



# Leveraging 16S rRNA Microbiome Sequencing Data to Identify Bacterial Signatures for Irritable Bowel Syndrome

## OPEN ACCESS

### Edited by:

Yolanda López-Vidal,  
 Universidad Nacional Autónoma  
 de México, Mexico

### Reviewed by:

Jane Adair Mullaney,  
 AgResearch Ltd, New Zealand  
 Junguk Hur,  
 University of North Dakota,  
 United States

### \*Correspondence:

Lijuan Chai  
 cjuan1258@163.com  
 Xiaoxuan Tian  
 tian\_xiaoxuan@tjutcm.edu.cn

†These authors have contributed  
 equally to this work

### Specialty section:

This article was submitted to  
 Microbiome in Health and Disease,  
 a section of the journal  
 Frontiers in Cellular and  
 Infection Microbiology

**Received:** 27 January 2021

**Accepted:** 29 April 2021

**Published:** 11 June 2021

### Citation:

Liu Y, Li W, Yang H, Zhang X, Wang W,  
 Jia S, Xiang B, Wang Y, Miao L,  
 Zhang H, Wang L, Wang Y,  
 Song J, Sun Y, Chai L and Tian X  
 (2021) Leveraging 16S rRNA  
 Microbiome Sequencing Data to  
 Identify Bacterial Signatures  
 for Irritable Bowel Syndrome.  
 Front. Cell. Infect. Microbiol. 11:645951.  
 doi: 10.3389/fcimb.2021.645951

Yuxia Liu<sup>1†</sup>, Wenhui Li<sup>1†</sup>, Hongxia Yang<sup>1</sup>, Xiaoying Zhang<sup>1</sup>, Wenxiu Wang<sup>1</sup>, Sitong Jia<sup>1</sup>,  
 Beibei Xiang<sup>2</sup>, Yi Wang<sup>1,3</sup>, Lin Miao<sup>1,3</sup>, Han Zhang<sup>1,4</sup>, Lin Wang<sup>5</sup>, Yujing Wang<sup>5</sup>,  
 Jixiang Song<sup>5</sup>, Yingjie Sun<sup>5</sup>, Lijuan Chai<sup>1,4\*</sup> and Xiaoxuan Tian<sup>1\*</sup>

<sup>1</sup> State Key Laboratory of Component-Based Chinese Medicine, Tianjin University of Traditional Chinese Medicine, Tianjin, China, <sup>2</sup> School of Chinese Materia Medica, Tianjin University of Traditional Chinese Medicine, Tianjin, China, <sup>3</sup> Key Laboratory of Pharmacology of Traditional Chinese Medical Formulae, Ministry of Education, Tianjin University of Traditional Chinese Medicine, Tianjin, China, <sup>4</sup> Laboratory of Pharmacology of Traditional Chinese Medical Formulae Co-Constructed by the Province-Ministry, Tianjin University of TCM, Tianjin, China, <sup>5</sup> Tianjin Zhongxin Pharmaceutical Group Co., Ltd. Le Ren Tang Pharmaceutical Factory, Tianjin, China

Irritable bowel syndrome (IBS) is a chronic gastrointestinal disorder characterized by abdominal pain or discomfort. Previous studies have illustrated that the gut microbiota might play a critical role in IBS, but the conclusions of these studies, based on various methods, were almost impossible to compare, and reproducible microorganism signatures were still in question. To cope with this problem, previously published 16S rRNA gene sequencing data from 439 fecal samples, including 253 IBS samples and 186 control samples, were collected and processed with a uniform bioinformatic pipeline. Although we found no significant differences in community structures between IBS and healthy controls at the amplicon sequence variants (ASV) level, machine learning (ML) approaches enabled us to discriminate IBS from healthy controls at genus level. Linear discriminant analysis effect size (LEfSe) analysis was subsequently used to seek out 97 biomarkers across all studies. Then, we quantified the standardized mean difference (SMDs) for all significant genera identified by LEfSe and ML approaches. Pooled results showed that the SMDs of nine genera had statistical significance, in which the abundance of *Lachnospirillum*, *Dorea*, *Erysipelatoclostridium*, *Prevotella 9*, and *Clostridium sensu stricto 1* in IBS were higher, while the dominant abundance genera of healthy controls were Ruminococcaceae UCG-005, *Holdemanella*, *Coprococcus 2*, and *Eubacterium coprostanoligenes* group. In summary, based on six published studies, this study identified nine new microbiome biomarkers of IBS, which might be a basis for understanding the key gut microbes associated with IBS, and could be used as potential targets for microbiome-based diagnostics and therapeutics.

**Keywords:** gut microbiome, biomarkers, 16S rRNA, machine learning algorithm, irritable bowel syndrome

## INTRODUCTION

Irritable bowel syndrome (IBS) is a common gastrointestinal disorder characterized by chronic, recurrent episodes of abdominal discomfort and pain with altered bowel habits. It has affected approximately 12% of the global population and over 16% of the United States population (Schmulson et al., 2006; Poulsen et al., 2017; Ma et al., 2020). According to defecation pattern, IBS patients can be divided into four main subtypes, including IBS with constipation (IBS-C), IBS with diarrhea (IBS-D), IBS with mixed bowel habits (IBS-M), and unclassified IBS. Currently, the Rome criteria (Rome IV) has defined IBS as recurrent abdominal pain with at least one day of abdominal pain per week for the past three months and two or more of the following symptoms: pain related to defecation, change in fecal frequency, change in fecal shape, and any of these not less than six months before diagnosis (Hellstrom and Benno, 2019; Asghar et al., 2020; Palsson et al., 2020). IBS has become a significant disease burden in terms of increased absenteeism from school or work and reduced health-related quality of life, although it is not a fatal disease (Zhen Lu et al., 2006; Ma et al., 2020; Yao et al., 2020). IBS is a complex and heterogeneous disease with many factors involved in its etiology and pathogenesis (i.e., mucosal immune hyperactivity, food intolerances, distortions in the gut microbiome) (Quigley, 2005; Mullin et al., 2014; Moloney et al., 2016; Gonzalez-Castro et al., 2017).

An increasing number of studies have reported a critical role for the gut microbiome in health and IBS, respectively (Kassinen et al., 2007; Krogius-Kurikka et al., 2009; Carroll et al., 2011; Rajilic-Stojanovic et al., 2011; Jeffery et al., 2012), but trends in the microbial biomarkers associated with IBS in these reports are inconsistent. At the phylum level, some studies showed higher abundance of Proteobacteria in IBS (Carroll et al., 2012; Pozuelo et al., 2015), whereas other research showed no difference of Proteobacteria relative to healthy ones (Chung et al., 2016; Dior et al., 2016). At the genus level, *Faecalibacterium*, *Blautia* (Pozuelo et al., 2015; Tap et al., 2017), *Veillonella* (Tana et al., 2010; Rigsbee et al., 2012), and *Ruminococcus* (Pozuelo et al., 2015) were identified to be positively associated with IBS and *Methanogens* (Rajilic-Stojanovic et al., 2011) was depletion in IBS. A study indicated that *Lachnospirillum* was significantly associated with the clinical symptoms of IBS (Zhu et al., 2019), while there was little concern in other studies. Thus, it remains uncertain whether there are highly reproducible microbial signatures that can differentiate IBS subjects from healthy controls across cohorts and study designs.

Meta-analysis has been used as a tool to establish and validate associations between the intestinal microbiome and diseases across populations or cohorts (Walters et al., 2014; Sze and Schloss, 2016; Duvall et al., 2017; Sze and Schloss, 2018; Armour et al., 2019). Several meta-analyses have been performed based on studies linking the gut microbiome to IBS using bacterial culture and qPCR, or 16S rRNA gene technology (Liu et al., 2017; Zhuang et al., 2017; Wang et al., 2020). Nevertheless, these existing meta-analysis studies still have limitations. The bacteria cultivation and qPCR technologies (Kassinen et al., 2007; Krogius-Kurikka et al., 2009; Rajilic-Stojanovic et al., 2011) focused on only a limited number of microorganism species (Gerritzen et al., 2011). Analysis based on

this evidence can only evaluate whether there are significant changes within “common” bacteria when IBS occurs, rather than looking for key bacterial characteristics without preconceived assumptions. For example, these articles highlighted only lower *Lactobacillus* and *Bifidobacterium* and higher *Escherichia coli* levels in IBS compared to the healthy group. This indicates that research based on non-targeted microbial identification methods, such as 16S rRNA gene technology, deserve more attention in meta-analysis. Besides, previous meta-analyses were almost always based on the final results of independent studies, rather than the raw sequencing data, which could not alleviate the incompatibility among multi-cohort efforts due to the variation in bioinformatics pipelines (Pollock et al., 2018). Meta-analyses based on raw sequencing data have successfully characterized the microbiome signatures in colorectal cancer, tumors (Shah et al., 2018; Sze and Schloss, 2018; Thomas et al., 2019), and obesity (Finucane et al., 2014; Walters et al., 2014; Sze and Schloss, 2016). Therefore, it is thus necessary to conduct such analyses to identify significant differences in the gut microbiome between IBS and healthy controls, which have not been reported to date.

Here, we collected 16S rRNA gene sequence data of stool samples (n=439) from six studies (Saulnier et al., 2011; Pozuelo et al., 2015; Labus et al., 2017; Zhuang et al., 2018; Lo Presti et al., 2019; Zhu et al., 2019). A unified pipeline was used to process raw sequencing data to investigate whether biomarkers describing bacterial communities or community-specific microbiota profiles could more accurately identify IBS and healthy controls. The results of our study showed that alterations in bacterial communities are indeed associated with IBS and that a subset of the bacterial profiles may be considered as potential biomarkers for identifying the presence of IBS.

## MATERIALS AND METHODS

### Data Sets Collection

We followed previously published methods for the meta-analyses of microbiome data (Duvall et al., 2017; Sze and Schloss, 2018). Raw data and metadata for the included cohorts were downloaded from the Sequence Read Archive (SRA). We collected 16S rRNA gene sequence data of 439 stool samples from six previously published studies, whose sequencing methods were performed using Illumina sequencing or 454 sequencing. We excluded studies that were reviewed or meta-analyzed, focused on cultivation and qPCR techniques, or were used only as abstracts for conference papers. Studies without controls or with fewer than five case patients were also excluded. Any studies that failed to provide either publicly available sequences or metadata were excluded. The reuse of these published data in our meta-analysis complied with all relevant ethical regulations. Of the included studies, two were from the US, two were from China, one was from Italy, and one was from Spain. We manually curated metadata tables for the public case-controls.

### Data Processing

Raw sequence data and metadata were obtained from the Sequence Read Archive (SRA) in NCBI. To avoid bias caused

by different bioinformatic analysis pipelines, the sequence read pools for each study were filtered and analyzed using the same custom script based on the QIIME2-2020.2 (Hall and Beiko, 2018; Bolyen et al., 2019). Each dataset was imported and assembled in QIIME2-2020.2 against single-end sequences for 454 sequencing or paired-end sequences for Illumina-sequencing. Denoising was implemented using DADA2 (divisive amplicon denoising algorithm 2) which discards chimeras and erroneous sequences (Callahan et al., 2016). The individual real biological sequences, referred to as amplicon sequence variants (ASV) (Callahan et al., 2017) were retrieved as a higher resolution version of the operational taxonomic unit (OTU) table than those generated by traditional methods, as well as their frequencies. All remained sequences had a length of  $\geq 200$  bp and an average sequence quality score of  $\geq 20$ . Subsequently, the resulting representative set of sequences was aligned and classified using the SILVA database (Quast et al., 2013; Yilmaz et al., 2014). ASVs with fewer than 10 reads were removed. We normalized the relative abundance of each ASV by dividing its value by the total number of reads for each sample. We then collapsed the ASVs to the genus level by summing their respective relative abundances, discarding any ASVs which were unannotated at the genus level. All statistical analyses were performed on the relative abundance data at this genus level.

## Community Analysis

Both alpha and beta diversity analyses were performed within each dataset through the q2-diversity plugin in QIIME2. Alpha diversity metrics (i.e., evenness, observed-OTUs, and Shannon) were calculated based on ASV level. The non-parametric Kruskal-Wallis test (Sculco, 2001; Ren et al., 2018) was used to test alpha diversity metric dissimilarities, and the p-value was corrected using the Benjamini-Hochberg method (Ferreira, 2007; Abbas et al., 2013). The Bray-Curtis distance was used to measure beta diversity metrics. We explored the community structure of the samples with PERMANOVA (Kelly et al., 2015) using the beta-group-significance command. Rarefaction was performed on the feature table before calculating the distances in QIIME2.

## Statistical Analysis

Two classification algorithms were used to classify healthy and IBS individuals. We built an AdaBoost classifiers (Montazeri et al., 2016) function with 1000 estimators utilizing Python's scikit-learn module (Pedregosa et al., 2011), and random forest classification models using the RandomForest package (Lebanov et al., 2020). All models were built with 10-fold cross-validation using data at the genus level. Based on the test results of cross-validation, the interpolated area under the receiver operating characteristic (ROC) curve (AUC) was calculated. AUC represents the area under the ROC curve, which can evaluate the classification ability of machine learning models. The higher the AUC, the better the model at correctly classifying instances, and 0.5 is the decision threshold.

Univariate analysis based on the relative abundances of genera was performed using the linear discriminant analysis effect size (LEfSe) method (Segata et al., 2011). We focused on those biomarkers that had significant difference between cases

and controls per dataset, and then combined these results across all studies. Finally, we performed fixed-effects model based on all important features and estimated the effect size standardized mean differences (SMDs). The entire analytical workflow is shown in **Figure 1**.

## RESULTS

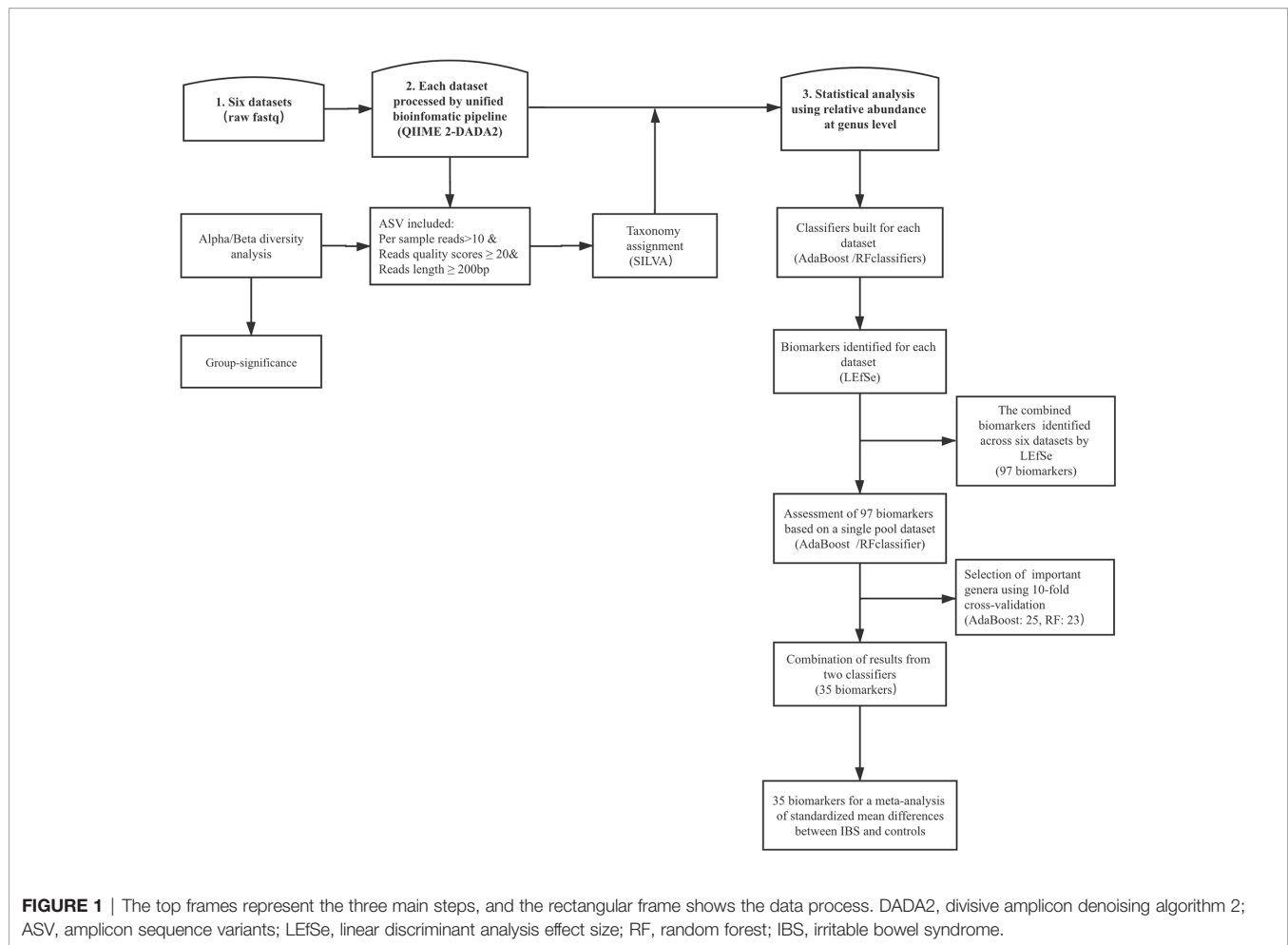
A total of 106 studies were identified in PubMed (**Figure S1**). A total of 31 studies that were reviews or animal studies were excluded. Of the 75 full-text articles reviewed for eligibility, 52 studies that did not meet certain eligibility requirements were discarded, and 23 studies had datasets with inclusion criteria. However, 17 datasets were excluded due to unavailable raw data or incomplete metadata. We finally collected and re-processed the original stool 16S rRNA sequence data published in six studies, including 253 IBS samples and 186 control samples (Saulnier et al., 2011; Pozuelo et al., 2015; Labus et al., 2017; Zhuang et al., 2018; Lo Presti et al., 2019; Zhu et al., 2019). A summary of the included studies, data characteristics, and metadata is presented in **Table 1**.

### No Consistent Significant Shifts in Community Diversity Analysis

We firstly assessed whether there was a divergence in microbial community composition and distribution at the ASV level between healthy controls and IBS patients. For the alpha diversity metrics, evenness, observed-OTUs, and Shannon were calculated. The result of the K-W test showed that only two of the six studies had statistically significant differences between the two groups (Pozuelo et al., 2015; Zhu et al., 2019), according to the Shannon and observed-OTUs indices (**Table S1**). To test for differences in microbiota profiles between IBS disease stages and healthy controls, we performed PERMANOVA analysis within each dataset based on the Bray-Curtis distance (**Table S2**). The results demonstrated that the statistically different studies were consistent with those found in alpha diversity analysis (Pozuelo et al., 2015; Zhu et al., 2019), and a study had a p-value marginally greater than 0.05 (Saulnier et al., 2011). These results implied that from a microbial perspective, the correlates of IBS might be some of the key taxa found, rather than the the entire community structure depicted at the level of ASVs or OTUs.

### Machine Learning Algorithm Showed Microbiological Changes at the Genus Level

Next, we wondered whether the relationship between intestinal microbiota and IBS would be recapitulated when rearranging microbiological data at a higher taxonomic level. To utilize the biodiversity data more effectively, two sensitive ML algorithms, AdaBoost and random forest classifiers, were employed at the genus level, to estimate whether the gut microbiota was altered in IBS, and to distinguish cases from controls, respectively. As shown in **Figure 2A**, both classifiers, especially the AdaBoost, provided AUC values greater than 0.8 in at least four datasets.



These results indicated that although the ASV results shown above were not expected, IBS patients carry an altered gut microbiota that can be correctly classified by genus-level gut microbiome data using the ML method.

## Identified Significant Microbiome Biomarkers

The linear discriminant analysis effect size (LEfSe) tool was utilized to search for bacterial biomarkers within each dataset. After pooling respective biomarkers from six datasets, a total of 97 biomarkers (LDA SCORE > 2) were identified (**Figure 2B**). However, the great majority of genera were dataset-specific, with only *Parabacteroides* genera being shared in four of the six studies and four genera being identified in 3/6 studies, i.e., *Lachnospiraceae\_NK4A136\_group*, *Butyrivimonas*, *Dorea*, and *Lachnoclostridium*.

Then, the AdaBoost and random forest classifiers were built based on the 97 biomarkers to further assess the combined biomarkers selected by LEfSe analysis. We observed an impressive classification with an AUC value of 0.86 (random forest) and 0.77 (AdaBoost), separately (**Figures S2, S3**). To obtain the essential taxa, after performing feature selection and pruning random forest trees, the AdaBoost and random forest classifiers

retained the 23 and 25 most important features, respectively (**Figure 3**). The most crucial genus was *Parabacteroides*, and other features included genera such as *Dorea*, Ruminococcaceae *UCG-005*, and *Prevotella 9*. Thus, to avoid the loss of essential genera information, we pooled the results from the two classifiers, and got 35 important features.

A fixed-effect model was used to aggregate these pieces of evidence in order to accurately assess the degree of consistency of bacterial biomarkers of IBS disease across datasets. Of the 35 important genera mentioned above, we found nine genera had statistically significant difference between the two groups across studies (**Figure 4**). Three of these significant genera showed a significant decrease in the impurity of the random forest tree, such as *Erysipelatoclostridium* (SMD=-0.18, 95% CI: -0.32 to -0.03; P= 0.0150), *Clostridium sensu stricto 1* (SMD=-0.19, 95% CI: -0.33 to -0.05; P= 0.0091), and *Coprococcus 2* (SMD=0.18, 95% CI 0.04 to 0.32; P= 0.014). *Holdemanella* (SMD=0.22, 95% CI 0.08 to 0.36; P= 0.0025) contributed only to the AdaBoost classifier. And five genera were ranked as highly important in both AdaBoost and random forest classifiers: the genera *Dorea* (SMD=-0.21, 95% CI: -0.35 to -0.06; P= 0.0050), *Prevotella 9* (SMD=-0.15, 95% CI: -0.30 to -0.01; P= 0.0399), *Lachnoclostridium* (SMD=-0.25, 95% CI: -0.40 to -0.11; P= 0.0006), Ruminococcaceae *UCG-005* (SMD=0.18, 95% CI

**TABLE 1** | Size and characteristics of the IBS 16S rRNA datasets included in this study.

Dataset	Groups (n)	Age (average ± sd.)	BMI (average ± sd.)	Sex F/M	Country	Data storage (NCBI SRA)	Amplification region	Sequencing method
Zhuang_2018 (Zhuang et al., 2018)	HC (13) IBS-D (30)	30.54 ± 6.75 32.1 ± 8.11	20.84 ± 1.46 20.61 ± 3.26	8/5 21/9	China	SRP150089	V3-V4	Illumina-MiSeq
Presti_2019 (Lo Presti et al., 2019)	HC (47), IBS (44) IBS-D (16) IBS-C (18) IBS-A (10)	54, 48 NA	23, 24 NA	20/27, 30/14 NA	Italy	SRP110018	V1-V3	454 GS Junior
Pozuelo_2015 (Pozuelo et al., 2015)	HC (66), IBS (113) IBS-D (54) IBS-C (32) IBS-M (27)	37.6 ± 13, 42.6 ± 13 41.9 ± 13 39.4 ± 10.8	23.7 ± 3.4, 23.7 ± 4 25 ± 4.6 23.3 ± 3.8	40/26, 80/33 29/25 31/2	Spain	SRP050404	V4	Illumina-MiSeq
Zhu_2019 (Zhu et al., 2019)	HC (15) IBS (15)	28.27 ± 1.56 47.67 ± 14.24	NA	7/8 1/2	China	SRP222428	V4	Illumina-HiSeq
Saulnier_2011 (Saulnier et al., 2011)	HC (22), IBS (22) IBS-D (1) IBS-C (13) IBS-U (7) other (1)	9.32 ± 1.52, 9.41 ± 1.04 9.38 ± 1.19 10 9.26 ± 1.89 9	NA	11/11, 8/ 14 5/8 0/1 3/4 0/1	US	SRP002457	V1-V3 V3-v5	454 GS FLX
Labus_2017 (Labus et al., 2017)	HC (23), IBS (29) IBS-D (10) IBS-C (11) IBS-M (5) IBS-A (1) IBS-U (2)	26.0 ± 6.48, 26.1 ± 5.72 NA	NA	14/9, 21/ 8 NA	US	SRP099239	V3-V5	454 GS FLX

NA, data not available; IBS, irritable bowel syndrome; HC, healthy controls; IBS-D, diarrhea-predominant IBS; IBS-C, constipation-predominant IBS; IBS-M, mixed IBS; IBS-A, alternating IBS; IBS-U, unsubtyped IBS.

0.04 to 0.32;  $P=0.0140$ ), and *Eubacterium coprostanoligenes* group (SMD=0.23, 95% CI: 0.09 to 0.38;  $P=0.0013$ ).

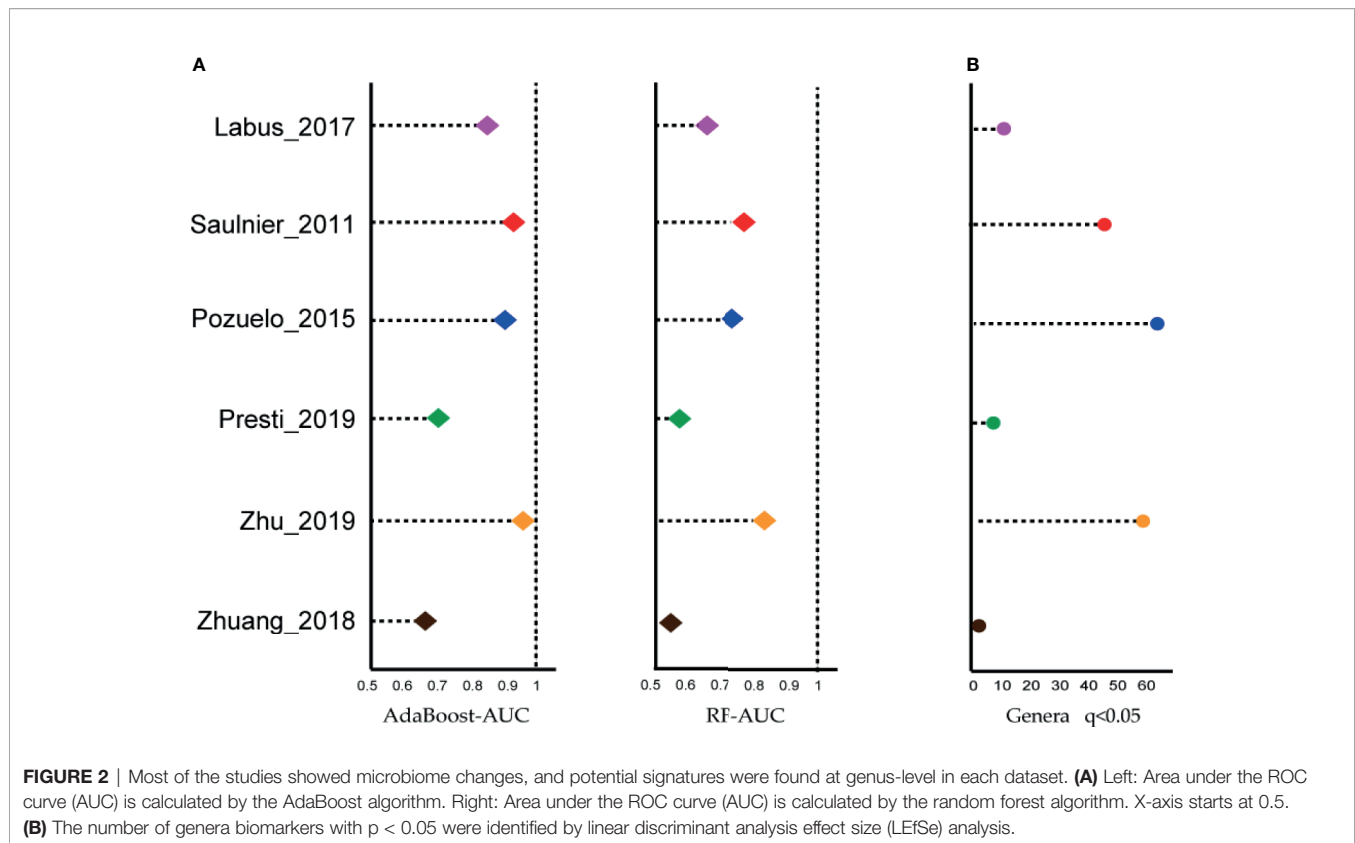
The results of the overall analysis provide robust evidence for a quantitative genus distribution of microorganisms associated with IBS, such as the nine microbiome biomarkers. The genera *Erysipelatoclostridium*, *Clostridium sensu stricto 1*, *Dorea*, *Prevotella 9*, and *Lachnoclostridium* were significantly elevated in IBS patients relative to controls, while *Coprococcus 2*, *Holdemanella*, Ruminococcaceae UCG-005, and *Eubacterium coprostanoligenes* group had a richer abundant in the controls compared to the IBS group. The variation trend for three genera, *Dorea* (Saulnier et al., 2011), *Lachnoclostridium*, and *Clostridium sensu stricto 1* (Zhu et al., 2019), were largely consistent with the initially reported results. The other statistically significant genera have hardly been highlighted in previous IBS microbiome studies or meta-analyses, and their biological functions related to IBS have not been reported so far. In despite of the inevitable heterogeneity, the nine genera signatures were important in classification. Therefore, our findings on repeatable microbial signatures for IBS might be helpful to design non-invasive diagnostic tools.

## DISCUSSION

Previous studies on dysbiosis of the gut microbiome in IBS have been reported (Zhuang et al., 2017; Pittayanon et al., 2019;

Wang et al., 2020). However, our research is the first meta-analysis based on raw sequencing data with the aid of two ML procedures. By collecting data from stool samples across studies, and analyzing them in a unified pipeline, we reassessed the role of the microbiome in IBS, and investigated several novel microbiota bacteria at the genus level that have been barely highlighted in previous studies. Importantly, our findings consolidate and extend previous reports that the gut microbiota was significantly associated with IBS and may provide a potential tool for non-invasive diagnosis. The development of non-invasive microbial biomarkers will make it possible for the diagnosis and guidance of drug use (Yu et al., 2017; Liang et al., 2020). For example, patients who are characterized by depletion of health-associated microbes could take treatments of probiotics or other interventions to enrich these taxa. Studies committed to colorectal cancer have taken a step further towards non-invasive microbial biomarkers in fecal samples. Therefore, it will be beneficial to excavate IBS-related microbial non-invasive tools.

The disease tends to have an altered microbiome profile. However, we observed inconsistent results of community diversity analysis between previous studies. Some studies report no significant difference in Shannon diversity (Labus et al., 2017; Tap et al., 2017; Jeffery et al., 2020), and others show reduced microbiota richness in the IBS group (Durban et al., 2012; Lo Presti et al., 2019). A recent meta-analysis from Lin also had an uncertain result for changes in  $\alpha$ -diversity (Wang et al., 2020),



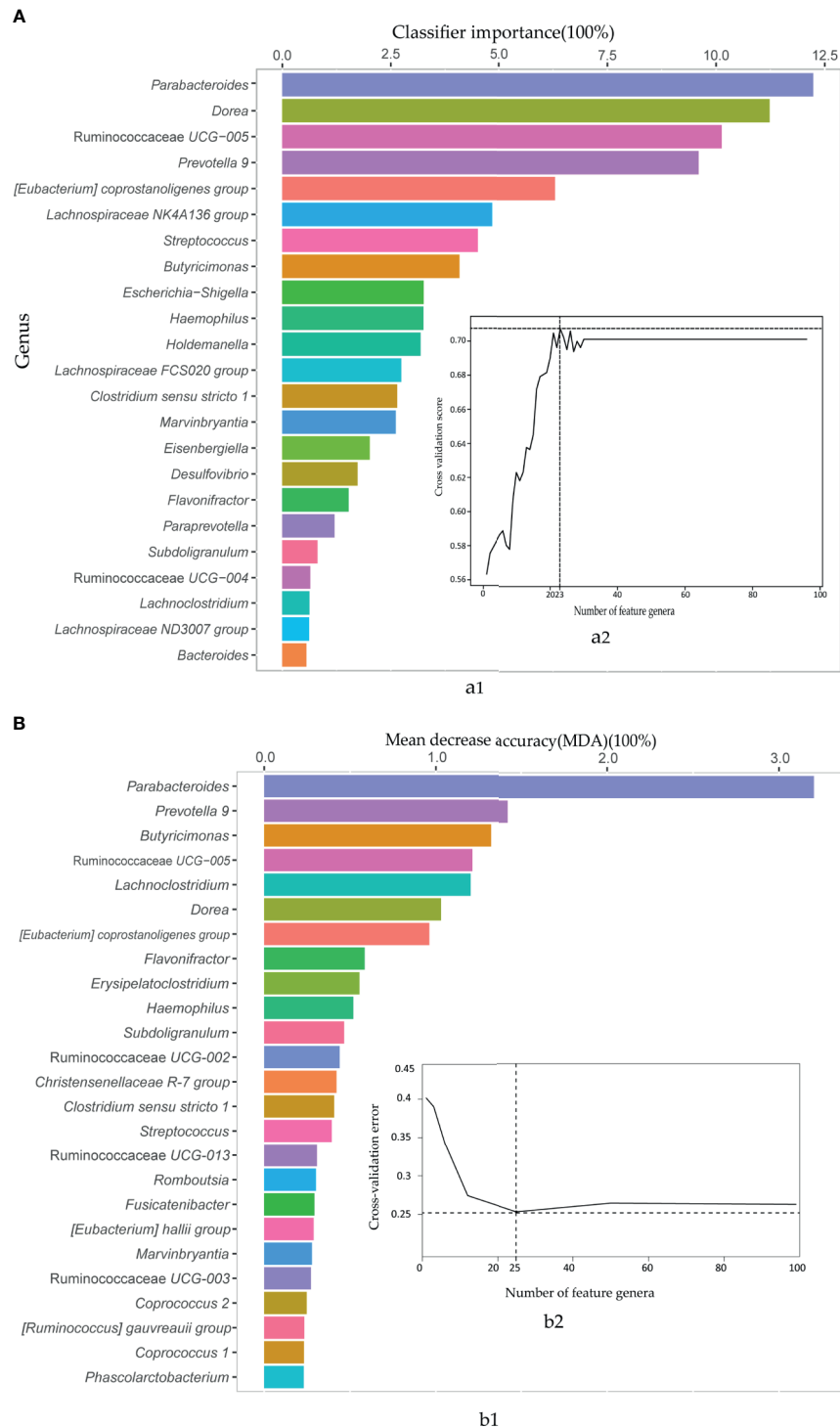
because the included studies reported conflicting findings, or too few studies were available. Fortunately, although our results also showed that the significance of  $\alpha$ -diversity or  $\beta$ -diversity differences is not a universal phenomenon, the two ML classifiers based on global genus-level data can comprehensibly classify IBS disease from healthy status. It implies that the ML algorithm might play a more significant role in the discovery of microbial profile differences.

Actually, the approach based on the machine learning algorithm has been involved in multiple medical fields, such as some case-control and cohort studies (Koyner et al., 2018; Kanda et al., 2019; Pasolli et al., 2019; Inaguma et al., 2020). Compared to classic approaches, ML procedures allow researchers to reduce the complexity of extensive data into specific classifications, shorten the computational cost time, and improve accuracy. A previous study reported no differences in fecal microbiota abundance or composition between healthy and IBS groups through classic ecologic approaches, but identified a microbial signature for IBS severity while using ML (Tap et al., 2017). Furthermore, although the results of the two algorithms differed, the use of complementary variables for the analysis allowed us to take full advantage of each analytical technique to provide more comprehensive and accurate information (Miller and Moore, 2014; Stadler and Mukherjee, 2017). There are also some inevitable limitations of ML models (Brynjolfsson and Mitchell, 2017). For example, when they do make errors, diagnosing and correcting them can be difficult because it will require going through the underlying complexities of the

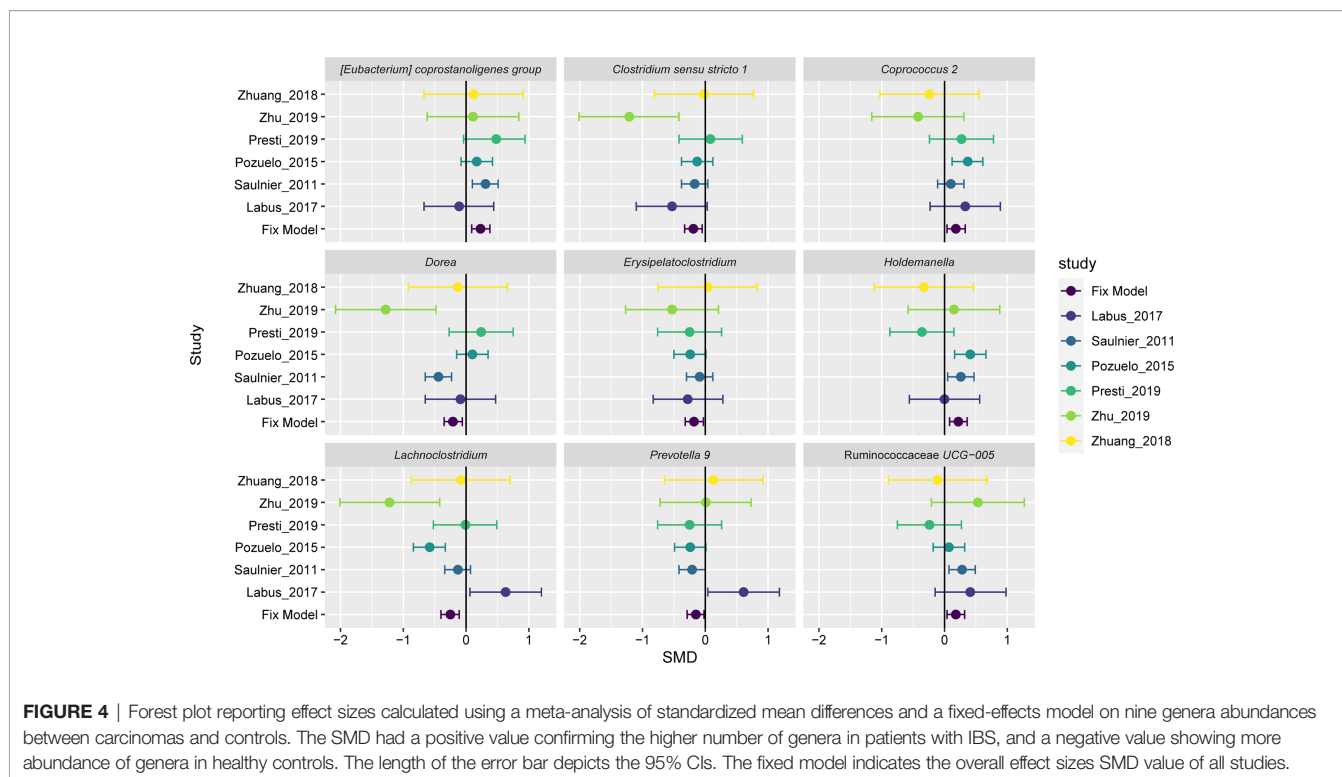
algorithms and associated processes; the bigger the data, the longer it will take. However, ML has been evolving, and our evidence confirms that this combined approach should be credible and reliable in the research of microbial markers.

Our study suggested that biomarkers selected by LEfSe could correctly classify IBS and healthy controls, and then nine key signatures were screened by ML algorithm and meta-analysis. The biological functions of the species belonging to them have been previously indicated. For instance, the *Prevotella* strains are known to be associated with chronic inflammatory conditions (Ley, 2016; Larsen, 2017). The species *Holdemanella bififormis* acting as a histone deacetylase inhibitor (HDACi) affects the activation of calcineurin and nuclear factor of activated T cells (NFAT)c3, which leads to the inhibition of tumor cell growth (Zagato et al., 2020). Both the presence of inflammation and the increased mast cell density play a critical role in the disease process of IBS, and has clinical significance (Ng et al., 2018). Besides, future studies should be designed to explore the exact biological function of these taxa, and to understand any potential role of these genera in the progression of IBS.

The genus *Parabacteroides* showed the largest contribution in the two classifiers. Previous authors also argued that its abundance correlated with 15Phe allele dosage in the sucrose-isomaltase gene, which is strongly associated with increased risk of IBS (Rangel et al., 2015). Nevertheless, there was no significant difference between two groups by meta-analysis across six studies. This finding should be viewed in the context of considerable variations in clinical and individual characteristics



**FIGURE 3** | The AdaBoost and random forest classifiers identified important feature genera in distinguishing IBS from healthy subjects. **(A)** The important features selected by the AdaBoost classifier. a1. The top 23 genus biomarkers were ranked in descending order of the most relevant features to the model. a2. 10-fold cross-validation score on the Y-axis and the number of features on the X-axis. **(B)** The important features contributed to random forest corresponding to AdaBoost. b1. 25 important genera sorted in descending order based on mean decrease accuracy (MDA). b2. 10-fold cross-validation error on the Y-axis and the number of features on the X-axis.



of included studies. For example, the researchers found differences between obese and lean individuals in this genus (Wu et al., 2019). This again hints that the ML algorithm plays a good role in disease-related microbiome analysis.

Another interesting discovery is that the genus *Brachyspira* (yet not important for any classifiers) was only present in the pediatric dataset included in our analysis. It was observed in patients with IBS that *Brachyspira* attached to the colonocyte apical membrane and linked with mast cell activation, mild mucosal inflammation, and changes of molecular pathways related to bacterial uptake (Jabbar et al., 2020). However, *Brachyspira* was not identified in the original study (Saulnier et al., 2011), possibly due to the different representative sequences or an outdated reference database (Quast et al., 2013; Duvall et al., 2017; Mancabelli et al., 2017; Gibbons et al., 2018; Armour et al., 2019). In this meta-analysis, the generation of representative sequences was based on ASVs, rather than OTUs. In general, ASVs have inherent biological significance to the DNA sequence, and provide more comprehensive inference from large marker-gene datasets, relative to *de novo* OTUs. Existing ASVs provide better sensitivity and precision than the OTU methods (Callahan et al., 2017; Caruso et al., 2019). Consequently, further applications of ASV methods and database updates should be recommended.

We adhered to a uniform pipeline for conducting the meta-analysis. However, heterogeneities are common and inevitable in meta-analyses. This was also observed in three other meta-analyses that including larger studies than ours. The inconsistency in the microbial profiles in the stools may be due to complicated and various factors among studies, for instance, individual experimental design (Sample storage, DNA extraction technique, 16S rRNA

target region) (Lozupone et al., 2013; Mancabelli et al., 2017), host-related covariates (diet, antibiotics, inflammation, age, geography, temporal, culture, and race) (Wu et al., 2011; Yatsunenko et al., 2012; Gorvitovskaia et al., 2016; Staudacher and Whelan, 2016; Rej et al., 2018; Zhong et al., 2019), and disease assessment. *Erysipelatoclostridium* has been reported to be positively associated with dairy intake (Shen et al., 2019). Abundance of *Holdemanella* has been observed to correlate with gender (Min et al., 2019). Therefore, the availability of comprehensive metadata might help to stratify the analysis and attenuate heterogeneity in further meta-analyses. Unfortunately, we failed to assess how our microbial findings were associated with the IBS subtype due to the lack of sufficient metadata. To make better use of the data collected in the original studies, data were unified and pooled to analyze instead of being segregated by subgroup to gain greater power. Meanwhile, numerous studies will need to be included to lessen substantial heterogeneities for the meta-analysis. And we highlight the significance of making raw data and associated patient metadata publicly available to enable more comprehensive analyses in the future.

Our meta-analysis with its deficiency adds new knowledge that the nine genera play an important role in distinguishing healthy individuals from the IBS group. Besides, we built ML classification models using these nine biomarkers. The AUC value of RF was 0.77, but the AdaBoost classifier had an AUC value about 0.6. This phenomenon might be caused by heterogeneity and other confounding factors. For instance, IBS is a complex and heterogeneous disease with many factors involved in its etiology and pathogenesis. Despite the unfortunate performance of these nine genera, our results provide new insight into the microbiome



dysbiosis of IBS. Further studies require targeted functional analysis of these nine signatures in IBS. Besides, we demonstrate that meta-analysis combined with machine learning algorithms may be a responsible approach in microbiome and disease research. Furthermore, meta-analysis can allow for further stratification of disease subtypes and microbiome disruption if potential influencing factors such as unavailable raw sequencing data or metadata, ambiguous and incomplete metadata can be addressed. In summary, this work demonstrates the feasibility of using a unified bioinformatics approach to pursue new findings in the broader field of clinically relevant microbiome research and enhances the value of separate analyses. As the field evolves, researchers should utilize an increasing number of replicated case-control studies to effectively translate putative microbiome ideals into clinical practice.

## DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: NCBI, accession numbers: SRP002457, PRJNA566284, SRP150089, PRJNA391149, PRJNA268708, PRJNA373876.

## AUTHOR CONTRIBUTIONS

YL: writing-original draft preparation. WL, HY, XZ, and WW: methodology and visualization. LM, YW, and HZ: writing—

review and editing. SJ, LC, and XT: supervision. LW, YJW, JS, and YS: project administration. All authors contributed to the article and approved the submitted version.

## FUNDING

This study was funded by the Science and Technology Key Program of Tianjin (No.19ZYPTJC00060).

## ACKNOWLEDGMENTS

This work is supported by the State Key Laboratory of Component-Based Chinese Medicine and the Tianjin State Key Laboratory of Modern Chinese Medicine, Tianjin University of Traditional Chinese Medicine, Tianjin, 300193, China. The language was revised by Professors Dan Wang and Xiaoxuan Tian, and the grammar was corrected by the Grammarly software.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcimb.2021.645951/full#supplementary-material>

## REFERENCES

- Abbas A., Kong X. B., Liu Z., Jing B. Y., and Gao X. (2013). Automatic Peak Selection by a Benjamini-Hochberg-Based Algorithm. *PLoS One* 8, e53112. doi: 10.1371/journal.pone.0053112
- Armour C. R., Nayfach S., Pollard K. S., and Sharpton T. J. (2019). A Metagenomic Meta-Analysis Reveals Functional Signatures of Health and Disease in the Human Gut Microbiome. *mSystems* 4 (4), e00332–18. doi: 10.1128/mSystems.00332-18
- Asghar Z., Thoufeeq M., Kurien M., Ball A. J., Rej A., David Tai F. W., et al. (2020). Diagnostic Yield of Colonoscopy in Patients With Symptoms Compatible With Rome IV Functional Bowel Disorders. *Clin. Gastroenterol. Hepatol.* S1542-3565 (20), 31221–0. doi: 10.1016/j.cgh.2020.08.062
- Boyen E., Rideout J. R., Dillon M. R., Bokulich N. A., Abnet C. C., Al-Ghalith G. A., et al. (2019). Reproducible, Interactive, Scalable and Extensible Microbiome Data Science Using QIIME 2. *Nat. Biotechnol.* 37, 852–857. doi: 10.1038/s41587-019-0209-9
- Brynjolfsson E., and Mitchell T. (2017). What can Machine Learning do? Workforce Implications. *Science* 358, 1530–1534. doi: 10.1126/science.aap8062
- Callahan B. J., McMurdie P. J., and Holmes S. P. (2017). Exact Sequence Variants Should Replace Operational Taxonomic Units in Marker-Gene Data Analysis. *ISME J.* 11, 2639–2643. doi: 10.1038/ismej.2017.119
- Callahan B. J., McMurdie P. J., Rosen M. J., Han A. W., Johnson A. J., and Holmes S. P. (2016). DADA2: High-Resolution Sample Inference From Illumina Amplicon Data. *Nat. Methods* 13, 581–583. doi: 10.1038/nmeth.3869
- Carroll I. M., Ringel-Kulka T., Keku T. O., Chang Y. H., Packey C. D., Sartor R. B., et al. (2011). Molecular Analysis of the Luminal- and Mucosal-Associated Intestinal Microbiota in Diarrhea-Predominant Irritable Bowel Syndrome. *Am. J. Physiol. Gastrointest Liver Physiol.* 301, G799–G807. doi: 10.1152/ajpgi.00154.2011
- Carroll I. M., Ringel-Kulka T., Siddle J. P., and Ringel Y. (2012). Alterations in Composition and Diversity of the Intestinal Microbiota in Patients With Diarrhea-Predominant Irritable Bowel Syndrome. *Neurogastroenterol Motil.* 24, 521–530, e248. doi: 10.1111/j.1365-2982.2012.01891.x
- Caruso V., Song X., Asquith M., and Karstens L. (2019). Performance of Microbiome Sequence Inference Methods in Environments With Varying Biomass. *mSystems* 4 (1), e00163–18. doi: 10.1128/mSystems.00163-18
- Chung C. S., Chang P. F., Liao C. H., Lee T. H., Chen Y., Lee Y. C., et al. (2016). Differences of Microbiota in Small Bowel and Faeces Between Irritable Bowel Syndrome Patients and Healthy Subjects. *Scand. J. Gastroenterol.* 51, 410–419. doi: 10.3109/00365521.2015.1116107
- Dior M., Delagrevie H., Duboc H., Jouet P., Coffin B., Brot L., et al. (2016). Interplay Between Bile Acid Metabolism and Microbiota in Irritable Bowel Syndrome. *Neurogastroenterol Motil.* 28, 1330–1340. doi: 10.1111/nmo.12829
- Durban A., Abellan J. J., Jimenez-Hernandez N., Salgado P., Ponce M., Ponce J., et al. (2012). Structural Alterations of Faecal and Mucosa-Associated Bacterial Communities in Irritable Bowel Syndrome. *Environ. Microbiol. Rep.* 4, 242–247. doi: 10.1111/j.1758-2229.2012.00327.x
- Duvallet C., Gibbons S. M., Gurry T., Irizarry R. A., and Alm E. J. (2017). Meta-Analysis of Gut Microbiome Studies Identifies Disease-Specific and Shared Responses. *Nat. Commun.* 8, 1784. doi: 10.1038/s41467-017-01973-8
- Ferreira J. A. (2007). The Benjamini-Hochberg Method in the Case of Discrete Test Statistics. *Int. J. Biostat* 3, Article 11. doi: 10.2202/1557-4679.1065
- Finucane M. M., Sharpton T. J., Laurent T. J., and Pollard K. S. (2014). A Taxonomic Signature of Obesity in the Microbiome? Getting to the Guts of the Matter. *PLoS One* 9, e84689. doi: 10.1371/journal.pone.0084689
- Gerritzen A., Wittke J. W., and Wolff D. (2011). Rapid and Sensitive Detection of Shiga Toxin-Producing Escherichia Coli Directly From Stool Samples by Real-Time PCR in Comparison to Culture, Enzyme Immunoassay and Vero Cell Cytotoxicity Assay. *Clin. Lab.* 57, 993–998.
- Gibbons S. M., Duvallet C., and Alm E. J. (2018). Correcting for Batch Effects in Case-Control Microbiome Studies. *PLoS Comput. Biol.* 14, e1006102. doi: 10.1371/journal.pcbi.1006102
- Gonzalez-Castro A. M., Martinez C., Salvo-Romero E., Fortea M., Pardo-Camacho C., Perez-Berezo T., et al. (2017). Mucosal Pathobiology and Molecular Signature of Epithelial Barrier Dysfunction in the Small Intestine in Irritable Bowel Syndrome. *J. Gastroenterol. Hepatol.* 32, 53–63. doi: 10.1111/jgh.13417

- Gorvitovskaia A., Holmes S. P., and Huse S. M. (2016). Interpreting Prevotella and Bacteroides as Biomarkers of Diet and Lifestyle. *Microbiome* 4, 15. doi: 10.1186/s40168-016-0160-7
- Hall M., and Beiko R. G. (2018). 16S Rna Gene Analysis With QIIME2. *Methods Mol. Biol.* 1849 113–129. doi: 10.1007/978-1-4939-8728-3\_8
- Hellstrom P. M., and Benno P. (2019). The Rome IV: Irritable Bowel Syndrome - a Functional Disorder. *Best Pract. Res. Clin. Gastroenterol.* 40–41, 101634. doi: 10.1016/j.bpg.2019.101634
- naguma D., Kitagawa A., Yanagiya R., Koseki A., Iwamori T., Kudo M., et al. (2020). Increasing Tendency of Urine Protein is a Risk Factor for Rapid Egfr Decline in Patients With CKD: A Machine Learning-Based Prediction Model by Using a Big Database. *PLoS One* 15, e0239262. doi: 10.1371/journal.pone.0239262
- Jabbar K. S., Dolan B., Eklund L., Wising C., Ermund A., Johansson A., et al. (2020). Association Between Brachyspira and Irritable Bowel Syndrome With Diarrhoea. *Gut* 70 (6), 1117–1129. doi: 10.1136/gutjnl-2020-321466
- Jeffery I. B., Das A., O'herlihy E., Coughlan S., Cisek K., Moore M., et al. (2020). Differences in Faecal Microbiomes and Metabolomes of People With vs Without Irritable Bowel Syndrome and Bile Acid Malabsorption. *Gastroenterology* 158, 1016–1028.e1018. doi: 10.1053/j.gastro.2019.11.301
- Jeffery I. B., O'toole P. W., Ohman L., Claesson M. J., Deane J., Quigley E. M., et al. (2012). An Irritable Bowel Syndrome Subtype Defined by Species-Specific Alterations in Faecal Microbiota. *Gut* 61, 997–1006. doi: 10.1136/gutjnl-2011-301501
- Kanda E., Kanno Y., and Katsukawa F. (2019). Identifying Progressive CKD From Healthy Population Using Bayesian Network and Artificial Intelligence: A Worksite-Based Cohort Study. *Sci. Rep.* 9, 5082. doi: 10.1038/s41598-019-41663-7
- Kassinen A., Krogius-Kurikka L., Makivuokko H., Rinttilä T., Paulin L., Corander J., et al. (2007). The Faecal Microbiota of Irritable Bowel Syndrome Patients Differs Significantly From That of Healthy Subjects. *Gastroenterology* 133, 24–33. doi: 10.1053/j.gastro.2007.04.005
- Kelly B. J., Gross R., Bittinger K., Sherrill-Mix S., Lewis J. D., Collman R. G., et al. (2015). Power and Sample-Size Estimation for Microbiome Studies Using Pairwise Distances and PERMANOVA. *Bioinformatics* 31, 2461–2468. doi: 10.1093/bioinformatics/btv183
- Koyner J. L., Carey K. A., Edelson D. P., and Churpek M. M. (2018). The Development of a Machine Learning Inpatient Acute Kidney Injury Prediction Model. *Crit. Care Med.* 46, 1070–1077. doi: 10.1097/CCM.0000000000003123
- Krogius-Kurikka L., Lyra A., Malinen E., Aarnikunnas J., Tuimala J., Paulin L., et al. (2009). Microbial Community Analysis Reveals High Level Phylogenetic Alterations in the Overall Gastrointestinal Microbiota of Diarrhoea-Predominant Irritable Bowel Syndrome Sufferers. *BMC Gastroenterol.* 9, 95. doi: 10.1186/1471-230X-9-95
- Labus J. S., Hollister E. B., Jacobs J., Kirbach K., Oezguen N., Gupta A., et al. (2017). Differences in Gut Microbial Composition Correlate With Regional Brain Volumes in Irritable Bowel Syndrome. *Microbiome* 5, 49. doi: 10.1186/s40168-017-0260-z
- Larsen J. M. (2017). The Immune Response to Prevotella Bacteria in Chronic Inflammatory Disease. *Immunology* 151, 363–374. doi: 10.1111/imm.12760
- Lebanov L., Tedone L., Ghiasvand A., and Paull B. (2020). Random Forests Machine Learning Applied to Gas Chromatography - Mass Spectrometry Derived Average Mass Spectrum Data Sets for Classification and Characterisation of Essential Oils. *Talanta* 208, 120471. doi: 10.1016/j.talanta.2019.120471
- Ley R. E. (2016). Gut Microbiota in 2015: Prevotella in the Gut: Choose Carefully. *Nat. Rev. Gastroenterol. Hepatol* 13, 69–70. doi: 10.1038/nrgastro.2016.4
- Liang J. Q., Li T., Nakatsu G., Chen Y. X., Yau T. O., Chu E., et al. (2020). A Novel Faecal Lachnospirillum Marker for the non-Invasive Diagnosis of Colorectal Adenoma and Cancer. *Gut* 69, 1248–1257. doi: 10.1136/gutjnl-2019-318532
- Liu H. N., Wu H., Chen Y. Z., Chen Y. J., Shen X. Z., and Liu T. T. (2017). Altered Molecular Signature of Intestinal Microbiota in Irritable Bowel Syndrome Patients Compared With Healthy Controls: A Systematic Review and Meta-Analysis. *Dig Liver Dis.* 49, 331–337. doi: 10.1016/j.dld.2017.01.142
- Lo Presti A., Zorzi F., Del Chierico F., Altomare A., Cocca S., Avola A., et al. (2019). Faecal and Mucosal Microbiota Profiling in Irritable Bowel Syndrome and Inflammatory Bowel Disease. *Front. Microbiol.* 10, 1655. doi: 10.3389/fmicb.2019.01655
- Lozupone C. A., Stombaugh J., Gonzalez A., Ackermann G., Wendel D., Vazquez-Baeza Y., et al. (2013). Meta-Analyses of Studies of the Human Microbiota. *Genome Res.* 23, 1704–1714. doi: 10.1101/gr.151803.112
- Ma C., Congly S. E., Novak K. L., Belletrutti P. J., Raman M., Woo M., et al. (2021). Epidemiologic Burden and Treatment of Chronic Symptomatic Functional Bowel Disorders in the United States: A Nationwide Analysis. *Gastroenterology* 160 (1), 88–98.e4. doi: 10.1053/j.gastro.2020.09.041
- Mancabelli L., Milani C., Lugli G. A., Turroni F., Cocconi D., Van Sinderen D., et al. (2017). Identification of Universal Gut Microbial Biomarkers of Common Human Intestinal Diseases by Meta-Analysis. *FEMS Microbiol. Ecol.* 93. doi: 10.1093/femsec/fix153
- Miller W. A., 3rd, and Moore P. B. (2014). Computational Study of Intramolecular Heterocyclic Ring Formation With Cyclic Phosphazenes. *Int. J. Eng. Res. Technol. (Ahmedabad)* 3, 1575–1582.
- Min Y., Ma X., Sankaran K., Ru Y., Chen L., Baiocchi M., et al. (2019). Sex-Specific Association Between Gut Microbiome and Fat Distribution. *Nat. Commun.* 10, 2408. doi: 10.1038/s41467-019-10440-5
- Moloney R. D., Johnson A. C., O'mahony S. M., Dinan T. G., Greenwood-Van Meerveld B., and Cryan J. F. (2016). Stress and the Microbiota-Gut-Brain Axis in Visceral Pain: Relevance to Irritable Bowel Syndrome. *CNS Neurosci. Ther.* 22, 102–117. doi: 10.1111/cns.12490
- Montazeri M., Montazeri M., Montazeri M., and Beigzadeh A. (2016). Machine Learning Models in Breast Cancer Survival Prediction. *Technol. Health Care* 24, 31–42. doi: 10.3233/THC-151071
- Mullin G. E., Shepherd S. J., Chander Roland B., and Ireton-Jones C. (2014). Irritable Bowel Syndrome: Contemporary Nutrition Management Strategies. *JPEN J. Parenter Enteral Nutr.* 38, 781–799. doi: 10.1177/0148607114545329
- Ng Q. X., Soh A. Y. S., Loke W., Lim D. Y., and Yeo W. S. (2018). The Role of Inflammation in Irritable Bowel Syndrome (IBS). *J. Inflammation Res.* 11, 345–349. doi: 10.2147/JIR.S174982
- Palsson O. S., Whitehead W., Tornblom H., Sperber A. D., and Simren M. (2020). Prevalence of Rome IV Functional Bowel Disorders Among Adults in the United States, Canada, and the United Kingdom. *Gastroenterology* 158, 1262–1273.e1263. doi: 10.1053/j.gastro.2019.12.021
- Pasolli E., Asnicar F., Manara S., Zolfo M., Karcher N., Armanini F., et al. (2019). Extensive Unexplored Human Microbiome Diversity Revealed by Over 150,000 Genomes From Metagenomes Spanning Age, Geography, and Lifestyle. *Cell* 176, 649–662.e620. doi: 10.1016/j.cell.2019.01.001
- Pedregosa F., Vairoquaux G., Gramfort A., Michel V., Thirion B., Grisel O., et al. (2011). Scikit-Learn: Machine Learning in Python. *J. Of Mach. Learn. Res.* 12, 2825–2830.
- Pittayanon R., Lau J. T., Yuan Y., Leontiadis G. I., Tse F., Surette M., et al. (2019). Gut Microbiota in Patients With Irritable Bowel Syndrome—a Systematic Review. *Gastroenterology* 157, 97–108. doi: 10.1053/j.gastro.2019.03.049
- Pollock J., Glendinning L., Wisedchanwet T., and Watson M. (2018). The Madness of Microbiome: Attempting to Find Consensus “Best Practice” for 16S Microbiome Studies. *Appl. Environ. Microbiol.* 84 (7), e02627–17. doi: 10.1128/AEM.02627-17
- Poulsen C. H., Eplöv L. F., Hjorthoj C., Eliassen M., Skovbjerg S., Dantoft T. M., et al. (2017). Irritable Bowel Symptoms and the Development of Common Mental Disorders and Functional Somatic Syndromes Identified in Secondary Care - a Long-Term, Population-Based Study. *Clin. Epidemiol.* 9, 393–402. doi: 10.2147/CLEP.S141344
- Pozuelo M., Panda S., Santiago A., Mendez S., Accarino A., Santos J., et al. (2015). Reduction of Butyrate- and Methane-Producing Microorganisms in Patients With Irritable Bowel Syndrome. *Sci. Rep.* 5, 12693. doi: 10.1038/srep12693
- Quast C., Pruesse E., Yilmaz P., Gerken J., Schweer T., Yarza P., et al. (2013). The SILVA Ribosomal RNA Gene Database Project: Improved Data Processing and Web-Based Tools. *Nucleic Acids Res.* 41, D590–D596. doi: 10.1093/nar/gks1219
- Quigley E. M. (2005). Disturbances of Motility and Visceral Hypersensitivity in Irritable Bowel Syndrome: Biological Markers or Epiphenomenon. *Gastroenterol. Clin. North Am.* 34221–233, vi. doi: 10.1016/j.gtc.2005.02.010
- Rajilic-Stojanovic M., Biagi E., Heilig H. G., Kajander K., Kekkonen R. A., Tims S., et al. (2011). Global and Deep Molecular Analysis of Microbiota Signatures in Faecal Samples From Patients With Irritable Bowel Syndrome. *Gastroenterology* 141, 1792–1801. doi: 10.1053/j.gastro.2011.07.043
- Rangel I., Sundin J., Fuentes S., Repsilber D., De Vos W. M., and Brummer R. J. (2015). The Relationship Between Faecal-Associated and Mucosal-Associated Microbiota in Irritable Bowel Syndrome Patients and Healthy Subjects. *Aliment Pharmacol. Ther.* 42, 1211–1221. doi: 10.1111/apt.13399

- Rej A., Avery A., Ford A. C., Holdoway A., Kurien M., Mckenzie Y., et al. (2018). Clinical Application of Dietary Therapies in Irritable Bowel Syndrome. *J. Gastrointest Liver Dis.* 27, 307–316. doi: 10.15403/jgld.2014.1121.273.avy
- Ren W. L., Wen Y. J., Dunwell J. M., and Zhang Y. M. (2018). Pkwmeb: Integration of Kruskal-Wallis Test With Empirical Bayes Under Polygenic Background Control for Multi-Locus Genome-Wide Association Study. *Heredity (Edinb)* 120, 208–218. doi: 10.1038/s41437-017-0007-4
- Rigsbee L., Agans R., Shankar V., Kenche H., Khamis H. J., Michail S., et al. (2012). Quantitative Profiling of Gut Microbiota of Children With Diarrhea-Predominant Irritable Bowel Syndrome. *Am. J. Gastroenterol.* 107, 1740–1751. doi: 10.1038/ajg.2012.287
- Saulnier D. M., Riehle K., Mistretta T. A., Diaz M. A., Mandal D., Raza S., et al. (2011). Gastrointestinal Microbiome Signatures of Pediatric Patients With Irritable Bowel Syndrome. *Gastroenterology* 141, 1782–1791. doi: 10.1053/j.gastro.2011.06.072
- Schmulson M., Ortiz O., Santiago-Lomeli M., Gutierrez-Reyes G., Gutierrez-Ruiz M. C., Robles-Diaz G., et al. (2006). Frequency of Functional Bowel Disorders Among Healthy Volunteers in Mexico City. *Dig Dis.* 24, 342–347. doi: 10.1159/000092887
- Sculco A. D. (2001). Spine Stats. The Kruskal-Wallis and Wilcoxon-Mann Whitney Tests. *Spine J.* 1, 153. doi: 10.1016/S1529-9430(01)00069-9
- Segata N., Izard J., Waldron L., Gevers D., Miropolsky L., Garrett W. S., et al. (2011). Metagenomic Biomarker Discovery and Explanation. *Genome Biol.* 12, R60. doi: 10.1186/gb-2011-12-6-r60
- Shah M. S., Desantis T. Z., Weinmaier T., Mcmurdie P. J., Cope J. L., Altrichter A., et al. (2018). Leveraging Sequence-Based Faecal Microbial Community Survey Data to Identify a Composite Biomarker for Colorectal Cancer. *Gut* 67, 882–891. doi: 10.1136/gutjnl-2016-313189
- Shen X., Wang M., Zhang X., He M., Li M., Cheng G., et al. (2019). Dynamic Construction of Gut Microbiota may Influence Allergic Diseases of Infants in Southwest China. *BMC Microbiol.* 19, 123. doi: 10.1186/s12866-019-1489-4
- Stadler N., and Mukherjee S. (2017). Two-Sample Testing in High Dimensions. *J. Of R. Stat. Soc. Ser. B-Statistical Method.* 79, 225–246. doi: 10.1111/rssb.12173
- Staudacher H. M., and Whelan K. (2016). Altered Gastrointestinal Microbiota in Irritable Bowel Syndrome and its Modification by Diet: Probiotics, Prebiotics and the Low FODMAP Diet. *Proc. Nutr. Soc.* 75, 306–318. doi: 10.1017/S0029665116000021
- Sze M. A., and Schloss P. D. (2016). Looking for a Signal in the Noise: Revisiting Obesity and the Microbiome. *mBio* 7, e01018–16. doi: 10.1128/mBio.01018-16
- Sze M. A., and Schloss P. D. (2018). Leveraging Existing 16S Rrna Gene Surveys to Identify Reproducible Biomarkers in Individuals With Colorectal Tumors. *mBio* 9 (3), e00630–18. doi: 10.1128/mBio.00630-18
- Tana C., Umesaki Y., Imaoka A., Handa T., Kanazawa M., and Fukudo S. (2010). Altered Profiles of Intestinal Microbiota and Organic Acids may be the Origin of Symptoms in Irritable Bowel Syndrome. *Neurogastroenterol Motil.* 22512-519, e114–e515. doi: 10.1111/j.1365-2982.2009.01427.x
- Tap J., Derrien M., Tornblom H., Brazeilles R., Cools-Portier S., Dore J., et al. (2017). Identification of an Intestinal Microbiota Signature Associated With Severity of Irritable Bowel Syndrome. *Gastroenterology* 152, 111–123 e118. doi: 10.1053/j.gastro.2016.09.049
- Thomas A. M., Manghi P., Asnicar F., Pasolli E., Armanini F., Zolfo M., et al. (2019). Metagenomic Analysis of Colorectal Cancer Datasets Identifies Cross-Cohort Microbial Diagnostic Signatures and a Link With Choline Degradation. *Nat. Med.* 25, 667–678. doi: 10.1038/s41591-019-0405-7
- Walters W. A., Xu Z., and Knight R. (2014). Meta-Analyses of Human Gut Microbes Associated With Obesity and IBD. *FEBS Lett.* 588, 4223–4233. doi: 10.1016/j.febslet.2014.09.039
- Wang L., Alammar N., Singh R., Nanavati J., Song Y., Chaudhary R., et al. (2020). Gut Microbial Dysbiosis in the Irritable Bowel Syndrome: A Systematic Review and Meta-Analysis of Case-Control Studies. *J. Acad. Nutr. Diet* 120, 565–586. doi: 10.1016/j.jand.2019.05.015
- Wu G. D., Chen J., Hoffmann C., Bittinger K., Chen Y. Y., Keilbaugh S. A., et al. (2011). Linking Long-Term Dietary Patterns With Gut Microbial Enterotypes. *Science* 334, 105–108. doi: 10.1126/science.1208344
- Wu T. R., Lin C. S., Chang C. J., Lin T. L., Martel J., Ko Y. F., et al. (2019). Gut Commensal Parabacteroides Goldsteini Plays a Predominant Role in the Anti-Obesity Effects of Polysaccharides Isolated From *Hirsutiella Sinensis*. *Gut* 68, 248–262. doi: 10.1136/gutjnl-2017-315458
- Yao X., Yang Y., Zhang S., Shi Y., Zhang Q., and Wang Y. (2020). The Impact of Overlapping Functional Dyspepsia, Belching Disorders and Functional Heartburn on Anxiety, Depression and Quality of Life of Chinese Patients With Irritable Bowel Syndrome. *BMC Gastroenterol.* 20, 209. doi: 10.1186/s12876-020-01357-1
- Yatsunenko T., Rey F. E., Manary M. J., Trehan I., Dominguez-Bello M. G., Contreras M., et al. (2012). Human Gut Microbiome Viewed Across Age and Geography. *Nature* 486, 222–227. doi: 10.1038/nature11053
- Yilmaz P., Parfrey L. W., Yarza P., Gerken J., Pruesse E., Quast C., et al. (2014). The SILVA and “All-Species Living Tree Project (LTP)” Taxonomic Frameworks. *Nucleic Acids Res.* 42, D643–D648. doi: 10.1093/nar/gkt1209
- Yu J., Feng Q., Wong S. H., Zhang D., Liang Q. Y., Qin Y., et al. (2017). Metagenomic Analysis of Faecal Microbiome as a Tool Towards Targeted non-Invasive Biomarkers for Colorectal Cancer. *Gut* 66, 70–78. doi: 10.1136/gutjnl-2015-309800
- Zagato E., Pozzi C., Bertocchi A., Schioppa T., Saccheri F., Guglietta S., et al. (2020). Endogenous Murine Microbiota Member *Faecalibaculum Rodentium* and its Human Homologue Protect From Intestinal Tumour Growth. *Nat. Microbiol.* 5, 511–524. doi: 10.1038/s41564-019-0649-5
- Zhen Lu W., Ann Gwee K., and Yu Ho K. (2006). Functional Bowel Disorders in Rotating Shift Nurses may be Related to Sleep Disturbances. *Eur. J. Gastroenterol. Hepatol* 18, 623–627. doi: 10.1097/00042737-200606000-00008
- Zhong H., Penders J., Shi Z., Ren H., Cai K., Fang C., et al. (2019). Impact of Early Events and Lifestyle on the Gut Microbiota and Metabolic Phenotypes in Young School-Age Children. *Microbiome* 7, 2. doi: 10.1186/s40168-018-0608-z
- Zhuang X., Tian Z., Li L., Zeng Z., Chen M., and Xiong L. (2018). Faecal Microbiota Alterations Associated With Diarrhea-Predominant Irritable Bowel Syndrome. *Front. Microbiol.* 9, 1600. doi: 10.3389/fmicb.2018.01600
- Zhuang X., Xiong L., Li L., Li M., and Chen M. (2017). Alterations of Gut Microbiota in Patients With Irritable Bowel Syndrome: A Systematic Review and Meta-Analysis. *J. Gastroenterol. Hepatol* 32, 28–38. doi: 10.1111/jgh.13471
- Zhu S., Liu S., Li H., Zhang Z., Zhang Q., Chen L., et al. (2019). Identification of Gut Microbiota and Metabolites Signature in Patients With Irritable Bowel Syndrome. *Front. Cell Infect. Microbiol.* 9, 346. doi: 10.3389/fcimb.2019.00346

**Conflict of Interest:** Authors LW, YW, JS and YS were employed by company Tianjin Zhongxin Pharmaceutical Group Co., Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Liu, Li, Yang, Zhang, Wang, Jia, Xiang, Wang, Miao, Zhang, Wang, Wang, Song, Sun, Chai and Tian. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.