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

Running title: The fecal microbiome as a tool for monitoring and predicting response outcomes in Ustekinumab-treated, anti-TNF-alpha refractory 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 ...

Abstract

Abstract: The 16S rRNA gene from patient stool samples was sequenced using the Illumina MiSeq platform. The resulting sequences were curated and assigned to taxonomic groups using the mothur software package to determine the bacterial communities and relative abundance of bacterial species present in these patients. The relative abundance among the fecal microbiota, patient demographic data, and clinical metadata were used as input to a random forest machine-learning algorithm to predict disease severity and response to treatment with UST.

Fecal microbial diversity at baseline significantly correlates with markers for disease severity, such as Crohn's Disease Activity Index (CDAI), stool frequency, and disease duration. Additionally, the overall community structure of the microbiome was significantly different based on stool frequency, CRP, fecal lactoferrin, fecal calprotectin, corticosteroid use, disease duration, and tissue involvement. Baseline fecal microbiome community structures and species diversity were significantly different among responders and non-responders to UST treatment. Faecalibacterium, among other taxa, was significantly more abundant in responders/remitters. Additionally, the microbiome of clinical responders changed over time, in contrast to nonresponsive subjects. Using AUC-RF, differences in the baseline microbiome and clinical metadata were able to predict response to UST, especially remission, with some AUCs approaching 0.85.

Importance: Crohn's disease (CD) is a global health issue characterized by patches of ulceration and inflammation along the gastrointestinal tract. Individuals with CD have reduced microbial diversity in their guts, compared to healthy individuals. It remains unclear if this reduced diversity is a result or cause of pathogenesis. We investigated the relationship between the fecal microbiome and clinical phenotypes in subjects with moderate to severe CD treated with Ustekinumab (UST) in a Phase 2b study to determine whether the fecal microbiome at baseline is predictive of disease severity and therapeutic response, as well as if the fecal microbiota changes due to therapy.

The ability to predict and monitor response to treatment using the microbiome will likely provide another clinical tool in treating CD patients. Additionally, the observed baseline differences in fecal microbiota and changes due to therapeutic response will allow further investigation into the microbes important in CD pathogenesis as well as establishing and maintaining CD remission. Finally, beneficial microbes

associated with response to treatment could be developed as probiotics to increase the likelihood of response while undergoing treatment.

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

Introduction

Crohn's disease (CD), an incurable inflammatory bowel disease (IBD), is a global health issue with increasing incidence. CD affects approximately 3 million people worldwide, causing large economic and healthcare utilization impacts on society (1–3). CD is characterized by patches of ulceration and inflammation affecting the entire bowel wall along the gastrointestinal tract, most commonly in the ileum and colon. Individuals with CD experience frequent diarrhea, abdominal pain, fatigue, and weight loss resulting in significant health care costs, lower quality of life, and economic impacts due to loss of productivity (2, 4, 5). Current treatments for CD include antibiotics, anti-inflammatory drugs, immunomodulators, surgery, and biologic agents targeting tumor necrosis factor alpha (TNF-α), such as Infliximab (Remicade). Within 10 years of diagnosis, approximately half of individuals with CD will require surgery and the majority will experience escalating immunosuppressive treatment (6). 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 (5, 7). Faster induction of remission following diagnosis reduces the risk of irreversible intestinal damage and disability (7–9) (7-9). Anti-TNF-α therapy in combination with thiopurines has emerged as the preferred treatment for CD, but up to half of individuals with CD fail to respond or lose response to anti-TNF-α therapy (6, 7). Ustekinumab (UST), a monoclonal antibody directed against the shared p40 subunit of IL-12 and IL-23, has been proposed as an alternative therapy for these patients (10). While clinical trials have demonstrated that UST is a viable option for the treatment of CD (7, 10–12) (7, 10-12), some patients within these trials were non-responsive to UST, which may be explained by differences in the patients' gut microbiomes.

The precise etiology of CD remains unknown, but host genetics, environmental exposure, and the gut microbiome appear involved (1, 13). Genome-wide association studies of individuals with CD identified several susceptibility genes including NOD2, a receptor involved in bacterial killing and innate immunity. Defects in NOD2 function affects microbial sensing, the regulation of IL-23 driven Th17 responses, and indirect modulation of the gut microbiome (5, 14). The gut microbiome has also been shown to play a key role in inflammation, immunity, and IBD (15). 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 (14,

16–19) . Additionally, previous studies have shown that the gut microbiome can be predictive of disease severity in new-onset, pediatric CD patients (19, 20). It remains to be determined, however, whether the microbiome can predict response to therapy in CD (14). Additionally, the effect of biologic treatment on the gut microbiome is not well understood. If the fecal microbiome can be used as a theraprognostic tool to non-invasively determine and monitor disease severity as well as predict response to specific treatment modalities, then more targeted treatment could result in reduced adverse effects of less effective therapies and faster achievement of remission.

Our lab was approached to analyze the gut microbiomes of individuals who participated in a Phase II clinical trial to determine the efficacy of UST in treating CD (10). Using stool samples taken prior to the start of the study, 16S rRNA gene sequence data from these patients will allow us to determine associations between clinical metadata, disease severity, and the fecal microbiome and whether clinical responders have a microbiome that is distinct from non-responders at baseline. Preliminary results generated with fecal samples from a subset of study participants and sequenced using the Roche 454 platform suggest that the fecal microbiota of moderate to severe CD patients refractory to anti-TNF-α may differentiate individuals who will respond to treatment with UST; however, large interpersonal variation limited the power of our findings. This study attempts to overcome many of the limitations in our preliminary analysis by increasing our sample size to the full patient cohort and using the Illumina MiSeq platform to improve our sequencing depth. We demonstrate that the fecal microbiome is associated with baseline clinical metadata and that these associations and differences are useful in predicting disease severity and treatment outcome.

Results

Characteristics of Study Population We studied the fecal microbiota in a subset of TNF-α refractory CD patients who took park in the CERTIFI clinical trial described in (10). Briefly, patients with a history of moderate to severe CD were randomly assigned to a treatment group in the induction phase of the study. Subjects provided a stool sample at screening (Week 0), Week 4 and Week 6. At Week 8 patients were re-randomized into maintenance therapy groups. A final stool sample was provided at Week 22. Response to therapy was evaluated at week 4, 6, 8, and 22 based on change in CDAI. Samples from subjects that completed the clinical trial and had complete clinical metadata were included

in our analysis. We used 16s rRNA gene sequencing to analyze the microbiome from 306 fecal samples provided prior to treatment as well as 258 Week 4, 289 Week 6, and 205 Week 22 post-treatment fecal samples, for a total of 1058 samples. Demographic and baseline disease characteristics are summarized in supplemental table 1.

Comparison of microbiome at screening based on clinical variables To determine if there were any significant associations between microbial diversity and clinical variables of interest, we compared the microbiome with clinical data at Week 0. We determined species richness (α-diversity) using the inverse Simpson metric and assessed associations between species richness and clinical data using Spearman's rank correlation, Wilcoxon rank-sum, or Kruskal-Wallis rank-sum tests. Associations between the overall community structure (ß-diversity) and clinical data were determined using the thetaYC distance metric as input to the adonis PERMANOVA function within the vegan R package (21). As seen in table 1, we observed a correlation between CDAI and species richness, with higher CDAI correlating to lower richness. The overall community structure was not different based on CDAI. When looking at CDAI subscores, we observed a significant association between species richness and the frequency of loose stools per week. The overall community structure was also significantly different based on weekly loose stool frequency. No significant association was observed between CRP and fecal calprotectin and species richness, while higher fecal lactoferrin weakly correlates with higher richness. The overall community structure was significantly different based on CRP, fecal calprotectin, and fecal lactoferrin. No significant differences in the microbiome were observed for BMI, weight, or sex. Overall community structure was different based on age. The overall community structure was also different based on the tissue affected. Species richness and the overall community structure were significantly different based on corticosteroid use. The community structure was significantly different based on disease duration and a significant correlation was seen between species richness and disease duration, with lower richness corresponding to longer disease.

consider including LEfSe data for quick look at discriminate OTU based on sig dif clinical variables like in Gevers and Zackular papers?

Comparison of clinical responders and non-responders Next, We wanted to see if there were associations between the microbiome at baseline and response to treatment. For this study, response was defined as a 30% decrease from CDAI at baseline and remission defined as a CDAI below 150. Of the

306 screening samples analyzed, 232 were from subjects receiving UST and 74 from subjects receiving placebo. Baseline fecal microbiome community structures and species diversity were different among responders and non-responders to UST treatment. Based on response at the primary endpoint of the study, 6 weeks after IV induction, there was no difference in species richness between response groups, but there was a significant difference in the overall community structure of the entire cohort. This difference in community structure was not significant in treatment vs. placebo groups. Week 6 remitters were significantly different from non-remitters in both species richness (0.0005) and overall community structure (0.017). When looking at treated vs. untreated Week 6 remitters, the treated group had significant differences in both species richness and community structure while untreated remitters we not different from untreated non-remitters. At the secondary endpoint, 22 weeks after IV-induction and 14 weeks after maintenance dosing, there was no difference in species richness between response groups, but there was a significant difference in the overall community structure of the entire cohort. Week 22 remitters were significantly different from non-remitters in both species richness (0.57) and overall community structure (0.016). However, these differences were not seen when the cohort was broken down by induction group. This could be due to changes in maintenance treatment.

The microbiome by treatment and response over time One major question with regards to biologic treatment of IBD and the microbiome is whether treatment has an effect on the microbiome. We explored this question 2 different ways. We included subjects that had stool samples at all 4 time points and another analysis using subjects who provided samples at weeks 0, 4, and 6. We used PERMANOVA stratified on each subject, as a proxy for a repeated measures ANOVA, to determine if the microbiome changed over time. We found that taken together treatment does not affect the microbiome. No significant difference was seen based on visit when looking at all groups and response status at week 4, 6, or 8 over the first 3 time points, but there was a significant interaction between response at week 22 and visit ($p=0.001$) and between relative response, induction group, and visit ($p=0.0445$).

This led us to examining just the week 22 responders vs. non-responders across visit. No significant difference over time was observed in non-responders. When we segregated week 22 responders, we saw a significant change in community structure over time. There was also a significant difference based on treatment group, but no significant interaction. When looking at treated vs. untreated responder groups, we observed a significant difference based on visit in the treated, Week 22 responder and in

untreated responders across the first 3 visits prior to maintenance phase.

When looking at time in all subjects across all 4 time points we observed a significant interaction between visit and response, however no interaction between visit, treatment group, and response. In all subjects there was a significant difference in community structure based on response at Week 22. In treated subjects, we observed a significant interaction between response and visit, as well as a significant difference in community structure based on response at Week 22. No significant difference was observed in untreated responders across all 4 time points.

Prediction of response based on the microbiome at screening

Another major question in IBD and the microbiome is if response can be predicted using the microbiome. To address this we used AUCRF to develop a random forest classification model to differentiate responders from non-responders, as well as remitters from non-remitters, based on the relative abundance of fecal microbiome community members, clinical metadata, and combined microbiome and clinical data (22, 23). We ran these models for response and remission at Week 4, 6, 8, and 22 of the study. The optimal models for response and remission at the primary endpoint (Week 6) are shown in Figure 1. Using only clinical metadata to predict response, the model predicted response with an AUC of 0.694 with a specificity of 0.765 and a sensitivity of 0.575. Using only microbiome data, the model predicted response with an AUC of 0.701 with a specificity of 0.62 and a sensitivity of 0.736. When combining clinical metadata with the microbiome, the model predicted response with an AUC of 0.707 with a specificity of 0.615 and a sensitivity of 0.736. With respect to Week 6 remission, using solely clinical metadata we achieved AUC of 0.676 with a specificity of 0.457 and a sensitivity of 0.871. Using only fecal microbiome data we achieved an AUC of 0.811 with a specificity of 0.621 and a sensitivity of 0.968. When combining clinical metadata with the microbiome AUC of 0.846 with a specificity of 0.663 and a sensitivity of 0.935.

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 significantly more abundant in responders/remitters. Their abundances can be seen in figure 4.

In addition to predicting future response, we wanted to determine if the microbiome could be used to monitor response to therapy. Again we used AUC-RF in order to determine if the fecal microbiome at Week 6 could be used to determine response or remission at Week 6. As seen in Supplemental Figure 1, using the microbiome alone we achieved an AUC of 0.696 for response with a sensitivity of 0.641 and a specificity of 0.711. For remission we had an AUC of 0.838 with a sensitivity of 0.767 and specificity of 0.816. Again we were better able to distinguish remitters from non-remitters than responders/non-responders. The clinical data were more reliable for determining disease activity at Week 6.

Discussion

Our results examine the fecal microbiome of a subset of patients who participated in the CERTIFI trials to determine if the microbiome can predict response to therapy and if therapy has any effect on the microbiome. Several previous studies have looked at fecal and mucosal microbiomes in pediatric patients with new-onset and established disease and with established disease in adults (19, 24, 25). Unlike these studies, our patients were mostly Caucasian adults in their late thirties to early forties who failed to respond or lost response to anti-TNF-α biologic treatment. We were able to find associations between the fecal microbiome of these patients and CDAI, stool frequency, fecal calprotectin, fecal lactoferrin, serum CRP, corticosteroid use, tissue involvement, and duration of disease.

The association of the microbiome with clinically relevant biomarkers and disease activity metrics indicates that the microbiome may also function as a biomarker for CD activity. Given that serum CRP, calprotectin, and lactoferrin are used as biomarkers to measure intestinal inflammation and CD severity, it is interesting to see that the microbial community structure is different among patients based on these markers (26, 27). This supports the idea that the microbiome could be useful as a biomarker for measuring disease activity in patients, especially when considered in relation to these biomarkers (25). 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 (19, 24, 25). 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. Given that higher stool frequency is associated with looser stool

consistency, this finding appears consistent with the association between loose stools and lower diversity (28).

We also observed differences in the microbiome in relation to other clinical variables. The microbial community structure was different 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 (29). This 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 (30). As corticosteroid use appears to impact diversity, corticosteroids may be useful when trying to positively impact the microbiome during biologic therapy and increase the possibility of response to CD therapies.

Unlike other studies, these patients had a CD diagnosis for an average of 12 years (Supplemental Table 1) (19, 24, 25). We observed that 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. One could hypothesize earlier biologic intervention may 'preserve' microbiome that promotes remission and reduces the likelihood of relapse. Publications have come out in support of earlier biologic intervention, as it appears to increase the likelihood of inducing remission and mucosal healing (31–33). 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 (34).

One 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 seeing the effects of change in disease activity and health rather than any effects from treatment. This interpretation is consistent with studies using the microbiome to distinguish between remission and active CD (25). 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.

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 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, 6 weeks after beginning treatment. This is consistent with other studies again suggesting the microbiome could be useful in detecting remission versus active disease (25).

While using the presented model may not be useful clinically to predict response to therapy at this time, it is useful for hypothesis generation about the biology of CD as it relates to the microbiome. 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 tot eh fecal microbiome (26, 27). Faecalibacterium was the most frequently occurring OTU in our models. It is associated with health and has been shown to be low in CD patients (14, 17, 29, 35). Remission was much more likely in individuals who had measurable Faecalibacterium present at baseline. This supports the hypothesis that Faecalibacterium impacts CD. Escherichia/Shigella also occurred frequently in our models. This OTU is associated with inflammation and has been shown to negatively impact CD (35). Fusobacterium also appeared in our predictive models and is associated with CD and CRC, something CD patients are more likely to get (35). These observations and the positive/negative associations of these microbes and CD 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.

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 effect 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. While this prediction is not clinically useful as of yet, 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 phase II clinical study of approximately 500 patients to assess the safety and efficacy of UST for treating anti-TNF-α refractory CD patients (10). 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 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 (22, 36). 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 (27). Sequences were curated as described previously using the mothur software package (37). 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 (38). After curation, we obtained between 1 and 130,074 sequences per sample (median 13786), with a median length of 253 bp. To limit effects of uneven sampling, we rarefied the dataset to 3,000 sequences per sample. Parallel sequencing of a mock community revealed an error rate of 0.017 %. Sequences were clustered into operational taxonomic units (OTU), as previously described (39). 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 (40). All fastq files and the MIMARKS

spreadsheet with de-identified clinical metadata are available at TBD.

Gut microbiome biomarker discovery analysis

Mothur as well as the R software package were used for our data analysis. Alpha diversity metrics (e.g. Shannon, Inverse Simpson) were calculated for each sample in the dataset, and compared using non-parametric statistical tests (i.e Kruskal-Wallace and Wilcox Test) (41, 42). Beta diversity was calculated the distance between samples using the theta YC metric, which takes into account the types of bacteria and their abundance to calculate the differences between the communities (43). These distance matrices were visualized by generating non-metric dimensional scaling (NMDS) plots of the distances. Overlap between sets of communities was assessed using the non-parametric analysis of molecular variance (AMOVA) and homogeneity of variance (HOMOVA) tests (21, 44). Differentially abundant OTUs were selected using the biomarker discovery algorithm, LEfSe [linear discriminant analysis (LDA) effect size] for each pairwise comparison of clinical groups (45). In short, This method uses the Wilcox non-parametric test to identify OTUs where there is a P-value less than 0.05 and then applies a LDA step to identify the effect sizes that are the most meaningful (i.e. greater than 2.0). We also used the relative abundance of each OTU across the samples and clinical metadata as input to the AUC-Random forest package available to identify phylotypes/clinical variables that would allow us to distinguish between various treatment and response groups (46).

Tables

Supplemental Table 1: Summary of clinical metadata of chort at baseline

Clinical Variable	Treated	Placebo	Total
Age (years)	38(13)	40(14)	39(13)
Sex Male(Female)	85(147)	32(42)	117(189)
Race Caucasian(Other)	213(19)	69(5)	282(24)
Cortico Steroid Use Yes(No)	93(139)	39(35)	132(174)
BMI	26(6.7)	25(4.9)	25(6.3)
Disease Duration (years)	12(8.4)	13(10)	12(8.8)
CDAI mean(sd)	330(62)	310(69)	320(64)
Bowel Stricture Yes(No)	29(203)	8(66)	37(269)
Tissue Involvement Colon/Ileocolic/Ileal	67/119/46	18/29/27	85/148/73

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

Clinical Variable	Summary	Species Richness (alpha−diversity)	Community Structure (beta−diversity)
CDAI	Min=154, Median=319, Max=483	rho = −0.2 (0.005)	0.3
Loose Stool Frequency (per week)	Min=2, Median=51, Max=100	rho = −0.2 (7e−04)	0.01
C−Reactive Protein (mg/L serum)	Min=0.1, Median=11.7, Max=199	rho = 0.06 (0.3)	0.02
Fecal Calprotectin (µg/g)	Min=14, Median=582.5, Max=26070	rho = 0.08 (0.1)	0.002
Fecal Lactoferrin (µg/g)	Min=0.25, Median=83.78, Max=3141	rho = 0.1 (0.03)	0.001
BMI	Min=15, Median=24, Max=55.3	rho = 0.07 (0.2)	0.2
Weight (kg)	Min=40, Median=69, Max=150	rho = 0.07 (0.2)	0.08
Age (years)	Min=18, Median=37, Max=76	rho = −0.05 (0.4)	0.02
Sex (F/M)	189/117	0.5	0.3
Corticosteroid Use (Y/N)	132/174	2e−04	0.004
Disease Duration (years)	Min=0.48, Median=10.44, Max=44.92	rho = −0.2 (7e−05)	0.001
Tissue Involvement	Colon=85, Colon−Ileum=148, Ileum=73	0.1	0.001

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

Status Week	Summary	Species Richness (Alpha−diversity)	Community Structure (beta−diversity)
Week 6 Response (Yes, No)	196, 110	0.11	0.012
Treated	150, 82	0.2	0.066
Placebo	46, 28	0.37	0.073
Week 6 Remission (Yes, No)	263, 43	5e−04	0.017
Treated	201, 31	0.002	0.022
Placebo	62, 12	0.11	0.26
Week 22 Response (Yes, No)	186, 120	0.57	0.016
Treated	144, 88	0.63	0.12
Placebo	42, 32	0.87	0.058
Week 22 Remission (Yes, No)	250, 56	0.019	0.007
Treated	196, 36	0.056	0.089
Placebo	54, 20	0.24	0.11

Supplemental Table 2: Table of taxa that appear frequently in predictive models at different response weeks, use baseline and post-treat time points, also pooled

	Occurance (out of 24)
Faecalibacterium (OTU7)	8
Blautia (OTU37)	6
Clostridiales (OTU155)	5
Pasteurellaceae (OTU24)	4
Lachnospiraceae (OTU51)	4
Ruminococcaceae (OTU197)	4
Clostridiales (OTU330)	4

Figures

Figure 1: Prediction of RESPONSE/REMISSION in treated subjects using all clinical metadata, baseline microbiome alone, and combined A. Response ROCs B. Response Model Performance vs. reality C. Top predictive taxa and abundance based on response D. REMISSION ROCs E. REMISSION Model Performance vs. reality F. Top predictive taxa and abundance based on remission

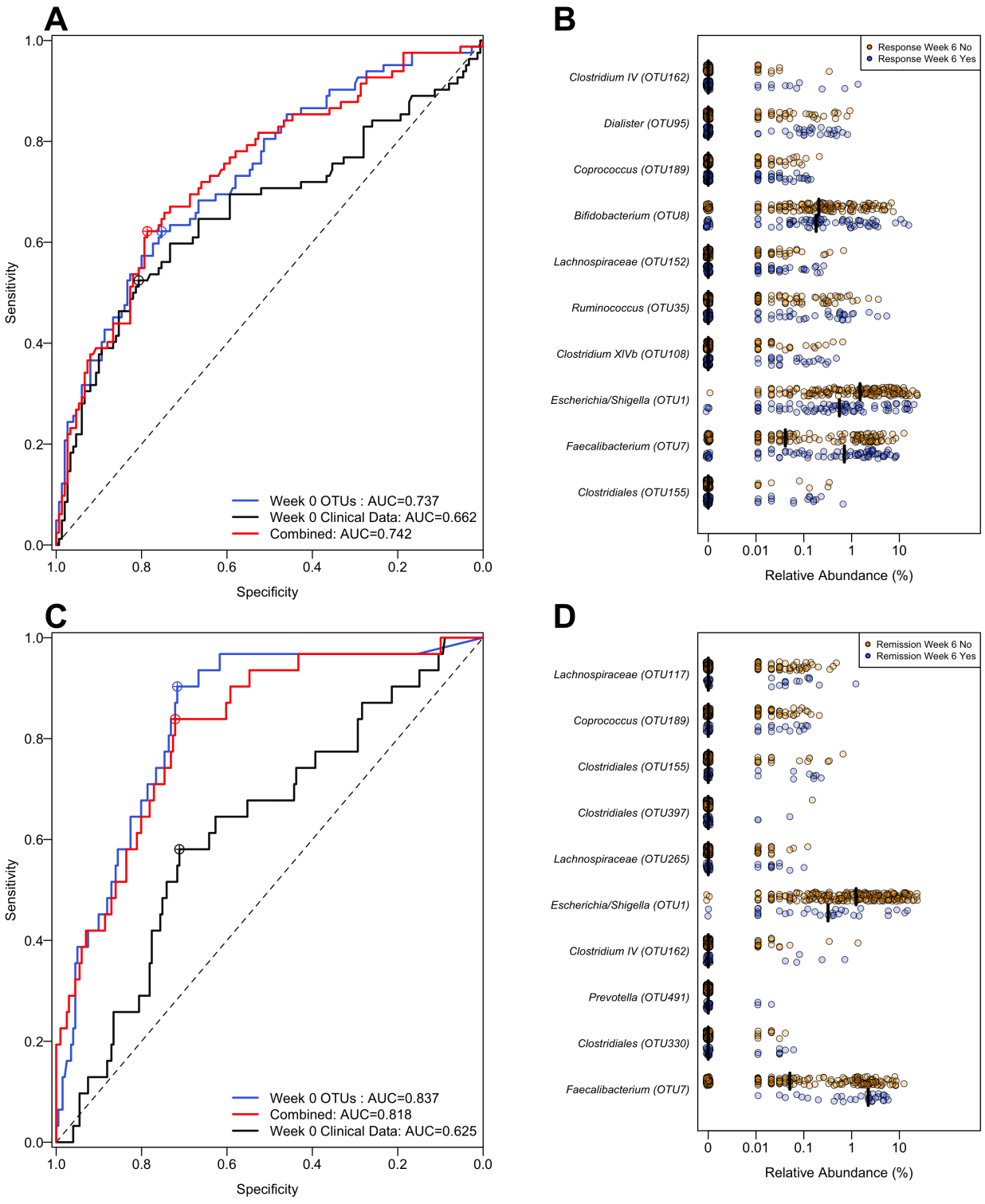

Figure 1: Prediction of RESPONSE/REMISSION in treated subjects using all clinical metadata, baseline microbiome alone, and combined. A. Response ROCs. B. Response Model Performance vs. reality. C. Top predictive taxa and abundance based on response. D. REMISSION ROCs. E. REMISSION Model Performance vs. reality. F. Top predictive taxa and abundance based on remission

Figure 2: Lefse data supporting the abundance/importance data in the predictive models Abundance strip charts of differential taxa based on A) response and B) remission.

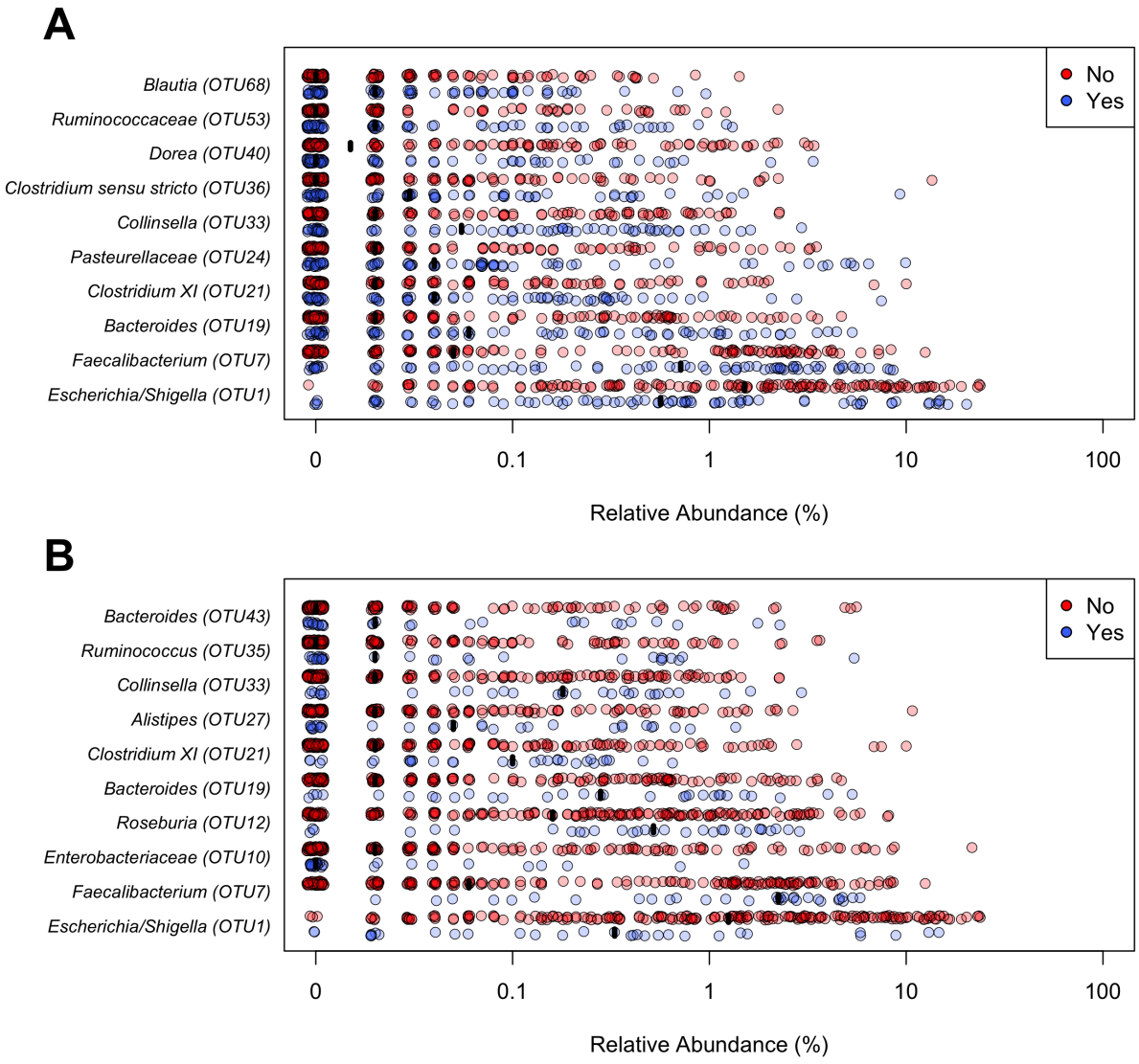

Figure 2: Figure 2: Lefse data supporting the abundance/importance data in the predictive models. Abundance strip charts of differential taxa based on A) response and B) remission.

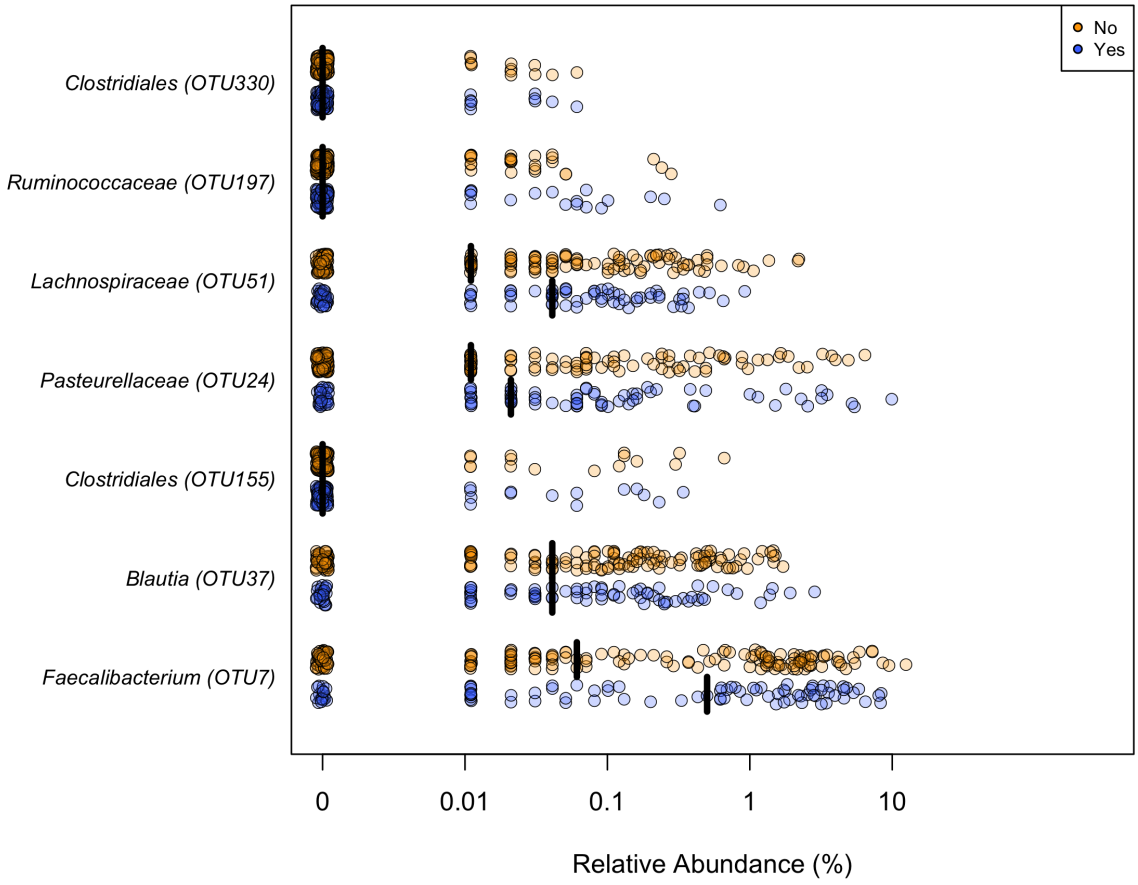

Figure 3: Abundance of frequently predictive OTUs in responders and remitters

348 **Figure 3: Abundance of frequently predictive OTUs in responders and remitters**

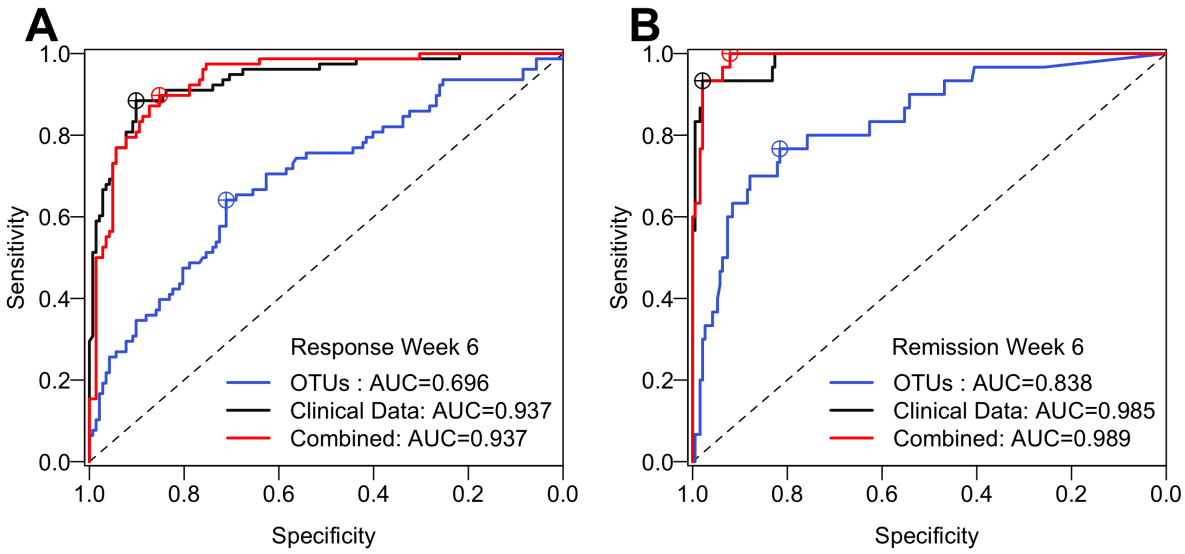

Figure 4: Supplemental Figure 1: Determining Week 6 disease status using Week 6 samples

Supplemental Figure 1: Determining Week 6 disease status using Week 6 samples

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. Mantzaris GJ, Viazis N, Polymeros D, Papamichael K, Bamias G, Koutroubakis IE. 2015. Clinical profiles of moderate and severe crohn's disease patients and use of anti-tumor necrosis factor agents: Greek expert consensus guidelines. Ann Gastroenterol 28:417–25.

5. 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.

6. Boyapati R, Satsangi J, Ho GT. 2015. Pathogenesis of crohn's disease. F1000Prime Rep 7:44.

7. 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.

8. 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.

9. 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.

10. 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.

11. 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.

12. 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.

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

14. 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.

15. Haag LM, Siegmund B. 2015. Intestinal microbiota and the innate immune system - a crosstalk in crohn's disease pathogenesis. Front Immunol 6:489.

16. 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.

17. 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.

18. Haberman Y, Tickle TL, Dexheimer PJ, Kim MO, Tang D, Karns R, Baldassano RN, Noe JD,

403 Rosh J, Markowitz J, Heyman MB, Griffiths AM, Crandall WV, Mack DR, Baker SS, Huttenhower C,
404 Keljo DJ, Hyams JS, Kugathasan S, Walters TD, Aronow B, Xavier RJ, Gevers D, Denson LA. 2014.
405 Pediatric crohn disease patients exhibit specific ileal transcriptome and microbiome signature. J Clin
406 Invest 124:3617–33.

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

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

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

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

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

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

428 25. Tedjo DI, Smolinska A, Savelkoul PH, Masclee AA, Schooten FJ van, Pierik MJ, Penders J, Jonkers
429 DMAE. 2016. The fecal microbiota as a biomarker for disease activity in crohn's disease. Scientific

Reports, Published online: 13 October 2016; doi:101038/srep35216.

26. 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.

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

28. 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.

29. 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.

30. 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.

31. 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.

32. 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.

33. 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.

34. 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.

35. Sartor RB, Wu GD. 2016. Roles for intestinal bacteria, viruses, and fungi in pathogenesis of

455 inflammatory bowel diseases and therapeutic approaches. Gastroenterology.

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

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

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

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

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

466 41. Sokal RR, Rohlf FJ. 1995. Biometry: The principles and practice of statistics in biological research,
467 3rd ed. Freeman, New York.

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

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

471 44. Schloss PD. 2008. Evaluating different approaches that test whether microbial communities have
472 the same structure. ISME J 2:265–75.

473 45. Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, Huttenhower C. 2011. Metage-
474 nomic biomarker discovery and explanation. Genome Biol 12:R60.

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