The mediating role of metabolites between gut microbiome and Hirschsprung disease: a bidirectional two-step Mendelian randomization study

Zhe Wang[1], Bingjun Gao[1], Xiao Liu[2] and Aiwu Li[1]*

[1]Department of Pediatric Surgery, Qilu Hospital of Shandong University, Jinan, China, [2]Department of Neurosurgery, Qilu Hospital of Shandong University, Jinan, China

Background: Gut microbiome (GM) was observed to be associated with the incidence of Hirschsprung disease (HD). However, the effect and mechanism of GM in HD is still unclear. To investigate the relationship between GM and HD and the effect of metabolites as mediators, a bidirectional two-step Mendelian randomization (MR) study was conducted.

Methods: The study selected instrument variables (IVs) from summary-level genome-wide association studies (GWAS). The MiBioGen consortium provided the GWAS data for GM, while the GWAS data for metabolites and HD were obtained from the GWAS Catalog consortium. Two-sample MR analyses were performed to estimate bidirectional correlations between IVs associated with GM and HD. Then, genetic variants related to 1,400 metabolite traits were selected for further mediation analyses using the Product method.

Results: This study found that seven genus bacteria had a significant causal relationship with the incidence of HD but not vice versa. 27 metabolite traits were significantly correlated with HD. After combining the significant results, three significant GM-metabolites-HD lines have been identified. In the *Peptococcus*-Stearoyl sphingomyelin (d18:1/18:0)-HD line, the Stearoyl sphingomyelin (d18:1/18:0) levels showed a mediation proportion of 14.5%, while in the *Peptococcus*-lysine-HD line, the lysine levels had a mediation proportion of 12.9%. Additionally, in the *Roseburia*-X-21733-HD line, the X-21733 levels played a mediation proportion of 23.5%.

Conclusion: Our MR study indicates a protective effect of *Peptococcus* on HD risk that is partially mediated through serum levels of stearoyl sphingomyelin (d18:1/18:0) and lysine, and a risk effect of *Roseburia* on HD that is partially mediated by X-21733 levels. These findings could serve as novel biomarkers and therapeutic targets for HD.

KEYWORDS

Mendelian randomization, gut microbiome, Hirschsprung disease, metabolites, two-step, mediation proportion

1 Introduction

Hirschsprung disease (HD) is a congenital disorder of enteric nervous system characterized by the absence of enteric ganglia in the distal part of the colon, leading to abnormal contractions (1). Until now, the only efficient treatment for HD is surgery to remove the aganglionic segment of the colon and reconstruct the healthy, innervated intestine to the anus (2). However,

approximately one-third of HD patients still experience complications after surgery, including fecal soiling, obstructive symptoms, and Hirschsprung-associated enterocolitis (HAEC) (3). The cause of these complications is not well understood. Various studies have suggested possible mechanisms such as genetic mutation (4), dysfunction of the immune system in the gut (5), and an imbalance of the gut microbiome (6).

Gut microbiome (GM) is a distinct microbial ecosystem within the digestive system that plays a critical role in regulating human health and disease (7). The imbalance and dysfunction of microbial composition in GM, resulting from a metabolic disorder, can lead to several diseases (8). GM is deemed an extra "organ" with more genes than the host and is influenced by genetics and the environment (9). Several studies have shown that GM is associated with a range of illnesses, such as diabetes (10), cardiovascular disorders (11), neurological diseases (12), and immune diseases (13). Genome-wide association studies (GWAS) have challenged the view that the gut microbiota is solely an environmental factor (14). However, the extent of its genetic influence remains a topic of debate.

In order to investigate the potential mediation effect that could bridge the GM to the incidence of HD, metabolites are considered. Metabolites refer to tiny molecules produced either as a result or at the end of various metabolic reactions (15). Multiple factors, such as genetics, lifestyle, disease, and GM, can impact metabolite levels in the body (16, 17). The changes in metabolites can impact the incidence of disease. The high heritability of metabolites makes them convenient to study and potential therapeutic targets for disease (18). Based on these findings, we hypothesize that metabolites mediate the effect of GM on HD. However, it is difficult to establish the causal relationship and avoid confounding bias in an observational study. Therefore, an innovative methodology should be adopted to investigate the causal relationship.

Mendelian randomization (MR) is an effective method that can avoid the confounding bias and evaluate the effect of exposures on outcomes (19). MR analysis assumes that during conception, random assignment of alleles to offspring follows Mendel's law of inheritance, mirroring the concept of a randomized controlled trial (20). Using single nucleotide polymorphisms (SNPs) associated with specific IVs as proxies for exposures, MR analysis can assess the causal effects between exposures and outcomes (21). When the effect between exposures and outcomes is confusing, a mediation MR analysis can help to understand the cause better and identify intermediate variables that could be potential targets for intervention (22).

Our study aimed to investigate the relationship between GM and HD and explore the potential role of metabolites as mediators by conducting a bidirectional two-step MR analysis.

2 Materials and methods

2.1 Study design

We used summary-level GWAS data to explore the genetic relationship between GM and HD in an MR framework. The effect of an exposure on an outcome consists of both direct and indirect effect (22). Our research involved a two-step MR analysis to investigate the causal relationship between GM and metabolites, followed by an assessment of the causal impact of metabolites on HD. To calculate the mediation effects of metabolites, we used the Product method, as previously described in other literature (23). Three fundamental assumptions must be met when selecting IVs, similar to a conventional MR study (24). Firstly, SNPs should show a strong correlation with exposures and meet the criterion for genome-wide significance threshold. Secondly, the selected IVs should be independent of any confounding factors affecting the study results. Lastly, the IVs should only affect the outcome through exposures and ensure that other lines or routes are not responsible for the effect on the outcome. The design of our bidirectional two-step MR study was illustrated in Figure 1.

2.2 Data sources

From the largest GWAS database on GM, the MiBioGen consortium (https://mibiogen.gcc.rug.nl), we obtained the summary-level data for GM, including 16S rRNA fecal microbiome data from 18,340 individuals (25). In this cohort, genetic loci impacting the covariate-adjusted abundance of bacterial taxa were adjusted for age, sex, technical covariates, and genetic principal components (25). Using these genetic markers, we can estimate the causal relationship between GM and complex traits through MR approaches (26).

Summary-level GWAS data for 1,400 metabolite traits (1,091 metabolites and 309 metabolite ratios) were extracted from the Canadian Longitudinal Study on Aging (CLSA) cohort of 8,299 individuals (15). The study found associations between 690 metabolites at 248 sites and 143 metabolite ratios at 69 sites. By integrating metabolite-gene and gene expression data, the study identified 94 effector genes for 109 metabolites and 48 metabolite ratios (15).

The GWAS Catalog consortium (https://www.ebi.ac.uk/gwas) provided GWAS data for HD, including 170 HD patients and 4,717 controls of European ancestry as the exposures (27). The diagnosis of HD was made using the Tenth Revision of the International Classification of Diseases (ICD-10) codes (28). In this cohort, the SNP heritability of HD has been estimated to be approximately 88% based on GWAS studies (27). Moreover, they have used Lasso regression to develop a potential genetic predictor for HD (27).

We conducted our study using publicly available GWAS summary data that has been approved by relevant ethics and institutional review boards. Therefore, we did not require ethical approval for our study. Details of the GWASs data are shown in Table 1.

2.3 Selection of instrument variables

Suitable SNPs for IVs are genetic variants that serve as proxies for exposures to investigate causal relationships at the genetic level in MR
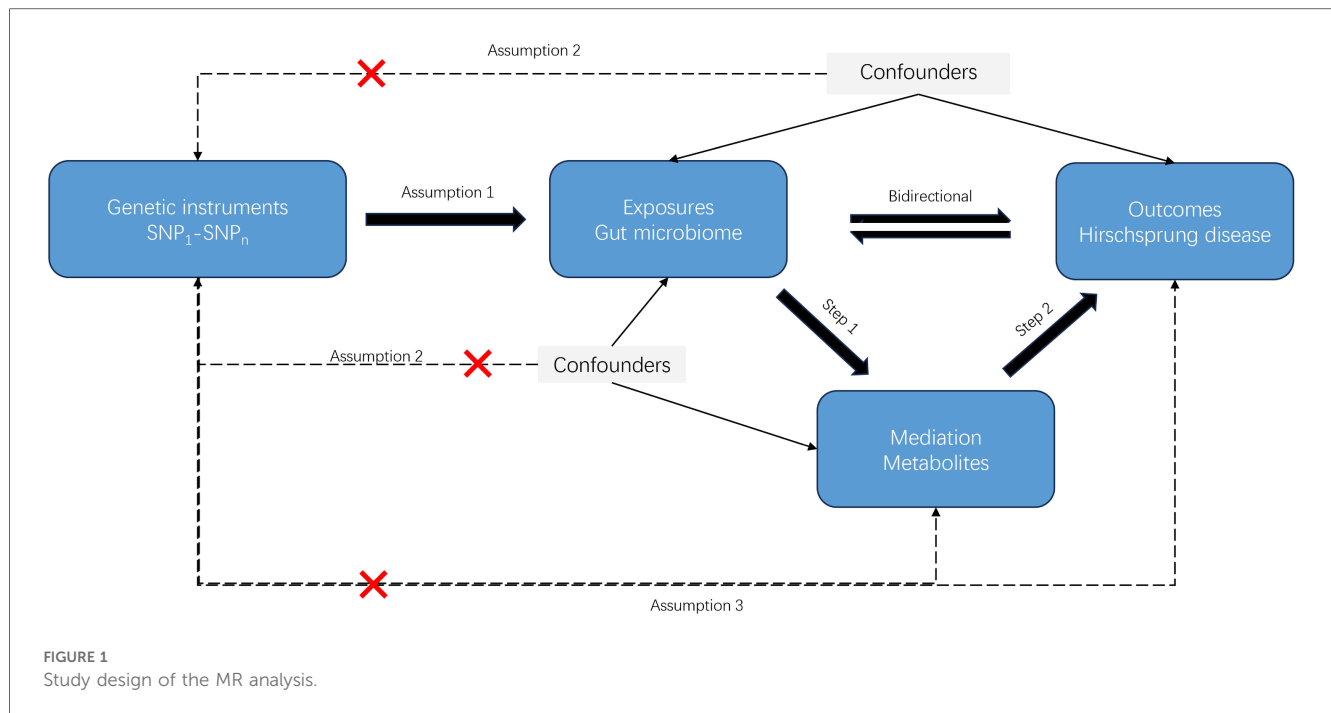
FIGURE 1
Study design of the MR analysis.

studies (29). The selection of IVs was based on several principles: First, since only a small number of SNPs met the statistical significance threshold of $p < 5e-08$ for GM, metabolites, and HD, different threshold values were set for each measure (30). For GM and metabolites, the genome-wide significance threshold was set at $p < 1e-05$, while for HD, it was set at $p < 5e-06$. Second, SNPs with an r^2 greater than 0.001 within a 10,000 kb distance were excluded due to linkage disequilibrium (LD) (31). Third, through the PhenoScanner website (http://www.phenoscanner.medschl.cam.ac.uk/) and as well as the IEU database (https://gwas.mrcieu.ac.uk/), we removed SNPs associated with inflammatory bowel disease, malignant neoplasm of colon, and irritable bowel syndrome in order to minimize the impact of confounding factors. Fourth, any SNPs that had a palindromic structure were excluded automatically from the analysis. Lastly, we calculated the F-statistics to assess the genetic liability of their genetic instruments using the following formula (32):

$$F = (beta/se)^2 \tag{1}$$

(se: standard error). From traditional experience, when the F statistic is less than 10, we usually consider the instrumental variable used as a weak instrumental variable, which may produce a certain bias to the result (33). To avoid weak instrument bias, only SNPs with F statistics greater than ten were

chosen for the study. SNP characteristics are listed in Supplementary Data Sheet 2.

2.4 Statistical analysis

To explore the causal relationship between GM and HD, a bidirectional two-sample MR analysis was conducted. Then, a two-step MR analysis was conducted to evaluate the potential mediating effect via metabolites in this causal relationship. The inverse-variance weighted (IVW) method was utilized as the primary analysis, considering it the most reliable when there were no signs of directional pleiotropy (34). In the absence of horizontal pleiotropy, the IVW test was used as the primary method for calculating the causal effect values to obtain unbiased estimates. Additionally, weighted mode (34), MR-Egger (35), weighted median (36), and simple mode (37) were used as complementary approaches to assess the consistency of the results.

The Cochran's Q method was utilized with a p-value lower than 0.05 indicating heterogeneity (38). Cochran's Q is a kind of heterogeneity statistic for the IVW model. If the Q statistic much larger than its degrees of freedom, this provides evidence for heterogeneity and invalid IVs. The MR-Egger regression analysis was used for identification of potential horizontal pleiotropy (39), while the MR pleiotropy residual sum and outlier (MR-PRESSO) analysis (40) was conducted to minimize possible confounding factors. To investigate the effect of a single SNP on causal associations, we conducted leave-one-out sensitivity tests by removing each SNP one at a time (41). Additionally, scatter plots and funnel plots were generated to assess the robustness of the MR results.

The mediation proportion of potential mediators in the total effect of genetically determined GM on HD risk was calculated

TABLE 1 Details of GWAS data in this MR study.

Trait	Consortium	Ethnicity	Sample size
GM	MiBioGen	European	18,340
Metabolites	CLSA	European	8,299
HD	GWAS Catalog	European	4,887

GWAS, genome-wide association study; GM, gut microbiome; HD, Hirschsprung disease; CLSA, Canadian longitudinal study on aging cohort.

using the Product method (22). The mediation proportions were determined using the following formula (42):

$$P(\%) \ = \ (\beta_1 \times \beta_2)/\beta_0 \qquad (2)$$

P(%): the proportions of mediation; β_0 stands for the total effect obtained from the primary analysis, β_1 stands for the effect of GM on mediators, and β_2 stands for the effect of mediators on HD.

The statistical analysis was conducted using R packages "TwoSampleMR" and "MRPRESSO" in R software (version 3.4.2, the R Foundation for Statistical Computing, Vienna, Austria). The findings were presented in a combined format of odds ratio (OR) and a 95% confidence interval (CI).

3 Results

3.1 Selection of instrument variables

After performing the clumping process for LD, 1,425 SNPs were identified as IVs associated with GM traits ($p < 1e\text{-}05$). We only considered the genus level data, which contains the most relevant information about the GM and overlaps with higher classification levels (43). For metabolites, we analyzed 33,571 SNPs, with each metabolite corresponding to an average of 24 SNPs. We selected 18 SNPs as IVs for HD ($p < 5e\text{-}06$). After harmonizing with the metabolites data, only nine SNPs were left. The average F-statistics for GM, metabolites, and HD were 21.73, 29.88, and 38.34, respectively. All F-statistics exceeded 10, indicating avoidance of potential instrumental bias. Finally, we removed the palindromic SNPs simultaneously. The details of selected SNPs are shown in Supplementary Data Sheet 2.

3.2 Bidirectional two-sample MR analyses of the total effect between GM and HD

Seven genus bacteria were found to have a causal effect on an increased incidence of HD. Our forward MR analyses revealed that IVs associated with an elevated *Peptococcus* (OR = 0.366, 95% CI = 0.185–0.727, $p = 0.002$), *RuminococcaceaeNK4A214group* (OR = 0.159, 95% CI = 0.038–0.658, $p = 0.011$), *Ruminococcus2* (OR = 0.32, 95% CI = 0.112–0.912, $p = 0.033$), *ErysipelotrichaceaeUCG003* (OR = 0.37, 95% CI = 0.143–0.957, $p = 0.039$), and *Paraprevotella* (OR = 0.449, 95% CI = 0.206–0.977, $p = 0.043$) were responsible for decreased susceptibility to HD. The participants of *Eggerthella* (OR = 2.66, 95% CI = 1.234–5.737, $p = 0.013$) and *Roseburia* (OR = 5.387, 95% CI = 1.076–26.96, $p = 0.04$) brought a higher incidence of HD (Figure 2). Supplementary Table S1 provides a clear illustration of the causal effects of GM on HD.

On the other hand, the reverse MR analysis showed that there was no impact of genetic predisposition to HD on pre-identified GM traits. The results of IVW estimates are illustrated in Figure 3 and Supplementary Table S2.

3.3 Two-step MR analyses

Our study utilized a two-step MR analysis approach. In the first step, we conducted a two-sample MR analysis to investigate the causal relationships between metabolites and the incidence of HD. This analysis revealed a significant link between 27 metabolite traits and HD (Supplementary Table S3). In the second step, we used the univariable MR method to investigate the relationship between the seven types of GM and 27 metabolite traits. Our findings identified three significant lines of evidence, including two types of bacteria and three metabolites. After combining these significant results, we obtained three significant GM-metabolites-HD lines (Figure 4).

3.3.1 Causal effects of GM on metabolites

According to our MR analysis, *Peptococcus* had a positive correlation with stearoyl sphingomyelin (d18:1/18:0) levels (OR = 1.137, 95% CI = 1.051–1.274, $p = 0.026$) as well as lysine levels (OR = 1.134, 95% CI = 1.009–1.275, $p = 0.035$). We found a negative correlation between *Roseburia* and X-21733 levels (OR = 0.803, 95% CI = 0.675–0.955, $p = 0.013$). Figure 5A displays the causal effects of GM traits on metabolites using the IVW method. Details are shown in Supplementary Table S3.

3.3.2 Causal effects of metabolites on HD

We observed that 27 metabolite traits were significantly correlated with incidence of HD (Supplementary Table S4). Three metabolite traits were identified to be the significant mediators, including stearoyl sphingomyelin (d18:1/18:0) levels (OR = 0.28, 95% CI = 0.088–0.887, $p = 0.03$), lysine levels (OR = 0.314, 95% CI = 0.117–0.844, $p = 0.022$), and X-21733 levels (OR = 0.165, 95% CI = 0.047–0.578, $p = 0.005$). The significant effect of metabolites on HD is illustrated in Figure 5B.

3.3.3 Mediating effects of metabolites on GM-HD effect

In this two-step MR analysis, we identified the metabolites that could potentially mediate the causal effect of GM on HD. In the first step, we found that stearoyl sphingomyelin (d18:1/18:0) levels, lysine levels, and X-21733 levels could be affected by GM. In the second step, we observed that these same metabolites had a significant role in mediating the effect of GM on HD. We used the Product method to calculate the mediation proportion. In the *Peptococcus*-Stearoyl sphingomyelin (d18:1/18:0)-HD line, the Stearoyl sphingomyelin (d18:1/18:0) levels showed a mediation proportion of 14.5%, while in the *Peptococcus*-lysine-HD line, the lysine levels had a mediation proportion of 12.9%. In the *Roseburia*-X-21733-HD line, the X-21733 levels played a mediation proportion of 23.5%.

3.4 Sensitivity analyses

We conducted Cochran's Q analysis and found no significant heterogeneity in the estimates. We also detected the p-value associated with the MR-Egger intercept, and found no potential
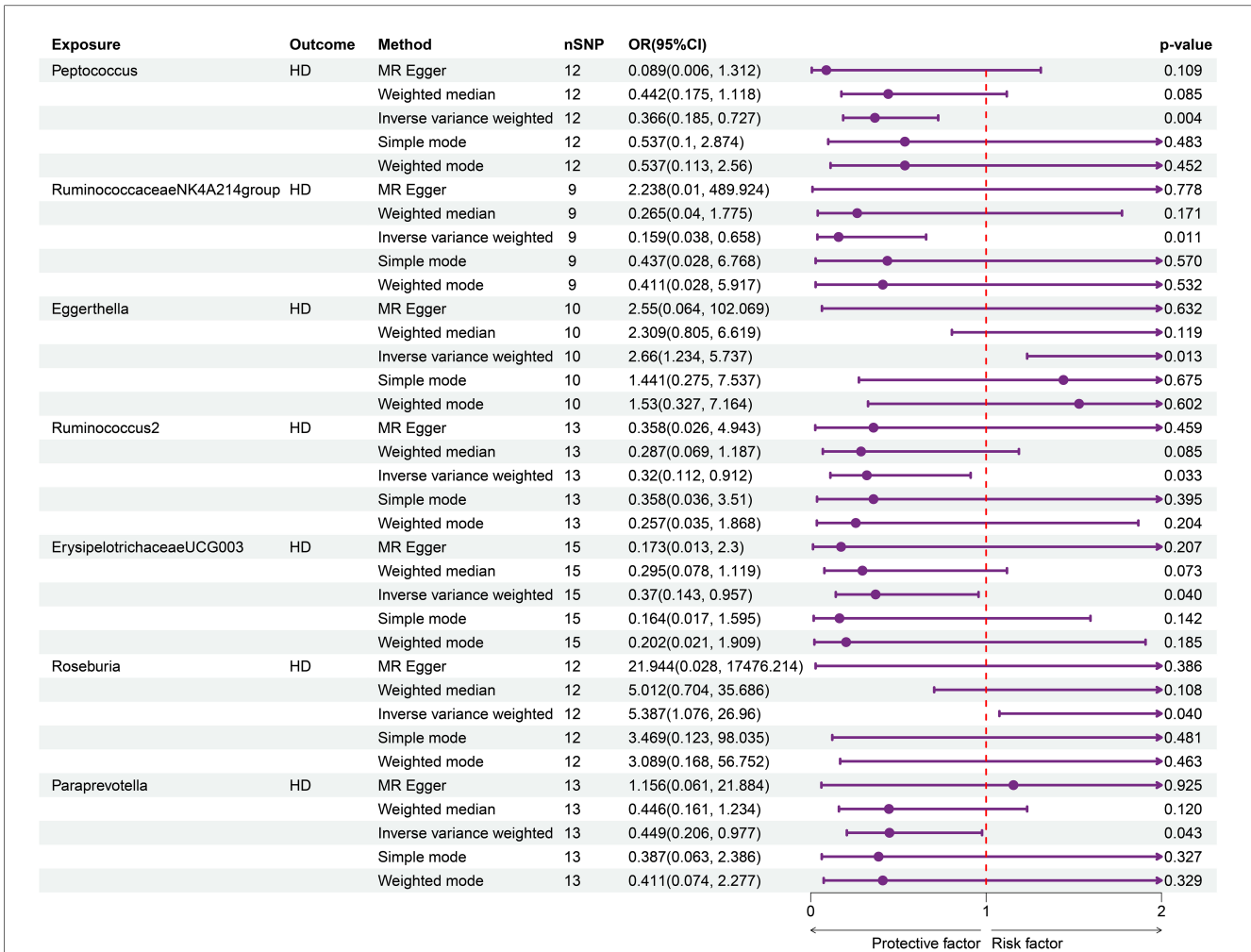
Exposure	Outcome	Method	nSNP	OR(95%CI)	p-value
Peptococcus	HD	MR Egger	12	0.089(0.006, 1.312)	0.109
		Weighted median	12	0.442(0.175, 1.118)	0.085
		Inverse variance weighted	12	0.366(0.185, 0.727)	0.004
		Simple mode	12	0.537(0.1, 2.874)	0.483
		Weighted mode	12	0.537(0.113, 2.56)	0.452
RuminococcaceaeNK4A214group	HD	MR Egger	9	2.238(0.01, 489.924)	0.778
		Weighted median	9	0.265(0.04, 1.775)	0.171
		Inverse variance weighted	9	0.159(0.038, 0.658)	0.011
		Simple mode	9	0.437(0.028, 6.768)	0.570
		Weighted mode	9	0.411(0.028, 5.917)	0.532
Eggerthella	HD	MR Egger	10	2.55(0.064, 102.069)	0.632
		Weighted median	10	2.309(0.805, 6.619)	0.119
		Inverse variance weighted	10	2.66(1.234, 5.737)	0.013
		Simple mode	10	1.441(0.275, 7.537)	0.675
		Weighted mode	10	1.53(0.327, 7.164)	0.602
Ruminococcus2	HD	MR Egger	13	0.358(0.026, 4.943)	0.459
		Weighted median	13	0.287(0.069, 1.187)	0.085
		Inverse variance weighted	13	0.32(0.112, 0.912)	0.033
		Simple mode	13	0.358(0.036, 3.51)	0.395
		Weighted mode	13	0.257(0.035, 1.868)	0.204
ErysipelotrichaceaeUCG003	HD	MR Egger	15	0.173(0.013, 2.3)	0.207
		Weighted median	15	0.295(0.078, 1.119)	0.073
		Inverse variance weighted	15	0.37(0.143, 0.957)	0.040
		Simple mode	15	0.164(0.017, 1.595)	0.142
		Weighted mode	15	0.202(0.021, 1.909)	0.185
Roseburia	HD	MR Egger	12	21.944(0.028, 17476.214)	0.386
		Weighted median	12	5.012(0.704, 35.686)	0.108
		Inverse variance weighted	12	5.387(1.076, 26.96)	0.040
		Simple mode	12	3.469(0.123, 98.035)	0.481
		Weighted mode	12	3.089(0.168, 56.752)	0.463
Paraprevotella	HD	MR Egger	13	1.156(0.061, 21.884)	0.925
		Weighted median	13	0.446(0.161, 1.234)	0.120
		Inverse variance weighted	13	0.449(0.206, 0.977)	0.043
		Simple mode	13	0.387(0.063, 2.386)	0.327
		Weighted mode	13	0.411(0.074, 2.277)	0.329

0   1   2

← Protective factor   Risk factor →

FIGURE 2
Forward MR analysis of GM on HD. Elevated Peptococcus (OR = 0.366, 95% CI = 0.185−0.727, p = 0.002), RuminococcaceaeNK4A214group (OR = 0.159, 95% CI = 0.038−0.658, p = 0.011), Ruminococcus2 (OR = 0.32, 95% CI = 0.112−0.912, p = 0.033), ErysipelotrichaceaeUCG003 (OR = 0.37, 95% CI = 0.143−0.957, p = 0.039), and Paraprevotella (OR = 0.449, 95% CI = 0.206−0.977, p = 0.043) were linked with decreased susceptibility to HD.

Exposure	Outcome	nSNP	OR(95%CI)	p.value
HD	Ruminococcus2	9	1.003(0.993,1.013)	0.586
HD	ErysipelotrichaceaeUCG003	9	0.993(0.981,1.005)	0.229
HD	Roseburia	9	0.999(0.989,1.008)	0.766
HD	Paraprevotella	9	1.004(0.988,1.021)	0.606
HD	Peptococcus	9	1.005(0.989,1.023)	0.525
HD	RuminococcaceaeNK4A214group	9	1.008(0.996,1.02)	0.177
HD	Eggerthella	9	1.014(0.996,1.032)	0.124

0.8   1   1.2
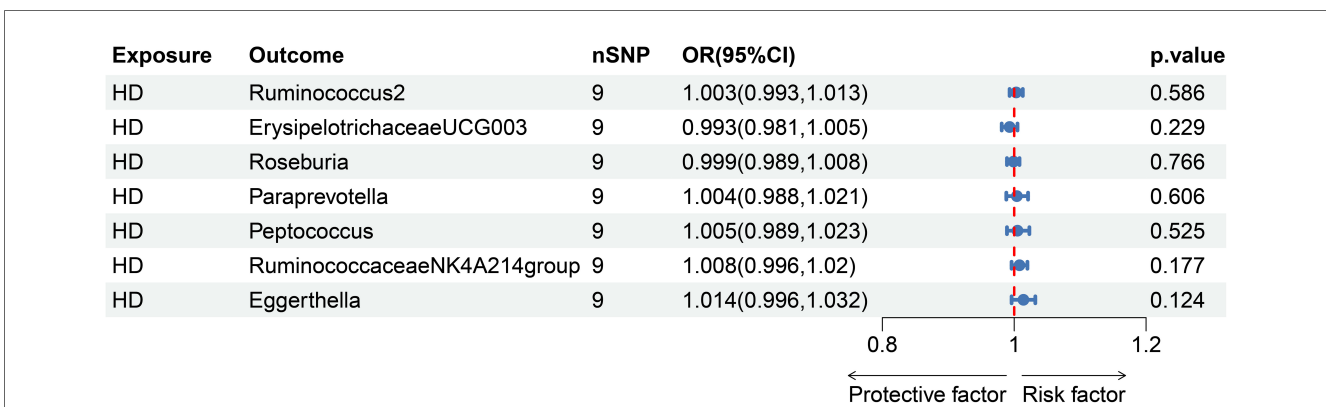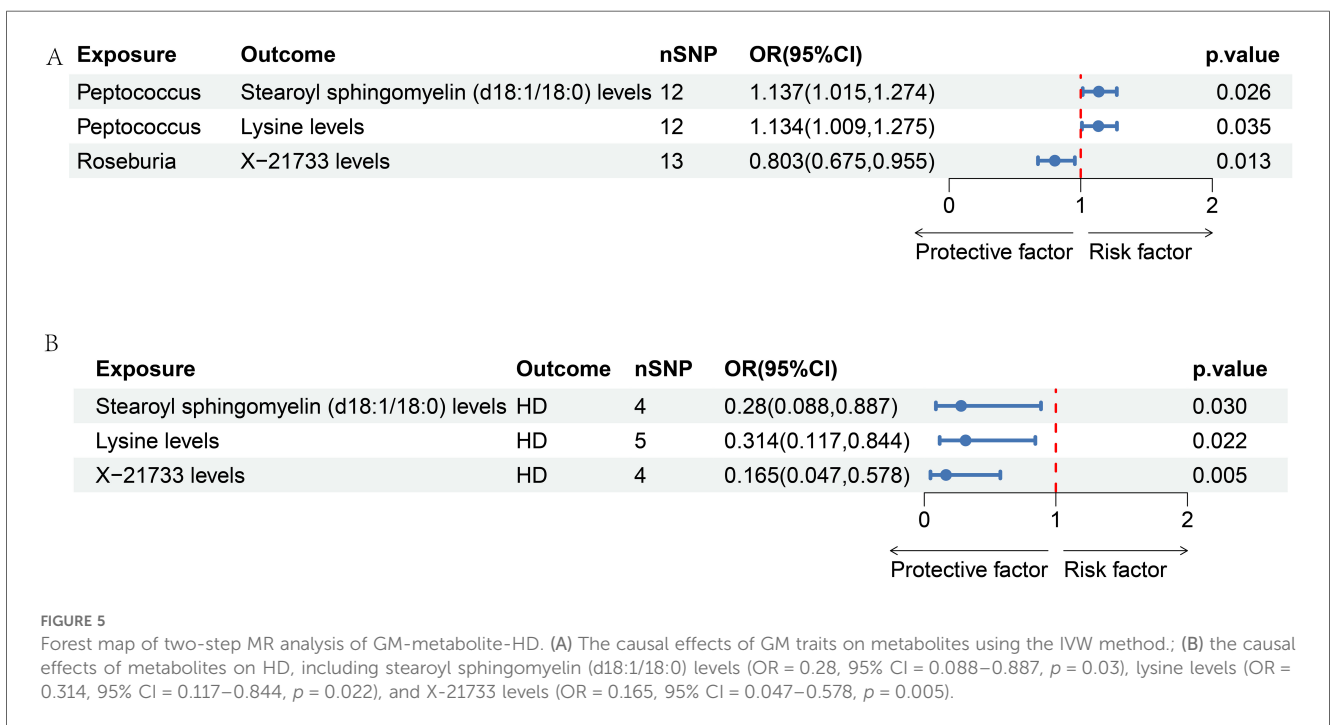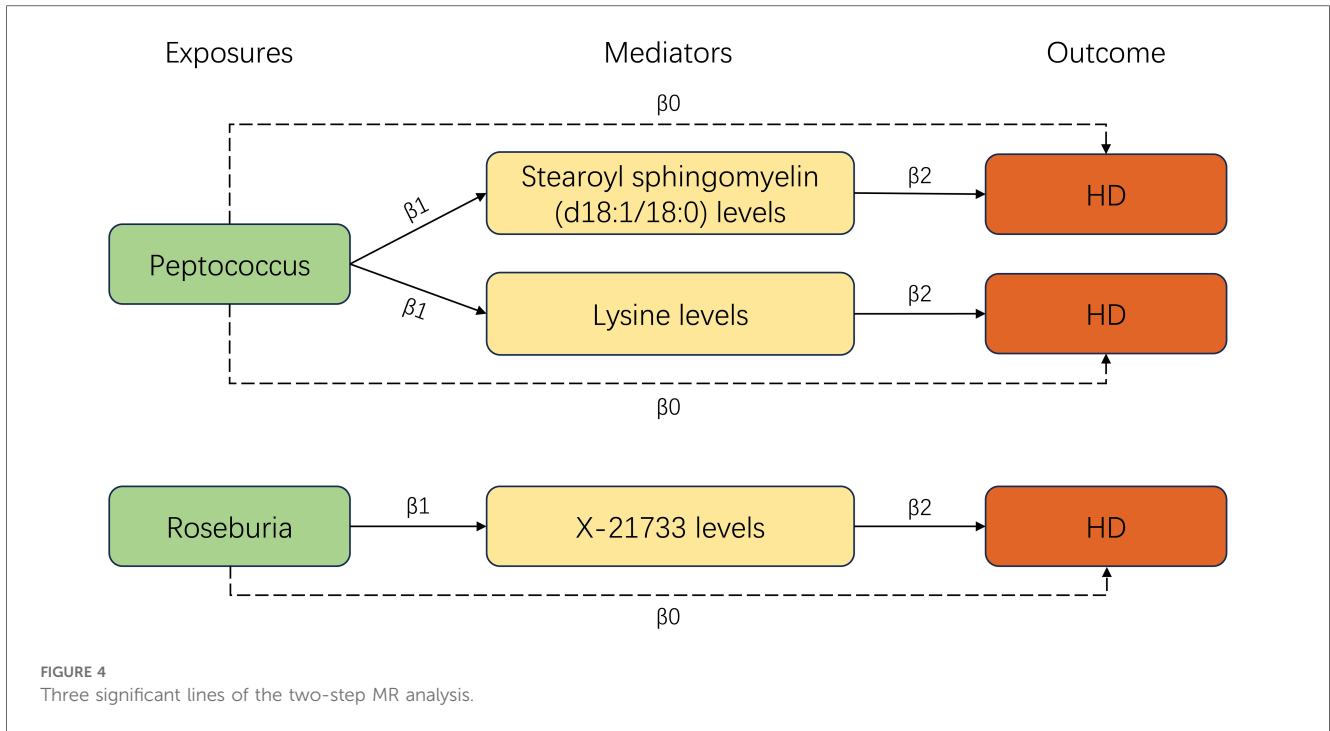
← Protective factor   Risk factor →

FIGURE 3
Reverse MR analysis of HD on GM.

pleiotropic effects. Moreover, we utilized multiple estimating methods to evaluate the causal effects, and the scatter plots involving these estimating methods were shown in Supplementary Figure S1. The consistency results of the sensitivity analyses reinforced the causal reasoning of the primary analyses. Additionally, forest plots, funnel plots, and leave-one-

FIGURE 4
Three significant lines of the two-step MR analysis.

FIGURE 5
Forest map of two-step MR analysis of GM-metabolite-HD. (A) The causal effects of GM traits on metabolites using the IVW method.; (B) the causal effects of metabolites on HD, including stearoyl sphingomyelin (d18:1/18:0) levels (OR = 0.28, 95% CI = 0.088–0.887, p = 0.03), lysine levels (OR = 0.314, 95% CI = 0.117–0.844, p = 0.022), and X-21733 levels (OR = 0.165, 95% CI = 0.047–0.578, p = 0.005).

out plots of each IV showed that the results of MR analysis were consistent and robust (Supplementary Figures S2–4).

4 Discussion

This study explored the causal relationship between GM and HD, and assessed the mediating effect of serum metabolite traits between GM and HD in a comprehensive bidirectional two-step MR analysis.

Our study revealed that genetically variants in *Ruminococcus2*, *ErysipelotrichaceaeUCG003*, *Roseburia*, *Paraprevotella*, *Peptococcus*, *RuminococcaceaeNK4A214group*, and *Eggerthella* were significantly associated with the incidence of HD. The results did not indicate bidirectional causal relationship. Our two-step MR analyses revealed that *Roseburia* contributed to an increased risk of HD, which was partially mediated by lower levels of X-21733. As the original paper stated that X-21733 is an unknown molecule in serum (15), we have excluded it from further discussion. On the other hand, *Peptococcus*

was found to be associated with a reduced risk of developing HD, partially mediated by levels of stearoyl sphingomyelin (d18:1/18:0) and lysine.

Studies conducted on animals and humans have confirmed a connection between the diversity of GM and HD. The human gastrointestinal tract commonly acts as a natural ecosystem and supplier of nutrients for microbiota. In return, microbiota aids in gut development, strengthen the immunity, and improve the defense mechanisms by generating different metabolites (44). Cheng et al. found decreased fecal microbiota alpha diversity in HD mice after microsurgical pull-through surgery on Ednrb knock-out mice (45). A study on children with HD showed reduced gut microbiota richness after surgery compared to control group (46). HD may require different lengths of surgical therapy depending on the degree of aganglionic involvement, which in turn may affect gut microbiota homeostasis (47). However, studies on the mechanism of GM underlying the development of HD were limited. Through our MR analyses, we found lines linking GM and HD by the mediation of metabolites.

The genus *Roseburia* belongs to the phylum *Firmicutes*, class *Clostridia*, order *Clostridiales*, and family *Lachnospiraceae* (48). There are five well-characterized *Roseburia* species: *Roseburia intestinalis*, *Roseburia hominis*, *Roseburia inulinivorans*, *Roseburia faecis*, and *Roseburia cecicola* (49). *Roseburia intestinalis* is known to protect against several inflammatory diseases by minimizing intestinal inflammatory reaction (48). The expression of colonic mucosal melatonin is positively related to *Roseburia hominis* (50). *Roseburia inulinivorans*, a newly discovered motile member of the *Firmicutes*, helps form butyrate from different dietary polysaccharide substrates present in the human large intestine (51). *Roseburia faecis,* a Gram-positive anaerobic bacterium that produces butyrate, has been tested for its usefulness in treating irritable bowel syndrome induced by repeated water avoidance stress in rat models (52). To date, no research has been carried out to explore the presence of the genus *Roseburia* in individuals with HD. Our MR analysis found that *Roseburia* may be involved in HD as a risk factor due to the mediation effect of serum X-21733. However, the information regarding serum X-21733 is unknown, and further molecular experiments are required to elucidate this mystery.

There has been debate regarding the precise role of *Peptococcus* in the human gut. Bourgault et al. conducted a study revealing *Peptococcus* as the most frequently occurring species of anaerobic Gram-positive cocci in significant infections (53). Moreover, it appears to be particularly pathogenic in infections of bones and joints or when present with foreign bodies (53). However, Gu et al. conducted research that demonstrated the protective effect of *Peptococcus* on inflammatory disorders of the breast (54). According to their study, there is a strong positive correlation between *Peptococcus* and the presence of valeric acid and butyrate, which regulate the inflammatory response and participate in the energy supply of tight junction proteins (54). Our MR analyses supported the protective effect of *Peptococcus* on HD, which may be mediated by the levels of stearoyl sphingomyelin (d18:1/18:0) and lysine. In the future, success may come from therapeutics targeting *Roseburia* or *Peptococcus*, including probiotics, fecal microbiota transplantation, prebiotics, and synbiotics.

Other bacteria were not associated with the development of HD. *Ruminococcus gnavus* is a Gram-positive strict anaerobe bacterium that forms chains. The bacteremia has been associated with an acute flare of ulcerative colitis (55). *Erysipelotrichaceae* has effect on the metabolism of lipid, warranting additional investigation into the metabolic profiles of these organisms. *Erysipelotrichaceae* may be correlated to inflammation and auto-immunogenic procedure (56).

Metabolomics is an advanced technology with great promise for uncovering the underlying mechanisms linked to various disease processes (57). Yang et al. identified 21 HD biomarkers in serum samples using metabolomic analysis (58). Sphingomyelin (d18:1/18:0) is a sphingolipid commonly found in cell membranes, particularly in the membranous myelin sheath surrounding nerve cell axons (59). Plekhova et al. discovered that lysine catabolism was increased in HD patients. This has been associated with inadequate bacterial butyrate production in the large bowel (60). Our study has found that serum levels of sphingomyelin (d18:1/18:0), lysine, and X-21733 were involved in mediating the GM-HD effect. However, the structure and function of X-21733 is still unknown. Our MR analyses have confirmed that metabolites are involved in the mechanism of HD development. To fully understand the underlying pathologies, further experiments are necessary.

This study has several advantages. Firstly, we used summary data from the GWAS consortium in our MR analyses, which increases statistical power since genetic variation is unaffected by confounding factors. Secondly, we combined bidirectional and two-step MR analyses to explore the potential mediation effects of metabolites, producing more precise and reliable results. This innovative approach provides a comprehensive understanding of the causal relationship between GM and HD.

There are a few limitations to consider in this study. Firstly, even if all three assumptions are fulfilled, it is impossible to avoid weak instrumental bias. However, GWAS data with larger sample sizes can help reduce this bias. Secondly, since the GWAS participants were only of European descent, it may not be appropriate to generalize the findings. Therefore, similar studies need to be conducted across multi-ethnic groups. Thirdly, MR analysis typically reveals exposure over a lifetime, and the presence of canalization may lead to an overestimation of effect size. Therefore, it is recommended that further randomized controlled trials be conducted to examine the effect.

5 Conclusions

In conclusion, our study indicated a causal relationship between *Peptococcus* and the incidence of HD that is partially mediated by serum levels of stearoyl sphingomyelin (d18:1/18:0) and lysine. Levels of *Peptococcus* in feces and stearoyl sphingomyelin (d18:1/18:0) and lysine levels in serum could serve as novel biomarkers and therapeutic targets for HD.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

Author contributions

Funding

Acknowledgments

Conflict of interest

The 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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fped.2024.1371933/full#supplementary-material

References

1. Ambartsumyan L, Smith C, Kapur RP. Diagnosis of hirschsprung disease. *Pediatr Dev Pathol.* (2020) 23:8–22. doi: 10.1177/1093526619892351

2. Heuckeroth RO. Hirschsprung disease—integrating basic science and clinical medicine to improve outcomes. *Nat Rev Gastroenterol Hepatol.* (2018) 15:152–67. doi: 10.1038/nrgastro.2017.149

3. Gosain A. Established and emerging concepts in hirschsprung's-associated enterocolitis. *Pediatr Surg Int.* (2016) 32:313–20. doi: 10.1007/s00383-016-3862-9

4. Hugot JP, Chamaillard M, Zouali H, Lesage S, Cézard JP, Belaiche J, et al. Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. *Nature.* (2001) 411:599–603. doi: 10.1038/35079107

5. Cheng Z, Dhall D, Zhao L, Wang HL, Doherty TM, Bresee C, et al. Murine model of Hirschsprung-associated enterocolitis. I: phenotypic characterization with development of a histopathologic grading system. *J Pediatr Surg.* (2010) 45:475–82. doi: 10.1016/j.jpedsurg.2009.06.009

6. Ward NL, Pieretti A, Dowd SE, Cox SB, Goldstein AM. Intestinal aganglionosis is associated with early and sustained disruption of the colonic microbiome. *Neurogastroenterol Motil.* (2012) 24:874–e400. doi: 10.1111/j.1365-2982.2012.01937.x

7. Heintz-Buschart A, Wilmes P. Human gut microbiome: function matters. *Trends Microbiol.* (2018) 26:563–74. doi: 10.1016/j.tim.2017.11.002

8. Fan Y, Pedersen O. Gut microbiota in human metabolic health and disease. *Nat Rev Microbiol.* (2021) 19:55–71. doi: 10.1038/s41579-020-0433-9

9. Marchesi JR, Adams DH, Fava F, Hermes GDA, Hirschfield GM, Hold G, et al. The gut microbiota and host health: a new clinical frontier. *Gut.* (2016) 65:330–9. doi: 10.1136/gutjnl-2015-309990

10. Lau WL, Tran T, Rhee CM, Kalantar-Zadeh K, Vaziri ND. Diabetes and the gut microbiome. *Semin Nephrol.* (2021) 41:104–13. doi: 10.1016/j.semnephrol.2021.03.005

11. Jie Z, Xia H, Zhong S-L, Feng Q, Li S, Liang S, et al. The gut microbiome in atherosclerotic cardiovascular disease. *Nat Commun.* (2017) 8:845. doi: 10.1038/s41467-017-00900-1

12. Cryan JF, O'Riordan KJ, Sandhu K, Peterson V, Dinan TG. The gut microbiome in neurological disorders. *Lancet Neurol.* (2020) 19:179–94. doi: 10.1016/S1474-4422(19)30356-4

13. Wiertsema SP, van Bergenhenegouwen J, Garssen J, Knippels LMJ. The interplay between the gut microbiome and the immune system in the context of infectious diseases throughout life and the role of nutrition in optimizing treatment strategies. *Nutrients.* (2021) 13:886. doi: 10.3390/nu13030886

14. Liu X, Tang S, Zhong H, Tong X, Jie Z, Ding Q, et al. A genome-wide association study for gut metagenome in Chinese adults illuminates complex diseases. *Cell Discov.* (2021) 7:9. doi: 10.1038/s41421-020-00239-w

15. Chen Y, Lu T, Pettersson-Kymmer U, Stewart ID, Butler-Laporte G, Nakanishi T, et al. Genomic atlas of the plasma metabolome prioritizes metabolites implicated in human diseases. *Nat Genet.* (2023) 55:44–53. doi: 10.1038/s41588-022-01270-1

16. Bar N, Korem T, Weissbrod O, Zeevi D, Rothschild D, Leviatan S, et al. A reference map of potential determinants for the human serum metabolome. *Nature.* (2020) 588:135–40. doi: 10.1038/s41586-020-2896-2

17. Lee W-J, Hase K. Gut microbiota-generated metabolites in animal health and disease. *Nat Chem Biol.* (2014) 10:416–24. doi: 10.1038/nchembio.1535

18. Long T, Hicks M, Yu H-C, Biggs WH, Kirkness EF, Menni C, et al. Whole-genome sequencing identifies common-to-rare variants associated with human blood metabolites. *Nat Genet.* (2017) 49:568–78. doi: 10.1038/ng.3809

19. Smith G D, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum Mol Genet.* (2014) 23:R89–98. doi: 10.1093/hmg/ddu328

20. Smith GD, Ebrahim S. "Mendelian randomization": can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol.* (2003) 32:1–22. doi: 10.1093/ije/dyg070

21. Kilaru V, Iyer SV, Almli LM, Stevens JS, Lori A, Jovanovic T, et al. Genome-wide gene-based analysis suggests an association between neuroligin 1 (NLGN1) and post-traumatic stress disorder. *Transl Psychiatry.* (2016) 6:e820. doi: 10.1038/tp.2016.69

22. Carter AR, Sanderson E, Hammerton G, Richmond RC, Davey Smith G, Heron J, et al. Mendelian randomisation for mediation analysis: current methods and challenges for implementation. *Eur J Epidemiol.* (2021) 36:465–78. doi: 10.1007/s10654-021-00757-1

23. Pasanen A, Karjalainen MK, FinnGen Zhang G, Tiensuu H, Haapalainen AM, et al. Meta-analysis of genome-wide association studies of gestational duration and spontaneous preterm birth identifies new maternal risk loci. *PLoS Genet.* (2023) 19:e1010982. doi: 10.1371/journal.pgen.1010982

24. Davies NM, Holmes MV, Davey Smith G. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. *Br Med J.* (2018) 362:k601. doi: 10.1136/bmj.k601

25. Kurilshikov A, Medina-Gomez C, Bacigalupe R, Radjabzadeh D, Wang J, Demirkan A, et al. Large-scale association analyses identify host factors influencing human gut microbiome composition. *Nat Genet.* (2021) 53:156–65. doi: 10.1038/s41588-020-00763-1

26. Jia J, Dou P, Gao M, Kong X, Li C, Liu Z, et al. Assessment of causal direction between gut Microbiota-dependent metabolites and cardiometabolic health: a bidirectional Mendelian randomization analysis. *Diabetes*. (2019) 68:1747–55. doi: 10.2337/db19-0153

27. Fadista J, Lund M, Skotte L, Geller F, Nandakumar P, Chatterjee S, et al. Genome-wide association study of Hirschsprung disease detects a novel low-frequency variant at the RET locus. *Eur J Hum Genet*. (2018) 26:561–9. doi: 10.1038/s41431-017-0053-7

28. Manchikanti L, Falco FJE, Hirsch JA. Ready or not! here comes ICD-10. *J Neurointerv Surg*. (2013) 5:86–91. doi: 10.1136/neurintsurg-2011-010155

29. Chen J, Zhou H, Jin H, Liu K. Role of inflammatory factors in mediating the effect of lipids on nonalcoholic fatty liver disease: a two-step, multivariable Mendelian randomization study. *Nutrients*. (2022) 14:4434. doi: 10.3390/nu14204434

30. Wang C, Zhu D, Zhang D, Zuo X, Yao L, Liu T, et al. Causal role of immune cells in schizophrenia: Mendelian randomization (MR) study. *BMC Psychiatry*. (2023) 23:590. doi: 10.1186/s12888-023-05081-4

31. Good BH. Linkage disequilibrium between rare mutations. *Genetics*. (2022) 220: iyac004. doi: 10.1093/genetics/iyac004

32. Lawlor DA, Harbord RM, Sterne JAC, Timpson N, Davey Smith G. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. *Stat Med*. (2008) 27:1133–63. doi: 10.1002/sim.3034

33. Feng R, Lu M, Xu J, Zhang F, Yang M, Luo P, et al. Pulmonary embolism and 529 human blood metabolites: genetic correlation and two-sample Mendelian randomization study. *BMC Genom Data*. (2022) 23:69. doi: 10.1186/s12863-022-01082-6

34. 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. doi: 10.1002/gepi.21965

35. Cho Y, Haycock PC, Sanderson E, Gaunt TR, Zheng J, Morris AP, et al. Exploiting horizontal pleiotropy to search for causal pathways within a Mendelian randomization framework. *Nat Commun*. (2020) 11:1010. doi: 10.1038/s41467-020-14452-4

36. Walker VM, Davies NM, Hemani G, Zheng J, Haycock PC, Gaunt TR, et al. Using the MR-base platform to investigate risk factors and drug targets for thousands of phenotypes. *Wellcome Open Res*. (2019) 4:113. doi: 10.12688/wellcomeopenres.15334.1

37. Xu J, Zhang S, Tian Y, Si H, Zeng Y, Wu Y, et al. Genetic causal association between iron Status and osteoarthritis: a two-sample Mendelian randomization. *Nutrients*. (2022) 14:3683. doi: 10.3390/nu14183683

38. Hoaglin DC. Misunderstandings about Q and "Cochran's Q test" in meta-analysis. *Stat Med*. (2016) 35:485–95. doi: 10.1002/sim.6632

39. Bowden J, Del Greco M F, Minelli C, Davey Smith G, Sheehan NA, Thompson JR. Assessing the suitability of summary data for two-sample Mendelian randomization analyses using MR-egger regression: the role of the I2 statistic. *Int J Epidemiol*. (2016) 45:1961–74. doi: 10.1093/ije/dyw252

40. Verbanck M, Chen C-Y, Neale B, Do R. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. *Nat Genet*. (2018) 50:693–8. doi: 10.1038/s41588-018-0099-7

41. Cheng H, Garrick DJ, Fernando RL. Efficient strategies for leave-one-out cross validation for genomic best linear unbiased prediction. *J Anim Sci Biotechnol*. (2017) 8:38. doi: 10.1186/s40104-017-0164-6

42. Relton CL, Davey Smith G. Two-step epigenetic Mendelian randomization: a strategy for establishing the causal role of epigenetic processes in pathways to disease. *Int J Epidemiol*. (2012) 41:161–76. doi: 10.1093/ije/dyr233

43. Clausen DS, Willis AD. Evaluating replicability in microbiome data. *Biostatistics*. (2022) 23:1099–114. doi: 10.1093/biostatistics/kxab048

44. Chantakhow S, Khorana J, Tepmalai K, Boonchooduang N, Chattipakorn N, Chattipakorn SC. Alterations of gut bacteria in Hirschsprung disease and Hirschsprung-associated enterocolitis. *Microorganisms*. (2021) 9:2241. doi: 10.3390/microorganisms9112241

45. Cheng Z, Zhao L, Dhall D, Ruegger PM, Borneman J, Frykman PK. Bacterial microbiome dynamics in post pull-through Hirschsprung-associated enterocolitis (HAEC): an experimental study employing the endothelin receptor B-null mouse model. *Front Surg*. (2018) 5:30. doi: 10.3389/fsurg.2018.00030

46. Neuvonen MI, Korpela K, Kyrklund K, Salonen A, de Vos W, Rintala RJ, et al. Intestinal microbiota in Hirschsprung disease. *J Pediatr Gastroenterol Nutr*. (2018) 67:594–600. doi: 10.1097/MPG.0000000000001999

47. Prato A P, Bartow-McKenney C, Hudspeth K, Mosconi M, Rossi V, Avanzini S, et al. A metagenomics study on Hirschsprung's disease associated enterocolitis: biodiversity and gut microbial homeostasis depend on resection length and patient's clinical history. *Front Pediatr*. (2019) 7:326. doi: 10.3389/fped.2019.00326

48. Nie K, Ma K, Luo W, Shen Z, Yang Z, Xiao M, et al. Roseburia intestinalis: a beneficial gut organism from the discoveries in genus and species. *Front Cell Infect Microbiol*. (2021) 11:757718. doi: 10.3389/fcimb.2021.757718

49. Tamanai-Shacoori Z, Smida I, Bousarghin L, Loreal O, Meuric V, Fong SB, et al. Roseburia spp.: a marker of health? *Future Microbiol*. (2017) 12:157–70. doi: 10.2217/fmb-2016-0130

50. Song L, He M, Sun Q, Wang Y, Zhang J, Fang Y, et al. Roseburia hominis increases intestinal melatonin level by activating p-CREB-AANAT pathway. *Nutrients*. (2021) 14:117. doi: 10.3390/nu14010117

51. Scott KP, Martin JC, Chassard C, Clerget M, Potrykus J, Campbell G, et al. Substrate-driven gene expression in Roseburia inulinivorans: importance of inducible enzymes in the utilization of inulin and starch. *Proc Natl Acad Sci U S A*. (2011) 108(Suppl 1): 4672–9. doi: 10.1073/pnas.1000091107

52. Choi SI, Kim N, Nam RH, Jang JY, Kim EH, Ha S, et al. The protective effect of Roseburia faecis against repeated water avoidance stress-induced irritable bowel syndrome in a wister rat model. *J Cancer Prev*. (2023) 28:93–105. doi: 10.15430/JCP.2023.28.3.93

53. Bourgault AM, Rosenblatt JE, Fitzgerald RH. Peptococcus magnus: a significant human pathogen. *Ann Intern Med*. (1980) 93:244–8. doi: 10.7326/0003-4819-93-2-244

54. Gu Y, Hou M, Chu J, Wan L, Yang M, Shen J, et al. The cause and effect of gut microbiota in development of inflammatory disorders of the breast. *Eur J Med Res*. (2023) 28:324. doi: 10.1186/s40001-023-01281-6

55. Martínez de Victoria Carazo J, Vinuesa García D, Serrano-Conde Sánchez E, Peregrina Rivas JA, Ruíz Rodríguez AJ, Hernández Quero J. Ruminococcus gnavus bacteremia: literature review and a case report associated with acute flare of ulcerative colitis in an immunocompromised patient. *Anaerobe*. (2023) 82:102762. doi: 10.1016/j.anaerobe.2023.102762

56. Kaakoush NO. Insights into the role of erysipelotrichaceae in the human host. *Front Cell Infect Microbiol*. (2015) 5:84. doi: 10.3389/fcimb.2015.00084

57. Carneiro G, Radcenco AL, Evaristo J, Monnerat G. Novel strategies for clinical investigation and biomarker discovery: a guide to applied metabolomics. *Horm Mol Biol Clin Investig*. (2019) 38(3):20180045. doi: 10.1515/hmbci-2018-0045

58. Yang S, Yang H, Huang Y, Chen G, Shen C, Zheng S. Serum metabolomic signatures of Hirschsprung's disease based on GC-MS and LC-MS. *J Proteome Res*. (2023) 22:2307–18. doi: 10.1021/acs.jproteome.3c00008

59. Koal T, Klavins K, Seppi D, Kemmler G, Humpel C. Sphingomyelin SM(d18:1/18:0) is significantly enhanced in cerebrospinal fluid samples dichotomized by pathological amyloid-β42, tau, and phospho-tau-181 levels. *J Alzheimers Dis*. (2015) 44:1193–201. doi: 10.3233/JAD-142319

60. Plekhova V, De Paepe E, Van Renterghem K, Van Winckel M, Hemeryck LY, Vanhaecke L. Disparities in the gut metabolome of post-operative Hirschsprung's disease patients. *Sci Rep*. (2021) 11:16167. doi: 10.1038/s41598-021-95589-0