Deep Sequencing of Crohn's Disease Lamina Propria Phagocytes Identifies Pathobionts and Correlates With Pro-Inflammatory Gene Expression

Gillian E. Jacobsen, PhD,[*,†,‡] Eddy E. Gonzalez, PhD,[‡] Payton Mendygral, BS,[‡]
Katerina M. Faust, PhD,[‡] Hajar Hazime, PhD,[*,‡] Irina Fernandez, MS,[‡] Ana M. Santander, BS,[‡]
Maria A. Quintero, MPH,[‡] Chunsu Jiang, MD,[‡] Oriana M. Damas, MD,[‡]
Amar R. Deshpande, MD,[‡] David H. Kerman, MD,[‡] Siobhan Proksell, MD,[‡]
Morgan Sendzischew Shane, MD,[‡] Daniel A. Sussman, MD,[‡] Bassel Ghaddar, PhD,[§]
Trevor Cickovsk, PhD,[¶] and Maria T. Abreu, MD[*,‡,ǁ]

[*]Department of Microbiology and Immunology, University of Miami Miller School of Medicine, Miami, FL, USA
[†]Medical Scientist Training Program, University of Miami Miller School of Medicine, Miami, FL, USA
[‡]Department of Medicine, Crohn's and Colitis Center, University of Miami Miller School of Medicine, Miami, FL, USA
[§]Center for Systems and Computational Biology, Rutgers Cancer Institute of New Jersey, Rutgers University, New Brunswick, NJ, USA
[¶]Bioinformatics Research Group (BioRG), Knight Foundation School of Computing and Information Sciences, Florida International University, Miami, FL, USA
[ǁ]F. Widjaja Inflammatory Bowel Disease Institute, Cedars-Sinai Medical Center, Los Angeles, CA, USA
Address correspondence to: Maria T. Abreu, MD, Cedars-Sinai Medical Center, 10 George Burns Road, Davis Building 4063, Los Angeles, CA, USA. Phone: (310) 423-7723 (maria.abreu@cshs.org).

Background: Crohn's disease (CD) is characterized by an inflammatory response to gut microbiota. Macrophages and dendritic cells play an active role in CD inflammation. Specific microbiota have been implicated in the pathogenesis of ileal CD. We investigated the phagocyte-associated microbiome using an unbiased sequencing approach to identify potential pathobionts and elucidate the host response to these microbes.

Methods: We collected ileal and colonic mucosal biopsies from CD patients and controls without inflammatory bowel disease (IBD), isolated lamina propria phagocytes (CD11b[+] cells), and performed deep RNA sequencing (*n* = 37). Reads were mapped to the human genome for host gene expression analysis and a prokaryotic database for microbiome taxonomic and metatranscriptomic profiling. Results were confirmed in a second IBD cohort (*n* = 17). Lysed lamina propria cells were plated for bacterial culturing; isolated colonies underwent whole genome sequencing (*n* = 11).

Results: Crohn's disease ileal phagocytes contained higher relative abundances of *Escherichia coli*, *Ruminococcus gnavus*, and *Enterocloster* spp. than those from controls. CD phagocyte-associated microbes had increased expression of lipopolysaccharide (LPS) biosynthesis pathways. Phagocytes with a higher pathobiont burden showed increased expression of pro-inflammatory and antimicrobial genes, including *PI3* (antimicrobial peptide) and *BPIFB1* (LPS-binding molecule). *E. coli* isolated from the CD lamina propria had more flagellar motility and antibiotic resistance genes than control-derived strains.

Conclusions: Lamina propria resident phagocytes harbor bacterial strains that may act as pathobionts in CD. Our findings shed light on the role of pathobionts and the immune response in CD pathogenesis and suggest new targets for therapies.

Lay Summary

We investigated the phagocyte-associated microbiome in Crohn's disease patients. Diverse pathobionts within lamina propria phagocytes displayed shared virulence factors (including increased expression of lipopolysaccharide biosynthesis and antimicrobial resistance pathways) and were linked to an inflammatory profile, potentially perpetuating disease.

Key Words: inflammatory bowel disease, microbiome, metatranscriptome, lipopolysaccharide

Introduction

Crohn's disease (CD), a subtype of inflammatory bowel disease (IBD), is an immune-mediated disorder of the gastrointestinal tract in which chronic inflammation occurs in the absence of a known pathogen. The disease is thought to arise from an aberrant inflammatory response to gut microbiota.[1] Compared to the microbiome of healthy individuals, the IBD microbiome is characterized by a decreased abundance of *Firmicutes* and butyrate-producing bacteria (eg, *Faecalibacterium prausnitzii*) and increased *Proteobacteria* and *Actinobacteria*.[1,2] Specific pathobionts have also been implicated in IBD pathogenesis; these species are native to the gut environment but may cause inflammation in susceptible individuals. For example, adherent invasive *Escherichia*

Key Messages

What is already known?

- Crohn's disease (CD) is characterized by an inflammatory response to gut microbiota, but identification of specific pathobionts and their relationship to inflammatory pathways has been elusive.

What is new here?

- We found that lamina propria phagocytes harbored specific bacterial species that may act as pathobionts in CD and exhibited increased expression of pro-inflammatory and antimicrobial genes in response.

How can this study help patient care?

- Our findings suggest new targets for therapies, such as targeting lipopolysaccharide-interacting genes and tailoring treatments to each patients' microbiome.

coli (AIEC) and *Clostridium innocuum* have been identified in ileal biopsies and resections from CD patients.[3,4] However, each of these pathobionts has only been observed in a subset of CD patients,[3,4] and studies have not definitively proven the causal role of these microbes in CD pathogenesis.

The exact relationship between microbes and the immune system in IBD pathogenesis remains unclear partially due to the limited methods for studying host-microbe interactions in IBD. 16S rRNA gene sequencing provides limited taxonomic information and no functional information, while shotgun metagenomics is challenging to perform on tissue samples against the high background levels of host genes. Furthermore, most microbiome studies have used stool samples, which are not completely representative of the microbiome at the gut mucosal interface or within the mucosa.[3,5] Therefore, the majority of IBD microbiome studies have been unable to identify specific causal organisms and have been limited in their conclusions.

Within the gut mucosa, phagocytic immune cells such as macrophages and dendritic cells interface with microbes at the luminal surface and within the lamina propria. In IBD, the function of these cells may be impaired: many CD susceptibility genes (including *ATG16L1*, *CARD9*, and *NOD2*) are related to microbial recognition and clearance.[6,7] Furthermore, certain pathobionts may have the capacity to cross the epithelial barrier and reside inside lamina propria cells to escape detection and clearance by the immune system. For example, AIEC can colonize macrophages in vitro and potentially trigger the formation of non-caseating granulomas in CD.[8] We posited that studying gut mucosal phagocytes and their associated pathobionts may provide a key to understanding CD pathogenesis.

In a previous study, we demonstrated that the lamina propria phagocyte-associated microbiome is distinct from the mucosa-associated microbiome.[9] We found that IBD-specific dysbiotic trends such as increased abundance of *Proteobacteria* and *Actinobacteria* were magnified in the phagocyte-associated microbiome and that *Enterobacteriaceae* were associated with the expression of pro-inflammatory genes such as IL1B, chemokine (C-X-C motif ligand 8 [CXCL8], and NOD-, LRR-, and pyrin-domain containing protein 3 [NLRP3]). However, the methods used (16S rRNA gene sequencing and

NanoString) were inherently limited, and the lack of healthy control samples restricted our conclusions.

In the current study, we aimed to identify bacteria at the species level and phagocyte responses unique to CD by comparing samples from CD patients to those from non-IBD controls. We used RNA sequencing to capture host and microbial sequences, which we subsequently analyzed to determine microbial composition, microbial functions, host gene expression, and correlations between these. We found that several taxa had increased abundance in CD phagocytes, particularly in the ileum, and that CD-associated microbes had shared functions including lipopolysaccharide (LPS) synthesis and antimicrobial resistance (AMR). We corroborated several of our findings in a second cohort using an RNA sequencing dataset from a previous study. In addition, we isolated *E. coli* strains from both CD patients and controls and found that CD strains were enriched for flagellar motility and AMR genes. Our findings help illuminate the role of microbial-immune interactions in CD pathogenesis and suggest targets for future therapies.

Materials and Methods

Subject Enrollment and Sample Collection

CD patients and non-IBD controls were enrolled through our ongoing biospecimen collection at the University of Miami Crohn's and Colitis Center (Institutional Review Board approval: 20081100). Notably, we focused on recruiting patients with ileal disease; patients could have ileal disease with or without colonic involvement. Non-IBD controls were undergoing screening colonoscopies and had no gastrointestinal symptoms (bleeding, diarrhea) and no personal history of gastrointestinal disease, autoimmune disease, or cancer. Samples were collected from inflamed sites when possible (ie, patients with active CD). Site inflammation was determined from the endoscopist's interpretation and the pathologist's histology report. Four to 6 gut mucosal biopsies were taken from the terminal ileum and/or ascending colon using jumbo forceps. Biopsies were placed in HypoThermosol solution (MilliporeSigma) and brought to the laboratory for same-day processing.

Sample Processing and RNA Extraction

Gut mucosal biopsies were washed and suspended in Dulbecco's modified Eagle medium (DMEM) with 8 μg/mL of gentamicin, 10 mM dithiothreitol, and 0.5 mM ethylenediaminetetraacetic acid. Samples were incubated at room temperature with shaking for 20 minutes, after which the supernatant, containing the majority of epithelial cells from the biopsies, was removed. The remaining biopsy tissue was washed and resuspended in DMEM with 8 μg/mL of gentamicin, 250 μg/mL of Liberase (MilliporeSigma), and 10 μg/mL DNase I (Lucigen). Samples were incubated at 37 °C with shaking for 45 minutes to dissociate the tissue, further mechanically dissociated by repeat pipetting, and filtered with a 70-μm strainer to produce single-cell suspensions. Samples were then labeled with CD11b MicroBeads (Miltenyi Biotec) for positive selection with magnetic-activated cell separation on LS columns (Miltenyi Biotec). RNase-free pellet pestles were used to gently lyse CD11b⁺ cells and intracellular microbes in RLT buffer with 10% β-mercaptoethanol. RNA extraction was performed using the AllPrep DNA/

RNA Kit (QIAGEN). Our previous study[10] showed that this method largely enriches phagocytes, including granulocytes, macrophages, monocytes, and dendritic cells.

RNA Sequencing

Sufficient quantities of RNA for sequencing were obtained from 37 samples. Libraries were prepped and sequenced at the University of Miami's John P. Hussman Institute for Human Genomics Sequencing Core. Total RNA was prepped with the TECAN Universal Plus Total RNA-Seq with NuQuant®, Human AnyDeplete using 5-60 ng via Qubit and 19 cycles of polymerase chain reaction (PCR). Libraries were sequenced on 6 lanes of Illumina NovaSeqX, and>100M PE150 reads were generated per sample.

RNA Sequencing Bioinformatics Analysis

Of the sequenced samples, 2 samples from one CD patient were excluded from the study due to a confounding acute infection. The remaining 15 ileum and 20 colon samples were included in the analysis. An additional 17 ileal samples from a prior study[10] (GSE183620) were analyzed in the confirmatory cohort.

Host and microbial RNA sequencing analyses were performed using CLC Genomics Workbench v23.0 and v24.0 (QIAGEN). Raw FASTQ files were trimmed, quality checked, and mapped to the human genome (hg38; Ensembl v106.1) to generate human transcript and gene expression tracks. Unmapped reads were saved for microbial analysis. Human gene expression tracks were normalized to reads per kilobase of exon model per million mapped reads. The Differential Expression for RNA-Seq tool, which uses multi-factorial statistics based on a negative binomial general linear model, was used to generate lists of differentially expressed genes and create volcano plots for various comparisons (eg, CD vs control). The Gene Set Test tool was used to determine if Gene Ontology (GO) terms were over-represented in the differentially expressed genes using a hypergeometric test. The significance threshold for differential gene expression and gene set tests was set at a false discovery rate (FDR)-adjusted P-value < .05, with a $\log_2$ fold change of 2 or more.

Reads that did not map to the human genome ("unmapped reads") were mapped to the QIAGEN Microbial Insights–Prokaryotic Taxonomy Database (QMI-PTDB) Genus v2.0 for taxonomic profiling. This curated database containing all genera is optimized for balanced taxonomic representation and contains one representative genome per bacterial or archaeal species. The CLC Genomics Workbench software uses a microbial mapping tool that has a low false positive rate compared to other popular tools.[11] Differential relative abundance was calculated using the Differential Abundance Analysis tool based on a generalized linear model. Significance was determined using the Wald test, with the significance threshold set at an FDR-adjusted P-value < .05. Differential abundance in comparisons between the CD and control groups was calculated at each taxonomic level for ileum and colon samples separately.

For functional profiling, a sequence list of detected taxa was subset from the QMI-PTDB and annotated with GO terms. The number of reads mapping to GO terms or protein family (Pfam) domains was calculated using the Build Functional Profile tool. GO functional profiles or Pfam domains were statistically compared between CD and control groups using the Differential Abundance Analysis tool. The significance threshold was set at an FDR-adjusted P-value < .05, with a $\log_2$ fold change of 1.5 or more.

AMR marker gene profiling was performed using the ShortBRED tool using the QMI-AMR Peptide Marker Database (7.0). The Differential Abundance Analysis tool was used to compare AMR gene counts in CD and control groups.

E. coli Isolation and DNA Extraction

A portion of lamina propria cells were lysed using a mammalian cell lysis buffer (1% Triton X-100 in water) and plated on *Klebsiella*-selective media (MilliporeSigma) to select for *Enterobacteriaceae* (*Escherichia*, *Klebsiella*, *Salmonella*, etc.). Plates were incubated at 37 °C for 24 hours. Individual colonies were counted, and one colony per positive sample was selected for whole genome sequencing. The selected colonies were grown overnight in brain heart infusion broth (MilliporeSigma) and then centrifuged at $3000 \times g$ for 5 minutes to form pellets. Pellets were lysed in RLT buffer (QIAGEN), and DNA was extracted using the DNeasy Blood and Tissue Kit (QIAGEN) optimized for Gram-negative bacteria. The final resulting isolate DNA from 9 different patient samples was frozen and sent on dry ice overnight to CosmosID Inc. for whole genome sequencing.

Whole Genome Sequencing and Bioinformatics Analysis of *E. coli* Isolates

Whole genome sequencing was performed by CosmosID Inc. Briefly, DNA samples were quantified with Qubit 4 fluorometer and QubitTM dsDNA HS Assay Kit (Thermo Fisher Scientific). Libraries were prepared with IDT Unique Dual Indexes using the Nextera XT DNA Library Preparation Kit (Illumina) and purified using AMpure magnetic beads (Beckman Coulter, Brea, CA) in EB buffer (QIAGEN). Libraries were then sequenced on Illumina NovaSeq PE150.

CosmosID Inc. performed de novo genome assembly and genome quality assessment and created a single nucleotide polymorphism (SNP) tree based on core genome phylogeny. Briefly, raw fastq files were trimmed and processed using BBDuk. The trimmed fastqs were assembled using SPAdes, and assembly completeness was evaluated using CheckM's lineage_wf function. The assembled contigs for each isolate were then processed with Parsnp to align to the core genome and FastTree2 to construct a phylogenomic tree.

CLC Genomics Workbench versions 23.0 and 24.0 (QIAGEN) were used for further analysis of *E. coli* isolates, including functional profiling and AMR gene set analysis, as described above in the RNA sequencing bioinformatics analysis.

Statistical Analysis

Comparisons between CD and control groups were performed using unpaired *t*-tests (for normally distributed continuous data) or Mann-Whitney U tests (non-normally distributed continuous data). Categorical demographic data were compared using chi-square tests and Fisher's exact tests. The significance threshold for these tests was set at P-value < .05.

Other bioinformatics analysis tools from the CLC Genomics Workbench software described above used custom statistical calculations curated for each tool (Differential Expression for RNA-Seq, Gene Set Test, and Differential Abundance Analysis). The significance threshold for AMR

gene differential abundance was set at *P*-value < .1 due to the short list of genes/phenotypes compared. The significance threshold for all other bioinformatics analyses was set at an FDR-adjusted *P*-value < .05.

Results

Patient and Sample Characteristics

CD patients and non-IBD controls undergoing screening colonoscopies were enrolled during endoscopy at the University of Miami Health system. The majority of patients were Hispanic, reflecting the large Hispanic population in our study area. From each patient, we collected 6 biopsies per sample from the ileum and the colon. CD patients underwent colonoscopy for disease monitoring; control patients underwent colonoscopy for routine colorectal cancer screening. We included CD patients with different phenotypes (fibrostenotic, inflammatory, or perforating) and current treatments (none, mesalamines, anti-tumor necrosis factor [TNF], anti-IL12/23, or Janus kinase [JAK] inhibitors) to capture the full range of disease. Control patients were selected based on the absence of gastrointestinal symptoms (bleeding, diarrhea) and no personal history of gastrointestinal disease, autoimmune disease, or cancer. None of the CD or control patients had recent (<3 months) antibiotic use.

Gut mucosal biopsies from the ileum and the colon of each patient were digested into single cells and treated with gentamicin to deplete extracellular bacteria; gentamicin does not penetrate mammalian cell membranes, and therefore intracellular bacteria should remain intact with this method. To enrich for cells potentially harboring microbes, we selected CD11b as a surface marker; this protein is involved in phagocytosis,[12] and the majority of CD11b+-enriched cells consist of phagocytes, as demonstrated by our previous study.[10] We isolated CD11b+ cells from 50 gut mucosal biopsy samples from 15 CD patients and 10 controls; of these, only 37 samples yielded sufficient RNA for sequencing. An overview of our methods can be seen in Figure S1 and the characteristics of the final patients and samples included in RNA sequencing are shown in Table 1.

The CD and control groups did not significantly differ in any demographic characteristics except for age. The control group had a slightly higher mean age because colorectal cancer screening is typically recommended for individuals>45 years old, whereas CD can be diagnosed at any age, and young adults have the highest incidence rate.[13]

Prokaryotic RNA Is Detectable in Lamina Propria Phagocytes

RNA samples from CD11b+ cells were deep sequenced at>100 million reads per sample. CLC Genomics Workbench (QIAGEN) was used to map reads to the human genome for host analysis; the remaining reads were mapped to a prokaryote genomic database for microbiome taxonomic and metatranscriptomic profiling. In all CD and control samples, we detected an average of 94.36% human reads and 0.25% prokaryotic reads (Figure 1A). CD and control samples did not differ in the total number of reads captured (Figure 1B), but the percentage of prokaryotic reads was slightly lower (and inversely, the percentage of human reads was slightly higher) in CD samples (Figure 1C and D). This may indirectly reflect decreased absolute bacterial abundance in

CD samples, a feature of CD dysbiosis.[14] The difference in read percentages did not affect the downstream results because each sample had many prokaryotic reads (mean ± SD: 602 539 ± 289 337) and Shannon alpha diversity plateaued in the majority of samples, suggesting that we sequenced to a sufficient depth to capture most microbial sequences (Figure 1E). Thus, we were able to subsequently characterize the phagocyte-associated microbiome and metatranscriptome in CD and control samples.

Potential Pathobionts Are Enriched in Lamina Propria Phagocytes of CD Patients

Through taxonomic profiling, we detected an average of 410 ± 184 taxa in ileal samples and 318 ± 132 taxa in colonic samples. The observed number of taxa did not differ significantly based on group (CD patient vs healthy control) (Figure 1F and G). We then calculated alpha and beta diversity metrics. In terms of alpha diversity (Shannon entropy, Chao1, Simpson's index), there were no significant differences between CD and control samples (Figure 1H and I, Figure S2A). Although the CD and control groups differed in mean age, we did not see any trends in alpha diversity based on age (Figure S2B). We also grouped samples according to sex and found that samples from male patients displayed a higher alpha diversity (Shannon entropy) than samples from female patients; this was true in the ileum but not in the colon (Supplementary Data). However, this sex difference does not confound the results of diagnostic group comparisons (CD vs control) because males and females were evenly distributed in each group. Grouping by other biological/clinical variables (age, body mass index [BMI], ethnicity, race, medication, disease phenotype) yielded no significant differences in alpha diversity (Supplementary Data). When visualizing beta diversity (Bray-Curtis dissimilarity) (Figure 1J), we found that ileal samples from the control group appeared to cluster in the center, while those from the CD group were more dispersed along PCo2. Among the ileal samples, no trends in beta diversity by age were observed (Figure S3). No trends were observed based on other variables (Figures S4 and S5). Colon samples were not observed to cluster by diagnosis nor any other variables (Figure 1K, Figures S4 and S5).

We then performed differential abundance analysis to detect taxa at each taxonomic level with higher relative abundance in CD patients or controls. Due to known differences in baseline microbiota composition,[15] ileal and colonic samples were analyzed separately. Among the species that were present in at least half of CD or control ileal samples, 7 species were significantly (FDR-adjusted *P*-value > .05) more abundant in CD patients and 3 were more abundant in controls (Figure 2A). Three of the 7 species elevated in CD patients (*E. coli*, *Ruminococcus gnavus*, and *Enterocloster clostridioformis*) are gut-resident bacteria that have been previously linked to CD.[16,17] For example, *R. gnavus* is a prevalent gut microbe that increases in abundance during CD flares.[16] The reads classified as unknown *Enterobacteriaceae* may be *E. coli* reads or shared genes with other related genera, such as *Klebsiella* or *Salmonella*; as they are shared, they cannot be pinpointed to a specific species. Similarly, the unknown *Enterocloster* sp. may be *E. clostridioformis* reads or another closely related species, such as *Enterocloster bolteae* (which is associated with hepatitis B[18]). One of the 3 species high in control samples was an unknown species of *Peptostreptococcaceae*, a family

Table 1. Patient characteristics.

Total (N = 20)	CD (n = 11)	Control (n = 9)	P-value
Sex			
Female	4 (36%)	5 (56%)	.6534 (ns)
Male	7 (64%)	4 (44%)	
Ethnicity			
Hispanic	9 (82%)	3 (33%)	.0648 (ns)
Non-Hispanic	2 (18%)	6 (67%)	
Race			
Asian	0	1 (11%)	.0737 (ns)
Black	0	2 (22%)	
White	11 (100%)	6 (67%)	
Birth country			
United States	7 (64%)	7 (78%)	
Other	4 (36%)	2 (22%)	
Age (years)			
Mean	38.09	60.33	.0008 (a)
Body mass index (kg/m²)			
Mean	24.39	30.65	.0583 (ns)
Medications			
Antibiotics	0	0	
Probiotics	1 (9%)	0	
PPIs	0	0	
CD treatment			
None	1 (9%)	9 (100%)	
Aminosalicylates	1 (9%)	—	
Immunomodulators	1 (9%)	—	
Steroids	1 (9%)	—	
Biologics	9 (81%)	—	
Anti-TNF	4 (36%)	—	
Anti-integrin	0	—	
Anti-IL-12/23	3 (27%)	—	
JAK inhibitor	2 (18%)	—	
Disease phenotype			
Inflammatory (B1)	8 (72%)	—	
Fibrostenotic (B2)	7 (64%)	—	
Penetrating/perforating (B3)	3 (27%)	—	
Perianal [b]	1 (9%)	—	
Location of disease			
Ileum only (L1)	4 (36%)	—	
Colon only (L2)	1 (9%)	—	
Ileocolon (L3)	6 (55%)	—	
Upper GI (L4) [b]	0	—	
Ileal endoscopy			
Actively inflamed	5 (46%)	—	
Previously involved, uninflamed	2 (18%)	—	
Never involved	4 (36%)	—	
Ileal histology			
Inflamed	5 (46%)	—	
Colon endoscopy [c]			
Actively inflamed	1 (9%)	—	
Previously involved, uninflamed	3 (27%)	—	

Table 1. Continued

Total (N = 20)	CD (n = 11)	Control (n = 9)	P-value
Never involved	6 (55%)	—	
Colon histology			
Inflamed	0	—	

[a]P<.005, ns = not significant.
[b]Modifier to the above categories, not category-exclusive.
[c]One patient lacked a colon biopsy.

that produces indole metabolites that regulate intestinal barrier function.[19] An unknown member of this family was previously shown to be decreased in CD.[20] The rest of the differential species (*YIM-78166* sp., *Thermosynechococcaceae*, *Corallococcus exiguus*) are either ubiquitous microbes found in soil or aquatic environments or otherwise undescribed.[21,22]

We also found some significant differences in the abundance of taxa at the family level in the ileum (Figure 2A). Most of these differences aligned with the species-level trends (eg, *Enterobacteriaceae* to *E. coli*, *Lachnospiraceae* to *Enterocloster* sp.). *Dermabacteraceae* abundance was higher in CD samples; this family includes skin-dwelling opportunistic pathogens and soil and aquatic microbes.[23] Overall, most of the taxa with increased abundance in CD samples are potential pathobionts, while those with increased abundance in controls are either protective or of unknown relevance.

In the colon, only 2 species that were observed in at least half of CD samples were significantly more abundant in CD samples compared to controls: *Symbiobacter mobilis* and an unknown species of *Halomonadaceae* (Figure 2B). The *Halomonadaceae* family of bacteria is enriched in colorectal cancer tissue,[24] implicating it as a potential pathobiont. *Symbiobacter mobilis* is an aquatic organism of unknown significance to human health.[25] Only one species was found in higher abundance in controls: an unknown *Amphibacillaceae* species. *Amphibacillus* is known to degrade indigo dyes[26] but has unknown significance to human health. No significant trends were observed at the family level in colonic samples.

Some differentially abundant taxa were present in only a few samples (Supplementary Data) and therefore were not included in the above statistical analyses. For example, *Bacteroides fragilis* and *Bacteroides ovatus*, pathobionts associated with CD,[27,28] were highly abundant in 3 or 4 different CD ileal and colonic samples. *Salmonella enterica*, an intracellular gut pathogen that may trigger IBD onset in susceptible individuals,[29] was also observed in 2 CD ileal samples. Importantly, none of these species were detected in control samples. The diversity of pathogens and pathobionts found in CD samples but not in control samples suggests that it is not one specific pathobiont but rather a group of pathobionts that contribute to CD pathogenesis by crossing the epithelial barrier and triggering inflammation.

CD Phagocyte-Associated Microbes Express More LPS Synthesis and AMR Genes

We hypothesized that the microbes significantly more abundant in CD phagocytes share functions that cause them to trigger inflammation or evade immune detection. In order to compare CD-associated and control-associated microbes, we performed
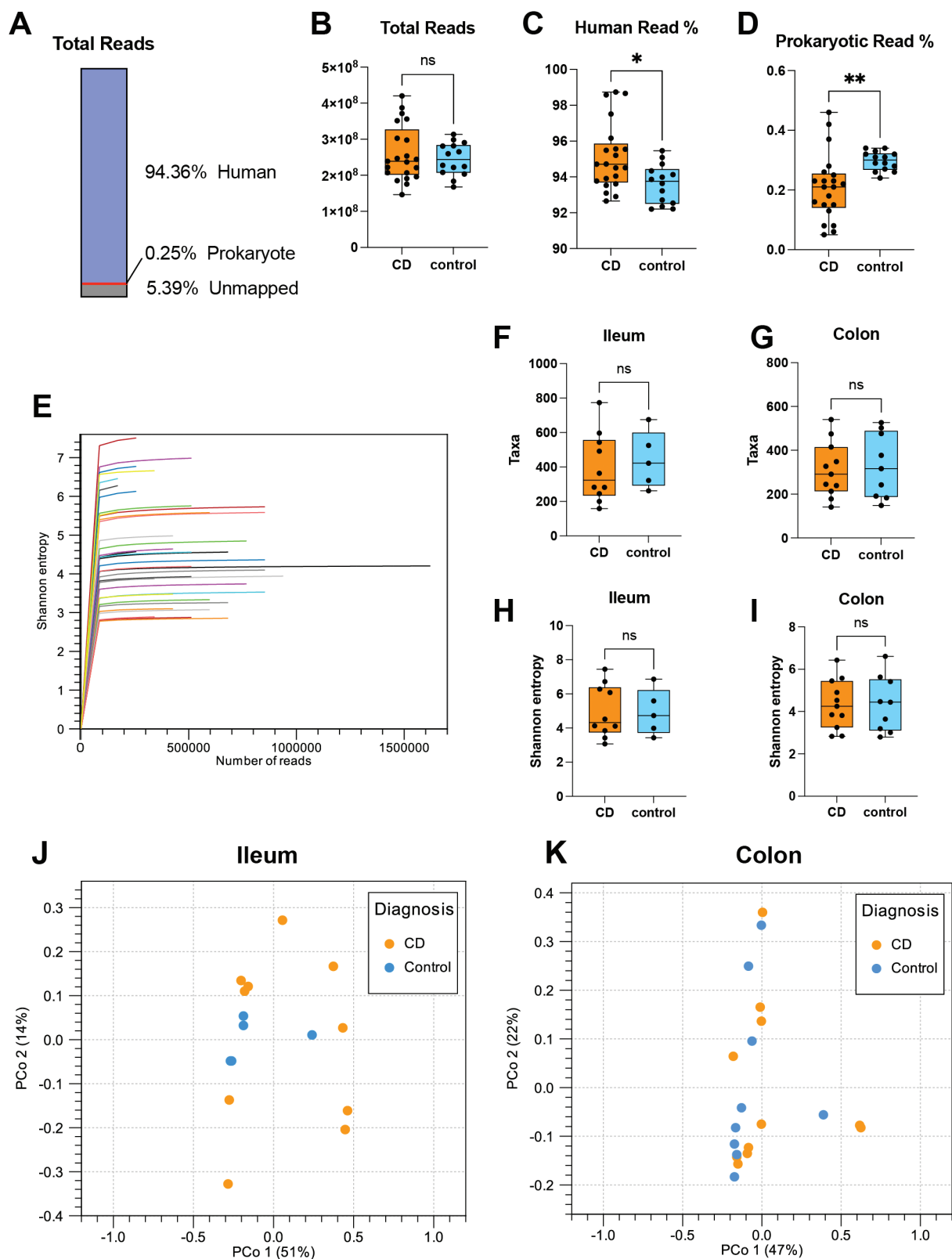
Figure 1. Prokaryotic reads and diversity in Crohn's disease (CD) vs control phagocytes. A, Average percentage of human, prokaryotic, and unmapped reads across all samples ($n = 35$). B, The numbers of total reads in CD ($n = 21$) vs control samples ($n = 14$). C and D, The percentage of human and prokaryotic reads in CD vs control samples. E, The alpha diversity (Shannon entropy) calculated at multiple read depths per sample. Each line represents one sample, and the line terminates at the maximum number of reads achieved for that sample. F and G, The number of taxa detected in CD vs control samples from the ileum ($n = 10$ CD vs 5 control) and the colon ($n = 11$ CD vs 9 control). H and I, The alpha diversity (Shannon entropy) of CD vs control samples from the ileum and colon. J and K, The beta diversity, based on Bray-Curtis dissimilarity, in principal coordinate scatter plots for ileal and colon samples. *P-value < .05, ** P-value < 0.01, ns, not significant.
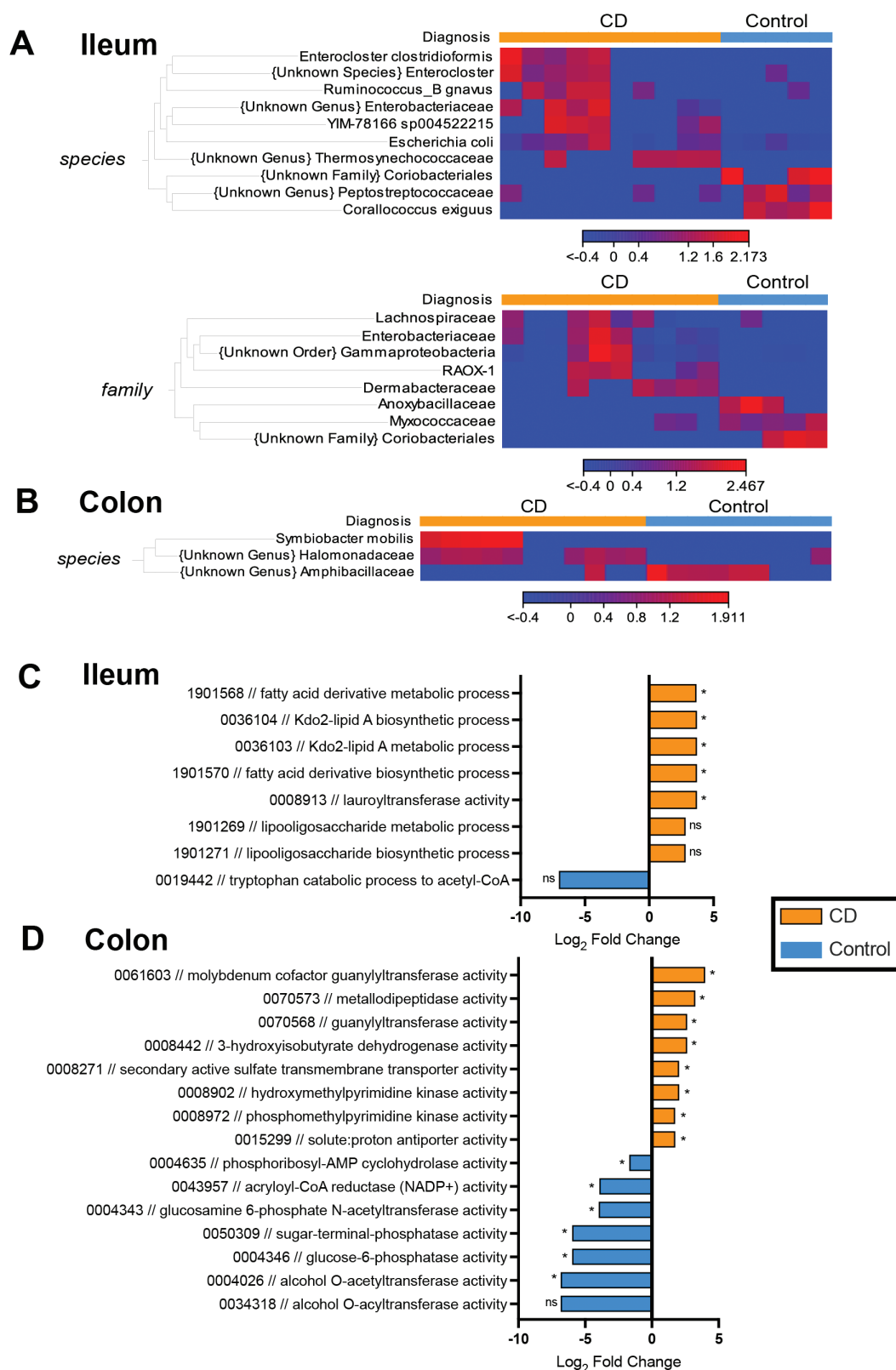
Figure 2. Differentially abundant taxa and microbial pathways in Crohn's disease (CD) vs control phagocytes. A and B, Heat maps, based on z-scores, display significantly (FDR P-value < .05) differentially abundant taxa present in more than half of samples in either group (n = 10 CD, n = 5 control). Each column represents a separate sample, and samples are grouped by diagnosis (CD vs control). A, Ileal samples with differentially abundant taxa at the species and family levels. B, Colon samples with differentially abundant taxa at the species level (no taxa at the family level were significant). C and D, Gene Ontology (GO) biological functions enriched in CD vs control samples from (C) the ileum (n = 10 CD vs n = 5 control) and (D) the colon (n = 11 CD vs n = 9 controls), plotted based on the log$_2$ fold change between groups. *FDR-adjusted P-value < .05. FDR, false discovery rate; ns, not significant.

microbial functional profiling (ie, metatranscriptomics) to determine which GO functions were actively expressed. Ileal and colonic samples were assessed separately.

In the ileum, 5 GO functions were significantly enriched (FDR-adjusted P-value < .05) in CD samples: lauroyltransferase activity, Kdo2-lipid A biosynthetic and metabolic processes, and fatty acid derivative metabolic and biosynthetic processes (Figure 2C). Lauroyltransferase activity and Kdo2-lipid A biosynthetic/metabolic processes are both involved in LPS synthesis.[30] LPS is a potent inducer of inflammation via Toll-like receptor 4 (TLR4) signaling in innate immune cells; TLR4 signaling is known to be increased in IBD.[31] Fatty acid derivative metabolic/biosynthetic processes also could be involved in LPS synthesis, since LPS is a glycolipid. In addition, there was a non-significant trend of increased lipooligosaccharide (LOS) metabolic process in CD samples; LOS is another bacterial surface molecule that induces inflammation.[32] In ileal control samples, the highest GO function (numerically but not significantly higher than in CD samples) was tryptophan catabolic process to acetyl-CoA. Microbial-derived tryptophan catabolites are known to strengthen gut epithelial barrier function and are decreased in IBD patients.[33] Overall, CD-associated microbes appear to have increased synthesis of pro-inflammatory compounds and reduced production of beneficial catabolites compared to control-associated microbes.

Although the colon had fewer significant trends in taxa abundance, it had many more GO functions that significantly differed between CD and control samples (Figure 2D). Overall, the GO functions highly expressed by CD microbes were involved in transmembrane transport (solute:proton antiporter activity, secondary active sulfate transmembrane transporter activity), thiamin biosynthesis (hydroxymethylpyrimidine kinase activity, phosphomethylpyrimidine kinase activity), and protein catabolism (3-hydroxyisobutyrate dehydrogenase activity, metallodipeptidase activity). In contrast, the GO functions enriched in controls were involved in sugar or alcohol metabolism (glucosamine 6-phosphate N-acetyltransferase activity, alcohol O-acetyltransferase activity, sugar-terminal-phosphatase activity, and acryloyl-CoA reductase [NADP+] activity) and histidine biosynthesis (phosphoribosyl-adenosine monophosphate cyclohydrolase activity). These metabolic changes may contribute to pathogenicity or otherwise reflect the differences in nutrient absorption in patients with CD vs controls.

We next evaluated AMR gene expression in these microbes, detecting a total of 25 genes in the ileum and colon samples. Statistical comparisons between CD and control groups were not possible since each of these genes was found in only 1-3 samples. However, most of the AMR genes were detected in CD-associated microbes and not shared with control-associated microbes (Figure S6A). In the ileum, all AMR genes were either CD-only or shared. In particular, CD samples had more gene counts encoding for antibiotic inactivation enzymes (Figure S6B). CD-associated microbes in both the ileum and colon expressed genes encoding APH(3′)-Ia, which confers resistance to aminoglycosides, and various beta-lactamases (Figure S6C). Overall, in addition to having more AMR genes, CD-associated microbes had a wider array of AMR products.

Replication of Pathobiont Findings in an Additional CD Cohort

To strengthen our findings, we set out to corroborate these results in an additional dataset. In our prior study of lamina propria phagocytes,[10] we similarly isolated CD11b$^+$ cells from the ileum and colon of CD and ulcerative colitis (UC) patients and performed RNA sequencing. Since the ileum is not affected in UC patients, we reasoned that we could compare ileal samples from CD and UC patients in this prior cohort ("past cohort") to see if we observed similar microbial taxonomic and functional profiles to the CD vs control samples in the current study cohort ("current cohort"). Notably, our prior study used penicillin/streptomycin instead of gentamicin and sequenced at a shallower depth (40 million reads); thus, we expected to obtain fewer microbial reads from these samples.

After removing human reads and performing taxonomic profiling, we obtained over 100 000 prokaryotic reads from 8 CD and 9 UC ileal samples in the past cohort (mean ± SD: 531 080 ± 552 653 reads). While the average number of prokaryotic reads in ileal samples did not significantly differ between the past cohort and the current cohort, the number of detected taxa was significantly lower in the past cohort (248 ± 100) than in the current cohort (410 ± 184) (P = .0038) (Figure S7A and B). Nevertheless, we detected all 10 of the differentially abundant species from the current cohort in this past cohort. We also ensured that the number of prokaryotic reads and detected taxa was not significantly different between the CD and UC groups (Figure S7C and D). We then performed differential abundance analysis between CD and UC samples in the past cohort. While none of the 10 species showed a significant difference in abundance in the past cohort (FDR-adjusted P-value < .05), many of them showed similar patterns to what we saw in the current cohort (Figure S7E). In particular, *Enterobacteriaceae*, *E. coli*, and *R. gnavus* were detected in more CD samples than UC samples. However, none of the bacteria that were more abundant in control samples in our current cohort showed the same pattern in UC samples in the past cohort. These results suggest that certain pathobionts such as *E. coli* and *R. gnavus* are more abundant in ileal phagocytes from CD patients compared to those from control or UC patients. In addition, there did not appear to be a common protective species found in ileal phagocytes from controls or UC patients.

We next probed the microbial GO functional profiles of these past cohort samples. Although we detected fewer taxa and differentially abundant species in the past cohort than in the current cohort, we observed 15 differentially expressed GO functions between the CD- and UC-associated microbes. All 15 of these GO functions were enriched in the CD microbes (Figure S7F), including all 5 GO functions that were enriched in CD microbes from the current cohort. Most of the GO functions are involved in LPS and LOS synthesis. They also included membrane lipid and fatty acid derivative metabolic/biosynthetic processes; these processes may also be involved in LPS/LOS synthesis. Interestingly, viral RNA genome replication was enriched in CD samples. Overall, these results support our current findings regarding CD-associated microbes and their functions.

Lamina Propria Phagocytes With a High Burden of Pathobionts Exhibit a Unique Pro-Inflammatory Response

After characterizing the phagocyte-associated microbes and their functions, we investigated host phagocyte RNA expression to determine if phagocytes from CD patients exhibited different responses to associated pathobionts than controls.

In the ileum, we found 344 differentially expressed genes in CD vs control samples (Figure 3A). As might be expected, many pro-inflammatory innate immune genes (such as *IL1R2*, *MPO*, *S100A8*, *CXCL5*, and *CXCR2*) had increased expression in CD samples. In addition, expression of genes that influence gut microbiota composition in IBD (*DUOX2*, *DUOXA2*, *MUC1*, and *MUC5B*) was increased in CD samples; these genes are upregulated in the intestinal epithelial cells of IBD patients[34–36] and were likely highly expressed in the small portion of captured epithelial cells. In the control samples, we found that genes with upregulated expression included *NPY* and *TMEM119*, both involved in neuronal function. We then performed a gene set test to identify enriched GO functions among the differentially expressed genes in the ileum, which yielded 177 significant (FDR-adjusted *P*-value < .05) GO terms. The top biological functions included cell surface receptor signaling pathway, defense response, response to external biotic stimulus, and cytokine-mediated signaling pathway (Figure 3B). These findings suggest that the phagocytes in CD patients are actively exhibiting a pro-inflammatory response to perceived pathogens.

In the colon samples, 53 genes were differentially expressed between the CD and control groups (Figure 3C). Compared to those in the ileum, colon phagocytes from CD patients expressed fewer pro-inflammatory genes, likely because the CD patients in our study primarily had ileal disease (Table 1). The top genes expressed in colon CD samples included *APOB*, the primary apolipoprotein of chylomicrons and low-density lipoproteins,[37] and *DEFA6*, an antimicrobial cytotoxic peptide.[38] However, there were no significantly enriched GO terms among the differentially expressed genes in CD vs control samples, suggesting that there were no major differences in inflammatory or metabolic processes in colon phagocytes. We then asked if CD phagocytes expressed any shared genes between the ileum and colon. Among the genes that overlapped in these comparisons (Figure 3D), *DEFA1B* encodes defensin, which is found in mammalian phagocytes[39] and triggers pro-inflammatory cytokine production.[40] It also inhibits bacterial cell wall synthesis, serving as an antibacterial defense mechanism.[41] *AKR1C1* encodes an aldo-keto reductase with a high affinity for binding bile acids.[42] *PI3* is an antimicrobial peptide that is upregulated upon exposure to LPS.[43]

As CD ileal samples were associated with more potential pathobionts, we further delved into host gene expression in response to these pathobionts. We chose 3 species—*E. coli*, *R. gnavus*, and *E. clostridioformis*—that were more abundant in ileal CD samples and most likely to be pathobionts due to previous associations with CD or other diseases.[16,17,44] We categorized ileal samples according to the abundance of each species (low, medium, or high abundance) and compared gene expression between these sample groups. Remarkably, we found 575 differentially expressed genes when comparing samples with high vs low *E. coli* abundance (Figure 4A). The top GO functions enriched in these differentially expressed genes included response to oxygen-containing compound, regulation of mitogen-activated protein kinase (MAPK) cascade, immune response-activating cell surface receptor signaling, positive regulation of nuclear factor-kappa-B (NF-κB) transcription factor activity, and pattern recognition receptor signaling pathway (Figure 4B). These pathways could reflect increased TLR4 signaling or other pattern recognition

receptor signaling in the *E. coli*-high phagocytes. Other pathways involved in phagocytosis, granulocyte activation, and T cell co-stimulation were also enriched. Top genes in the *E. coli*-low samples included C10orf99, which enables chemokine activity,[45] AQP8, an aquaporin, and FRAS1, an extracellular matrix protein.[46] Overall, there seemed to be fewer pro-inflammatory or otherwise innate immunity genes in the *E. coli*-low samples. Similarly, samples with high vs low levels of *R. gnavus* and *E clostridioformis* exhibited 384 and 121 differentially expressed genes, respectively (Figure 4C). Across all 3 high vs low comparisons, we found that 276 genes were shared between at least 2 comparisons, and 60 genes were common among all 3 comparisons. Among these 60 shared genes, the top innate immune genes in pathobiont-high samples included *BPIFB1*, *CXCL5*, *TREM1*, *CXCR2*, and *CXCR1*. Interestingly, *BPIFB1*, a secreted protein that can bind LPS, was the most highly expressed shared gene but has not yet been implicated in CD pathogenesis. Similarly, *ACOD1* (which encodes aconitate decarboxylase 1) was highly expressed in pathobiont-high samples. This protein is an immunometabolic regulator and mediator of sepsis, activated in response to LPS.[47] Altogether, the differential gene expression in pathobiont-high samples suggests that phagocytes mount an LPS-induced inflammatory response that is proportional to their burden of pathobionts.

E. coli Isolated From Lamina Propria Cells Are Viable and Contain Antimicrobial Resistance Genes

Next, we attempted to culture microbes from lamina propria cells to determine if these microbes were viable and to enable further study of CD vs control isolates. We collected samples from 21 CD and 20 control patients, isolated lamina propria cells, and treated these cells with gentamicin to remove extracellular bacteria. We then plated cell lysates on media containing bile salts to select for *E. coli* and other *Enterobacteriaceae* species, since these taxa were among the most abundant species observed in CD patients. Colonies grew from 5 CD samples (from 4 CD patients) and 6 control samples (from 3 control patients). We did not observe any distinguishing characteristics that differed between these 4 CD patients that grew colonies and the other CD patients. We picked colonies from each sample and performed whole genome sequencing of 11 isolates in total.

All isolates were determined to be *E. coli* based on core genome phylogeny. We generated an SNP phylogenetic tree of the isolates to visualize differences in predicted strain phylogeny (Figure 5A). We observed that isolates from the same patient clustered together, but CD and control isolates did not cluster by diagnosis. Among the isolates, some of the closest related strains were *E. coli* O157H7 strain TR01, a shiga toxin-producing strain,[48] and *E. coli* strain JJ2434, a uropathogenic strain.[49]

We next compared the GO functional profiles of the CD and control *E. coli* isolates (Figure 5B). Interestingly, CD isolates displayed more GO functions involved in cell motility, particularly flagellum-dependent motility, than control isolates. CD isolates also exhibited increased DNA transposition and chromate transport GO functions. In control isolates, enriched GO functions were nonspecific cellular processes (eg, positive regulation of biological process or regulation of cellular component organization). These differential GO
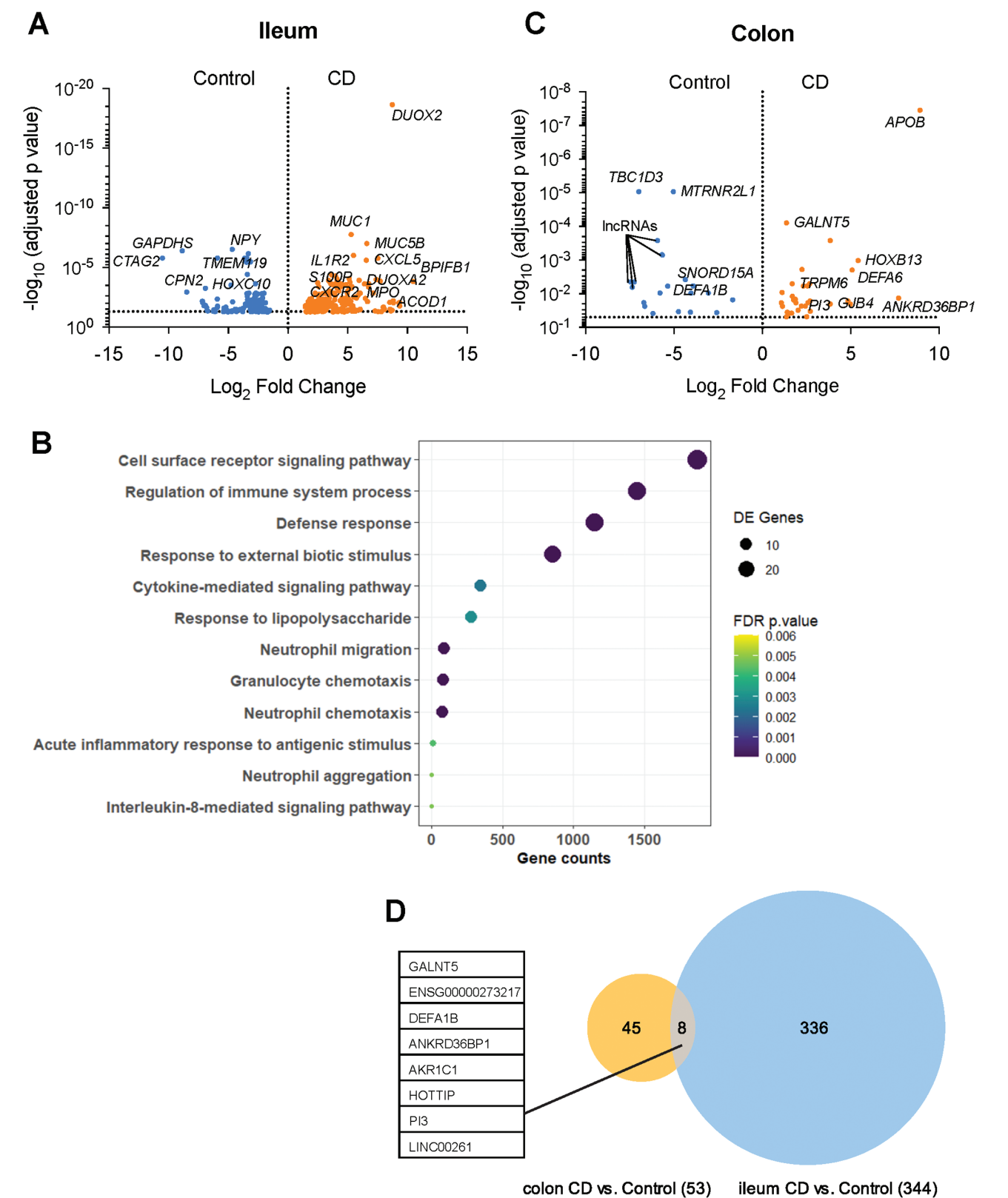
Figure 3. Differentially expressed genes and pathways in Crohn's disease (CD) vs control phagocytes. A, Volcano plot of all genes detected in ileal samples, plotted based on the log$_2$ fold change and $-\log_{10}$(P-value) in CD vs control samples (n = 10 CD vs n = 5 control). B, Ten Gene Ontology (GO) functions related to innate immunity selected from the 177 significant (FDR-adjusted P-value < .05) GO biological functions enriched in differentially expressed genes from CD vs control ileal samples. C, Volcano plot of all genes detected in colon samples, plotted based on the log$_2$ fold change and $-\log_{10}$(P-value) in CD vs control samples. D, Differentially expressed genes between CD patients and controls that overlapped between ileal and colon samples. The 8 shared genes are shown. FDR, false discovery rate.
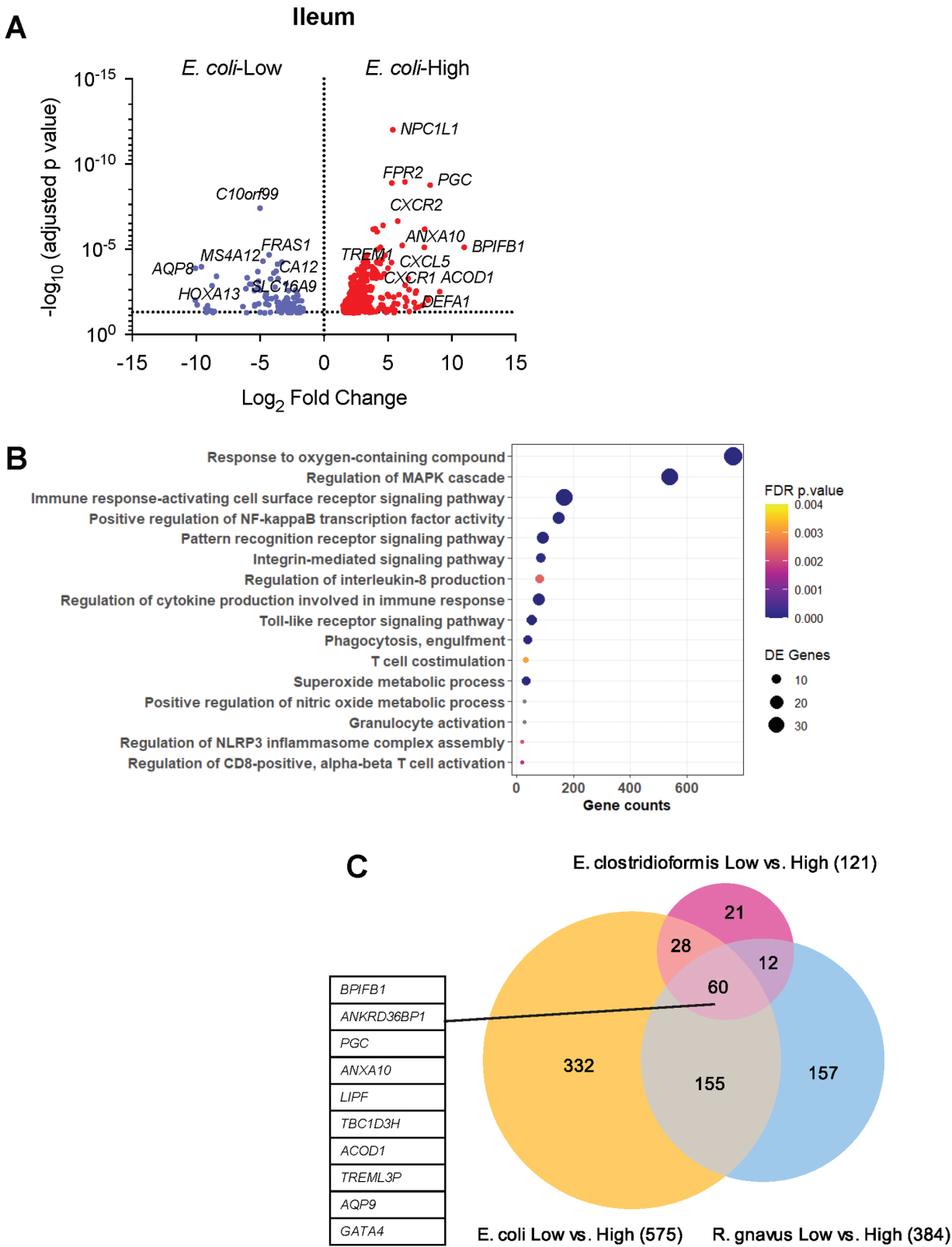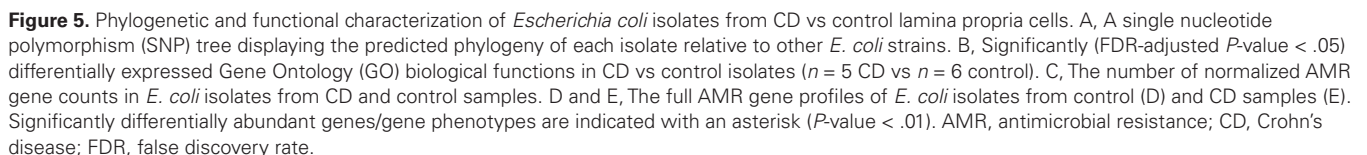
A

Ileum

B

C

Figure 4. Differentially expressed genes and pathways in phagocytes with a high vs low relative abundance of pathobionts. Ileal samples were dichotomized based on the abundance of *Escherichia coli*, *Enterocloster clostridioformis*, and *Ruminococcus gnavus* into high vs low groups for each species. A, Volcano plot of all genes detected in ileal samples, plotted based on the log$_2$ fold change and −log$_{10}$(*P*-value) in *E. coli*-high (*n* = 5) vs *E. coli*-low samples (*n* = 5). B, Ten Gene Ontology (GO) functions related to innate immunity selected from the 62 significant (FDR-adjusted *P*-value < .05) GO biological functions enriched in differentially expressed genes from *E. coli*-high vs *E. coli*-low ileal samples. C, Differentially expressed genes between high- and low-abundance ileal samples that overlapped across *E. clostridioformis*, *E. coli*, and *R. gnavus* comparisons. The top 10 shared genes are shown. FDR, false discovery rate.

Figure 5. Phylogenetic and functional characterization of *Escherichia coli* isolates from CD vs control lamina propria cells. A, A single nucleotide polymorphism (SNP) tree displaying the predicted phylogeny of each isolate relative to other *E. coli* strains. B, Significantly (FDR-adjusted *P*-value < .05) differentially expressed Gene Ontology (GO) biological functions in CD vs control isolates (*n* = 5 CD vs *n* = 6 control). C, The number of normalized AMR gene counts in *E. coli* isolates from CD and control samples. D and E, The full AMR gene profiles of *E. coli* isolates from control (D) and CD samples (E). Significantly differentially abundant genes/gene phenotypes are indicated with an asterisk (*P*-value < .01). AMR, antimicrobial resistance; CD, Crohn's disease; FDR, false discovery rate.

functions suggest that the CD isolates were more invasive and had more horizontal gene transfer than the control isolates.

Subsequently, we quantified AMR genes among the isolates. A total of 63 AMR genes were detected across all samples; 6 were unique to CD and 3 were unique to control samples (Figure 5C). CD samples had significantly higher expression of genes (*P*-value < .1) coding for antibiotic inactivation enzymes, particularly ANT(3″)-Ia, an aminoglycoside inactivation enzyme (Figure 5D and E). The CD isolates also displayed more genes encoding subunits of tet(B), an efflux pump that confers resistance to tetracycline. The 3 genes unique to control samples were found in low copy number in only 1-2 samples and therefore not significantly more abundant in control isolates than in CD isolates.

Finally, we examined the ability of *E. coli* strains isolated from inside the CD11b⁺ cells of CD patients to invade intestinal epithelial cells. Of the 3 strains that we assessed, 2 demonstrated the ability to invade colonic organoids derived from healthy people (Figure S8). Thus, *E. coli* strains can be isolated from the macrophages of CD patients, remain viable inside these cells, and retain the ability to invade epithelial cells—serving as further proof of the principle that the strains identified in CD patients are pathobionts.

Discussion

There is substantial evidence that gut microbiota are involved in IBD pathogenesis. However, the search for specific pathobionts has been hindered by limited methods for studying the relevant microbiota in IBD. In this study, we used an unbiased RNA sequencing approach to detect microbes within lamina propria phagocytes from CD patients and non-IBD controls. RNA sequencing, by definition, allowed us to identify bacteria that were actively transcribing genes and to characterize host gene responses. We hypothesized that specific species could reside within lamina propria phagocytes and act as pathobionts to trigger inflammation in CD. Indeed, we identified microbial RNA in CD11b⁺ cells (primarily macrophages, neutrophils, and dendritic cells[10]) from the gut lamina propria of both CD patients and non-IBD controls. We found a number of potential pathobionts that were present exclusively in CD phagocytes or more abundant in CD phagocytes compared to controls. A few of these species, such as *E. coli* and *R. gnavus*, have been previously associated with IBD—others have not, lending credence to our agnostic approach to discovering potential CD-related pathobionts.

Among the potential pathobionts that we detected in CD samples, we expected to find *E. coli*. There is substantial literature on the potential role of *E. coli*, particularly AIEC (found in 29% of CD patients[44]), in CD pathogenesis. In vitro experiments have demonstrated that AIEC can adhere to and invade intestinal epithelial cells, as well as invade and survive within macrophages.[8] Our data corroborate these in vitro findings in clinical samples, as our clinical isolates of *E. coli* from CD phagocytes were invasive in colonoid monolayers. We also detected increased *R. gnavus* in CD phagocytes; this species is a prevalent gut microbe that increases in abundance during CD flares.[16] Another potential pathobiont that was increased in CD phagocytes, *E. clostridioformis*, has been associated with CD in at least one study.[17] In addition, *Enterocloster* and *Halomonadaceae* spp., found in ileum and colon CD samples, respectively, contribute to dysbiosis and

have been detected in the colorectal cancer tumor microenvironment.[24,50] Overall, many of the potential pathobionts we found have been previously associated with either IBD or other gastrointestinal pathologies, which is reassuring considering issues with reproducibility in microbiome study methods and results.[51]

We focused on the bacteria that were differentially abundant across multiple CD samples, but we noted that CD samples were more heterogeneous than controls and that certain pathogenic bacteria were unique to specific CD patients. An elevated variability in the microbiome in CD patients compared to controls has been reported by previous studies.[52–54] For example, 2 CD ileal samples had a high abundance of *S. enterica* and another CD sample had a high abundance of *Clostridioides difficile*, neither of which were found in other CD or control samples. These are known pathogens with limited evidence of their role in CD but could be involved in the initial triggering of inflammation or exacerbation of symptoms in susceptible individuals.[29,55] We hypothesize that the variability among CD patients is due to differences in the pathobionts associated with CD that may be specific to individuals. Previous studies have reported that IBD patients display reduced microbial diversity but greater microbial variability than healthy controls[52–54]; however, the variability in phagocyte-associated microbiome in health and disease has not been characterized. These observations demonstrate the diversity between patients and highlight the need for personalized microbiome profiling to fully understand each patient's disease pathogenesis. Even if individuals possess unique driving pathogens, targeting these pathogens could potentially reduce the inflammatory response.

Specific species may act as pathobionts in CD; alternatively, certain metabolic functions or virulence factors (found in multiple species) may be the root cause of inflammation in CD. In the ileum, we observed increased LPS synthesis in CD-associated microbes in both the current cohort and our past cohort. Serotype-specific LPS may selectively trigger inflammation,[51] accounting for differences in commensal and pathogenic Gram-negative bacteria. Furthermore, pathogens have been noted to alter LPS expression to enhance virulence and evade immune detection.[56,57] AIEC LPS may play a role in resistance to complement and degradation by phagocytes.[58] Chronic LPS exposure may shift innate and adaptive immune cells in the lamina propria to a pro-inflammatory state in IBD.[56] Thus, upregulation of LPS production could contribute to resistance against innate immune responses and perpetuate chronic inflammation in CD. In the colon, it was less clear if any genes expressed by CD-associated microbes contributed to virulence. Given that our cohort had more active ileal disease, it was unsurprising that there were fewer clear trends in colon samples.

We next sought to demonstrate whether active pathobiont colonization occurs in the lamina propria. Our successful isolation of viable *E. coli* from this location in a subset of patients suggests that this species and other pathobionts may actively colonize the lamina propria, particularly in CD patients, as we observed a higher relative abundance of pathobiont sequences in this group. In these *E. coli* isolates, we found that flagellar motility and AMR genes were enriched in isolates from CD samples, factors that may support increased virulence. Flagella facilitate AIEC invasion[59]; indeed, we observed that 2 out of 3 isolates could invade epithelial cells. CD-associated AIEC

is also known to be more antibiotic-resistant than strains found in healthy controls.[58] In fact, one study[57] found that CD AIEC were more resistant to penicillin/beta lactam and aminoglycoside antibiotics, aligning with our findings. It is unclear whether antimicrobial-resistant microbes contribute to CD, but studies have shown that antibiotic exposure may increase the risk of developing IBD.[56]

Characterization of potential pathobionts reveals only part of the mechanisms involved in the perpetuation of CD pathogenesis. The host response to these pathobionts is equally important. We hypothesized that CD phagocytes would have a different inflammatory response to microbes than control samples. Indeed, we found that CD11b+ cells had a proinflammatory signature in the ileum of CD patients compared to controls. We believe this difference can be at least partially attributed to the microbiome, since pathobiont-high samples exhibited a more pronounced pro-inflammatory and antimicrobial gene signature. No such pro-inflammatory gene signature was observed in the colon, as expected given our cohort displayed less active colonic disease. Overall, the data suggest that CD phagocytes were mounting a response to an attempt to clear perceived pathogens in the ileum.

Researchers have long known that CD patients carry serum IgG antibodies to specific microbes in higher abundance than healthy controls.[60,61] Many of these microbes are those identified within phagocytes in our study, including *E. coli*, *R. gnavus*, *Lachnospiraceae*, *B. fragilis*, and *B. ovatus*.[28,62] In fact, antimicrobial antibodies are found in serum prior to CD development and can identify patients who will develop CD years before diagnosis.[63] It is plausible that the pathobionts we identified can cross the epithelial barrier and trigger an immune response that leads to antibody formation and predicts CD development. However, future studies that specifically focus on newly diagnosed, treatment-naïve patients would be necessary to further elucidate this hypothesis.

So how can we target pathobionts to treat IBD? Antibiotics and fecal microbiota transplants have had mixed results in clinical trials as IBD treatments.[59,64–66] In fact, antibiotics may even trigger disease,[67] likely because they are relatively nonspecific, killing beneficial microbiota in conjunction with pathogens and subsequently compromising the epithelial barrier.[64] Antibiotic-resistant microbes are also a major public health concern. We showed increased AMR genes, particularly in CD-associated microbes, evidence that further limits the potential for antibiotic effectiveness in CD. Bacteriophage therapy, on the other hand, could eventually be used to target specific pathobionts. Phages can control AIEC and *Klebsiella pneumoniae* and subsequently ameliorate colitis in mouse models.[68,69] A phase I clinical trial in healthy volunteers demonstrated that phage therapy targeting *K. pneumoniae* was safe, with no adverse side effects or ripple effects on microbiome composition.[69] This proof-of-concept study suggests a new avenue for targeting pathobionts: tailoring phage treatment to each patients' microbiome for personalized medicine.

Additionally, we discovered a few potential host targets in our study. For example, *PI3* was highly expressed in CD ileum and colon. This gene encodes peptidase inhibitor 3 (also known as elafin), an antimicrobial peptide that affects a broad spectrum of bacteria and fungi. However, it also downregulates LPS-induced innate immune signaling by inhibiting AP-1 and

NF-κB activation.[43] High circulating levels of elafin are associated with CD intestinal strictures.[70] *BPIFB1*, a secreted protein that binds LPS, was also very highly expressed in the pathobiont-high ileal samples. We believe these LPS-interacting genes should be further investigated as potential new targets for therapy.

Our study also has some limitations that should be noted. First, our controls (age range: 52-70) were older than our CD patients (age range: 23-68). As mentioned above, this difference is because our controls were drawn from individuals undergoing colorectal cancer screening (typically recommended for individuals>45 years), whereas CD can be diagnosed at any age. Age can influence the gut microbiome composition, although the most pronounced age-related effects are typically observed in individuals over 70 years old.[71–73] Thus, while age may influence the gut microbiome, we believe that this influence would be relatively small in our sample. Second, our sample had a high proportion of Hispanic CD patients, with some born outside the United States. However, given that previous studies have reported that Hispanic immigrants rapidly adopt an American diet[74] and that dietary acculturation shifts the gut microbial composition towards that of US-born Whites,[75] we do not believe that the ethnicity of patients substantially influenced the results. Third, our CD patients were on diverse therapies for their disease. It is unclear how the medications contributed to the variation in the gut microbiome (including alpha and beta diversity) between CD patients. Our small sample size precluded analyses of medication subgroups. Future studies should characterize the association between specific pathobionts and IBD phenotypes such as fibrostenotic disease vs internal perforating disease.

Conclusions

We investigated whether patients with established CD, especially ileal CD, have a phagocyte-associated microbiome that differs from healthy people and examined the relationship between the microbiota and the host immune response. The microbiota associated with phagocytes in patients with established CD had characteristics of pathobionts. Even if inflammation is controlled on medications, the cessation of medications generally results in disease relapse.[76] These putative pathobionts have virulence factors that may allow them to breach the epithelial barrier and survive within lamina propria cells, triggering chronic inflammation. For example, we found that *E. coli*, *R. gnavus*, and *Enterocloster* spp. were enriched in the lamina propria phagocytes of CD patients. However, no single species was found in all CD cases, suggesting that different pathobionts may contribute to CD depending on the individual. We also found that increased LPS production may be a virulence factor common to multiple CD-associated species. Indeed, CD phagocytes exhibited increased expression of proinflammatory genes, including proteins that bind LPS, most likely in an attempt to counteract these species. Flagellar motility genes and AMR genes may also contribute to virulence, as seen in the *E. coli* that we isolated from the lamina propria of CD patients. Our study provides avenues for further research to fully elucidate the relationship between pathobionts and CD and whether this relationship can be targeted to treat CD.

Supplementary Data

Supplementary data is available at *Inflammatory Bowel Diseases* online.

Acknowledgments

We thank the patients and healthy subjects for providing data for this study.

Funding

Conflicts of Interest

M.T.A. has received research funding from The Leona M. and Harry B. Helmsley Charitable Trust and the Crohn's and Colitis Foundation. She has served as a consultant for or is on the advisory board of the following companies: AbbVie Inc., Amgen, Bristol Myers Squibb, Celsius Therapeutics, Eli Lilly and Company, Gilead Sciences, Janssen Pharmaceuticals, Matera Prima, and Pfizer Pharmaceutical. M.T.A. has served as a teacher, lecturer, or speaker for the following companies: Janssen Pharmaceuticals and Takeda Pharmaceuticals. All other authors declare that they have no conflicts of interest.

Data Availability

The RNA sequencing data generated in this study have been deposited to the National Center for Biotechnology Information (NCBI)'s Gene Expression Omnibus (GEO) with accession number GSE267465. The bacterial whole genome sequencing data generated in this study will be deposited to NCBI's GenBank.

Ethics Approval and Consent to Participate

The protocol for this study was approved by the University of Miami Institutional Review Board (approval #: 20081100). All subjects provided written informed consent before participating in this study.

References

1. Qiu P, Ishimoto T, Fu L, Zhang J, Zhang Z, Liu Y. The gut microbiota in inflammatory bowel disease. *Front Cell Infect Microbiol.* 2022;12:733992. doi:10.3389/fcimb.2022.733992

2. Ni J, Wu GD, Albenberg L, Tomov VT. Gut microbiota and IBD: causation or correlation? *Nat Rev Gastroenterol Hepatol.* 2017;14(10):573-584. doi:10.1038/nrgastro.2017.88

3. Buisson A, Vazeille E, Fumery M, et al.; CEALIVE & REMIND study group. Faster and less invasive tools to identify patients with ileal colonization by adherent-invasive *E. coli* in Crohn's disease. *United European Gastroenterol J.* 2021;9(9):1007-1018. doi:10.1002/ueg2.12161

4. Ha CWY, Martin A, Sepich-Poore GD, et al. Translocation of viable gut microbiota to mesenteric adipose drives formation of creeping fat in humans. *Cell.* 2020;183(3):666-683.e17. doi:10.1016/j.cell.2020.09.009

5. Nowicki C, Ray L, Engen P, et al. Comparison of gut microbiome composition in colonic biopsies, endoscopically-collected and at-home-collected stool samples. *Front Microbiol.* 2023;14:1148097. doi:10.3389/fmicb.2023.1148097

6. Imhann F, Vich Vila A, Bonder MJ, et al. Interplay of host genetics and gut microbiota underlying the onset and clinical presentation of inflammatory bowel disease. *Gut.* 2018;67(1):108-119. doi:10.1136/gutjnl-2016-312135

7. Larabi A, Barnich N, Nguyen HTT. New insights into the interplay between autophagy, gut microbiota and inflammatory responses in IBD. *Autophagy* 2020;16(1):38-51. doi:10.1080/15548627.2019.1635384

8. Meconi S, Vercellone A, Levillain F, et al. Adherent-invasive *Escherichia coli* isolated from Crohn's disease patients induce granulomas in vitro. *Cell Microbiol.* 2007;9(5):1252-1261. doi:10.1111/j.1462-5822.2006.00868.x

9. Dheer R, Davies JM, Quintero MA, et al. Microbial signatures and innate immune gene expression in lamina propria phagocytes of inflammatory bowel disease patients. *Cell Mol Gastroenterol Hepatol.* 2020;9(3):387-402. doi:10.1016/j.jcmgh.2019.10.013

10. Jacobsen EG, Fernández I, Quintero AM, et al. Lamina propria phagocyte profiling reveals targetable signaling pathways in refractory inflammatory bowel disease. *Gastro Hep Adv.* 2022;1(3):380-392. doi:10.1016/j.gastha.2022.01.005

11. Couto N, Schuele L, Raangs CE, et al. Critical steps in clinical shotgun metagenomics for the concomitant detection and typing of microbial pathogens. *Sci Rep.* 2018;8(1):13767.doi:10.1038/s41598-018-31873-w

12. Uribe-Querol E, Rosales C. Phagocytosis: our current understanding of a universal biological process. *Front Immunol.* 2020;11:1066. doi:10.3389/fimmu.2020.01066

13. Shivashankar R, Tremaine WJ, Harmsen WS, Loftus EV. Incidence and prevalence of Crohn's disease and ulcerative colitis in Olmsted County, Minnesota From 1970 through 2010. *Clin Gastroenterol Hepatol.* 2017;15(6):857-863. doi:10.1016/j.cgh.2016.10.039

14. Vieira-Silva S, Sabino J, Valles-Colomer M, et al. Quantitative microbiome profiling disentangles inflammation- and bile duct obstruction-associated microbiota alterations across PSC/IBD diagnoses. *Nat Microbiol.* 2019;4(11):1826-1831. doi:10.1038/s41564-019-0483-9

15. Vaga S, Lee S, Ji B, et al. Compositional and functional differences of the mucosal microbiota along the intestine of healthy individuals. *Sci Rep.* 2020;10(1):14977. doi:10.1038/s41598-020-71939-2

16. Henke MT, Kenny DJ, Cassilly CD, Vlamakis H, Xavier RJ, Clardy J. Ruminococcus gnavus, a member of the human gut microbiome associated with Crohn's disease, produces an inflammatory polysaccharide. *Proc Natl Acad Sci U S A.* 2019;116(26):12672-12677. doi:10.1073/pnas.1904099116

17. Franzosa EA, Sirota-Madi A, Avila-Pacheco J, et al. Gut microbiome structure and metabolic activity in inflammatory bowel disease. *Nat Microbiol.* 2019;4(2):293-305. doi:10.1038/s41564-018-0306-4

18. Magdy Wasfy R, Mbaye B, Borentain P, et al. Ethanol-producing *Enterocloster bolteae* is enriched in chronic hepatitis B-associated gut dysbiosis: a case-control culturomics study. *Microorganisms* 2023;11(10):2437. doi:10.3390/microorganisms11102437

19. Schirmer M, Garner A, Vlamakis H, Xavier RJ. Microbial genes and pathways in inflammatory bowel disease. *Nat Rev Microbiol.* 2019;17(8):497-511. doi:10.1038/s41579-019-0213-6

20. Pascal V, Pozuelo M, Borruel N, et al. A microbial signature for Crohn's disease. *Gut.* 2017;66(5):813-822. doi:10.1136/gutjnl-2016-313235

21. Prondzinsky P, Berkemer SJ, Ward LM, McGlynn SE. The thermosynechococcus genus: wide environmental distribution, but a highly conserved genomic core. *Microbes Environ.* 2021;36(2):ME20138. doi:10.1264/jsme2.ME20138

22. Zhang X, Yao Q, Cai Z, Xie X, Zhu H. Isolation and identification of myxobacteria from saline-alkaline soils in Xinjiang, China. *PLoS One.* 2013;8(8):e70466. doi:10.1371/journal.pone.0070466

23. Stackebrandt E. 2014. The Family Dermabacteraceae. In: Rosenberg E, DeLong EF, Lory S, Stackebrandt E, Thompson F, eds. *The Prokaryotes.* Springer; 2014. doi:10.1007/978-3-642-30138-4_170

24. Wang Q, Ye J, Fang D, et al. Multi-omic profiling reveals associations between the gut mucosal microbiome, the metabolome, and host DNA methylation associated gene expression in patients with colorectal cancer. *BMC Microbiol.* 2020;20(Suppl 1):83. doi:10.1186/s12866-020-01762-2

25. Liu Z, Muller J, Li T, et al. Genomic analysis reveals key aspects of prokaryotic symbiosis in the phototrophic consortium "Chlorochromatium aggregatum". *Genome Biol.* 2013;14(11):R127. doi:10.1186/gb-2013-14-11-r127

26. Li S, Shi Y, Huang H, Tong Y, Wu S, Wang Y. Fermentation blues: analyzing the microbiota of traditional indigo vat dyeing in Hunan, China. *Microbiol Spectr* 2022;10(4):e0166322. doi:10.1128/spectrum.01663-22

27. Becker HEF, Jamin C, Bervoets L, et al. Higher prevalence of *Bacteroides fragilis* in Crohn's disease exacerbations and strain-dependent increase of epithelial resistance. *Front Microbiol.* 2021;12:598232. doi:10.3389/fmicb.2021.598232

28. Saitoh S, Noda S, Aiba Y, et al. *Bacteroides ovatus* as the predominant commensal intestinal microbe causing a systemic antibody response in inflammatory bowel disease. *Clin Diagn Lab Immunol.* 2002;9(1):54-59. doi:10.1128/cdli.9.1.54-59.2002

29. Schultz BM, Paduro CA, Salazar GA, et al. A potential role of *Salmonella* infection in the onset of inflammatory bowel diseases. *Front Immunol.* 2017;8:191. doi:10.3389/fimmu.2017.00191

30. Six DA, Carty SM, Guan Z, Raetz CRH. Purification and mutagenesis of LpxL, the lauroyltransferase of *Escherichia coli* lipid A biosynthesis. *Biochemistry.* 2008;47(33):8623-8637. doi:10.1021/bi800873n

31. Burgueno JF, Abreu MT. Epithelial Toll-like receptors and their role in gut homeostasis and disease. *Nat Rev Gastroenterol Hepatol.* 2020;17(5):263-278. doi:10.1038/s41575-019-0261-4

32. Gulati S, Shaughnessy J, Ram S, Rice PA. Targeting lipooligosaccharide (LOS) for a gonococcal vaccine. *Front Immunol.* 2019;10:321. doi:10.3389/fimmu.2019.00321

33. Roager HM, Licht TR. Microbial tryptophan catabolites in health and disease. *Nat Commun.* 2018;9(1):3294. doi:10.1038/s41467-018-05470-4

34. Burgueno JF, Fritsch J, Gonzalez EE, et al. Epithelial TLR4 signaling activates DUOX2 to induce microbiota-driven tumorigenesis. *Gastroenterology.* 2021;160(3):797-808.e6. doi:10.1053/j.gastro.2020.10.031

35. Leoncini G, Cari L, Ronchetti S, et al. Mucin expression profiles in ulcerative colitis: new insights on the histological mucosal healing. *Int J Mol Sci* . 2024;25(3):1858. doi:10.3390/ijms25031858

36. Li J, Simmons AJ, Hawkins CV, et al. Identification and multimodal characterization of a specialized epithelial cell type associated with Crohn's disease. *Nat Commun.* 2024;15:7204. doi:10.1038/s41467-024-51580-7

37. Jang SJ, Tuan WL, Hsu LA, et al. Pleiotropic effects of APOB variants on lipid profiles, metabolic syndrome, and the risk of diabetes mellitus. *Int J Mol Sci.* 2022;23(23):14963. doi:10.3390/ijms232314963

38. Barthold L, Heber S, Schmidt CQ, et al. Human alpha-defensin-6 neutralizes *Clostridioides difficile* toxins TcdA and TcdB by direct binding. *Int J Mol Sci.* 2022;23(9):4509. doi:10.3390/ijms23094509

39. Ganz T, Selsted EM, Lehrer IR. Defensins. *Eur J Haematol.* 1990;44(1):1-8. https://onlinelibrary.wiley.com/doi/10.1111/j.1600-0609.1990.tb00339.x

40. Wang F, Qiao L, Lv X, et al. Alarmin human alpha defensin HNP1 activates plasmacytoid dendritic cells by triggering NF-kappaB and IRF1 signaling pathways. *Cytokine.* 2016;83:53-60. doi:10.1016/j.cyto.2016.03.015

41. de Leeuw E, Li C, Zeng P, et al. Functional interaction of human neutrophil peptide-1 with the cell wall precursor lipid II. *FEBS Lett.* 2010;584(8):1543-1548. doi:10.1016/j.febslet.2010.03.004

42. Stolz A, Hammond L, Lou H, Takikawa H, Ronk M, Shively JE. cDNA cloning and expression of the human hepatic bile acid-binding protein. A member of the monomeric reductase gene family. *J Biol Chem.* 1993;268(14):10448-10457. doi:10.1016/S0021-9258(18)82220-7

43. Butler MW, Robertson I, Greene CM, O'Neill SJ, Taggart CC, McElvaney NG. Elafin prevents lipopolysaccharide-induced AP-1 and NF-kappaB activation via an effect on the ubiquitin-proteasome pathway. *J Biol Chem.* 2006;281(46):34730-34735. doi:10.1074/jbc.M604844200

44. Kamali Dolatabadi R, Feizi A, Halaji M, Fazeli H, Adibi P. The prevalence of adherent-invasive *Escherichia coli* and its association with inflammatory bowel diseases: a systematic review and meta-analysis. *Front Med (Lausanne).* 2021;8:730243. doi:10.3389/fmed.2021.730243

45. Ocon B, Pan J, Dinh TT, et al. A mucosal and cutaneous chemokine ligand for the lymphocyte chemoattractant receptor GPR15. *Front Immunol.* 2017;8:1111. doi:10.3389/fimmu.2017.01111

46. Midro AT, Stasiewicz-Jarocka B, Borys J, et al. Two unrelated families with variable expression of Fraser syndrome due to the same pathogenic variant in the FRAS1 gene. *Am J Med Genet A.* 2020;182(4):773-779. doi:10.1002/ajmg.a.61495

47. Wu R, Liu J, Wang N, et al. Aconitate decarboxylase 1 is a mediator of polymicrobial sepsis. *Sci Transl Med.* 2022;14(659):eabo2028. doi:10.1126/scitranslmed.abo2028

48. Gally DL, Stevens MP. Microbe profile: *Escherichia coli* O157:H7 - notorious relative of the microbiologist's workhorse. *Microbiology (Reading)* 2017;163(1):1-3. doi:10.1099/mic.0.000387

49. Min H, Baek K, Lee A, Seok Y-J, Choi Y. Genomic characterization of four *Escherichia coli* strains isolated from oral lichen planus biopsies. *J Oral Microbiol.* 2021;13(1):1905958. doi:10.1080/20002297.2021.1905958

50. Fidelle M, Rauber C, Alves Costa Silva C, et al. A microbiota-modulated checkpoint directs immunosuppressive intestinal T cells into cancers. *Science.* 2023;380(6649):eabo2296. doi:10.1126/science.abo2296

51. Stephens M, von der Weid PY. Lipopolysaccharides modulate intestinal epithelial permeability and inflammation in a species-specific manner. *Gut Microbes* 2020;11(3):421-432. doi:10.1080/19490976.2019.1629235

52. Clooney GA, Eckenberger J, Laserna-Mendieta E, et al. Ranking microbiome variance in inflammatory bowel disease: a large longitudinal intercontinental study. *Gut.* 2021;70(3):499-510. doi:10.1136/gutjnl-2020-321106

53. Halfvarson J, Brislawn JC, Lamendella R, et al. Dynamics of the human gut microbiome in inflammatory bowel disease. *Nat Microbiol.* 2017;2:17004. doi:10.1038/nmicrobiol.2017.4

54. Lloyd-Price J, Arze C, Ananthakrishnan NA, et al. Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases. *Nature.* 2019;569(7758):655-662. doi:10.1038/s41586-019-1237-9

55. Balram B, Battat R, Al-Khoury A, et al. Risk factors associated with clostridium difficile infection in inflammatory bowel disease: a systematic review and meta-analysis. *J Crohns Colitis.* 2019;13(1):27-38. doi:10.1093/ecco-jcc/jjy143

56. Candelli M, Franza L, Pignataro G, et al. Interaction between Lipopolysaccharide and gut microbiota in inflammatory bowel diseases. *Int J Mol Sci* . 2021;22(12):6242. doi:10.3390/ijms22126242

57. Abdelhalim KA, Uzel A, Gulsen Unal N. Virulence determinants and genetic diversity of adherent-invasive *Escherichia coli* (AIEC) strains isolated from patients with Crohn's disease. *Microb Pathog*. 2020;145:104233. doi:10.1016/j.micpath.2020.104233

58. Ohno M, Hasegawa M, Hayashi A, et al. Lipopolysaccharide O structure of adherent and invasive *Escherichia coli* regulates intestinal inflammation via complement C3. *PLoS Pathog*. 2020;16(10):e1008928. doi:10.1371/journal.ppat.1008928

59. Costello SP, Hughes PA, Waters O, et al. Effect of fecal microbiota transplantation on 8-week remission in patients with ulcerative colitis: a randomized clinical trial. *JAMA*. 2019;321(2):156-164. doi:10.1001/jama.2018.20046

60. Vermeire S, Rutgeerts P. Antibody responses in Crohn's disease. *Gastroenterology*. 2004;126(2):601-604. doi:10.1053/j.gastro.2003.12.020

61. Shome M, Song L, Williams S, et al. Serological profiling of Crohn's disease and ulcerative colitis patients reveals anti-microbial antibody signatures. *World J Gastroenterol*. 2022;28(30):4089-4101. doi:10.3748/wjg.v28.i30.4089

62. Bourgonje AR, Roo-Brand G, Lisotto P, et al. Patients with inflammatory bowel disease show IgG immune responses towards specific intestinal bacterial genera. *Front Immunol*. 2022;13:842911. doi:10.3389/fimmu.2022.842911

63. Torres J, Petralia F, Sato T, et al. Serum biomarkers identify patients who will develop inflammatory bowel diseases up to 5 years before diagnosis. *Gastroenterology*. 2020;159(1):96-104. doi:10.1053/j.gastro.2020.03.007

64. Ledder O, Turner D. Antibiotics in IBD: still a role in the biological era? *Inflamm Bowel Dis*. 2018;24(8):1676-1688. doi:10.1093/ibd/izy067

65. Torres J, Bonovas S, Doherty G, et al. ECCO guidelines on therapeutics in Crohn's disease: medical treatment. *J Crohns Colitis*. 2020;14(1):4-22. doi:10.1093/ecco-jcc/jjz180

66. Sokol H, Landman C, Seksik P, et al.; Saint-Antoine IBD Network. Fecal microbiota transplantation to maintain remission in Crohn's disease: a pilot randomized controlled study. *Microbiome* 2020;8(1):12. doi:10.1186/s40168-020-0792-5

67. Lee JY, Cevallos SA, Byndloss MX, et al. High-fat diet and antibiotics cooperatively impair mitochondrial bioenergetics to trigger dysbiosis that exacerbates pre-inflammatory bowel disease. *Cell Host Microbe*. 2020;28(2):273-284.e6. doi:10.1016/j.chom.2020.06.001

68. Galtier M, De Sordi L, Sivignon A, et al. Bacteriophages targeting adherent invasive *Escherichia coli* strains as a promising new treatment for Crohn's disease. *J Crohns Colitis*. 2017;11(7):840-847. doi:10.1093/ecco-jcc/jjw224

69. Federici S, Kredo-Russo S, Valdes-Mas R, et al. Targeted suppression of human IBD-associated gut microbiota commensals by phage consortia for treatment of intestinal inflammation. *Cell*. 2022;185(16):2879-2898.e24. doi:10.1016/j.cell.2022.07.003

70. Wang J, Ortiz C, Fontenot L, et al. High circulating elafin levels are associated with Crohn's disease-associated intestinal strictures. *PLoS One*. 2020;15(4):e0231796. doi:10.1371/journal.pone.0231796

71. Wilmanski T, Diener C, Rappaport N, et al. Gut microbiome pattern reflects healthy ageing and predicts survival in humans. *Nat Metab*. 2021;3(2):274-286. doi:10.1038/s42255-021-00348-0

72. Xu C, Zhu H, Qiu P. Aging progression of human gut microbiota. *BMC Microbiol*. 2019;19(1):236. doi:10.1186/s12866-019-1616-2

73. Cuesta-Zuluaga LDJ, Kelley TS, Chen Y, et al. Age- and sex-dependent patterns of gut microbial diversity in human adults. *mSystems* 2019;4(4):e00261-19. doi:10.1128/mSystems.00261-19

74. Damas MO, Estes D, Avalos D, et al. Hispanics coming to the US adopt us cultural behaviors and eat less healthy: implications for development of inflammatory bowel disease. *Dig Dis Sci*. 2018;63(11):3058-3066. doi:10.1007/s10620-018-5185-2

75. Peters AB, Yi SS, Beasley MJ, et al. US nativity and dietary acculturation impact the gut microbiome in a diverse US population. *ISME J*. 2020;14(7):1639-1650. doi:10.1038/s41396-020-0630-6

76. Reenaers C, Mary J-Y, Nachury M, et al.; Groupe d'Etude Therapeutique des Affections Inflammatoires du tube Digestif. Outcomes 7 years after infliximab withdrawal for patients with Crohn's disease in sustained remission. *Clin Gastroenterol Hepatol*. 2018;16(2):234-243.e2. doi:10.1016/j.cgh.2017.09.061