European Journal of Clinical Nutrition www.nature.com/ejcn

ARTICLE

Check for updates

Prevention of Non Communicable Diseases

GWAS-associated bacteria and their metabolites appear to be causally related to the development of inflammatory bowel disease

Zhenhuang Zhuang[1,7], Nan Li[1,2,7], Jiayi Wang[3], Ruotong Yang[1], Wenxiu Wang[1], Zhonghua Liu[4] and Tao Huang[1,5,6]✉

BACKGROUND: Accumulating evidence has suggested that the imbalance of gut microbiota is commonly observed in patients with inflammatory bowel disease (IBD). However, it remains unclear whether dysbiosis is a cause or consequence of chronic intestinal inflammation. We aimed to investigate the causal relationships of gut microbiota and metabolites with IBD, including ulcerative colitis (UC) and Crohn's disease (CD).
METHODS: We applied two-sample Mendelian randomization using summary statistics from the gut microbiota genetic consortium ($n = 1812$), the Framingham Heart Study ($n = 2076$) and the International IBD Genetics Consortium ($n = 86{,}640$).
RESULTS: Using the genetic approach, the increase in *OTU10032 unclassified Enterobacteriaceae* was associated with higher risks of IBD (OR, 1.03; 95% CI, 1.00–1.06; $P = 0.033$) and CD (1.04; 1.01–1.08; $P = 0.015$). Importantly, an *Enterobacteriaceae*-related metabolite taurine was positively associated with risks of IBD (1.04; 1.01–1.08; $P = 0.016$) and UC (1.05; 1.01–1.10; $P = 0.024$). Notably, we also found betaine, a downstream product of *Enterobacteriaceae* metabolism, was causally associated with a higher risk of CD (1.10; 1.02–1.18; $P = 0.008$). In addition, increased *Erysipelotrichaceae family* were causally related to lower risks of IBD (0.88; 0.78–0.98; $P = 0.026$) and UC (0.86; 0.75–0.99; $P = 0.042$), and *Actinobacteria class* (0.80; 0.65–0.98; $P = 0.028$) and *Unclassified Erysipelotrichaceae* (0.79; 0.64–0.98; $P = 0.036$) were associated with lower risks of UC and CD, respectively.
CONCLUSIONS: Our finding provided new insights into the key role of gut metabolites such as taurine and betaine in host-microbiota interactions of IBD pathogenesis, indicating that host-microbe balance strongly influences inflammatory conditions.

European Journal of Clinical Nutrition; https://doi.org/10.1038/s41430-022-01074-w

INTRODUCTION

Ulcerative colitis (UC) and Crohn's disease (CD), collectively known as inflammatory bowel disease (IBD), are chronic inflammatory disorders of the gastrointestinal tract, resulting from alterations in intestinal microbes and the immune system [1, 2]. The risk of IBD was genetically related to host pathways that implicated the potential role for abnormal immune responses to gut microbiota and metabolites [3–5]. Therefore, documentation of host-microbiota interactions in IBD pathogenesis can inform on novel targets for clinical prevention and treatment [6].

Recent animal studies have suggested that an abnormally composed microbiota (known as "dysbiosis") is commonly observed in IBD, but generalizing these results from animal models to humans has proven challenging [7, 8]. Notably, several observational studies showed that IBD status was associated with the alterations of gut microbiota, especially the increased *Enterobacteriaceae* family [9–11], which was not substantially different between IBD cases and healthy controls in another study [12]. In contrast, this study revealed a 10-fold decrease in total bacterial load in the IBD-subset, as well as members of the *Lachnospiraceae* and *Bacteroidetes* diminished in quantity in the IBD-subset samples [12]. The conflicting findings in gut microbial dysbiosis of IBD were possibly due to confounding or reverse causation in observational studies.

The Mendelian randomization (MR) approach, which used genetic variant as instrumental variable in epidemiological study, has been widely accepted to explore the potential causal effect of exposure on diseases [13]. The reasons for using MR as a more reliable causal inference method than conventional observational studies include Mendel's laws and the fact that MR method is analogous to a randomized controlled trial (RCT) where genotypes of germline genetic variation are defined at conception and are generally not associated with conventional confounders of observational studies [14, 15]. Previous genetic studies have

[1]Department of Epidemiology & Biostatistics, School of Public Health, Peking University, 100191 Beijing, China. [2]Institute of Reproductive and Child Health/Chinese National Health Commission Key Laboratory of Reproductive Health, Peking University Health Science Center, Dongguan, China. [3]Department of Pharmacy, Peking University First Hospital, Beijing, China. [4]Department of Statistics and Actuarial Science, The University of Hong Kong, Hong Kong, China. [5]Center for Intelligent Public Health, Academy for Artificial Intelligence, Peking University, 100191 Beijing, China. [6]Key Laboratory of Molecular Cardiovascular Sciences (Peking University), Ministry of Education, 100191 Beijing, China. [7]These authors contributed equally: Zhenhuang Zhuang, Nan Li. ✉email: huangtaotao@pku.edu.cn

Received: 8 September 2021 Revised: 31 December 2021 Accepted: 10 January 2022
Published online: 19 January 2022

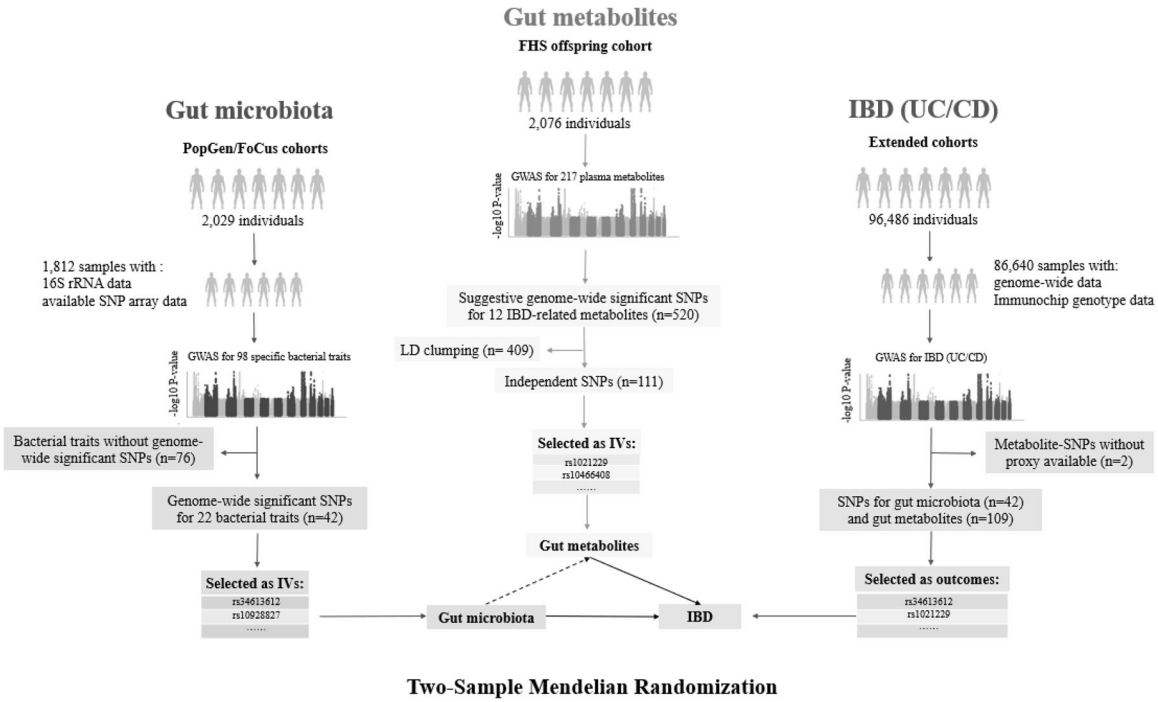

Fig. 1 Flowchart of the data selection process. Figure 1 is a flowchart of the data selection process, highlighting for each step of our study design.

Table 1. Description of gut microbiota, metabolites, IBD, UC, and CD.

	Consortium or study	Sample size	Populations	Journal	Year
Gut					
Gut microbiota	PopGen/FoCus[21]	1812 individuals	European	Nat Genet.	2016
Gut metabolite	FHS[34]	2076 individuals	European	Cell Metab.	2013
Disease					
IBD	International IBD Genetics Consortium[35]	12,882 cases and 21,770 controls	European	Nat Genet.	2015
UC		6968 cases and 20,464 controls			
CD		5956 cases and 14,927 controls			

Abbreviations: *FoCus* Food-chain plus; *FHS* Framingham heart study.

demonstrated that the host genetic variants can influence the composition and abundance of gut microbiota, allowing us to infer the relationship of gut microbiota with IBD based on this MR method [16–18].

Therefore, in the present study, we performed a two-sample MR analysis using summary data from genome-wide association studies (GWASs) to examine the causal association of gut microbiota and metabolites with IBD, including UC and CD.

METHODS

Study design overview

We applied the two-sample MR design to assess the causal relationships of core gut microbiota with IBD, UC, and CD, and the flowchart of the data selection process can be seen in Fig. 1. Generally, the MR method requires that the genetic instruments are associated with the exposure of interest (assumption 1), and genetic instruments are not associated with confounders (assumption 2) and genetic instruments influence risk of the outcome only through the exposure but not through any causal pathways (assumption 3) [19]. Given that assumption 1 and 2 is empirically verifiable, careful consideration of potential violations of assumption 3, which due to such as population stratification, linkage disequilibrium (LD), or horizontal pleiotropy, is essential to minimize bias [20]. An SNP that violates these assumptions is regarded as an invalid instrumental variable whose inclusion in MR analyses may bias the results. Ethical approval and informed consent from all participants for each study included in the MR analysis can be found in the original articles.

Data sources and instruments

Gut microbiota. We leveraged summary statistics from a recent GWAS of gut microbiota conducted among two independent but geographically matched cohorts of European ancestry ($n = 1812$) using 16 S rRNA gene sequencing (Table 1) [21], involving a total of 38 and 374 identified phyla and genera respectively. Water, alcohol, and all other highly correlated nutritional variables were further used in the GWAS analysis as covariates. Then we selected a "core measurable microbiota" defined by this GWAS after removing rare bacteria, including 40 operational taxonomic units (OTUs) and 58 taxa ranging from the genus to the phylum level. Ultimately, the GWAS identified 54 genome-wide significant associations involving 40 loci and 22 bacterial traits ($P < 5 \times 10^{-8}$; Supplementary Table 1).

Gut microbial metabolites. Considering the central roles of intestinal metabolites in host-microbiota interactions, we further chose several metabolites which were potentially associated with IBD according to previous studies, including niacinamide, pantothenic acid [22], taurine [23–25], propionic acid [26], indole-3-propionate [27], carnosine [3], ribose, serotonin [28], TMAO, betaine, choline, and carnitine [29–32]. We searched PubMed for GWASs of the gut metabolites and drew on summary-level

data from a recent GWAS of the human metabolome conducted among 2,076 participants of the Framingham Heart Study (Table 1) [33]. For each metabolite, we selected single nucleotide polymorphisms (SNPs) at thresholds for suggestive genome-wide significance ($P < 1 \times 10^{-5}$) from the GWAS.

Intestinal diseases

For disease outcomes, we selected summary statistics from the first trans-ethnic GWAS of IBD [34]. Corresponding effect estimates on the risk of IBD, UC, and CD were obtained from the International IBD Genetics Consortium of 86,640 European individuals, which is the largest GWAS published to date for IBD [34]. Notably, the IBD diagnosis was based on accepted radiologic, endoscopic, and histopathologic evaluation. All included cases fulfill clinical criteria for IBD. In addition, association tests were carried out using 15, 7, or 10 principal components for IBD, UC, or CD, respectively, as covariates, chosen from the first 20 principal components. The methods of these GWAS are described in detail elsewhere [21, 33, 34].

Statistical analysis

Linkage disequilibrium assessment and pleiotropy assessment. To verify that the SNPs selected in this study met the assumption 1 and 2, we examined that genetic association with each microbiota or metabolite, and further measured LD between all the SNPs for each trait, and finally selected independent genetic variants [35]. We chose the variant with the lowest *P* value for association with each trait if genetic variants are in LD. We used MR-Egger regression to assess the presence of pleiotropic effects, in which the SNP's effect upon each exposure is plotted against its effect upon outcomes, and an intercept distinct from the origin provides evidence for pleiotropic effects [36]. Therefore, the associations we identified in the present study did not violate assumption 3.

Mendelian randomization analysis. Analyses were performed using R version 3.5.3 (R Project for Statistical Computing). We applied the online tutorial (https://mrcieu.github.io/TwoSampleMR/) as a guidance to conduct the appropriate steps of data management prior to analysis. A Bonferroni-adjusted *p* value of 7.6×10^{-4} ($p = 0.05/66$) or 1.4×10^{-3} ($p = 0.05/36$) was used as the threshold for statistical significance for the association between gut microbiota or metabolites and IBD subtypes, respectively. $P \leq 0.05$ but above the Bonferroni corrected significance threshold was considered as suggestive of evidence for a potential association. Then we run the command of "clump" to retain only independent SNPs for use in MR analysis with default settings. Next, data for each of the chosen SNPs was extracted from summary statistics of the outcome (i.e., IBD). Data harmonization was conducted for each combination of bacterial trait and IBD subtype to verify the presence of corresponding effect alleles.

MR analyses were then performed for the methods list, including four different methods: inverse variance weighted (IVW), weighted mode, weighted median, and MR Egger. The IVW method is the default and simplest method, which provides a combined estimate of the causal estimate from each SNP. IVW is equivalent to a two-stages least squares or allele score analysis using individual-level data, and is hence considered here as conventional MR [37]. In the present MR analyses, weighted median, weighted median, and MR-Egger methods were considered as sensitivity analyses for MR investigations with multiple genetic variants. We also conducted MR analyses by single SNP and ran leave one out analysis examining heterogeneity and horizontal pleiotropy. Detailed information on MR methods and its assumptions, as well as additional tests, has been previously described [37–40]. Power calculations for MR were conducted based on the website: mRnd (http://cnsgenomics.com/shiny/mRnd/). Statistical code is available on request from the corresponding author at huangtao@bjmu.edu.cn.

RESULTS

Characteristics of selected SNPs

Table 1 shows the sample size, population, and publication year of gut microbiota, metabolites and IBD subtypes. The characteristics of the selected SNPs for each gut microbiota or metabolite are presented in Supplementary Tables 1, 2. Furthermore, 2 metabolite-associated SNPs were excluded from the association analysis without available proxy in the GWAS for IBD. In all, we identified 42 genome-wide significant SNPs for 22 bacterial traits and 109 suggestive genome-wide significant SNPs for 12 microbial metabolites. This process was described in Fig. 1. We found that all SNPs for the same trait show strong association (F-statistic > 10, the strength of the instrument) (Supplementary Table 1, 2).

Causal effects of gut microbiota and metabolites on intestinal diseases

We found that the host-genetic-driven increase in *OTU10032 unclassified Enterobacteriaceae* was related to higher risks of IBD (per relative abundance: OR, 1.03; 95% CI, 1.00–1.06; $P = 0.033$) and CD (1.04; 1.01–1.08; $P = 0.015$) (Figs. 2 and Supplementary Table 3, Supplementary Fig. 1, 2). Importantly, an *Enterobacteriaceae*-related metabolite taurine was positively associated with

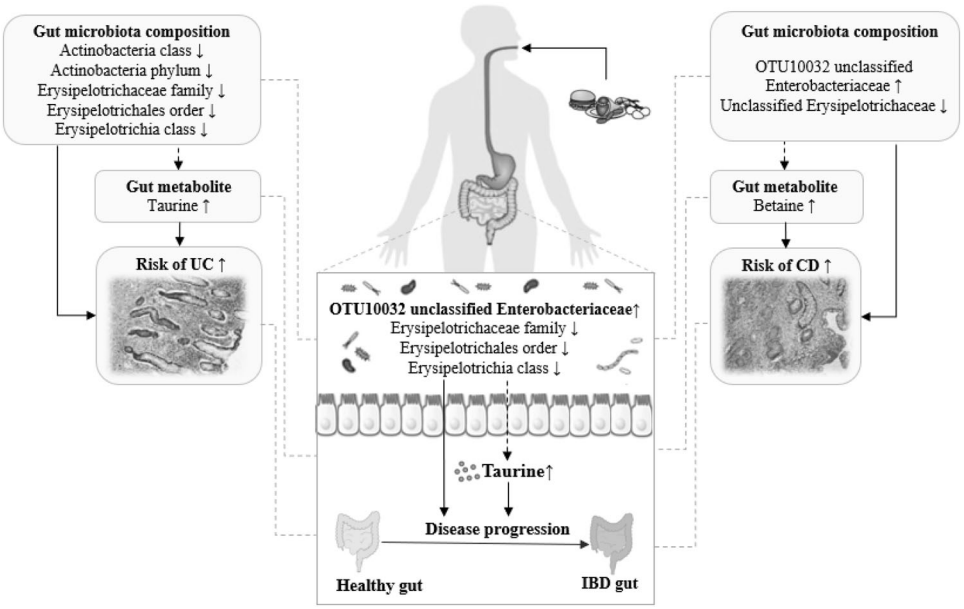

Fig. 2 Potential pathways underpinning the relationship between gut microbiota and inflammatory bowel disease. Although the exact mechanism underpinning the association between gut microbiota and risk of IBD is unknown, we speculate a potential mechanism that microbial metabolites play essential roles in the inflammatory process.

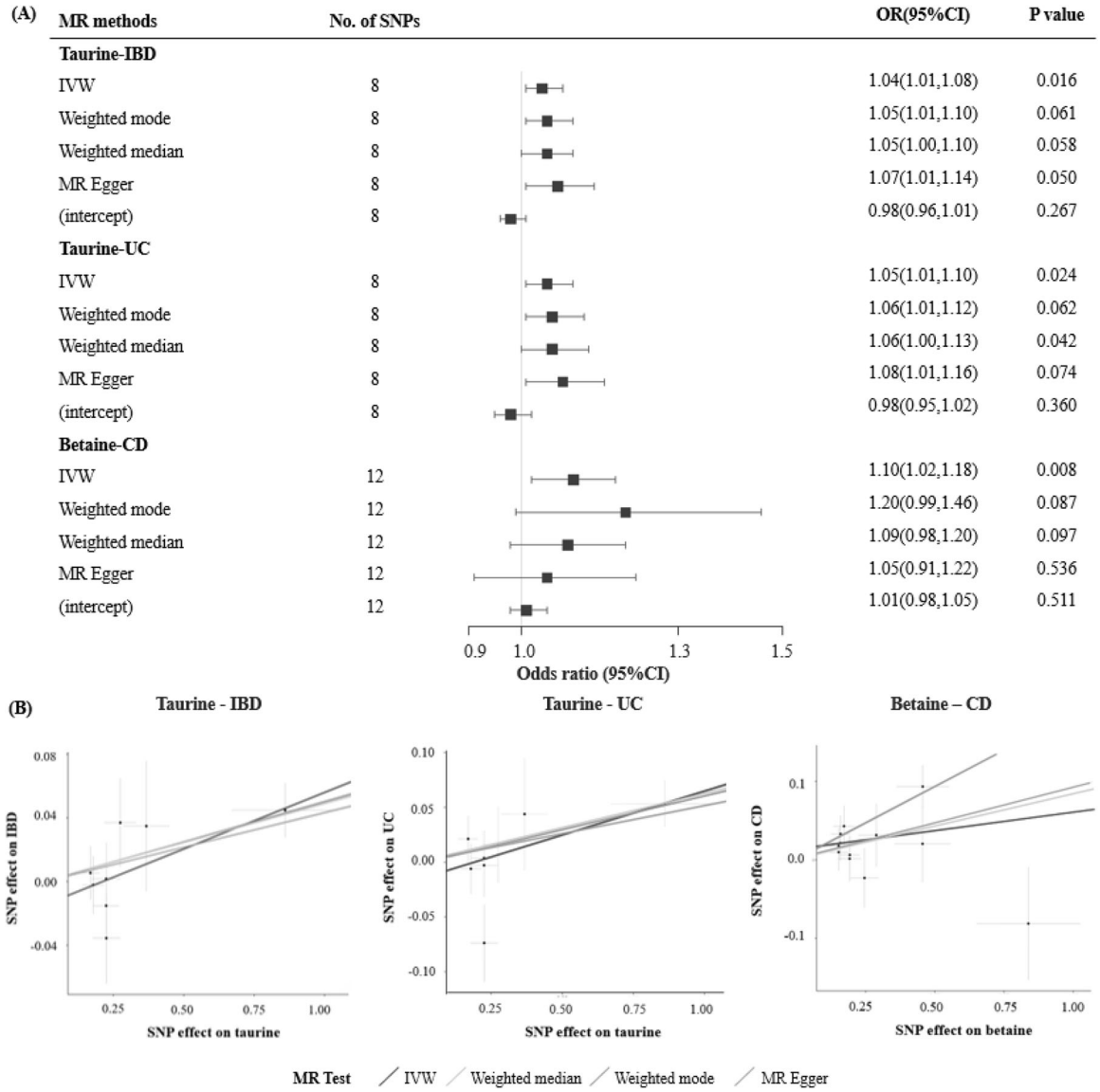

Fig. 3 Mendelian randomization study for gut metabolites and risk of IBD. A MR-derived associations between genetically predicted plasma gut metabolites levels (10 units increase) with risk of diverse forms of IBD. **B** Associations of genetically predicted taurine with risk of IBD (left), of taurine with risk of UC (middle), of betaine with risk of CD (right).

IBD (per 10 units: 1.04; 1.01–1.08; $P = 0.016$) and UC (1.05; 1.01–1.10; $P = 0.024$) (Figs. 2 and 3 and Supplementary Fig. 3, 4). We also found betaine, which might serve as a downstream product of *Enterobacteriaceae* metabolism, was causally associated with a higher risk of CD (1.10; 1.02–1.18; $P = 0.008$) (Figs. 2, 3, Supplementary Fig. 5). The results of sensitivity analyses that used the weighted mode and weighted median methods showed directionally consistent trends. Furthermore, the MR results of single genetic variant suggested that the associations of taurine with IBD and UC were driven by a genetic variant (rs13088785) and did not remain significant when we excluded this variant in leave-one-out analysis (Supplementary Fig. 3 and 4 and Supplementary Table 4). The odds ratio was changed to 0.99 (0.92–1.06; $P = 0.694$) and 0.97 (0.86–1.09, $P = 0.586$) for associations of taurine with IBD and UC after we excluded the variant (rs13088785), respectively. In MR-Egger analysis, there was no evidence of directional pleiotropy (all $P \geq 0.26$), while the precisions of the causal estimates and intercepts were low.

In addition, genetically increased *Erysipelotrichaceae family*, *Erysipelotrichales order*, and *Erysipelotrichia class* were related to lower risks of IBD (0.88; 0.78–0.98; $P = 0.026$) and UC (0.86; 0.75–0.99; $P = 0.042$) (Fig. 2, Supplementary Table 3). We also identified *Actinobacteria class* and *Actinobacteria phylum* to be associated with a lower risk of UC (0.80; 0.65–0.98; $P = 0.028$), and *Unclassified Erysipelotrichaceae* to be associated with a lower risk of CD (0.79; 0.64–0.98; $P = 0.036$) (Fig. 2, Supplementary Table 3). No significant result was found for any of other selected gut microbiota or metabolites with IBD subtypes.

DISCUSSION

Using genetic instruments, the present MR study assessed the potential causal association of gut microbiome and its metabolites with IBD, including UC and CD. We provided evidence of causal relationships between genetically increased *OTU10032 unclassified Enterobacteriaceae* with higher risks of IBD and CD. Importantly, several *Enterobacteriaceae*-related metabolites such as taurine and betaine have also been positively associated with the risk of IBD subtypes. In addition, other species including *Erysipelotrichaceae family* and *Actinobacteria class* were associated with lower risks of

IBD subtypes. Our findings implicated the key role of gut metabolites such as taurine and betaine in host-microbiota crosstalk of IBD pathogenesis, underscoring the importance of modulating host-microbe balance in the prevention of IBD.

Recently, the associations among altered gut microbiota, microbial metabolites, and IBD status have become increasingly clear, but still poorly understood. Animal studies have indicated that inflammation may cause the dysbiosis of gut microbiota and the overgrowth of *Enterobacteriaceae*, even in mice that are not genetically predisposed to immunopathologic responses [41, 42]. Another animal experiment has shown that the presence of *Enterobacteriaceae* correlates with colitis in T-bet $-/-$ $\times$ Rag2 $-/-$ ulcerative colitis (TRUC) mice and that TRUC derived strains in conjunction with an endogenous microbial community can incite colitis in wild type mice [42]. While *Enterobacteriaceae* was overrepresented, *Erysipelotrichia* was underrepresented in dogs with IBD [43, 44]. Additionally, the facultative anaerobic taxa *Enterobacteriaceae* increased whereas *Actinobacteria* decreased in cats with chronic enteropathies, which showed patterns of dysbiosis similar to those in IBD persons [8]. It was noted that microbial metabolites including taurine and TMAO were also associated with intestinal dysbiosis and thereby potentially induced the development of colitis [23, 45]. The previous study has suggested the administration of taurine reduced the inflammatory parameters in this rat model of IBD by increasing the defenses against oxidative insult [46]. Additionally, a recent review has implied that betaine is known to function physiologically as an important osmoprotectant and methyl group donor, which has anti-inflammatory functions in numerous diseases [47]. Taken together, these results point the way towards a comprehensive understanding of host-microbiota crosstalk in IBD pathogenesis across different mammalian models. However, the ability to generalize the results of these animal experiments to people has proven limited. Therefore, further population-based researches are needed to explore the causal relationship between gut microbiota and IBD.

Our MR findings had some consistency with those of observational studies (nå 1000), indicating changes in bacteria, such as increased *Enterobacteriaceae*, *Actinobacteria*, and decreased *Erysipelotrichales*, were strongly correlated with IBD, although we found *Actinobacteria* was inversely associated with the disease [10, 11]. In addition, small case-control studies showed that increased amounts of *Enterobacteriaceae* were observed in patients with CD, indicating inflammatory environment of the ileum might favor the growth of this bacterial clade [9, 48]. Notably, as for gut metabolites, taurine levels increased and taurine and hypotaurine metabolism enriched in UC patients according to a case-control study ($n = 110$) [49]. However, our results were less consistent with a culture-independent rRNA sequence analysis showing no association of the family *Enterobacteriaceae* with IBD status while members of the Lachnospiraceae and Bacteroidetes diminished in quantity in the IBD-subset samples, possibly owing to that confounding or reverse causation is inevitable in observational design [12]. Furthermore, an observational study ($n = 479$) suggested that plasma TMAO levels were lower in IBD cases compared with healthy controls, which was also different from our conclusions [29]. In fact, these observational studies of the gut microbiota might be susceptible to reverse causation and unmeasured confounders like dietary factors. Rodriguez et al. showed that the basal diet of mice determined the long-term composition of their gut microbiome and the mouse phenotypes to a greater extent than the transfer of largely different fecal microbiomes obtained from lean or obese human donors [50]. In order to overcome some of the intricate host-microbe-diet interaction problems, a simplified experimental mouse model has also been developed [51]. Notably, our MR analyses using microbiota-associated genetic variants as unconfounded surrogates for lifelong exposure of gut microbiota supported the claim that specific bacterial traits were causal for IBD, emphasizing the importance of nutritional interventions in the prevention and treatment of IBD.

Prominent changes in the gut microbiota play a key role in the development of IBD, but the underlying mechanisms remain unknown. Recently, a family-based observational study ($n = 90$) showed the IBD-associated microbial and metabolomics states included increases in *Enterobacteriaceae* and taurine levels which were highly correlated, suggesting that they represented an integrated ecosystem [52]. Subsequently, microbial metabolites caused colitis by regulating NLRP6 inflammasome signaling, epithelial IL-18 secretion and downstream antimicrobial peptides [23]. Of interest, we found that *Enterobacteriaceae* and taurine was causally associated with IBD in this MR analysis, which confirmed the hypothesis. A case-control study showed that intestinal inflammation strongly correlated with various faecal metabolites such as taurine in UC patients, which had some consistency with our findings [49].

In addition, an correlation analysis showed that *Enterobacteriaceae* was significantly associated with the cut-Kp gene cluster, which might predict higher TMAO levels [53]. However, we found no association of TMAO with any IBD subtypes. As we know, some *Enterobacteriaceae* species oxidize choline via betaine aldehyde to the osmotic protectant betaine, indicating an association between *Enterobacteriaceae* and betaine [54]. Therefore, we assumed the potential mechanism might be that dietary betaine is metabolized in the liver by intestinal bacteria to produce TMA [55], which is reported to be a key factor of IBD in mice model [53]. Interestingly, our MR results suggested that *Enterobacteriaceae family* was a cause of IBD, whereas *Enterobacteriaceae*-related betaine was also causally associated with CD, further verifying the assumption. In fact, it is essential to determine whether there are causal relationships of the family *Enterobacteriaceae* and its downstream metabolites with IBD, which will be the focus of current and future metagenomic studies on intestinal microbial ecology in various populations of IBD patients.

Major strengths of the present study include taking advantage of the large sample size of publicly available summarized data from GWASs, which has sufficient power for estimating reliable and lifelong causality. This MR study systematically examined the causal effect of gut microbiota and its metabolites on IBD, including UC and CD. Importantly, we applied four different MR approaches and additional sensitivity analysis to demonstrate the robustness of causal estimations.

However, this study has several limitations. First, our findings were not robust to Bonferroni-adjusted significance, but this MR analysis serves as a hypothesis-driven study testing epidemiologically established associations based on enough physiological evidence. Second, bi-directional MR studies designed to assess whether IBD subtypes influence the gut microbiota might be informative. However, IBD-associated SNPs were not available in the GWAS for gut microbiota, which prevented us from conducting reverse MR analysis. Third, the findings might be influenced by weak instrument bias, although we selected genome-wide significant SNPs associated with 22 bacterial traits, while our F statistics suggested instruments are valid and reliable. In addition, power was limited for the analysis of gut microbiota on IBD risk, so larger GWASs are necessary to distinguish associations with small effect sizes from null associations. Fourth, we could not exclude the possibility that such associations might share genetic basis rather than causal relationship since human genome influenced both gut microbiota and intestinal disease. Fifth, another potential source of bias was population stratification, which was reduced in our study because the selected GWASs were restricted to individuals of European ancestry. Sixth, the 16 S rRNA gene sequencing only allowed resolution from the genus to the phylum level instead of a more specific level, thus the results biased when some specific species influenced the risk of IBD.

Finally, since we conducted two-sample MR analysis using summary-level GWASs without individual data, it is difficult for us to examine the influence of diet in the present study. Further investigation for the role of diet between gut microbiota and IBD risk is warranted.

In summary, our findings supported causal relationships between gut microbiome and IBD subtypes, indicating the important roles of microbial metabolites such as taurine and betaine in microbiota-host crosstalk in IBD. Further population-based studies regarding the potential mechanisms of gut microbiota in the development of IBD are required.

DATA AVAILABILITY

All data used in the present study were obtained from genome-wide association study summary statistics which were publicly released by genetic consortia.

REFERENCES

1. Alatab S, Sepanlou SG, Ikuta K, Vahedi H, Bisignano C, Safiri S, et al. The global, regional, and national burden of inflammatory bowel disease in 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet Gastroenterol Hepatol. 2020;5:17–30.
2. Ni J, Wu GD, Albenberg L, Tomov VT. Gut microbiota and IBD: causation or correlation? Nat Rev Gastroenterol Hepatol. 2017;14:573–84.
3. Schirmer M, Garner A, Vlamakis H, Xavier RJ. Microbial genes and pathways in inflammatory bowel disease. Nat Rev Microbiol. 2019;17:497–511.
4. Chen L, Wilson JE, Koenigsknecht MJ, Chou WC, Montgomery SA, Truax AD, et al. NLRP12 attenuates colon inflammation by maintaining colonic microbial diversity and promoting protective commensal bacterial growth. Nat Immunol. 2017;18:541–51.
5. Lamas B, Richard ML, Leducq V, Pham HP, Michel ML, Da Costa G, et al. CARD9 impacts colitis by altering gut microbiota metabolism of tryptophan into aryl hydrocarbon receptor ligands. Nat Med. 2016;22:598–605.
6. Caruso R, Lo BC, Núñez G. Host-microbiota interactions in inflammatory bowel disease. Nat Rev Immunol. 2020;20:411–26.
7. Lupp C, Robertson ML, Wickham ME, Sekirov I, Champion OL, Gaynor EC, et al. Host-mediated inflammation disrupts the intestinal microbiota and promotes the overgrowth of Enterobacteriaceae. Cell Host Microbe. 2007;2:119–29.
8. Marsilio S, Pilla R, Sarawichitr B, Chow B, Hill SL, Ackermann MR, et al. Characterization of the fecal microbiome in cats with inflammatory bowel disease or alimentary small cell lymphoma. Sci Rep. 2019;9:19208.
9. Morgan XC, Tickle TL, Sokol H, Gevers D, Devaney KL, Ward DV, et al. Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment. Genome Biol. 2012;13:R79.
10. Zhou Y, Xu ZZ, He Y, Yang Y, Liu L, Lin Q, et al. Gut microbiota offers universal biomarkers across ethnicity in inflammatory bowel disease diagnosis and infliximab response prediction. mSystems. 2018;3:e00188–17.
11. Gevers D, Kugathasan S, Denson LA, Vázquez-Baeza Y, Van Treuren W, Ren B, et al. The treatment-naive microbiome in new-onset Crohn's disease. Cell Host Microbe. 2014;15:382–92.
12. 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 USA. 2007;104:13780–5.
13. Emdin CA, Khera AV, Kathiresan S. Mendelian randomization. JAMA. 2017;318:1925–6.
14. Smith GD, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? Int J Epidemiol. 2003;32:1–22.
15. Smith GD, Lawlor DA, Harbord R, Timpson N, Day I, Ebrahim S. Clustered environments and randomized genes: a fundamental distinction between conventional and genetic epidemiology. PLoS Med. 2007;4:e352.
16. Sanna S, van Zuydam NR, Mahajan A, Kurilshikov A, Vich Vila A, Võsa U, et al. Causal relationships among the gut microbiome, short-chain fatty acids and metabolic diseases. Nat Genet. 2019;51:600–5.
17. Turpin W, Espin-Garcia O, Xu W, Silverberg MS, Kevans D, Smith MI, et al. Association of host genome with intestinal microbial composition in a large healthy cohort. Nat Genet. 2016;48:1413–7.
18. Goodrich JK, Davenport ER, Beaumont M, Jackson MA, Knight R, Ober C, et al. Genetic determinants of the gut microbiome in UK twins. Cell Host Microbe. 2016;19:731–43.
19. VanderWeele TJ, Tchetgen Tchetgen EJ, Cornelis M, Kraft P. Methodological challenges in mendelian randomization. Epidemiology. 2014;25:427–35.
20. Hartwig FP, Borges MC, Horta BL, Bowden J, Davey, Smith G. Inflammatory biomarkers and risk of schizophrenia: a 2-sample mendelian randomization study. JAMA Psychiatry. 2017;74:1226–33.
21. Wang J, Thingholm LB, Skiecevičienė J, Rausch P, Kummen M, Hov JR, et al. Genome-wide association analysis identifies variation in vitamin D receptor and other host factors influencing the gut microbiota. Nat Genet. 2016;48:1396–406.
22. 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.
23. Levy M, Thaiss CA, Zeevi D, Dohnalová L, Zilberman-Schapira G, Mahdi JA, et al. Microbiota-modulated metabolites shape the intestinal microenvironment by regulating NLRP6 inflammasome signaling. Cell. 2015;163:1428–43.
24. Walker A, Schmitt-Kopplin P. The role of fecal sulfur metabolome in inflammatory bowel diseases. Int J Med Microbiol: IJMM. 2021;311:151513.
25. Guzior DV, Quinn RA. Review: microbial transformations of human bile acids. Microbiome. 2021;9:140.
26. Dorrestein PC, Mazmanian SK, Knight R. Finding the missing links among metabolites, microbes, and the host. Immunity. 2014;40:824–32.
27. Roager HM, Licht TR. Microbial tryptophan catabolites in health and disease. Nat Commun. 2018;9:3294.
28. Kwon YH, Wang H, Denou E, Ghia J-E, Rossi L, Fontes ME, et al. Modulation of gut microbiota composition by serotonin signaling influences intestinal immune response and susceptibility to colitis. Cell Mol Gastroenterol Hepatol. 2019;7:709–28.
29. Wilson A, Teft WA, Morse BL, Choi Y-H, Woolsey S, DeGorter MK, et al. Trimethylamine-N-oxide: a novel biomarker for the identification of inflammatory bowel disease. Dig Dis Sci. 2015;60:3620–30.
30. Krautkramer KA, Fan J, Bäckhed F. Gut microbial metabolites as multi-kingdom intermediates. Nat Rev Microbiol. 2021;19:77–94.
31. Koh A, Bäckhed F. From association to causality: the role of the gut microbiota and its functional products on host metabolism. Mol Cell. 2020;78:584–96.
32. Schicho R, Shaykhutdinov R, Ngo J, Nazyrova A, Schneider C, Panaccione R, et al. Quantitative metabolomic profiling of serum, plasma, and urine by (1)H NMR spectroscopy discriminates between patients with inflammatory bowel disease and healthy individuals. J Proteome Res. 2012;11:3344–57.
33. Rhee EP, Ho JE, Chen MH, Shen D, Cheng S, Larson MG, et al. A genome-wide association study of the human metabolome in a community-based cohort. Cell Metab. 2013;18:130–43.
34. Liu JZ, van Sommeren S, Huang H, Ng SC, Alberts R, Takahashi A, et al. Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. Nat Genet. 2015;47:979–86.
35. Lawlor DA, Harbord RM, Sterne JA, Timpson N, Davey, Smith G. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. Stat Med. 2008;27:1133–63.
36. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. Int J Epidemiol. 2015;44:512–25.
37. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. Genet Epidemiol. 2013;37:658–65.
38. Yavorska OO, Burgess S. MendelianRandomization: an R package for performing Mendelian randomization analyses using summarized data. Int J Epidemiol 2017;46:1734–9.
39. Bowden J, Davey, Smith G, Haycock PC, Burgess S. Consistent estimation in mendelian randomization with some invalid instruments using a weighted median estimator. Genet Epidemiol. 2016;40:304–14.
40. Hartwig FP, Davey Smith G, Bowden J. Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. Int J Epidemiol. 2017;46:1985–98.
41. Lupp C, Robertson ML, Wickham ME, Sekirov I, Champion OL, Gaynor EC, et al. Host-mediated inflammation disrupts the intestinal microbiota and promotes the overgrowth of Enterobacteriaceae. Cell Host Microbe. 2007;2:204.
42. Garrett WS, Gallini CA, Yatsunenko T, Michaud M, DuBois A, Delaney ML, et al. Enterobacteriaceae act in concert with the gut microbiota to induce spontaneous and maternally transmitted colitis. Cell Host Microbe. 2010;8:292–300.
43. Cassmann E, White R, Atherly T, Wang C, Sun Y, Khoda S, et al. Alterations of the ileal and colonic mucosal microbiota in canine chronic enteropathies. PloS One. 2016;11:e0147321.
44. Minamoto Y, Otoni CC, Steelman SM, Büyükleblebici O, Steiner JM, Jergens AE, et al. Alteration of the fecal microbiota and serum metabolite profiles in dogs with idiopathic inflammatory bowel disease. Gut Microbes. 2015;6:33–47.
45. Dalla Via A, Gargari G, Taverniti V, Rondini G, Velardi I, Gambaro V, et al. Urinary TMAO levels are associated with the taxonomic composition of the gut microbiota and with the choline TMA-lyase gene (cutC) harbored by enterobacteriaceae. Nutrients. 2019;12:E62.

46. Son M, Ko JI, Kim WB, Kang HK, Kim BK. Taurine can ameliorate inflammatory bowel disease in rats. Adv Exp Med Biol. 1998;442:291–8.
47. Zhao G, He F, Wu C, Li P, Li N, Deng J, et al. Betaine in inflammation: mechanistic aspects and applications. Front Immunol. 2018;9:1070.
48. Willing BP, Dicksved J, Halfvarson J, Andersson AF, Lucio M, Zheng Z, et al. A pyrosequencing study in twins shows that gastrointestinal microbial profiles vary with inflammatory bowel disease phenotypes. Gastroenterology. 2010;139:1844–54.e1.
49. Kolho K-L, Pessia A, Jaakkola T, de Vos WM, Velagapudi V. Faecal and serum metabolomics in paediatric inflammatory bowel disease. J Crohns Colitis. 2017;11:321–34.
50. Rodriguez DM, Benninghoff AD, Aardema NDJ, Phatak S, Hintze KJ. Basal diet determined long-term composition of the gut microbiome and mouse phenotype to a greater extent than fecal microbiome transfer from lean or obese human donors. Nutrients. 2019;11:1630.
51. Kovatcheva-Datchary P, Shoaie S, Lee S, Wahlström A, Nookaew I, Hallen A, et al. Simplified intestinal microbiota to study microbe-diet-host interactions in a mouse model. Cell Rep. 2019;26:3772–83.e6.
52. Jacobs JP, Goudarzi M, Singh N, Tong M, McHardy IH, Ruegger P, et al. A disease-associated microbial and metabolomics state in relatives of pediatric inflammatory bowel disease patients. Cell Mol Gastroenterol Hepatol. 2016;2:750–66.
53. Murdoch TB, Fu H, MacFarlane S, Sydora BC, Fedorak RN, Slupsky CM. Urinary metabolic profiles of inflammatory bowel disease in interleukin-10 gene-deficient mice. Anal Chem. 2008;80:5524–31.
54. Lamark T, Styrvold OB, Strøm AR. Efflux of choline and glycine betaine from osmoregulating cells of Escherichia coli. FEMS Microbiol Lett. 1992;75:149–54.
55. Chhibber-Goel J, Gaur A, Singhal V, Parakh N, Bhargava B, Sharma A. The complex metabolism of trimethylamine in humans: endogenous and exogenous sources. Expert Rev Mol Med. 2016;18:e8.

ACKNOWLEDGEMENTS

The PopGen 2.0 network (P2N) is supported by a grant from the German Federal Ministry for Education and Research (01EY1103). We thank Drs. Andre Franke and Wolfgang Lieb for sharing the GWAS summary data for beta diversity and bacterial abundance from published paper [21]. The study was supported by grants from the Peking University Start-up Grant (BMU2018YJ002), the National Key R&D Program of China (2020YFC2003401) and High-performance Computing Platform of Peking University. The funding organization had no role in the preparation of the manuscript.

AUTHOR CONTRIBUTIONS

ZZ, NL, and TH designed the research. ZZ and TH had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. ZZ, NL, ZL, and TH wrote the paper and performed the data analysis. All authors contributed to the statistical analysis, critically reviewed the manuscript during the writing process, and approved the final version to be published. ZZ and TH are the guarantors for the study.

COMPETING INTERESTS

The authors declare no competing interests.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Contributing studies received ethical approval from their respective institutional review boards. Informed consent was obtained from all participants of contributing studies.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41430-022-01074-w.

Correspondence and requests for materials should be addressed to Tao Huang.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.