

<sup>1</sup>      The fecal microbiome as a tool for monitoring and  
<sup>2</sup>      predicting response outcomes in Ustekinumab-treated,  
<sup>3</sup>      anti-TNF-alpha refractory Crohn's Disease patients.

<sup>4</sup>      Running title: The fecal microbiome as a tool for monitoring and predicting response outcomes in  
<sup>5</sup>      Ustekinumab-treated, anti-TNF-alpha refractory Crohn's Disease patients.

<sup>6</sup>      Matthew K. Doherty<sup>2</sup>, Tao Ding<sup>2α</sup>, Charlie Koumpouras<sup>2</sup>, Shannon Telesco<sup>1</sup>, Calixte Monast<sup>1</sup>, and  
<sup>7</sup>      Patrick D. Schloss<sup>2†</sup>

<sup>8</sup>      † To whom correspondence should be addressed: pschloss@umich.edu

<sup>9</sup>      1. Janssen Pharmaceutical Companies of Johnson & Johnson, Spring House, PA, USA

<sup>10</sup>      2. Department of Microbiology and Immunology, University of Michigan, Ann Arbor, MI, USA

<sup>11</sup>      α Currently at ...

12 **Abstract**

13 *Background:* Crohn's disease (CD) is a global health issue characterized by patches of ulceration and  
14 inflammation along the gastrointestinal tract. Individuals with CD have reduced microbial diversity  
15 in their guts, compared to healthy individuals. It remains unclear if this reduced diversity is a result  
16 or cause of pathogenesis. We investigated the relationship between the fecal microbiome and clinical  
17 phenotypes in subjects with moderate to severe CD treated with Ustekinumab (UST) in a Phase 2b study  
18 to determine whether the fecal microbiome at baseline is predictive of disease severity and therapeutic  
19 response, as well as if the fecal microbiota changes due to therapy. (1–3)

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

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

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

<sup>40</sup> beneficial microbes associated with response to treatment could be developed as probiotics to increase  
<sup>41</sup> the likelihood of response while undergoing treatment.

<sup>42</sup> **Keywords:** Crohn's Disease, fecal microbiome, biologics, prediction

43 **Introduction**

44 Crohn's disease (CD), an incurable inflammatory bowel disease (IBD), is a global health issue  
45 with increasing incidence. CD affects approximately 3 million people worldwide, causing large  
46 economic and healthcare utilization impacts on society ???{Ananthakrishnan\_2015, ???{Floyd\_2015,  
47 ???{Molodecky\_2015 (1-3). CD is characterized by patches of ulceration and inflammation affecting  
48 the entire bowel wall along the gastrointestinal tract, most commonly in the ileum and colon. Individuals  
49 with CD experience frequent diarrhea, abdominal pain, fatigue, and weight loss resulting in significant  
50 health care costs, lower quality of life, and economic impacts due to loss of productivity (2, 4, 5).

51 Current treatments for CD include antibiotics, anti-inflammatory drugs, immunomodulators, surgery,  
52 and biologic agents targeting tumor necrosis factor alpha (TNF- $\alpha$ ), such as Infliximab (Remicade).

53 Within 10 years of diagnosis, approximately half of individuals with CD will require surgery and the  
54 majority will experience escalating immunosuppressive treatment (6). Currently, individuals with CD  
55 are treated based on disease location and risk of complications using escalating immunosuppressive  
56 treatment and/or surgery with the goal of achieving and sustaining remission (5, 7). Faster induction  
57 of remission following diagnosis reduces the risk of irreversible intestinal damage and disability (7-9).

58 Anti-TNF- $\alpha$  therapy in combination with thiopurines has emerged as the preferred treatment for CD,  
59 but up to half of individuals with CD fail to respond or lose response to anti-TNF- $\alpha$  therapy (6, 7).  
60 Ustekinumab (UST), a monoclonal antibody directed against the shared p40 subunit of IL-12 and  
61 IL-23, has been proposed as an alternative therapy for these patients (10). While clinical trials have  
62 demonstrated that UST is a viable option for the treatment of CD (7, 10-12), some patients within  
63 these trials were non-responsive to UST, which may be explained by differences in the patients' gut  
64 microbiomes.

65 The precise etiology of CD remains unknown, but host genetics, environmental exposure, and the gut  
66 microbiome appear involved (1, 13). Genome-wide association studies of individuals with CD identified  
67 several susceptibility genes including NOD2, a receptor involved in bacterial killing and innate immunity.

68 Defects in NOD2 function affects microbial sensing, the regulation of IL-23 driven Th17 responses,  
69 and indirect modulation of the gut microbiome (5, 14). The gut microbiome has also been shown to  
70 play a key role in inflammation, immunity, and IBD (15). Individuals with CD have reduced microbial  
71 diversity in their guts, compared to healthy individuals, with a lower relative abundance of Firmicutes

72 and an increased relative abundance of Enterobacteriacae and Bacteroides, at the phylum level (14,  
73 16-19). Additionally, previous studies have shown that the gut microbiome can be predictive of disease  
74 severity in new-onset, pediatric CD patients (19, 20). It remains to be determined, however, whether the  
75 microbiome can predict response to therapy in CD (14). Additionally, the effect of biologic treatment on  
76 the gut microbiome is not well understood. If the fecal microbiome can be used as a theraprognostic tool  
77 to non-invasively determine and monitor disease severity as well as predict response to specific treatment  
78 modalities, then more targeted treatment could result in reduced adverse effects of less effective therapies  
79 and faster achievement of remission.

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

## 93 Results

94 **Characteristics of Study Population** We studied the fecal microbiota in a subset of TNF- $\alpha$  refractory  
95 CD patients who took park in the CERTIFI clinical trial described in Sandborn et al 2012 (10). Briefly,  
96 patients with a history of moderate to severe CD were randomly assigned to a treatment group in the  
97 induction phase of the study. Subjects provided a stool sample at screening (Week 0), Week 4 and  
98 Week 6. At Week 8 patients were re-randomized into maintenance therapy groups. A final stool sample  
99 was provided at Week 22. Response to therapy was evaluated at week 4, 6, 8, and 22 based on change

100 in CDAI. Samples from subjects that completed the clinical trial and had complete clinical metadata  
101 were included in our analysis. We used 16s rRNA gene sequencing to analyze the microbiome from 306  
102 fecal samples provided prior to treatment as well as 258 Week 4, 289 Week 6, and 205 Week 22 post-  
103 treatment fecal samples, for a total of 1058 samples. Demographic and baseline disease characteristics  
104 are summarized in supplemental table 1.

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

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

127 **Comparison of clinical responders and non-responders** Next, We wanted to see if there were asso-  
128 ciations between the microbiome at baseline and response to treatment. For this study, response was

129 defined as a 30% decrease from CDAI at baseline and remission defined as a CDAI below 150. Of the  
130 306 screening samples analyzed, 232 were from subjects receiving UST and 74 from subjects receiving  
131 placebo. Baseline fecal microbiome community structures and species diversity were different among  
132 responders and non-responders to UST treatment. Based on response at the primary endpoint of the  
133 study, 6 weeks after IV induction, there was no difference in species richness between response groups,  
134 but there was a significant difference in the overall community structure of the entire cohort. This dif-  
135 ference in community structure was not significant in treatment vs. placebo groups. Week 6 remitters  
136 were significantly different from non-remitters in both species richness (0.0005) and overall community  
137 structure (0.017). When looking at treated vs. untreated Week 6 remitters, the treated group had  
138 significant differences in both species richness and community structure while untreated remitters we  
139 not different from untreated non-remitters. At the secondary endpoint, 22 weeks after IV-induction and  
140 14 weeks after maintenance dosing, there was no difference in species richness between response groups,  
141 but there was a significant difference in the overall community structure of the entire cohort. Week  
142 22 remitters were significantly different from non-remitters in both species richness (0.57) and overall  
143 community structure (0.016). However, these differences were not seen when the cohort was broken  
144 down by induction group. This could be due to changes in maintenance treatment.

145 **The microbiome by treatment and response over time** One major question with regards to biologic  
146 treatment of IBD and the microbiome is whether treatment has an effect on the microbiome. We  
147 explored this question 2 different ways. We included subjects that had stool samples at all 4 time points  
148 and another analysis using subjects who provided samples at weeks 0, 4, and 6. We used PERMANOVA  
149 stratified on each subject, as a proxy for a repeated measures ANOVA, to determine if the microbiome  
150 changed over time. We found that taken together treatment does not affect the microbiome. No  
151 significant difference was seen based on visit when looking at all groups and response status at week 4,  
152 6, or 8 over the first 3 time points, but there was a significant interaction between response at week 22  
153 and visit ( $p=0.001$ ) and between relative response, induction group, and visit ( $p=0.0445$ ).

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

158 groups, we observed a significant difference based on visit in the treated, Week 22 responder and in  
159 untreated responders across the first 3 visits prior to maintenance phase.

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

166 **Prediction of response based on the microbiome at screening**

167 Another major question in IBD and the microbiome is if response can be predicted using the microbiome.  
168 To address this we used AUCRF to develop a random forest classification model to differentiate respon-  
169 ders from non-responders, as well as remitters from non-remitters, based on the relative abundance of  
170 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  
171 optimal models for response and remission at the primary endpoint (Week 6) are shown in Figure 1.  
172 Using only clinical metadata to predict response, the model predicted response with an AUC of 0.665  
173 with a specificity of 0.813 and a sensitivity of 0.512. Using only microbiome data, the model predicted  
174 response with an AUC of 0.714 with a specificity of 0.82 and a sensitivity of 0.512. When combining  
175 clinical metadata with the microbiome, the model predicted response with an AUC of 0.682 with a  
176 specificity of 0.76 and a sensitivity of 0.561. With respect to Week 6 remission, using solely clinical  
177 metadata we achieved AUC of 0.637 with a specificity of 0.786 and a sensitivity of 0.452. Using only  
178 fecal microbiome data we achieved an AUC of 0.832 with a specificity of 0.627 and a sensitivity of 0.968.  
179 When combining clinical metadata with the microbiome AUC of 0.788 with a specificity of 0.697 and a  
180 sensitivity of 0.806.

182 Across all weeks and responses, prediction with clinical metadata alone did not perform as well as models  
183 using the fecal microbiome at screening. Also, combining microbiome data with clinical metadata did not  
184 consistently improve prediction compared to microbiome data alone. Additionally we found several OTUs  
185 occurred frequently across models including *Faecalibacterium*, among other taxa that were significantly

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

187 In addition to predicting future response, we wanted to determine if the microbiome could be used to  
188 monitor response to therapy. Again we used AUC-RF in order to determine if the fecal microbiome  
189 at Week 6 could be used to determine response or remission at Week 6. As seen in Supplemental  
190 Figure 1, using the microbiome alone we achieved an AUC of 0.696 for response with a sensitivity of  
191 0.641 and a specificity of 0.711. For remission we had an AUC of 0.838 with a sensitivity of 0.767  
192 and specificity of 0.816. Again we were better able to distinguish remitters from non-remitters than  
193 responders/non-responders. The clinical data were more reliable for determining disease activity at Week  
194 6.

## 195 Discussion

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

204 The association of the microbiome with clinically relevant biomarkers and disease activity metrics indi-  
205 cates that the microbiome may also function as a biomarker for CD activity. Given that serum CRP,  
206 calprotectin, and lactoferrin are used as biomarkers to measure intestinal inflammation and CD severity,  
207 it is interesting to see that the microbial community structure is different among patients based on  
208 these markers (26, 27). This supports the idea that the microbiome could be useful as a biomarker for  
209 measuring disease activity in patients, especially when considered in relation to these biomarkers (25).  
210 Higher CDAI was associated with lower microbial diversity. This appears to be consistent with other  
211 studies on the microbiome in individuals with CD compared to healthy individuals and studies looking  
212 at active disease compared to remission (19, 24, 25). However, these differences may have been driven

213 by weekly stool frequency, one component of the CDAI, where higher stool frequency is also negatively  
214 associated with microbial diversity. Given that higher stool frequency is associated with looser stool  
215 consistency, this finding appears consistent with the association between loose stools and lower diversity  
216 (28).

217 We also observed differences in the microbiome in relation to other clinical variables. The microbial  
218 community structure was different based on disease localization. These results are consistent with a  
219 study by Naftali et al finding distinct microbiotas for ileal versus colonic CD using mucosal tissue (29).  
220 This study also found that corticosteroid use impacts the composition of the human fecal microbiome.  
221 This supports data seen in the mouse model where corticosteroid injections altered the fecal mouse  
222 microbiome (30). As corticosteroid use appears to impact diversity, corticosteroids may be useful when  
223 trying to positively impact the microbiome during biologic therapy and increase the possibility of response  
224 to CD therapies.

225 Unlike other studies, these patients had a CD diagnosis for an average of 12 years (Supplemental Table  
226 1) (19, 24, 25). We observed that that longer disease duration is associated with a reduction in fecal  
227 microbial diversity. This decreased diversity may be due to the long duration of inflammatory conditions  
228 in the gut. One could hypothesize earlier biologic intervention may ‘preserve’ microbiome that promotes  
229 remission and reduces the likelihood of relapse. Publications have come out in support of earlier biologic  
230 intervention, as it appears to increase the likelihood of inducing remission and mucosal healing (31-33).  
231 However, the cost of biologics for patients is hindrance to early biologic intervention. Using aptamers in  
232 place of monoclonal antibodies may reduce this cost and make earlier intervention possible. Aptamers  
233 are short strands of DNA or RNA capable of specifically binding small molecules, proteins, and whole  
234 cells. Anti-TNF aptamers have been published that could potentially be used to test this in the mouse  
235 model (34).

236 One important question for the microbiome and IBD is whether or not the microbiome is affected by  
237 treatment with biologics. This study attempted to answer that question by looking at the microbiome  
238 of our CD subjects across multiple time points during treatment. While we were unable to see direct  
239 effects of the drug on the fecal microbiome, we observed that the microbiome of clinical responders  
240 changed over time, in contrast to nonresponsive subjects. This was observed for responsive patients  
241 regardless of induction treatment, leading us to think we are seeing the effects of change in disease

242 activity and health rather than any effects from treatment. This interpretation is consistent with studies  
243 using the microbiome to distinguish between remission and active CD (25). We did however observe a  
244 significant difference in community structure based on treatment and cannot eliminate the possibility of  
245 a direct effect on the microbiome in treated responders.

246 Another important question in for the importance of the microbiome in IBD is whether response to  
247 therapy can be predicted with the microbiome. We attempted to address this by developing a random-  
248 forest model that used relative microbial abundance data and/or clinical metadata for input. We found  
249 we were better able to predict remission status compared to response status. Response may be less  
250 predictable due to the “floating target” nature of a relative decrease in CDAI compared to the hard  
251 threshold for remission (CDAI<150). We were also better able to distinguish remission/non-remission  
252 than response/non-response, 6 weeks after beginning treatment. This is consistent with other studies  
253 again suggesting the microbiome could be useful in detecting remission versus active disease (25).

254 While using the presented model may not be useful clinically to predict response to therapy at this time,  
255 it is useful for hypothesis generation about the biology of CD as it relates to the microbiome. Some of  
256 the frequently occurring factors in our predictive models have already been linked to CD pathogenesis.  
257 As far as clinical biomarkers, fecal lactoferrin and fecal calprotectin occurred in the majority of models  
258 where clinical metadata was combined with the microbiome, supporting their importance as biomarkers  
259 for CD activity, especially in relation to the fecal microbiome (26, 27). *Faecalibacterium* was the most  
260 frequently occurring OTU in our models. It is associated with health and has been shown to be low  
261 in CD patients (14, 17, 29, 35). Remission was much more likely in individuals who had measurable  
262 *Faecalibacterium* present at baseline. This supports the hypothesis that *Faecalibacterium* impacts CD.  
263 *Escherichia/Shigella* also occurred frequently in our models. This OTU is associated with inflammation  
264 and has been shown to negatively impact CD (35). *Fusobacterium* also appeared in our predictive  
265 models and is associated with CD and CRC, something CD patients are more likely to get (35). These  
266 observations and the positive/negative associations of these microbes and CD allow us to hypothesize  
267 on ways to alter the microbiome to increase the likelihood therapeutic response. Prior to the initiation  
268 of therapy, patients could get a fecal microbiome analysis. The community data could then be used  
269 to direct the patient to undergo a round of antibiotics to target and reduce the levels of *Escherichia* in  
270 the patient’s gut. Alternatively, the microbes found to be positively associated with response could be

<sup>271</sup> formulated into a daily probiotic patients could take while receiving therapy with the goal of increasing  
<sup>272</sup> the likelihood of remission and mucosal healing.

<sup>273</sup> With this study we sought to gain a more detailed understanding of if and how biologic treatment affects  
<sup>274</sup> the microbiome, to determine whether the microbiome can be used to identify patients who will respond  
<sup>275</sup> to therapy, and to gain a better understanding of the interaction between the human gut microbiome  
<sup>276</sup> and CD pathogenesis in adult patients. We found the fecal microbiome to be useful in uncovering  
<sup>277</sup> associations between the microbiome and aspects of CD severity metrics and treatment outcomes. We  
<sup>278</sup> also demonstrated that the microbiome of treated responders changed over time, though it is not yet  
<sup>279</sup> possible to determine any direct effect of treatment on the microbiome. Finally, we were able to show  
<sup>280</sup> that the microbiome could be useful in predicting response to therapy, especially clinical remission,  
<sup>281</sup> compared to clinical metadata alone in our unique patient cohort. While this prediction is not clinically  
<sup>282</sup> useful as of yet, altering the weighting or binning of important factors in the model could make prediction  
<sup>283</sup> of response or remission more reliable. This could eventually allow for pre-screening of patients with  
<sup>284</sup> stool samples to predict successful treatment or better direct treatment. If the fecal microbiome can  
<sup>285</sup> be used as a therapgnostic tool to non-invasively predict response to specific treatment modalities  
<sup>286</sup> or inform treatment, then more personalized treatment could result in faster achievement of remission,  
<sup>287</sup> thereby increasing patients' quality of life and reducing economic and healthcare impacts.

288 **Methods**

289 **Study Design and Sample Collection**

290 Janssen Research and Development conducted a phase II clinical study of approximately 500 patients to  
291 assess the safety and efficacy of UST for treating anti-TNF- $\alpha$  refractory CD patients (10). Participants  
292 provided a stool sample prior to the initiation of the study and were then divided into 4 groups of  
293 125 individuals receiving placebo or 1, 3, or 6 mg/kg doses of UST by IV. Additional stool samples  
294 were provided at week 4. At week 6 an additional stool sample was collected, patients were scored for  
295 their response to UST based on CD Activity Index (CDAI), and divided into groups receiving either  
296 subcutaneous injection of UST or placebo at weeks 8 and 16 as maintenance therapy. Finally, at 22  
297 weeks patients provided an additional stool sample and were then scored using CDAI for their response  
298 to therapy. Frozen fecal samples were shipped to the University of Michigan and stored at -80°C prior  
299 to DNA extraction

300 **DNA extraction and 16S rRNA gene sequencing**

301 Microbial genomic DNA was extracted using the PowerSoil-htp 96 Well Soil DNA Isolation Kit (MoBio  
302 Laboratories) using an EPMotion 5075 pipetting system, as previously described (22, 36). The V4 region  
303 of the 16S rRNA gene from each sample was amplified and sequenced using the Illumina MiSeq Personal  
304 Sequencing platform as described elsewhere (27). Sequences were curated as described previously using  
305 the mothur software package (28) (37). Briefly, we reduced sequencing and PCR errors, aligned the  
306 resulting sequences to the SILVA 16S rRNA sequence database (29), and removed any chimeric sequences  
307 flagged by UCHIME (30) (38). After curation, we obtained between 1 and 130,074 sequences per sample  
308 (median 13786), with a median length of 253 bp. To limit effects of uneven sampling, we rarefied the  
309 dataset to 3,000 sequences per sample. Parallel sequencing of a mock community revealed an error rate  
310 of 0.017 %. Sequences were clustered into operational taxonomic units (OTU), as previously described  
311 (39). Briefly, OTUs were clustered at a 97% similarity cutoff and the relative abundance was calculated  
312 for OTUs in each sample. All sequences were classified using a naive Bayesian classifier trained against  
313 the RDP training set (version 11) and OTUs were assigned a classification based on which taxonomy had  
314 the majority consensus of sequences within a given OTU (31) (40). All fastq files and the MIMARKS

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

316 **Gut microbiome biomarker discovery analysis**

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

332 **Refs**

333 **References**

- 334 1. Ananthakrishnan AN. Epidemiology and risk factors for IBD. Nat Rev Gastroenterol Hepatol.  
335 2015;12(4):205-17. doi: 10.1038/nrgastro.2015.34. PubMed PMID: 25732745.
- 336 2. Floyd DN, Langham S, Severac HC, Levesque BG. The economic and quality-of-life burden of  
337 Crohn's disease in Europe and the United States, 2000 to 2013: a systematic review. Digestive  
338 diseases and sciences. 2015;60(2):299-312. Epub 2014/09/27. doi: 10.1007/s10620-014-3368-z.  
339 PubMed PMID: 25258034.
- 340 3. Molodecky NA, Soon IS, Rabi DM, Ghali WA, Ferris M, Chernoff G, Benchimol EI, Panaccione

- 341 R, Ghosh S, Barkema HW, Kaplan GG. Increasing incidence and prevalence of the inflammatory  
342 bowel diseases with time, based on systematic review. *Gastroenterology*. 2012;142(1):46-54.e42;  
343 quiz e30. Epub 2011/10/18. doi: 10.1053/j.gastro.2011.10.001. PubMed PMID: 22001864.
- 344 4. Mantzaris GJ, Viazis N, Polymeros D, Papamichael K, Bamias G, Kourloubakis IE. Clinical profiles  
345 of moderate and severe Crohn's disease patients and use of anti-tumor necrosis factor agents:  
346 Greek expert consensus guidelines. *Annals of gastroenterology : quarterly publication of the*  
347 *Hellenic Society of Gastroenterology*. 2015;28(4):417-25. Epub 2015/10/02. PubMed PMID:  
348 26424173; PMCID: PMC4585386.
- 349 5. Randall CW, Vizuete JA, Martinez N, Alvarez JJ, Garapati KV, Malakouti M, Taboada CM. From  
350 historical perspectives to modern therapy: a review of current and future biological treatments  
351 for Crohn's disease. *Therap Adv Gastroenterol*. 2015;8(3):143-59. Epub 2015/05/08. doi:  
352 10.1177/1756283x15576462. PubMed PMID: 25949527; PMCID: PMC4416294.
- 353 6. Boyapati R, Satsangi J, Ho GT. Pathogenesis of Crohn's disease. *F1000prime reports*. 2015;7:44.  
354 Epub 2015/06/23. doi: 10.12703/p7-44. PubMed PMID: 26097717; PMCID: PMC4447044.
- 355 7. Wils P, Bouhnik Y, Michetti P, Flourié B, Brixi H, Bourrier A, Allez M, Duclos B, Grimaud JC,  
356 Buisson A, Amiot A, Fumery M, Roblin X, Peyrin-Biroulet L, Filippi J, Bouguen G, Abitbol V, Cof-  
357 fin B, Simon M, Laharie D, Pariente B. Subcutaneous Ustekinumab Provides Clinical Benefit for  
358 Two-Thirds of Patients With Crohn's Disease Refractory to Anti-Tumor Necrosis Factor Agents.  
359 *Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gas-*  
360 *troenterological Association*. 2015. Epub 2015/10/04. doi: 10.1016/j.cgh.2015.09.018. PubMed  
361 PMID: 26432476.
- 362 8. Colombel JF, Reinisch W, Mantzaris GJ, Kornbluth A, Rutgeerts P, Tang KL, Oortwijn A, Beve-  
363 lander GS, Cornillie FJ, Sandborn WJ. Randomised clinical trial: deep remission in biologic and  
364 immunomodulator naive patients with Crohn's disease - a SONIC post hoc analysis. *Alimentary*  
365 *pharmacology & therapeutics*. 2015;41(8):734-46. Epub 2015/03/03. doi: 10.1111/apt.13139.  
366 PubMed PMID: 25728587.
- 367 9. Baert F, Moortgat L, Van Assche G, Caenepeel P, Vergauwe P, De Vos M, Stokkers P, Hommes  
368 D, Rutgeerts P, Vermeire S, D'Haens G. Mucosal healing predicts sustained clinical remission in  
369 patients with early-stage Crohn's disease. *Gastroenterology*. 2010;138(2):463-8; quiz e10-1. Epub  
370 2009/10/13. doi: 10.1053/j.gastro.2009.09.056. PubMed PMID: 19818785.

- 371 10. Sandborn WJ, Gasink C, Gao LL, Blank MA, Johanns J, Guzzo C, Sands BE, Hanauer SB, Targan  
372 S, Rutgeerts P, Ghosh S, de Villiers WJ, Panaccione R, Greenberg G, Schreiber S, Lichtiger S,  
373 Feagan BG. Ustekinumab induction and maintenance therapy in refractory Crohn's disease. N  
374 Engl J Med. 2012;367(16):1519-28. Epub 2012/10/19. doi: 10.1056/NEJMoa1203572. PubMed  
375 PMID: 23075178.
- 376 11. Sandborn WJ, Feagan BG, Fedorak RN, Scherl E, Fleisher MR, Katz S, Johanns J, Blank M,  
377 Rutgeerts P. A randomized trial of Ustekinumab, a human interleukin-12/23 monoclonal antibody,  
378 in patients with moderate-to-severe Crohn's disease. Gastroenterology. 2008;135(4):1130-41.  
379 Epub 2008/08/19. doi: 10.1053/j.gastro.2008.07.014. PubMed PMID: 18706417.
- 380 12. Kopylov U, Afif W, Cohen A, Bitton A, Wild G, Bessisow T, Wyse J, Al-Taweel T, Szilagyi A,  
381 Seidman E. Subcutaneous ustekinumab for the treatment of anti-TNF resistant Crohn's disease—  
382 the McGill experience. Journal of Crohn's & colitis. 2014;8(11):1516-22. Epub 2014/07/06. doi:  
383 10.1016/j.crohns.2014.06.005. PubMed PMID: 24996483.
- 384 13. Sartor RB. Mechanisms of disease: pathogenesis of Crohn's disease and ulcerative colitis. Nature  
385 clinical practice Gastroenterology & hepatology. 2006;3(7):390-407. Epub 2006/07/05. doi:  
386 10.1038/ncpgasthep0528. PubMed PMID: 16819502.
- 387 14. Wright EK, Kamm MA, Teo SM, Inouye M, Wagner J, Kirkwood CD. Recent advances in charac-  
388 terizing the gastrointestinal microbiome in Crohn's disease: a systematic review. Inflamm Bowel  
389 Dis. 2015;21(6):1219-28. Epub 2015/04/07. doi: 10.1097/mib.0000000000000382. PubMed  
390 PMID: 25844959; PMCID: PMC4450900.
- 391 15. Haag LM, Siegmund B. Intestinal Microbiota and the Innate Immune System - A Crosstalk  
392 in Crohn's Disease Pathogenesis. Front Immunol. 2015;6:489. Epub 2015/10/07. doi:  
393 10.3389/fimmu.2015.00489. PubMed PMID: 26441993; PMCID: PMC4585200.
- 394 16. Manichanh C, Rigottier-Gois L, Bonnaud E, Gloux K, Pelletier E, Frangeul L, Nalin R, Jarrin  
395 C, Chardon P, Marteau P, Roca J, Dore J. Reduced diversity of faecal microbiota in Crohn's  
396 disease revealed by a metagenomic approach. Gut. 2006;55(2):205-11. Epub 2005/09/29. doi:  
397 10.1136/gut.2005.073817. PubMed PMID: 16188921; PMCID: PMC1856500.
- 398 17. Hansen R, Russell RK, Reiff C, Louis P, McIntosh F, Berry SH, Mukhopadhyay I, Bisset WM,  
399 Barclay AR, Bishop J, Flynn DM, McGrogan P, Loganathan S, Mahdi G, Flint HJ, El-Omar EM,  
400 Hold GL. Microbiota of de-novo pediatric IBD: increased Faecalibacterium prausnitzii and reduced

- 401 bacterial diversity in Crohn's but not in ulcerative colitis. *Am J Gastroenterol.* 2012;107(12):1913-  
402 22. Epub 2012/10/10. doi: 10.1038/ajg.2012.335. PubMed PMID: 23044767.
- 403 18. Haberman Y, Tickle TL, Dexheimer PJ, Kim MO, Tang D, Karns R, Baldassano RN, Noe JD,  
404 Rosh J, Markowitz J, Heyman MB, Griffiths AM, Crandall WV, Mack DR, Baker SS, Huttenhower  
405 C, Keljo DJ, Hyams JS, Kugathasan S, Walters TD, Aronow B, Xavier RJ, Gevers D, Denson LA.  
406 Pediatric Crohn disease patients exhibit specific ileal transcriptome and microbiome signature. *J  
407 Clin Invest.* 2014;124(8):3617-33. Epub 2014/07/09. doi: 10.1172/jci75436. PubMed PMID:  
408 25003194; PMCID: PMC4109533.
- 409 19. Gevers D, Kugathasan S, Denson LA, Vazquez-Baeza Y, Van Treuren W, Ren B, Schwager E,  
410 Knights D, Song SJ, Yassour M, Morgan XC, Kostic AD, Luo C, Gonzalez A, McDonald D, Haber-  
411 man Y, Walters T, Baker S, Rosh J, Stephens M, Heyman M, Markowitz J, Baldassano R, Griffiths  
412 A, Sylvester F, Mack D, Kim S, Crandall W, Hyams J, Huttenhower C, Knight R, Xavier RJ. The  
413 treatment-naive microbiome in new-onset Crohn's disease. *Cell Host Microbe.* 2014;15(3):382-  
414 92. Epub 2014/03/19. doi: 10.1016/j.chom.2014.02.005. PubMed PMID: 24629344; PMCID:  
415 PMC4059512.
- 416 20. Wang F, Kaplan JL, Gold BD, Bhasin MK, Ward NL, Kellermayer R, Kirschner BS, Heyman  
417 MB, Dowd SE, Cox SB, Dogan H, Steven B, Ferry GD, Cohen SA, Baldassano RN, Moran CJ,  
418 Garnett EA, Drake L, Otu HH, Mirny LA, Libermann TA, Winter HS, Korolev KS. Detecting  
419 Microbial Dysbiosis Associated with Pediatric Crohn Disease Despite the High Variability of the  
420 Gut Microbiota. *Cell reports.* 2016. Epub 2016/01/26. doi: 10.1016/j.celrep.2015.12.088.  
421 PubMed PMID: 26804920.
- 422 21. Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara  
423 RB, Simpson GL, Solymos P, Stevens MHH, Szoecs E, Wagner H. vegan: Community Ecology  
424 Package. R package version 2.4-12016. doi: <https://CRAN.R-project.org/package=vegan>.
- 425 22. Baxter NT, Ruffin MT, Rogers MA, Schloss PD. Microbiota-based model improves the sensitivity  
426 of fecal immunochemical test for detecting colonic lesions. *Genome medicine.* 2016;8(1):37.  
427 Epub 2016/04/09. doi: 10.1186/s13073-016-0290-3. PubMed PMID: 27056827; PMCID:  
428 PMC4823848.
- 429 23. Calle ML, Urrea V, Boulesteix A-L, Malats N. AUC-RF: A New Strategy for Genomic Profiling  
430 with Random Forest. *Human Heredity.* 2011;72(2):121-32. doi: 10.1159/000330778. PubMed

- 431 PMID: 21996641.
- 432 24. Papa E, Docktor M, Smillie C, Weber S, Preheim SP, Gevers D, Giannoukos G, Ciulla D, Tabbaa  
433 D, Ingram J, Schauer DB, Ward DV, Korzenik JR, Xavier RJ, Bousvaros A, Alm EJ. Non-invasive  
434 mapping of the gastrointestinal microbiota identifies children with inflammatory bowel disease.  
435 PLoS One. 2012;7(6):e39242. Epub 2012/07/07. doi: 10.1371/journal.pone.0039242. PubMed  
436 PMID: 22768065; PMCID: PMC3387146.
- 437 25. Tedjo DI, Smolinska A, Savelkoul PH, Masclee AA, Schooten FJv, Pierik MJ, Penders J, Jonkers  
438 DMAE. The fecal microbiota as a biomarker for disease activity in Crohn's disease. Scientific  
439 Reports, Published online: 13 October 2016; | doi:101038/srep35216. 2016. doi: doi:10.1038/  
440 srep35216.
- 441 26. Boon GJ, Day AS, Mulder CJ, Gearry RB. Are faecal markers good indicators of mucosal healing  
442 in inflammatory bowel disease? World J Gastroenterol. 2015;21(40):11469-80. Epub 2015/11/03.  
443 doi: 10.3748/wjg.v21.i40.11469. PubMed PMID: 26523111; PMCID: PMC4616222.
- 444 27. Chang S, Malter L, Hudesman D. Disease monitoring in inflammatory bowel disease. World  
445 J Gastroenterol. 2015;21(40):11246-59. Epub 2015/11/03. doi: 10.3748/wjg.v21.i40.11246.  
446 PubMed PMID: 26523100; PMCID: PMC4616202.
- 447 28. Vandepitte D, Falony G, Vieira-Silva S, Tito RY, Joossens M, Raes J. Original article: Stool  
448 consistency is strongly associated with gut microbiota richness and composition, enterotypes and  
449 bacterial growth rates. Gut. 2016;65(1):57-62. doi: 10.1136/gutjnl-2015-309618. PubMed  
450 PMID: 26069274.
- 451 29. Naftali T, Reshef L, Kovacs A, Porat R, Amir I, Konikoff FM, Gophna U. Distinct Microbiotas  
452 are Associated with Ileum-Restricted and Colon-Involving Crohn's Disease. Inflamm Bowel Dis.  
453 2016;22(2):293-302. Epub 2016/01/12. doi: 10.1097/mib.0000000000000662. PubMed PMID:  
454 26752462.
- 455 30. Huang EY, Inoue T, Leone VA, Dalal S, Touw K, Wang Y, Musch MW, Theriault B, Higuchi K,  
456 Donovan S, Gilbert J, Chang EB. Using corticosteroids to reshape the gut microbiome: implica-  
457 tions for inflammatory bowel diseases. Inflamm Bowel Dis. 2015;21(5):963-72. Epub 2015/03/05.  
458 doi: 10.1097/mib.0000000000000332. PubMed PMID: 25738379; PMCID: PMC4402247.
- 459 31. Monteleone G, Neurath MF, Ardizzone S, Di Sabatino A, Fantini MC, Castiglione F, Scribano ML,  
460 Armuzzi A, Caprioli F, Sturniolo GC, Rogai F, Vecchi M, Atreya R, Bossa F, Onali S, Fichera M,

- 461 Corazza GR, Biancone L, Savarino V, Pica R, Orlando A, Pallone F. Mongersen, an oral SMAD7  
462 antisense oligonucleotide, and Crohn's disease. *N Engl J Med.* 2015;372(12):1104-13. Epub  
463 2015/03/19. doi: 10.1056/NEJMoa1407250. PubMed PMID: 25785968.
- 464 32. Monteleone G, Di Sabatino A, Ardizzone S, Pallone F, Usiskin K, Zhan X, Rossiter G, Neurath MF.  
465 Impact of patient characteristics on the clinical efficacy of mongersen (GED-0301), an oral Smad7  
466 antisense oligonucleotide, in active Crohn's disease. *Alimentary pharmacology & therapeutics.*  
467 2016;43(6):717-24. Epub 2016/01/15. doi: 10.1111/apt.13526. PubMed PMID: 26766141;  
468 PMCID: PMC4849204.
- 469 33. Ardizzone S, Bevivino G, Monteleone G. Mongersen, an oral Smad7 antisense oligonucleotide, in  
470 patients with active Crohn's disease. *Therap Adv Gastroenterol* 2016. p. 527-32.
- 471 34. Orava EW, Jarvik N, Shek YL, Sidhu SS, Gariepy J. A short DNA aptamer that recognizes  
472 TNFalpha and blocks its activity in vitro. *ACS Chem Biol.* 2013;8(1):170-8. Epub 2012/10/11.  
473 doi: 10.1021/cb3003557. PubMed PMID: 23046187.
- 474 35. Sartor RB, Wu GD. Roles for Intestinal Bacteria, Viruses, and Fungi in Pathogenesis of Inflam-  
475 matory Bowel Diseases and Therapeutic Approaches. *Gastroenterology.* 2016. Epub 2016/10/23.  
476 doi: 10.1053/j.gastro.2016.10.012. PubMed PMID: 27769810.
- 477 36. Zackular JP, Rogers MA, Ruffin MTt, Schloss PD. The human gut microbiome as a screening  
478 tool for colorectal cancer. *Cancer prevention research (Philadelphia, Pa).* 2014;7(11):1112-21.  
479 Epub 2014/08/12. doi: 10.1158/1940-6207.capr-14-0129. PubMed PMID: 25104642; PMCID:  
480 PMC4221363.
- 481 37. Schloss PD, Gevers D, Westcott SL. Reducing the effects of PCR amplification and sequencing  
482 artifacts on 16S rRNA-based studies. *PLoS One.* 2011;6(12):e27310. Epub 2011/12/24. doi:  
483 10.1371/journal.pone.0027310. PubMed PMID: 22194782; PMCID: PMC3237409.
- 484 38. Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R. UCHIME improves sensitivity and speed  
485 of chimaera detection. *Bioinformatics.* 2011;27(16):2194-200. Epub 2011/06/28. doi: btr381 [pii]  
486 10.1093/bioinformatics/btr381. PubMed PMID: 21700674; PMCID: 3150044.
- 487 39. Schloss PD, Westcott SL. Assessing and improving methods used in operational taxonomic unit-  
488 based approaches for 16S rRNA gene sequence analysis. *Applied and environmental microbiology.*  
489 2011;77(10):3219-26. doi: 10.1128/AEM.02810-10. PubMed PMID: 21421784; PMCID:  
490 3126452.

- 491 40. Wang Q, Garrity GM, Tiedje JM, Cole JR. Naive Bayesian classifier for rapid assignment of  
492 rRNA sequences into the new bacterial taxonomy. *Applied and environmental microbiology*.  
493 2007;73(16):5261-7. doi: 10.1128/AEM.00062-07. PubMed PMID: 17586664.
- 494 41. Sokal RR, Rohlf FJ. *Biometry: the principles and practice of statistics in biological research*. 3rd  
495 ed. New York: Freeman; 1995. xix, 887 p.
- 496 42. Magurran AE. *Measuring biological diversity*. Malden, Ma.: Blackwell Pub.; 2004.
- 497 43. Yue JC, Clayton MK. A similarity measure based on species proportions. *Communications in  
498 Statistics-Theory and Methods*. 2005;34(11):2123-31. PubMed PMID: ISI:000233019600005.
- 499 44. Schloss PD. Evaluating different approaches that test whether microbial communities have the  
500 same structure. *ISME J*. 2008;2(3):265-75. PubMed PMID: 18239608.
- 501 45. Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, Huttenhower C. Metagenomic  
502 biomarker discovery and explanation. *Genome biology*. 2011;12(6):R60. doi: 10.1186/gb-2011-  
503 12-6-r60. PubMed PMID: 21702898; PMCID: 3218848.
- 504 46. Breiman L. Random forests. *Machine Learning*. 2001;45(1):5-32. PubMed PMID:  
505 ISI:000170489900001.

506 **Tables**507 **Supplemental Table 1: Summary of clinical metadata of cohort 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

508

509 **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	Spearman p=0.005 ( $\rho=-0.2$ )	p=0.3
Loose Stool Frequency (per week)	Min=2, Median=51, Max=100	Spearman p=7e-04 ( $\rho=-0.2$ )	p=0.01
C-Reactive Protein (mg/L serum)	Min=0.1, Median=11.7, Max=199	Spearman p=0.3 ( $\rho=0.06$ )	p=0.04
Fecal Calprotectin ( $\mu\text{g/g}$ )	Min=14, Median=582.5, Max=26070	Spearman p=0.1 ( $\rho=0.08$ )	p=0.001
Fecal Lactoferrin ( $\mu\text{g/g}$ )	Min=0.25, Median=83.78, Max=3141	Spearman p=0.03 ( $\rho=0.1$ )	p=0.001
BMI	Min=15, Median=24, Max=55.3	Spearman p=0.2 ( $\rho=0.07$ )	p=0.2
Weight (kg)	Min=40, Median=69, Max=150	Spearman p=0.2 ( $\rho=0.07$ )	p=0.09
Age (years)	Min=18, Median=37, Max=76	Spearman p=0.4 ( $\rho=-0.05$ )	p=0.02
Sex	F=189, M=117	Wilcoxon p=0.5	p=0.2
Corticosteroid Use	No=174, Yes=132	Wilcoxon p=2e-04	p=0.003
Disease Duration (years)	Min=0.48, Median=10.44, Max=44.92	Spearman p=7e-05 ( $\rho=-0.2$ )	p=0.001
Tissue Involvement	Colon=85, Colon–ileum=148, Ileum=73	Kruskal–Wallis p=0.1	p=0.001

510

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

Status Week	Summary	Species Richness (Alpha-diversity)	Community Structure (beta-diversity)
ReIRSPwk6	No=196, Yes=110	p=0.11	p=0.012
Treated	No=150, Yes=82	p=0.2	p=0.066
Placebo	No=46, Yes=28	p=0.37	p=0.073
REMISSwk6	No=263, Yes=43	p=5e-04	p=0.017
Treated	No=201, Yes=31	p=0.002	p=0.022
Placebo	No=62, Yes=12	p=0.11	p=0.26
ReIRSPwk22	No=186, Yes=120	p=0.57	p=0.016
Treated	No=144, Yes=88	p=0.63	p=0.12
Placebo	No=42, Yes=32	p=0.87	p=0.058
REMISSwk22	No=250, Yes=56	p=0.019	p=0.007
Treated	No=196, Yes=36	p=0.056	p=0.089
Placebo	No=54, Yes=20	p=0.24	p=0.11

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

	Occurance (out of 24)
<i>Faecalibacterium</i> (OTU7)	19
<i>Ruminococcus</i> (OTU35)	16
<i>Clostridium_IV</i> (OTU222)	15
<i>Clostridium_XI</i> (OTU21)	11
<i>Ruminococcaceae</i> (OTU53)	9
<i>Coprococcus</i> (OTU189)	9
<i>Bacteria</i> (OTU440)	9
<i>Escherichia/Shigella</i> (OTU1)	8
<i>Bifidobacterium</i> (OTU8)	8

515

516 **Figures**

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

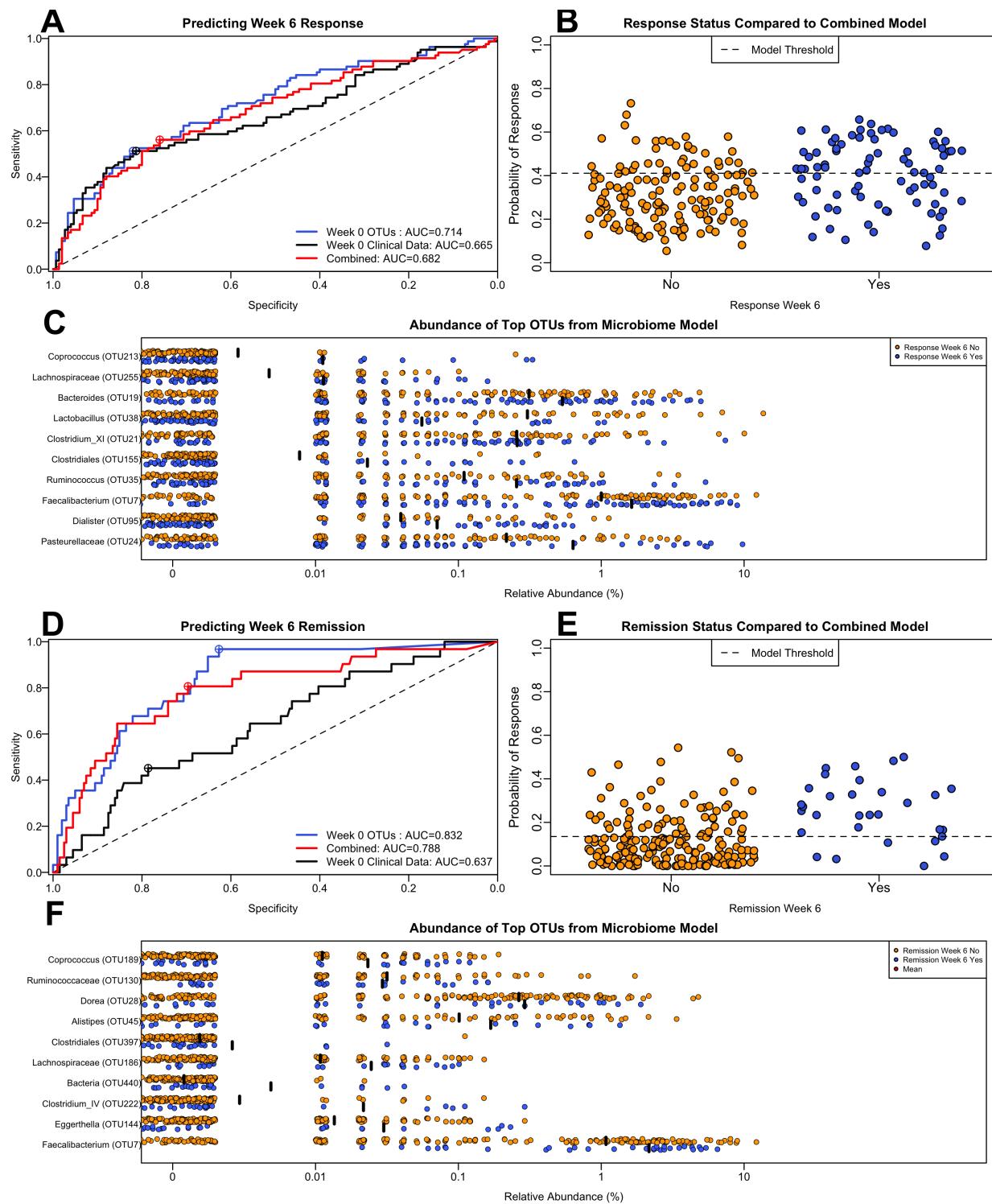


Figure 1: Prediction of RESPONSE/REMISSION in treated subjects using all clinical metadata, baseline microbiome alone, and combined. A. Response ROCs. B. Response Model Performance vs. reality. C. Top predictive taxa and abundance based on response. D. REMISSION ROCs. E. REMISSION Model Performance vs.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

521 **Figure 2: Lefse data supporting the abundance/importance data in the predictive models**

522 Abundance strip charts of differential taxa based on A) response and B) remission.

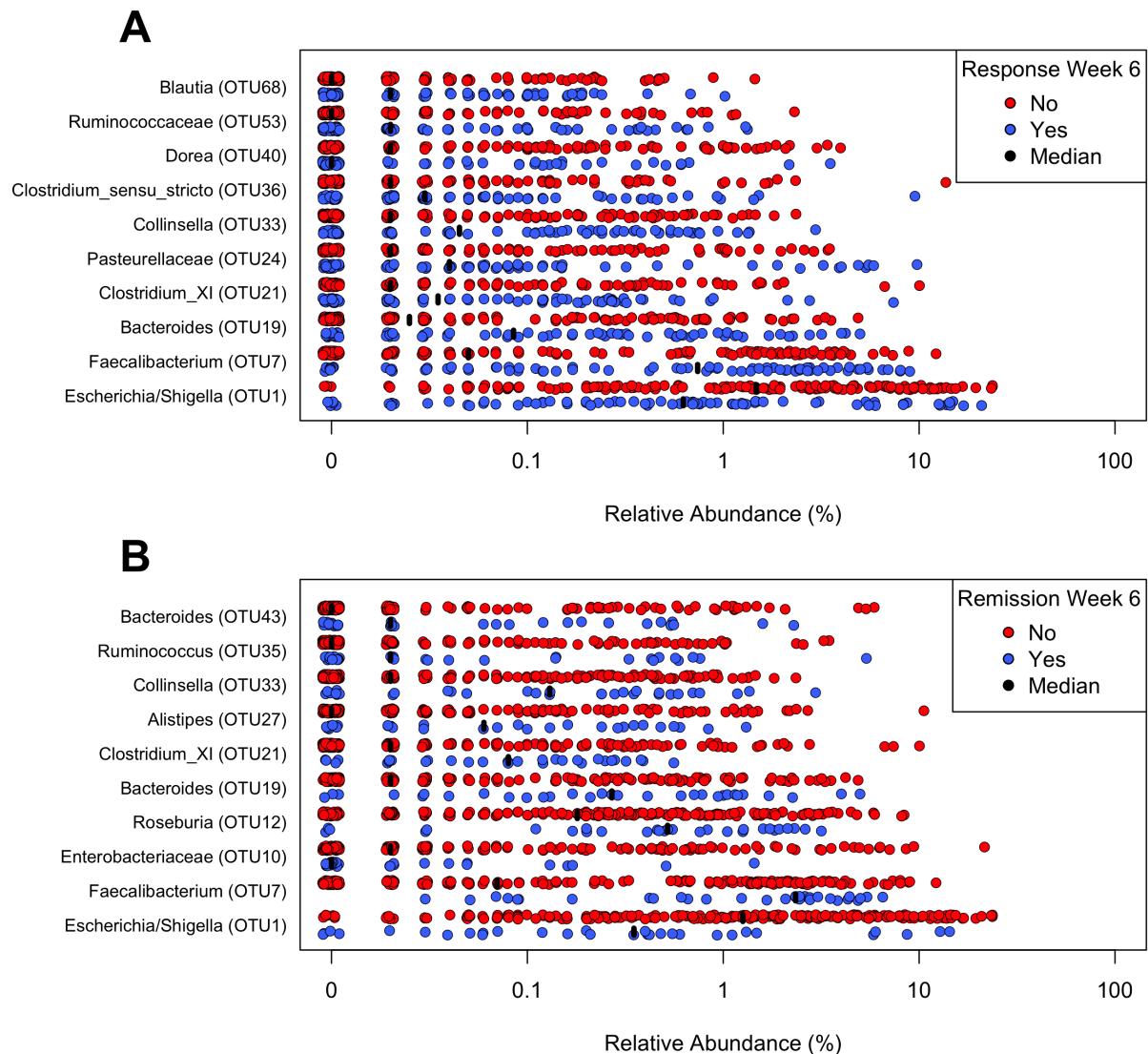


Figure 2: Figure 2: Lefse data supporting the abundance/importance data in the predictive models. Abundance strip charts of differential taxa based on A) response and B) remission.

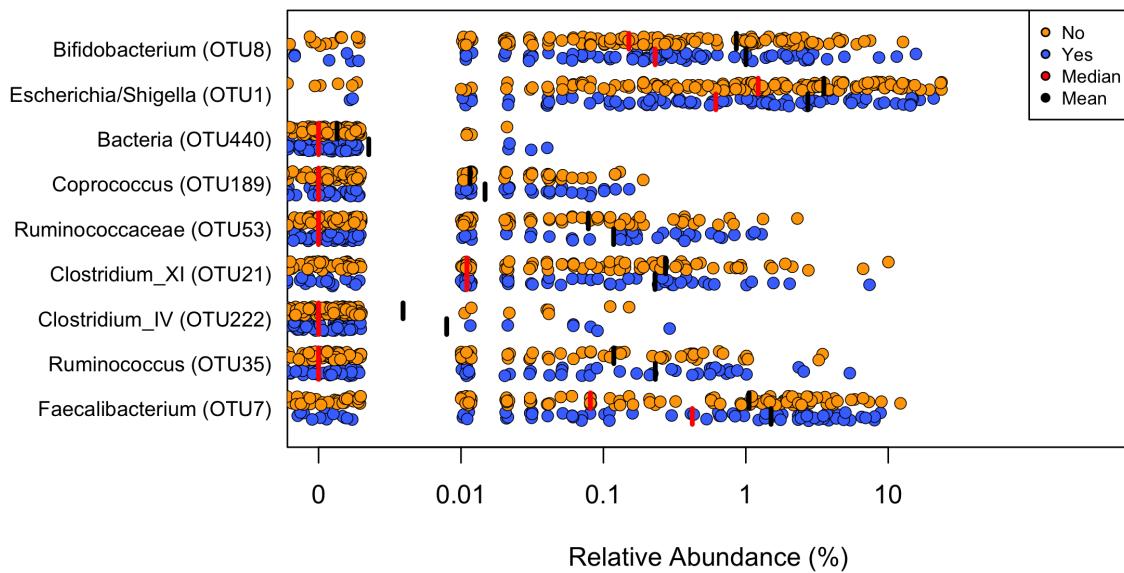


Figure 3: Abundance of frequently predictive OTUs in responders and remitters

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

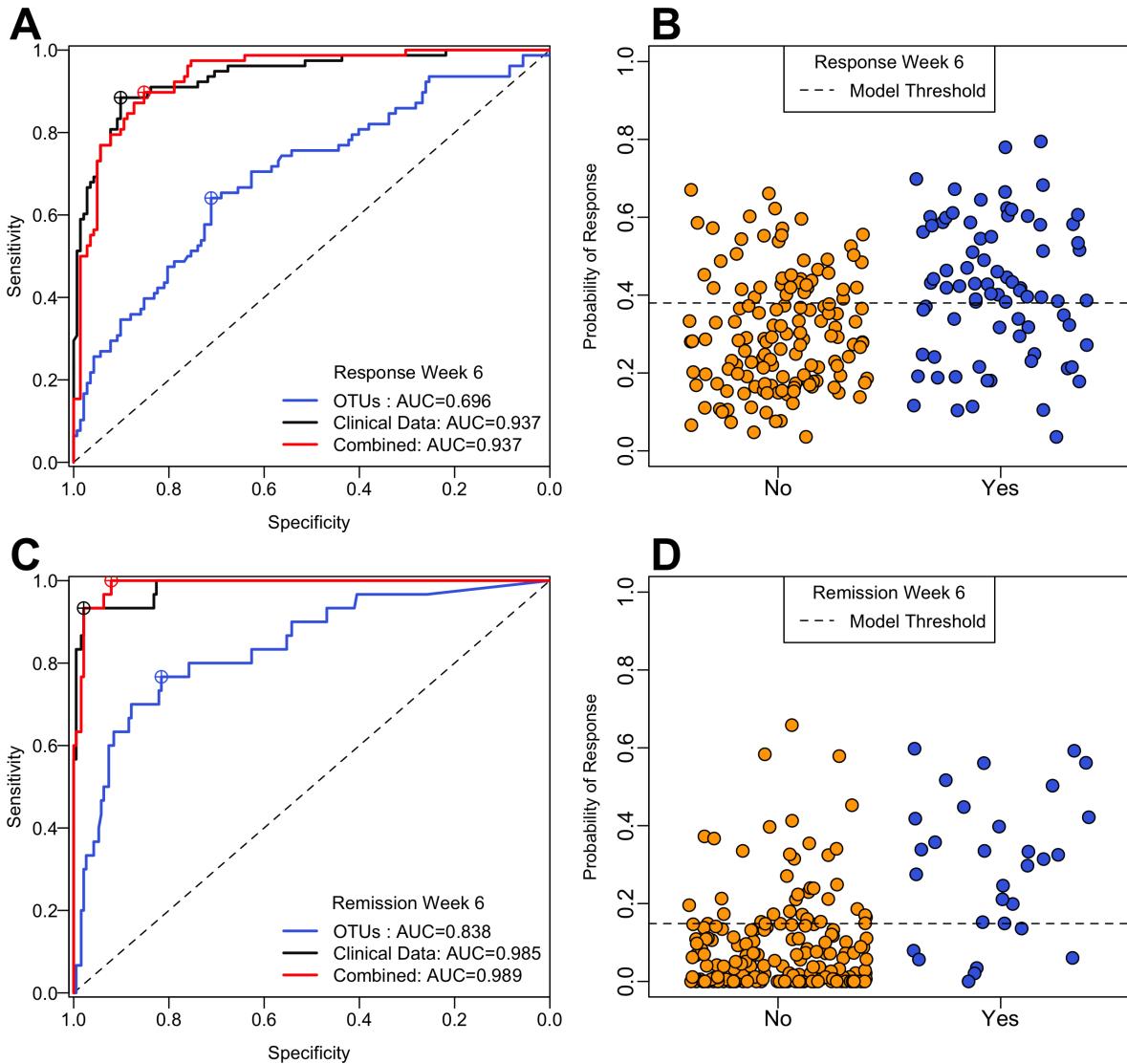


Figure 4: Supplemental Figure 1: Determining Week 6 disease status using Week 6 samples

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

- 525 1. Ananthakrishnan AN. 2015. Epidemiology and risk factors for IBD. *Nature Reviews Gastroenterology & Hepatology* 12:205–217.
- 526
- 527 2. Floyd DN, Langham S, Séverac HC, Levesque BG. 2014. The economic and quality-of-life burden of
- 528 crohn's disease in europe and the united states, 2000 to 2013: A systematic review. *Digestive Diseases and Sciences* 60:299–312.
- 529
- 530 3. Molodecky NA, Soon IS, Rabi DM, Ghali WA, Ferris M, Chernoff G, Benchimol EI, Panaccione R,
- 531 Ghosh S, Barkema HW, Kaplan GG. 2012. Increasing incidence and prevalence of the inflammatory

<sup>532</sup> bowel diseases with time, based on systematic review. *Gastroenterology* 142:46–54.e42.