Published in final edited form as:

Cell Host Microbe. 2017 May 10; 21(5): 603–610.e3. doi:10.1016/j.chom.2017.04.010.

Gut microbiome function predicts response to anti-integrin biologic therapy in Inflammatory Bowel diseases

Ashwin N Ananthakrishnan^{*,1,2}, Chengwei Luo^{*,3}, Vijay Yajnik^{1,2}, Hamed Khalili^{1,2}, John J Garber¹, Betsy W Stevens², Thomas Cleland¹, and Ramnik J Xavier^{1,2,3,4}

¹Division of Gastroenterology, Massachusetts General Hospital, Boston, MA 02114, USA

²Harvard Medical School, Boston, MA 02115, USA

³Broad Institute of MIT and Harvard, Cambridge MA 02142, USA

⁴Center for Microbiome Informatics and Therapeutics, Massachusetts Institute of Technology, Cambridge MA 02142, USA

Abstract

The gut microbiome plays a central role in inflammatory bowel diseases (IBD) pathogenesis and propagation. To determine if the gut microbiome may predict responses to IBD therapy, we conducted a prospective study with Crohn's disease (CD) or ulcerative colitis (UC) patients initiating anti-integrin therapy (vedolizumab). Disease activity and stool metagenomes at baseline, and weeks 14, 30, and 54 after therapy initiation were assessed. Community α -diversity was significantly higher, and *Roseburia inulinivorans* and a *Burkholderiales* species were more abundant at baseline among CD patients achieving week 14 remission. Several significant associations were identified with microbial function; 13 pathways including branched chain amino acid synthesis were significantly enriched in baseline samples from CD patients achieving remission. A neural network algorithm, vedoNet, incorporating microbiome and clinical data, and provided highest classifying power for clinical remission. We hypothesize that the trajectory of early microbiome changes may be a marker of response to IBD treatment.

Graphical abstract

Corresponding Author and Lead contact: Ramnik J Xavier, MD, Division of Gastroenterology, Massachusetts General Hospital, 55 Fruit Street, Boston, MA 02214, xavier@molbio.mgh.harvard.edu, Tel 617-724-6007, Fax 617-724-6832.

*These authors contributed equally.

Author contributions:

A.A., V.Y and R.J.X. conceived and designed the study, A.A., R.J.X., V.Y, H.K., J.G, B.S., and T.C. enrolled the patients and conducted sampling and clinical measures. C.L. designed and performed microbiome analysis, conceived and implemented the algorithm. A.A. and C.L. wrote the paper and all authors read, discussed, and approved the final manuscript.

Keywords

microbiome; vedolizumab; treatment response; roseburia; butyrate; remission

INTRODUCTION

Biologic monoclonal antibody therapy is the cornerstone of treatment of inflammatory diseases including inflammatory bowel diseases (IBD; Crohn's disease (CD), ulcerative colitis (UC)), rheumatoid arthritis (RA), and psoriasis (PsA) (Baumgart and Sandborn, 2012; Ordas et al., 2012; Ramiro et al., 2016; Singh et al., 2016) which affect over 10 million individuals in the United States (Brezinski et al., 2015; Cross et al., 2014; Molodecky et al., 2012). In each of these diseases, biologic therapy reduces disease-related morbidity (Baumgart and Sandborn, 2012; Ordas et al., 2012; Ramiro et al., 2016; Singh et al., 2016). Trials of therapeutic strategies have shown that early initiation of such therapy is associated with greater response (Castro-Rueda and Kavanaugh, 2008; D'Haens et al., 2008; Upchurch and Kay, 2012). Consequently, treatment paradigms have evolved from a step-up strategy to favor up-front biologic therapy to prevent damage. The availability of diverse therapeutic targets has brought forward the importance of personalizing treatment which require *a priori* predicting response to each mechanism of action. Initial attempts to do so relying on clinical factors yielded disappointing results (Siegel and Melmed, 2009). Genetics also performs imperfectly in predicting therapeutic response (Siegel and Melmed, 2009). Genomic expression profiles of target organs (intestine in IBD, articular cartilage in RA) demonstrated initial promise but predictive ability remains modest (Arijs et al., 2009), highlighting the need to identify novel determinants of response.

The past decade has highlighted the central role of the gut microbiome in many immune-mediated diseases (Becker et al., 2015; Forbes et al., 2016; Knights et al., 2013; Kostic et al., 2014). In IBD, the gut microbiome demonstrates reduced diversity, expansion of pro-

inflammatory bacteria like *Enterobacteriaceae* and *Fusobacteriaceae* and depletion of phyla with anti-inflammatory effects such as *Firmicutes* (Becker et al., 2015; Knights et al., 2013; Kostic et al., 2014). Clinical observations of resolution of intestinal inflammation with fecal diversion and exacerbation following restoration of luminal continuity further support this concept (Rutgeerts et al., 1991). In RA, altered T-lymphocyte response due to segmented filamentous bacteria (SFB) in the gut plays an important role (Wu et al., 2010). A similar reduction in diversity and depletion of *Ruminococcus* is seen in psoriatic arthritis as in IBD (Eppinga et al., 2014). Thus, given its role in the pathogenesis of these immune-mediated diseases, taxonomic and functional composition of the gut microbiome may influence likelihood of response to immuno-modulatory therapy for these diseases. An effect of the microbiome on therapy response has been demonstrated previously whereby inactivation of digoxin by *Eggerthella lenta* resulted in altered drug pharmacokinetics and reduced serum concentration (Haiser et al., 2013). Whether a similar effect may be seen with biologic therapy has not been defined previously.

Using a prospectively recruited cohort of patients with IBD initiating gut-selective anti-integrin therapy with vedolizumab as a proof of concept, we performed this study to (1) define the relationship between microbial metagenomic structure and function and clinical remission with vedolizumab induction; (2) to identify longitudinal trajectory of changes in the microbiome with maintenance treatment; and (3) develop a comprehensive predictive model incorporating clinical and microbiome-related data to accurately classify treatment response.

RESULTS

Study population

The study included 85 patients with IBD (43 UC, 42 CD) with a mean disease duration of 13 years at the start of therapy. Just under half of the patients were on concomitant therapy with immunomodulators (42%). Most had previously failed an anti-TNF agent. The mean HBI and SCCAI at baseline were 6 and 5.9 respectively with a mean CRP of 13.2 mg/L (range 0.1 – 140). At week 14, 31 patients met our primary outcome of clinical remission. At week 54 (n=71), 35% of patients remained in remission. Patients who attained remission were likely to have had disease for a shorter duration, more likely to have a diagnosis of CD and less likely to have had prior anti-TNF exposure ($p < 0.05$ for all) (Table S1).

Baseline metagenomic composition and remission at week 14

Community alpha-diversity at baseline was significantly higher in CD patients who achieved remission at week 14 ($q < 0.1$, student's t -test; Figure 1a). This did not achieve statistical significance in UC though we noted a wider range of baseline community diversity for those in remission ($p = 0.031$, F test; Figure 1a). This effect was only observed at the species level suggesting that fine taxonomic differences differentiate remission and non-remission groups (Supplementary Figure 1A). For CD patients, beta-diversity measured was lower at baseline within the remission group compared with the non-remission group (Figure 1b). There was no difference in the microbiome dysbiosis index between remitters and non-remitters either in CD or UC (Supplemental Figure 1B).

Principle component analysis (PCoA) did not differentiate remitters from non-remitters across different taxonomic ranks (Figures 1c–d, Supplementary Figure 2), possibly due to similar baseline relative abundance of the top 15 most abundant species among remitters and non-remitters (Figure 1e–f). However, two species demonstrated a statistically significant difference in relative abundance at baseline between week 14 remitters and non-remitters. Both *Roseburia inulinivorans* and a *Burkholderiales* species were significantly more abundant at baseline among CD patients achieving week 14 remission compared to non-remitters ($q=0.0914$ for *R. inulinivorans*, $q=0.0614$ for *Burkholderiales* sp.; Figure 2a).

Thirteen pathways were significantly enriched ($q < 0.1$) in baseline samples from the CD patients achieving remission compared to non-remitters including branched chain amino acid (BCAA) biosynthesis pathways involved in biosynthesis of *L*-citrulline (log₂ fold difference: 0.812, $q=0.0957$), *L*-isoleucine from threonine (log₂ fold difference: 0.482, $q=0.02$), and arginine and polyamine (log₂ fold difference 1.073, $q=0.0828$) (Figure 2b). Two and three pathways were respectively significantly enriched and depleted among UC patients achieving remission ($q<0.1$; Figure 2b).

Longitudinal trajectory of the microbiome

In CD, only five taxa were significantly different in relative abundance between baseline and follow-up (24 CD patients with paired samples; 10 achieving remission). These were *Bifidobacterium longum*, *Eggerthella*, *Ruminococcus gnavus*, *Roseburia inulinivorans*, and *Veillonella parvula* (Figure 3b). All these taxa decreased in relative abundance in patients achieving remission. In UC (17 patients with paired samples; 11 achieving remission), only one taxon, *Streptococcus salivarius*, significantly changed in relative abundance with an increase in abundance in patients not achieving remission (Figure 3b).

In contrast to these few changes in microbial composition, there were significantly greater metagenomic alterations in microbial function. In CD, 17 pathways were significantly reduced on follow-up at week 14 compared to baseline, of which 15 were noted only in patients achieving remission (Figure 3a). These included a decrease in several tricarboxylic acid cyclic (TC) pathways (I and V types) and nicotinamide adenine dinucleotide (NAD) salvage pathway, suggesting decreased oxidative stress in patients achieving remission. O-antigen building blocks biosynthesis in *E. coli* was also decreased but not accompanied by a corresponding reduction in abundance of *E. coli*. In patients not achieving remission, only two pathways - *L*-arginine biosynthesis via *N*-acetyl-*L*-citrulline pathway and tetrapyrrole biosynthesis from glutamate pathway – were decreased by week 14. The changes were less striking in UC. Three pathways - polyamine biosynthesis, non-oxidative pentose phosphate pathway, and sucrose degradation – increased in relative abundance among patients achieving remission (Figure 3a). In contrast, gluconeogenesis, uridine monophosphatate (UMP) biosynthesis, and putrescine biosynthesis decreased in relative abundance among those not achieving remission (Figure 3a).

Finally, we examined if there was a difference in the direction of change for taxa or pathways between those achieving remission and those not. The only significantly different species in CD was *Roseburia inulinivorans* which decreased in abundance in those achieving remission while increasing in those who had not ($q=0.013$; Benjamini-Hochberg adjusted χ^2

test). Hexitol degradation and glycolysis pathways also demonstrated different directions of change in CD patients achieving remission compared to those not. In UC, while no significant differential changes in taxa were noted, palmitate and stearate biosynthesis pathways increased in relative abundance in patients achieving remission and decreased in those not achieving remission ($q=0.087$ and 0.045 , respectively; Benjamini-Hochberg adjusted χ^2 test).

Persistence of changes in the microbiome at 1 year

Eight patients (3 CD; 5 UC) provided stool samples at baseline, weeks 14, 30, and 54 while thirteen patients (5 CD; 8 UC) had samples available at baseline, weeks 14 and 54. Persistency of treatment effect was observable at both week 30 and week 54 for among both the remission and non-remission groups though only the changes in the remission group demonstrated a statistical difference compared to random sampling (Figure 3d). Specifically, patients achieving remission at week 14 demonstrated highly significant persistency in the microbial composition at week 30 ($P=0.00039$) and a weaker effect at week 54 ($P=0.019$), suggesting that attainment of remission at week 14 is associated with durable changes in the microbiome.

Neural network algorithms (vedoNet) to predict treatment response

Several different neural network models were evaluated to predict clinical remission at week 14 (Figures 4a–b, see **Methods** for model details). Baseline clinical data alone was insufficient in predicting remission at week 14 (AUC 0.619). In contrast, use of available microbial taxa (vedoNet.tx; AUC 0.715) and pathways (vedoNet.pw, AUC 0.738) resulted in improved predictive ability (Figure 4b). Taxonomic profiles at the level of genus, family, or class performed less well than information at the species level (Supplementary Figure 3). A model incorporating clinical data, taxonomy, and pathway relative abundance without any pre-selection of variables performed better than each individual model (vedoNet.hybrid, AUC 0.776) (Figure 4b). Finally, a manually curated list of 40 microbiome variables (**Methods**, Supplementary Table 1) provided the highest classifying power (AUC=0.872; Figure 4b), successfully achieving > 80% true positive discovery rate with a less than 25% false negative discovery rate. For other models to achieve the same true positive rate, the false negative rates were over 50% (vedoNet.tx and vedoNet.pw) or 60% (clinical data and vedoNet.hybrid). vedoNet was implemented in Python with Keras library; the codes, models, and tutorials can be found at www.bitbucket.org/luo-chengwei/vedoNet. We repeated the analysis stratifying by type of IBD at baseline, allowing for incorporation of disease-specific phenotypic information. This did not significantly improve the predictive value of the vedoNet model for either CD (AUC 0.881) or UC (AUC 0.853). In analysis stratified by MD-index (above or below median), among high MD-index subjects, vedoNet's sensitivity was 0.75 and specificity was 0.769; while in low MD-index subjects, vedoNet's sensitivity and specificity were 0.818 and 0.85 respectively.

Strain-level analysis

To investigate if the strain variability can play a role in determining the outcome of the treatment, we employed a strain-level resolution approach to identify any strain-specific

signal showed significant association with outcome. We sought to focus on pathways that differentiated remission group from non-remission group, since differences in genes within these pathways might offer essential information in determining treatment. We found that among CD baseline samples, those who entered remission at week 14 possessed a cluster of unique SNPs (FDR<0.1) located in L-arginine biosynthesis pathways. Such SNPs are predominantly contributed by *Bifidobacterium longum* (21/41=51.22% of total unique SNPs) and *Dialister invisus* (11/41=26.83% of total unique SNPs); among UC baseline samples, those who entered remission at week 14 possessed a more diversified group-specific SNP profile that spread among UMP biosynthesis pathway and pentose phosphate pathway (Supplemental Figure 4). The major differentiating species contributing to the stratifying SNPs include *Bifidobacterium longum*, *Ruminococcus torques*, and *E. coli*.

Validation in an anti-TNF cohort

Twenty patients (14 CD, 6 UC) were included in the anti-TNF validation cohort with a mean disease duration of 7 years. The baseline C-reactive protein levels were 16.9 mg/dL and mean HBI and SCCAI were 5 and 6 respectively. At week 14, 13 patients (65%) achieved clinical remission. Supplemental Table 4 presents the results of the various classification algorithms in classifying remitters and non-remitters. vedoNet was able to accurately identify 11 out of the 13 patients achieving remission and was superior to models utilizing clinical data (10/13) or microbial taxa alone (9/13).

DISCUSSION

The gut microbiome is a key determinant of initiation and propagation of luminal inflammation in IBD(Becker et al., 2015; Forbes et al., 2016; Gevers et al., 2014; Knights et al., 2013; Kostic et al., 2014). Here, we describe the microbial composition and structure from a large cohort of IBD patients initiating vedolizumab therapy(Shelton et al., 2015). We demonstrate associations between baseline taxonomic composition and functional pathway abundance and clinical remission at 14 weeks and demonstrate the utility of predictive models incorporating both clinical and microbiome data in predicting clinical remission. We also hypothesize that trajectory of early changes in the microbiome may be a marker of response to treatment in IBD.

There have been few studies of longitudinal changes in the gut microbiome with drug treatment in IBD. Shaw *et al.* characterized 19 children with CD and 4 with UC, showing dysbiosis at baseline that correlated with luminal inflammatory burden(Shaw et al., 2016). An improvement in fecal diversity was seen with clinical response in UC but not CD. In our study, a more diverse microbial composition at baseline predicted week 14 clinical remission. A less diverse microbiome has been consistently linked to the development of IBD; factors such as antibiotics that reduce gut diversity increase risk of IBD(Becker et al., 2015; Forbes et al., 2016; Gevers et al., 2014; Knights et al., 2013; Kostic et al., 2014; Lewis et al., 2015; Singh et al., 2009; Ungaro et al., 2014). Thus, a more diverse microbiome at baseline may reflect prevalent microbes and/or metabolites with anti-inflammatory effect on colonic inflammation and a less disrupted mucosal barrier, leading to greater treatment response. Restoration of gut diversity with has been reported previously with anti-TNF

therapy (Lewis et al., 2015; Shaw et al., 2016) though a more diverse microbiome has not been previously shown to be predictive of treatment response. One could hypothesize that this difference may be due to the systemic effect of anti-TNF therapy compared to the inhibition of gut-specific leukocyte trafficking by vedolizumab. Microbiome derived signals maybe more relevant to response to agents that block T cell traffic as compared to antibodies that neutralize specific cytokines. The complexity in predicting treatment response using gut microbial structure is highlighted by poor separation between the remitters and non-remitters on simple PCoA, consistent with a few previous studies (Shaw et al., 2016). Our more adaptive and informative neural network based approach performed significantly better with an AUC of 0.87 and suggested added and complementary value to both clinical and microbial parameters.

Taxonomically, the relative abundance of *R. inulinivorans* and *Burkholderiales* at baseline was predictive of week 14 remission. *R. inulinivorans* is a relatively low abundance gram-positive organism, certain strains of which encode genes for pro-inflammatory flagellin proteins that stimulate interleukin-8 production (Neville et al., 2013). *R. inulinivorans* also produces butyrate and propionate both of which have anti-inflammatory effects through a variety of mechanisms including reinforcing the integrity of the colonic epithelial barrier, reducing oxidative stress, and decreasing inflammation through inhibition of nuclear factor κ B (NF- κ B) activation by histone deacetylation (Canani et al., 2011; Hamer et al., 2008; Inan et al., 2000). Butyrate also inhibits inflammation through inhibition of the IFN γ /STAT1 signaling pathways associated with chronic inflammation and enhances apoptosis of colonic T-cells (Hamer et al., 2008; Zimmerman et al., 2012).

In contrast to the relatively few changes between remitters and non-remitters at the species or genus level, differences in functional pathways were more striking. Pathways related to BCAA biosynthesis including citrulline, isoleucine, arginine and polyamine were enriched at baseline in CD patients who achieved week 14 remission. BCAA may reduce colonic inflammation through a variety of mechanisms. Arginine and isoleucine supplementation results in upregulation of human beta defensin 1 (hBD-1) in colon cells; reduced beta-defensin expression is associated with colonic inflammation in IBD (Ramasundara et al., 2009). In a C57BL/6 mouse dextran sodium sulfate (DSS) colitis model, arginine supplementation reduced intestinal inflammation and cytokine production (Coburn et al., 2012). Arginine is also a precursor for nitric oxide (NO) and endothelial NO is important for maintenance of intestinal perfusion and barrier integrity while NO produced by the inducible nitric-oxide synthase has direct anti-bacterial activity and is an important regulator of host defense (Kolios et al., 2004). NO may also reduce damage from oxidative stress and through inhibition of NF- κ B translocation (Kolios et al., 2004). Glycosaminoglycan (GAG) degradation pathways were also enriched in those achieving remission. In mice models, intestinal flora mediated degradation of GAG resulted in metabolites with a cytotoxic effect on intestinal epithelium; inhibition of this degradation with antibiotics ameliorated colitis (Lee et al., 2009). Week 14 remission was associated with a reduction in several functional pathways up regulated at baseline. For example, the NAD salvage pathway decreased by week 14 among those achieving remission suggesting that clinical improvement was associated with a reduction in luminal oxidative stress. These discoveries

together suggest that functional rather than taxonomic differences may be important determinants of treatment outcome.

Interestingly, we also observed that responders at week 14 demonstrated greater persistence of their microbial changes at 1 year compared to non-responders suggesting that early changes in the microbiome could be an indicator of clinical response. Similar clinical observations have been noted in CD and UC. In the ACT trial of infliximab in UC, endoscopic response by week 8 was associated with a lower rate of colectomy at week 54 (Colombel et al., 2011). In parallel, early reduction in fecal calprotectin has been associated with improved long-term outcomes in patients with IBD (Pavlidis et al., 2016). Thus, early microbiome changes may be an added marker of sensitivity to treatment and initial response.

There are several implications to our findings. Our study demonstrates the ability to predict response to anti-integrin treatment using the gut microbiome, highlighting the role not just of microbial taxonomy but more importantly functional pathways that may be relevant to treatment. Similar analyses for drugs with different mechanisms of action may offer the ability to *a priori* select agents with higher likelihood of response based on gut microbial composition. Advances in technology increasingly allow rapid sequencing of the microbiome through methods such as paper-based assays (Pardee et al., 2014), and one could envision such methods being incorporated into routine clinical care. The persistence of microbial changes in those achieving remission at week 14 further highlights the importance of early clinical response in predicting long-term outcome with treatment, and potentially a mechanism thereof. Identification of predictive microbial signals could also allow for refinement of novel probiotics that may deliver specific anti-inflammatory taxa or strains, or stimulate anti-inflammatory metabolic pathways that may be of benefit in ameliorating gut inflammation.

We readily acknowledge several limitations to our study. This was a single center cohort of predominantly refractory patients. Remission relied on clinical indicators rather than biochemical, fecal, or endoscopic outcomes. Few patients provided stools at each of the time points through 1 year, limiting our statistical power. Diet was not routinely assessed in all patients, and consequently its effect on the gut microbiome cannot be excluded. While we performed validation of our model in an independent cohort of 20 patients initiating anti-TNF therapy, we acknowledge that more robust examination in larger cohorts is essential prior to application to clinical practice. Further experimental studies are important to determine the full mechanistic implications of the metagenomic pathways and bacteria identified and how they may be harnessed to improve response to existing therapy in IBD.

In conclusion, we describe associations between gut microbial taxonomic composition and function and response to anti-integrin therapy in CD and UC. Early clinical remission could be predicted by microbial functional composition at baseline with a weaker influence at the level of the species or genus. The association between abundance of butyrate producing bacteria and enrichment of branched chain amino acid biosynthesis pathways at baseline in remitters, and reduction in oxidative stress pathways with therapy response provides support for an important role of these pathways in the propagation and resolution of intestinal

inflammation. The pathways and microbes thus identified could potentially serve as targets for newer therapies and shed further light on the pathogenesis and progression of these complex diseases.

STAR*METHODS

Contact for reagent and resource sharing

Further information and requests for reagents or resources may be directed to and will be fulfilled by lead contact Ramnik J. Xavier (xaviermolbio.mgh.harvard.edu)

Experimental model and subject details

Vedolizumab cohort and outcomes—This study was nested within a longitudinal prospective IBD cohort at Massachusetts General Hospital (Prospective Registry of IBD Study at MGH (PRISM)). Details of this cohort have been published previously (Ananthakrishnan et al., 2014; Shelton et al., 2015). In brief, the PRISM registry is open to all adult patients with IBD seeking care at the MGH Crohn's and Colitis center. This nested study was a prospective inception cohort of patients initiating vedolizumab for refractory luminal CD or UC, often in the setting of prior anti-tumor necrosis factor α (anti-TNF) failure. Characteristics of the included patients initiating vedolizumab therapy is presented in Supplemental Table 1. All patients initiating vedolizumab as part of their routine clinical care were eligible for inclusion without an *a priori* fixed sample size for recruitment. Patients with an ileostomy or J-pouch were excluded as disease activity scores could not be reliably calculated. Most patients had failed more than one anti-TNF therapy previously. Patients received intravenous vedolizumab 300mg at weeks 0, 2, 6, and every 8 weeks thereafter. At weeks 0 (baseline), 6, 14, 30, and 54, patients provided stool for metagenomic sequencing. At each infusion, disease activity was assessed using the Harvey Bradshaw index for CD (Harvey and Bradshaw, 1980) and simple clinical colitis activity index for UC (Walmsley et al., 1998). Hemoglobin, serum albumin, C-reactive protein, erythrocyte sedimentation rate, white blood cell and platelet count were obtained at each infusion. Our primary study outcome was clinical remission at week 14, defined as HBI < 4 or SCCAI < 2. A reduction in either the HBI or SCCAI by ≥ 3 points indicated clinical response, consistent with the cut-offs used in clinical trials (Harvey and Bradshaw, 1980; Walmsley et al., 1998).

Validation cohort—External validation of the results of our predictive model was performed in an independent cohort of 20 patients with moderate-to-severe CD or UC initiating therapy with an anti-TNF biologic therapy (infliximab or adalimumab). Similar to the vedolizumab cohort, disease activity using the HBI or SCCAI was collected along with stool for metagenomic sequencing at baseline and at week 14. We examined the ability of the final predictive model developed in the vedolizumab cohort to classify anti-TNF remitters and non-remitters at week 14.

Ethics Statement—Both studies were approved by the Institutional Review Board of Partners Healthcare and all patients provided informed consent.

Method Details

Microbiome community profiling and sequencing—RNA and DNA purification from stool aliquots was performed according to protocols optimized in the Human Microbiome Project (Group et al., 2009; Integrative, 2014). In brief, participating patients provided stool in storage tubes containing RNA later. Stool samples were stored at 4°C for less than 24 hours and then stored at −80°C until DNA extraction. Genomic DNA extraction from stool was performed using the Qiagen AllPrep MiniKit (Valencia, CA, USA) as per manufacturer's instructions. Illumina based DNA shotgun sequencing was performed at the Broad Institute (Cambridge, MA) to characterize rare taxa and understand relationships between community membership and community function.

Quantification and Statistical Analysis

Metagenomic analysis—Metagenomic reads were quality trimmed using trimmomatic v0.36 with default settings, retaining post-trimming reads with both ends longer than 60bp (Bolger et al., 2014). Samples were minimized for reads originating in human genomes using BMTagger (<ftp://ftp.ncbi.nlm.nih.gov/pub/agarwala/bmtagger/>, 07 March 2011, version 3.101). MetaPhlan2 (Truong et al., 2015) was employed to taxonomically profile each sample using default settings with Bowtie v2.2.4 as search engine (Langmead and Salzberg, 2012). Pathway relative abundance of each sample was quantified by HUMAn v2.0 (Abubucker et al., 2012) using DIAMOND (Buchfink et al., 2015) with package-shipped ChocoPhlAn and EC-filtered UniRef90 databases. Fisher's Alpha was calculated for the taxonomic profiles at phylum, class, order, family, genus, and species levels for each sample using normalized read count from MetaPhlan2's marker gene mapping. Bray-Curtis dissimilarity (BCD) was calculated in intra-group and inter-group fashions on the same taxonomic ranks. The top most abundant taxonomic groups and pathways were selected by the median relative abundance across all samples. Student's *t*-test was carried out to test if any taxonomic group, pathway, or diversity metrics at baseline significantly differed between those achieving remission at week 14 and those who did not; *p*-values were corrected for multiple testing using the Benjamini-Hochberg procedure. The *F*-test was used to examine if the two normal distributions significantly differed in their variances. We also calculated a microbiome dysbiotic index (MD-index) as the logarithm of the ratio between the relative abundance sum of IBD-increased taxonomic groups and the relative abundance sum of the IBD-decreased taxonomic groups defined by Gevers *et al.* Differences in the MD-index between remitters and non-remitters were compared.

Analysis of longitudinal trajectory of the gut microbiome—For IBD patients with both baseline and week 14 stool samples available, we calculated the log₂ fold change (FC) in taxa and pathway relative abundance. The change was defined as significant if a taxon or pathway experienced consistent ≥ 1.5 FC (log₂FC values are ±0.58) in over 80% of the samples. For subjects with follow-up samples available at later time points (weeks 30 and 54), we designed a persistent index to measure the degree of persistency of the effect of treatment on taxa or pathways. This index was defined as the difference in the degree of the later follow-up sample mimicking the week 14 sample accounting for baseline differences. Formally, the index, *P*, was defined as:

$$P = [\text{BCD}(f, b) - \text{BCD}(f, k)] / \text{BCD}(k, b),$$

where $\text{BCD}(x, y)$ represents the Bray-Curtis dissimilarity between sample x and y , and b , k , and f defined the baseline, week 14, and later follow-up (week 30 or week 54) samples. If sample f was identical to sample k representing maximum persistency, then $P=1$. If sample f was identical to sample b representing no persistency, then $P=-1$. A randomized profile was generated by decoupling the taxa and relative abundance and re-associating them at random. P was calculated for the randomized samples and independent student's t -test was applied to analyze differences between random and observed profiles.

Neural network predictor, vedoNet—A neural network structure-based predictor was constructed using baseline information to predict remission at week 14. This consisted of an input layer, a convolution layer with a softmax dropout layer, a rectified linear unit (ReLU), and an output unit to classify if the input data could predict week 14 remission. Random parameters drawn from standard norm $\sim \mathcal{N}(0,1)$ were assigned to the initial neural network, and the network was trained and tested on 5-fold cross validation. The input variables including both microbiome data as well as clinical information including type of IBD, age at diagnosis, gender, smoking history, baseline disease activity, and laboratory parameters (CRP, WBC count, ESR, platelet count, hemoglobin, and albumin) (Supplemental Table 2). Different microbiome models were tested and compared, including purely taxon profiles-based (vedoNet.tx), pathway-based (vedoNet.pw), as well as mixture of both, with knowledge-guided input variable selection (vedoNet). For vedoNet.tx, MetaPhlan v2.0 profiles at species, genus, and class levels were respectively employed in model construction; for vedoNet.pw, normalized HUMAnN2 output pathway profiles were used in model construction. Lastly, for vedoNet, we selected the relative abundance of *Roseburia inulinivorans*, *Burkholderiales*, *Eggerthella*, *Bifidobacterium longum*, *Ruminococcus gnavus*, *Veillonella parvula*, *Lactobacillus salivarius*, and the relative abundance of pathways shown in Supplementary Table 3 as input for model building. These phyla were selected based on the fold change difference between baseline and week 14 among the remission and non-remission groups. The number of neural units for the input layers varied to accommodate the different input vector lengths in each respective model.

The area under curve (AUC) of the receiver operating characteristic (ROC) curve served as the main indicator of vedoNet's performance. Subjects with complete follow up information at week 14 and a stool sample at baseline were divided into five batches at random for the 5-fold cross validation procedure. One-hundred such combinations were generated to train and test models. The training process deployed stochastic gradient descent (SGD) using categorical cross entropy as cost function. The model structure with the highest AUC was selected and re-trained in 1,000 iterations using SGD and calibrated with all samples.

Strain level analysis—To obtain strain-level resolution on the differentiating pathways, we mapped every baseline sample's reads onto the reference genes using Bowtie2 (Langmead and Salzberg, 2012) and then piled up the reads using SAMTools (Li et al., 2009) mpileup function with default settings, controlling read mapping quality to be no less than 15. We

first called SNPs with reference-free, aggregated allele frequencies; positions with higher than 10× relative abundance in both sample groups (remission and non-remission) and minor allele with at least 2× coverage and frequency >0.1 were identified as SNPs. Based on week 14 followup remission status, we calculated the likelihood that the two groups allele frequencies were drawn from different background. To quantify the SNPs' uniqueness in differentiating the two groups, Chi-2 test was carried out and further corrected for FDR using Benjamini-Hochberg approach. SNPs sites with $q < 0.1$ were selected as unique SNP sites.

Data and Software availability

The data from the study are available at <https://www.ncbi.nlm.nih.gov/sra/>.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Source of funding: Ananthakrishnan is supported by a grant from the National Institutes of Health (DK097142) and Crohn's and Colitis Foundation. Luo and Xavier are supported by the Helmsley Charitable Trust, Crohn's and Colitis Foundation, and the National Institutes of Health (DK43351, DK92405). This work is also supported by the National Institutes of Health (NIH) (P30 DK043351) to the Center for Study of Inflammatory Bowel Diseases and the Center for Microbiome Informatics and Therapeutics at MIT.

Conflicts of Interest: Ananthakrishnan has served on the scientific advisory boards for Abbvie, Takeda, Exact Sciences, and Merck and has received grant support from Merck and Amgen.

The authors thank participating patients and research staff at the MGH Crohn's and Colitis center and the Broad Institute.

References

- Abubucker S, Segata N, Goll J, Schubert AM, Izard J, Cantarel BL, Rodriguez-Mueller B, Zucker J, Thiagarajan M, Henrissat B, et al. Metabolic reconstruction for metagenomic data and its application to the human microbiome. *PLoS Comput Biol*. 2012; 8:e1002358. [PubMed: 22719234]
- Ananthakrishnan AN, Huang H, Nguyen DD, Sauk J, Yajnik V, Xavier RJ. Differential effect of genetic burden on disease phenotypes in Crohn's disease and ulcerative colitis: analysis of a North American cohort. *Am J Gastroenterol*. 2014; 109:395–400. [PubMed: 24419484]
- Arijs I, Li K, Toedter G, Quintens R, Van Lommel L, Van Steen K, Leemans P, De Hertogh G, Lemaire K, Ferrante M, et al. Mucosal gene signatures to predict response to infliximab in patients with ulcerative colitis. *Gut*. 2009; 58:1612–1619. [PubMed: 19700435]
- Baumgart DC, Sandborn WJ. Crohn's disease. *Lancet*. 2012; 380:1590–1605. [PubMed: 22914295]
- Becker C, Neurath MF, Wirtz S. The Intestinal Microbiota in Inflammatory Bowel Disease. *ILAR J*. 2015; 56:192–204. [PubMed: 26323629]
- Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014; 30:2114–2120. [PubMed: 24695404]
- Brezinski EA, Dhillon JS, Armstrong AW. Economic Burden of Psoriasis in the United States: A Systematic Review. *JAMA Dermatol*. 2015; 151:651–658. [PubMed: 25565304]
- Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. *Nat Methods*. 2015; 12:59–60. [PubMed: 25402007]
- Canani RB, Costanzo MD, Leone L, Pedata M, Meli R, Calignano A. Potential beneficial effects of butyrate in intestinal and extraintestinal diseases. *World J Gastroenterol*. 2011; 17:1519–1528. [PubMed: 21472114]

- Castro-Rueda H, Kavanaugh A. Biologic therapy for early rheumatoid arthritis: the latest evidence. *Curr Opin Rheumatol*. 2008; 20:314–319. [PubMed: 18388524]
- Coburn LA, Gong X, Singh K, Asim M, Scull BP, Allaman MM, Williams CS, Rosen MJ, Washington MK, Barry DP, et al. L-arginine supplementation improves responses to injury and inflammation in dextran sulfate sodium colitis. *PLoS One*. 2012; 7:e33546. [PubMed: 22428068]
- Colombel JF, Rutgeerts P, Reinisch W, Esser D, Wang Y, Lang Y, Marano CW, Strauss R, Oddens BJ, Feagan BG, et al. Early mucosal healing with infliximab is associated with improved long-term clinical outcomes in ulcerative colitis. *Gastroenterology*. 2011; 141:1194–1201. [PubMed: 21723220]
- Cross M, Smith E, Hoy D, Carmona L, Wolfe F, Vos T, Williams B, Gabriel S, Lassere M, Johns N, et al. The global burden of rheumatoid arthritis: estimates from the global burden of disease 2010 study. *Ann Rheum Dis*. 2014; 73:1316–1322. [PubMed: 24550173]
- D'Haens G, Baert F, van Assche G, Caenepeel P, Vergauwe P, Tuijnman H, De Vos M, van Deventer S, Stitt L, Donner A, et al. Early combined immunosuppression or conventional management in patients with newly diagnosed Crohn's disease: an open randomised trial. *Lancet*. 2008; 371:660–667. [PubMed: 18295023]
- Eppinga H, Konstantinov SR, Peppelenbosch MP, Thio HB. The microbiome and psoriatic arthritis. *Curr Rheumatol Rep*. 2014; 16:407. [PubMed: 24474190]
- Forbes JD, Van Domselaar G, Bernstein CN. The Gut Microbiota in Immune-Mediated Inflammatory Diseases. *Front Microbiol*. 2016; 7:1081. [PubMed: 27462309]
- Gevers D, Kugathasan S, Denson LA, Vazquez-Baeza Y, Van Treuren W, Ren B, Schwager E, Knights D, Song SJ, Yassour M, et al. The treatment-naïve microbiome in new-onset Crohn's disease. *Cell Host Microbe*. 2014; 15:382–392. [PubMed: 24629344]
- Group NHW, Peterson J, Garges S, Giovanni M, McInnes P, Wang L, Schloss JA, Bonazzi V, McEwen JE, Wetterstrand KA, et al. The NIH Human Microbiome Project. *Genome Res*. 2009; 19:2317–2323. [PubMed: 19819907]
- Haiser HJ, Gootenberg DB, Chatman K, Sirasani G, Balskus EP, Turnbaugh PJ. Predicting and manipulating cardiac drug inactivation by the human gut bacterium *Eggerthella lenta*. *Science*. 2013; 341:295–298. [PubMed: 23869020]
- Hamer HM, Jonkers D, Venema K, Vanhoutvin S, Troost FJ, Brummer RJ. Review article: the role of butyrate on colonic function. *Aliment Pharmacol Ther*. 2008; 27:104–119. [PubMed: 17973645]
- Harvey RF, Bradshaw JM. A simple index of Crohn's-disease activity. *Lancet*. 1980; 1:514. [PubMed: 6102236]
- Inan MS, Rasoulopour RJ, Yin L, Hubbard AK, Rosenberg DW, Giardina C. The luminal short-chain fatty acid butyrate modulates NF-kappaB activity in a human colonic epithelial cell line. *Gastroenterology*. 2000; 118:724–734. [PubMed: 10734024]
- Integrative, H.M.P.R.N.C. The Integrative Human Microbiome Project: dynamic analysis of microbiome-host omics profiles during periods of human health and disease. *Cell Host Microbe*. 2014; 16:276–289. [PubMed: 25211071]
- Knights D, Lassen KG, Xavier RJ. Advances in inflammatory bowel disease pathogenesis: linking host genetics and the microbiome. *Gut*. 2013; 62:1505–1510. [PubMed: 24037875]
- Kolios G, Valatas V, Ward SG. Nitric oxide in inflammatory bowel disease: a universal messenger in an unsolved puzzle. *Immunology*. 2004; 113:427–437. [PubMed: 15554920]
- Kostic AD, Xavier RJ, Gevers D. The microbiome in inflammatory bowel disease: current status and the future ahead. *Gastroenterology*. 2014; 146:1489–1499. [PubMed: 24560869]
- Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012; 9:357–359. [PubMed: 22388286]
- Lee HS, Han SY, Ryu KY, Kim DH. The degradation of glycosaminoglycans by intestinal microflora deteriorates colitis in mice. *Inflammation*. 2009; 32:27–36. [PubMed: 19067146]
- Lewis JD, Chen EZ, Baldassano RN, Otley AR, Griffiths AM, Lee D, Bittinger K, Bailey A, Friedman ES, Hoffmann C, et al. Inflammation, Antibiotics, and Diet as Environmental Stressors of the Gut Microbiome in Pediatric Crohn's Disease. *Cell Host Microbe*. 2015; 18:489–500. [PubMed: 26468751]

- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. Genome Project Data Processing, S. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009; 25:2078–2079. [PubMed: 19505943]
- Molodecky NA, Soon IS, Rabi DM, Ghali WA, Ferris M, Chernoff G, Benchimol EI, Panaccione R, Ghosh S, Barkema HW, et al. Increasing incidence and prevalence of the inflammatory bowel diseases with time, based on systematic review. *Gastroenterology*. 2012; 142:46–54. e42. quiz e30. [PubMed: 22001864]
- Neville BA, Sheridan PO, Harris HM, Coughlan S, Flint HJ, Duncan SH, Jeffery IB, Claesson MJ, Ross RP, Scott KP, et al. Pro-inflammatory flagellin proteins of prevalent motile commensal bacteria are variably abundant in the intestinal microbiome of elderly humans. *PLoS One*. 2013; 8:e68919. [PubMed: 23935906]
- Ordas I, Eckmann L, Talamini M, Baumgart DC, Sandborn WJ. Ulcerative colitis. *Lancet*. 2012; 380:1606–1619. [PubMed: 22914296]
- Pardee K, Green AA, Ferrante T, Cameron DE, DaleyKeyser A, Yin P, Collins JJ. Paper-based synthetic gene networks. *Cell*. 2014; 159:940–954. [PubMed: 25417167]
- Pavlidis P, Gulati S, Dubois P, Chung-Faye G, Sherwood R, Bjarnason I, Hayee B. Early change in faecal calprotectin predicts primary non-response to anti-TNF α therapy in Crohn's disease. *Scand J Gastroenterol*. 2016; 51:1447–1452. [PubMed: 27400728]
- Ramasundara M, Leach ST, Lemberg DA, Day AS. Defensins and inflammation: the role of defensins in inflammatory bowel disease. *J Gastroenterol Hepatol*. 2009; 24:202–208. [PubMed: 19215333]
- Ramiro S, Smolen JS, Landewe R, van der Heijde D, Dougados M, Emery P, de Wit M, Cutolo M, Oliver S, Gossec L. Pharmacological treatment of psoriatic arthritis: a systematic literature review for the 2015 update of the EULAR recommendations for the management of psoriatic arthritis. *Ann Rheum Dis*. 2016; 75:490–498. [PubMed: 26660203]
- Rutgeerts P, Goobes K, Peeters M, Hiele M, Penninckx F, Aerts R, Kerremans R, Vantrappen G. Effect of faecal stream diversion on recurrence of Crohn's disease in the neoterminal ileum. *Lancet*. 1991; 338:771–774. [PubMed: 1681159]
- Shaw KA, Bertha M, Hofmekler T, Chopra P, Vatanen T, Srivatsa A, Prince J, Kumar A, Sauer C, Zwick ME, et al. Dysbiosis, inflammation, and response to treatment: a longitudinal study of pediatric subjects with newly diagnosed inflammatory bowel disease. *Genome Med*. 2016; 8:75. [PubMed: 27412252]
- Shelton E, Allegretti JR, Stevens B, Lucci M, Khalili H, Nguyen DD, Sauk J, Giallourakis C, Garber J, Hamilton MJ, et al. Efficacy of Vedolizumab as Induction Therapy in Refractory IBD Patients: A Multicenter Cohort. *Inflamm Bowel Dis*. 2015; 21:2879–2885. [PubMed: 26288002]
- Siegel CA, Melmed GY. Predicting response to Anti-TNF Agents for the treatment of crohn's disease. *Therap Adv Gastroenterol*. 2009; 2:245–251.
- Singh JA, Saag KG, Bridges SL Jr, Akl EA, Bannuru RR, Sullivan MC, Vaysbrot E, McNaughton C, Osani M, Shmerling RH, et al. 2015 American College of Rheumatology Guideline for the Treatment of Rheumatoid Arthritis. *Arthritis Rheumatol*. 2016; 68:1–26.
- Singh S, Graff LA, Bernstein CN. Do NSAIDs, antibiotics, infections, or stress trigger flares in IBD? *Am J Gastroenterol*. 2009; 104:1298–1313. quiz 1314. [PubMed: 19337242]
- Truong DT, Franzosa EA, Tickle TL, Scholz M, Weingart G, Pasolli E, Tett A, Huttenhower C, Segata N. MetaPhlAn2 for enhanced metagenomic taxonomic profiling. *Nat Methods*. 2015; 12:902–903. [PubMed: 26418763]
- Ungaro R, Bernstein CN, Garry R, Hviid A, Kolho KL, Kronman MP, Shaw S, Van Kruiningen H, Colombel JF, Atreja A. Antibiotics associated with increased risk of new-onset Crohn's disease but not ulcerative colitis: a meta-analysis. *Am J Gastroenterol*. 2014; 109:1728–1738. [PubMed: 25223575]
- Upchurch KS, Kay J. Evolution of treatment for rheumatoid arthritis. *Rheumatology (Oxford)*. 2012; 51(Suppl 6):vi28–36. [PubMed: 23221584]
- Walmsley RS, Ayres RC, Pounder RE, Allan RN. A simple clinical colitis activity index. *Gut*. 1998; 43:29–32. [PubMed: 9771402]

- Wu HJ, Ivanov II, Darce J, Hattori K, Shima T, Umesaki Y, Littman DR, Benoist C, Mathis D. Gut-residing segmented filamentous bacteria drive autoimmune arthritis via T helper 17 cells. *Immunity*. 2010; 32:815–827. [PubMed: 20620945]
- Zimmerman MA, Singh N, Martin PM, Thangaraju M, Ganapathy V, Waller JL, Shi H, Robertson KD, Munn DH, Liu K. Butyrate suppresses colonic inflammation through HDAC1-dependent Fas upregulation and Fas-mediated apoptosis of T cells. *Am J Physiol Gastrointest Liver Physiol*. 2012; 302:G1405–1415. [PubMed: 22517765]

Figure 1. The differences of baseline stool samples between remission group and non-remission group

(a) The alpha-diversity measured in Fisher's alpha in remission and non-remission groups, segregated by diagnosis; (b) the beta-diversity measured by Bray-Curtis dissimilarity in intra- and inter-group fashion in remission and non-remission groups among CD and UC patients; (c–d) PCoA plots of baseline samples for CD and UC patients; and (e–f) the top 15 most abundant species in baseline samples for CD and UC patients. (box marks the interquartile range (IQR), the whiskers mark the range between lower quartile-1.5 IQR and higher quartile+1.5 IQR, and dots mark the outliers; *, $q < 0.1$; **, $q < 0.01$; ***, $q < 0.001$; ns, not significant).

Figure 2. The significantly differentiated taxa and pathways between remission and non-remission groups in baseline samples

(a) Two taxa, *Burkholderiales* and *Roseburia inulinivorans*, were significantly more abundant in CD remission baseline samples; and (b) pathways that were significantly differentiated between remission and non-remission groups in baseline samples for CD (left) and UC (right) patients ($q < 0.1$). (Pathway codes: A, super-pathway of arginine and polyamine biosynthesis; B, super-pathway of branched amino acid biosynthesis; C, Calvin-Benson-Bassham cycle; D, L-citrulline biosynthesis; E, dTDP-L-rhamnose biosynthesis; F, super-pathway of N-acetylglucosamine, N-acetylmannosamin and N-acetylneuraminic acid degradation; G, super-pathway of β -D-glucuronide and D-glucuronate degradation; H, super-pathway of hexitol degradation; I, L-isoleucine biosynthesis; J, super-pathway of polyamine biosynthesis; K, L-histidine degradation; L, GDP-mannose biosynthesis; M, acetyl-CoA fermentation to butanoate; N, colonic acid building blocks biosynthesis; O, lipid IVA biosynthesis; P, N10-formyl-tetrahydrofolate biosynthesis; Q, pentose phosphate pathway; R, pyruvate fermentation to acetate and lactate).

Figure 3. Longitudinal changes in taxa and pathways between remission and non-remission groups

(a–b) Log2 fold change (log2FC) in CD (a) and UC (b) patients' microbiome pathways that represented significant change at week 14 follow-up in comparison with baseline samples, divided into remission and non-remission groups (FDR<0.1); (c) log2FC of species that represented significant change at week 14 follow-up in comparison with baseline sample (left panel, CD; right panel, UC), divided into remission and non-remission groups (FDR<0.1); and (d) the persistency index, P , for subjects with a later follow-up (wk30 or wk54) available. Horizontal bars indicate the t -test performed on respect group pair and the significance level (p <0.05; **, p <0.01; ***, p <0.001; ns, not significant).

(Pathway codes: A, super-pathway of arginine and polyamine biosynthesis; B, super-pathway of branched amino acid biosynthesis; C, Calvin-Benson-Bassham cycle; D, L-citrulline biosynthesis; E, dTDP-L-rhamnose biosynthesis I; F, super-pathway of N-acetylglucosamine, N-acetylmannosamin and N-acetylneuraminate degradation; G, super-pathway of β -D-glucuronide and D-glucuronate degradation; H, super-pathway of hexitol degradation; I, L-isoleucine biosynthesis I; J, super-pathway of polyamine biosynthesis I; K, L-histidine degradation III; L, GDP-mannose biosynthesis; M, acetyl-CoA fermentation to butanoate II; N, colonic acid building blocks biosynthesis; O, lipid IVA biosynthesis; P, N10-formyl-tetrahydrofolate biosynthesis; Q, pentose phosphate pathway; R, pyruvate fermentation to acetate and lactate II.).

Figure 4. The architecture, training, and performance of vedoNet

(a) The vedoNet and associated other model variates (vedoNet.tx, vedoNet.hybrid, etc) are based on a neural network structure with an input layer, a few hidden layers with softmax dropout and rectified linear unit, and a binary output layer to classify if input data will support treatment outcome as remission or non-remission. The input data is a vector with two parts: the clinical metadata, and the microbiome profile which varies for different models (pathways, taxa, or a combination of both). The training deployed a 5-fold cross validation scheme, which resampled the subjects without replacement for test set and train set.