Gastroenterology and Hepatology From Bed to Bench.
2025 **RIGLD**, Research Institute for Gastroenterology and Liver Diseases

ORIGINAL ARTICLE

p-ISSN: 2008-2258 e-ISSN: 2008-4234

Microbiome-microRNA interactions in inflammatory bowel disease: insights from metagenomic and transcriptomic data analysis

Nesa Kazemifard[1†], Mohammad Hossein Norouzi-Beirami[2†], Shaghayegh Baradaran Ghavami[1], Leila Ghanbari-Maman[3], Mohammad Reza Zali[4], Shabnam Shahrokh[1], Kaveh Kavousi[3*]

[1]Basic and Molecular Epidemiology of Gastrointestinal Disorders Research Center, Research Institute for Gastroenterology and Liver Diseases, Shahid Beheshti University of Medical Sciences, Tehran, Iran
[2]Department of Computer Engineering, Islamic Azad University, Osku Branch, Osku, Iran
[3]Laboratory of Complex Biological systems and Bioinformatics (CBB), Department of Bioinformatics, Institute of Biochemistry and Biophysics (IBB), University of Tehran, Tehran, Iran
[4]Gastroenterology and Liver Diseases Research Center, Research Institute for Gastroenterology and Liver Diseases, Shahid Beheshti University of Medical Sciences, Tehran, Iran
[†] These authors contributed equally to this work.

ABSTRACT

Background: Inflammatory Bowel Disease (IBD) is a chronic inflammation of the gastrointestinal tract, the precise origins of which remain not fully elucidated. This study investigates the complex relationship between gut metagenomics and host transcriptomics in IBD patients, focusing on Ulcerative Colitis (UC) and Crohn's Disease (CD).

Method: One proposed theory suggests that microRNAs produced by the host may significantly influence IBD development by impacting the gut microbiota. Conversely, the gut microbiome may regulate the expression of host microRNAs, leading to dysfunction in the intestinal epithelium. An enrichment analysis was conducted to pinpoint associated pathways. To unravel this intricate interplay, the study utilized data from the IBDMDB database, selecting samples from adult individuals.

Result: The dataset comprised 50 paired metagenomic and host transcriptomic samples, including 8 controls, 18 UCs, and 24 CDs. Computational analyses and network constructions were applied to identify relationships between bacterial species, microRNAs, and other transcripts.

Conclusion: This research offers valuable insights into the dynamic relationship between the gut microbiome and human transcriptomics in IBD, providing a deeper understanding of potential disease mechanisms. Furthermore, it sheds light on the complex tripartite network connecting bacterial species, microRNAs, and transcripts, contributing to a comprehension of IBD pathogenesis and the identification of novel therapeutic targets.

Keywords: Inflammatory bowel disease, Gut microbiome, microRNAs, Metagenomics, Transcriptomics.

(Please cite as: **Kazemifard N, Norouzi-Beirami MH, Baradaran Ghavami SH, Ghanbari-Maman L, Zali MR, Shahrokh SH, et al. Microbiome-microRNA Interactions in Inflammatory Bowel Disease: Insights from Metagenomic and Transcriptomic Data Analysis. Gastroenterol Hepatol Bed Bench 2025;18(SI):85-96. https://doi.org/10.22037/ghfbb.v18i4.3186).**

Introduction

Inflammatory bowel disease (IBD), comprising the two main types of Crohn's disease (CD) and ulcerative colitis (UC), represents a chronic relapsing disorder with a complex etiology, characterized by inflammation in the intestinal tract (1). The incidence of IBD, including in pediatric populations, has risen steadily in recent decades, transforming IBD into a pressing public health issue of global significance (2-4). The intricate multifactorial nature of IBD, combining genetic predisposition, immunologic factors, gut microbiota dysbiosis, and environmental elements, poses significant challenges to effective prevention and treatment (5, 6).

Received: 21 Aug 2025 *Accepted*: 09 Oct 2025 *Published:* 21 Dec 2025
Reprint or Correspondence: **Kaveh Kavousi,** Laboratory of Complex Biological systems and Bioinformatics (CBB), Department of Bioinformatics, Institute of Biochemistry and Biophysics (IBB), University of Tehran, Tehran, Iran.
E-mail: kkavousi@ut.ac.ir
ORCID ID: 0000-0002-1906-3912

Recent findings highlight an overactive immune response to variations in gut microbiota, in genetically susceptible hosts, as a principal driving factor in IBD (7,8). This emphasizes the imperative for further exploration of the interplay between the gut microbiome and the host in the etiology and pathogenesis of IBD. microRNAs, small non-coding RNAs, have emerged as critical regulators of gene expression at the post-transcriptional level, and their dysregulation has been linked to the pathophysiology of immune diseases like IBD (9-12). These microRNAs can either enhance or inhibit immune and inflammatory signaling, thereby regulating the immune response pathways associated with chronic inflammation in IBD (13, 14).

Comprising a vast population of microorganisms, with bacteria being the most abundant, the gut microbiome plays a vital role in immune regulation, suggesting that its composition could influence immune responses to environmental triggers (15, 16). Emerging evidence indicates that the gut microbiome can modulate host genome expression, notably microRNAs, with subsequent implications for immune pathways in IBD. Certain differentially expressed microRNAs, such as mir-1226, mir-515-5p, and mir-548ab, have shown a distinct impact on gut microbial populations in conditions of inflammation and cancer (17, 18). However, the intricacies of the interaction between microRNAs and the microbiome in the pathogenesis of IBD remain to be fully elucidated.

Advancements in the field of metagenomics over the past decade have greatly enriched our understanding of the role of microbial communities in human health and disease, despite the challenges posed by the vast complexity and diversity inherent to these communities (19). While taxonomic profiling of the microbiome provides crucial information about the relative abundance of different microbial species, a comprehensive analytical approach should also consider functional and strain-level profiling for a holistic understanding of microbial community dynamics (20).

Differential abundance analysis has emerged as a powerful tool for identifying changes in microbial community structures across different conditions or over time. To facilitate this complex analysis, bioinformatics pipelines like CAMAMED have been developed to integrate various analytical approaches (21). Yet, it is essential to appreciate the importance of appropriate statistical method selection and proper result interpretation, especially when employing nonparametric methods. Furthermore, transcriptomic analysis, including methods implemented in DESeq2, offers valuable insights into the functional potential of microbial communities, which has broad implications across health and environmental studies (22). Achieving an accurate alignment of transcriptomic data is a critical step in these analyses, with additional tools like HTSeq available for further processing of high-throughput sequencing data (23). To conclude, the advent of metagenomics has ushered in an era of both unparalleled opportunities and formidable challenges. By enabling a deeper exploration of the intricate relationships between microbial communities and human health, metagenomics can pave the way for critical breakthroughs in our understanding of the dynamic and multifaceted world of the microbiome.

Concerning these considerations, the primary objective of the present study was to scrutinize the interaction between gut microbiota and microRNAs in the context of IBD. One target was to unravel the intricate dynamics of the gut microbiome-host interplay, with a specific emphasis on identifying a distinct microRNA-bacterial taxa interaction pattern between CD and UC. Moreover, we leveraged state-of-the-art metagenomic and transcriptomic analyses to find pieces of evidence related to the pathogenesis and progression of IBD. As a result, this study not only furthers our understanding of the complex etiology of IBD but also lays the foundation for future research aimed at developing targeted, precision medicine approaches to IBD management. Our findings underscore the critical role of bioinformatics tools and pipelines in deciphering the multifaceted dynamics of microbial communities, thereby offering novel insights into IBD, a disease of global public health significance.

Methods

Metagenomics and host Transcriptomics datasets

This research aims to scrutinize the interplay between human gut metagenomics and host transcriptomics profiles in the context of Inflammatory Bowel Disease (IBD). Our principal objective involves to examine gut bacterial shifts and their associations

with alterations in human microRNA expression in IBD. To achieve this, it is essential to utilize individuals that have both gut metagenomics and host transcriptomics sequenced. In this study, we utilized the IBDMDB database (https://ibdmdb.org/). We applied some filters to select the proper related data. First, the samples were extracted from individuals who had both metagenome and host transcriptome samples. Second, the age of the individual is greater than or equal to 18 years. Third, they have not used antibiotics in the last six months. Fourth, transcriptomic samples were obtained from the sigmoid colon and rectum, focusing on inflamed tissue within the first 4 weeks of the disease. In cases with multiple samples during this period, the one with the shorter duration was chosen.

Following these criteria, we identified 50 samples with paired metagenomic and host transcriptomic data, comprising 8 controls (non-IBD), 18 UCs, and 24 CDs samples (details in Supplementary Table S1). Metagenomic samples were sequenced using the Illumina platform and paired-end sequencing method. Host transcriptome samples are presented as raw read counts, mapped to human genes, reflecting the number of reads associated with each human gene.

Metagenomic samples analysis

Various tools were employed for the comprehensive analysis of metagenomic samples. Initially, quantitative and qualitative assessments of these samples were conducted using FastQC v0.11.9 software. This tool, which takes FASTQ format sequence files as input, provides crucial information such as the total number of sequences, identification of sequences with poor quality, and details regarding GC content. Notably, due to the satisfactory quality of the reads in the samples, no trimming was applied.

MetaPhlAn3 (24) was used to obtain the taxonomic profile of metagenomic samples. This tool takes paired-end FASTQ sequences as input and provides the percent abundance of each microorganism at different taxonomic levels. In this study, we focused on bacteria at the species level. The resulting taxonomic profile, in its raw form, required normalization. Given the inherent sparsity of metagenomic data and potential biases in sampling and sequencing, it was imperative to employ an appropriate normalization method (25). To address this, the metagenomeSeq package (26) was employed, utilizing the cumulative sum scaling (CSS) algorithm (27). Proven to be highly effective for sparse raw abundances (21, 28). This package normalizes and mitigates bias in metagenomic data.

Finally, to pinpoint microorganisms exhibiting significant changes among the non-IBD, UC, and CD groups, the Kruskal-Wallis statistical test was applied. The Benjamini-Hochberg method was subsequently employed to obtain the false discovery rate (29) ensuring robust and accurate results.

Host transcriptomic samples analysis

In this study, human transcriptomic profiles from colon tissue samples associated with IBD were utilized, establishing a one-to-one relationship with metagenomic samples from the IBDMDB database (30). These samples are in the form of a raw data matrix that is mapped to human gene catalog. Each row of this matrix represents the number of reads mapped to each transcript in the human gene catalog. DESeq2 (31) facilitated frequency normalization, proving to be an efficient tool for expression data analysis (21). After data normalization, the Kruskal–Wallis test and Benjamini–Hochberg correction were applied to identify transcripts showing significant differences among the non-IBD, UC, and CD groups.

Co-abundance and co-expression network analysis

In section 2-2, using the Kruskal-Wallis statistical test on metagenomic samples, significant changes in microorganisms between healthy and diseased states were determined independently for non-IBD vs. UC and non-IBD vs. CD groups. Analogous analyses were performed in section 2-3 on host transcriptomic samples. With 50 metagenomic samples (8 non-IBDs, 18 UCs, and 24 CDs) and corresponding host transcriptomic samples, a bipartite co-abundance network (consistent presence and similar abundance patterns of microbial species network) for UC data and CD data was constructed separately. In this section, we do not use non-IBD samples and only used them to candidate species and transcripts in the previous stages as the nodes of the networks. The Pearson correlation coefficient of the frequency of transcripts and species is calculated, and if it is greater than or equal to the threshold, an edge is drawn between two nodes. Finally, we extract the co-abundance sub-network related to microRNAs and species for UC and CD samples.

In sections 2-3, the Kruskal-Wallis statistical test on host transcriptome samples identified significant

transcript changes between healthy and diseased states. Considering the regulatory role of microRNAs, a co-expression network (correlations between gene expression levels across samples) in the UC group was created such that transcripts significantly changing between non-IBD and UC (adjusted p-value ≤ 0.05) were categorized into microRNAs and others, and their Pearson correlation coefficient was calculated. A similar process was applied to the CD group.

The resultant UC and CD networks—co-abundance between species and microRNAs, and co-expression between microRNAs and other transcripts—were then interconnected. This integration produced a tripartite network hierarchically correlating species, microRNAs, and other transcripts.

Pathway Enrichment

For the identification of pathways associated with microRNAs, a pathway enrichment analysis was conducted on correlated transcripts from the previous step using the EnrichR online tool (www.maayanlab.cloud/Enrichr/). EnrichR, a web-based tool, offers summaries of collective gene list functions, providing diverse visualizations for enrichment results. The Kyoto Encyclopedia of Genes and Genomes (KEGG) database, encompassing genomes, enzymatic pathways, and biological chemicals, was employed for pathway enrichment. Results with an adjusted p-value ≤ 0.05 were deemed statistically significant.

Result

Statistical analysis of samples related to IBD disease

The objective of our research is to unveil a meaningful association between the human gut microbiome and the expression of human microRNAs. To achieve this, the comprehensive resource of the IBDMDB database was utilized. (30). In the IBDMDB, multiple datasets pertaining to IBD are available. The study specifically focused on datasets encompassing gut metagenomics and human transcriptomics samples. A fundamental criterion for sample selection was the coexistence of both metagenomic and host transcriptomic data for a given individual. Additional filters were also applied, including age restrictions (age ≥ 18), a six-month antibiotic-free period, and the initial 4-week timeframe of the disease onset. When multiple samples were available for an individual, the one from the earliest week was chosen. Concerning the host transcriptomic samples, only those derived from inflamed areas of the sigmoid colon or rectum were selected.

After adhering to these specific filtration parameters, a collection of 50 paired metagenomic and host transcriptomic samples were procured. This collection consisted of 8 control (non-IBD), 18 UC, and 24 CD samples. Metagenomic samples were in the format of paired-end FASTQ sequences, while host transcriptomic samples were represented as raw abundance matrices, detailing the count of reads mapped to each human mRNA per sample. The quality of the metagenomic samples was meticulously evaluated using FastQC software. The sequencing quality met the acceptable standard for inclusion in the analysis, with metagenomic samples boasting an average of 14.2 million paired reads. Likewise, for the host transcriptomic samples, there was an average of 37.7 million reads mapped to human mRNAs.

Transcripts were aligned to hg19 using TopHat version 2.0.14 (32) and then quantified using htseq-count version 0.6.1 (33). The quality of the samples was evaluated by determining the ratio of reads that mapped to exonic versus non-exonic regions. A sample was deemed to have failed quality control if over 40% of the reads were identified to correspond with non-exonic regions. (30).

To extract the taxonomic profiling of metagenomic samples MetaPhlAn3 tool (24) was used. This process led to the identification of 244 bacterial species with non-zero values at the species level. For normalizing the taxonomic profile, we used the metagenomeSeq package (26) in the R software.

In the next step, we used the Kruskal-Wallis statistical test and the Benjamini-Hochberg method to identify the species that altered significantly in the two healthy and diseased groups, and the species with adjusted p-value ≤ 0.05 were identified. This analysis was done separately for the two groups of non-IBD and UC as well as non-IBD and CD, which obtained 53 and 47 species respectively with adjusted. p-value ≤ 0.05. Figure 1(a) shows the top 10 species for each group.

Erysipelotrichaceae_bacterium (adjusted p-value 0.0016), *Citrobacter_freundii* (adjusted p-value 0.0431) and *Clostridium_clostridioforme* (adjusted p-value 0.0017) were the most significant increased taxa in CD

patients with log2FC 6.54, 4.73 and 4.66 respectively. On the other hand, *Ruminococcus_champanellensis* (adjusted p-value 0.0006), *Dysgonomonas_mossii* (adjusted p-value 0.0008) and *Butyrivibrio_crossotus* (adjusted p-value 1.74E-05) were the most reduced taxa in CD patients with log2FC -20.32, -19.36 and -10.41. In UC patients, *Bacteroides_fluxus* (adjusted p-value 9.70E-05), *Synergistes* (adjusted p-value 6.28E-05) and *Mitsuokella_multacida* (adjusted p-value 0.0005) showed the most increased abundance with log2FC 60.50, 57.33 and 55.66. Moreover, *Bacteroides_barnesiae* (adjusted p-value 0.00409), *Ruminococcus_champanellensis* (adjusted p-value 8.64E-05), and *Coprococcus_eutactus* (adjusted p-value 1.09E-05) significantly decreased in UC patients with log2FC -29.25, -23.36 and -7.56. These findings demonstrate key microbial shifts in the context of IBD.

As explained in the preceding section, the host transcriptome dataset is a raw abundance matrix, indicating the number of reads mapped to human transcriptome catalog in each sample. This matrix encompasses 47,024 RNA coding genes (transcripts). Initial steps involve selecting the 50 samples outlined in the materials section and filtering out transcripts with fewer than five mapped reads. This process yields 38,258 transcripts with at least five reads mapped across the 50 samples. Subsequently, normalization of the resulting 38,258*50 matrix is executed using the DESeq2 package (31) in the R software. Moving forward, the Kruskal-Wallis statistical test and the Benjamini-Hochberg method are employed to identify transcripts with significant changes, considering adjusted p-value ≤ 0.05. This analysis is conducted separately for non-IBD vs. UC and non-IBD vs. CD groups, resulting in the detection of 4,435 and 3,003 transcripts, respectively, with adjusted p-value ≤ 0.05. Within these analyzed transcripts, 14 microRNAs are identified for the non-IBD and UC groups, and 9 microRNAs for the non-IBD and CD groups. Figure 1(b) presents the detected microRNAs alongside adjusted p-value and log2FC. Notably, some microRNAs are common between UC and CD groups compared to non-IBD group. MIR199A1 (adjusted p-value 0.010, log2FC 20.46) and MIR4525 (adjusted p-value 0.026, log2FC -18.25) stand out as the most upregulated and downregulated microRNAs in UC patients, respectively. Similarly, MIR125B1 (adjusted p-value 0.009, log2FC 19.58) and MIR596 (adjusted p-value 0.012, log2FC -2.78) exhibit the most prominent upregulation and downregulation in CD patients.

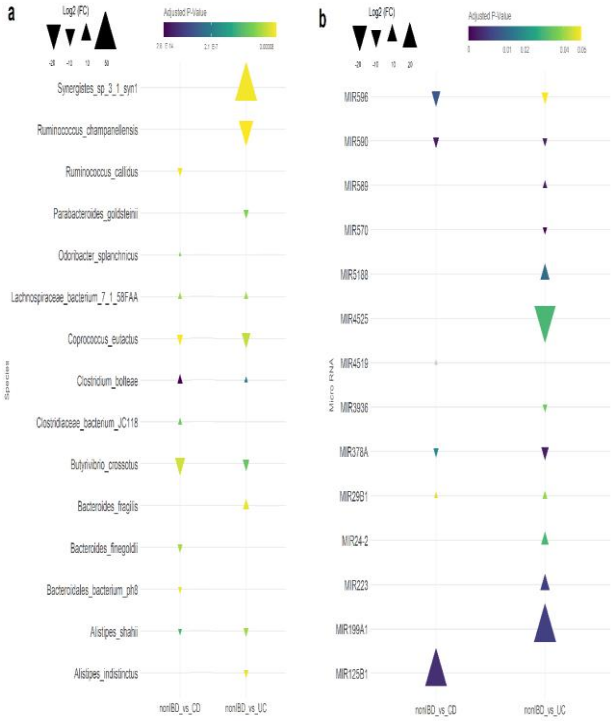

Figure 1. alterations of significant (adjusted p-value ≤ 0.05) species and microRNAs between non-IBD and subtypes of IBD (a) Top 10 species that have change significantly between the non-IBD with Ulcerative colitis (UC) and Crohn's disease (CD) groups based on adjusted p-value, (**b**) The microRNAs that have change significantly (adjusted p-value ≤ 0.05) between the non-IBD with UC and CD groups. The log2 fold change (log2FC) is calculated as log2((UC or CD)/non-IBD). (Log2 fold-changes are represented by triangles, where the size reflects the magnitude of the change and the direction indicates whether bacterial species or microRNAs are upregulated or downregulated. Triangle color corresponds to the adjusted p-value, with dark purple denoting more statistically significant changes. Only the top 10 most significant changes are displayed in this figure.)

Co-abundance and co-expression network between metagenomic and host transcriptomic data

As we explained in section 3-1, after analyzing the metagenomic and host transcriptomic samples, 53 and 47 species, 4435 and 3003 transcripts, and 14 and 9 microRNA were found, respectively, which were significant changes between two groups of non-IBD and UC as well as non-IBD and CD. Now we create a

bipartite co-abundance network between species and candidate transcripts. This network is created independently for UC and CD data. If the Pearson correlation coefficient is greater than or equal to 0.7 or less than or equal to -0.7, an edge is created between the species and transcript. This network has 53 and 4435 nodes (for species and transcripts, respectively) and 4004 edges for UC. Also, the network related to CD has 47 and 3003 nodes and 2879 edges. A very important point in these networks is that no negative correlation between species and transcripts was observed for UC and CD patients. Now we separate the sub-network related to species and microRNAs. Table 1 shows the relationship between species and microRNAs and their Pearson correlation coefficient for UC and CD .The interesting point is that there is no common species and microRNAs between the UC and CD groups, but in the CD group, most of the species, have a significant correlation with several microRNAs.

Also, for host transcriptomic data, there are 4435 and 3003 transcripts, which are significantly changed between the two groups of non-IBD and UC, as well as non-IBD and CD, respectively. We divide these transcripts into two groups, microRNA and other transcripts, respectively, and draw a co-expression network with a Pearson correlation coefficient greater than or equal to 0.8 and less than or equal to -0.8. Then we separate the sub-network related to micro RNAs (details in Supplementary Table S2).

As depicted in Figure 2 (b), the UC network is split into 3 separate parts including a microRNA and its co-expressed transcripts and its co-abundant species. But in CD group (Figure 2 (a)), we face a more complex network with denser relationships.

Pathway enrichment

Based on enrichment results for microRNAs' correlated genes, IBD related pathways including IL-17, TNF and PI3K-Akt signaling pathways were enriched for MIR29B1 and MIR125B1 in CD patients. Furthermore, B cell receptor, NOD-like receptor, Toll-like receptor and IL-17 signaling pathways for MIR29B1, Cytokine-cytokine receptor interaction, PI3K-Akt, TNF and IL-17 signaling pathways for MIR223, and Chemokine, B cell receptor, Cytokine-cytokine receptor, C-type lectin receptor and TNF signaling pathways for MIR589 were enriched in UC patients (adjusted p-value <0.05).

Discussion

With the development of high-throughput sequencing technologies, host-metagenome interactions were studied more comprehensively, and new insights about its regulatory effect on host immunity and inflammatory responses were achieved (34). Clinical microbiome

Table 1. This table presents the Pearson correlation coefficients reflecting associations between microbial species and microRNAs in ulcerative colitis (UC) and Crohn's disease (CD) groups.

UC			CD		
Species	MicroRNA	Pearson correlation coefficient	Species	MicroRNA	Pearson correlation coefficient
Anaerotruncus_colihominis	MIR199A1	0.78	Alistipes_onderdonkii	MIR125B1	0.72
Bilophila_wadsworthia		0.71	Bacteroidales_bacterium_ph8		0.97
Burkholderiales_bacterium_1_1_47		0.74	Bacteroides_xylanisolvens		0.82
Clostridium_bolteae		0.84	Oxalobacter_formigenes		0.73
Clostridium_scindens		0.84	Ruminococcus_callidus		0.95
Clostridium_sp_KLE_1755		0.84	Alistipes_onderdonkii	MIR29B1	0.83
Eggerthella_lenta		0.84	Bacteroidales_bacterium_ph8		0.97
Eggerthella_unclassified		0.81	Bacteroides_xylanisolvens		0.80
Lachnospiraceae_bacterium_5_1_57FAA		0.84	Ruminococcus_callidus		0.94
Lachnospiraceae_bacterium_6_1_63FAA		0.84	Alistipes_sp_HGB5	MIR596	0.99
Lachnospiraceae_bacterium_9_1_43BFAA		0.84			
Alistipes_finegoldii	MIR24-2	0.90			
Bacteroides_stercoris		0.75			
Clostridium_asparagiforme		0.75			
Oscillibacter_unclassified		0.87			
Clostridium_bartlettii	MIR3936	0.81			
Streptococcus_vestibularis		0.81			

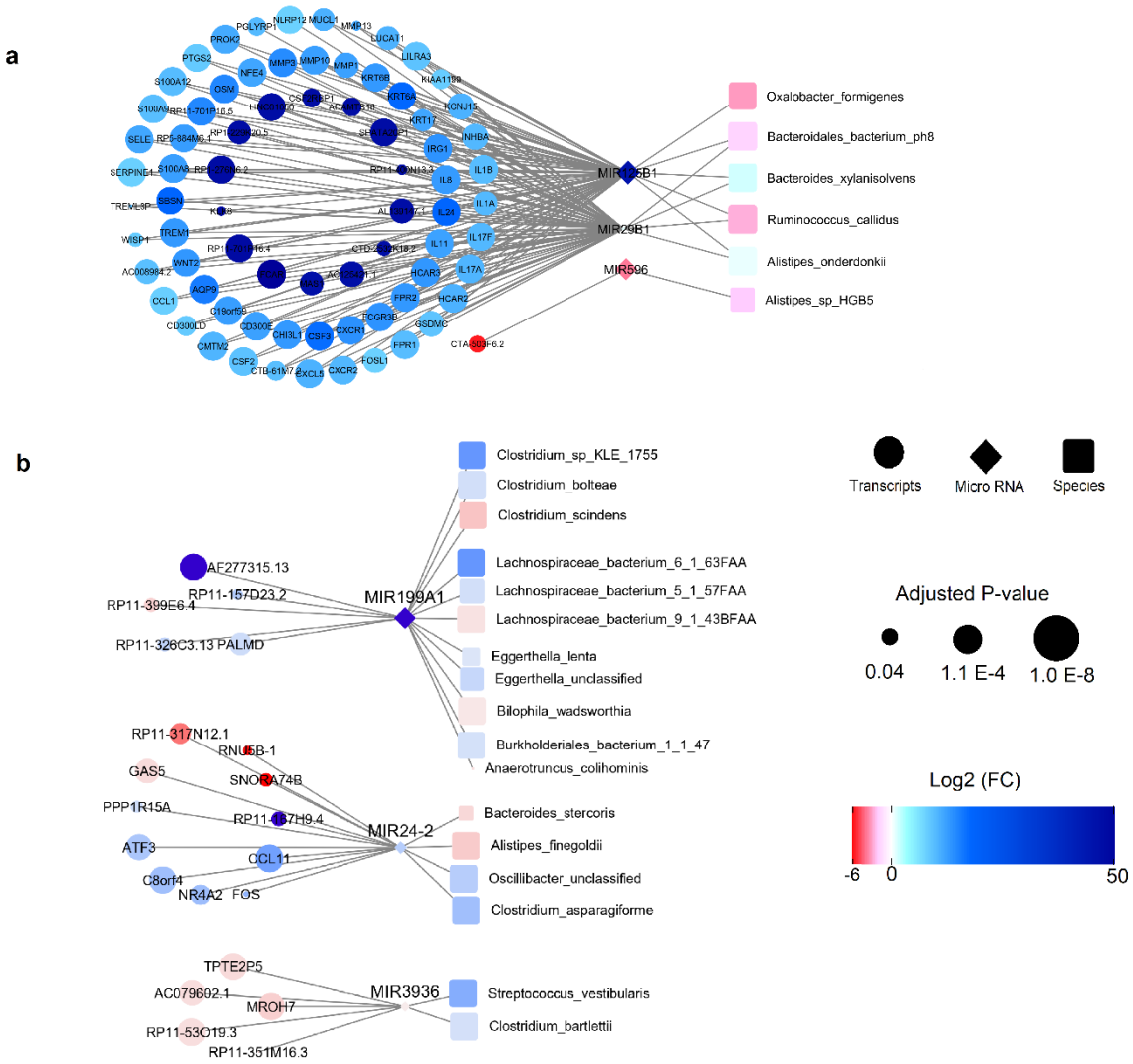

Figure 2: Three layer co-abundant network between microbial taxa, microRNA and other transcript (a) for CD patients and (b) for UC patients. The edges represent the co-abundant relationship between the two adjacent nodes. In the networks the edges show the relationship with more than 0.9 co-abundance coefficient. (Since the number of other transcripts in CD network (a), was very large, only transcripts with |Log2FC|>=4 was included in the network). (The shape of each node represents its type: bacterial species (squares), microRNAs (rhombuses), and mRNA transcripts (circles). Node size corresponds to the significance level (adjusted p-value), while color indicates the log2 fold change (Log2 FC): red denotes a decrease, and blue indicates an increase.)

analysis with 16S rRNA gene sequencing provided an in-depth characterization of the microbiome and its composition in diseased and healthy subjects (35). In this study, we analyzed metagenome data from stool samples alongside their matched colon tissue transcriptome profile in IBD patients. We aimed to discover the gut microbiome composition changes and their association with dysregulated transcriptomes, specifically microRNAs -as a part of transcription regulators in IBD. Learning about the gut microbiome and host microRNA relationship could lead to a better understanding of metagenome impact on cellular processes and transcription by microRNAs and vice versa.

Based on our metagenome data analysis, a shift in a gut microbiome and bacterial abundance have been observed in CD and UC patients compared to controls. A recent study found that UC patients who responded to adalimumab had higher levels of *Faecalibacterium prausnitzii* and *Bacteroides uniformis*, while non-responders showed increased *Escherichia coli* and *Ruminococcus gnavus*. These microbial differences suggest potential biomarkers for predicting treatment

response (36). Moreover, another study reported that, *Erysipelotrichaceae_bacterium* significantly increased in CD patients. *Erysipelotrichaceae* are highly enriched in colorectal cancer and presented as an inflammation-related bacteria. However, controversial reports about its abundance in IBD patients have been disclosed (37). Schaubeck and colleagues observed an increase in the abundance of *Erysipelotrichaceae* in TNF-driven Crohn's disease (CD)-like mice model (38). On the other hand, Dey and colleagues found lower levels of *Erysipelotrichaceae* in recurrent CD patients (39). *Ruminococcus_champanellensis* with the most abundance depletion in CD and also a great decrease in UC is a member of Firmicutes phylum and shows cellulose-degrading activity (40). The ability of pectin degradation in these bacteria could lead to anti-inflammatory activity and promote IL-10 production (41). Thus, the decreased number of *Ruminococcus_champanellensis* with anti-inflammatory effects may worsen the gut chronic inflammation in IBD patients. *Bacteroides fluxus* is a gram-negative anaerobic bacillus from *Bacteroides* spp. and our analysis indicates it's around 60-fold increase in fecal samples of UC patients. There isn't enough evidence around its role in the gut, however, a case study reported that *Bacteroides fluxus* is associated with abdominal infection (42). Furthermore, *Coprococcus eutactus*, known as a suggested probiotic with the potential of reducing pro-inflammatory cytokines and promoting anti-inflammatory factors such as IL-10 (43), showed a significant decrease in UC fecal samples compared to controls.

The connection between microRNAs and gut microbiota has been reported previously. Due to these studies, in dysbiotic conditions, changes in microRNAs expression profile affect Intestinal Epithelial Cell (IEC) regulation and epithelial barrier function (44). In contrast, IECs can produce secretory microRNAs in the gut lumen and target bacterial mRNA, regulating their expression and function (44). Based on bioinformatics analysis results, in colorectal cancer (CRC), several dysregulated microRNAs are correlated with the microbiome abundance in the tumor microenvironment and regulate genes that control cellular interactions with microbes (18). Also, *Fusobacterium nucleatum*, *Escherichia coli*, enterotoxigenic *Bacteroides fragilis*, and *Faecalibacterium prausnitzii* reported to be associated with microRNA expression and tumor progression in CRC (45). This study also reported a positive correlation between some microbiome changes and alterations in microRNA expression in CD and UC patients. In our analysis, *Alistipes finegoldii* showed a strong correlation (Pearson correlation coefficient: 0.9) with MIR24-2 in UC patients. *Alistipes finegoldii* was reported as a driver microbiome for colitis and carcinogenesis (46) and MIR-24-2 has a regulatory effect on MAPK signaling pathway as reported in previous studies (47). On the other hand, MIR199A, which resulted in the most upregulated microRNA in UC patients, has the potential role in regulating WNT, ERK and hypoxia-inducible factor (Hif)-1α pathways involved in differentiation, inhibiting cell proliferation and apoptosis respectively (48-50). MIR199A correlated with *Bilophila wadsworthia*, which can cause inflammation in mice models (51) and *Clostridium_bolteae* abundance in UC patients. *Clostridium_bolteae* is known as an autism-associated Bacterium (52) however studies showed its role in insulin resistance, dyslipidemia and inflammation in obese people (53). In contrast, another study reported the anti-inflammatory role of *Clostridium bolteae* AHG0001 cell-free culture supernatant by suppressing NF-κkB factor in organoid culture (54). *Eggerthella_lenta* enriched in IBD patients gut microbiome and worsens colitis in mice model (55) and showed a positive correlation with MIR199A in UC. In CD patients, MIR125B1, MIR29B1 and MIR596 correlation with microbiome changes were detected. MIR125B1 with the highest overexpression in CD, can control cell proliferation, differentiation and immunity by regulating NF-κB and Wnt pathways (56) and is considered a potential biomarker in UC (57). MIR125B1 targets the 3'UTR of TLR4, one of the important cell surface molecules in detecting microbiome composition by the host (58). Besides, MIR29 family including MIR29B1 showed a critical role in immune response regulation in several studies and regulates Th1 differentiation and IFNϒ pathway (59). Our enrichment results suggested MIR125B1 and MIR29B1 impact on controlling inflammation by regulation of IL-17 and TNF signaling. *Alistipes_onderdonkii* is abundant in CD patients and associated with MIR125B1 and MIR29B1 expression, based on this study. There is contrasting evidence about the role of *Alistipes* and they are considered as a

protective or cancer progression factor in several diseases and inflammatory conditions (60) and their correlation with MIR596 was also detected. Moreover, *Alistipes onderdonkii* was introduced as a CRC tumor tissue-specific bacterial biomarker (61). *Bacteroides_xylanisolvens* as another correlated species with the MIR125B1 and MIR29B1 in CD patients is under investigation as a next generation probiotic with xylanolytic activity (62).

A three-layered network between the bacterial taxa, microRNA and related transcripts showed the indirect regulatory effect of dysbiosis on cell function and pathways. The findings presented in Figure 2 reveal a consistent three-layered network structure in CD patients. Our analysis indicates that nearly all bacterial species display connections with dysregulated microRNAs, which subsequently regulate a set of mRNAs. In contrast, our investigation of ulcerative UC patients yielded three distinct networks featuring unique bacterial species, microRNAs, and related mRNAs. These results suggest that the two conditions may differ in their underlying molecular mechanisms. Further research is necessary to elucidate the precise nature of these differences. Metagenome-correlated microRNAs were in charge for controlling various genes including immune response related genes. In CD patients, *Alistipes onderdonkii*, *Oxalobacter formigenes*, *Ruminococcus callidus*, Bacteroides xylanisolvens and *Bacteroidales bacterium_ph8* / MIR125B1 MIR125B1 and MIR29B1 MIR29B1 axis controls IL17A, IL17F, IL8 and IL1B expression. These cytokines are known to trigger inflammation signaling and drive chronic inflammation in UC and CD patients and previous studies reported the link between gut microbiome composition and cytokine production in gut (63). Furthermore, this axis correlates with MMP1, MMP10 and MMP3 which is assumed to be involve in fistulae formation in CD patients (64). In UC patients, each microRNA correlates with a unique set of bacterial taxa and transcripts which most of them are non-code RNAs. These networks clearly represent the axis between bacterial taxa/microRNA/transcripts and the regulatory effect of metagenome and microRNAs on cell function and inflammation pathways.

Although our study has yielded some valuable insights, it is important to note that we encountered certain limitations along the way. For instance, we were unable to procure an adequate number of samples with matching transcriptome and metagenome data, which may have limited the robustness of our results. Moreover, in this study, the method used for microRNA detection was based on total transcriptome analysis. However, the database did not utilize a specific library for detecting microRNAs, which could have provided a wider range of microRNAs and improved the accuracy of the results. By not using a specialized library, some microRNAs may have been missed, and the overall variability of the detected microRNAs may have been limited. Additionally, the presence of abundant ribosomal and mRNAs can further dilute the detection of microRNAs. Therefore, including a specific library for microRNA detection could potentially enhance the study's findings and provide a more comprehensive understanding of the microRNA expression profile. Nonetheless, our research has yielded valuable insights into the pathogenesis of IBD by establishing an innovative connection between the gut microbiome, microRNA, and transcriptome in CD and UC patients. Our findings serve as a solid foundation for future studies in this area and have the potential to greatly enhance our understanding of the underlying mechanisms of IBD. The unique relationship that we have uncovered between these factors provides a promising avenue for further exploration.

Based on the insights from previous studies and results obtained from our bioinformatics analysis, it is suggested that the interaction between microRNA and the gut microbiome may play a significant role in the pathogenesis of IBD. It seems that changes in the gut microbiome composition alongside microRNA expression alterations could lead to dysregulation of proliferation, inflammation, and differentiation pathways and cause IBD or worsen the inflammation condition in the gut. A compelling topic for additional exploration could involve investigating the cause-and-effect dynamics within this relationship. It remains unclear whether changes in microRNAs are a precursor to alterations in the microbiota, or conversely, if dysbiosis could trigger microRNA dysregulation and subsequent transcriptomic alterations. However, this interaction between microRNA and the microbiome could be a possible biomarker for predicting gut microbiome or microRNA expression shifts. Also, distinct correlated microRNA and microbiome changes in CD and UC could have the potential to use as a marker for

distinguishing CD and UC. However, further investigations are warranted to validate these in silico findings in vitro and provide a clearer understanding of the precise interactions between bacteria and microRNA. By doing so, we can better comprehend the complex mechanisms underlying IBD and, potentially, pave the way for novel diagnostic and therapeutic strategies.

Acknowledgement

The authors would like to acknowledge the support of the Research Institute for Gastroenterology and Liver Diseases (RIGLD), Shahid Beheshti University of Medical Sciences, Tehran, Iran and the Department of Bioinformatics, Institute of Biochemistry and Biophysics (IBB), University of Tehran, Iran. This research was supported by RIGLD (grant number: IR.SBMU.RIGLD.REC.1400.0.002).

Data Availability

The metagenomic and transcriptomic data have been downloaded from the IBDMDB database (https://ibdmdb.org) with some criteria which is mentioned in the materials and methods section and the ID of patients are reported in the Supplementary Table_S1.

Conflict of interests

There is no conflict of interest for authors of this article.

References

1. Mulder DJ, Noble AJ, Justinich CJ, Duffin JM. A tale of two diseases: the history of inflammatory bowel disease. J Crohns Colitis 2014;8:341-8.

2. Barnes EL, Kappelman MD. Editorial: Increasing incidence of pediatric inflammatory bowel disease in France: Implications for etiology, diagnosis, prognosis, and treatment. Am J Gastroenterol 2018;113:273-275.

3.Rostami Nejad M, Rostami K, Yamaoka Y, Mashayekhi R, Molaei M, Dabiri H, et al. Clinical and histological presentation of Helicobacter pylori and gluten related gastroenteropathy. Arch Iran Med 2011;14:115-8.

4. Ashtari S, Pourhoseingholi MA, Rostami K, Aghdaei HA, Rostami-Nejad M, Busani L, et al. Prevalence of gluten-related disorders in Asia-Pacific region: a systematic review. J Gastrointestin Liver Dis 2019;28:95-105.

5. Zuo T, Ng SC. The gut microbiota in the pathogenesis and therapeutics of inflammatory bowel disease. Front Microbiol 2018;9:2247.

6.Alhagamhmad MH, Day AS, Lemberg DA, Leach ST. An update of the role of nutritional therapy in the management of Crohn's disease. J Gastroenterol 2012;47:872-82.

7. Belkaid Y, Hand TW. Role of the microbiota in immunity and inflammation. Cell 2014;157:121-41.

8. 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;104:13780-5.

9. Dai R, Ahmed SA. MicroRNA, a new paradigm for understanding immunoregulation, inflammation, and autoimmune diseases. Transl Res 2011;157:163-79.

10. Gebert LFR, MacRae IJ. Regulation of microRNA function in animals. Nat Rev Mol Cell Biol 2019;20:21-37.

11. Amini-Farsani Z, Yadollahi-Farsani M, Arab S, Forouzanfar F, Yadollahi M, Asgharzade S. Prediction and analysis of microRNAs involved in COVID-19 inflammatory processes associated with the NF-kB and JAK/STAT signaling pathways. Int Immunopharmacol 2021;100:108071.

12. Lou C, Li Y. Functional role of microRNA-135a in colitis. J Inflamm 2018;15:7.

13. Bocchetti M, Ferraro MG, Ricciardiello F, Ottaiano A, Luce A, Cossu AM, et al. The role of microRNAs in development of colitis-associated colorectal cancer. Int J Mol Sci 2021;22:3967.

14. Wani S, Man Law IK, Pothoulakis C. Role and mechanisms of exosomal miRNAs in IBD pathophysiology. Am J Physiol Gastrointest Liver Physiol 2020;319:G646-G654.

15. Turnbaugh PJ, Ley RE, Hamady M, Fraser-Liggett CM, Knight R, Gordon JI. The human microbiome project. Nature 2007;449:804-10.

16. Kazemifard N, Dehkohneh A, Baradaran Ghavami S. Probiotics and probiotic-based vaccines: A novel approach for improving vaccine efficacy. Front Med 2022;9:940454.

17.Masi L, Capobianco I, Magrì C, Marafini I, Petito V, Scaldaferri F. MicroRNAs as innovative biomarkers for inflammatory bowel disease and prediction of colorectal cancer. Int J Mol Sci 2022;23:7991.

18. Yuan C, Burns MB, Subramanian S, Blekhman R. Interaction between host MicroRNAs and the gut microbiota in colorectal cancer. mSystems 2018;3:e00205-17.

19. Zhang L, Chen F, Zeng Z, Xu M, Sun F, Yang L, et al. Advances in metagenomics and its application in environmental microorganisms. Front Microbiol 2021;12:766364.

20. Beghini F, McIver LJ, Blanco-Míguez A, Dubois L, Asnicar F, Maharjan S, et al. Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3. Elife 2021;10:e65088.

21. Norouzi-Beirami MH, Marashi SA, Banaei-Moghaddam AM, Kavousi K. CAMAMED: a pipeline for composition-aware mapping-based analysis of metagenomic data. NAR Genom Bioinform 2021;3:lqaa107.

22. Quijada NM, Dzieciol M, Schmitz-Esser S, Wagner M, Selberherr E. Metatranscriptomic analyses unravel dynamic changes in the microbial and metabolic transcriptional profiles in artisanal austrian hard-cheeses during ripening. Front Microbiol 2022;13:813480.

23. Deshpande D, Chhugani K, Chang Y, Karlsberg A, Loeffler C, Zhang J, et al. RNA-seq data science: From raw data to effective interpretation. Front Genet 2023;14:997383.

24. Beghini F, McIver LJ, Blanco-Míguez A, Dubois L, Asnicar F, Maharjan S, et al. Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3. Elife 2021;10:e65088.

25. Norouzi-Beirami MH, Marashi SA, Banaei-Moghaddam AM, Kavousi K. Beyond taxonomic analysis of microbiomes: a functional approach for revisiting microbiome changes in colorectal cancer. Front Microbiol 2020;10:3117.

26. Paulson JN, Stine OC, Bravo HC, Pop M. Differential abundance analysis for microbial marker-gene surveys. Nat Methods 2013;10:1200-2.

27. Tsilimigras MC, Fodor AA. Compositional data analysis of the microbiome: fundamentals, tools, and challenges. Ann Epidemiol 2016;26:330-5.

28. Lee C, Lee S, Park T. A comparison study of statistical methods for the analysis metagenome data. In: 2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Kansas City, MO, USA, 2017, pp. 1777-1781.

29. Corder GW, Foreman DI. Nonparametric statistics for non-statisticians. John Wiley Sons, Hoboken 2009:99-105.

30. 'https://ibdmdb.org/'.

31. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 2014;15:550.

32. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol 2013;14:R36.

33. Anders S, Pyl PT, Huber W. HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics 2015;31:166-9.

34. Hooper LV, Littman DR, Macpherson AJ. Interactions between the microbiota and the immune system. Science 2012;336:1268-73.

35. Chiu CY, Miller SA. Clinical metagenomics. Nat Rev Genet 2019;20:341-355.

36. Ghavami SB, Moshiri A, Bonaretti C, Farmani M, Squillario M, Di Marco E, et al. Predictive microbial markers distinguish responders and non-responders to adalimumab: a step toward precision medicine in ulcerative colitis. Microorganisms 2025;13:1941.

37. Kaakoush NO. Insights into the role of erysipelotrichaceae in the human host. Front Cell Infect Microbiol 2015;5:84.

38. 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;65:225-37.

39. Dey N, Soergel DA, Repo S, Brenner SE. Association of gut microbiota with post-operative clinical course in Crohn's disease. BMC Gastroenterol 2013;13:131.

40. Chassard C, Delmas E, Robert C, Lawson PA, Bernalier-Donadille A. Ruminococcus champanellensis sp. nov., a cellulose-degrading bacterium from human gut microbiota. Int J Syst Evol Microbiol 2012;62:138-143.

41. Chung WSF, Meijerink M, Zeuner B, Holck J, Louis P, Meyer AS, et al. Prebiotic potential of pectin and pectic oligosaccharides to promote anti-inflammatory commensal bacteria in the human colon. FEMS Microbiol Ecol 2017;93.

42. Cobo F, Pérez-Carrasco V, Gómez-Vicente E, Martín-Hita L, García-Salcedo JA, Navarro-Marí JM. First case of abdominal infection caused by bacteroides fluxus. Anaerobe 2021;69:102363.

43. Yang R, Shan S, Shi J, Li H, An N, Li S, et al. Coprococcus eutactus, a Potent Probiotic, Alleviates Colitis via Acetate-Mediated IgA Response and Microbiota Restoration. J Agric Food Chem 2023.

44. Behrouzi A, Ashrafian F, Mazaheri H, Lari A, Nouri M, Riazi Rad F, et al. The importance of interaction between MicroRNAs and gut microbiota in several pathways. Microb Pathog 2020;144:104200.

45. Xing J, Liao Y, Zhang H, Zhang W, Zhang Z, Zhang J, et al. Impacts of MicroRNAs induced by the gut microbiome on regulating the development of colorectal cancer. Front Cell Infect Microbiol 2022;12:804689.

46. Yang Y, Jobin C. Novel insights into microbiome in colitis and colorectal cancer. Curr Opin Gastroenterol 2017;33:422-427.

47. Meng L, Yuan L, Ni J, Fang M, Guo S, Cai H, et al. Mir24-2-5p suppresses the osteogenic differentiation with Gnai3 inhibition presenting a direct target via inactivating JNK-p38 MAPK signaling axis. Int J Biol Sci 2021;17:4238-4253.

48. Alexander MS, Kawahara G, Motohashi N, Casar JC, Eisenberg I, Myers JA, et al. MicroRNA-199a is induced in dystrophic muscle and affects WNT signaling, cell proliferation, and myogenic differentiation. Cell Death Differ 2013;20:1194-208.

49. Hou J, Lin L, Zhou W, Wang Z, Ding G, Dong Q, et al. Identification of miRNomes in human liver and hepatocellular carcinoma reveals miR-199a/b-3p as therapeutic target for hepatocellular carcinoma. Cancer Cell 2011;19:232-43.

50. Rane S, He M, Sayed D, Vashistha H, Malhotra A, Sadoshima J, et al. Downregulation of miR-199a derepresses hypoxia-inducible factor-1alpha and Sirtuin 1 and recapitulates hypoxia preconditioning in cardiac myocytes. Circ Res 2009;104:879-86.

51. Feng Z, Long W, Hao B, Ding D, Ma X, Zhao L, et al. A human stool-derived Bilophila wadsworthia strain caused systemic inflammation in specific-pathogen-free mice. Gut Pathog 2017;9:59.

52. Pequegnat B, Monteiro MA. Carbohydrate scaffolds for the study of the autism-associated bacterium, clostridium bolteae. Curr Med Chem 2019;26:6341-6348.

53. Campisciano G, de Manzini N, Delbue S, Cason C, Cosola D, Basile G, et al. The obesity-related gut bacterial and viral dysbiosis can impact the risk of colon cancer development. Microorganisms 2020;8:431.

54. Giri R, Hoedt EC, Khushi S, Salim AA, Bergot AS, Schreiber V, et al. Secreted NF-κB suppressive microbial metabolites modulate gut inflammation. Cell Rep 2022;39:110646.

55. Alexander M, Ang QY, Nayak RR, Bustion AE, Sandy M, Zhang B, et al. Human gut bacterial metabolism drives Th17 activation and colitis. Cell Host Microbe 2022;30:17-30.e9.

56. Wang Y, Zeng G, Jiang Y. The emerging roles of miR-125b in cancers. Cancer Manag Res 2020;12:1079-1088.

57. Coskun M, Bjerrum JT, Seidelin JB, Troelsen JT, Olsen J, Nielsen OH. miR-20b, miR-98, miR-125b-1*, and let-7e* as new potential diagnostic biomarkers in ulcerative colitis. World J Gastroenterol 2013;19:4289-99.

58. Belville C, Ponelle-Chachuat F, Rouzaire M, Gross C, Pereira B, Gallot D, Sapin V, Blanchon L. Physiological TLR4 regulation in human fetal membranes as an explicative mechanism of a pathological preterm case. Elife 2022;11:e71521.

59. Horita M, Farquharson C, Stephen LA. The role of miR-29 family in disease. J Cell Biochem 2021;122:696-715.

60. Parker BJ, Wearsch PA, Veloo ACM, Rodriguez-Palacios A. The genus alistipes: gut bacteria with emerging implications to inflammation, cancer, and mental health. Front Immunol 2020;11:906.

61. Hasan R, Bose S, Roy R, Paul D, Rawat S, Nilwe P, et al. Tumor tissue-specific bacterial biomarker panel for colorectal cancer: Bacteroides massiliensis, Alistipes species, Alistipes onderdonkii, Bifidobacterium pseudocatenulatum, Corynebacterium appendicis. Arch Microbiol 2022;204:348.

62. Tan H, Zhai Q, Chen W. Investigations of Bacteroides spp. towards next-generation probiotics. Food Res Int 2019;116:637-644.

63. Schmitt H, Neurath MF, Atreya R. Role of the IL23/IL17 pathway in Crohn's disease. Front Immunol 2021;12:622934.

64. Kirkegaard T, Hansen A, Bruun E, Brynskov J. Expression and localisation of matrix metalloproteinases and their natural inhibitors in fistulae of patients with Crohn's disease. Gut 2004;53:701-9.