



Altered gut bacterial and metabolic signatures and their interaction in inflammatory bowel disease

Guangcai Liang

School of Chemical Engineering and Technology, Tianjin University, Yaguan Road 135, Jinnan District, Tianjin, 300350, China

ARTICLE INFO

Keywords:

Inflammatory bowel disease
Gut microbiome
Metabolic pathways
Machine learning
Diagnosis

ABSTRACT

Dysregulation of the gut microbiome has been implicated in the progression of many diseases. This study explored the role of microbial and metabolic signatures, and their interaction between the Human inflammatory bowel disease (IBD) and healthy controls (HCs) based on the combination of machine learning and traditional statistical analysis, using data collected from the Human Microbiome Project (HMP) and the Integrative Human Microbiome Project (iHMP). It was showed that the microbial and metabolic signatures of IBD patients were significantly different from those of HCs. Compared to HCs, IBD subjects were characterized by 25 enriched species and 6 depleted species. Furthermore, a total of 17 discriminative pathways were identified between the IBD and HC groups. Those differential pathways were mainly involved in amino acid, nucleotide biosynthesis, and carbohydrate degradation. Notably, co-occurrence network analysis revealed that non-predominant bacteria *Ruminococcus obeum* and predominant bacteria *Faecalibacterium prausnitzii* formed the same broad and strong co-occurring relationships with pathways. Moreover, the essay identified a combinatorial marker panel that could distinguish IBD from HCs. Receiver Operating Characteristic (ROC) and Decision Curve Analysis (DCA) confirmed the high accuracy (AUC = 0.966) and effectiveness of the model. Meanwhile, an independent cohort used for external validation also showed the identical high efficacy (AUC = 0.835). These findings showed that the gut microbes may be relevant to the pathogenesis and pathophysiology, and offer universal utility as a non-invasive diagnostic test in IBD.

1. Introduction

IBD, the most common complication in adults, encompassing both Crohn's disease (CD) and ulcerative colitis (UC), is a weakening and multifactorial complex disease characterized by abnormalities in the immune system, intestinal homeostasis disorders, and changes in gut microbiota [1]. Periods of clinical remission and disease flare-ups is commonly observed in IBD patients and seriously affects the survival and quality of life [2].

Currently, the underlying molecular basis of IBD remains mostly obscure, although several hypotheses have attempted to explain its pathophysiological mechanisms. However, accumulating evidence suggests that the gut microbiota in the pathogenesis and therapeutics as crucial factors in the Pathogenesis of IBD: The germfree state reduces severity of symptoms in animal models of chronic intestinal inflammation [3,4]; the reactivity to intestinal microbes have been changed in humans with IBD [5], and the antibodies in the blood are widely known biomarker in CD and UC, but the antibodies are only present as the

intestine microbiota altered [6]. Although the cause of IBD remain largely unclear, strategies to reduction of microbiota exposure (such as elemental diets or antibiotic therapy) can have beneficial effects against particular cases of IBD [7–9].

In the study, metagenomic datasets generated by the HMP [10,11] and iHMP [12] were used to investigate role of microbial signatures and metabolic signatures and their interaction between the IBD and HCs. Furthermore, the data was sought to reveal the relationship between specific microbial composition, abundance and metabolic pathways, as well as how regulated severity of IBD, either directly or indirectly by adjusting the gut microbiota, the co-occurrence network analysis were performed. Moreover, to apply the expected differences to clinical practice, a risk classification method involving machine-learning algorithms to was proposed to assess IBD risk based on the biological significance and accuracy contribution of bacterial genus in the cohort, and the accuracy of model was further verified using an external independent cohort, which showed the identical high accuracy.

Peer review under responsibility of KeAi Communications Co., Ltd.

E-mail address: lianguangcai@tju.edu.cn.

<https://doi.org/10.1016/j.synbio.2021.10.003>

Received 18 August 2021; Received in revised form 30 September 2021; Accepted 9 October 2021

2405-805X/© 2021 The Authors. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC

BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

2. Methods

2.1. Metagenomic sequence analysis

Metagenomic cohort datasets with a total 28 patients with IBD and 35 healthy individuals and another an independent cohort including 357 patients with IBD and 81 healthy individuals for external validation were obtained through the NIH Human Microbiome Project database (<http://portal.hmpdacc.org/>). Taxonomic profiles of the microbiome was analyzed using MetaPhlan2 pipeline [13] phylogenetic clade identification (default parameters). Functional and pathway composition was performed with HUMAnN2 pipeline [14] using the UniRef90 database and MetaCyc database. Alpha diversity analysis, nonparametric multi-dimensional scaling (NMDS) and principal co-ordinates analysis (PCoA) were conducted and visualized using the vegan and ggplot2 packages in R. In addition, the permutational multivariate analysis of variance (PERMANOVA) and were used to test group differences. The

linear discriminant analysis effect size (LEfSe) [15] was used to identify the differential taxonomic classification of microbial between the two groups with a linear discriminant analysis (LDA) score >2 .

2.2. Statistical analysis

Statistical analyses were carried out using R v4.0.2 and STAMP [16]. Random forest classifier (R package) was utilized to identify biomarkers, and construct model for identifying potential microorganisms-disease associations between IBD and HCs [17]. Internal validation was performed to tune and validate the model by 5 times repeat, 10-fold cross-validation. The microbial multivariate associations analysis was performed via logistic regression. The receiver operating characteristic (ROC) curve was used to obtain accuracy of the constructed score model, then the result used the area under the ROC curve (AUC) and Decision Curve Analysis (DCA) [18] as the metrics for evaluation. The statistical analysis were carried out using R with level of statistical significance:

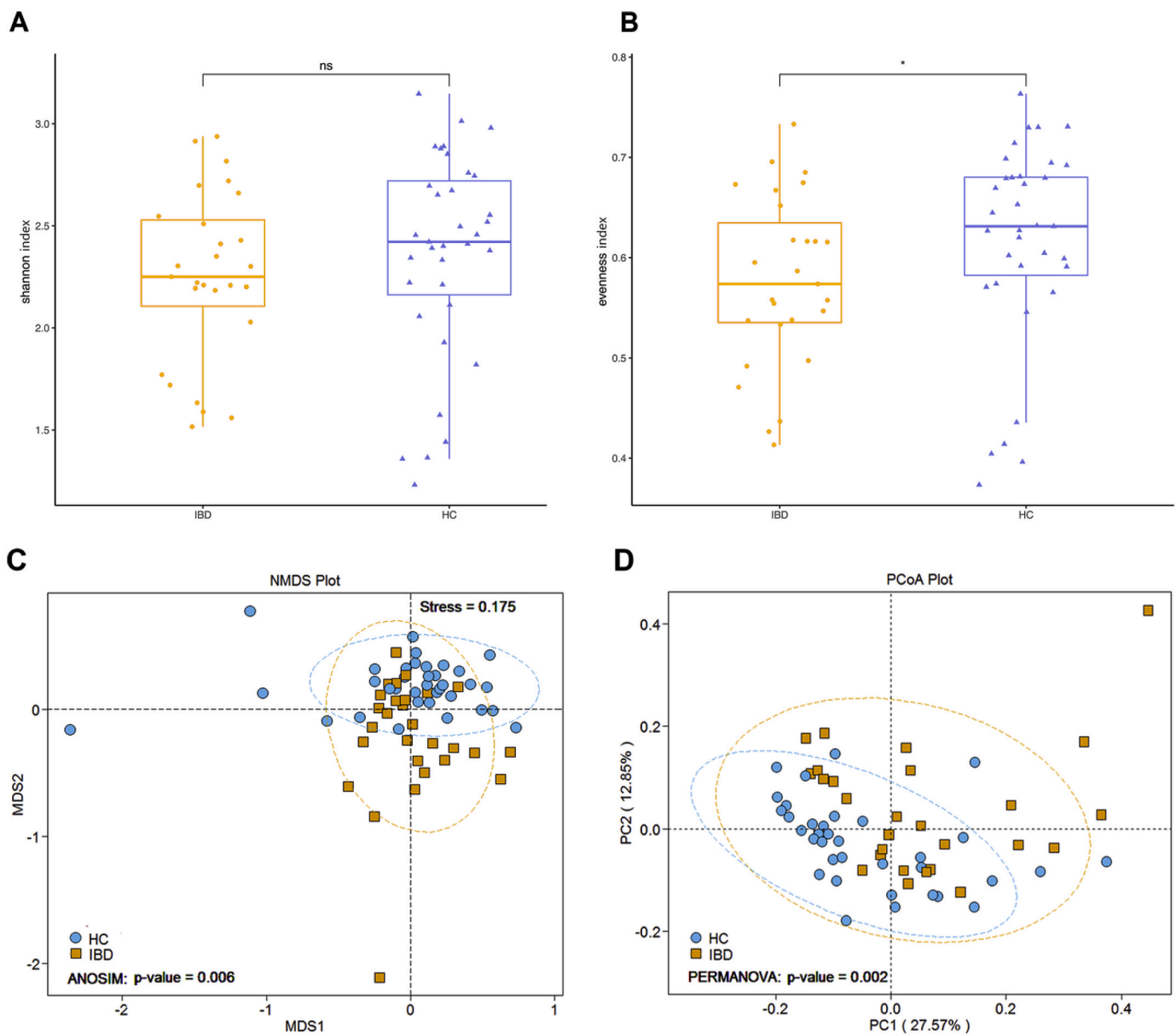


Fig. 1. Altered gut microbiome signatures in IBD versus HCs. (A) There were no significance shannon index. differences between the two groups (Wilcoxon, ns: $P > 0.05$). (B) Species evenness between the two groups were significantly different (Wilcoxon, *: $P < 0.05$). (C) NMDS analysis (ANOSIM, $P = 0.006 < 0.01$) and (D) PCoA analysis (PERMANOVA, $P = 0.002 < 0.01$) showed bacterial communities between the two groups were extremely significantly distinct based on Bray-Curtis distance.

p-value < 0.05. The co-occurrence network was constructed based on the relative abundance of bacterial species and MetaCyc pathways using the Spearman's correlation coefficient ($r > 0.6$ or < -0.6 ; $P < 0.05$). The generated co-occurrence network was visualized in Cytoscape [19].

3. Results

3.1. Gut microbiome differences between IBD and HC subjects

Here, metagenomic sequencing was used to compare the microbial compositions of patients with IBD and HCs. Firstly, alpha diversity

analysis showed that there was no significant difference in shannon index, but species evenness in IBD patients is significantly lower relative to HCs (Fig. 1A and B). To explore whether the overall gut bacterial phenotypes of two groups were dissimilar. Beta diversity was evaluated using Principal coordinates analysis (PCoA) and Non-Metric Multidimensional Scaling (NMDS) according to the Bray–Curtis distance, which showed that bacterial signatures in the IBD and HCs were extremely significantly distinct (NMDS: ANOSIM, $P = 0.006 < 0.01$, Fig. 1C; PCoA: PERMANOVA, $P = 0.002 < 0.01$, Fig. 1D). Next, a total of 36 discriminative bacterial species between the MDD and HC groups using Linear Discriminant Analysis Effect Size (LEfSe) were identified. Compared

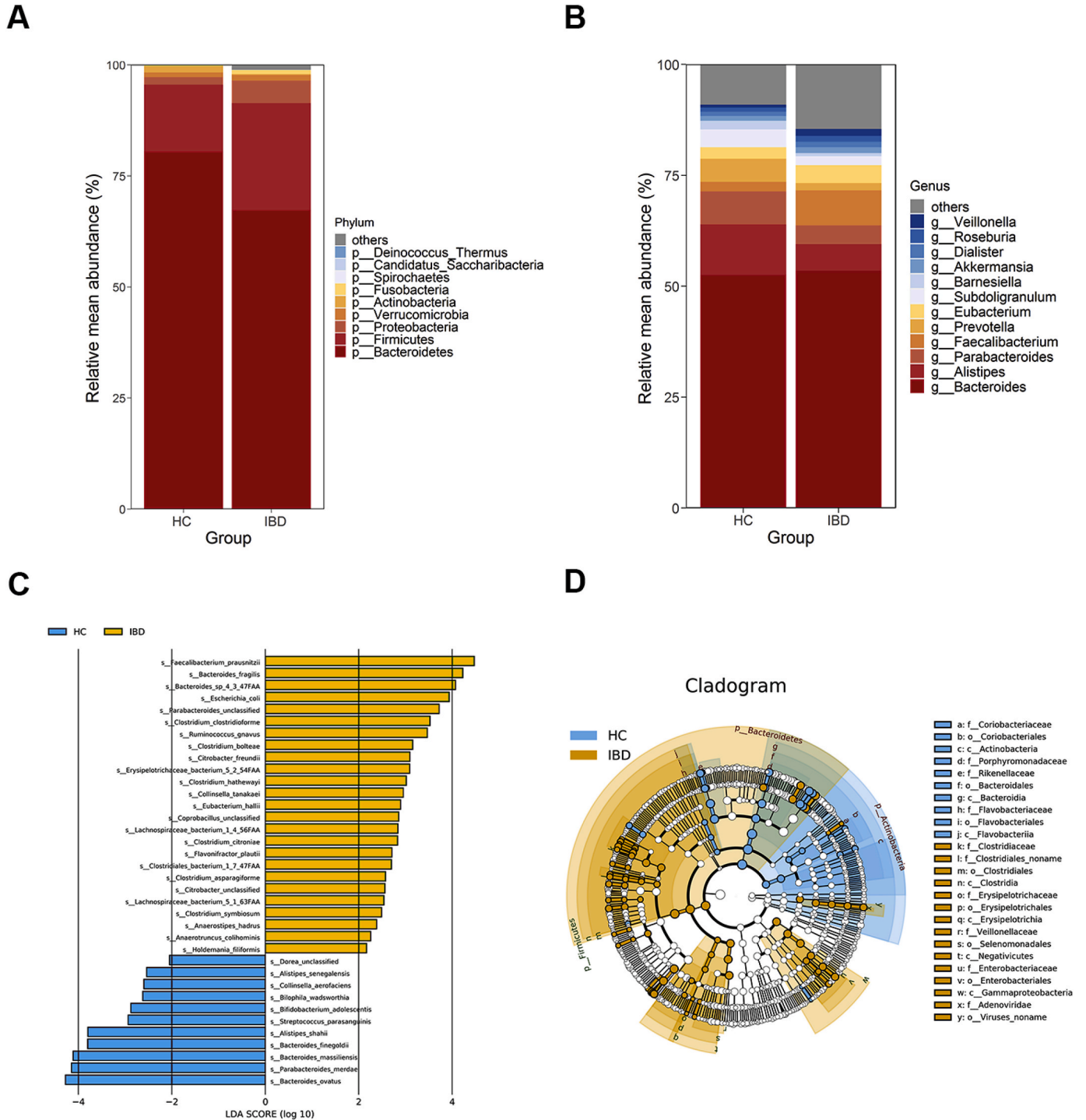


Fig. 2. Microbial composition of intestinal microbiota between IBD and HCs. Relative mean abundance of the dominant phyla (A) and genera (>1%) (B) in these two groups. (C) Linear Discriminant Analysis Effect Size (LEfSe) analysis for differentially abundant species. (D) Taxonomic cladogram of Least Discriminant Analysis (LDA) coupled with effective size measurement showing differentially microbiota hierarchy.

with HCs, IBD subjects were characterized by 25 enriched species mainly belonging to the family *Clostridiaceae* (6 species: *Clostridium asparagiforme*, *Clostridium bolteae*, *Clostridium citroniae*, *Clostridium clostridioforme*, *Clostridium hathewayi*, and *Clostridium symbiosum*), *Lachnospiraceae* (4 species: *Anaerostipes hadrus*, *Lachnospiraceae bacterium_1_4_56FAA*, *Lachnospiraceae bacterium_5_1_63FAA*, and *Ruminococcus gnavus*), *Erysipelotrichaceae* (3 species: *Coprobacillus unclassified*, *Erysipelotrichaceae bacterium_5_2_54FAA*, and *Holdemania filiformis*), and *Enterobacteriaceae* (3 species: *Citrobacter freundii*, *Citrobacter unclassified*, *Escherichia coli*), and by 6 depleted species mainly belonging to the family *Bacteroidaceae* (3 species: *Bacteroides finegoldii*, *Bacteroides massiliensis*, and *Bacteroides ovatus*) and *Rikenellaceae* (2 species: *Alistipes senegalensis* and *Alistipes shahii*) (Fig. 2C). The majority of the higher and lower species in IBD be-longed to the phylum

Firmicutes (72%, 18 of 25) and *Bacteroidetes* (55%, 6 of 11) relative to HCs, respectively, as well as, taxonomic cladogram showed the dominant phyla identified in IBD and HCs were *Firmicutes* and *Bacteroidetes*, respectively (Fig. 2D). Meanwhile, compared to HCs in the overall gut microbial communities, the abundance of the phylum *Firmicutes*, *Proteobacteria*, and *Verrucomicrobia* and the genus *Faecalibacterium*, *Eubacterium* markedly increase, whereas the phylum *Bacteroidetes* and the genus *Alistipes*, *Parabacteroides*, *Prevotella*, *Subdoligranulum*, *Barnesiella* and *Akkermansia* exhibited a significant decrease (Fig. 2A and B).

3.2. Alterations of gut microbial functional signatures in IBD

Then, a total of 17 discriminative pathways based on MetaCyc databases between the MDD and HC groups using Humann2 pipeline and

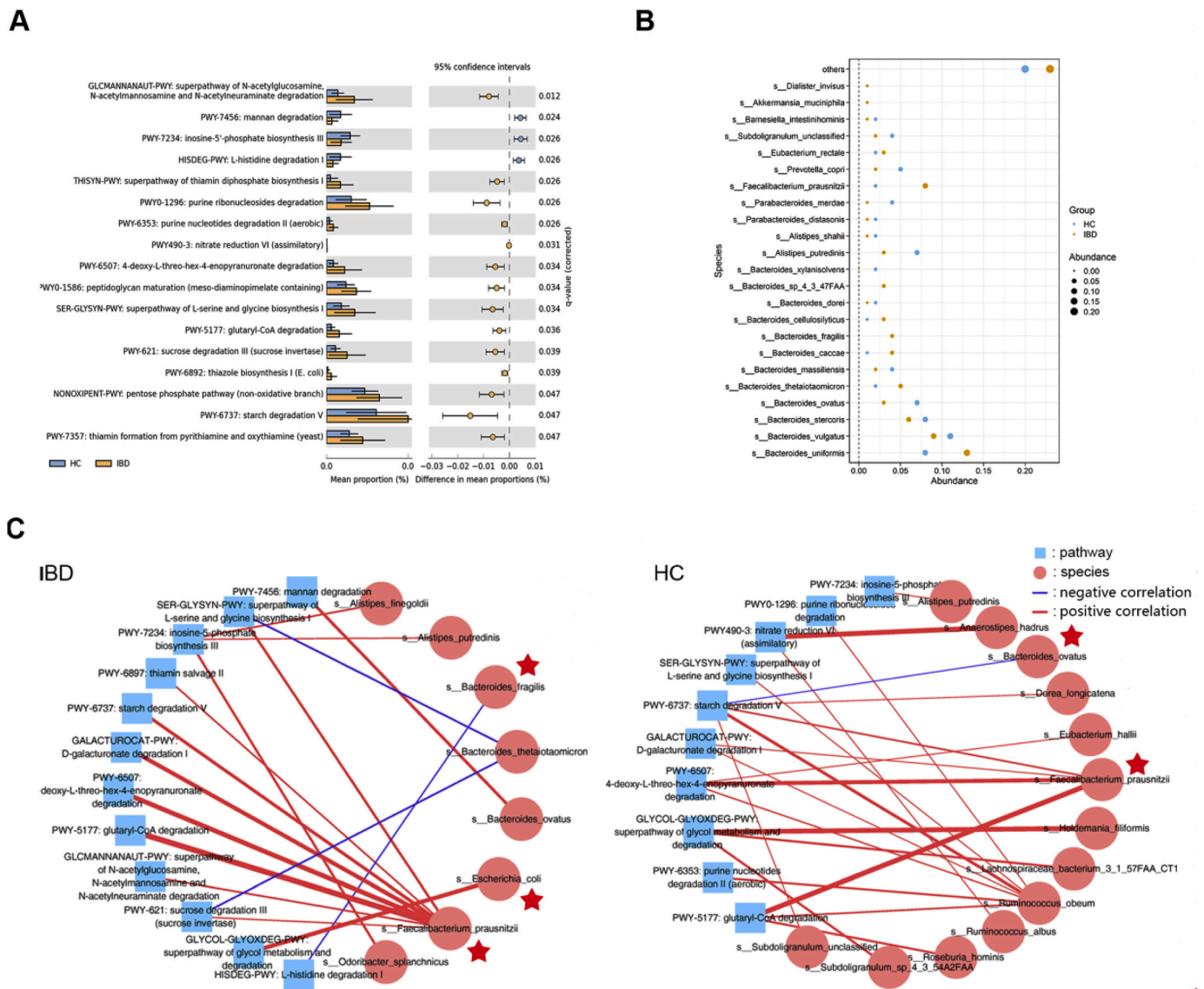
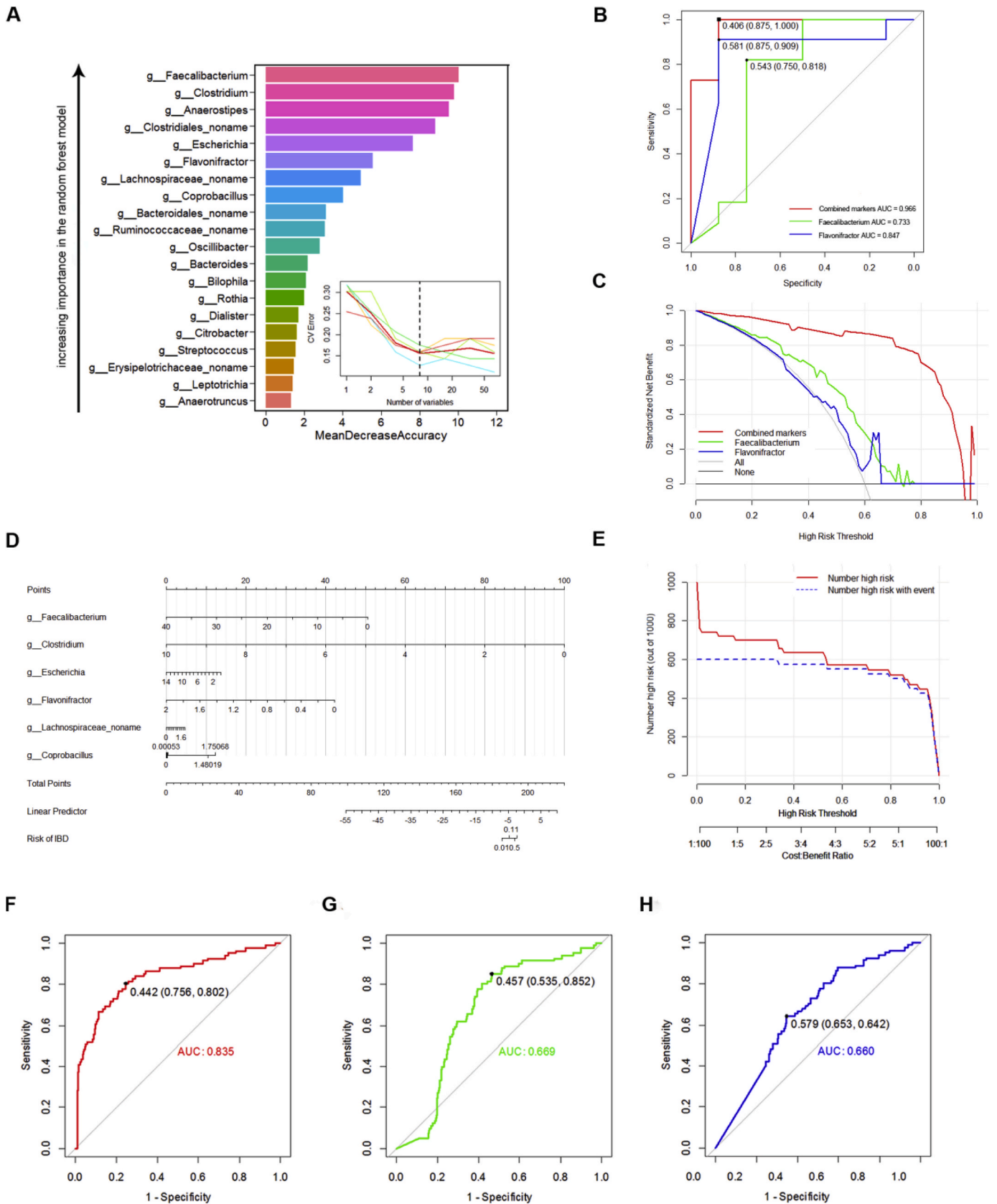


Fig. 3. MetaCyc pathway based on humann2 pipeline that discriminate IBD from HCs, the highly abundance bacterial species (>1%), and co-occurrence network constructed from the relative abundances of differential microbial taxa and distinct pathway in IBD subjects versus HCs. (A) STAMP analysis identified that relative abundances of 17 MetaCyc pathway differentiating between the two groups. Compared with HC, the IBD group was characterized by 14 up-regulated metabolites and 3 down-regulated pathway. Those pathways were mainly involved in amino acid biosynthesis, carbohydrate degradation, cofactor, carrier, and vitamin Biosynthesis, nucleoside and nucleotide biosynthesis. (Corrected $P < 0.05$). (B) Bubble chart of 23 dominant bacterial species with minimum 1% mean relative abundance. These species mainly belonging to the genus *Bacteroides* (12 species), *Alistipes* (2 species), and *Parabacteroides* (2 species). (C) Co-occurrence coefficients among microbiome components at the species level, MetaCyc pathways and the cohort were calculated by spearman correlation analysis, and networks ($P < 0.05$; Spearman's correlation > 0.6 or < -0.6 , the line thickness indicates the level of the correlation coefficient) are depicted using Cytoscape. Overall, the species *Faecalibacterium prausnitzii*, *Ruminococcus obeum* formed broad and strong co-occurring relationships with pathways in two groups. Blue square and red circle indicated pathway and species-level, respectively. Red star represent the dominant species displayed in the (B).



(caption on next page)

Fig. 4. A universal model for IBD based on Host-Microbiome And Disease Risk Classification Method (HADRCM). (A) The top 20 bacterial genera ranked in descending order of importance to the accuracy in the diagnosis model were identified by applying Random forest classifier. The insert showed that the number of genera was selected as disease markers using five times repeated, 10-fold cross validation. (B) Multivariate logistic regression was conducted to accurately classify IBD and HCs using 6 filtered variables in the final models. The models for single marker and combined markers were evaluated by the Receiver-Operator Characteristic (ROC) curve and the area under the ROC curve. This combinatorial marker including 6 markers yielded more robust diagnostic performance over that of separate microbial markers (AUC = 0.966). (C) The Decision Curve Analysis (DCA) showed the standardized net benefit against different decision thresholds for the diagnosis models. (D) Nomogram for predicting IBD risk based on combinatorial biomarkers. (E) Clinical Impact Curve exhibited 1,000 people Risk stratification based on SIMPLE model. The x-axis showed the risk threshold and ratio of cost and benefit. The number of high risk is almost consistent with true positive number when risk threshold close to 0.6. Receiver-operating characteristic (ROC) curves for performance of models in an external independent cohort. ROC curves using the combination of 6 biomarkers (F) and 2 individual markers (G, H) were plotted for the datasets, and the areas under the ROC curves (AUCs) were calculated.

STAMP software were identified (Fig. 3A). Compared with HCs, IBD patients were enriched in 14 pathways and depleted in 3 pathways (Corrected $P < 0.05$). The down-regulated pathways, including polysaccharide degradation, proteinogenic amino acid degradation, and purine nucleotide biosynthesis, would probably play a vital role in IBD. Those differential pathways were mainly involved in four biological processes: amino acid biosynthesis, carbohydrate degradation, cofactor, carrier, and vitamin Biosynthesis, nucleoside and nucleotide biosynthesis.

3.3. Co-occurrence network analysis of functional pathways and gut bacteria components

Afterwards, the study explored the potential correlations of abundances of these differential gut bacterial species and functional pathways. Overall, co-occurrence analysis showed that bacterial species widely formed strong co-occurring relationships with pathways (Fig. 3C). The species dominant *Faecalibacterium prausnitzii* in the IBD, and *Ruminococcus obeum* and *Faecalibacterium prausnitzii* in the HCs formed broad and strong positive relationships with many pathways in the two groups, meanwhile *Bacteroides fragilis* and *Bacteroides thetaiotaomicron*, and *Bacteroides ovatus* displayed a negative correlations in IBD and HCs, respectively, whereas compare to HCs, *Bacteroides ovatus* of the IBD group showed a strong positive correlations with another pathway. In addition, the dominant species *Faecalibacterium prausnitzii*, *Bacteroides fragilis*, *Escherichia coli*, *Bacteroides ovatus* displayed in Fig. 3B were found reside in co-occurring network, in contrast, the lowly abundance species, such as *Ruminococcus obeum* (mean abundance 0.04% in HCs), remain strongly significant and comparable across all groups. These findings suggest that different intestinal microbiota and signaling pathways or molecules may form synergistic and niche-related co-occurring network in patients with IBD.

3.4. Combinatorial biomarkers model with high accuracy for discriminating IBD from HC

Random forest classifier was used to identify potential biomarker discriminate subjects with IBD from HCs. A five times repeated, 10-fold cross validation was used to identify the more representative bacterial genus, which can showed the most significant deviations in IBD and HCs. The top 8 genera selected as disease markers were used to construct the model via logistic regression (Fig. 4A). The datasets were divided into training and validation sets using a 70–30% split, the samples from the validation set were used to confirm the diagnostic performance independently. The model of combined 6 markers (the genus *Faecalibacterium*, *Clostridium*, *Anaerostipes*, *Clostridiales noname*, *Escherichia*, *Flavonifractor*, *Lachnospiraceae noname*, *Coprobacillus*) yielded more robust diagnostic performance (AUC = 0.966) over that of separate microbial markers (*Faecalibacterium* and *Flavonifractor*) in the final models with a c-statistic of 0.965 (Fig. 4B). The decision curve was used to evaluate the benefit of the combined markers and single one method, which displayed combined markers model provided superior net benefit and reduction than that of treating everyone or treating no one, with a probability threshold of 0.6 or greater (Fig. 4C). A nomogram for assessment of IBD risk based on combinatorial biomarkers was finally

shown in Fig. 4D, which would be more feasible and economical in clinical practice and could be convenient to predict the risk of IBD by assigning points for variable by drawing a line upward from the corresponding variable to the line of Points and summing the points associated with the line of Total Points. The performance of model was further conformed by Clinical Impact Curve (CIC) with 1,000 people Risk stratification based on self SIMPLE model (Fig. 4E). Here, the above mentioned methods was defined as Host-Microbiome And Disease Risk Classification Method (HADRCM).

To further validate our models above, an external cohort was used to confirm the diagnostic performance independently. Consequently, compared to 2 individual marker, the marker panel of these six biomarkers could still effectively discriminate between the two groups with AUC of 0.835, however, a relatively poor diagnostic performance with AUC of 0.669 and 0.660 were achieved using the 2 individual marker, respectively (Fig. 4F–H). These above mentioned showed using HADRCM could yield a robust model with high accuracy for cohort study and would be more feasible, economical, and convenient in clinical practice.

4. Discussion

This study outlined landscapes and interaction networks of differential microbe and pathways in the IBD and HCs group, Moreover, it provided a novel method named as HADRCM based on machine-learning and statistical analysis. Meanwhile, using HADRCM, a combinatorial marker panel that could distinguish IBD from HC subjects with high accuracy was identified and independently validated, and nomogram of the model could be conveniently applied in clinical practice. Furthermore, because the two subtypes (CD and UC) of IBD have different microbial communities, this approach based on theoretical basis can be also applicable to distinguish IBD subtypes. Our findings suggest that gut microbiota disturbances could potentially contribute to IBD pathogenesis by modulating the host's amino acid, nucleotide, cofactor, vitamin and carbohydrate metabolism, which provides a new avenue by which to understand the basis of IBD.

Compared with the 16s rRNA amplicon sequencing, metagenomic sequencing can provide species level and more valuable information, especially potential functional mechanisms in the microbiota. Our findings were consistent with that of previous research, in which a decrease of alpha diversity in IBD patients was observed [20], however, the species level was not reported. Here, 25 differential bacterial species responsible for this discrimination were identified, which were primarily from the phyla *Firmicutes* (72.2%), however depleted species mainly belonging to the phyla *Bacteroidetes* (55%). The data presented in this study confirm previously reported changes of phyla *Firmicutes* and *Bacteroidetes* in IBD patients versus HCs [21,22]. *Bacteroidetes* play significant roles in gut microbiota–host interactions, especially on metabolic pathways [23]. Consistent with previous reports, many *Bacteroidetes* species were significantly correlated amino acid, nucleotide, cofactor, vitamin and carbohydrate metabolism. Therefore, the down-regulation *Bacteroidetes* species may account for lower amino acid levels in IBD.

In the co-expression network analysis, *Faecalibacterium prausnitzii* was found to form strong co-occurring relationships with lots of pathways assigned to carbohydrate and amino acid metabolism and was the

dominant species in both groups. Furthermore, previous gut microbiome studies suggested *Faecalibacterium* have special effects in IBD [24–26].

Currently, machine learning is increasingly used in the disease diagnostics field [27,28]. However, many studies only apply a single algorithm or simple statistics, and only obtain low accuracy. And it is rare and difficult to be applied in actual clinical medical diagnosis. In this study, random forest, logistic regression and statistical analysis were combined to construct diagnostics model and confirmed the excellent performance using ROC and DCA. The union of the methods was defined as HADRCM. Using HADRCM, some microbial markers could effectively discriminate the IBD individuals from HCs. Moreover, a combinatorial marker panel could distinguish IBD individuals from HCs with AUC of 0.966, which still generate a vigorous performance with AUC of 0.835 in an independent external cohort. Ultimately, this output nomogram can provide the non-invasive diagnosis proof for clinicians. Hence, the HADRCM would be a promising alternative in the cohort study.

There are several limitations of this study: (i) The specific microbial species mechanisms need to be further clarified in animal studies. (ii) All data were collected from HMP, which has a relatively wide geographical representation and there is no information similar to BMI of the samples. (iii) Strain-level have no been explored and compared in the cohort.

Despite these limitations, the results from this cohort emphasize the importance of the gut microbiota as a risk factor accounting for the pathogenesis of IBD and the diagnostic implication and accuracy of combination of biomarkers as a non-invasive clinical evaluation indicator for IBD, meanwhile, the union of various machine learning and statistical analysis would be a very promising approach for translational medicine, precision medicine and next generation medicine.

Data availability

All data could be obtained from HMP and iHMP databases and the accession numbers are made available in the [Supplementary Table 1](#).

Disclosure of potential conflicts of interest

No potential conflicts of interest were disclosed.

CRediT authorship contribution statement

Guangcai Liang: Conceptualization, Methodology, Software, Data curation, Writing – original draft, preparation, Visualization, Validation, Writing – review & editing, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.synbio.2021.10.003>.

References

- Xavier RJ, Podolsky DK. Unravelling the pathogenesis of inflammatory bowel disease. *Nature* 2007;448:427–34. <https://doi.org/10.1038/nature06005>.
- Ordas I, Eckmann L, Talamini M, Baumgart DC, Sandborn W. J. Ulcerative colitis. *Lancet* 2012;380:1606–19. [https://doi.org/10.1016/S0140-6736\(12\)60150-0](https://doi.org/10.1016/S0140-6736(12)60150-0).
- Blumberg R, Powrie F. Microbiota, disease, and back to health: a metastable journey. *Sci Transl Med* 2012;4:137rv7. <https://doi.org/10.1126/scitranslmed.3004184>.
- Strober W, Fuss IJ, Blumberg RS. The immunology of mucosal models of inflammation. *Annu Rev Immunol* 2002;20:495–549. <https://doi.org/10.1146/annurev.immunol.20.100301.064816>.
- Hegazy AN, West NR, Stubbington MJT, Wendt E, Suijker KIM, Datsi A, et al. Circulating and tissue-resident CD4(+) T cells with reactivity to intestinal microbiota are abundant in healthy individuals and function is altered during inflammation. *Gastroenterology* 2017;153:1320–37. <https://doi.org/10.1053/j.gastro.2017.07.047>. e16.
- 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;569:655–62. <https://doi.org/10.1038/s41586-019-1237-9>.
- Gecse KB, Bemelman W, Kamm MA, Stoker J, Khanna R, Ng SC, et al. A global consensus on the classification, diagnosis and multidisciplinary treatment of perianal fistulising Crohn's disease. *Gut* 2014;63:1381–92. <https://doi.org/10.1136/gutjnl-2013-306709>.
- Cummings J. H. & Kong, S. C. Probiotics, prebiotics and antibiotics in inflammatory bowel disease. *Novartis Found Symp* 263, 99–111; discussion -4, 211–8 (2004).
- Zachos M, Tondeur M, Griffiths AM. Enteral nutritional therapy for induction of remission in Crohn's disease. *Cochrane Database Syst Rev* 2007. <https://doi.org/10.1002/14651858.CD000542.pub2>. CD000542.
- Huttenhower C, Gevers D, Knight R, Abubucker S, White O. Structure, function and diversity of the healthy human microbiome. *Nature* 2012;486:207–14. <https://doi.org/10.1038/nature11234>.
- Human Microbiome Project C. A framework for human microbiome research. *Nature* 2012;486:215–21. <https://doi.org/10.1038/nature11209>.
- Proctor LM, Creasy HH, Fettweis JM, Lloyd-Price J, Mahurkar A, Zhou W, et al. The integrative human microbiome Project. *Nature* 2019;569:641–8. <https://doi.org/10.1038/s41586-019-1238-8>.
- Truong DT, Franzosa EA, Tickle TL, Scholz M, Weingart G, Pasolli E, et al. MetaPhlan2: enhanced metagenomic taxonomic profiling. *Nat Methods* 2015;12:902–3. <https://doi.org/10.1038/nmeth.3589>.
- Franzosa EA, McIver LJ, Rahnava G, Thompson LR, Schirmer M, Weingart G, et al. Species-level functional profiling of metagenomes and metatranscriptomes. *Nat Methods* 2018;15:962–8. <https://doi.org/10.1038/s41592-018-0176-y>.
- Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, et al. Metagenomic biomarker discovery and explanation. *Genome Biol* 2011;12:R60. <https://doi.org/10.1186/gb-2011-12-6-r60>.
- Parks DH, Tyson GW, Hugenholtz P, Beiko RG. STAMP: statistical analysis of taxonomic and functional profiles. *Bioinformatics* 2014;30:3123–4. <https://doi.org/10.1093/bioinformatics/btu494>.
- Bento AP, Gaulton A, Hersey A, Bellis LJ, Chambers J, Davies M, et al. Classification and regression by randomForest. *R News* 2002;23:18–22.
- Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making* 2006;26:565–74. <https://doi.org/10.1177/0272989X06295361>.
- Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 2003;13:2498–504. <https://doi.org/10.1101/gr.1239303>.
- Dovrolis N, Michalopoulos G, Theodoropoulos GE, Arvanitidis K, Gazouli M. The interplay between mucosal microbiota composition and host gene-expression is linked with infliximab response in inflammatory bowel diseases. *Microorganisms* 2020;8:438. <https://doi.org/10.3390/microorganisms8030438>.
- Singh P, Teal TK, Marsh TL, Tiedje JM, Mosci R, Jernigan K, et al. Intestinal microbial communities associated with acute enteric infections and disease recovery. *Microbiome* 2015;3:45. <https://doi.org/10.1186/s40168-015-0109-2>.
- Santorù ML, Piras C, Murgia A, Palmas V, Camboni T, Liggi S, et al. Cross sectional evaluation of the gut-microbiome metabolome axis in an Italian cohort of IBD patients. *Sci Rep* 2017;7:9523. <https://doi.org/10.1038/s41598-017-10034-5>.
- Maier E, Anderson RC, Roy NC. Understanding how commensal obligate anaerobic bacteria regulate immune functions in the large intestine. *Nutrients* 2014;7:45–73. <https://doi.org/10.3390/nu7010045>.
- Heinken A, Ravcheev DA, Baldini F, Heirendt L, Thiele I. Systematic assessment of secondary bile acid metabolism in gut microbes reveals distinct metabolic capabilities in inflammatory bowel disease. *Microbiome* 2019;7. <https://doi.org/10.1186/s40168-019-0689-3>.
- Sokol H, Seksik P, Furet JP, Firmesse O, Nion-Larmurier I, Beaugerie L, et al. Low counts of *Faecalibacterium prausnitzii* in colitis microbiota. *Inflamm Bowel Dis* 2009;15:1183–9. <https://doi.org/10.1002/ibd.20903>.
- Sokol H, Pigneur B, Watterlot L, Lakhdari O, Bermudez-Humaran LG, Grataudoux JJ, et al. *Faecalibacterium prausnitzii* is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. *Proc Natl Acad Sci U S A* 2008;105:16731–6. <https://doi.org/10.1073/pnas.0804812105>.
- Oh TG, Kim SM, Caussy C, Fu T, Guo J, Bassirian S, et al. A universal gut-microbiome-derived signature predicts cirrhosis. *Cell Metab* 2020;32:878–888 e6. <https://doi.org/10.1016/j.cmet.2020.06.005>.
- Ren Z, Li A, Jiang J, Zhou L, Yu Z, Lu H, et al. Gut microbiome analysis as a tool towards targeted non-invasive biomarkers for early hepatocellular carcinoma. *Gut* 2019;68:1014–23. <https://doi.org/10.1136/gutjnl-2017-315084>.