanborne et al 2012. (A) Diagram of experimetnal design and (B) stool sampling, treatment, and response evalution timeline.

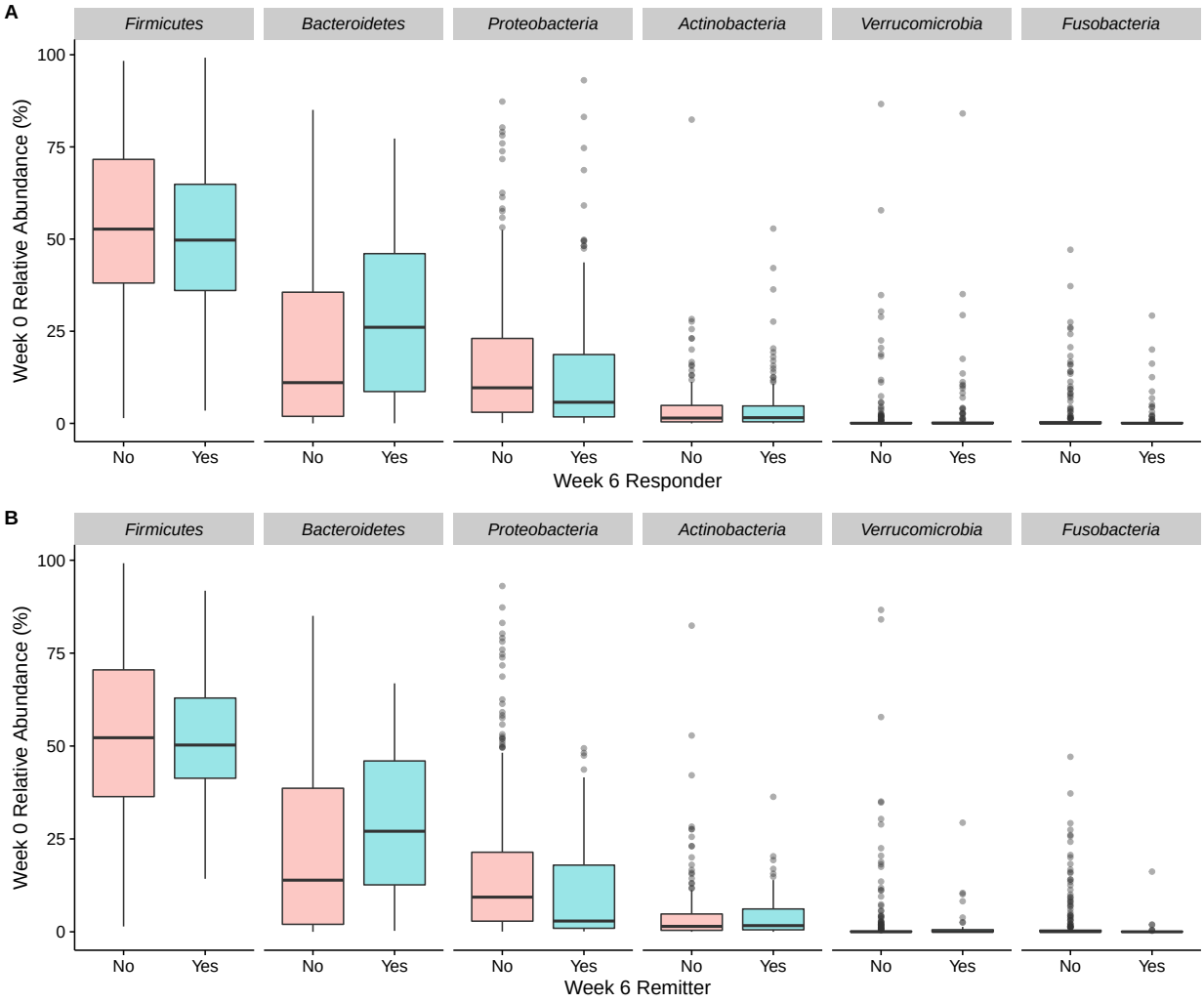


IV = Intravenous; SC = Subcutaneous; ↑ = Study agent administration

PE = Primary Endpoint; R = Randomization; RR = Rerandomization only for subjects receiving ustekinumab induction therapy

* Subjects receiving placebo at Week 0 who are not in clinical response at Week 6 will receive ustekinumab 270 mg SC and 90 mg SC at Weeks 8 and 16, respectively.

Figure 2: Phyla from week 0 stool samples in subjects treated with UST by week 6 outcome (A) Response and (B) remission status.



Supplemental Figure 1: Phyla from week 0 stool samples in subjects treated and maintained with UST by week 22 outcome (A) Response and (B) remission status.

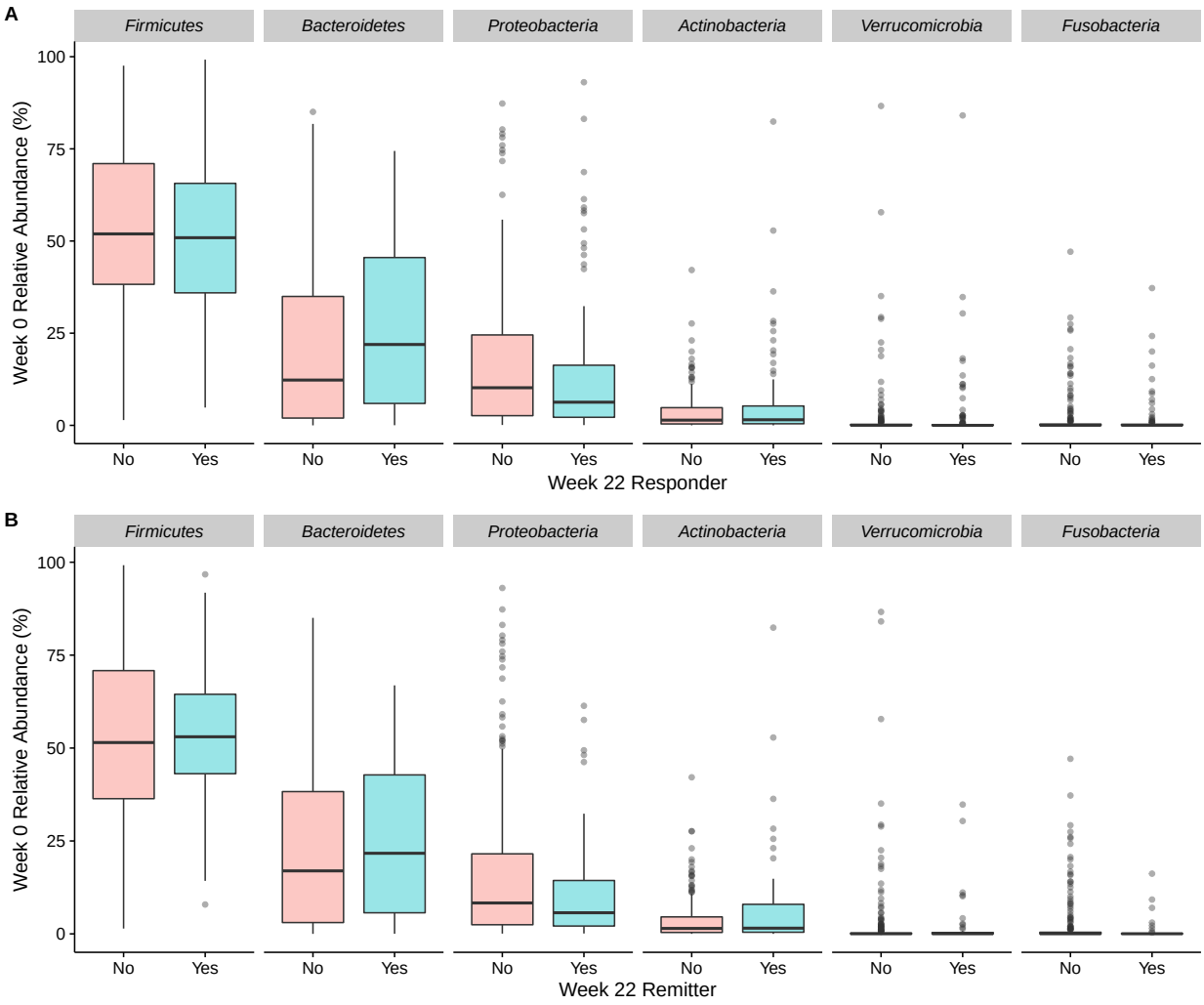


Figure 3: Differential taxa in week 0 stool samples from subjects treated with UST, based on week 6 remission status

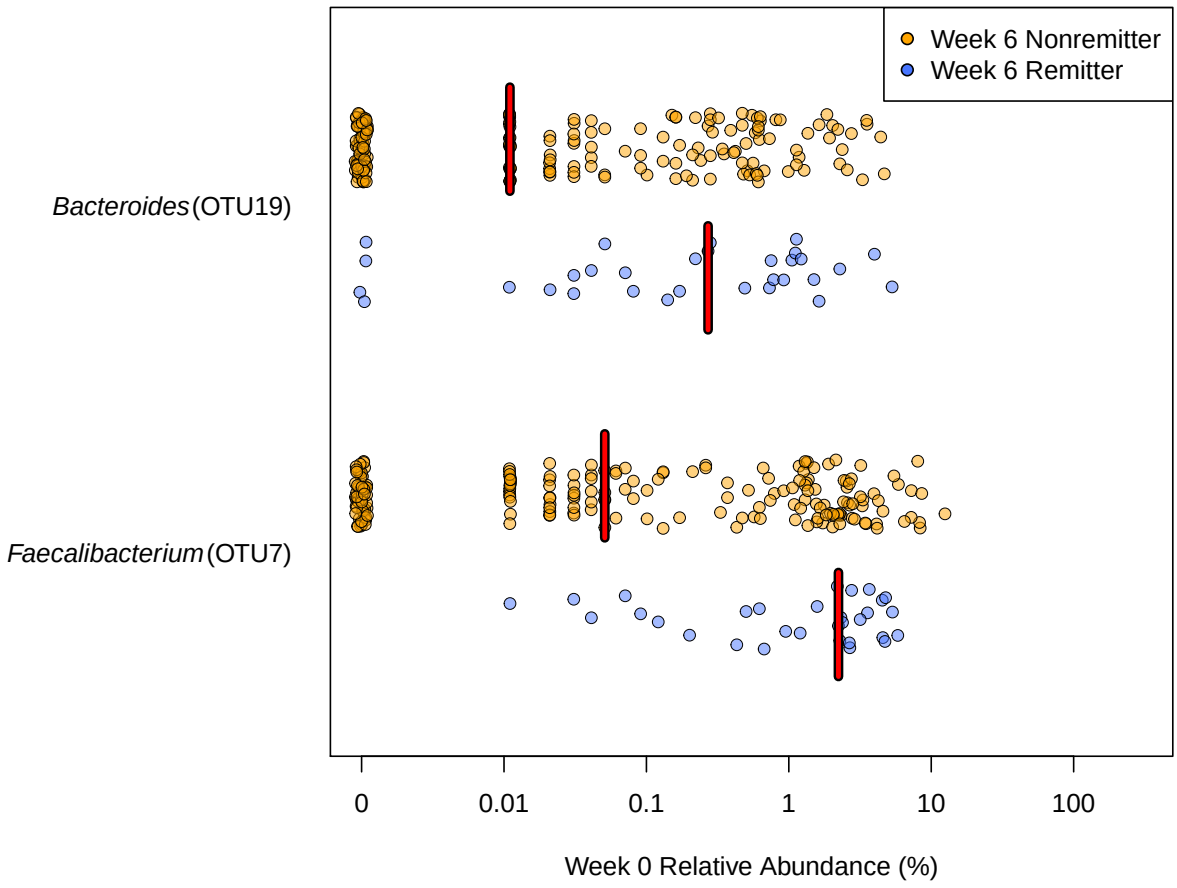


Figure 4: Change in alpha diversity over time by induction treatment and week 22 response status.

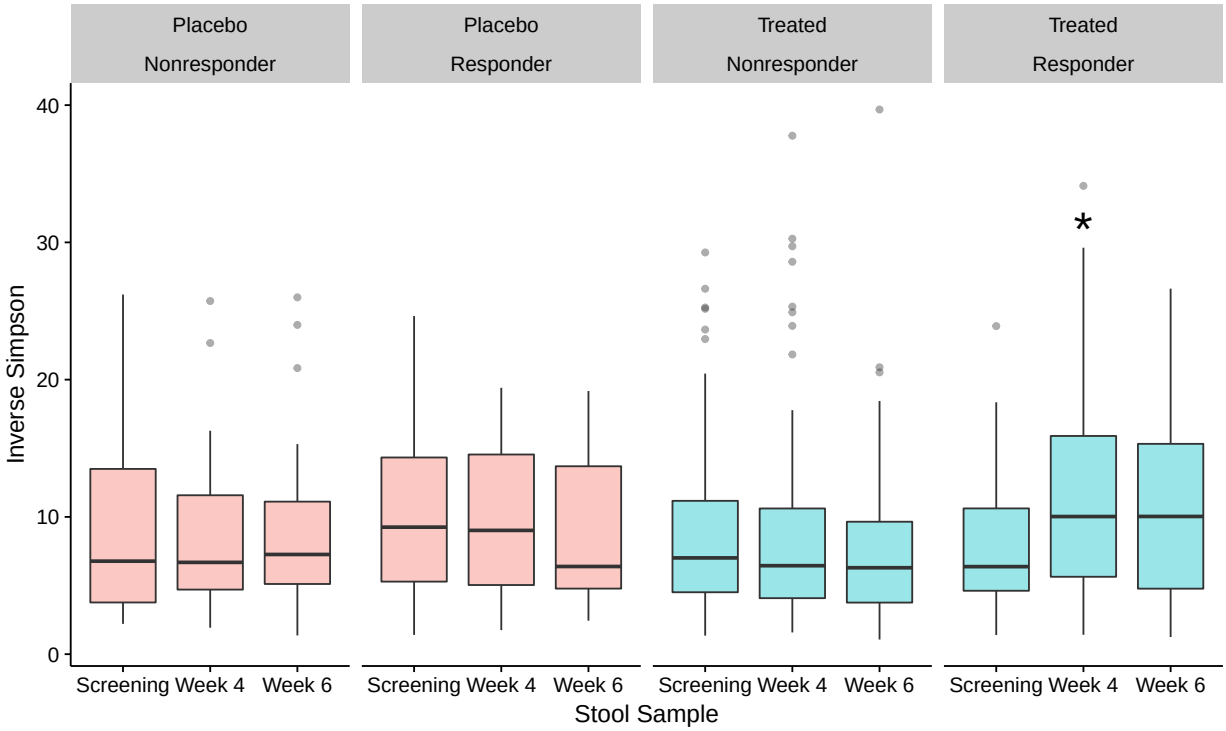


Figure 5: Classification of week 6 response or remission status using week 6 stool samples from subjects treated with UST (A) ROCs for week 6 outcome based on the microbiome. (B) Top predictive taxa from week 6 stool for remission status at week 6.

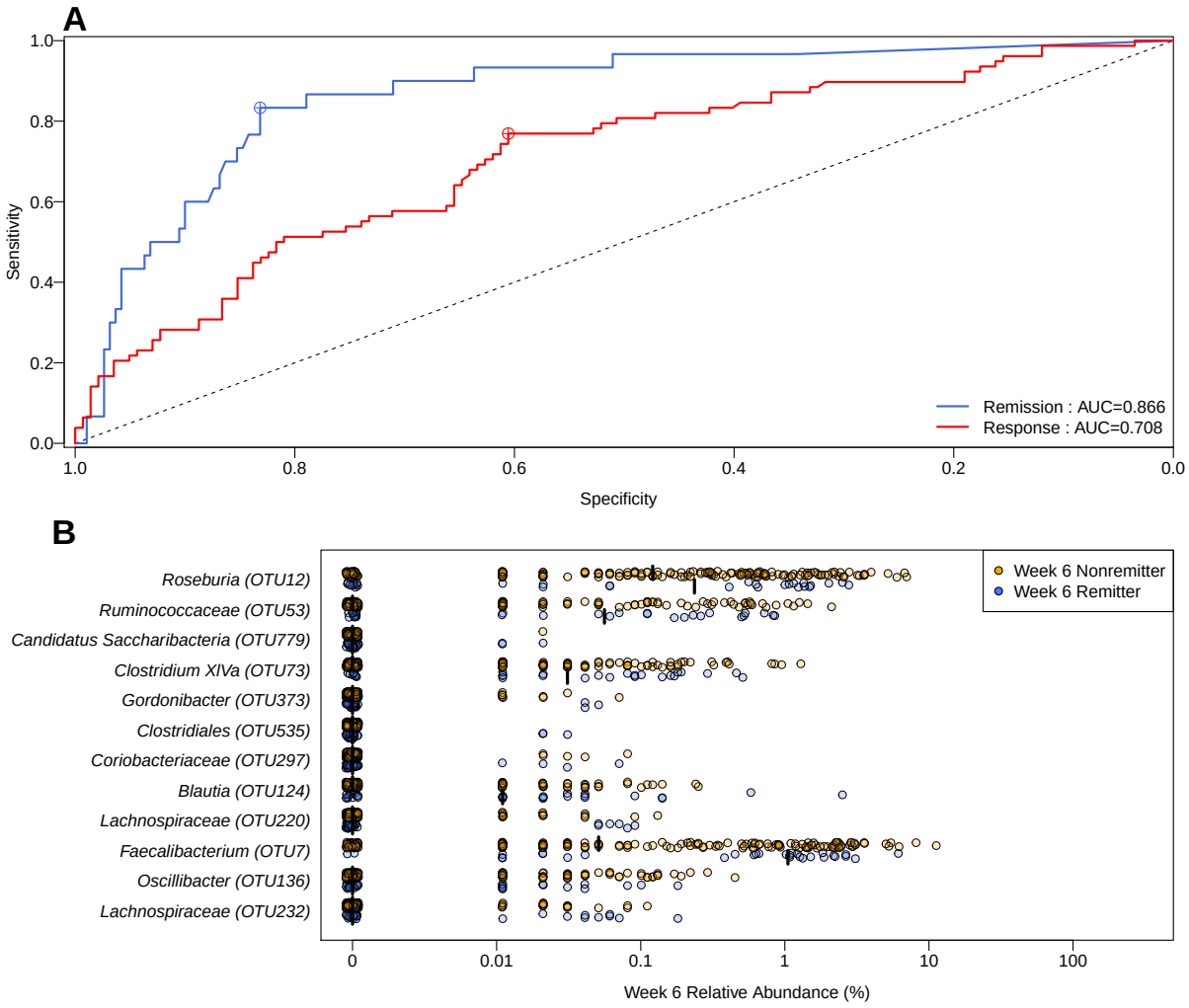
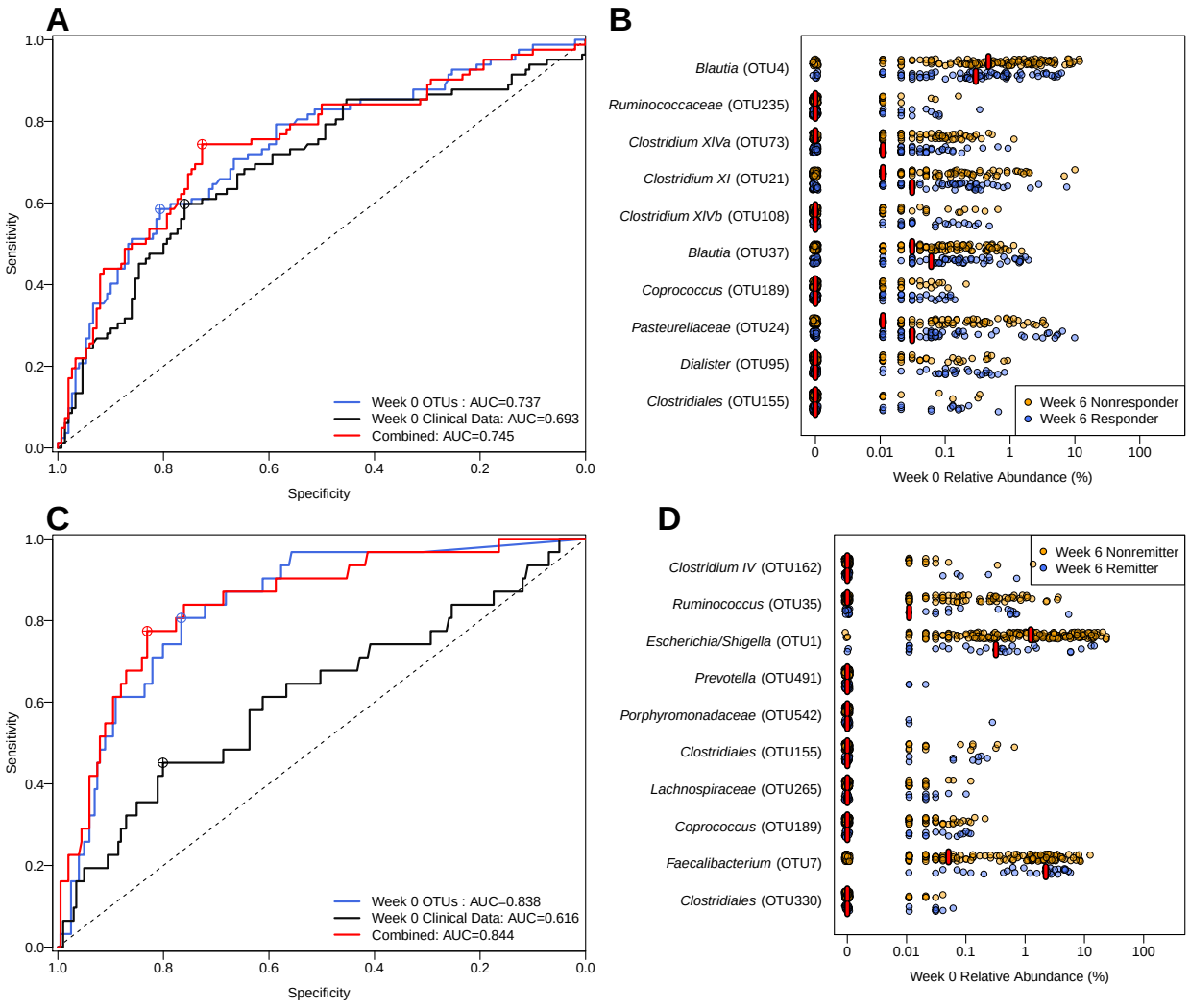
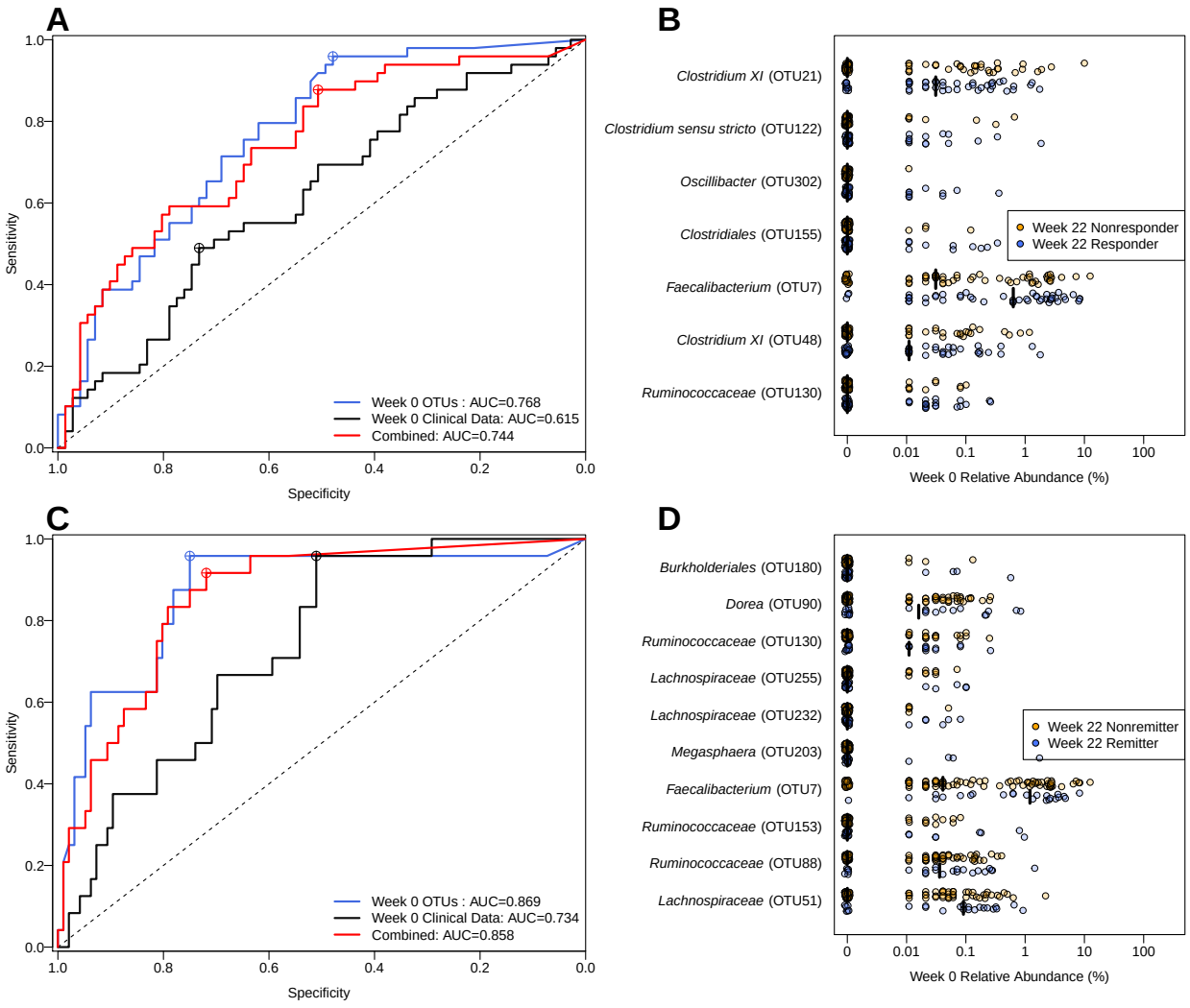


Figure 6: Prediction of week 6 disease status in subjects treated with UST, using week 0 samples ROCs for (A) response and (C) remission using microbiome data, clinical metadata, and the combined model. Top predictive taxa for the microbiome model based on MDA for (B) response and (D) remission.



Supplemental Figure 2: Predicting week 22 disease status in subjects treated and maintained with UST, using week 0 samples ROCs for (A) response and (C) remission using microbiome data, clinical metadata, and the combined model. Top predictive taxa for the microbiome model based on MDA for (B) response and (D) remission.



References

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

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

3. Molodecky NA, Soon IS, Rabi DM, Ghali WA, Ferris M, Chernoff G, Benchimol EI, Panaccione R, Ghosh S, Barkema HW, Kaplan GG. 2012. Increasing incidence and prevalence of the inflammatory bowel diseases with time, based on systematic review. Gastroenterology 142:46–54.e42; quiz e30.

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

5. 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.

6. 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.

7. 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.

8. Sartor RB. 2006. Mechanisms of disease: Pathogenesis of crohn's disease and ulcerative colitis. Nat

Clin Pract Gastroenterol Hepatol 3:390–407.

9. 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.

10. Manichanh C, Rigottier-Gois L, Bonnaud E, Gloux K, Pelletier E, Frangeul L, Nalin R, Jarrin C, Chardon P, Marteau P, Roca J, Dore J. 2006. Reduced diversity of faecal microbiota in crohn's disease revealed by a metagenomic approach. Gut 55:205–11.

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

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

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

14. 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.

15. 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.

16. Schubert AM, Sinani H, Schloss PD. 2015. Antibiotic-induced alterations of the murine gut

microbiota and subsequent effects on colonization resistance against clostridium difficile. MBio 6:e00974.

17. Schubert AM, Rogers MAM, Ring C, Mogle J, Petrosino JP, Young VB, Aronoff DM, Schloss PD. 2014. Microbiome data distinguish patients with clostridium difficile infection and non-c. difficile-associated diarrhea from healthy controls. mBio 5.

18. 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.

19. 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.

20. 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.

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

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

23. Sandborn WJ, Gasink C, Gao LL, Blank MA, Johanns J, Guzzo C, Sands BE, Hanauer SB, Targan S, Rutgeerts P, Ghosh S, Villiers WJ de, Panaccione R, Greenberg G, Schreiber S, Lichtiger S, Feagan BG. 2012. Ustekinumab induction and maintenance therapy in refractory crohn's disease. N Engl J Med 367:1519–28.

24. 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.

25. 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–

459 the McGill experience. J Crohns Colitis 8:1516–22.

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

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

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

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

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

473 31. Papa E, Docktor M, Smillie C, Weber S, Preheim SP, Gevers D, Giannoukos G, Ciulla D, Tabbaa D,
474 Ingram J, Schauer DB, Ward DV, Korzenik JR, Xavier RJ, Bousvaros A, Alm EJ. 2012. Non-invasive
475 mapping of the gastrointestinal microbiota identifies children with inflammatory bowel disease. PLoS
476 One 7:e39242.

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

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

482 34. Huang EY, Inoue T, Leone VA, Dalal S, Touw K, Wang Y, Musch MW, Theriault B, Higuchi
483 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.

35. 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.

36. 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.

37. 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.

38. 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.

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

40. Schloss PD, Gevers D, Westcott SL. 2011. Reducing the effects of PCR amplification and sequencing artifacts on 16S rRNA-based studies. PLoS One 6:e27310.

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

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

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

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

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

46. Yue JC, Clayton MK. 2005. A similarity measure based on species proportions. Communications in

510 Statistics-Theory and Methods 34:2123–2131.

511 47. Schloss PD. 2008. Evaluating different approaches that test whether microbial communities have
512 the same structure. ISME J 2:265–75.

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

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

518 50. Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: A practical and powerful
519 approach to multiple testing. Journal of the Royal Statistical Society Series B (Methodological) 57:289–
520 300.

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