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

Running title: Microbiota of Ustekinumab-treated Crohn's 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: Patrick D. Schloss, 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

$\alpha$ Currently at Department of Biology, New York University, New York, NY, USA.

# Abstract

The fecal microbiota is a rich source of biomarkers that have previously been shown to be predictive of numerous disease states. Less well studied is whether these biomarkers can be predictive of response to therapy. This study sought to predict the therapeutic response of Ustekinumab (UST) treated Crohn's disease (CD) patients. Using stool samples collected over the course of 22 weeks, the composition of these patients' fecal bacterial communities was characterized by sequencing the 16S rRNA gene. Patients in remission could be distinguished from those with active disease 6 weeks post treatment induction using random forest models trained on the composition of their baseline microbiome and clinical metadata (AUC $= 0.844$, specificity $= 0.831$, sensitivity $= 0.774$). The most predictive OTUs, ubiquitous among patients, were affiliated with *Faecalibacterium* and *Escherichia/Shigella*. Among patients in remission 6 weeks post induction, the median baseline community diversity was 1.7 times higher than treated patients with active disease ($p = 0.020$). Their baseline community structures were also different ($p = 0.017$). Two OTUs affiliated with *Faecalibacterium* ($p = 0.0026$) and *Bacteroides* ($p = 0.022$) were significantly more abundant at baseline in patients who were in remission 6 weeks post induction than those with active CD . The diversity of UST treated clinical responders increased over the 22 weeks of the study, in contrast to nonresponsive patients ($p = 0.005$). The fecal microbiota at baseline was also associated with markers for disease severity.

*Importance:* The observed baseline differences in fecal microbiota and changes due to therapeutic response support using the microbiota as a biomarker for predicting a patient's response to UST. Finding prognostic biomarkers that give clinicians the ability to predict response to CD treatment at diagnosis will increase the likelihood of faster induction and maintenance of remission. OTUs associated with remission post treatment induction, especially *Faecalibacterium*, could be biomarkers for successful UST treatment of TNF-$\alpha$ refractory CD patients. More broadly, these results suggest the fecal microbiota could be a useful non-invasive biomarker for directing or monitoring the treatment of gastrointestinal diseases.

**Keywords: IBD, microbiome, biologics, prediction, biomarkers, remission, Stelara, machine learning**

## Introduction

The microbiome has been correlated with a variety of diseases and has shown promise as a predictive tool for disease outcome for gingivitis (1), cardiovascular disease (2), *Clostridium difficile* infection (3, 4), and colorectal cancer (5, 6). Additionally, the microbiome has been shown to alter the efficacy of vaginal microbicides (7), cardiac drugs (8), and cancer treatments (9, 10). These results strongly suggest that it is possible to use elements of the microbiome as a prognostic biomarker. In relation to inflammatory bowel disease (IBD), previous studies have shown that the bacterial gut microbiota correlates with disease severity in new-onset, pediatric Crohn's disease (CD) patients (11, 12). Additionally, recent studies have shown promise for using the gut microbiota as tool to predict therapeutic response in treating IBD (13, 14). It remains to be determined, however, whether the composition of the fecal gut microbiota can predict and monitor response to CD therapy. Considering the involvement of the immune system and previous evidence for involvement of the microbiome, it is likely that response to immunological CD therapy can be predicted using microbiome data.

CD is a global health concern causing large economic and healthcare impacts (15, 16). The disease is characterized by patches of ulceration and inflammation along the entire gastrointestinal tract, with most cases involving 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 (17, 18). Faster induction of remission following diagnosis reduces the risk of irreversible intestinal damage and disability (18–20). Ideally, clinicians would be able to determine personalized treatment options for CD patients at diagnosis that would result in faster achievement of remission (21). Therefore, recent research has been focused on identifying noninvasive biomarkers to monitor CD severity and predict therapeutic response (22–24).

The precise etiology of CD remains unknown, but host genetics, environmental exposure, and the gut microbiome appear to be involved (15, 25). Individuals with CD have reduced microbial diver-

sity in their guts, compared to healthy individuals, with a lower relative abundance of *Firmicutes* and an increased relative abundance of *Enterobacteraciae* and *Bacteroides* (11, 26–29). Additionally, genome-wide association studies of individuals with CD identified several susceptibility loci, including loci involved in the IL-23 signaling pathway, which could impact the gut microbiota composition and function (17, 26, 30–33). If the fecal microbiota can be used to monitor disease severity and predict response to specific treatment modalities, then clinicians could use it as a noninvasive tool for prescribing therapies that result in faster remission (34).

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 (18, 35–37). Given the potential impact of IL-23 on the microbiota (30–33), we hypothesized that response to UST could be predicted or influenced by differences in patients' gut microbiota and that UST treatment may alter the fecal microbiota. The effects of biologic treatment of IBD on the microbiota are not yet well described, but are hypothesized to be indirect, as these drugs act on host factors (17). We analyzed the fecal microbiota of patients who participated in a double-blinded, placebo-controlled Phase II clinical trial that demonstrated the safety and efficacy of UST for treating CD (35). The original study found that UST induction treatment had an increased rate of response as well as increased rates of response and remission with UST maintenance therapy, compared to placebo. We quantified the association between the fecal microbiota and disease severity. Finally, we tested whether clinical responders had a microbiota that was distinct from non-responders and if the fecal microbiota changed in patients treated with UST using 16S rRNA gene sequence data from these patients' stool samples. Our study demonstrates that these associations are useful in predicting and monitoring UST treatment outcome and suggest the fecal microbiota may be a broadly useful source of biomarkers for predicting response to treatment.

# Results

## Study design

We characterized the fecal microbiota in a subset of anti-TNF-$\alpha$ refractory CD patients, with moderate to severe CD, who took part in the double-blinded, CERTIFI clinical trial that demonstrated the efficacy of UST in treating CD (35). Demographic and baseline disease characteristics of this subset are summarized in Table 1. Patients were randomly assigned to a treatment group in the induction phase of the study and at week 8 patients were re-randomized into maintenance therapy groups based on their induction response (Figure 1A). In our study response was defined as a decrease in a patient's initial Crohn's Disease Activity Index (CDAI) greater than 30%. Remission was defined as a CDAI below 150 points. The CDAI is the standard instrument for evaluating clinical symptoms and disease activity in CD (38, 39). The CDAI weights patient reported stool frequency, abdominal pain, and general well being over a week, in combination with weight change, hematocrit, opiate usage for diarrhea, and the presence of abdominal masses or other complications to determine the disease severity score (38, 39). Patients provided stool samples at baseline (screening) and at 4, 6, and 22 weeks post induction for analysis using 16S rRNA gene sequencing (Figure 1B).

## Prediction of remission following treatment

We investigated whether the composition of the baseline fecal microbiota could predict therapeutic remission (CDAI $< 150$) 6 weeks post induction. To test this hypothesis, we generated prognostic random forest (RF) models to classify patients in remission from those with active CD 6 weeks post induction based on the relative abundance of the fecal microbiota at baseline, clinical metadata at baseline, and the combination of microbiota and clinical data. We determined the optimal model based the largest area under the curve (AUC) of the receiver operating characteristic (ROC) curve for the RF model (6, 40). Clinical data included components of the CDAI, biomarkers for inflammation, and patient metadata described further in the methods section. We ran these models on 232 baseline stool samples from patients induced with UST. Clinical data alone resulted in an AUC of 0.616 (specificity $= 0.801$, sensitivity $= 0.452$) (Figure 2A). Using only fecal microbiota data the model had an AUC of 0.838 (specificity $= 0.766$, sensitivity $= 0.806$).

Finally, when combining clinical metadata with the microbiota we achieved an AUC of 0.844 (specificity = 0.831, sensitivity = 0.774) for remission 6 weeks post induction. Prediction with clinical metadata alone did not perform as well as models using the baseline fecal microbiome ($p = 0.0011$) or the combined model ($p = 0.00087$); however, there was not a significant difference between the baseline fecal microbiota model and the combined model ($p = 0.84$).

Optimal predictors were determined based on their mean decrease in accuracy (MDA) in the ability of the model to classify remission from active CD (Figure 2B). The majority of OTUs identified as optimal predictors in our model for remission had low abundance. However, two OTUs were differentially abundant for patients in remission 6 weeks post induction. The relative abundance of *Escherichia/Shigella* (OTU1) was lower in remitters (median = 1.07%, IQR = 0.033-3.7) compared to patients with active CD (median = 4.13%, IQR = 0.667-15.4). Also, the relative abundance of *Faecalibacterium* (OTU7) was not only higher in remitters (median = 7.43%, IQR = 1.43-11.9) than patients with active CD (median = 0.167%, IQR = 0-5.1), it was also present prior to the start of treatment in every patient who was in remission 6 weeks post induction.

**Prediction of response following treatment**

We also hypothesized that the composition of the baseline fecal microbiota could predict therapeutic response (CDAI decrease >30%) 6 weeks post induction. To test this hypothesis, we again used RF models to classify responders from non-responders 6 weeks post induction. Clinical data alone resulted in an AUC of 0.693 (specificity = 0.76, sensitivity = 0.598) (Figure 2C). Using only microbiota data, the model predicted response with an AUC of 0.737 (specificity = 0.807, sensitivity = 0.585). When combining clinical metadata with the microbiome, the model predicted response with an AUC of 0.745 (specificity = 0.727, sensitivity = 0.744). These models were not significantly different in their ability to predict response ($p > 0.05$ for each comparison). Optimal predictors were again determined based on their MDA in the ability of the model to classify response (Figure 2D). Also, our baseline fecal microbiota model was significantly better

able to classify remission compared to response (p = 0.043), whereas this was not true for the combined model (p = 0.055).

## Comparison of baseline microbiota based on therapeutic outcome

As our random forest models identified OTUs abundant across our cohort that were important in classification of outcome, we further investigated differences in the baseline microbiota that could serve as potential biomarkers for successful UST treatment. We compared the baseline microbiomes of all 306 patients who provided a baseline sample based on treatment group and treatment outcome 6 weeks post induction. There was no significant difference in diversity based on response 6 weeks post induction, however the baseline $\beta$-diversity was significantly different by response (p = 0.012). No phyla were significantly different by treatment and response (Fig. S1) and no OTUs were significantly different based on UST response or among patients receiving placebo for induction, regardless of response and remission status.

Patients in remission 6 weeks post induction with UST had significantly higher baseline diversity, based on the inverse Simpson index, than patients with active CD (respective median values = 11.6 (IQR = 4.66-13.9), 6.95 (IQR = 4.4-11.8), p = 0.020). The baseline community structure was also significantly different based on remission status in patients 6 weeks post induction (p = 0.017). Finally, 2 OTUs were significantly more abundant in patients in remission 6 weeks post induction compared to patients with active CD; *Bacteroides* (OTU19) (p = 0.022) and *Faecalibacterium* (OTU7) (p = 0.0026) (Figure 3).

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

Based on the associations we identified between baseline microbial diversity and response, we hypothesized that there were associations between the microbiota and clinical variables at baseline that could support the use of the microbiota as a non-invasive biomarker for disease activity (34). To test this hypothesis, we compared the baseline microbiota with clinical data at baseline for all 306 samples provided at baseline (Supplemental Table 1). We observed small, but significant correlations for lower $\alpha$-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). Corticosteroid use was associated with 1.45 times higher $\alpha$-diversity (p = 0.001). No significant associations were observed between $\alpha$-diversity and CRP, fecal calprotectin, or fecal lactoferrin. However, the $\beta$-diversity was significantly different based on CRP (p = 0.033), fecal calprotectin (p = 0.006), and fecal lactoferrin (p = 0.004). The $\beta$-diversity was also significantly different based on weekly loose stool frequency (p= 0.024), age (p = 0.033), the tissue affected (p = 0.004), corticosteroid use (p =0.010), and disease duration (p = 0.004). No significant differences in the microbiota were observed for BMI, weight, or sex.

**The diversity of the microbiota changes in UST responders**

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

**The microbiota post induction can distinguish between treatment outcomes**

Having demonstrated the microbiome changes in patients who responded to UST treatment, we hypothesized that the microbiota could be used to monitor response to UST therapy by classifying patients based on disease activity (34). We again constructed a random forest classification model to distinguish between patients by UST treatment outcome based on their fecal microbiota (6, 40).

The study design resulted in only 75 week twenty-two stool samples from patients induced and maintained with UST, so we focused our analysis on the 220 week 6 stool samples from patients induced with UST. We were again better able to distinguish patients in remission from patients with active CD compared to responders from non-responders ($p = 0.0019$; Figure 5A). Our model using week 6 stool samples to determine response 6 weeks post induction could classify patients who responded from non-responders with an AUC of 0.708 (sensitivity $= 0.769$, specificity $= 0.606$). For determining remission 6 weeks post UST induction, the model had an AUC of 0.866 (sensitivity $= 0.833$, specificity $= 0.832$) when classifying patients in remission from patients with active CD. OTUs that were important for these classifications again included *Faecalibacterium* (OTU7), as well as *Blautia* (OTU124), *Clostridium XIVa* (OTU73), *Ruminococcaceae* (OTU53), and *Roseburia* (OTU12). These all had higher median relative abundance in patients in remission 6 weeks post induction than those with active disease (Figure 5B).

## Discussion

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

The development of prognostic and predictive models for disease or treatment outcome is anticipated to have a significant impact on clinical decision-making in healthcare (41). Prognostic models are statistical tools that predict outcome based on more than one aspect of patient data

(41). These models will help clinicians decide on the correct course of disease treatment or interventions for disease prevention with their patients. Additionally, patients will benefit with more individualized care that will potentially reduce adverse effects and result in faster recovery, reduce expenses from ineffective therapies, or increase quality of life by preventing disease in patients with high risk.

*The presented prognostic model is useful for biomarker discovery and hypothesis generation about the biology of CD as it relates to the microbiome. Similar models could be further developed into a clinically useful prognostic tool. Faecalibacterium* frequently occurred in our models. It is associated with health, comprising up to 5% of the relative abundance in healthy individuals, and has been shown to be rare in CD patients (26, 28, 42, 43). Each patient in remission six weeks post UST induction had measurable *Faecalibacterium* present at baseline. This supports the hypothesis that *Faecalibacterium* impacts CD pathogenesis. *Escherichia/Shigella* also occurred frequently in our models. This OTU is associated with inflammation and has been shown to be associated with CD pathogenesis (43). Many other taxa observed in our analysis had low abundance or were absent in the majority of patients. However, in many cases these taxa are related and may serve similar ecologic and metabolic roles in the gut environment. We hypothesize that these microbes may have genes that perform similar metabolic functions. Performing metagenomics on stool samples in future studies, especially in patients who achieve remission, could reveal these functions.

We were better able to predict whether a patient would go into remission rather than respond to treatment. We hypothesize that this is due to the subjective nature of the patient-reported CDAI factors and the relative nature of the response criteria compared to the threshold used to determine remission status. We defined response as a decrease in a patient's baseline CDAI of 30% or more, while remission was defined as a CDAI below 150. The original study used a decrease in CDAI of 100 points for their measure of response, but we felt using the relative percent change better represented a meaningful difference in disease activity and patient quality of life

(35). Additionally, the field appears to be moving away from CDAI and towards more objectively quantifiable biomarkers for inflammation as wells as endoscopic verification of mucosal healing. (19).

We observed several associations between the microbiota and clinical variables that could impact how CD is monitored and treated in the future. Serum CRP, fecal calprotectin, and fecal lactoferrin are widely used as biomarkers to measure inflammation and CD severity. We found that the microbial community structure is different among patients based on these markers. This supports the hypothesis that the fecal microbiota could function as a biomarker for measuring disease activity in patients, especially in concert with established inflammatory biomarkers (34, 44, 45). We also found that higher CDAI scores were associated with lower microbial diversity. This is consistent with other studies on the microbiota in individuals with CD compared to healthy individuals and studies looking at active disease compared to remission (11, 34, 46). However, the CDAI sub score of weekly stool frequency may have driven these differences (Supplementary Table 1), which is consistent with previous studies (47). We also observed differences in the microbial community structure based on disease localization, which is consistent with a study by Naftali et al (42). Our study also found that corticosteroid use impacts the composition of the human fecal microbiota, which is consistent with observations in mouse models (48). 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 (49–51).

We observed that the $\alpha$-diversity of clinical responders increased over time, in contrast to nonresponsive patients. This observation could be due to lower inflammation and changes in disease

activity corresponding to improved health in patients who responded to UST. We also addressed whether response to therapy can be predicted with the microbiota by developing a random-forest model that used relative microbial abundance data and/or clinical metadata for input. Again, we were better able to predict remission/non-remission than response/non-response. These findings are again consistent with other studies suggesting the microbiota could be useful as a biomarker in detecting remission versus active disease (34).

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

## Methods

### Study Design and Sample Collection

Janssen Research and Development conducted a placebo-controlled, phase II clinical study of approximately 500 patients to assess the safety and efficacy of UST for treating anti-TNF-$\alpha$ refractory, moderate to severe CD patients (35) (Figure 1). Institutional review board approval was acquired at each participating study center and patients provided written informed consent (35). Patient data was de-identified for our study. 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 treatment groups. Additional stool samples were provided 4 weeks post induction. At 6 weeks post induction an additional stool sample was collected, patients were scored for their response to UST based on CDAI, and then divided into groups receiving either subcutaneous injection of UST or placebo at weeks 8 and 16 as maintenance therapy. Response was defined as a decrease in a patient's initial CDAI of 30% or more. This value was determined by using the approximate percent change in CDAI from mild-moderate CD (220) to remission (150). Remission is defined as a CDAI below the threshold of 150. 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) and an EPMotion 5075 pipetting system (5, 6). The V4 region of the 16S rRNA gene from each sample was amplified and sequenced using the Illumina MiSeq™ platform (45). Sequences were curated as described previously using the mothur software package (v.1.34.4) (52, 53). Briefly, we curated the sequences to reduce sequencing and PCR errors (54), aligned the resulting sequences to the SILVA 16S rRNA sequence database (55), and used

UCHIME to remove any chimeric sequences (56). Sequences were clustered into operational taxonomic units (OTU) at a 97% similarity cutoff using the average neighbor algorithm (57). All sequences were classified using a naive Bayesian classifier trained against the RDP training set (version 14) and OTUs were assigned a classification based on which taxonomy had the majority consensus of sequences within a given OTU (58).

Following sequence curation using the mothur software package (52), we obtained a median of 13,732 sequences per sample (IQR = 7,863-21,978). Parallel sequencing of a mock community had an error rate of 0.017%. To limit effects of uneven sampling, we rarefied the dataset to 3,000 sequences per sample. Samples from patients that completed the clinical trial and had complete clinical metadata were included in our analysis. Of these samples, 306 were provided prior to treatment as well as 258 provided at week four, 289 at week six, and 205 at week twenty-two post-treatment, for a total of 1,058 samples. All fastq files and the MIMARKS spreadsheet with de-identified clinical metadata are available at **SRA**.

## Gut microbiota biomarker discovery and statistical analysis

R v.3.3.2 (2016-10-31) and mothur were used to analyze the data (59). To assess $\alpha$-diversity, the inverse Simpson index was calculated for each sample in the dataset. Spearman correlation tests were performed to compare the inverse Simpson index and continuous clinical data. Wilcoxon rank sum tests were performed for pairwise comparisons and Kruskal-Wallis rank sum tests for comparisons with more than two groups (60, 61). To measure $\beta$-diversity, the distance between samples was calculated using the thetaYC metric, which takes into account the types of bacteria and their abundance to calculate the differences between the communities (62). 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 (63), as well as the adonis function in the R package vegan (v.2.4.2) (64). Change in $\alpha$-diversity over time based on week twenty-two response was assessed using a Friedman test on patients who provided a sample at each time point (65). The Friedman test is a function in the R package stats

(v.3.3.2). Multiple comparisons following a Friedman test were performed using the friedmanmc function in the package pgirmess (v.1.6.5) (66). Change in beta-diversity over time by treatment group and response was assessed using the adonis function in vegan stratified by patient. We used the relative abundance of each OTU, inverse Simpson index, age, sex, current medications, BMI, disease duration, disease location, fecal calprotectin, fecal lactoferrin, C-reactive protein, bowel stricture, and CDAI sub scores as input into our RF models constructed with the AUCRF R package (v.1.1) (67), in order to identify phylotypes or clinical variables that distinguish between various treatment and response groups, as well as to predict or determine response outcome (68). Optimal predictors were determined based on their mean decrease in accuracy (MDA) of the model to classify patients. Differentially abundant OTUs and phyla were selected through comparison of clinical groups using Kruskal-Wallis and Wilcox tests, where appropriate, to identify OTUs/phyla where there was a p-value less than 0.05 following a Benjamini-Hochberg correction for multiple comparisons (69). Other R packages used in our analysis included ggplot2 v.2.2.1 (70), dplyr v.0.5.0 (71), pROC v.1.9.1 (72), knitr v.1.15.1 (73), gridExtra v.2.2.1 (74), devtools v.1.12.0 (75), knitcitations v.1.0.7 (76), scales v.0.4.1 (77), tidyr v.0.6.1 (78), Hmisc v.4.0.2 (79), and cowplot v.0.7.0 (80). A reproducible version of this analysis and manuscript are available at https://github.com/SchlossLab/Doherty_CDprediction_mBio_2017.

## Funding information.

Janssen Research and Development generously supported this study.

## Acknowledgements

We would like to thank N. Baxter, K. Iverson, K. Flynn, M. Sze, and N. Lesniak, as well as our collaborators at Janssen, for discussion of the work and comments on previous versions of the manuscript.

# Tables

## Table 1: Summary of clinical metadata of cohort 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 cohort at baseline**

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

# Figures

**Figure 1: Experimental design as adapted from Sandborne et al 2012.** (A) Participants were divided into treatment groups receiving placebo or UST by IV for induction. At week 8, patients were divided into groups receiving either subcutaneous injection of UST or placebo at weeks 8 and 16 as maintenance therapy, based on response at week 6. Finally, at 22 weeks patients were scored using CDAI for their response to therapy. (B) Stool sampling, treatment, and response evalution timeline. ↑, treatment administration; IV, intravenous; PE, primary endpoint; R, randomization; RR, rerandomization (only for subjects receiving UST induction therapy); SC, subcutaneous.

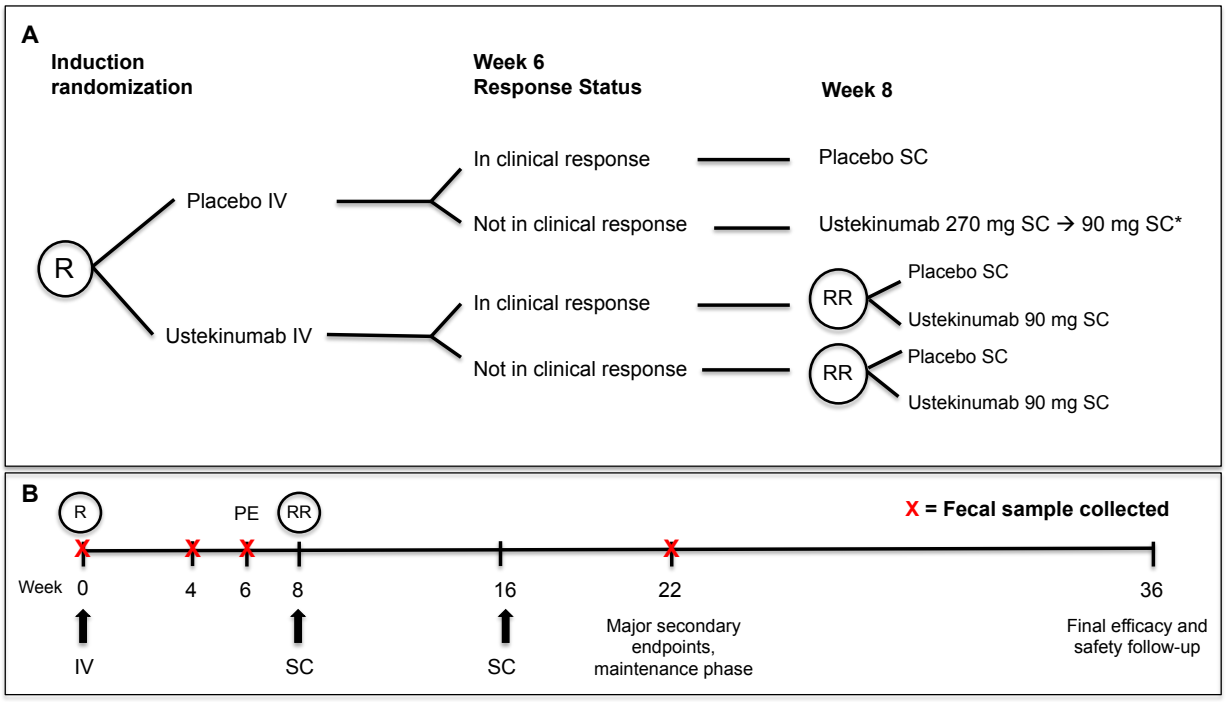
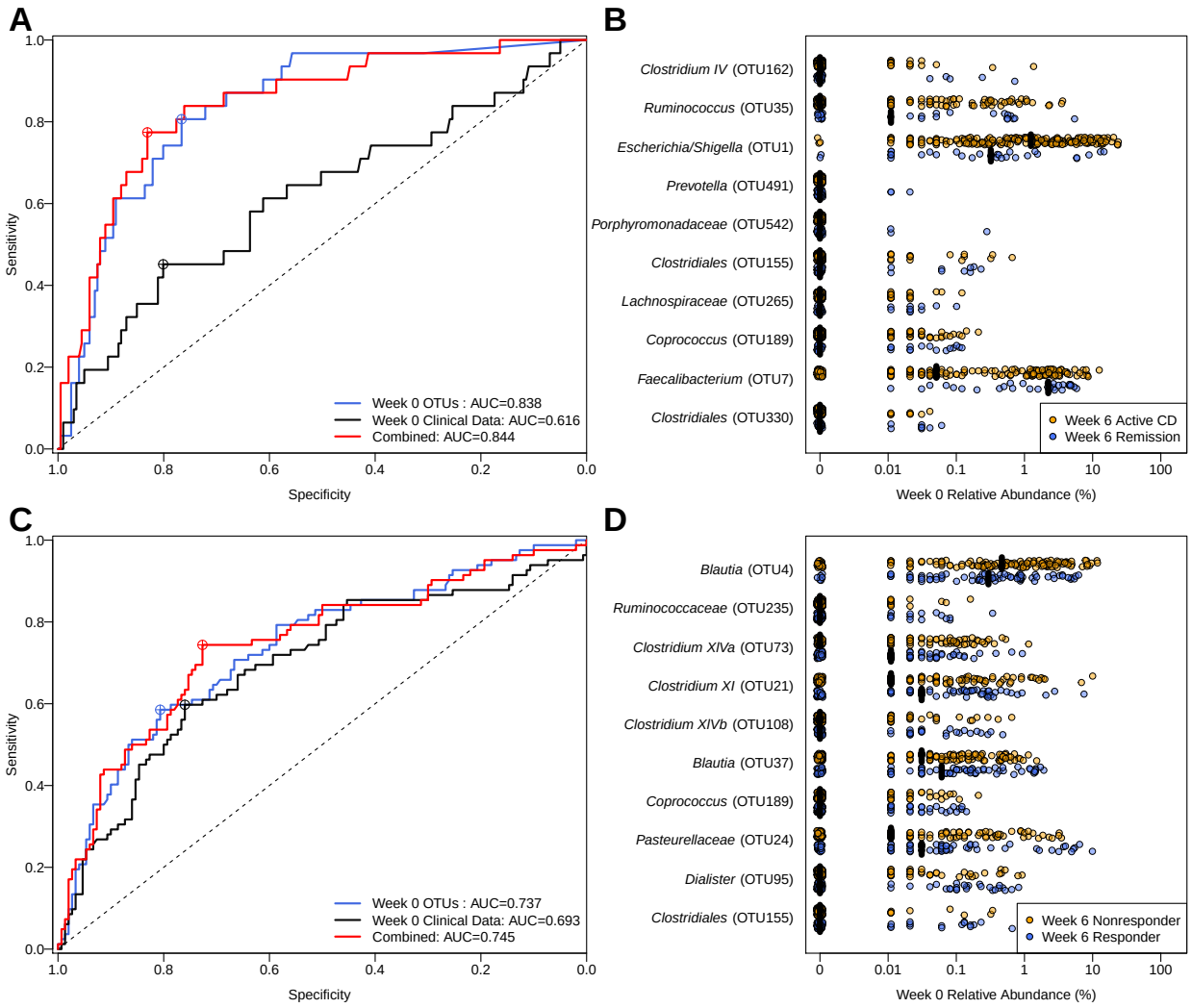


**Figure 2: Prediction of week 6 treatment outcome in patients treated with UST, using baseline samples** Receiver operating characteristic (ROC) curves for (A) response and (C) remission using microbiota data (blue), clinical metadata (black), and a combined model (red). Top predictive OTUs for the microbiota model based on mean decrease in accuracy (MDA) for (B) response and (D) remission. Black bars represent the median relative abundance.



**Supplemental Figure 1: Phyla from baseline stool samples in patients treated with UST by week six outcome** The relative abundance of each phylum in UST teated patients were compared based on (A) response and (B) remission status using a Wilcoxon rank sum test and to identify phyla where there was a p-value less than 0.05 following a Benjamini-Hochberg correction for multiple comparisons. No comparisons were significant. Whiskers represent the range and boxes represent the 25-75% interquartile range of the median (black bar).

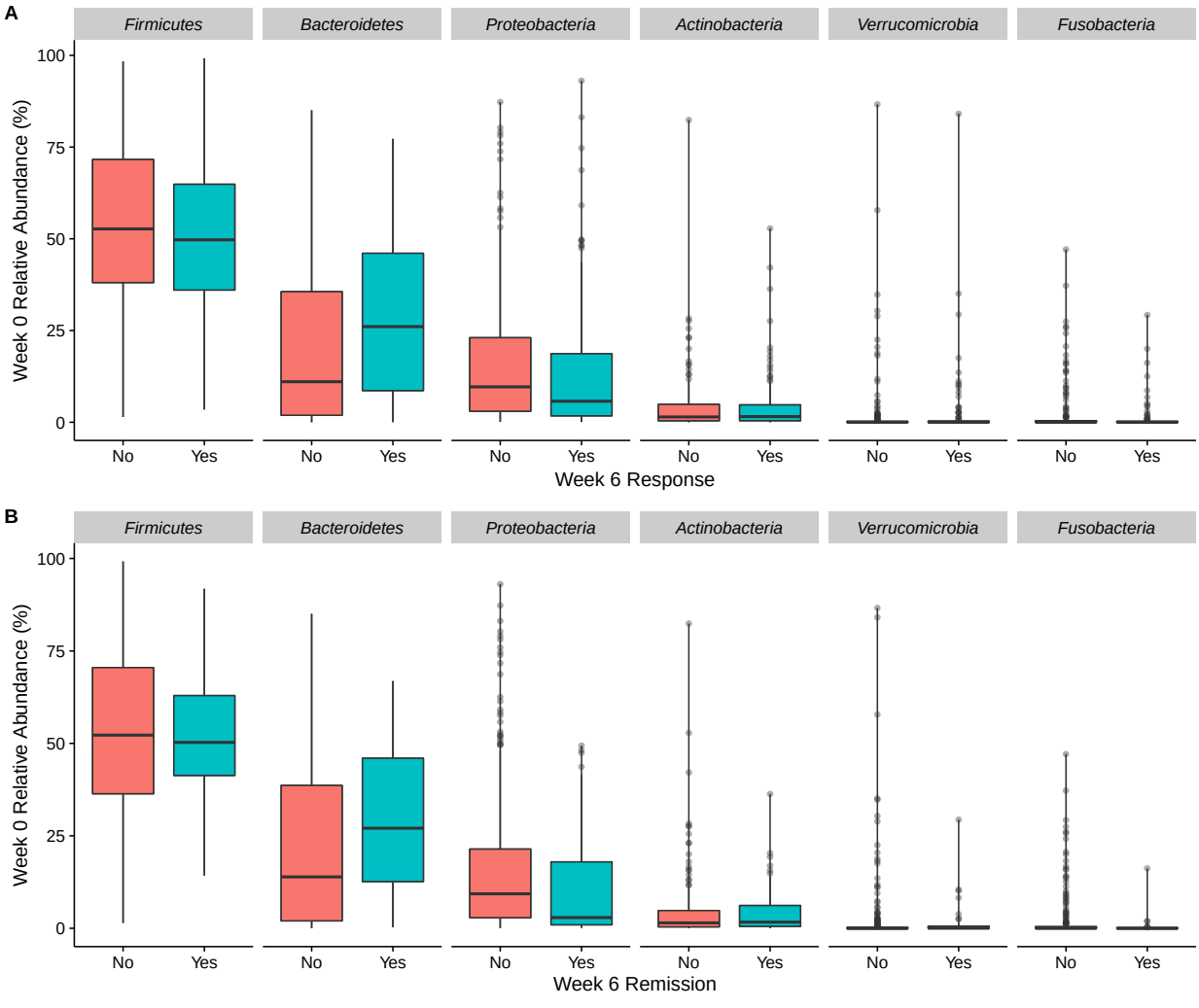


**Figure 3: Differential taxa in baseline stool samples from patients treated with UST, based on week six remission status** The baseline relative abundance of each OTU was compared between patients in remission and those with active CD 6 weeks post induction using a Wilcoxon rank sum test followed by a Benjamini-Hochberg correction for multiple comparisons. This identified 2 OTUs with significantly different relative abundance at baseline ($p < 0.05$). Black bars represent the median relative abundance.

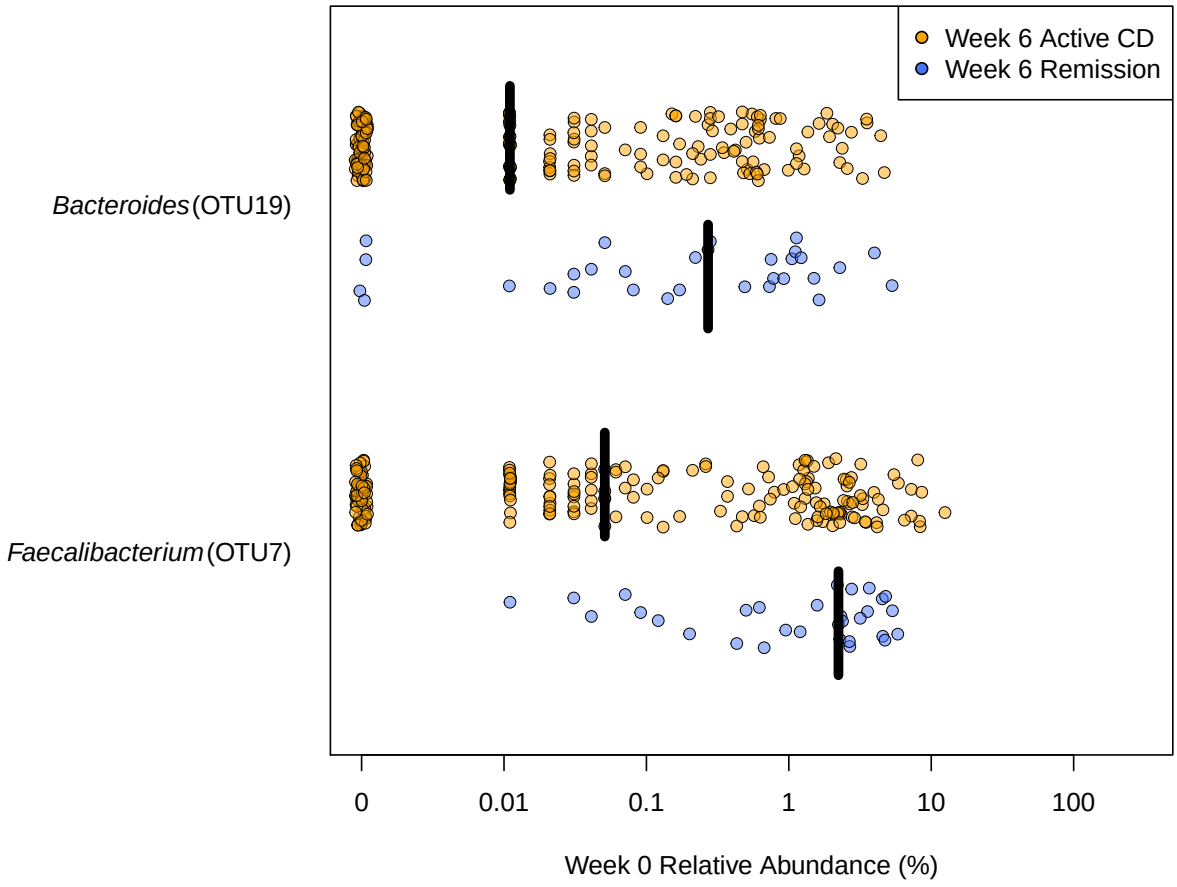


**Figure 4: Change in alpha diversity over time by induction treatment and week twenty-two response status.** The $\alpha$-diversity of 48 patients induced and maintained with UST and 14 patients induced and maintained with placebo was assessed at each time point. Friedman test were perfomed within each teatment and responder group. Whiskers represent the range. * indicates week 22 is significantly different from baseline ($p < 0.05$).

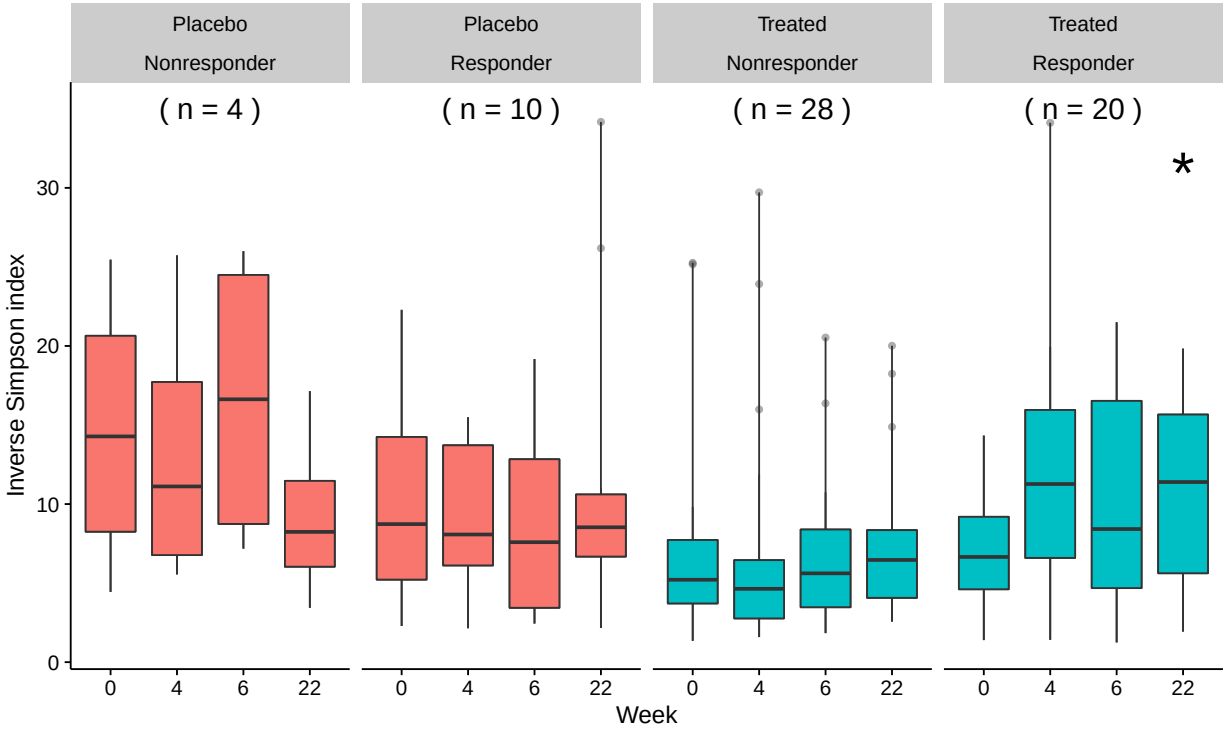
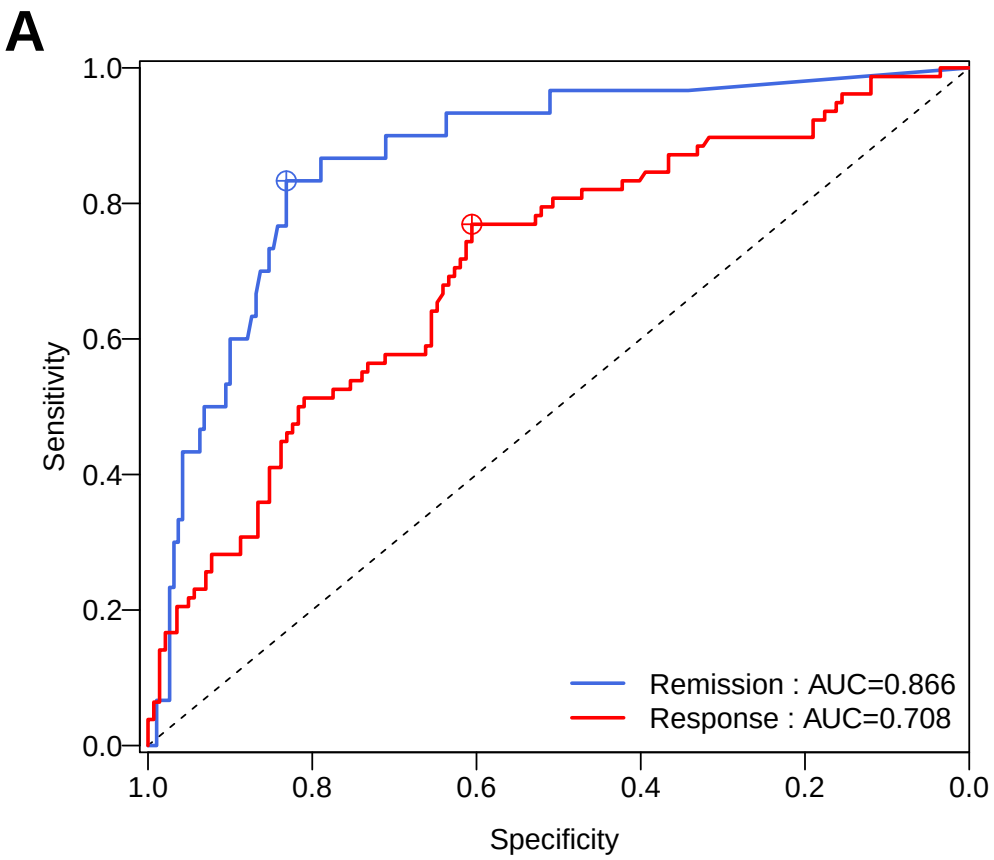
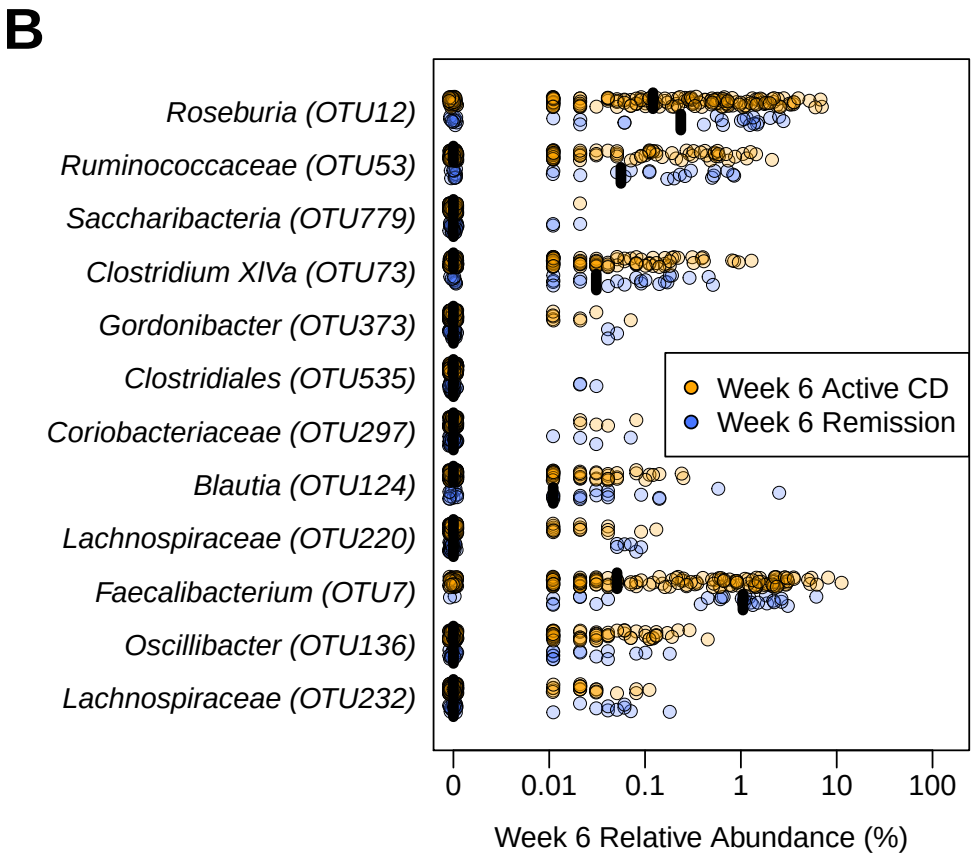


**Figure 5: Classification of week 6 response or remission status using week 6 stool samples from patients treated with UST** (A) ROC curves for week 6 outcome based on the week 6 microbiota. (B) Predictive OTUs from week 6 stool for remission status at 6 weeks post induction, based on mean decrease in accuracy. Black bars represent the median relative abundance.


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


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

409

# References


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

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

3. 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.

4. 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.

5. 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.

6. 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.

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

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

9. Sivan A, Corrales L, Hubert N, Williams JB, Aquino-Michaels K, Earley ZM, Benyamin FW, Lei

YM, Jabri B, Alegre ML, Chang EB, Gajewski TF. 2015. Commensal bifidobacterium promotes antitumor immunity and facilitates anti-pd-l1 efficacy. Science 350:1084–9.

10. Vetizou M, Pitt JM, Daillere R, Lepage P, Waldschmitt N, Flament C, Rusakiewicz S, Routy B, Roberti MP, Duong CP, Poirier-Colame V, Roux A, Becharef S, Formenti S, Golden E, Cording S, Eberl G, Schlitzer A, Ginhoux F, Mani S, Yamazaki T, Jacquelot N, Enot DP, Berard M, Nigou J, Opolon P, Eggermont A, Woerther PL, Chachaty E, Chaput N, Robert C, Mateus C, Kroemer G, Raoult D, Boneca IG, Carbonnel F, Chamaillard M, Zitvogel L. 2015. Anticancer immunotherapy by ctla-4 blockade relies on the gut microbiota. Science 350:1079–84.

11. 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.

12. 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.

13. 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.

14. Shaw KA, Bertha M, Hofmekler T, Chopra P, Vatanen T, Srivatsa A, Prince J, Kumar A, Sauer C, Zwick ME, Satten GA, Kostic AD, Mulle JG, Xavier RJ, Kugathasan S. 2016. Dysbiosis, inflammation, and response to treatment: A longitudinal study of pediatric subjects with newly

459 diagnosed inflammatory bowel disease. Genome Med 8:75.

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

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

465 17. Randall CW, Vizuete JA, Martinez N, Alvarez JJ, Garapati KV, Malakouti M, Taboada CM.
466 2015. From historical perspectives to modern therapy: A review of current and future biological
467 treatments for crohn's disease. Therap Adv Gastroenterol 8:143–59.

468 18. Wils P, Bouhnik Y, Michetti P, Flourie B, Brixi H, Bourrier A, Allez M, Duclos B, Grimaud
469 JC, Buisson A, Amiot A, Fumery M, Roblin X, Peyrin-Biroulet L, Filippi J, Bouguen G, Abitbol
470 V, Coffin B, Simon M, Laharie D, Pariente B. 2015. Subcutaneous ustekinumab provides clinical
471 benefit for two-thirds of patients with crohn's disease refractory to anti-tumor necrosis factor
472 agents. Clin Gastroenterol Hepatol.

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

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

481 21. Lichtenstein GR. 2010. Emerging prognostic markers to determine crohn's disease natural
482 history and improve management strategies: A review of recent literature. Gastroenterol Hepatol

(N Y) 6:99–107.

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

23. 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.

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

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

26. 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.

27. 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.

28. 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.

29. 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.

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

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

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

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

34. Tedjo DI, Smolinska A, Savelkoul PH, Masclee AA, Schooten FJ van, Pierik MJ, Penders J, Jonkers 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.

35. 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.

36. 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.

37. 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.

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

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

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

41. Vogenberg FR. 2009. Predictive and prognostic models: Implications for healthcare decision-making in a modern recession. Am Health Drug Benefits 2:218–22.

42. 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.

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

44. 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.

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

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

47. Vandeputte D, Falony G, Vieira-Silva S, Tito RY, Joossens M, Raes J. 2016. Original

555 article: Stool consistency is strongly associated with gut microbiota richness and composition,
556 enterotypes and bacterial growth rates. Gut 65:57–62.

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

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

564 50. Monteleone G, Di Sabatino A, Ardizzone S, Pallone F, Usiskin K, Zhan X, Rossiter G, Neurath
565 MF. 2016. Impact of patient characteristics on the clinical efficacy of mongersen (GED-0301),
566 an oral smad7 antisense oligonucleotide, in active crohn's disease. Aliment Pharmacol Ther
567 43:717–24.

568 51. Ardizzone S, Bevivino G, Monteleone G. 2016. Mongersen, an oral smad7 antisense oligonu-
569 cleotide, in patients with active crohn's disease. Therap Adv Gastroenterol 9:527–32.

570 52. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA,
571 Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Van Horn DJ, Weber CF.
572 2009. Introducing mothur: Open-source, platform-independent, community-supported software
573 for describing and comparing microbial communities. Appl Environ Microbiol 75:7537–41.

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

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

579 55. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO. 2013.

580 The silva ribosomal rna gene database project: Improved data processing and web-based tools.
581 Nucleic Acids Res 41:D590–6.

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

584 57. Schloss PD, Westcott SL. 2011. Assessing and improving methods used in operational
585 taxonomic unit-based approaches for 16S rRNA gene sequence analysis. Appl Environ Microbiol
586 77:3219–26.

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

589 59. R Core Team. 2016. R: A language and environment for statistical computing. R Foundation
590 for Statistical Computing, Vienna, Austria.

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

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

594 62. Yue JC, Clayton MK. 2005. A similarity measure based on species proportions. Communica-
595 tions in Statistics-Theory and Methods 34:2123–2131.

596 63. Schloss PD. 2008. Evaluating different approaches that test whether microbial communities
597 have the same structure. ISME J 2:265–75.

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

601 65. Friedman M. 1937. The use of ranks to avoid the assumption of normality implicit in the

602 analysis of variance. Journal of the American Statistical Association 32:675–701.

603 66. Giraudoux P. 2016. Pgirmess: Data analysis in ecology.

604 67. Urrea V, Calle M. 2012. AUCRF: Variable selection with random forest and the area under
605 the curve.

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

607 69. Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: A practical and powerful
608 approach to multiple testing. Journal of the Royal Statistical Society Series B (Methodological)
609 57:289–300.

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

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

612 72. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J-C, Müller M. 2011. PROC:
613 An open-source package for r and s+ to analyze and compare roc curves. BMC Bioinformatics
614 12:77.

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

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

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

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

620 77. Wickham H. 2016. Scales: Scale functions for visualization.

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

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

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