COMPARATIVE ANALYSIS OF GUT MICROBIAL DIVERSITY IN HEALTHY INDIVIDUALS AND PATIENTS WITH INFLAMMATORY BOWEL DISEASE

Md. Saheed Askar

Department of Microbiology, Government Medical College, Ramanathapuram, Tamil Nadu, India. Email: saheedaskar@gmail.com

Abstract

Gut microbial dysbiosis is common in inflammatory bowel disease (IBD), such as Crohn's disease and ulcerative colitis. This study aimed to compare the diversity and composition of the gut microbiome of healthy people and those with IBD. Shotgun metagenomic data that have been publicly made available from the Human Microbiome Project 2 IBD Multi-omics Database were analysed. After removal of duplicates, 1,373 stool samples from 116 individuals (27 healthy, 56 Crohn's patients, and 33 ulcerative colitis patients) were retained. Shannon index, Simpson index, inverse Simpson and observed species were used to assess alpha diversity. Differences among microbial communities were determined using Bray–Curtis dissimilarity, principal coordinate analysis, PERMANOVA and differential-abundance testing. Observed species richness was significantly higher in healthy subjects compared to those with IBD, with a median of 50 species and 40 species, respectively ($p = 0.003$). There were no significant differences between Shannon, Simpson or inverse Simpson indices across diagnostic groups. The beta diversity analysis revealed that there were significant overlaps among the groups, and the diagnostic model accounted for just 1.8% of the community variation ($p = 0.395$). The taxa Ruminococcus bicirculans, Gemmiger formicilis and Coprococcus eutactus were depleted in IBD while the taxa Flavonifractor plautii and Clostridium symbiosum were increased. Only Roseburia sp. The CAG 182 was still significant even after correction for the false discovery rate. IBD was mainly correlated with a decrease in microbial richness and taxon-specific changes, yet there was significant microbiome inter-personal diversity.

Keywords: *gut microbiome; inflammatory bowel disease; microbial diversity; dysbiosis; metagenomics*

1. Introduction

The human gut microbiome is a highly diverse collection of microorganisms that is vital for maintaining host physiology, metabolism, immune regulation and protection against pathogenic microorganisms (1,2). Microbiome technologies, and metagenomic sequencing technology in particular, have significantly contributed to understanding the diversity of microorganisms and the complexity of the intestinal ecosystem and its interactions with the host (1,2). A healthy gut microbiota is responsible for maintaining intestinal homeostasis via nutrient metabolism, production of beneficial metabolites, maintaining epithelial barrier integrity and regulating immune responses in healthy humans (3,4). The community composition is dependent on the host genetics, diet, age and environment and therefore displays different microbial configurations, called enterotypes (5).

Inflammatory bowel disease (IBD), the most common form of which is Crohn's disease (CD), and ulcerative colitis (UC), is a chronic relapsing inflammatory disorder of the gastrointestinal tract with Immune Dysregulation and chronic inflammation of the intestinal tract (6). The growing evidence that changes in "gut microbiome" are closely linked to the onset, progression, and clinical outcomes of disease makes microbial profiling an important part of the new research into IBD (7).

Microbiological dysbiosis is a hallmark of IBD, where there is a decrease in microbial diversity, loss of beneficial commensal organisms and an increase in pro-inflammatory taxa (8,9). Beneficial short-chain fatty acid (SCFA) producing bacteria, such as members of the genera Faecalibacterium and Roseburia, are often decreased and opportunistic bacteria that have the potential to drive inflammatory processes are relatively abundant (3). The microbial changes affect the function of the epithelial barrier and immune regulation and drive chronic intestinal inflammation (6). Further, multi-omics studies have identified that microorganisms affect the metabolism of their hosts and immune signaling, highlighting the intricate relationships between microorganisms and disease pathology (10). Recent studies have also shown that other gut microorganisms, including intestinal protists like Blastocystis, can also have an effect on the microbial environment and inflammation, which is currently being investigated (11).

Much of the research has documented changes in the microbiome of IBD, but the use of varying sequencing technologies, workflows and bioinformatic pipelines has made it difficult to directly compare the diversity of microbes across different studies (12). Standardized comparative analyses at high-resolution level of microbial profiles of healthy versus diseased populations can be achieved using large publicly available metagenomic resources obtained through the Integrative Human Microbiome Project (13). An in-depth assessment of alpha diversity, beta diversity and taxonomy within the same analytical framework, however, is still fairly limited.

This study is undertaken to investigate the differences in the diversity and composition of gut microbiota between healthy people and people suffering from inflammatory bowel disease (IBD) by analyzing publicly available shotgun metagenomic sequencing data of their gut microbiota. Specific aims were to compare alpha diversity between healthy controls and IBD patients, test for differences in beta diversity and overall structure of the microbial community, identify bacterial taxa that are differentially abundant between healthy and IBD, determine potential microbial biomarkers associated with IBD, and characterize the microbial composition at several taxonomic levels using standardized metagenomic analyses.

2. Materials and Methods

2.1 Study Design

This comparative longitudinal bioinformatics analysis was performed to examine differences in the diversity and taxonomic makeup of the gut microbiome of healthy controls and inflammatory bowel disease (IBD) patients. Patients in the IBD group were those with Crohn's disease and ulcerative colitis. The statistical analyses took multiple stool samples from the same participants into consideration.

2.2 Data Source

The data were from the Human Microbiome Project 2 Inflammatory Bowel Disease Multi-omics Database, which is publicly available. The dataset included species-level relative-abundance data of stool metagenomic samples along with participant ID, sample ID, sampling week, diagnosis and faecal calprotectin (14). The dataset consisted of 3,387 records and 571 variables, with 566 variables being microbial species. Of these stool samples, 116 participants had a total of 1,373 samples, after duplicate sample records were removed. No additional ethical approval was needed for this secondary analysis as the data were publicly available and de-identified.

2.3 Study Population

Twenty-seven healthy participants, 56 Crohn's disease and 33 ulcerative colitis participants were included in the final dataset. The numbers of healthy, Crohn's disease and ulcerative colitis samples were 350, 647 and 376, respectively, with 1,373 eligible samples. Samples were only included if they had a valid participant identifier, sample identifier, diagnosis, sampling week and complete microbial abundance profile. Duplicate records, samples without diagnostic information and samples without abundance information or with invalid abundance data were removed.

2.4 Study Variables

Disease status (healthy or IBD) was the main outcome variable. Secondary comparisons were made between healthy individuals, Crohn's patients and ulcerative colitis patients. Five hundred and sixty-six species of the microbial world were analysed in terms of their relative abundance. Where species-level data were available they were aggregated to genus level. Fecal calprotectin was added as a variable to measure intestinal inflammation, and participant ID and sampling week were added to control for repeated longitudinal measurements. The alpha diversity was measured based

on Shannon index, Simpson index, inverse Simpson index and observed species richness. Bray–Curtis dissimilarity and Jaccard distance were used to assess beta diversity which was subsequently analysed by principal coordinate analysis.

2.5 Data Preprocessing

The data were checked for inconsistencies in abundances, invalid diagnostic labels, zero-variance taxa, and duplicate data. A single record was kept for each unique sample identifier, which decreased the number of samples from 3,387 to 1,373. Seven species, which had zero abundance in all the samples, were eliminated. Taxa were maintained if found at an abundance of 0.01% or more and in at least 1% of the samples. Percentages of relative abundance were converted to proportions. No rarefactions were made since the data set consisted of normalized relative abundance values and not raw counts from the sequencing data. Presence-absence data were created to use for richness and Jaccard-distance analyses. If necessary, a small pseudocount was added before logarithmic or centred log-ratio transformation.

2.6 Statistical Analysis

Frequencies, percentages, means, standard deviations, medians and interquartile ranges were used for summarising participant and sample characteristics, as appropriate. The assumptions of the normal distribution were evaluated graphically and by the Shapiro–Wilk test. The Mann–Whitney U test was used to compare the alpha-diversity indices between healthy and IBD groups. Kruskal–Wallis test and Dunn's post hoc analysis were used for comparisons between healthy, CD and UC groups. Longitudinal analyses were performed using linear mixed-effects models, with diagnosis and sampling week as fixed effects and participant ID as a random intercept. Spearman correlation and mixed-effects modelling were used to evaluate the correlation between faecal calprotectin and microbial diversity. Permutational multivariate analysis of variance was used to test the differences between beta-diversity, permutational with participants. The homogeneity of multivariate dispersion was checked prior to interpretation of the group differences. A Benjamini–Hochberg false discovery rate was performed and false discovery rate-corrected p-values < 0.05 were deemed statistically significant.

2.7 Bioinformatics Analysis

Species-level abundance matrix was filtered and alpha diversity indices were calculated. Bray–Curtis and Jaccard distance matrices were used to evaluate differences in microbial community composition. To visualize clustering by diagnostic group, principal coordinate analysis (PCA) was used. The summarization of the dominant and differentially abundant taxa was performed at species and genus levels. Differential abundance analysis was performed using ANCOM-BC or another compositional analysis appropriate for relative abundance data and repeated observations. Log-transformed and standardized abundance values were used for the construction of heatmaps and hierarchical clustering. To avoid information leakage at the level of the participants, penalized logistic regression or random forest models were used to identify candidate microbial biomarkers.

3. Results

3.1 Study Sample Characteristics

One record per external sample identifier was kept and 1,373 stool samples from 116 participants were included. A total of 27 healthy subjects, 56 patients with Crohn's disease and 33 patients with ulcerative colitis (UC) were studied. The most representative number of samples was from Crohn's disease. There were 455 samples with faecal calprotectin levels. Faecal calprotectin levels were significantly higher in both the IBD groups compared with healthy individuals, reflecting higher levels of intestinal inflammation in patients with Crohn's disease and with ulcerative colitis as shown in Table 1.

Table 1. Participant and sample characteristics according to diagnostic group

Diagnostic group	Participants, n	Stool samples, n	Study-week range	Calprotectin samples, n	Median faecal calprotectin
Healthy	27	350	0–57	108	17.02
Crohn’s disease	56	647	0–52	220	118.59
Ulcerative colitis	33	376	0–57	127	107.80
Total	116	1,373	0–57	455	—

3.2 Distribution of Stool Samples

The number of stool samples varied between diagnostic groups (647 from Crohn's disease, 376 from ulcerative colitis and 350 from healthy subjects). As shown in Figure 1, the largest sample group was Crohn's disease; however, there were repeated samples for this group that were handled in the statistical analysis at the participant level.

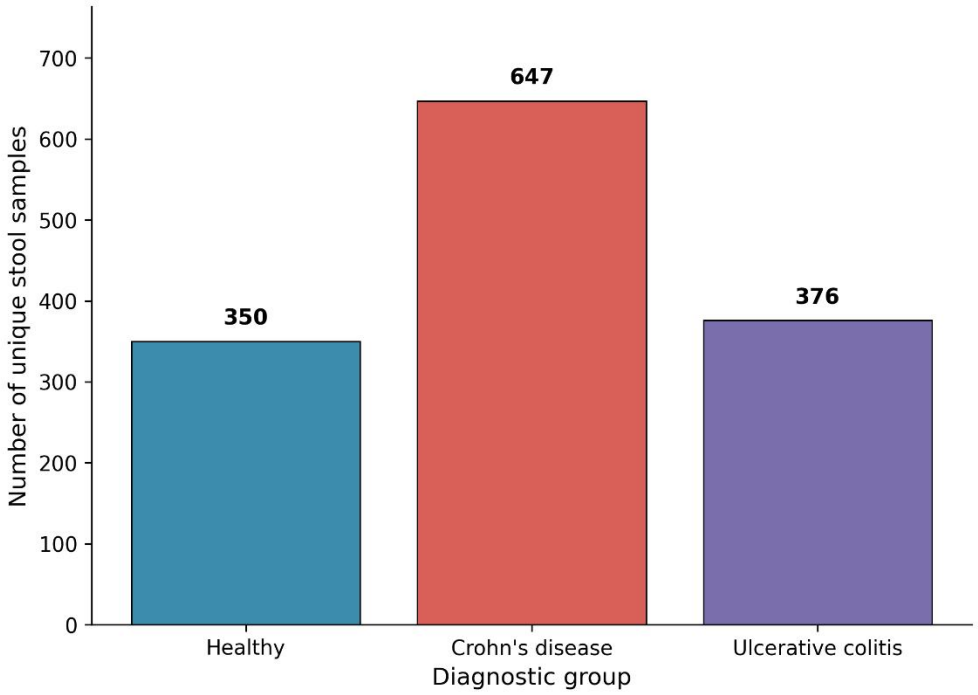

Figure 1. Distribution of unique stool samples across diagnostic groups

3.3 Comparison of Alpha Diversity

The analysis of alpha-diversity was performed with the use of the first available sample of each participant. The median Shannon and Simpson diversity values were lower in the IBD groups compared to the healthy individuals, but did not differ statistically. Conversely, the observed species richness was significantly different among the three diagnostic groups. Table 2 illustrates that the median observed species richness was the highest in healthy people. When the two diseases (Crohn and ulcerative colitis) were pooled to form one group of IBD, the observed richness was much lower in IBD compared to healthy individuals (median: 40 vs. 50 species; Mann-Whitney p = 0.003). These results suggest that IBD was more significantly related to decreased microbial richness than to total Shannon or Simpson diversity.

Table 2. Comparison of gut microbial alpha diversity among diagnostic groups

Diversity measure	Healthy, median (IQR)	Crohn's disease, median (IQR)	Ulcerative colitis, median (IQR)	Kruskal–Wallis H	p-value
Shannon index	2.366 (1.969–2.496)	2.255 (1.756–2.482)	2.070 (1.751–2.494)	2.293	0.318
Simpson index	0.835 (0.790–0.862)	0.836 (0.743–0.865)	0.821 (0.672–0.868)	1.039	0.595
Inverse Simpson index	6.076 (4.772–7.272)	6.064 (3.771–7.268)	5.574 (3.046–7.551)	0.970	0.616
Observed species	50.0 (40.0–59.0)	38.5 (31.0–46.3)	40.0 (34.0–51.0)	9.592	0.008

IQR: interquartile range.

3.4 Distribution of Shannon Diversity

There was significant within-group variation in the Shannon diversity index. The median Shannon index of healthy participants was greater than that of participants with Crohn disease or ulcerative colitis, but there was a significant overlap in distributions. Figure 2 depicts a negative change in the median Shannon diversity between the healthy individuals and the IBD groups. Nevertheless, such a wide overlap was not unexpected given that the overall difference was not significant as indicated in Table 2.

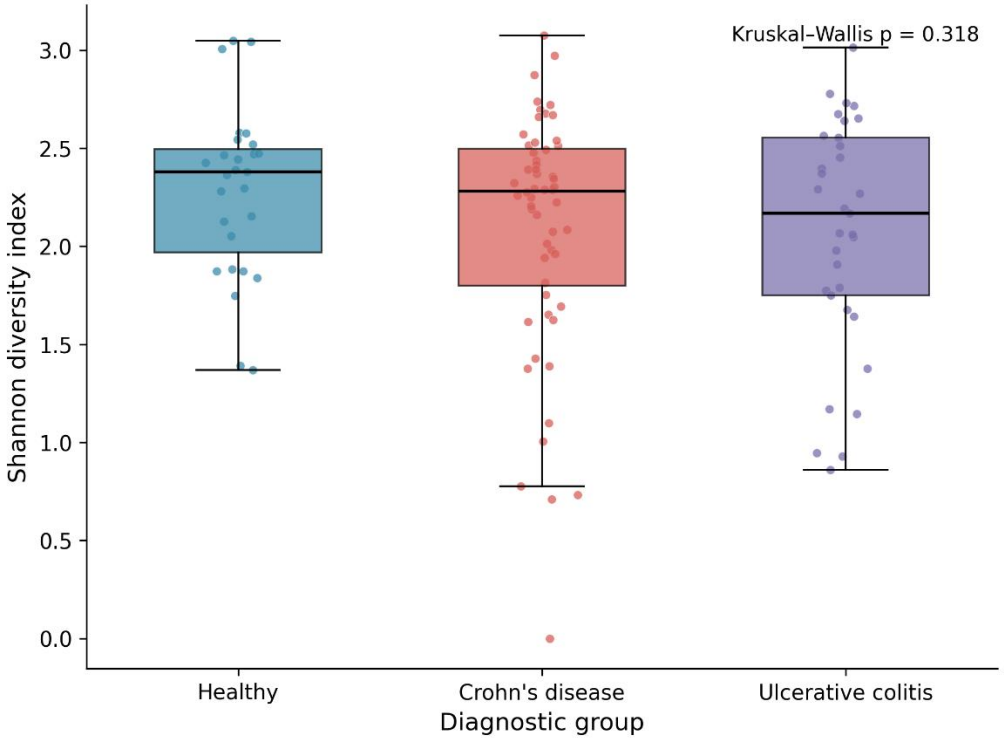

Figure 2. Shannon microbial diversity in healthy individuals and participants with inflammatory bowel disease

3.5 Dominant Gut Microbial Taxa

Bacteroides, Faecalibacterium, Prevotella, Alistipes and Parabacteroides species dominated the gut microbial community. The most common species in all diagnostic groups was Bacteroides vulgatus. Table 3 shows that the relative abundance of Faecalibacterium prausnitzii, Bacteroides stercoris, and Alistipes putredinis was lower in both IBD subgroups as compared to healthy people. On the other hand, Bacteroides fragilis was more prevalent in Crohn's disease and ulcerative colitis.

Table 3. Mean relative abundance of the ten dominant microbial species

Microbial species	Healthy (%)	Crohn's disease (%)	Ulcerative colitis (%)	Overall (%)
Bacteroides vulgatus	15.09	16.52	17.59	16.49
Bacteroides uniformis	8.70	9.87	9.79	9.58
Faecalibacterium prausnitzii	9.59	7.10	8.06	7.95
Prevotella copri	7.67	5.50	8.46	6.85
Bacteroides stercoris	7.32	5.20	5.77	5.86
Bacteroides dorei	4.43	5.19	3.46	4.52
Alistipes putredinis	3.83	2.74	2.39	2.90
Bacteroides fragilis	1.36	3.00	2.64	2.51
Bacteroides ovatus	1.63	2.36	3.40	2.49
Parabacteroides distasonis	1.74	2.84	2.35	2.45

3.6 Beta Diversity and Microbial Community Structure

BrayCurtis dissimilarity was used to conduct principal coordinate analysis on the microbial profiles averaged across participants. The first and second principal coordinate axis accounted 15.8 and 11.6 percent of the variation in the positive-coordinate, respectively. As Figure 3 shows, there is a lot of overlap between the groups of healthy individuals, the people with Crohn's disease, and the people with ulcerative colitis. The overall difference in the microbial community composition at a participant level did not show a statistically significant difference between the three groups

(pseudo-F = 1.04, R2 = 0.018, permutation p = 0.395). The variation in community composition was thus attributed to diagnosis explaining about 1.8%.

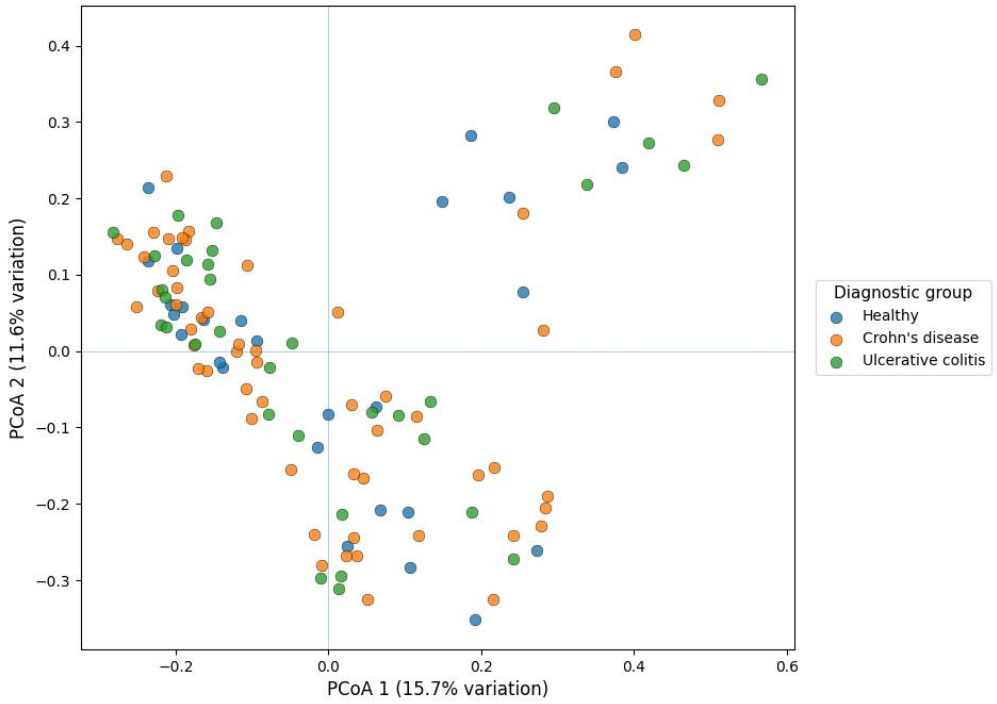

Figure 3. Principal coordinate analysis of gut microbial composition based on Bray–Curtis dissimilarity

3.7 Differentially Abundant Microbial Species

Relative abundances of the participants were compared between the groups of healthy and mixed IBD. With benjamini-Hochberg correction, Roseburia sp. only. CAG 182 was statistically significant with an FDR of 0.05. Some other taxa had nominal differences, but failed to be significant after multiple-testing correction. A rank-biserial value of negative value suggests an increased abundance or prevalence in healthy persons, and a positive value suggests an increased abundance in IBD.

Table 4 reveals that Gemmiger formicilis, Ruminococcus bicirculans, Coprococcus eutactus, and Firmicutes bacterium CAG 83 were more likely to be abundant in healthy people. Conversely, Flavonifractor plautii and Clostridium symbiosum were more abundant in the IBD group. Nevertheless, these results can be viewed as exploratory since only a single taxon was significant following FDR correction.

Table 4. Microbial species showing the strongest abundance differences between healthy and IBD groups

Microbial species	Healthy median (%)	IBD median (%)	Rank-biserial effect	Unadjusted p	FDR-adjusted p
Roseburia sp. CAG 182	0.0000	0.0000	−0.273	<0.001	0.040
Coprococcus eutactus	0.0000	0.0000	−0.251	0.003	0.275
Gemmiger formicilis	0.0047	0.0000	−0.358	0.003	0.275
Turicibacter sanguinis	0.0012	0.0000	−0.321	0.004	0.275
Clostridium symbiosum	0.0006	0.0171	0.341	0.007	0.316
Flavonifractor plautii	0.1184	0.4045	0.338	0.008	0.316
Ruminococcus bicirculans	0.2279	0.0223	−0.325	0.009	0.316
Firmicutes bacterium CAG 83	0.1029	0.0085	−0.303	0.015	0.316

3.8 Comparative Abundance of Key Differential Taxa

The mean relative abundances of the taxa showing the strongest group differences were examined to illustrate the

direction and magnitude of microbial alterations. Figure 4 illustrates lower mean abundances of Gemmiger formicilis, Coprococcus eutactus, and Ruminococcus bicirculans in participants with IBD. Conversely, Flavonifractor plautii, Clostridium symbiosum, and Turicibacter sanguinis were relatively enriched in the IBD group.

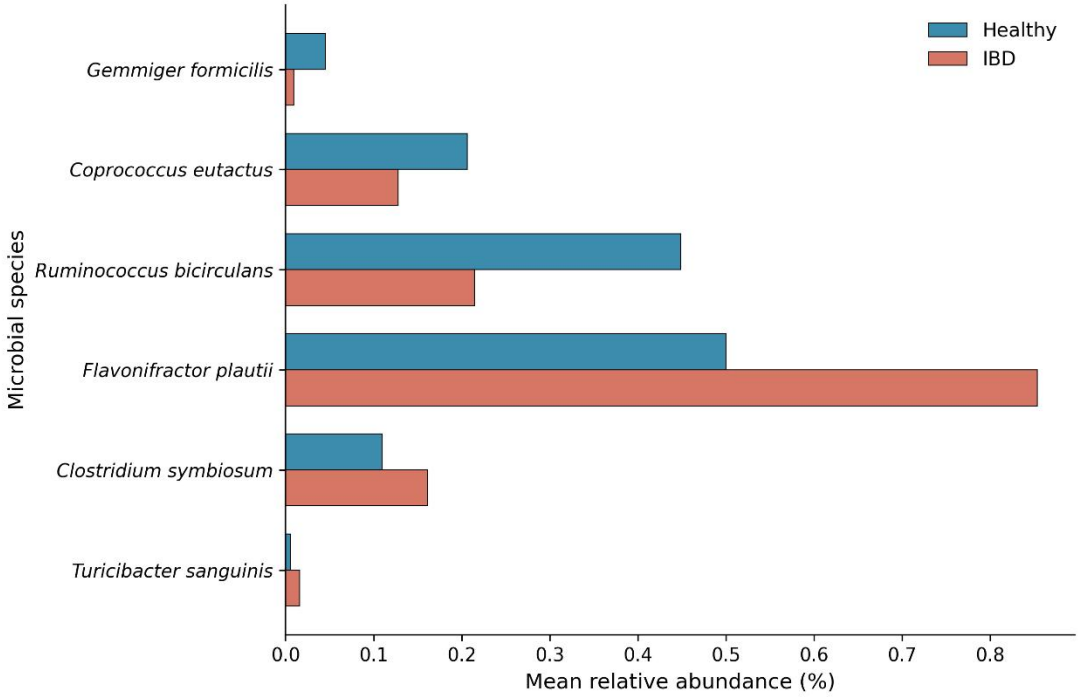

Figure 4. Relative abundance of selected microbial species in healthy and IBD participants

In general, the findings suggest that inflammatory bowel disease was defined by decreased microbial species richness and changes in the selected taxa of bacteria. But the wide overlap of Shannon diversity and beta-diversity profiles indicates that there is a great deal of interpersonal heterogeneity in the gut microbiome of healthy and IBD subjects.

4. Discussion

The current experiment found clear variations in microbial diversity and taxonomic composition of the gut in healthy people and patients with inflammatory bowel disease. Both participants with Crohn's and ulcerative colitis demonstrated decreased observed species richness, less relative abundance of a variety of beneficial commensal taxa, and enrichment of chosen bacteria related to intestinal inflammation. Nevertheless, no significant difference between groups was found in Shannon and Simpson diversity indices and beta-diversity analysis showed significant overlap, which meant that there was a great deal of interpersonal heterogeneity in both healthy and diseased groups.

The richness that is observed in IBD is much lower implying a shrinkage of the microbial ecosystem, not a uniform shrinkage across all of the measures of diversity. This can cause loss of richness, compromising potentially redundant functions and the capacity of the microbiome to sustain intestinal homeostasis. Stable microbial communities and disrupted microbial networks in Crohn disease and specifically during relapsing and treatment-refractory disease have also been reported in the past (15). However, diet, geography and lifestyle, as well as host characteristics, can affect microbial diversity, which can be one reason why the distributions of the diversity overlap in the current analysis (16,17).

The reduced representation of beneficial taxa including Gemmiger formicilis, Coprococcus eutactus and Ruminococcus bicirculans an indication of loss of organisms that ferment fibre and produce short-chain fatty acids. On the other hand, the relative increase in Flavonifractor plautii and Clostridium symbiosum in IBD represents a higher proportion of the taxa that are well adapted to an inflammatory intestinal environment. Large comparative studies of IBD and other gastrointestinal disorders have reported similar changes in their compositional and functional alterations (18). Alterations in microbial transcriptional activity also suggest that dysbiosis does not just occur through taxonomic changes but also through changes in the functioning of the microbes (19).

Microbial dysbiosis can also cause IBD by compromising epithelial barrier functions, dysregulated immune responses, disrupted bile acid metabolism, and low levels of anti-inflammatory metabolites. The role of host microbiota interactions is becoming one of the key elements in intestinal inflammation and disease development (20,21). Nevertheless, it is still not clear whether dysbiosis is a causative factor or a follow-up effect of inflammation since the inflammatory conditions also reconstruct the intestinal environment and prefer the growth of opportunistic organism (22).

The current results can be correlated with past data indicating the depletion of microbial richness and disruption of microbial networks in IBD. However, the fact that the separation of diagnostic groups in beta-diversity analysis is limited implies that the disease-related microbiome signatures are not universal but heterogeneous. Other chronic intestinal diseases have been characterized by similar host–microbiota maladaptation, demonstrating that microbial imbalance can contribute to interactions between host immunity and disease across disease states (23).

The microbiome profiling of the gut used to aid in disease classification, risk stratification, treatment monitoring, and

the identification of microbial biomarkers. Microbiota-targeted approaches, such as probiotics, dietary change, faecal microbiota transplantation, and interventions with the microbiota-bile acid axis can be used to supplement traditional IBD therapies (24). The use of approaches aimed at restoring regulatory immune activity by means of manipulation of resident microbiota is promising in particular (25).

One of the strongest points was the utilization of a big publicly available metagenomic dataset with species-level resolution and re-sampling. Nevertheless, the study was limited by the unequal sizes of groups, missing values of faecal calprotectin, unavailable dietary and medication information, the use of relative-abundance data instead of raw count data. The impact of nutrition and the environment was not entirely evaluated, but these aspects have a significant impact on the animal and human microbial communities (26).

Future research ought to combine metagenomics and transcriptomics, metabolomics, dietary data and immune markers and longitudinal clinical outcomes. Validation of populations is also needed since microbiome composition differs with geography and culture. Further studies on microbiota-based therapy in gastrointestinal and metabolic diseases can help to better understand the potential of microbial modulation in clinical practice.

5. Conclusion

The study has compared gut microbial diversity and community composition among healthy subjects and patients with inflammatory bowel disease, using a massive publicly available shotgun metagenomic dataset. The results revealed that IBD was related to a substantially lower microbial species richness and changes in abundance of various bacterial taxa implicated in intestinal homeostasis. Despite significant overlap of Shannon and Simpson diversity indices and community structure as a whole in the comparison between healthy and IBD groups, the reduction in species richness and loss of useful commensal bacteria observed justify the existence of microbial dysbiosis in IBD. The fact that selected taxa related to the intestinal inflammation become enriched also underscores the complexity of ecological changes that accompany the progression of disease. The findings support the notion that microbial changes in the gut are a significant aspect of IBD pathophysiology and that metagenomic profiling at the species level can be used to offer insights into microbial signatures of disease. Heterogeneity of the participants observed also indicates that no universal microbial profile is representative of IBD and thus, microbiome assessment should be done on an individual basis. Incorporation of metagenomic, metabolomic, transcriptomic, dietary, and clinical data in future studies can help enhance the knowledge of host-microbiome interactions and enable the identification of dependable microbial biomarkers. All in all, this report adds to the existing body of literature that suggests microbiome-based solutions to monitoring diseases, making a precise diagnosis, and creating specific treatment plans to manage patients with inflammatory bowel disease.

Reference

1. Almeida A, Mitchell AL, Boland M, Forster SC, Gloor GB, Tarkowska A, et al. A new genomic blueprint of the human gut microbiota. Nature. 2019;568(7753):499–504.
2. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol. 2019 Aug;37(8):852–7. doi:10.1038/s41587-019-0209-9
3. Lavelle A, Sokol H. Gut microbiota-derived metabolites as key actors in inflammatory bowel disease. Nature reviews Gastroenterology & hepatology. 2020;17(4):223–37.
4. Rinninella E, Raoul PC, Cintoni M, Franceschi F, Miggiano GAD, Gasbarrini A, et al. What is the healthy gut microbiota composition? A changing ecosystem across age, environment, diet, and diseases. Microorganisms. 2019;2019(7):N-A.
5. Costea PI, Hildebrand F, Arumugam M, Bäckhed F, Blaser MJ, Bushman FD, et al. Enterotypes in the landscape of gut microbial community composition. Nature microbiology. 2018;3(1):8–16.
6. Nishida A, Inoue R, Inatomi O, Bamba S, Naito Y, Andoh A. Gut microbiota in the pathogenesis of inflammatory bowel disease. Clin J Gastroenterol. 2018 Feb;11(1):1–10. doi:10.1007/s12328-017-0813-5
7. Haiser H, Porter J, Reinker S. Gut microbiome structure and metabolic activity in inflammatory bowel disease. Nature Microbiology. 2019;4(2):293–305.
8. Halfvarson J, Brislawn CJ, Lamendella R, Vázquez-Baeza Y, Walters WA, Bramer LM, et al. Dynamics of the human gut microbiome in inflammatory bowel disease. Nature microbiology. 2017;2(5):17004.
9. Pascal V, Pozuelo M, Borruel N, Casellas F, Campos D, Santiago A, et al. A microbial signature for Crohn's disease. Gut. 2017;66(5):813–22.
10. Lloyd-Price J, Arze C, Ananthakrishnan AN, Schirmer M, Avila-Pacheco J, Poon TW, et al. Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases. Nature. 2019;569(7758):655–62.
11. Aykur M, Malatyalı E, Demirel F, Cömert-Koçak B, Gentekaki E, Tsaousis AD, et al. Blastocystis: a mysterious member of the gut microbiome. Microorganisms. 2024;12(3):461.
12. Prodan A, Tremaroli V, Brolin H, Zwinderman AH, Nieuwdorp M, Levin E. Comparing bioinformatic pipelines for microbial 16S rRNA amplicon sequencing. PloS one. 2020;15(1):e0227434.
13. Proctor LM, Creasy HH, Fettweis JM, Lloyd-Price J, Mahurkar A, Zhou W, et al. The integrative HMP (iHMP) research network consortium. The integrative human microbiome project Nature. 2019;569(7758):641–8.
14. Hussain Q. Human Gut Microbiome Atlas (HMP2) [Internet]. 2026 [cited 2026 Jul 29]. Available from: https://www.kaggle.com/datasets/qasimhu/human-gut-microbiome-atlas-hmp2
15. Yilmaz B, Juillerat P, Øyås O, Ramon C, Bravo FD, Franc Y, et al. Microbial network disturbances in relapsing refractory Crohn's disease. Nature medicine. 2019;25(2):323–36.
16. Sheng Y, Wang J, Gao Y, Peng Y, Li X, Huang W, et al. Combined analysis of cross-population healthy adult human microbiome reveals consistent differences in gut microbial characteristics between Western and non-

Western countries. Comput Struct Biotechnol J. 2024 Dec;23:87–95. doi:10.1016/j.csbj.2023.11.047
17. Zhang L, Tuoliken H, Li J, Gao H. Diet, gut microbiota, and health: a review. Food Sci Biotechnol. 2025 Jun;34(10):2087–99. doi:10.1007/s10068-024-01759-x
18. 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 Dec 19;10(472):eaap8914. doi:10.1126/scitranslmed.aap8914
19. Schirmer M, Franzosa EA, Lloyd-Price J, McIver LJ, Schwager R, Poon TW, et al. Dynamics of metatranscription in the inflammatory bowel disease gut microbiome. Nature microbiology. 2018;3(3):337–46.
20. Caruso R, Lo BC, Núñez G. Host–microbiota interactions in inflammatory bowel disease. Nature Reviews Immunology. 2020;20(7):411–26.
21. Zuo T, Ng SC. The gut microbiota in the pathogenesis and therapeutics of inflammatory bowel disease. Frontiers in microbiology. 2018;9:2247.
22. Khan I, Ullah N, Zha L, Bai Y, Khan A, Zhao T, et al. Alteration of gut microbiota in inflammatory bowel disease (IBD): cause or consequence? IBD treatment targeting the gut microbiome. Pathogens. 2019;8(3):126.
23. Janney A, Powrie F, Mann EH. Host–microbiota maladaptation in colorectal cancer. Nature. 2020;585(7826):509–17.
24. Li L, Liu T, Gu Y, Wang X, Xie R, Sun Y, et al. Regulation of gut microbiota-bile acids axis by probiotics in inflammatory bowel disease. Frontiers in Immunology. 2022;13:974305.
25. Mishima Y, Sartor RB. Manipulating resident microbiota to enhance regulatory immune function to treat inflammatory bowel diseases. J Gastroenterol. 2020 Jan;55(1):4–14. doi:10.1007/s00535-019-01618-1
26. Saar R, Dodd S. Small animal microbiomes and nutrition [Internet]. John Wiley & Sons; 2023 [cited 2026 Jul 29]. Availablefrom:https://books.google.com/books?hl=en&lr=&id=UzzOEAAAQBAJ&oi=fnd&pg=PR17&dq=Mart%C3%ADn,+R.,+Miquel,+S.,+Ulmer,+J.,+et+al.+(2022).+Gut+microbiota+and+inflammatory+bowel+disease:+Current+knowledge+and+future+perspectives.+Microorganisms,+10(6),+1158.&ots=fZYmiWvV94&sig=cIe5_-enMKwpeDZBdOA8QfVOn1s