Mucosal host–microbe interactions associate with clinical phenotypes in inflammatory bowel disease

Shixian Hu[1,2,3*], Arno R. Bourgonje[1,*], Ranko Gacesa[1,2], Bernadien H. Jansen[1], Johannes R. Björk[1], Amber Bangma[1], Iwan J. Hidding[1], Hendrik M. van Dullemen[1], Marijn C. Visschedijk[1], Klaas Nico Faber[1], Gerard Dijkstra[1], Hermie J. M. Harmsen[4], Eleonora A. M. Festen[1], Arnau Vich Vila[1,2], Lieke M. Spekhorst[1,5,#], Rinse K. Weersma[1,#]

[1] Department of Gastroenterology and Hepatology, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands

[2] Department of Genetics, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands

[3] Institute of Precision Medicine, the First Affiliated Hospital of Sun Yat-Sen University, Sun Yat-Sen University, Guangzhou, Guangdong, China

[4] Department of Medical Microbiology, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands

[5] Department of Gastroenterology and Hepatology, Medisch Spectrum Twente, Enschede, the Netherlands

[*, #] *These authors contributed equally.*

Correspondence:

Prof. R.K. Weersma, M.D., Ph.D.

Department of Gastroenterology and Hepatology

University of Groningen, University Medical Center Groningen

P.O. Box 30.001, 9700 RB Groningen, the Netherlands

E-mail: r.k.weersma@umcg.nl

Tel: + 31 50 361 26 20; Fax: +31 50 361 93 06

Abstract

Dysregulation of gut mucosal host–microbe interactions is a central feature of inflammatory bowel disease (IBD). To study tissue-specific interactions, we performed transcriptomic (RNA-seq) and microbial (16S-rRNA-seq) profiling of 696 intestinal biopsies derived from 353 patients with IBD and controls. Analysis of transcript-bacteria interactions identified six distinct groups of inflammation-related pathways that were associated with intestinal microbiota, findings we could partially validate in an independent cohort. An increased abundance of *Bifidobacterium* was associated with higher expression of genes involved in fatty acid metabolism, while *Bacteroides* was associated with increased metallothionein signaling. In fibrostenotic Crohn's disease, a transcriptional network dominated by immunoregulatory genes associated with *Lachnoclostridium* bacteria in non-stenotic tissue. In patients using TNF-α-antagonists, a transcriptional network dominated by fatty acid metabolism genes associated with *Ruminococcaceae*. Mucosal microbiota composition was associated with enrichment of specific intestinal cell types. Overall, we identify multiple host–microbe interactions that may guide microbiota-directed precision medicine.

Keywords: inflammatory bowel disease, gene expression, mucosal microbiota, microbiome, host–microbe interactions.

Main

Inflammatory bowel diseases (IBD), which encompass Crohn's disease (CD) and ulcerative colitis (UC), are chronic inflammatory diseases of the gastrointestinal tract [1]. The pathogenesis of IBD is thought to be caused by a complex interplay between inherited and environmental factors, gut microbiota and the host immune system [2,3]. Alterations in gut microbiota composition and functionality are commonly observed in patients with IBD, including decreased microbial diversity, decreased abundances of butyrate-producing bacteria and increased proportions of pathobionts [4-8].

Interactions between host genetics and the gut microbiome have been studied in both healthy subjects and patients with IBD. For example, we previously focused on host genome–gut microbiota interactions in the context of IBD [9]. However, in order to disentangle disease mechanisms that might underlie the etiology and progression of IBD, there should be a greater focus on mucosal gene expression studies [10]. Modulation of host mucosal gene expression by gut microbiota or effects of gene expression on microbial fitness may expose mechanisms that contribute to IBD pathogenesis, knowledge that could be utilized to explore novel therapeutic targets [11,12]. Most studies, however, employ fecal sampling for microbiota characterization, which precludes analysis of local interactions and their immediate impact on host intestinal expression signatures. Such studies examining mucosal gene expression–microbiome associations in the context of IBD previously identified microbial groups associated with host transcripts from immune-mediated and inflammatory pathways [12-15]. In a longitudinal host–microbe interaction study, the chemokine genes *CXCL6* and *CCL20* were negatively associated with the relative abundances of *Eubacterium rectale* and *Streptococcus*, suggesting that these bacteria are more susceptible to the actions of these chemokines [13]. Another study found an inverse association between host expression of *DUOX2*, which produces reactive oxygen species (ROS), and the relative abundance of *Ruminococcaceae*, an association that may suggest ROS-mediated antibacterial effects [16]. However, few studies to date have been able to carry out comprehensive integrated analysis of IBD-associated interaction factors among mucosa-attached microbiota and host intestinal-gene expression.

Here we analyzed 696 fresh-frozen intestinal biopsies derived from 337 patients with IBD and 16 non-IBD controls for which we generated both mucosal transcriptomic and microbial characterization using bulk RNA-sequencing and 16S rRNA gene sequencing, respectively. We further combined both datasets to comprehensively investigate mutual mucosal host-microbe interactions and integrated these with the extensive clinical characteristics collected. Following this approach, we aimed to investigate mucosal host–microbe interactions while disentangling disease-, location- and inflammation-specific associations (a graphical representation of the study workflow is presented in **Figure 1**). Most importantly, we could study the associations between mucosal host–microbe interactions and clinical phenotypes of patients with IBD. Finally, we also sought to replicate our main results in data from a smaller independent, publicly available cohort [13].

Results

Cohort description

Demographic and clinical characteristics of the study population are presented in **Table 1**. In total, we included 640 intestinal biopsies from 337 patients with IBD and 56 intestinal biopsies from 16 non-IBD controls. Biopsies were derived from the colon (64.4%) and ileum (35.6%), and patients with CD and UC were equally represented among inflamed (CD: 53.8%, UC: 46.2%) and non-inflamed (CD: 55.4%, UC: 44.6%) biopsies. Mean age and the proportion of smokers were higher among controls (P<0.01 and P=0.01, respectively). Among biopsies derived from patients with IBD, the proportion of steroid users was higher among patients from whom inflamed biopsies were collected (P<0.01). Remaining patient characteristics were evenly distributed among groups without significant differences.

Table 1. Demographic and clinical characteristics of the study population compared between the inflamed and non-inflamed dataset.

Variable	Total	IBD		Non-IBD Controls	*P*-value
		Inflamed biopsies	Non-inflamed biopsies		
	n = 696	*n* = 212	*n* = 428	*n* = 56	
Biopsy inflammation, *n* (%)					
Inflamed	212 (30.5)	212 (100)	-	-	
Non-inflamed	428 (61.5)	-	428 (100)	-	
Biopsy location, *n* (%)					<0.01
Ileum	248 (35.6)	66 (31.1)	173 (40.4)	9 (16.1)	
Colon	448 (64.4)	146 (68.9)	255 (59.6)	47 (83.9)	
Diagnosis or control, *n* (%)					0.74

CD	351 (50.4)	114 (53.8)	237 (55.4)	-	
UC	289 (41.5)	98 (46.2)	191 (44.6)	-	
Controls	56 (8.0)	-	-	56 (100)	
Age at biopsy (years)	43.1 ± 15.3	43.0 ± 15.8	42.3 ± 15.5	45.5 ± 10.8	<0.01
Sex, *n* (%)					
Male	322 (46.3)	92 (43.4)	188 (43.9)	42 (75.0)	
Female	374 (53.7)	120 (56.6)	240 (56.1)	14 (25.0)	
BMI (kg/m^2)	25.7 ± 4.5	25.7 ± 4.6	25.7 ± 4.7	24.7 ± 2.5	
Current smoking, *n* (%)					0.01
Yes	145 (20.8)	37 (17.5)	88 (20.6)	20 (35.7)	
No	551 (79.2)	175 (82.5)	340 (79.4)	36 (64.3)	
Montreal classification					
Montreal Age (A), *n* (%)	638 (99.7)	212 (100)	426 (99.5)	-	0.95
A1 (≤16 years)	73 (11.4)	23 (10.8)	50 (11.7)	-	
A2 (17–40 years)	397 (62.0)	132 (62.3)	265 (62.2)	-	
A3 (>40 years)	168 (26.3)	57 (26.9)	111 (26.1)	-	
Montreal Location (L), *n* (%)	333 (94.9)	108 (94.7)	225 (94.9)	-	0.83
L1 (ileal disease)	66 (19.8)	18 (16.7)	48 (21.3)	-	
L2 (colonic disease)	51 (15.3)	17 (15.7)	34 (15.1)	-	
L3 (ileocolonic disease)	172 (51.7)	58 (53.7)	114 (50.7)	-	
L1 + L4	12 (3.6)	3 (2.8)	9 (4.0)	-	
L2 + L4	6 (1.8)	3 (2.8)	3 (1.3)	-	
L3 + L4	26 (7.8)	9 (8.3)	17 (7.6)	-	

Montreal Behavior (B), *n* (%)	333 (94.9)	108 (94.7)	225 (94.9)	-	0.31
B1 (non-stricturing, non-penetrating)	146 (43.8)	53 (49.1)	93 (41.3)	-	
B2 (stricturing)	59 (17.7)	20 (18.5)	39 (17.3)	-	
B3 (penetrating)	30 (9.0)	9 (8.3)	21 (9.3)	-	
B1 + P (perianal disease)	35 (10.5)	10 (9.3)	25 (11.1)	-	
B2 + P (perianal disease)	48 (14.4)	15 (13.9)	33 (14.7)	-	
B3 + P (perianal disease)	15 (4.5)	1 (0.9)	14 (6.2)	-	
Montreal Extension (E), *n* (%)	246 (85.1)	83 (84.7)	163 (85.3)	-	0.95
E1 (proctitis)	19 (7.7)	7 (8.4)	12 (7.4)	-	
E2 (left-sided colitis)	75 (30.5)	25 (30.1)	50 (30.7)	-	
E3 (pancolitis)	152 (61.8)	51 (61.4)	101 (62.0)	-	
Montreal Severity (S), *n* (%)	207 (71.6)	68 (69.4)	139 (72.8)		0.70
S0 (remission)	11 (5.3)	3 (4.4)	8 (5.8)		
S1 (mild)	28 (13.5)	9 (13.2)	19 (13.7)		
S2 (moderate)	109 (52.7)	33 (48.5)	76 (54.7)		
S3 (severe)	59 (28.5)	23 (33.8)	36 (25.9)		
Medication use					
Aminosalicylates, *n* (%)	271 (42.3)	91 (42.9)	180 (42.1)	-	0.87
Thiopurines, *n* (%)	210 (32.8)	66 (31.1)	144 (33.6)	-	0.53
Steroids, *n* (%)	262 (40.9)	105 (49.5)	157 (36.7)	-	<0.01
Methotrexate, *n* (%)	44 (6.9)	18 (8.5)	26 (6.1)	-	0.32

TNF-α-antagonists, *n* (%)[†]	113 (17.7)	35 (16.5)	78 (18.4)	-	0.58
Clinical disease activity					
HBI	324 (92.3)	104 (91.2)	220 (92.8)	-	0.18
Remission (<5)	205 (63.3)	60 (57.7)	145 (65.9)	-	
Active disease (≥5)	119 (36.7)	44 (42.3)	75 (34.1)	-	
SCCAI	257 (88.9)	84 (85.7)	173 (90.6)	-	0.14
Remission (≤2)	152 (59.1)	44 (52.4)	108 (62.4)	-	
Active disease (>2)	105 (40.9)	40 (47.6)	65 (37.6)	-	
Surgical history					
Ileocecal resection, *n* (%)	132 (20.6)	40 (18.9)	92 (21.5)	-	0.47
Colon resection (or partial), *n* (%)	146 (22.8)	55 (25.9)	91 (21.3)	-	0.19
Small intestinal (partial) resection, *n* (%)	82 (12.8)	29 (13.7)	53 (12.4)	-	0.71

Data are presented as proportions *n* with corresponding percentages (%), mean ± standard deviation (SD) or as median [interquartile range, IQR] in case of continuous variables. *P*-values ≤ 0.05 were considered statistically significant. [†]Use of TNF-α-antagonists included use of infliximab, adalimumab, golimumab and certolizumab pegol. Abbreviations: BMI, body-mass index; CD, Crohn's disease; HBI, Harvey-Bradshaw Index; IBD, inflammatory bowel disease; TNF-α, tumor necrosis factor alpha; SCCAI, Simple Clinical Colitis Activity Index; UC, ulcerative colitis.

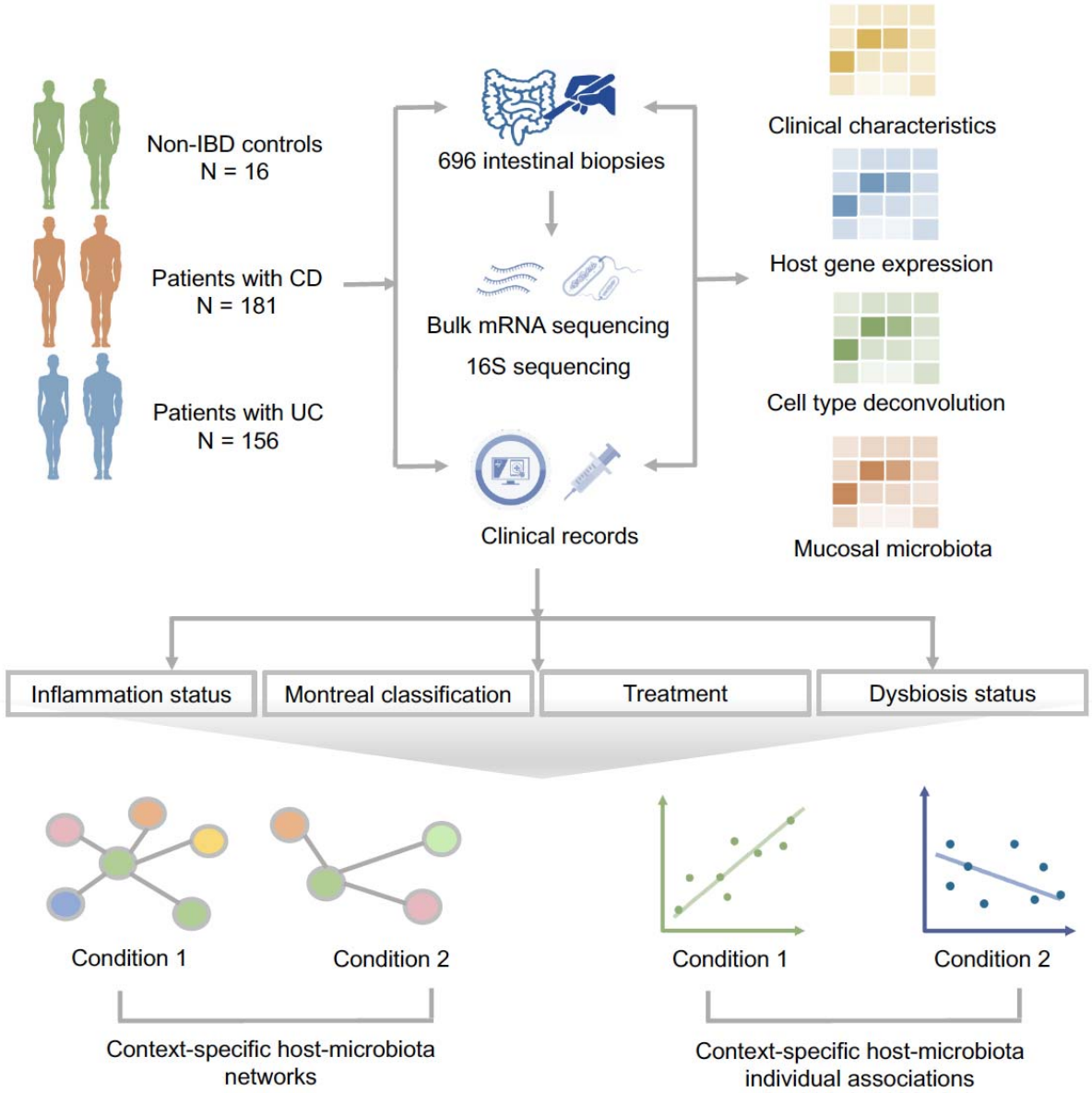


Figure 1. Methodological workflow of the study. The study cohort consisted of 337 patients with IBD (CD: *n*=181, UC: *n*=156) and 16 non-IBD controls, from whom 696 intestinal biopsies were collected (IBD: *n*=640, controls: *n*=56) and processed to perform bulk mucosal mRNA-sequencing and 16S gene rRNA sequencing. Detailed phenotypic data were extracted from clinical records for all study participants. In total, 251 ileal biopsies (CD: *n*=186, UC: *n*=56, controls: *n*=9) and 445 colonic biopsies (CD: *n*=165, UC: *n*=233, controls: *n*=47) were included: 212 biopsies derived from inflamed regions and 484 from non-inflamed regions. Mucosal gene expression and bacterial abundances were systematically analyzed in relation to different (clinical) phenotypes: presence of tissue inflammation, Montreal disease classification,

medication use (e.g. TNF-α-antagonists) and dysbiotic status. Pathway-based clustering and network analysis (Sparse-CCA and centrLCC analysis) and individual pairwise gene–taxa associations were investigated to identify host–microbiota interactions in different contexts. We then analyzed the degree to which mucosal microbiota could explain the variation in intestinal cell type–enrichment (estimated by deconvolution of bulk RNA-seq data). To confirm our main findings, we used publicly available mucosal 16S and RNA-seq datasets for external validation [13].

Mucosal gene expression reflects tissue specificity, inflammatory status and disease subtypes

Principal component analysis (PCA) showed that gene transcriptional patterns could be stratified by biopsy location (ileum vs. colon), inflammatory status (non-inflamed vs. inflamed) and IBD subtype (CD vs. UC) in the first two components (**Fig. 2A**), consistent with previous observations [13]. Tissue location and inflammatory status were significantly associated with the first two PCs (biopsy location, ileum vs. colon: $P_{Wilcoxon}=2.87x10^{-12}$; biopsy inflammatory status, $P=7.15x10^{-27}$), whereas disease/control status (CD vs. UC vs. controls) was associated with the second PC ($P=2.14x10^{-16}$).

Inflammation-associated gene expression showed overlap between inflamed biopsies from ileal CD, colonic CD and UC (**Fig. 2B**). Differential expression analyses between non-IBD controls, non-inflamed and inflamed biopsies in all these three groups revealed 3157, 3486, and 6710 differentially expressed genes (DEGs), respectively (FDR<0.05) (**Supplementary Table S1**). These DEGs fall mainly within interleukin signaling, neutrophil degranulation and extracellular matrix (ECM) organization pathways ($FDR_{Fisher}<0.05$, **Extended Data Fig. S1**). Overlapping results from all three differential expression analyses identified 1437 shared DEGs, including *DUOX2*, *MUC1*, *JAK2*, *OSM* and *IL17A* (**Fig. 2C**). We also observed an enrichment of these DEGs in IBD-associated genomic loci ($P_{Fisher}=9.6x10^{-9}$) [2].

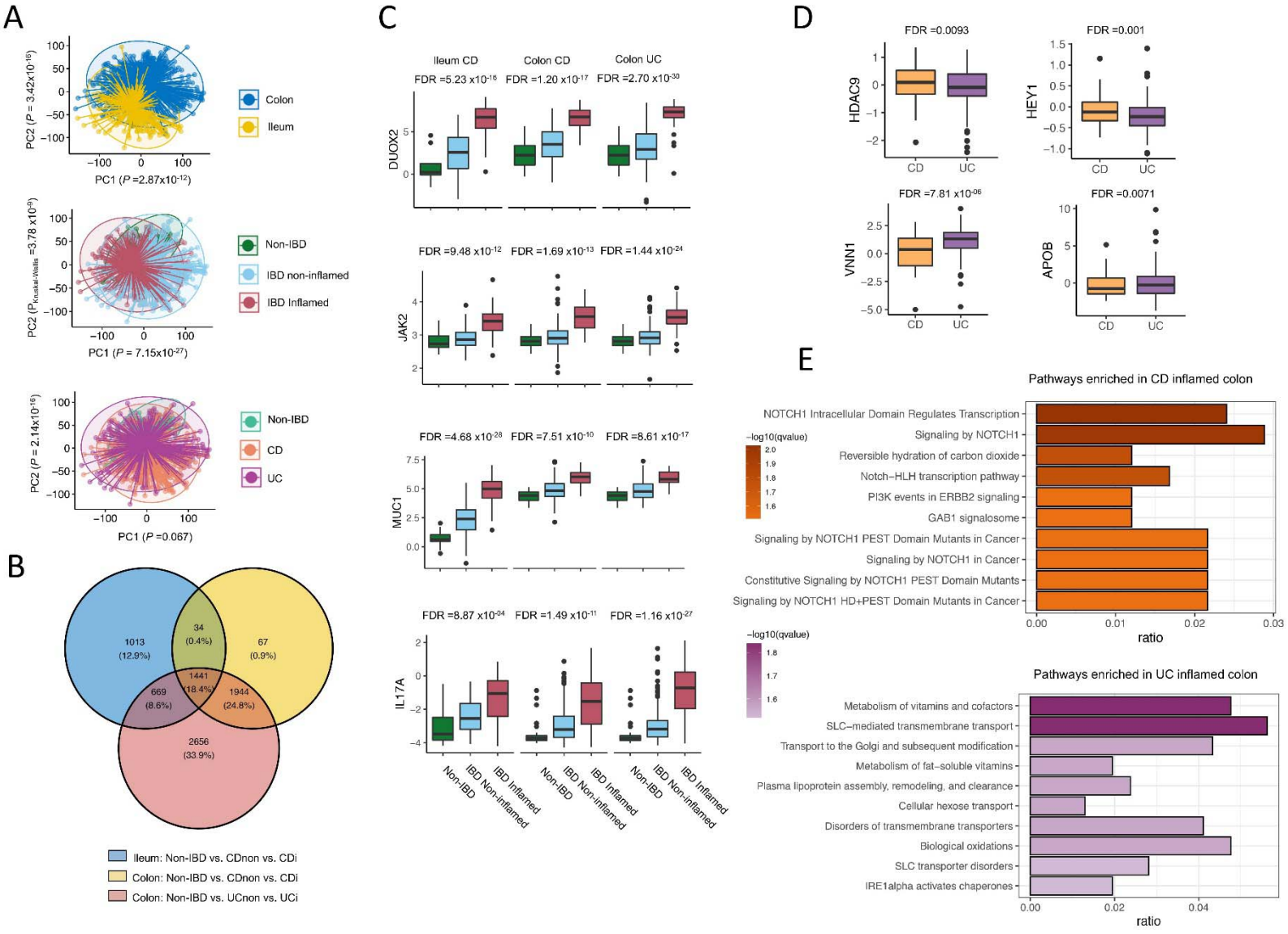


Figure 2. Mucosal host gene expression patterns in intestinal tissue from patients with IBD and controls. a, Principal component analysis, labeled by tissue location (ileum/colon), inflammatory status (non-inflamed/inflamed) and disease diagnosis (control/CD/UC), shows that variation in host gene expression can be significantly explained by tissue location and inflammatory status. **b**, Venn diagram of inflammation-associated genes from three comparisons: 1) ileal tissue from controls vs. non-inflamed tissue from patients with CD vs. inflamed tissue from patients with CD, 2) colonic tissue from controls vs. non-inflamed tissue from patients with CD vs. inflamed tissue from patients with CD and 3) colonic tissue from controls vs. non-inflamed tissue from patients with UC vs. inflamed tissue from patients with UC (all FDR <0.05). **c**, Relevant examples of four inflammation-associated genes, *DUOX2*, *JAK2*, *MUC1* and *IL17A*, illustrating the presence of tissue inflammation (FDR <0.05). **d**, Relevant examples of inflammation-associated genes differentially expressed between patients with CD and UC (keeping tissue location and inflammatory status constant) showing higher expression of *HDAC9* (histone deacetylase 9) and *HEY1* (hairy/enhancer-of-split related with YRPW motif protein 1) in patients with CD and higher expression of *VNN1* (pantetheinase) and *APOB* (apolipoprotein B) in patients with UC. **e**, Analysis of pathways associated with either the presence of CD (orange) or UC (purple) demonstrates that genes upregulated in CD are mainly associated with Notch-1 signaling, whereas pathways upregulated in UC are mainly related to vitamin and cofactor metabolism, SLC-mediated transmembrane transport and intracellular protein modification. Pathways were annotated using the Reactome pathway database. CDi, inflamed tissue from patients with Crohn's disease. CD-non, non-inflamed tissue from patients with Crohn's disease. FDR, false discovery rate. PC, principal component. UCi, inflamed tissue from patients with ulcerative colitis. UC-non, non-inflamed tissue from patients with ulcerative colitis.

We then investigated the genes differentially expressed between inflamed colonic tissue from patients with CD and UC. In total, 1466 genes were differentially abundant, of which 733 (50%) were overrepresented in CD and 733 (50%) in UC (FDR<0.05) (**Supplementary Table S2**). Pathway enrichment analysis showed the Notch-1 signaling pathway (e.g. *HDAC9* and *HEY1*, **Fig. 2D**) to be highly upregulated in CD compared to UC, whereas vitamin, cofactor and lipoprotein metabolism pathways (e.g. *VNN1* and *APOB*, **Fig. 2D**) were more pronounced in UC (**Fig. 2E**), which corroborates previous findings [17-20]. Cell type–deconvolution revealed that plasma cells, endothelial cells and Th2-lymphocytes were significantly increased in UC compared with CD (FDR<0.05, **Supplementary Table S3**), suggesting that distinct immunological mechanisms are involved in CD and UC.

Mucosal microbiota composition is highly personalized

The most common bacterial phylum observed across all tissue samples was Bacteroidetes (CD: 58%, UC: 58%, controls: 66%), followed by Firmicutes (CD: 27%, UC: 33%, controls: 23%) and Proteobacteria (CD: 14%, UC: 8%, controls: 9%). Interestingly, the overall mucosa-attached microbial composition was similar between colonic and ileal biopsies and independent of inflammation (**Extended Data Fig. S2**). Only seven bacterial taxa were differentially abundant between patients and controls (**Supplementary Tables S4-5**), consistent with previous findings [13,21,22].

Shannon diversity was significantly lower in samples from patients with CD compared to UC and non-IBD controls (P=2.75x10^{-16} and P=0.03, respectively, **Fig. 3A**). This difference was still present when comparing only colonic biopsies from patients with CD to those from UC, indicating that this difference was not solely attributable to ileal CD (**Extended Data Fig. S3**). Differences in microbial communities between tissue samples were evaluated by quantifying the Aitchison's distance (**Fig. 3B-C**). We obtained comparable findings when we externally validated our results using data derived from the HMP2 cohort (**Extended Data Fig. S4**) [13].

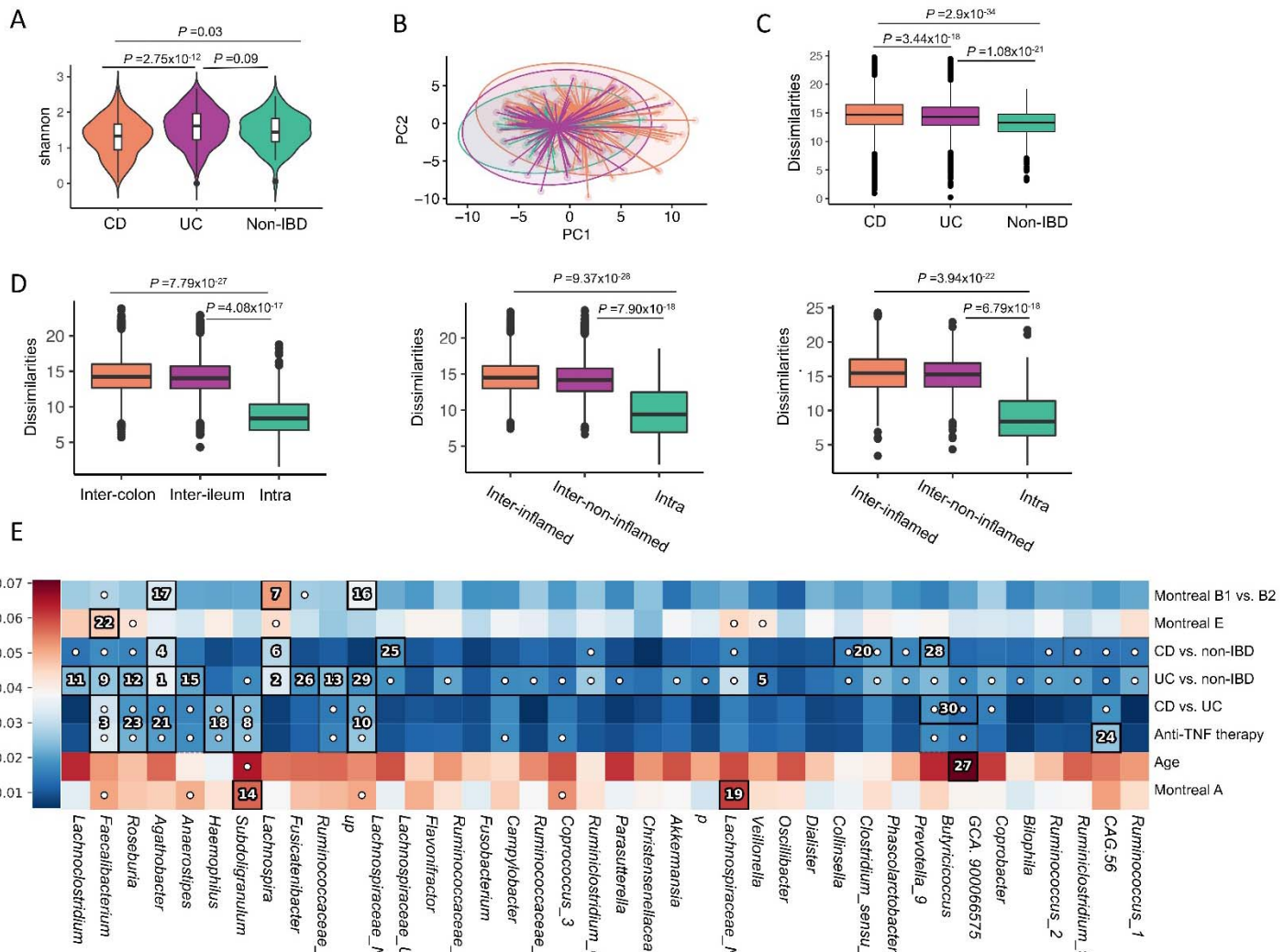


Figure 3. **Overall characterization of mucosa-attached microbiota in patients with IBD and controls. a**, Microbial alpha-diversity (Shannon index) was lowest in patients with CD (n=351) compared to patients with UC (n=289) and non-IBD controls (n=56). **b**, PCA plot based on Aitchison's distances demonstrates the microbial dissimilarity of the mucosa-attached microbiota (colors as in **a**). **c**, Microbial dissimilarity (Aitchison's distances) comparison between non-IBD control, CD and UC. Microbial dissimilarity is highest in biopsies from patients with CD, followed by patients with UC and non-IBD controls. **d**, Microbial dissimilarity is higher in samples from different individuals when compared to paired samples from the same individual, which includes paired inflamed–non-inflamed tissue from ileum and colon (left panel, inter-colon: *n*=11,430, inter-ileum: *n*=7,377, intra: *n*=203), paired colonic tissue samples from inflamed and non-inflamed areas (middle panel, inter-inflamed: *n*=7,372, inter-non-inflamed: *n*=8,369, intra: *n*=166) and paired ileal tissue samples from inflamed and non-inflamed areas (right panel, inter-inflamed: *n*=1,590, inter-non-inflamed: *n*=1,592, intra: *n*=73). **e**, Hierarchical analysis performed using an end-to-end statistical algorithm (HAllA) indicates the main phenotypic factors that correlate with intestinal mucosal microbiota composition. Heatmap color palette indicates normalized mutual information. Numbers or dots in cells identify significant pairs of features (phenotypic factors vs. bacterial taxa) in patients with IBD and controls. CD, Crohn's disease. PCA, principal coordinate analysis. UC, ulcerative colitis.

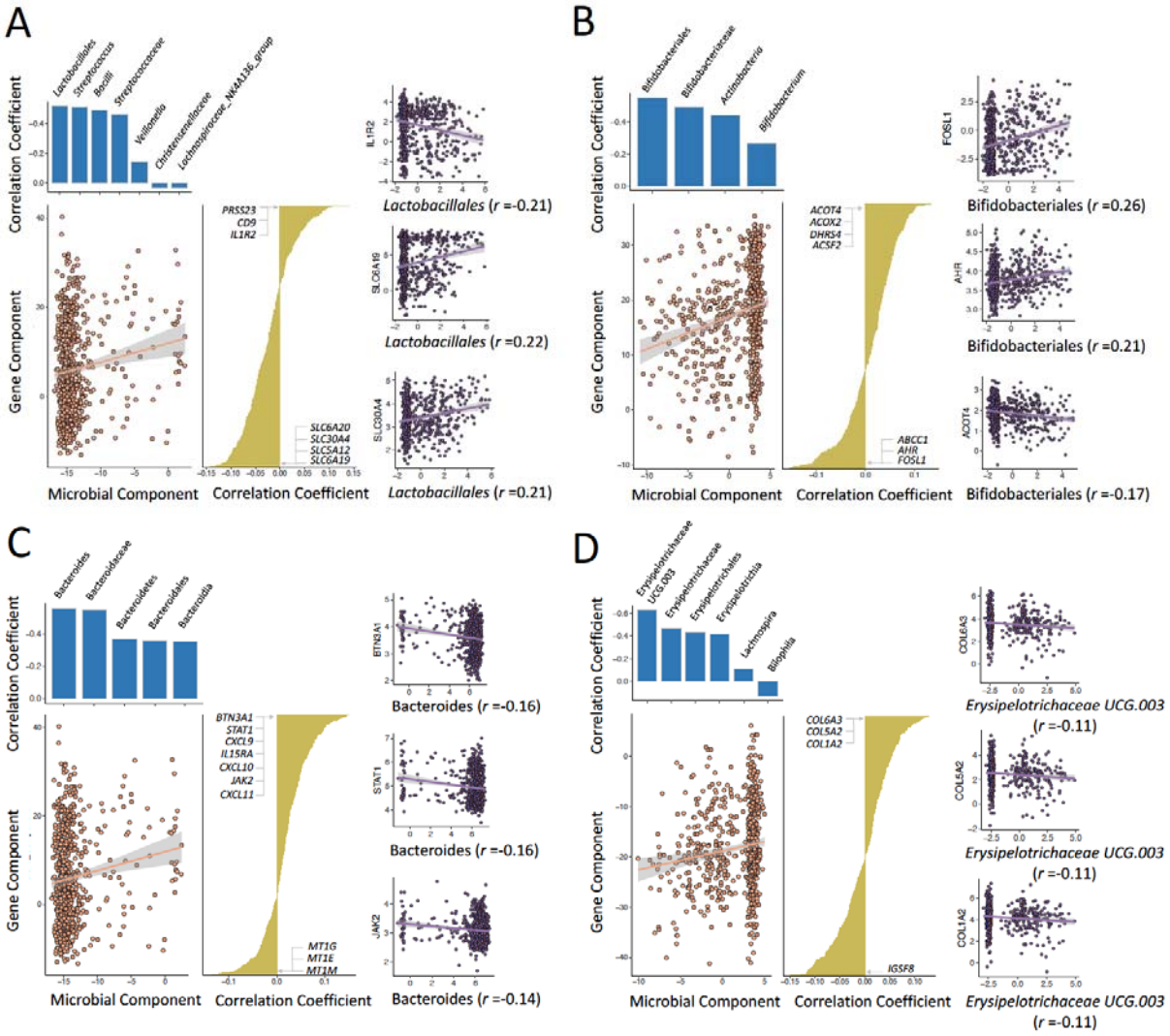
Intra-individual microbial dissimilarity was lowest in all our comparative analyses of paired tissue samples (**Fig. 3D**). Hierarchical clustering analysis performed on paired samples demonstrated a clear tendency of these samples to cluster together, a finding that we could also replicate in the HMP2 cohort data (**Extended Data Fig. S5A**) [13]. Overall, our data demonstrate that the composition of the mucosal microbiota is highly personalized and that inter-individual variability dominates over the effects of tissue location or inflammatory status.

We then aimed to identify phenotypic factors that shape the composition of the mucosal microbiota using Hierarchical All-against-All association (HAllA) analysis. This allowed us to study the relative associations between microbial taxa and phenotypic factors and disease characteristics (**Fig. 3E**, **Supplementary Table S6**). Analysis at bacterial genus level revealed that the main factors correlating with mucosal microbiota composition are stricturing disease in CD (fibrostenotic CD, Montreal B2), usage of TNF-α-antagonists, age at time of sampling, age of onset and the comparisons of patients with CD vs. controls, UC vs. controls and CD vs. UC. In contrast, inflammatory status and tissue location did not show a significant effect, and this was also the case within the HMP2 cohort data (**Extended Data Fig. S5B**). These findings are in line with several previous observations from which age at diagnosis, age at sampling and TNF-α-antagonist use emerged as critical determinants of mucosal microbiota composition [22].

Distinct host–microbe interaction modules are identified in relation to IBD

To capture the main microbial taxa associated with inflammation-associated gene expression, we combined the data and performed sparse canonical correlation analysis (sparse-CCA) on 1437 inflammation-associated genes and 131 microbial taxa. This approach enabled us to identify gene pathways and groups of microbiota and their potential correlations. In total, we found six distinct pairings of groups of genes with bacterial taxa to be significantly correlated with each other (FDR<0.05, **Supplementary Tables S7-S18**). To prioritize the individual genes and bacteria involved in the sparse-CCA analysis, we performed individual pairwise gene–bacteria associations, which revealed 312 significant gene–bacteria pairs, with most pairs (94.17%) overlapping with

231 the sparse-CCA results. We then replicated these associations in the HMP2 cohort
232 (Spearman correlation ρ=0.16, P=0.005, **Supplementary Table S19, Extended Data**
233 **Fig. S6, Methods**). Further details on the most intriguing individual pairwise gene–
234 bacteria associations are discussed in **Box 1**.



235

236 **Figure 4. Mucosal host–microbe interaction modules in the context of IBD.** Sparse canonical
237 correlation analysis (sparse-CCA) was performed to identify distinct correlation modules of mucosal gene
238 expression vs. mucosal microbiota through the identification of sparse linear combinations of two
239 separate distance matrices that are highly correlated. Using 1437 inflammation-related genes and 131
240 microbial taxa as input, we identified six distinct pairs of significantly correlated gene–microbe
241 components (FDR<0.05). **a**, A diverse group of mainly lactic acid producing bacteria (LAB) represented
242 by order Lactobacillales, genus *Streptococcus*, class Bacilli, family *Streptococcaceae* and, to a lesser

extent, genus *Veillonella*, family *Christensenallaceae* and the *Lachnospiraceae* NK4A136 group is associated with host pathways predominantly related to solute transport and liver metabolism. **b**, The abundance of mucosal *Bifidobacterium* bacteria is inversely associated with host fatty acid metabolism pathways (e.g. *ACOT4*, *ACOX2* and *ACSF2*) and positively associated with expression of specific genes, including *FOSL1*, *AHR* and *ABCC1/MRP1*. **c**, Mucosal *Bacteroides* bacteria inversely correlate with expression of genes representing host interleukin signaling pathways (e.g. *STAT3*, *JAK2*, *CXCL9* and *IL15RA*) but positively correlate with expression of genes representing metal ion response and metallothionein pathways (e.g. the metal ion response transcription factors *MT1G*, *MT1E* and *MT1M*). **d**, Mucosal *Erysipelotrichaceae* abundance inversely associates with expression of genes involved in collagen biosynthesis and collagen trimerization (e.g. *COL1A2*, *COL4A1* and *COL5A2*). Details of the two other significantly correlated pairs of components are presented in **Box 2**.

Mucosal lactic acid-producing bacteria positively correlate with nutrient uptake and solute transport

In the first significant pair of correlated components (component pair 1, P=5.72x10^{-14}, FDR<0.05), the bacterial component is represented by bacteria from order Lactobacillales, family *Streptococcaceae*, class Bacilli and genus *Streptococcus* and, to a lesser extent, genus *Veillonella*, family *Christensenallaceae* and the *Lachnospiraceae* NK4A136 group (**Supplementary Tables S7-S8**). This bacterial component is mainly represented by lactic acid producing bacteria (LABs, including Lactobacillales, Bacilli, *Streptococcaceae*, *Streptococcus*) that actively participate in physiological food digestion, particularly carbohydrate fermentation, with lactic acid being their main metabolic product. Many of these bacterial groups are associated with genes involved in pathways related to solute transport and liver metabolism, including SLC-mediated transmembrane transport of bile salts, organic acids, metal ions and amine compounds; amino acid transport; biological oxidation; cytochrome P450 enzymes and the ephrin signaling pathway (involved in the migration of intestinal epithelial cells along the crypt-villus axis).

LABs are widely present in commercially available probiotics, and their beneficial effects on intestinal epithelial health are well-recognized [23]. SLC transporters mediate the bidirectional passage of nutrients such as sugars, amino acids, vitamins, electrolytes and drugs across the intestinal epithelium [24]. Although SLC transporters are often found to be dysregulated in patients with IBD (particularly CD), their expression may be

stimulated and subsequently restored by commensal probiotic bacteria [25-27]. Taken together, however, we foresee that this host–microbe interaction component might not be IBD-specific as the genes and bacteria involved have important physiological functions in nutrient digestion and absorption.

Mucosa-residing Bifidobacterium species show significant interplay with host fatty acid metabolism and bile acid transport pathways

The second pair of significantly associated components (component pair 3, P=1.89x10^{-8}, FDR<0.05) is predominantly represented by bifidobacteria (**Supplementary Tables S9-S10**). The top associated pathways are represented by genes involved in fatty acid metabolism, including fatty acid biosynthesis (e.g. *ACOT4* and *ACSF2*), arachidonic acid metabolism (e.g. *CYP2J2* and *EPHX2*) and genes involved in peroxisomal protein import and fatty acid synthesis (e.g. *PEX5* and *ACOT4*), and these genes are all inversely associated with the bacterial component. In contrast, the genes *AHR* (encoding for the aryl hydrocarbon receptor) and *ABCC1* (encoding multidrug resistance protein 1) are positively correlated with the bacterial component. The inverse associations between bifidobacteria and the expression of genes involved in adipogenesis are consistent with findings from animal and small-scale human studies that investigated the effects of treatment with *Bifidobacterium* species on fatty acid metabolism [28-32]. Our findings may reflect the anti-inflammatory and anti-lipogenic role of bifidobacteria, which has previously been demonstrated in experimental settings, and may support the therapeutic potential of microbiome-directed interventions in attenuating or preventing colitis [31,32].

Mucosal Bacteroides associate with host interleukin signaling and metal ion response pathways

The third pair of significantly correlated components (component pair 7, P=1.28x10^{-4}, FDR<0.05) is represented by Bacteroidetes. Twenty-four different pathways were significantly associated with this microbial component (**Supplementary Tables S11-S12**). A number of interferon signaling pathways (e.g. IFN-α, IFN-β and IFN-γ as well as the IL-2, IL-4, IL-6, IL-10, IL-12 and IL-13 signaling pathways) are all inversely associated with the microbial component. In addition, metal ion response and

metallothionein pathways (e.g. metal ion transcription factors *MT1A*, *MT1E*, *MT1F*, *MT1G* and others) are positively associated with the microbial component. Taken together, these observations could suggest a predominance of potentially beneficial *Bacteroides* species associated with this component. Previous studies have shown that *Bacteroides* can exert either beneficial, mutualistic, or pathogenic effects on the host, depending on local interactions, intestinal location and nutrient availability [21,33]. The co-occurrence of *Bacteroides* with lower expression of interleukin signaling pathways is supported by experimental work that found potential anti-inflammatory and protective roles for these bacteria in the context of intestinal inflammation [34,39,40]. Still, the relative contributions of each of these species, as well as their behavior in the context of intestinal inflammation, remains elusive, although our data might reflect an overrepresentation of anti-inflammatory members [34-39,43]. The positive associations between *Bacteroides* and expression of metal ion response genes and metallothioneins (MTs) are intriguing in the context of IBD because aberrant MT homeostasis and intracellular zinc metabolism have been implicated in disease pathophysiology [44-48].

Mucosal Erysipelotrichaceae bacteria interact with collagen biosynthesis pathways

In the fourth pair of significantly correlated components (component pair 8, P=1.22x10^{-4}, FDR<0.05), the microbial component, represented by the family *Erysipelotrichaceae*, is inversely associated with the expression of genes belonging to a wide range of ECM and collagen genes that are involved in collagen biosynthesis, integrin cell surface interactions, collagen chain trimerization, collagen fibril cross-linking, collagen fibril assembly, ECM proteoglycans, collagen degradation and related pathways (**Supplementary Tables S13-S14**). Similar to *Bacteroides*, the precise role of *Erysipelotrichaeae* in the context of IBD has not yet been fully elucidated. Some studies found lower abundances of *Erysipelotrichaceae* in patients with new-onset CD [49] and postoperative active CD [50], whereas others reported higher levels of *Erysipelotrichaceae* in the context of ileitis [51] and TNF-regulated CD-like transmural inflammation [52]. These inconsistencies have been suggested to be due to *Erysipelotrichaceae* behaving differently in response to intestinal inflammation, but they may also reflect incomplete characterization of the precise species that belong to the

family of *Erysipelotrichaceae* [53]. Interactions between *Erysipelotrichaceae* and ECM/collagen remodeling pathways have not yet been reported in the context of IBD, but they would be particularly relevant because fibrosis occurs in a large fraction of patients with CD and *Erysipelotrichaceae* bacteria have been associated with fibrotic conditions beyond IBD [54-60].

Patients with fibrostenotic CD exhibit a *Lachnoclostridium*-associated gene network involved in immune regulation

In pairwise comparative analyses, patients with fibrostenotic CD (Montreal B2, *n*=107) and patients using TNF-α-antagonists (*n*=113) exhibited several differentially abundant microbial taxa. We therefore analyzed microbiota-associated host mucosal gene interactions in these phenotypes (**Figure 5**). Pairwise comparisons between patients with non-stricturing, non-penetrating disease vs. fibrostenotic CD revealed 2639 differentially abundant genes that were enriched in cellular energy metabolism and immune system pathways (FDR<0.05, **Supplementary Table S20**). When comparing microbial taxa, abundances of mucosal *Faecalibacterium*, *Erysipelotrichaceae_UCG-003* and *Coprococcus_3* were lower in fibrostenotic CD, whereas abundances of *Lachnoclostridium* and *Flavonifractor* were elevated in these patients (FDR<0.05). We hypothesized that these altered bacterial abundances and gene expression patterns may also translate into altered microbiota–gene networks relating to fibrostenotic CD. In patients with non-stricturing, non-penetrating CD, we observed 1508 individual gene–bacteria associations (corresponding to 84 different pathway–bacteria associations), whereas we found 541 individual associations (corresponding to 40 different pathway–bacteria associations) in patients with fibrostenotic CD. Comparing each bacteria-associated gene cluster between patients with non-stricturing, non-penetrating and fibrostenotic CD (FDR <0.05, **Methods, Supplementary Table S21**) identified four distinct networks represented by mucosal *Lachnoclostridium*, *Coprococcus*, *Erysipelotrichaceae* and *Flavonifractor*. The most significantly altered connections were associated with *Lachnoclostridium*, which was associated with 955 genes in patients with non-stricturing, non-penetrating CD, and these connections were mainly involved in

cell activation pathways such as vesicle-mediated cellular transport and membrane trafficking (**Fig. 5A**). In total, 148 genes were associated with *Lachnoclostridium* in patients with fibrostenotic CD (FDR<0.05), and these genes were involved in cellular immunoregulatory interactions and adaptive immune system pathways (e.g. *CD8A*, *CLEC2B* and *CXCR5*), tyrosine kinase signaling (e.g. *FGF16*), opioid signaling and G alpha (s) signaling events (mediated via cAMP-dependent protein kinases, e.g. *POMC*, *GNG7* and *GNG11*) and vesicle-mediated transport (e.g. *APOE*, *COLEC12* and *KIF3B*) (**Fig. 5A-B**).

Earlier studies had shown that *Lachnoclostridium* bacteria are generally increased in patients with (complicated) CD, e.g. postoperative CD [61], ASCA-positive CD [62] and active granulomatous colitis [63]. Recently, *Lachnoclostridium* was also associated with non-invasive diagnosis of colorectal adenoma and colorectal cancer [64,65]. These associations may potentially explain associations with genes involved in cellular proliferation and activation pathways. Increased abundances of *Lachnoclostridium* have been observed in relation to pulmonary fibrosis and its progression [66] but not in relation to intestinal fibrosis. In contrast, reduced abundances of *Faecalibacterium* and *Eubacterium* species (belonging to the *Erysipelotrichaceae* family) have previously been associated with luminal narrowing in patients with pediatric ileal CD [67]. Our results suggest that it is not only increased *Lachnoclostridium* abundances that may play a role in fibrostenotic CD, host immune-regulatory expression patterns may also vary along with these bacterial shifts. Notably, as the tissues investigated in our study were not derived from fibrotic regions, our findings show that these gene expression signatures are already present in non-stenotic intestinal tissue.

Use of TNF-α-antagonists is associated with *Ruminococcaceae*-associated gene interactions related to fatty acid metabolism

Subsequently, we investigated the impact of TNF-α-antagonist use on mucosal host–microbe interactions. Pairwise comparisons revealed that TNF-α-antagonist use was significantly associated with three different bacterial taxa, *Faecalibacterium*, *Ruminococcaceae_UCG-002*, and *Ruminococcaceae_UCG-005* (all showing reduced

abundances in users), and 513 different genes (FDR<0.05, **Supplementary Table S22**). For instance, one of the top genes associated with TNF-α-antagonist use was *CXCL13*, which encodes B cell attracting chemokine 1. By comparing each taxa-associated gene cluster between patients using and not using TNF-α-antagonists, we identified a single cluster represented by mucosal *Ruminococcaceae_UCG-002* that was significantly altered in users vs. non-users (FDR<0.05, **Supplementary Table S23**). *Ruminococcaceae_UCG-002* bacteria were associated with 135 genes in non-users, and these genes were mainly enriched in cell cycle–associated pathways (e.g. *PRIM1* and *PRIM2*), including mitosis-, prometaphase- and DNA-replication-associated genes (**Fig. 5C-D**). However, the *Ruminococcaceae_UCG-002*-associated genes in TNF-α-antagonist users (FDR<0.05) were predominantly involved in lipid/fatty acid metabolism (e.g. *ACAA1*, *ACSL5* and *PDK4*), glycerophospholipid biosynthesis and phospholipid metabolism. *Ruminococcaceae* comprise multiple distinct bacterial genera. Some of these are part of the healthy gut microbiome [68], but others are potentially pathogenic and commonly enriched in IBD [13,69]. The *Ruminococcaceae UCG_002* group is classified under the *Oscillospiraceae* family, which consists of obligate anaerobic bacteria including *Faecalibacterium prausnitzii*. Depending on their micro-environment, *Ruminococcaceae* UCG_002 bacteria can produce short-chain fatty acids due to their fiber-metabolizing capacity [70-72]. The inverse associations between *Ruminococcaceae*_UCG_002 and genes involved in (peroxisomal) fatty acid oxidation in patients using TNF-α-antagonists might reflect a beneficial therapeutic modulation, i.e. a reduction of fatty acid oxidation and lipotoxicity, and possibly even attenuation of microbiota-induced intestinal inflammation [73-85].

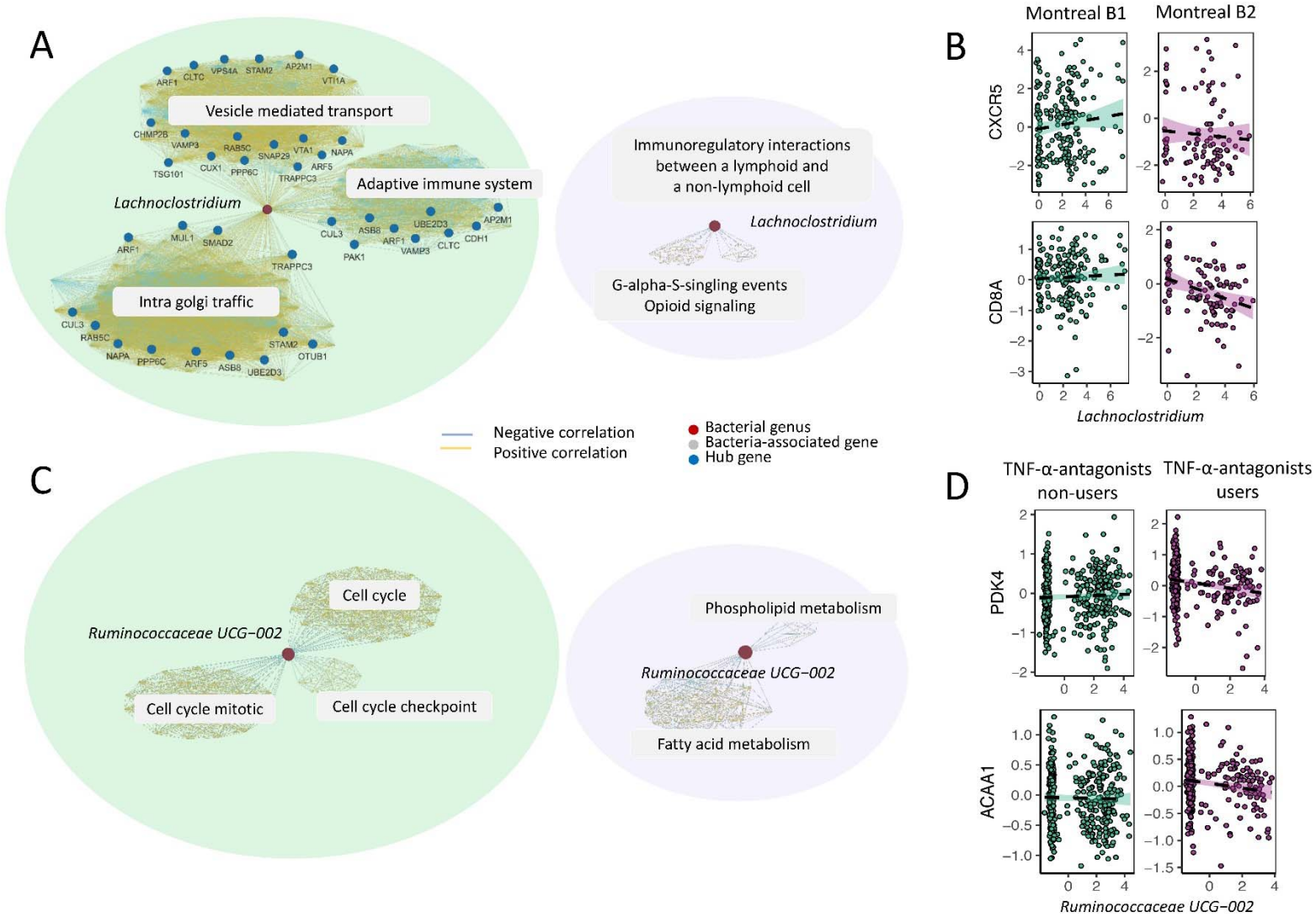


Figure 5. Fibrostenotic CD and TNF-α-antagonist usage significantly alter mucosal host–microbe interactions in the context of IBD.

CentrLCC-network analyses were performed to characterize altered mucosal host–microbe interactions between different patient phenotypes.

Overall, fibrostenotic CD (Montreal B2 vs. non-stricturing, non-penetrating CD, i.e. Montreal B1) and use of TNF-α-antagonists (vs. non-users) demonstrated significant modulation of observed mucosal host–microbe associations. **a**, Network graphs showing microbiota–gene association networks in patients with non-stricturing, non-penetrating CD (Montreal B1) (left) and patients with fibrostenotic CD (Montreal B2) (right). When comparing these patient groups, 5 bacterial taxa and 2639 host genes were significantly different (FDR<0.05). Four of the five bacterial taxa were significantly altered in fibrostenotic CD vs. non-stricturing, non-penetrating CD, and *Lachnoclostridium* was the top bacteria involved (covering 63% of total associations in non-stricturing, non-penetrating CD and decreasing to 27% in fibrostenotic CD). In general, patients with fibrostenotic CD were characterized by a loss of *Lachnoclostridium*–gene interactions. Red dots indicate gut microbiota. Blue dots indicate hub genes. Gray fields indicate the main pathways represented by the associated genes. Yellow lines indicate positive associations between gene expression and bacterial abundances. Light blue lines indicate negative associations. **b**, Key examples of *Lachnoclostridium*–gene interactions that were significantly altered in patients with fibrostenotic CD compared to patients with non-stricturing, non-penetrating CD, including genes involved in immunoregulatory interactions between lymphoid and non-lymphoid cells and tyrosine kinase signaling (*CD8A* and *CXCR5*). **c**, Network graphs showing microbiota–gene interaction networks in patients not using TNF-α-antagonists (left) vs. patients using TNF-α-antagonists (right). When comparing both groups, 3 bacterial groups and 513 genes were differentially abundant (FDR<0.05). Among these, a single bacterial group, represented by *Ruminococcaceae*_UCG_002, was altered in interactions with host genes in patients using TNF-α-antagonists. **d**, Key examples of *Ruminococcaceae*–UCG_002–gene interactions significantly altered in TNF-α-antagonists users vs. non-users. These genes were involved in general biological processes such as the cell cycle but also included genes involved in fatty acid metabolism (*PDK4* and *ACAA1*).

Mucosal host–microbe interactions depend on individual dysbiotic status

As patients with IBD have microbial dysbiosis compared to healthy individuals, we hypothesized that the strength and/or direction of the individual gene–bacteria interactions may depend on the microbial community (eubiosis vs dysbiosis). We therefore performed PCA on the microbiota data and calculated dysbiosis scores for all patients and controls, as represented by the median Aitchison's distances to non-IBD controls (**Fig. 6A**). Patients with IBD demonstrated higher dysbiosis scores compared to controls (CD vs. non-IBD: P=5.1x10^{-8}, UC vs. non-IBD: P=0.0015), but there was no clear difference between patients with CD and UC (**Fig. 6B**). When comparing patients with IBD above and below the 90th percentile of dysbiosis scores [13], 204 individual gene–bacteria interactions showed significant dependence on microbial dysbiosis (**Fig. 6C, Supplementary Table S24**) (FDR<0.05). We also performed permutation tests, which confirmed that the significant interactions were not observed by chance (**Methods**, FDR<0.05). In one example of these interactions, expression of the *PLAUR* gene encoding for the urokinase plasminogen activator surface receptor was positively associated with *Lachnospiraceae* abundance, but this shifted to an inverse association when only considering individuals with a high degree of mucosal dysbiosis (90–100%) (P=1.69x10^{-6}). The Ly6/PLAUR domain containing protein 8 (Lypd8) also functions as an antimicrobial peptide and has previously been shown to be capable of protecting the host from invading pathogenic flagellated bacteria [86]. Another example is the positive association between *S100A8*, which encodes S100 calcium-binding protein A8 (also known as calgranulin A), and *Lachnospiraceae*, which showed a negative association in individuals with high dysbiosis (P=1.78x10^{-5}). S100A8 has a wide variety of functions in regulating inflammatory processes and forms a heterodimer with S100A9, also known as calprotectin, which is used a as biomarker for inflammatory activity in IBD. Its known antimicrobial activity towards bacteria via chelation of zinc ions, which are essential for microbial growth, may therefore be disrupted in a dysbiotic environment [44]. Similar to the two previous examples, the observed association between the expression of *IL1RN* (encoding for the interleukin-1 receptor antagonist protein) and *Lachnospiraceae* shifted from positive to negative (P=4.10x10^{-5}), indicating that the natural protection against the proinflammatory effects of IL-1β, which associates with *Lachnospiraceae* abundance,

may be lost in circumstances of high microbial dysbiosis. Similarly, expression of the *CXCL17* gene encoding for a mucosal chemokine protein known to exert broad antimicrobial activity [87] positively correlated with *Lachnospiraceae* abundance, which was clearly different among individuals with higher dysbiosis scores.

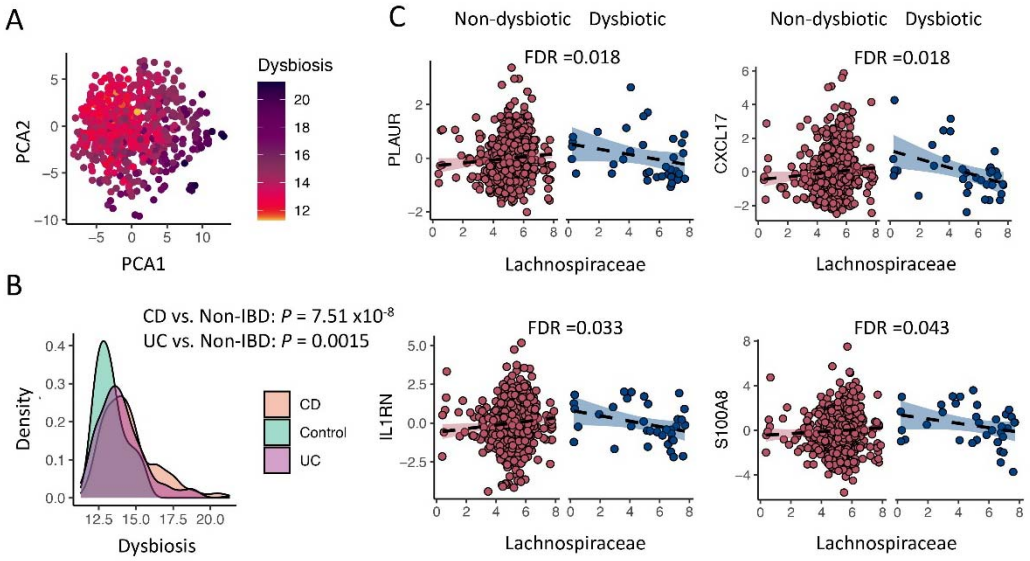


Figure 6. Mucosal host–microbe interactions depend on individual dysbiotic status. a, PCA of mucosal 16S rRNA sequencing data shows that degree of mucosal dysbiosis explains a large part of microbial variation. **b**, Dysbiosis scores were generally higher among patients with CD and UC compared to controls. **c**, Key examples of individual gene–bacteria interactions that demonstrate a directional shift upon higher dysbiosis (90–100%) as compared to patients with lower dysbiosis scores (0–90%). Mucosal *Lachnospiraceae* bacteria positively associate with the expression of the *PLAUR*, *CXCL17*, *IL1RN* and *S100A8* genes, whose gene products all have beneficial antimicrobial activity towards pathogenic bacteria. CD, Crohn's disease. PCA, principal component analysis. UC, ulcerative colitis.

Mucosal microbiota associate with variation in intestinal cell type–enrichment

Subsequently, we aimed to evaluate which intestinal cell types are involved in mucosal host–microbe interactions (**Figure 7, Extended Data Fig. S7**). Deconvolution of host gene expression data revealed that the mucosal microbiota was significantly associated with several cell types, but most evidently with intestinal epithelial cells, M1 macrophages, NK cells and mucosal eosinophils. Tissue inflammatory status and

location also strongly contributed to the variation in most intestinal cell types. Mucosal microbiota that were significantly associated with intestinal epithelial cell enrichment typically belonged to the Firmicutes phylum, including *Agathobacter*, *Dialister*, *Lachnospira*, *Lachnoclostridium* and *Ruminococcaceae* (**Supplementary Table S25**).

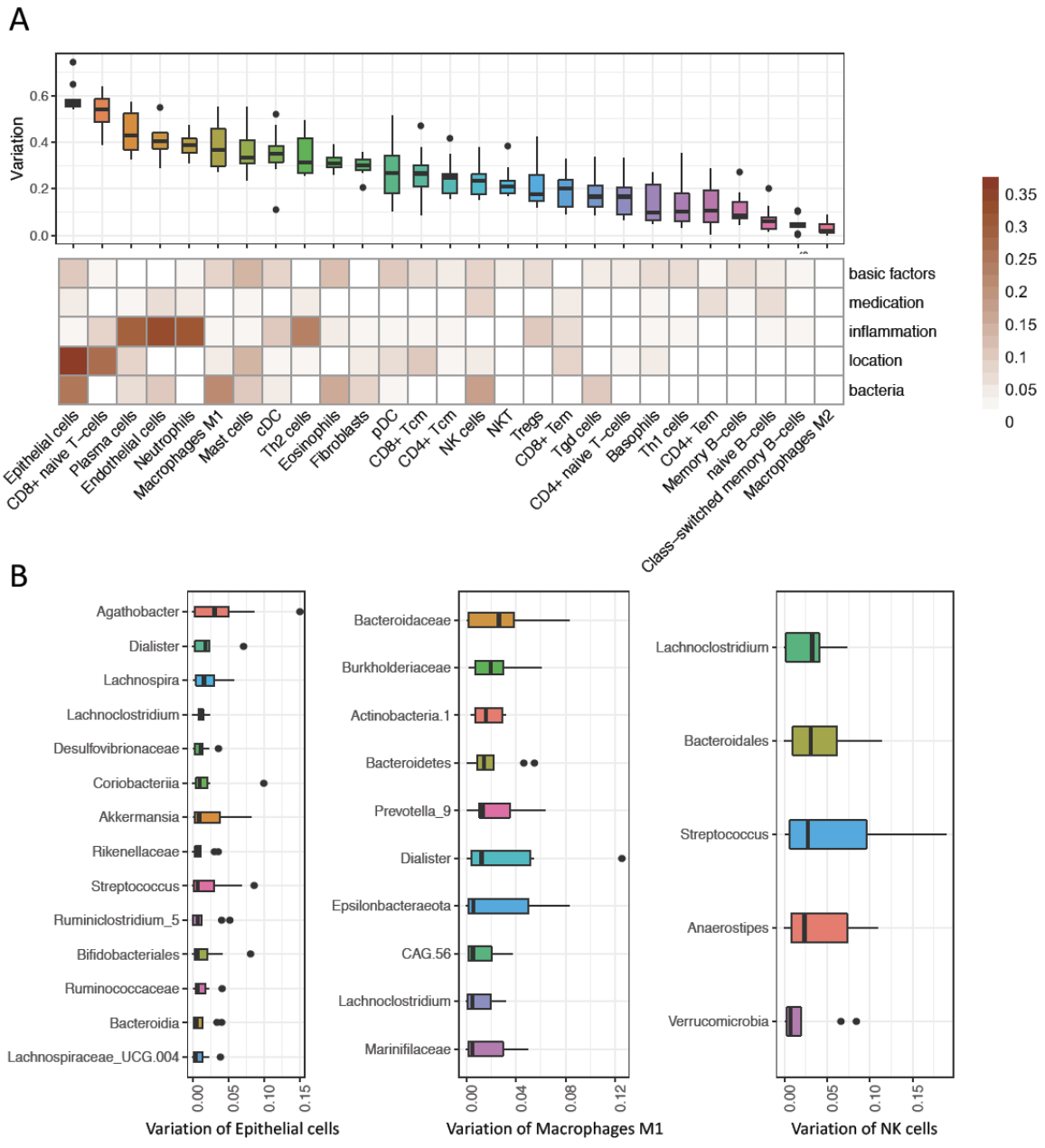


Figure 7. Mucosal microbiota associate with distinct intestinal mucosal cell types. **a**, Boxplots show the amount of variation in intestinal cell type–enrichment that could be explained by mucosal microbiota. Heatmap shows the contribution of other fitted models in explaining intestinal cell type–enrichment,

including ‘basic factors’ (age, sex and BMI), medication use, tissue inflammatory status and tissue location. Mucosal microbiota contributed most to the variation in enrichment of intestinal epithelial cells, M1-macrophages, NK cells and eosinophils. **b**, Boxplots showing the contribution of the main bacterial taxa that explain the variation in mucosal enrichment of intestinal epithelial cells, M1-macrophages and NK cells—the cell types that interacted most strongly with the mucosal microbiota.

Discussion

In this study, we show distinct mucosal host–microbe interactions in intestinal tissue from patients with IBD. Mucosal gene expression patterns in IBD are mainly determined by tissue location and inflammatory status and systematically demonstrate upregulation of distinct inflammation-associated genes, even in endoscopically non-inflamed tissue. Subsequently, we observed that the mucosal microbiota composition in patients is marked by high inter-individual variability. The main focus of our analyses, however, was integrative analysis of both data entities, which allowed us to comprehensively uncover many host–microbe associations, both on component level and as individual associations in IBD. Furthermore, we identify specific transcriptional networks that are significantly altered in patients with fibrostenotic CD and patients using TNF-α-antagonists and observe that these associations depend on the degree of mucosal dysbiosis. Finally, we show that mucosal microbiota are significantly associated with intestinal cell type composition, in particular with epithelial cells, macrophages and NK cells.

Tissue location and inflammatory status have the greatest impact on the variation in mucosal gene expression patterns. Enriched genes are mainly represented by those involved in pathophysiological pathways relevant to IBD, e.g. interleukin and interferon signaling and ECM remodeling. Patients with CD and UC show striking differences, e.g. Notch-1 signaling pathways are upregulated in CD, while genes involved in nutrient absorption and lipid metabolism are downregulated. Activation of Notch-1 signaling has been associated with improved mucosal barrier function, driven by lamina propria-residing $CD4^+$-T-lymphocytes that induce intestinal epithelial cell differentiation [17]. Notch-1 signaling more efficiently spreads within CD intestinal epithelia, as compared to UC or control epithelia. Notch-1 is not only implicated in IBD, it also confers protection against the development of colorectal carcinoma via p53 signaling, thereby promoting cell cycle arrests and cellular apoptosis [18,88,89]. Since UC patients with long-lasting colonic inflammation have a higher risk of developing IBD-associated colorectal carcinoma, we hypothesize that downregulation of Notch-1 in these patients may potentially be involved in carcinogenesis.

Analysis of mucosal microbiota in patients with IBD reveals reduced alpha-diversity, microbial dissimilarity and marked intra-individual variability that is particularly strong in CD but still present to a lesser extent in UC. Given the large heterogeneity in IBD and the fact that compositional differences are largely attributable to individual phenotypic factors, cautious interpretation is warranted when associating mucosal microbial profiles to disease phenotypes or outcomes, rendering them inappropriate for diagnostic purposes. These observations corroborate those of previously published mucosal 16S studies in IBD [13,21,22]. Moreover, our findings align with a recent prospective meta-analysis study that concluded there is sparse evidence for additional population structure in mucosal microbiomes in IBD, e.g. microbiota-driven discrete disease subtypes within IBD [90].

Sparse-CCA analysis was performed to capture the key pathway–bacteria interactions. These include numerous inverse associations between bifidobacteria and expression of genes involved in fatty acid metabolism, which align well with previously published data from animal studies demonstrating anti-inflammatory and anti-lipogenic effects of *Bifidobacterium* treatment on chemically-induced intestinal inflammation [29,31,32]. For example, treatment with *Bifidobacterium adolescentis* IM38 attenuated high fat diet–induced colitis in mice by inhibiting lipopolysaccharide production, NF-κB activation and TNF-expression in colonic epithelial cells [31]. Likewise, treatment with *Bifidobacterium infantis*, with or without a combination of inulin-type fructans, ameliorated DSS-induced colitis in rats, as evidenced by decreased expression of IL-1β, malondialdehyde (MDA, a lipid peroxidation marker), decreased bacterial translocation and increased production of short-chain fatty acids [32]. In line with our findings, this supports the ongoing quest for efficacious probiotic (bifidobacteria-containing) interventions in patients with IBD [91,92]. In addition, we observe a *Ruminococcaceae*-UCG-002-associated network of genes involved in (peroxisomal) fatty acid oxidation and lipotoxicity, which are inversely associated with these bacteria in patients using TNF-α-antagonists. Interestingly, multiple studies have observed that *Ruminococcaceae* increase after anti-TNF therapy in patients with CD and UC [73,75-77]. One of these studies specifically identified an association between the *Ruminococcaceae*_UCG-002 group and responsiveness to TNF-α-antagonists, albeit not in relation to host gene expression patterns [75].

Strikingly, many of the network-associated genes we observe are controlled by the PPAR-γ transcription factor, a butyrate sensor that may result in reduced lipotoxicity and reduced intestinal inflammation through prevention of overgrowth of potentially pathogenic bacteria [79-85]. These findings underscore the potential relevance of PPAR-γ as a therapeutic target in IBD [85].

We also observed an intriguing inverse relationship between *Erysipelotrichaceae* and intestinal ECM remodeling pathways, which may support the notion that intestinal fibrosis in IBD is highly linked to microbial composition [60,93,94]. Interestingly, a decreased relative abundance of *Erysipelotrichaceae* has previously been observed in patients with collagenous colitis [55] and cystic fibrosis–related lung fibrosis [56-58], as well as in mice with liver fibrosis and hepatocyte-specific *NOD2* deletions [59]. In CD, several bacterial species belonging to *Erysipelotrichaceae*, including *Clostridium innocuum* and *Erysipeloclostridium ramosum*, have been associated with the expansion of mesenteric adipose tissue (“creeping fat”), a unique feature of CD [60]. Creeping fat in CD has previously been characterized by higher abundances of *Erysipelotrichaceae* compared to adjacent mesenteric adipose tissue and underlying mucosal tissue and is accompanied by higher expression of ECM- and collagen-related genes. *C. innocuum* translocated to mesenteric fat, promoted fibrosis and stimulated tissue-remodeling in patients with CD, resulting in an adipose tissue barrier that may prevent systemic translocation of intestinal bacteria [60]. This phenomenon could potentially explain the inverse associations we observe between expression of ECM remodeling and mucosal *Erysipelotrichaceae*. In our differential network analyses, we observe a substantial decrease of *Lachnoclostridium*-associated genes in patients with fibrostenotic CD that are mainly associated with cellular immunoregulatory interactions and adaptive immune system pathways. These findings suggest that *Lachnoclostridium*-associated immunoregulatory expression patterns may play a role in fibrostenotic CD. Although little is known about the exact role of *Lachnoclostridium* in IBD, these bacteria were recently strongly associated with the development of colorectal cancer and with pulmonary fibrosis [64-66].

Another key host–microbe interaction module pertains to *Bacteroides*, which inversely correlates with interleukin signaling and positively associates with metal stress response transcription factors encoding for MTs. To maintain cellular redox balance, MTs detoxify heavy metal ions and scavenge ROS, thereby attenuating oxidative stress. Previous studies have shown that MTs may prevent experimental colitis or act as danger signals by mediating immune cell infiltration in the intestine [45,46]. Although experimental evidence seems to be inconclusive, there is ample evidence indicating a role for aberrant MT homeostasis in IBD [47]. This mechanism depends on the intracellular accumulation of zinc, which induces autophagy under chronic NOD2-stimulation. In IBD, the mucosal microbiota may contribute to the regulation of MT expression, intracellular zinc homeostasis and autophagy, thereby regulating intracellular bacterial clearance by intestinal macrophages. Findings from this study may support a putative role for *Bacteroides* in modulating MT activation, thereby contributing to intracellular redox homeostasis, zinc levels, macrophage autophagy, or even host defense against pathogens. Importantly, MTs and zinc regulation constitute potential therapeutic targets in IBD [44-47, 95-97].

Individual gene–bacteria association analysis revealed distinct mucosal host–microbe interactions that largely overlap with those from the sparse-CCA analysis, but these provide more granular insight into the observed associations. Key examples of individual host gene–bacteria interactions are listed in **Box 1**. Amongst others, we demonstrate several host–microbe interactions that are putatively involved in immunological tolerance and prevention of autoimmunity (e.g. bifidobacteria and *FOSL1*/*KLF2* expression), colorectal carcinogenesis (e.g. *Anaerostipes* and *SMAD4*, *Akkermansia* and *YDJC*) and inflammatory signaling (e.g. *Oscillibacter* and *OSM* expression). Notably, many of these associations are dependent on fibrostenotic disease, TNF-α-antagonist use and the degree of mucosal dysbiosis. In addition, deconvolution of the mucosal RNA-seq data reveal cell type–specific patterns of microbial interactions that warrant further study, for example through single-cell RNA-seq studies.

Mucosal host–microbiota interactions have been investigated previously in both cohort (e.g. the HMP2 and Irish IBD) and experimental studies [12-16]. Alongside several observations consistent with previous findings, we identify many novel host–microbe interactions. Differences in sample size, patient phenotypes and sample handling may be at least partially responsible for these observations. In our study, large groups of gene–bacteria associations are revealed that cover a wide range of molecular mechanisms potentially relevant in the context of IBD, including immune response pathways, cellular processes and a variety of metabolic pathways. Moreover, our study features the largest sample size so far [12-15], and this enabled us to perform an integrative analysis with respect to the large disease heterogeneity and identify novel host–microbiota crosstalk related to different clinical characteristics. However, several limitations also warrant recognition. As our study is of cross-sectional origin, we cannot assess the longitudinal dynamics of host–microbe interactions to discover signatures for therapy responsiveness or disease prognosis. Consequently, our associative results cannot establish potential causality between microbial abundances and host gene expression. Functional experiments are thus required to validate the biological relevance of the individual host–microbe interactions, as well as their behavior in microbial ecosystems. Finally, bowel preparation prior to the endoscopic procedure or cross-contamination between biopsy sites during endoscopy can affect the mucosal microbiota composition [21,50,98].

Our results demonstrate a complex and heterogeneous interplay between mucosal microbiota and mucosal gene expression patterns that is concomitant with the strong impact of specific patient traits in a large cohort of patients with IBD. Our findings may guide development of mechanistic studies (e.g. host–microbe co-culture systems) that could provide functional confirmation of relevant pathophysiological gene–bacteria interactions and serve as a resource for rational selection of therapeutic targets in IBD. This study presents a large-scale, comprehensive landscape of intestinal host–microbe interactions in IBD that could aid in guiding drug development and provide a rationale for microbiota-targeted therapy as a strategy to control disease course. Future studies are warranted to focus on the integration of host–microbe interaction modules in prospective clinical trials investigating their utility for predicting disease course and

651 responsiveness to treatment and for stratifying patients to facilitate therapeutic decision-
652 making.

Methods

Study population

Patients with an established diagnosis of IBD were included at the outpatient clinic of the University Medical Center Groningen (UMCG) based on their participation in the 1000IBD project, for which detailed phenotypic information and multi-omics profiles had been generated [99]. Patients included in this study were at least 18 years old and were enrolled from 2003–2019. Diagnosis of IBD was based upon clinical, laboratory, endoscopic and histopathological criteria, with the latter criteria also used to determine the inflammatory status of collected biopsies. Detailed phenotypic data were collected for all patients, including age, sex, BMI (body weight divided by squared height), smoking status, Montreal disease classification, medication usage, history of surgery, clinical disease activity and histological disease activity, and all were assessed at time of sampling. Montreal disease classification was recorded from the closest visit to the outpatient clinic at time of sampling. Clinical disease activity was established using the Harvey-Bradshaw Index (HBI) for patients with CD and the Simple Clinical Colitis Activity Index (SCCAI) for patients with UC. We further included 17 healthy non-IBD controls (*n*=59 biopsies) who underwent endoscopy because of clinical suspicion of intestinal disease or within the context of colon cancer screening. All participants provided written informed consent prior to sample collection. This study was approved by the Institutional Review Board (IRB) of the UMCG, Groningen, the Netherlands (in Dutch: 'Medisch Ethische Toetsingscommissie', METc; IRB nos. 2008/338 and 2016/424) and was conducted according to the principles of the Declaration of Helsinki (2013).

Mucosal RNA-sequencing

711 intestinal biopsies from 420 patients with IBD were collected. These were immediately snap-frozen in liquid nitrogen by an endoscopy nurse or research technician present during the endoscopic procedure. Biopsy inflammatory status was

assessed based on histological examination by certified pathologists. Biopsies were stored at -80℃ until further processing.

RNA isolation was performed using the AllPrep DNA/RNA mini kit (Qiagen, reference number: 80204) according to manufacturer's instructions. Homogenization of intestinal biopsies was performed in RLT lysis buffer including β-mercaptoethanol using the Qiagen Tissue Lyser with stainless steel beads (diameter 5 mm, reference number: 69989). For the first sample batch, sample preparation was executed using the BioScientific NEXTflex^{TM} Rapid Directional RNA-Seq Kit (Perkin-Elmer). Paired-end sequencing of RNA was performed using the Illumina NextSeq500 sequencer (Illumina). For the second sample batch, sample preparation was performed for construction of the Eukaryotic Transcriptome Library (Novogene). Paired-end sequencing of RNA was performed using the Illumina HiSeq PE250 platform. Sequencing was performed in two different batches, which necessitated pseudo-randomization (covering type of IBD diagnosis, biopsy location and disease activity) across plates to mitigate potential batch effects. The batch effects have been taken into account in all the analysis. On average, approximately 25 million reads were generated per sample.

Raw read quality was checked using FastQC with default parameters (ref v.0.11.7). Adaptors identified by FastQC were clipped using Cutadapt (ref v1.1) with default settings. Sickle (ref v1.200) was used to trim low-quality ends from the reads (length <25 nucleotides, quality <20). Reads were aligned to the human genome (human_glk_v37) using HISAT (ref v0.1.6) (with maximum allowance of two mismatches), and read sorting was performed using SAMtools (ref v0.1.19). SAMtools flagstat and Picard tools (ref v2.9.0) were used to obtain mapping statistics. Six samples with low percentage read alignment (< 90%) were removed. Gene expression was estimated using HTSeq (ref v0.9.1), based on Ensemble version 75 annotation, resulting in a RNA expression dataset featuring 15,934 genes. Expression data on gene level were normalized using a trimmed mean of *M* values, and *clr* transformation was applied, resulting in 826 mucosal RNA-seq samples.

Mucosal 16S rRNA gene sequencing

Total DNA extraction of intestinal biopsies using 0.25 g of sample was performed as described previously, with minor modifications [100]. Microbial composition of intestinal biopsies was determined by Illumina MiSeq paired-end sequencing of the V3-V4 hypervariable region of the 16S rRNA gene (MiSeq Benchtop Sequencer, Illumina Inc., San Diego, USA). Amplification of bacterial DNA was performed by PCR using modified 341F and 806R primers with a six-nucleotide barcode on the 806R primer for multiplexing [101,102]. Sequences of both primers can be found in **Supplementary Table S1**. Both primers contain an Illumina MiSeq adapter sequence, which is necessary for flow cell–binding in the MiSeq machine. A detailed overview of the PCR, DNA clean-up and MiSeq library preparation using a 2x300 cartridge can be found in the **Supplementary Methods**. Read trimming and filtering was done using *Trimmomatic* (0.33) to obtain an average read quality of 25 and a minimum length of 50. Quality was further checked using R package *DADA2* (v1.03) with the following parameters: truncLen=c(240,160), maxN=0, maxEE=c(2,2), truncQ=2 and rm.phix=TRUE. After error correction and chimera removal, the amplicon sequence variants were assigned to the *silva* database (v.132). Samples with >2,000 mapped reads were used for further analysis, resulting in 755 mucosal 16S samples. After accounting for overlap between mucosal RNA-seq and mucosal 16S data, 696 intestinal biopsies from 337 different patients and 16 non-IBD controls were available for host–microbiota interaction analyses.

Statistical analysis

Descriptive statistics

Descriptive data are presented as means ± standard deviation (SD), medians [interquartile range, IQR] or proportions *n* with corresponding percentages (%). Between-group comparisons were performed using Mann-Whitney *U*-tests, Pearson's chi-squared tests or Fisher's exact tests (if *n* observations were <10). Nominal *P*-values ≤ 0.05 were considered statistically significant.

Mucosal gene expression analysis

Sample gene expression dissimilarity was calculated using Aitchison's distances. General linear models were used to assess the associations between mucosal gene expression and clinical phenotypes while controlling for potential confounders, which were determined from our previous study (medication included the use of aminosalicylates, thiopurines and steroids) [103]. In particular, to assess the effect of mucosal inflammation on gene expression, we re-coded the inflammation status in an ordinal fashion as 0, 1 or 2 to represent biopsies from non-IBD controls, biopsies from non-inflamed tissue of patients with IBD and biopsies from inflamed areas of patients with IBD, respectively. Intestinal inflammatory status was thus treated as a continuous variable to account for presence of residual inflammation in biopsies marked as being taken from non-inflamed areas in the intestines. A correction for multiple hypotheses testing was applied using an FDR threshold of 5%.

1) Inflammation-associated genes were identified in three comparisons: (1) CD colonic inflamed tissue vs. CD colonic non-inflamed tissue vs. non-IBD colonic tissue, (2) CD ileocecal inflamed tissue vs. CD ileocecal non-inflamed tissue vs. non-IBD ileocecal tissue and (3) UC colonic inflamed tissue vs. UC colonic non-inflamed tissue vs. non-IBD colonic tissue:

Gene ~ intercept + inflammation + age + sex + BMI + medication + batch

2) Clinical phenotype–associated genes were identified using the following model:

Gene ~ intercept + Montreal/anti-TNF therapy + age + sex + BMI + inflammation + tissue location + medication + batch

Microbial characterization

Microbial richness and evenness was determined by calculating the Shannon index representing alpha-diversity of the gut microbiota. Microbial dissimilarity of samples was determined by calculating Aitchison's distances after *clr* transformation using the R package *Compositions* (v2.02). Analysis of paired samples from the same individuals was performed while comparing microbial features between inflammation status,

disease location and disease subtype using paired Wilcoxon tests. Factors potentially influencing mucosal microbiota were determined using Hierarchical All-against-All significance testing (HAllA) [104]. Associations between microbial features and biopsy inflammatory status, IBD diagnosis, disease location (biopsy origin) and clinical phenotypes were performed using general linear models (see below). Per sample, the mucosal dysbiosis score was defined as the median Aitchison distance from that sample to a reference sample set of non-IBD controls. Dysbiotic status was defined as being at the 90th percentile of this score [13].

1) Associations between microbial taxa and biopsy inflammation/location:

Taxa ~ intercept + inflammation + location + age + sex + BMI + medication + batch + surgical resection

2) Associations between microbial taxa and clinical phenotypes:

Taxa ~ intercept + Montreal/anti-TNF therapy + inflammation + location + age + sex + BMI + medication + batch + surgical resection

Gene–microbiota interaction analysis

We first focused on host inflammation-related genes (*n*=1,437) to investigate their potential associations with mucosal microbiota. Group-level correlations between gene expression and mucosal microbiota were performed using sparse-CCA using the residuals of genes and microbiota after correcting for age, gender, BMI, inflammation, tissue location and surgical resection separately. Sparse-CCA identifies the PCs from two related datasets that maximize the correlation between the two components. A set of enriched host pathways for all significant components was combined while adjusting for multiple comparisons using the FDR approach. Individual pairwise gene–microbiota associations were assessed by fitting a general linear model while adjusting for age, sex, BMI, inflammation status, tissue location, sequencing batch and medication use (including the use of aminosalicylates, thiopurines and steroids, see below). A gene–microbiota network analysis was visualized using the R package *ggview*.

1) Individual gene–bacteria associations were determined using the following model:

Gene ~ intercept + taxa + inflammation + location + age + sex + BMI + medication + batch

Second, we focused on host–microbiota interactions associated with fibrostenotic CD and usage of TNF-α-antagonists. Genes and taxa that were differentially abundant between clinical phenotypes were selected and then served as input for CentrLCC-network analysis using the *NetCoMi* R package. Hub nodes were defined as those with an eigenvector centrality value above the empirical 95% quantile of all eigenvector centralities in the network. This analysis was done in different groups separately (e.g. users and non-users of TNF-α-antagonists). To assess whether the taxa-associated gene networks were altered between groups, the associated genes for each taxa node were ranked within the total geneset background based on Z-scores. The Wilcoxon test was used to compare the two gene rank lists for each taxa.

Third, we assessed whether gene–microbiota associations depend on intestinal dysbiosis by modeling these associations using an additional interaction term in linear models. The dysbiosis score was treated as a continuous value. To determine whether these interactions were observed by chance, we also performed permutation tests that randomly shuffled the dysbiosis score 100 times across all samples, and then repeated the interaction models. On average, only three FDR-adjusted significant results were obtained for each round of permutation testing, suggesting that the rate of total false positives was approximately ~ 0.014 (3/204).

2) *Gene ~ intercept + taxa + dysbiosis + taxa * dysbiosis + inflammation + location + age + sex + BMI + medication + batch*

Fourth, enrichment of specific intestinal cell types was inferred from the RNA-seq data using the *Xcell* package in R. The effects of tissue location, inflammatory status and type of IBD diagnosis on expression levels of mucosal cell types were assessed using linear models, adjusting for age, sex, BMI, batch and medication usage. Subsequently, we used the *glmnet* R package to investigate the variation of cell type–enrichment that

could be explained by the mucosal microbiota using *lasso* regression while employing a nested 10-fold cross-validation using six models:

1) *Cell enrichment ~ age + gender + BMI + batch*
2) *Cell enrichment ~ medication (aminosalicylates, thiopurines, steroids, biologicals)*
3) *Cell enrichment ~ inflammation*
4) *Cell enrichment ~ tissue location*
5) *Cell enrichment ~ bacteria abundance*
6) *Cell enrichment ~ full factors mentioned above*

The percentage of explained variance (R^2) was calculated to estimate the variation in cell type–enrichment explained by the mucosal microbiota. All analyses were corrected for multiple testing using a FDR significance threshold of 0.05. All gene pathway enrichment analyses were conducted using the Reactome database from MsigDB [105,106].

Replication in the HMP2 dataset

RNA-seq and 16S raw data were obtained from https://ibdmdb.org and reprocessed using the same pipeline in this study. After harmonizing with the phenotype file, we included 152 intestinal biopsies from the 85 patients with CD, 46 patients with UC and 45 non-IBD controls. First, gene expression and mucosal microbiota patterns were compared separately between this study and HMP2. Second, given the limited overlap in clinical phenotypes between the two cohorts, we restricted the replication analysis to inflammation-related host–microbiota interactions. Individual gene–microbiota associations were calculated using the same linear models used in this study while adjusting for age, gender, tissue location and inflammation status. Spearman correlation coefficients were used to assess the concordance between the Z-scores of gene–microbiota associations from the two studies.

References

1. Chang JT. Pathophysiology of Inflammatory Bowel Diseases. *N Engl J Med* 2020;383(27):2652-2664.

2. de Lange KM, Moutsianas L, Lee JC, Lamb CA, Luo Y, Kennedy NA, et al. Genome-wide association study implicates immune activation of multiple integrin genes in inflammatory bowel disease. *Nat Genet* 2017;49(2):256-261.

3. de Souza HSP, Fiocchi C, Iliopoulos D. The IBD interactome: an integrated view of aetiology, pathogenesis and therapy. *Nat Rev Gastroenterol Hepatol* 2017;14(12):739-749.

4. Franzosa EA, Sirota-Madi A, Avila-Pacheco J, Fornelos N, Haiser HJ, Reinker S, et al. Gut microbiome structure and metabolic activity in inflammatory bowel disease. *Nat Microbiol* 2019;4(2):293-305.

5. Vich Vila A, Imhann F, Collij V, Jankipersadsing SA, Gurry T, Mujagic Z, et al. Gut microbiota composition and functional changes in inflammatory bowel disease and irritable bowel syndrome. *Sci Transl Med* 2018;10(472):eaap8914.

6. Frank DN, St Amand AL, Feldman RA, Boedeker EC, Harpaz N, Pace NR. Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases. *Proc Natl Acad Sci U S A* 2007;104(34):13780-5.

7. Kostic AD, Xavier RJ, Gevers D. The microbiome in inflammatory bowel disease: current status and the future ahead. *Gastroenterology* 2014;146(6):1489-99.

8. Kurilshikov A, Wijmenga C, Fu J, Zhernakova A. Host Genetics and Gut Microbiome: Challenges and Perspectives. *Trends Immunol* 2017;38(9):633-647.

9. Hu S, Vich Vila A, Gacesa R, Collij V, Stevens C, Fu JM, et al. Whole exome sequencing analyses reveal gene-microbiota interactions in the context of IBD. *Gut* 2021;70(2):285-296.

10. Cohen LJ, Cho JH, Gevers D, Chu H. Genetic Factors and the Intestinal Microbiome Guide Development of Microbe-Based Therapies for Inflammatory Bowel Diseases. *Gastroenterology* 2019;156(8):2174-2189.

882 11. Davison JM, Lickwar CR, Song L, Breton G, Crawford GE, Rawls JF. Microbiota
883 regulate intestinal epithelial gene expression by suppressing the transcription factor
884 Hepatocyte nuclear factor 4 alpha. *Genome Res* 2017;27(7):1195-1206.

885 12. Priya S, Burns MB, Ward T, Mars RAT, Adamowicz B, Lock EF, et al. Identification
886 of shared and disease-specific host gene-microbiome interactions across human
887 diseases using multi-omic integration. *Nat Microbiol* 2021; doi: 10.1038/s41564-022-
888 01121-z.

889 13. Lloyd-Price J, Arze C, Ananthakrishnan AN, Schirmer M, Avila-Pacheco J, Poon
890 TW, et al. Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases.
891 *Nature* 2019;569(7758):641-648.

892 14. Häsler R, Sheibani-Tezerji R, Sinha A, Barann M, Rehman A, Esser D, et al.
893 Uncoupling of mucosal gene regulation, mRNA splicing and adherent microbiota
894 signatures in inflammatory bowel disease. *Gut* 2017;66(12):2087-2097.

895 15. Morgan XC, Kabakchiev B, Waldron L, Tyler AD, Tickle TL, Milgrom R, et al.
896 Associations between host gene expression, the mucosal microbiome, and clinical
897 outcome in the pelvic pouch of patients with inflammatory bowel disease. *Genome Biol*
898 2015;16(1):67.

899 16. Lipinski S, Till A, Sina C, Arlt A, Grasberger H, Schreiber S, et al. DUOX2-derived
900 reactive oxygen species are effectors of NOD2-mediated antibacterial responses. *J Cell*
901 *Sci* 2009;122(Pt 19):3522-30.

902 17. Dahan S, Rabinowitz KM, Martin AP, Berin MC, Unkeless JC, Mayer L. Notch-1
903 signaling regulates intestinal epithelial barrier function, through interaction with CD4+ T
904 cells, in mice and humans. *Gastroenterology* 2011;140(2):550-9. doi:
905 10.1053/j.gastro.2010.10.057.

906 18. Garg P, Jeppsson S, Dalmasso G, Ghaleb AM, McConnell BB, Yang VW, et al.
907 Notch1 regulates the effects of matrix metalloproteinase-9 on colitis-associated cancer
908 in mice. *Gastroenterology* 2011;141(4):1381-92. doi: 10.1053/j.gastro.2011.06.056.

19. Weisshof R, Chermesh I. Micronutrient deficiencies in inflammatory bowel disease. *Curr Opin Clin Nutr Metab Care* 2015;18(6):576-81. doi: 10.1097/MCO.0000000000000226.

20. Walters JRF, Tasleem AM, Omer OS, Brydon WG, Dew T, le Roux CW. A new mechanism for bile acid diarrhea: defective feedback inhibition of bile acid biosynthesis. *Clin Gastroenterol Hepatol* 2009;7(11):1189-94. doi: 10.1016/j.cgh.2009.04.024.

21. Ryan FJ, Ahern AM, Fitzgerald RS, Laserna-Mendieta EJ, Power EM, Clooney AG, et al. Colonic microbiota is associated with inflammation and host epigenomic alterations in inflammatory bowel disease. *Nat Commun* 2020;11(1):1512. doi: 10.1038/s41467-020-15342-5.

22. Yılmaz B, Juillerat P, Øyås O, Ramon C, Bravo FD, Franc Y, et al. Microbial network disturbances in relapsing refractory Crohn's disease. *Nat Med* 2019;25(2):323-326. doi: 10.1038/s41591-018-0308-z.

23. Saez-Lara MJ, Gomez-Llorente, Plaza-Diaz J, Gil A. The role of probiotic lactic acid bacteria and bifidobacteria in the prevention and treatment of inflammatory bowel disease and other related diseases: a systematic review of randomized human clinical trials. *Biomed Res Int* 2015;2015:505878. doi: 10.1155/2015/505878.

24. Hediger MA, Clémençon B, Burrier RE, Bruford EA. The ABCs of membrane transporters in health and disease (SLC series): introduction. *Mol Aspects Med* 2013;34(2-3):95-107. doi: 10.1016/j.mam.2012.12.009.

25. Pérez-Torres S, Iglesias I, Llopis M, Lozano JJ, Antolín M, Guarner F, et al. Transportome Profiling Identifies Profound Alterations in Crohn's Disease Partially Restored by Commensal Bacteria. *J Crohns Colitis* 2016;10(7):850-9. doi: 10.1093/ecco-jcc/jjw042.

26. Kotka M, Lieden A, Pettersson S, Trinchieri V, Masci A, D'Amato M. Solute carriers (SLC) in inflammatory bowel disease: a potential target of probiotics? *J Clin Gastroenterol* 2008;42 Suppl 3 Pt 1:S133-5. doi: 10.1097/MCG.0b013e31815f5ab6.

27. Deleu S, Machiels K, Raes J, Verbeke K, Vermeire S. Short chain fatty acids and its producing organisms: An overlooked therapy for IBD? *EBioMedicine* 2021;66:103293. doi: 10.1016/j.ebiom.2021.103293.

28. Salazar N, Neyrinck AM, Bindels LB, Druart C, Ruas-Madiedo P, Cani PD, et al. Functional Effects of EPS-Producing *Bifidobacterium* Administration on Energy Metabolic Alterations of Diet-Induced Obese Mice. *Front Microbiol* 2019;10:1809. doi: 10.3389/fmicb.2019.01809.

29. Sun S, Luo L, Liang W, Yin Q, Guo J, Rush AM, et al. *Bifidobacterium* alters the gut microbiota and modulates the functional metabolism of T regulatory cells in the context of immune checkpoint blockade. *Proc Natl Acad Sci U S A* 2020;117(44):27509-27515. doi: 10.1073/pnas.1921223117.

30. Ray M, Hor PK, Ojha D, Soren JP, Singh SN, Mondal KC. Bifidobacteria and its rice fermented products on diet induced obese mice: analysis of physical status, serum profile and gene expressions. *Benef Microbes* 2018;9(3):441-452. doi: 10.3920/BM2017.0056.

31. Lim SM, Kim DH. Bifidobacterium adolescentis IM38 ameliorates high-fat diet-induced colitis in mice by inhibiting NF-κB activation and lipopolysaccharide production by gut microbiota. *Nutr Res* 2017;41:86-96. doi: 10.1016/j.nutres.2017.04.003.

32. Osman N, Adawi D, Molin G, Ahrne S, Berggren A, Jeppsson B. Bifidobacterium infantis strains with and without a combination of oligofructose and inulin (OFI) attenuate inflammation in DSS-induced colitis in mice. *BMC Gastroenterol* 2006;6:31. doi: 10.1186/1471-230X-6-31.

33. Bloom SM, Bijanki VN, Nava GM, Sun L, Malvin NP, Donermeyer DL, et al. Commensal Bacteroides species induce colitis in host-genotype-specific fashion in a mouse model of inflammatory bowel disease. *Cell Host Microbe* 2011;9(5):390-403. doi: 10.1016/j.chom.2011.04.009.

34. Hooper LV, Wong MH, Thelin A, Hansson L, Falk PG, Gordon JI. Molecular analysis of commensal host-microbial relationships in the intestine. Science 2001;291(5505):881-4. doi: 10.1126/science.291.5505.881.

965 35. Kelly D, Campbell JI, King TP, Grant G, Jansson EA, Coutts AGP, et al. Commensal
966 anaerobic gut bacteria attenuate inflammation by regulating nuclear-cytoplasmic
967 shuttling of PPAR-gamma and RelA. *Nat Immunol* 2004;5(1):104-12. doi:
968 10.1038/ni1018.

969 36. Wrzosek L, Miquel S, Noordine ML, Bouet S, Chevalier-Curt MJ, Robert V, et al.
970 Bacteroides thetaiotaomicron and Faecalibacterium prausnitzii influence the production
971 of mucus glycans and the development of goblet cells in the colonic epithelium of a
972 gnotobiotic model rodent. *BMC Biol* 2013;11:61. doi: 10.1186/1741-7007-11-61.

973 37. Delday M, Mulder I, Logan ET, Grant G. Bacteroides thetaiotaomicron Ameliorates
974 Colon Inflammation in Preclinical Models of Crohn's Disease. *Inflamm Bowel Dis*
975 2019;25(1):85-96. doi: 10.1093/ibd/izy281.

976 38. Uronis JM, Mühlbauer M, Herfarth HH, Rubinas TC, Jones GS, Jobin C. Modulation
977 of the intestinal microbiota alters colitis-associated colorectal cancer susceptibility.
978 *PLoS One* 2009;4(6):e6026. doi: 10.1371/journal.pone.0006026.

979 39. Hudcovic T, Kozáková H, Kolínská J, Štěpánková R, Hrnčíř T, Tlaskalová-
980 Hogenová H. Monocolonization with Bacteroides ovatus protects immunodeficient SCID
981 mice from mortality in chronic intestinal inflammation caused by long-lasting dextran
982 sodium sulfate treatment. *Physiol Rep* 2009;58(1):101-110. doi:
983 10.33549/physiolres.931340.

984 40. Chang YC, Ching YH, Chiu CC, Liu JY, Hung SW, Huang WC, et al. TLR2 and
985 interleukin-10 are involved in Bacteroides fragilis-mediated prevention of DSS-induced
986 colitis in gnotobiotic mice. *PLoS One* 2017;12(7):e0180025. doi:
987 10.1371/journal.pone.0180025.

988 41. Chu H, Khosravi A, Kusumawardhani IP, Kwon AH, Vasconcelos AC, Cunha LD, et
989 al. Gene-microbiota interactions contribute to the pathogenesis of inflammatory bowel
990 disease. *Science* 2016;352(6289):1116-20. doi: 10.1126/science.aad9948.

991 42. Mazmanian SK, Round JL, Kasper DL. A microbial symbiosis factor prevents
992 intestinal inflammatory disease. *Nature* 2008;453(7195):620-5. doi:
993 10.1038/nature07008.

994 43. Mills RH, Dulai PS, Vázquez-Baeza Y, Sauceda C, Daniel N, Gerner RR, et al.
995 Multi-omics analyses of the ulcerative colitis gut microbiome link Bacteroides vulgatus
996 proteases with disease severity. *Nat Microbiol* 2022;7(2):262-276. doi: 10.1038/s41564-
997 021-01050-3.

998 44. Dong G, Chen H, Qi M, Dou Y, Wang Q. Balance between metallothionein and
999 metal response element binding transcription factor 1 is mediated by zinc ions (review).
1000 *Mol Med Rep* 2015;11(3):1582-6. doi: 10.3892/mmr.2014.2969.

1001 45. Tsuji T, Naito Y, Takagi T, Kugai M, Yoriki H, Horie R, et al. Role of metallothioneins
1002 in murine experimental colitis. *Int J Mol Med* 2013;31(5):1037-46. doi:
1003 10.3892/ijmm.2013.1294.

1004 46. Devisscher L, Hindryckx P, Lynes MA, Waeytens A, Cuvelier C, De Vos F, et al.
1005 Role of metallothioneins as danger signals in the pathogenesis of colitis. *J Pathol*
1006 2014;233(1):89-100. doi: 10.1002/path.4330.

1007 47. Waeytens A, De Vos M, Laukens D. Evidence for a potential role of metallothioneins
1008 in inflammatory bowel diseases. *Mediators Inflamm* 2009;2009:729172. doi:
1009 10.1155/2009/729172.

1010 48. Lahiri A, Abraham C. Activation of pattern recognition receptors upregulates
1011 metallothioneins, thereby increasing intracellular accumulation of zinc, autophagy, and
1012 bacterial clearance by macrophages. *Gastroenterology* 2014;147(4):835-46. doi:
1013 10.1053/j.gastro.2014.06.024.

1014 49. Gevers D, Kugathasan S, Denson LA, Vázquez-Baeza Y, Van Treuren W, Ren B, et
1015 al. The treatment-naive microbiome in new-onset Crohn's disease. *Cell Host Microbe*
1016 2014;15(3):382-392. doi: 10.1016/j.chom.2014.02.005.

1017 50. Dey N, Soergel DA, Repo S, Brenner SE. Association of gut microbiota with post-
1018 operative clinical course in Crohn's disease. *BMC Gastroenterol* 2013;13:131. doi:
1019 10.1186/1471-230X-13-131.

1020 51. Craven M, Egan CE, Dowd SE, McDonough SP, Dogan B, Denkers EY, et al.
1021 Inflammation drives dysbiosis and bacterial invasion in murine models of ileal Crohn's
1022 disease. *PLoS One* 2012;7(7):e41594. doi: 10.1371/journal.pone.0041594.

52. Schaubeck M, Clavel T, Calasan J, Lagkouvardos I, Haange SB, Jehmlich N, et al. Dysbiotic gut microbiota causes transmissible Crohn's disease-like ileitis independent of failure in antimicrobial defence. *Gut* 2016;65(2):225-37. doi: 10.1136/gutjnl-2015-309333.

53. Kaakoush NO. Insights into the Role of Erysipelotrichaceae in the Human Host. *Front Cell Infect Microbiol* 2015;5:84. doi: 10.3389/fcimb.2015.00084.

54. van Haaften WT, Blokzijl T, Hofker HS, Olinga P, Dijkstra G, Bank RA, et al. Intestinal stenosis in Crohn's disease shows a generalized upregulation of genes involved in collagen metabolism and recognition that could serve as novel anti-fibrotic drug targets. *Therap Adv Gastroenterol* 2020;13:1756284820952578. doi: 10.1177/1756284820952578.

55. Carstens A, Dicksved J, Nelson R, Lindqvist MM, Andreasson A, Bohr J, et al. The Gut Microbiota in Collagenous Colitis Shares Characteristics With Inflammatory Bowel Disease-Associated Dysbiosis. *Clin Transl Gastroenterol* 2019;10(7):e00065. doi: 10.14309/ctg.0000000000000065.

56. Nielsen S, Needham B, Leach ST, Day AS, Jaffe A, Thomas T, et al. Disrupted progression of the intestinal microbiota with age in children with cystic fibrosis. *Sci Rep* 2016;6:24857. doi: 10.1038/srep24857.

57. Vernocchi P, Del Chierico F, Russo A, Majo F, Rossitto M, Valerio M, et al. Gut microbiota signatures in cystic fibrosis: Loss of host CFTR function drives the microbiota enterophenotype. *PLoS One* 2018;13(12):e0208171. doi: 10.1371/journal.pone.0208171.

58. Coffey MJ, Nielsen S, Wemheuer B, Kaakoush NO, Garg M, Needham B, et al. Gut Microbiota in Children With Cystic Fibrosis: A Taxonomic and Functional Dysbiosis. *Sci Rep* 2019;9(1):18593. doi: 10.1038/s41598-019-55028-7.

59. Cavallari JF, Pokrajac NT, Zlitni S, Foley KP, Henriksbo BD, Schertzer JD. NOD2 in hepatocytes engages a liver-gut axis to protect against steatosis, fibrosis, and gut dysbiosis during fatty liver disease in mice. *Am J Physiol Endocrinol Metab* 2020;319(2):E305-E314. doi: 10.1152/ajpendo.00181.2020.

1052 60. Ha CWY, Martin A, Sepich-Poore GD, Shi B, Wang Y, Gouin K, et al. Translocation
1053 of Viable Gut Microbiota to Mesenteric Adipose Drives Formation of Creeping Fat in
1054 Humans. *Cell* 2020;183(3):666-683.e17. doi: 10.1016/j.cell.2020.09.009.

1055 61. Zhuang X, Tian Z, Li N, Mao R, Li X, Zhao M, et al. Gut Microbiota Profiles and
1056 Microbial-Based Therapies in Post-operative Crohn's Disease: A Systematic Review.
1057 *Front Med (Lausanne)* 2021;7:615858. doi: 10.3389/fmed.2020.615858.

1058 62. Frau A, Ijaz UZ, Slater R, Jonkers D, Penders J, Campbell BJ, et al. Inter-kingdom
1059 relationships in Crohn's disease explored using a multi-omics approach. *Gut Microbes*
1060 2021;13(1):1930871. doi: 10.1080/19490976.2021.1930871.

1061 63. Davrandi M, Harris S, Smith PJ, Murray CD, Lowe DM. The Relationship Between
1062 Mucosal Microbiota, Colitis, and Systemic Inflammation in Chronic Granulomatous
1063 Disorder. *J Clin Immunol* 2022;42:312-324. doi: 10.1007/s10875-021-01165-6.

1064 64. Liang JQ, Li T, Nakatsu G, Chen YX, Yau TO, Chu E, et al. A novel faecal
1065 *Lachnoclostridium* marker for the non-invasive diagnosis of colorectal adenoma and
1066 cancer. *Gut* 2020;69(7):1248-1257. doi: 10.1136/gutjnl-2019-318532.

1067 65. Zhao L, Cho WC, Nicolls MR. Colorectal Cancer-Associated Microbiome Patterns
1068 and Signatures. *Front Genet* 2021;12:787176. doi: 10.3389/fgene.2021.787176.

1069 66. Zhou Y, Chen L, Sun G, Li Y, Huang R. Alterations in the gut microbiota of patients
1070 with silica-induced pulmonary fibrosis. *J Occup Med Toxicol* 2019;14:5. doi:
1071 10.1186/s12995-019-0225-1.

1072 67. Ta AD, Ollberding NJ, Karns R, Haberman Y, Alazraki AL, Hercules D, et al.
1073 Association of Baseline Luminal Narrowing With Ileal Microbial Shifts and Gene
1074 Expression Programs and Subsequent Transmural Healing in Pediatric Crohn Disease.
1075 *Inflamm Bowel Dis* 2021;27(11):1707-1718. doi: 10.1093/ibd/izaa339.

1076 68. Liu S, Zhao W, Lan P, Mou X. The microbiome in inflammatory bowel diseases:
1077 from pathogenesis to therapy. *Protein Cell* 2021;12(5):331-345. doi: 10.1007/s13238-
1078 020-00745-3.

69. Hall AB, Yassour M, Sauk J, Garner A, Jiang X, Arthur T, et al. A novel Ruminococcus gnavus clade enriched in inflammatory bowel disease patients. *Genome Med* 2017;9(1):103. doi: 10.1186/s13073-017-0490-5.

70. Long C, Venema K. Pretreatment of Rapeseed Meal Increases Its Recalcitrant Fiber Fermentation and Alters the Microbial Community in an *in vitro* Model of Swine Large Intestine. *Front Microbiol* 2020;11:588264. doi: 10.3389/fmicb.2020.588264.

71. Baxter NT, Schmidt AW, Venkataraman A, Kim KS, Waldron C, Schmidt TM. Dynamics of Human Gut Microbiota and Short-Chain Fatty Acids in Response to Dietary Interventions with Three Fermentable Fibers. *mBio* 2019;10(1):e02566-18. doi: 10.1128/mBio.02566-18.

72. Morrison DJ, Preston T. Formation of short chain fatty acids by the gut microbiota and their impact on human metabolism. *Gut Microbes* 2016;7(3):189-200. doi: 10.1080/19490976.2015.1134082.

73. Wang Y, Gao X, Ghozlane A, Hu H, Li X, Xiao Y, et al. Characteristics of Faecal Microbiota in Paediatric Crohn's Disease and Their Dynamic Changes During Infliximab Therapy. *J Crohns Colitis* 2018;12(3):337-346. doi: 10.1093/ecco-jcc/jjx153.

74. Ventin-Holmberg R, Eberl A, Saqib S, Korpela K, Virtanen S, Sipponen T, et al. Bacterial and Fungal Profiles as Markers of Infliximab Drug Response in Inflammatory Bowel Disease. *J Crohns Colitis* 2021;15(6):1019-1031. doi: 10.1093/ecco-jcc/jjaa252.

75. Dovrolis N, Michalopoulos G, Theodoropoulos GE, Arvanitidis K, Kolios G, Sechi LA, et al. The Interplay between Mucosal Microbiota Composition and Host Gene-Expression is Linked with Infliximab Response in Inflammatory Bowel Diseases. *Microorganisms* 2020;8(3):438. doi: 10.3390/microorganisms8030438.

76. Aden K, Rehman A, Waschina S, Pan WH, Walker A, Lucio M, et al. Metabolic Functions of Gut Microbes Associate With Efficacy of Tumor Necrosis Factor Antagonists in Patients With Inflammatory Bowel Diseases. *Gastroenterology* 2019;157(5):1279-1292.e11. doi: 10.1053/j.gastro.2019.07.025.

77. Schierova D, Roubalova R, Kolar M, Stehlikova Z, Rob F, Jackova Z, et al. Fecal Microbiome Changes and Specific Anti-Bacterial Response in Patients with IBD during Anti-TNF Therapy. *Cells* 2021;10(11):3188. doi: 10.3390/cells10113188.

78. Nepelska M, de Wouters T, Jacouton E, Béguet-Crespel F, Lapaque N, Doré J, et al. Commensal gut bacteria modulate phosphorylation-dependent PPARγ transcriptional activity in human intestinal epithelial cells. *Sci Rep* 2017;7:43199. doi: 10.1038/srep43199.

79. Byndloss MX, Olsan EE, Rivera-Chávez F, Tiffany CR, Cevallos SA, Lokken KL, et al. Microbiota-activated PPAR-γ signaling inhibits dysbiotic Enterobacteriaceae expansion. *Science* 2017;357(6351):570-575. doi: 10.1126/science.aam9949.

80. Abbot EL, McCormack JG, Reynet C, Hassall DG, Buchan KW, Yeaman SJ. Diverging regulation of pyruvate dehydrogenase kinase isoform gene expression in cultured human muscle cells. *FEBS J* 2005;272(12):3004-14. doi: 10.1111/j.1742-4658.2005.04713.x.

81. Bowman TA, O'Keeffe KR, D'Aquila T, Yan QW, Griffin JD, Killion EA, et al. Acyl CoA synthetase 5 (ACSL5) ablation in mice increases energy expenditure and insulin sensitivity and delays fat absorption. *Mol Metab* 2016;5(3):210-220. doi: 10.1016/j.molmet.2016.01.001.

82. Kumar J, Rani K, Datt C. Molecular link between dietary fibre, gut microbiota and health. *Mol Biol Rep* 2020;47(8):6229-6237. doi: 10.1007/s11033-020-05611-3.

83. Ahmadian R, Suh JM, Hah N, Liddle C, Atkins AR, Downes M, et al. PPARγ signaling and metabolism: the good, the bad and the future. *Nat Med* 2013;19(5):557-66. doi: 10.1038/nm.3159.

84. Dubuquoy L, Rousseaux C, Thuru L, Peyrin-Biroulet L, Romano O, Chavatte P, et al. PPARgamma as a new therapeutic target in inflammatory bowel diseases. *Gut* 2006;55(9):1341-9. doi: 10.1136/gut.2006.093484.

85. Peyrin-Biroulet L, Beisner J, Wang G, Nuding S, Oommen ST, Kelly D, et al.Peroxisome proliferator-activated receptor gamma activation is required for

maintenance of innate antimicrobial immunity in the colon. *Proc Natl Acad Sci U S A* 2010;107(19):8772-7. doi: 10.1073/pnas.0905745107.

86. Okumura R, Kurakawa T, Nakano T, Kayama H, Kinoshita M, Motooka D, et al. Lypd8 promotes the segregation of flagellated microbiota and colonic epithelia. *Nature* 2016;532(7597):117-21. doi: 10.1038/nature17406.

87. Burkhardt AM, Tai KP, Flores-Guiterrez JP, Vilches-Cisneros N, Kamdar K, Barbosa-Quintana O, et al. CXCL17 is a mucosal chemokine elevated in idiopathic pulmonary fibrosis that exhibits broad antimicrobial activity. *J Immunol* 2012;188(12):6399-406. doi: 10.4049/jimmunol.1102903.

88. Jensen J, Pedersen EE, Galante P, Hald J, Heller RS, Ishibashi M, et al. Control of endodermal endocrine development by Hes-1. *Nat Genet* 2000;24(1):36-44. doi: 10.1038/71657.

89. Katoh M, Katoh M. Notch signaling in gastrointestinal tract (review). *Int J Oncol* 2007;30(1):247-51.

90. Ma S, Shungin D, Mallick H, Schirmer M, Nguyen LH, Kolde R, et al. Population Structure Discovery in Meta-Analyzed Microbial Communities and Inflammatory Bowel Disease. *bioRxiv* 2020; doi: 10.1101/2020.08.31.261214.

91. Limketkai BN, Akobeng AK, Gordon M, Adepoju AA. Probiotics for induction of remission in Crohn's disease. *Cochrane Database Syst Rev* 2020;7(7):CD006634. doi: 10.1002/14651858.CD006634.pub3.

92. Jakubczyk D, Leszczyńska K, Górska S. The Effectiveness of Probiotics in the Treatment of Inflammatory Bowel Disease (IBD)- A Critical Review. *Nutrients* 2020;12(7):1973. doi: 10.3390/nu12071973.

93. Rieder F, Fiocchi C, Rogler G. Mechanisms, Management, and Treatment of Fibrosis in Patients With Inflammatory Bowel Diseases. *Gastroenterology* 2017;152(2):340-350.e6. doi: 10.1053/j.gastro.2016.09.047.

94. Rieder F. The gut microbiome in intestinal fibrosis: environmental protector or provocateur? *Sci Transl Med* 2013;5(190):190ps10. doi: 10.1126/scitranslmed.3004731.

95. Siva S, Rubin DT, Gulotta G, Wroblewski K, Pekow J. Zinc Deficiency is Associated with Poor Clinical Outcomes in Patients with Inflammatory Bowel Disease. *Inflamm Bowel Dis* 2017;23(1):152-157. doi: 10.1097/MIB.0000000000000989.

96. Ananthakrishnan AN, Khalili H, Song M, Higuchi LM, Richter JM, Chan AT. Zinc intake and risk of Crohn's disease and ulcerative colitis: a prospective cohort study. *Int J Epidemiol* 2015;44(6):1995-2005. doi: 10.1093/ije/dyv301.

97. Dai H, Wang L, Li L, Huang Z, Ye L. Metallothionein 1: A New Spotlight on Inflammatory Diseases. *Front Immunol* 2021;12:739918. doi: 10.3389/fimmu.2021.739918.

98. Jalanka J, Salonen A, Salojärvi J, Ritari J, Immonen O, Marciani L, et al. Effects of bowel cleansing on the intestinal microbiota. *Gut* 2015;64(10):1562-8. doi: 10.1136/gutjnl-2014-307240.

99. Imhann F, van der Velde KJ, Barbieri R, Alberts R, Voskuil MD, Vich Vila A, et al. The 1000IBD project: multi-omics data of 1000 inflammatory bowel disease patients; data release 1. *BMC Gastroenterol* 2019;19(1):5.

100. Heida FH, van Zoonen AGJF, Hulscher JBF, Te Kiefte BJC, Wessels R, Kooi EMW, et al. A Necrotizing Enterocolitis-Associated Gut Microbiota Is Present in the Meconium: Results of a Prospective Study. *Clin Infect Dis* 2016;62(7):863-870.

101. Bartram AK, Lynch MDJ, Stearns JC, Moreno-Hagelsieb G, Neufeld JD. Generation of multimillion-sequence 16S rRNA gene libraries from complex microbial communities by assembling paired-end illumina reads. *Appl Environ Microbiol* 2011;77(11):3846-52.

102. Bokulich NA, Joseph CM, Allen G, Benson AK, Mills DA. Next-generation sequencing reveals significant bacterial diversity of botrytized wine. *PLoS One* 2012;e36357.

103. Hu S, Uniken Venema WT, Westra HJ, Vich Vila A, Barbieri R, Voskuil MD, et al. Inflammation status modulates the effect of host genetic variation on intestinal gene expression in inflammatory bowel disease. *Nat Commun* 2021;12(1):1122.

104. Ghazi AR, Sucipto K, Rahnavard G, Franzosa EA, McIver LJ, Lloyd-Price J, et al. High-sensitivity pattern discovery in large, paired multi-omic datasets. *bioRxiv* 2021; doi: https://doi.org/10.1101/2021.11.11.468183

105. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 2005;102(43):15545-50.

106. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi, Mesirov JP, Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst* 2015;1(6):417-425.

Acknowledgements

All authors would like to express their gratitude towards all participants of the 1000IBD cohort. In addition, the authors would like to thank Kate McIntyre (Scientific Editor, Department of Genetics, University Medical Center Groningen) for language editing, and Ren Mao for his kind suggestions on figure styling and structural improvements (Department of Gastroenterology, The First Affiliated Hospital of Sun Yat-Sen University, Sun Yat-Sen University, Guangzhou, Guangdong, China).

Data availability

The datasets used and/or analyzed for the current study are available from the corresponding author on reasonable request. The data for the Groningen 1000IBD cohort can be requested with the accession number EGAS00001002702 (IDs: EGAD00001003991, EGAD00001008214, and EGAD00001008215).

Code availability

All analytic code used for this study can be found at the following link: https://github.com/GRONINGEN-MICROBIOME-CENTRE/Groningen-Microbiome/tree/master/Projects/IBD_biopsy_project.

Funding

RKW is supported by the Seerave foundation and the Netherlands Organization for Scientific Research. LMS is supported by Takeda. ARB is supported by an MD-PhD trajectory grant (grant no. 17-57) from the Junior Scientific Masterclass (JSM) of the

University of Groningen the Netherlands. This study was funded by Takeda Development Center Americas, Inc.

Conflicts of Interest

RKW acted as consultant for Takeda, received unrestricted research grants from Takeda, Johnson & Johnson, Tramedico and Ferring and received speaker fees from MSD, Abbvie, and Janssen Pharmaceuticals. GD received an unrestricted research grant from Takeda and speaker fees from Pfizer and Janssen Pharmaceuticals. All other authors declare no competing interests.

Authors' contributions

Conceptualization: RKW. Investigation: SH, ARB, RG, BHJ, RM, AB, IJH, JRB, HJMH, AVV, LMS and RKW. Methodology: SH, ARB, RG, BHJ, RM, IJH, JRB, HJMH, AVV, LMS and RKW. Funding acquisition: RKW. Supervision: EAMF, AVV, LMS and RKW. Writing - original manuscript: SH, ARB and LMS. Writing - review and editing: all authors.

Supplementary Methods

Polymerase chain reaction (PCR), DNA clean-up, and MiSeq library preparation for mucosal 16 microbiota characterization

The PCR procedure consisted of the following conditions: an initial cycle of 94°C for 3 min followed by 32 cycles of 94°C for 45 sec, 50°C for 60 sec and 72°C for 90 sec, with a final extension of 72°C for 10 min. Agarose gel electrophoresis confirmed the presence of the PCR product (band at ~465 bp) in successfully amplified samples. Subsequently, DNA samples were thoroughly cleaned by mixing the remainder of the PCR product with 25 µL Agencourt AMPure XP beads (Beckman Coulter, Brea, California, USA) followed by an incubation of 5 min at room temperature. Beads were separated from the mixture by placing the samples within a magnetic bead separator for 2 min. After discarding the cleared solution, beads were washed twice by resuspending them in 200 µL fresh 80% ethanol, followed by an incubation of 30 sec in the magnetic bead separator, and again discarding the cleared solution. The pellet was dried for 15 min and resuspended in 52.5 µL 10 mM Tris HCl buffer (pH 8.5). Fifty (50) µL of this solution was subsequently brought into a new tube. DNA concentrations were measured using a Qubit® 2.0 fluorometer (Thermo-Fisher Scientific, Waltham, Massachusetts, USA). To ensure similar library representations across samples, 2 nM dilutions of each sample were prepared accordingly. A library was created by pooling 5 µl of each diluted sample. Subsequently, 10 µL of the sample pool and 10 µL 0.2 M NaOH were mixed and incubated for 5 min to allow denaturation of the sample DNA. 980 µL of the HT1 buffer of the MiSeq 2x300 cartridge was then added to this mixture. Next, a denatured diluted PhiX solution was created by combining 2 µL 10 nM PhiX library with 3 µL 10 mM Tris HCl buffer (pH 8.5) with 0.1% Tween-20. 5 µL 0.2 M NaOH was added to this mixture and incubated for 5 min at room temperature. This 10 µL mixture was eventually mixed with 990 µL HT1 buffer. From the diluted sample pool, 150 µL was combined with 50 µL of the diluted PhiX solution, which was further diluted by the addition of 800 µL HT1 buffer. Finally, 600 µL of the prepared library solution was loaded into the sample loading reservoir of the 2x300 MiSeq cartridge for 16S rRNA amplicons sequencing (MiSeq Benchtop Sequencer, Illumina, San Diego, California,

USA). Samples with low DNA concentrations after clean-up (quality score < 0.9) were discarded by PANDAseq to increase quality of sequence read-outs.

Supplementary Table S1. Nucleotide sequences of primers used for library construction for bacterial 16S rRNA gene (Illumina) sequencing.

Primer	Sequence
V3_F_modified	aatgatacggcgaccaccgagatctacactctttccctacacgacgctcttccgatctNNNNCCTACGGGAGGCAGCAG
V4_1R	caagcagaagacggcatacgagat**CGTGAT**gtgactggagttcagacgtgtgctcttccgatctGGACTACHVGGGTWTCTAAT
V4_2R	caagcagaagacggcatacgagat**ACATCG**gtgactggagttcagacgtgtgctcttccgatctGGACTACHVGGGTWTCTAAT
V4_3R	caagcagaagacggcatacgagat**GCCTAA**gtgactggagttcagacgtgtgctcttccgatctGGACTACHVGGGTWTCTAAT
V4_4R	caagcagaagacggcatacgagat**TGGTCA**gtgactggagttcagacgtgtgctcttccgatctGGACTACHVGGGTWTCTAAT
V4_5R	caagcagaagacggcatacgagat**CACTGT**gtgactggagttcagacgtgtgctcttccgatctGGACTACHVGGGTWTCTAAT
V4_6R	caagcagaagacggcatacgagat**ATTGGC**gtgactggagttcagacgtgtgctcttccgatctGGACTACHVGGGTWTCTAAT
V4_7R	caagcagaagacggcatacgagat**GATCTG**gtgactggagttcagacgtgtgctcttccgatctGGACTACHVGGGTWTCTAAT

V4_8R	caagcagaagacggcatacgagat**TCAAGT**gtgactggagttcagacgtgtgctcttccgatct GGACTACHVGGGTWTCTAAT
V4_9R	caagcagaagacggcatacgagat**CTGATC**gtgactggagttcagacgtgtgctcttccgatct GGACTACHVGGGTWTCTAAT
V4_10R	caagcagaagacggcatacgagat**AAGCTA**gtgactggagttcagacgtgtgctcttccgatct GGACTACHVGGGTWTCTAAT
V4_11R	caagcagaagacggcatacgagat**GTAGCC**gtgactggagttcagacgtgtgctcttccgatct GGACTACHVGGGTWTCTAAT
V4_12R	caagcagaagacggcatacgagat**TACAAG**gtgactggagttcagacgtgtgctcttccgatct GGACTACHVGGGTWTCTAAT
V4_13R	caagcagaagacggcatacgagat**CGTACT**gtgactggagttcagacgtgtgctcttccgatct GGACTACHVGGGTWTCTAAT
V4_14R	caagcagaagacggcatacgagat**GACTGA**gtgactggagttcagacgtgtgctcttccgatct GGACTACHVGGGTWTCTAAT
V4_15R	caagcagaagacggcatacgagat**GCTCAA**gtgactggagttcagacgtgtgctcttccgatct GGACTACHVGGGTWTCTAAT
V4_16R	caagcagaagacggcatacgagat**TCGCTT**gtgactggagttcagacgtgtgctcttccgatct GGACTACHVGGGTWTCTAAT

1272 Bold uppercase letter highlight the index sequences. Lowercase letters indicate adapter sequences
1273 necessary for flow-cell binding. Underlined lowercase letters indicate binding sites for Illumina sequencing
1274 primers. Regular uppercase letters indicate the V3 and V4 region primers (341F for the forward primer,
1275 806R for the reverse primer). The inclusion of four maximally degenerated bases ("NNNN") maximizes
1276 the diversity during the first four bases of the run, which is important for unique cluster identification and
1277 base-calling accuracy.

1278

Supplementary Table Index

All Supplementary Tables have been uploaded separately for peer-review.

Supplementary Table S1. Differential gene expression analyses between non-inflamed and inflamed biopsies from ileal CD (group 1), colonic CD (group 2) and UC (group 3).

Supplementary Table S2. Differential gene expression analysis between inflamed biopsies from patients with CD (reference) and patients with UC.

Supplementary Table S3. Differential expression analysis of deconvoluted cell types between inflamed colonic biopsies of patients with CD (reference) and patients with UC.

Supplementary Table S4. Relative abundances of mucosal bacterial groups in different groups (CD, UC and non-IBD controls) and biopsy locations (ileum or colon).

Supplementary Table S5. Comparison of relative abundances of bacterial groups between non-IBD controls and ileal CD (group 1), colonic CD (group 2) and UC (group 3).

Supplementary Table S6. Hierarchical analysis performed using an end-to-end statistical algorithm (HAllA) demonstrating the main associations between mucosal bacterial groups and clinical phenotypes.

Supplementary Table S7. Genes and bacteria contained in component pair 1 from sparse-CCA analysis (FDR<0.05).

Supplementary Table S8. Pathway annotation of genes involved in component pair 1 from sparse-CCA analysis (FDR<0.05).

Supplementary Table S9. Genes and bacteria contained in component pair 3 from sparse-CCA analysis (FDR<0.05).

Supplementary Table S10. Pathway annotation of genes involved in component pair 3 from sparse-CCA analysis (FDR<0.05).

Supplementary Table S11. Genes and bacteria contained in component pair 7 from sparse-CCA analysis (FDR<0.05).

Supplementary Table S12. Pathway annotation of genes involved in component pair 7 from sparse-CCA analysis (FDR<0.05).

Supplementary Table S13. Genes and bacteria contained in component pair 8 from sparse-CCA analysis (FDR<0.05).

Supplementary Table S14. Pathway annotation of genes involved in component pair 8 from sparse-CCA analysis (FDR<0.05).

Supplementary Table S15. Genes and bacteria contained in component pair 5 from sparse-CCA analysis (FDR<0.05).

Supplementary Table S16. Pathway annotation of genes involved in component pair 5 from sparse-CCA analysis (FDR<0.05).

Supplementary Table S17. Genes and bacteria contained in component pair 9 from sparse-CCA analysis (FDR<0.05).

Supplementary Table S18. Pathway annotation of genes involved in component pair 9 from sparse-CCA analysis (FDR<0.05).

Supplementary Table S19. Individual pairwise gene-bacteria associations.

Supplementary Table S20. Genes and bacteria associated with fibrostenotic CD/Montreal B2 (reference: Montreal B1).

Supplementary Table S21. Microbiota-associated gene clusters in patients with non-stricturing, non-penetrating disease (Montreal B1) and fibrostenotic CD (Montreal B2) including microbiota-associated pathway annotation and cluster comparisons.

Supplementary Table S22. Genes and bacteria associated with TNF-α-antagonists use (reference: non-users).

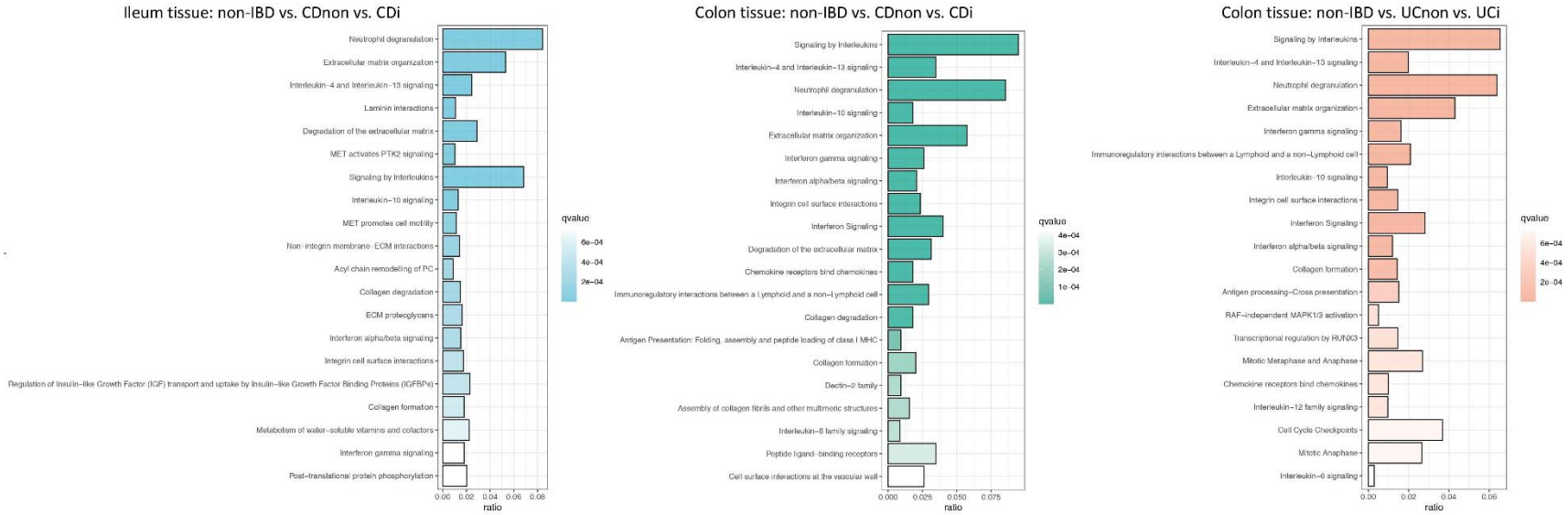
Supplementary Table S23. Microbiota-associated gene clusters in patients not using and using TNF-α-antagonists including microbiota-associated pathway annotation and cluster comparisons.

Supplementary Table S24. Individual pairwise gene-bacteria associations and their interaction with the degree of mucosal dysbiosis (Lloyd-Price *et al.*, *Nature* 2019).

Supplementary Table S25. Mucosal microbiota and other phenotypic factors explaining variation in mucosal cell type enrichment in patients with IBD.

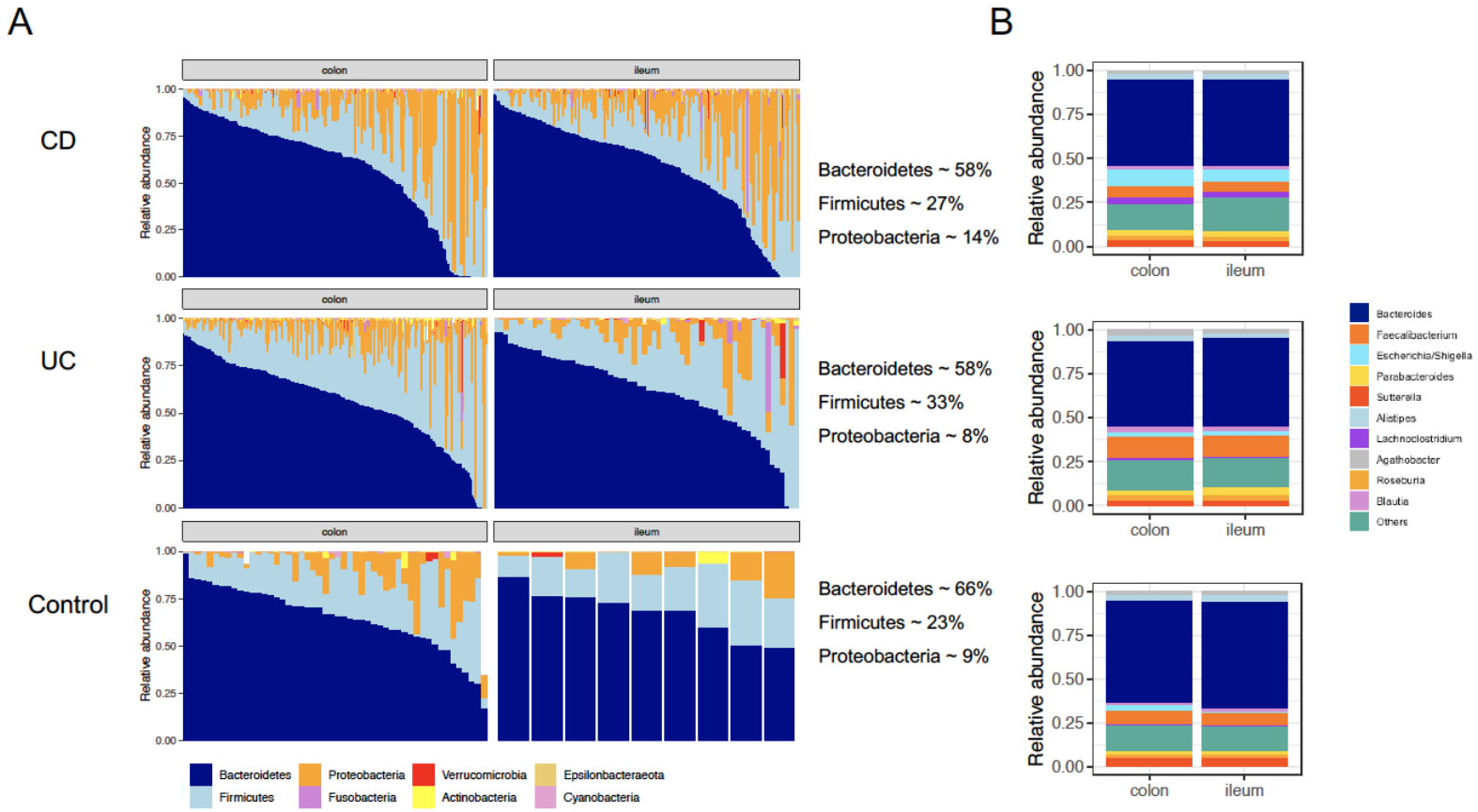
Extended Data Figures

Extended Data Fig. S1



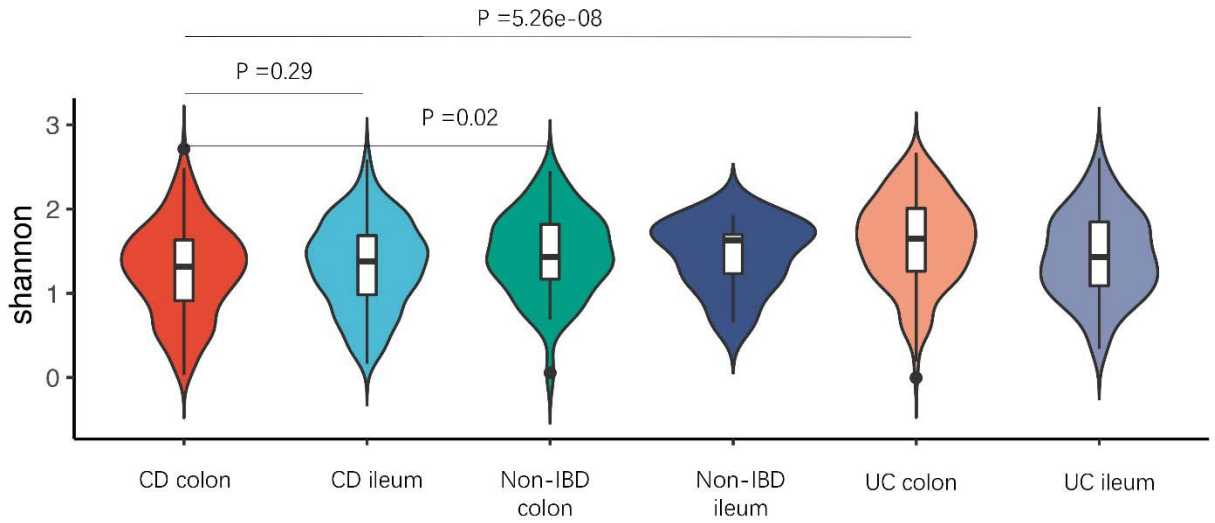
Extended Data Fig. S1. **Analysis of pathways associated with each comparative gene expression analysis.** The main pathways associated with inflamed ileal tissue in patients with CD (blue) include neutrophil degranulation, extracellular matrix (ECM) organization and IL-4/IL-13-signaling. Similar pathways were overexpressed in inflamed colonic tissue from patients with CD (green), but with a more prominent contribution from interleukin signaling pathways. Interleukin signaling pathways were also dominantly expressed in inflamed colonic tissue from patients with UC (orange), with other pathways expressed including neutrophil degranulation, ECM pathways, interferon gamma signaling and immunoregulatory interactions between lymphoid and non-lymphoid cells. Pathways were annotated using the Reactome pathway database. Abbreviations: CDi, inflamed tissue from patients with Crohn's disease; CD-non, non-inflamed tissue from patients with Crohn's disease; UCi, inflamed tissue from patients with ulcerative colitis; UC-non, non-inflamed tissue from patients with ulcerative colitis.

Extended Data Fig. S2



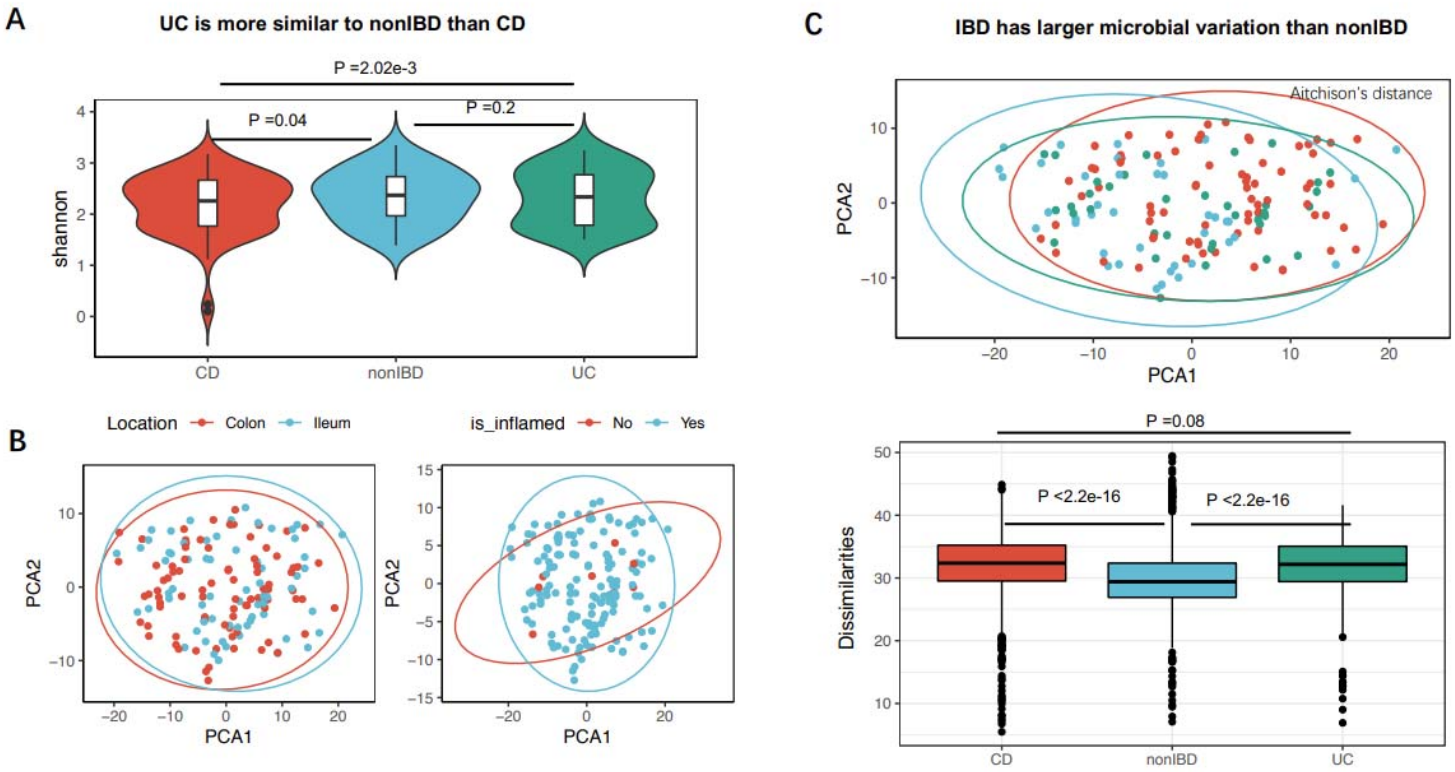
Extended Data Fig. S2. **Mucosal 16S rRNA gene sequencing characterization demonstrates distinct compositional differences in relative abundances** on (**A**) bacterial phylum level and (**B**) bacterial genus level. Abbreviations: CD, Crohn's disease; UC, ulcerative colitis.

Extended Data Fig. S3



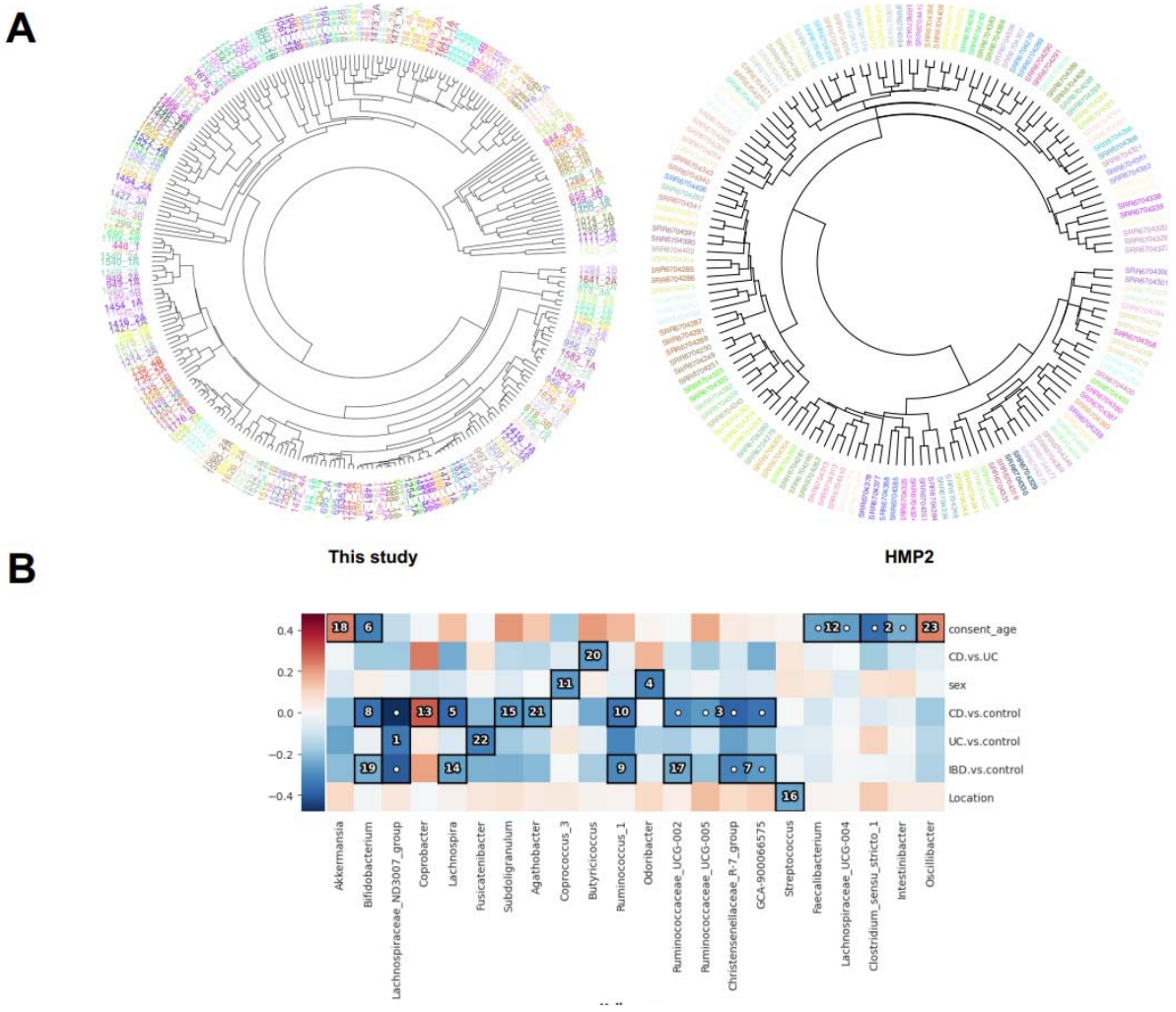
Extended Data Fig. S3. **Microbial alpha-diversity (Shannon index) is significantly lower in colonic biopsies from patients with CD compared to colonic biopsies derived from patients with UC or controls.** This indicates that this difference is not solely attributable to ileal biopsies from patients with CD.

Extended Data Fig. S4



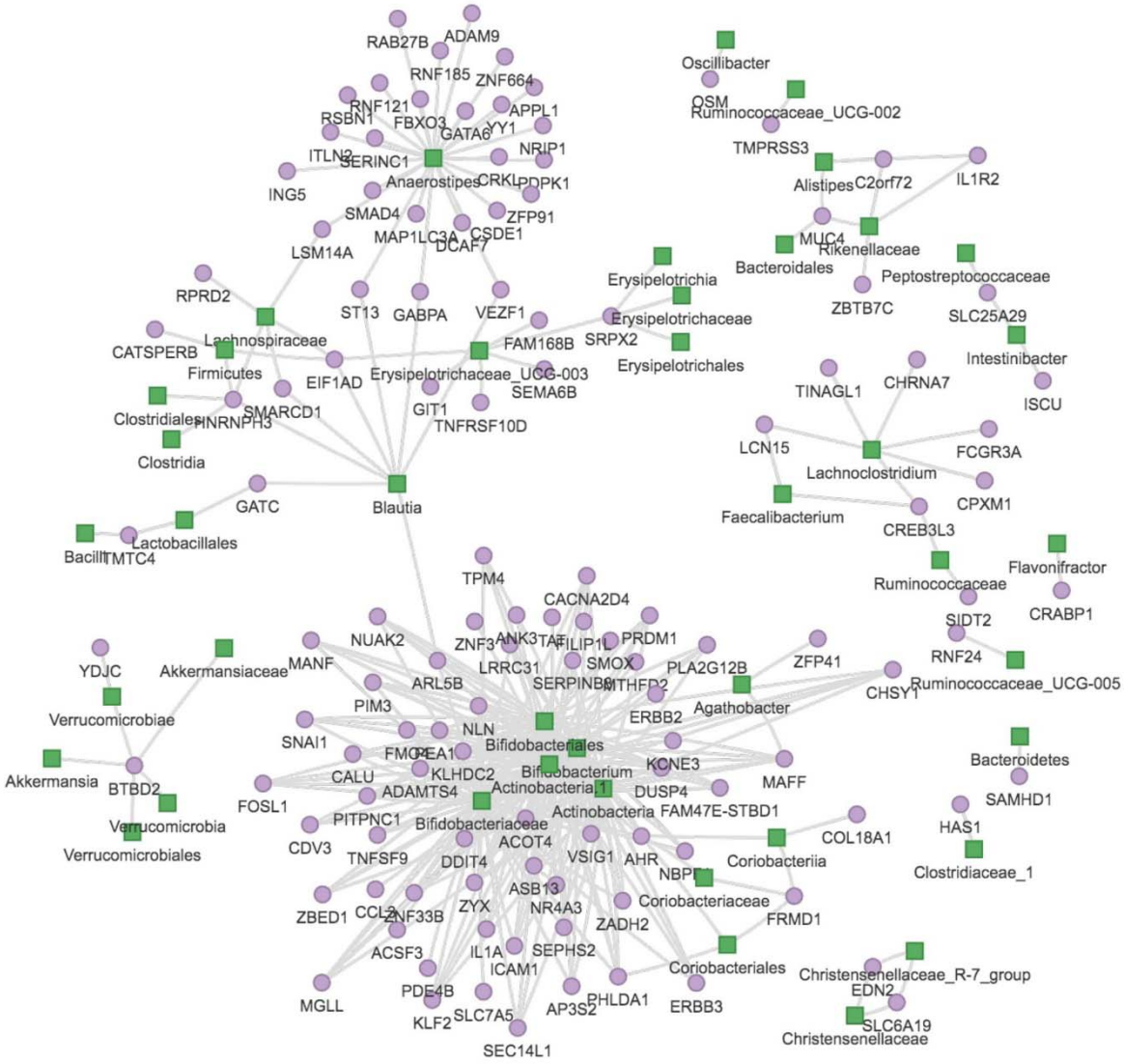
Extended Data Fig. S4. **Replication of overall mucosal microbiota characterization in patients with IBD and non-IBD controls.** Replication was performed in data derived from the HMP2 cohort study [13]. **a**, Microbial alpha-diversity (Shannon index) was lowest in ptaients with CD (n=85) compared to patients with UC (n=46) and non-IBD controls (n=45). **b**, PCA plots based on Aitchison's distances and stratified by tissue location and inflammatory status (colors as in **a**). **c**, PCA plot showing microbial dissimilarity (Aitchison's distances) in CD, UC and non-IBD controls. **d**, Microbial dissimilarity is highest in samples from patients with CD, followed by patients with UC and non-IBD controls. CD, Crohn's disease; PCA, principal component analysis; UC, ulcerative colitis.

Extended Data Fig. S5



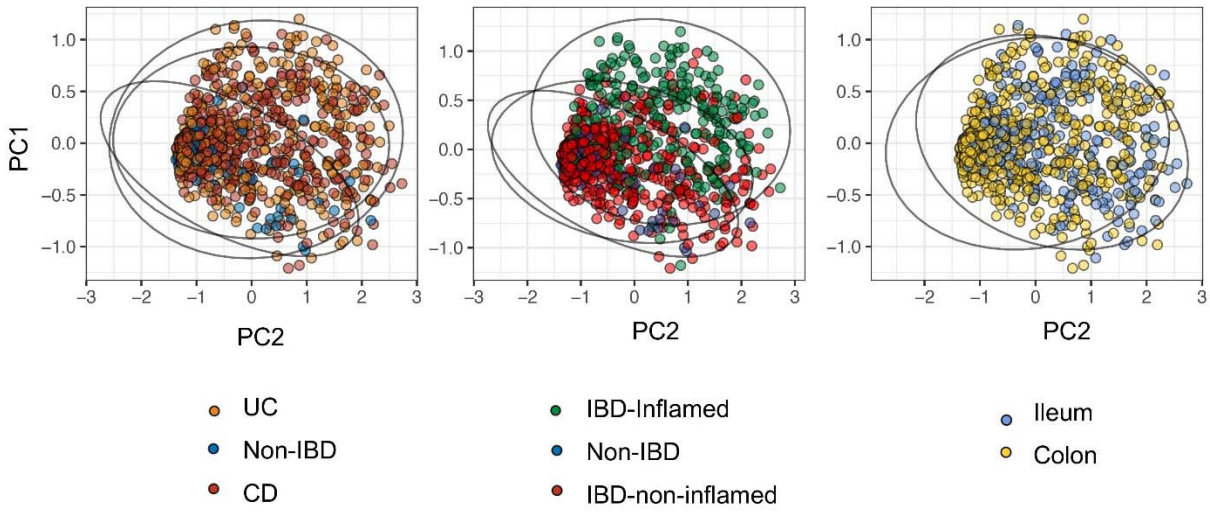
Extended Data Fig. S5. **Composition of the mucosal microbiota is highly personalized and influenced by disease parameters and clinical factors in patients with IBD and controls.** (**A**) Hierarchical clustering analysis demonstrating that tissue samples from the same individual (paired samples) clearly cluster together (colors indicate unique individuals). (**B**) Hierarchical analysis performed using an end-to-end statistical algorithm (HAllA) showing the main phenotypic factors that correlate with intestinal mucosal microbiota composition. Heatmap color palette indicates normalized mutual information. Numbers and dots in cells identify the significant pairs of features (phenotypic factors vs. bacterial taxa) in patients with IBD and controls. Abbreviations: BMI, body-mass index; CD, Crohn's disease; UC, ulcerative colitis.

Extended Data Fig. S6



Extended Data Fig. S6. Network graph displaying significant individual gene–bacteria interactions. Green squares indicate bacterial groups. Purple dots indicate host gene expression. Each connecting line indicates statistically significant gene–bacteria associations after adjustment for age, sex, batch, medication use, tissue inflammatory status and tissue location. Most individual gene–bacteria associations (94%) overlap with the results from the sparse-CCA analysis (**Figure 4**).

Extended Data Fig. S7



Extended Data Fig. S7. **Principal component analysis (PCA) plots demonstrating variation in cell type–enrichment labeled by diagnosis, biopsy inflammatory status and intestinal location**. Each dot represents one tissue sample. Left: Patients with IBD, both CD and UC, show significantly different intestinal cell type composition compared to controls. Middle: Tissue inflammatory status induces shifts in cell type composition, showing differences between non-inflamed IBD tissue vs. control tissue and inflamed IBD tissue vs. control tissue. Right: Tissue location (ileum vs. colon) also demonstrates distinct variation in cell type composition.

Supplementary Results

Box 1. Individual mucosal gene–bacteria associations and their potential biological implications in IBD.

Mucosal bifidobacteria positively associate with aryl hydrocarbon receptor (AHR) and ABC-transporter (ABCC1) expression levels

The positive association between *AHR* expression and bifidobacteria could be explained by the fact that *Bifidobacterium* spp. can produce aromatic lactic acids such as indole-3-lactic acid (out of aromatic amino acids like tryptophan) via aromatic lactate dehydrogenase, which in turn activates the host aryl hydrocarbon receptor [1,2]. Activation of the aryl hydrocarbon receptor, a crucial regulator of intestinal homeostasis and immune responses, leads to a reduction of inflammation in intestinal epithelial cells [3] and confers immunoprotective effects [4].

Another intriguing observation is the positive association between bifidobacteria and host expression of the *ABCC1* gene. *ABCC1* is a member of the ATP-binding cassette transporters (ABC transporters, and also known as multidrug resistance-associated protein 1, MRP1) that has multiple physiological functions, but it may also confer pathophysiological sequelae, especially in the context of cancer [5]. Under physiological circumstances, it detoxifies endogenously generated toxic substances (as well as xenobiotics), protects against oxidative stress, transports leukotrienes and lipids and may facilitate the cellular export and body distribution of vitamin B_{12} [6]. Interestingly, several *Bifidobacterium* species (e.g. *B. animalis*, *B. longum* and *B. infantis*) can synthesize vitamin B_{12}, which is subsequently absorbed in the large intestine via unknown mechanisms [7-9].

Mucosal bifidobacteria associate with FOSL1, a subunit of the AP-1 transcription factor

Associations between mucosal *Bifidobacterium* bacteria and expression of *FOSL1* genes were amongst the top significant individual gene–bacteria interactions. Fos-related antigen 1 (FRA1), encoded by *FOSL1*, is a subunit of the activator protein 1 (AP-1) transcription factor. In the intestine, the AP-1 transcription factor is commonly

activated in response to inflammatory stimuli and has been implicated in IBD [10]. More specifically, an interaction may exist between AP-1 activity and the glucocorticoid receptor, which may be part of the anti-inflammatory effects of steroid treatment [11]. In steroid-resistant patients with CD, AP-1 activation was primarily observed in the nuclei of intestinal epithelial cells, whereas this activation was restricted to lamina propria macrophages in steroid-sensitive patients [10]. This suggests a differing cellular activation pattern of AP-1 activation in steroid-resistant patients where the expression of this transcription factor may interfere with the activity of the glucocorticoid response. In an experimental study in which pregnant mice were supplemented with butyrate, FOS genes, including *Fosl1*, were observed to be downregulated in the colon and associated with protection against experimentally-induced colitis [12]. Although there are currently no reports of potential immune-modulating effects for Fosl1, it has 85% homology with Fosl2, another AP-1 transcription factor. A recent study demonstrated that Fosl2 is important in T-reg development and control of autoimmunity. Interestingly, several GWASs have reported associations of a SNP located in the promoter region of *FOSL2* with IBD [13-15], and the presence of this SNP was also shown to correlate with *FOSL2* expression in blood cells of patients with IBD [16]. In the context of T-regs, *FOSL2* also appears to be important as it is a determinant of a highly suppressive subpopulation of T-regs in humans that are particularly enriched in the lamina propria of patients with CD, supporting wound healing in the intestinal mucosa [17]. Although speculative, bifidobacteria, as well as their metabolites such as butyrate, may potentially confer immune-modulating properties via interaction with *FOSL1* expression.

Mucosal bifidobacteria positively associate with Krüppel-like factor 2 (KLF2) expression

Krüppel-like factor 2 (encoded by *KLF2*) is a negative regulator of intestinal inflammation, and its expression is found to be reduced in patients with IBD [18]. *KLF2* also negatively regulates differentiation of adipocytes and strongly inhibits PPAR-γ expression, which prevents differentiation of preadipocytes into adipocytes and thereby prevents adipogenesis [19]. *KLF2* also plays an important role in endothelial physiology, where it may act as a molecular switch by regulating endothelial cell function in inflammatory disease states [20]. Interestingly, *KLF2* modifies the trafficking of T-regs,

as increased *KLF2* expression in T-regs promotes the induction of peripheral immunological tolerance, whereas, in the absence of its expression, T-regs are unable to effectively migrate to secondary lymphoid tissues [21]. Indeed,, it was demonstrated in mouse experiments that mice developed IBD in the presence of *KLF2*-deficient T-regs, which were unable to prevent colitis by disrupted co-trafficking of effector and regulatory T cells. In light of these considerations, mucosal bifidobacteria may confer beneficial immune-modulating properties by upregulating *KLF2* expression, thereby stimulating T-reg migration and contributing to immunological self-tolerance in the context of IBD.

Mucosal Anaerostipes bacteria positively associate with host SMAD4 expression

Anaerostipes, which belong to the *Lachnospiraceae* family, are anaerobic bacteria that are well-known butyrate-producers. Butyrate serves as the primary energy source for colonic epithelial cells and is characterized by anti-inflammatory and anti-carcinogenic properties. *SMAD4* is an important intracellular effector of the TGF-β superfamily of proteins. These proteins have important functions in alleviating intestinal inflammation and maintenance of gut mucosal homeostasis. Haploinsufficiency of *SMAD4* in mice and humans has been associated with an increased susceptibility to colonic inflammation [22]. In patients with CD, reduced epithelial protein levels of SMAD4 were observed that was associated with disease activity, indicating defective mucosal TGF-β signaling during active intestinal inflammation. In an experimental animal study, mice with an epithelial deletion of *Smad4* presented with macroscopic invasive adenocarcinoma of the distal colon and rectum 3 months after DSS-induced colitis [23]. Indeed, *SMAD4* mutations in humans are linked to juvenile polyposis syndrome and associated with poor disease outcome in several types of cancer [24-27]. Using RNA-seq analysis, a strong inflammatory expression profile was observed after *SMAD4* deletion, with expression of various inflammatory cytokines and chemokines, including CCL20. In addition, it was demonstrated that CCL20 could be repressed by *SMAD4* in colonic epithelial cells, proving that TGF-β signaling could block the induction of CCL20 expression to protect against the development of colitis-associated cancer.

In an experimental study involving human hepatic stellate cells, butyrate was demonstrated to be protective against diet-induced nonalcoholic steatohepatitis and liver fibrosis via suppression of TGF-β signaling pathways in which SMAD proteins are involved. Although butyrate mainly showed antifibrotic effects via reduction of non-canonical TGF-β signaling cascades, there was also a significant increase in the expression of SMAD4 with the addition of butyrate on top of TGF-β treatment [28]. We found *Anaerostipes* bacteria to also be strongly associated with expression of *ZNF644*, a zinc finger protein that is positively regulated by intracellular zinc concentrations. Depletion of intracellular zinc levels, or even zinc deficiency, may have destabilizing effects on SMAD proteins and thereby impair the TGF-β signaling pathway [29].

Mucosal Verrucomicrobia bacteria inversely associate with expression of the IBD susceptibility gene YDJC

We observed significant inverse associations between Verrucomicrobia bacteria, of which *Akkermansia muciniphila* is a well-known member, and the expression of the *YDJC* gene, which encodes for the YdjC chitooligosaccharide deacetylase homolog (YdjC) protein. This gene has been identified as a shared susceptibility gene for CD, UC and psoriasis [13,30,31]. *YDJC* was originally identified as a celiac disease–associated susceptibility locus, but some SNPs were also associated with CD as well as with pediatric-onset CD [32]. YdjC catalyzes the deacetylation of acetylated carbohydrates, an important reaction in the degradation of oligosaccharides [33]. *YDJC* expression has been associated with tumor progression in studies of lung cancer [34,35]. The observed inverse association between *Akkermansia* and *YDJC* expression may suggest a potential protective role of *Akkermansia*, as decreased *YDJC* expression may mitigate its pro-carcinogenic effects. Despite the association between *YDJC* and the susceptibility to IBD on a genetic level, its precise functional role remains largely unknown [32].

Mucosal Alistipes bacteria positively associate with MUC4 expression

The bacterial genus *Alistipes*, belonging to family *Rikenellaceae* and phylum Bacteroidetes, is a recently discovered bacterial species, of which many have been isolated from the human gut microbiome. The role of *Alistipes* in health and disease is

still unclear. Some evidence indicates that it may confer protective effects to the host, but other studies report pathogenic effects, e.g. in colorectal cancer development. A key factor believed to determine the relative abundance of *Alistipes* is the dysbiotic state of the gut microbiome [36]. In IBD, there is also conflicting data about the pathogenicity of *Alistipes* species. *Alistipes finegoldii* has been demonstrated to exert anti-inflammatory effects in experimental models of colitis [37]. Likewise, another study found an increased abundance of *Alistipes* in *NOD2*-knockout mice that had less severe (TNBS-induced) colitis compared to wild-type mice [38]. It has also been reported that *Alistipes* abundance could increase after taking probiotic supplements, which in turn may protect against hepatocellular cancer growth in an experimental setting [39]. However, metagenomic studies have shown that *Alistipes* abundances were increased in mouse models of spontaneous CD-like ileitis terminalis as compared to wild-type mice, suggesting that *Alistipes* species may also play a pathogenic role by eliciting segmental ileitis [40,41].

MUC4 encodes for mucin 4, a protein found in the glycocalyx present on the intestinal epithelium. Deletion or knockouts of *Muc4* have demonstrated protective effects in mouse models, as shown by lower levels of proinflammatory factors and resistance against DSS-induced colitis. It is still unclear how this protective mechanism of *MUC4* deletion works, but it has been hypothesized that it may trigger the concomitant upregulation of other mucin proteins (e.g. *MUC*1-3) as these genes have been observed to be highly expressed in *Muc4*-knockout mice with DSS-induced colitis [42,43]. Based on this, we speculate that the positive association between *Alistipes* abundance and *MUC4* expression may imply a potential pathogenic role of *Alistipes* in the context of IBD-associated dysbiosis. However, in our data, we did not observe a significant interaction via dysbiotic status between *Alistipes* abundance and *MUC4* expression.

Mucosal Oscillibacter bacteria positively associate with OSM expression

Oscillibacter-like bacteria, which include *Oscillibacter* and *Oscillospira*, are commonly detected in human gut microbial communities, although their exact physiological role is not fully understood. Previously, it was reported that *Oscillibacter* may be a potentially important bacterium in mediating high fat diet–induced intestinal dysfunction, which was

supported by a negative association between *Oscillibacter* and intestinal barrier function parameters [44]. Similarly, the abundance of *Oscillibacter* has been reported as a key bacterial group associated with colitis development in DSS-induced colitis in mice and with prenatal stress in rodents [45,46]. However, a recent study linking gut microbiota profiles to sulfur metabolism in patients with CD demonstrated that *Oscillibacter* abundance was enriched in patients with inactive compared to active disease but diminished in patients with IBD compared to controls [47,48]. Thus, similar to *Bacteroides* and *Alistipes*, the exact functional role of *Oscillibacter* in the context of IBD remains elusive, but it will likely depend on gut microbial dysbiosis and the intestinal (inflammatory) environment. The *OSM* gene encodes for the oncostatin M protein, a well-known inflammatory mediator in IBD that drives intestinal inflammation, mainly *via* activation of JAK-STAT and PI3K-Akt pathways [49]. Besides induction of other inflammatory events, it primarily triggers the production of various cytokines, chemokines and adhesion molecules that contribute to intestinal inflammation [50]. In addition, OSM is a marker for non-responsiveness to TNF-α-antagonists in patients with IBD [51]. Considering these findings, the positive association between *OSM* expression and *Oscillibacter* abundance we observe supports a potentially pathogenic role for this bacterial species in IBD.

Mucosal Blautia bacteria associate with host ST13 expression levels

Hsc70-interacting protein, encoded by the *ST13* gene, mediates the assembly of the human glucocorticoid receptor, which requires involvement of intracellular chaperone proteins such as heat shock proteins HSP70 and HSP90 [52]. Reduced expression of ST13 has been observed in patients with colorectal cancer, suggesting that ST13 may constitute a candidate tumor-suppressor gene [53,54]. The positive association we observe between mucosal *Blautia* abundance and *ST13* gene expression may therefore point to a protective anti-carcinogenic role for *Blautia* in the intestines.

Supplementary References to Box 1

1. Wong CB, Tanaka A, Kuhara T, Xiao JZ. Potential effects of Indole-3-Lactic Acid, a Metabolite of Human Bifidobacteria, on NGF-induced Neurite Outgrowth in PC12 Cells. *Microorganisms* 2020;8(3):398. doi: 10.3390/microorganisms8030398.

2. Laursen MF, Sakanaka M, von Burg N, Mörbe U, Andersen D, Moll JM, et al. Bifidobacterium species associated with breastfeeding produce aromatic lactic acids in the infant gut. *Nat Microbiol* 2021;6(11):1367-1382. doi: 10.1038/s41564-021-00970-4.

3. Ehrlich AM, Pacheco AR, Henrick BM, Taft D, Xu G, Huda MN, et al. Indole-3-lactic acid associated with Bifidobacterium-dominated microbiota significantly decreases inflammation in intestinal epithelial cells. *BMC Microbiol* 2020;20(1):357. doi: 10.1186/s12866-020-02023-y.

4. Henrick BM, Rodriguez L, Lakshmikanth T, Pou C, Henckel E, Arzoomand A, et al. Bifidobacteria-mediated immune system imprinting early in life. *Cell* 2021;184(15):3884-3898.e11. doi: 10.1016/j.cell.2021.05.030.

5. Bakos E, Homolya L. Portrait of multifaceted transporter, the multidrug resistance-associated protein 1 (MRP1/ABCC1). *Pflugers Arch* 2007;453(5):621-41. doi: 10.1007/s00424-006-0160-8.

6. He SM, Li R, Kanwar JR, Zhou SF. Structural and functional properties of human multidrug resistance protein 1 (MRP1/ABCC1). *Curr Med Chem* 2011;18(3):439-81. doi: 10.2174/092986711794839197.

7. Yoshii K, Hosomi K, Sawane K, Kunisawa J. Metabolism of Dietary and Microbial Vitamin B Family in the Regulation of Host Immunity. *Front Nutr* 2019;6:48. doi: 10.3389/fnut.2019.00048.

8. Beedholm-Ebsen R, van de Wetering K, Hardlei T, Nexø E, Borst P, Moestrup SK. Identification of multidrug resistance protein 1 (MRP1/ABCC1) as a molecular gate for cellular export of cobalamin. *Blood* 2010;115(8):1632-9. doi: 10.1182/blood-2009-07-232587.

9. Lee JH, O'Sullivan DJ. Genomic insights into bifidobacteria. *Microbiol Mol Biol Rev* 2010;74(3):378-416. doi: 10.1128/MMBR.00004-10.

10. Bantel H, Schmitz ML, Raible A, Gregor M, Schulze-Osthoff K. Critical role of NF-kappaB and stress-activated protein kinases in steroid unresponsiveness. *FASEB J* 2002;16(13):1832-4. doi: 10.1096/fj.02-0223fje.

11. Karin M, Chang L. AP-1--glucocorticoid receptor crosstalk taken to a higher level. *J Endocrinol* 2001;169(3):447-51. doi: 10.1677/joe.0.1690447.

12. Barbian ME, Owens JA, Naudin CR, Denning PW, Patel RM, Jones RM. Butyrate supplementation to pregnant mice elicits cytoprotection against colonic injury in the offspring. *Pediatr Res* 2021; doi: 10.1038/s41390-021-01767-1.

13. Jostins L, Ripke S, Weersma RK, Duerr RH, McGovern DP, Hui KY, et al. Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* 2012;491(7422):119-24. doi: 10.1038/nature11582.

14. Liu JZ, van Sommeren S, Huang H, Ng SC, Alberts R, Takahashi A, et al. Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. *Nat Genet* 2015;47(9):979-986. doi: 10.1038/ng.3359.

15. Ye BD, McGovern DPB. Genetic variation in IBD: progress, clues to pathogenesis and possible clinical utility. *Exp Rev Clin Immunol* 2016;12(10):1091-107. doi: 10.1080/1744666X.2016.1184972.

16. Di Narzo AF, Peters LA, Argmann C, Stojmirovic A, Perrigoue J, Li K, et al. Blood and Intestine eQTLs from an Anti-TNF-Resistant Crohn's Disease Cohort Inform IBD Genetic Association Loci. *Clin Transl Gastroenterol* 2016;7(6):e177. doi: 10.1038/ctg.2016.34.

17. Povoleri GAM, Nova-Lamperti E, Scottà C, Fanelli G, Chen YC, Becker PD, et al. Human retinoic acid-regulated CD161+ regulatory T cells support wound repair in intestinal mucosa. *Nat Immunol* 2018;19(12):1403-1414. doi: 10.1038/s41590-018-0230-z.

18. Wang ZL, Wang YD, Wang K, Li JA, Li L. KFL2 participates in the development of ulcerative colitis through inhibiting inflammation via regulating cytokines. *Eur Rev Med Pharmacol Sci* 2018;2(15):4941-4948. doi: 10.26355/eurrev_201808_15633.

19. Banerjee SS, Feinberg MW, Watanabe M, Gray S, Haspel RL, Denkinger DJ, et al. The Krüppel-like factor KLF2 inhibits peroxisome proliferator-activated receptor-gamma expression and adipogenesis. *J Biol Chem* 2003;278(4):2581-4. doi: 10.1074/jbc.M210859200.

20. Banerjee SS, Lin Z, Atkins GB, Greif DM, Rao RM, Kumar A, et al. KLF2 Is a novel transcriptional regulator of endothelial proinflammatory activation. *J Exp Med* 2004;199(10):1305-15. doi: 10.1084/jem.20031132.

21. Pabbisetty SK, Rabacal W, Volanakis EJ, Parekh VV, Olivares-Villagómez D, Cendron D, et al. Peripheral tolerance can be modified by altering KLF2-regulated Treg migration. *Proc Natl Acad Sci U S A* 2016;113(32):E4662-70. doi: 10.1073/pnas.1605849113.

22. Szigeti R, Pangas SA, Nagy-Szakal D, Dowd SE, Shulman RJ, Olive AP, et al. SMAD4 haploinsufficiency associates with augmented colonic inflammation in select humans and mice. *Ann Clin Lab Sci* 2012;42(4):401-8.

23. Means AL, Freeman TJ, Zhu J, Woodbury LG, Marincola-Smith P, Wu C, et al. Epithelial Smad4 Deletion Up-Regulates Inflammation and Promotes Inflammation-Associated Cancer. *Cell Mol Gastroenterol Hepatol* 2018;6(3):257-276. doi: 10.1016/j.jcmgh.2018.05.006.

24. Wang Y, Xue Q, Zheng Q, Jin Y, Shen X, Yang M, et al. SMAD4 mutation correlates with poor prognosis in non-small cell lung cancer. *Lab Invest* 2021;101(4):463-476. doi: 10.1038/s41374-020-00517-x.

25. Mizuno T, Cloyd JM, Vicente D, Omichi K, Chun YS, Kopetz SE, et al. SMAD4 gene mutation predicts poor prognosis in patients undergoing resection for colorectal liver metastases. *Eur J Surg Oncol* 2018;44(5):684-692. doi: 10.1016/j.ejso.2018.02.247.

26. Miyaki M, Iijima T, Konishi M, Sakai K, Ishii A, Yasuno M, et al. Higher frequency of Smad4 gene mutation in human colorectal cancer with distant metastasis. *Oncogene* 1999;18(20):3098-103. doi: 10.1038/sj.onc.1202642.

27. Lin LH, Chang KW, Cheng HW, Liu CJ. SMAD4 Somatic Mutations in Head and Neck Carcinoma Are Associated With Tumor Progression. *Front Oncol* 2019;9:1379. doi: 10.3389/fonc.2019.01379.

28. Gart E, van Duyvenvoorde W, Toet K, Caspers MPM, Verschuren L, Nielsen MJ, et al. Butyrate Protects against Diet-Induced NASH and Liver Fibrosis and Suppresses Specific Non-Canonical TGF-β Signaling Pathways in Human Hepatic Stellate Cells. *Biomedicines* 2021;9(12):1954. doi: 10.3390/biomedicines9121954.

29. Dong S, Tian Q, Zhu T, Wang K, Lei G, Liu Y, et al. SLC39A5 dysfunction impairs extracellular matrix synthesis in high myopia pathogenesis. *J Cell Mol Med* 2021;25(17):8432-8441. doi: 10.1111/jcmm.16803.

30. Ellinghaus D, Ellinghaus E, Nair RP, Stuart PE, Esko T, Metspalu A, et al. Combined analysis of genome-wide association studies for Crohn disease and psoriasis identifies seven shared susceptibility loci. *Am J Human Genet* 2012;90(4):636-47. doi: 10.1016/j.ajhg.2012.02.020.

31. Ye BD, Choi H, Hong M, Yun WJ, Low HQ, Haritunians T, et al. Identification of Ten Additional Susceptibility Loci for Ulcerative Colitis Through Immunochip Analysis in Koreans. *Inflamm Bowel Dis* 2016;22(1):13-9. doi: 10.1097/MIB.0000000000000584.

32. Parmar AS, Lappalainen M, Paavola-Sakki P, Halme L, Färkkilä M, Turunen U, et al. Association of celiac disease genes with inflammatory bowel disease in Finnish and Swedish patients. *Genes Immun* 2012;13(6):474-80. doi: 10.1038/gene.2012.21.

33. Verma SC, Mahadevan S. The chbG gene of the chitobiose (chb) operon of Escherichia coli encodes a chitooligosaccharide deacetylase. *J Bacteriol* 2012;194(18):4959-71. doi: 10.1128/JB.00533-12.

34. Kim EJ, Park MK, Kang GJ, Byun HJ, Kim HJ, Yu L, et al. YDJC Induces Epithelial-Mesenchymal Transition via Escaping from Interaction with CDC16 through Ubiquitination of PP2A. *J Oncol* 2019;2019:3542537. doi: 10.1155/2019/3542537.

35. Kim EJ, Park MK, Byun HJ, Kang GJ, Yu L, Kim HJ, et al. YdjC chitooligosaccharide deacetylase homolog induces keratin reorganization in lung cancer cells: involvement of

1681 interactions between YDJC and CDC16. *Oncotarget* 2018;9(33):22915-22928. doi:
1682 10.18632/oncotarget.25145.

1683 36. Parker BJ, Wearsch PA, Veloo ACM, Rodriguez-Palacios A. The Genus Alistipes:
1684 Gut Bacteria With Emerging Implications to Inflammation, Cancer, and Mental Health.
1685 *Front Immunol* 2020;11:906. doi: 10.3389/fimmu.2020.00906.

1686 37. Dziarski R, Park SY, Kashyap DR, Dowd SE, Gupta D. Pglyrp-Regulated Gut
1687 Microflora Prevotella falsenii, Parabacteroides distasonis and Bacteroides eggerthii
1688 Enhance and Alistipes finegoldii Attenuates Colitis in Mice. *PLoS One*
1689 2016;11(1):e0146162. doi: 10.1371/journal.pone.0146162.

1690 38. Butera A, Di Paola M, Pavarini L, Strati F, Pindo M, Sanchez M, et al. Nod2
1691 Deficiency in mice is Associated with Microbiota Variation Favouring the Expansion of
1692 mucosal $CD4^+$ LAP^+ Regulatory Cells. *Sci Rep* 2018;8(1):14241. doi: 10.1038/s41598-
1693 018-32583-z.

1694 39. Li J, Sung CY, Lee N, Ni Y, Pihlajamäki J, Panagiotou G, et al. Probiotics modulated
1695 gut microbiota suppresses hepatocellular carcinoma growth in mice. *Proc Natl Acad Sci*
1696 *U S A* 2016;113(9):E1306-15. doi: 10.1073/pnas.1518189113.

1697 40. Rodriguez-Palacios A, Kodani T, Kaydo L, Pietropaoli D, Corridoni D, Howell S, et
1698 al. Stereomicroscopic 3D-pattern profiling of murine and human intestinal inflammation
1699 reveals unique structural phenotypes. *Nat Commun* 2015;6:7577. doi:
1700 10.1038/ncomms8577.

1701 41. Rodriguez-Palacios A, Harding A, Menghini P, Himmelman C, Retuerto M,
1702 Nickerson KP, et al. The Artificial Sweetener Splenda Promotes Gut Proteobacteria,
1703 Dysbiosis, and Myeloperoxidase Reactivity in Crohn's Disease-Like Ileitis. *Inflamm*
1704 *Bowel Dis* 2018;24(5):1005-1020. doi: 10.1093/ibd/izy060.

1705 42. Grondin JA, Kwon YH, Far PM, Haq S, Khan WI. Mucins in Intestinal Mucosal
1706 Defense and Inflammation: Learning From Clinical and Experimental Studies. *Front*
1707 *Immunol* 2020;11:2054. doi: 10.3389/fimmu.2020.02054.

1708 43. Das S, Rachagani S, Sheinin Y, Smith LM, Gurumurthy CB, Roy HK, et al. Mice
1709 deficient in Muc4 are resistant to experimental colitis and colitis-associated colorectal
1710 cancer. *Oncogene* 2016;35(20):2645-54. doi: 10.1038/onc.2015.327.

1711 44. Lam YY, Ha CWY, Campbell CR, Mitchell AJ, Dinudom A, Oscarsson J, et al.
1712 Increased gut permeability and microbiota change associate with mesenteric fat
1713 accumulation and metabolic dysfunction in diet-induced obese mice. *PLoS One*
1714 2012;7(3):e34233. doi: 10.1371/journal.pone.0034233.

1715 45. Peng Y, Yan Y, Wan P, Chen D, Ding Y, Ran L, et al. Gut microbiota modulation
1716 and anti-inflammatory properties of anthocyanins from the fruit of Lycium ruthenicum
1717 Murray in dextran sodium sulfate-induced colitis in mice. *Free Radic Biol Med*
1718 2019;136:96-108. doi: 10.1016/j.freeradbiomed.2019.04.005.

1719 46. Golubeva AV, Crampton S, Desbonnet L, Edge D, O'Sullivan O, Lomasney KW, et
1720 al. Prenatal stress-induced alterations in major physiological systems correlate with gut
1721 microbiota composition in adulthood. *Psychoneuroendocrinology* 2015;60:58-74. doi:
1722 10.1016/j.psyneuen.2015.06.002.

1723 47. Metwaly A, Dunkel A, Waldschmitt N, Raj ACD, Lagkouvardos I, Corraliza AM, et al.
1724 Integrated microbiota and metabolite profiles link Crohn's disease to sulfur metabolism.
1725 *Nat Commun* 2020;11(1):4322. doi: 10.1038/s41467-020-17956-1.

1726 48. Ryan FJ, Ahern AM, Fitzgerald RS, Laserna-Mendieta EJ, Power EM, Clooney AG,
1727 et al. Colonic microbiota is associated with inflammation and host epigenomic
1728 alterations in inflammatory bowel disease. *Nat Commun* 2020;11(1):1512. doi:
1729 10.1038/s41467-020-15342-5.

1730 49. Chollangi S, Mather T, Rodgers KK, Ash JD. A unique loop structure in oncostatin M
1731 determines binding affinity toward oncostatin M receptor and leukemia inhibitory factor
1732 receptor. *J Biol Chem* 2012;287(39):32848-59. doi: 10.1074/jbc.M112.387324.

1733 50. Verstockt S, Verstockt B, Machiels K, Vancamelbeke M, Ferrante M, Cleynen I, et
1734 al. Oncostatin M Is a Biomarker of Diagnosis, Worse Disease Prognosis, and
1735 Therapeutic Nonresponse in Inflammatory Bowel Disease. *Inflamm Bowel Dis*
1736 2021;27(10):1564-1575. doi: 10.1093/ibd/izab032.

51. West NR, Hegazy AN, Owens BMJ, Bullers SJ, Linggi B, Buonocore S, et al. Oncostatin M drives intestinal inflammation and predicts response to tumor necrosis factor-neutralizing therapy in patients with inflammatory bowel disease. *Nat Med* 2017;23(5):579-589. doi: 10.1038/nm.4307.

52. Place SP. Single-point mutation in a conserved TPR domain of Hip disrupts enhancement of glucocorticoid receptor signaling. *Cell Stress Chaperones* 2011;16(4):469-74. doi: 10.1007/s12192-010-0254-2.

53. Bai R, Shi Z, Zhang JW, Li D, Zhu YL, Zheng S. ST13, a proliferation regulator, inhibits growth and migration of colorectal cancer cell lines. *J Zhejiang Univ Sci B* 2012;13(11):884-893. doi: 10.1631/jzus.B1200037.

54. Wang LB, Zheng S, Zhang SZ, Peng JP, Ye F, Fang SC, et al. Expression of ST13 in colorectal cancer and adjacent normal tissues. *World J Gastroenterol* 2005;11(3):336-9. doi: 10.3748/wjg.v11.i3.336.

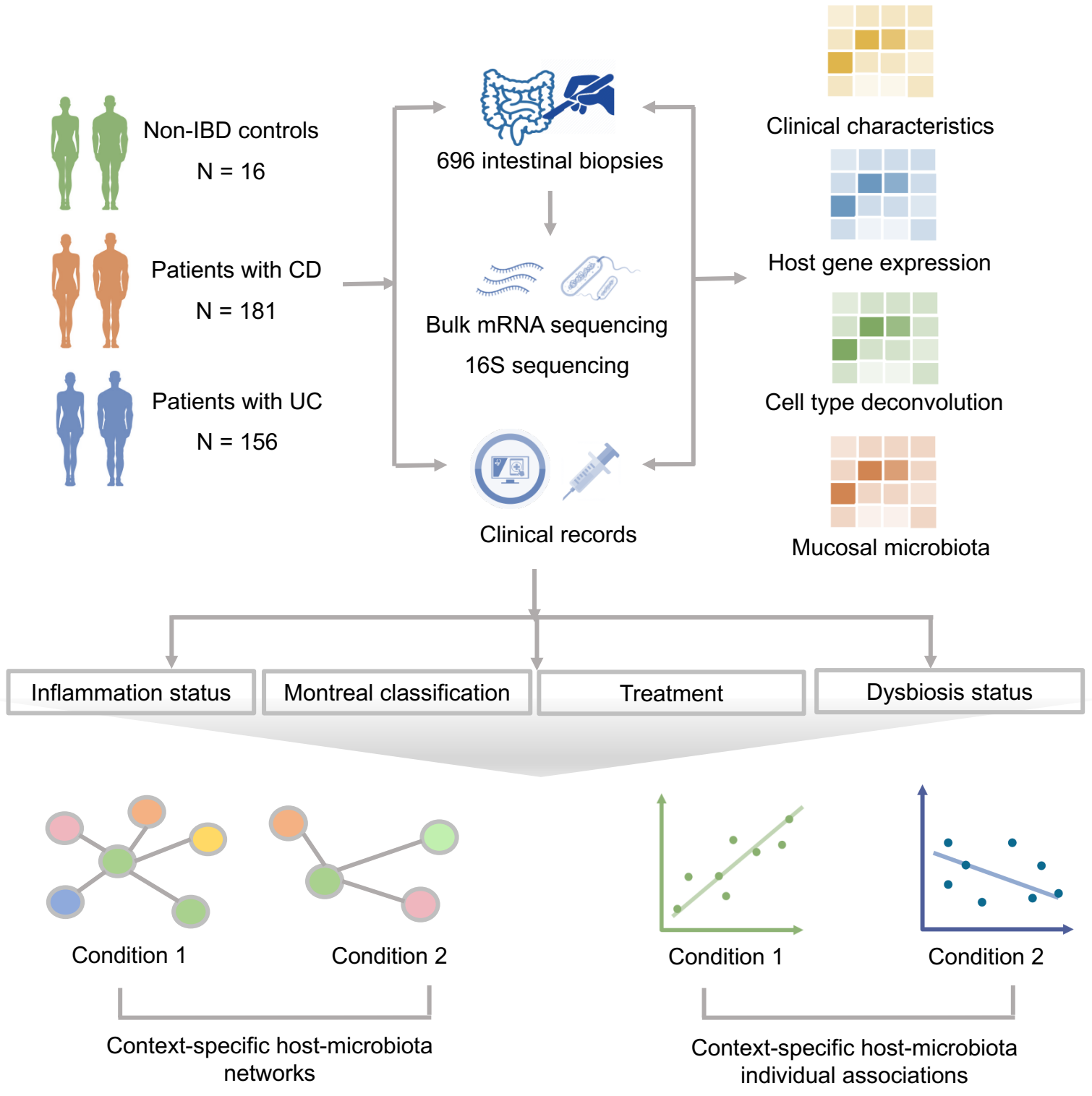
Box 2. Miscellaneous component pairs from sparse-CCA analysis.

The microbial part of the fifth pair of components (component pair 5, P=2.87 x10^{-8}, FDR<0.05) was formed by *Christensenellaceae*, *Ruminococcaceae*, *Lachnospiraceae* (NK4A136 group), Coriobacteria and the genera *Coprococcus* and *Ruminoclostridium*, which are all inversely associated with pathways representing SLC-mediated transmembrane transport (e.g. transport of bile acids and organic acids, metal ions and amine compounds) as well as biological oxidation and fat metabolism pathways including arachidonic acid metabolism and (glycero)phospholipid biosynthesis (**Supplementary Tables S15-S16**).

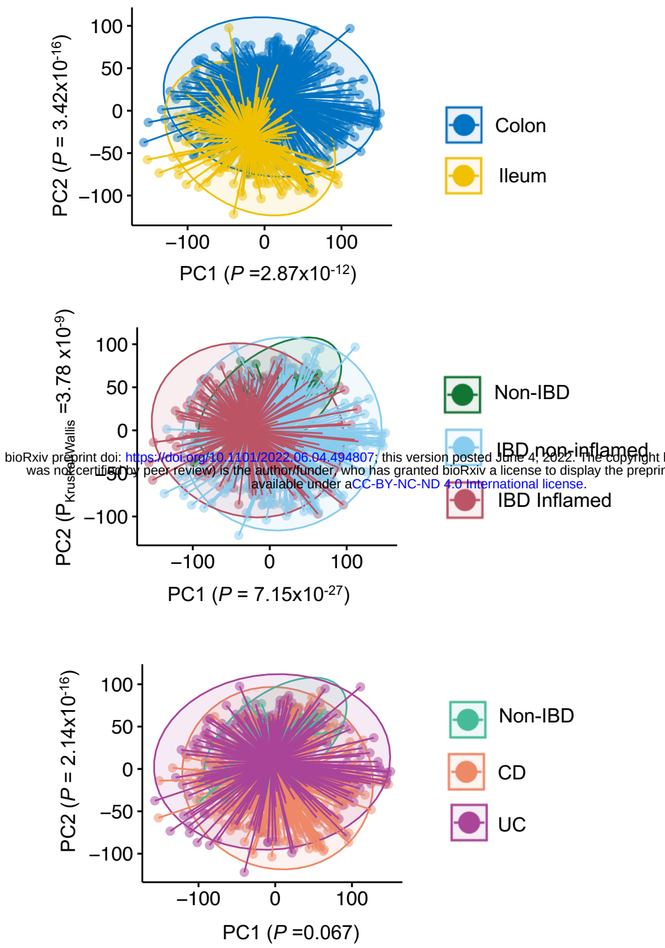
In the sixth pair of components (component pair 9, P=9.65 x10^{-7}, FDR<0.05), the microbial component was primarily composed of bifidobacteria (i.e. order Bifidobacteriales, family *Bifidobacteriaceae* and genus *Bifidobacterium*), which were inversely associated with pathways representing phospholipid synthesis (e.g. phosphatidic acid synthesis) and NR1H2/NR1H3 or liver X receptor (LXR)-mediated

signaling (**Supplementary Tables S17-S18**). NR1H3 (LXR-α) and NR1H2 (LXR-β) are ligand-activated transcription factors stimulated by endogenously produced oxysterols, which are in turn produced by oxidation of cholesterol, enzymatic reactions or alimentary processes [1]. Under physiological conditions, oxysterols are formed proportional to the cellular cholesterol content and thereby stimulate LXRs (acting as cholesterol sensors) to alter gene expression and activate protective mechanisms to prevent cholesterol overload in the cell. This occurs via inhibition of intestinal cholesterol absorption, activation of cholesterol efflux from cells to HDL (via ABCA1 and ABCG1 transporters) and activation of the hepatic conversion of cholesterol to bile acids and stimulation of biliary cholesterol and bile acid excretion. In addition, LXR-agonists enhance *de novo* synthesis of fatty acids by stimulating the expression of the lipogenic transcription factor SREBP-1c, which may result in elevated plasma triglycerides and hepatic steatosis. LXRs are also involved in modulation of innate and adaptive immune responses and regulate diverse aspects of inflammatory gene expression in macrophages. The ability of LXRs to coordinate metabolic and immune response constitutes an attractive therapeutic target for treatment of IBD.

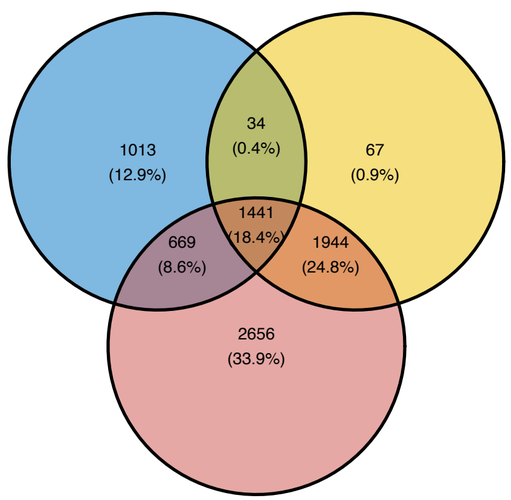
1. Jakobsson T, Treuter E, Gustafsson JA, Steffensen KR. Liver X receptor biology and pharmacology: new pathways, challenges and opportunities. *Trends Pharmacol Sci* 2012;33(7):394-404. doi: 10.1016/j.tips.2012.03.013.


Non-IBD controls
N = 16
Patients with CD
N = 181
Patients with UC
N = 156
696 intestinal biopsies
Bulk mRNA sequencing
16S sequencing
Clinical records
Clinical characteristics
Host gene expression
Cell type deconvolution
Mucosal microbiota
Inflammation status
Montreal classification
Treatment
Dysbiosis status
Condition 1
Condition 2
Condition 1
Condition 2
Context-specific host-microbiota networks
Context-specific host-microbiota individual associations

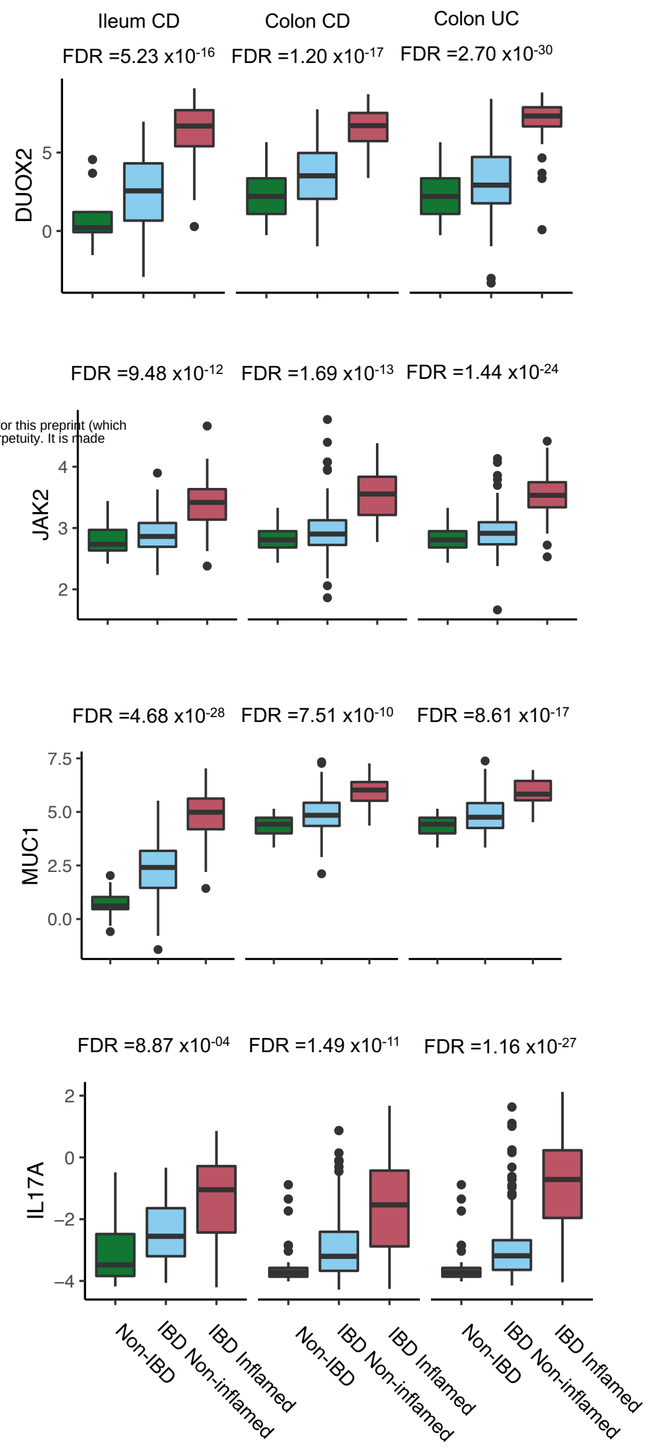
A



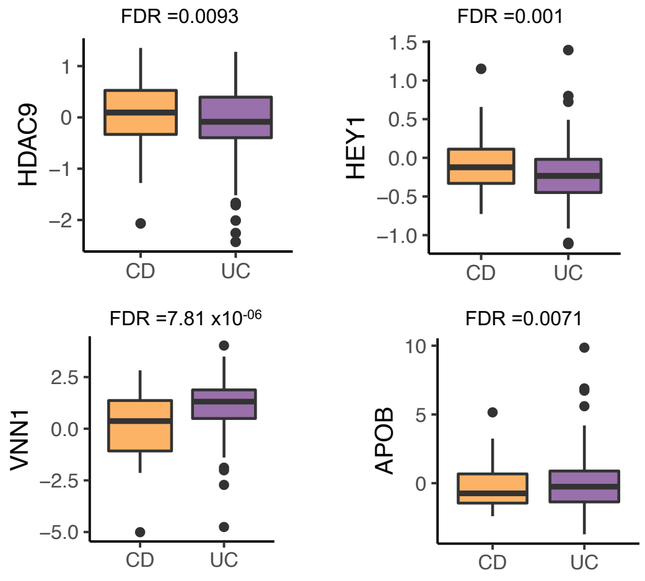
B



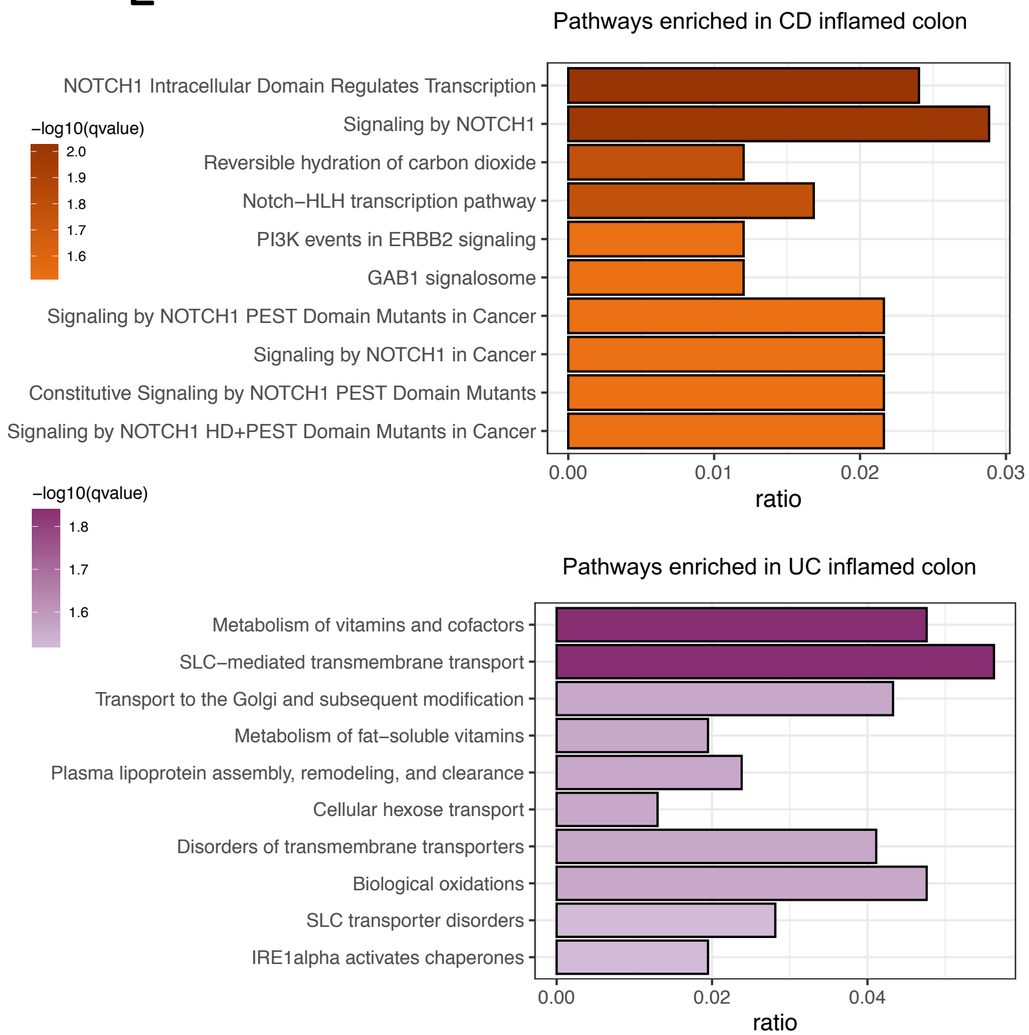
C



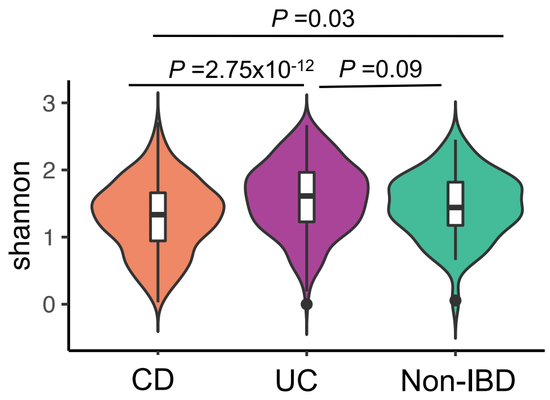
D



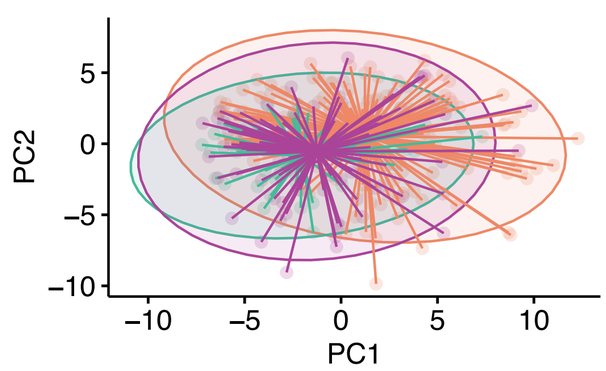
E



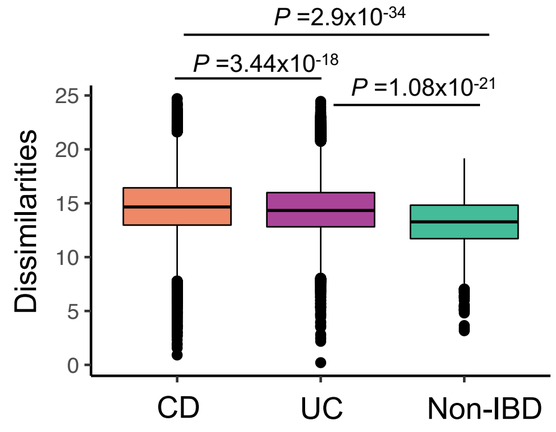
A



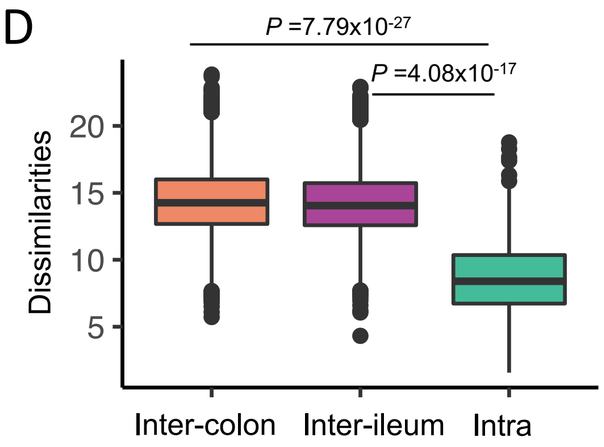
B

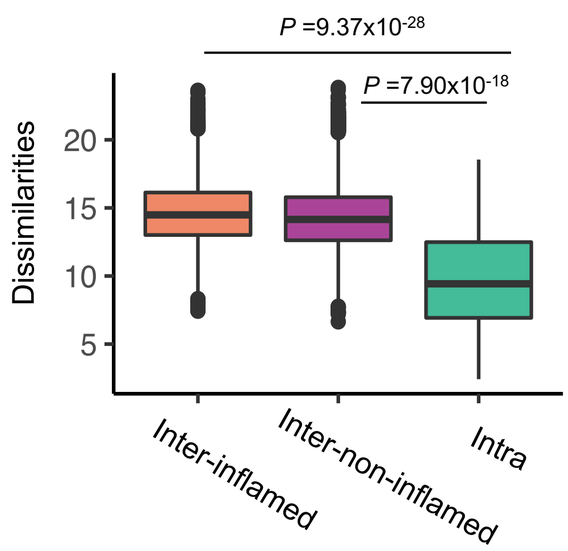


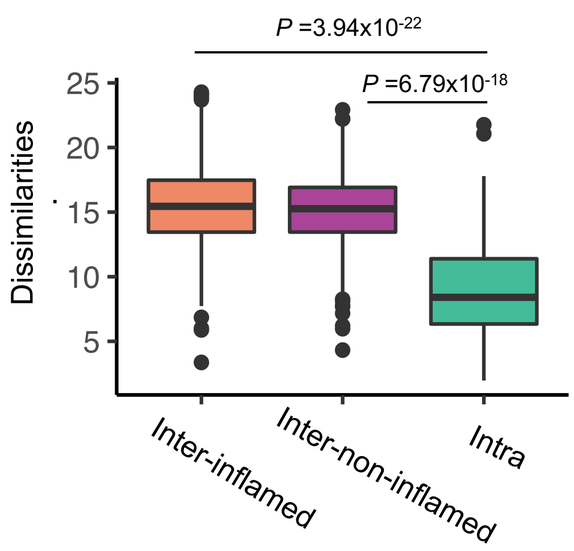
C



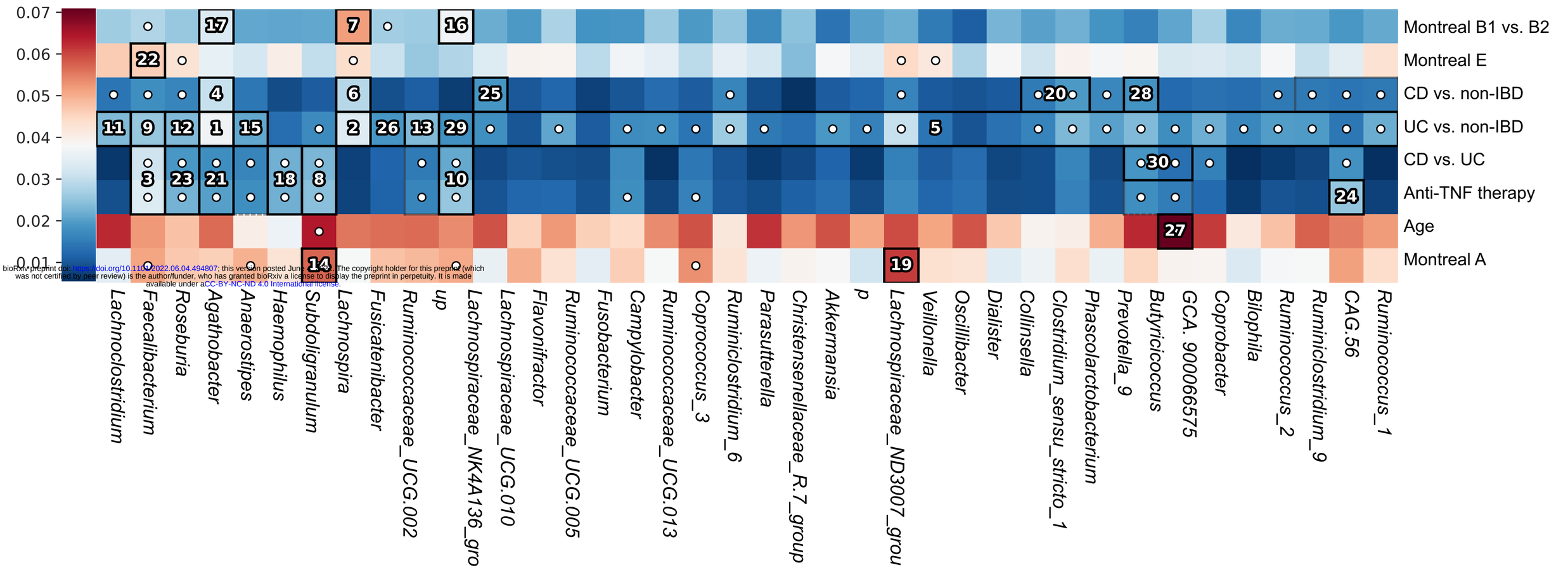
D

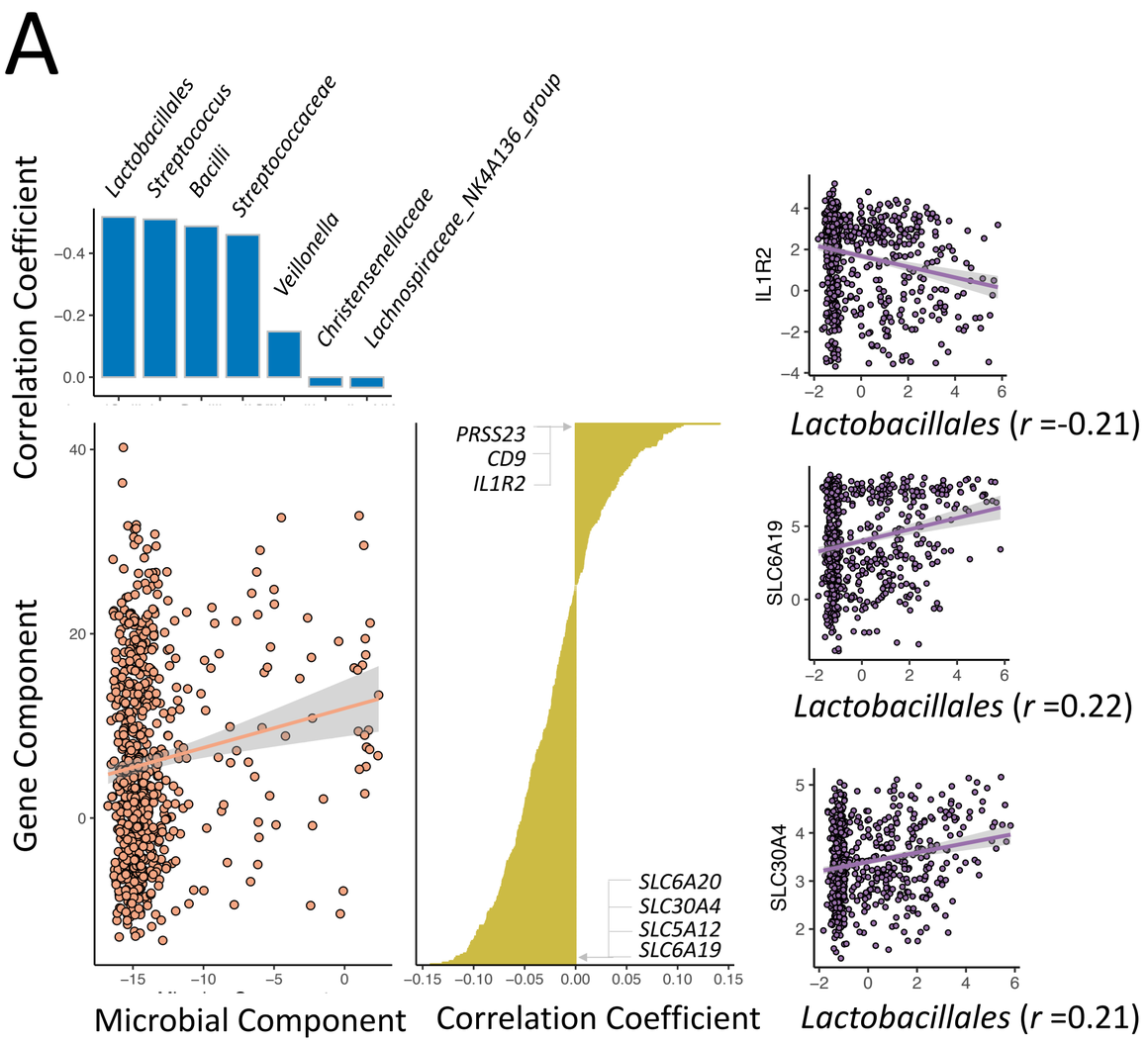


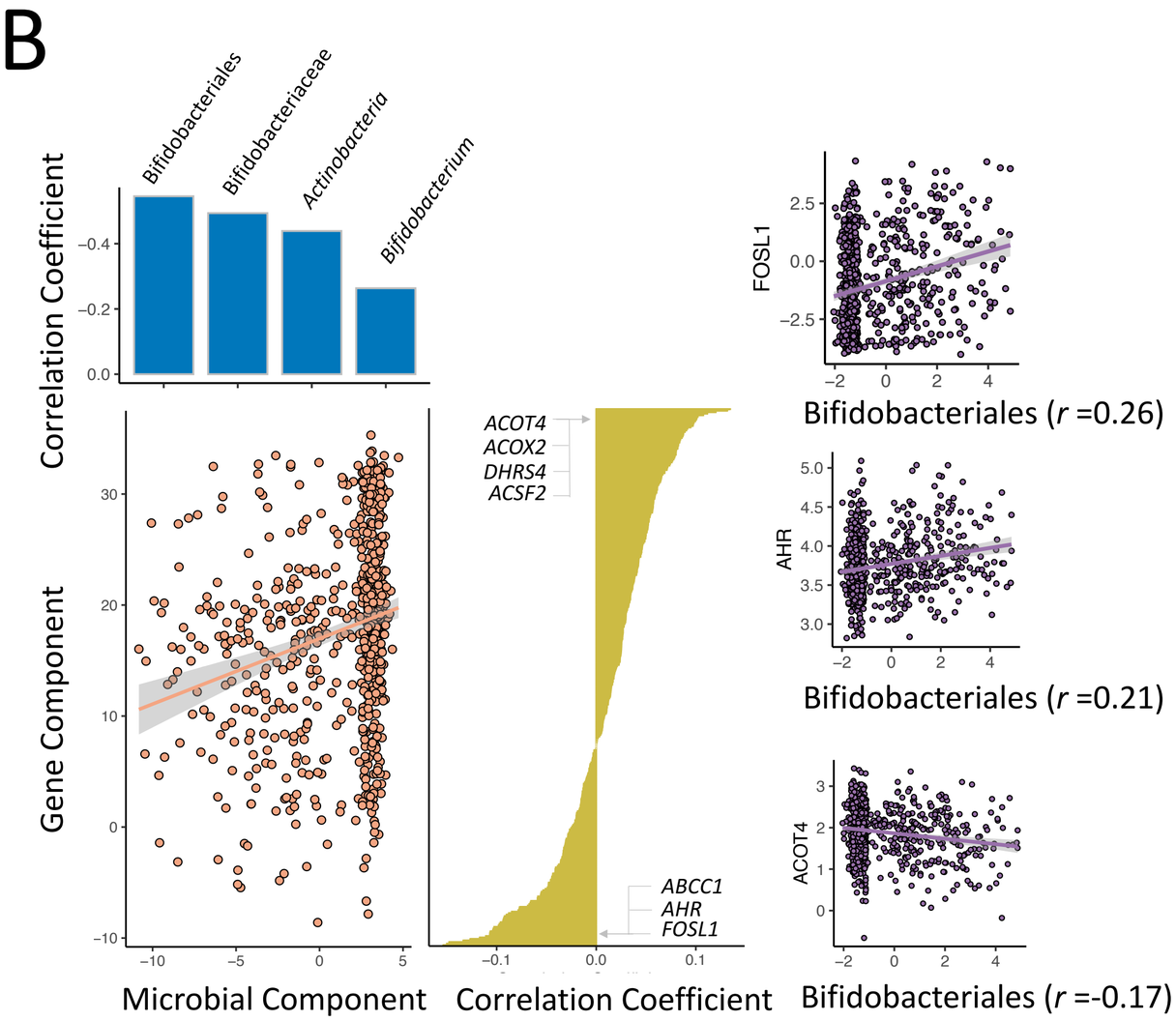


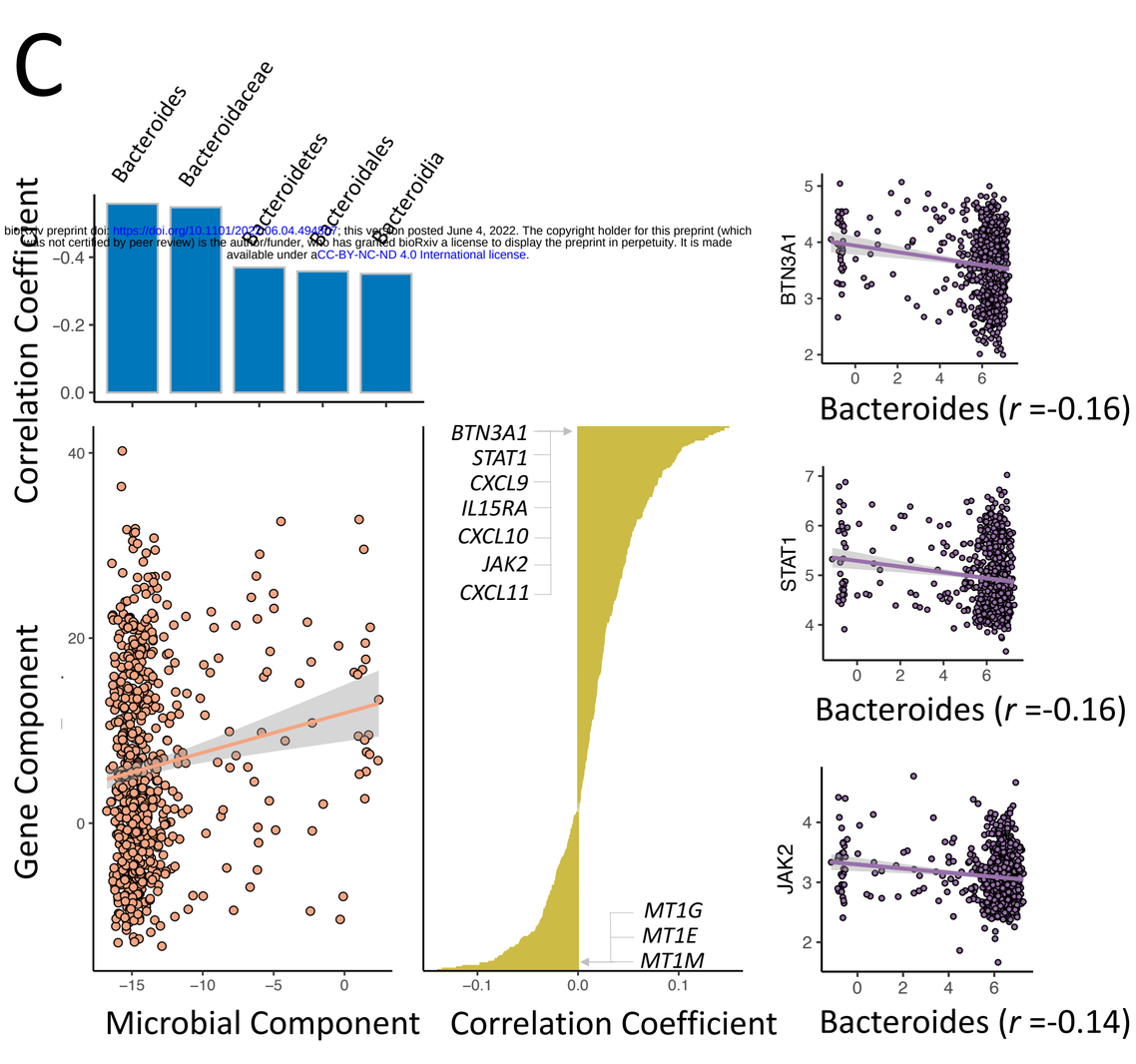


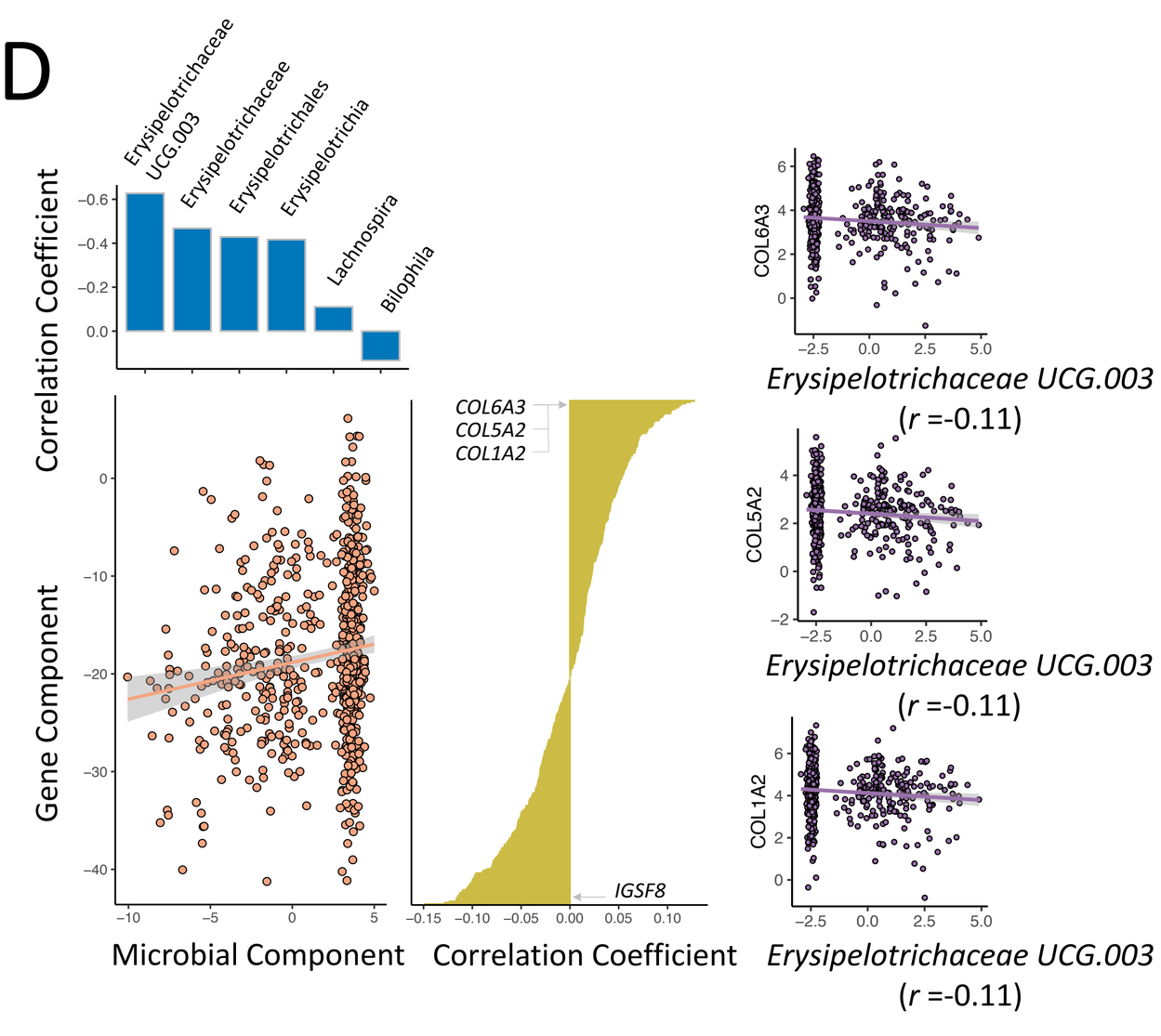
E



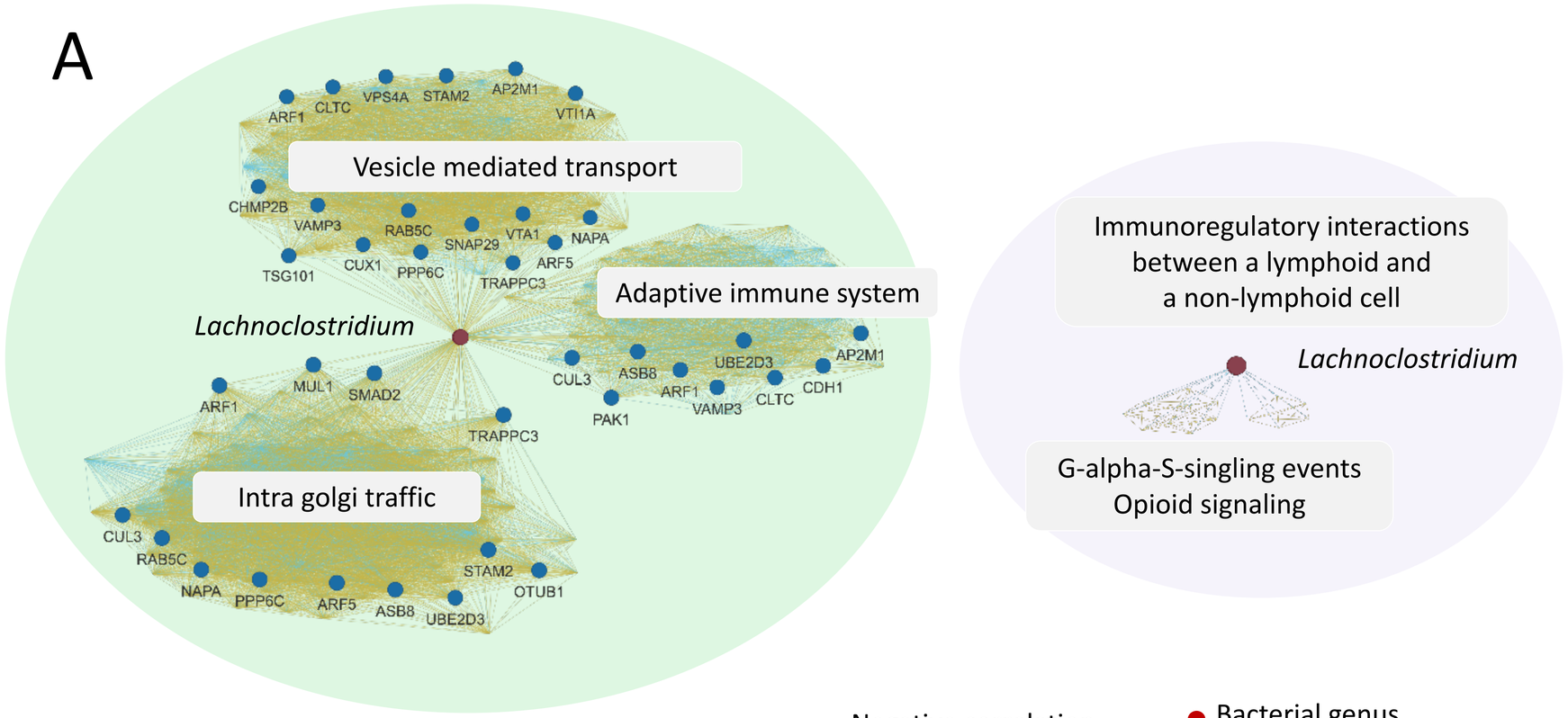




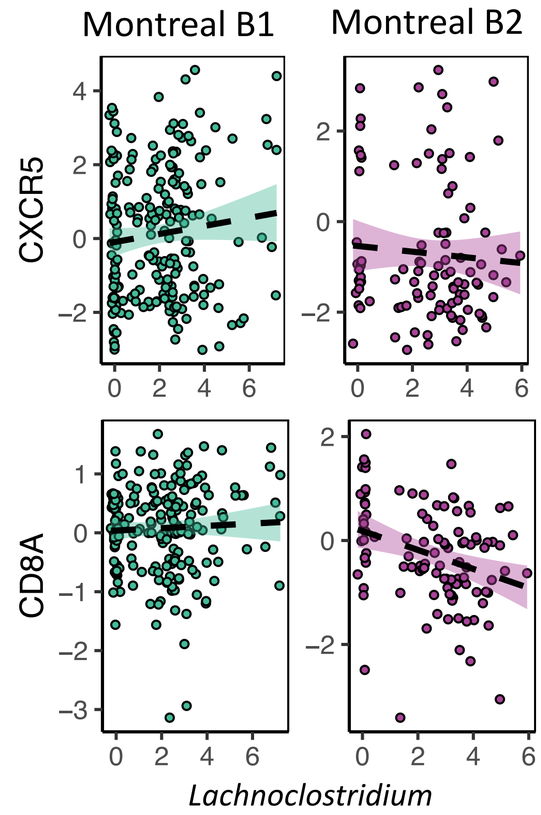




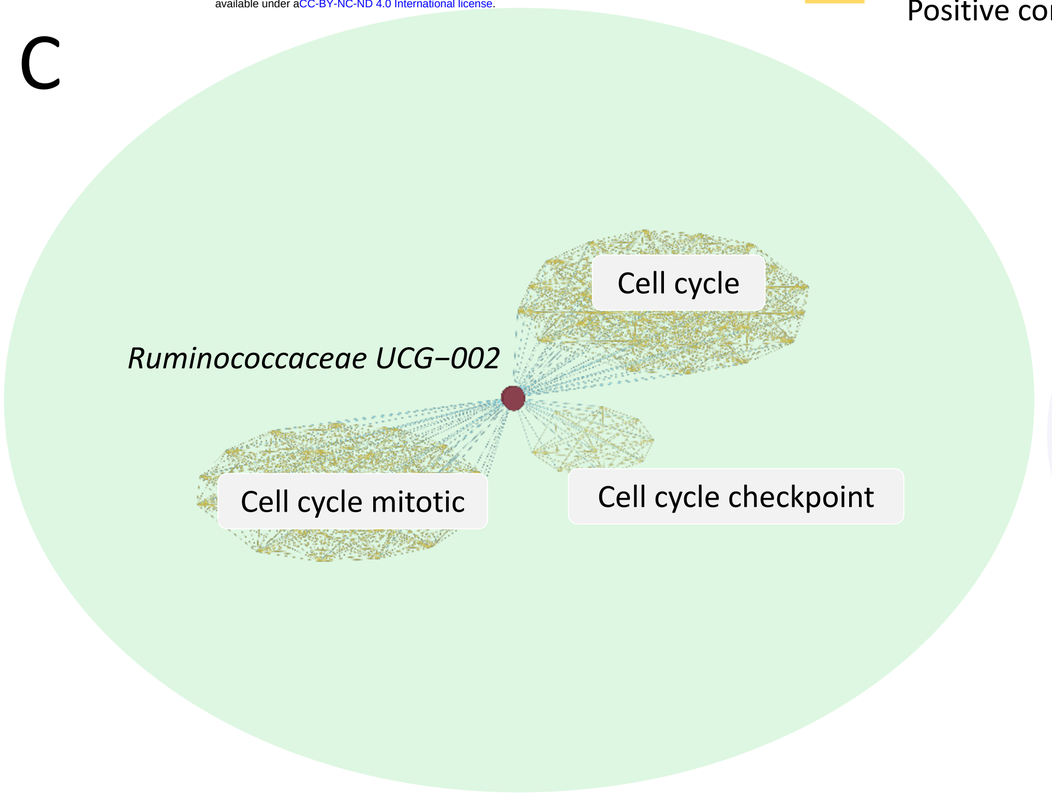
A

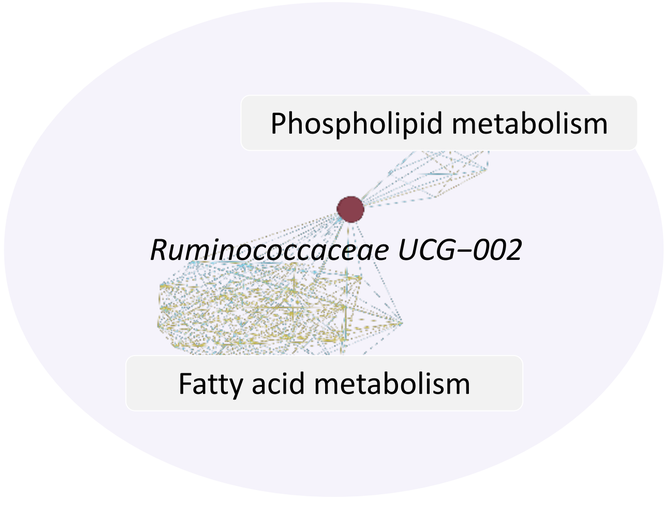


B



C

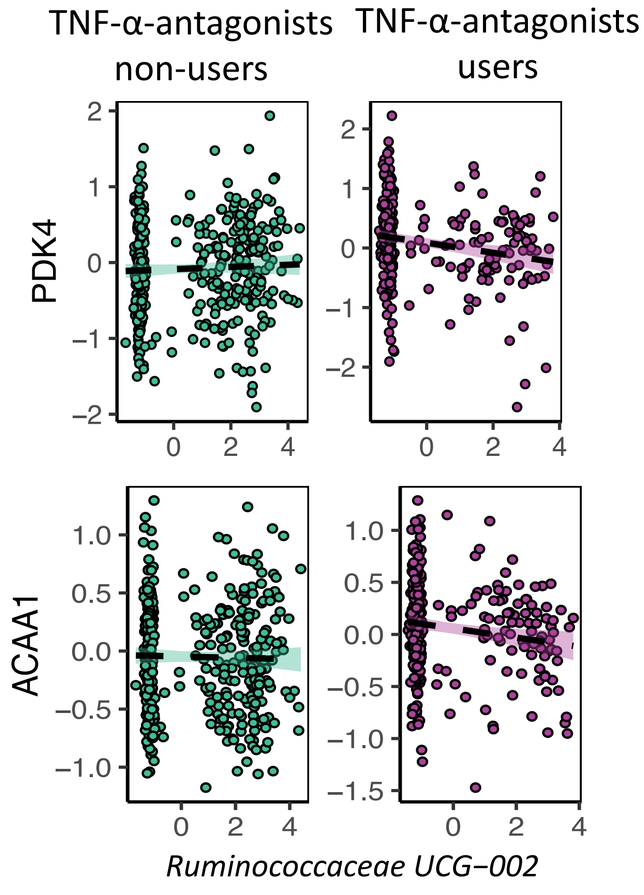


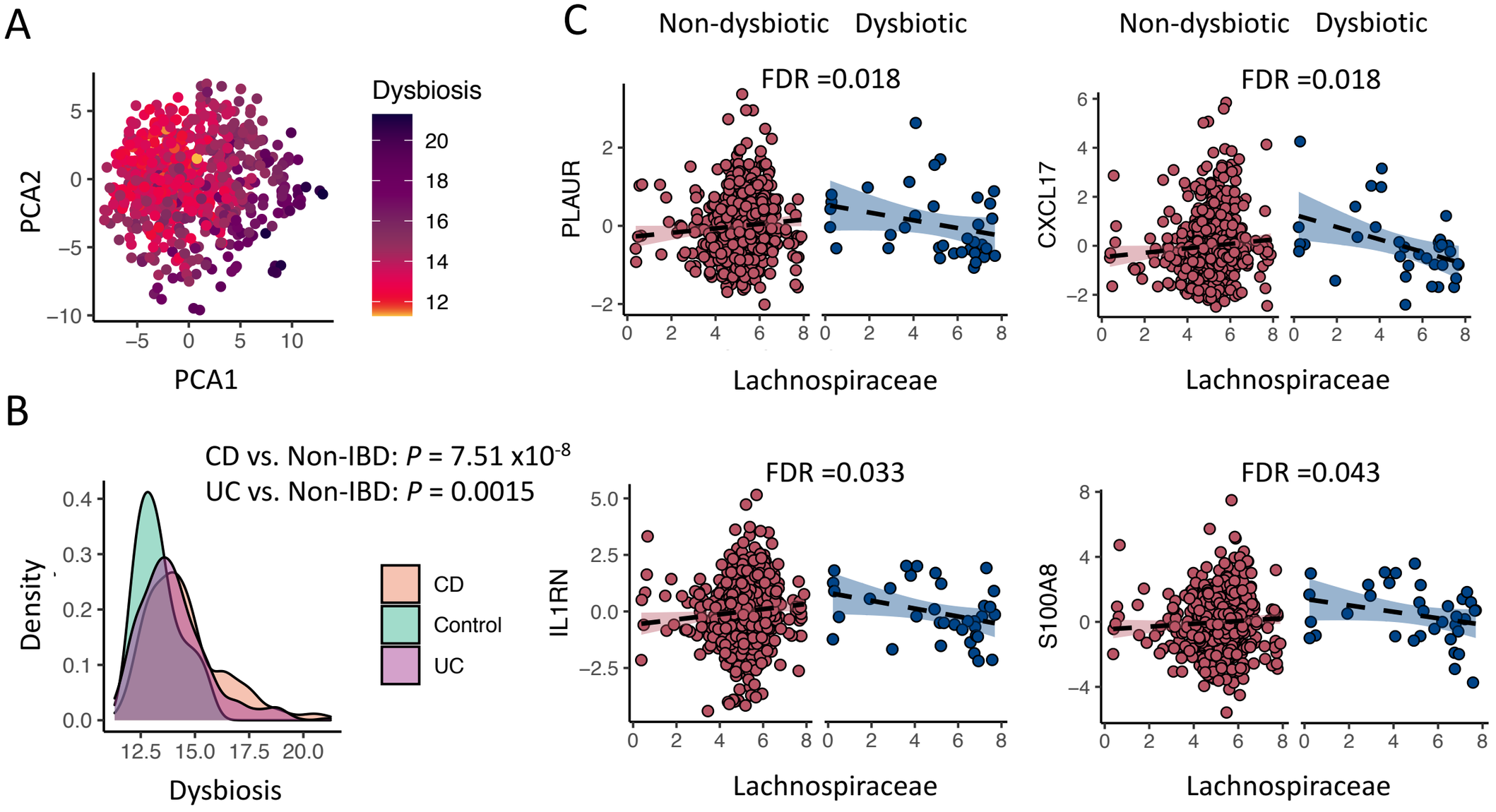


Negative correlation
Positive correlation

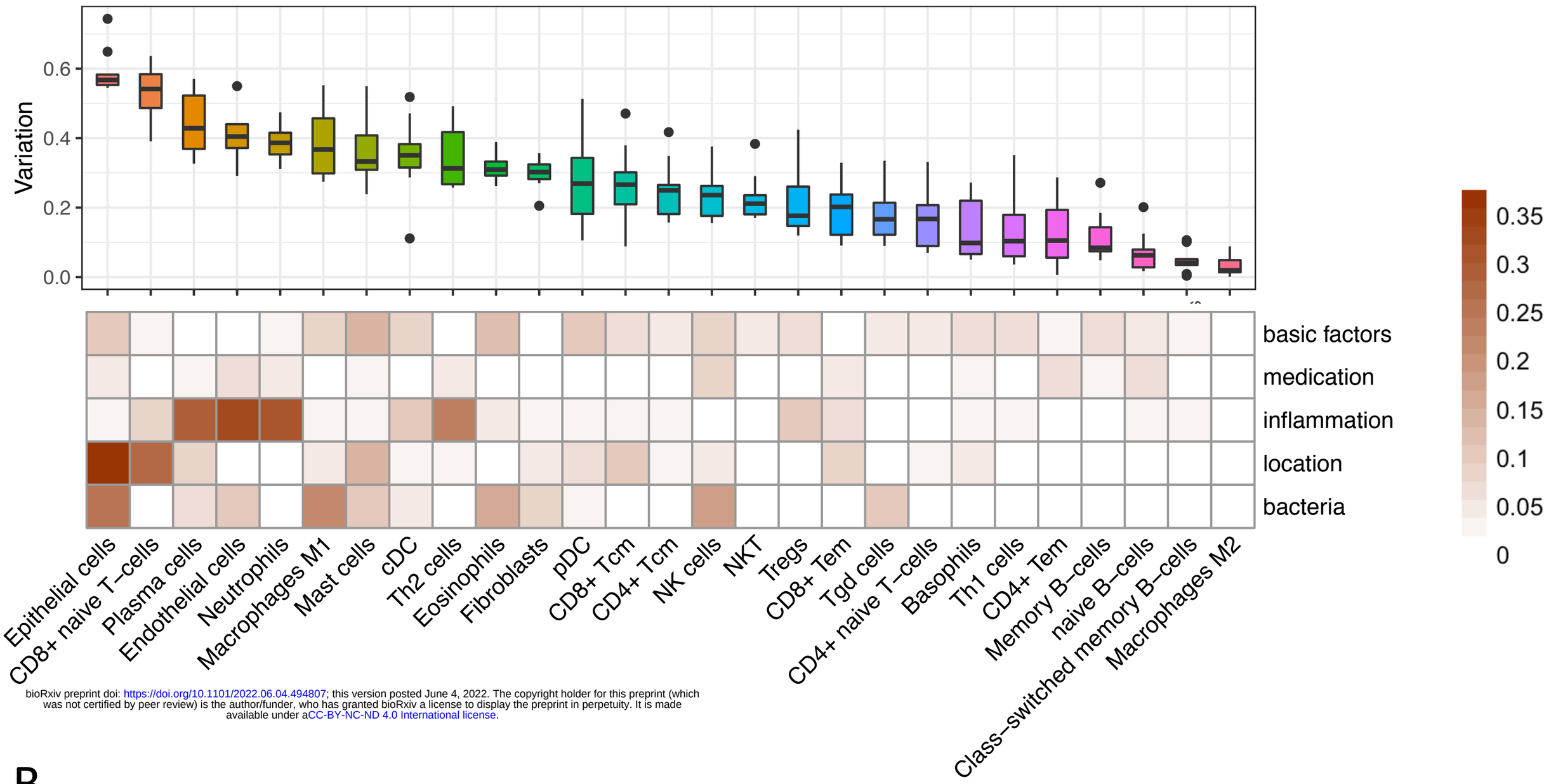
Bacterial genus
Bacteria-associated gene
Hub gene

D

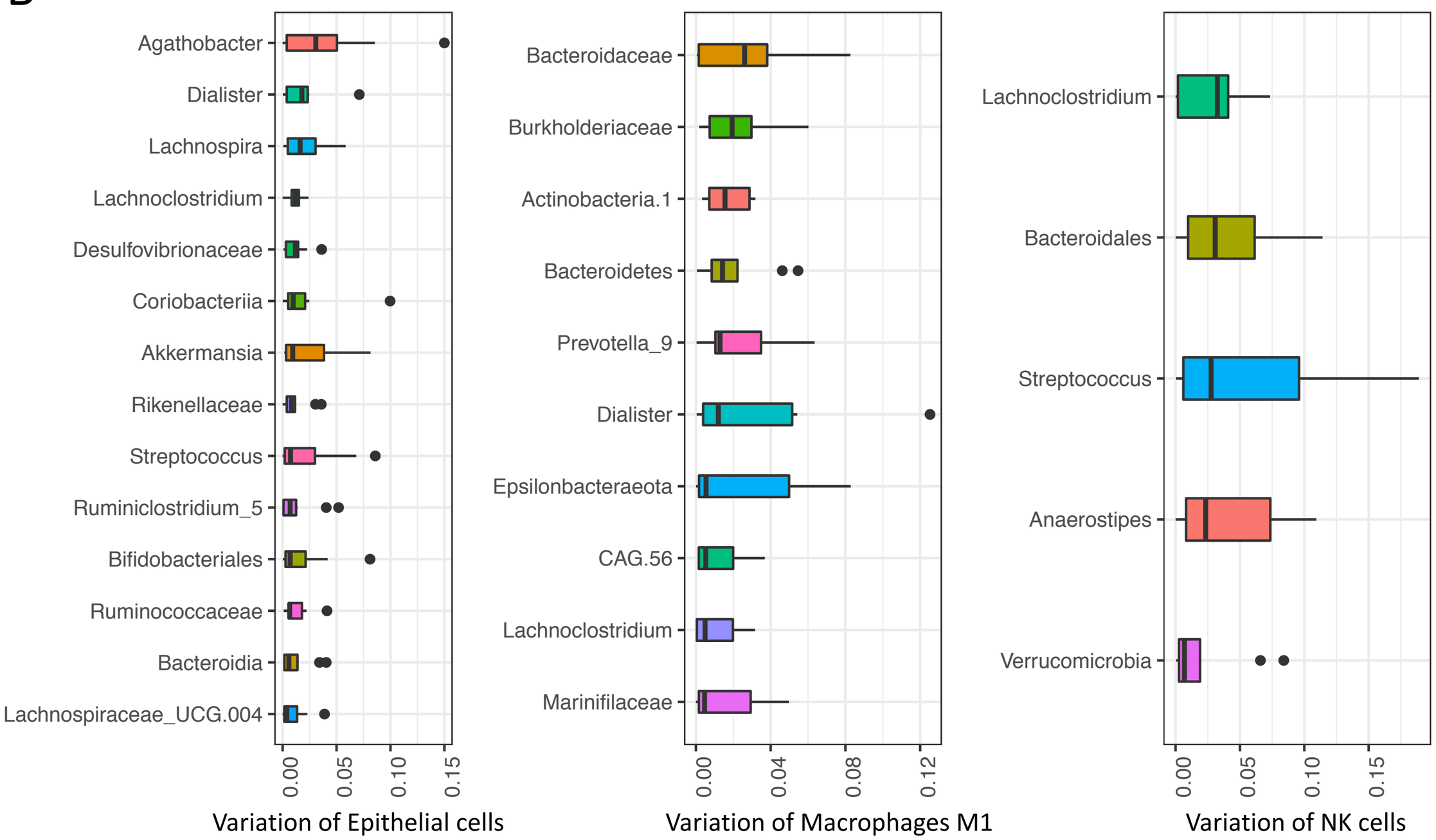



A
PCA2
PCA1
Dysbiosis
20
18
16
14
12
B
CD vs. Non-IBD: P = 7.51 x10-8
UC vs. Non-IBD: P = 0.0015
Density
Dysbiosis
CD
Control
UC
C
Non-dysbiotic
Dysbiotic
FDR =0.018
PLAUR
Lachnospiraceae
FDR =0.018
CXCL17
Lachnospiraceae
FDR =0.033
IL1RN
Lachnospiraceae
FDR =0.043
S100A8
Lachnospiraceae

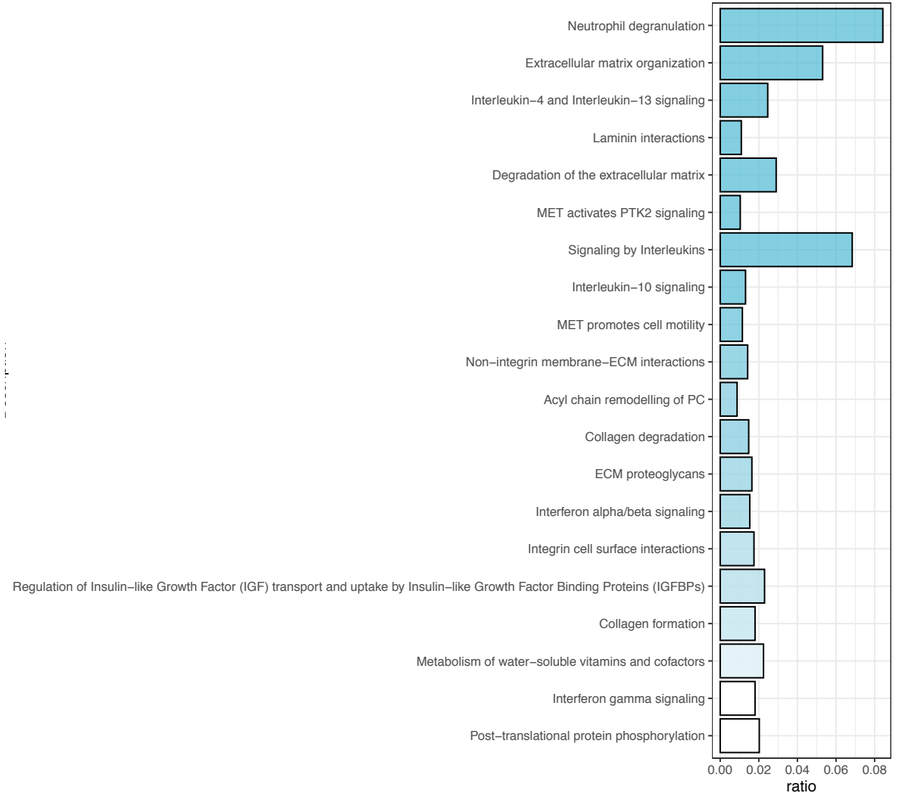
A



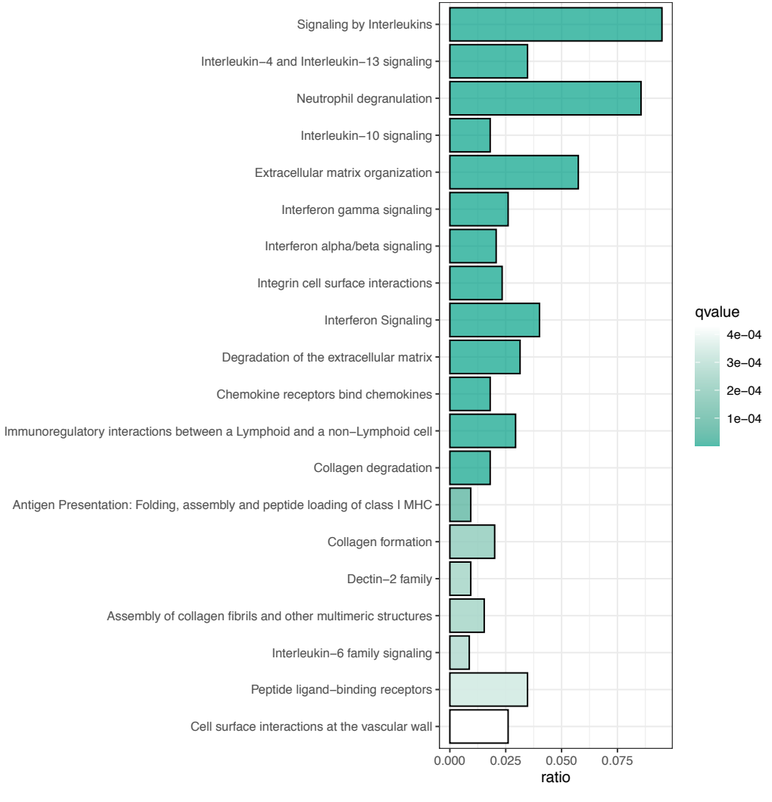
B



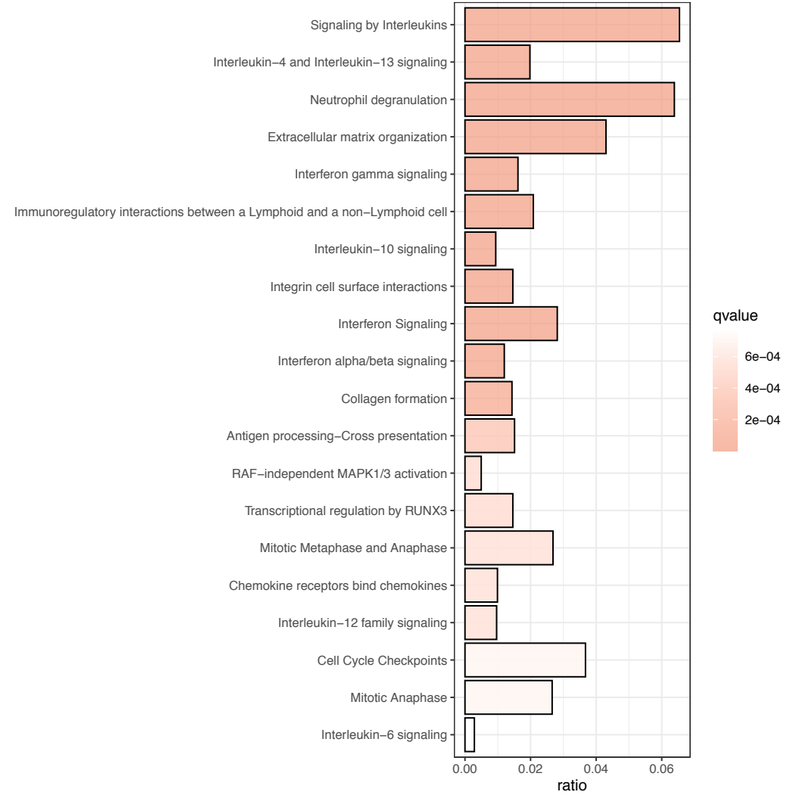
Ileum tissue: non-IBD vs. CDnon vs. CDi



Colon tissue: non-IBD vs. CDnon vs. CDi



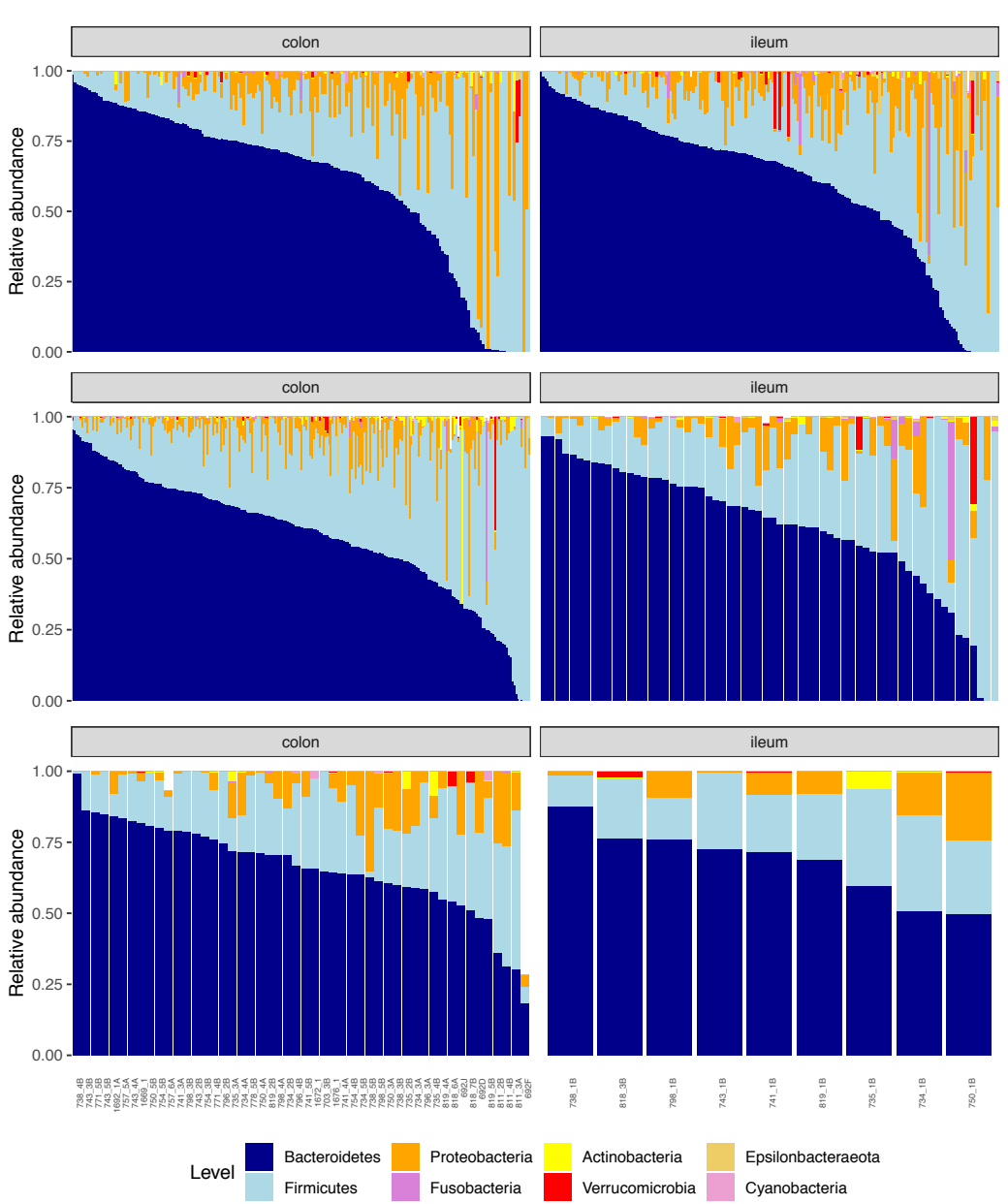
Colon tissue: non-IBD vs. UCnon vs. UCi

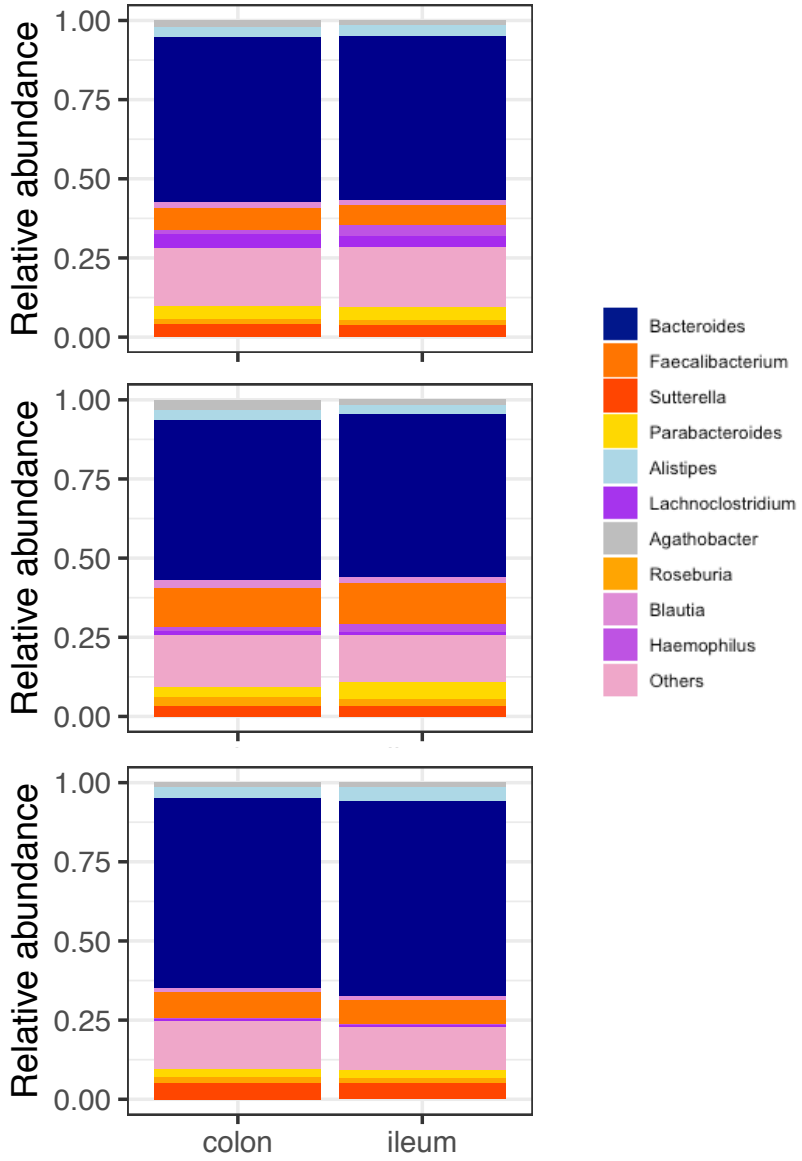


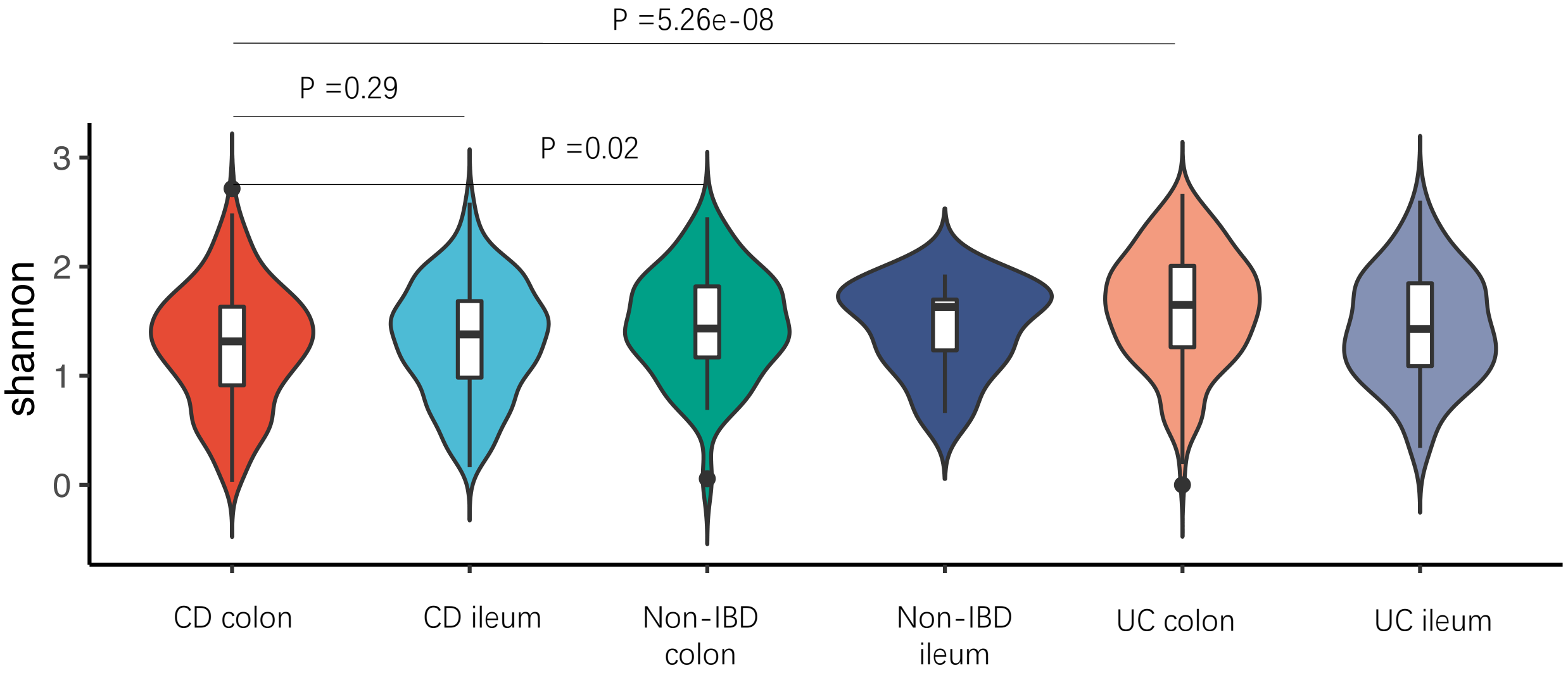
CD

UC

Control

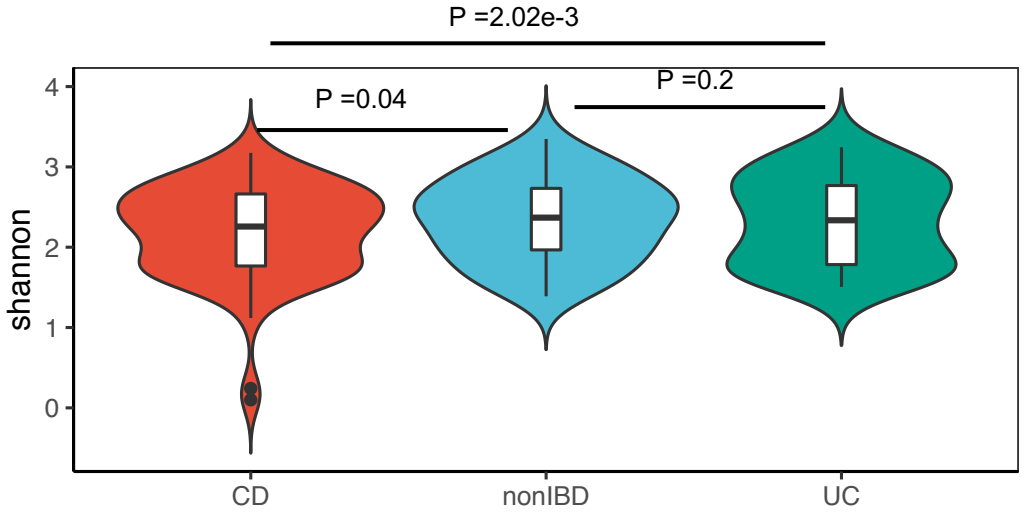




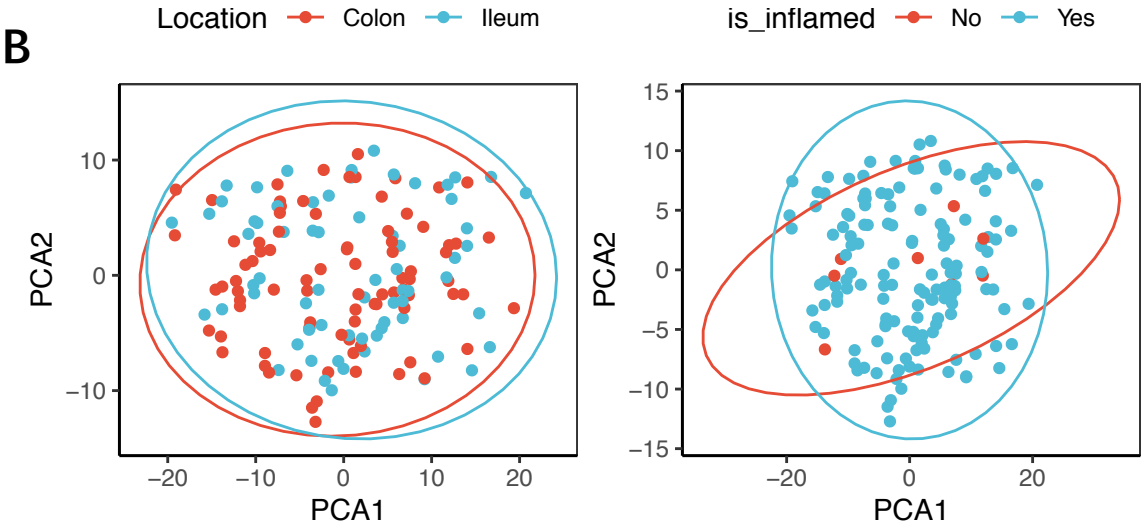

P =5.26e-08
P =0.29
P =0.02
shannon
3
2
1
0
CD colon
CD ileum
Non-IBD colon
Non-IBD ileum
UC colon
UC ileum

CD/UC/nonIBD microbial patterns (HMP2 validation)

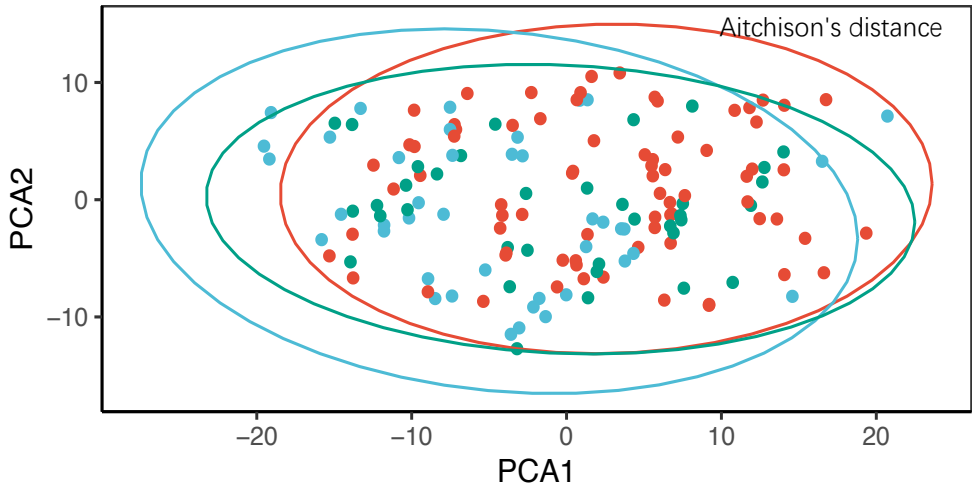
A

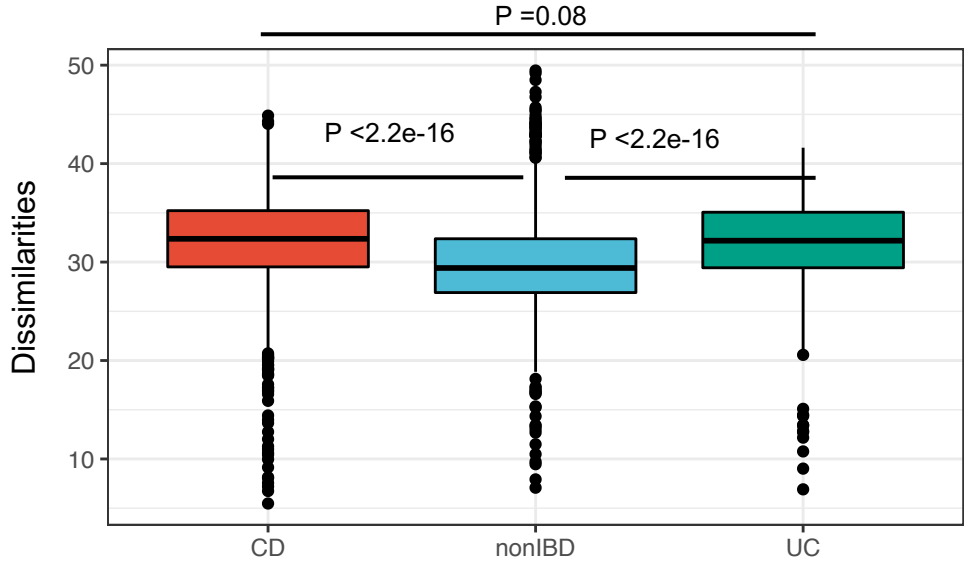


B

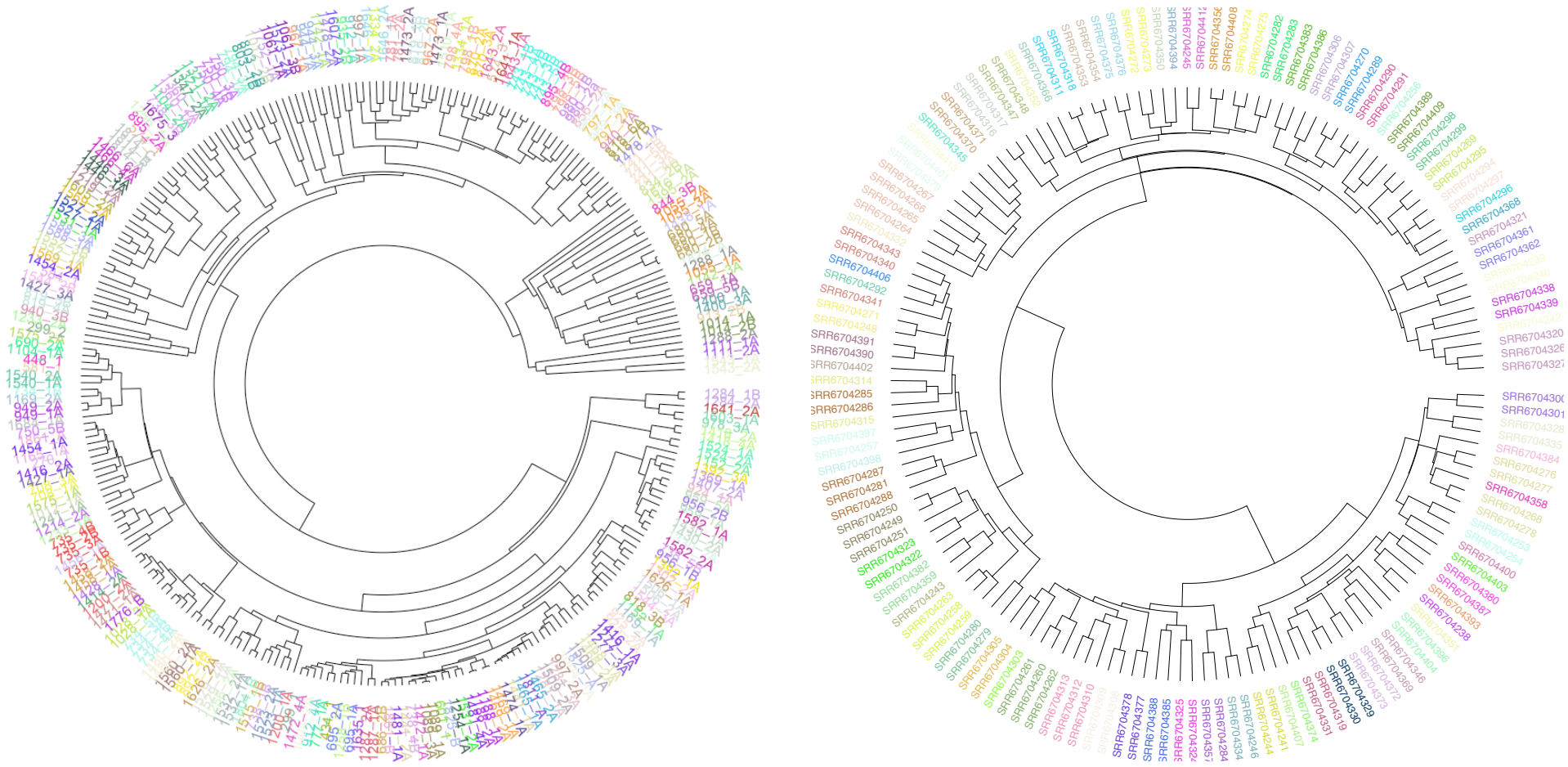


C

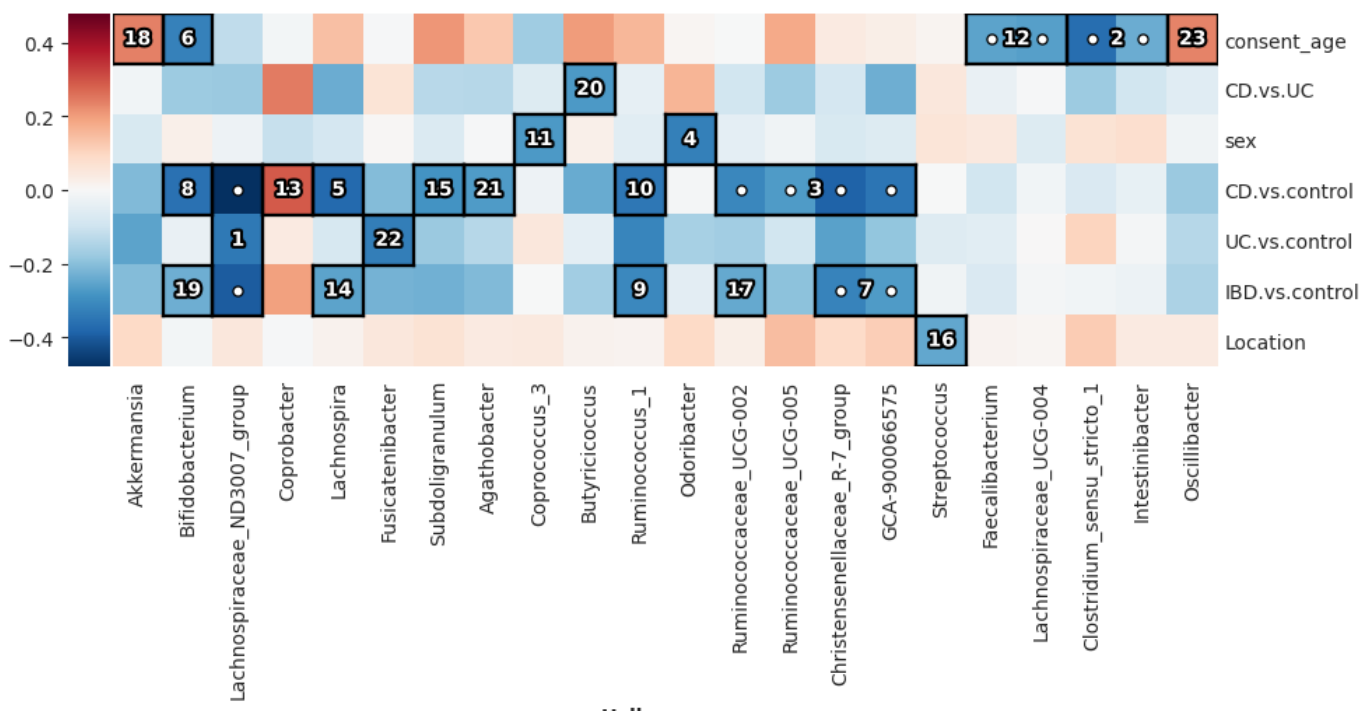


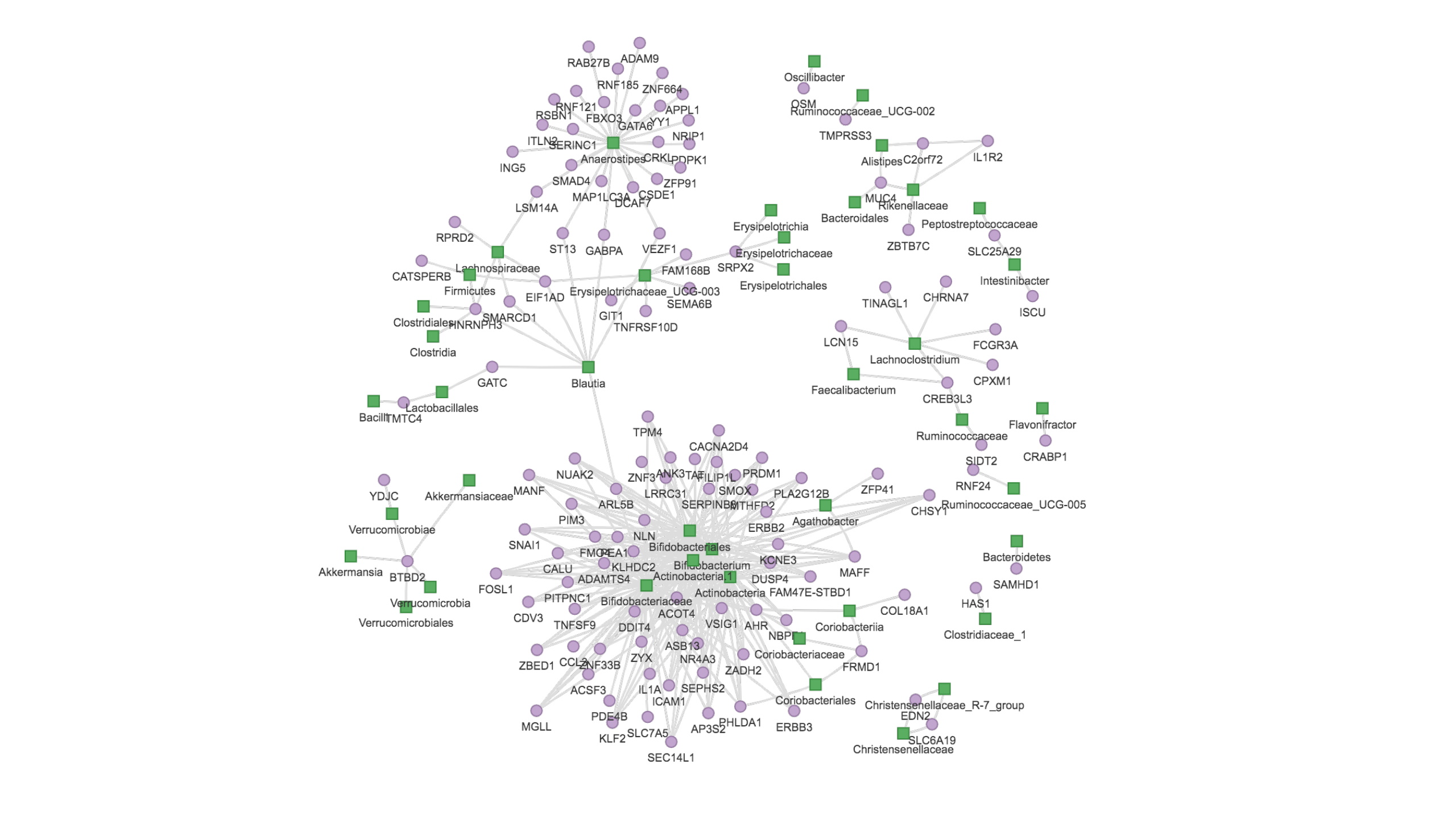


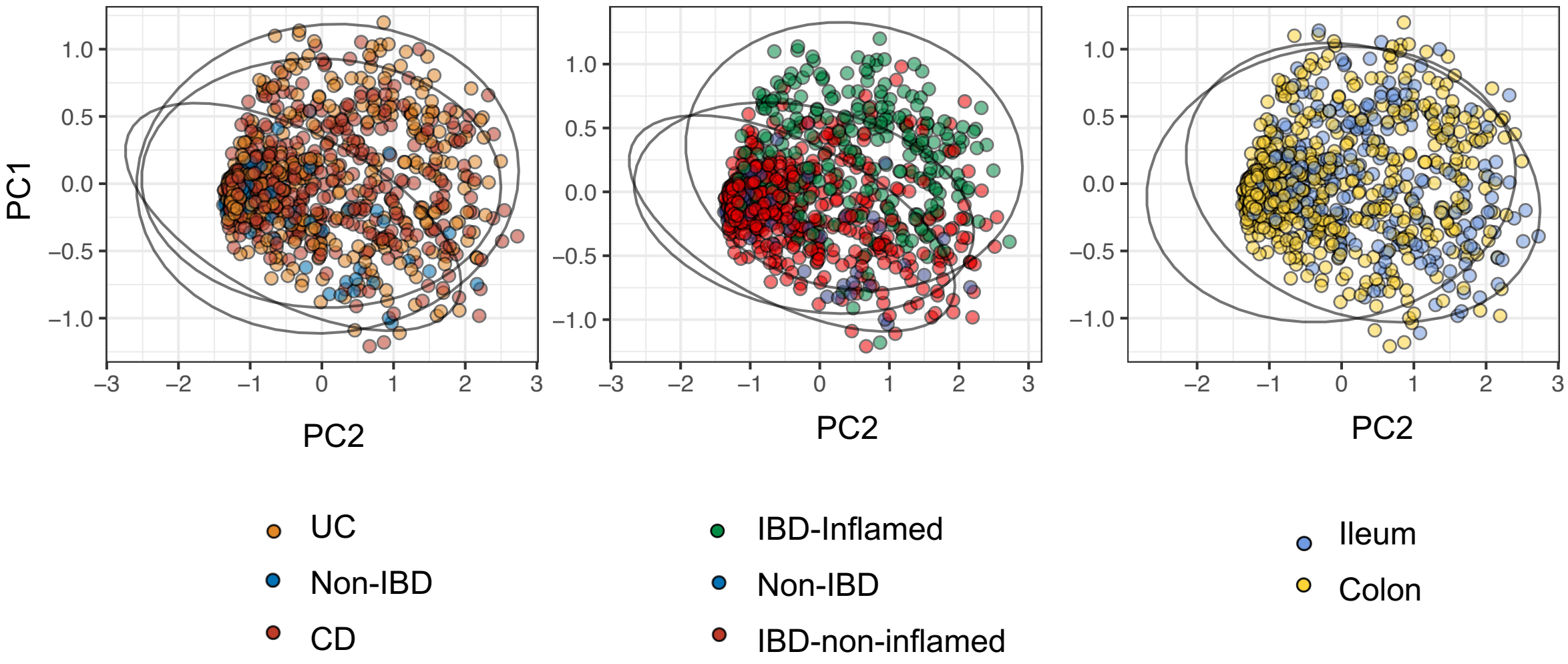
A



B




RAB27B
ADAM9
RNF185
ZNF664
RNF121
APPL1
RSBN1
FBXO3
GATA6
YY1
NRIP1
ITLN2
SERINC1
ING5
Anaerostipes
CRKL
PDPK1
SMAD4
ZFP91
MAP1LC3A
CSDE1
LSM14A
DCAF7
RPRD2
ST13
GABPA
VEZF1
CATSPERB
Lachnospiraceae
FAM168B
Firmicutes
EIF1AD
Erysipelotrichaceae_UCG-003
SEMA6B
SMARCD1
GIT1
Clostridiales
HNRNPH3
TNFRSF10D
Clostridia
GATC
Blautia
Lactobacillales
Bacilli
TMTC4
Oscillibacter
OSM
Ruminococcaceae_UCG-002
TMPRSS3
Alistipes
C2orf72
IL1R2
MUC4
Rikenellaceae
Bacteroidales
ZBTB7C
Peptostreptococcaceae
SLC25A29
Erysipelotrichia
Erysipelotrichaceae
SRPX2
Erysipelotrichales
Intestinibacter
TINAGL1
CHRNA7
ISCU
LCN15
FCGR3A
Lachnoclostridium
CPXM1
Faecalibacterium
CREB3L3
Flavonifractor
Ruminococcaceae
CRABP1
SIDT2
RNF24
Ruminococcaceae_UCG-005
CHSY1
TPM4
CACNA2D4
NUAK2
ZNF3
ANK3TAF
FILIP1L
PRDM1
ZFP41
MANF
LRRC31
SMOX
PLA2G12B
YDJC
Akkermansiaceae
ARL5B
SERPINB6
MTHFD2
Agathobacter
PIM3
ERBB2
Verrucomicrobiae
NLN
SNAI1
Bifidobacteriales
Bacteroidetes
Akkermansia
KLHDC2
KCNE3
CALU
Bifidobacterium
MAFF
BTBD2
ADAMTS4
Actinobacteria_1
DUSP4
SAMHD1
FOSL1
Actinobacteria
FAM47E-STBD1
PITPNC1
Bifidobacteriaceae
HAS1
Verrucomicrobia
ACOT4
COL18A1
CDV3
VSIG1
AHR
Coriobacteriia
TNFSF9
DDIT4
Verrucomicrobiales
Clostridiaceae_1
ASB13
ZBED1
ZYX
NR4A3
Coriobacteriaceae
FRMD1
ZADH2
ACSF3
IL1A
SEPHS2
Coriobacteriales
Christensenellaceae_R-7_group
ICAM1
EDN2
MGLL
PDE4B
PHLDA1
ERBB3
AP3S2
SLC6A19
KLF2
SLC7A5
Christensenellaceae
SEC14L1


PC1
PC2
PC2
PC2
UC
Non-IBD
CD
IBD-Inflamed
Non-IBD
IBD-non-inflamed
Ileum
Colon