Gut microbiota and insomnia: Mendelian randomization and network pharmacology to predict potential intervention with traditional Chinese medicine

Keywords
gut microbiota, Network pharmacology, Causality, Insomnia, Mendelian randomization, traditional Chinese medicine

Abstract

Introduction
To identify gut microbiota (GM) with genetic causal effects on insomnia using Mendelian randomization (MR) and predict potential traditional Chinese medicine (TCM) for GM-targeted intervention in insomnia.

Material and methods
Summary data from genome-wide association studies (GWAS) on GM and insomnia were obtained from the IEU OpenGWAS database. The R 4.4.1 software, particularly the TwoSampleMR package, was utilized to assess the genetic correlation between GM and insomnia, primarily using the inverse-variance weighted (IVW) method. Functional enrichment analysis was conducted on the genes adjacent to the instrumental variables to explore the signaling pathways through which related gut microbiota may mediate insomnia. The CTD and Coremine databases were combined to predict TCM with potential regulatory effects on the genes adjacent to the instrumental variables, and their properties, meridian tropism, and efficacy information were compiled.

Results
The MR analysis revealed that Ruminococcaceae and Marvinbryantia were associated with an increased risk of insomnia, while Pasteurellaceae, Olsenella, the Ruminococcus gnavus group, Mollicutes RF9, and Pasteurellales were associated with a decreased risk. The genes adjacent to the instrumental variables were mainly enriched in signaling pathways such as neuroactive ligand-receptor interaction and the mTOR.

Conclusions
The MR analysis identified seven gut microbiota, represented by Ruminococcaceae and Marvinbryantia, that may mediate the occurrence and development of insomnia through signaling pathways such as mTOR and neuroactive ligand-receptor interaction. It predicted potential TCMs that act on gut microbiota to intervene in insomnia. This study provided a reference for exploring TCM prevention and treatment strategies for insomnia from the perspective of gut microbiota.

Gut microbiota and insomnia: Mendelian randomization and network pharmacology to predict potential intervention with traditional Chinese medicine

Xu WANG[1,2,3]，Jie ZHANG[1,3], Xiao-yu SONG[1,2,3], Jian-ping DUAN[1,2,3], Ying ZHANG[4]*

1 Dept of Encephalopathy, the First Affiliated Hospital of Henan University of Chinese Medicine, Zhengzhou 450000, Henan, China

2 the First Clinical Medical College, Henan University of Chinese Medicine, Zhengzhou 450000, Henan, China

3 Collaborative Innovation Center of Prevention and Treatment of Major Diseases by Chinese and Western Medicine,Henan Province, Zhengzhou 450000, Henan, China

4 Henan University of Chinese Medicine, Zhengzhou 450000, Henan, China

** Correspondence:*

Jie ZHANG

Jie830@163.com

Abstract

Objective

To identify gut microbiota (GM) with genetic causal effects on insomnia using Mendelian randomization (MR) and predict potential traditional Chinese medicine (TCM) for GM-targeted intervention in insomnia.

Methods

Summary data from genome-wide association studies (GWAS) on GM and insomnia were obtained from the IEU OpenGWAS database. The R 4.4.1 software, particularly the TwoSampleMR package, was utilized to assess the genetic correlation between GM and insomnia, primarily using the inverse-variance weighted (IVW) method. Functional enrichment analysis was

conducted on the genes adjacent to the instrumental variables to explore the signaling pathways through which related gut microbiota may mediate insomnia. The CTD and Coremine databases were combined to predict TCM with potential regulatory effects on the genes adjacent to the instrumental variables, and their properties, meridian tropism, and efficacy information were compiled.

Results

The MR analysis revealed that Ruminococcaceae and Marvinbryantia were associated with an increased risk of insomnia, while Pasteurellaceae, Olsenella, the Ruminococcus gnavus group, Mollicutes RF9, and Pasteurellales were associated with a decreased risk. The genes adjacent to the instrumental variables were mainly enriched in signaling pathways such as neuroactive ligand-receptor interaction and the mammalian target of rapamycin (mTOR). Representative TCM with high mapping frequencies included Panax ginseng, Curcuma aromatica, Salviae Miltiorrhizae Radix, Zingiber officinale, Glycyrrhizae Radix et Rhizoma, Aucklandiae Radix, Magnoliae Officinalis Cortex, Scutellariae Radix, Ganoderma lucidum, and Poria cocos.

Conclusion

The MR analysis identified seven gut microbiota, represented by Ruminococcaceae and Marvinbryantia, that may mediate the occurrence and development of insomnia through signaling pathways such as mTOR and neuroactive ligand-receptor interaction. It predicted potential traditional Chinese medicines that act on gut microbiota to intervene in insomnia. This study provided a reference for exploring TCM prevention and treatment strategies for insomnia from the perspective of gut microbiota.

Keywords: gut microbiota; Insomnia; Mendelian randomization; Network pharmacology; Causality; traditional Chinese medicine

1.Introduction

Insomnia is a common sleep disorder characterized by difficulties falling asleep or maintaining sufficient sleep duration, dreaming so much, waking up easily, and returning sleep difficulty, or even staying awake throughout the night[1]. Statistical data indicate that the incidence of insomnia among adults in China reaches as high as 38.2%, affecting over 300 million Chinese individuals with sleep disorders[2]. Chronic insomnia poses a significant burden on patients[3]. Although hypnotic medications can temporarily alleviate insomnia symptoms, they may lead to over-reliance on such drugs without addressing the underlying issue[4]. The pathophysiological mechanisms of insomnia are complex and influenced by factors such as sleep environment[5], dietary habits[6], endocrine status[7], psychological issues[8], and circadian rhythms[9]. Increasing evidence suggests that sleep may be related to gut microbiota (GM)[10]. GM is crucial for maintaining bodily homeostasis, including nutrient absorption, metabolism, and toxin degradation. Emerging research suggests a potential bidirectional relationship between sleep regulation and GM functionality through the "microbiome-gut-brain axis"[11]. Clinical studies have confirmed that GM diversity is positively correlated with sleep efficiency and total sleep time, and negatively correlated with wake after sleep onset (WASO)[12]. Animal experimental research has shown that sleep restriction and increased WASO can lead to decreased microbial richness and diversity[13]. The mechanism of the "microbiome-gut-brain axis" between GM and the nervous system reveals the complex interactions between the gut and the brain[14]. GM dysregulation not only affects the metabolic system but may also impact brain function, leading to insomnia, through immune regulation[15], neuroendocrinology[10], and other pathways.

Relevant studies have indicated that individuals with gastrointestinal diseases are more prone to insomnia than the general population[16]. However, identifying specific bacterial species that affect the onset of insomnia among the complex GM remains a research hotspot and challenge. Mendelian randomization (MR), as an advanced causal inference method, uses single-

nucleotide polymorphisms (SNPs) as instrumental variables (IVs) for exposure to explore the causal relationship between exposure factors and outcome events[17]. It minimizes biases caused by confounding factors and reverse causality, thereby avoiding reverse causality and common errors in various epidemiological studies. Therefore, this study employs MR analysis to explore the potential causal relationship between GM and insomnia and conducts functional enrichment analysis on the genes adjacent to the IVs to investigate the signaling pathways through which related GM may mediate the occurrence of insomnia. Finally, combining the CTD and Coremine databases, we predict Traditional Chinese Medicine (TCM) with potential regulatory effects on the genes adjacent to the IVs, aiming to provide a potential theoretical basis for the integrated traditional Chinese and Western medicine treatment of insomnia.

2.Materials and methods

2.1 Study design

A two-sample MR study was designed to estimate the potential causal link between GM and insomnia. The SNPs were selected as IVs and stick to three essential premises as follows[18]: (1) SNPs should be intensely linked to GM as exposure; (2) SNPs should not be linked to confounding factors; and (3) SNPs should not be linked to insomnia as outcome directly. Subsequently, the adjacent genes of the instrumental variables were obtained, and functional enrichment analysis was conducted to identify the key biological pathways through which the gut microbiota mediates the occurrence of insomnia. Furthermore, potential traditional Chinese medicines that regulate this pathway were predicted (**Figure 1**).

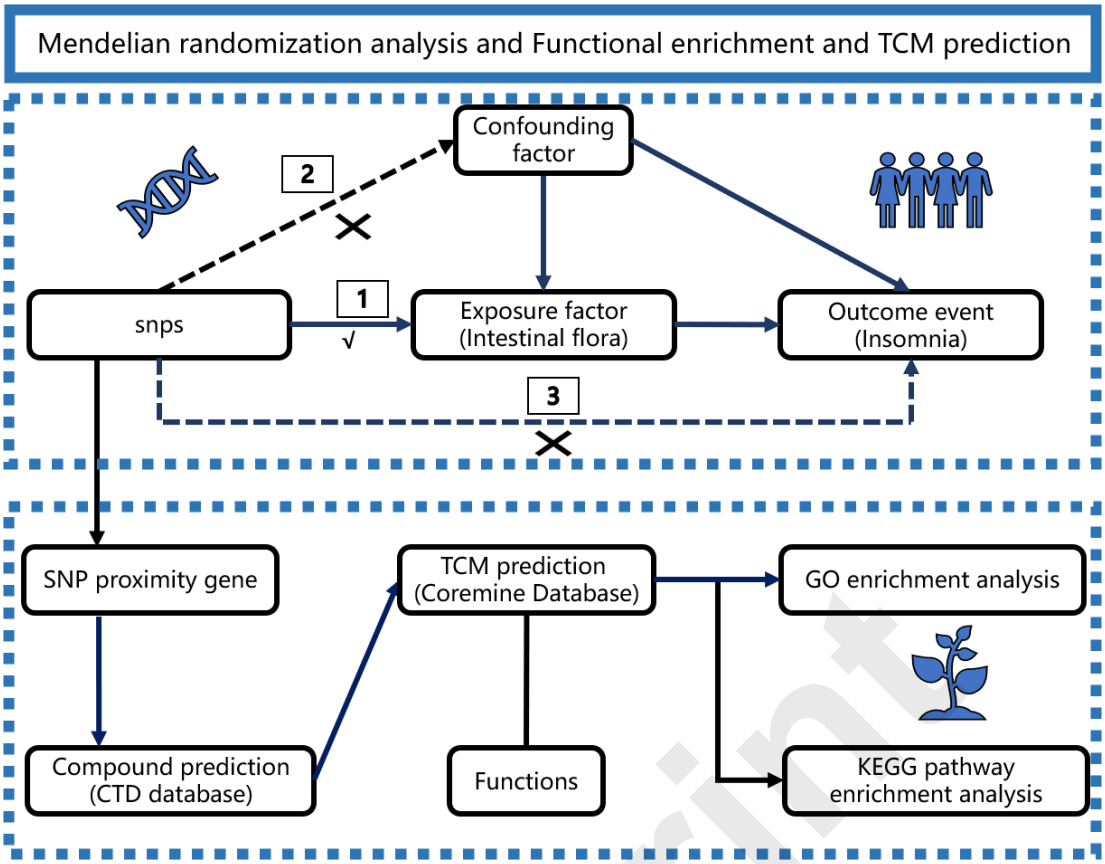


Figure 1. Study design of MR analysis and Functional enrichment and TCM prediction

2.2 GWAS summary statistics

The GWAS data for GM were sourced from the IEU OpenGWAS database (https://gwas.mrcieu.ac.uk/). Utilizing the ao function of the TwoSampleMR software package in R, we extracted 418 GWAS summary datasets for 211 gut microbiota traits from the IEU OpenGWAS database. Similarly, the insomnia data were also obtained from the IEU OpenGWAS database, encompassing 1,402 patients and 485,225 healthy subjects (controls). Both GWAS datasets included populations of European descent, sharing the same genetic background, which allows for the conduct of MR studies.

2.3 Ethical approval

All summary-level datasets in our study were obtained from de-identified public data/studies. Ethical approval and informed consent were previously obtained from the ethics committee. Thus, the requirement for ethical approval was waived for this study.

2.4 SNPs selection

Firstly, we conducted a screening process to identify SNPs that were highly correlated with exposure at a genome-wide significance level ($p < 5\times10^{-8}$). Secondly, we implemented a criterion ($r^2 < 0.001$, kb=10000) to choose SNPs that were free from dependence on linkage disequilibrium (LD). Thirdly, we excluded SNPs that were not present in the insomnia dataset and palindromic SNPs which have the potential to introduce bias. All of the SNPs for instrumental variables were uploaded to PhenoScanner to identify confounding SNPs associated with insomnia. Based on the assumption of the MR analysis, SNPs used as instrumental variables should be strongly associated with exposure. Subsequently, we ensured the harmonization of exposure and outcome data, confirming that the effect of the SNP on the exposure corresponded to the same allele as its effect on the outcome. Following this, we assessed the possibility of weak instrumental bias by calculating F-statistics, and excluded SNPs with F-statistics less than 10. The F statistic was calculated using the formula $F = beta^2/se^2$. Finally, we employed the MR-PRESSO method to identify outlier SNPs. After removing the outliers, the remaining SNPs were utilized for subsequent MR analysis.

2.5 Two-sample Mendelian analysis

Three popular MR methods were employed to assess causal effects: inverse variance weighted (IVW), weighted median and MR-Egger[19,20]. IVW, a reliable and robust MR method in the absence of horizontal pleiotropy[21], combines the Wald estimates of individual SNP to derive overall estimates of the effect of GM on insomnia risk. Consequently, the IVW method is broadly acknowledged as the most effective approach to assess causality. Odds ratios (ORs) were utilized to express the effects of GM on insomnia risk. If the result of the IVW method is significant ($p < 0.05$), it can be considered positive even if other methods yield nonsignificant results, provided that the ORs of those methods line up in the identical direction without heterogeneity or pleiotropy. Two types of IVW approaches, namely the fixed and random effect model,

were employed to account for existing heterogeneity. Cochran's Q test was used to assess the heterogeneity in the IVW method and MR-Egger regression, with a P-value < 0.05 considered statistically significant[22]. Unlike IVW, the MR-Egger method includes an intercept term designed to test for horizontal pleiotropy. A non-zero intercept term indicates that not all genetic variants are valid instruments, thereby biasing IVW estimates. When the instrument strength independent of direct effect (InSIDE) assumption is met, the MR–Egger method can offer an approximation of the causal impact of horizontal pleiotropy. The weighted median method offers a robust effect estimate, even in the presence of unbalanced horizontal pleiotropy (e.g., when 50% of instrumental SNPs are invalid)[23]. Finally, the MR-PRESSO method encompasses three detection functions[24]: horizontal pleiotropic detection, horizontal pleiotropic correction (after outlier removal), as well as assessment of differences in the results of causality estimation before and after correction.

2.6 Statistical analysis

Heterogeneity was assessed by employing Cochran's Q test, where a p-value > 0.05 indicated the absence of heterogeneity. The MR-Egger regression test was utilized to identify horizontal pleiotropy, where a zero-intercept suggests the absence of pleiotropy ($p > 0.05$).

2.7 Reverse MR analysis

To explore the potential causal relationship between insomnia and GM, a reverse MR analysis was carried out, wherein insomnia served as the exposure and GM as the outcome, employing SNPs associated with insomnia as IVs.

All statistical analyses were conducted using R software (version 4.2.3) with the "TwoSampleMR" (version 0.5.6), "MRPRESSO" (version 1.0), and "MendelianRandomization" (version 0.7.0) packages.

2.8 Instrumental variables near gene function enrichment and traditional Chinese medicine prediction

Firstly, using the Rstudio software package, we identified the proximal genes of the instrumental variables based on the SNP identification numbers, as well

as their respective chromosomal sequences and loci. Subsequently, we utilized the CTD database (https://ctdbase.org/)[25] to search for the chemical components corresponding to these proximal genes. By reviewing the literature, we eliminated chemical components with lower support numbers. Then, using Coremine data (https://coremine.com/medical/)[26], we identified TCM that were significantly associated with the aforementioned chemical components, with a screening threshold of $P < 0.05$. Finally, we employed the David tool (https://david.ncifcrf.gov) to perform Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG)[27] enrichment analyses on the proximal genes of the instrumental variables. All enriched pathways and functional terms were filtered using an FDR-adjusted $P < 0.05$ to ensure biological relevance.

3.Results

3.1 SNP screen results

Based on the established screening criteria, we identified a total of 2,559 SNPs significantly and independently associated with GM. All SNPs exhibited an F-statistic >10, indicating the absence of weak instrument bias. SNPs associated with the outcome were excluded using the PhenoScanner database (http://www.phenoscanner.medschl.cam.ac.uk/), thereby removing confounded SNPs. Finally, MR-PRESSO was employed to detect outliers and correct for horizontal pleiotropy. Where horizontal pleiotropy was detected among the instrumental variables, outliers were subsequently removed.

3.2 The MR analysis results

The results of the MR analysis, primarily using the IVW method, showed that the odds ratios (ORs) for Ruminococcaceae (IVW: OR=1.578; 95%CI: 1.074 – 2.317; P=0.020) and Marvinbryantia (IVW: OR=1.537; 95%CI: 1.062 – 2.225; P=0.023) were both greater than 1, indicating that they increase the risk of

insomnia. Conversely, the ORs for Pasteurellaceae (IVW: OR=0.764; 95%CI: 0.599 – 0.975; P=0.030), Olsenella (IVW: OR=0.781; 95%CI: 0.641 – 0.951; P=0.014), the Ruminococcus gnavus group (IVW: OR=0.746; 95%CI: 0.588 – 0.946; P=0.016), Mollicutes RF9 (IVW: OR=0.706; 95%CI: 0.525 – 0.949; P=0.021), and Pasteurellales (IVW: OR=0.764; 95%CI: 0.599 – 0.975; P=0.030) were all less than 1, suggesting that they are protective factors against insomnia, reducing the risk of its development. Table 1 presents the detailed results of the MR analysis. The forest plot (Figure 2) and Circus plot (Figure 3) display of the MR analysis results, which lists the positive result data in detail.

Table 1 Positive results of MR Analysis of GM and Insomnia

Exposure	Method	nSNP	P	OR	or_lci95	or_uci95
Pasteurellaceae	IVW	14	0.030	0.764	0.599	0.975
Pasteurellaceae	MR Egger	14	0.059	0.568	0.334	0.966
Pasteurellaceae	Weighted median	14	0.138	0.763	0.534	1.091
Pasteurellaceae	Weighted mode	14	0.309	0.767	0.470	1.253
Pasteurellaceae	Simple mode	14	0.584	0.846	0.472	1.516
Ruminococcaceae	IVW	10	0.020	1.578	1.074	2.317
Ruminococcaceae	Weighted median	10	0.175	1.433	0.852	2.407
Ruminococcaceae	Weighted mode	10	0.302	1.385	0.773	2.484
Ruminococcaceae	Simple mode	10	0.456	1.339	0.643	2.789
Ruminococcaceae	MR Egger	10	0.814	1.107	0.488	2.508
Marvinbryantia	IVW	10	0.023	1.537	1.062	2.225
Marvinbryantia	Weighted median	10	0.063	1.570	0.976	2.528
Marvinbryantia	Weighted mode	10	0.142	1.697	0.890	3.235
Marvinbryantia	Simple mode	10	0.144	1.638	0.895	2.998
Marvinbryantia	MR Egger	10	0.845	1.167	0.260	5.234
Olsenella	IVW	11	0.014	0.781	0.641	0.951
Olsenella	Weighted median	11	0.021	0.730	0.558	0.954
Olsenella	Simple mode	11	0.065	0.658	0.443	0.977
Olsenella	Weighted mode	11	0.073	0.685	0.473	0.991
Olsenella	MR Egger	11	0.115	0.547	0.278	1.077
Ruminococcus gnavus group	IVW	12	0.016	0.746	0.588	0.946
Ruminococcus gnavus group	Weighted median	12	0.291	0.842	0.612	1.158
Ruminococcus gnavus group	MR Egger	12	0.329	0.532	0.159	1.776
Ruminococcus gnavus group	Weighted mode	12	0.699	0.897	0.526	1.532

Ruminococcus gnavus group	Simple mode	12	0.701	0.893	0.508	1.570
Mollicutes RF9	IVW	13	0.021	0.706	0.525	0.949
Mollicutes RF9	Weighted median	13	0.029	0.627	0.412	0.954
Mollicutes RF9	MR Egger	13	0.043	0.327	0.126	0.851
Mollicutes RF9	Weighted mode	13	0.091	0.543	0.282	1.042
Mollicutes RF9	Simple mode	13	0.114	0.550	0.277	1.093
Pasteurellales	IVW	14	0.030	0.764	0.599	0.975
Pasteurellales	MR Egger	14	0.059	0.568	0.334	0.966
Pasteurellales	Weighted median	14	0.131	0.763	0.537	1.084
Pasteurellales	Weighted mode	14	0.300	0.767	0.474	1.241
Pasteurellales	Simple mode	14	0.603	0.846	0.457	1.565

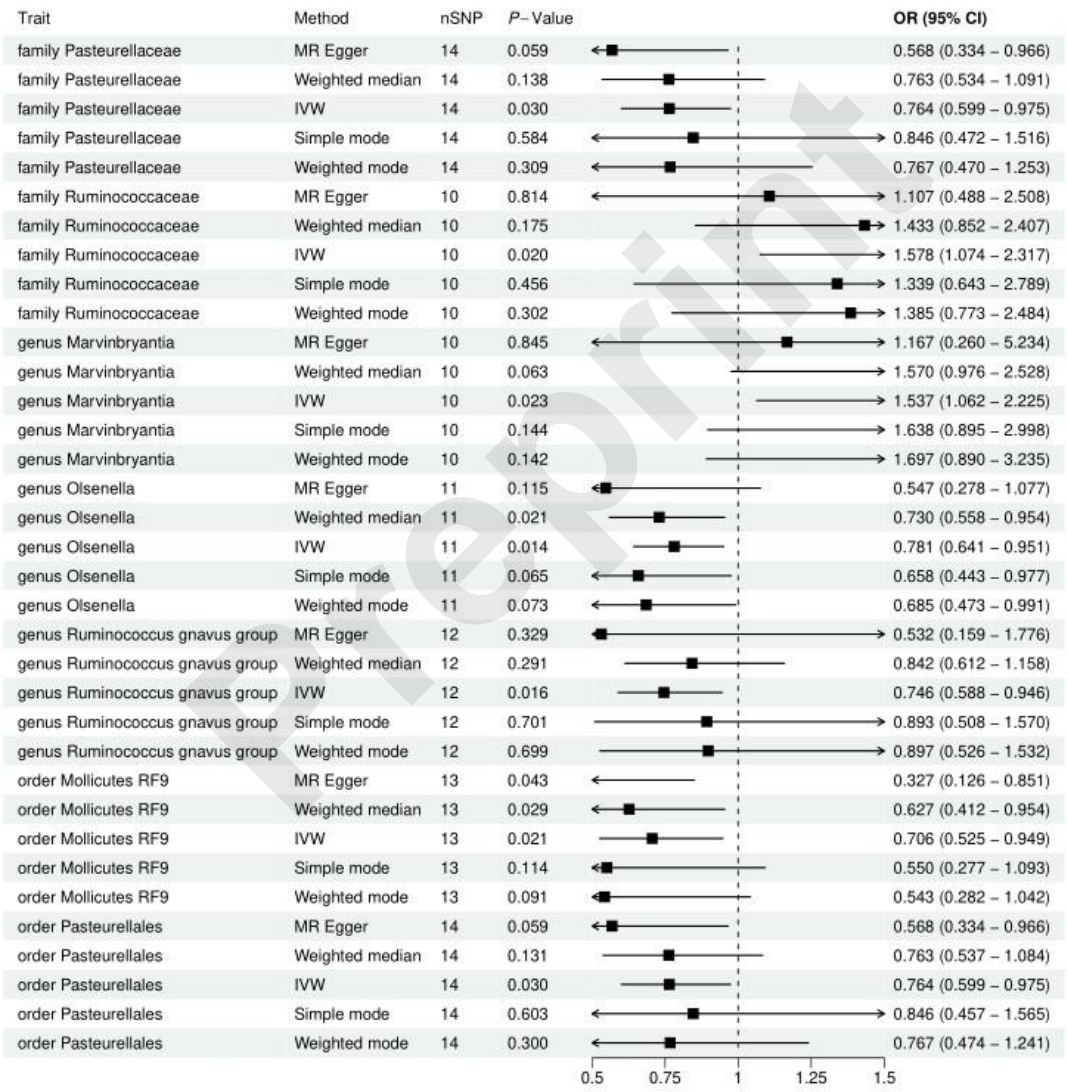


Figure 2 MR forest graph analysis results.From left to right, exposure factors, study methods, instrumental variables, P value (P value < 0.05 is defined as a positive result) and OR value (OR value > 1 is a risk factor and < 1 is a protective

factor) are represented.

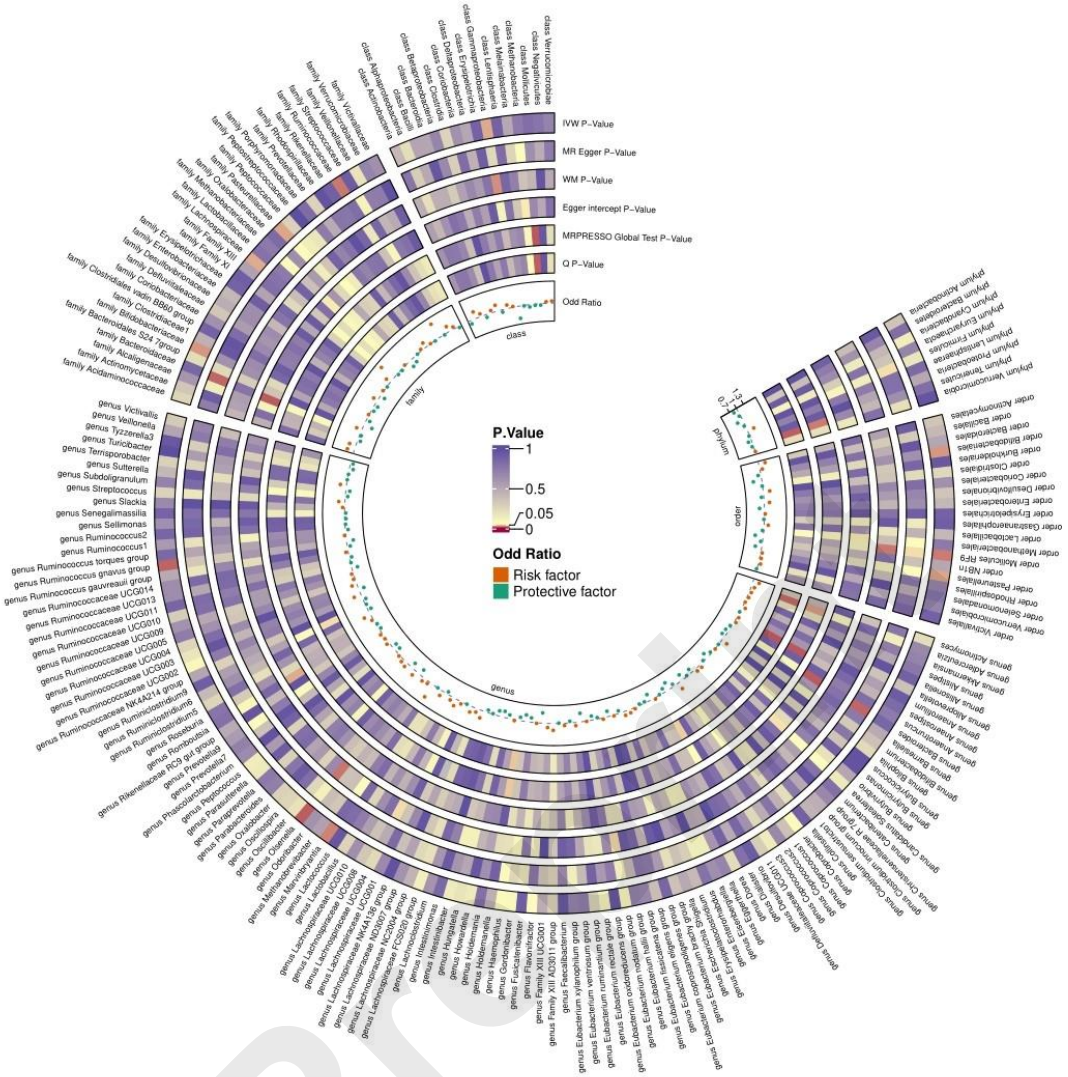


Figure 3 Circus plot of MR analyses of gut GM on insomnia

3.3 Sensitivity analysis

Both the Cochran's Q test for IVW and MR-Egger regression indicated no heterogeneity among the SNPs (Table 2). The intercept of the MR-Egger regression was nearly zero, suggesting the absence of horizontal pleiotropy. The funnel plot indicated minimal influence of potential bias on the causal relationship. Results from the "leave-one-out" analysis showed that after sequentially excluding each SNP, the IVW analysis results of the remaining SNPs were similar to those obtained when all SNPs were included, with no SNPs found to have a significant impact on the causal association, indicating

robustness of the results (Figure 4、Figure 5).

Table 2 Quality control results of GM with causal relationship with Insomnia

Exposure	heterogeneity			pleiotropy_test		MR-PRESSO	
	Method	Q	P	Egger	P	global.test	P
Pasteurellaceae	IVW	12.453	0.491	0.039	0.342	14.526	0.517
Ruminococcaceae	IVW	3.279	0.952	0.038	0.364	3.882	0.967
Marvinbryantia	IVW	5.807	0.759	0.024	0.720	6.723	0.802
Olsenella	IVW	7.146	0.712	0.049	0.309	8.637	0.724
Ruminococcusgnavus group	IVW	11.006	0.443	0.037	0.587	12.988	0.464
Mollicutes RF9	IVW	11.189	0.513	0.065	0.125	13.387	0.538
Pasteurellales	IVW	12.453	0.491	0.039	0.242	14.526	0.522

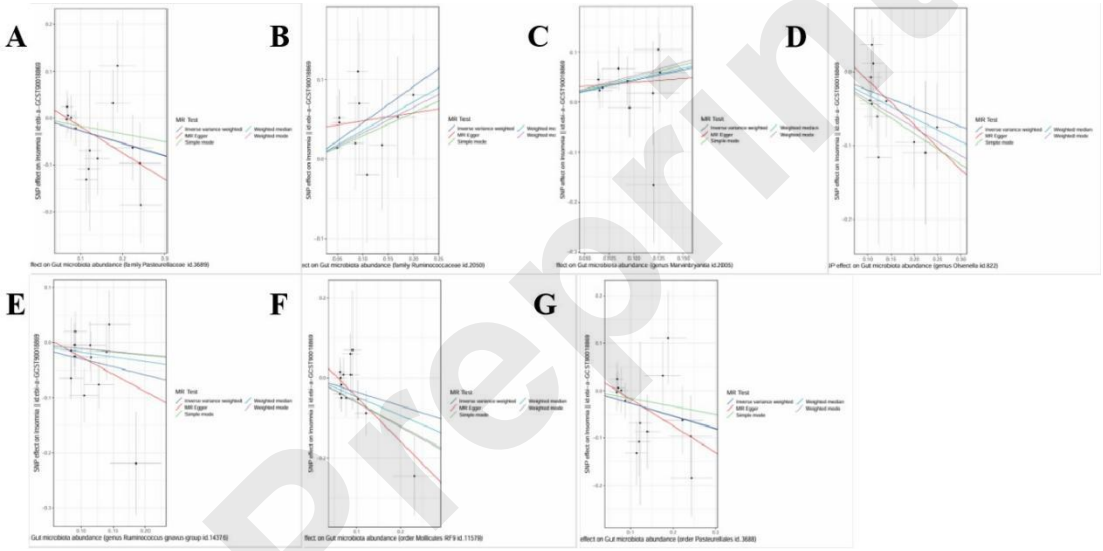


Figure 4 Scatter plot of GM and insomnia. （A） Pasteurellaceae, (B) Ruminococcaceae, (C) Marvinbryantia, (D) Olsenella, (E) Ruminococcusgnavus group, (F) Mollicutes RF9, (G) Pasteurellales.

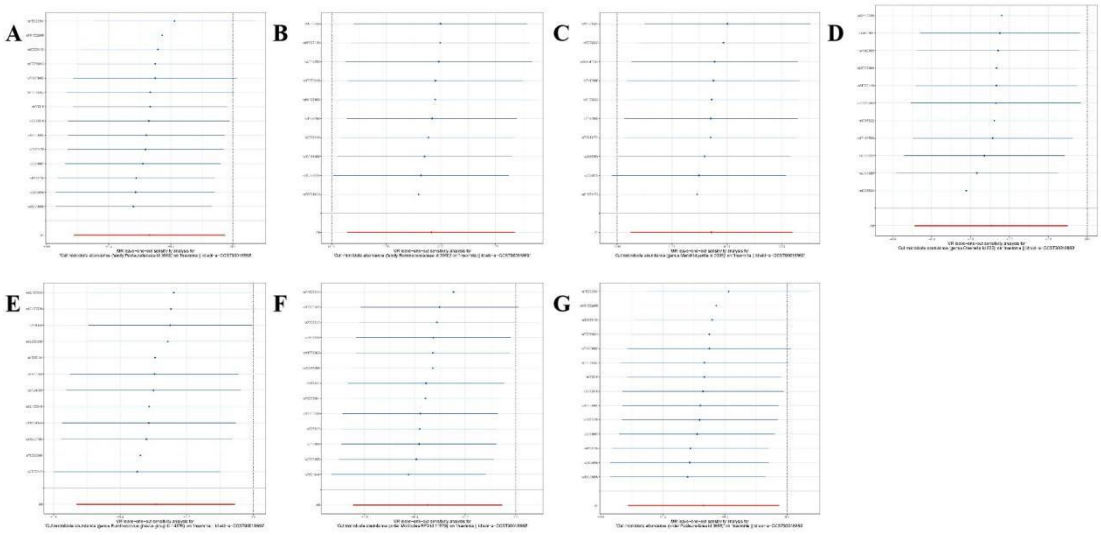


Figure 5 Leave-one-out analysis plot of GM and insomnia.（A）Pasteurellaceae, (B) Ruminococcaceae, (C) Marvinbryantia, (D) Olsenella, (E) Ruminococcusgnavus group, (F) Mollicutes RF9, (G) Pasteurellales.

3.4 Statistical power Calculating

In this study, statistical power was calculated as 0.91 by specifying parameters including sample size, Type I error rate (α), case proportion (K), variance explained by the instrumental variables (R^2), and odds ratio (OR). This value significantly exceeds the conventional threshold of 80% required for adequate power in typical studies. The analysis demonstrates that sufficient power to reliably identify disease-associated genetic variants was maintained in this study, even under imbalanced case-control ratios.

3.5 Functional Enrichment Analysis of Instrumental Variable-Adjacent Genes and Prediction of Potential Traditional Chinese Medicines

Based on the SNP numbers, their respective chromosomal sequences, and loci, 166 genes corresponding to 84 SNPs were identified. These genes were submitted to the CTD database, yielding 111 chemical compounds represented by lipopolysaccharide, quercetin, and others. Subsequently, 336 TCM were obtained from the Coremine database, which were significantly associated with 80 of these chemical compounds. Cytoscape was used to visualize the top-ranked nodes in terms of degree value within the "gene-chemical component-TCM" mapping network, as shown in Figure 6. Representative TCM with high

mapping frequencies include Camellia sinensis root, Ginseng, Radix Curcumae, Salviae Miltiorrhizae Radix, Dried Ginger, Glycyrrhizae Radix and Rhizoma, Aucklandiae Radix, Magnoliae Officinalis Cortex, Scutellariae Radix, Ganodermae Lucidum, and Poria. Finally, functional enrichment analysis was conducted on the genes that have mapping associations with the predicted TCM and chemical components. GO enrichment analysis revealed that these genes are primarily enriched in biological processes such as phosphorylation, mRNA splicing, and hippocampal development; cellular components such as cytoplasm, protein-containing complexes, spliceosomal complexes, and endocytic vesicles; and molecular functions such as protein binding, ATP binding, and protein serine/threonine kinase activity, as shown in Figure 7A. KEGG pathway enrichment analysis indicated that these genes are mainly enriched in pathways such as Neuroactive ligand-receptor interaction and mTOR signaling, as shown in Figure 7B.

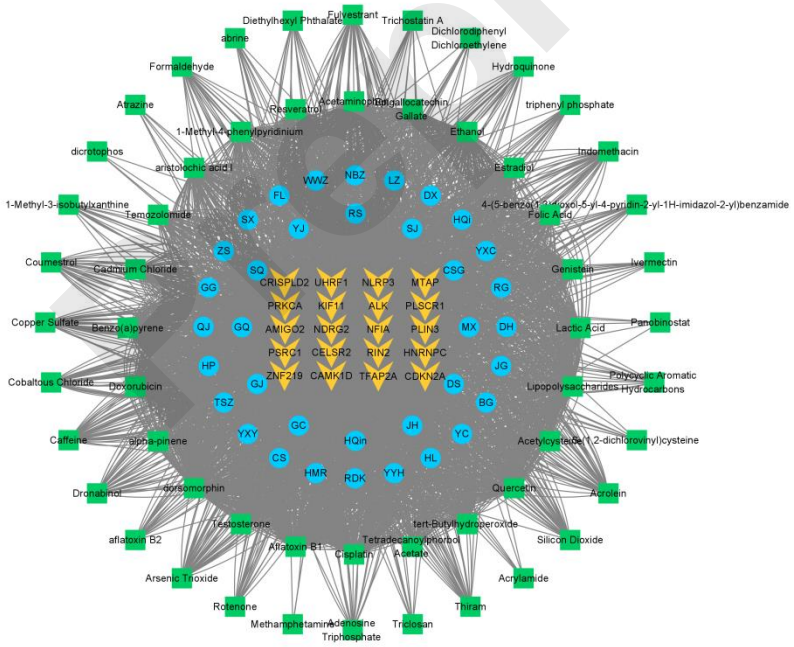
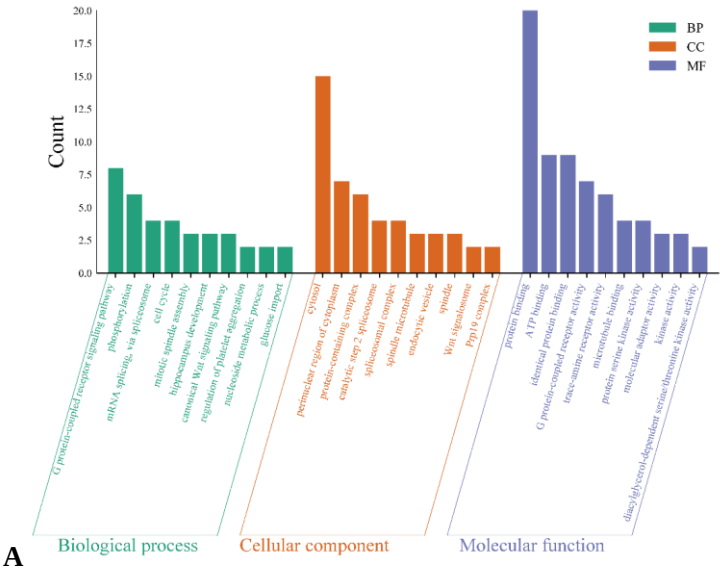
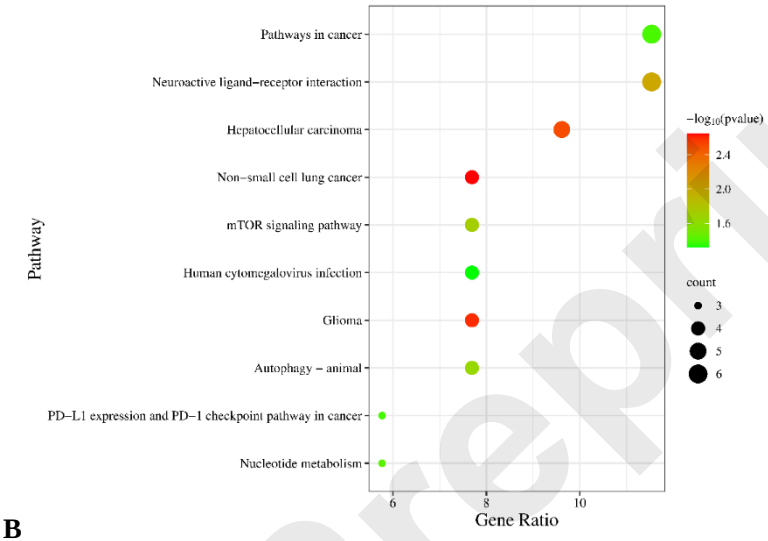


Figure 6 Mapping network diagram of "Gene-chemical composition-Chinese medicine herb" The yellow triangle is the gene, the blue circle is the Chinese medicine, and the green square is the chemical composition.



A



B

Figure 6 Functional enrichment analysis. (A) GO enrichment analysis, (B) KEGG pathway enrichment analysis.

Discussion

The "microbiota-gut-brain axis" represents a complex bidirectional communication system, primarily composed of neural, immune, metabolic, and endocrine pathways, which tightly links the gut and the brain. Abnormal activation of the hypothalamic-pituitary-adrenal (HPA) axis leads to increased cortisol release. Cortisol interacts with immune cells and controls the release of cytokines. Cortisol also disrupts the gut microbiota balance, damages the intestinal barrier, and increases proinflammatory cytokines that interfere with

sleep, potentially leading to insomnia. The vagus nerve, a crucial nerve for gut-brain communication, can receive and transmit various signals from the gut, such as signals from intestinal immune cells, bacterial metabolites (like short-chain fatty acids), and neurotransmitters (like 5-hydroxytryptamine), ultimately affecting brain function and sleep[28].

Numerous studies have indicated a close relationship between the gut microbiota and the occurrence and development of insomnia[29]. Recent Mendelian Randomization Analysis studies have suggested that gut microbiota may influence sleep through various metabolic pathways[30]. Therefore, regulating the gut microbiota may be an effective therapeutic strategy for improving insomnia symptoms, but the specific gut microbiota that affect insomnia remain unclear. Grosicki et al. discovered that Blautia and Ruminococcus abundance and α-diversity increased, while Prevotella abundance decreased, in individuals with better sleep quality[31]. Additionally, some animal experimental results are inconsistent. For example, in mice with fragmented sleep, Firmicutes increased while Bacteroidetes decreased[32]. In contrast, Maki et al. found that during chronic sleep deprivation, Bacteroidetes decreased compared to the control group, with reduced Firmicutes abundance, F:B ratio, and α-diversity[33]. In these observational studies, the association between the gut microbiota and insomnia is susceptible to confounding factors such as age, environment, dietary patterns, and lifestyle, limiting the establishment of a clear causal relationship between the gut microbiota and insomnia. In such cases, MR emerges as a novel method to explore the causal relationship between the gut microbiota and insomnia. Based on GWAS summary data for the gut microbiota and insomnia, our MR analysis identified seven gut microbiota, represented by Ruminococcaceae and the Ruminococcus gnavus group, that have significant genetic causal associations with insomnia. Ruminococcus has been shown to increase in abundance in individuals with higher sleep quality, aligning with our finding that the Ruminococcus gnavus group is associated with a reduced risk of insomnia[31]. Ruminococcaceae can

improve sleep quality in insomnia patients (negatively correlated with PSQI and ISI scores)[34] and has beneficial effects on intestinal barrier function, being lower in mice with circadian rhythm disorders[35]. The results of our study, which found an increased risk of insomnia associated with Ruminococcaceae, contrast with some observational studies. Possible reasons include the vulnerability of observational study results to confounding factors, which may introduce bias due to confounding or reverse causality. Our study provides new insights and support for the impact of the gut microbiota on insomnia.

TCM applies the holistic view and syndrome differentiation and treatment concepts in clinical practice, leveraging the multi-target and multi-pathway effects of Chinese materia medica, offering unique advantages in preventing and treating insomnia[36,37]. Many Chinese medicine compound prescriptions, herb pairs and single Chinese herb have the effect of tranquillizing spirit and stabilizing mind, thus playing a clinical effect in treating insomnia[38]. Basic evidence showed that TCM could improve the sleep quality of insomnia model animals. Its mechanism is related to the improvement of gut microbiota disorder[39]. This study utilizes CTD and Coremine Database to further predict, based on the genes adjacent to SNPs, potential traditional Chinese medicines that may intervene in relevant signaling pathways and thereby affect the gut microbiota to treat insomnia, mainly including ginseng, poria cocos, dried ginger, licorice, magnolia bark, scutellaria root, and ganoderma lucidum. Ginsenoside Rb1 (Rb1) of ginseng exerts neuroprotective effects through regulation of Lactobacillus helveticus abundance[40]. Studies have shown that saponin compounds in ginseng can significantly improve sleep quality indices, shorten sleep latency, and extend sleep duration in mice[41]. Poria cocos has a sweet and mild taste, entering the heart, spleen, and kidney meridians. It strengthens the spleen, removes dampness, calms the mind, and treats insomnia caused by spleen and stomach qi deficiency and damp-phlegm obstruction. The water-soluble polysaccharide,which is the main component of Poria cocos decoction, could significantly improve species richness and

diversity in the intestinal flora of rats with chronic sleep deprivation[42]. Studies have found that chemical components such as α-Pinene in compound essential oils for calming the mind have sedative and hypnotic effects, reducing sleep latency and extending sleep duration[43]. The orexin system-mediated mTOR signaling pathway is an important part of the downstream signaling network of orexin[44]. Inhibiting mTOR can reduce orexin overexpression, thereby alleviating insomnia episodes, aligning with our study findings.

Domestic and international studies suggest that the gut microbiota plays a significant role in the occurrence and development of insomnia. However, observational study results are susceptible to confounding factors. Given the unique advantages of MR in inferring causal effects, our study innovatively employed MR analysis to explore gut microbiota with significant causal associations with insomnia, minimizing the possibility of bias due to residual factors, confounding factors, or reverse causality. Several sensitivity analyses were conducted to satisfy the core assumptions of MR, ultimately identifying highly correlated and independent genetic variants with the phenotype, excluding those associated with potential confounding factors, and ensuring the accuracy of the results. MR studies are affected by pleiotropy, and our study used MR-Egger regression to ensure robustness. Based on this, predicting potential interventional Chinese materia medica through MR instrumental variable adjacent genes is of great significance in the prevention and treatment of insomnia, providing a reference for subsequent research on new anti-insomnia Chinese materia medica. This study has some limitations: ① All GWAS data selected were from European populations, so the applicability of our conclusions to other populations needs further verification. ② Single genetic factors cannot explain all phenotypic variations, and our study could not consider environmental factors. ③ As MR analysis is a research method based on genetic inference of causality, it can only provide potential causal relationships and cannot determine the specific biological pathways leading to these causal relationships. Future studies can use a combination of

bioinformatics analysis and experiments to explain and validate potential molecular mechanisms.

Conclusion

In this study, we foubd that the MR analysis indicate that there is a bidirectional causal relationship between the gut microbiota and insomnia. We have predicted potential traditional Chinese medicines that act on gut microbiota to intervene in insomnia. The findings of this research offer valuable perspectives on the mechanism and clinical investigation of insomnia caused by gut microbiota.

Funding

This study was supported by the Henan Province Traditional Chinese Medicine Science Research Special Fund Project（2022JDZX004） and the National Nature ScienceFoundation of China（81904264）.

Data availability

The data used in this study are publicly available. The summary statistics for GM and insomnia were acquired and extracted from the IEU OpenGWAS database (https://gwas.mrcieu.ac.uk/).

Declarations

Ethics approval and consent to participate

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

References

[1] Sutton E L. Insomnia[J]. Ann Intern Med, 2021, 174(3): Itc33-itc48.

[2] Morin C M, Inoue Y, Kushida C, et al. Endorsement of European guideline for the diagnosis and treatment of insomnia by the World Sleep Society[J]. Sleep Med, 2021, 81: 124-126.

[3] Nicholson K, Rodrigues R, Anderson K K, et al. Sleep behaviours and multimorbidity occurrence in middle-aged and older adults: findings from the Canadian Longitudinal Study on Aging (CLSA)[J]. Sleep Med, 2020, 75: 156-162.

[4] Kim Y S, Lee B K, Kim C S, et al. Sedum kamtschaticum Exerts Hypnotic Effects via the Adenosine A(2A) Receptor in Mice[J]. Nutrients, 2024, 16(16).

[5] Duan J, Li Q, Yin Z, et al. Outdoor Artificial Light at Night and Insomnia-Related Social Media Posts[J]. JAMA Netw Open, 2024, 7(11): e2446156.

[6] Yao L, Liang K, Huang L, et al. Longitudinal Associations between Healthy Eating Habits, Resilience, Insomnia, and Internet Addiction in Chinese College Students: A Cross-Lagged Panel Analysis[J]. Nutrients, 2024, 16(15).

[7] Luo Y, Yu L, Zhang P, et al. Larger hypothalamic subfield volumes in patients with chronic insomnia disorder and relationships to levels of corticotropin-releasing hormone[J]. J Affect Disord, 2024, 351: 870-877.

[8] Jin H, Wu P, Wang Z, et al. Association between Life's Essential 8 and depression: a population-based study[J]. Arch Med Sci, 2025, 21(2): 505-513.

[9] Sarsembayeva D, Hartman C A, Cardoso Melo R D, et al. Nonlinear associations between insomnia symptoms and circadian preferences in the general population: Symptom-specific and lifespan differences in men and women[J]. Sleep Health, 2024, 10(2): 171-181.

[10] Jiang Z, Zhuo L B, He Y, et al. The gut microbiota-bile acid axis links the positive association between chronic insomnia and cardiometabolic diseases[J]. Nat Commun, 2022, 13(1): 3002.

[11] Sgro M, Kodila Z N, Brady R D, et al. Synchronizing our clocks as we age: the influence of the brain-gut-immune axis on the sleep-wake cycle across the lifespan[J]. Sleep, 2022, 45(3).

[12] Smith R P, Easson C, Lyle S M, et al. Gut microbiome diversity is associated with sleep physiology in humans[J]. PLoS One, 2019, 14(10): e0222394.

[13] Bowers S J, Vargas F, González A, et al. Repeated sleep disruption in mice leads to persistent shifts in the fecal microbiome and metabolome[J]. PLoS One, 2020, 15(2): e0229001.

[14] Wang Z, Wang Z, Lu T, et al. The microbiota-gut-brain axis in sleep disorders[J]. Sleep Med Rev, 2022, 65: 101691.

[15] Wang Q, Chen B, Sheng D, et al. Multiomics Analysis Reveals Aberrant Metabolism and Immunity Linked Gut Microbiota with Insomnia[J]. Microbiol Spectr, 2022, 10(5): e0099822.

[16] Salwen-Deremer J K, Sun M. Management of Sleep and Fatigue in Gastrointestinal Patients[J]. Gastroenterol Clin North Am, 2022, 51(4): 829-847.

[17] Yavorska O O, Burgess S. MendelianRandomization: an R package for performing Mendelian randomization analyses using summarized data[J]. Int J Epidemiol, 2017, 46(6): 1734-1739.
[18] Skrivankova V W, Richmond R C, Woolf B a R, et al. Strengthening the Reporting of Observational Studies in Epidemiology Using Mendelian Randomization: The STROBE-MR Statement[J]. Jama, 2021, 326(16): 1614-1621.
[19] Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression[J]. Int J Epidemiol, 2015, 44(2): 512-25.
[20] Burgess S, Thompson S G. Interpreting findings from Mendelian randomization using the MR-Egger method[J]. Eur J Epidemiol, 2017, 32(5): 377-389.
[21] Slob E a W, Burgess S. A comparison of robust Mendelian randomization methods using summary data[J]. Genet Epidemiol, 2020, 44(4): 313-329.
[22] Greco M F, Minelli C, Sheehan N A, et al. Detecting pleiotropy in Mendelian randomisation studies with summary data and a continuous outcome[J]. Stat Med, 2015, 34(21): 2926-40.
[23] Bowden J, Davey Smith G, Haycock P C, et al. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator[J]. Genet Epidemiol, 2016, 40(4): 304-14.
[24] Verbanck M, Chen C Y, Neale B, et al. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases[J]. Nat Genet, 2018, 50(5): 693-698.
[25] Davis A P, Wiegers T C, Sciaky D, et al. Comparative toxicogenomics database's 20th anniversary: update 2025[J]. Nucleic Acids Res, 2024.
[26] Liu Y, Zhao W, Hu W, et al. Exploring the relationship between anal fistula and colorectal cancer based on Mendelian randomization and bioinformatics[J]. J Cell Mol Med, 2024, 28(15): e18537.
[27] Kanehisa M, Furumichi M, Sato Y, et al. KEGG: biological systems database as a model of the real world[J]. Nucleic Acids Res, 2025, 53(D1): D672-d677.
[28] Martin C R, Osadchiy V, Kalani A, et al. The Brain-Gut-Microbiome Axis[J]. Cell Mol Gastroenterol Hepatol, 2018, 6(2): 133-148.
[29] Neroni B, Evangelisti M, Radocchia G, et al. Relationship between sleep disorders and gut dysbiosis: what affects what?[J]. Sleep Med, 2021, 87: 1-7.
[30] Xie F, Feng Z, Xu B. Metabolic Characteristics of Gut Microbiota and Insomnia: Evidence from a Mendelian Randomization Analysis[J]. Nutrients, 2024, 16(17).
[31] Grosicki G J, Riemann B L, Flatt A A, et al. Self-reported sleep quality is associated with gut microbiome composition in young, healthy individuals: a pilot study[J]. Sleep Med, 2020, 73: 76-81.
[32] Poroyko V A, Carreras A, Khalyfa A, et al. Chronic Sleep Disruption Alters Gut Microbiota, Induces Systemic and Adipose Tissue Inflammation and Insulin Resistance in Mice[J]. Sci Rep, 2016, 6: 35405.
[33] Maki K A, Burke L A, Calik M W, et al. Sleep fragmentation increases blood pressure and is associated with alterations in the gut microbiome and fecal metabolome in rats[J]. Physiol Genomics, 2020, 52(7): 280-292.

[34] Zeng H, Xu J, Zheng L, et al. Traditional Chinese herbal formulas modulate gut microbiome and improve insomnia in patients with distinct syndrome types: insights from an interventional clinical study[J]. Front Cell Infect Microbiol, 2024, 14: 1395267.
[35] Liu J L, Xu X, Rixiati Y, et al. Dysfunctional circadian clock accelerates cancer metastasis by intestinal microbiota triggering accumulation of myeloid-derived suppressor cells[J]. Cell Metab, 2024, 36(6): 1320-1334.e9.
[36] Wang H, Qin X, Gui Z, et al. The effect of Bailemian on neurotransmitters and gut microbiota in p-chlorophenylalanine induced insomnia mice[J]. Microb Pathog, 2020, 148: 104474.
[37] Wu C, Dou J, Song X, et al. Gut microbiota: a new target for the prevention and treatment of insomnia using Chinese herbal medicines and their active components[J]. Front Pharmacol, 2025, 16: 1572007.
[38] Yue S, He T, Li B, et al. Effectiveness of Yi-Zhi-An-Shen granules on cognition and sleep quality in older adults with amnestic mild cognitive impairment: protocol for a randomized, double-blind, placebo-controlled trial[J]. Trials, 2019, 20(1): 518.
[39] Si Y, Wei W, Chen X, et al. A comprehensive study on the relieving effect of Lilium brownii on the intestinal flora and metabolic disorder in p-chlorphenylalanine induced insomnia rats[J]. Pharm Biol, 2022, 60(1): 131-143.
[40] Chen H, Shen J, Li H, et al. Ginsenoside Rb1 exerts neuroprotective effects through regulation of Lactobacillus helveticus abundance and GABA(A) receptor expression[J]. J Ginseng Res, 2020, 44(1): 86-95.
[41] Shao J, Zheng X, Qu L, et al. Ginsenoside Rg5/Rk1 ameliorated sleep via regulating the GABAergic/serotoninergic signaling pathway in a rodent model[J]. Food Funct, 2020, 11(2): 1245-1257.
[42] Zhang D D, Li H J, Zhang H R, et al. Poria cocos water-soluble polysaccharide modulates anxiety-like behavior induced by sleep deprivation by regulating the gut dysbiosis, metabolic disorders and TNF-α/NF-κB signaling pathway[J]. Food Funct, 2022, 13(12): 6648-6664.
[43] Zhong Y, Zheng Q, Hu P, et al. Sedative and hypnotic effects of compound Anshen essential oil inhalation for insomnia[J]. BMC Complement Altern Med, 2019, 19(1): 306.
[44] Yan G, Li F, Tao Z, et al. Effects of Vestibular Damage on the Sleep and Expression Level of Orexin in the Hypothalamus of Rats and Its Correlation with Autophagy and Akt Tumor Signal Pathway[J]. J Oncol, 2022, 2022: 2514555.

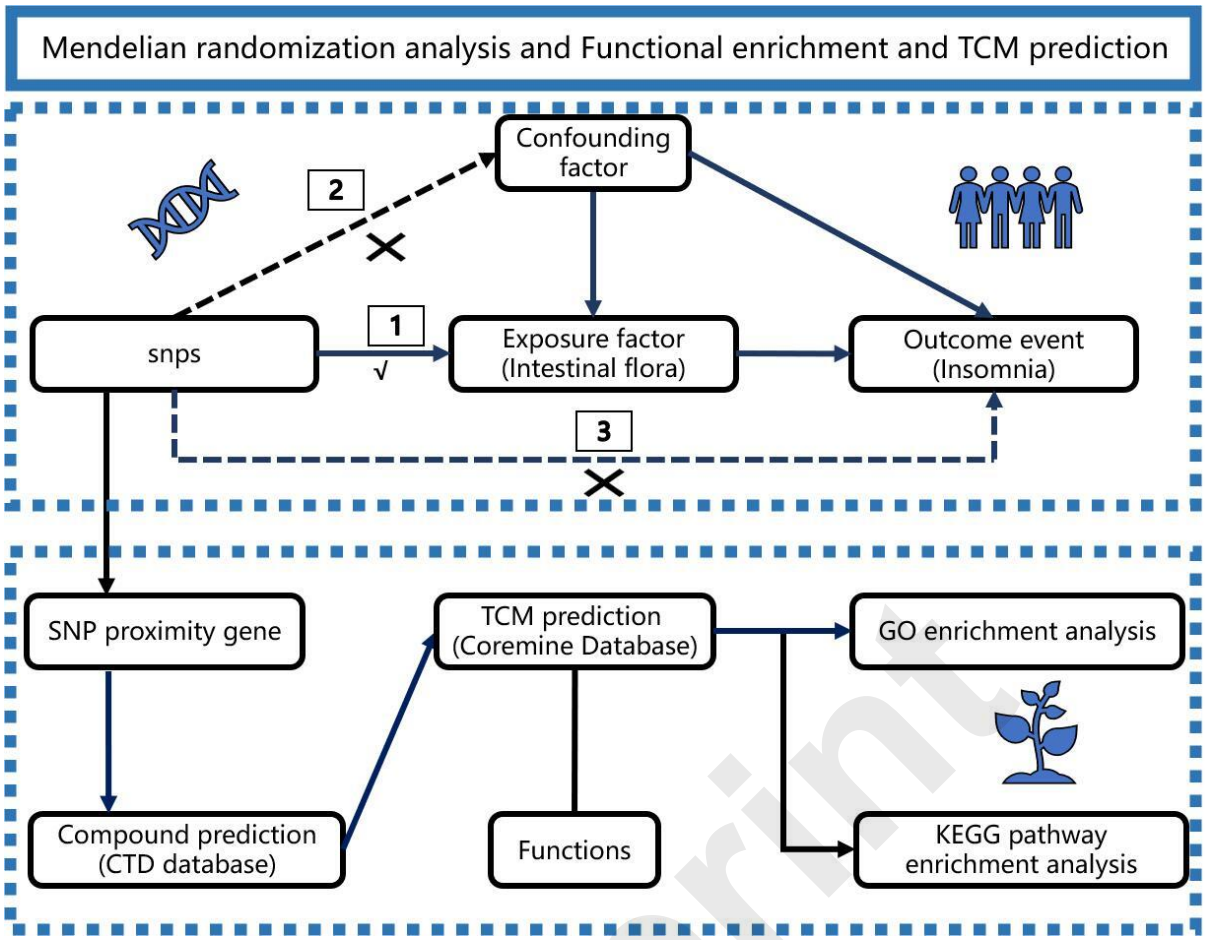

Mendelian randomization analysis and Functional enrichment and TCM prediction
Confounding factor
2
snps
1
√
Exposure factor (Intestinal flora)
Outcome event (Insomnia)
3
SNP proximity gene
TCM prediction (Coremine Database)
GO enrichment analysis
Compound prediction (CTD database)
Functions
KEGG pathway enrichment analysis

Table 1 Positive results of MR Analysis of GM and Insomnia

Exposure	Method	nSNP	P	OR	or_lci95	or_uci95
Pasteurellaceae	IVW	14	0.030	0.764	0.599	0.975
Pasteurellaceae	MR Egger	14	0.059	0.568	0.334	0.966
Pasteurellaceae	Weighted median	14	0.138	0.763	0.534	1.091
Pasteurellaceae	Weighted mode	14	0.309	0.767	0.470	1.253
Pasteurellaceae	Simple mode	14	0.584	0.846	0.472	1.516
Ruminococcaceae	IVW	10	0.020	1.578	1.074	2.317
Ruminococcaceae	Weighted median	10	0.175	1.433	0.852	2.407
Ruminococcaceae	Weighted mode	10	0.302	1.385	0.773	2.484
Ruminococcaceae	Simple mode	10	0.456	1.339	0.643	2.789
Ruminococcaceae	MR Egger	10	0.814	1.107	0.488	2.508
Marvinbryantia	IVW	10	0.023	1.537	1.062	2.225
Marvinbryantia	Weighted median	10	0.063	1.570	0.976	2.528
Marvinbryantia	Weighted mode	10	0.142	1.697	0.890	3.235
Marvinbryantia	Simple mode	10	0.144	1.638	0.895	2.998
Marvinbryantia	MR Egger	10	0.845	1.167	0.260	5.234
Olsenella	IVW	11	0.014	0.781	0.641	0.951
Olsenella	Weighted median	11	0.021	0.730	0.558	0.954
Olsenella	Simple mode	11	0.065	0.658	0.443	0.977
Olsenella	Weighted mode	11	0.073	0.685	0.473	0.991
Olsenella	MR Egger	11	0.115	0.547	0.278	1.077
Ruminococcus gnavus group	IVW	12	0.016	0.746	0.588	0.946
Ruminococcus gnavus group	Weighted median	12	0.291	0.842	0.612	1.158
Ruminococcus gnavus group	MR Egger	12	0.329	0.532	0.159	1.776
Ruminococcus gnavus group	Weighted mode	12	0.699	0.897	0.526	1.532
Ruminococcus gnavus group	Simple mode	12	0.701	0.893	0.508	1.570
Mollicutes RF9	IVW	13	0.021	0.706	0.525	0.949
Mollicutes RF9	Weighted median	13	0.029	0.627	0.412	0.954
Mollicutes RF9	MR Egger	13	0.043	0.327	0.126	0.851
Mollicutes RF9	Weighted mode	13	0.091	0.543	0.282	1.042
Mollicutes RF9	Simple mode	13	0.114	0.550	0.277	1.093
Pasteurellales	IVW	14	0.030	0.764	0.599	0.975
Pasteurellales	MR Egger	14	0.059	0.568	0.334	0.966
Pasteurellales	Weighted median	14	0.131	0.763	0.537	1.084
Pasteurellales	Weighted mode	14	0.300	0.767	0.474	1.241
Pasteurellales	Simple mode	14	0.603	0.846	0.457	1.565

Table 2 Quality control results of GM with causal relationship with Insomnia

Exposure	heterogeneity			pleiotropy_test		MR-PRESSO	
	Method	Q	P	Egger	P	global.test	P
Pasteurellaceae	IVW	12.453	0.491	0.039	0.342	14.526	0.517
Ruminococcaceae	IVW	3.279	0.952	0.038	0.364	3.882	0.967
Marvinbryantia	IVW	5.807	0.759	0.024	0.720	6.723	0.802
Olsenella	IVW	7.146	0.712	0.049	0.309	8.637	0.724
Ruminococcusgnavus group	IVW	11.006	0.443	0.037	0.587	12.988	0.464
Mollicutes RF9	IVW	11.189	0.513	0.065	0.125	13.387	0.538
Pasteurellales	IVW	12.453	0.491	0.039	0.242	14.526	0.522

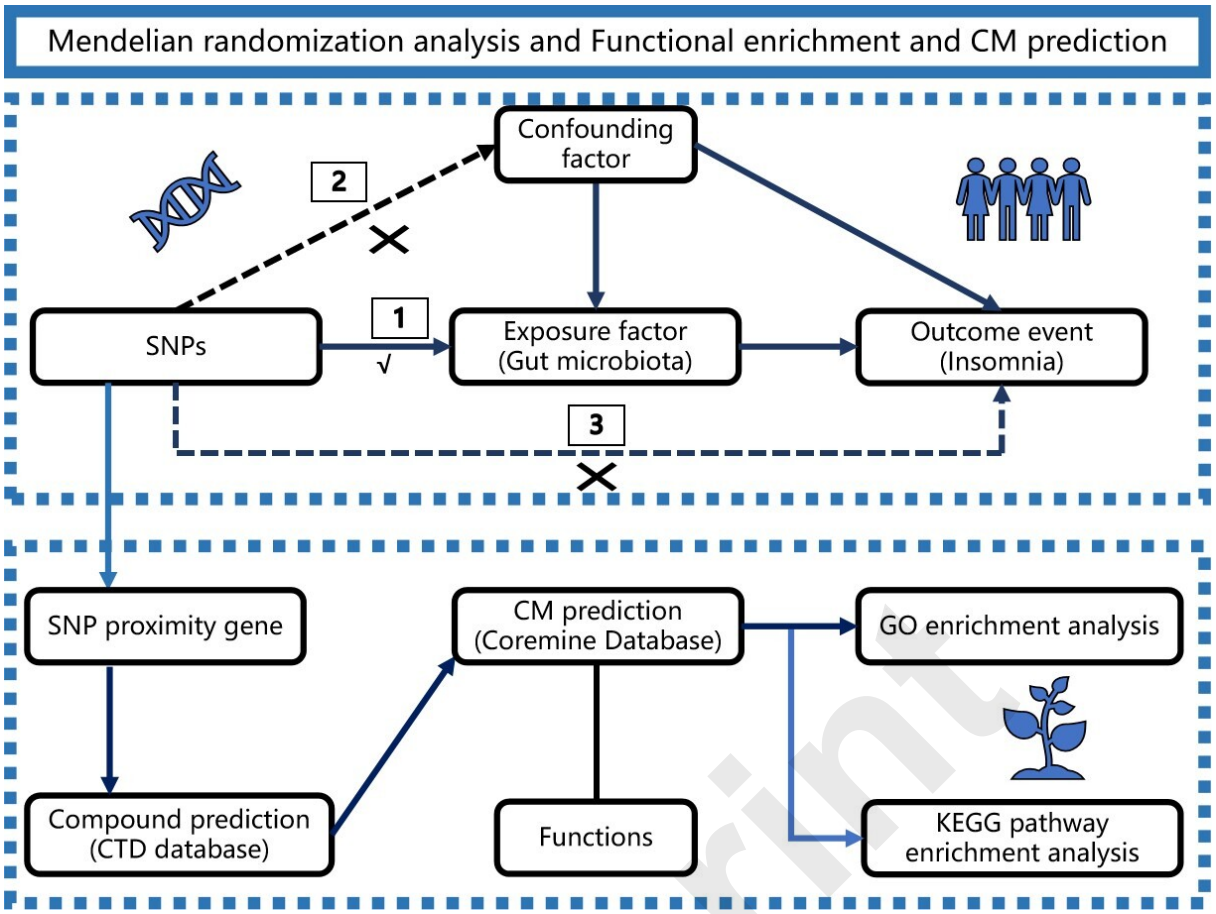
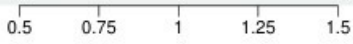
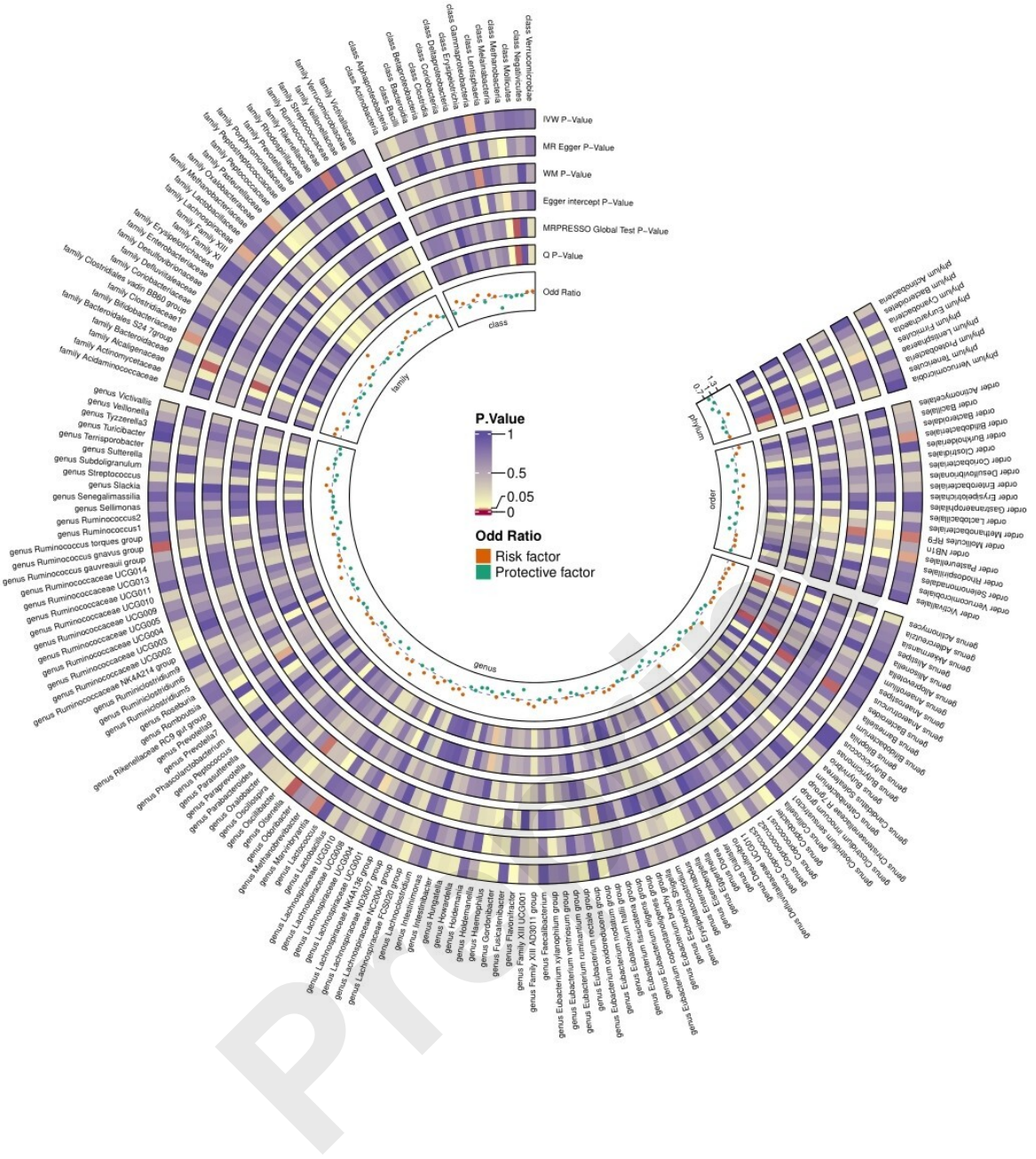
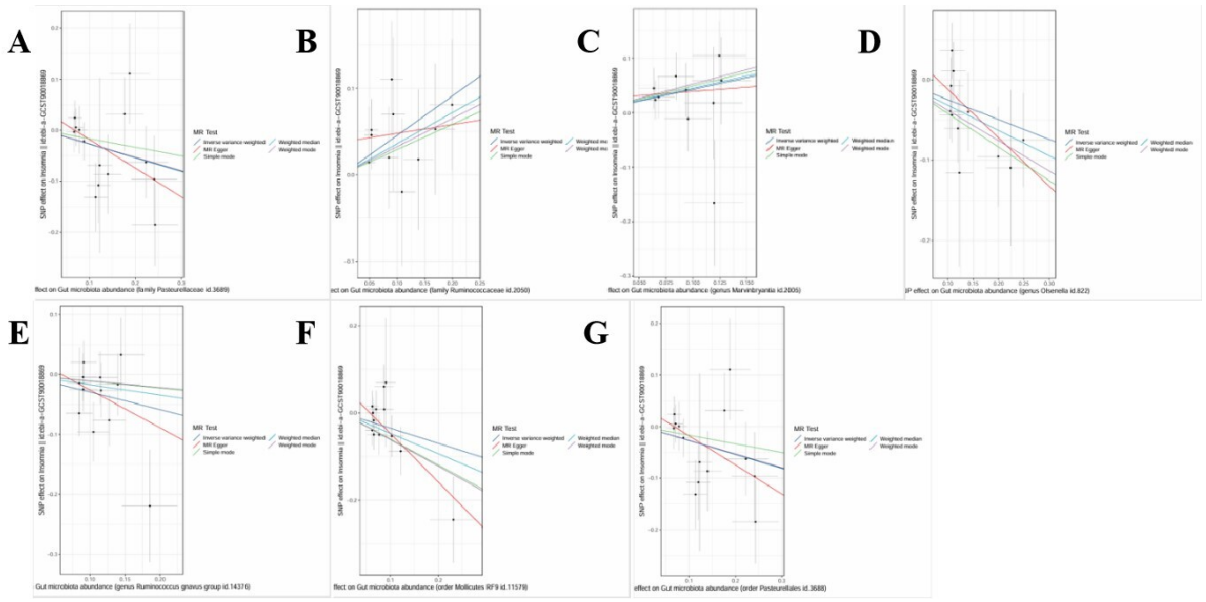


Figure 1 Study design of MR analysis and Functional enrichment and TCM prediction

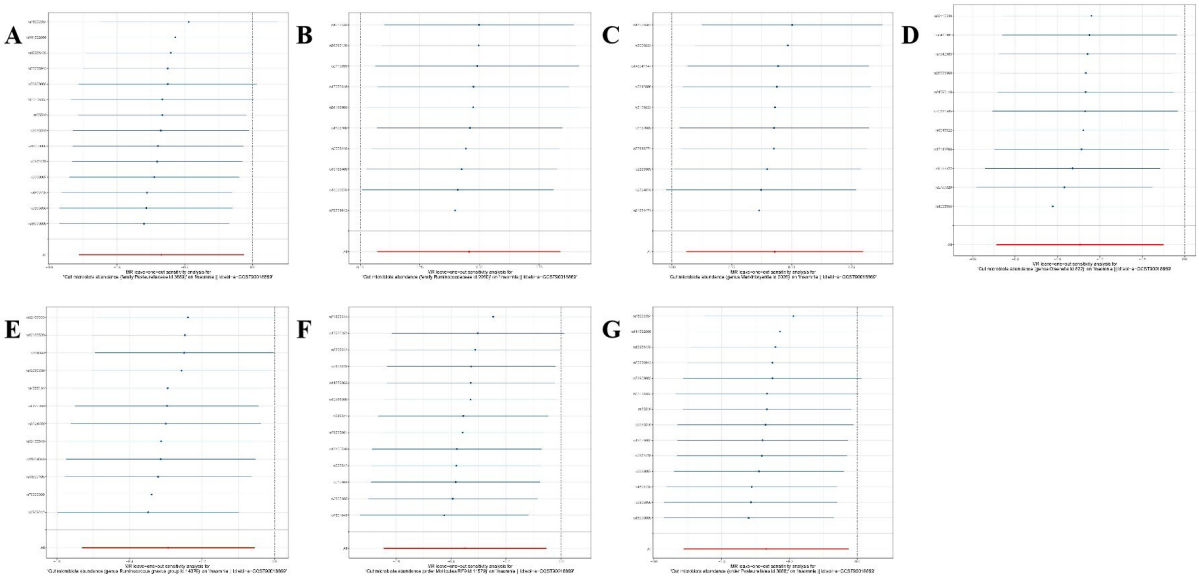
Trait	Method	nSNP	*P*− Value	OR (95% CI)
family Pasteurellaceae	MR Egger	14	0.059	0.568 (0.334 – 0.966)
family Pasteurellaceae	Weighted median	14	0.138	0.763 (0.534 – 1.091)
family Pasteurellaceae	IVW	14	0.030	0.764 (0.599 – 0.975)
family Pasteurellaceae	Simple mode	14	0.584	0.846 (0.472 – 1.516)
family Pasteurellaceae	Weighted mode	14	0.309	0.767 (0.470 – 1.253)
family Ruminococcaceae	MR Egger	10	0.814	1.107 (0.488 – 2.508)
family Ruminococcaceae	Weighted median	10	0.175	1.433 (0.852 – 2.407)
family Ruminococcaceae	IVW	10	0.020	1.578 (1.074 – 2.317)
family Ruminococcaceae	Simple mode	10	0.456	1.339 (0.643 – 2.789)
family Ruminococcaceae	Weighted mode	10	0.302	1.385 (0.773 – 2.484)
genus Marvinbryantia	MR Egger	10	0.845	1.167 (0.260 – 5.234)
genus Marvinbryantia	Weighted median	10	0.063	1.570 (0.976 – 2.528)
genus Marvinbryantia	IVW	10	0.023	1.537 (1.062 – 2.225)
genus Marvinbryantia	Simple mode	10	0.144	1.638 (0.895 – 2.998)
genus Marvinbryantia	Weighted mode	10	0.142	1.697 (0.890 – 3.235)
genus Olsenella	MR Egger	11	0.115	0.547 (0.278 – 1.077)
genus Olsenella	Weighted median	11	0.021	0.730 (0.558 – 0.954)
genus Olsenella	IVW	11	0.014	0.781 (0.641 – 0.951)
genus Olsenella	Simple mode	11	0.065	0.658 (0.443 – 0.977)
genus Olsenella	Weighted mode	11	0.073	0.685 (0.473 – 0.991)
genus Ruminococcus gnavus group	MR Egger	12	0.329	0.532 (0.159 – 1.776)
genus Ruminococcus gnavus group	Weighted median	12	0.291	0.842 (0.612 – 1.158)
genus Ruminococcus gnavus group	IVW	12	0.016	0.746 (0.588 – 0.946)
genus Ruminococcus gnavus group	Simple mode	12	0.701	0.893 (0.508 – 1.570)
genus Ruminococcus gnavus group	Weighted mode	12	0.699	0.897 (0.526 – 1.532)
order Mollicutes RF9	MR Egger	13	0.043	0.327 (0.126 – 0.851)
order Mollicutes RF9	Weighted median	13	0.029	0.627 (0.412 – 0.954)
order Mollicutes RF9	IVW	13	0.021	0.706 (0.525 – 0.949)
order Mollicutes RF9	Simple mode	13	0.114	0.550 (0.277 – 1.093)
order Mollicutes RF9	Weighted mode	13	0.091	0.543 (0.282 – 1.042)
order Pasteurellales	MR Egger	14	0.059	0.568 (0.334 – 0.966)
order Pasteurellales	Weighted median	14	0.131	0.763 (0.537 – 1.084)
order Pasteurellales	IVW	14	0.030	0.764 (0.599 – 0.975)
order Pasteurellales	Simple mode	14	0.603	0.846 (0.457 – 1.565)
order Pasteurellales	Weighted mode	14	0.300	0.767 (0.474 – 1.241)



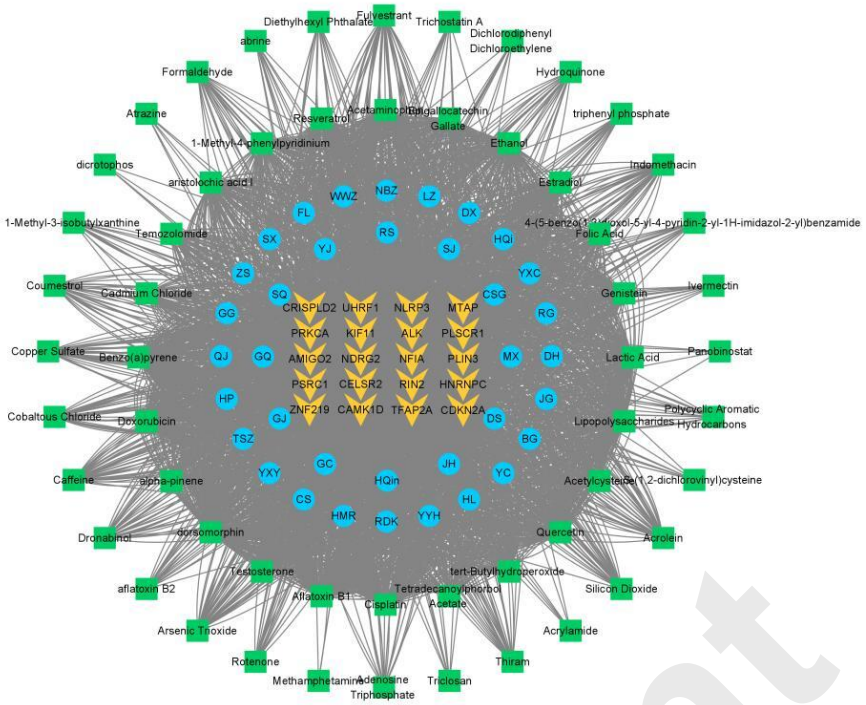

IVW P−Value
MR Egger P−Value
WM P−Value
Egger intercept P−Value
MRPRESSO Global Test P−Value
Q P−Value
Odd Ratio
class
family
genus
order
phylum
P.Value
1
0.5
0.05
0
Odd Ratio
Risk factor
Protective factor

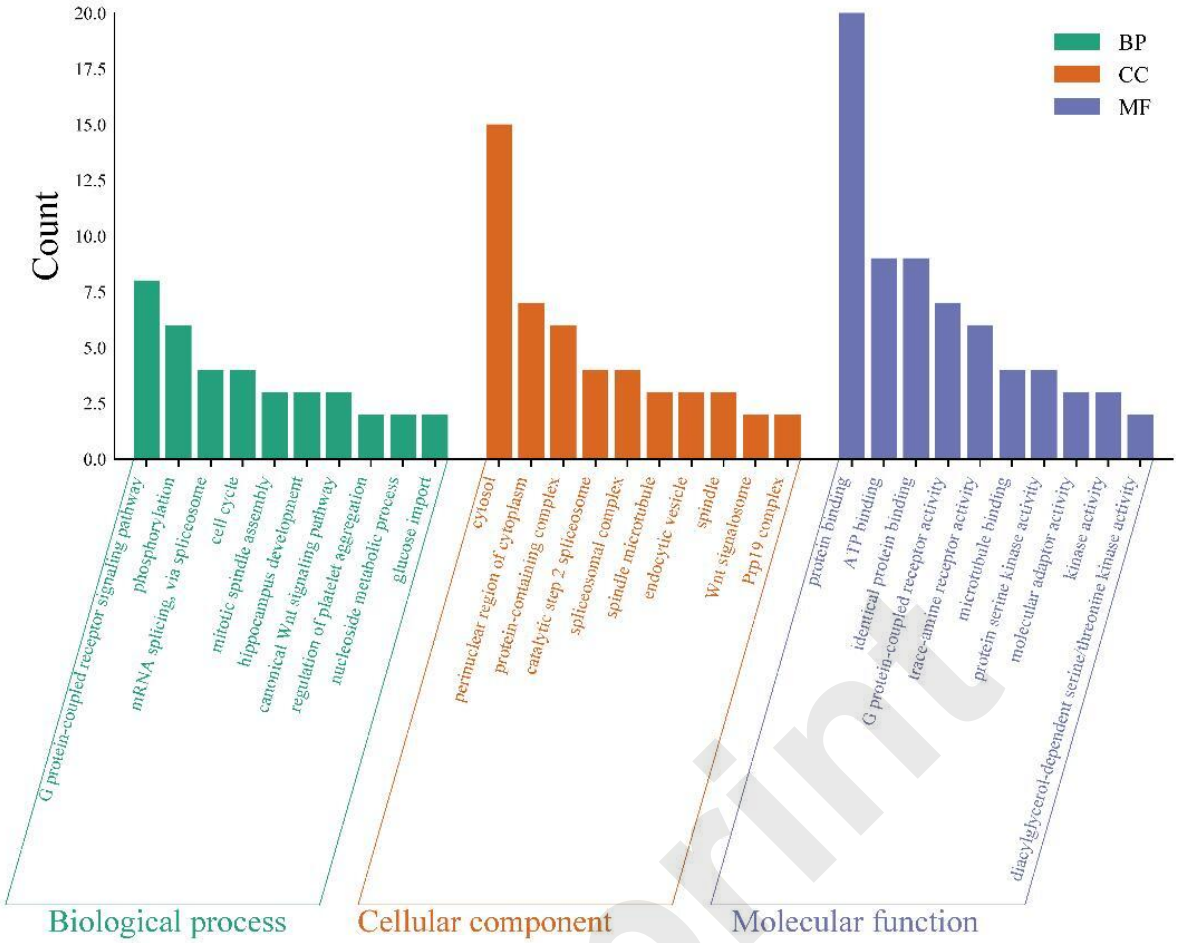


Scatter plot of GM and insomnia

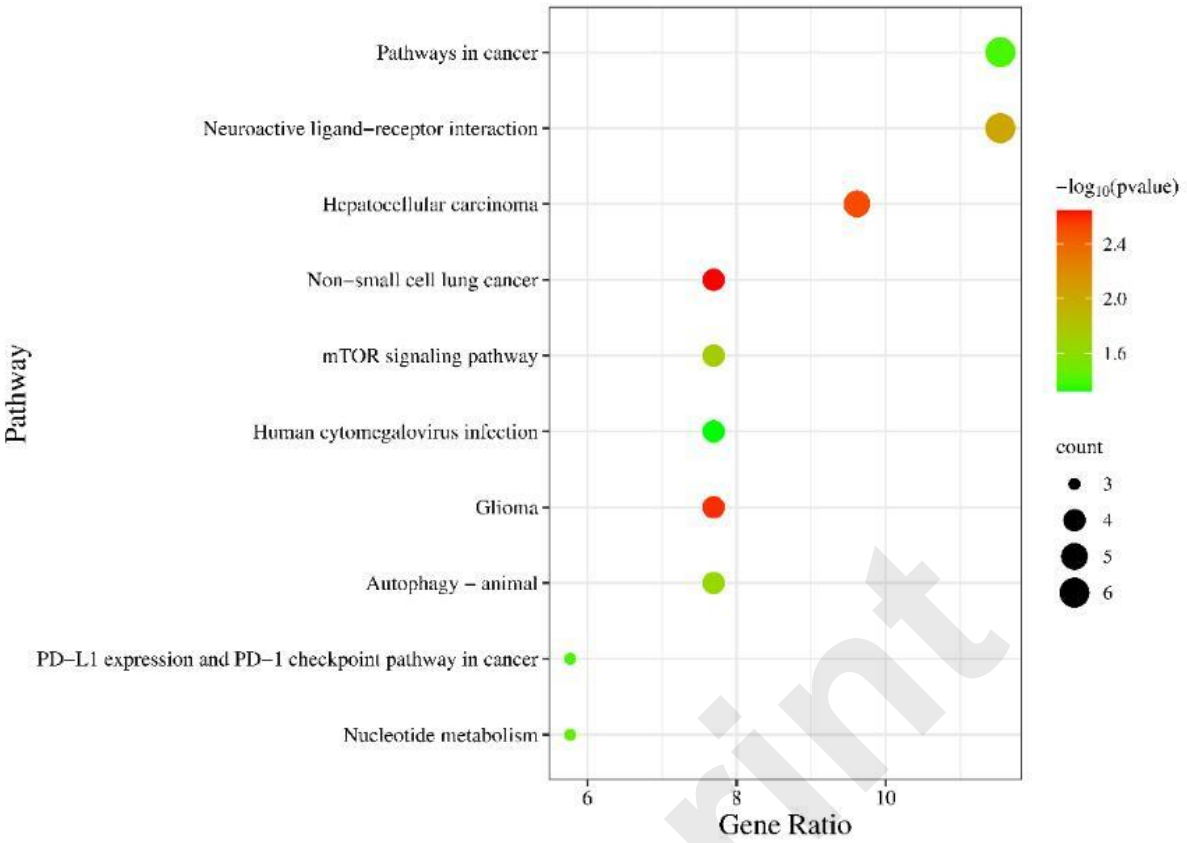


Leave-one-out analysis plot of GM and insomnia


Fulvestrant
Diethylhexyl Phthalate
Trichostatin A
abrine
Dichlorodiphenyl Dichloroethylene
Formaldehyde
Hydroquinone
Atrazine
Resveratrol
Acetaminophen
Epigallocatechin Gallate
triphenyl phosphate
1-Methyl-4-phenylpyridinium
Ethanol
dicrotophos
Indomethacin
aristolochic acid I
Estradiol
1-Methyl-3-isobutylxanthine
Temozolomide
4-(5-benzo[1,3]dioxol-5-yl-4-pyridin-2-yl-1H-imidazol-2-yl)benzamide
Folic Acid
Coumestrol
Cadmium Chloride
Genistein
Ivermectin
Copper Sulfate
Benzo(a)pyrene
Panobinostat
Lactic Acid
Cobaltous Chloride
Doxorubicin
Polycyclic Aromatic Hydrocarbons
Lipopolysaccharides
Caffeine
alpha-pinene
S-(1,2-dichlorovinyl)cysteine
Acetylcysteine
Dronabinol
dorsomorphin
Quercetin
Acrolein
Testosterone
tert-Butylhydroperoxide
aflatoxin B2
Aflatoxin B1
Tetradecanoylphorbol Acetate
Silicon Dioxide
Cisplatin
Arsenic Trioxide
Acrylamide
Rotenone
Thiram
Methamphetamine
Adenosine Triphosphate
Triclosan
NBZ
WWZ
LZ
FL
DX
RS
SX
YJ
SJ
HQi
ZS
YXC
SQ
CSG
GG
RG
QJ
GQ
MX
DH
HP
JG
GJ
DS
TSZ
BG
GC
JH
YXY
HQin
YC
CS
HL
HMR
RDK
YYH
CRISPLD2
UHRF1
NLRP3
MTAP
PRKCA
KIF11
ALK
PLSCR1
AMIGO2
NDRG2
NFIA
PLIN3
PSRC1
CELSR2
RIN2
HNRNPC
ZNF219
CAMK1D
TFAP2A
CDKN2A



Functional enrichment analysis1



Functional enrichment analysis2