frontiers | Research Topics

Mendelian randomization: the study of causal relationships between tumors and underlying pathogenic factors

Edited by
Ziheng Wang and Guangsheng Pei

Published in
Frontiers in Genetics

ISSN 1664-8714
ISBN 978-2-8325-6348-9
DOI 10.3389/978-2-8325-6348-9

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Mendelian randomization: the study of causal relationships between tumors and underlying pathogenic factors

Topic editors

Ziheng Wang — University of Macau, China

Guangsheng Pei — University of Texas MD Anderson Cancer Center, United States

Citation

Wang, Z., Pei, G., eds. (2025). *Mendelian randomization: the study of causal relationships between tumors and underlying pathogenic factors*.

Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-6348-9

Table of contents

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 23 November 2023
DOI 10.3389/fgene.2023.1285274

OPEN ACCESS

EDITED BY
Xuanye Cao,
University of Texas MD Anderson Cancer Center, United States

REVIEWED BY
Xiang Zhang,
Bot Auto, United States
Wen Jiang,
AstraZeneca, United States

*CORRESPONDENCE
Haijun Wang,
wnavy200299@163.com
Wei Su,
weisu_xxmu@163.com

†These authors have contributed equally to this work

RECEIVED 29 August 2023
ACCEPTED 07 November 2023
PUBLISHED 23 November 2023

CITATION
Cui K, Song N, Fan Y, Zeng L, Shi P, Wang Z, Su W and Wang H (2023), A two-sample Mendelian randomization analysis: causal association between chemokines and pan-carcinoma. *Front. Genet.* 14:1285274. doi: 10.3389/fgene.2023.1285274

A two-sample Mendelian randomization analysis: causal association between chemokines and pan-carcinoma

Kai Cui[1,2†], Na Song[1,2†], Yanwu Fan[1], Liqun Zeng[1], Pingyu Shi[1], Ziwei Wang, Wei Su[2*] and Haijun Wang[1,2*]

[1]Department of Pathology, Xinxiang Medical University, Xinxiang, China, [2]Department of Pathology, Xinxiang Key Laboratory of Tumor Precision Medicine, The First Affiliated Hospital of Xinxiang Medical University, Xinxiang, China

Objective: According to the 2020 data from the World Health Organization (WHO), cancers stand as one of the foremost contributors to global mortality. Revealing novel cancer risk factors and protective factors is of paramount importance in the prevention of disease occurrence. Studies on the relationship between chemokines and cancer are ongoing; however, due to the coordination of multiple potential mechanisms, the specific causal association remains unclear.

Methods: We performed a bidirectional Mendelian randomization analysis to explore the causal association between serum chemokines and pan-carcinoma. All data is from the GWAS catalog and IEU Open GWAS database. The inverse-variance weighted (IVW) method is primarily employed for assessing the statistical significance of the findings. In addition, the significance threshold after the multiple hypothesis test (Bonferroni) was 0.0013, and the evidence of a potential association was considered if the p-value < 0.05, but remained greater than Bonferroni's threshold.

Results: The results indicate that CCL1 (odds ratio, OR = 1.18), CCL2 (OR = 1.04), CCL8 (OR = 1.36), CCL14 (Colorectal, OR = 1.08, Small intestine, OR = 0.77, Lung, OR = 1.11), CCL15 (OR = 0.85), CCL18 (Breast, OR = 0.95, Prostate, OR = 0.96), CCL19 (Lung, OR = 0.66, Prostate, OR = 0.92), CCL20 (Lung, OR = 0.53, Thyroid, OR = 0.76), CCL21 (OR = 0.62), CCL22 (OR = 2.05), CCL23 (OR = 1.31), CCL24 (OR = 1.06), CCL27 (OR = 1.49), CCL28 (OR = 0.74), CXCL5 (OR = 0.95), CXCL9 (OR = 3.60), CXCL12 (Breast, OR = 0.87, Small intestine, OR = 0.58), CXCL13 (Breast, OR = 0.93, Lung, OR = 1.29), CXCL14 (Colon, OR = 1.40) and CXCL17 (OR = 1.07) are potential risk factors for cancers. In addition, there was a reverse causal association between CCL1 (OR = 0.94) and CCL18 (OR = 0.94) and breast cancer. Sensitivity analysis results were similar. The results of the other four MR Methods were consistent with the main results, and the leave-one-out method showed that the results were not driven by a Single nucleotide polymorphism (SNP). Moreover, there was no heterogeneity and pleiotropy in our analysis.

Conclusion: Based on the two-sample MR Analysis method, we found that chemokines might be upstream factors of cancer pathogenesis. These results might provide new insights into the future use of chemokines as potential targets

for cancer prevention and treatment. Our results also provide important clues for tumor prevention, and changes of serum chemokine concentration may be recognized as one of the features of precancerous lesions in future clinical trials.

KEYWORDS

causal association, pan-carcinoma, GWAS, Mendelian randomization, chemokine

Introduction

A substantial number of new cancer cases are diagnosed annually, and most of them die from the disease. A significant proportion of cancer patients, such as those with pancreatic cancer, were diagnosed at an advanced stage due to a poor prognosis, high mortality rates, and rapid disease progression (Halbrook et al., 2023). Fortunately, due to the progress and improvement of treatment methods, there has been a significant reduction in the incidence of cervical cancer among vaccinated women. Similarly, advancements in immunotherapy and targeted therapy have led to a significant reduction in mortality rates for melanoma, kidney cancer, and other types of cancer. However, the incidence of breast, uterine, and prostate cancers continues to exhibit an upward trend year after year (Siegel et al., 2023). In order to reduce the incidence of cancer, the discovery of risk factors in precancerous lesions is particularly important. So far, prospective studies have identified several factors that can interfere with cancer risk (Kliemann et al., 2023; Lagou and Karagiannis, 2023; Wang et al., 2023). For example, processed food intake and obesity can influence changes in a range of cancer risk indicators. In addition, a meta-analysis investigated the complexity of aging and cancer risk (López-Otín et al., 2023). A growing number of factors are proving to be associated with cancer risk. The discovery of risk factors may provide potential value for cancer prevention.

In recent years, more and more studies have confirmed the potential value of chemokines for cancer progression and treatment (Märkl et al., 2022; Propper and Balkwill, 2022). Chemokines are a class of cytokines that transport immune cells and are associated with lymphoid tissue (Schulz et al., 2016; Cambier et al., 2023). In cancer, however, they promoted the migration of immunosuppressive cells, such as Tregs, M2 macrophages, and so on (Moreno Ayala et al., 2023; Zhou et al., 2023). Furthermore, chemokines promoted cancer progression by mediating tumor-related pathways such as PI3K/AKT and ERK1/2 (Zhao et al., 2017). However, it should be noted that not all chemokines are implicated in tumor progression; indeed, certain chemokines exhibit anti-tumor effects (Korbecki et al., 2020). Some studies had found that high-expression chemokines are more sensitive to cancer immunotherapy (Limagne et al., 2022). Non-small cell lung cancers with high CXCL10 expression had a better response when treated with immune checkpoint suppression. In addition, the chemokine CXCL10 recruited $CD4^+$ and $CD8^+$ T cells to the tumor via $CCR6^+$ type 3 innate lymphoid cells (Bruchard et al., 2022). Surprisingly, CXCL10 also promoted tumor cell migration in mouse models (Hirth et al., 2020), and CXCL10 secreted by mesenchymal stem cells promoted tumor growth (Timaner et al., 2018). In addition, the relationship between other chemokines and tumors is particularly complex. Curiously, if there is a causal association between chemokines and cancers. Although several meta-analyses had been conducted to explore causal associations between chemokines and cancer (Cho and Kim, 2013; Liu et al., 2018), there had not been a systematic comprehensive study.

The above studies are fuzzy about the association between chemokines and cancers, which may be influenced by environmental and other factors. Therefore, it is necessary to conduct a Mendelian randomization (MR) study between chemokines and tumors. MR uses genetic variation as instrumental variables (IVs) to measure potential causal associations between exposures and outcomes (Cheng et al., 2022). Single nucleotide polymorphisms (SNPs) were obtained from genome-wide association studies. The advantage of MR is to establish a causal association between exposures and outcomes from a genetic perspective, excluding other external environmental and confounding factors (Timaner et al., 2018). So, the association between chemokines and the risk of 14 types of malignancies were evaluated using two-sample MR Analysis in our study.

Materials and methods

Study design

We conducted a two-sample MR Analysis between cancers and chemokines using publicly available online data. Including GWAS Catalog (https://www.ebi.ac.uk/gwas/) and IEU OpenGWAS (https://gwas.mrcieu.ac.uk/) (Buniello et al., 2019). These databases have received ethical approval and informed consent, so no additional instructions are required. Three preconditions must be met when performing MR analysis (Guyatt et al., 2023). First, the association hypothesis: IVs must be strongly associated with chemokines, and F-value is considered as measure indicator of association. Second, the independence hypothesis: IVs and confounding factors were independent of each other. In short, chemokines IVs were not associated with other factors that had a causal association with the tumors. Third, the exclusivity hypothesis: IVs influence tumors only through chemokines (Figure 1).

Exposure and outcome data

Chemokines data came from a study on serum proteins in the GWAS catalog. To explore associations between genetic variants and serum proteins, Gudjonsson et al. (2022) conducted a GWAS study involving 5,368 European individuals. We downloaded 38 serum chemokine protein-associated SNPs from the study as exposure factors.

The outcome factors were 14 cancers, including breast cancer, lung cancer, gastrointestinal cancer and some other site-specific tumors. IVs for Breast Cancer were derived from the Breast Cancer Association Consortium [BCAC (Oncoarray, $N = 106,776$) (iCOGS,

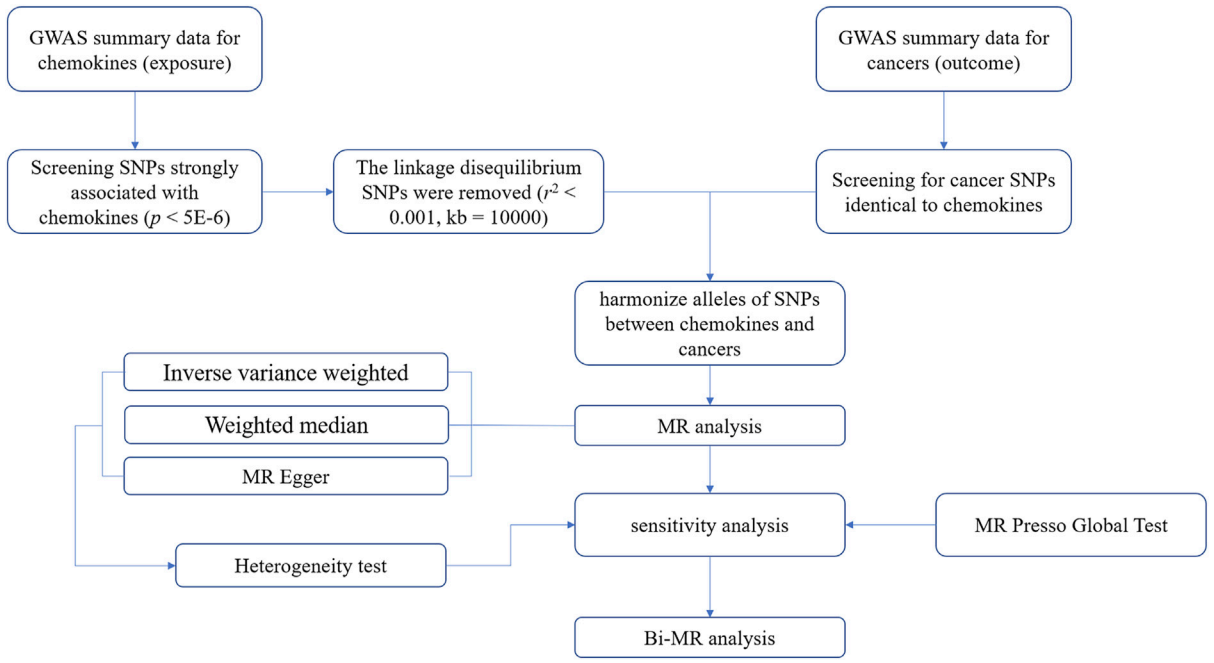


FIGURE 1
A flowchart for analyzing causal associations between chemokines and tumors based on Mendelian randomization (MR).

N = 89,677), FinnGen database (N = 123,579) (Michailidou et al., 2017; Kurki et al., 2023). IVs for prostate cancer were obtained from Schumacher's GWAS data (N = 140,254) (Schumacher et al., 2018). Malignant neoplasm of ovary (N = 123,579) and all other tumor IVs (N = 218,792) were derived from the FinnGen database. All sources of tumor GWAS information are provided in the Supplementary Table S1.

Instrumental variable selection

First, the selection of instrumental variables cannot violate the first hypothesis of Mendelian randomization, so we used the threshold of p < 5E-8 to screen the IVs strongly related to serum chemokines. However, some chemokines did not have SNPs with this threshold, and then the threshold of significance was eased to p < 5E-6 (Luo et al., 2022; Yu et al., 2023). And the SNPs with F values less than 10 were excluded (Luo et al., 2021). SNPs with F statistic >10 are considered to be strongly associated with exposure. Secondly, there may be linkage disequilibrium (LD) between SNPs. The LD phenomenon implies non-random transmission of different alleles to offspring, and it is crucial to maintain LD between various SNPs prior to conducting MR analysis (Yarmolinsky et al., 2023). To eliminate LD, the TwoSample MR package was employed in this study with specific parameters set as r^2 = 0.001 and kb = 10,000. The variable r^2 represents the association of LD between SNPs, while kb represents the region range of LD between SNPs. Third, information about the SNPs in the outcome was matched according to the SNPs screened during exposure. In this process, in the absence of SNP information, substitute proxy SNPs are not utilized. Finally, SNPs with palindromic structure were removed.

MR analysis

To determine the causal association between serum chemokines and cancers, a two-sample MR Analysis was performed. A total of three common MR Analysis methods have been used, including inverse-variance weighted (IVW) (Huang et al., 2022), MR-Egger regression (Wu et al., 2020), weighted median (Li et al., 2022), weighted mode and simple mode methods are supplemented. According to the survey, the IVW test exhibits superior advantages compared to additional methods (Lin et al., 2021). And it has been used as the primary MR Analysis method in most studies (Yang M. et al., 2023; Yang Y. et al., 2023; Ding et al., 2023; Li et al., 2023). Similarly, IVW was used as the main test method in our study, while other methods were used as references. In addition, the MR-Egger regression test and MR-Presso were used to verify the existence of horizontal pleiotropy, and p-value < 0.05 is considered to be horizontal pleiotropy. To ensure the validity of our findings, we conducted leave-one-out sensitivity analysis to ascertain whether a single SNP is responsible for driving the results. Based on the causal relationship between 38 chemokines and cancer, the more conservative Bonferroni method was used to correct for significance results. Before correction, p < 0.05 was a significant result, and after correction, p < 0.0013 was a significant result. Results with p < 0.05 but higher than 0.0013 were considered for potential causal associations (Sedgwick, 2014; Larsson et al., 2017). All statistical tests were performed in two-sample MR and MR-PRESSO packages. Moreover, Heterogeneity test results were significant (p < 0.05), which was considered to be heterogeneity among IVs.

To investigate the bidirectional causal relationship between cancer and chemokines, we performed a bi-directional Mendelian

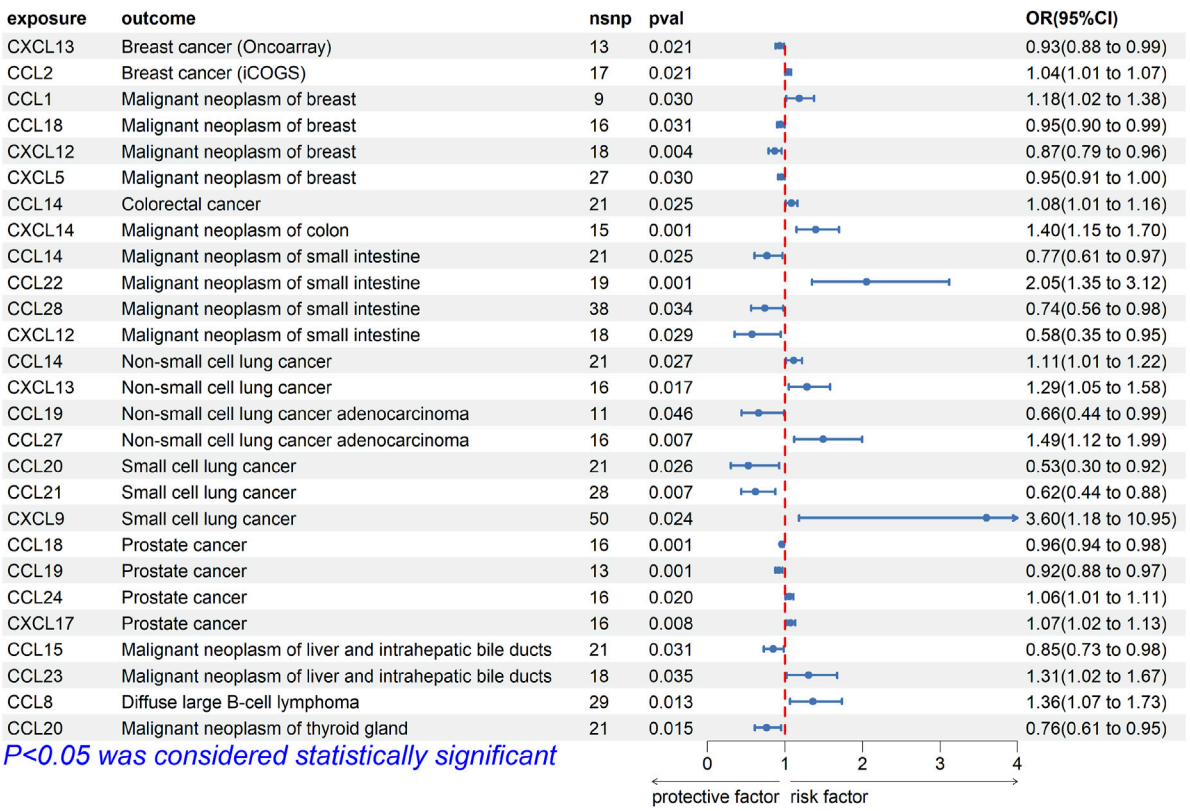


FIGURE 2
Forest plot for the causal association of chemokines on the risk of tumors derived from IVW. The OR value > 0 is considered a risk factor for tumor. The OR value < 0 is considered a protective factor for tumor. OR, odds ratio; CI, confidence interval.

randomization (bi-MR) analysis. Cancers were used as exposure variable and chemokines as outcome variable.

Result

SNP data

First, significant SNPs were screened by *p*-value. Some chemokines did not detect SNPs with *p*-values less than 5E-8, in addition, CCL24 had less than 3 SNPs below this threshold. So, a significance threshold of 5E-6 was set. After significance screening (p < 5E-6) and LD filtering (r^2 = 0.001, kb = 10,000), a total of 828 SNPs of serum chemokine proteins were obtained. F values of 828 SNPs were calculated, and the values were > 10, suggesting that there was no weak instrument bias. Information on all SNPs with a threshold of 5E-6 was shown in Supplementary Table S2 (including F values). IVW test was used as the main MR Analysis method for all chemokines. The statistical results between 38 chemokines and pan-carcinoma were shown in Supplementary Table S3. Similar results were obtained for all sensitivity analyses. The results of heterogeneity analysis and pleiotropy analysis were shown in Supplementary Table S4.

Bi-MR Analysis was performed for all results that met the significance threshold. To ensure sufficient SNPs were available for MR Analysis, the SNPs threshold was set at 5E-8 for breast cancer (excluding finn-b-C3_BREAST) and prostate cancer, and 5E-6 for other malignancies.

In addition, for the results of significance, the *p*-values of the heterogeneity test were all > 0.05 and there was no pleiotropy, Including MR-egger and MR-Presso methods. Moreover, the leave-one-out sensitivity analysis did not find that causality was determined by a single SNP (Supplementary Table S5).

Breast cancer

For breast cancer, we investigated the causal association between chemokines and the disease using three breast cancer GWAS datasets. The results of MR analysis showed significant causal association between CXCL13 [OR (95%CI), 0.93 (0.88–0.99), p = 0.021] and breast cancer (ieu-a-1129), CCL2 [OR (95%CI), 1.04 (1.01–1.07), p = 0.021] and breast cancer (ieu-a-1130), and CCL1 [OR, (95%CI), 1.18 (1.02–1.38), p = 0.030], CCL18 [OR (95%CI), 0.95 (0.90–0.99), p = 0.031], CXCL5 [OR (95%CI), 0.95 (0.91–1.00), p = 0.030], CXCL12 [OR (95%CI), 0.87 (0.79–0.96), p = 0.004] and breast cancer (finn-b-C3_BREAST) (Figure 2). The results were not consistent across different GWAS data, which might be due to different IVs.

Cochrane's Q test did not provide evidence of heterogeneity between CCL1 (p = 0.227), CCL2 (p = 0.982), CCL18 (p = 0.221), CXCL5 (p = 0.533), CXCL12 (p = 0.977), and CXCL13 (p = 0.520)

and breast cancer. The intercept of MR-Egger test did not detect pleiotropy of SNPs for CCL1 (p = 0.779), CCL2 (p = 0.519), CCL18 (p = 0.141), CXCL5 (p = 0.639), CXCL12 (p = 0.800) and CXCL13 (p = 0.643). The MR-Presso test did not detect abnormal SNPs and there was no pleiotropy between SNPs (CCL1 p = 0.223, CCL2 p = 0.943, CCL18 p = 0.413, CXCL5 p = 0.497, CXCL12 p = 0.968, CXCL13 p = 0.569). These results suggest that the serum proteins CCL18, CXCL5, CXCL12, and CXCL13 are protective factors for breast cancer, while CCL1 and CCL2 are risk factors for breast cancer.

Intestinal cancer

For intestinal cancer, we investigated the causal association between chemokines and the disease. The results of MR analysis showed that significant causal association between CCL14 [OR (95% CI), 1.083 (1.010–1.161), p = 0.03] and colorectal cancer (finn-b-C3_COLORECTAL), CXCL14 [OR (95%CI), 1.397 (1.150–1.698), p = 7.98E-04] and colon cancer (finn-b-C3_COLON), CCL22 [OR (95% CI), 2.051 (1.350–3.116), p = 7.58E-04], CCL28 [OR (95%CI), 0.741 (0.562–0.977), p = 0.03], CCL14 [OR (95%CI), 0.766 (0.607–0.967), p = 0.03] and CXCL12 [OR (95%CI), 0.576 (0.351–0.946), p = 0.03] and small intestinal malignant neoplasm (finn-b-C3_SMALL_INTESTINE) (Figure 2).

Cochrane's Q test did not provide evidence of heterogeneity between CCL14 (p = 0.768, p = 0.808), CCL22 (p = 0.502), CCL28 (p = 0.175), CXCL12 (p = 0.530) and CXCL14 (p = 0.534) and intestinal cancer. The intercept of MR-Egger test did not detect pleiotropy of SNPs for CCL14 (p = 0.713, p = 0.399), CCL22 (p = 0.220), CCL28 (p = 0.278), CXCL12 (p = 0.938) and CXCL14 (p = 0.867). The MR-Presso test did not detect abnormal SNPs and there was no pleiotropy between SNPs (CCL14 p = 0.858, p = 0.817, CCL22 p = 0.577, CCL28 p = 0.182, CXCL12 p = 0.549, CXCL14 p = 0.547). These results suggest that the serum proteins CCL14 is a protective factor for colorectal cancer, CXCL14 is a risk factor for colon cancer, CCL14, CCL28 and CXCL12 are protective factors for malignant neoplasm of small intestine, while CCL22 is a risk factor for malignant neoplasm of small intestine.

Lung cancer

For lung cancer, we investigated the causal association between chemokines and the disease. The results of MR analysis showed that significant causal association between CCL14 [OR (95%CI), 1.111 (1.018–1.220), p = 0.03] and CXCL13 [OR (95%CI), 1.286 (1.047–1.579), p = 0.02] and non-small cell lung cancer (finn-b-C3_LUNG_NONSMALL), CCL27 [OR (95%CI), 1.493 (1.118–1.994), p = 0.007] and CCL19 [OR (95%CI), 0.660 (0.438–0.993), p = 0.046] and adenocarcinoma (finn-b-C3_NSCLC_ADENO), CXCL9 [OR (95%CI), 3.597 (1.182–10.953), p = 0.02], CCL20 [OR (95%CI), 0.527 (0.300–0.925), p = 0.03], and CCL21 [OR (95%CI), 0.619 (0.438–0.877), p = 0.01] and small cell lung cancer (finn-b-C3_SCLC) (Figure 2).

Cochrane's Q test did not provide evidence of heterogeneity between CCL14 (p = 0.916), CCL19 (p = 0.446), CCL20 (p = 0.130), CCL21 (p = 0.878), CCL27 (p = 0.762), CXCL9 (p = 0.340) and CXCL13 (p = 0.770) and lung cancer. The intercept of MR-Egger test did not detect pleiotropy of SNPs for CCL14 (p = 0.247), CCL19 (p = 0.696), CCL20 (p = 0.732), CCL21 (p = 0.373), CCL27 (p = 0.963), CXCL9 (p = 0.455) and CXCL13 (p = 0.686). The MR-Presso test did not detect abnormal SNPs and there was no pleiotropy between SNPs (CCL14 p = 0.923, CCL19 p = 0.481, CCL20 p = 0.124, CCL21 p = 0.898, CCL27 p = 0.764, CXCL9 p = 0.239, CXCL13 p = 784). These results suggest that the serum proteins CCL14, CCL27, and CXCL13 are risk factors for non-small cell lung cancer, while CCL19 is a protective factor for non-small cell lung cancer, CXCL9 is a risk factor for small cell lung cancer, while CCL20 and CCL21 are protective factors for small cell lung cancer.

Other cancer

For prostate cancer, we investigated the causal association between chemokines and the disease. The results of MR analysis showed that significant causal association between CCL18 [OR (95% CI), 0.961 (0.939–0.984), p = 1.13E-03], CCL19 [OR (95%CI), 0.920 (0.875–0.968), p = 1.20E-03], CCL24 [OR (95%CI), 1.058 (1.009–1.109), p = 0.02] and CXCL17 [OR (95%CI), 1.074 (1.020–1.132), p = 7.83E-03] and prostate cancer (ebi-a-GCST006085) (Figure 2).

Cochrane's Q test did not provide evidence of heterogeneity between CCL18 (p = 0.767), CCL19 (p = 0.093), CCL24 (p = 0.935), CXCL17 (p = 0.210) and prostate cancer. The intercept of MR-Egger test did not detect pleiotropy of SNPs for CCL18 (p = 0.429), CCL19 (p = 0.790), CCL24 (p = 0.886) and CXCL17 (p = 0.886). The MR-Presso test did not detect abnormal SNPs and there was no pleiotropy between SNPs (CCL18 p = 657, CCL19 p = 0.090, CCL24 p = 0.937, CXCL17 p = 0.228). These results suggest that the serum proteins CCL24, CXCL17 are risk factors for prostate cancer, while CCL18 and CCL19 are protective factors for prostate cancer.

For liver cancer, we investigated the causal association between chemokines and the disease. The results of MR analysis showed that significant causal association between CCL15 [OR (95%CI), 0.848 (0.730–0.985), p = 0.03] and CCL23 [OR (95%CI), 1.306 (1.020–1.673), p = 0.03] and malignant neoplasm of liver (finn-b-C3_LIVER_INTRAHEPATIC_BILE_DUCTS) (Figure 2).

Cochrane's Q test did not provide evidence of heterogeneity between CCL15 (p = 0.698) and CCL23 (p = 0.978) and liver cancer. The intercept of MR-Egger test did not detect pleiotropy of SNPs for CCL15 (p = 0.440) and CCL23 (p = 0.672). The MR-Presso test did not detect abnormal SNPs and there was no pleiotropy between SNPs (CCL15 p = 0.784, CCL23 p = 0.994). These results suggest that the serum proteins CCL23 is a risk factor for malignant neoplasm of liver, while CCL15 is a protective factor for malignant neoplasm of liver.

For Diffuse large B-cell lymphoma (DLBL), we investigated the causal association between chemokines and the disease. The results of MR analysis showed that significant causal association between CCL8 [OR (95%CI), 1.360 (1.065–1.734), p = 0.01] and Diffuse large B-cell lymphoma (finn-b-C3_DLBCL) (Figure 2). Cochrane's Q test did not provide evidence of heterogeneity between CCL8 (p = 0.293) and Diffuse large B-cell lymphoma. The intercept of MR-Egger test did not detect pleiotropy of SNPs for CCL8 (p = 0.099). The

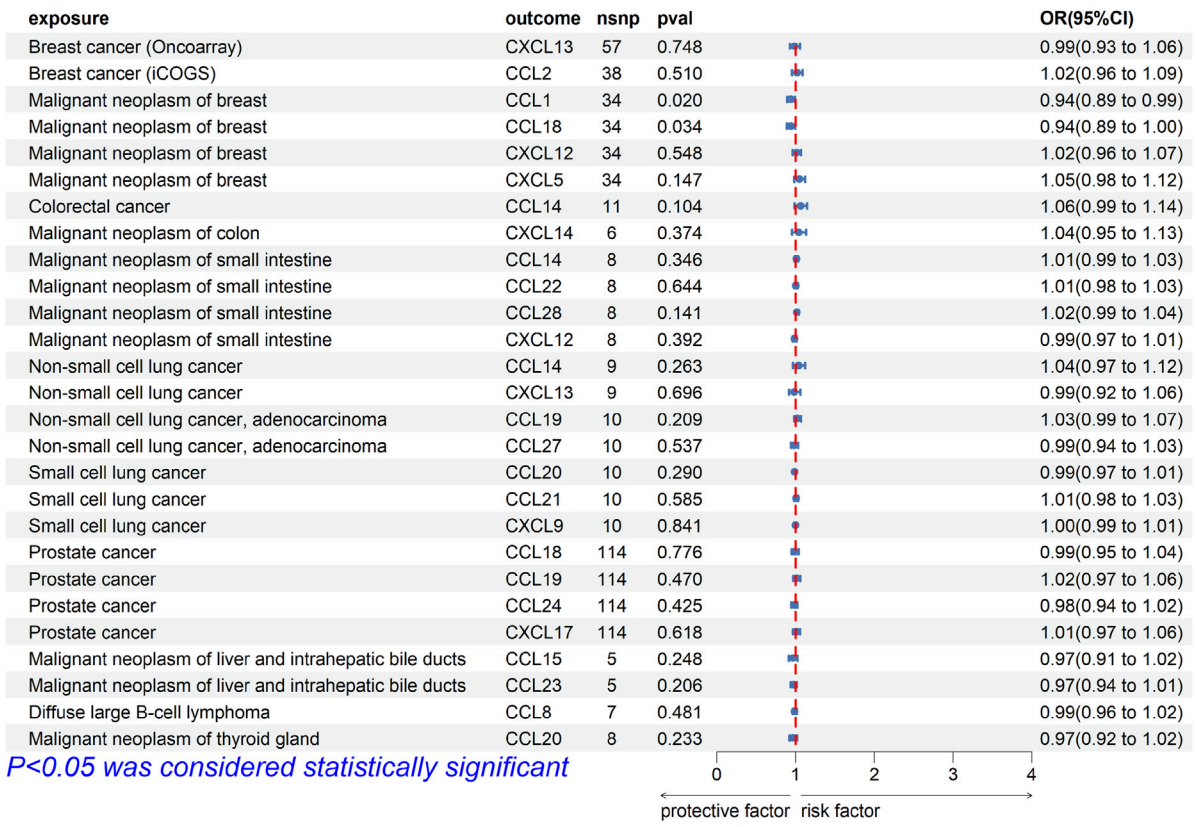


FIGURE 3
Forest plot for the reverse causal association of chemokines on the risk of tumors derived from IVW. The OR value > 0 is considered a risk factor for tumor. The OR value < 0 is considered a protective factor for tumor. OR, odds ratio; CI, confidence interval.

MR-Presso test did not detect abnormal SNPs and there was no pleiotropy between SNPs (p = 0.286). These results suggest that the serum proteins CCL8 is a risk factor for DLBL, For thyroid cancer, we investigated the causal association between chemokines and the disease. The results of MR analysis showed that significant causal association between CCL20 [OR (95%CI), 0.763 (0.614–0.949), p = 0.02] and Malignant neoplasm of thyroid gland (finn-b-C3_THYROID_GLAND) (Figure 2). Cochrane's Q test did not provide evidence of heterogeneity between CCL20 (p = 0.357) and thyroid cancer. The intercept of MR-Egger test did not detect pleiotropy of SNPs for CCL20 (p = 0.887). The MR-Presso test did not detect abnormal SNPs and there was no pleiotropy between SNPs (p = 0.442). These results suggest that the serum proteins CCL20 is a protective factor for malignant neoplasm of thyroid gland.

In addition, the causal association between chemokines and other tumors had also been analyzed, such as malignant tumors of the brain, stomach, pancreas, kidney, ovary, skin, and acute lymphoblastic leukemia. However, there was no causal association between them.

Bi-causal effects between chemokines and tumor risk

To explore whether there was reverse causality in the significant results obtained, we regarded cancer as the exposure factor, chemokines as the outcome, and cancer-related SNPs (p < 5E-8 or p < 5E-6) as the IVs. In bi-MR, the causal association between CCL1 [OR (95%CI), 0.94 (0.89–0.99), p = 0.020] and CCL18 [OR (95%CI), 0.94 (0.89–1.00), p = 0.034] and breast cancer (finn-b-C3_BREAST) was found (Figure 3). Cochrane's Q test did not provide evidence of heterogeneity between CCL1 (p = 0.675) and CCL18 (p = 0.336) and breast cancer. The intercept of MR-Egger test did not detect pleiotropy of SNPs for CCL1 (p = 0.382) and CCL18 (p = 0.964). The MR-Presso test did not detect abnormal SNPs and there was no pleiotropy between SNPs (CCL1 p = 0.595, CCL18 p = 0.297). In addition, MR Analysis showed no causal association between other significance result (p > 0.05).

Discussion

This was the first comprehensive MR analysis to investigate the causal association between chemokines and pan-carcinoma. In two-sample MR Analysis, we initially investigated the causal association between CCL and CXC chemokines and breast, intestinal, lung, and other cancers. Based on the genetic variation of serum protein chemokines and cancers in the publicly available database, it was found that causal association between chemokines and cancer susceptibility. Interestingly, there were also causal association between cancers and partial chemokines. These clues suggest that some chemokines are upstream to drive or hinder the development of cancers.

In our results, there was causal association between 20 chemokines and cancers, as shown in Figure 2. Some studies on chemokines were consistent with our findings. For CCL chemokines, previous studies had shown that CCL1 mainly recruits Tregs to change the tumor microenvironment and promote the progression of breast cancer stem cells (Xu et al., 2017; Kuehnemuth et al., 2018). Surprisingly, there was causal association between CCL1 and breast cancer in bi-MR analysis. The immunosuppressive mechanism of CCL1-recruited Tregs had been widely recognized, but Tregs were also key regulators of $CD8^+$ T cells initiation (Pace et al., 2012). In addition, tissue-resident memory T cells were marker of good prognosis for early triple-negative breast cancer (Byrne et al., 2020). At the same time, other studies had shown that CCL1 was also present in human memory CD8 T cells (Brinza et al., 2016). Although CCL1 promotes tumor progression through Tregs, Tregs may also trigger the accumulation of $CD8^+$ T cells. It could be concluded that CCL1 was upstream of breast cancer, and that breast cancer might also act on CCL1 through negative feedback. For other chemokines, according to recent research reports, CCL2 recruits monocytes to generate vascular endothelial growth factors, thereby facilitating breast cancer cell extravasation (Qian et al., 2011). There was a potential association between CCL8 and DLBL, where CCL8 was involved in the polarization of M2 macrophages and affected patient survival (Lou et al., 2022). Both CCL22 and CCL23 were immunosuppressive chemokines derived from macrophages, which had a unique role in inhibiting anti-tumor immunity (Kamat et al., 2022; Lecoq et al., 2022). In addition, CCL24 was involved in the biological process of cancer through various functions such as angiogenesis and M2 macrophage polarization (Lim, 2021). Moreover, there were some evidences that CCL27 was associated with development of tumors (Martínez-Rodríguez and Monteagudo, 2021). As shown in MR Results, CCL1 and CCL2 were risk factors for breast cancer. CCL8 was a risk factor for DLBL. CCL22 was a risk factor for small intestine malignancy. CCL23 was a risk factor for liver and bile duct malignancy. CCL24 was a risk factor for prostate cancer, and CCL27 was a risk factor for non-small cell lung cancer. Interestingly, CCL14 was a risk factor in lung cancer and colorectal cancer and a protective factor in small intestine tumors. As a chemokine that activates immune cells, studies had found that CCL14 was strongly correlated with a variety of anti-tumor immune cells, including $CD8^+$ T cells, in cancers (Gu et al., 2020). However, other studies had shown that the CCL14 chemokine signaling pathway promotes cancer progression, and inhibiting the expression of CCL14 could reduce the ability of breast cancer to metastasize (Li et al., 2011). Therefore, CCL14 might have different causal associations between different cancers.

Besides, six factors had inverse causal associations with cancer, including CCL15, CCL18, CCL19, CCL20, CCL21, and CCL28. The study found that the chemokine CCL15 recruits CCR^+ $CD14^+$ monocytes in hepatocellular carcinoma, driving multiple tumor-promoting factors (Liu et al., 2019). In addition, CCL18 had been reported as a cancer risk factor in both breast and prostate cancer (Chen et al., 2011; Xu et al., 2014). However, there was an inverse association between CCL15 and CCL18 and cancers in our analysis. For other chemokines, the study had demonstrated that CCL19 exerts a potential stimulatory effect on the response of $CD8^+$ T cells (Yan et al., 2021). In non-small cell lung cancer, CCL19 and CXCL11 reduced the receptor activator of nuclear factor-κB ligand/osteoprotegerin ratio, an indicator of osteoclast stimulation (Kim et al., 2015). In addition, CCL19 and CCL21 migrated dendritic cells in prostate cancer to inhibit cancer progression (Youlin et al., 2018). The same trend was found in our analysis. CCL19 was a protective factor for prostate and lung cancer. The roles of CCL20 and CCL28 in small intestinal and thyroid cancer remain insufficiently investigated, while our findings demonstrate their potential as protective factors, which might provide valuable insights for future research endeavors.

Among CXC chemokines, the causal association with cancer had three positive factors and two negative factors. CXCL9 derived Th1 responses and limited Th2 infiltration, and it was associated with favorable prognosis in small cell lung cancer (Yang L. et al., 2023), however, other studies have reported that CXCL9 binds to CXCR3 in tumors to promote EMT and cancer cell migration (Neo and Lundqvist, 2020). In addition, multiple meta-analyses showed that CXCL12 expression improved the prognosis of breast cancer patients, which was consistent with our results that CXCL12 had a reverse causal association with breast cancer (Samarendra et al., 2017; Liu et al., 2018). Moreover, the studies had shown that CXCL13 drives an anti-tumor immune response to limit tumor progression in mouse breast cancer cells (Ma et al., 2021). And TFH cells that produce CXCL13 played a key role in reversing the immunosuppressive environment induced by Tregs (Gu-Trantien et al., 2017). However, another study suggested that CXCL13 may inhibit tumor growth in breast cancer through CXCR5/ERK signaling (Xu et al., 2018). Simultaneously, in the context of lung cancer, CXCL13 was considered to be a carcinogenic cytokine with significantly enhanced expression levels and facilitating cancer cell invasion through the epithelial-mesenchymal transition (EMT) process (Kazanietz et al., 2019). In our results, CXCL13 was a protective factor in breast cancer and a risk factor in lung cancer. The expression of CXCL14 in colorectal cancer tissues was correlated with TNM stage and poor prognosis. In addition, the invasion ability of cancer cells was also regulated by CXCL14 expression (Zeng et al., 2013), which suggests the pathogenicity of CXCL14 in colorectal cancer. One study reported ventral prostate hyperplasia in estrogen receptor $\beta-/-$ mice with a possible increased incidence of prostate cancer, while genetic analysis found a significant increase in CXCL17 (Wu et al., 2017). Therefore, CXCL17 might also be a potential carcinogen of prostate cancer.

Based on previous studies, some of our results were supported, but partial studies were not consistent with our results, for example, the causal association of CCL15 and CCL18 with tumors incidence. In cancer, the cancer-promoting mechanism of CCL15 was mainly dependent on the monocytes it recruits, and MR Analysis was to analyze the causal association between CCL15 and cancer alone, without involving other factors. The cancer-promoting effect of monocytes recruited by CCL15 might mask the causal association between CCL15 and cancer. In addition, CCL18 mainly recruits Tregs, Th2 and immunosuppressive cells. The effect of immunosuppressive effects on tumors may be larger than the causal association between CCL18 and cancer. In our study, the chemokine concentrations we used were located in the serum, and different locations of cytokine proteins might also cause different causal associations.

In previous analyses, there had not been a comprehensive study to analyze the causal association between chemokines and cancer. One of our strengths is to extract the genetic variation of CCL and CXC chemokines and cancers from a public database for MR Analysis. Based on our analysis, a variety of chemokines were risk factors and protective factors for cancers, and there was no heterogeneity and pleiotropy. Sensitivity analysis also obtained similar results, indicating that our results are credible and accurate. Despite the inherent advantage in MR analysis, it was important to acknowledge its limitations as well. First, we only analyzed the GWAS data of chemokines in serum, and did not analyze the chemokine concentrations in other liquid/tissue samples, which might be biased due to different sites. Secondly, chemokine and cancer GWAS data were obtained from publicly available databases, and subgroup analyses were not possible due to the lack of detailed clinical patient information. Third, the GWAS data are from European populations, and the results may not apply to non-European populations. Finally, the results of this study should still be treated with caution, and more investigations and studies should be conducted to verify the results and consider their application to clinical trial diagnosis.

Conclusion

In summary, since the causal association between chemokines and cancer remains uncertain, and there had not been a comprehensive study to analyze the causal association between chemokines and cancer in previous studies, hereon, we performed a comprehensive two-sample MR Analysis.

As mentioned above, our results showed that causal associations of some chemokines were consistent with previous studies, including CCL2, CCL14, CCL27, CCL19, CCL21, CXCL13, CXCL14 and CXCL17. These chemokines possess the potential to serve as serum diagnostic markers. However, a large number of clinical trials are needed to verify them. In addition, some results were interesting. CCL23 was a risk factor in liver cancer and a protective factor in biliary tract cancer. In colorectal cancer, CCL14 was a risk factor, while in small intestine tumors it was a protective factor. In addition, as widely cognitive tumor-suppressor factor, CXCL9 in small cell lung cancer might be a risk factor. Further study of the underlying mechanisms of these chemokines may provide new insights into targeted therapies for tumors. Moreover, our results also provided new potential targets for tumors, including CCL8, CCL20, CCL28 and CXCL12.

Chemokines in MR results might contribute to tumor prevention and targeted therapy. At present, the detection of serum and plasma markers is crucial for cancer prevention and diagnosis. Based on our results, the serum chemokine concentrations may become new serum markers and parts of chemokines may become the potential therapy targets. Therefore, our results might provide new insights into the future use of chemokines as potential targets for cancer prevention and treatment.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding authors.

Author contributions

KC: Formal Analysis, Methodology, Software, Validation, Writing–original draft. NS: Data curation, Formal Analysis, Methodology, Validation, Writing–original draft, Writing–review and editing. YF: Data curation, Formal Analysis, Methodology, Validation, Writing–original draft. LZ: Data curation, Formal Analysis, Writing–original draft. PS: Data curation, Formal Analysis, Resources, Validation, Writing–original draft. ZW: Data curation, Formal Analysis, Validation, Visualization, Writing–original draft. WS: Data curation, Resources, Writing–review and editing. HW: Conceptualization, Data curation, Funding acquisition, Investigation, Project administration, Supervision, Writing–original draft, Writing–review and editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by the Program for Young Key Teachers in Colleges and Universities in Henan Province (2020GGJS150 and 2021GGJS104), and Key Medical Science and Technology Research Program Project of Henan Province (20232028).

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2023.1285274/full#supplementary-material

SUPPLEMENTARY TABLE S1
Overview of the source of cancer data.

SUPPLEMENTARY TABLE S2
Overview of chemokine SNP data.

SUPPLEMENTARY TABLE S3
All results of Mendelian randomization analysis between chemokines and tumors.

SUPPLEMENTARY TABLE S4
Heterogeneity analysis and pleiotropy analysis.

SUPPLEMENTARY TABLE S5
Mendelian randomization leave-one-out sensitivity analysis.

References

Brinza, L., Djebali, S., Tomkowiak, M., Mafille, J., Loiseau, C., Jouve, P. E., et al. (2016). Immune signatures of protective spleen memory CD8 T cells. *Sci. Rep.* 6, 37651. doi:10.1038/srep37651

Bruchard, M., Geindreau, M., Perrichet, A., Truntzer, C., Ballot, E., Boidot, R., et al. (2022). Recruitment and activation of type 3 innate lymphoid cells promote antitumor immune responses. *Nat. Immunol.* 23, 262–274. doi:10.1038/s41590-021-01120-y

Buniello, A., MacArthur, J. A. L., Cerezo, M., Harris, L. W., Hayhurst, J., Malangone, C., et al. (2019). The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res.* 47, D1005-D1012–d1012. doi:10.1093/nar/gky1120

Byrne, A., Savas, P., Sant, S., Li, R., Virassamy, B., Luen, S. J., et al. (2020). Tissue-resident memory T cells in breast cancer control and immunotherapy responses. *Nat. Rev. Clin. Oncol.* 17, 341–348. doi:10.1038/s41571-020-0333-y

Cambier, S., Gouwy, M., and Proost, P. (2023). The chemokines CXCL8 and CXCL12: molecular and functional properties, role in disease and efforts towards pharmacological intervention. *Cell Mol. Immunol.* 20, 217–251. doi:10.1038/s41423-023-00974-6

Chen, J., Yao, Y., Gong, C., Yu, F., Su, S., Chen, J., et al. (2011). CCL18 from tumor-associated macrophages promotes breast cancer metastasis via PITPNM3. *Cancer Cell* 19, 541–555. doi:10.1016/j.ccr.2011.02.006

Cheng, Q., Zhang, X., Chen, L. S., and Liu, J. (2022). Mendelian randomization accounting for complex correlated horizontal pleiotropy while elucidating shared genetic etiology. *Nat. Commun.* 13, 6490. doi:10.1038/s41467-022-34164-1

Cho, Y. A., and Kim, J. (2013). Association of polymorphisms in the MCP-1 and CCR2 genes with the risk of cancer: a meta-analysis. *Cytokine* 64, 213–220. doi:10.1016/j.cyto.2013.06.319

Ding, M., Zhang, Z., Chen, Z., Song, J., Wang, B., and Jin, F. (2023). Association between periodontitis and breast cancer: two-sample Mendelian randomization study. *Clin. Oral Investig.* 27, 2843–2849. doi:10.1007/s00784-023-04874-x

Gu, Y., Li, X., Bi, Y., Zheng, Y., Wang, J., Li, X., et al. (2020). CCL14 is a prognostic biomarker and correlates with immune infiltrates in hepatocellular carcinoma. *Aging (Albany NY)* 12, 784–807. doi:10.18632/aging.102656

Gudjonsson, A., Gudmundsdottir, V., Axelsson, G. T., Gudmundsson, E. F., Jonsson, B. G., Launer, L. J., et al. (2022). A genome-wide association study of serum proteins reveals shared loci with common diseases. *Nat. Commun.* 13, 480. doi:10.1038/s41467-021-27850-z

Gu-Trantien, C., Migliori, E., Buisseret, L., de Wind, A., Brohée, S., Garaud, S., et al. (2017). CXCL13-producing TFH cells link immune suppression and adaptive memory in human breast cancer. *JCI Insight* 2, e91487. doi:10.1172/jci.insight.91487

Guyatt, A., John, C., Williams, A. T., Shrine, N., Reeve, N. F., Sayers, I., et al. (2023). Mendelian randomisation of eosinophils and other cell types in relation to lung function and disease. *Thorax* 78, 496–503. doi:10.1136/thoraxjnl-2021-217993

Halbrook, C. J., Lyssiotis, C. A., Pasca di Magliano, M., and Maitra, A. (2023). Pancreatic cancer: advances and challenges. *Cell* 186, 1729–1754. doi:10.1016/j.cell.2023.02.014

Hirth, M., Gandla, J., Höper, C., Gaida, M. M., Agarwal, N., Simonetti, M., et al. (2020). CXCL10 and CCL21 promote migration of pancreatic cancer cells toward sensory neurons and neural remodeling in tumors in mice, associated with pain in patients. *Gastroenterology* 159, 665–681. doi:10.1053/j.gastro.2020.04.037

Huang, D., Lin, S., He, J., Wang, Q., and Zhan, Y. (2022). Association between COVID-19 and telomere length: a bidirectional Mendelian randomization study. *J. Med. Virol.* 94, 5345–5353. doi:10.1002/jmv.28008

Kamat, K., Krishnan, V., and Dorigo, O. (2022). Macrophage-derived CCL23 upregulates expression of T-cell exhaustion markers in ovarian cancer. *Br. J. Cancer* 127, 1026–1033. doi:10.1038/s41416-022-01887-3

Kazanietz, M. G., Durando, M., and Cooke, M. (2019). CXCL13 and its receptor CXCR5 in cancer: inflammation, immune response, and beyond. *Front. Endocrinol. (Lausanne)* 10, 471. doi:10.3389/fendo.2019.00471

Kim, H. J., Park, J., Lee, S. K., Kim, K. R., Park, K. K., and Chung, W. Y. (2015). Loss of RUNX3 expression promotes cancer-associated bone destruction by regulating CCL5, CCL19 and CXCL11 in non-small cell lung cancer. *J. Pathol.* 237, 520–531. doi:10.1002/path.4597

Kliemann, N., Rauber, F., Bertazzi Levy, R., Viallon, V., Vamos, E. P., Cordova, R., et al. (2023). Food processing and cancer risk in Europe: results from the prospective EPIC cohort study. *Lancet Planet Health* 7, e219–e232. doi:10.1016/s2542-5196(23)00021-9

Korbecki, J., Kojder, K., Simińska, D., Bohatyrewicz, R., Gutowska, I., Chlubek, D., et al. (2020). CC chemokines in a tumor: a review of pro-cancer and anti-cancer properties of the ligands of receptors CCR1, CCR2, CCR3, and CCR4. *Int. J. Mol. Sci.* 21, 8412. doi:10.3390/ijms21218412

Kuehnemuth, B., Piseddu, I., Wiedemann, G. M., Lauseker, M., Kuhn, C., Hofmann, S., et al. (2018). CCL1 is a major regulatory T cell attracting factor in human breast cancer. *BMC Cancer* 18, 1278. doi:10.1186/s12885-018-5117-8

Kurki, M. I., Karjalainen, J., Palta, P., Sipilä, T. P., Kristiansson, K., Donner, K. M., et al. (2023). FinnGen provides genetic insights from a well-phenotyped isolated population. *Nature* 613, 508–518. doi:10.1038/s41586-022-05473-8

Lagou, M. K., and Karagiannis, G. S. (2023). Obesity-induced thymic involution and cancer risk. *Semin. Cancer Biol.* 93, 3–19. doi:10.1016/j.semcancer.2023.04.008

Larsson, S. C., Traylor, M., Malik, R., Dichgans, M., Burgess, S., Markus, H. S., et al. (2017). Modifiable pathways in Alzheimer's disease: Mendelian randomisation analysis. *Bmj* 359, j5375. doi:10.1136/bmj.j5375

Lecoq, I., Kopp, K. L., Chapellier, M., Mantas, P., Martinenaite, E., Perez-Penco, M., et al. (2022). CCL22-based peptide vaccines induce anti-cancer immunity by modulating tumor microenvironment. *Oncoimmunology* 11, 2115655. doi:10.1080/2162402x.2022.2115655

Li, N., Wang, Y., Wei, P., Min, Y., Yu, M., Zhou, G., et al. (2023). Causal effects of specific gut microbiota on chronic kidney diseases and renal function-A two-sample mendelian randomization study. *Nutrients* 15, 360. doi:10.3390/nu15020360

Li, P., Wang, H., Guo, L., Gou, X., Chen, G., Lin, D., et al. (2022). Association between gut microbiota and preeclampsia-eclampsia: a two-sample Mendelian randomization study. *BMC Med.* 20, 443. doi:10.1186/s12916-022-02657-x

Li, Q., Shi, L., Gui, B., Yu, W., Wang, J., Zhang, D., et al. (2011). Binding of the JmjC demethylase JARID1B to LSD1/NuRD suppresses angiogenesis and metastasis in breast cancer cells by repressing chemokine CCL14. *Cancer Res.* 71, 6899–6908. doi:10.1158/0008-5472.Can-11-1523

Lim, S. J. (2021). CCL24 signaling in the tumor microenvironment. *Adv. Exp. Med. Biol.* 1302, 91–98. doi:10.1007/978-3-030-62658-7_7

Limagne, E., Nuttin, L., Thibaudin, M., Jacquin, E., Aucagne, R., Bon, M., et al. (2022). MEK inhibition overcomes chemoimmunotherapy resistance by inducing CXCL10 in cancer cells. *Cancer Cell* 40, 136–152.e12. doi:10.1016/j.ccell.2021.12.009

Lin, Z., Deng, Y., and Pan, W. (2021). Combining the strengths of inverse-variance weighting and Egger regression in Mendelian randomization using a mixture of regressions model. *PLoS Genet.* 17, e1009922. doi:10.1371/journal.pgen.1009922

Liu, H., Li, Z., Deng, M., Liu, Q., Zhang, T., Guo, W., et al. (2018). Prognostic and clinicopathological value of CXCL12/SDF1 expression in breast cancer: a meta-analysis. *Clin. Chim. Acta* 484, 72–80. doi:10.1016/j.cca.2018.05.041

Liu, L. Z., Zhang, Z., Zheng, B. H., Shi, Y., Duan, M., Ma, L. J., et al. (2019). CCL15 recruits suppressive monocytes to facilitate immune escape and disease progression in hepatocellular carcinoma. *Hepatology* 69, 143–159. doi:10.1002/hep.30134

López-Otín, C., Pietrocola, F., Roiz-Valle, D., Galluzzi, L., and Kroemer, G. (2023). Meta-hallmarks of aging and cancer. *Cell Metab.* 35, 12–35. doi:10.1016/j.cmet.2022.11.001

Lou, X., Zhao, K., Xu, J., Shuai, L., Niu, H., Cao, Z., et al. (2022). CCL8 as a promising prognostic factor in diffuse large B-cell lymphoma via M2 macrophage interactions: a bioinformatic analysis of the tumor microenvironment. *Front. Immunol.* 13, 950213. doi:10.3389/fimmu.2022.950213

Luo, J., le Cessie, S., Blauw, G. J., Franceschi, C., Noordam, R., and van Heemst, D. (2022). Systemic inflammatory markers in relation to cognitive function and measures of brain atrophy: a Mendelian randomization study. *Geroscience* 44, 2259–2270. doi:10.1007/s11357-022-00602-7

Luo, J., le Cessie, S., van Heemst, D., and Noordam, R. (2021). Diet-derived circulating antioxidants and risk of coronary heart disease: a mendelian randomization study. *J. Am. Coll. Cardiol.* 77, 45–54. doi:10.1016/j.jacc.2020.10.048

Ma, Q., Chen, Y., Qin, Q., Guo, F., Wang, Y. S., and Li, D. (2021). CXCL13 expression in mouse 4T1 breast cancer microenvironment elicits antitumor immune response by regulating immune cell infiltration. *Precis. Clin. Med.* 4, 155–167. doi:10.1093/pcmedi/pbab020

Märkl, F., Huynh, D., Endres, S., and Kobold, S. (2022). Utilizing chemokines in cancer immunotherapy. *Trends Cancer* 8, 670–682. doi:10.1016/j.trecan.2022.04.001

Martínez-Rodríguez, M., and Monteagudo, C. (2021). CCL27 signaling in the tumor microenvironment. *Adv. Exp. Med. Biol.* 1302, 113–132. doi:10.1007/978-3-030-62658-7_9

Michailidou, K., Lindström, S., Dennis, J., Beesley, J., Hui, S., Kar, S., et al. (2017). Association analysis identifies 65 new breast cancer risk loci. *Nature* 551, 92–94. doi:10.1038/nature24284

Moreno Ayala, M. A., Campbell, T. F., Zhang, C., Dahan, N., Bockman, A., Prakash, V., et al. (2023). CXCR3 expression in regulatory T cells drives interactions with type I dendritic cells in tumors to restrict CD8(+) T cell antitumor immunity. *Immunity* 56, 1613–1630.e5. doi:10.1016/j.immuni.2023.06.003

Neo, S. Y., and Lundqvist, A. (2020). The multifaceted roles of CXCL9 within the tumor microenvironment. *Adv. Exp. Med. Biol.* 1231, 45–51. doi:10.1007/978-3-030-36667-4_5

Pace, L., Tempez, A., Arnold-Schrauf, C., Lemaitre, F., Bousso, P., Fetler, L., et al. (2012). Regulatory T cells increase the avidity of primary CD8+ T cell responses and promote memory. *Science* 338, 532–536. doi:10.1126/science.1227049

Propper, D. J., and Balkwill, F. R. (2022). Harnessing cytokines and chemokines for cancer therapy. *Nat. Rev. Clin. Oncol.* 19, 237–253. doi:10.1038/s41571-021-00588-9

Qian, B. Z., Li, J., Zhang, H., Kitamura, T., Zhang, J., Campion, L. R., et al. (2011). CCL2 recruits inflammatory monocytes to facilitate breast-tumour metastasis. *Nature* 475, 222–225. doi:10.1038/nature10138

Samarendra, H., Jones, K., Petrinic, T., Silva, M. A., Reddy, S., Soonawalla, Z., et al. (2017). A meta-analysis of CXCL12 expression for cancer prognosis. *Br. J. Cancer* 117, 124–135. doi:10.1038/bjc.2017.134

Schulz, O., Hammerschmidt, S. I., Moschovakis, G. L., and Förster, R. (2016). Chemokines and chemokine receptors in lymphoid tissue dynamics. *Annu. Rev. Immunol.* 34, 203–242. doi:10.1146/annurev-immunol-041015-055649

Schumacher, F. R., Al Olama, A. A., Berndt, S. I., Benlloch, S., Ahmed, M., Saunders, E. J., et al. (2018). Association analyses of more than 140,000 men identify 63 new prostate cancer susceptibility loci. *Nat. Genet.* 50, 928–936. doi:10.1038/s41588-018-0142-8

Sedgwick, P. (2014). Multiple hypothesis testing and Bonferroni's correction. *Bmj* 349, g6284. doi:10.1136/bmj.g6284

Siegel, R. L., Miller, K. D., Wagle, N. S., and Jemal, A. (2023). Cancer statistics, 2023. *CA Cancer J. Clin.* 73, 17–48. doi:10.3322/caac.21763

Timaner, M., Letko-Khait, N., Kotsofruk, R., Benguigui, M., Beyar-Katz, O., Rachman-Tzemah, C., et al. (2018). Therapy-educated mesenchymal stem cells enrich for tumor-initiating cells. *Cancer Res.* 78, 1253–1265. doi:10.1158/0008-5472.Can-17-1547

Wang, F., Tan, F., Shen, S., Wu, Z., Cao, W., Yu, Y., et al. (2023). Risk-stratified approach for never- and ever-smokers in lung cancer screening: a prospective cohort study in China. *Am. J. Respir. Crit. Care Med.* 207, 77–88. doi:10.1164/rccm.202204-0727OC

Wu, F., Huang, Y., Hu, J., and Shao, Z. (2020). Mendelian randomization study of inflammatory bowel disease and bone mineral density. *BMC Med.* 18, 312. doi:10.1186/s12916-020-01778-5

Wu, W. F., Maneix, L., Insunza, J., Nalvarte, I., Antonson, P., Kere, J., et al. (2017). Estrogen receptor β, a regulator of androgen receptor signaling in the mouse ventral prostate. *Proc. Natl. Acad. Sci. U. S. A.* 114, E3816-E3822–e3822. doi:10.1073/pnas.1702211114

Xu, L., Liang, Z., Li, S., and Ma, J. (2018). Signaling via the CXCR5/ERK pathway is mediated by CXCL13 in mice with breast cancer. *Oncol. Lett.* 15, 9293–9298. doi:10.3892/ol.2018.8510

Xu, Y., Dong, X., Qi, P., Ye, Y., Shen, W., Leng, L., et al. (2017). Sox2 communicates with Tregs through CCL1 to promote the stemness property of breast cancer cells. *Stem Cells* 35, 2351–2365. doi:10.1002/stem.2720

Xu, Y., Zhang, L., Sun, S. K., and Zhang, X. (2014). CC chemokine ligand 18 and IGF-binding protein 6 as potential serum biomarkers for prostate cancer. *Tohoku J. Exp. Med.* 233, 25–31. doi:10.1620/tjem.233.25

Yan, Y., Zhao, W., Liu, W., Li, Y., Wang, X., Xun, J., et al. (2021). CCL19 enhances CD8(+) T-cell responses and accelerates HBV clearance. *J. Gastroenterol.* 56, 769–785. doi:10.1007/s00535-021-01799-8

Yang, L., Zhang, Z., Dong, J., Zhang, Y., Yang, Z., Guo, Y., et al. (2023a). Multi-dimensional characterization of immunological profiles in small cell lung cancer uncovers clinically relevant immune subtypes with distinct prognoses and therapeutic vulnerabilities. *Pharmacol. Res.* 194, 106844. doi:10.1016/j.phrs.2023.106844

Yang, M., Wan, X., Zheng, H., Xu, K., Xie, J., Yu, H., et al. (2023b). No evidence of a genetic causal relationship between ankylosing spondylitis and gut microbiota: a two-sample mendelian randomization study. *Nutrients* 15, 1057. doi:10.3390/nu15041057

Yang, Y., Ma, X., Pang, W., and Jiang, C. (2023c). Causal associations of PM2.5 and gdm: a two-sample mendelian randomization study. *Toxics* 11, 171. doi:10.3390/toxics11020171

Yarmolinsky, J., Bouras, E., Constantinescu, A., Burrows, K., Bull, C. J., Vincent, E. E., et al. (2023). Genetically proxied glucose-lowering drug target perturbation and risk of cancer: a Mendelian randomisation analysis. *Diabetologia* 66, 1481–1500. doi:10.1007/s00125-023-05925-4

Youlin, K., Weiyang, H., Simin, L., and Xin, G. (2018). Prostaglandin E2 inhibits prostate cancer progression by countervailing tumor microenvironment-induced impairment of dendritic cell migration through lxrα/CCR7 pathway. *J. Immunol. Res.* 2018, 5808962. doi:10.1155/2018/5808962

Yu, M., Li, Y., Li, B., and Ge, Q. (2023). Inflammatory biomarkers and delirium: a Mendelian randomization study. *Front. Aging Neurosci.* 15, 1221272. doi:10.3389/fnagi.2023.1221272

Zeng, J., Yang, X., Cheng, L., Liu, R., Lei, Y., Dong, D., et al. (2013). Chemokine CXCL14 is associated with prognosis in patients with colorectal carcinoma after curative resection. *J. Transl. Med.* 11, 6. doi:10.1186/1479-5876-11-6

Zhao, J., Ou, B., Han, D., Wang, P., Zong, Y., Zhu, C., et al. (2017). Tumor-derived CXCL5 promotes human colorectal cancer metastasis through activation of the ERK/Elk-1/Snail and AKT/GSK3β/β-catenin pathways. *Mol. Cancer* 16, 70. doi:10.1186/s12943-017-0629-4

Zhou, C., Weng, J., Liu, C., Liu, S., Hu, Z., Xie, X., et al. (2023). Disruption of SLFN11 deficiency-induced CCL2 signaling and macrophage M2 polarization potentiates anti-PD-1 therapy efficacy in hepatocellular carcinoma. *Gastroenterology* 164, 1261–1278. doi:10.1053/j.gastro.2023.02.005

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 13 December 2023
DOI 10.3389/fgene.2023.1299930

OPEN ACCESS

EDITED BY
Eleonore Fröhlich,
Medical University of Graz, Austria

REVIEWED BY
Le Liu,
Southern Medical University, China
Kaijian Hou,
Shantou University, China

*CORRESPONDENCE
Yong Meng,
mengyong@nwu.edu.cn

†These authors have contributed equally to this work and share first authorship

RECEIVED 23 September 2023
ACCEPTED 04 December 2023
PUBLISHED 13 December 2023

CITATION
Quan Z, Zhang X, Wang S and Meng Y (2023), Causal analysis of the gut microbiota in differentiated thyroid carcinoma: a two-sample Mendelian randomization study.
Front. Genet. 14:1299930.
doi: 10.3389/fgene.2023.1299930

Causal analysis of the gut microbiota in differentiated thyroid carcinoma: a two-sample Mendelian randomization study

Zheng Quan[1†], Xiaoyu Zhang[2†], Shilong Wang[3] and Yong Meng[1*]

[1]Department of Oncology Surgery, The Affiliated Hospital of Northwest University, Xi'an, China,
[2]Department of Intensive Care Unit, The First Affiliated Hospital of Xi'an Jiaotong University, Xi'an, China,
[3]Department of Surgical Oncology, The First Affiliated Hospital of Xi'an Jiaotong University, Xi'an, China

Objective: Numerous studies have highlighted an association between the gut microbiota (GM) and thyroid tumors. Employing Mendelian randomization methodology, we seek to elucidate the causal link between the gut microbiota and thyroid neoplasms.

Methods: We procured data from the Mibiogen database encompassing 211 distinct gut microbiota taxa, alongside extensive genome-wide association studies (GWAS) summary data for differentiated thyroid carcinoma (DTC). Our principal analytical approach involved the application of the Inverse-Variance Weighted method (IVW) within the framework of Mendelian randomization. Simultaneously, we conducted sensitivity analyses to assess result heterogeneity, horizontal pleiotropy, and outcome stability.

Results: IVW analysis revealed a dual role of the GM in thyroid carcinoma. The phylum Actinobacteria (OR, 0.249 [95% CI, 0.121–0.515]; $p < 0.001$) was associated with a decreased risk of DTC. Conversely, the genus Ruminiclostridium9 (OR, 11.276 [95% CI, 4.406–28.860]; $p < 0.001$), class Mollicutes (OR, 5.902 [95% CI, 1.768–19.699]; $p = 0.004$), genus RuminococcaceaeUCG004 (OR, 3.831 [95% CI, 1.516–9.683]; $p = 0.005$), genus Paraprevotella (OR, 3.536 [95% CI, 1.330–9.401]; $p = 0.011$), and phylum Tenericutes (OR, 5.902 [95% CI, 1.768–19.699]; $p = 0.004$) were associated with an increased risk of DTC.

Conclusion: Our findings underscore that the presence of genus Ruminiclostridium9, class Mollicutes, genus RuminococcaceaeUCG004, genus Paraprevotella, and phylum Tenericutes is associated with an elevated risk of DTC, whereas the presence of the phylum Actinobacteria is linked to a decreased risk. These discoveries enhance our comprehension of the relationship between the GM and DTC.

KEYWORDS

gut microbiota, differentiated thyroid carcinoma, heterogeneity assessment, Mendelian randomization study, causal relationship

1 Introduction

Thyroid carcinoma, a common endocrine neoplasm of the head and neck, has experienced a steady increase in incidence, currently ranking as the fifth most prevalent cancer globally (Bray et al., 2018; WHO, 2020). Projections suggest that following its current trajectory, thyroid malignancies will become the fourth most common cancer in the United States by 2030 (Gonçalves et al., 2017). In 2020, global age-standardized incidence rates for thyroid cancer were 10.1 cases per 100,000 females and 3.1 cases per 100,000 males, with corresponding mortality rates of 0.5 and 0.3 cases per 100,000, respectively (Pizzato et al., 2022). An epidemiological survey covering 24% of the population in China (Yao et al., 2023) revealed that in 2019, the age-standardized incidence and mortality rates for thyroid cancer were 2.05 and 0.39 per 100,000, respectively. Over the last 30 years, the International Agency for Research on Cancer has noted an increasing incidence of thyroid cancer in diverse populations worldwide. In the United States, from 1970 to 2013, the annual growth rate of thyroid cancer incidence was reported to be 3% (Lim et al., 2017). Fortunately, the mortality rates for both males and females in most countries exhibit a stable or declining trend (Huang et al., 2023). DTC accounts for over 90% of all pathological diagnoses, with papillary carcinoma being its predominant histological subtype (Gu et al., 2014).The etiology of thyroid tumors is multifactorial, involving chromosomal mutations, genetic predisposition (Bonnefond and Davies, 2014), estrogen levels (Luo et al., 2016), ionizing radiation exposure (Bonnefond and Davies, 2014), autoimmune thyroid disorders (Khatami, 2009), and other factors. However, the risk factors for thyroid tumors are not fully understood, requiring additional research to uncover their pathogenic mechanisms.

Currently, the National Comprehensive Cancer Network (NCCN) guidelines recommend primary surgical intervention for DTC, reserving radioactive iodine-131 treatment for specific patient subsets (Haddad et al., 2022). Considering the global prevalence of thyroid tumors and their healthcare burden, as well as their profound impact on the wellbeing of affected individuals, our research is dedicated to uncovering the etiological underpinnings of this affliction.

Humans have coexisted with microorganisms throughout their existence, hosting a diverse array of microbes within various bodily niches, including the oral cavity, respiratory tract, gastrointestinal tract, genitourinary tract, and skin. Among these, the gut harbors the most intricate microbial ecosystem (Jiang et al., 2022). The human gut, in particular, teems with an assembly of microbial denizens numbering in the billions, with bacteria occupying the central stage (Gill et al., 2006). Such a vast consortium of GM also fulfills distinctive roles. Presently, microbiota are acknowledged for their substantial contributions to vitamin synthesis (B-complex vitamins, folate, vitamin K, among others) (Gu et al., 2016; Fang et al., 2017), facilitation of dietary fiber digestion, and regulation of immune responses (Bastiaanssen et al., 2019). Beyond these functions, microbiota also exhibit intricate associations with various diseases, encompassing gastrointestinal disorders, psychiatric illnesses, respiratory maladies, autoimmune conditions, and significantly, diverse malignancies, including lung, breast, colorectal, and esophageal cancers (Stebbings et al., 2002; Wang et al., 2016; Ishaq et al., 2021; Vitale et al., 2021).

A plethora of evidence has pointed toward the association between GM and thyroid malignancies, including thyroid carcinoma (Stebbings et al., 2002; Ishaq et al., 2021). (Ejtahed et al., 2020) Thyroid cancer patients exhibit dysbiosis in the gut microbiota, characterized by a reduction in the relative abundance of Faecalibacterium prausnitzii. Interestingly, an increase in the abundance of Faecalibacterium prausnitzii is observed after Radioactive Iodine Therapy (RAIT) (Fernandes et al., 2023). Lu et al. have identified significant alterations in the composition of the gut microbiota in thyroid cancer patients, with the *Bacteroides* enterotype emerging as the predominant bacterial type (Lu et al., 2022). Furthermore, gene sequencing results indicate a higher abundance of Firmicutes (Liu et al., 2021). In a study encompassing 74 patients, high-throughput sequencing was utilized to compare the microbial structural characteristics of 20 thyroid carcinoma patients, 18 thyroid nodule patients, and 36 healthy controls. The results underscored a close relationship between thyroid carcinoma, thyroid nodules, and altered microbiota (Zhang et al., 2019). Despite numerous indications suggesting an association between GM and thyroid malignancies, our understanding of this relationship remains incomplete, as these studies have not yet established causal links between thyroid tumors and GM.

Mendelian randomization represents an analytical framework harnessing genetic variation as instrumental variables (IV) to infer causal relationships between specific risk factors (i.e., exposures) and particular phenotypes (i.e., outcomes). Genetic variation, in this context, predominantly alludes to single nucleotide polymorphisms, signifying variations in specific nucleotides within the genetic material. In the realm of clinical investigation, myriad confounding factors often obscure the precision of our conclusions, rendering causal inferences tentative at best. The intrinsic merit of Mendelian randomization lies in its capacity to circumvent the impracticability of randomized controlled trials, such as the random allocation of microbiota to study individuals. Instead, this method leverages the natural grouping of single nucleotide polymorphisms (SNPs) and employs statistical techniques to ascertain the influence of SNPs on exposure and outcomes. nsequently, it helps to clarify causal associations between exposures and outcomes. Notably, owing to the even distribution of SNP loci, Mendelian randomization outcomes remain comparatively impervious to the interference of confounding factors, thereby conferring results akin to those derived from randomized controlled trials (Davies et al., 2018).

We utilize Mendelian randomization to disentangle the impact of confounding factors, enabling a precise evaluation of the causal relationship between GM and DTC., and Figure 1 provides an overview of the main research approach in this paper. This study's overarching goal resides in employing Mendelian randomization as a methodological prism, utilizing genetic variation as instrumental variables to elucidate the causal nexus between microbiota and DTC. In doing so, we aspire to contribute novel evidence to the etiological and therapeutic paradigms within the domain of thyroid pathology.

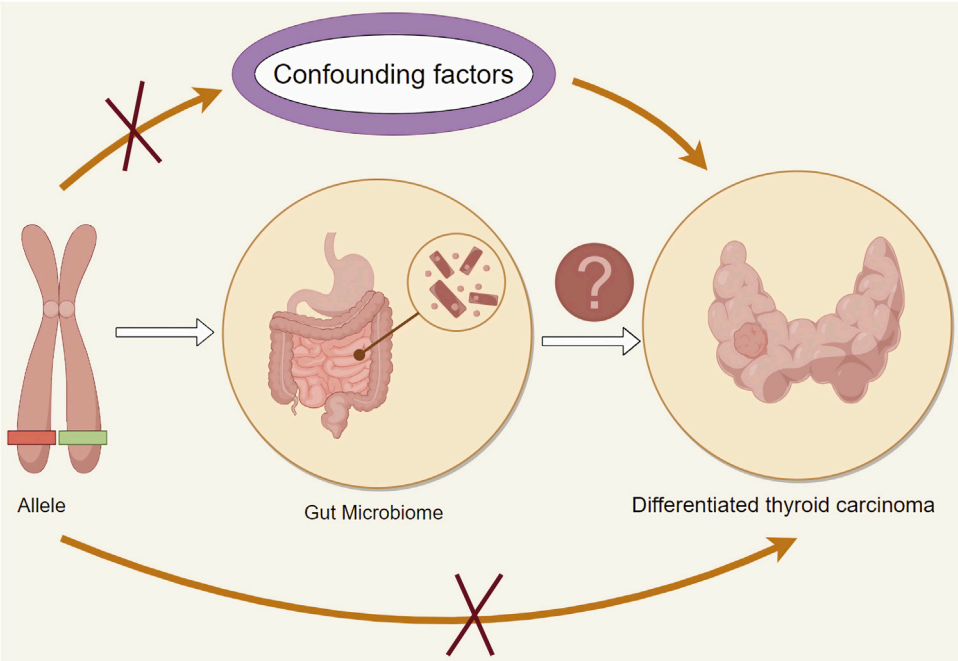


FIGURE 1
Study design concept and framework.

2 Methods

2.1 Ethics statement

As the data used in this study comes from publicly available databases, after obtaining ethical approval from the Affiliated Hospital of Northwest University, Xi'an NO.3 hospital, the committee deemed formal ethical approval unnecessary. This decision was predicated upon the utilization of publicly accessible data devoid of identifiable patient information.

2.2 Study design

In this study, we define exposure as the GM, the outcome as malignant thyroid neoplasms, and instrumental variables as single nucleotide polymorphism (SNP) loci. In accordance with this premise, we discerned GM significantly associated with malignant thyroid neoplasms and subsequently proceeded with Mendelian randomization analysis. Throughout the course of MR analysis, we adhere to the following assumptions.

1. Associational Hypothesis: That is, SNP loci under investigation and GM demonstrate robust correlations. In our study, the significance threshold for the associational hypothesis is set at $P < 1E\text{-}5$.
2. Independence of SNPs and Confounding Factors: Among the SNPs ultimately incorporated into the MR study, those SNP loci exhibiting associations with either tumors or GM were excluded.
3. Exclusivity of Instrumental Variables' Impact on Outcomes through Exposure: Instrumental variables should solely affect outcomes through the exposure and remain inert to other pathways, such as confounding. In essence, there should be no pleiotropy.

2.3 Data collection

The data pertaining to GM emanates from the consortium's whole-genome association study summary data, Mibiogen. This dataset encapsulates 211 distinct taxonomic groups within the GM, spanning six taxonomic levels: kingdom, phylum, class, order, family, and genus. Access to this data can be procured from the website (https://mibiogen.gcc.rug.nl/). Notably, eight unidentified bacterial species were excluded from subsequent microbiota SNP locus analyses, leaving a total of 203 microbial SNPs for further investigation.

Thyroid tumor data, on the other hand, was sourced from Aleksandra Köhler et al.'s prospective study (Köhler et al., 2013), which conducted a comprehensive genome-wide association study encompassing 701 patients afflicted with DTC. Diagnoses of thyroid tumors were ascertained through pathological results furnished by the Cisanello Hospital in Pisa, a prominent Italian referral center for thyroid disorders.

2.4 Variable selection

To assess the correlation between instrumental variables and microbiota, a filtration process for employed instrumental variables was enacted, involving the following steps:

1. The GM data, having been downloaded from the Mibiogen website, was validated and subsequently imported into R for analysis.
2. SNP loci exhibiting a stronger correlation with exposure were identified and filtered out based on a threshold of P < 1E-5 in the initial filtration process.
3. Instrumental variables demonstrating linkage disequilibrium were excluded. A standard of r2 > 0.001 and a physical distance (Kb) of 10,000 were employed for the removal of SNP loci exhibiting r^2 values exceeding 0.001 with the most significant SNP within a 10,000 Kb range.
4. F-statistics were computed to assess the strength of instrumental variables. The calculation of F-statistics was predicated on beta values and standard errors (SE) for SNPs and exposure. In our study, all instrumental variables exhibited F-statistics exceeding 10.
5. SNP loci not conforming to the independence assumption were eliminated. We accessed the Phenoscanner database to identify secondary phenotypes associated with each SNP, verifying their correlation with confounding factors. SNP loci associated with both exposure outcomes were discarded based on criteria of *p*-value: < 1E-8 and r̂2 > 0.8. Naturally, SNP loci directly correlated with thyroid tumors were also excluded. Following this comprehensive filtration, we gathered information regarding instrumental variables in outcomes and amalgamated effect sizes, commencing the MR analysis.

2.5 Statistical analyses

In the course of Mendelian randomization analysis, six distinct methodologies were employed, namely, IVW, IVW random-effects model, MR-Egger, MR-Egger bootstrap, Weighted Median, and Simple Median. Of these, the IVW analysis results, which calculated both the unadjusted *p*-values and the False Discovery Rate (FDR)-corrected *p*-values for each SNP locus, served as the primary analytical approach for this study. Multiple sensitivity analyses were additionally conducted, serving three primary objectives: firstly, to assess the robustness of the outcomes; secondly, to evaluate the potential presence of biases, including pleiotropy and data heterogeneity; and thirdly, to appraise the scenario where a specific instrumental variable unduly influenced the outcome.

To quantify heterogeneity in individual causal effects, Cochran's Q was computed and subjected to examination, with a significance threshold of $p \leq 0.05$ indicating the presence of pleiotropy. Within the context of heterogeneity testing, MR-Egger's intercept and the Mendelian randomization residual sum and outlier (MR-PRESSO) method were employed. If the *p*-value exceeded 0.05, it indicated the absence of horizontal pleiotropy. All results underwent comprehensive visualization. To evaluate the scenario where a specific instrumental variable significantly impacted the outcome, a leave-one-out analysis was conducted by systematically excluding each SNP locus and observing the remaining SNPs' Mendelian randomization. Finally, a reverse Mendelian randomization analysis was performed to ascertain the causal direction.

All statistical analyses were executed using the R programming language (https://www.r-project.org, R version 4.2.1). Statistical significance was deemed at $p < 0.05$. The initial date of analysis commenced in May 2023.

3 Results

3.1 Instrumental variable selection and initial MR results

Following our criteria, an initial set of 14,569 instrumental variable loci was established. Supplementary Table S1 provides a comprehensive breakdown of all microbiota details. By matching SNPs with thyroid tumor data, we obtained a subset of 3,302 SNPs.

Initial Mendelian randomization analysis yielded insights into the relationships between 203 GM and thyroid function, as presented in Supplementary Table S2. Based on the IVW-derived *p*-values, an initial selection identified 13 microbiota entities, namely,: Genus Ruminiclostridium 9 (ID:11357), Class Mollicutes (ID 3920), Phylum Tenericutes (ID:3919), Genus Ruminococcaceae UCG004 (ID:11362), Genus Paraprevotella (ID:962), Genus Ruminococcaceae UCG003 (ID:11361), Family Victivallaceae (ID: 2,255), Genus Candidatus Soleaferrea (ID:11350), Phylum Actinobacteria (ID:400), However, through rigorous heterogeneity testing, horizontal pleiotropy assessments, and the exclusion of loci indirectly or directly related to thyroid tumor diseases, we ultimately distilled the selection down to six GM entities and 47 SNP loci as instrumental variables: Genus Ruminiclostridium 9 (ID:11357), Class Mollicutes (ID:3920), Genus Ruminococcaceae UCG004 (ID:11362), Genus Paraprevotella (ID:962), Phylum Actinobacteria (ID:400), Phylum Tenericutes (ID:3919).

The secondary features of these aforementioned SNPs were queried using PhenoScanner and are documented in Supplementary Table S3. It is noteworthy that these features have been confirmed as non-pleiotropic factors contributing to thyroid tumor etiology.

3.2 Detailed Mendelian randomization analysis results

We conducted Mendelian randomization analysis on the final set of six GM and 47 SNP loci. IVW results revealed significant associations as follows: genus Ruminiclostridium9(OR, 11.276, [95% CI, 4.406–28.860]; $p < 0.001$), class Mollicutes (OR, 5.902, [95% CI, 1.768–19.699]; $p = 0.004$), genus RuminococcaceaeUCG004(OR, 3.831, [95% CI, 1.516–9.683]; $p = 0.005$), genus Paraprevotella (OR, 3.536, [95% CI, 1.330–9.401]; $p = 0.011$), phylum Tenericutes (OR, 5.902 [95% CI, 1.768–19.699]; $p = 0.004$)exhibited an elevated risk of thyroid tumors, whereas phylum actinobacteria (OR, 0.249 [95% CI, 0.121–0.515]; $p < 0.001$)demonstrated a decreased risk of thyroid tumors. These findings are graphically depicted in the forest plot (Figure 2). Additional MR analysis outcomes are presented in Supplementary Table S4. In our findings, IVW and IVW-MRE methods demonstrated consistency, while other methodologies may yield differing results compared to IVW and IVW-MRE. We posit that the IVW approach derives an overall effect by weight-averaging estimates across distinct loci, whereas IVW-MRE additionally accounts for measurement error, employing a random-effects model to estimate the total effect. Discrepancies among other methods may be attributed to sample size, skewness, or heterogeneity. In case of inconsistencies, pay particular attention to

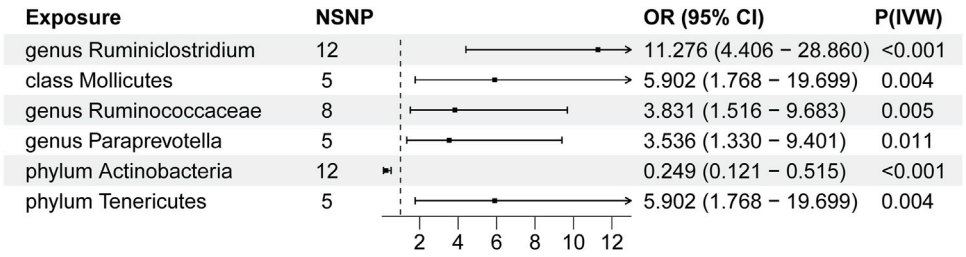


FIGURE 2
Forest Plot of Primary MR Analysis Results. NSNP = number of SNPs; OR = odds ratio; P = significance *p*-value; IVW = Inverse-Variance Weighted method.

the results from IVW and IVW-MRE methods, considering the unique attributes of alternative methodologies.

To visualize our findings comprehensively, we presented all results in a scatter plot (Figure 3), where each data point represents an SNP, the upper and lower lines delineate confidence intervals, and the horizontal and vertical axes respectively denote the SNP's effects on GM and thyroid tumor outcomes. The colored lines signify the fitting effects of the MR. Intriguingly, IVW and IVW-MRE methods exhibited remarkable consistency across all results.

3.3 Sensitivity analysis

Initially, we conducted heterogeneity checks on the results obtained from the selected six bacteria and 47 SNP loci. We observed that all I2 values for the microbiota were <50%, and the *p*-values obtained using two different methods were both greater than 0.05 (Table 1). This suggests that our results exhibited minimal heterogeneity. Additionally, the MR-PRESSO outlier test did not identify any anomalies (Table 1). To assess the horizontal pleiotropy of SNP loci, we utilized the global test from MRPRESSO and the MR-Egger intercept test. The *p*-values for all tests of horizontal pleiotropy exceeded 0.3, signifying that the impact of instrumental variables on thyroid cancer is unlikely to be influenced by factors other than the microbiota (Table 1).

In addition to MR-Egger and MR-PRESSO, we conducted individual SNP MR estimates (Supplementary Figure S1) and systematically removed individual SNPs to compute the remaining SNPs' Mendelian randomization effects (Figure 4). Leave-one-out analysis, where each SNP is excluded in turn, showed that all error bars were consistently on the right or left of zero in Figure 4. This indicates minimal variation in the overall error bars, signifying high robustness in the results. The causal estimates were not driven by any single SNP. In the forest plot of Supplementary Figure S1, each solid horizontal line represents the results estimated using the Wald ratio method for individual SNPs, while the red line represents the composite outcome, reflecting the risk of thyroid cancer for each microbiota under the IVW method.

3.4 Reverse MR analysis

In the reverse MR analysis, exposure and outcome were interchanged. However, we did not observe any significant causal relationship between the outcome and exposure, except in the case of the Phylum Tenericutes using the MR Egger method, where the *p*-value exceeded 0.05 (Table 2).

We conducted a global test for horizontal pleiotropy using the MR-PRESSO method, and no evidence of horizontal pleiotropy was detected among the SNPs. Moreover, no outliers were detected in the outlier test. Hence, we have reasonable grounds to posit that these various bacteria are causative factors for thyroid tumors, rather than being outcomes of the condition.

4 Discussion

In spite of the prior research on the relationship between the GM and thyroid cancer, the concept of the association between DTC and the GM remains relatively uncharted (Samimi and Haghpanah, 2020). To the best of our knowledge, this study represents the first causal investigation into the link between the GM and DTC. Through a Mendelian randomization analysis involving two-sample datasets, we report that five microbiota entities, namely, genus Ruminiclostridium 9 ($p < 0.001$), class Mollicutes ($p = 0.004$), genus Ruminococcaceae UCG004 ($p = 0.005$), genus Paraprevotella ($p = 0.011$), and phylum Tenericutes ($p = 0.004$), are associated with an increased risk of developing non-differentiated thyroid carcinoma. Additionally, phylum Actinobacteria ($p < 0.001$) appears to be associated with a lower the risk of non-differentiated thyroid carcinoma. We utilized various sensitivity analysis techniques to affirm the reliability and robustness of our findings.

The GM is known to be influenced by a multitude of factors. Notably, infants born via cesarean section exhibit lower diversity in their GM(31). Throughout one's life, the GM remains under the continual influence of various factors including diet, medication, genetics, environment, disease, and the use of antibiotics (Maslowski and Mackay, 2011). This substantiates the significance of the GM as a potential therapeutic target for diseases.

Most of the research concerning the GM has been concentrated on gastrointestinal diseases, such as colorectal cancer and inflammatory bowel disease. In the context of colorectal cancer, gut bacteria may promote tumorigenesis by influencing bile acid secretion or undergoing changes in their taxonomic composition. Decreased bile acid secretion results in gut dysbiosis, accelerating inflammation and DNA damage, thereby directly contributing to tumorigenesis (Louis et al., 2014; Ridlon et al., 2014). Furthermore,

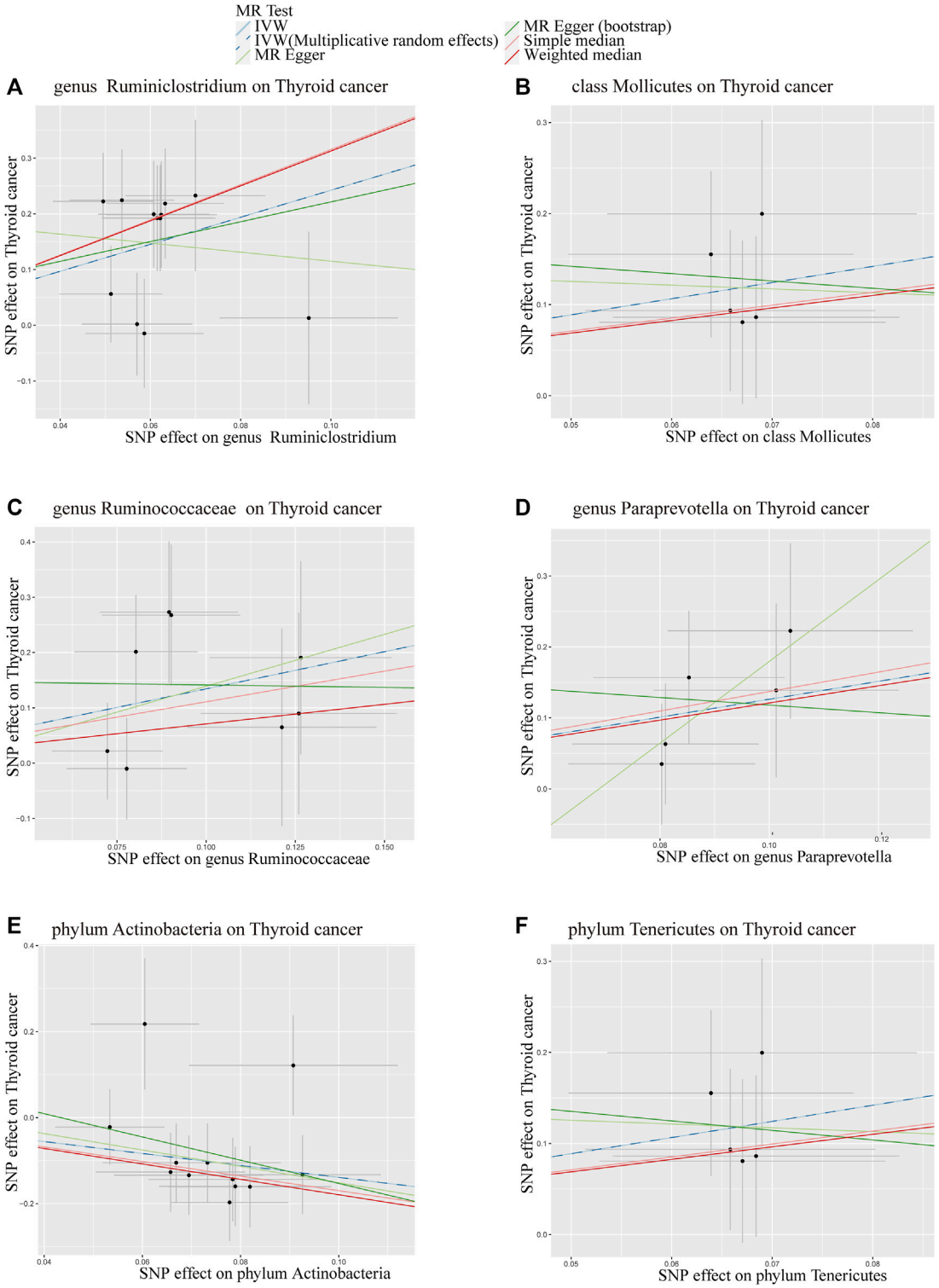


FIGURE 3
Scatter Plot of MR Analysis for the Influence of 6 Gut Microbiota on Thyroid Tumors. **(A)** = genus Ruminiclostridium; **(B)** = class Mollicutes; **(C)** = genus Ruminococcaceae; **(D)** = genus Paraprevotella; **(E)** = phylum Actinobacteria; **(F)** = phylum Tenericutes; MR Test = Statistical analysis methods.

an increase in *Fusobacterium* nucleatum, a type of gut bacteria, can promote colon tumorigenesis through its metabolites and cytotoxicity, (Wu et al., 2004; Wu et al., 2009), thereby affecting signaling pathways such as E-cadherin, NF-κB, and STAT3 (Maslowski and Mackay, 2011; Rautava et al., 2012). These studies suggest that the role of microbiota in tumorigenesis seems to be associated with DNA damage and the regulation of local inflammation via metabolic products.

TABLE 1 Heterogeneity and horizontal pleiotropy test results for gut microbiota.

Gut microbes	Heterogeneity test				MR-PRESSO				MR-Egger	
	IVW	*p*-value	MR-Egger	P	Global test	P	Outlier-corrected	P	Egger intercept test	P
Genus Ruminiclostridium	11.357	0.414	10.404	0.406	1.795	0.904	NA	NA	0.196	0.361
Class Mollicutes	1.180	0.881	1.171	0.760	1.794	0.892	NA	NA	0.146	0.930
Genus Ruminococcaceae	6.376	0.497	6.324	0.388	12.336	0.517	NA	NA	−0.048	0.831
Genus Paraprevotella	1.656	0.799	0.762	0.859	13.459	0.474	NA	NA	−0.396	0.414
Phylum Actinobacteria	10.804	0.460	10.761	0.376	8.457	0.501	NA	NA	0.039	0.846
Phylum Tenericutes	1.180	0.881	1.171	0.760	2.621	0.82	NA	NA	0.146	0.930

Heterogeneity tests were conducted using Cochran Q for both IVW, and MR-Egger, while horizontal pleiotropy was assessed using MR-PRESSO, and the Egger intercept test.

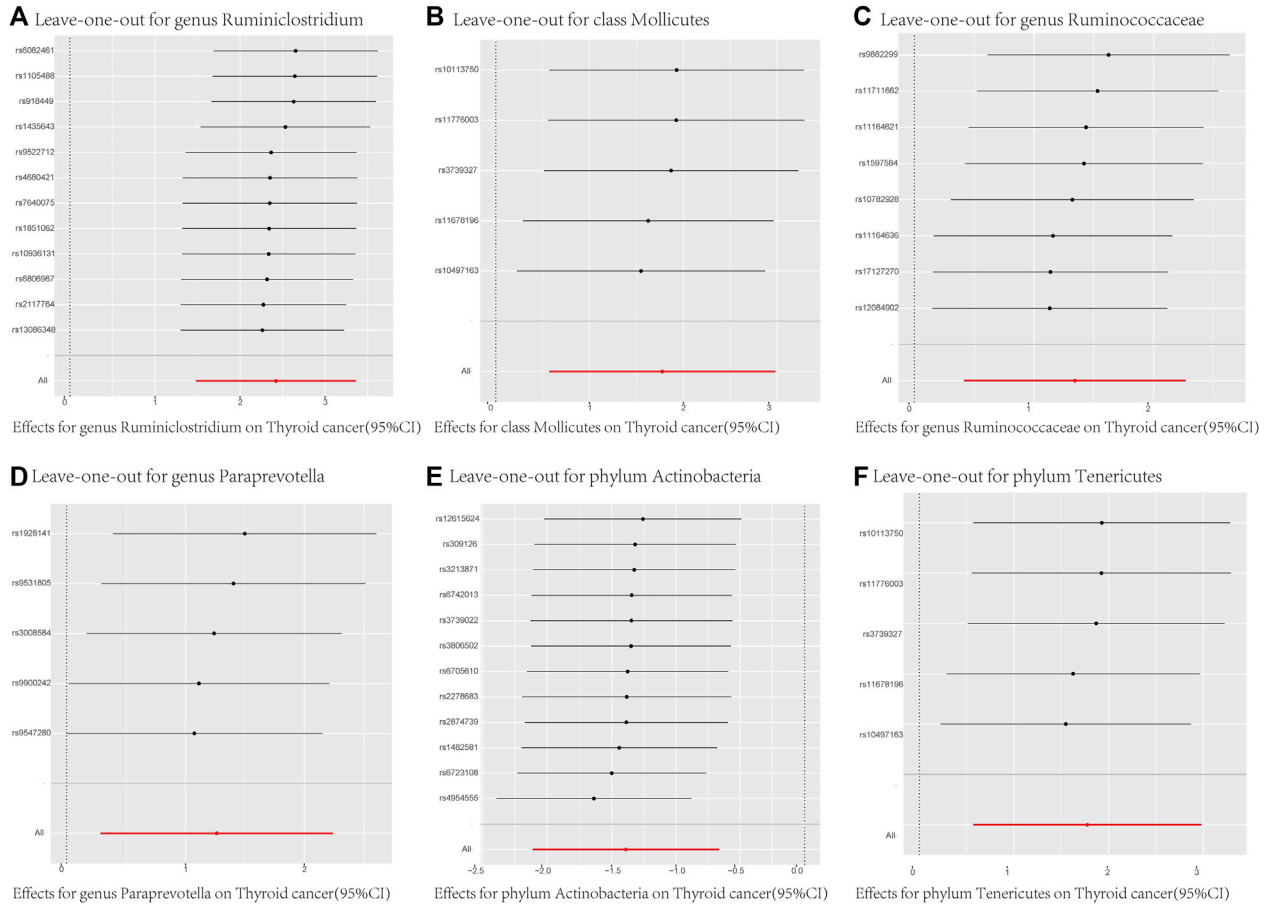


FIGURE 4
MR Analysis Results Using Leave-One-Out for Different SNP Loci of the 6 Gut Microbiota. **(A)** = genus Ruminiclostridium; **(B)** = class Mollicutes; **(C)** = genus Ruminococcaceae; **(D)** = genus Paraprevotella; **(E)** = phylum Actinobacteria; **(F)** = phylum Tenericutes; MR Test = Statistical analysis methods.

Microbes can impact thyroid disease through various pathways. The GM may influence the secretion of thyroid-stimulating hormones via the hypothalamus-pituitary axis, thereby playing a role in thyroid diseases (Fröhlich and Wahl, 2019). For instance, in studies of patients with Hashimoto's thyroiditis (HT), a form of hypothyroidism, researchers observed dysbiosis in the patients' GM, along with overgrowth of certain microbes (Ishaq et al., 2017; Zhao et al., 2018). This influence is thought to occur because thyroid-related nutrients need to be acquired through the gut (Knezevic et al., 2020). Furthermore, in Graves' disease patients, there is a higher abundance of Bacteroidaceae and Prevotellaceae and a lower abundance of Veillonellaceae, Enterobacteriaceae, and

10.3389/fgene.2023.1299930

TABLE 2 Reverse MR analysis of the main results.

Outcome	method	NSNP	b	se	*p*-value	*p*-value (MR-PRESSO)
Phylum Tenericutes	Inverse variance weighted	328	−0.000413	0.0006304	0.512	0.561
	MR Egger	328	−0.002639	0.0011036	0.017	
	Simple mode	328	−0.001691	0.0019129	0.377	
	Weighted median	328	−0.000413	0.0009814	0.674	
	Weighted mode	328	0.0002101	0.0013032	0.872	
Phylum Actinobacteria	Inverse variance weighted	329	0.0003186	0.0005009	0.525	0.977
	MR Egger	329	2.73E-05	0.0008769	0.975	
	Simple mode	329	−0.001726	0.00158	0.275	
	Weighted median	329	−0.000538	0.0007755	0.488	
	Weighted mode	329	−0.001229	0.0011322	0.278	
Genus Ruminococcaceae	Inverse variance weighted	328	−0.000573	0.0006887	0.405	0.751
	MR Egger	328	0.0008787	0.0012057	0.467	
	Simple mode	328	0.0003512	0.0020584	0.865	
	Weighted median	328	−3.73E-05	0.0010136	0.971	
	Weighted mode	328	0.0007602	0.0014312	0.596	
Genus Ruminiclostridium	Inverse variance weighted	329	0.0002747	0.0005414	0.612	0.126
	MR Egger	329	0.0009264	0.0009481	0.329	
	Simple mode	329	−0.000906	0.0016819	0.591	
	Weighted median	329	0.0004644	0.0007995	0.561	
	Weighted mode	329	0.00038	0.0010821	0.726	
Genus Paraprevotella	Inverse variance weighted	328	−0.000161	0.0007957	0.840	0.781
	MR Egger	328	0.0004046	0.0013927	0.772	
	Simple mode	328	0.001209	0.0021467	0.574	
	Weighted median	328	−0.000415	0.0011592	0.720	
	Weighted mode	328	−0.000828	0.001428	0.563	
Class Mollicutes	Inverse variance weighted	328	−0.000413	0.0006304	0.512	0.555
	MR Egger	328	−0.002639	0.0011036	0.017	
	Simple mode	328	−0.001691	0.0016845	0.316	
	Weighted median	328	−0.000413	0.0009433	0.661	
	Weighted mode	328	0.0002101	0.001257	0.867	

NSNP = number of SNPs; b = effect size; se = standard error; pval = *p*-value.

Lachnospiraceae compared to healthy individuals (Ishaq et al., 2018). In the realm of researching the relationship between the GM and thyroid cancer, the first study proposing a potential connection between thyroid cancer and the GM was published in 2017. Shen et al. employed gas chromatography-time-of-flight mass spectrometry to analyze the serum of thyroid cancer patients with and without distant metastases. They found elevated levels of serum ammonia, pyruvic acid, and γ-aminobutyric acid in patients with distant metastases, suggesting a potential connection to differences in GM or diet (Shen et al., 2017). The first study to specifically investigate the relationship between thyroid cancer and the GM was conducted by Jing Feng in 2019 (Feng et al., 2019). Feng reported higher gut microbial richness and α-diversity in patients with TC compared to healthy controls, but the sample size was limited (30 patients from Harbin, China, *versus* 35 healthy controls), and the study design was observational. In other observational studies, GM taxa that were found to be increased in abundance in thyroid cancer patients included Clostridiaceae, Nesterenkonia, and *Streptococcus* (Zhang et al., 2019), while decreased taxa primarily included *Lactobacillus* (Zhang et al., 2019), *Bacteroides*, *Clostridium*, and Prevotella. These studies shed light on the potential link between the GM and thyroid cancer.

Our study highlights the significance of five specific microbiota entities—genus Ruminiclostridium 9, class Mollicutes, genus Ruminococcaceae UCG004, genus Paraprevotella, phylum Tenericutes, and phylum Actinobacteria—in relation to DTC.

Genus Ruminiclostridium 9, formerly known as Ruminoclostridium, is a characteristic bacterium found in the colon. Research indicates that interactions among colonic microorganisms are more intricate compared to the duodenum (Wu et al., 2020). This bacterium is associated with various diseases; it can reduce the risk of Alzheimer's disease (OR 0.969, 95% CI 0.943–0.996, $p = 0.009$) (Ning et al., 2022). Furthermore, it has a positive correlation with cognitive function (Guo et al., 2021). In tumor immunology studies, Ruminiclostridium has shown a negative correlation with $CD8^+$ T cells (Singh et al., 2023). In an autoimmune model of the central nervous system, medium-chain fatty acids (MCFAs) produced by this genus counteract the anti-inflammatory effects of short-chain fatty acids (SCFAs) by enhancing TH1 and TH17 cell differentiation (Haghikia et al., 2015).Therefore, certain genera, such as Ruminiclostridium 9, may contribute to the initiation of DTC through analogous mechanisms, potentially by reducing short-chain fatty acidsand modulating the immune response.

Class Mollicutes, representing the smallest self-replicating bacteria without cell walls, belongs to the phylum Tenericutes (Chernova et al., 2021). Mollicutes exhibit versatility, as they have been associated with both decreased relative abundance in severe depression patients (Zhu et al., 2019) and as a risk factor for Graves' disease (Cao et al., 2023). Phylum Tenericutes has also been linked to various diseases, such as breast cancer (Niccolai et al., 2023), Crohn's disease (Russo et al., 2022), lower risk of intrahepatic cholestasis of pregnancy (Li et al., 2023), and polycystic ovary syndrome (Li et al., 2023). However, research on the mechanisms underlying these associations remains limited. Genus Ruminococcaceae, like Ruminiclostridium, is present in the colonic mucosa. A decrease in Ruminococcaceae has been associated with various inflammatory bowel diseases, including ulcerative colitis and Crohn's disease (Sokol et al., 2008; Joossens et al., 2011; Morgan et al., 2012). This bacterium produces short-chain fatty acids (SCFAs) and other small molecules, which serve as an energy source for colonic epithelial cells. A deficiency in these SCFAs may lead to disturbances and dysfunction in colonic mucosa (Young and Schmidt, 2004; Wong et al., 2006). SCFAs, particularly butyrate, are known to influence immune regulation and possess anti-inflammatory properties (Köhling et al., 2017), and butyrate can inhibit the activity and life cycle of cancer cells (Davie, 2003). This knowledge, seemingly at odds with our study results, suggests the involvement of unknown mechanisms. Interestingly, it has also been identified as a potential etiological factor contributing to Graves' Disease (GD) (Cao et al., 2023). This association may be attributed to the ability of short-chain fatty acids (SCFA) to inhibit histone deacetylases (HDAC) and activate the re-expression of transport proteins in thyroid cancer cells, thereby inducing the differentiation of tumor cells and enhancing iodine uptake (Zhou et al., 2018; Rathod et al., 2020). Genus Paraprevotella has been the subject of several studies. It is a polymorphic, anaerobic, non-spore-forming Gram-negative rod isolated from human feces (Morotomi et al., 2009). This bacterium exhibits complex effects. On one hand, it is more abundant in individuals with genetic longevity (Liu et al., 2023) and correlates with small intestinal mucosal healing in Crohn's disease patients (Hattori et al., 2020). On the other hand, it is more abundant in patients with heart failure and depression (Gutiérrez-Calabrés et al., 2020) and positively correlates with the severity of depression (Liśkiewicz et al., 2021). Previous studies have indicated its utility in distinguishing untreated primary hypothyroidism patients from healthy individuals.Currently, Paraprevotella is considered to have a positive correlation with plasma butyrate and valerate concentrations and contributes to the regulation of colonic motility, possibly through regulating fecal butyrate levels and serum IL-8 concentrations. The Paraprevotella strain proves to be an efficient pancreatic protease-degrading symbiont. The autolysis of pancreatic proteases facilitates bacterial invasion and destruction, potentially culminating in inflammation and injury, thereby creating a conducive environment for the onset of thyroid cancer.In summary, Paraprevotella is associated with various diseases, but its potential impact on human health remains unclear (Morotomi et al., 2009), and its specific mechanisms are yet to be explored. Phylum Actinobacteria is one of the most diverse bacterial phyla in nature (Lewin et al., 2016). Due to its diversity, it has both positive and negative effects on human health. Actinobacteria includes many bacteria that produce antibiotics, such as *Streptomyces*, which is known for synthesizing eptomycin, kanamycin, chloramphenicol, and erythromycin (Bérdy, 2005). It also encompasses several human health-threatening pathogens, such as *Mycobacterium tuberculosis*, which causes pulmonary tuberculosis. In the order Bifidobacteriales, members like Bifidobacterium are known for their beneficial effects on host health. A lower abundance of these bacteria has been associated with various diseases (Binda et al., 2018). Some researchers propose that species of Bifidobacterium in the human gut may contribute to host health by exerting antibacterial activity against pathogens and possibly by aiding in the development and function of the immune system as defensive symbionts (Servin, 2004). This aligns with our research findings. We also postulate that Actinobacteria may indirectly inhibit the occurrence of thyroid cancer by suppressing the processes of inflammation and oxidative stress.

What mechanisms might bacteria employ to induce DTC? Carcinogenesis primarily relies on two mechanisms: DNA damage and cell apoptosis, as well as the immune surveillance against tumor growth (Docimo et al., 2020; Liu et al., 2022). Intestinal bacteria can impact tumor proliferation through both of these mechanisms. New mechanisms involving bacteria and their metabolites or toxins causing direct DNA damage and carcinogenic mutations have been identified. For instance, infection with *Enterococcus* can lead to an increase in hydroxyl free radical production, resulting in DNA damage (Wang and Huycke, 2007). Additionally, oxidative stress can disrupt the homeostasis of the host gut microbiota (Riaz Rajoka et al., 2021). Therefore, drugs targeting both oxidative stress and gut microbiota may hold prognostic and therapeutic significance for thyroid cancer.Concerning immune surveillance, the gut microbiota significantly regulates the balance within the body and the development of immune cells. It modulates both innate and adaptive immune systems, particularly outside the intestinal tract (Maslowski and Mackay, 2011), making it a potential immune modulator (Pitt et al., 2016). As it is widely known, more than 70% of the entire immune system is associated with the

gastrointestinal lymphoid tissues, and the gut microbiota can regulate immune balance and cell development (Maslowski and Mackay, 2011; Docimo et al., 2020). Certain metabolic products can also induce autoimmune reactions, leading to an imbalance in endocrine homeostasis and the occurrence of autoimmune diseases (Liu et al., 2022). Microbial dysbiosis can stimulate CD8 (+) T cells to promote chronic inflammation and early T-cell exhaustion, thereby diminishing the immune capacity against tumors (Yu et al., 2020). Therefore, the gut microbiota holds significant potential as a biomarker for predicting immune-related adverse events (Von Itzstein et al., 2020) and may influence the progression of thyroid cancer through immune modulation.In the future, we can develop novel strategies for diagnosing, predicting prognosis, actively monitoring, and intervening in DTC by studying the correlation between different bacterial enterotypes and thyroid cancer.This study also presents certain limitations. Firstly, we solely analyzed common microbiota and SNP loci, leaving unidentified microbiota and SNPs unexamined, thereby limiting our analysis. Secondly, our GWAS data was derived from individuals of European descent, and we did not identify other patient characteristics, which may restrict the generalizability of our conclusions to other populations, given the varying compositions of microbiota across different countries and ethnicities. Additionally, while we employed different SNP loci as "natural groupings," it is essential to acknowledge the numerous unknown or potential interactions among distinct SNP loci. Finally, it is essential to note that this study is an observational investigation, lacking foundational research to mechanistically substantiate the observed outcomes. In the future, in addition to requiring more Genome-Wide Association Studies (GWAS) and microbiota data, further in-depth research is warranted to elucidate the mechanisms underlying the association between gut microbiota and thyroid tumor development. These studies will contribute to enhancing our understanding of the causal link between thyroid cancer and the gut microbiota.

In conclusion, our findings indicate a bidirectional role of GM in thyroid cancer. Genus Ruminiclostridium9, class Mollicutes, genus RuminococcaceaeUCG004, genus Paraprevotella, and phylum Tenericutes were associated with an increased risk of undifferentiated thyroid cancer, while phylum Actinobacteria ($p < 0.001$) was associated with a reduced risk of undifferentiated thyroid cancer. However, the underlying processes involved are intricate, necessitating further mechanistic research.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

Ethics statement

The requirement of ethical approval was waived by The Medical Ethics Committee of Xi'an Third Hospital for the studies involving humans because Utilization of publicly accessible data devoid of identifiable patient information. The studies were conducted in accordance with the local legislation and institutional requirements. The ethics committee/institutional review board also waived the requirement of written informed consent for participation from the participants or the participants' legal guardians/next of kin because Utilization of publicly accessible data devoid of identifiable patient information. Written informed consent was not obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article because utilization of publicly accessible data.

Author contributions

ZQ: Writing–original draft, Writing–review and editing. XZ: Writing–original draft, Writing–review and editing. SW: Conceptualization, Software, Supervision, Writing–review and editing. YM: Conceptualization, Investigation, Writing–review and editing.

Funding

The authors declare financial support was received for the research, authorship, and/or publication of this article. This work was funded by the Xi'an NO. 3 Hospital Institutional Medical Research Fund (Grant NO. Y2023yxyj09), Clinical Research Award of the First Affiliated Hospital of Xi'an Jiaotong University (Grant NO. XJTU1AF2021CRF-009).

Acknowledgments

We extend our gratitude to Mibiogene and Kohler A for providing the thyroid data, which have advanced the etiological research of malignant thyroid tumors in the world.

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2023.1299930/full#supplementary-material

References

Bastiaanssen, T. F. S., Cowan, C. S. M., Claesson, M. J., Dinan, T. G., and Cryan, J. F. (2019). Making sense of the microbiome in psychiatry. *Int. J. Neuropsychopharmacol.* 22 (1), 37–52. doi:10.1093/ijnp/pyy067

Bérdy, J. (2005). Bioactive microbial metabolites: a personal view. *J. Antibiot.* 58 (1), 1–26. doi:10.1038/ja.2005.1

Binda, C., Lopetuso, L. R., Rizzatti, G., Gibiino, G., Cennamo, V., and Gasbarrini, A. (2018). Actinobacteria: a relevant minority for the maintenance of gut homeostasis. *Dig. Liver Dis.* 50 (5), 421–428. doi:10.1016/j.dld.2018.02.012

Bonnefond, S., and Davies, T. F. (2014). Thyroid cancer—risks and causes. *Oncol. Hematol. Rev. (US)* 10 (2), 144. doi:10.17925/ohr.2014.10.2.144

Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., and Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* 68 (6), 394–424. doi:10.3322/caac.21492

Cao, J., Wang, N., Luo, Y., Ma, C., Chen, Z., Chenzhao, C., et al. (2023). A cause–effect relationship between Graves' disease and the gut microbiome contributes to the thyroid–gut axis: a bidirectional two-sample Mendelian randomization study. *Front. Immunol.* 14, 977587. doi:10.3389/fimmu.2023.977587

Chernova, O. A., Chernov, V. M., Mouzykantov, A. A., Baranova, N. B., Edelstein, I. A., and Aminov, R. I. (2021). Antimicrobial drug resistance mechanisms among Mollicutes. *Int. J. Antimicrob. Agents* 57 (2), 106253. doi:10.1016/j.ijantimicag.2020.106253

Davie, J. R. (2003). Inhibition of histone deacetylase activity by butyrate. *J. Nutr.* 133 (7), 2485S–2493S. doi:10.1093/jn/133.7.2485S

Davies, N. M., Holmes, M. V., and Davey Smith, G. (2018). Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. *BMJ* 362, k601. doi:10.1136/bmj.k601

Docimo, G., Cangiano, A., Romano, R. M., Pignatelli, M. F., Offi, C., Paglionico, V. A., et al. (2020). The human microbiota in endocrinology: implications for pathophysiology, treatment, and prognosis in thyroid diseases. *Front. Endocrinol.* 11, 586529. doi:10.3389/fendo.2020.586529

Ejtahed, H. S., Angoorani, P., Soroush, A. R., Siadat, S. D., Shirzad, N., Hasani-Ranjbar, S., et al. (2020). Our little friends with big roles: alterations of the gut microbiota in thyroid disorders. *EMIDDT* 20 (3), 344–350. doi:10.2174/1871530319666190930110605

Fang, H., Kang, J., and Zhang, D. (2017). Microbial production of vitamin B12: a review and future perspectives. *Microb. Cell Fact.* 16 (1), 15. doi:10.1186/s12934-017-0631-y

Feng, J., Zhao, F., Sun, J., Lin, B., Zhao, L., Liu, Y., et al. (2019). Alterations in the gut microbiota and metabolite profiles of thyroid carcinoma patients. *Int. J. Cancer* 144 (11), 2728–2745. doi:10.1002/ijc.32007

Fernandes, A., Oliveira, A., Carvalho, A. L., Soares, R., and Barata, P. (2023). Faecalibacterium prausnitzii in differentiated thyroid cancer patients treated with radioiodine. *Nutrients* 15 (12), 2680. doi:10.3390/nu15122680

Fröhlich, E., and Wahl, R. (2019). Microbiota and thyroid interaction in health and disease. *Trends Endocrinol. Metabolism* 30 (8), 479–490. doi:10.1016/j.tem.2019.05.008

Gill, S. R., Pop, M., DeBoy, R. T., Eckburg, P. B., Turnbaugh, P. J., Samuel, B. S., et al. (2006). Metagenomic analysis of the human distal gut microbiome. *Science* 312 (5778), 1355–1359. doi:10.1126/science.1124234

Gonçalves, C. F. L., de Freitas, M. L., and Ferreira, A. C. F. (2017). Flavonoids, thyroid iodide uptake and thyroid cancer-A review. *Int. J. Mol. Sci.* 18 (6), 1247. doi:10.3390/ijms18061247

Gu, Q., and Li, P. (2016). "Biosynthesis of vitamins by probiotic bacteria," in *Probiotics and prebiotics in human nutrition and health*. Editors V. Rao and L. G. Rao (IntechOpen) Available from: https://www.intechopen.com/about-intechopen (Accessed September 4, 2023).

Gu, Y., Yu, Y., Ai, L., Shi, J., Liu, X., Sun, H., et al. (2014). Association of the ATM gene polymorphisms with papillary thyroid cancer. *Endocrine* 45 (3), 454–461. doi:10.1007/s12020-013-0020-1

Guo, M., Peng, J., Huang, X., Xiao, L., Huang, F., and Zuo, Z. (2021). Gut microbiome features of Chinese patients newly diagnosed with Alzheimer's disease or mild cognitive impairment. *JAD* 80 (1), 299–310. doi:10.3233/JAD-201040

Gutiérrez-Calabrés, E., Ortega-Hernández, A., Modrego, J., Gómez-Gordo, R., Caro-Vadillo, A., Rodríguez-Bobada, C., et al. (2020). Gut microbiota profile identifies transition from compensated cardiac hypertrophy to heart failure in hypertensive rats. *Hypertension* 76 (5), 1545–1554. doi:10.1161/HYPERTENSIONAHA.120.15123

Haddad, R. I., Bischoff, L., Ball, D., Bernet, V., Blomain, E., Busaidy, N. L., et al. (2022). Thyroid carcinoma, version 2.2022, NCCN clinical practice guidelines in oncology. *J. Natl. Compr. Canc Netw.* 20 (8), 925–951. doi:10.6004/jnccn.2022.0040

Haghikia, A., Jörg, S., Duscha, A., Berg, J., Manzel, A., Waschbisch, A., et al. (2015). Dietary fatty acids directly impact central nervous system autoimmunity via the small intestine. *Immunity* 43 (4), 817–829. doi:10.1016/j.immuni.2015.09.007

Hattori, S., Nakamura, M., Yamamura, T., Maeda, K., Sawada, T., Mizutani, Y., et al. (2020). The microbiome can predict mucosal healing in small intestine in patients with Crohn's disease. *J. Gastroenterol.* 55 (12), 1138–1149. doi:10.1007/s00535-020-01728-1

Huang, J., Ngai, C. H., Deng, Y., Pun, C. N., Lok, V., Zhang, L., et al. (2023). Incidence and mortality of thyroid cancer in 50 countries: a joinpoint regression analysis of global trends. *Endocrine* 80 (2), 355–365. doi:10.1007/s12020-022-03274-7

Ishaq, H. M., Mohammad, I. S., Guo, H., Shahzad, M., Hou, Y. J., Ma, C., et al. (2017). Molecular estimation of alteration in intestinal microbial composition in Hashimoto's thyroiditis patients. *Biomed. Pharmacother.* 95, 865–874. doi:10.1016/j.biopha.2017.08.101

Ishaq, H. M., Mohammad, I. S., Shahzad, M., Ma, C., Raza, M. A., Wu, X., et al. (2018). Molecular alteration analysis of human gut microbial composition in Graves' disease patients. *Int. J. Biol. Sci.* 14 (11), 1558–1570. doi:10.7150/ijbs.24151

Ishaq, H. M., Mohammad, I. S., Sher Muhammad, K., Li, H., Abbas, R. Z., Din Sindhu, Z. U., et al. (2021). Gut microbial dysbiosis and its association with esophageal cancer. *J. Appl. Biomed.* 19 (1), 1–13. doi:10.32725/jab.2021.005

Jiang, W., Lu, G., Gao, D., Lv, Z., and Li, D. (2022). The relationships between the gut microbiota and its metabolites with thyroid diseases. *Front. Endocrinol.* 13, 943408. doi:10.3389/fendo.2022.943408

Joossens, M., Huys, G., Cnockaert, M., De Preter, V., Verbeke, K., Rutgeerts, P., et al. (2011). Dysbiosis of the faecal microbiota in patients with Crohn's disease and their unaffected relatives. *Gut* 60 (5), 631–637. doi:10.1136/gut.2010.223263

Khatami, M. (2009). Inflammation, aging, and cancer: tumoricidal versus tumorigenesis of immunity: a common denominator mapping chronic diseases. *Cell Biochem. Biophys.* 55 (2), 55–79. doi:10.1007/s12013-009-9059-2

Knezevic, J., Starchl, C., Tmava Berisha, A., and Amrein, K. (2020). Thyroid-Gut-Axis: how does the microbiota influence thyroid function? *Nutrients* 12 (6), 1769. doi:10.3390/nu12061769

Köhler, A., Chen, B., Gemignani, F., Elisei, R., Romei, C., Figlioli, G., et al. (2013). Genome-wide association study on differentiated thyroid cancer. *J. Clin. Endocrinol. Metabolism* 98 (10), E1674–E1681. doi:10.1210/jc.2013-1941

Köhling, H. L., Plummer, S. F., Marchesi, J. R., Davidge, K. S., and Ludgate, M. (2017). The microbiota and autoimmunity: their role in thyroid autoimmune diseases. *Clin. Immunol.* 183, 63–74. doi:10.1016/j.clim.2017.07.001

Lewin, G. R., Carlos, C., Chevrette, M. G., Horn, H. A., McDonald, B. R., Stankey, R. J., et al. (2016). Evolution and ecology of *actinobacteria* and their bioenergy applications. *Annu. Rev. Microbiol.* 70 (1), 235–254. doi:10.1146/annurev-micro-102215-095748

Li, C., Li, N., Liu, C., and Yin, S. (2023). Causal association between gut microbiota and intrahepatic cholestasis of pregnancy: mendelian randomization study. *BMC Pregnancy Childbirth* 23 (1), 568. doi:10.1186/s12884-023-05889-8

Lim, H., Devesa, S. S., Sosa, J. A., Check, D., and Kitahara, C. M. (2017). Trends in thyroid cancer incidence and mortality in the United States, 1974-2013. *JAMA* 317 (13), 1338–1348. doi:10.1001/jama.2017.2719

Liśkiewicz, P., Kaczmarczyk, M., Misiak, B., Wroński, M., Bąba-Kubiś, A., Skonieczna-Żydecka, K., et al. (2021). Analysis of gut microbiota and intestinal integrity markers of inpatients with major depressive disorder. *Prog. Neuro-Psychopharmacology Biol. Psychiatry* 106, 110076. doi:10.1016/j.pnpbp.2020.110076

Liu, C. J., Chen, S. Q., Zhang, S. Y., Wang, J. L., Tang, X. D., Yang, K. X., et al. (2021). The comparison of microbial communities in thyroid tissues from thyroid carcinoma patients. *J. Microbiol.* 59 (11), 988–1001. doi:10.1007/s12275-021-1271-9

Liu, Q., Sun, W., and Zhang, H. (2022). Interaction of gut microbiota with endocrine homeostasis and thyroid cancer. *Cancers* 14 (11), 2656. doi:10.3390/cancers14112656

Liu, X., Zou, L., Nie, C., Qin, Y., Tong, X., Wang, J., et al. (2023). Mendelian randomization analyses reveal causal relationships between the human microbiome and longevity. *Sci. Rep.* 13 (1), 5127. doi:10.1038/s41598-023-31115-8

Louis, P., Hold, G. L., and Flint, H. J. (2014). The gut microbiota, bacterial metabolites and colorectal cancer. *Nat. Rev. Microbiol.* 12 (10), 661–672. doi:10.1038/nrmicro3344

Lu, G., Yu, X., Jiang, W., Luo, Q., Tong, J., Fan, S., et al. (2022). Alterations of gut microbiome and metabolite profiles associated with anabatic lipid dysmetabolism in thyroid cancer. *Front. Endocrinol.* 13, 893164. doi:10.3389/fendo.2022.893164

Luo, J., Hendryx, M., Manson, J. E., Liang, X., and Margolis, K. L. (2016). Hysterectomy, oophorectomy, and risk of thyroid cancer. *J. Clin. Endocrinol. Metab.* 101 (10), 3812–3819. doi:10.1210/jc.2016-2011

Maslowski, K. M., and Mackay, C. R. (2011). Diet, gut microbiota and immune responses. *Nat. Immunol.* 12 (1), 5–9. doi:10.1038/ni0111-5

Morgan, X. C., Tickle, T. L., Sokol, H., Gevers, D., Devaney, K. L., Ward, D. V., et al. (2012). Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment. *Genome Biol.* 13 (9), R79. doi:10.1186/gb-2012-13-9-r79

Morotomi, M., Nagai, F., Sakon, H., and Tanaka, R. (2009). Paraprevotella clara gen. nov., sp. nov. and Paraprevotella xylaniphila sp. nov., members of the family "Prevotellaceae" isolated from human faeces. *Int. J. Syst. Evol. Microbiol.* 59 (8), 1895–1900. doi:10.1099/ijs.0.008169-0

Niccolai, E., Baldi, S., Nannini, G., Gensini, F., Papi, L., Vezzosi, V., et al. (2023). Breast cancer: the first comparative evaluation of oncobiome composition between males and females. *Biol. Sex. Differ.* 14 (1), 37. doi:10.1186/s13293-023-00523-w

Ning, J., Huang, S. Y., Chen, S. D., Zhang, Y. R., Huang, Y. Y., and Yu, J. T. (2022). Investigating casual associations among gut microbiota, metabolites, and neurodegenerative diseases: a mendelian randomization study. *J. Alzheimers Dis.* 87 (1), 211–222. doi:10.3233/JAD-215411

Pitt, J. M., Vétizou, M., Daillère, R., Roberti, M. P., Yamazaki, T., Routy, B., et al. (2016). Resistance mechanisms to immune-checkpoint blockade in cancer: tumor-intrinsic and -extrinsic factors. *Immunity* 44 (6), 1255–1269. doi:10.1016/j.immuni.2016.06.001

Pizzato, M., Li, M., Vignat, J., Laversanne, M., Singh, D., La Vecchia, C., et al. (2022). The epidemiological landscape of thyroid cancer worldwide: GLOBOCAN estimates for incidence and mortality rates in 2020. *Lancet Diabetes Endocrinol.* 10 (4), 264–272. doi:10.1016/S2213-8587(22)00035-3

Rathod, M., Kelkar, M., Valvi, S., Salve, G., and De, A. (2020). FOXA1 regulation turns benzamide HDACi treatment effect-specific in BC, promoting NIS gene-mediated targeted radioiodine Therapy. *Mol. Ther. - Oncolytics* 19, 93–104. doi:10.1016/j.omto.2020.08.015

Rautava, S., Luoto, R., Salminen, S., and Isolauri, E. (2012). Microbial contact during pregnancy, intestinal colonization and human disease. *Nat. Rev. Gastroenterol. Hepatol.* 9 (10), 565–576. doi:10.1038/nrgastro.2012.144

Riaz Rajoka, M. S., Thirumdas, R., Mehwish, H. M., Umair, M., Khurshid, M., Hayat, H. F., et al. (2021). Role of food antioxidants in modulating gut microbial communities: novel understandings in intestinal oxidative stress damage and their impact on host health. *Antioxidants* 10 (10), 1563. doi:10.3390/antiox10101563

Ridlon, J. M., Kang, D. J., Hylemon, P. B., and Bajaj, J. S. (2014). Bile acids and the gut microbiome. *Curr. Opin. Gastroenterology* 30 (3), 332–338. doi:10.1097/MOG.0000000000000057

Russo, E., Cinci, L., Di Gloria, L., Baldi, S., D'Ambrosio, M., Nannini, G., et al. (2022). Crohn's disease recurrence updates: first surgery vs. surgical relapse patients display different profiles of ileal microbiota and systemic microbial-associated inflammatory factors. *Front. Immunol.* 13, 886468. doi:10.3389/fimmu.2022.886468

Samimi, H., and Haghpanah, V. (2020). Gut microbiome and radioiodine-refractory papillary thyroid carcinoma pathophysiology. *Trends Endocrinol. Metabolism* 31 (9), 627–630. doi:10.1016/j.tem.2020.03.005

Servin, A. L. (2004). Antagonistic activities of lactobacilli and bifidobacteria against microbial pathogens. *FEMS Microbiol. Rev.* 28 (4), 405–440. doi:10.1016/j.femsre.2004.01.003

Shen, C. T., Zhang, Y., Liu, Y. M., Yin, S., Zhang, X. Y., Wei, W. J., et al. (2017). A distinct serum metabolic signature of distant metastatic papillary thyroid carcinoma. *Clin. Endocrinol.* 87 (6), 844–852. doi:10.1111/cen.13437

Singh, A., Ashar, H., Butcher, J. T., and Ranjan, A. (2023). Age-associated changes in the gut microbiome impact efficacy of tumor immunomodulatory treatments. *Exp. Gerontol.* 181, 112268. doi:10.1016/j.exger.2023.112268

Sokol, H., Pigneur, B., Watterlot, L., Lakhdari, O., Bermúdez-Humarán, L. G., Gratadoux, J. J., et al. (2008). *Faecalibacterium prausnitzii* is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. *Proc. Natl. Acad. Sci. U. S. A.* 105 (43), 16731–16736. doi:10.1073/pnas.0804812105

Stebbings, S., Munro, K., Simon, M. A., Tannock, G., Highton, J., Harmsen, H., et al. (2002). Comparison of the faecal microflora of patients with ankylosing spondylitis and controls using molecular methods of analysis. *Rheumatology* 41 (12), 1395–1401. doi:10.1093/rheumatology/41.12.1395

Vitale, G., Dicitore, A., Barrea, L., Sbardella, E., Razzore, P., Campione, S., et al. (2021). From microbiota toward gastro-enteropancreatic neuroendocrine neoplasms: are we on the highway to hell? *Rev. Endocr. Metab. Disord.* 22 (3), 511–525. doi:10.1007/s11154-020-09589-y

Von Itzstein, M. S., Khan, S., and Gerber, D. E. (2020). Investigational biomarkers for checkpoint inhibitor immune-related adverse event prediction and diagnosis. *Clin. Chem.* 66 (6), 779–793. doi:10.1093/clinchem/hvaa081

Wang, H., Naghavi, M., Allen, C., Barber, R. M., Bhutta, Z. A., Carter, A., et al. (2016). Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980–2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet* 388 (10053), 1459–1544. doi:10.1016/S0140-6736(16)31012-1

Wang, X., and Huycke, M. M. (2007). Extracellular superoxide production by *Enterococcus faecalis* promotes chromosomal instability in mammalian cells. *Gastroenterology* 132 (2), 551–561. doi:10.1053/j.gastro.2006.11.040

WHO (2020). IARC. Available from: https://gco.iarc.fr/today/data/factsheets/populations/620-portugal-fact-sheets.pdf (Accessed September 4, 2023).

Wong, J. M. W., De Souza, R., Kendall, C. W. C., Emam, A., and Jenkins, D. J. A. (2006). Colonic health: fermentation and short chain fatty acids. *J. Clin. Gastroenterology* 40 (3), 235–243. doi:10.1097/00004836-200603000-00015

Wu, M., Li, P., Li, J., An, Y., Wang, M., and Zhong, G. (2020). The differences between luminal microbiota and mucosal microbiota in mice. *J. Microbiol. Biotechnol.* 30 (2), 287–295. doi:10.4014/jmb.1908.08037

Wu, S., Powell, J., Mathioudakis, N., Kane, S., Fernandez, E., and Sears, C. L. (2004). Bacteroides fragilis enterotoxin induces intestinal epithelial cell secretion of interleukin-8 through mitogen-activated protein kinases and a tyrosine kinase-regulated nuclear factor-kappaB pathway. *Infect. Immun.* 72 (10), 5832–5839. doi:10.1128/IAI.72.10.5832-5839.2004

Wu, S., Rhee, K. J., Albesiano, E., Rabizadeh, S., Wu, X., Yen, H. R., et al. (2009). A human colonic commensal promotes colon tumorigenesis via activation of T helper type 17 T cell responses. *Nat. Med.* 15 (9), 1016–1022. doi:10.1038/nm.2015

Yao, C. Z., Zhang, M., Zeng, Y. K., Zhang, Y. Y., Wu, X., Xiong, W. J., et al. (2023). Analysis and prediction of thyroid cancer morbidity and mortality trends in China. *Zhonghua Liu Xing Bing Xue Za Zhi* 44 (6), 917–923. doi:10.3760/cma.j.cn112338-20221010-00869

Young, V. B., and Schmidt, T. M. (2004). Antibiotic-associated diarrhea accompanied by large-scale alterations in the composition of the fecal microbiota. *J. Clin. Microbiol.* 42 (3), 1203–1206. doi:10.1128/jcm.42.3.1203-1206.2004

Yu, A. I., Zhao, L., Eaton, K. A., Ho, S., Chen, J., Poe, S., et al. (2020). Gut microbiota modulate CD8 T cell responses to influence colitis-associated tumorigenesis. *Cell Rep.* 31 (1), 107471. doi:10.1016/j.celrep.2020.03.035

Zhang, J., Zhang, F., Zhao, C., Xu, Q., Liang, C., Yang, Y., et al. (2019). Dysbiosis of the gut microbiome is associated with thyroid cancer and thyroid nodules and correlated with clinical index of thyroid function. *Endocrine* 64 (3), 564–574. doi:10.1007/s12020-018-1831-x

Zhao, F., Feng, J., Li, J., Zhao, L., Liu, Y., Chen, H., et al. (2018). Alterations of the gut microbiota in Hashimoto's thyroiditis patients. *Thyroid* 28 (2), 175–186. doi:10.1089/thy.2017.0395

Zhou, L., Zhang, M., Wang, Y., Dorfman, R. G., Liu, H., Yu, T., et al. (2018). Faecalibacterium prausnitzii produces butyrate to maintain Th17/treg balance and to ameliorate colorectal colitis by inhibiting histone deacetylase 1. *Inflamm. Bowel Dis.* 24 (9), 1926–1940. doi:10.1093/ibd/izy182

Zhu, H. Z., Liang, Y. D., Ma, Q. Y., Hao, W. Z., Li, X. J., Wu, M. S., et al. (2019). Xiaoyaosan improves depressive-like behavior in rats with chronic immobilization stress through modulation of the gut microbiota. *Biomed. Pharmacother.* 112, 108621. doi:10.1016/j.biopha.2019.108621

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 13 December 2023
DOI 10.3389/fgene.2023.1280378

OPEN ACCESS

EDITED BY
Ziheng Wang,
University of Macau, China

REVIEWED BY
Rafael Katayama,
Federal University of São Paulo, Brazil
Tomas Drgon,
United States Food and Drug Administration, United States

*CORRESPONDENCE
Haoyu Wang,
surgeonwanghaoyu@mail.xjtu.edu.cn

RECEIVED 20 August 2023
ACCEPTED 28 November 2023
PUBLISHED 13 December 2023

CITATION
Wang H, Chen Z, Dang X and Wang H (2023), Rheumatoid arthritis and gastroesophageal reflux disease: a bidirectional and multivariable two-sample Mendelian randomization study. *Front. Genet.* 14:1280378. doi: 10.3389/fgene.2023.1280378

Rheumatoid arthritis and gastroesophageal reflux disease: a bidirectional and multivariable two-sample Mendelian randomization study

Haifan Wang, Zhihao Chen, Xiaoqian Dang and Haoyu Wang*

Department of Orthopaedics, The Second Affiliated Hospital of Xi'an Jiaotong University, Xi'an, Shaanxi, China

Aims/hypothesis: The association between gastroesophageal reflux disease (GERD) and rheumatoid arthritis (RA) has been reported by many observational studies in the Asian population. This study aimed to examine the bidirectional causal effects between GERD and RA by two-sample Mendelian randomization (MR) analyses using genetic evidence.

Methods: Two-sample Mendelian randomization analyses were performed to determine the causal effect of GERD (129,080 cases vs. 602,604 control participants) on RA (6,236 cases vs. 147,221 control participants) and RA on GERD, respectively. The inverse-variance weighted (IVW) method was used as the primary analysis. Weighted median and MR-Egger regression were taken as supplementary analyses. Cochran's Q test evaluated the heterogeneity. Horizontal pleiotropy was detected by estimating the intercept term of MR-Egger regression. Furthermore, multivariable MR analyses were performed to exclude the influence of confounding factors, including the years of schooling, BMI, and time spent watching television, between GERD and RA.

Result: Both univariate MR (UVMR) and multivariable MR (MVMR) provided valid evidence that RA was causally and positively influenced by GERD (UVMR: OR = 1.49, 95% CI = 1.25–1.76, $p = 6.18*10^{-6}$; MVMR: OR = 1.69, 95% CI = 1.24–2.31, $p = 8.62*10^{-4}$), whereas GERD was not influenced by RA (UVMR: OR = 1.03, 95% CI = 1.00–1.06, $p = 0.042$; MVMR: OR = 1.04, 95% CI = 1.00–1.07, $p = 0.0271$).

Conclusion: Our comprehensive bidirectional MR analysis found that for the European population, GERD can induce the occurrence of RA (OR = 1.69, $p < 0.00125$), whereas RA only has no significant influence on GERD. In particular, patients with GERD are suffering a 69% increased risk of RA occurrence, which means GERD is a substantial risk factor for RA.

KEYWORDS

rheumatoid arthritis, gastroesophageal reflux disease, Mendelian randomization, education, BMI

Introduction

Rheumatoid arthritis (RA) mainly affects the joints, with extra-articular tissues being involved (Smolen et al., 2016). The incidence of RA is estimated to be approximately 0.1–0.5 per 1,000 person/year, which varies according to the ethnic group (Tobón et al., 2010). The extra-articular manifestations revealed the existence of systemic inflammation in RA (Smolen et al., 2022). Apart from inherited susceptibility, low socioeconomic status, periodontal diseases, and microbiome are also the risk factors of RA (Millar et al., 2013; Scher et al., 2016; Li et al., 2017). Those risk factors may influence systemic inflammation to induce RA.

Gastroesophageal reflux disease (GERD) is a common disease that results from the reverse flow of stomach acid into the esophagus (Vakil et al., 2006). Nearly one-fifth of North American people are suffering from GERD, which is nearly four times more prevalent than in Asian populations. This disease causes distressing symptoms, such as heartburn, inappetence, nausea, and susceptibility to pharyngitis, and some other diseases (Punjabi et al., 2015). Mechanically, GERD is primarily caused by abnormal physiology and anatomy changes in the stomach and esophagus. These changes include an increased pressure gradient between the abdomen and thorax, dysmotility of the esophagus, hiatus musculature, and/or the stomach. As a result, the normal reflux barrier of the LES breaks down (Mikami and Murayama, 2015).

In addition to anatomical and physiological factors, inflammation occurring in the stomach and esophagus plays a crucial role in the development of GERD (Punjabi et al., 2015; Souza et al., 2017; Surdea-Blaga et al., 2019). Consistently, an association between GERD and systemic inflammatory diseases has been reported (Chen, 2015; Linz et al., 2017), particularly for RA (Cryer et al., 2011; Nampei et al., 2013; Lin et al., 2017). In detail, some observational research found a higher incidence of GERD in patients with RA (Cryer et al., 2011; Lin et al., 2017). In Japan, more than 2-fold higher incidence of GERD in patients with RA than that in normal people has been reported (Lin et al., 2017). Meanwhile, it has also been reported that patients with GERD exhibit a nearly 3-fold higher risk of RA than those in the control group in the Taiwan population (Lin et al., 2017). The bidirectional association between GERD and RA has been identified. One cohort study performed on the Asian population reported an HR of 1.49 for RA in patients with GERD and 1.46 for GERD in patients with RA (Kim et al., 2021).

Mendelian randomization (MR) analysis utilizes single-nucleotide polymorphisms (SNPs) to find the causality between risk factors and outcomes (Lawlor et al., 2008). The superiority of MR lies in the fact that SNPs are determined before the intervention of the environment. Thus, they can be used as proxies for phenotypes and diseases (Davey Smith and Hemani, 2014). Due to its reduced susceptibility to reverse causation and confounding, MR conclusions are considered more reliable than those of conventional observational studies (Davey Smith and Hemani, 2014).

The prerequisites for MR are based on three assumptions: first, IVs should be strongly associated with exposure; second, IVs should influence the outcome only through the exposure (no horizontal pleiotropy); and third, IVs should not be associated with confounders. The two-sample MR tests the causality based on the GWAS data risk factors, and outcomes are measured in their respective samples (Boef et al., 2015). After searching in reference databases, we found that MR has not been applied to explore the causal effects between GERD and RA.

However, the relationship between GERD and RA has not been observed in European population and evaluated using the MR method. In this study, univariate and multivariable bidirectional two-sample MR analyses were performed to test the reciprocal causal relationship between GERD and RA.

Methods

Study design

As shown in Figure 1, first, the univariate bidirectional MR analysis of the causal relationship between GERD and RA was performed. When exposure was set as GERD, RA was considered to be the outcome. When exposure was set as RA, GERD was considered to be the outcome. Furthermore, confounding factors between GERD and RA were retrieved. In detail, the mutual potential exposure relationship between GERD and RA was found by batching the processing TwoSample-MR R script. Then, three confounding factors were included through the automatic exposure finding R script. Finally, three confounding factors are selected: years of schooling, BMI, and time spent watching television.

For UVMR, the exposure SNPs ($p < 5 \times 10^{-8}$, r2 < 0.001, F > 10 for GERD and $p < 1 \times 10^{-5}$, r2<0.001, F > 10 for RA) were selected as instrumental variables. Furthermore, sensitivity and pleiotropy analyses were performed to ensure the robustness of the results. For multivariable MR (MVMR), the selection threshold of the mv_extract_exposures function in the TwoSample-MR package was set as default.

Data source

All data involved are publicly available. The RA GWAS data including 153,457 European individuals (6,236 cases vs.

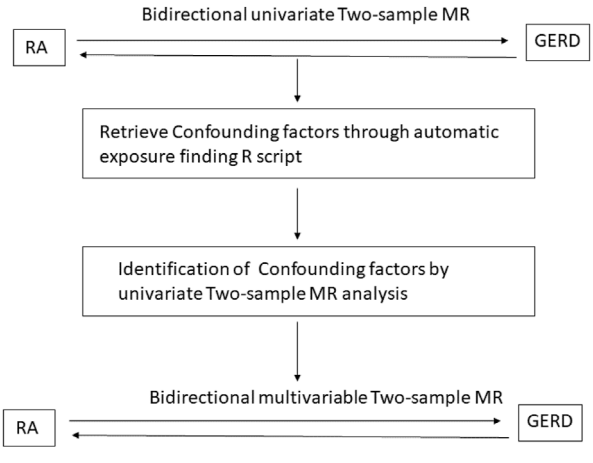


FIGURE 1
Workflow of this study design is shown in the sketch map.

147,221 control participants) retrieved from the IEU database (id: finn-b-M13_RHEUMA) were originally derived from the Finn-gen Consortium [Trait: Duration of vigorous activity – IEU OpenGWAS project. 2021. https://gwas.mrcieu.ac.uk/datasets/ukb-b-13932/ (1 December 2020, date last accessed), n.d.]. The GERD GWAS data including 129,080 cases and 602,604 control of European individuals were also retrieved from the IEU database (id: ebi-a-GCST90000514) originally derived from the EBI Consortium. The F-statistic of SNPs was calculated by the formula to select strong IVs ($F = R2\times (N-2)/(1-R2)$ (Chen et al., 2022). Then, we selected SNPs with an F-statistic larger than 10 to prevent potential weak instrument bias. Three confounding factors from GWAS data were retrieved from the IEU website: body mass index (id:ieu-a-94), years of schooling (id:ieu-a-1239), and time spent watching TV (id:ukb-b-5192).

Instrumental variable selection

SNPs were filtrated by the TwoSampleMR packages of R software. Genome-wide SNPs that are closely associated with education duration were acquired by the extract_instruments function (thresholds were set as $p < 5 \times 10^{-8}$, r2 < 0.001, window size = 10000 kb for GERA, and $p < 1 \times 10^{-5}$, r2<0.001, window size = 10000 kb for RA).

Statistical analyses

The inverse variance-weighted (IVW) method was considered as the main MR analysis to initially estimate the causal relationship of education duration on joint pain and sciatica with lumbago. The IVW method's robustness depends on IV's pleiotropy. Furthermore, another two MR analyses, namely, weighted median (WM) and MR-Egger, were selected as supplementary analyses to detect causalities. The WM method can estimate unbiased causality, with more than 50% of the weight coming from valid instrumental variables (Bowden et al., 2016), whereas MR-Egger estimates consistently account for pleiotropy when all IVs are invalid with the lowest power (Bowden et al., 2015). Our MR estimates of the risk of GERD or RA were presented as follows: odds ratio (OR), 95% confidence interval [CI]. A two-sided value of $p < 0.05$ is considered statistically significant for UVMR and $p < 0.0125$ for MVMR (four exposures).

Sensitivity analysis

Cochran's Q test, MR-Egger intercept tests, leave-one-out (LOO) analyses, and funnel plots were performed to examine the presence of pleiotropy in the results. In particular, Cochran's Q test was applied to evaluate heterogeneity, which was detected if the *p*-value was less than 0.05. The horizontal pleiotropy of both UVMR and MVMR was appraised by estimating the intercept term derived from MR-Egger regression. The LOO analysis was performed to detect any pleiotropy driven by a single SNP. All these MR analyses were performed using the TwoSampleMR package in R.

Results

Univariate MR result of GERD on RA

The UVMR results of education duration on joint pain are shown in Figures 2, 3. A total of 75 SNPs were selected as instrumental variables. Given the IVW method, RA was casually influenced by GERD (OR = 1.49, 95% confidence interval [CI] = 1.25–1.76, $p = 6.18*10^{-6}$), suggesting that patients with GERD are suffering a 49% increased risk of RA occurrence. This result was consistent with the weighted median (OR = 1.49, 95% CI = 1.20–1.86, $p = 4.00*10^{-4}$). Heterogeneity was not found in the effect of GERD on RA using Cochran's Q test ($p = 0.794$), and directional pleiotropy is not existent in the SNPs associated with GERD *via* MR-Egger regression (intercept = −0.023 $p = 0.3531$). The result of leave-one-out analyses shows that the global effect of GERD on RA was not dependent on any single IV. The symmetrical funnel plots suggested that there was no significant bias in SNP selection.

Univariate MR result of RA on GERD

The UVMR results of RA on GERD are shown in Figure 3. Only 15 SNPs were available as instrumental variables. Even the IVW method results show a *p*-value that is slightly lower than 0.05; effect value β is very small (0.03–0.04) in all three methods. For the IVW method, the GERD was slightly casually influenced by RA (OR = 1.03, 95% [CI] = 1.00–1.06, $p = 0.0844$), suggesting that both weighted median (OR = 1.03, 95% CI = 1.00–1.06, $p = 0.2793$) and MR Egger (OR = 1.04, 95% CI = 0.97–1.12, $p = 0.2793$) analyses did not support the above results. Heterogeneity was not found in the effect of RA on GERD using Cochran's Q test ($p = 0.0.52$), and directional pleiotropy is non-existent in SNPs associated with RA *via* MR-Egger regression (intercept = −0.0016, $p = 0.7486$). In other words, there is no strong enough evidence to support that RA can induce the occurrence of GERD.

Confounding factors and their effects

For further MVMR analysis, confounding factors are retrieved through automatic exposure finding the R script, and the two-sample MR results of those confounding factors are shown in Figure 4. It implies that all the three factors that were selected are potential confounding factors with positive two-sample UVMR analysis results.

Multivariable MR analysis between GERD and RA

The bidirectional MVMR results are shown in Figure 5. After removing the influence of confounding factors, the causal relationship of GERD on RA still exists (OR = 1.69, 95% CI = 1.24–2.31, $p < 0.0125$). Consistent with UVMR, RA only has slight effects on GERD (OR = 1.04, 95% CI = 1.00–1.07, $p = 0.0271$). It is worth mentioning that the years of schooling is an effective protective factor for GERD (OR = 0.36, 95% CI = 0.31–0.41, $p < 0.0125$), but BMI (OR = 1.27, 95% CI = 1.18–1.37, $p < 0.0125$) and

10.3389/fgene.2023.1280378

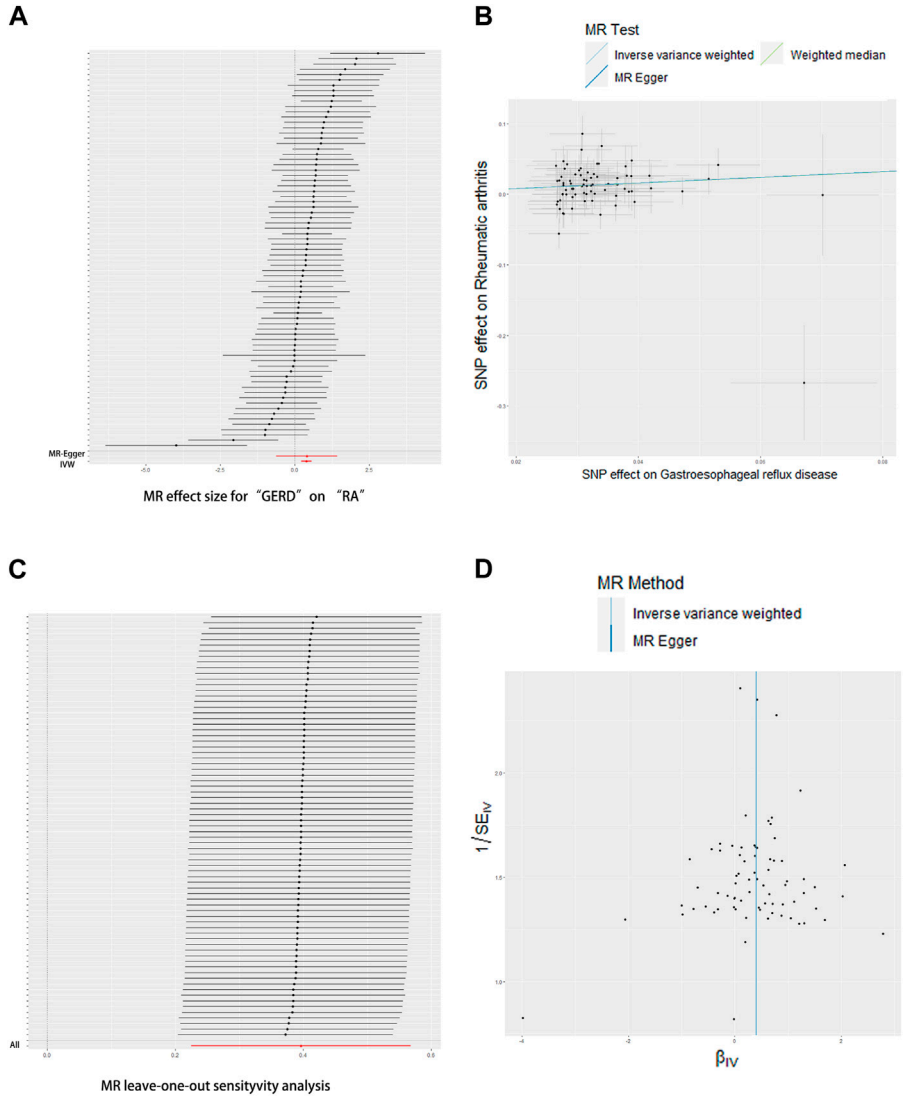


FIGURE 2
Effects of GERD on RA. **(A)** Forrest plot, **(B)** scatterplot, **(C)** leave-one-out plot, and **(D)** funnel plot.

Outcome	Exposure	nSNP	β	SE	Method	OR(95%CI)	P
RA	GERD	75	0.39627773	0.08767013	Inverse variance weighted	1.49(1.25 to 1.76)	6.181049e-06
		75	0.39947676	0.11285577	Weighted median	1.49(1.20 to 1.86)	4.005661e-04
		75	0.40101729	0.52102587	MR Egger	1.49(0.54 to 4.15)	4.439799e-01
GERD	RA	15	0.03001066	0.01473625	Inverse variance weighted	1.03(1.00 to 1.06)	4.169819e-02
		15	0.02926265	0.01695598	Weighted median	1.03(1.00 to 1.06)	8.438320e-02
		15	0.04071048	0.03605970	MR Egger	1.04(0.97 to 1.12)	2.793118e-01

FIGURE 3
Univariate two-sample MR analysis between GERD and RA.

time spent watching TV (OR = 1.72, 95% CI = 1.33–2.24, $p < 0.0125$) are risk factors for GERD. Directional pleiotropy was not detected in both MVMR of GERD (intercept = −0.002, $p = 0.668$) on RA and RA on GERD (intercept = 0.001, $p = 0.130$).

Discussion

In order to examine the potential reciprocal causal relationship between GERD and RA, bidirectional two-sample Mendelian

Outcome	Exposure	Method	β	SE	nsnp	OR(95%CI)	P
GERD	Time spent watching television	Inverse variance weighted	1.4573033	0.09347769	80	4.29(3.58 to 5.16)	8.530000e-55
	Time spent watching television	Weighted median	1.2235304	0.09696121	80	3.40(2.81 to 4.11)	1.660000e-36
	Time spent watching television	MR Egger	2.0335127	0.45979542	80	7.64(3.10 to 18.82)	3.120000e-05
	Years of schooling	Inverse variance weighted	-1.2828368	0.03974094	245	0.28(0.26 to 0.30)	1.340000e-228
	Years of schooling	Weighted median	-1.2430003	0.04381422	245	0.29(0.26 to 0.31)	4.770000e-177
	Years of schooling	MR Egger	-1.0412783	0.16570313	245	0.35(0.26 to 0.49)	1.510000e-09
	BMI	Inverse variance weighted	0.3817575	0.05075744	10	1.46(1.33 to 1.62)	5.430000e-14
	BMI	Weighted median	0.3839172	0.05252976	10	1.47(1.32 to 1.63)	2.700000e-13
	BMI	MR Egger	0.3644464	0.19753968	10	1.44(0.98 to 2.12)	1.022695e-01
RA	Time spent watching television	Inverse variance weighted	0.7928204	0.24601120	110	2.21(1.36 to 3.58)	1.269883e-03
	Time spent watching television	Weighted median	0.5578978	0.27465074	110	1.75(1.02 to 2.99)	4.222466e-02
	Time spent watching television	MR Egger	0.3778821	1.16053416	110	1.46(0.15 to 14.19)	7.453491e-01
	Years of schooling	Inverse variance weighted	-0.5028612	0.10311223	305	0.60(0.49 to 0.74)	1.080000e-06
	Years of schooling	Weighted median	-0.5027957	0.14562060	305	0.60(0.45 to 0.80)	5.548450e-04
	Years of schooling	MR Egger	-0.2470868	0.40367069	305	0.78(0.35 to 1.72)	5.409306e-01
	BMI	Inverse variance weighted	0.3072191	0.13783106	10	1.36(1.04 to 1.78)	2.581698e-02
	BMI	Weighted median	0.2675541	0.18292699	10	1.31(0.91 to 1.87)	1.435693e-01
	BMI	MR Egger	0.2066596	0.49833439	10	1.23(0.46 to 3.27)	6.892584e-01

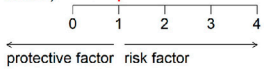


FIGURE 4
Univariate two-sample MR analysis of confounding factors' effects on RA and GERD.

Outcome	Exposure	β	SE	nSNP	OR(95%CI)	P
GERD	Rheumatoid arthritis	0.0369552002	0.01672355	2	1.04(1.00 to 1.07)	2.712110e-02
	Years of schooling	-1.0303860957	0.06927440	215	0.36(0.31 to 0.41)	4.863417e-50
	Body mass index	0.2419840178	0.03743548	4	1.27(1.18 to 1.37)	1.019512e-10
	Time spent watching television	0.5440836354	0.13296823	45	1.72(1.33 to 2.24)	4.279789e-05
RA	GERD	0.5264113848	0.15799048	34	1.69(1.24 to 2.31)	8.624955e-04
	Years of schooling	0.2174523952	0.26267055	215	1.24(0.74 to 2.08)	4.077542e-01
	Body mass index	-0.0001101332	0.11635157	4	1.00(0.80 to 1.26)	9.992448e-01
	Time spent watching television	-0.0886942078	0.38969014	43	0.92(0.43 to 1.96)	8.199558e-01

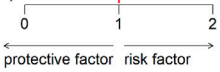


FIGURE 5
Result of MVMR analysis between GERD and RA.

randomization analyses were conducted. The results suggested that GERD can induce the occurrence of RA, whereas RA has no significant impact on GERD. In particular, individuals with GERD are at a 69% higher risk of developing RA, highlighting GERD as a significant risk factor for this condition. In addition, the impact of three confounding variables, namely, educational attainment, BMI, and duration of television viewing, on both GERD and RA has also been identified.

The association between GERD and RA has been reported by many observational studies, as mentioned in the Introduction part. Those observational studies reported bidirectional causal relationships between GERD and RA. However, our bidirectional two-sample Mendelian randomization analyses only identified the causal effects of GERD on RA and not *vice versa*.

There are some potential common risk factors for both GERD and RA. Inclusion bias may arise when the observational studies fail to exclude those mutual factors (Thombs et al., 2011). In order to incorporate confounding factors effectively, this study conducted a preliminary two-sample MR analysis with GERD and RA as separate outcomes. This step aimed to screen for common risk factors before proceeding to the MVMR analysis. Considering the evidence level of observational studies, MR analysis, and the superiority of MVMR (Zhou et al., 2023), results of this study are influenced by fewer confounding factors. When cross-sectional and cohort studies found the reciprocal association between GERD and RA, those known and unknown confounding factors are often not excluded.

For instance, a cohort study identifying the reciprocal association between GERD and RA only matched factors, including age, group, sex, income group, income group, and region of residence (Kim et al., 2021). However, smoking, BMI, diabetes, *etc.*, are not taken into account, which can result in selection bias (Punjabi et al., 2015). The control group selected in those studies often exhibits fewer risk factors of GERD that are not caused by RA. Coincidentally, the HR (1.49) of RA in patients with GERD calculated by one cohort study was approximate to our UVMR result, but the HR (1.46) of GERD in patients with RA is not found in our study (Kim et al., 2021). Recently, a meta-analysis research based on cohort studies has also reported an OR of 1.98 for GERD in patients with RA (Thongpiya et al., 2023). It is important to note that the observational studies utilized in the meta-analysis were predominantly conducted on the Asian population, whereas our MR analysis is based on the European population. In addition, the prevalence of GERD in North America is nearly 4-fold in the Asian population (Chen, 2015). This difference in population may also

account for the difference between our study and former observational research.

On the other hand, patients with RA tend to take more NSAIDS than their control groups because of the symptom of RA (Burmester and Pope, 2017; Aletaha and Smolen, 2018; Ben Mrid et al., 2022). Meanwhile, the common adverse reaction of NSAIDS is gastrointestinal, which may induce GERD (Altman et al., 2015; García-Rayado et al., 2018; Bindu et al., 2020). As a result, it is natural to hypothesize about the causal effects of RA on GERD. However, our results found no significant effect of RA on GERD, unlike the obvious effect reported by previous observational studies.

Similarly, for the causal effect of GERD on RA, ORs 1.49 and 1.69 calculated in this study were smaller than the HRs reported by other observational studies. For instance, a nearly 3-fold risk of RA susceptibility in patients with GERD than their control groups was reported by a nested case–control study in the Asian population (Lin et al., 2017). Even this study has considered many known confounding factors, including hypertension, diabetes, smoking, hyperlipidemia, obesity, stroke, and coronary heart disease; when those factors ($p > 0.1$ between patients with GERD and their controls) combined, inclusion bias may also derive.

Mechanistically speaking, the physiological and anatomical changes, which may not appear directly associated with RA, have the potential to increase the risk of developing RA due to the persistent inflammation and immune dysregulation observed in GERD. However, further basic and clinical studies are required to substantiate these assumptions. In addition, the MVMR analysis also revealed that years of schooling is an effective protective factor for GERD (OR = 0.36, 95% CI = 0.31–0.41, $p < 0.0125$). Conversely, BMI (OR = 1.27, 95% CI = 1.18–1.37, $p < 0.0125$) and time spent watching TV (OR = 1.72, 95% CI = 1.33–2.24, $p < 0.0125$) are identified as risk factors for GERD. The education duration has been consistently reported as a protective factor for many diseases and phenotypes, including low back pain, RA, and lifestyle (Jiang et al., 2015; Saper et al., 2017; Davies et al., 2019; Kari et al., 2020; Zhao et al., 2022), while excessive high BMI and longer time spent watching TV have been found to be detrimental to health in many studies (Antonopoulos et al., 2016; Caballero, 2019; Raichlen et al., 2022; Sun et al., 2023). Although the effects of BMI and education on RA are detected in the search of confounding factors by UVMR searching, the MVMR analysis did not find significant effects of them on RA. This finding may support the notion that the previously reported effects of BMI and education on RA were also generated by bias factors.

The strength of this study is that the confounding factors are included in the identification of bidirectional causal relationships between GERD and RA through MR analysis. To the best of our knowledge, it is the first time to investigate their association using genetic evidence, despite previous observational studies reporting a bidirectional association between RA and GERD (Miura et al., 2014).

However, there are several limitations to this study. The GWAS data used in this study were derived from the European population, whereas most previous observational studies on the topic are based on the Asian population. Whether our findings are generalizable to non-European populations still needs to be confirmed. In addition, even though the sample of GWAS data used in this study was large, more extensive and new GWAS data may produce different conclusions.

Conclusion

In conclusion, our bidirectional MR analysis found that for the European population, GERD can induce the occurrence of RA (OR = 1.69, $p < 0.00125$), whereas RA only has no significant influence on GERD (OR = 1.04, $p > 0.0125$). In particular, European GERD patients are suffering a 69% increased risk of RA occurrence, which means GERD is a substantial risk factor for RA.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material; further inquiries can be directed to the corresponding author.

Ethics statement

Ethical approval was not required for the study involving humans in accordance with the local legislation and institutional requirements. Written informed consent to participate in this study was not required from the participants or the participants' legal guardians/next of kin in accordance with the national legislation and the institutional requirements.

Author contributions

HiW: formal analysis, methodology, software, validation, visualization, and writing–original draft. ZC: methodology, editing, formal analysis. XD: supervision and writing–review and editing. HoW: conceptualization, investigation, project administration, resources, and writing–review and editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This work was supported by grants from the NSFC (82202437).

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

References

Aletaha, D., and Smolen, J. S. (2018). Diagnosis and management of rheumatoid arthritis: a review. *JAMA* 320, 1360–1372. doi:10.1001/jama.2018.13103

Altman, R., Bosch, B., Brune, K., Patrignani, P., and Young, C. (2015). Advances in NSAID development: evolution of diclofenac products using pharmaceutical technology. *Drugs* 75, 859–877. doi:10.1007/s40265-015-0392-z

Antonopoulos, A. S., Oikonomou, E. K., Antoniades, C., and Tousoulis, D. (2016). From the BMI paradox to the obesity paradox: the obesity-mortality association in coronary heart disease. *Obes. Rev.* 17, 989–1000. doi:10.1111/obr.12440

Ben Mrid, R., Bouchmaa, N., Ainani, H., El Fatimy, R., Malka, G., and Mazini, L. (2022). Anti-rheumatoid drugs advancements: new insights into the molecular treatment of rheumatoid arthritis. *Biomed. Pharmacother.* 151, 113126. doi:10.1016/j.biopha.2022.113126

Bindu, S., Mazumder, S., and Bandyopadhyay, U. (2020). Non-steroidal anti-inflammatory drugs (NSAIDs) and organ damage: a current perspective. *Biochem. Pharmacol.* 180, 114147. doi:10.1016/j.bcp.2020.114147

Boef, A. G. C., Dekkers, O. M., and le Cessie, S. (2015). Mendelian randomization studies: a review of the approaches used and the quality of reporting. *Int. J. Epidemiol.* 44, 496–511. doi:10.1093/ije/dyv071

Bowden, J., Davey Smith, G., and Burgess, S. (2015). Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int. J. Epidemiol.* 44, 512–525. doi:10.1093/ije/dyv080

Bowden, J., Davey Smith, G., Haycock, P. C., and Burgess, S. (2016). Consistent estimation in mendelian randomization with some invalid instruments using a weighted median estimator. *Genet. Epidemiol.* 40, 304–314. doi:10.1002/gepi.21965

Burmester, G. R., and Pope, J. E. (2017). Novel treatment strategies in rheumatoid arthritis. *Lancet (London, Engl.* 389, 2338–2348. doi:10.1016/S0140-6736(17)31491-5

Caballero, B. (2019). Humans against obesity: who will win? *Adv. Nutr.* 10, S4–S9. doi:10.1093/advances/nmy055

Chen, D., Zhang, Y., Yidilisi, A., Xu, Y., Dong, Q., and Jiang, J. (2022). Causal associations between circulating adipokines and cardiovascular disease: a mendelian randomization study. *J. Clin. Endocrinol. Metab.* 107, e2572–e2580. doi:10.1210/clinem/dgac048

Chen, Y. (2015). Gastroesophageal reflux disease and non-digestive tract diseases. *Expert Rev. Gastroenterol. Hepatol.* 9, 685–692. doi:10.1586/17474124.2015.1012495

Cryer, B., Luo, X., Assaf, A. R., Sands, G., and Mardekian, J. (2011). Persistence with non-selective NSAIDs and celecoxib among patients with gastroesophageal reflux disease and osteoarthritis or rheumatoid arthritis. *Curr. Med. Res. Opin.* 27, 295–302. doi:10.1185/03007995.2010.542056

Davey Smith, G., and Hemani, G. (2014). Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum. Mol. Genet.* 23, R89–R98. doi:10.1093/hmg/ddu328

Davies, N. M., Hill, W. D., Anderson, E. L., Sanderson, E., Deary, I. J., and Smith, G. D. (2019). Multivariable two-sample mendelian randomization estimates of the effects of intelligence and education on health. *Elife* 8, 439900–e44022. doi:10.7554/eLife.43990

García-Rayado, G., Navarro, M., and Lanas, A. (2018). NSAID induced gastrointestinal damage and designing GI-sparing NSAIDs. *Expert Rev. Clin. Pharmacol.* 11, 1031–1043. doi:10.1080/17512433.2018.1516143

Jiang, X., Sandberg, M. E. C., Saevarsdottir, S., Klareskog, L., Alfredsson, L., and Bengtsson, C. (2015). Higher education is associated with a better rheumatoid arthritis outcome concerning for pain and function but not disease activity: results from the EIRA cohort and Swedish rheumatology register. *Arthritis Res. Ther.* 17, 317. doi:10.1186/s13075-015-0836-6

Kari, J. T., Viinikainen, J., Böckerman, P., Tammelin, T. H., Pitkänen, N., Lehtimäki, T., et al. (2020). Education leads to a more physically active lifestyle: evidence based on Mendelian randomization. *Scand. J. Med. Sci. Sport.* 30, 1194–1204. doi:10.1111/sms.13653

Kim, Y. H., Lee, J. W., Kim, Y., Bae, J. S., Kim, Y. J., Min, C., et al. (2021). Bidirectional association between GERD and rheumatoid arthritis: two longitudinal follow-up studies with a national sample cohort. *BMJ Open* 11, 1249–1257. doi:10.1136/bmjopen-2020-046283

Lawlor, D. A., Harbord, R. M., Sterne, J. A. C., Timpson, N., and Davey Smith, G. (2008). Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. *Stat. Med.* 27, 1133–1163. doi:10.1002/sim.3034

Li, R., Tian, C., Postlethwaite, A., Jiao, Y., Garcia-Godoy, F., Pattanaik, D., et al. (2017). Rheumatoid arthritis and periodontal disease: what are the similarities and differences? *Int. J. Rheum. Dis.* 20, 1887–1901. doi:10.1111/1756-185X.13240

Lin, H., Xirasagar, S., Lee, C., and Huang, C. (2017). The association between gastro-oesophageal reflux disease and subsequent rheumatoid arthritis occurrence: a nested case – control study from Taiwan. *BMJ Open*. 1–7. doi:10.1136/bmjopen-2017-016667

Linz, D., Hohl, M., Vollmar, J., Ukena, C., Mahfoud, F., and Böhm, M. (2017). Atrial fibrillation and gastroesophageal reflux disease: the cardiogastric interaction. *Europace* 19, 16–20. doi:10.1093/europace/euw092

Mikami, D. J., and Murayama, K. M. (2015). Physiology and pathogenesis of gastroesophageal reflux disease. *Surg. Clin. North Am.* 95, 515–525. doi:10.1016/j.suc.2015.02.006

Millar, K., Lloyd, S. M., McLean, J. S., Batty, G. D., Burns, H., Cavanagh, J., et al. (2013). Personality, socio-economic status and inflammation: cross-sectional, population-based study. *PLoS One* 8, e58256. doi:10.1371/journal.pone.0058256

Miura, Y., Fukuda, K., Maeda, T., and Kurosaka, M. (2014). Gastroesophageal reflux disease in patients with rheumatoid arthritis. *Mod. Rheumatol.* 24, 291–295. doi:10.3109/14397595.2013.843749

Nampei, A., Shi, K., Ebina, K., Tomita, T., Sugamoto, K., Yoshikawa, H., et al. (2013). Prevalence of gastroesophageal reflux disease symptoms and related factors in patients with rheumatoid arthritis. *J. Clin. Biochem. Nutr.* 52, 179–184. doi:10.3164/jcbn.12-83

Punjabi, P., Hira, A., Prasad, S., Wang, X., and Chokhavatia, S. (2015). Review of gastroesophageal reflux disease (GERD) in the diabetic patient. *J. Diabetes* 7, 599–609. doi:10.1111/1753-0407.12279

Raichlen, D. A., Klimentidis, Y. C., Sayre, M. K., Bharadwaj, P. K., Lai, M. H. C., Wilcox, R. R., et al. (2022). Leisure-time sedentary behaviors are differentially associated with all-cause dementia regardless of engagement in physical activity. *Proc. Natl. Acad. Sci. U. S. A.* 119, e2206931119. doi:10.1073/pnas.2206931119

Saper, R. B., Lemaster, C., Delitto, A., Sherman, K. J., Herman, P. M., Sadikova, E., et al. (2017). Yoga, physical therapy, or education for chronic low back pain: a randomized noninferiority trial. *Ann. Intern. Med.* 167, 85–94. doi:10.7326/M16-2579

Scher, J. U., Littman, D. R., and Abramson, S. B. (2016). Microbiome in inflammatory arthritis and human rheumatic diseases. *Arthritis Rheumatol. Hob. N.J.)* 68, 35–45. doi:10.1002/art.39259

Smolen, J. S., Aletaha, D., and McInnes, I. B. (2016). Rheumatoid arthritis. *Lancet* 388, 2023–2038. doi:10.1016/S0140-6736(16)30173-8

Smolen, J. S., Feist, E., Fatenejad, S., Grishin, S. A., Korneva, E. V., Nasonov, E. L., et al. (2022). Olokizumab versus placebo or adalimumab in rheumatoid arthritis. *N. Engl. J. Med.* 387, 715–726. doi:10.1056/nejmoa2201302

Souza, R. F., Bayeh, L., Spechler, S. J., Tambar, U. K., and Bruick, R. K. (2017). A new paradigm for GERD pathogenesis. Not acid injury, but cytokine-mediated inflammation driven by HIF-2α: a potential role for targeting HIF-2α to prevent and treat reflux esophagitis. *Curr. Opin. Pharmacol.* 37, 93–99. doi:10.1016/j.coph.2017.10.004

Sun, Y., Chen, C., Yu, Y., Zhang, H., Tan, X., Zhang, J., et al. (2023). Replacement of leisure-time sedentary behavior with various physical activities and the risk of dementia incidence and mortality: a prospective cohort study. *J. Sport Heal. Sci.* 12, 287–294. doi:10.1016/j.jshs.2022.11.005

Surdea-Blaga, T., Negrutiu, D. E., Palage, M., and Dumitrascu, D. L. (2019). Food and gastroesophageal reflux disease. *Curr. Med. Chem.* 26, 3497–3511. doi:10.2174/0929867324666170515123807

Thombs, B. D., Arthurs, E., El-Baalbaki, G., Meijer, A., Ziegelstein, R. C., and Steele, R. J. (2011). Risk of bias from inclusion of patients who already have diagnosis of or are undergoing treatment for depression in diagnostic accuracy studies of screening tools for depression: systematic review. *BMJ* 343, d4825. doi:10.1136/bmj.d4825

Thongpiya, J., Charoenngam, N., Ponvilawan, B., Yingchoncharoen, P., Jaroenlapnopparat, A., and Ungprasert, P. (2023). Increased prevalence of gastroesophageal reflux disease among patients with rheumatoid arthritis: a systematic review and meta-analysis. *Curr. Rheumatol. Rev.* 19, 496–503. doi:10.2174/1573397119666230414103624

Tobón, G. J., Youinou, P., and Saraux, A. (2010). The environment, geo-epidemiology, and autoimmune disease: rheumatoid arthritis. *Autoimmun. Rev.* 9, A288–A292. doi:10.1016/j.autrev.2009.11.019

Trait: Duration of vigorous activity – IEU OpenGWAS project (2021). gwas. Available at: https://gwas.mrcieu.ac.uk/datasets/ukb-b-13932/ (Accessed December 1, 2020).

Vakil, N., van Zanten, S. V., Kahrilas, P., Dent, J., Jones, R., and Global Consensus Group (2006). The Montreal definition and classification of gastroesophageal reflux disease: a global evidence-based consensus. *Am. J. Gastroenterol.* 101, 1900–1920. doi:10.1111/j.1572-0241.2006.00630.x

Zhao, S. S., Holmes, M. V., Zheng, J., Sanderson, E., and Carter, A. R. (2022). The impact of education inequality on rheumatoid arthritis risk is mediated by smoking and body mass index: mendelian randomization study. *Rheumatol. (United Kingdom)* 61, 2167–2175. doi:10.1093/rheumatology/keab654

Zhou, W., Cai, J., Li, Z., and Lin, Y. (2023). Association of atopic dermatitis with autoimmune diseases: a bidirectional and multivariable two-sample mendelian randomization study. *Front. Immunol.* 14, 1132719. doi:10.3389/fimmu.2023.1132719

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 20 December 2023
DOI 10.3389/fgene.2023.1260367

OPEN ACCESS

EDITED BY
Xuanye Cao,
University of Texas MD Anderson Cancer Center, United States

REVIEWED BY
Ziheng Wang,
University of Macau, China
Yang Liu,
University of Texas MD Anderson Cancer Center, United States

*CORRESPONDENCE
Min Yang,
20204041@zcmu.edu.cn
Jun Shi,
shijundoctor@163.com

RECEIVED 17 July 2023
ACCEPTED 05 December 2023
PUBLISHED 20 December 2023

CITATION
Zhang Y-C, Lu C-D, Li Q-Y, Shi J-N, Shi J and Yang M (2023), Association between glycemic traits and melanoma: a mendelian randomization analysis. *Front. Genet.* 14:1260367. doi: 10.3389/fgene.2023.1260367

Association between glycemic traits and melanoma: a mendelian randomization analysis

Yun-Chao Zhang[1], Cen-Di Lu[2], Quan-Yao Li[1], Jin-Na Shi[3], Jun Shi[4]* and Min Yang[5]*

[1]Department of Oncology, Yueyang Hospital of Integrated Traditional Chinese and Western Medicine, Shanghai University of Traditional Chinese Medicine, Shanghai, China, [2]Department of Neurosurgery, The Second Affiliated Hospital of Zhejiang Chinese Medical University, Xinhua Hospital of Zhejiang Province, Hangzhou, Zhejiang, China, [3]Department of General Practice, KangQiao Campus of the Second Affiliated Hospital of Zhejiang Chinese Medical University, Xinhua Hospital of Zhejiang Province, Hangzhou, Zhejiang, China, [4]Department of Traditional Chinese Medicine, Shanghai Fourth People's Hospital Affiliated to Tongji University School of Medicine, Shanghai, China, [5]Department of Oncology, The Second Affiliated Hospital of Zhejiang Chinese Medical University, Xinhua Hospital of Zhejiang Province, Hangzhou, Zhejiang, China

Background: The causation of Glycemic Traits and risks of Melanoma remains unknown. We used Mendelian Randomization (MR) to assess the links between Glycemic Traits and Melanoma.

Method: Pooled data from Genome-Wide Association Studies (GWAS) were utilized to examine the relationships that exist between Fasting Insulin (n = 26), 2-h Glucose (n = 10), Fasting Glucose (n = 47), HbA1c (n = 68), and Type-2 Diabetes (n = 105) and Melanoma. We evaluated the correlation of these variations with melanoma risk using Two-Samples MR.

Result: In the IVW model, Fasting Glucose (OR = 0.99, 95%CI = 0.993–0.998, $p < 0.05$, IVW), Type-2 Diabetes (OR = 0.998, 95%CI = 0.998–0.999, $p < 0.01$, IVW) and HbA1c (OR = 0.19, 95%CI = 0.0415–0.8788, $p < 0.05$, IVW) was causally associated with a lower risk of Melanoma. In all models analyzed, there was no apparent causal relationship between Fasting Insulin and Melanoma risk. There was no obvious causal difference in the IVW analysis of 2-h Glucose and Melanoma, but its $p < 0.05$ in MR Egger (OR = 0.99, 95%CI = 0.9883–0.9984, $p < 0.05$, MR Egger), and the direction was consistent in other MR analyses, suggesting that there may be a causal relationship.

Conclusion: The results of this study suggest that a higher risk of Fasting Glucose, Type-2 Diabetes, 2-h Glucose, and HbA1c may be associated with a lower risk of Melanoma. However, no causal relationship between fasting insulin and melanoma was found. These results suggest that pharmacological or lifestyle interventions that regulate plasma glucose levels in the body may be beneficial in the prevention of melanoma.

KEYWORDS

two-sample mendelian randomization, glycemic traits, melanoma, genome-wide association study, causality

1 Introduction

Melanoma is a malignant tumor produced by the malignant transformation of melanocytes, which has a high probability of local spread and metastatic spread. Studies have shown that its incidence increases linearly in young and middle-aged people aged 25 to 50, and is high in people aged 57 (Carr et al., 2020). It is less common than other types of skin cancer but accounts for 73 percent of skin cancer-related deaths (Gershenwald and Guy, 2016). Studies have shown that in the next 10 years, the incidence and mortality of melanoma will continue to rise (Whiteman et al., 2016). The intervention effect of early surgical treatment and late radiotherapy and chemotherapy on the prognosis of patients is not satisfactory, and the results of several clinical trials have shown that the objective remission rate of patient's symptoms after treatment is less than 1% (O'Neill et al., 2006; Atzpodien et al., 2008; Bhatia et al., 2012). It is particularly important to look for risk factors to prevent the occurrence of Melanoma.

Several recent studies have shown that obesity is positively associated with the risk of melanoma (Dusingize et al., 2020; Larsson and Burgess, 2021). Obese adults have a higher prevalence of metabolic problems, such as insulin irregularities, hyperglycemia, and Type-2 Diabetes. Some researches has found that Type-2 Diabetes may be associated with an increased risk of Melanoma (Harding et al., 2015; Yuan et al., 2020). However, other cohort studies and case-control studies have found the opposite results (Qi et al., 2014; Malavolti et al., 2017). Previous prospective studies have shown that higher Fasting Glucose is closely related to the occurrence and development of Melanoma (Stattin et al., 2007). Recent studies have demonstrated a positive association between Glycemic Traits and the risk of colorectal cancer (Murphy et al., 2022a). But the MR studies of the associations between various Glycemic Traits and melanoma have not yet been reported.

This study used MR to explore the causal relationship between Glycemic Traits and Melanoma risk. MR is a comparable method to randomization in randomized controlled trials. When parents with two or more pairs of qualities cross when alleles are separated, genes on non-homologous chromosomes operate as free combinations, according to the law of independent assortment. Since germline genetic variation, and the random nature of allelic segregation is fixed at conception, MR analysis is less susceptible to traditional confounding and reverse causation. In this study, we used GWAS related to Fasting Glucose, Fasting Insulin, 2-h Glucose, HbA1c, and Type-2 Diabetes, GWAS data on Melanoma from risk on United Kingdom Biobank cohort study and FinnGen cohort study (Mahajan et al., 2018; Chen et al., 2021). Two-sample MR was used to explore the potential causal influence of the Glycemic Traits on the risk of Melanoma.

2 Methods

2.1 Study design

MR investigates the link between exposure and illness by employing genetic variation Single Nucleotide Polymorphisms (SNPs) as Instrumental Variables IV). IV was extracted from a disease-specific Genome-Wide Association Studies (GWAS) dataset for this investigation. The IV in this work should fulfill three criteria: there should be a high connection between IV and exposure, IV should only affect the outcomes through exposure, and IV should not have horizontal pleiotropy. Appropriate SNPs for usage as IVs must be strongly linked to malignancy ($p < 5 \times 10^{-8}$). To ensure independence, SNPs were restricted by low linkage disequilibrium (LD, $r^2 < 0.001$, window size = 10,000 kb) using clumping. By MR GWAS data. from different sources were analyzed to assess the causal relationship between glycemic signature and melanoma risk. Assess the strength of IV using the F statistic ($F = beta^2/se^2$), where β is the effect size of the allele and SE is the standard error (Feng et al., 2022). If F > 10, the correlation between IV and exposure was considered strong enough to protect the results of MR analysis from weak instrument bias. Meanwhile, it will ensure no confounders like UV radiation, light skin type, the presence of multiple atypical nevi, and a positive family history using the Phenoscanner (http://www.phenoscanner.medschl.cam.ac.uk/phenoscanner) website to have a search over each SNP (Rastrelli et al., 2014; Ugurel and Gutzmer, 2023).

2.2 Data source

Glycemic Traits data come from the largest GWAS to date (Glucose And Insulin-related Traits Consortium). A GWAS study of 2 h Glucose, Fasting Glucose, and Fasting Insulin included 63,396 (SNPs= 27,330,879), 200,622 (SNPs= 31,008,728), and 151,013 (SNPs= 29,664,438) participants of European ancestry, respectively (Chen et al., 2021). The GWAS for Type-2 Diabetes included 74,124 individuals with Type-2 Diabetes and 824,006 controls of European ancestry (SNPs= 21,000,000) (Mahajan et al., 2018). The HbA1c GWAS from the United Kingdom biobank (http://www.nealelab.is/uk-biobank) included 361,194 participants of European ancestry (SNPs= 1,048,575). The GWAS of Melanoma were obtained from United Kingdom biobank (https://www.ukbiobank.ac.uk/) and FinnGen (https://www.finngen.fi/en/access_results), the GWAS of United Kingdom biobank included 3,598 patients and 459,335 for controls (SNPs= 9,851,867) of European ancestry, the GWAS of FinnGen included 393 patients and 180,622 controls (SNPs= 16,380,337) of European ancestry. All study participants gave written informed consent, and the ethics committee approved all studies. SNPs data can be found in Supplementary Table S1.

2.3 Method selection

We estimated the relationship between Glycemic Traits and Melanoma risk using MR Egger, Inverse Variance Weighting (IVW), weighted Median, Simple Mode, and Weighted Mode MR methods. The IVW method assumes that all SNPs do not have horizontal pleiotropy (The impact of genetic variation on results is solely influenced by exposure of interest) and that all SNPs are effective tools. The fixed-effect inverse variance weighting (IVW) method was mainly used as the main analysis method (Kamiza et al., 2022). The intercept of the MR-Egger test was used to examine potential pleiotropic effects. Scatterplots are used to display the findings of several MR procedures. Odds ratios (ORs) and 95% confidence intervals (CIs) were used to represent the causal effects of overall and Melanoma. We utilized scatterplots

10.3389/fgene.2023.1260367

TABLE 1 Power calculation for Mendelian randomization analyses for glycemic traits in relation to Melanoma risk.

Glycemic trait	Remove the SNPs for selection criteria above	Remove the SNPs for palindromic and ambiguous structure	No. of SNPs for MR	Variance explained %	F statistic	MR egger Q statistic	MR egger *p*-value
Fasting insulin	26	0	26	0.0189	55	30.24	0.1768
2-h glucose	12	2	10	0.1728	58	6.65	0.5749
Fasting glucose	53	6	47	0.0106	117	57.99	0.0926
HbA1c	76	8	68	0.0096	108	88.38	0.0344
Type-2 diabetes	105	44	105	0.1854	89	112.72	0.2409

to show the genetic relationship between glycemic characteristics and melanoma risk, and funnel plots to visually analyze the consistency of MR estimations and potential related biases. R software was used for these analyses, where the "Two-Sample MR" and "MR-PRESSO" R packages were used.

2.4 Sensitivity analysis

Pleiotropy was investigated using the MR-Egger approach, which was used to determine if a single locus impacts numerous phenotypes. Second, the Leave-one-out sensitivity test was used to gradually remove the SNPs to ensure that the results were credible. The Cochran Q statistic was used to standardize heterogeneity analyses. In addition, MR PRESSO was used to detect and eliminate anomalous instrumental factors.

3 Results

3.1 MR assessment of glucose traits and melanoma risk

After a quality control process, we obtained 10 SNPs strongly associated with 2-h Glucose, 47 SNPs strongly associated with Fasting Glucose, 26 SNPs associated with Fasting Insulin, and 105 strongly associated with Type-2 Diabetes from GWAS and 68 SNPs closely related to Fasting Glucose. The F-statistics of these SNPs were all greater than 10, indicating that our instrumental variables were closely related to Glucose Traits. Furthermore, our instrumental variables were not directly associated with the risk of Melanoma (Table 1).

3.2 Mendelian randomization analysis of the association between glycemic traits and the risk of melanoma

IVW provides accurate estimates since the lack of heterogeneity and directional pleiotropy between exposure and outcome variables. Focusing primarily on the results of the IVW analysis, we assessed the causal relationship between these SPNs and melanoma risk with the Glycemic Traits (Table 2). The results showed that Fasting Glucose (OR = 0.99, 95%CI = 0.993–0.998, $p < 0.05$, IVW), Type-2 Diabetes (OR = 0.998, 95%CI = 0.998–0.999, $p < 0.01$, IVW) and HbA1c (OR= 0.19, 95%CI = 0.0415–0.8788, $p < 0.05$, IVW) was causally associated with a lower risk of Melanoma. Its orientation is consistent with several other MR analysis methods (Figures 1A–C). The 2-h Glucose and Fasting Insulin results showed no apparent causal relationship with Melanoma risk. Among them, no obvious causal difference was found in the IVW analysis of 2-h Glucose, but its $p < 0.05$ in MR Egger, and the direction was consistent in other MR analyses, suggesting that there may be a causal relationship (Figure 1D).

3.3 Sensitivity analysis

In MR Egger, the *p*-values of MR Egger intercepts in each instrumental variable of Glycemic Traits were greater than 0.05,

10.3389/fgene.2023.1260367

TABLE 2 Associated between the Glycemic Traits and risk of Melanoma using two-sample MR.

Glycemic traits	MR method	OR	95%CI	*p*-value
2-Hour Glucose				
	MR Egger	0.9911	0.9883–0.9984	<0.05
	Weighted Median	0.9980	0.9995–1.0005	0.11
	Inverse Variance Weighted	0.9989	0.9996–1.0001	0.31
	Simple Mode	0.9984	0.9994–1.0002	0.47
	Weighted Mode	0.9977	0.9984–1.0001	0.17
Fasting Glucose				
	MR Egger	0.9958	0.9901–1.0015	0.16
	Weighted Median	0.9978	0.9937–1.0018	0.30
	Inverse Variance Weighted	0.9968	0.9938–0.9998	<0.05
	Simple Mode	0.9942	0.9860–1.0025	0.13
	Weighted Mode	0.9971	0.9929–1.0012	0.16
Fasting Insulin				
	MR Egger	0.9917	0.9716–1.0121	0.43
	Weighted Median	1.0011	0.9941–1.0082	0.76
	Inverse Variance Weighted	0.9997	0.9943–1.0051	0.91
	Simple Mode	1.0032	0.9906–1.0159	0.62
	Weighted Mode	1.0025	0.9925–1.0127	0.67
Type-2 Diabetes				
	MR Egger	0.9986	0.9972–0.9999	<0.05
	Weighted Median	0.9985	0.9975–0.9995	<0.01
	Inverse Variance Weighted	0.9989	0.9983–0.9995	<0.01
	Simple Mode	0.9978	0.9959–0.9998	<0.05
	Weighted Mode	0.9984	0.9973–0.9996	<0.05
HbA1c				
	MR Egger	0.5199	0.029–9.3293	0.66
	Weighted Median	0.4510	0.0496–4.1015	0.46
	Inverse Variance Weighted	0.1910	0.0415–0.8788	<0.05
	Simple Mode	0.3793	0.004–35.9837	0.65
	Weighted Mode	0.8021	0.0863–7.4481	0.85

suggesting that the intercept does not exist, indicating fasting There was no horizontal pleiotropy for Fasting Glucose, 2-h Glucose, Fasting Insulin, HbA1c, and Type-2 Diabetes (Table 3). However, we found evidence of heterogeneity between HbA1c and melanoma risk with a *p*-value of 0.03442561 for the Q statistic. As we used the random-effects IVW as main result in MR of HbA1c, heterogeneity is acceptable (Burgess et al., 2019). No heterogeneity was found in the other analyses. Then we performed the Leave-one-out (Figures 2A–D) method and MR-PRESSO (Figures 3A–D) to identify and delete abnormal instrumental variables. The results showed that no abnormal instrumental variables were found, and the above results suggested that the MR analysis results were relatively stable.

4 Discussion

Melanoma is a malignant tumor caused by melanocytes that is also a very deadly disease due to its high metastatic potential,

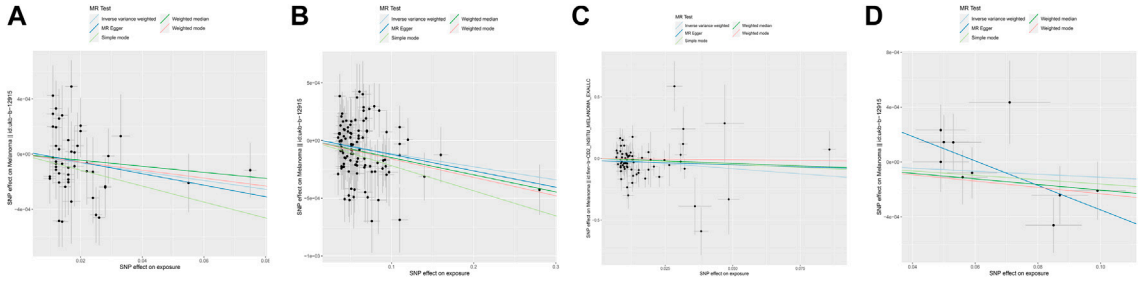


FIGURE 1
Scatter plot of genetic causality between Glycemic Traits and Melanoma using different MR methods. **(A)** Fasting Glucose **(B)** Type 2 Diabetes **(C)** HbA1c **(D)** 2-h Glucose. The dark blue line represents MR Egger, the light green line represents simple mode, the dark green line represents Weighted median, the light blue line represents IVW, and the red line represents Weighted mode.

TABLE 3 Estimates of Egger intercept to evaluate evidence for directional pleiotropy in MR association.

Glycemic traits	Egger intercept	SE of egger intercept	*p*-value
2-h Glucose	0.00054	0.00025	0.06
Fasting Glucose	0.00003	0.00006	0.69
Fasting Insulin	0.00013	0.00017	0.43
Type 2 Diabetes	0.00002	0.00005	0.64
HbA1c	−0.01877	0.02342	0.43

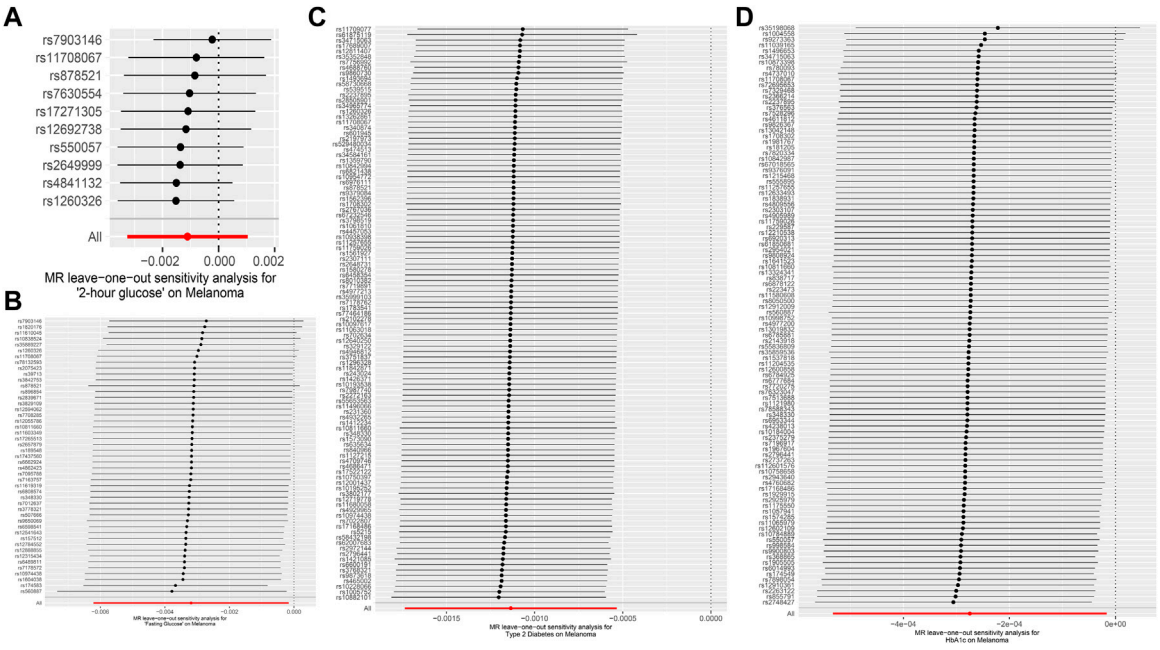


FIGURE 2
Forest map of Melanoma based on Glycemic Traits. **(A)** 2-h glucose **(B)** Fasting Glucose **(C)** Type 2 Diabetes **(D)** HbA1c. Black dots represent estimates of causal effects of Glycemic Traits on Melanoma (beta coefficients). The black line represents the estimated 95% confidence interval.

accounting for 75% of skin cancer deaths (Davis et al., 2019). People who have a family history of skin cancer, have a high amount of common or underdeveloped nevi, or are excessively exposed to UV light are at high risk for melanoma (Dummer et al., 2009; Guo et al., 2016). It is still unknown whether endocrine factors can also affect the occurrence of melanoma (Guo et al., 2016). This study used the

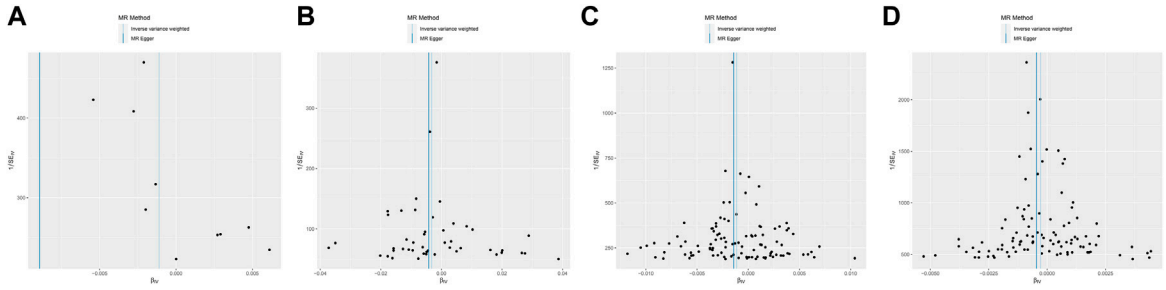


FIGURE 3
Funnel plot of Melanoma based on Glycemic Traits genetic variants. **(A)** 2-h glucose **(B)** Fasting Glucose **(C)** Type 2 Diabetes **(D)** HbA1c. Overall causal estimates (beta coefficients) of Glycemic Traits and Melanoma estimated by the IVW (light blue line) and MR-Egger (dark blue line) methods are shown.

MR method for the first time to explore the causal relationship between the characteristics of Fasting Glucose (including 2-h Glucose, Fasting Glucose, Fasting Insulin, Type-2 Diabetes, and HbA1c) and the risk of Melanoma from the perspective of genetics. Overall, we found that Fasting Glucose, Type-2 Diabetes, and higher HbA1c levels were negatively associated with the risk of Melanoma. The results of MR Egger suggest that 2-h Glucose may be negatively related to the risk of Melanoma. But there is no evidence that gene-predicted Fasting Insulin levels increase the risk of Melanoma. Due to the use of suitable genetic instrument tools (F-statistics>10 and r2<0.001) in this study, no significant SNP was detected in the retention method or MR-PRESSO, and the results were highly consistent among the 5 MR algorithms. Therefore, we believe that the results of this study are to some extent reliable. The results of this study suggest that people with higher Glycemic Traits levels may be at low risk for melanoma. Patients and high-risk populations may reduce the risk of melanoma by adjusted dietary structure regulating Glycemic Traits.

In previous studies, we have noticed that Glycemic Traits exhibit different causal relationships among different tumors. There is a positive correlation between plasma glucose index and disease risk in lung cancer, but there is no significant correlation with colon cancer (Murphy et al., 2022b; Du et al., 2022). A meta-analysis shows that glycemic index will increase the overall risk of cancer, increase the risk of breast cancer (Long et al., 2022), and is positively correlated with the risk of bladder cancer and gastric cancer (Zhu et al., 2020; Kim et al., 2022). Relevant studies have shown that higher HbA1c levels are associated with an increased risk of colorectal cancer, pancreatic cancer, respiratory cancer, and female reproductive tract cancer, and are not associated with an increased risk of breast cancer, gastrointestinal or urinary system malignancies (Hong et al., 2009; Lu et al., 2015; Hope et al., 2016; Murphy et al., 2022b), but are linearly associated with overall cancer-related deaths (Yoo et al., 2022). A study shows that Type-2 Diabetes will increase the risk of colorectal cancer (Murphy et al., 2022b). The cohort study and case-control study in Melanoma suggest that Type-2 Diabetes may be negatively related to the risk of Melanoma (Harding et al., 2015; Yuan et al., 2020), but the opposite results have appeared in other studies (Qi et al., 2014; Malavolti et al., 2017). The above studies suggest that the specific relationship between Glycemic Traits and Melanoma is still contradictory.

The results of this study suggest that 2-h Glucose (OR = 0.99, 95% CI = 0.984–0.998, $p < 0.05$, MR Egger), Fasting Glucose (OR = 0.99, 95% CI = 0.993–0.998, $p < 0.05$, IVW), Type-2 Diabetes (OR = 0.998, 95% CI = 0.998–0.999, $p < 0.01$, IVW) and higher HbA1c level (OR = 0.19, 95% CI = 0.0415–0.8788, $p < 0.05$, IVW) are all possible negatively related to the risk of Melanoma. This is different from the manifestation of Glycemic Traits in causal relationships with other tumors. This may be due to the reduction of melanocytes in Diabetes patients, whose melanin content is related to plasma glucose control in diabetes and obesity (Mackiewicz-Wysocka et al., 2014). Research shows that people with light skin are about 30 times more likely to suffer from Melanoma than people with dark skin (Doepner et al., 2022). This is because the pigmentation of human skin is determined by the transfer of mature melanin synthesized by epidermal Melanoma cells to the surrounding keratinocytes. Human Chromatophores synthesize two types of melanin, namely, eumelanin (EM) and phaeomelanin (PM). The content of eumelanin is directly related to skin pigmentation and has a photoprotective effect, which can protect the skin from ultraviolet rays, thereby reducing the incidence rate of Melanoma (Upadhyay et al., 2022). Moreover, a high level of plasma glucose is often associated with high BMI and obesity. Studies have shown that obesity is associated with elevated circulating estradiol levels due to the aromatase activity of adipose tissue converting androgens into estrogen compounds (Schneider et al., 1979). In primary Melanoma, there may be high expression of estrogen receptors β with anti-melanoma-proliferative, and sending non-classical estrogen signals through G protein-coupled receptors (de Giorgi et al., 2013; Marzagalli et al., 2015). This may explain the different causal relationships between Glycemic Traits in Melanoma and other tumors.

Insulin is a protein hormone secreted by cells stimulated by endogenous or exogenous substances such as glucose, lactose, ribose, arginine, glucagon, *etc.* Related research suggests that it may be related to the development of tumors. Lilalutide, an analog of glucagon-like peptide 1, is a molecule that regulates glucose by increasing insulin production and inhibiting glucagon secretion. It can significantly reduce the formation of NET in tumor mice by

improving the plasma glucose of patients, inhibiting tumor progression, and enhancing the anti-tumor effect of PD-1 inhibitors (Chen et al., 2022). Insulin is also an important cell growth factor that can promote cell growth, proliferation, and migration (Leitner et al., 1997). Higher fasting insulin is positively correlated with the risk of colorectal cancer (Murphy et al., 2022b). Extracellular vesicles secreted by breast cancer cells inhibit insulin secretion through miR-122, thereby damaging systemic glucose homeostasis and promoting tumor growth (Cao et al., 2022). However, in this study, the results suggest that there is no correlation between insulin level and the risk of Melanoma (OR = 0.99, 95% CI = 0.994–1.005, p = 0.91, IVW). Therefore, the causal relationship between Insulin and the risk of Melanoma requires more research.

There are still limitations to this study. Firstly, we restricted the relevance of the findings to other groups by concentrating on research subjects with European populations. Secondly, this study did not consider the impact of gender on MR analysis and did not conduct further subgroup analysis. Thirdly, to verify the findings, this study does not analyze additional data sources. Finally, This study used Mendelian Randomization analysis, which has the potential for weak instrument bias and pleiotropy. Future research is needed to address these limitations and to confirm the findings of this study.

Although previous observational studies can identify the relationship between Glucose, Type-2 Diabetes, Insulin and the risk of Melanoma, however, because of research confounding variables, a causal association cannot be established. In conclusion, this study used MR technology for the first time to analyze the causal relationship between Glycemic Traits and Melanoma and found that there is a negative correlation, and the underlying mechanism may provide valuable insights for carcinogenesis.

Data availability statement

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

Author contributions

Y-CZ: Conceptualization, Data curation, Project administration, Writing–original draft, Writing–review and editing. C-DL: Conceptualization, Formal Analysis, Investigation, Resources, Validation, Writing–review and editing. Q-YL: Data curation, Formal Analysis, Methodology, Project administration, Validation, Writing–original draft. J-NS: Conceptualization, Data curation, Investigation, Methodology, Writing–original draft. JS: Data curation, Funding acquisition, Investigation, Resources, Software, Supervision, Visualization, Writing–original draft, Writing–review and editing. MY: Conceptualization, Formal Analysis, Methodology, Project administration, Supervision, Writing–original draft, Writing–review and editing, Funding acquisition.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was funded by a Research Assistance Project of the Second Affiliated Hospital of Zhejiang Chinese Medicine University.

Acknowledgments

Thanks for the sharing of datasets provided by public databases.

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2023.1260367/full#supplementary-material

References

Atzpodien, J., Terfloth, K., Fluck, M., and Reitz, M. (2008). Cisplatin, gemcitabine and treosulfan is effective in chemotherapy-pretreated relapsed stage IV uveal melanoma patients. *Cancer Chemother. Pharmacol.* 62 (4), 685–688. doi:10.1007/s00280-007-0655-9

Bhatia, S., Moon, J., Margolin, K. A., Weber, J. S., Lao, C. D., Othus, M., et al. (2012). Phase II trial of sorafenib in combination with carboplatin and paclitaxel in patients with metastatic uveal melanoma: SWOG S0512. *PLoS One* 7 (11), e48787. doi:10.1371/journal.pone.0048787

Burgess, S., Davey Smith, G., Davies, N. M., Dudbridge, F., Gill, D., Glymour, M. M., et al. (2019). Guidelines for performing Mendelian randomization investigations: update for summer 2023. *Wellcome Open Res.* 4, 186. doi:10.12688/wellcomeopenres.15555.2

Cao, M., Isaac, R., Yan, W., Ruan, X., Jiang, L., Wan, Y., et al. (2022). Cancer-cell-secreted extracellular vesicles suppress insulin secretion through miR-122 to impair systemic glucose homeostasis and contribute to tumour growth. *Nat. Cell Biol.* 24 (6), 954–967. doi:10.1038/s41556-022-00919-7

Carr, S., Smith, C., and Wernberg, J. (2020). Epidemiology and risk factors of melanoma. *Surg. Clin. North Am.* 100 (1), 1–12. doi:10.1016/j.suc.2019.09.005

Chen, D., Liang, H., Huang, L., Zhou, H., and Wang, Z. (2022). Liraglutide enhances the effect of checkpoint blockade through the inhibition of neutrophil extracellular traps in murine lung and liver cancers. *FEBS Open Bio*. doi:10.1002/2211-5463.13499

Chen, J., Spracklen, C. N., Marenne, G., Varshney, A., Corbin, L. J., Luan, J., et al. (2021). The trans-ancestral genomic architecture of glycemic traits. *Nat. Genet.* 53 (6), 840–860. doi:10.1038/s41588-021-00852-9

Davis, L. E., Shalin, S. C., and Tackett, A. J. (2019). Current state of melanoma diagnosis and treatment. *Cancer Biol. Ther.* 20 (11), 1366–1379. doi:10.1080/15384047.2019.1640032

de Giorgi, V., Gori, A., Gandini, S., Papi, F., Grazzini, M., Rossari, S., et al. (2013). Oestrogen receptor beta and melanoma: a comparative study. *Br. J. Dermatol* 168 (3), 513–519. doi:10.1111/bjd.12056

Doepner, M., Lee, I., Natale, C. A., Brathwaite, R., Venkat, S., Kim, S. H., et al. (2022). Endogenous DOPA inhibits melanoma through suppression of CHRM1 signaling. *Sci. Adv.* 8 (35), eabn4007. doi:10.1126/sciadv.abn4007

Du, H., Zhang, T., Lu, X., Chen, M., Li, X., and Li, Z. (2022). Glycemic index, glycemic load, and lung cancer risk: a meta-analysis of cohort and case-control studies. *PloS One* 17 (9), e0273943. doi:10.1371/journal.pone.0273943

Dummer, R., Hauschild, A., and Pentheroudakis, G. (2009). Cutaneous malignant melanoma: ESMO clinical recommendations for diagnosis, treatment and follow-up. *Ann. Oncol.* 20 (Suppl. 4), 129–131. doi:10.1093/annonc/mdp152

Dusingize, J. C., Olsen, C. M., An, J., Pandeya, N., Law, M. H., Thompson, B. S., et al. (2020). Body mass index and height and risk of cutaneous melanoma: mendelian randomization analyses. *Int. J. Epidemiol.* 49 (4), 1236–1245. doi:10.1093/ije/dyaa009

Feng, R., Lu, M., Xu, J., Zhang, F., Yang, M., Luo, P., et al. (2022). Pulmonary embolism and 529 human blood metabolites: genetic correlation and two-sample Mendelian randomization study. *BMC Genom Data* 23 (1), 69. doi:10.1186/s12863-022-01082-6

Gershenwald, J. E., and Guy, G. P. (2016). Stemming the rising incidence of melanoma: calling prevention to action. *J. Natl. Cancer Inst.* 108 (1), djv381. doi:10.1093/jnci/djv381

Guo, J., Qin, S., Liang, J., Lin, T., Si, L., Chen, X., et al. (2016). Chinese guidelines on the diagnosis and treatment of melanoma (2015 edition). *Chin. Clin. Oncol.* 5 (4), 57. doi:10.21037/cco.2015.12.02

Harding, J. L., Shaw, J. E., Peeters, A., Cartensen, B., and Magliano, D. J. (2015). Cancer risk among people with type 1 and type 2 diabetes: disentangling true associations, detection bias, and reverse causation. *Diabetes Care* 38 (2), 264–270. doi:10.2337/dc14-1996

Hong, S. K., Lee, S. T., Kim, S. S., Min, K. E., Byun, S. S., Cho, S. Y., et al. (2009). Significance of preoperative HbA1c level in patients with diabetes mellitus and clinically localized prostate cancer. *Prostate* 69 (8), 820–826. doi:10.1002/pros.20932

Hope, C., Robertshaw, A., Cheung, K. L., Idris, I., and English, E. (2016). Relationship between HbA1c and cancer in people with or without diabetes: a systematic review. *Diabet. Med.* 33 (8), 1013–1025. doi:10.1111/dme.13031

Kamiza, A. B., Fatumo, S., Singini, M. G., Yeh, C. C., and Chikowore, T. (2022). Hepatitis B infection is causally associated with extrahepatic cancers: a Mendelian randomization study. *EBioMedicine* 79, 104003. doi:10.1016/j.ebiom.2022.104003

Kim, S. Y., Eun, C. S., Han, D. S., Kim, Y. S., Song, K. S., Choi, B. Y., et al. (2022). A high glycemic index and glycemic load increased the risk of gastric cancer: a case-control study in Korea. *Nutr. Res.* 105, 11–19. doi:10.1016/j.nutres.2022.06.001

Larsson, S. C., and Burgess, S. (2021). Causal role of high body mass index in multiple chronic diseases: a systematic review and meta-analysis of Mendelian randomization studies. *BMC Med.* 19 (1), 320. doi:10.1186/s12916-021-02188-x

Leitner, J. W., Kline, T., Carel, K., Goalstone, M., and Draznin, B. (1997). Hyperinsulinemia potentiates activation of p21Ras by growth factors. *Endocrinology* 138 (5), 2211–2214. doi:10.1210/endo.138.5.5240

Long, T., Liu, K., Long, J., Li, J., and Cheng, L. (2022). Dietary glycemic index, glycemic load and cancer risk: a meta-analysis of prospective cohort studies. *Eur. J. Nutr.* 61 (4), 2115–2127. doi:10.1007/s00394-022-02797-z

Lu, Y., García Rodríguez, L. A., Malgerud, L., González-Pérez, A., Martín-Pérez, M., Lagergren, J., et al. (2015). New-onset type 2 diabetes, elevated HbA1c, anti-diabetic medications, and risk of pancreatic cancer. *Br. J. Cancer* 113 (11), 1607–1614. doi:10.1038/bjc.2015.353

Mackiewicz-Wysocka, M., Araszkiewicz, A., Schlaffke, J., Kuczynski, S., Micek, I., and Zozulinska-Ziolkiewicz, D. (2014). Lower melanin content in the skin of type 1 diabetic patients and the risk of microangiopathy. *Exp. Clin. Endocrinol. Diabetes* 122 (4), 231–235. doi:10.1055/s-0034-1367003

Mahajan, A., Taliun, D., Thurner, M., Robertson, N. R., Torres, J. M., Rayner, N. W., et al. (2018). Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. *Nat. Genet.* 50 (11), 1505–1513. doi:10.1038/s41588-018-0241-6

Malavolti, M., Malagoli, C., Crespi, C. M., Brighenti, F., Agnoli, C., Sieri, S., et al. (2017). Glycaemic index, glycaemic load and risk of cutaneous melanoma in a population-based, case-control study. *Br. J. Nutr.* 117 (3), 432–438. doi:10.1017/S000711451700006X

Marzagalli, M., Casati, L., Moretti, R. M., Montagnani Marelli, M., and Limonta, P. (2015). Estrogen receptor β agonists differentially affect the growth of human melanoma cell lines. *PLoS One* 10 (7), e0134396. doi:10.1371/journal.pone.0134396

Murphy, N., Song, M., Papadimitriou, N., Carreras-Torres, R., Langenberg, C., Martin, R. M., et al. (2022a). Associations between glycemic traits and colorectal cancer: a mendelian randomization analysis. *J. Natl. Cancer Inst.* 114 (5), 740–752. doi:10.1093/jnci/djac011

Murphy, N., Song, M., Papadimitriou, N., Carreras-Torres, R., Langenberg, C., Martin, R. M., et al. (2022b). Associations between glycemic traits and colorectal cancer: a mendelian randomization analysis. *J. Natl. Cancer Inst.* 114 (5), 740–752. doi:10.1093/jnci/djac011

O'Neill, P. A., Butt, M., Eswar, C. V., Gillis, P., and Marshall, E. (2006). A prospective single arm phase II study of dacarbazine and treosulfan as first-line therapy in metastatic uveal melanoma. *Melanoma Res.* 16 (3), 245–248. doi:10.1097/01.cmr.0000205017.38859.07

Qi, L., Qi, X., Xiong, H., Liu, Q., Li, J., Zhang, Y., et al. (2014). Type 2 diabetes mellitus and risk of malignant melanoma: a systematic review and meta-analysis of cohort studies. *Iran. J. Public Health* 43 (7), 857–866.

Rastrelli, M., Tropea, S., Rossi, C. R., and Alaibac, M. (2014). Melanoma: epidemiology, risk factors, pathogenesis, diagnosis and classification. *Vivo* 28 (6), 1005–1011.

Schneider, G., Kirschner, M. A., Berkowitz, R., and Ertel, N. H. (1979). Increased estrogen production in obese men. *J. Clin. Endocrinol. Metab.* 48 (4), 633–638. doi:10.1210/jcem-48-4-633

Stattin, P., Björ, O., Ferrari, P., Lukanova, A., Lenner, P., Lindahl, B., et al. (2007). Prospective study of hyperglycemia and cancer risk. *Diabetes Care* 30 (3), 561–567. doi:10.2337/dc06-0922

Ugurel, S., and Gutzmer, R. (2023). Melanom. *J. Dtsch. Dermatol Ges.* 21 (4), 343–347. doi:10.1111/ddg.15053

Upadhyay, P. R., Starner, R. J., Swope, V. B., Wakamatsu, K., Ito, S., and Abdel-Malek, Z. A. (2022). Differential induction of reactive oxygen species and expression of antioxidant enzymes in human melanocytes correlate with melanin content: implications on the response to solar UV and melanoma susceptibility. *Antioxidants (Basel)* 11 (6), 1204. doi:10.3390/antiox11061204

Whiteman, D. C., Green, A. C., and Olsen, C. M. (2016). The growing burden of invasive melanoma: projections of incidence rates and numbers of new cases in six susceptible populations through 2031. *J. Invest. Dermatol* 136 (6), 1161–1171. doi:10.1016/j.jid.2016.01.035

Yoo, T. K., Lee, M. Y., Lee, S. A., Cheong, E. S., Seo, M. H., and Sung, K. C. (2022). Association of glycosylated hemoglobin level and cancer-related mortality in patients without diabetes. *J. Clin. Med.* 11 (19), 5933. doi:10.3390/jcm11195933

Yuan, S., Kar, S., Carter, P., Vithayathil, M., Mason, A. M., Burgess, S., et al. (2020). Is type 2 diabetes causally associated with cancer risk? Evidence from a two-sample mendelian randomization study. *Diabetes* 69 (7), 1588–1596. doi:10.2337/db20-0084

Zhu, H., Mo, Q., Shen, H., Wang, S., Liu, B., and Xu, X. (2020). Carbohydrates, glycemic index, and glycemic load in relation to bladder cancer risk. *Front. Oncol.* 10, 530382. doi:10.3389/fonc.2020.530382

frontiers | Frontiers in Oncology

TYPE Systematic Review
PUBLISHED 04 January 2024
DOI 10.3389/fonc.2023.1323812

OPEN ACCESS

EDITED BY
Xuanye Cao,
University of Texas MD Anderson Cancer Center, United States

REVIEWED BY
Wenchao Xu,
Huazhong University of Science and Technology, China
Hsiuying Wang,
National Yang Ming Chiao Tung University, Taiwan

*CORRESPONDENCE
Ping-yu Zhu
zhupingyu@nsmc.edu.cn
Xiao-ming Wang
wangnsmc@163.com

†These authors have contributed equally to this work

RECEIVED 01 November 2023
ACCEPTED 08 December 2023
PUBLISHED 04 January 2024

CITATION
Wang L, Deng J-y, Fan X-y, Yang D, Zhu P-y and Wang X-m (2024) Absence of causal relationship between Parkinson's disease and subsequent prostate cancer: evidence from meta-analysis and Mendelian randomization studies.
Front. Oncol. 13:1323812.
doi: 10.3389/fonc.2023.1323812

Absence of causal relationship between Parkinson's disease and subsequent prostate cancer: evidence from meta-analysis and Mendelian randomization studies

Li Wang[1†], Jing-ya Deng[2†], Xi-yue Fan[2], Dan Yang[2], Ping-yu Zhu[1*] and Xiao-ming Wang[2,3*]

[1]Department of Urology, Affiliated Hospital of North Sichuan Medical College, Nanchong, China, [2]Department of Neurology, Affiliated Hospital of North Sichuan Medical College, Nanchong, China, [3]Department of Neurology, Institute of Neurological Diseases of North Sichuan Medical College, Sichuan, Nanchong, China

Background: Numerous observational studies have investigated the risk of prostate cancer (PCa) in patients diagnosed with Parkinson's Disease (PD). However, the existence of a definitive association remains uncertain.

Methods: Systematic searches were performed on PubMed, Web of Science, Scopus, and Google Scholar for studies published up to October 1, 2023. For Mendelian randomized (MR) causal inference, we employed pooled data from the IPDGC and PRACTICAL Consortium. The inverse variance weighted (IVW) method served as the principal technique for estimating odds ratios (ORs) and 95% confidence intervals (CIs) for the associations under investigation.

Results: Cumulative analysis of nine studies revealed no significant association between patients diagnosed with PD and the subsequent incidence of PCa ([relative ratio] RR = 0.89, 95%CI = 0.73 to 1.08, P = 0.237). However, subgroup analyses indicated a reduced occurrence of PCa in Caucasian patients with PD (RR = 0.81, 95%CI = 0.69 to 0.95, P = 0.011). MR analyses failed to establish a significant link between increased genetic susceptibility to PD and the risk of PCa (IVW OR = 1.025, 95%CI = 0.997 to 1.054, P = 0.082). Sensitivity analyses further corroborated the robustness of these results.

Conclusion: Both observational meta-analysis and MR analysis based on genetic variation do not support an association between PD patients and the subsequent risk of PCa. Further research is warranted to unravel the potential underlying mechanisms linking these two diseases.

Systematic review registration: https://www.crd.york.ac.uk/PROSPERO/, identifier CRD42023473527.

KEYWORDS

Parkinson's disease, prostate cancer, Mendelian randomization, genetic variants, meta-analysis

1 Introduction

Parkinson's disease (PD) is the second most common neurodegenerative disorder and increases with age (1). In individuals with PD, there is a loss of dopaminergic neurons in the substantia nigra pars compacta, leading to resting tremors, rigidity, motor dysfunction, and postural instability (2). Numerous cellular pathways, including mitochondrial dysfunction, excitotoxicity, compromised autophagic processes, oxidative stress, the accumulation of misfolded proteins, and genetic mutations, have been postulated as interlinked contributors to the neurodegenerative processes observed in PD (3).

Epidemiological evidence reveals a noteworthy correlation between PD and cancer (4–6). One hallmark of tumors is unbridled cell proliferation and a deficiency in apoptosis, whereas individuals with PD exhibit an augmented inclination toward cellular apoptosis (7). Certain studies postulate shared genetic and biological pathways between PD and cancer. Conversely, males demonstrate greater susceptibility to PD, implying a hormonal regulatory influence on PD (8). On the other hand, prostate cancer (PCa), as the second most common malignancy worldwide, is regulated by sex hormones and ranks as the sixth leading cause of cancer-related deaths in males (9). Previous studies on the incidence of PCa in patients with PD have yielded contentious outcomes (4, 10, 11), and observational studies cannot infer a causal relationship between PD and prostate cancer, as this might be influenced by reverse causation or confounding factors.

Mendelian randomization (MR) emerges as a method of instrumental variable (IV) analysis that harnesses single nucleotide polymorphisms (SNPs) derived from genome-wide association studies (GWAS) as tools to deduce causal associations between two traits (12). MR approximates the inherent attributes of a RCT and exhibits a reduced susceptibility to the impact of covariates. Moreover, its operational simplicity and cost-effectiveness enhance its appeal (13). Consequently, we conducted an updated meta-analysis and integrated MR studies to investigate the causal relationship between PD and PCa.

2 Methods

2.1 Meta-analysis

This study adheres to the Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) guidelines (Supplementary Table 1) and has been registered with PROSPERO (CRD42023473527) (14).

2.2 Search strategy

We conducted a comprehensive search of the published literature for associations between PD and prostate cancer in MEDLINE via the Cochrane Library, PubMed, Web of Science, Scopus and Google Scholar databases, up to October 1, 2023. The following strings were constructed using a combination of medical subject terms and keywords: [(Parkinson OR Parkinson disease OR PD) AND (prostate cancer OR prostate carcinomas OR prostate neoplasm)].

2.3 Eligibility criteria

Inclusion criteria were defined as follows: (1) Population-based study of patients with diagnostic criteria for PD. (2) Cohort or case-control studies of PD diagnosis prior to PCa; (3) studies that reported either an odds ratio (OR), relative risk (RR), hazard ratio (HR), or standardized incidence ratio (SIR) along with the corresponding confidence interval (CI); (3) original research published in English. The exclusion criteria comprised: (1) studies lacking relevant exposures (PD) and outcomes (Pca); (2) studies without meta-analysis data; (3) reviews, letters, case reports or conference reports. If study populations overlap, select the newest or most informative published studies.

2.4 Data acquisition and quality evaluation

Two investigators (JY, WL) employed EndNote X9 to identify and remove duplicate records. They subsequently reviewed both the titles and full texts of the remaining records for further screening. Relevant data were extracted and recorded in an Excel spreadsheet, including the following information: first author, year of publication, geographical region, duration of follow-up, method of PD diagnosis, number of cases and controls, adjusted covariates, risk values for outcome estimates. Two reviewers (XY and YD) evaluated the risk of bias using the Cochrane Collaboration Risk of Bias in Non-Randomized Studies of Interventions (ROBINS-I) tool (15). Moreover, we assessed study quality using the Newcastle-Ottawa Scale for cohort and case-control studies, with scores

ranging from 0 to 9 (16). The included studies were categorized into two groups based on their mean quality score: a low-quality group (<7) and a high-quality group (≥7). In addition, the level of evidence (LOE) was graded according to the criteria of the Oxford Centre for Evidence-Based Medicine (17). In cases of disagreements, these were resolved through negotiation.

2.5 Statistical analysis

Given the low absolute incidence of prostate cancer, the four types of measurements were estimated to have similar RR values. In conjunction with previously published meta-analyses, we present the results using RR (18, 19). Due to the unavoidable high degree of heterogeneity between publications ($P < 0.05$, $I^2 > 50\%$), pooled effect sizes were calculated using random effects models. Otherwise, a fixed-effects model was used ($P > 0.5$, $I^2 < 50\%$). Egger's test and funnel plots were utilized to evaluate publication bias. Sensitivity analyses assess the reliability of results by removing each study in turn. Furthermore, we performed subgroup analyses considering time to cancer diagnosis, study type, study quality, population, and year of publication. Meta-analyses were conducted using Stata 16.0 and considered statistically significant at $p < 0.05$.

2.6 Mendelian randomization

The study rigorously adhered to the guidelines outlined in the Strengthening the Reporting of Observational Studies in Epidemiology Mendelian Randomization (STROBE-MR) framework (20). MR relies on three essential assumptions: IVs demonstrate strong correlation with PD, remain unaffected by confounding variables, and impact Pca solely through the exposure under investigation. The basic assumptions and MR design flow are depicted in Figure 1. Since publicly available pooled data were utilized, ethical approval was not necessary for this study.

2.7 Data source and SNP selection

Summary data for PD were obtained from the comprehensive GWAS meta-analysis conducted by the International Parkinson's Disease Genomics Consortium (IPDGC), encompassing 33,674 cases and 449,056 controls of European descent (21). GWAS data for Pca from Prostate Cancer Association Group to Investigate Cancer Associated Alterations in the Genome (PRACTICAL) Consortium (79,148 cases and 61,106 control cases) (22). To ensure the stability of the causal relationship between exposure and outcome, IVs were selected based on the following principles: (1) We established genome-wide significance thresholds for PD at $p < 5\times10^{-8}$. (2) Cluster analysis was conducted to address linkage disequilibrium (LD) among the selected IVs ($r^2 < 0.001$, kb = 10,000). (3) Only SNPs with a minor allele frequency (MAF) exceeding 0.01 were considered. (4) To mitigate bias from weak IVs. the strength of the IVs was quantified using the F value (β^2/SE), with those having $F < 10$ being excluded (23). Here, β represents the effect size of exposure and SE represents the standard error of the effect size. we also used Phenoscanner to examine potential confounders (such as body mass index, smoking, alcohol consumption and vitamin D supplementation) (24) (Table 1).

2.8 Statistical analysis

The primary analysis employed the robust inverse-variance weighted (IVW) method (25). This method has the strongest statistical efficacy, but it must be satisfied that all genetic variation is a valid instrumental variable, and therefore we employed the weighted median, MR-Egger regression, maximum likelihood and simple weighted mode methods as validation approaches (26, 27). Sensitivity analysis assumes a vital role in the assessment of heterogeneity and potential biases within MR studies. Firstly, heterogeneity was evaluated through the application of Cochran's Q test, which involved calculating the weighted sum of squared differences between specific variability estimates and the overall IVW estimate (28). To address potential outliers, the MR Pleiotropy RESidual Sum and Outlier (MR-PRESSO) method was employed during data analysis (29). Furthermore, MR-Egger regression was utilized, and intercepts were assessed to identify potential horizontal pleiotropy ($p < 0.05$ was judged significant). in addition, we performed a leave-one-out analysis to test the stability of the results. We evaluated heterogeneity among variant-specific causal estimates and pinpointed outliers through scatter and funnel plots. Finally, we identified potential

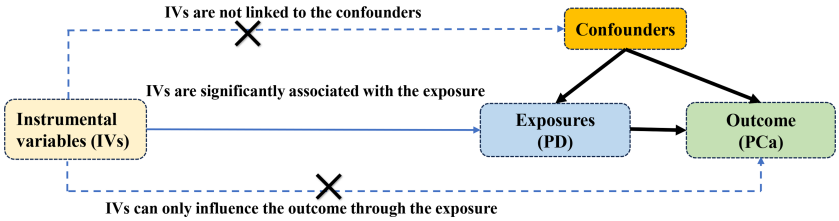


FIGURE 1
The three main assumptions of Mendelian randomization.

TABLE 1 Characteristics of the included GWAS summary studies in Mendelian randomization.

Trait	First author	Consortium	Sex/population	Sample size	Number of (cases/controls)	Year	GWAS ID
Exposure							
Parkinson's disease	Nalls MA	IPDGC	Male and female/European	482,730	33,674/449,056	2019	ieu-b-7
Outcome							
Prostate cancer	Schumacher	PRACTICAL	Male/European	140,254	79,148/61,106	2018	ebi-a-GCST006085
Confounders							
Obesity	Berndt SI	GIANT	Male and female/European	98,697	32,858/65,839	2013	ieu-a-90
Smoking status: Current	Neale	Neale Lab	Male and female/European	336,024	33,928/302,096	2017	ukb-a-225
Ever smoked	Ben Elsworth	MRC-IEU	Male and female/European	461,066	280,508/180,558	2018	ukb-b-20261
Former alcohol drinker	Ben Elsworth	MRC-IEU	Male and female/European	31,506	16,191/15,315	2018	ukb-b-12654
Triglycerides	Willer CJ	GLGC	Males and females/Mixed	177,861	NA	2013	ieu-a-302
Vitamin D supplements	Ben Elsworth	MRC-IEU	Male and female/European	460,351	17,879/442,472	2018	ukb-b-12648

GWASs, genome-wide association studies; IPGDC, International Parkinson's Disease Genomics Consortium; PRACTICAL, Prostate Cancer Association Group to Investigate Cancer Associated Alterations in the Genome; GIANT, genetic investigation of anthropometric traits consortium; MRC-IEU, MRC Integrative Epidemiology Unit; GLGC, Global Lipids Genetics Consortium; NA, not available.

bidirectional links between SNPs related to the PD and PCa using the MR Steiger Filtering Test (30). In addition, we performed multivariate MR (MVMR) analyses to observe the effect of confounding factors on PCa.

Statistical analyses were executed using R version 4.2.2 with the "TwoSampleMR" and "MRPRESSO" packages. Odds ratios (ORs) with 95% confidence intervals (CIs) were used to quantify the MR analysis, and statistical significance was defined as $P < 0.05$.

3 Result

3.1 Meta–analysis results

3.1.1 Study characteristics and quality evaluation

After a rigorous examination of online databases, 9 articles (5, 10, 31–37)(8 cohort and 1 case-control) from 2007 to 2019 were included in the final analysis. Figure 2 illustrates the selection process, and Table 2 provides detailed information on the included literature. 6 studies received high-quality ratings. however, all studies were at low to moderate risk of bias (Supplementary Table 2).

3.1.2 PCa risk in PD

Pooled analyses overall showed no significant association between patients with PD and the subsequent risk of PCa (RR = 0.89; 95% CI: 0.73 to 1.08; p = 0.237) (Figure 3A). This result held true across different types of studies (Figure 3B). Interestingly, within the Caucasian population, patients with PD were found to have a lower risk of PCa (RR = 0.81; 95% CI: 0.69 to 0.95; p = 0.011) (Figure 3C).

3.1.3 Sensitivity analysis

Summarized effects remain stable through the successive exclusion of each study (Supplementary Figure 1). Furthermore, evidence of significant bias was not found in funnel plots or through Egger's (p = 0.963) and Begg's test (p = 0.297) (Supplementary Figure 2).

3.2 Mendelian randomization results

The *a priori* calculation of statistical power was conducted meticulously (38). By setting α at 5%, we attained a substantial statistical power exceeding 80% in scenarios where the expected OR concerning PCa were either at or below 1.04 within the context of PD.

3.2.1 Effect of PD on PCa

The associations between the 21 designated SNPs and PCa are meticulously delineated in Supplementary Table 3. The range of

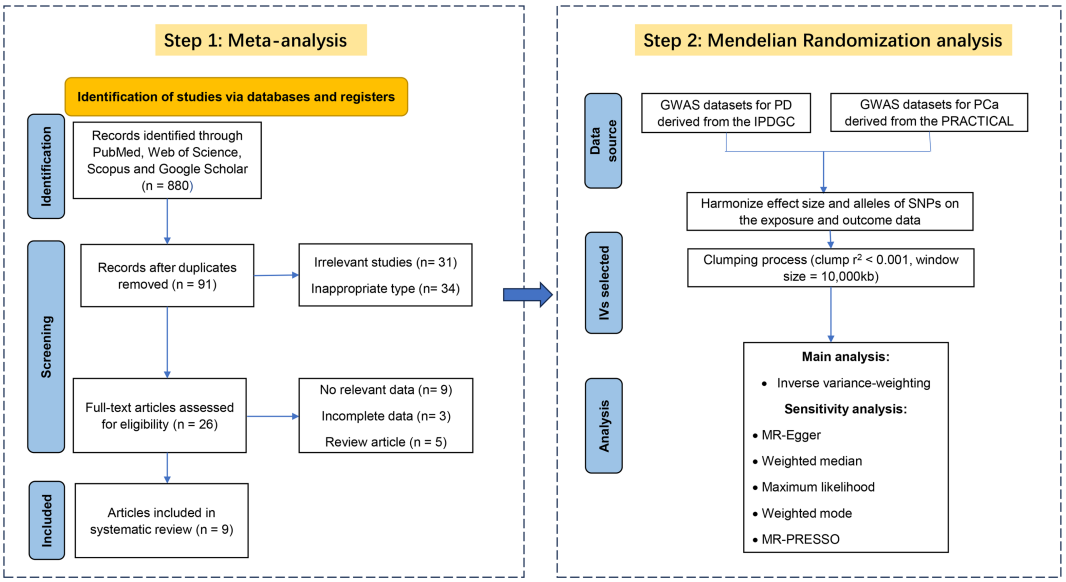


FIGURE 2
Flow chart for meta-analysis and Mendelian randomization analysis.

variance expounded upon by these SNPs in relation to the exposure variables extended from 0.004 to 0.02. Furthermore, the IVs demonstrated robust statistical significance ($F > 10$). After a rigorous Steiger filtering process, no signs of reverse causality were found. There is no apparent association between genetic predisposition to PD and the occurrence of PCa (OR = 1.025; 95% CI: 0.997 to 1.054; P = 0.082), which is consistent with the overall effect results of the meta-analysis. No heterogeneity was observed in the sensitivity analysis, and there was no horizontal pleiotropy detected in the MR-Egger analysis. Additionally, the MR-PRESSO test did not identify any outliers (Global test p = 0.315) (Figure 4; Supplementary Figure 3). Results between genetic susceptibility to PD and PCa remained robust in MVMR adjusted for relevant confounders (Table 3).

4 Discussion

This study has undertaken a comprehensive assessment of the risk of PCa in patients diagnosed with PD. The results of cumulative analysis and MR analysis have confirmed the lack of significant correlation between PD and PCa under genetic prediction. The co-occurrence of two distinct diseases within the same individual may stem from shared environmental or genetic factors. Previous studies have yielded conflicting evidence regarding the relationship between PD and cancer (5, 37), and several potential explanatory mechanisms have been proposed.

PD, a neurodegenerative disorder, is characterized by the demise of dopaminergic neurons, distinguishing it from PCa, which is typified by unrestricted cellular proliferation and a lack of apoptosis. Interestingly, cells in PD patients exhibit a greater propensity to undergo apoptosis, which may serve as a defensive mechanism against cancer progression.

Smoking is recognized as a significant risk factor for various types of tumors while seemingly reducing the risk of developing PD (39). Nicotine has been observed to stimulate the release of dopamine and demonstrate neuronal protection in various experimental models (40). Although PCa is not typically associated with smoking, earlier investigations have reported a decreased risk of PCa among individuals with PD (4, 41). It's worth noting that patients diagnosed with PD typically have higher mortality rates than the general population. Furthermore, those who do survive are less likely to die from subsequent cancers (42).

One of the therapeutic strategies for individuals with PD involves increasing dopamine levels within the central nervous system, thereby stimulating the sympathetic nerves. Concurrently, anticholinergic drugs might act on parasympathetic nerves to alleviate symptoms (43). The stroma of the prostate is heavily innervated by branches of the autonomic nervous system, which play a significant role in the growth and sustenance of the prostate gland (44). A study by Magnon et al. (45) discovered that sympathetic neurons foster tumor genesis at an early stage, while parasympathetic fibers drive the dissemination of cancer.

TABLE 2 Baseline characteristics of the included studies.

Author (year)	Design	Country	Mean or median follow-up (years)	Disease ascertainment	Sample size	Adjustment for covariates	Outcomes	NOS	LOE
Fois (31) (2010)	Cohort	UK	3.2	Coded	4,355 cases	Age, Sex, Time period in single calendar years and district of residence	RR	7	2b
Lo (32) (2010)	Cohort	UK	4.3	Medical record and clinical	692 cases; 761 controls	Age, sex, cigarette smoking, alcohol consumption, BMI	OR	7	2b
Wirdefeldt (33) (2014)	Cohort	Sweden	NA	Coded	11,786 cases; 58,930 controls	Age, sex, urbanization	HR	6	2b
Becker (34) (2010)	Case-control	UK	NA	Medical records	466 cases; 1864 controls	Age, sex, calendar time, BMI, smoking status	OR	9	2b
Driver (35) (2007)	Cohort	USA	5.2	Self- report PD diagnosis	572 cases; 478 controls	Smoking history, alcohol use, physical activity, BMI	RR	8	2b
Rugbjerg (36) (2012)	Cohort	Denmark	5.7	Coded	20343 cases	NA	SIR	6	2b
Ong (10) (2014)	Cohort	UK	12	Coded	219,194 cases; 9,015,614 controls	Age, sex, calendar year of first recorded admission, region of residence, quintile of patients' Index of Deprivation score	RR	8	2b
Lin (5) (2015)	Cohort	Taiwan	7	Coded	62,023 cases; 124,046 controls	Age, sex	HR	6	2b
Park (37) (2019)	Cohort	Korea	6	Coded	52,009 cases; 260,045 controls	Age, sex, hypertension, diabetes mellitus, hyperlipidemia, income	HR	8	2b

BMI, body mass index; HR, hazard ratio; OR, odds ratio; PD, Parkinson disease; RR, relative risk; SIR, standardized incidence ratio; NA, not applicable.

Consequently, medications targeting branches of the autonomic nervous system could potentially offer therapeutic advantages.

Levodopa and other dopaminergic drugs may be administered following a diagnosis of PD. Current studies indicate that L-Dopa decarboxylase (DDC) is an androgen receptor co-activator, its expression increases with the progression of the disease, and its co-expresses with receptors in prostate cancer cells. The related drugs enhance anti-tumor activity by inhibiting the DDC pathway (46). Interestingly, our findings indicate that Caucasian populations exhibit a lower prevalence of PCa following the onset of PD. Lin et al. (5)discovered that Taiwanese men diagnosed with PD had an elevated risk of PCa, a phenomenon attributed to a confluence of distinctive genetic backgrounds, habits, and/or environmental exposures. However, in MR analyses conducted on European populations, no significant causal association was observed between PD and the risk of subsequent PCa occurrence. This appears to suggest that the results of meta-analyses may have been influenced by bias and confounding factors.

4.1 Strength and limitation

Our study possesses several strengths. Firstly, we adhered strictly to PRISMA guidelines in our literature screening and conducted subgroup analyses and bias assessments. Secondly, our MR study adhered to the three key hypotheses and utilized a two-sample approach to explore the causal relationship between PD and PCa. Sensitivity analyses confirmed the reliability of our results, while MVMR analyses helped to eliminate confounding bias. Despite these strengths, our study is not without limitations. For one, the MR analysis validated results solely for the European population, which might have resulted in a more homogeneous association. Furthermore, we did not perform a gender-stratified analysis, which may have introduced some bias. Moreover, the results of the meta-analysis were inevitably highly heterogeneous. Finally, the insufficient sample size may lead to instability in subgroup effects, and future studies with larger sample sizes are needed to enhance the reliability of the results.

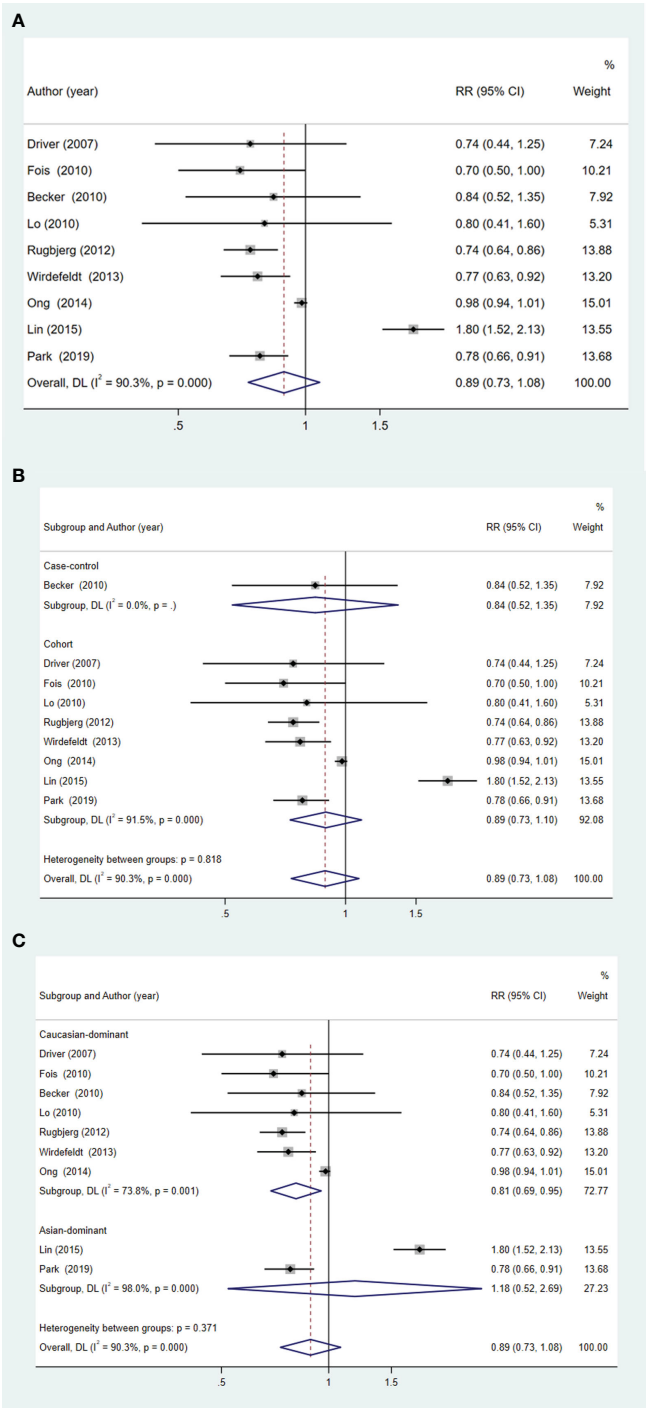


FIGURE 3
Forest plot of PCa risk in patients with PD and subgroup analysis. **(A)** overall effect; **(B)** subgroup analysis of study type; **(C)** subgroup analysis of different ethnicities.

Effect	SNPs	Methods	OR(95% CI)	Pvalue	Heterogeneity (P)	MR Egger (intercept, P)	MR−PRESSO global test p
PD on PCa	21	Inverse variance weighted	1.025 (0.997 to 1.054)	0.082	0.335	0.009, 0.121	0.315
		MR Egger	0.975 (0.913 to 1.042)	0.465	0.431		
		Weighted median	1.001 (0.961 to 1.043)	0.949			
		Simple median	1.016 (0.973 to 1.061)	0.467			
		Maximum likelihood	1.026 (0.999 to 1.054)	0.063			
		MR-PRESSO(Raw)	1.025 (0.997 to 1.054)	0.097			

FIGURE 4
MR analysis results from PD to PCa risk.

TABLE 3 Complete MVMR results of PD in relevant prostate cancer risk factors.

Adjustments	Methods	Parkinson's Disease					
		SNPs	Causal effect		Heterogeneity	Pleiotropy	
			OR (95% CI)	*P*	*P**	Intercept	*P*
Obesity	IVW	22	1.05 (0.99-1.13)	0.11	0.066	0.005	0.199
	Median		1.01 (0.93-1.09)	0.83			
	Egger		1.05 (0.99-1.12)	0.12			
Smoking status: Current	IVW	35	1.00 (0.96-1.04)	0.87	< 0.001	0.003	0.522
	Median		1.01 (0.97-1.05)	0.75			
	Egger		0.98 (0.91-1.06)	0.65			
Ever smoked	IVW	90	1.01 (0.98-1.04)	0.60	< 0.001	0.002	0.290
	Median		1.02 (0.98-1.07)	0.27			
	Egger		0.99 (0.95-1.04)	0.81			
Former alcohol drinker	IVW	19	1.01 (0.96-1.06)	0.69	< 0.001	0.012	0.358
	Median		1.01 (0.96-1.05)	0.81			
	Egger		0.93 (0.78-1.11)	0.45			
Triglycerides	IVW	53	0.99 (0.90-1.10)	0.92	< 0.001	0.003	0.365
	Median		0.99 (0.92-1.08)	0.95			
	Egger		0.99 (0.89-1.10)	0.79			
Vitamin D supplements	IVW	19	1.01 (0.96-1.06)	0.70	< 0.001	0.010	0.445
	Median		1.04 (0.99-1.08)	0.10			
	Egger		0.95 (0.80-1.13)	0.54			

*Heterogeneity P < 0.05 indicated potential heterogeneity existing in the IVW model, and the median method was suggested for causal inference in this situation. MVMR, multivariable Mendelian randomization; SNP, single nucleotide polymorphisms; IVW, inverse variance weighted; OR, odds ratio; CI, confidence interval; PD Parkinson's Disease.

5 Conclusion

This comprehensive MR and meta-analysis did not demonstrate an association between PD and PCa risk. The potential biological pathways contributing to the co-morbidity between these two diseases certainly warrant further exploration.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding authors.

Author contributions

LW: Writing – original draft. JD: Data curation, Writing – original draft. XF: Methodology, Writing – original draft. DY: Formal Analysis, Writing – original draft. PZ: Methodology, Supervision, Writing – review & editing. XW: Funding acquisition, Supervision, Validation, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was funded by the Technology Department Project of Sichuan Science (2023NSFSC0622).

Acknowledgments

We are grateful to the consortium that provided all the public GWAS data. We would like to express our sincere gratitude to the Technology Department Project of Sichuan Science (2023NSFSC0622) for providing financial support for this study.

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/fonc.2023.1323812/full#supplementary-material

References

1. Tysnes O-B, Storstein A. Epidemiology of Parkinson's disease. *J Neural Transm (Vienna)* (2017) 124(8):901–5. doi: 10.1007/s00702-017-1686-y

2. Mazzoni P, Shabbott B, Cortés JC. Motor control abnormalities in Parkinson's disease. *Cold Spring Harbor Perspect Med* (2012) 2(6):a009282. doi: 10.1101/cshperspect.a009282

3. Zeng XS, Geng WS, Jia JJ, Chen L, Zhang PP. Cellular and molecular basis of neurodegeneration in Parkinson disease. *Front Aging Neurosci* (2018) 10:109. doi: 10.3389/fnagi.2018.00109

4. Jespersen CG, Nørgaard M, Borre M. Parkinson's disease and risk of prostate cancer: A Danish population-based case-control study, 1995-2010. *Cancer Epidemiol* (2016) 45:157–61. doi: 10.1016/j.canep.2016.11.002

5. Lin PY, Chang SN, Hsiao TH, Huang BT, Lin CH, Yang PC. Association between Parkinson disease and risk of cancer in Taiwan. *JAMA Oncol* (2015) 1(5):633–40. doi: 10.1001/jamaoncol.2015.1752

6. Ording AG, Veres K, Horváth-Puhó E, Glymour MM, Rørth M, Henderson VW, et al. Alzheimer's and Parkinson's diseases and the risk of cancer: A cohort study. *J Alzheimer's Dis JAD* (2019) 72(4):1269–77. doi: 10.3233/JAD-190867

7. Garber K. Parkinson's disease and cancer: the unexplored connection. *J Natl Cancer Institute* (2010) 102(6):371–4. doi: 10.1093/jnci/djq081

8. Bourque M, Dluzen DE, Di Paolo T. Neuroprotective actions of sex steroids in Parkinson's disease. *Front Neuroendocrinol* (2009) 30(2):142–57. doi: 10.1016/j.yfrne.2009.04.014

9. Rebello RJ, Oing C, Knudsen KE, Loeb S, Johnson DC, Reiter RE, et al. Prostate cancer. *Nat Rev Dis Primers* (2021) 7(1):9. doi: doi:10.1038/s41572-020-00243-0

10. Ong EL, Goldacre R, Goldacre M. Differential risks of cancer types in people with Parkinson's disease: a national record-linkage study. *Eur J Cancer (Oxford Engl 1990)* (2014) 50(14):2456–62. doi: 10.1016/j.ejca.2014.06.018

11. Peretz C, Gurel R, Rozani V, Gurevich T, El-Ad B, Tsamir J, et al. Cancer incidence among Parkinson's disease patients in a 10-yrs time-window around disease onset: A large-scale cohort study. *Parkinsonism Related Disord* (2016) 28:68–72. doi: 10.1016/j.parkreldis.2016.04.028

12. Sanderson E, Glymour MM, Holmes MV, Kang H, Morrison J, Munafò MR, et al. Mendelian Randomization. *Nature Reviews Methods Primers* (2022) 2(1):6. doi: 10.1038/s43586-021-00092-5

13. Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum Mol Genet* (2014) 23(R1):R89–98. doi: 10.1093/hmg/ddu328

14. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ (Clinical Res ed)* (2021) 372:n71. doi: 10.1136/bmj.n71

15. Sterne JA, Hernán MA, Reeves BC, Savović J, Berkman ND, Viswanathan M, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ (Clinical Res ed)* (2016) 355:i4919. doi: 10.1136/bmj.i4919

16. Wells GA, Shea B, O'Connell D, Peterson J, Welch V, Losos M, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. (2000).

17. Howick J. The Oxford 2011 levels of evidence. (2011).

18. Lin L, Chu H, Murad MH, Hong C, Qu Z, Cole SR, et al. Empirical comparison of publication bias tests in meta-analysis. *J Gen Internal Med* (2018) 33(8):1260–7. doi: 10.1007/s11606-018-4425-7

19. Fang T, Zhang Z, Zhou H, Wu W, Zou L. Multiple sclerosis and breast cancer risk: a meta-analysis of observational and Mendelian randomization studies. *Front Neuroinform* (2023) 17:1154916. doi: 10.3389/fninf.2023.1154916

20. Skrivankova VW, Richmond RC, Woolf BAR, Yarmolinsky J, Davies NM, Swanson SA, et al. Strengthening the reporting of observational studies in epidemiology using mendelian randomization: the STROBE-MR statement. *Jama* (2021) 326(16):1614–21. doi: 10.1001/jama.2021.18236

21. Nalls MA, Blauwendraat C, Vallerga CL, Heilbron K, Bandres-Ciga S, Chang D, et al. Identification of novel risk loci, causal insights, and heritable risk for Parkinson's disease: a meta-analysis of genome-wide association studies. *Lancet Neurol* (2019) 18(12):1091–102. doi: 10.1016/S1474-4422(19)30320-5

22. Schumacher FR, Al Olama AA, Berndt SI, Benlloch S, Ahmed M, Saunders EJ, et al. Association analyses of more than 140,000 men identify 63 new prostate cancer susceptibility loci. *Nat Genet* (2018) 50(7):928–36. doi: 10.1038/s41588-018-0142-8

23. Burgess S, Small DS, Thompson SG. A review of instrumental variable estimators for Mendelian randomization. *Stat Methods Med Res* (2017) 26(5):2333–55. doi: 10.1177/0962280215597579

24. Pernar CH, Ebot EM, Wilson KM, Mucci LA. The epidemiology of prostate cancer. *Cold Spring Harbor perspectives in medicine* (2018) 8,12 a030361. doi: 10.1101/cshperspect.a030361

25. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. *Genet Epidemiol* (2013) 37(7):658–65. doi: 10.1002/gepi.21758

26. 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(4):304–14. doi: 10.1002/gepi.21965

27. Hartwig FP, Davey Smith G, Bowden J. Robust inference in summary data Mendelian randomization *via* the zero modal pleiotropy assumption. *Int J Epidemiol* (2017) 46(6):1985–98. doi: 10.1093/ije/dyx102

28. Burgess S, Bowden J, Fall T, Ingelsson E, Thompson SG. Sensitivity analyses for robust causal inference from Mendelian randomization analyses with multiple genetic variants. *Epidemiol (Cambridge Mass)* (2017) 28(1):30–42. doi: 10.1097/EDE.0000000000000559

29. Verbanck M, Chen CY, 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(5):693–8. doi: 10.1038/s41588-018-0099-7

30. Hemani G, Tilling K, Davey Smith G. Orienting the causal relationship between imprecisely measured traits using GWAS summary data. *PloS Genet* (2017) 13(11):e1007081. doi: 10.1371/journal.pgen.1007081

31. Fois AF, Wotton CJ, Yeates D, Turner MR, Goldacre MJ. Cancer in patients with motor neuron disease, multiple sclerosis and Parkinson's disease: record linkage studies. *J Neurol Neurosurg Psychiatry* (2010) 81(2):215–21. doi: 10.1136/jnnp.2009.175463

32. Lo RY, Tanner CM, Van Den Eeden SK, Albers KB, Leimpeter AD, Nelson LM. Comorbid cancer in Parkinson's disease. *Movement Disord Off J Movement Disord Society* (2010) 25(12):1809–17. doi: 10.1002/mds.23246

33. Wirdefeldt K, Weibull CE, Chen H, Kamel F, Lundholm C, Fang F, et al. Parkinson's disease and cancer: A register-based family study. *Am J Epidemiol* (2014) 179(1):85–94. doi: 10.1093/aje/kwt232

34. Becker C, Brobert GP, Johansson S, Jick SS, Meier CR. Cancer risk in association with Parkinson disease: a population-based study. *Parkinsonism Related Disord* (2010) 16(3):186–90. doi: 10.1016/j.parkreldis.2009.11.005

35. Driver JA, Logroscino G, Buring JE, Gaziano JM, Kurth T. A prospective cohort study of cancer incidence following the diagnosis of Parkinson's disease. *Cancer Epidemiol Biomarkers Prev* (2007) 16(6):1260–5. doi: 10.1158/1055-9965.EPI-07-0038

36. Rugbjerg K, Friis S, Lassen CF, Ritz B, Olsen JH. Malignant melanoma, breast cancer and other cancers in patients with Parkinson's disease. *Int J Cancer* (2012) 131(8):1904–11. doi: 10.1002/ijc.27443

37. Park JH, Kim DH, Park YG, Kwon DY, Choi M, Jung JH, et al. Cancer risk in patients with Parkinson's disease in South Korea: A nationwide, population-based cohort study. *Eur J Cancer (Oxford Engl 1990)* (2019) 117:5–13. doi: 10.1016/j.ejca.2019.04.033

38. Brion MJ, Shakhbazov K, Visscher PM. Calculating statistical power in Mendelian randomization studies. *Int J Epidemiol* (2013) 42(5):1497–501. doi: 10.1093/ije/dyt179

39. Li X, Li W, Liu G, Shen X, Tang Y. Association between cigarette smoking and Parkinson's disease: A meta-analysis. *Arch Gerontol Geriatr* (2015) 61(3):510–6. doi: 10.1016/j.archger.2015.08.004

40. Quik M. Smoking, nicotine and Parkinson's disease. *Trends Neurosci* (2004) 27 (9):561–8. doi: 10.1016/j.tins.2004.06.008

41. Freedman DM, Wu J, Chen H, Engels EA, Enewold LR, Freedman ND, et al. Associations between cancer and Parkinson's disease in U.S. elderly adults. *Int J Epidemiol* (2016) 45(3):741–51. doi: 10.1093/ije/dyw016

42. Vanacore N, Spila-Alegiani S, Raschetti R, Meco G. Mortality cancer risk in parkinsonian patients: a population-based study. *Neurology* (1999) 52(2):395–8. doi: 10.1212/WNL.52.2.395

43. Kobylecki C. Update on the diagnosis and management of Parkinson's disease. *Clin Med (Lond)* (2020) 20(4):393–8. doi: 10.7861/clinmed.2020-0220

44. McVary KT, Razzaq A, Lee C, Venegas MF, Rademaker A, McKenna KE. Growth of the rat prostate gland is facilitated by the autonomic nervous system. *Biol Reprod* (1994) 51(1):99–107. doi: 10.1095/biolreprod51.1.99

45. Magnon C, Hall SJ, Lin J, Xue X, Gerber L, Freedland SJ, et al. Autonomic nerve development contributes to prostate cancer progression. *Sci (New York NY)* (2013) 341 (6142):1236361. doi: 10.1126/science.1236361

46. Wafa LA, Cheng H, Plaa N, Ghaidi F, Fukumoto T, Fazli L, et al. Carbidopa abrogates L-dopa decarboxylase coactivation of the androgen receptor and delays prostate tumor progression. *Int J Cancer* (2012) 130(12):2835–44. doi: 10.1002/ijc.26287

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 04 January 2024
DOI 10.3389/fgene.2023.1308263

OPEN ACCESS

EDITED BY
Ziheng Wang,
University of Macau, China

REVIEWED BY
Ruizheng Sun,
Central South University, China
Huihui Chen,
Central South University, China
Ying Liao,
Sun Yat-sen University Cancer Center (SYSUCC), China
Kaijian Hou,
Shantou University, China

*CORRESPONDENCE
Daqing Sun,
sdqchris2019@tmu.edu.cn
Chao Sun,
sunchao0330@hotmail.com

†These authors have contributed equally to this work

RECEIVED 06 October 2023
ACCEPTED 01 December 2023
PUBLISHED 04 January 2024

CITATION
Wang S, Yin F, Guo Z, Li R, Sun W, Wang Y, Geng Y, Sun C and Sun D (2024), Association between gut microbiota and glioblastoma: a Mendelian randomization study.
Front. Genet. 14:1308263.
doi: 10.3389/fgene.2023.1308263

Association between gut microbiota and glioblastoma: a Mendelian randomization study

Song Wang[1†], Fangxu Yin[1†], Zheng Guo[1†], Rui Li[1], Wei Sun[1], Yuchao Wang[1], Yichen Geng[2], Chao Sun[3*] and Daqing Sun[1*]

[1]Department of Pediatric Surgery, Tianjin Medical University General Hospital, Tianjin, China, [2]Nursing College of Binzhou Medical University, Yantai, Shandong, China, [3]Department of Orthopedic Surgery, Tianjin Medical University General Hospital, Tianjin, China

Background: Glioblastoma (GBM) is the most prevalent malignant brain tumor, significantly impacting the physical and mental wellbeing of patients. Several studies have demonstrated a close association between gut microbiota and the development of GBM. In this investigation, Mendelian randomization (MR) was employed to rigorously evaluate the potential causal relationship between gut microbiota and GBM.

Methods: We utilized summary statistics derived from genome-wide association studies (GWAS) encompassing 211 gut microbiota and GBM. The causal association between gut microbiota and GBM was scrutinized using Inverse Variance Weighted (IVW), MR-Egger, and Weighted Median (WM) methods. Cochrane's Q statistic was employed to conduct a heterogeneity test. MR-Pleiotropic Residuals and Outliers (MR-PRESSO) were applied to identify and eliminate SNPs with horizontal pleiotropic outliers. Additionally, Reverse MR was employed to assess the causal relationship between GBM and pertinent gut microbiota.

Results: The MR study estimates suggest that the nine gut microbiota remain stable, considering heterogeneity and sensitivity methods. Among these, the *family.Peptostreptococcaceae* and *genus.Eubacterium brachy group* were associated with an increased risk of GBM, whereas *family.Ruminococcaceae*, *genus.Anaerostipes*, *genus.Faecalibacterium*, *genus.LachnospiraceaeUCG004*, *genus.Phascolarctobacterium*, *genus.Prevotella7*, and *genus.Streptococcus* were associated with a reduced risk of GBM. Following Benjamini and Hochberg (BH) correction, *family.Ruminococcaceae* (OR = 0.04, 95% CI: 0.01–0.19, FDR = 0.003) was identified as playing a protective role against GBM.

Conclusion: This groundbreaking study is the first to demonstrate that *family.Ruminococcaceae* is significantly associated with a reduced risk of GBM. The modulation of *family_Ruminococcaceae* for the treatment of GBM holds considerable potential clinical significance.

KEYWORDS

glioblastoma, gut microbiota, gut-brain axis, Mendelian randomization, causality

1 Introduction

Glioblastoma (GBM) is one of the most prevalent types of malignant brain tumors, with an annual incidence ranging from 3 to 6.4 per 100,000 individuals. It constitutes approximately 23.3% of central nervous system tumors and 78.3% of malignant brain tumors. The 5-year mortality rate ranks second only to that of pancreatic cancer and lung cancer (Sung et al., 2021; Ostrom et al., 2023). Typically arising from glial cells or precursor cells, its clinical manifestations encompass increased intracranial pressure, neurological and cognitive impairment, as well as seizures (Omuro and DeAngelis, 2013). According to the World Health Organization (WHO) classification, gliomas are categorized into four grades, with a direct correlation between higher grade and poorer prognosis. Notably, GBM stands out as the most malignant subtype. Characterized by a suppressive immune microenvironment and a grim prognosis, GBM stands as one of the most challenging tumors, prone to recurrence and imposing a substantial societal burden (Chen et al., 2021).

There is mounting evidence that the immunosuppressive environment of GBM is not only mediated by the immunosuppressive cells and molecules described above but also has many connections to the gut microbiota that contribute to the development of GBM (5). The human gut microbiota contains microbes with diverse properties and functions. Imbalance in the gut microbiota refers to the inability of bacteria in the human environment to maintain a dynamic balance, resulting in an imbalance of gut microbiota. Bacteria in the human environment are unable to maintain homeostasis, leading to inflammation and immunosuppression, and the gut microbiota is particularly responsive to the presence of tumors (Ferreiro et al., 2018; Sepich-Poore et al., 2021). In recent years, the role of the gut microbiota in tumors has been extensively studied. In neurodegenerative diseases and tumors of the central nervous system (CNS), the gut microbiota establishes interactions between the gut and the CNS in complex and as yet unclear ways (Fung et al., 2017).

Given the ethical issues and costs associated with clinical trials, determining causation becomes challenging (Bothwell and Podolsky, 2016). Many studies investigating the relationship between the gut microbiota and tumors have primarily employed case-control designs, introducing difficulty in establishing the temporal sequence between changes in the composition of the gut microbiota and the onset of tumors (de Clercq et al., 2021; Bellerba et al., 2022; Reichard et al., 2022). In light of these challenges, Mendelian randomization (MR) emerges as a robust approach, utilizing single nucleotide polymorphisms (SNPs) as instrumental variables (IV) derived from genome-wide association studies (GWAS) to ascertain causality between exposure and outcome (Sekula et al., 2016). Consequently, our present study employs Mendelian randomization methods to analyze the causal association between gut microbiota and glioblastoma multiforme (GBM), providing insights for potential clinical interventions for GBM.

2 Materials and methods

2.1 Study population

As illustrated in Figure 1, our study outlines the two-sample MR investigation employed to explore the causal association between the gut microbiota and GBM. Subsequently, rigorous quality controls, including heterogeneity and gene pleiotropy tests, were executed to validate the dependability of the causal findings. In enhancing the precision of causal effect estimation, adherence to three crucial assumptions is imperative when utilizing SNPs as IVs in MR analysis (Sung et al., 2021): IVs must be closely aligned with the exposure factor; (Ostrom et al., 2023) IVs should exhibit no correlation with confounding factors; (Omuro and DeAngelis, 2013); IVs must exclusively influence outcomes through exposure, avoiding other pathways (Figure 2).

The main exposure factor in our study is the gut microbiota, and we investigate human genetics within the context of studying the gut microbiota. This investigation is conducted as part of an international consortium known as MiBioGen (Kurilshikov et al., 2021). Our study encompasses data from the human gut microbiota of 18,340 European individuals derived from 24 population-based cohorts. After adjustment for age, sex, technical covariates, and genetic principal components, spearman's correlation analysis was performed to identify genetic loci that affected the covariate-adjusted abundance of bacterial taxa. Following the exclusion of 15 genera lacking specific species names, we identified 196 bacterial taxa, comprising 9 phyla, 16 orders, 20 orders, 32 families, and 119 genera.

The outcome variable we focus on is GBM, and the GWAS dataset associated with GBM came from a publicly available GWAS meta-analysis that included 91 cases and 218,701 controls of European ancestry (Sudlow et al., 2015). The GWAS meta-analysis, a prospective cohort study, systematically gathers comprehensive genetic and phenotypic data from approximately 500,000 individuals across the UK. Each participant contributes a wealth of phenotypic and health-related information. Genome-wide genotype data were collected for all participants by linking health and medical records to provide comprehensive follow-up information.

2.2 Selection of instrumental variables

Single nucleotide polymorphisms (SNPs) are the most frequently utilized genetic variations in MR Analysis, which mainly refers to the DNA sequence diversity caused by a change in a single nucleotide at the genomic level. In this study, SNPs significantly associated with the relative abundance of 196 gut microbiota were selected as the available instrumental variables (IVs). Previous studies have shown that the inclusion of multiple instrumental variables enhances the explanatory power of the observed variation and enhances the accuracy and reliability of the analyzed results. Therefore, in this study, the selection of IVs was based on the results of correlation analysis where significance was determined at $P < 1 \times 10^{-5}$. The criteria for linkage disequilibrium were set at $R^2 < 0.001$ and a genetic distance of 10,000 kb, whereby

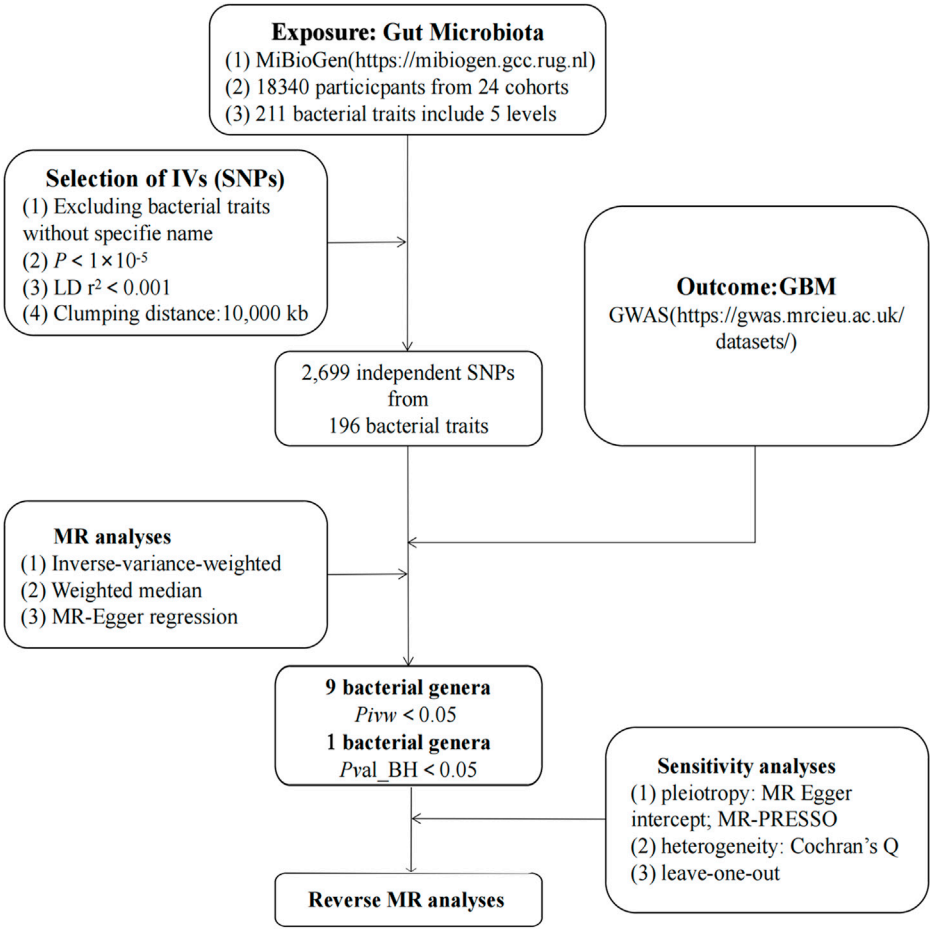


FIGURE 1
The study design of the present MR study of the associations of gut microbiota and GBM. Abbreviations: GBM, glioblastoma; LD, linkage disequilibrium, which used to measure the correlations between SNPs; IVW, Inverse Variance Weighted, the main analyses to evaluate the relationship between exposure and outcome; MR-PRESSO, Mendelian Randomization Pleiotropy RESidual Sum and Outlier, a method test the pleiotropic biases in the SNPs and correct the pleiotropic effects; MR, Mendelian randomization; SNP, single nucleotide polymorphism, as instrumental variables for the exposures and outcomes.

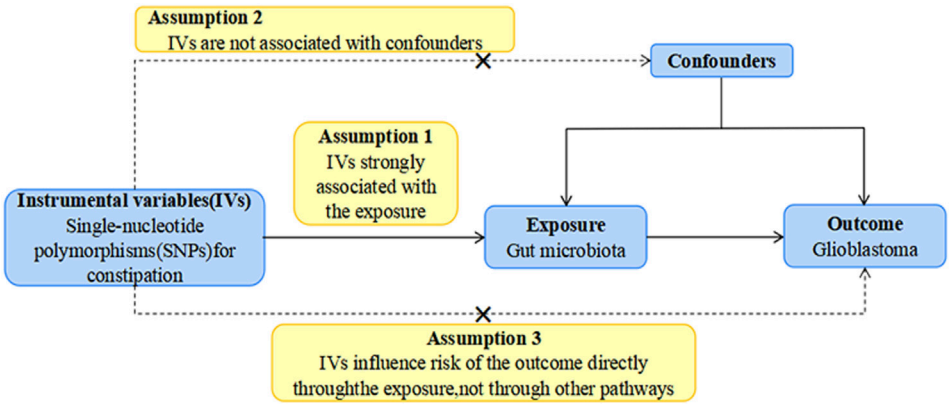


FIGURE 2
The study design of the present MR study of the associations of gut microbiota and GBM. Abbreviations: GBM, glioblastoma; LD, linkage disequilibrium, which used to measure the correlations between SNPs; IVW, Inverse Variance Weighted, the main analyses to evaluate the relationship between exposure and outcome; MR-PRESSO, Mendelian Randomization Pleiotropy RESidual Sum and Outlier, a method test the pleiotropic biases in the SNPs and correct the pleiotropic effects; MR, Mendelian randomization; SNP, single nucleotide polymorphism, as instrumental variables for the exposures and outcomes.

highly correlated SNPs were excluded to ensure the independence of the included SNPs from each other. Finally, SNPs associated with the relative abundance of gut microbiota were projected to the GWAS pooled data of GBM, and the corresponding statistical parameters were extracted. Utilizing statistical parameters associated with identical loci in the relative abundance of gut microbiota and the GWAS results for GBM, the data were harmonized. This harmonization ensured that the effect values for both exposure and outcome corresponded to the same effect allele.

2.3 Statistical analysis

In this study, Inverse Variance Weighted (IVW), MR-Egger, Weighted Median (WME) were used to estimate the dependent effects. The IVW method operates under the assumption that all genetic variants are valid IVs. It employs the ratio method to calculate the causal effect values for individual instrumental variables, subsequently summarizing each estimate through a weighted linear regression to derive the total effect value. Notably, the main divergence between the MR-Egger method and the IVW method lies in the regression, which takes into account the presence of an intercept term. Conversely, the WME method strategically leverages the intermediate effects of all available genetic variants, obtaining estimates by weighting the inverse variance of the correlation of each SNP with the outcome.

Since the IVW method exhibits higher test efficacy compared to other MR methods, we chose it as the preferred method for estimating causal effects in this study. Additionally, for enhanced result interpretation, the study transformed Beta (β) values obtained from the results into Odds Ratios (OR), while simultaneously calculating the 95% confidence intervals (CI). To assess the association of effect estimates for causality, which might be influenced by weak instrumental bias, the strength of IV was evaluated using the F statistic. This statistic was calculated using the following equation: $F = R^2 (n-k-1)/k (1-R^2)$, where R^2 represents the variance explained by IV (for each gut microbiota), and n is the sample size. The value of R^2 was estimated using the minor allele frequency (MAF), and b values were determined by the equation: $R2 = 2 \times MAF \times (1-MAF) \times b^2$.

In addition, for the purpose of further testing the stability and reliability of the results, quality control included sensitivity analysis and heterogeneity testing, as well as a gene multiplicity test. Sensitivity analysis was performed using the leave-one-out method, where the combined effect values of the remaining SNPs were calculated by sequentially deleting individual SNPs, and the effect of each SNP on the results was assessed. Heterogeneity testing was conducted using the Cochran Q test to determine the heterogeneity of the SNPs, aiming to assess the possible bias in the estimation of the causal effect due to the measurement error of SNPs caused by different analysis platforms, experimental conditions, and analyzing populations. Horizontal gene pleiotropy tests were employed to assess whether IVs affected outcomes through pathways other than exposure, utilizing intercept terms from MR-Egger regression. Finally, reverse MR was performed to analyze whether there was a reverse causal relationship between GBM and meaningful gut microbiota. MR analyses and quality control for this study were conducted using version 4.0.3 of R and additionally version 0.5.6 of the TwoSampleMR software package.

3 Results

3.1 Two-sample Mendelian randomization

The results of this study involving gut microbiota associated with GBM are presented in Supplementary Table S1. After a series of quality control steps, 136 independent SNPs from 9 gut microbiota were associated with GBM. The F-statistics for the gut microbiota ranged from 14.58 to 88.42, and all met the threshold of greater than 10, suggesting that they are unlikely to be affected by weak instrumental bias (Supplementary Table S2). Briefly, we identified nine gut microbiota associated with GBM. After undergoing BH correction, the *family.Ruminococcaceae* was found to play a protective role against GBM (Table 1). Details of the IVs used are listed in Supplementary Table S3.

3.2 Causal effects of gut microbiota on GBM

Nine gut microbiota were screened for correlation with GBM according to the IVW (Figure 3). Among them, *family.Peptostreptococcaceae* (OR: 3.83, 95% CI: 1.02–14.35, p = 0.046) and *genus.Eubacterium brachy group* (OR: 2.85, 95% CI: 1.16–7.01, p = 0.023) were found to increase the risk of GBM, while *family.Ruminococcaceae* (OR: 0.04, 95% CI: 0.01–0.19, p = 9.51E-05), *genus.Anaerostipes* (OR: 0.16, 95% CI: 0.03–0.83, p = 0.029), *genus.Faecalibacterium* (OR: 0.16, 95% CI: 0.04–0.65, p = 0.011), *genus.Lachnospiraceae UCG004* (OR: 0.20, 95% CI: 0.04–0.96, p = 0.045), *genus.Phascolarctobacterium* (OR: 0.16, 95% CI: 0.03–0.76, p = 0.021), *genus.Prevotella7* (OR: 0.30, 95% CI: 0.13–0.68, p = 0.004), and *genus.Streptococcus* (OR: 0.21, 95% CI:0.05–0.97, p = 0.046) showed a negative correlation with GBM. However, only *family.Ruminococcaceae* was found to be negatively associated with the risk of GBM after strict BH correction (P_{FDR} = 0.003).

The WME method has suggested that *family.Peptostreptococcaceae* (OR: 6.42, 95% CI: 1.09–37.71, p = 0.040) and *genus.Eubacterium brachy group* (OR: 4.70, 95% CI: 1.46–15.14, p = 0.009) are associated with an increased risk of GBM, while *family.Ruminococcaceae* (OR: 0.08, 95% CI: 0.01–0.79, p = 0.031) and *genus.Prevotella7* (OR: 0.28, 95% CI: 0.10–0.84, p = 0.023) show a negative correlation with GBM. However, there was no observed association between *genus.Anaerostipes*, *genus.Faecalibacterium*, *genus.Lachnospiraceae UCG004*, *genus.Phascolarctobacterium*, *genus.Streptococcus* and GBM (Figures 3, 4).

Additionally, the MR-Egger regression intercept did not show evidence of pleiotropy of the gut microbiota with GBM (All intercept $p > 0.05$) (Table 2; Supplementary Table S3). MRPRESSO regression did not identify outliers (All intercept $p > 0.05$).The results of heterogeneity analysis confirmed the accuracy of the findings (Table 2; Supplementary Table S4). Meanwhile, the data's robustness was further confirmed by the leave-one-out results, which demonstrated a consistent negative association between

TABLE 1 Effect estimation of the association between meaningful gut microbiota and risk of GBM in MR analysis. Abbreviations: GBM, glioblastoma; MR, Mendelian randomization analysis; SNPs, Number of single nucleotide polymorphism. CI, confidence interval; OR, odds ratio; P_{FDR}, *p*-value was calculated by the Benjamini-Hochberg method.

Gut microbiota	Outcome	SNPs	Methods	OR (95% CI)	*p*-value	P_{FDR}
family.Peptostreptococcaceae	**GBM**					
		13	MR-Egger	2.34 (0.10–52.17)	0.603	
		13	Weighted median	6.42 (1.09–37.71)	0.040	
		13	IVW	3.83 (1.02–14.35)	0.046	0.566
family.Ruminococcaceae	**GBM**					
		9	MR-Egger	0.02 (3.89E-4-0.72)	0.070	
		9	Weighted median	0.08 (0.01–0.79)	0.031	
		9	IVW	0.04 (0.01–0.19)	9.51E-5	0.003
genus.Anaerostipes	**GBM**					
		13	MR-Egger	2.94 (0.01–1058.94)	0.727	
		13	Weighted median	0.34 (0.04–2.78)	0.312	
		13	IVW	0.16 (0.03–0.83)	0.029	0.680
genus.Eubacterium brachy group	**GBM**					
		10	MR-Egger	0.96 (0.03–36.27)	0.984	
		10	Weighted median	4.70 (1.46–15.14)	0.009	
		10	IVW	2.85 (1.16–7.01)	0.023	0.680
genus.Faecalibacterium	**GBM**					
		18	MR-Egger	0.31 (0.02–4.96)	0.434	
		18	Weighted median	0.18 (0.02–1.52)	0.115	
		18	IVW	0.16 (0.04–0.65)	0.011	0.632
genus.Lachnospiraceae UCG004	**GBM**					
		24	MR-Egger	0.09 (1.42E-4-61.80)	0.491	
		24	Weighted median	0.32 (0.04–2.53)	0.281	
		24	IVW	0.20 (0.04–0.96)	0.045	0.758
genus.Phascolarctobacterium	**GBM**					
		12	MR-Egger	0.03 (1.20E-5-75.19)	0.414	
		12	Weighted median	0.65 (0.08–5.60)	0.694	
		12	IVW	0.19 (0.04–0.93)	0.041	0.680
genus.Prevotella7	**GBM**					
		27	MR-Egger	0.82 (0.01–92.06)	0.936	
		27	Weighted median	0.28 (0.10–0.84)	0.023	
		27	IVW	0.30 (0.13–0.68)	0.004	0.458
genus.Streptococcus	**GBM**					
		10	MR-Egger	0.02 (6.59E-5- 3.57)	0.159	
		10	Weighted median	0.27 (0.03–2.16)	0.215	
		10	IVW	0.21 (0.05–0.97)	0.046	0.758

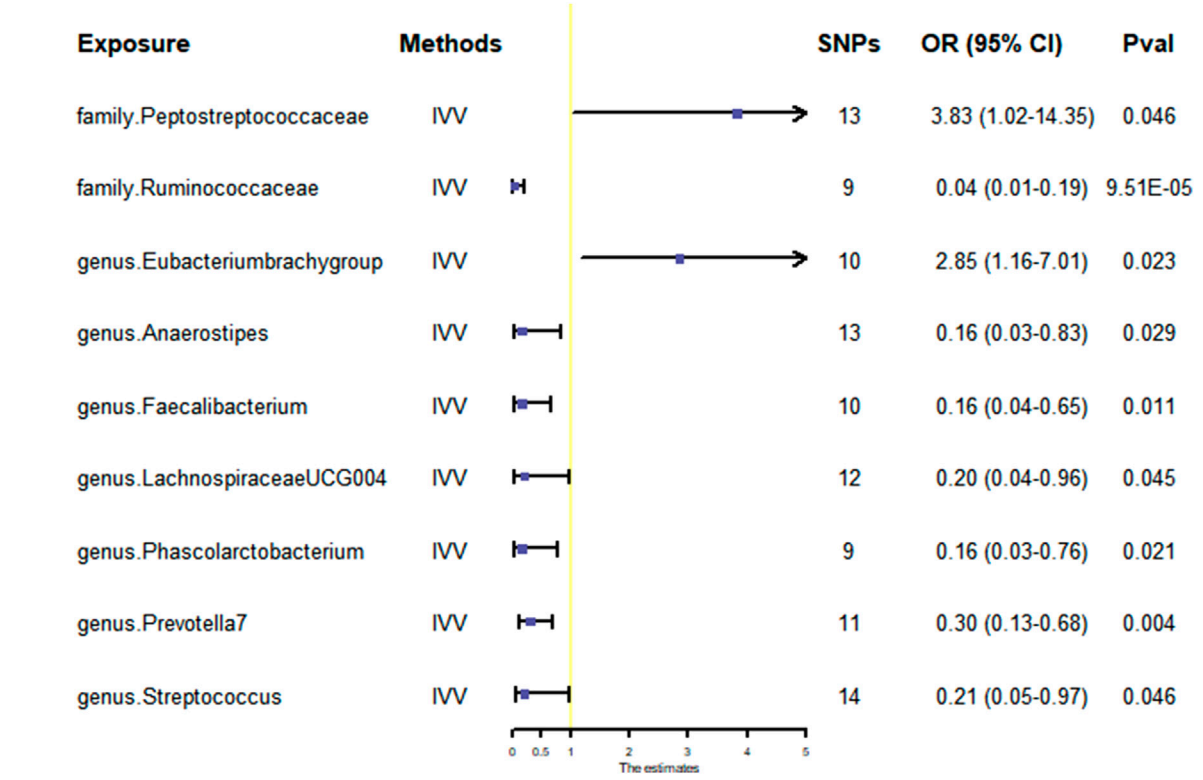


FIGURE 3
Scatter plots for the causal association between gut microbiota and GBM. Abbreviations: GBM, glioblastoma; OR, odds ratio; CI, confidence interval.

family_Ruminococcaceae and GBM risk (Figure 5; Supplementary Table S5).

3.3 Inverse MR analysis

In the reverse MR, GBM was selected as an exposure factor. However, the results of the MR study did not support a causal relationship between GBM and altered gut microbiota (IVW, OR = 1.012, 95% CI: 0.807–1.268, p = 0.921) (Supplementary Table S6).

4 Discussion

Our study is the first to identify the existence of a direct causal association between gut microbiota and GBM, indicating that an elevated abundance of gut microbiota, such as the *family.Ruminococcaceae*, is associated with a reduced risk of developing GBM. *Ruminococcus* was one of the first gastrointestinal bacteria to be discovered and plays a crucial role in metabolism (Mizrahi et al., 2021). A study on the inflammatory properties of the *family.Ruminococcaceae* found that it produces metabolites in the form of glucomannan polysaccharides, and that these polysaccharides can prime immune system cells (Teng et al., 2022). During the development of GBM, when the BBB is disrupted in the body and circulating immune cells are suppressed in a immunosuppressive environment, gut microbiota such as C. tumefaciens can further enhance the stimulation of immune system cell production. Thus, this bacterium may be a potential protective factor in the development of GBM.

Genus Faecalibacterium has been reported as one of the major butyrate producers found in the intestine (Lopez-Siles et al., 2017). *In vitro* studies have demonstrated that butyrate exhibits antitumor effects, such as inhibiting tumor growth by reducing tumor necrosis factor (TNF) secretion in intestinal epithelial cells and inducing differentiation and apoptosis of tumor cells. Butyrate, as a short-chain fatty acid, serves as a histone deacetylase (HDAC) inhibitor, thereby impeding the activity and life cycle of cancer cells (Modoux et al., 2022). Moreover, butyrate, as a short-chain fatty acid and HDAC inhibitor, enhances CPT1A activity to promote induced regulatory T-cell (iTreg) differentiation. iTreg plays a pivotal role in immunosuppression and maintaining immune homeostasis in brain tissue (He et al., 2022). *Genus.Anaerostipes* also belongs to butyrate-producing bacteria and exhibits anti-inflammatory and immunomodulatory functions (Zhang et al., 2016). Within the *genus_LachnospiraceaeUCG004* can reduce tumorigenesis by modulating the function of tumor immunosurveillance (Carasso et al., 2021). However, further studies are needed to explore its potential in terms of GBM risk protection.Therefore, we suggest that these gut microbiota may play a role in GBM development by modulating immunity.

A growing body of evidence underscores the pivotal role of the gut microbiota in tumor therapy, highlighting its key involvement in both local gut immunity and systemic immunity (Park et al., 2022). A robust microbiota employs direct and indirect mechanisms to

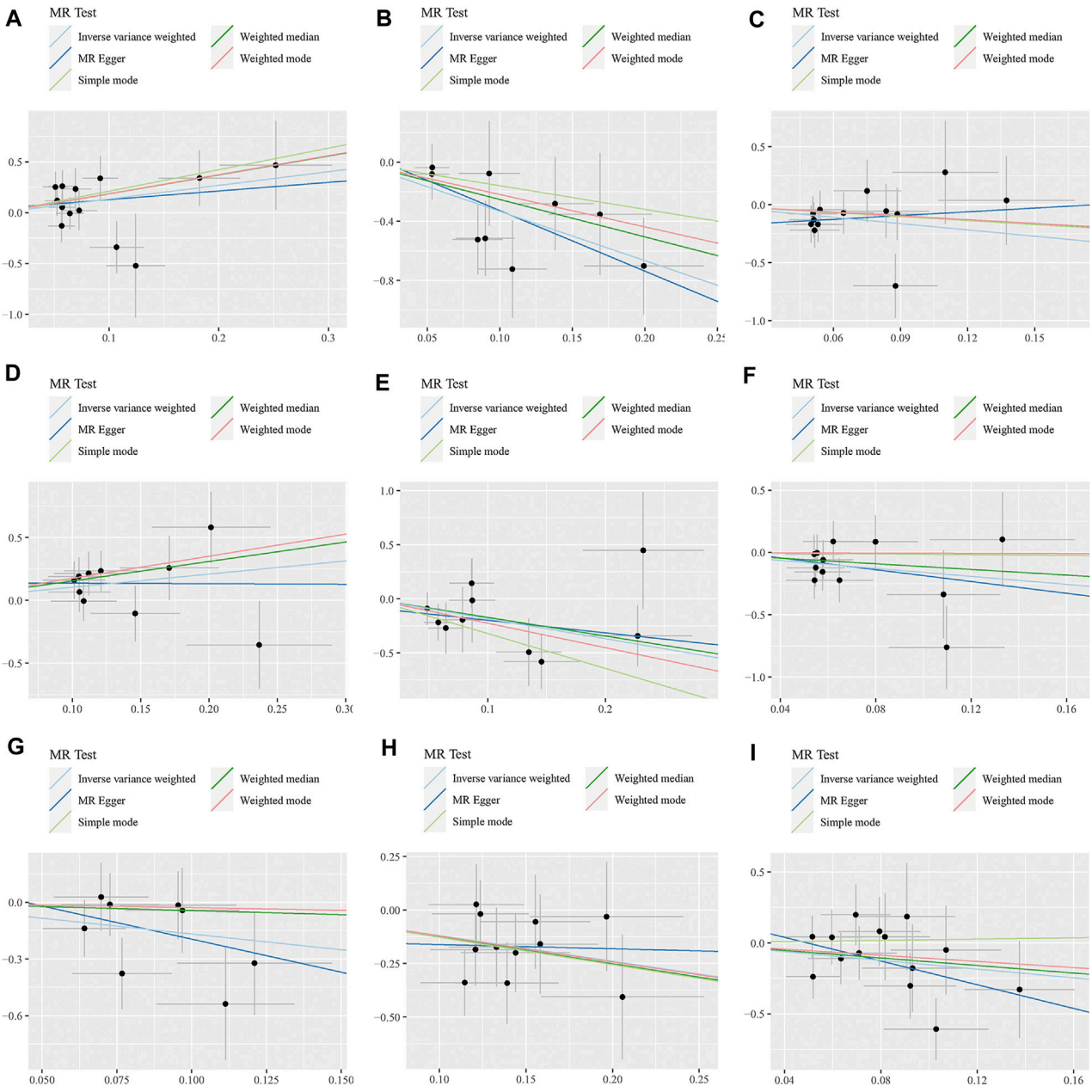


FIGURE 4
Scatter plots for the causal association between 9 gut microbiota and GBM. **(A)** A. *family.Peptostreptococcaceae;* **(B)** *family.Ruminococcaceae;* **(C)** *genus.Anaerostipes;* **(D)** *genus.Eubacterium brachy group;* **(E)** *genus.Faecalibacterium;* **(F)** *genus.Lachnospiraceae UCG004;* **(G)** *genus.Phascolarctobacterium;* **(H)** *genus.Prevotella7;* **(I)** *genus.Streptococcus.*

resist the colonization and invasion of harmful microorganisms, emerging as an integral component of the human defense against external threats. With this in mind, we focused on exploring whether changes in gut flora abundance are linked to the development of GBM as the central theme of this MR. The brain, characterized by a unique immune environment, establishes a crucial link between the gut microbiome and brain tumors through the gut-brain axis. The principal immune privilege in this connection arises from the presence of the blood-brain barrier (BBB), a highly specialized membrane barrier comprised of endothelial cells. The BBB regulates the entry of soluble substances, including antibodies, metabolites, signaling molecules, and immune cells, into the CNS (Obermeier et al., 2013). Experimental studies have elucidated bidirectional communication pathways linking the gut and the brain, encompassing diverse mechanisms such as neural, endocrine, and inflammatory pathways. These pathways are subject to modulation by alterations in gut wall integrity and BBB permeability. Comparable mechanisms are observed between the gut flora and GBM. Notably, when GBM manifests, it disrupts the BBB, facilitating the infiltration of immune cells from the body into the brain parenchyma. Within this specific microenvironment, these immune cells might experience a context where their

TABLE 2 Heterogeneity and sensitivity analyses of MR. Abbreviations: MR, Mendelian randomization analysis; SNPs, Number of single nucleotide polymorphism; GBM, Glioblastoma; IVW, Inverse Variance Weighted; MR-PRESSO, Mendelian Randomization Pleiotropy RESidual Sum and Outlier.

Gut microbiota	Outcome	Methods	Q	*P*	Intercept	*P*	MR-PRESSO
family.Peptostreptococcaceae	**GBM**						
		IVW	12.152	0.434	0.044	0.735	0.530
		MR-Egger	12.019	0.362			
family.Ruminococcaceae	**GBM**						
		IVW	5.115	0.745	0.082	0.670	0.600
		MR-Egger	4.917	0.670			
genus.Anaerostipes	**GBM**						
		IVW	8.193	0.770	−0.191	0.338	0.800
		MR-Egger	7.190	0.783			
genus.Eubacterium brachy group	**GBM**						
		IVW	8.011	0.533	0.141	0.562	0.850
		MR-Egger	7.645	0.469			
genus.Faecalibacterium	**GBM**						
		IVW	7.745	0.560	−0.083	0.586	0.540
		MR-Egger	7.422	0.492			
genus.Lachnospiraceae UCG004	**GBM**						
		IVW	8.505	0.668	0.051	0.817	0.720
		MR-Egger	8.448	0.585			
genus.Phascolarctobacterium	**GBM**						
		IVW	5.796	0.670	0.207	0.530	0.660
		MR-Egger	5.359	0.616			
genus.Prevotella7	**GBM**						
		IVW	5.465	0.858	−0.142	0.683	0.870
		MR-Egger	5.287	0.809			
genus.Streptococcus	**GBM**						
		IVW	11.041	0.607	0.207	0.343	0.640
		MR-Egger	10.066	0.610			

functionality becomes suppressed (Oberoi et al., 2016; Zhou et al., 2017). This immunosuppression potentially hampers the efficacy of GBM immunotherapy. Hence, there arises a critical consideration: balancing the composition and abundance of gut microbiota could attenuate immunosuppression within the microenvironment surrounding GBM. This modulation may, in turn, potentiate specific therapeutic effects of GBM.

Gut microbiota may regulate astrocyte activity through microbial metabolism that activates the astrocytic aromatic hydrocarbon receptor (AHR). It has been demonstrated that gut commensal microbiota degrade ichthyosine, producing metabolites that reach the CNS and activate the AHR in astrocytes, thereby limiting CNS inflammation (Rothhammer et al., 2018). Aromatic hydrocarbon receptor signaling intricately regulates peripheral T cell differentiation. Additionally, peripheral T cells recruited to the CNS exert control over astrocytic and microglial responses (Rothhammer and Quintana, 2019). Gramarzki et al. reported that aromatic hydrocarbon receptors in GBM cells drive TGF-B expression. Moreover, they highlighted that aromatic hydrocarbon receptor signaling promotes an immunosuppressive microenvironment in GBM (Gramatzki et al., 2009). These findings collectively suggest that gut microbiota may wield a pivotal role in GBM immune evasion by modulating AHR and, consequently, glioma development. Furthermore, they propose the potential of gut microbiota as therapeutic targets for GBM. The microbiota can regulate local and systemic intestinal immunity, particularly in the induction and maturation of immune cells in the nervous system. Gut microbiota dysregulation has been reported to down-regulate granulocyte macrophage colony-stimulating factor (GM-CSF) signal transduction, leading to significant expression of reactive

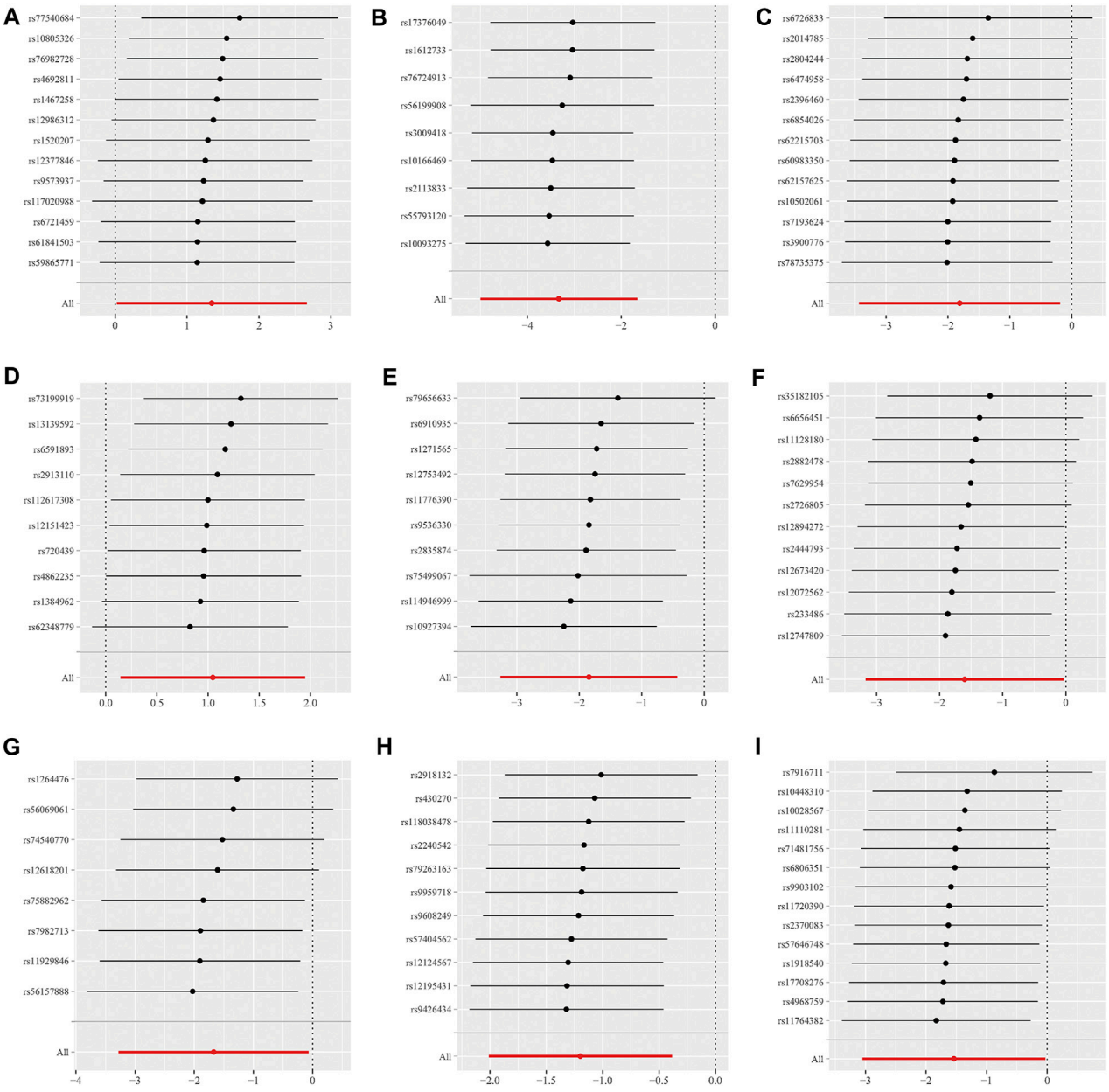


FIGURE 5
Scatter plots for the causal association between 9 gut microbiota and GBM. **(A)** *family.Peptostreptococcaceae;* **(B)** *family.Ruminococcaceae;* **(C)** *genus.Anaerostipes;* **(D)** *genus.Eubacterium brachy group;* **(E)** *genus.Faecalibacterium;* **(F)** *genus.Lachnospiraceae UCG004;* **(G)** *genus.Phascolarctobacterium;* **(H)** *genus.Prevotella7;* **(I)** *genus.Streptococcus.*

oxygen species (ROS) in activated immature myelocytes, thereby increasing the inhibitory activity of MDSC against T cells (Deh et al., 2019). In addition, dysregulation of the gut microbiota affects the balance between anti-inflammatory Tregs and pro-inflammatory Th17 cells (Chen and Tang, 2021), downregulates Foxp3 expression on tumor cells (Fan et al., 2022), and leads to inhibition of glioma cell growth and apoptosis.

Changes in the gut microbiota composition alter gut immune-brain communication and promote GBM development by creating a tumor-tolerant microenvironment in the CNS (DAlessandro et al., 2020). Recent studies have shown that after the development of GBM, a significant increase in the structure of the bacterial flora is observed, with a significant increase in *Bacteroidetes*, a decrease in the level of *Bacteroidetes* thickeniensis, an increase in the number of *Ackermannia* and *Verrucomicrobia*, and a decrease in the intestinal metabolites propionic, butyric, and acetic acids (Dono et al., 2020). Disruption of the gut microbiota further alters the tumor microenvironment and affects the antitumor efficacy of chemotherapy (Viaud et al., 2013; Daillère et al., 2016). The effects of chemotherapy have been shown to be remarkable in the treatment of tumors. Notably, the microbiota changes differently at different stages after temozolomide treatment. Specifically, there is an increase in the number of *Ackermannia*, *Bifidobacterium*, and *Verrucomicrobium* 7 days after the first

temozolomide treatment. Additionally, an increase in the number of *Ackermannia* is observed in patients who responded positively to immunotherapy with PD-1 blockade, suggesting its potential role in mediating the tumor response to immunotherapy (Routy et al., 2018).

The strength of this study lies in the identification of a causal relationship, providing potential gut microbiota candidates for subsequent functional studies. However, several limitations should be considered: (Sung et al., 2021): the MR analysis utilized GWAS data from a European population, necessitating replication in diverse populations; (Ostrom et al., 2023); the study included a limited range of gut microbiota; obtaining GWAS data from additional gut microbiota was crucial for a more comprehensive exploration of their association with GBM; (Omuro and DeAngelis, 2013); while MR is a highly efficient causal analysis method, validating the potential causal link between gut microbiota and GBM requires animal experiments. Finally, (Chen et al., 2021), the causal relationship between gut microbiota and GBM is multifaceted; exploring the etiology and pathogenesis demands a multi-perspective investigation.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding authors.

Author contributions

SW: Conceptualization, Data curation, Formal Analysis, Funding acquisition, Writing–original draft. FY: Investigation, Methodology, Resources, Writing–review and editing. ZG: Software, Supervision, Writing–original draft. RL: Conceptualization, Investigation, Writing–original draft. WS: Investigation, Resources, Visualization, Writing–original draft. YW: Conceptualization, Formal Analysis, Writing–original draft. YG: Methodology, Writing–original draft. CS: Project administration, Investigation, Methodology, Writing–original draft. DS: Funding acquisition, Project administration, Resources, Visualization, Writing–review and editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by the General Program of the National Natural Science Foundation of China (82070554 and 81770537); the Major Scientific and Technological Special Project for Public Health in Tianjin (21ZXGWSY00080); the Tianjin Medical University General Hospital Clincal Research Program (22ZYYLCCG06) to DS.

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/fgene.2023.1308263/full#supplementary-material

SUPPLEMENTARY TABLE S1
Effect estimates of the associations between 196 gut microbiota and risk of GBM in MR analyses.

SUPPLEMENTARY TABLE S2
The F number and R2 detect the intensity of the IVs between 196 gut microbiota and risk of GBM in MR analyses.

SUPPLEMENTARY TABLE S3
MR-Egger regression analysis between 196 gut microbiota and risk of GBM in MR analyses.

SUPPLEMENTARY TABLE S4
Testing for heterogeneity between 196 gut microbiota and risk of GBM in MR analyses.

SUPPLEMENTARY TABLE S5
Leave one out between 196 gut microbiota and risk of GBM in MR analyses.

SUPPLEMENTARY TABLE S6
Effect estimates of the associations between GBM and risk of gut microbiota in the reverse MR analyses.

References

Bellerba, F., Serrano, D., Johansson, H., Pozzi, C., Segata, N., NabiNejad, A., et al. (2022). Colorectal cancer, Vitamin D and microbiota: a double-blind Phase II randomized trial (ColoViD) in colorectal cancer patients. *Neoplasia* 34, 100842. doi:10.1016/j.neo.2022.100842

Bothwell, L. E., and Podolsky, S. H. (2016). The emergence of the randomized, controlled trial. *N. Engl. J. Med.* 375, 501–504. doi:10.1056/NEJMp1604635

Carasso, S., Fishman, B., Lask, L. S., Shochat, T., Geva-Zatorsky, N., and Tauber, E. (2021). Metagenomic analysis reveals the signature of gut microbiota associated with human chronotypes. *Faseb J.* 35, e22011. doi:10.1096/fj.202100857RR

Chen, F., Wendl, M. C., Wyczalkowski, M. A., Bailey, M. H., Li, Y., and Ding, L. (2021). Moving pan-cancer studies from basic research toward the clinic. *Nat. Cancer* 2, 879–890. doi:10.1038/s43018-021-00250-4

Chen, P., and Tang, X. (2021). Gut microbiota as regulators of Th17/treg balance in patients with myasthenia gravis. *Front. Immunol.* 12, 803101. doi:10.3389/fimmu.2021.803101

Daillère, R., Vétizou, M., Waldschmitt, N., Yamazaki, T., Isnard, C., Poirier-Colame, V., et al. (2016). Enterococcus hirae and barnesiella intestinihominis facilitate

cyclophosphamide-induced therapeutic immunomodulatory effects. *Immunity* 45, 931–943. doi:10.1016/j.immuni.2016.09.009

DAlessandro, G., Antonangeli, F., Marrocco, F., Porzia, A., Lauro, C., Santoni, A., et al. (2020). Gut microbiota alterations affect glioma growth and innate immune cells involved in tumor immunosurveillance in mice. *Eur. J. Immunol.* 50, 705–711. doi:10.1002/eji.201948354

de Clercq, N. C., van den Ende, T., Prodan, A., Hemke, R., Davids, M., Pedersen, H. K., et al. (2021). Fecal microbiota transplantation from overweight or obese donors in cachectic patients with advanced gastroesophageal cancer: a randomized, double-blind, placebo-controlled, phase II study. *Clin. Cancer Res.* 27, 3784–3792. doi:10.1158/1078-0432.Ccr-20-4918

Dehhaghi, M., Tan, V., Heng, B., Mohammadipanah, F., and Guillemin, G. J. (2019). Protective effects of myxobacterial extracts on hydrogen peroxide-induced toxicity on human primary astrocytes. *Neuroscience* 399, 1–11. doi:10.1016/j.neuroscience.2018.11.033

Dono, A., Patrizz, A., McCormack, R. M., Putluri, N., Ganesh, B. P., Kaur, B., et al. (2020). Glioma induced alterations in fecal short-chain fatty acids and neurotransmitters. *CNS Oncol.* 9, Cns57. doi:10.2217/cns-2020-0007

Fan, Y., Su, Q., Chen, J., Wang, Y., and He, S. (2022). Gut microbiome alterations affect glioma development and Foxp3 expression in tumor microenvironment in mice. *Front. Oncol.* 12, 836953. doi:10.3389/fonc.2022.836953

Ferreiro, A., Crook, N., Gasparrini, A. J., and Dantas, G. (2018). Multiscale evolutionary dynamics of host-associated microbiomes. *Cell* 172, 1216–1227. doi:10.1016/j.cell.2018.02.015

Fung, T. C., Olson, C. A., and Hsiao, E. Y. (2017). Interactions between the microbiota, immune and nervous systems in health and disease. *Nat. Neurosci.* 20, 145–155. doi:10.1038/nn.4476

Gramatzki, D., Pantazis, G., Schittenhelm, J., Tabatabai, G., Köhle, C., Wick, W., et al. (2009). Aryl hydrocarbon receptor inhibition downregulates the TGF-beta/Smad pathway in human glioblastoma cells. *Oncogene* 28, 2593–2605. doi:10.1038/onc.2009.104

He, J., Chu, Y., Li, J., Meng, Q., Liu, Y., Jin, J., et al. (2022). Intestinal butyrate-metabolizing species contribute to autoantibody production and bone erosion in rheumatoid arthritis. *Sci. Adv.* 8, eabm1511. doi:10.1126/sciadv.abm1511

Kurilshikov, A., Medina-Gomez, C., Bacigalupe, R., Radjabzadeh, D., Wang, J., Demirkan, A., et al. (2021). Large-scale association analyses identify host factors influencing human gut microbiome composition. *Nat. Genet.* 53, 156–165. doi:10.1038/s41588-020-00763-1

Lopez-Siles, M., Duncan, S. H., Garcia-Gil, L. J., and Martinez-Medina, M. (2017). Faecalibacterium prausnitzii: from microbiology to diagnostics and prognostics. *Isme J.* 11, 841–852. doi:10.1038/ismej.2016.176

Mizrahi, I., Wallace, R. J., and Moraïs, S. (2021). The rumen microbiome: balancing food security and environmental impacts. *Nat. Rev. Microbiol.* 19, 553–566. doi:10.1038/s41579-021-00543-6

Modoux, M., Rolhion, N., Lefevre, J. H., Oeuvray, C., Nádvorník, P., Illes, P., et al. (2022). Butyrate acts through HDAC inhibition to enhance aryl hydrocarbon receptor activation by gut microbiota-derived ligands. *Gut Microbes* 14, 2105637. doi:10.1080/19490976.2022.2105637

Obermeier, B., Daneman, R., and Ransohoff, R. M. (2013). Development, maintenance and disruption of the blood-brain barrier. *Nat. Med.* 19, 1584–1596. doi:10.1038/nm.3407

Oberoi, R. K., Parrish, K. E., Sio, T. T., Mittapalli, R. K., Elmquist, W. F., and Sarkaria, J. N. (2016). Strategies to improve delivery of anticancer drugs across the blood-brain barrier to treat glioblastoma. *Neuro Oncol.* 18, 27–36. doi:10.1093/neuonc/nov164

Omuro, A., and DeAngelis, L. M. (2013). Glioblastoma and other malignant gliomas: a clinical review. *Jama* 310, 1842–1850. doi:10.1001/jama.2013.280319

Ostrom, Q. T., Price, M., Neff, C., Cioffi, G., Waite, K. A., Kruchko, C., et al. (2023). CBTRUS statistical report: primary brain and other central nervous system tumors diagnosed in the United States in 2016-2020. *Neuro Oncol.* 25, iv1–iv99. doi:10.1093/neuonc/noad149

Park, E. M., Chelvanambi, M., Bhutiani, N., Kroemer, G., Zitvogel, L., and Wargo, J. A. (2022). Targeting the gut and tumor microbiota in cancer. *Nat. Med.* 28, 690–703. doi:10.1038/s41591-022-01779-2

Reichard, C. A., Naelitz, B. D., Wang, Z., Jia, X., Li, J., Stampfer, M. J., et al. (2022). Gut microbiome-dependent metabolic pathways and risk of lethal prostate cancer: prospective analysis of a PLCO cancer screening trial cohort. *Cancer Epidemiol. Biomarkers Prev.* 31, 192–199. doi:10.1158/1055-9965.Epi-21-0766

Rothhammer, V., Borucki, D. M., Tjon, E. C., Takenaka, M. C., Chao, C. C., Ardura-Fabregat, A., et al. (2018). Microglial control of astrocytes in response to microbial metabolites. *Nature* 557, 724–728. doi:10.1038/s41586-018-0119-x

Rothhammer, V., and Quintana, F. J. (2019). The aryl hydrocarbon receptor: an environmental sensor integrating immune responses in health and disease. *Nat. Rev. Immunol.* 19, 184–197. doi:10.1038/s41577-019-0125-8

Routy, B., Le Chatelier, E., Derosa, L., Duong, C. P. M., Alou, M. T., Daillère, R., et al. (2018). Gut microbiome influences efficacy of PD-1-based immunotherapy against epithelial tumors. *Science* 359, 91–97. doi:10.1126/science.aan3706

Sekula, P., Del Greco, M. F., Pattaro, C., and Köttgen, A. (2016). Mendelian randomization as an approach to assess causality using observational data. *J. Am. Soc. Nephrol.* 27, 3253–3265. doi:10.1681/asn.2016010098

Sepich-Poore, G. D., Zitvogel, L., Straussman, R., Hasty, J., Wargo, J. A., and Knight, R. (2021). The microbiome and human cancer. *Science* 371, eabc4552. doi:10.1126/science.abc4552

Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., et al. (2015). UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med.* 12, e1001779. doi:10.1371/journal.pmed.1001779

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., et al. (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* 71, 209–249. doi:10.3322/caac.21660

Teng, Y., Mu, J., Xu, F., Zhang, X., Sriwastva, M. K., Liu, Q. M., et al. (2022). Gut bacterial isoamylamine promotes age-related cognitive dysfunction by promoting microglial cell death. *Cell Host Microbe* 30, 944–960.e8. doi:10.1016/j.chom.2022.05.005

Viaud, S., Saccheri, F., Mignot, G., Yamazaki, T., Daillère, R., Hannani, D., et al. (2013). The intestinal microbiota modulates the anticancer immune effects of cyclophosphamide. *Science* 342, 971–976. doi:10.1126/science.1240537

Zhang, Q., Wu, Y., Wang, J., Wu, G., Long, W., Xue, Z., et al. (2016). Accelerated dysbiosis of gut microbiota during aggravation of DSS-induced colitis by a butyrate-producing bacterium. *Sci. Rep.* 6, 27572. doi:10.1038/srep27572

Zhou, W., Chen, C., Shi, Y., Wu, Q., Gimple, R. C., Fang, X., et al. (2017). Targeting glioma stem cell-derived pericytes disrupts the blood-tumor barrier and improves chemotherapeutic efficacy. *Cell Stem Cell* 21, 591–603.e4. doi:10.1016/j.stem.2017.10.002

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 08 January 2024
DOI 10.3389/fgene.2023.1340245

OPEN ACCESS

EDITED BY
Ziheng Wang,
University of Macau, China

REVIEWED BY
Zhengxiao Wei,
Public Health Clinical Center of Chengdu, China
Lin Chen,
Fujian Medical University, China

*CORRESPONDENCE
Gong-Chang Yu,
yugongchang@sdfmu.edu.cn
Ying Yin,
563298098@qq.com
Fan-Jie Liu,
50180479@qq.com
Bin Shi,
bshi@sdfmu.edu.cn

RECEIVED 17 November 2023
ACCEPTED 18 December 2023
PUBLISHED 08 January 2024

CITATION
Jia X-M, Deng T-T, Su H, Shi H-J, Qin H, Yu G-C, Yin Y, Liu F-J and Shi B (2024), Genetic causality and site-specific relationship between sarcopenia and osteoarthritis: a bidirectional Mendelian randomization study.
Front. Genet. 14:1340245.
doi: 10.3389/fgene.2023.1340245

Genetic causality and site-specific relationship between sarcopenia and osteoarthritis: a bidirectional Mendelian randomization study

Xue-Min Jia[1,2], Ting-Ting Deng[3], Hang Su[4], Hao-Jun Shi[5], Hao Qin[2], Gong-Chang Yu[1]*, Ying Yin[6]*, Fan-Jie Liu[1,2]* and Bin Shi[1,2]*

[1]Shandong First Medical University and Shandong Academy of Medical Sciences, Jinan, Shandong, China, [2]Neck-Shoulder and Lumbocrural Pain Hospital of Shandong First Medical University, Shandong First Medical University and Shandong Academy of Medical Sciences, Jinan, Shandong, China, [3]College of Traditional Chinese Medicine, Shandong University of Traditional Chinese Medicine, Jinan, China, [4]College of Rehabilitation, Shandong University of Traditional Chinese Medicine, Jinan, Shandong, China, [5]School of TCM, Macau University of Science and Technology, Macau, China, [6]Affiliated Hospital of Shandong University of Traditional Chinese Medicine, Jinan, Shandong, China

Background: Previous studies demonstrated a controversial relationship between sarcopenia (SP) and osteoarthritis (OA) and their genetic causality is unclear. Thus, we conducted a Mendelian randomization (MR) analysis to evaluate the possible causal association between sarcopenia-related traits (appendicular lean mass (ALM), grip strength, usual walking pace) and OA.

Method: We used pooled genetic data from the UK Biobank for ALM(n = 450,243), left-hand grip strength (n = 461,026), right-hand grip strength (n = 461,089) and usual walking pace (n = 459,915). Moreover, summary statistics for OA were obtained from the latest study conducted by the Genetics of Osteoarthritis Consortium, including all OA (n = 826,690), hand OA (n = 303,7782), hip OA (n = 353,388) and knee OA (n = 396,054). The primary method for estimating causal effects was the inverse-variance weighted (IVW) method, with the utilizing of false discovery rate adjusted *p* values (P_{FDR}). Additional MR methods such as MR-Egger regression, MR pleiotropy residual sum and outlier (MR-PRESSO), weighted median were employed as supplementary analyses.

Results: We discovered ALM (odds ratio (OR) = 1.103, 95% confidence interval (CI) = 1.052–1.156, P_{FDR} = 2.87E-04), hand grip strength (left, IVW OR = 0.823, 95% CI = 0.712 to 0.952, P_{FDR} = 0.020; right, OR = 0.826, 95% CI = 0.718 to 0.950, P_{FDR} = 0.020), and usual walking pace (OR = 0.339, 95% CI = 0.204 to 0.564, P_{FDR} = 2.38E-04) were causally associated with OA risk. In the reverse MR analysis, we identified a causal effect of OA on ALM (β = −0.258, 95% CI = −0.369 to 0.146, P_{FDR} = 0.6.07E-06), grip strength (left, β = −0.064, 95% CI = −0.104 to 0.024, P_{FDR} = 0.002; right, β = −0.055, 95% CI = −0.095 to 0.014, P_{FDR} = 0.008), and usual walking pace (β = −0.104, 95% CI = −0.147 to 0.061, P_{FDR} = 1.61E-05).

Conclusion: This present study suggests an obvious causality of SP on OA, with condition exhibiting site-specific effects, while evidence was also provided for the causal effect of OA on SP.

KEYWORDS

sarcopenia, osteoarthritis, Mendelian randomization, degenerative musculoskeletal diseases, causal relationship

Introduction

Sarcopenia (SP) and osteoarthritis (OA) as degenerative musculoskeletal diseases (DMD) emerged as major challenges for the aging population (Yin et al., 2023). SP is a muscle disease (muscle failure) (Cruz-Jentoft et al., 2019), characterized by an accelerated loss of muscle mass and function (Cruz-Jentoft and Sayer, 2019). Currently, about 50 million people worldwide suffer from SP, and its prevalence increases with age (Hida et al., 2014). Some studies have shown that the global prevalence of SP in people over 60 years old ranges from 10.00% to 27.00%, and in people over 80 years old is as high as 50.00% (Therakomen et al., 2020; Petermann-Rocha et al., 2022). OA is a degenerative disease with clinical manifestations of chronic pain, joint stiffness, and swelling (Cho et al., 2021), which afflicts more than 500 million people worldwide and has become the leading cause of chronic pain and disability in older adults (Wen and Xiao, 2022). As SP and OA are often diagnosed as comorbidities clinically, epidemiologic studies are increasingly examining the relationship between these two prevalent diseases.

Several studies have demonstrated a significant interaction between SP and OA. SP or its related traits are likely to be associated with outcomes in OA (predominantly OA of the knee). James S Andrews et al. showed that ALM and grip strength may be related to the development of KOA in older men (Andrews et al., 2021). On the other hand, a systematic review and meta-analysis showed that the prevalence of sarcopenia was more than twice as high in patients with osteoarthritis of the knee compared with controls (Pegreffi et al., 2023). However, there are also some studies do not agree with the aforementioned notion, and they consider the interconnection of the two to be highly controversial (Jones et al., 2021; Mezian et al., 2021; Tzartza et al., 2023). In addition, for ethical and practical purposes, a causal association between the two diseases could not be proved by a randomized controlled trial (RCT).

Mendelian randomization (MR) is a data analysis technique used to evaluate etiologic inferences in epidemiologic studies. The technique utilizes genetic variation as an instrumental variable (IV) to estimate reliable causal associations between exposures and outcomes (Burgess et al., 2019; Richmond and Davey Smith, 2022). Based on a two-sample bidirectional MR framework, we examined the potential causality of all OA, hand OA, hip OA, and knee OA with SP-related traits (appendicular lean mass (ALM), hand grip strength (left), hand grip strength (right), and usual walking pace).

Materials and methods

Study design overview

Figure 1 illustrates the design of our bidirectional MR study. We first estimated the causal effect of SP-related traits on OA and then assessed the causal effect of OA on SP-related traits. Genetic variants were considered as IVs only if they met the following three strict core assumptions. First, genetic variants were highly correlated with exposure. Second, genetic variants are not associated with confounding factors. Finally, genetic variation cannot act directly on the outcome, its effect on the outcome can only be reflected by exposure.

Data sources for sarcopenia-related traits

All summary-level genetic data for three SP-related traits were obtained from the UK Biobank (UKB). UKB is a large-scale repository of biomedical genes and information resources, containing about half a million people. The repository includes samples of volunteers' genetic information, lifestyle choices, and pedigree data (Sudlow et al., 2015). Identification of ALM, hand grip strength, and usual walking pace as consensus diagnostic criteria for SP was based on the report of the European Working Group on Sarcopenia in Older People (EWGSOP) (Cruz-Jentoft et al., 2010). In terms of the ALM, pooled data were analyzed for 450,243 UKB cohort participants, and ALM-related values were quantified and adjusted (Pei et al., 2020). Grip strength data were obtained from UKB's summary of hand grip strength (left and right) for 461,026 and 461,089 European ethnicity (Mitchell et al., 2019), calibrated to hand size, and adjusted for factors such as age and sex. Finally, for the "usual walking pace", data was similarly summarized from the GWAS summary data collected by the UKB for 459,915 European populations (Mitchell et al., 2019).

Data sources for all and site-specific OA

Summary statistics for all OA and its specific sites were obtained from the Genetics of Osteoarthritis (GO) Consortium's GWAS meta-analysis involving 826,690 individuals (177,517 patients with OA and 649,173 controls) from nine different populations (Boer et al., 2021). OA was defined by the GO based on self-reported status, hospital diagnosis, the 10th edition of the International Classification of Diseases (ICD-10) code, or TREAT-OA Consortium-defined radiology definition. The study identified 100 independently associated risk variants in 11 OA phenotypes and is the most recent and comprehensive GWAS analysis known for OA. We selected four phenotypes for analysis based on research needs: all OA (n = 826,690), hand OA (n = 303,782), hip OA (n = 353,388) and knee OA (n = 396,054).

Table 1 shows additional details such as phenotypes of all study participants.

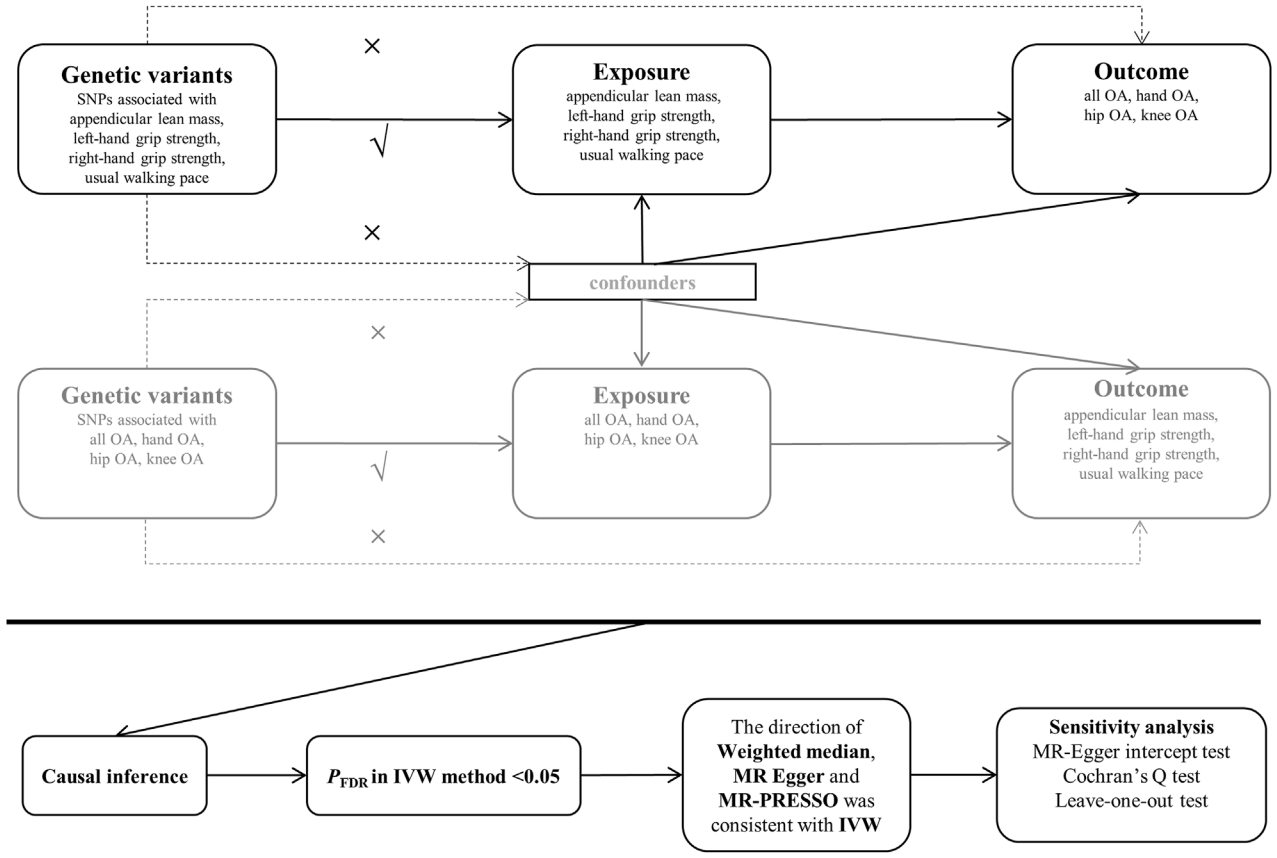


FIGURE 1
Generalized diagram of the study process. Abbreviations: OA: osteoarthritis; MR: Mendelian randomization;P_{FDR}: The inverse-variance weighted (IVW) method with false discovery rate (FDR) adjusted *p* values.

TABLE 1 Data sources used in this study.

Phenotype	Sample size (cases/controls)	Population	GWAS ID or PMID
ALM	450,243	European	ebi-a-GCST90000025
hand grip strength (left)	461,026	European	ukb-b-7478
hand grip strength (right)	461,089	European	ukb-b-10215
usual walking pace	459,915	European	ukb-b-4711
All OA	826,690 (177,517/649,173)	Mixed	34,822,786
Hand OA	303,782 (20,901/282,881)	Mixed	34,822,786
Hip OA	353,388 (36,445/316,943)	Mixed	34,822,786
Knee OA	396,054 (62,497/333,557)	Mixed	34,822,786

Abbreviations: ALM: appendicular lean mass; OA, osteoarthritis.

Selection of IVs

We first ensured that the SNPs of the three SP-related traits were genome-wide significant (p < 5E-08). Furthermore, to exclude the SNPs of strong linkage disequilibrium, we carried out an aggregation with r^2 < 0.001 and kb = 10,000 processes. Then, we also assessed the strength of IVs based on calculated R^2 [R^2 = 2 * MAF*(1-MAF) *beta (Cruz-Jentoft et al., 2019)] and F-statistics [F = R (Cruz-Jentoft et al., 2019)/(1-R^2) * (N-K-1)/K] for each SNP (Burgess et al., 2011). In order to reduce bias, the IVs with F-statistics less than 10 were excluded (Hemani et al., 2018a), and the SNPs and their related data such as F-statistics and R^2 for subsequent causal analysis are listed in Supplementary Table S1. In the reverse MR analysis, the SNP screening process was consistent with the aforementioned procedure.

TABLE 2 Sensitivity and IVW analysis of the causal effects of sarcopenia-related traits on osteoarthritis.

Exposures	Outcomes	No. of IVs	Heterogeneity test	MR Egger		IVW		
			Cochran's Q (*P*)	Intercept	*P*	OR (95% CI)	*P*	P_{FDR}
Appendicular lean mass	**All OA**	188	314.35 (<0.001)	0.002	0.160	1.103 (1.052,1.156)	5.38E-05	2.87E-04
Hand grip strength (left)	**All OA**	40	32.320 (0.767)	−0.001	0.827	0.823 (0.712,0.952)	0.009	0.020
Hand grip strength (right)	**All OA**	44	32.080 (0.889)	−0.001	0.706	0.826 (0.718,0.950)	0.007	0.020
Usual walking pace	**All OA**	18	63.293 (<0.001)	0.008	0.666	0.339 (0.204,0.564)	2.98E-05	2.38E-04
Appendicular lean mass	**Hand OA**	196	229.148 (0.047)	−4.86E-04	0.852	1.123 (1.014,1.243)	0.026	0.052
Hand grip strength (left)	**Hand OA**	52	84.381 (0.002)	−0.013	0.213	0.817 (0.535,1.250)	0.352	0.402
Hand grip strength (right)	**Hand OA**	59	63.282 (0.295)	−0.008	0.326	0.768 (0.550,1.073)	0.122	0.195
Usual walking pace	**Hand OA**	18	28.491 (0.040)	0.015	0.663	1.026 (0.412,2.556)	0.956	0.956
Appendicular lean mass	**Hip OA**	194	320.495 (<0.001)	0.004	0.070	1.095 (1.001,1.197)	0.048	0.085
Hand grip strength (left)	**Hip OA**	44	31.552 (0.902)	0.003	0.608	1.159 (0.885,1.518)	0.282	0.348
Hand grip strength (right)	**Hip OA**	50	37.682 (0.880)	0.001	0.847	1.194 (0.930,1.534)	0.165	0.240
Usual walking pace	**Hip OA**	16	13.983 (0.527)	0.013	0.546	0.474 (0.272,0.827)	0.008	0.020
Appendicular lean mass	**Knee OA**	205	472.488 (<0.001)	0.001	0.606	1.246 (1.151,1.350)	6.07E-08	9.71E-07
Hand grip strength (left)	**Knee OA**	50	94.915 (<0.001)	−0.013	0.051	1.182 (0.902,1.548)	0.226	0.301
Hand grip strength (right)	**Knee OA**	59	135.377 (<0.001)	−0.004	0.530	0.991 (0.753,1.305)	0.950	0.956
Usual walking pace	**Knee OA**	18	63.585 (<0.001)	0.010	0.744	0.299 (0.136,0.659)	0.003	0.011

Abbreviations: MR, mendelian randomization; IVW, inverse variance weighted; IVs, instrumental variables; P_{FDR}, false discovery rate (FDR) adjusted *p* values; CI, confidence interval; OA, osteoarthritis.

MR analysis

This study used R software (Version 4.3.0) and the two-sample MR package (Version 0.5.7) for data analysis (Yavorska and Burgess, 2017). We used the random-effects IVW method as the primary analytical method for MR estimation. This method was considered the most reliable in the absence of indications of directional pleiotropy in the selected IVs (Holmes et al., 2017). In addition, several sensitivity analyses were carried out, including weighted median (WM), MR-Egger, MR pleiotropy residual sum, and outlier (MR-PRESSO) test. Within this framework, the WM selected the median estimate to calculate the causal effects (Burgess et al., 2017). The MR-Egger regression method effectively tests the null causality hypothesis and gives consistent estimates of causality even if no genetic variation is valid. Additionally, the MR-Egger regression method is robust to horizontal pleiotropy (Bowden et al., 2015). The MR-PRESSO method detects pleiotropy, removes potentially pleiotropic IVs (outliers), and provides outlier-adjusted estimates (Verbanck et al., 2018). In order to correct the problem of multiple testing relatively gently, the *p*-value was adjusted using the false discovery rate (FDR) which is also called *q* value in the main IVW MR analysis, and the significance of causal inference was set to less than 0.05 (Chen et al., 2021). In addition, the statistical power was calculated with an online sample size and power calculator for MR (https://sb452.shinyapps.io/power/), and the results are shown in Supplementary Table S4.

Heterogeneity was assessed using Cochran's Q test, and $p < 0.05$ was considered statistically significant (Bowden et al., 2016). Moreover, we performed pleiotropic tests using MR-Egger intercept test and MR-PRESSO global testing to ensure that IVs do not influence the risk of outcome through other confounding factors or other biological pathways unrelated to exposure (Hemani et al., 2018b). Furthermore, we performed leave-one-out analyses to ensure the reliability of associations with individual SNPs.

Results

Overview

Based on the inverse-variance weighted (IVW) method, we observed significant evidence of a bi-directionally causal relationship between SP and OA. After FDR correction, most of the meaningful results were still retained. Additionally, the results were corroborated using other MR analysis methods.

The causal effects of SP-related traits on OA

After accounting for the independence of the genetic variation, none of the IVs for ALM, hand grip strength (left), hand grip strength (right), and usual walking pace that we obtained were without linkage disequilibrium (kb > 10,000 and $r^2 < 0.001$). Furthermore, all IVs reached genome-wide significance ($p < 5E-08$). Additionally, IVs with an F-statistic of less than 10 were considered weak instruments and were omitted from the MR

TABLE 3 Weighted median, MR Egger and MR PRESSO analysis of the causal effects of sarcopenia-related traits on osteoarthritis.

Exposures	Outcomes	No. of IVs	Weighted median		MR Egger		MR PRESSO	
			OR (95% CI)	*P*	OR (95% CI)	*P*	OR (95% CI)	*P*
Appendicular lean mass	**All OA**	188	1.096 (1.029,1.168)	0.004	1.025 (0.916,1.147)	0.670	1.098 (1.095,1.102)	1.08E-04
Hand grip strength (left)	**All OA**	40	0.896 (0.731,1.098)	0.290	0.888 (0.446,1.769)	0.737	0.823 (0.806,0.841)	0.006
Hand grip strength (right)	**All OA**	44	0.924 (0.754,1.131)	0.442	0.928 (0.500,1.722)	0.814	0.826 (0.811,0.841)	0.003
Usual walking pace	**All OA**	18	0.385 (0.236,0.629)	1.40E-04	0.131 (0.002,9.437)	0.366	0.296 (0.268,0.326)	4.85E-05
Appendicular lean mass	**Hand OA**	196	1.138 (0.984,1.315)	0.081	1.146 (0.906,1.448)	0.257	1.123 (1.115,1.131)	0.027
Hand grip strength (left)	**Hand OA**	52	0.934 (0.551,1.582)	0.799	2.540 (0.415,15.543)	0.318	0.817 (0.771,0.867)	0.356
Hand grip strength (right)	**Hand OA**	59	0.697 (0.439,1.107)	0.126	1.496 (0.384,5.822)	0.564	0.768 (0.736,0.802)	0.127
Usual walking pace	**Hand OA**	18	0.774 (0.272,2.202)	0.631	0.180 (0.001,416.109)	0.670	0.788 (0.658,0.945)	0.552
Appendicular lean mass	**Hip OA**	194	1.048 (0.932,1.179)	0.435	0.921 (0.750,1.132)	0.435	1.095 (1.088,1.102)	0.049
Hand grip strength (left)	**Hip OA**	44	1.185 (0.805,1.746)	0.390	0.863 (0.273,2.727)	0.803	1.159 (1.12,1.2)	0.216
Hand grip strength (right)	**Hip OA**	50	1.241 (0.865,1.78)	0.241	1.079 (0.379,3.076)	0.887	1.194 (1.157,1.232)	0.12
Usual walking pace	**Hip OA**	16	0.538 (0.252,1.147)	0.109	0.105 (0.001,12.883)	0.374	0.474 (0.415,0.542)	0.016
Appendicular lean mass	**Knee OA**	205	1.271 (1.155,1.399)	9.32E-07	1.194 (0.995,1.432)	0.057	1.256 (1.25,1.263)	1.04E-08
Hand grip strength (left)	**Knee OA**	50	0.964 (0.697,1.334)	0.824	3.382 (1.174,9.742)	0.029	1.124 (1.083,1.167)	0.385
Hand grip strength (right)	**Knee OA**	59	0.985 (0.744,1.305)	0.919	1.387 (0.472,4.078)	0.555	1.057 (1.022,1.093)	0.675
Usual walking pace	**Knee OA**	18	0.422 (0.204,0.871)	0.020	0.098 (0.001,74.976)	0.502	0.357 (0.306,0.416)	0.008

Abbreviations: MR, mendelian randomization; IVW, inverse variance weighted; IVs, instrumental variables; CI, confidence interval; OA, osteoarthritis.

analysis thus reducing the bias in the estimation of the IVs. Ultimately, we selected 583 SNPs, 127 SNPs, 144 SNPs, and 47 SNPs as IVs for ALM, hand grip strength (left), hand grip strength (right), and usual walking pace, respectively. Details of the IVs for SP-related traits are displayed in Supplementary Tables S1, S2.

As displayed in Table 2, a portion of Cochran's Q test was used to detect heterogeneity ($p < 0.05$), hence we used the IVW method in the random effects model. In the IVW model, adjusted for the FDR, genetically elevated SP-related traits levels were causally associated with all OA, ALM (Odds ratio (OR) = 1.103, 95% confidence interval (CI) = 1.052–1.156, P_{FDR} = 2.87E-04), hand grip strength (left, OR = 0.823, 95% CI = 0.712 to 0.952, P_{FDR} = 0.020; right, OR = 0.826, 95% CI = 0.718 to 0.950, P_{FDR} = 0.020), and usual walking pace (OR = 0.339, 95% CI = 0.204 to 0.564, P_{FDR} = 2.38E-04). The results of the WM and MR-PRESSO were consistent with the IVW (Table 3). No evidence of directional pleiotropy in the MR-Egger intercept test was observed for all IVs (p for intercept >0.05) (Table 2). Leave-one-out test indicated that SNPs without large effect sizes skewed the estimates (Supplementary Figure S1).

In the site-specific OA analysis, considering the results of the heterogeneity test, the IVW in the random effects model was used. In the IVW model, we observed that genetically determined levels of ALM were causally associated with knee OA (OR = 1.246, 95% CI = 1.151 to 1.350, P_{FDR} = 9.71E-07), usual walking pace was negatively correlated with hip OA (OR = 0.474, 95% CI = 0.272 to 0.827, P_{FDR} = 0.020) and knee OA (OR = 0.299, 95% CI = 0.136 to 0.659, P_{FDR} = 0.011) (Table 2). Moreover, WM and MR-PRESSO reached similar causal conclusions, and the MR-Egger effect estimate was in the same direction as IVW, which was considered supportive (Table 3). Further, no directional pleiotropy was detected for our selected IVs on MR-Egger analysis (Table 2). Leave-one-out test results were consistent with the above (Supplementary Figures S2–S4).

The causal effect of OA on SP-related traits

After similar screening criteria, 25 SNPs, 8 SNPs, 28 SNPs, and 22 SNPs were obtained as IVs for all OA, hand OA, hip OA, and knee OA, respectively (Supplementary Tables S1, S3). As Table 4 illustrates, heterogeneity was observed ($p < 0.05$) between the IVs of all the selected OA and the traits associated with SP. Therefore, we used the random effects model of IVW. The results of our analysis showed that the onset and progression of OA could lead to the worsening of sarcopenia-related traits [ALM (IVW β = −0.258, 95% CI = −0.369 to −0.146, P_{FDR} = 0.6.07E-06), grip strength (left, β = −0.064, 95% CI = −0.104 to −0.024, P_{FDR} = 0.002; right, β = −0.055, 95% CI = −0.095 to −0.014, P_{FDR} = 0.008), and usual walking pace (β = −0.104, 95% CI = −0.147 to −0.061, P_{FDR} = 1.61E-05)]. The WM and MR-PRESSO suggested similar findings, and the beta values of the MR-Egger are in the same direction (Table 5). MR-Egger intercept test did not demonstrate directional pleiotropy for our selected IVs (Table 4). The results of the leave-one-out- test displayed that SNPs without large effect sizes resulted in biases in the estimates (Supplementary Figure S5).

In the site-specific analysis, the results of using the IVW method in the random effects model pointed to a negative causal relationship between hand OA and grip strength (left, β = −0.102, 95%

TABLE 4 Sensitivity and IVW analysis of the causal effects of osteoarthritis on sarcopenia-related traits.

Exposures	Outcomes	No. of IVs	Heterogeneity test	MR Egger		IVW		
			Cochran's Q (*P*)	Intercept	*P*	Beta (95% CI)	*P*	P_{FDR}
All OA	**Appendicular lean mass**	11	72.193 (<0.001)	−0.002	0.891	−0.258 (-0.369, −0.146)	6.07E-06	3.24E-05
All OA	**Hand grip strength (left)**	20	53.312 (<0.001)	−0.002	0.676	−0.064 (-0.104, −0.024)	0.002	0.004
All OA	**Hand grip strength (right)**	19	48.792 (<0.001)	−0.001	0.758	−0.055 (-0.095, −0.014)	0.008	0.014
All OA	**Usual walking pace**	22	104.602 (<0.001)	−0.002	0.673	−0.104 (-0.147, −0.061)	2.01E-06	1.61E-05
Hand OA	**Appendicular lean mass**	7	186.506 (<0.001)	0.010	0.919	0.013 (-0.095,0.122)	0.81	0.810
Hand OA	**Hand grip strength (left)**	8	84.682 (<0.001)	−0.007	0.684	−0.102 (-0.149, −0.055)	2.26E-05	7.23E-05
Hand OA	**Hand grip strength (right)**	8	86.489 (<0.001)	−0.005	0.759	−0.104 (-0.152, −0.057)	1.84E-05	7.23E-05
Hand OA	**Usual walking pace**	8	26.509 (<0.001)	−0.001	0.937	−0.021 (-0.044,0.001)	0.062	0.100
Hip OA	**Appendicular lean mass**	22	366.557 (<0.001)	−0.004	0.654	0.047 (-0.006,0.100)	0.08	0.117
Hip OA	**Hand grip strength (left)**	28	266.255 (<0.001)	0.003	0.564	0.013 (-0.015,0.041)	0.359	0.414
Hip OA	**Hand grip strength (right)**	28	309.543 (<0.001)	0.003	0.566	0.011 (-0.019,0.041)	0.474	0.506
Hip OA	**Usual walking pace**	22	45.030 (0.002)	0.001	0.666	−0.025 (-0.038, −0.013)	9.81E-05	2.62E-04
Knee OA	**Appendicular lean mass**	15	316.000 (<0.001)	0.013	0.51	−0.179 (-0.281, −0.076)	0.001	0.001
Knee OA	**Hand grip strength (left)**	19	309.382 (<0.001)	0.003	0.841	0.029 (-0.092,0.034)	0.362	0.414
Knee OA	**Hand grip strength (right)**	19	342.572 (<0.001)	0.003	0.825	0.041 (-0.107,0.025)	0.228	0.304
Knee OA	**Usual walking pace**	20	70.118 (<0.001)	0.007	0.066	−0.064 (-0.087, −0.04)	8.55E-08	1.37E-06

Abbreviations: MR, mendelian randomization; IVW, inverse variance weighted; IVs, instrumental variables; P_{FDR}, false discovery rate (FDR) adjusted *p* values; CI, confidence interval; OA, osteoarthritis.

CI = −0.149 to −0.055, P_{FDR} = 7.23E-05; right, β = -0.104, 95% CI = −0.152 to −0.057, P_{FDR} = 7.23E-05). Hip OA (β = −0.025, 95% CI = −0.038 to −0.013, P_{FDR} = 2.62E-04) and knee OA (β = −0.064, 95% CI = −0.087 to −0.04, P_{FDR} = 1.37E-06), remained significantly causally and negatively correlated with usual walking pace (Table 4). In addition, we identified that only knee OA was causally related to ALM in the available site-specific cohorts (β = −0.179, 95% CI = −0.281 to −0.076, P_{FDR} = 0.001), the results of all other MR analysis methods were similar to those of the IVW (Table 5). Heterogeneity for the associations between the selected IVs of the knee OA and ALM was not observed, as in the study above no directional pleiotropy was detected (Table 4). The results of the leave-one-out test were also consistent with the aforementioned results (Supplementary Figures S6–S8).

Discussion

We investigated the potential causality between SP-related traits and OA. Following the FDR correction, the results demonstrated a possible causal relationship between SP-related traits and all OA, the usual walking pace and hip OA, the usual walking pace and knee OA, the ALM and knee OA, respectively. In addition, we found evidence of their previous reverse causation with each other. The results also indicated a causal relationship between hand OA and hand grip strength.

Firstly, SP increases the risk of OA. In recent years, studies have demonstrated a correlation of varying intensity between SP and OA (Jin et al., 2017; Vlietstra et al., 2019; Dalle and Koppo, 2020; Godziuk et al., 2021). Decreased muscle strength is the main feature of SP and previous experiments have demonstrated that decreased muscle strength or muscle weakness is a risk factor for the development and progression of OA (Tanaka et al., 2019; Xu et al., 2020). Andrews et al. discovered that in men only, the likelihood of knee OA was reduced for each standard deviation (SD) reduction in ALM (OR per SD reduction: 0.68; 95% CI: 0.47–0.97) (Andrews et al., 2021). Similarly, another study noted significant ALM or total lean mass and increased fat mass (FM) were associated with radiographic knee OA (Azuma et al., 2017). Optimizing medial femoral size is important in clinical management to reduce the progression of OA and subsequent knee arthroplasty, which in part reflects the important impact of SP on the development of OA (Wang et al., 2012).

In addition, OA has been noted as a risk factor for increasing the incidence of SP (Kemmler et al., 2015; Dharmakulsakti et al., 2022). A previous foundational experimental study provided some mechanistic insights, indicating that knee osteoarthritis induced by anterior cruciate ligament transection promotes remodeling and atrophy in the neuromuscular junctions of the quadriceps and tibialis anterior muscles. These changes were associated with signs of inflammation and alterations in muscle gene and protein expression (Cunha et al., 2019). Besides, some studies have indicated a strong relationship between OA and low skeletal muscle mass (Berenbaum and van den Berg, 2015; Jeon et al., 2019). A significant association between OA of the knee and

10.3389/fgene.2023.1340245

TABLE 5 Weighted median, MR Egger and MR PRESSO analysis of the causal effects of osteoarthritis on sarcopenia-related traits.

Exposures	Outcomes	No. of IVs	Weighted median		MR Egger		MR-PRESSO	
			Beta (95% CI)	*P*	Beta (95% CI)	*P*	Beta (95% CI)	*P*
All OA	**Appendicular lean mass**	11	−0.168 (-0.245, −0.092)	1.66E-05	−0.193 (-1.104,0.719)	0.689	−0.218 (-0.245, −0.191)	0.002
All OA	**Hand grip strength (left)**	20	−0.085 (-0.124, −0.046)	1.71E-05	−0.001 (-0.292,0.290)	0.993	−0.064 (-0.073, −0.055)	0.006
All OA	**Hand grip strength (right)**	19	−0.053 (-0.095, −0.012)	0.012	−0.009 (-0.297,0.278)	0.949	−0.055 (-0.064, −0.046)	0.016
All OA	**Usual walking pace**	22	−0.112 (-0.148, −0.076)	1.56E-09	−0.027 (-0.382,0.328)	0.884	−0.102 (-0.109, −0.094)	2.09E-05
Hand OA	**Appendicular lean mass**	7	0.014 (-0.019,0.047)	0.403	−0.112 (−2.417,2.192)	0.928	0.01 (0.003,0.017)	0.411
Hand OA	**Hand grip strength (left)**	8	−0.088 (-0.114, −0.062)	3.46E-11	−0.014 (-0.420,0.391)	0.948	−0.09 (-0.094, −0.085)	0.001
Hand OA	**Hand grip strength (right)**	8	−0.087 (-0.113, −0.061)	5.04E-11	−0.037 (-0.450,0.376)	0.866	−0.088 (-0.097, −0.079)	0.002
Hand OA	**Usual walking pace**	8	−0.018 (-0.034, −0.001)	0.032	−0.013 (−0.209,0.183)	0.899	−0.021 (-0.024, −0.018)	0.006
Hip OA	**Appendicular lean mass**	22	0.03 (0.002,0.059)	0.036	0.100 (−0.134,0.333)	0.413	0.029 (0.024,0.033)	0.023
Hip OA	**Hand grip strength (left)**	28	0.02 (0.004,0.036)	0.015	−0.026 (−0.159,0.107)	0.706	0.015 (0.012,0.017)	0.035
Hip OA	**Hand grip strength (right)**	28	0.016 (-0.001,0.032)	0.061	−0.031 (−0.175,0.113)	0.678	0.012 (0.009,0.016)	0.176
Hip OA	**Usual walking pace**	22	−0.023 (-0.038, −0.009)	0.001	−0.039 (−0.104,0.026)	0.248	−0.025 (-0.028, −0.022)	0.001
Knee OA	**Appendicular lean mass**	15	−0.083 (-0.139, −0.027)	0.004	−0.439 (−1.198,0.321)	0.278	−0.182 (-0.199, −0.165)	0.001
Knee OA	**Hand grip strength (left)**	19	0.001 (-0.036,0.039)	0.945	−0.090 (−0.679,0.499)	0.768	−0.032 (-0.041, −0.023)	0.127
Knee OA	**Hand grip strength (right)**	19	−0.039 (-0.075, −0.003)	0.035	−0.111 (−0.731,0.509)	0.729	−0.051 (-0.06, −0.042)	0.031
Knee OA	**Usual walking pace**	20	−0.048 (-0.07, −0.027)	7.94E-06	−0.202 (−0.343, −0.062)	0.011	−0.06 (-0.064, −0.056)	4.83E-06

Abbreviations: MR, mendelian randomization; IVW, inverse variance weighted; IVs, instrumental variables; CI, confidence interval; OA, osteoarthritis.

walking pace was observed ($p < 0.001$, OR:0.073) (Nakamura and Ogata, 2016). A recent study conducted a more comprehensive analysis, revealing that the OA group had statistically significantly worse SP parameters than the control group, with lower appendicular skeletal muscle mass ($p = 0.041$), impaired performance on the 40-m fast walk test ($p = 0.020$), and reduced right ($p < 0.01$) and left ($p < 0.01$) hand grip strength. The findings suggest an early onset of sarcopenia in these individuals (de Almeida et al., 2020).

According to previous studies, bone, muscle, and fat tissue are connected and interact with each other through molecules. And SP seems to have a bidirectional relationship with the maintenance or destruction of joint structures (Spanoudaki et al., 2023). SP and OA elevate the risk of mutual development, which is one of the most important reasons for their frequent coexistence (Peng and Zeng, 2022). A new concept of "sarcopenic knee OA" has also been proposed (Iijima and Aoyama, 2021). The study by Jiyong Yang et al. identified a common network of genetic interactions between KOA and SP, including 14 common differentially expressed genes, 4 hub genes, and 10 potential chemical compounds, among other important findings, which have updated the research results of the mechanism between OA and SP (Yang et al., 2023).

These previous studies support our view to some extent. Our study's methodology thus has some advantages. Firstly, the MR method can effectively avoid the drawbacks of traditional observational research methods such as residual confounding uncertainty and reverse causality. Secondly, the IVs for SP-related traits were obtained from the existing large GWAS, and the IVs for OA were obtained from the most recent GWAS. This allowed for a more precise assessment of effect sizes than would be possible with individual-level data or findings from studies with limited sample sizes. Lastly, we performed an analysis of the specific relationships between individual SP-related traits and the different sites of OA, which led to a more comprehensive understanding of the potential link between them.

Nevertheless, our study also has some limitations. Firstly, due to the lack of data on the large GWAS concerning SP, we could only use the related traits to analyze the relationship with OA. Secondly, selecting SNPs from the different large-sample GWAS studies may increase the risk of sample overlap between exposure and outcome variables, which may bias the results. Furthermore, due to the complexity of biological systems, bidirectional MR assumes that causality happens in one direction, and feedback loops may exist between the exposure and the outcome, which might affect the accuracy of the results.

Finally, since a majority of the participants included in the study were of European ancestry and were not representative of other racial groups, further magnetic resonance studies are needed to verify causality.

Conclusion

This present study suggests an obvious causality of SP on OA, with condition exhibiting site-specific effects, while evidence was also provided for the causal effect of OA on SP. It presents some evidence of reciprocal interaction between SP and OA, which may facilitate the development of novel treatment strategies for both diseases. However, the causal relationship between the two conditions still necessitates further investigation and substantiation through a multitude of studies.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding authors.

Author contributions

X-MJ: Writing–original draft, Writing–review and editing. T-TD: Methodology, Supervision, Validation, Writing–review and editing. HS: Resources, Visualization, Writing–review and editing. H-JS: Validation, Writing–review and editing. HQ: Validation, Writing–review and editing. G-CY: Conceptualization, Validation, Writing–review and editing. YY: Writing–review and editing. F-JL: Methodology, Supervision, Writing–review and editing. BS: Funding acquisition, Resources, Supervision, Validation, Writing–review and editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. The key R&D Program of Shandong Province (2022CXGC020510) and the Academic Promotion Program of Shandong First Medical University (2019QL003) funded this study.

Acknowledgments

We would like to thank Editage (www.editage.cn) for English language editing.

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2023.1340245/full#supplementary-material

References

Andrews, J. S., Gold, L. S., Nevitt, M., Heagerty, P. J., and Cawthon, P. M. (2021). Appendicular lean mass, grip strength, and the development of knee osteoarthritis and knee pain among older adults. *ACR open Rheumatol.* 3 (8), 566–572. doi:10.1002/acr2.11302

Azuma, K., Sera, Y., Shinjo, T., Takayama, M., Shiomi, E., Momoshima, S., et al. (2017). Maintenance of the paraspinal muscles may protect against radiographic knee osteoarthritis. *Open access rheumatology Res. Rev.* 9, 151–158. doi:10.2147/oarrr.S130688

Berenbaum, F., and van den Berg, W. B. (2015). Inflammation in osteoarthritis: changing views. *Osteoarthritis Cartilage. Nov* 23 (11), 1823–1824. doi:10.1016/j.joca.2015.09.012

Boer, C. G., Hatzikotoulas, K., Southam, L., Stefánsdóttir, L., Zhang, Y., Coutinho de Almeida, R., et al. (2021). Deciphering osteoarthritis genetics across 826,690 individuals from 9 populations. *Cell* 184 (18), 4784–4818. doi:10.1016/j.cell.2021.07.038

Bowden, J., Davey Smith, G., and Burgess, S. (2015). Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int. J. Epidemiol.* 44 (2), 512–525. doi:10.1093/ije/dyv080

Bowden, J., Del Greco, M. F., Minelli, C., Davey Smith, G., Sheehan, N. A., and Thompson, J. R. (2016). 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.* 45 (6), 1961–1974. doi:10.1093/ije/dyw220

Burgess, S., Bowden, J., Fall, T., Ingelsson, E., and Thompson, S. G. (2017). Sensitivity analyses for robust causal inference from mendelian randomization analyses with multiple genetic variants. *Epidemiol. Camb. Mass)* 28 (1), 30–42. doi:10.1097/ede.0000000000000559

Burgess, S., Davey Smith, G., Davies, N. M., Dudbridge, F., Gill, D., Glymour, M. M., et al. (2019). Guidelines for performing Mendelian randomization investigations. *Wellcome open Res.* 4, 186. doi:10.12688/wellcomeopenres.15555.2

Burgess, S., Thompson, S. G., and CRP CHD Genetics Collaboration (2011). Avoiding bias from weak instruments in Mendelian randomization studies. *Int. J. Epidemiol.* 40 (3), 755–764. doi:10.1093/ije/dyr036

Chen, X., Robinson, D. G., and Storey, J. D. (2021). The functional false discovery rate with applications to genomics. *Biostat. Oxf. Engl.* 22 (1), 68–81. doi:10.1093/biostatistics/kxz010

Cho, Y., Jeong, S., Kim, H., Kang, D., Lee, J., Kang, S. B., et al. (2021). Disease-modifying therapeutic strategies in osteoarthritis: current status and future directions. *Exp. Mol. Med.* 53 (11), 1689–1696. doi:10.1038/s12276-021-00710-y

Cruz-Jentoft, A. J., Baeyens, J. P., Bauer, J. M., Boirie, Y., Cederholm, T., Landi, F., et al. (2010). Sarcopenia: European consensus on definition and diagnosis: report of the European working group on sarcopenia in older people. *Age ageing* 39 (4), 412–423. doi:10.1093/ageing/afq034

Cruz-Jentoft, A. J., Bahat, G., Bauer, J., Boirie, Y., Bruyère, O., Cederholm, T., et al. (2019). Sarcopenia: revised European consensus on definition and diagnosis. *Age ageing* 48 (1), 16–31. doi:10.1093/ageing/afy169

Cruz-Jentoft, A. J., and Sayer, A. A. (2019). Sarcopenia. *Lancet (London, Engl.* 393 (10191), 2636–2646. doi:10.1016/s0140-6736(19)31138-9

Cunha, J. E., Barbosa, G. M., Castro, P., Luiz, B. L. F., Silva, A. C. A., Russo, T. L., et al. (2019). Knee osteoarthritis induces atrophy and neuromuscular junction remodeling in

the quadriceps and tibialis anterior muscles of rats. *Sci. Rep.* 9 (1), 6366. doi:10.1038/s41598-019-42546-7

Dalle, S., and Koppo, K. (2020). Is inflammatory signaling involved in disease-related muscle wasting? Evidence from osteoarthritis, chronic obstructive pulmonary disease and type II diabetes. *Exp. Gerontol.* 137, 110964. doi:10.1016/j.exger.2020.110964

de Almeida, A. C., Aily, J., Pedroso, M., Bergamaschi, J. R., de Carvalho Silva, G., and Mattiello, S. (2020). Middle-aged adults with knee osteoarthritis present worse sarcopenia parameters, symptoms and functional performance. *Osteoarthr. Cartil.* 28, S399–S400. doi:10.1016/j.joca.2020.02.623

Dharmakulsakti, P., Roopsawang, I., and Aree-Ue, S. (2022). Sarcopenia among older adults with knee osteoarthritis: a cross-sectional study of prevalence and its associated factors. *Pac. Rim Int. J. Nurs. Res.* 26 (1), 121–134.

Godziuk, K., Prado, C. M., Woodhouse, L. J., and Forhan, M. (2021). Associations between self-reported weight history and sarcopenic obesity in adults with knee osteoarthritis. *Obes. (Silver Spring, Md)* 29 (2), 302–307. doi:10.1002/oby.23074

Hemani, G., Bowden, J., and Davey Smith, G. (2018b). Evaluating the potential role of pleiotropy in Mendelian randomization studies. *Hum. Mol. Genet.* 27 (R2), R195–R208. doi:10.1093/hmg/ddy163

Hemani, G., Zheng, J., Elsworth, B., Wade, K. H., Haberland, V., Baird, D., et al. (2018a). The MR-Base platform supports systematic causal inference across the human phenome. *eLife* 7, e34408. doi:10.7554/eLife.34408

Hida, T., Harada, A., Imagama, S., and Ishiguro, N. (2014). Managing sarcopenia and its related-fractures to improve quality of life in geriatric populations. *Aging Dis.* 5 (4), 226–237. doi:10.14336/ad.2014.0500226

Holmes, M. V., Ala-Korpela, M., and Smith, G. D. (2017). Mendelian randomization in cardiometabolic disease: challenges in evaluating causality. *Nat. Rev. Cardiol.* 14 (10), 577–590. doi:10.1038/nrcardio.2017.78

Iijima, H., and Aoyama, T. (2021). Increased recurrent falls experience in older adults with coexisting of sarcopenia and knee osteoarthritis: a cross-sectional study. *BMC Geriatr.* 21 (1), 698. doi:10.1186/s12877-021-02654-4

Jeon, H., Lee, S. U., Lim, J. Y., Chung, S. G., Lee, S. J., and Lee, S. Y. (2019). Low skeletal muscle mass and radiographic osteoarthritis in knee, hip, and lumbar spine: a cross-sectional study. *Aging Clin. Exp. Res.* 31 (11), 1557–1562. doi:10.1007/s40520-018-1108-5

Jin, W. S., Choi, E. J., Lee, S. Y., Bae, E. J., Lee, T. H., and Park, J. (2017). Relationships among obesity, sarcopenia, and osteoarthritis in the elderly. *J. Obes. metabolic syndrome* 26 (1), 36–44. doi:10.7570/jomes.2017.26.1.36

Jones, T. L., Esa, M. S., Li, K. H. C., Krishnan, S. R. G., Elgallab, G. M., Pearce, M. S., et al. (2021). Osteoporosis, fracture, osteoarthritis and sarcopenia: a systematic review of circulating microRNA association. *Bone* 152, 116068. doi:10.1016/j.bone.2021.116068

Kemmler, W., Teschler, M., Goisser, S., Bebenek, M., von Stengel, S., Bollheimer, L. C., et al. (2015). Prevalence of sarcopenia in Germany and the corresponding effect of osteoarthritis in females 70 years and older living in the community: results of the FORMoSA study. *Clin. interventions aging* 10, 1565–1573. doi:10.2147/cia.S89585

Mezian, K., Angerová, Y., Kara, M., and Özçakar, L. (2021). Obesity paradox in sarcopenia and knee osteoarthritis: comment on the article by Andrews et al. *ACR open Rheumatol.* 3 (11), 812–813. doi:10.1002/acr2.11330

Mitchell, R., Elsworth, B., Mitchell, R., Raistrick, C. A., Paternoster, L., Hemani, G., et al. (2019). *MRC IEU UK Biobank GWAS pipeline version 2*. Bristol, UK: University of Bristol.

Nakamura, K., and Ogata, T. (2016). Locomotive syndrome: definition and management. *Clin. Rev. bone mineral metabolism* 14 (2), 56–67. doi:10.1007/s12018-016-9208-2

Pegreffi, F., Balestra, A., De Lucia, O., Smith, L., Barbagallo, M., and Veronese, N. (2023). Prevalence of sarcopenia in knee osteoarthritis: a systematic review and meta-analysis. *J. Clin. Med.* 12 (4), 1532. doi:10.3390/jcm12041532

Pei, Y. F., Liu, Y. Z., Yang, X. L., Zhang, H., Feng, G. J., Wei, X. T., et al. (2020). The genetic architecture of appendicular lean mass characterized by association analysis in the UK Biobank study. *Commun. Biol.* 3 (1), 608. doi:10.1038/s42003-020-01334-0

Peng, H., and Zeng, Y. (2022). Research progress on the correlation between sarcopenia and osteoarthritis. *Zhongguo xiu fu chong jian wai ke za zhi = Zhongguo xiufu chongjian waike zazhi = Chin. J. reparative Reconstr. Surg.* 36 (12), 1549–1557. doi:10.7507/1002-1892.202209015

Petermann-Rocha, F., Balntzi, V., Gray, S. R., Lara, J., Ho, F. K., Pell, J. P., et al. (2022). Global prevalence of sarcopenia and severe sarcopenia: a systematic review and meta-analysis. *J. cachexia, sarcopenia muscle* 13 (1), 86–99. doi:10.1002/jcsm.12783

Richmond, R. C., and Davey Smith, G. (2022). Mendelian randomization: concepts and scope. *Cold Spring Harb. Perspect. Med.* 12 (1), a040501. doi:10.1101/cshperspect.a040501

Spanoudaki, M., Giaginis, C., Mentzelou, M., Bisbinas, A., Solovos, E., Papadopoulos, K., et al. (2023). Sarcopenia and sarcopenic obesity and osteoarthritis: a discussion among muscles, fat, bones, and aging. *Life (Basel, Switz.* 13 (6), 1242. doi:10.3390/life13061242

Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., et al. (2015). UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med.* 12 (3), e1001779. doi:10.1371/journal.pmed.1001779

Tanaka, R., Hirohama, K., and Ozawa, J. (2019). Can muscle weakness and disability influence the relationship between pain catastrophizing and pain worsening in patients with knee osteoarthritis? A cross-sectional study. *Braz. J. Phys. Ther.* 23 (3), 266–272. doi:10.1016/j.bjpt.2018.08.011

Therakomen, V., Petchlorlian, A., and Lakananurak, N. (2020). Prevalence and risk factors of primary sarcopenia in community-dwelling outpatient elderly: a cross-sectional study. *Sci. Rep.* 10 (1), 19551. doi:10.1038/s41598-020-75250-y

Tzartza, C. L., Karapalis, N., Voulgaridou, G., Zidrou, C., Beletsiotis, A., Chatziprodromidou, I. P., et al. (2023). Comparative analysis on the effect of sarcopenia in patients with knee osteoarthritis before and after total knee arthroplasty. *Dis. (Basel, Switz.* 11 (1), 36. doi:10.3390/diseases11010036

Verbanck, M., Chen, C. Y., Neale, B., and Do, R. (2018). Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. *Nat. Genet.* 50 (5), 693–698. doi:10.1038/s41588-018-0099-7

Vlietstra, L., Stebbings, S., Meredith-Jones, K., Abbott, J. H., Treharne, G. J., and Waters, D. L. (2019). Sarcopenia in osteoarthritis and rheumatoid arthritis: the association with self-reported fatigue, physical function and obesity. *PloS one* 14 (6), e0217462. doi:10.1371/journal.pone.0217462

Wang, Y., Wluka, A. E., Berry, P. A., Siew, T., Teichtahl, A. J., Urquhart, D. M., et al. (2012). Increase in vastus medialis cross-sectional area is associated with reduced pain, cartilage loss, and joint replacement risk in knee osteoarthritis. *Arthritis rheumatism* 64 (12), 3917–3925. doi:10.1002/art.34681

Wen, C., and Xiao, G. (2022). Advances in osteoarthritis research in 2021 and beyond. *J. Orthop. Transl.* 32, A1–A2. doi:10.1016/j.jot.2022.02.011

Xu, J., She, G., Gui, T., Hou, H., Li, J., Chen, Y., et al. (2020). Knee muscle atrophy is a risk factor for development of knee osteoarthritis in a rat model. *J. Orthop. Transl.* 22, 67–72. doi:10.1016/j.jot.2019.10.003

Yang, J., Jiang, T., Xu, G., Wang, S., and Liu, W. (2023). Exploring molecular mechanisms underlying the pathophysiological association between knee osteoarthritis and sarcopenia. *Osteoporos. sarcopenia* 9 (3), 99–111. doi:10.1016/j.afos.2023.08.005

Yavorska, O. O., and Burgess, S. (2017). MendelianRandomization: an R package for performing Mendelian randomization analyses using summarized data. *Int. J. Epidemiol.* 46 (6), 1734–1739. doi:10.1093/ije/dyx034

Yin, P., Jiang, Y., Fang, X., Wang, D., Li, Y., Chen, M., et al. (2023). Cell-based therapies for degenerative musculoskeletal diseases. *Adv. Sci. (Weinheim, Baden-Wurttemberg, Ger.* 10, e2207050. doi:10.1002/advs.202207050

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 11 January 2024
DOI 10.3389/fgene.2023.1321484

OPEN ACCESS

EDITED BY
Ziheng Wang,
University of Macau, China

REVIEWED BY
Jiahao Cai,
Guangzhou Medical University, China
Xi Wu,
Shanghai Jiao Tong University, China
Vincent Salvatore Gallicchio,
Clemson University, United States

*CORRESPONDENCE
Li Gao,
gaoli0519@126.com

RECEIVED 14 October 2023
ACCEPTED 26 December 2023
PUBLISHED 11 January 2024

CITATION
Wu K, Sun Q, Liu D, Lu J, Wen D, Zang X and Gao L (2024), Genetically predicted circulating levels of cytokines and the risk of oral cavity and pharyngeal cancer: a bidirectional mendelian-randomization study.
Front. Genet. 14:1321484.
doi: 10.3389/fgene.2023.1321484

Genetically predicted circulating levels of cytokines and the risk of oral cavity and pharyngeal cancer: a bidirectional mendelian-randomization study

Kehan Wu[1], Qianhui Sun[2], Dongxu Liu[1], Jiayi Lu[1], Deyu Wen[1], Xiyan Zang[1] and Li Gao[1]*

[1]Department of Oral and Maxillofacial Surgery, The Second Affiliated Hospital of Harbin Medical University, Harbin, Heilongjiang, China, [2]Department of Cardiology, The Second Affiliated Hospital of Harbin Medical University, Harbin, Heilongjiang, China

Background: Epidemiological research has established associations between various inflammatory cytokines and the occurrence of oral cancer and oropharyngeal cancer (OCPC). We performed a Mendelian randomization (MR) analysis to systematically investigate the causal relationship between inflammatory cytokines and OCPC.

Methods: We performed a bidirectional two-sample MR analysis using OCPC from 12 studies (6,034 cases and 6,585 controls) and genome-wide association study (GWAS) results for 41 serum cytokines from 8,293 Finns, respectively. Inverse variance weighting was used as the primary MR method and four additional MR methods (MR Egger, Weighted median, Simple mode, Weighted mode) were used to examine genetic associations between inflammatory traits and OCPC, and Cochran's Q test, MR-Egger intercept, leave-one-out analysis, funnel plot, and multivariate MR (MVMR) analysis were used to assess the MR results.

Results: The results suggested a potential association between high gene expression of Macrophage inflammatory protein-1α (MIP1α/CCL3) and an increased risk of OCPC (Odds Ratio (OR): 1.71, 95% Confidence Interval (CI): 1.09–2.68, p = 0.019). Increasing the expression levels of the interleukin-7 (IL-7) gene by 1 standard deviation reduced the risk of OCPC (OR: 0.64, 95%CI: 0.48–0.86, p = 0.003). In addition, multivariate Mendelian randomization analysis also showed the same results (MIP1α/CCL3, OR: 1.002, 95% CI: 0.919–1.092, p = 0.044; IL-7, OR: 0.997, 95% CI: 0.994–0.999, p = 0.011). Conversely, there was a positive correlation between genetic susceptibility to OCPC and an increase in Interleukin-4 (IL-4) (OR: 1.04, 95%CI: 1.00–1.08, p = 0.027).

Conclusion: Our study systematically assessed the association between inflammatory cytokines and the risk of OCPC. We identified two upstream

regulatory factors (IL-7 and CCL3) and one downstream effector factor (IL-4) that were associated with OCPC, offering potential avenues for the development of novel treatments.

KEYWORDS

oral cavity and pharyngeal cancer, inflammation factors, cytokines, mendelian randomization, HNSCC, genome-wide association study

1 Introduction

Based on the latest estimates from Global Cancer Statistics 2020, Head and Neck Squamous Cell Carcinoma (HNSCC) ranks as the seventh most prevalent cancer worldwide, with an annual incidence of over 890,000 new cases and a mortality of over 450,000 (Sung et al., 2021). Among HNSCC cases, Oral Squamous Cell Carcinoma (OSCC) and Oropharyngeal Squamous Cell Carcinoma (OPSCC) are predominant, contributing to a global incidence of over 260,000 cases and over 128,000 deaths, respectively (Siegel et al., 2020). Despite some advances in the treatment of oral and oropharyngeal cancer (OCPC), the 10-year survival rate remains low at below 60%. Even if the treatment is successful, patients may still experience severe functional impairments, including compromised abilities in feeding, swallowing, and speech. Additionally, the recurrence rate remains high (Sung et al., 2021).

Earlier investigations in preclinical settings have indicated that inflammatory cytokines, such as TNF-α, IL-1β, and IL-6, promote the growth, invasion, and spread of cancer cells. Additionally, the transcription factors associated with these cytokines such as NF-kB and STAT3 show increased expression in most cancer types (Voronov et al., 2003; Pikarsky et al., 2004; Bierie and Moses, 2006; Luo et al., 2007). For example, inhibiting the activity or expression of IL-1β can prevent the occurrence of oral cancer by regulating specific key node genes in the tumor microenvironment (TME) (Wu et al., 2016). Moreover, mounting evidence indicates an increased risk of OCPC in the presence of inflammation, and inflammation commonly accompanies the development of OCPC (Leon et al., 2015). These observations suggest that the pharmacological targeting of additional inflammation biomarkers identified in the epidemiological literature through observational studies could offer a potentially effective approach for treating OCPC (Todoric et al., 2016). Nevertheless, the investigations conducted thus far have primarily concentrated on a limited range of inflammatory elements and have failed to acknowledge the impact of additional factors on the changes in inflammation levels. Hence, it is crucial to ascertain whether the changes in inflammatory factors cause the onset of tumors or if the tumors themselves modify the microenvironment, leading to differences in inflammatory factors. Due to the limited comprehension of the cause of OCPC, investigating the exact characteristics of the connections between inflammatory factors and OCPC has considerable clinical significance.

The dynamic nature of the inflammatory response suggests that a specific time point's measurement, whether high or low, might not precisely represent the overall trend of inflammatory factor variations. Epidemiological, genetic, and biological investigations have confirmed the link between inflammatory factors and OCPC. However, the outcomes derived from these investigations may be distorted by unforeseen confounding variables or reverse causal associations, thus complicating the establishment of clear causal relationships.

Observational studies may hinder an exhaustive comprehension of the connection between OCPC and inflammation due to their intrinsic research limitations. Thus, a detailed understanding of the role of circulating cytokines and their association with OCPC risk may aid in the development of prevention, prediction, and treatment strategies. Researchers may achieve a more comprehensive understanding of the connection between inflammation and OCPC by addressing these limitations.

Mendelian randomization (MR) is a commonly used tool in genetic epidemiology (10), which utilizes instrumental variable IV) variation derived from non-experimental data to assess the causal impact of exposure (e.g., circulating cytokines) on an outcome (e.g., OCPC) (Lawlor et al., 2008). Given the random allocation of alleles during meiosis, MR can mitigate confounding variables and reverse causality, thereby offering more robust evidence for causal inferences (Burgess et al., 2015). The use of two-sample MR analysis allows researchers to evaluate the associations between the instrument-exposure and instrument-outcome in two distinct population samples, thereby improving the applicability and efficiency of testing (Hartwig et al., 2016). In this study, we performed an analysis of the genome-wide association study (GWAS) summary data of 41 inflammatory cytokines to identify relevant genetic variations. These variations were then further investigated in relation to OCPC. Specifically, we examined the correlations between these genetic variations and OCPC by reversing the exposure and outcome. Our research results not only provided substantial Supplementary Data for previous epidemiological investigations but also offered fresh perspectives on the development and prevention of OCPC.

2 Methods

2.1 Study design

This study used a bidirectional MR approach to assess the causal relationship between the circulating cytokines and OCPCs. The study's overall design is illustrated in Figure 1. To assess the causal relationship between circulating cytokines and OCPCs, MR analysis was performed to test the following three hypotheses: 1) Genetic instruments have a strong association with exposure; 2) Genetic instruments are not affected by any potential known confounding factors; 3) The association between the genetic instrument and the outcome is solely influenced by the exposure (Smith and Ebrahim, 2003). Simultaneously, the reverse MR method is used to examine potential reverse causal effects. The data used in this study were

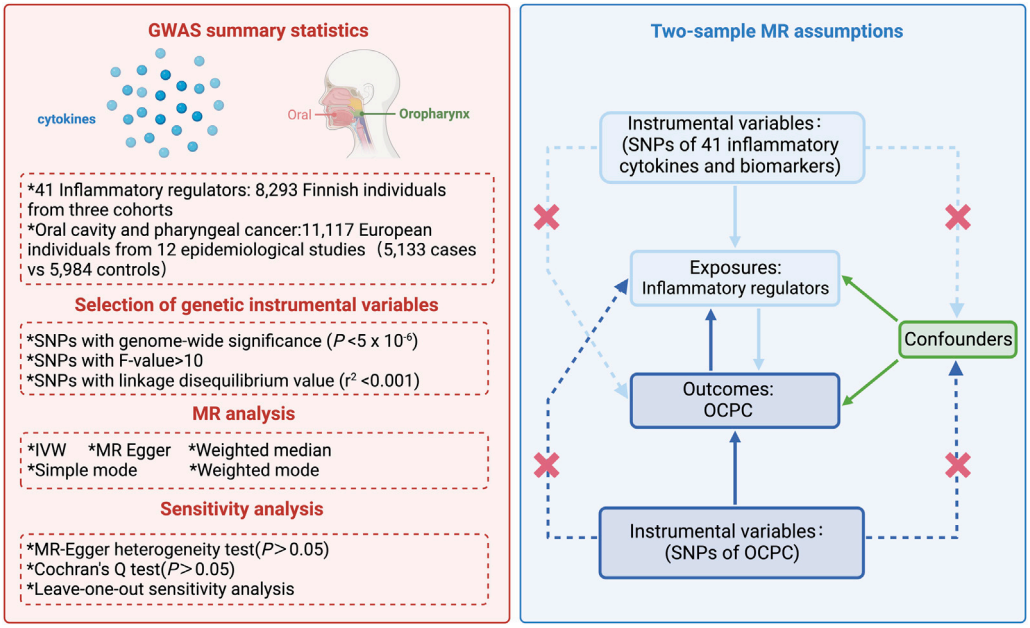


FIGURE 1
Illustration of the research design in the bidirectional Mendelian randomization (MR) analysis. Important instrumental variables, including 41 inflammatory cytokines and OCPC, were selected to explore the bidirectional causal relationships. This method is used to detect correlations (solid line) and violations of the Mendelian randomization assumption (dashed line). This figure was created using BioRender.com.

obtained from publicly available large-scale GWAS which can be accessed through the GWAS Catalog service (https://www.ebi.ac.uk/gwas/home). The studies included in the initial GWAS had obtained approval from their respective institutional review boards.

2.2 Genetic instrumental variables for inflammatory factors

The research used an extensive meta-analysis of GWAS on the circulating levels of 41 cytokines. This analysis combined information from three separate cohort studies: FINRISK 1997, FINRISK 2002, and the Young Finns Study on Cardiovascular Risk (YFS). This study included a total of 8,293 participants. The genomic and cytokine data were contributed by 4,608 individuals from FINRISK 1997, while another 1,705 participants from FINRISK2002 also provided their cytokine data. Cytokine quantification was performed by analyzing the EDTA-treated plasma of FINRISK 1997, the heparinized plasma of FINRISK 2002, and the serum samples from YFS. To control for potential confounding factors, such as age, sex, body mass index, and genetic variations, genetic associations were properly adjusted. Specifically, the top 10 genetic principal components were included using genomic control correction. Bio-Rad's pre-mixed Bio-Plex Pro Human Cytokine 27-plex Assay and 21-plex Assay were used to measure a total of 41 cytokines in YFS and FINRISK 2002, respectively. The measurements were performed by the Bio-Plex 200 reader equipped with Bio-Plex 6.0 software. However, it is important to note that, out of the 41 cytokines, 7 had to be excluded due to the presence of missing values exceeding 90% (Ahola-Olli et al., 2017; Kalaoja et al., 2021). A detailed overview of the GWAS data for the cytokines used in the MR analysis is presented in Table 1.

2.3 Genetic instrumental variables for OCPC

Using data from the most extensive GWAS to date, we investigated the impact of inflammatory elements on the vulnerability to OCPC. By extracting specific variations in single-nucleotide polymorphisms (SNPs) linked to exposure, we assessed a total of 6,034 cases and 6,585 controls enrolled in 12 different epidemiological studies (Lesseur et al., 2016). The Genetic Associations and Mechanisms of Oncology (GAME-ON) network is responsible for conducting these studies. Every participant has provided informed consent, and the relevant institutional review boards have approved this study. Furthermore, the study includes data from the European Prospective Investigation into Cancer and Nutrition (EPIC) study, and the Health and Nutrition 5,000 (HN5000) study, which specifically examines the health and lifestyle of 5,000 United Kingdom participants. Extensive information about the study, as well as the genotyping and imputation methods used, has been previously documented (Lesseur et al., 2016). The study sample comprised individuals from Europe (45.3%), North America (43.9%), and South America (10.8%). The study encompasses diverse types of cancer identified through the specific International Classification of Diseases, 10th Revision (ICD-10) codes, such as oral cancer (C02.0–C02.9, C03.0–C03.9, C04.0–C04.9, C05.0–C06.9) and oropharyngeal cancer (C01.9, C02.4, C09.0–C10.9). Exclusion criteria led to the removal of 954 individuals with hypopharyngeal cancer, unidentifiable codes, or overlapping

TABLE 1 Sample size for each cytokine analyzed in this study acquired from the GWAS.

Cytokines		Abbreviation	Sample size	Number
Chemokines	Cutaneous T cell attracting (CCL27)	CTACK	3,631	GCST004420
	Eotaxin (CCL11)	Eotaxin	8,153	GCST004460
	Growth regulated oncogene-α (CXCL1)	GROα	3,505	GCST004457
	Interferon gamma-induced protein 10 (CXCL10)	IP10	3,685	GCST004440
	Monocyte chemotactic protein-1 (CCL2)	MCP1	8,293	GCST004438
	Monocyte specific chemokine 3 (CCL7)	MCP3	843	GCST004437
	Monokine induced by interferon-gamma (CXCL9)	MIG	3,685	GCST004435
	Macrophage inflammatory protein-1α (CCL3)	MIP1α	3,522	GCST004434
	Macrophage inflammatory protein-1β (CCL4)	MIP1β	8,243	GCST004433
	Regulated on Activation, Normal T Cell Expressed and Secreted (CCL5)	RANTES	3,421	GCST004431
	Stromal cell-derived factor-1 alpha (CXCL12)	SDF1α	5,998	GCST004427
Growth factors	Beta nerve growth factor	βNGF	3,531	GCST004421
	Basic fibroblast growth factor	bFGF	7,565	GCST004459
	Granulocyte colony-stimulating factor	GCSF	7,904	GCST004458
	Hepatocyte growth factor	HGF	8,292	GCST004449
	Macrophage colony-stimulating factor	MCSF	840	GCST004436
	Platelet-derived growth factor BB	PDGFbb	8,293	GCST004432
	Stem cell factor	SCF	8,290	GCST004429
	Stem cell growth factor beta	SCGFβ	3,682	GCST004428
	Vascular endothelial growth factor	VEGF	7,118	GCST004422
Interleukins	Interleukin-10	IL-10	7,681	GCST004444
	Interleukin-12p70	IL-12p70	8,270	GCST004439
	Interleukin-13	IL-13	3,557	GCST004443
	Interleukin-16	IL-16	3,483	GCST004430
	Interleukin-17	IL-17	7,760	GCST004442
	Interleukin-18	IL-18	3,636	GCST004441
	Interleukin-1 receptor antagonist	IL1ra	3,638	GCST004447
	Interleukin-1-beta	IL-1β	3,309	GCST004448
	Interleukin-2	IL-2	3,475	GCST004455
	Interleukin-2 receptor, alpha subunit	IL2rα	3,677	GCST004454
	Interleukin-4	IL-4	8,124	GCST004453
	Interleukin-5	IL-5	3,364	GCST004452
	Interleukin-6	IL-6	8,189	GCST004446
	Interleukin-7	IL-7	3,409	GCST004451
	Interleukin-8 (CXCL8)	IL-8	3,526	GCST004445
	Interleukin-9	IL-9	3,634	GCST004450
Others	Interferon-gamma	IFN-γ	7,701	GCST004456
	Macrophage migration inhibitory factor (glycosylation-inhibiting factor)	MIF	3,494	GCST004423

(Continued on following page)

TABLE 1 (*Continued*) Sample size for each cytokine analyzed in this study acquired from the GWAS.

Cytokines		Abbreviation	Sample size	Number
	Tumor necrosis factor-alpha	TNFα	3,454	GCST004426
	Tumor necrosis factor-beta	TNFβ	1,559	GCST004425
	TNF-related apoptosis inducing ligand	TRAIL	8,186	GCST004424

cancers. To mitigate the effects of heterogeneity in these regions, GWAS studies were conducted exclusively on individuals with European ancestry, totaling 11,117 participants. Among these, there were 5,133 cases, comprising 2,700 cases of oral cancer and 2,433 cases of oropharyngeal cancer. In addition, there were 5,984 control cases (Supplementary Table S1).

2.4 Assessing common risk factors for OCPC

The cytokines-OCPC pathway is a potential target for cancer prevention and treatment. To identify the possible mediators of this pathway, we applied MR methods to examine the causal links between cytokines and common risk factors for OCPC. We used an existing database to investigate the following risk factors that have been widely recognized as associated with OCPC: smoking, drinking, body mass index, type 2 diabetes, hypertension, and HPV16/18 infection (Argiris et al., 2008; Leemans et al., 2018). Supplementary Table S2 lists the GWAS summary data for the above risk factors.

2.5 Selection of genetic instrumental variables

To ensure that the conclusion regarding the causal relationship between cytokines and the risk of OCPC is authentic and accurate, we implement various quality control measures in the selection of optimal genetic instrumental variables. Our approach involves the selection of SNPs that are closely associated with inflammatory factors and demonstrate genome-wide significance with a p-value lower than 5×10^{-8}. By focusing on these specific SNPs, we establish robust instrumental variables for our study (Burgess et al., 2011). We proceeded to eliminate linkage disequilibrium (LD). Our criteria for removal included an r^2 value below 0.001 and a distance of 5,000 kb. SNPs exceeding the r^2 threshold of 0.001, which includes the most significant SNP within a 5,000 kb range, were excluded. After aligning the chosen SNPs with the outcome data, we found that only ten out of the 41 systemic inflammation factors available exhibited two or more independent SNPs at a significance level of p-value less than 5×10^{-8}. Additionally, it was found that nine of them displayed three or more independent SNPs. Since inflammatory cytokines are a class of cytokines with multiple functions and interactions, and they may involve multiple genes and SNPs, a less stringent criterion is needed to capture their genetic variation. When selecting instrumental variables, we established A less stringent p-value threshold (5×10^{-6}) to capture more SNPs of inflammatory cytokines (Li et al., 2021; Pan et al., 2023). If there is an insufficient number of exposure-associated SNPs detected in the GWAS findings, proxy SNPs exhibiting high LD with a correlation coefficient ($r^2 > 0.90$) will be employed as alternative substitutes. These proxy SNPs can be accessed through LDlink (https://ldlink.nci.nih.gov/) as a resource (Machiela and Chanock, 2015). To exclude all SNPs associated with exposure and avoid potential pleiotropic effects, we performed a comprehensive investigation using the PhenoScanner V2 tool (http://www.phenoscanner.medschl.cam.ac.uk/) (Staley et al., 2016). Through the aforementioned steps, we acquired 41 cytokines associated with inflammation. As an instrumental variable with a significance threshold of $F < 10$ is considered weak, we will exclude it from our study. Detailed information on the identified SNPs is listed in Supplementary Tables S3 and S4. Furthermore, to uphold the fundamentals of MR, we shall examine the desired SNPs to exclude any that exhibit associations with the resulting outcomes.

2.6 MR statistical analysis

To explore the causal relationship between inflammatory regulators and OCPC, we utilized GWAS data and employed two-sample MR and multivariate MR methods. Statistical analysis was conducted using R software (v4.1.3) and the "MendelianRandomization", "MVMR" and "MRPRESSO" software packages. Multivariable MR (MVMR) is mainly used to evaluate the impact of multiple potential exposures on overall outcomes and identify potential risk factors (Tian and Burgess, 2023). The relationship between inflammatory factors and OCPC was examined using the inverse variance weighted (IVW) method. The instrumental variable method was used to estimate the mean effects of SNPs by regressing SNP-inflammatory factors on SNP-OCPC. Additionally, we utilized the weighted median estimator (WME) and the MR-Egger regression. WME is a statistical method that applies weights to the empirical distribution function of ratio estimates for SNPs within the study range, minimizing biases in estimating causal effects. MR-Egger regression employs weighted linear regression to estimate the effect of SNP-OCPC, considering SNP-inflammatory factors, and provides an evaluation of causal effects, even in the presence of invalid instruments (Bowden et al., 2015). If the directions of the β-values of other methods are the same, it may be interpreted as a positive result (Chen et al., 2020). Moreover, the "Leave-one-out" strategy was employed to visually illustrate whether a single SNP substantially influences the causal association. To ensure the dependability of the MR findings, we performed diverse evaluations of diversity and sensitivity. We employed Cochran's Q test to assess heterogeneity among SNPs. If no indications of heterogeneity were observed, we utilized a fixed-effect model; otherwise, a random-effects model was implemented. The Egger-intercept test was carried out for horizontal pleiotropy examination (Verbanck et al., 2018; Song et al., 2022). The results of

10.3389/fgene.2023.1321484

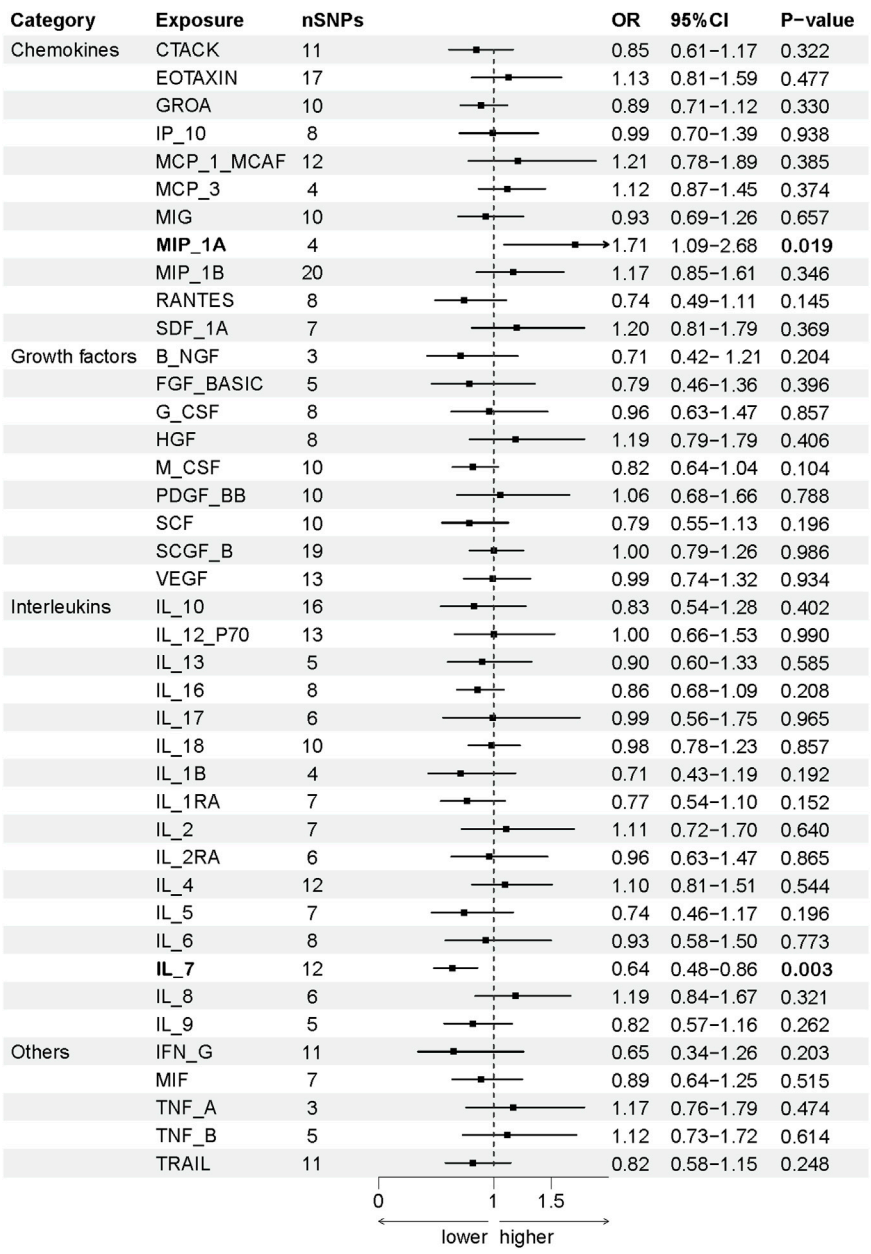


FIGURE 2
Using SNPs with a significance level of $p < 5 \times 10^{-6}$, we predicted the potential impact of inflammation regulatory factors across the genome on OCPC. To establish a causal link between circulating cytokine levels and OCPC, we conducted a two-sample Mendelian randomization (MR) analysis employing the IVW method. By estimating the odds ratio of OCPC for every 1-SD rise in predicted circulating cytokine levels, as determined by genetic prediction, we derived a 95% confidence interval (CI) value. This approach primarily determines the causal relationship between OCPC and the levels of circulating cytokines. A detailed overview of cytokines is provided in Table 1.

the study were presented using odds ratios (ORs) and their corresponding 95% confidence intervals (CIs). Any results with a *p*-value less than 0.05 were considered statistically significant. We applied the FDR method to correct for multiple testing, which is a method implemented by the *q*-value program. It is a commonly used method for multiple testing correction, which can control the proportion of false rejections of the null hypothesis among multiple hypothesis tests. We chose a *q*-value of less than 0.1 as the significance level, which is a reasonable choice because it can reduce the false positive rate while maintaining the statistical power. When the *p*-value is less than 0.05 but the *q*-value is greater than or equal to 0.1, we consider it a suggestive association result (Storey and Tibshirani, 2003; Li et al., 2022). The analysis of the MR study adhered to the guidelines set forth by the Strengthening the Reporting of Observational Studies in Epidemiology using Mendelian Randomization (STROBE-MR) statement, which emphasizes the scientific rigor and reporting standards for observational studies in epidemiology (Skrivankova et al., 2021).

TABLE 2 Our research examined whether there is a causal relationship between the levels of MIP1α/CCL3 and IL-7 in the circulatory system and OCPC. To achieve this, we employed genetic prediction techniques. To assess the variation in estimates of individual SNP effects, we used Cochran's Q test, and we employed the MR-Egger intercept test to evaluate horizontal pleiotropy.

Cytokines	Methods	MR results					Heterogeneity test			Horizontal pleiotropy test		
							Cochran's Q test			MR-egger intercept test		
		SNPs	β	SE	P	OR (95% CI)	Q	df	P	Intercept	SE	P
MIP1α/CCL3	IVW	4	0.537	0.229	0.019	1.711 (1.092–2.680)	2.188	3	0.534	0.071	0.118	0.607
	MR Egger	4	0.104	0.753	0.903	1.111 (0.254–4.852)	1.824	2	0.402			
	Weighted median	4	0.459	0.268	0.086	1.583 (0.937–2.674)						
	Simple mode	4	0.412	0.416	0.395	1.510 (0.668–3.412)						
	Weighted mode	4	0.369	0.378	0.390	1.446 (0.703–2.975)						
IL-7	IVW	12	−0.448	0.151	0.003	0.639 (0.475–0.859)	18.890	11	0.063	−0.055	0.073	0.468
	MR Egger	12	−0.191	0.374	0.621	0.826 (0.397–1.720)	17.875	10	0.057			
	Weighted median	12	−0.469	0.168	0.005	0.626 (0.450–0.870)						
	Simple mode	12	−0.394	0.319	0.242	0.674 (0.361–1.260)						
	Weighted mode	12	−0.586	0.246	0.036	0.557 (0.344–0.902)						

Abbreviations: SNP, single nucleotide polymorphism; β, effect size of SNP on exposure; SE, standard error; OR, odds ratio; CI, confidence interval; df, degrees of freedom; IVW, inverse variance weighted.

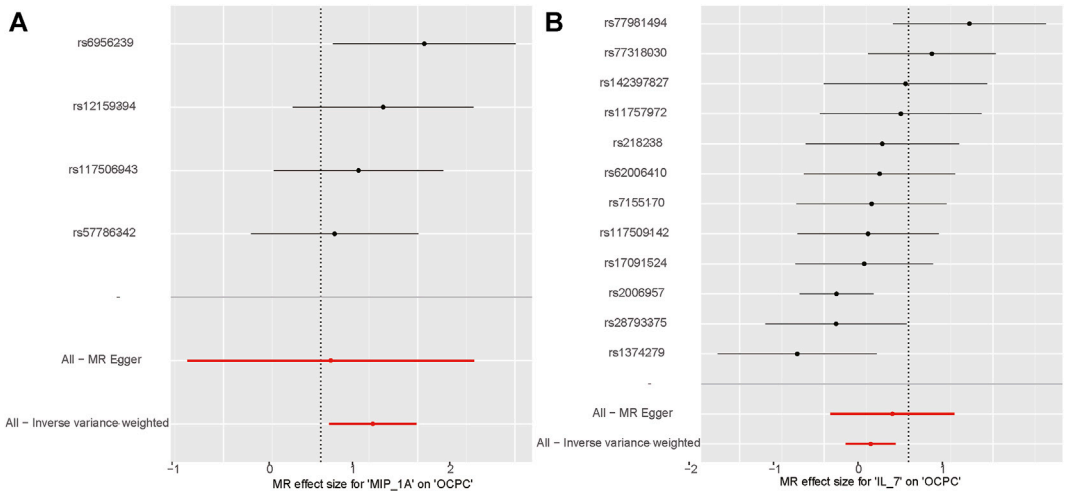


FIGURE 3
MIP1α/CCL3 and IL-7 related SNPs are depicted in the forest plot for OCPC risk. **(A)** Forest plot for MIP1α/CCL3 exposure. **(B)** Forest plot for IL-7 exposure. The *X*-axis represents the effect estimates of MIP1α/CCL3 and IL-7 on OCPC analyzed using MR. The *Y*-axis represents the MR effect values for each SNP site.

3 Results

3.1 Causal impact of systemic inflammation factors on the risk of OCPC

The association between systemic inflammation factors and OCPC was revealed through gene prediction, supported by the following findings (Figure 2). The IVW method uncovered a significant surge in OCPC risk, linked to higher levels of Macrophage inflammatory protein-1α (MIP1α/CCL3) (OR: 1.71, 95% CI: 1.09–2.68, $p = 0.019$). Both the MR-Egger heterogeneity test and Cochran's Q test failed to identify any signs of heterogeneity, indicating an absence of variation ($p > 0.05$). Additionally, by utilizing the IVW method, increased Interleukin-7 (IL-7) levels were associated with a decreased likelihood of OCPC(OR: 0.64, 95%CI: 0.48–0.86, $p = 0.003$). No evidence of heterogeneity or horizontal pleiotropy was discovered ($p > 0.05$) (Table 2). Figures 3, 4 show forest plots and scatter plots illustrating the genetic

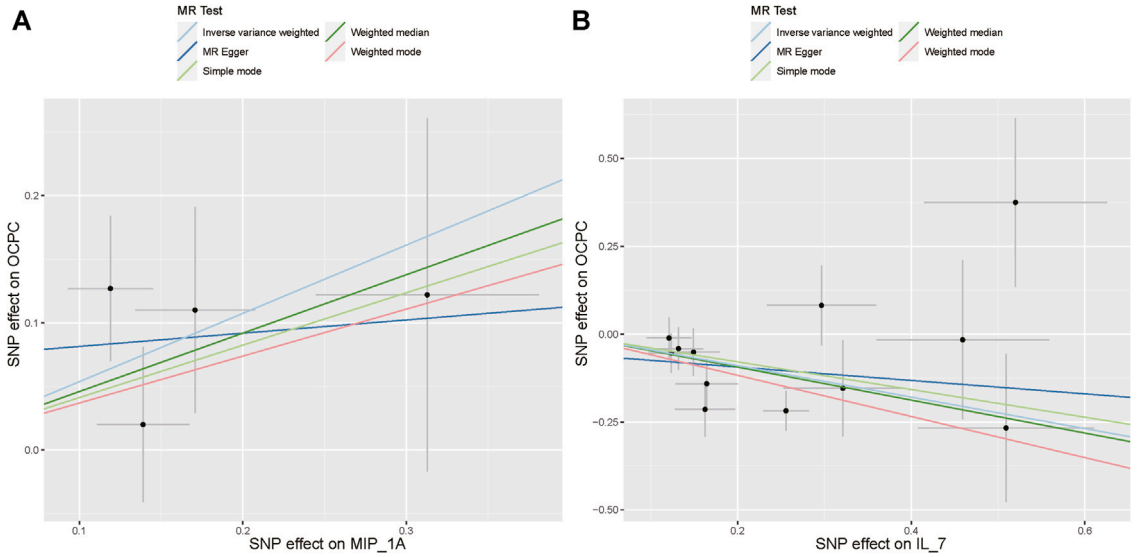


FIGURE 4
Using various MR methods, the scatter plots demonstrate the genetic correlation between SNP and OCPC for MIP1α/CCL3 and IL-7. **(A)** The exposure of MIP1α/CCL3 is illustrated in the scatter plot. **(B)** The exposure of IL-7 is illustrated in the scatter plot. Each scatter plot illustrates the associations between the alleles and the risk of the outcome, plotted against the association with one standard deviation of exposure. The effects are represented by the gray error bars, which indicate the 95% confidence intervals. The IVW, MR Egger, Weighted median, Simple mode, and Weighted mode were employed for the analysis. The estimated MR effect of each method can be determined by the slope of each line.

association of MIP1α/CCL3 and IL-7 SNPs with OCPC. The funnel plot demonstrates overall symmetry, indicating little evidence of heterogeneity (Supplementary Figure S1). A sensitivity analysis, using a leave-one-out approach, provided evidence of the lack of a single SNP with substantial impact on the overall effect. This verifies the dependability and consistency of the estimation of the causal effect (Supplementary Figure S3).

Supplementary Table S5 presents the MR outcomes concerning the prediction of genetic susceptibility to systemic inflammation factors and OCPC risk assessment. Furthermore, it contains summaries of the analyses conducted to assess heterogeneity, pleiotropy, and sensitivity.

3.2 Causal impact of OCPC on the risk of systemic inflammation factors

To evaluate the reverse causal effects, we conducted a study in which we identified 12 SNPs that have a significant and independent association with OCPC. The association reached a significance level of $p < 5 \times 10{-}^{6}$. We used a varying number of SNPs for each cytokine due to the lack of certain SNPs to be used universally. Detailed information about the number of SNPs used for each cytokine can be found in Supplementary Table S6. A suggestive association was observed between genetic susceptibility to OCPC and increased levels of Interleukin-4 (IL-4) based on the IVW method (OR: 1.04, 95%CI: 1.00–1.08, $p = 0.03$). No other significant associations were found, except for IL-4 (Figure 5). Furthermore, significant results of the MR and sensitivity analysis of OCPC and cytokines are shown in Table 3. Figures 6, 7 present forest plots and scatter plots illustrating the causal impact of OCPC-associated SNPs on IL-4.

3.3 Other factors and MVMR

In univariate MR analysis, while examining the causal relationship between cytokines and OCPC, we also found that after removing SNPs associated with confounding factors, there were still some SNPs that were not only associated with OCPC but also strongly associated with other risk factors. Therefore, assessing the causal relationship between cytokines and common risk factors for OCPC is beneficial to identifying interfering factors that may mediate the association between cytokines and OCPC. We combined all SNPs related to MIP1α/CCL3 and IL-7 as cytokine instrumental variables to find the greatest genetic confounding. Preliminary results show a potential causal relationship between cytokines and common OCPC risk factors (including smoking, drinking, body mass index, and hypertension) (Figure 8, Supplementary Table S7). To control for pleiotropic pathways, when we further applied the MVMR model, the results showed that cytokines still have a potential causal effect on OCPC (MIP1α/CCL3, OR: 1.002, 95% CI: 0.919–1.092, $p = 0.044$; IL-7, OR: 0.997, 95% CI: 0.994–0.999, $p = 0.011$) (Table 4).

4 Discussion

Numerous observational studies have found that the levels of circulating cytokines are related to the occurrence of OCPC. However, observational studies may be subject to bias due to inadequate sample size and confounding factors, resulting in skewed results. Moreover, there is insufficient genetic evidence to support this relationship in this field. Therefore, we utilized the latest GWAS data and adopted a systematic analysis approach to explore the causal effects of 41 different cytokines on OCPC. Unlike

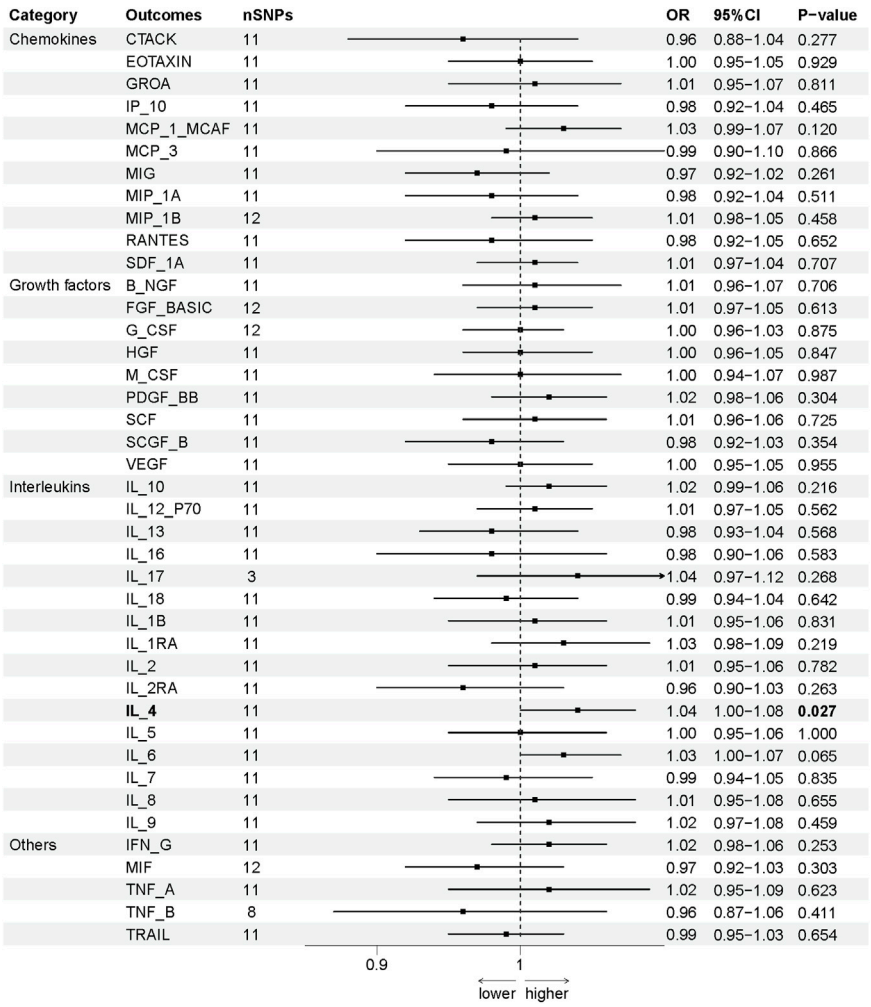


FIGURE 5
The influence of OCPC gene estimation on systemic inflammatory regulators (SNPs reaching $p < 5 \times 10^{-6}$) can lead to causal consequences. The connection between OCPC and levels of circulating cytokines can be predominantly ascertained through a two-sample Mendelian randomization (MR) analysis using the IVW technique. The estimated range of the 95% confidence interval (CI) reflects the odds ratio of circulating cytokines for each 1-SD rise in genetically anticipated OCPC levels. A detailed overview of cytokines can be found in Table 1.

TABLE 3 We evaluated the effects of genetic prediction (OCPC) on the levels of circulating IL-4. To assess heterogeneity between individual SNP estimates, we utilized Cochran's Q test. The MR-Egger intercept test was employed to investigate horizontal pleiotropy.

Cytokines	Methods	MR results					Heterogeneity test			Horizontal pleiotropy test		
							Cochran's Q test			MR-egger intercept test		
		SNPs	β	SE	*P*	OR (95% CI)	Q	df	*P*	Intercept	SE	*P*
IL-4	IVW	11	0.041	0.018	0.027	1.042 (1.005–1.080)	9.090	10	0.524	0.017	0.017	0.332
	MR Egger	11	0.000	0.044	0.996	1.000 (0.918–1.099)	8.040	9	0.530			
	Weighted median	11	0.008	0.026	0.741	1.009 (0.959–1.060)						
	Simple mode	11	−0.010	0.048	0.842	0.990 (0.900–1.099)						
	Weighted mode	11	−0.007	0.045	0.879	0.993 (0.909–1.099)						

Abbreviations: SNP, single nucleotide polymorphism; β, effect size of SNP on exposure; SE, standard error; OR, odds ratio; CI, confidence interval; df, degrees of freedom; IVW, inverse variance weighted.

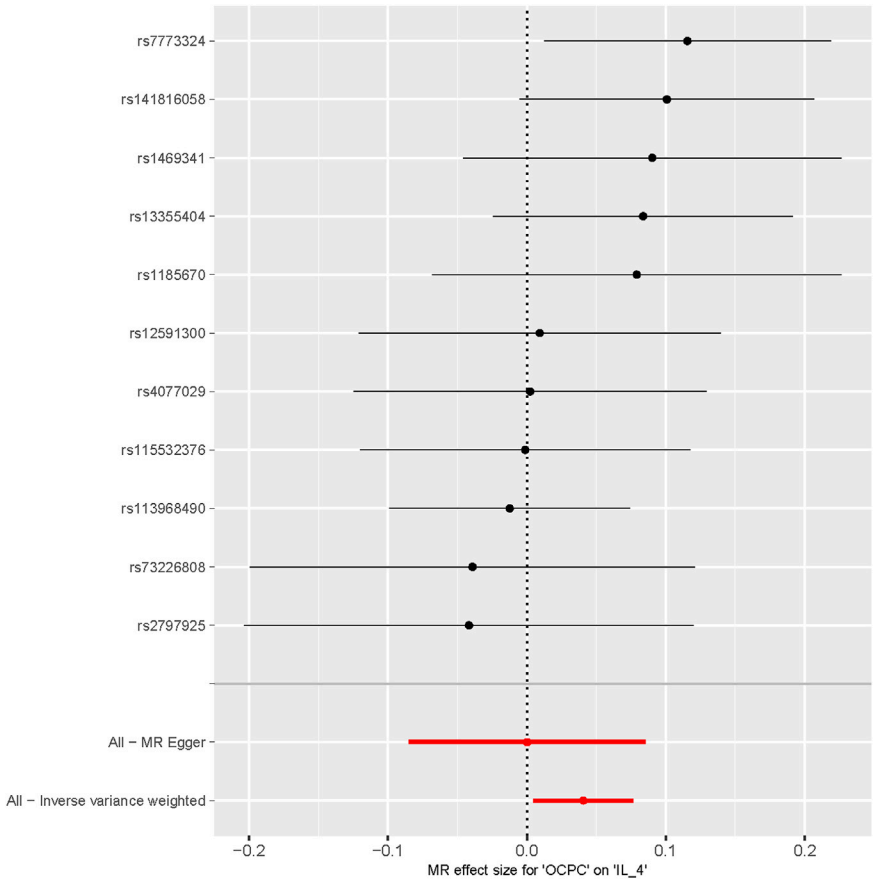


FIGURE 6
Forest plot of OCPC-related SNPs and IL-4 risk (The effect estimates of OCPC on IL-4 analyzed through MR are represented on the *X*-axis. Each SNP site's MR effect values are represented on the *Y*-axis).

traditional MR studies, we not only focused on single cytokines but also performed bidirectional MR analysis, which can simultaneously test whether two factors have a causal relationship and what the causal direction is. In this way, we can identify the cytokines before or after the disease pathway. In addition, we also used various MR methods to enhance the robustness of the study and minimize the interference of pleiotropic effects. The objective of this study was to systematically assess the potential causal relationship between 41 circulating cytokines and OCPC risk using bidirectional MR analysis.

The tumor microenvironment and the malignant properties of tumor cells are crucial factors in shaping the biological behavior of tumors, directly driving their growth, invasion, and metastasis (Hanahan and Coussens, 2012). Systemic cytokines are a group of molecules that play a broad role in controlling inflammation throughout the body. These cytokines maintain the balance between pro- and anti-inflammatory processes in the TME, thereby ensuring that the immune system functions effectively during infection, injury, or disease while avoiding excessive damage to body tissues (Zhang et al., 2020; Ozga et al., 2021; Bhat et al., 2022). Accumulating evidence indicates that chemokines may exert pro-tumor effects in different types of cancer (Soria and Ben-Baruch, 2008; Levina et al., 2009; Hwang et al., 2012). Our findings show a positive association between high levels of CCL3 and an elevated risk of OCPC. CCL3, a chemokine that belongs to the CC chemokine subfamily, is synthesized by monocytes/macrophages, lymphocytes, neutrophils, as well as various immune cells including eosinophils, mast cells, fibroblasts, and dendritic cells. It is also called as macrophage inflammatory protein-1α (MIP-1α) (Hanahan and Coussens, 2012). CCL3 plays a pivotal role in recruiting inflammatory cells under both homeostatic and pathological conditions. CCL3 may contribute to cancer progression by stimulating leukocyte accumulation, angiogenesis, and tumorigenesis. OCPC cells and tissues exhibit overexpression of CCL3, which correlates with poorer survival rates among OCPC patients (Silva et al., 2007; da Silva et al., 2017). CCL3 can stimulate cancer cell growth and migration (da Silva et al., 2017; Hsu et al., 2013), whereas blocking CCL3 can suppress tumor growth, and angiogenesis, and increase cell sensitivity to therapeutic drugs (Liao et al., 2016; Kim et al., 2017). These mechanisms could elucidate the role of CCL3 in driving the pathophysiology of OCPC.

Research has shown that inflammation in specific organs can affect the risk of cancer, and inflammatory factors also exist in the tumor microenvironment to alter the proliferation, survival, and metastasis of malignant cells. An association between OCPC and inflammatory cytokines, wound-healing genes, growth factors, and cell cycle genes has been reported (Perez-Sayans et al., 2009; Mantovani, 2010). Interestingly, our study suggests that an

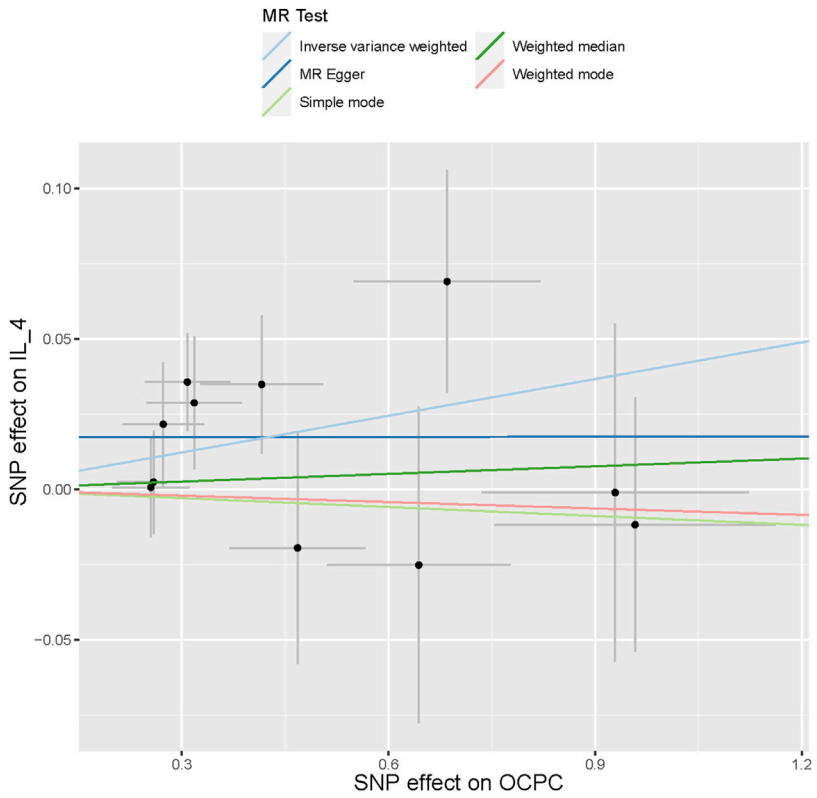


FIGURE 7
The scatter plot in this passage shows the genetic associations between SNPs related to OCPC and IL-4 using different MR methods. The gray error bars represent the 95% confidence intervals of the effects. These scatter plots represent the associations between each allele and the risk of the outcome, plotted against the association with one standard deviation of exposure. The analysis was conducted using IVW, MR Egger, Weighted median, Simple mode, and Weighted mode. The slope of each line represents the estimated MR effect of each method.

increase of 1 standard deviation (SD) in the predicted levels of the IL-7 gene is related to a decreased risk of OCPC. IL-7, an indispensable cytokine for the adaptive immune system, is primarily secreted by the bone marrow, thymus, and lymph nodes. IL-7 plays a crucial role in lymphocyte development and survival, contributing to the maintenance of immune self-stability in the body (Fry and Mackall, 2002; Jiang et al., 2005; Hong et al., 2012; Barata et al., 2019). This effect may be due to the potential role of active inflammatory responses in decreasing the incidence of OCPC. Due to the powerful biological effects of IL-7, specifically, its role in promoting T cell longevity, growth, replication, and preservation of memory, various scientific groups have used IL-7 as a molecular enhancer to improve the immune response generated by cancer vaccines and maintain long-lasting memory reactions against cancer (Zhao et al., 2022). In conclusion, our study suggests a possible relationship between IL-7 and inflammatory tumors in OCPC, which may have some clinical significance for exploring the pathogenesis of OCPC and reducing its incidence.

We also explored potential mediators in the cytokines-OCPC pathway. Previous MR studies have shown that MIP1α/CCL3 and IL-7 are associated with OCPC risk. In observational studies, smoking, drinking, body mass index, type 2 diabetes, hypertension, and HPV16/18 infection have been suggested as risk factors for the development of OCPC. To account for the above potential mediating factors, the MVMR model was further applied to examine the possibility of the observed confounders introducing horizontal pleiotropy. However, MVMR analysis showed that OCPC was no longer significantly associated with these factors at the conventional 5% level. Therefore, it seems unlikely that these factors play a substantial role in the pathway of inflammatory exposure. Notably, MIP1α/CCL3 and IL-7 still have estimated causal effects on OCPC even if mediating factors are excluded.

Bidirectional MR analysis during the OCPC stage suggested that OCPC is potentially associated with alterations in IL-4 levels in blood, despite the limited available evidence on their correlation. Hypotheses propose that elevated levels of IL-4 might signify the presence of analogous immune-suppressing mechanisms within the microenvironment of OCPC. These mechanisms potentially facilitate the expansion of tumors and evasion of immune vigilance. Previous studies have reported significant IL-4 expression in the tumor microenvironment (Setrerrahmane and Xu, 2017; Gao et al., 2021). Inhibiting it can alter inflammation and improve the tumor's response to immunotherapy (Ito et al., 2017). To sum up, the complex interactions between the immune system and the tumor are reflected by the varying IL-4 levels in the blood of patients with OCPC. These results highlight the importance of immune regulation and the imbalance of cytokines in the development of OCPC. Therefore, additional research is necessary to understand the exact mechanisms behind these changes in cytokines and how they impact the occurrence, advancement, and treatment strategies for OCPC.

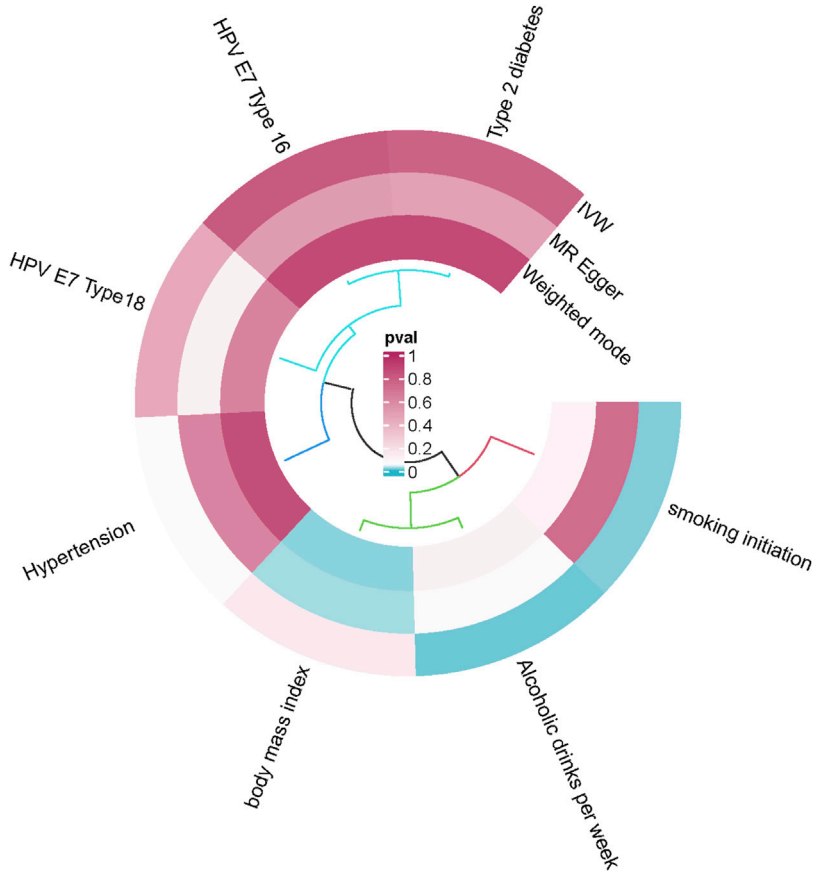


FIGURE 8
Causal effects of MIP1α/CCL3 and IL-7 on common risk factors for OCPC.

TABLE 4 MVMR results for OCPC.

Outcome	Exposure	SNPs, n	Or (95% CI)	*p*-value
OCPC	MVMR based on MIP1α/CCL3			
	MIP1α/CCL3	14	1.002 (0.919–1.092)	0.044
	Body mass index	455	1.000 (0.862–1.161)	0.076
	Smoking initiation	21	1.002 (0.575–1.182)	0.085
	Alcoholic drinks per week	6	1.002 (0.575–1.746)	0.283
	Hypertension	0	NA	NA
	MVMR based on IL-7			
	IL-7	43	0.997 (0.994–0.999)	0.011
	Body mass index	129	1.001 (0.783–1.279)	0.125
	Smoking initiation	14	1.001 (0.330–3.034)	0.566
	Alcoholic drinks per week	2	1.002 (0.254–3.949)	0.700
	Hypertension	0	NA	NA

Abbreviations: MVMR, multivariable Mendelian randomization; n, number; SNP, single-nucleotide polymorphism; OCPC, oral cavity and pharyngeal cancer; OR, odds ratio; CI, confidence interval.

This study is remarkable for its novel use of large-scale GWAS data to investigate multiple exposure factors related to OCPC risk, which enhances the stability and accuracy of estimating the effects. Second, the bidirectional MR design is employed to mitigate confounding factors and eliminate reverse causality. Third, the MVMR model was further applied to examine the possibility of introducing horizontal pleiotropy by common confounders of OCPC. Finally, this study comprehensively investigates the association between OCPC and 41 circulating cytokines, making it the most extensive MR study to date on this topic. They may provide oncologists and therapists with new perspectives and implications for designing more personalized and effective treatments for patients who may develop or suffer from OCPC. For example, host genotype could be exploited to enable more precise diagnosis and treatment by investigating inflammatory exposure factors in clinical patients. The use of monoclonal antibodies to reduce the concentration of certain inflammatory factors or exogenous supplementation of specific inflammatory factors can also be studied to prevent or treat OCPC. While the MR design offers advantages, this study also has limitations. First, although GWAS statistical data of European ancestry were used to mitigate population bias, the generalizability of our findings across different populations remains uncertain. Additional GWAS with larger sample sizes are necessary to validate and update the findings of this study. Moreover, the results of these studies may be confounded by various factors, including the production and interplay of cytokines. Despite conducting a MVMR analysis, it seems that the estimates of time-varying exposure may not accurately reflect the causal effects within a specific time period. When the exposure only affects the outcome at a few distinct time points and the risk factors in the MVMR analysis are the exposure values at these specific time points, it is possible to obtain a reliable estimation of the causal effect during these time points. However, if these time points are not correctly identified, estimates obtained from ambiguous models will be incorrect. This can mislead any inferences made about the magnitude, existence, or direction of the causal effects. Finally, it is important to note that all research findings require validation in clinical and basic research. As a result, caution should be exercised in interpreting potential causal relationships, and further investigation of potential physiopathological mechanisms is needed.

5 Conclusion

Our comprehensive MR analysis revealed potential causal relationships between 41 circulating cytokines and OCPC, providing new insights into their interactions. The following conclusions were drawn: MIP1α/CCL3 and IL-7 may be potential factors driving OCPC. And susceptibility to OCPC may also increase IL-4 levels in prognosis. However, due to the limitations of our study, including the relatively small number of OCPC cases and the use of only a single ancestry population, our findings should be validated in a larger cohort and the exact underlying biological mechanisms require further investigation.

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.

Ethics statement

Ethical approval was not required for the study involving humans in accordance with the local legislation and institutional requirements. Written informed consent to participate in this study was not required from the participants or the participants' legal guardians/next of kin in accordance with the national legislation and the institutional requirements.

Author contributions

KW: Conceptualization, Data curation, Methodology, Software, Validation, Visualization, Writing–original draft, Writing–review and editing. QS: Conceptualization, Methodology, Supervision, Validation, Visualization, Writing–original draft. DL: Conceptualization, Supervision, Validation, Writing–review and editing. JL: Formal Analysis, Software, Validation, Writing–review and editing. DW: Formal Analysis, Investigation, Validation, Writing–review and editing. XZ: Formal Analysis, Investigation, Supervision, Writing–review and editing. LG: Conceptualization, Project administration, Supervision, Writing–review and editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. Medical Wisdom Research Fund supported this research by the Heilongjiang Provincial Natural Science Foundation (LH2021H033).

Acknowledgments

The authors thank all investigators and participants contributing to the GWAS, YFS, and FINRISK studies.

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2023.1321484/full#supplementary-material

References

Ahola-Olli, A. V., Wurtz, P., Havulinna, A. S., Aalto, K., Pitkanen, N., Lehtimaki, T., et al. (2017). Genome-wide association study identifies 27 loci influencing concentrations of circulating cytokines and growth factors. *Am. J. Hum. Genet.* 100, 40–50. doi:10.1016/j.ajhg.2016.11.007

Argiris, A., Karamouzis, M. V., Raben, D., and Ferris, R. L. (2008). Head and neck cancer. *Lancet* 371, 1695–1709. doi:10.1016/S0140-6736(08)60728-X

Barata, J. T., Durum, S. K., and Seddon, B. (2019). Flip the coin: IL-7 and IL-7R in health and disease. *Nat. Immunol.* 20, 1584–1593. doi:10.1038/s41590-019-0479-x

Bhat, A. A., Nisar, S., Singh, M., Ashraf, B., Masoodi, T., Prasad, C. P., et al. (2022). Cytokine- and chemokine-induced inflammatory colorectal tumor microenvironment: emerging avenue for targeted therapy. *Cancer Commun. (Lond)* 42, 689–715. doi:10.1002/cac2.12295

Bierie, B., and Moses, H. L. (2006). TGF-beta and cancer. *Cytokine Growth Factor Rev.* 17, 29–40. doi:10.1016/j.cytogfr.2005.09.006

Bowden, J., Davey Smith, G., and Burgess, S. (2015). Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int. J. Epidemiol.* 44, 512–525. doi:10.1093/ije/dyv080

Burgess, S., Scott, R. A., Timpson, N. J., Davey Smith, G., Thompson, S. G., and Consortium, E.-I. (2015). Using published data in Mendelian randomization: a blueprint for efficient identification of causal risk factors. *Eur. J. Epidemiol.* 30, 543–552. doi:10.1007/s10654-015-0011-z

Burgess, S., Thompson, S. G., and Collaboration, C. C. G. (2011). Avoiding bias from weak instruments in Mendelian randomization studies. *Int. J. Epidemiol.* 40, 755–764. doi:10.1093/ije/dyr036

Chen, X., Kong, J., Diao, X., Cai, J., Zheng, J., Xie, W., et al. (2020). Depression and prostate cancer risk: a Mendelian randomization study. *Cancer Med.* 9, 9160–9167. doi:10.1002/cam4.3493

Da Silva, J. M., Moreira Dos Santos, T. P., Sobral, L. M., Queiroz-Junior, C. M., Rachid, M. A., Proudfoot, A. E. I., et al. (2017). Relevance of CCL3/CCR5 axis in oral carcinogenesis. *Oncotarget* 8, 51024–51036. doi:10.18632/oncotarget.16882

Fry, T. J., and Mackall, C. L. (2002). Interleukin-7: from bench to clinic. *Blood* 99, 3892–3904. doi:10.1182/blood.v99.11.3892

Gao, S., Hsu, T. W., and Li, M. O. (2021). Immunity beyond cancer cells: perspective from tumor tissue. *Trends Cancer* 7, 1010–1019. doi:10.1016/j.trecan.2021.06.007

Hanahan, D., and Coussens, L. M. (2012). Accessories to the crime: functions of cells recruited to the tumor microenvironment. *Cancer Cell* 21, 309–322. doi:10.1016/j.ccr.2012.02.022

Hartwig, F. P., Davies, N. M., Hemani, G., and Davey Smith, G. (2016). Two-sample Mendelian randomization: avoiding the downsides of a powerful, widely applicable but potentially fallible technique. *Int. J. Epidemiol.* 45, 1717–1726. doi:10.1093/ije/dyx028

Hong, C., Luckey, M. A., and Park, J. H. (2012). Intrathymic IL-7: the where, when, and why of IL-7 signaling during T cell development. *Semin. Immunol.* 24, 151–158. doi:10.1016/j.smim.2012.02.002

Hsu, C. J., Wu, M. H., Chen, C. Y., Tsai, C. H., Hsu, H. C., and Tang, C. H. (2013). AMP-activated protein kinase activation mediates CCL3-induced cell migration and matrix metalloproteinase-2 expression in human chondrosarcoma. *Cell Commun. Signal* 11, 68. doi:10.1186/1478-811X-11-68

Hwang, T. L., Lee, L. Y., Wang, C. C., Liang, Y., Huang, S. F., and Wu, C. M. (2012). CCL7 and CCL21 overexpression in gastric cancer is associated with lymph node metastasis and poor prognosis. *World J. Gastroenterol.* 18, 1249–1256. doi:10.3748/wjg.v18.i11.1249

Ito, S. E., Shirota, H., Kasahara, Y., Saijo, K., and Ishioka, C. (2017). IL-4 blockade alters the tumor microenvironment and augments the response to cancer immunotherapy in a mouse model. *Cancer Immunol. Immunother.* 66, 1485–1496. doi:10.1007/s00262-017-2043-6

Jiang, Q., Li, W. Q., Aiello, F. B., Mazzucchelli, R., Asefa, B., Khaled, A. R., et al. (2005). Cell biology of IL-7, a key lymphotrophin. *Cytokine Growth Factor Rev.* 16, 513–533. doi:10.1016/j.cytogfr.2005.05.004

Kalaoja, M., Corbin, L. J., Tan, V. Y., Ahola-Olli, A. V., Havulinna, A. S., Santalahti, K., et al. (2021). The role of inflammatory cytokines as intermediates in the pathway from increased adiposity to disease. *Obes. (Silver Spring)* 29, 428–437. doi:10.1002/oby.23060

Kim, J. H., Kim, W. S., Hong, J. Y., Ryu, K. J., Kim, S. J., and Park, C. (2017). Epstein-Barr virus EBNA2 directs doxorubicin resistance of B cell lymphoma through CCL3 and CCL4-mediated activation of NF-κB and Btk. *Oncotarget* 8, 5361–5370. doi:10.18632/oncotarget.14243

Lawlor, D. A., Harbord, R. M., Sterne, J. A., Timpson, N., and Davey Smith, G. (2008). Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. *Stat. Med.* 27, 1133–1163. doi:10.1002/sim.3034

Leemans, C. R., Snijders, P. J. F., and Brakenhoff, R. H. (2018). The molecular landscape of head and neck cancer. *Nat. Rev. Cancer* 18, 269–282. doi:10.1038/nrc.2018.11

Leon, X., Bothe, C., Garcia, J., Parreno, M., Alcolea, S., Quer, M., et al. (2015). Expression of IL-1α correlates with distant metastasis in patients with head and neck squamous cell carcinoma. *Oncotarget* 6, 37398–37409. doi:10.18632/oncotarget.6054

Lesseur, C., Diergaarde, B., Olshan, A. F., Wunsch-Filho, V., Ness, A. R., Liu, G., et al. (2016). Genome-wide association analyses identify new susceptibility loci for oral cavity and pharyngeal cancer. *Nat. Genet.* 48, 1544–1550. doi:10.1038/ng.3685

Levina, V., Nolen, B. M., Marrangoni, A. M., Cheng, P., Marks, J. R., Szczepanski, M. J., et al. (2009). Role of eotaxin-1 signaling in ovarian cancer. *Clin. Cancer Res.* 15, 2647–2656. doi:10.1158/1078-0432.CCR-08-2024

Liao, Y. Y., Tsai, H. C., Chou, P. Y., Wang, S. W., Chen, H. T., Lin, Y. M., et al. (2016). CCL3 promotes angiogenesis by dysregulation of miR-374b/VEGF-A axis in human osteosarcoma cells. *Oncotarget* 7, 4310–4325. doi:10.18632/oncotarget.6708

Li, P., Wang, H., Guo, L., Gou, X., Chen, G., Lin, D., et al. (2022). Association between gut microbiota and preeclampsia-eclampsia: a two-sample Mendelian randomization study. *BMC Med.* 20, 443. doi:10.1186/s12916-022-02657-x

Li, Y., Lu, J., Wang, J., Deng, P., Meng, C., and Tang, H. (2021). Inflammatory cytokines and risk of ischemic stroke: a mendelian randomization study. *Front. Pharmacol.* 12, 779899. doi:10.3389/fphar.2021.779899

Luo, J. L., Tan, W., Ricono, J. M., Korchynskyi, O., Zhang, M., Gonias, S. L., et al. (2007). Nuclear cytokine-activated IKKalpha controls prostate cancer metastasis by repressing Maspin. *Nature* 446, 690–694. doi:10.1038/nature05656

Machiela, M. J., and Chanock, S. J. (2015). LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. *Bioinformatics* 31, 3555–3557. doi:10.1093/bioinformatics/btv402

Mantovani, A. (2010). Molecular pathways linking inflammation and cancer. *Curr. Mol. Med.* 10, 369–373. doi:10.2174/156652410791316968

Ozga, A. J., Chow, M. T., and Luster, A. D. (2021). Chemokines and the immune response to cancer. *Immunity* 54, 859–874. doi:10.1016/j.immuni.2021.01.012

Pan, S., Wu, S., Wei, Y., Liu, J., Zhou, C., Chen, T., et al. (2023). Exploring the causal relationship between inflammatory cytokines and inflammatory arthritis: a Mendelian randomization study. *Cytokine* 173, 156446. doi:10.1016/j.cyto.2023.156446

Perez-Sayans, M., Somoza-Martin, J. M., Barros-Angueira, F., Reboiras-Lopez, M. D., Gandara Rey, J. M., and Garcia-Garcia, A. (2009). Genetic and molecular alterations associated with oral squamous cell cancer (Review). *Oncol. Rep.* 22, 1277–1282. doi:10.3892/or_00000565

Pikarsky, E., Porat, R. M., Stein, I., Abramovitch, R., Amit, S., Kasem, S., et al. (2004). NF-kappaB functions as a tumour promoter in inflammation-associated cancer. *Nature* 431, 461–466. doi:10.1038/nature02924

Setrerrahmane, S., and Xu, H. (2017). Tumor-related interleukins: old validated targets for new anti-cancer drug development. *Mol. Cancer* 16, 153. doi:10.1186/s12943-017-0721-9

Siegel, R. L., Miller, K. D., and Jemal, A. (2020). Cancer statistics, 2020. *CA Cancer J. Clin.* 70, 7–30. doi:10.3322/caac.21590

Silva, T. A., Ribeiro, F. L., Oliveira-Neto, H. H., Watanabe, S., Alencar Rde, C., Fukada, S. Y., et al. (2007). Dual role of CCL3/CCR1 in oral squamous cell carcinoma: implications in tumor metastasis and local host defense. *Oncol. Rep.* 18, 1107–1113. doi:10.3892/or.18.5.1107

Skrivankova, V. W., Richmond, R. C., Woolf, B. A. R., Davies, N. M., Swanson, S. A., Vanderweele, T. J., et al. (2021). Strengthening the reporting of observational studies in epidemiology using mendelian randomisation (STROBE-MR): explanation and elaboration. *BMJ* 375, n2233. doi:10.1136/bmj.n2233

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

Song, J., Li, A., Qian, Y., Liu, B., Lv, L., Ye, D., et al. (2022). Genetically predicted circulating levels of cytokines and the risk of cancer. *Front. Immunol.* 13, 886144. doi:10.3389/fimmu.2022.886144

Soria, G., and Ben-Baruch, A. (2008). The inflammatory chemokines CCL2 and CCL5 in breast cancer. *Cancer Lett.* 267, 271–285. doi:10.1016/j.canlet.2008.03.018

Staley, J. R., Blackshaw, J., Kamat, M. A., Ellis, S., Surendran, P., Sun, B. B., et al. (2016). PhenoScanner: a database of human genotype-phenotype associations. *Bioinformatics* 32, 3207–3209. doi:10.1093/bioinformatics/btw373

Storey, J. D., and Tibshirani, R. (2003). Statistical significance for genomewide studies. *Proc. Natl. Acad. Sci. U. S. A.* 100, 9440–9445. doi:10.1073/pnas.1530509100

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., et al. (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* 71, 209–249. doi:10.3322/caac.21660

Tian, H., and Burgess, S. (2023). Estimation of time-varying causal effects with multivariable Mendelian randomization: some cautionary notes. *Int. J. Epidemiol.* 52, 846–857. doi:10.1093/ije/dyac240

Todoric, J., Antonucci, L., and Karin, M. (2016). Targeting inflammation in cancer prevention and therapy. *Cancer Prev. Res. (Phila)* 9, 895–905. doi:10.1158/1940-6207.CAPR-16-0209

Verbanck, M., Chen, C. Y., Neale, B., and Do, R. (2018). Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. *Nat. Genet.* 50, 693–698. doi:10.1038/s41588-018-0099-7

Voronov, E., Shouval, D. S., Krelin, Y., Cagnano, E., Benharroch, D., Iwakura, Y., et al. (2003). IL-1 is required for tumor invasiveness and angiogenesis. *Proc. Natl. Acad. Sci. U. S. A.* 100, 2645–2650. doi:10.1073/pnas.0437939100

Wu, T., Hong, Y., Jia, L., Wu, J., Xia, J., Wang, J., et al. (2016). Modulation of IL-1β reprogrammes the tumor microenvironment to interrupt oral carcinogenesis. *Sci. Rep.* 6, 20208. doi:10.1038/srep20208

Zhang, W., Wang, H., Sun, M., Deng, X., Wu, X., Ma, Y., et al. (2020). CXCL5/CXCR2 axis in tumor microenvironment as potential diagnostic biomarker and therapeutic target. *Cancer Commun. (Lond)* 40, 69–80. doi:10.1002/cac2.12010

Zhao, Y., Wei, K., Chi, H., Xia, Z., and Li, X. (2022). IL-7: a promising adjuvant ensuring effective T cell responses and memory in combination with cancer vaccines? *Front. Immunol.* 13, 1022808. doi:10.3389/fimmu.2022.1022808

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 09 February 2024
DOI 10.3389/fonc.2023.1275913

OPEN ACCESS

EDITED BY
Xuanye Cao,
University of Texas MD Anderson Cancer Center, United States

REVIEWED BY
Kimberly Dieudonnee van der Willik,
Ikazia Ziekenhuisn Rotterdam, Netherlands
Pedro Barata Coelho,
University Hospital Center of Porto, Portugal
Wen Jiang,
AstraZeneca, United States

*CORRESPONDENCE
Zhengkui Sun
sunzhengkui@sohu.com

†These authors have contributed equally to this work

RECEIVED 22 August 2023
ACCEPTED 27 December 2023
PUBLISHED 09 February 2024

CITATION
Yu C, Xu J, Xu S, Huang Y, Tang L, Zeng X, Yu T, Chen W and Sun Z (2024) Appraising the causal association between Crohn's disease and breast cancer: a Mendelian randomization study.
Front. Oncol. 13:1275913.
doi: 10.3389/fonc.2023.1275913

Appraising the causal association between Crohn's disease and breast cancer: a Mendelian randomization study

Chengdong Yu[1†], Jiawei Xu[1†], Siyi Xu[1†], Yanxiao Huang[1], Lei Tang[1], Xiaoqiang Zeng[1], Tenghua Yu[2], Wen Chen[2] and Zhengkui Sun[2*]

[1]Jiangxi Medical College, Nanchang University, Nanchang, China, [2]Department of Breast Surgery, Jiangxi Cancer Hospital, Nanchang, China

Background: Previous research has indicated that there may be a link between Crohn's disease (CD) and breast cancer (BC), but the causality remains unclear. This study aimed to investigate the causal association between CD and BC using Mendelian randomization (MR) analysis.

Methods: The summary data for CD (5,956 cases/14,927 controls) was obtained from the International Inflammatory Bowel Disease Genetics Consortium (IIBDGC). And the summary data for BC (122,977 cases/105,974 controls) was extracted from the Breast Cancer Association Consortium (BCAC). Based on the estrogen receptor status, the cases were classified into two subtypes: estrogen receptor-positive (ER^+) BC and estrogen receptor-negative (ER^-) BC. We used the inverse variance weighted method as the primary approach for two-sample MR. MR-PRESSO method was used to rule out outliers. Heterogeneity and pleiotropy tests were carried out to improve the accuracy of results. Additionally, multivariable MR was conducted by adjusting for possible confounders to ensure the stability of the results.

Results: The two-sample MR indicated that CD increased the risks of overall (OR: 1.020; 95% CI: 1.010-1.031; p=0.000106), ER^+ (OR: 1.019; 95%CI: 1.006-1.034; p=0.006) and ER^- BC (OR: 1.019; 95%CI: 1.000-1.037; p=0.046) after removal of outliers by MR-PRESSO. This result was reliable in the sensitivity analysis, including Cochran's Q and MR-Egger regression. In multivariate MR analyses, after adjusting for smoking and drinking separately or concurrently, the positive association between CD and the risks of overall and ER^+ BC remained, but it disappeared in ER^- BC. Furthermore, reverse MR analysis suggested that BC did not have a significant impact on CD risk.

Conclusion: Our findings provide evidence for a possible positive association between CD and the risk of BC. However, further studies are needed to fully understand the underlying mechanisms and establish a stronger causal relationship.

KEYWORDS

Crohn's disease, breast cancer, Mendelian randomization, causal association, risk

1 Introduction

Crohn's disease (CD) is a chronic and progressive inflammatory disease characterized by alternating periods of remission and relapse (1, 2). CD primarily affects the gastrointestinal tract with extraintestinal manifestations and related immune dysregulation (3). Patients with CD are more susceptible to cancer, depression, and infection (4).

Breast cancer (BC) is the most prevalent malignancy among women globally, with 684,996 deaths reported in 2020, representing a substantial threat to their health (5–7). Patients with CD have an increased risk of digestive tract, skin, bladder, and lung cancers (8, 9), but the association between CD and BC remains unclear. Chronic inflammation characterized by sustained immune activation is associated with promoting the occurrence, growth, and progression of BC (10–12). Several researchers have investigated the association between CD and BC. Riegler et al. found first-degree relatives of patients with CD have a higher risk of developing BC (13). Further, a study by Pellino et al. showed that CD was an independent risk factor for BC (OR: 2.76; 95% CI: 1.2-6.2; p=0.017) (14). In contrast, Gong et al. reported no significant association between CD and BC risk (15). Hence, there is controversy regarding the relationship between CD and BC risk. In addition, immunosuppressive medications are the cornerstone of long-term maintenance treatment for CD (16). Due to the decreased immune surveillance, immunosuppression may potentially increase the risk of cancer (17). A retrospective study attributed the development of BC in CD patients to immunosuppressive therapy (18). Thus, the association between CD itself and BC needs to be further investigated. Moreover, assessing the true causal association between CD and BC is challenging due to the interference of common residual confounders and reverse causality in traditional observational studies.

To overcome these challenges and gain a more nuanced understanding of the causality between CD and BC, we turned to Mendelian randomization (MR). MR is a robust statistical method that harnesses genetic variants as instrumental variables (IVs) to explore causal connections between exposure and outcome (19, 20). By capitalizing on the natural random assortment of genetic variants during conception, MR effectively mimics the randomized controlled trial (RCT) setting, thereby mitigating issues like confounding and reverse causation that often plague observational studies (21–23).

2 Materials and methods

2.1 Study design

In order to assess the potential causal association between CD and BC, we conducted a two-sample MR study. The single nucleotide polymorphisms (SNPs) selected as IVs were required to adhere to three following key premises (24): (1) SNPs must be intensely linked to CD; (2) SNPs must not be linked to confounding factors; and (3) SNPs should not be directly linked to BC (Figure 1).

2.2 Data source

The summary data for CD (5,956 cases/14,927 controls) was obtained from the International Inflammatory Bowel Disease Genetics Consortium (IIBDGC) (25). And the summary data for BC (122,977 cases/105,974 controls) was extracted from the Breast Cancer Association Consortium (BCAC). Based on the estrogen receptor status, the cases were classified into two subtypes: estrogen receptor-positive (ER^+) BC and estrogen receptor-negative (ER^-) BC (26). Table 1 presents details of the exposure and outcomes.

2.3 SNP selection

First, we screened for SNPs that were strongly associated with exposure at a genome-wide significance level ($p < 5\times10^{-8}$). Second, we implemented a criterion ($r^2 < 0.001$, kb=10000) to select SNPs that were independent of linkage disequilibrium (LD) (27). Third, we excluded SNPs that were not found in the BC dataset and palindromic SNPs that may cause bias. Next, we harmonized the exposure and outcome data, ensuring that the effect of the SNP on the exposure corresponded to the same allele as the effect on the outcome. Subsequently, we evaluated the possibility of weak instrumental bias by calculating F-statistics and excluded SNPs with F-statistics less than 10 (28, 29). The F statistic was calculated as $F = beta^2/se^2$ (30, 31). Finally, the MR-PRESSO method was conducted to detect outlier SNPs (32), and after excluding these outlier SNPs, the remaining SNPs were used for subsequent MR analysis. Figure 2 shows the selection flowchart.

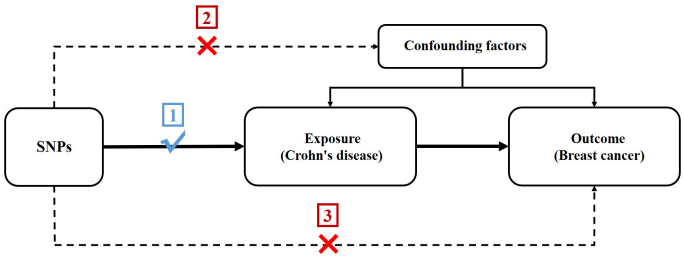


FIGURE 1
Schematic for the MR study design.

TABLE 1 Detailed information on the exposure and outcomes.

Exposure/Outcome	ncase	ncontrol	Sample size	Consortium	Ancestry
Crohn's disease	5,956	14,927	20,883	IIBDGC	European
Overall Breast cancer	122,977	105,974	228,951	BCAC	European
ER^+ Breast cancer	69,501	105,974	175,475	BCAC	European
ER^- Breast cancer	21,468	105,974	127,442	BCAC	European

BCAC, Breast Cancer Association Consortium; IIBDGC, International Inflammatory Bowel Disease Genetics Consortium.

2.4 Two-sample MR analysis

Three commonly used MR methods were applied to estimate causal effects: inverse variance weighted (IVW) (24), weighted median (33), and MR-Egger (34). The IVW method is considered to be the most effective method for assessing causality (35); therefore, the results were mainly based on the IVW method. We used odds ratios (ORs) to express the effects of CD on BC risk. If the result of the IVW method is significant ($p < 0.05$), even if no significant result is obtained by the other methods, it can be considered as a positive result as long as the ORs of the other methods are in the same direction and there is no heterogeneity or pleiotropy (36).

2.5 Sensitivity analysis

Cochran's Q test was employed to assess heterogeneity, with $p > 0.05$ representing the absence of heterogeneity (37). The MR-Egger regression test was applied to detect horizontal pleiotropy, with a zero intercept signifying the absence of pleiotropy ($p > 0.05$) (38).

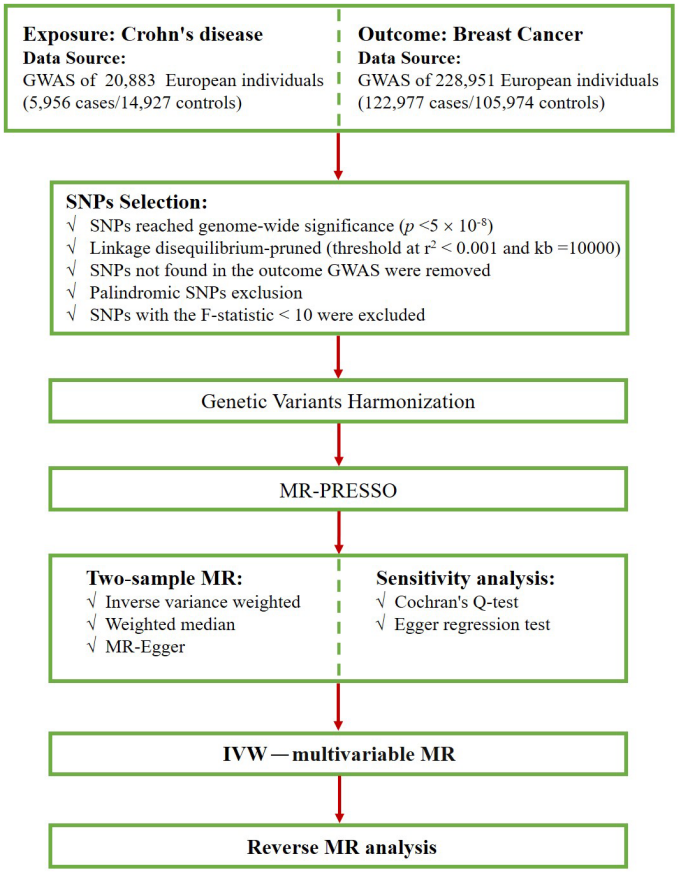


FIGURE 2
Flowchart of our MR study.

2.6 Multivariable MR analysis

Based on the search results on the PhenoScanner website and possible confounders between CD and BC, we performed multivariable MR (MVMR) analyses with the addition of smoking and drinking separately or together to adjust for causal impacts between exposure and outcome (39).

2.7 Reverse MR analysis

To explore whether BC has any causal effect on CD, we also conducted a reverse MR analysis (i.e., BC as the exposure and CD as the outcome) using SNPs related to BC as IVs.

2.8 Statistical analyses

All analyses were performed in R software (version 4.2.3) using the "TwoSampleMR" (version 0.5.6), "MRPRESSO" (version 1.0), and "MendelianRandomization" (version 0.7.0) packages (40).

3 Results

3.1 SNP selection

Initially, we extracted 53 genome-wide significant ($p<5\times10^{-8}$) SNPs associated with CD. No SNPs were ruled out due to LD. Next, during the extraction of information on IVs and outcome, we excluded rs11564236 due to the lack of corresponding outcome data. Additionally, we excluded one palindromic SNP (rs12692254) while harmonizing the exposure and outcome data. Furthermore, we removed rs7543234 from the analysis of overall BC due to its association with the outcome. Finally, potentially outlier SNPs were excluded using MR-PRESSO. Specifically, rs12194825, rs1873625, rs2188962, and rs3091315 were excluded from the analysis of overall BC; rs12194825, rs1873625, rs2188962, and rs7543234 were excluded from the analysis of ER^+ BC, and rs1873625 and rs3091315 were removed from the analysis of ER^- BC. The F-statistics of all SNPs were greater than 10. After removing these SNPs, 46 SNPs, 47 SNPs, and 49 SNPs were included in the analysis of overall, ER^+, and ER^- BC, respectively (Supplementary Sheet).

3.2 Analyses using the Two-sample MR

Using existing SNPs as IV, the results of the IVW method showed that CD was positively associated with the risks of overall (OR: 1.020; 95% CI: 1.010-1.031; p=0.000106), ER^+ (OR: 1.019; 95% CI: 1.006-1.034; p=0.006), and ER^- (OR: 1.019; 95% CI: 1.000-1.037; p=0.046) BC (Figure 3). The scatterplot depicts the causal estimates obtained from every SNP (Figure 4). Although the weighted median and MR-Egger methods did not obtain significant results ($p > 0.05$), the direction of the ORs was consistent with the IVW method (OR > 1). Furthermore, Cochran's Q and MR-Egger regression analyses demonstrated that there was no heterogeneity or horizontal pleiotropy affecting the stability of the results. The same result was also suggested by the symmetry of the funnel plots (Figure 5). Therefore, based on the significant IVW results ($p < 0.05$), we can conclude that there is a causal effect of CD on BC. The details of the results are presented in Table 2.

3.3 Analyses using the MVMR

After adjusting for current tobacco smoking and alcoholic drinks per week separately or together, MVMR analysis revealed that the positive association between CD and the risks of overall and ER^+ BC remained, but it disappeared in ER^- BC. In addition, no potential horizontal pleiotropy was discovered for the MR-Egger intercept (Table 3). Results of MVMR suggested that the observed effects of CD on overall and ER^+ BC were stable and not influenced by potential confounders.

3.4 Reverse MR analysis

In the reverse study (BC on the risk of CD), no genetic effects of overall BC (OR: 1.082; 95% CI: 0.989-1.183; p=0.085), ER^+ BC (OR: 1.039; 95% CI: 0.950-1.136; p=0.405), and ER^- BC (OR: 1.033; 95% CI: 0.924-1.156; p=0.567) on the risk of CD were detected (Table 4). In all of the analyses, MR-Egger regression did not show IVs had horizontal pleiotropy. Therefore, genetically predicted BC exerts no impact on the risk of CD.

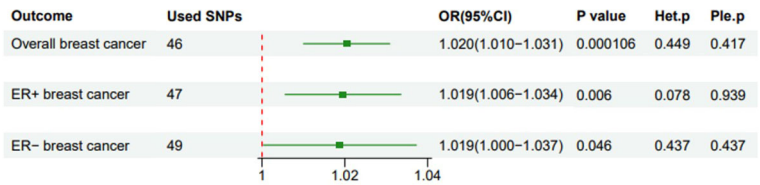


FIGURE 3
Forest plot of the impact of Crohn's disease on breast cancer risk using IVW method (after removing outliers). Het.p refers to the p-value for heterogeneity; Ple.p refers to the p-value for pleiotropy; OR, odds ratio; CI, confidence interval.

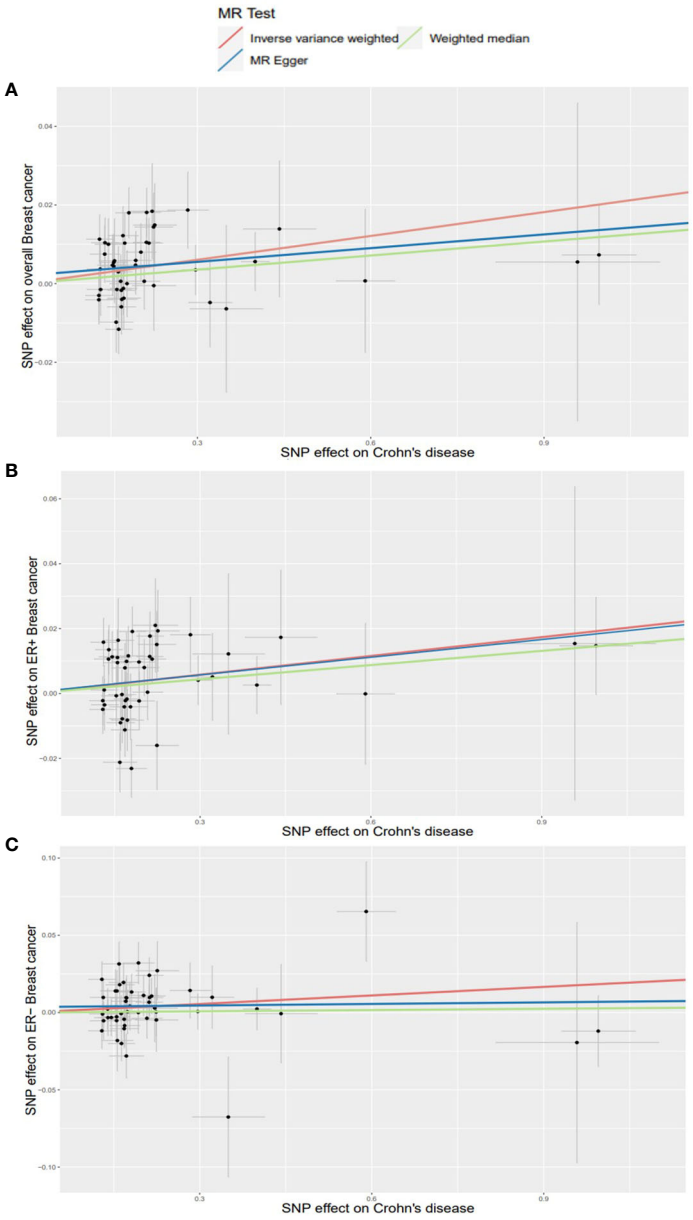


FIGURE 4
Scatterplots for effects of Crohn's disease on breast cancer risk (after removing outliers). **(A)** overall BC; **(B)** ER^+ BC; **(C)** ER^- BC.

4 Discussion

In this study, we carried out two-sample MR analyses to appraise the causal relationship of CD with overall, ER^+, and ER^- BC for the first time. The results showed that CD increased the risks of overall, ER^+, and ER^- BC. We further assessed the robustness of the results by MVMR analysis. However, in MVMR analysis, CD only increased the risks of overall and ER^+ BC, but not ER^- BC. This suggested a potential impact of smoking and alcohol drinking on the correlation between CD and ER^- BC. Additionally, reverse MR analysis revealed that BC did not have a significant impact on CD risk.

However, a recent MR study found no association between CD and BC risk (41). We analyzed possible reasons for the discrepancy. Their study included only 732 cases of CD, whereas our study included 5,956 cases. They used a significance threshold of $p<5\times10^{-6}$ for SNP selection, but re-running MR on their data at $p<5\times10^{-8}$ revealed a positive association of CD on BC risk ($p=0.016$). Furthermore, we conducted subtype analyses based on estrogen receptor status and performed MVMR to adjust for possible confounding factors.

This MR study provides some insights into the association between CD and BC. Some studies have also revealed an elevated risk of BC in patients with CD (13, 14). The result of a 20-year follow-up study indicated that CD patients have a higher risk of developing BC (42). In addition, a study from Denmark showed BC patients with CD have a more advanced stage and a worse chemotherapy prognosis than patients without CD (43).

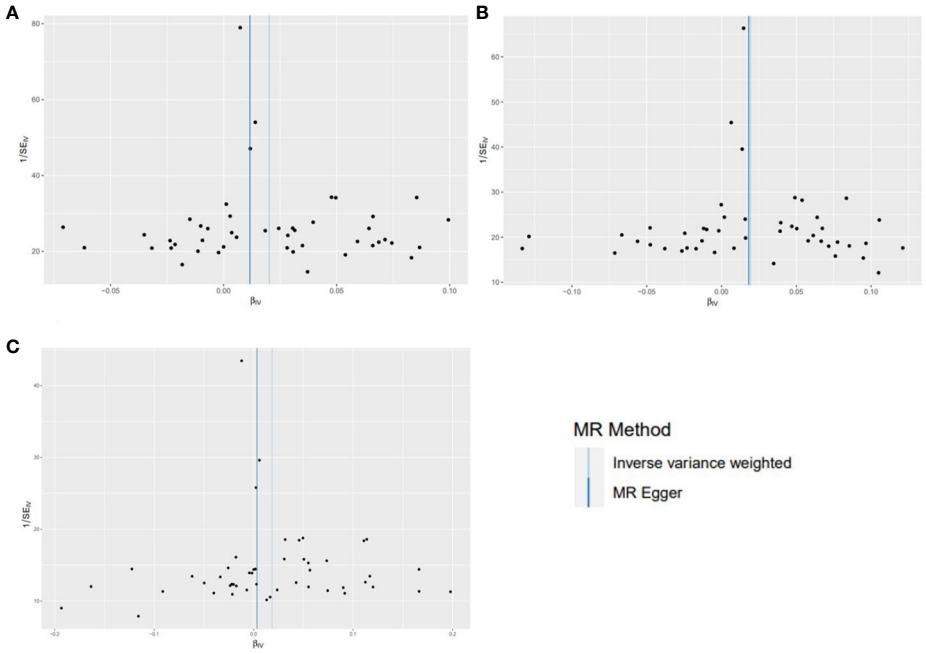


FIGURE 5
Funnel plots for effects of Crohn's disease on breast cancer risk (after removing outliers). **(A)** overall BC; **(B)** ER^+ BC; **(C)** ER^- BC.

Several possible factors may account for the association between CD and BC. Existing studies indicated that CD and BC may share common molecular mechanisms. Recent evidence suggested that there are 53 overlapping differentially expressed genes between the CD and BC, with enrichment analyses showing that both diseases are related to NF-κB signaling pathways and interleukin-17 (IL-17) (44). It has been shown that inflammation is involved in the process of development and progression of malignant tumors (45). T helper 17 (Th17) cells are important inflammatory mediators in CD, and when Th17 cells reach breast tumor tissues, they upregulate a variety of cytokines including IL-17 and tumor necrosis factor-α (TNF-α) (46). IL-17 can upregulate the expression of chemokine CXCL1 in BC cells. This chemokine increases the activation of the AKT/NF-κB signaling pathway to promote BC growth and metastasis (47). Furthermore, previous studies have indicated that TNF-α is involved in epithelial-mesenchymal transition (EMT), thereby promoting tumor metastasis (48). A study conducted on patients with inflammatory BC demonstrated a direct association between TNF-α and the presence of tumor cells expressing EMT markers (49). In addition, there is another potential point of association between CD and BC that lies in the involvement of estrogen and the G protein-coupled estrogen receptor (GPER) (50, 51). GPER has been shown to regulate intestinal function, inflammation, and immune responses, and promote the occurrence and progression of BC (52, 53).

There is growing interest in the role of the microbiome in health and disease. Studies in human subjects have revealed distinct differences in the gut microbiome between patients with CD and

TABLE 2 Assessing the effects of Crohn's disease on breast cancer risk (after removing outliers).

Outcome	MR-PRESSO Outliers	Used SNPs	method	OR (95%CI)	*P*	Het.p	Ple.p
Overall breast cancer	rs12194825, rs1873625, rs2188962, rs3091315	46	IVW	1.020(1.010-1.031)	0.000106	0.449	0.417
			weighted median	1.012(0.996-1.029)	0.152		
			MR-Egger	1.012(0.989-1.035)	0.326		
ER^+ breast cancer	rs12194825, rs1873625, rs2188962, rs7543234	47	IVW	1.019(1.006-1.034)	0.006	0.078	0.939
			weighted median	1.015(0.995-1.035)	0.144		
			MR-Egger	1.018(0.987-1.051)	0.261		
ER^- breast cancer	rs1873625, rs3091315	49	IVW	1.019(1.000-1.037)	0.046	0.437	0.437
			weighted median	1.003(0.976-1.031)	0.846		
	,		MR-Egger	1.003(0.962-1.046)	0.874		

Het.p refers to the p-value for heterogeneity; Ple.p refers to the p-value for pleiotropy; OR, odds ratio; CI, confidence interval.

TABLE 3 Assessing the effects of Crohn's disease on breast cancer using IVW multivariable MR.

Outcome	Adjustment	OR (95%CI)	*P*	Egger-Intercept	Int.p
Overall breast cancer	Current tobacco smoking	1.018(1.003-1.033)	0.019	<0.001	0.976
	Alcoholic drinks per week	1.019(1.001-1.037)	0.037	-0.001	0.724
	Adjust together	1.017(1.001-1.034)	0.037	-0.001	0.586
ER^+ breast cancer	Current tobacco smoking	1.023(1.005-1.040)	0.011	0.001	0.560
	Alcoholic drinks per week	1.021(1.002-1.041)	0.034	<0.001	0.960
	Adjust together	1.021(1.003-1.039)	0.025	<0.001	0.982
ER^- breast cancer	Current tobacco smoking	1.002(0.980-1.024)	0.864	-0.002	0.555
	Alcoholic drinks per week	1.003(0.978-1.028)	0.833	-0.005	0.153
	Adjust together	0.998(0.975-1.022)	0.869	-0.003	0.164

Int.p refers to the p-value derived from the Egger-intercept.

TABLE 4 Assessing the effects of breast cancer on Crohn's disease using IVW method.

Expose	Outcome	Used SNPs	OR (95%CI)	*P*	Egger-Intercept	Int.p
Overall breast cancer	Crohn's disease	132	1.082(0.989-1.183)	0.085	0.002	0.790
ER^+ breast cancer	Crohn's disease	97	1.039(0.950-1.136)	0.405	-0.006	0.431
ER^- breast cancer	Crohn's disease	34	1.033(0.924-1.156)	0.567	-0.018	0.293

Int.p refers to the p-value derived from the Egger-intercept.

healthy control subjects (54). Notably, the gut microbiome also affects the risk of developing BC (55). Dysbiosis of the intestinal flora has been found to have a direct effect on the dissemination of breast tumors (56, 57). The gut microbiome may also be involved in the correlation between CD and the risk of BC, and more relevant research is needed to confirm this in the future.

Research has demonstrated that chronic psychological stress can inhibit the anti-tumor effects of the immune system in CD (58). Intestinal inflammation in CD can activate the hypothalamic–pituitary–adrenal (HPA) axis through the opposite action of the brain-gut axis, thereby inducing anxiety and depression (59, 60). Several studies have shown that patients with BC also experience varying degrees of anxiety and depression (61). Hence, the mechanisms behind the effects of mental and emotional factors on CD and BC need to be further explored. The possible mechanisms for the effect of CD on BC risk are depicted in Figure 6.

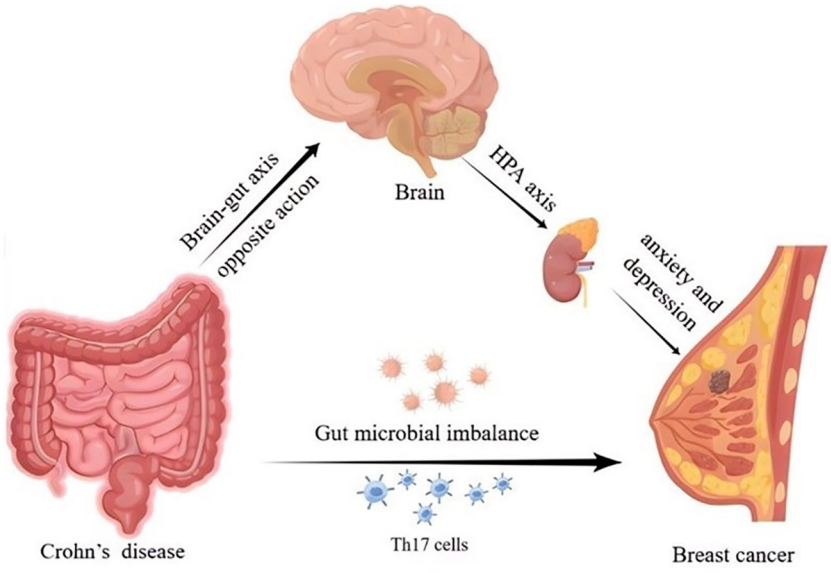


FIGURE 6
Diagram of possible mechanisms for the effect of Crohn's disease on breast cancer risk.

The strength of our study is that it explored the causality between CD and BC risk by MR analysis. Compared to previous observational studies that found CD to be associated with BC, MR studies are less susceptible to confounders and reverse causation. Besides, our study utilized a large sample size and SNPs from GWAS, providing sufficient statistical validity to estimate causality. Furthermore, sensitivity analyses enhanced the credibility of our results.

Nevertheless, this study has several limitations. First, the GWAS data for this study included only European populations, which limits the application of our findings to other populations. Hence, future studies are required to verify the applicability of our results to different populations. Second, we cannot stratify the analysis by sex due to the lack of sex-specific GWAS data. Third, the OR of CD on BC risk is relatively small, indicating that the enhanced risk is just modest. Therefore, we don't recommend that patients with CD be screened for BC more frequently or earlier than the routine screening. Finally, MR also has its limitations. (1) SNPs are generally considered to have lifetime effects, but in specific situations, the effects of SNPs may vary due to an individual's physiological status, environmental factors, or interactions with other genetic variations. If the genetic variants used in MR analysis change over time, it could potentially affect the validity of the causal estimates. (2) Additional adjustments for smoking and alcohol consumption may lead to collider bias. (3) The MR study can only analyze the causality and cannot explain the mechanism of CD on BC risk. Further research is necessary to investigate the mechanisms behind the link between CD and the risk of BC.

5 Conclusion

Our findings provide evidence for a potential positive association between CD and the risk of BC. However, further studies are needed to fully understand the underlying mechanisms and establish a stronger causal relationship.

Data availability statement

In this study, all GWAS data were extracted from the IEU Open GWAS project (https://gwas.mrcieu.ac.uk/).

Author contributions

CY: Conceptualization, Investigation, Methodology, Writing – original draft. JX: Data curation, Writing – original draft. SX: Data curation, Writing – review & editing. YH: Writing – review & editing. LT: Writing – review & editing. XZ: Writing – review & editing. TY: Formal Analysis, Writing – original draft. WC: Visualization, Writing – review & editing. ZS: Funding acquisition, Supervision, Validation, Writing – review & editing.

Funding

The author(s) declare financial support was received for theresearch, authorship, and/or publication of this article. This work was supported by the Wu Jieping Medical Foundation (320.6750.2020-20-17) and the Open Fund for Scientific Research of Jiangxi Cancer Hospital (2021K04).

Acknowledgments

Thanks to the MRC Integrative Epidemiology Unit (IEU) at the University of Bristol for their invaluable contribution to the development of the IEU open GWAS project. Thank them for deriving relevant GWAS data from IIBDGC and BCAC. We plotted Figure 6 by Figdraw, thanks to Home for Researchers for providing the plotting platform.

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2023.1275913/full#supplementary-material

References

1. Torres J, Mehandru S, Colombel JF, Peyrin-Biroulet L. Crohn's disease. *Lancet* (2017) 389(10080):1741–55. doi: 10.1016/S0140-6736(16)31711-1

2. Noor NM, Verstockt B, Parkes M, Lee JC. Personalised medicine in Crohn's disease. *Lancet Gastroenterol Hepatol* (2020) 5(1):80–92. doi: 10.1016/S2468-1253(19)30340-1

3. Roda G, Chien Ng S, Kotze PG, Argollo M, Panaccione R, Spinelli A, et al. Crohn's disease. *Nat Rev Dis Primers* (2020) 6(1):22. doi: 10.1038/s41572-020-0156-2

4. Veauthier B, Hornecker JR. Crohn's disease: diagnosis and management. *Am Fam Physician* (2018) 98(11):661–9.

5. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* (2021) 71(3):209–49. doi: 10.3322/caac.21660

6. Arnold M, Morgan E, Rumgay H, Mafra A, Singh D, Laversanne M, et al. Current and future burden of breast cancer: Global statistics for 2020 and 2040. *Breast* (2022) 66:15–23. doi: 10.1016/j.breast.2022.08.010

7. Loibl S, Poortmans P, Morrow M, Denkert C, Curigliano G. Breast cancer. *Lancet* (2021) 397(10286):1750–69. doi: 10.1016/S0140-6736(20)32381-3

8. Olen O, Erichsen R, Sachs MC, Pedersen L, Halfvarson J, Askling J, et al. Colorectal cancer in Crohn's disease: a Scandinavian population-based cohort study. *Lancet Gastroenterol Hepatol* (2020) 5(5):475–84. doi: 10.1016/S2468-1253(20)30005-4

9. Pedersen N, Duricova D, Elkjaer M, Gamborg M, Munkholm P, Jess T. Risk of extra-intestinal cancer in inflammatory bowel disease: meta-analysis of population-based cohort studies. *Am J Gastroenterol* (2010) 105(7):1480–7. doi: 10.1038/ajg.2009.760

10. Atsumi T, Singh R, Sabharwal L, Bando H, Meng J, Arima Y, et al. Inflammation amplifier, a new paradigm in cancer biology. *Cancer Res* (2014) 74(1):8–14. doi: 10.1158/0008-5472.CAN-13-2322

11. Cole SW. Chronic inflammation and breast cancer recurrence. *J Clin Oncol* (2009) 27(21):3418–9. doi: 10.1200/JCO.2009.21.9782

12. Danforth DN. The role of chronic inflammation in the development of breast cancer. *Cancers (Basel)* (2021) 13(15):3918. doi: 10.3390/cancers13153918

13. Riegler G, Caserta L, Castiglione F, Esposito I, Valpiani D, Annese V, et al. Increased risk of breast cancer in first-degree relatives of Crohn's disease patients. An IG-IBD study. *Dig Liver Dis* (2006) 38(1):18–23. doi: 10.1016/j.dld.2005.07.006

14. Pellino G, Sciaudone G, Patturelli M, Candilio G, De Fatico GS, Landino I, et al. Relatives of Crohn's disease patients and breast cancer: an overlooked condition. *Int J Surg* (2014) 12(Suppl 1):S156–8. doi: 10.1016/j.ijsu.2014.05.022

15. Gong C, Xu R, Zou P, Zhang Y, Wang X. Inflammatory bowel disease and risk of breast cancer: a meta-analysis of cohort studies. *Eur J Cancer Prev* (2022) 31(1):54–63. doi: 10.1097/CEJ.0000000000000667

16. Hazenberg H, de Boer NKH, Mulder CJJ, Mom SH, van Bodegraven AA, Tack Md Ph DG. Neoplasia and precursor lesions of the female genital tract in IBD: epidemiology, role of immunosuppressants, and clinical implications. *Inflammation Bowel Dis* (2018) 24(3):510–31. doi: 10.1093/ibd/izx062

17. Annese V, Beaugerie L, Egan L, Biancone L, Bolling C, Brandts C, et al. European evidence-based consensus: inflammatory bowel disease and Malignancies. *J Crohns Colitis* (2015) 9(11):945–65. doi: 10.1093/ecco-jcc/jjv141

18. Hamzaoglu H, Cooper J, Alsahli M, Falchuk KR, Peppercorn MA, Farrell RJ. Safety of infliximab in Crohn's disease: a large single-center experience. *Inflammation Bowel Dis* (2010) 16(12):2109–16. doi: 10.1002/ibd.21290

19. Richmond RC, Davey Smith G. Mendelian randomization: concepts and scope. *Cold Spring Harb Perspect Med* (2022) 12(1):a040501. doi: 10.1101/cshperspect.a040501

20. Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum Mol Genet* (2014) 23(R1):R89–98. doi: 10.1093/hmg/ddu328

21. Yarmolinsky J, Wade KH, Richmond RC, Langdon RJ, Bull CJ, Tilling KM, et al. Causal inference in cancer epidemiology: what is the role of Mendelian randomization? *Cancer Epidemiol Biomarkers Prev* (2018) 27(9):995–1010. doi: 10.1158/1055-9965.EPI-17-1177

22. Sekula P, Del Greco MF, Pattaro C, Kottgen A. Mendelian randomization as an approach to assess causality using observational data. *J Am Soc Nephrol* (2016) 27(11):3253–65. doi: 10.1681/ASN.2016010098

23. Emdin CA, Khera AV, Kathiresan S. Mendelian randomization. *JAMA* (2017) 318(19):1925–6. doi: 10.1001/jama.2017.17219

24. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. *Genet Epidemiol* (2013) 37(7):658–65. doi: 10.1002/gepi.21758

25. Liu JZ, van Sommeren S, Huang H, Ng SC, Alberts R, Takahashi A, et al. Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. *Nat Genet* (2015) 47(9):979–86. doi: 10.1038/ng.3359

26. Michailidou K, Lindstrom S, Dennis J, Beesley J, Hui S, Kar S, et al. Association analysis identifies 65 new breast cancer risk loci. *Nature* (2017) 551(7678):92–4. doi: 10.1038/nature24284

27. Nounu A, Kar SP, Relton CL, Richmond RC. Sex steroid hormones and risk of breast cancer: a two-sample Mendelian randomization study. *Breast Cancer Res* (2022) 24(1):66. doi: 10.1186/s13058-022-01553-9

28. Burgess S, Thompson SG. Use of allele scores as instrumental variables for Mendelian randomization. *Int J Epidemiol* (2013) 42(4):1134–44. doi: 10.1093/ije/dyt093

29. Pierce BL, Ahsan H, Vanderweele TJ. Power and instrument strength requirements for Mendelian randomization studies using multiple genetic variants. *Int J Epidemiol* (2011) 40(3):740–52. doi: 10.1093/ije/dyq151

30. Huang W, Xiao J, Ji J, Chen L. Association of lipid-lowering drugs with COVID-19 outcomes from a Mendelian randomization study. *Elife* (2021) 10:e73873. doi: 10.7554/eLife.73873

31. 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(1):69. doi: 10.1186/s12863-022-01082-6

32. Verbanck M, Chen CY, 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(5):693–8. doi: 10.1038/s41588-018-0099-7

33. 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(4):304–14. doi: 10.1002/gepi.21965

34. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol* (2015) 44(2):512–25. doi: 10.1093/ije/dyv080

35. Hartwig FP, Davey Smith G, Bowden J. Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. *Int J Epidemiol* (2017) 46(6):1985–98. doi: 10.1093/ije/dyx102

36. Wang S, Zhu H, Pan L, Zhang M, Wan X, Xu H, et al. Systemic inflammatory regulators and risk of acute-on-chronic liver failure: A bidirectional mendelian-randomization study. *Front Cell Dev Biol* (2023) 11:1125233. doi: 10.3389/fcell.2023.1125233

37. Cohen JF, Chalumeau M, Cohen R, Korevaar DA, Khoshnood B, Bossuyt PM. Cochran's Q test was useful to assess heterogeneity in likelihood ratios in studies of diagnostic accuracy. *J Clin Epidemiol* (2015) 68(3):299–306. doi: 10.1016/j.jclinepi.2014.09.005

38. Burgess S, Thompson SG. Interpreting findings from Mendelian randomization using the MR-Egger method. *Eur J Epidemiol* (2017) 32(5):377–89. doi: 10.1007/s10654-017-0255-x

39. Sanderson E. Multivariable Mendelian randomization and mediation. *Cold Spring Harb Perspect Med* (2021) 11(2):a038984. doi: 10.1101/cshperspect.a038984

40. Yavorska OO, Burgess S. MendelianRandomization: an R package for performing Mendelian randomization analyses using summarized data. *Int J Epidemiol* (2017) 46(6):1734–9. doi: 10.1093/ije/dyx034

41. Lu Y, Ma L. Investigation of the causal relationship between breast cancer and autoimmune diseases: A bidirectional mendelian randomization study. *Med (Baltimore)* (2023) 102(34):e34612. doi: 10.1097/MD.0000000000034612

42. Hovde O, Hoivik ML, Henriksen M, Solberg IC, Smastuen MC, Moum BA. Malignancies in patients with inflammatory bowel disease: results from 20 years of follow-up in the IBSEN study. *J Crohns Colitis* (2017) 11(5):571–7. doi: 10.1093/ecco-jcc/jjw193

43. Sogaard KK, Cronin-Fenton DP, Pedersen L, Sorensen HT, Lash TL. Survival in Danish patients with breast cancer and inflammatory bowel disease: a nationwide cohort study. *Inflammation Bowel Dis* (2008) 14(4):519–25. doi: 10.1002/ibd.20341

44. Zhou J, Yang R. Identification of key pathways and genes shared between Crohn's disease and breast cancer using bioinformatics analysis. *Oncol Lett* (2020) 20(4):119. doi: 10.3892/ol.2020.11981

45. Diakos CI, Charles KA, McMillan DC, Clarke SJ. Cancer-related inflammation and treatment effectiveness. *Lancet Oncol* (2014) 15(11):e493–503. doi: 10.1016/S1470-2045(14)70263-3

46. Brand S. Crohn's disease: Th1, Th17 or both? The change of a paradigm: new immunological and genetic insights implicate Th17 cells in the pathogenesis of Crohn's disease. *Gut* (2009) 58(8):1152–67. doi: 10.1136/gut.2008.163667

47. Ma K, Yang L, Shen R, Kong B, Chen W, Liang J, et al. Th17 cells regulate the production of CXCL1 in breast cancer. *Int Immunopharmacol* (2018) 56:320–9. doi: 10.1016/j.intimp.2018.01.026

48. Asiedu MK, Ingle JN, Behrens MD, Radisky DC, Knutson KL. TGFbeta/TNF(alpha)-mediated epithelial-mesenchymal transition generates breast cancer stem cells with a claudin-low phenotype. *Cancer Res* (2011) 71(13):4707–19. doi: 10.1158/0008-5472.CAN-10-4554

49. Cohen EN, Gao H, Anfossi S, Mego M, Reddy NG, Debeb B, et al. Inflammation mediated metastasis: immune induced epithelial-to-mesenchymal transition in inflammatory breast cancer cells. *PloS One* (2015) 10(7):e0132710. doi: 10.1371/journal.pone.0132710

50. Wlodarczyk M, Sobolewska-Wlodarczyk A, Cygankiewicz AI, Jacenik D, Piechota-Polanczyk A, Stec-Michalska K, et al. G protein-coupled receptor 30 (GPR30) expression pattern in inflammatory bowel disease patients suggests its key role in the inflammatory process. *A Preliminary Study J Gastrointestin Liver Dis* (2017) 26(1):29–35. doi: 10.15403/jgld.2014.1121.261.gpr

51. Filardo EJ. A role for G-protein coupled estrogen receptor (GPER) in estrogen-induced carcinogenesis: Dysregulated glandular homeostasis, survival and metastasis. *J Steroid Biochem Mol Biol* (2018) 176:38–48. doi: 10.1016/j.jsbmb.2017.05.005

52. Jacenik D, Beswick EJ, Krajewska WM, Prossnitz ER. G protein-coupled estrogen receptor in colon function, immune regulation and carcinogenesis. *World J Gastroenterol* (2019) 25(30):4092–104. doi: 10.3748/wjg.v25.i30.4092

53. Prossnitz ER, Barton M. The G protein-coupled oestrogen receptor GPER in health and disease: an update. *Nat Rev Endocrinol* (2023) 19(7):407–24. doi: 10.1038/s41574-023-00822-7

54. Glassner KL, Abraham BP, Quigley EMM. The microbiome and inflammatory bowel disease. *J Allergy Clin Immunol* (2020) 145(1):16–27. doi: 10.1016/j.jaci.2019.11.003

55. Kwa M, Plottel CS, Blaser MJ, Adams S. The intestinal microbiome and estrogen receptor-positive female breast cancer. *J Natl Cancer Inst* (2016) 108(8):djw029. doi: 10.1093/jnci/djw029

56. Buchta Rosean C, Bostic RR, Ferey JCM, Feng TY, Azar FN, Tung KS, et al. Preexisting commensal dysbiosis is a host-intrinsic regulator of tissue inflammation and tumor cell dissemination in hormone receptor-positive breast cancer. *Cancer Res* (2019) 79(14):3662–75. doi: 10.1158/0008-5472.CAN-18-3464

57. Zhang J, Xia Y, Sun J. Breast and gut microbiome in health and cancer. *Genes Dis* (2021) 8(5):581–9. doi: 10.1016/j.gendis.2020.08.002

58. Hauser W, Janke KH, Klump B, Hinz A. Anxiety and depression in patients with inflammatory bowel disease: comparisons with chronic liver disease patients and the general population. *Inflammation Bowel Dis* (2011) 17(2):621–32. doi: 10.1002/ibd.21346

59. Labanski A, Langhorst J, Engler H, Elsenbruch S. Stress and the brain-gut axis in functional and chronic-inflammatory gastrointestinal diseases: A transdisciplinary challenge. *Psychoneuroendocrinology* (2020) 111:104501. doi: 10.1016/j.psyneuen.2019.104501

60. Ge L, Liu S, Li S, Yang J, Hu G, Xu C, et al. Psychological stress in inflammatory bowel disease: Psychoneuroimmunological insights into bidirectional gut-brain communications. *Front Immunol* (2022) 13:1016578. doi: 10.3389/fimmu.2022.1016578

61. Hashemi SM, Rafiemanesh H, Aghamohammadi T, Badakhsh M, Amirshahi M, Sari M, et al. Prevalence of anxiety among breast cancer patients: a systematic review and meta-analysis. *Breast Cancer* (2020) 27(2):166–78. doi: 10.1007/s12282-019-01031-9

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 16 February 2024
DOI 10.3389/fgene.2024.1353118

OPEN ACCESS

EDITED BY
Ziheng Wang,
University of Macau, China

REVIEWED BY
Ning Ding,
Changsha Central Hospital, China
Qiang Li,
Nanjing Medical University, China
Hao Zhang,
Chongqing Medical University, China

*CORRESPONDENCE
Fuhong Huang,
fuhong_huang@163.com

RECEIVED 09 December 2023
ACCEPTED 01 February 2024
PUBLISHED 16 February 2024

CITATION
Wei Z, Liu Y, Mei X, Zhong J and Huang F (2024), Circulating micronutrient levels and their association with sepsis susceptibility and severity: a Mendelian randomization study. *Front. Genet.* 15:1353118.
doi: 10.3389/fgene.2024.1353118

Circulating micronutrient levels and their association with sepsis susceptibility and severity: a Mendelian randomization study

Zhengxiao Wei[1], Yingfen Liu[1], Xue Mei[2], Jing Zhong[1] and Fuhong Huang[3]*

[1]Department of Clinical Laboratory, Public Health Clinical Center of Chengdu, Chengdu, Sichuan, China, [2]Department of Infectious Diseases, Public Health Clinical Center of Chengdu, Chengdu, Sichuan, China, [3]Department of Ultrasound, Sichuan Provincial People's Hospital, University of Electronic Science and Technology of China, Chengdu, China

Background: Sepsis, a global health challenge, necessitates a nuanced understanding of modifiable factors for effective prevention and intervention. The role of trace micronutrients in sepsis pathogenesis remains unclear, and their potential connection, especially with genetic influences, warrants exploration.

Methods: We employed Mendelian randomization (MR) analyses to assess the causal relationship between genetically predicted blood levels of nine micronutrients (calcium, β-carotene, iron, magnesium, phosphorus, vitamin C, vitamin B6, vitamin D, and zinc) and sepsis susceptibility, severity, and subtypes. The instrumental variables for circulating micronutrients were derived from nine published genome-wide association studies (GWAS). In the primary MR analysis, we utilized summary statistics for sepsis from two independent databases (UK Biobank and FinnGen consortium), for initial and replication analyses. Subsequently, a meta-analysis was conducted to merge the results. In secondary MR analyses, we assessed the causal effects of micronutrients on five sepsis-related outcomes (severe sepsis, sepsis-related death within 28 days, severe sepsis-related death within 28 days, streptococcal septicaemia, and puerperal sepsis), incorporating multiple sensitivity analyses and multivariable MR to address potential heterogeneity and pleiotropy.

Results: The study revealed a significant causal link between genetically forecasted zinc levels and reduced risk of severe sepsis-related death within 28 days (odds ratio [OR] = 0.450; 95% confidence interval [CI]: 0.263, 0.770; $p = 3.58 \times 10^{-3}$). Additionally, suggestive associations were found for iron (increased risk of sepsis), β-carotene (reduced risk of sepsis death) and vitamin C (decreased risk of puerperal sepsis). No significant connections were observed for other micronutrients.

Conclusion: Our study highlighted that zinc may emerges as a potential protective factor against severe sepsis-related death within 28 days, providing theoretical support for supplementing zinc in high-risk critically ill sepsis patients. In the future, larger-scale data are needed to validate our findings.

KEYWORDS

micronutrients, Mendelian randomization, sepsis, susceptibility, severe sepsis-related death within 28 days, zinc

Introduction

Sepsis is a critical health concern marked by an exaggerated immune response to infection, presenting a global public health challenge (Cecconi et al., 2018). It routinely precipitates multi-organ dysfunction, with high incidence and mortality rates (Evans et al., 2021), thereby imparting a substantial encumbrance upon societal and global healthcare infrastructures. Notwithstanding, the susceptibility and severity of sepsis are influenced by a multitude of factors (Rhee et al., 2017), accentuating the imperativeness of discerning modifiable factors for the prevention, timely diagnosis, and efficacious intervention in sepsis.

In recent years, although some factors potentially influencing sepsis have been identified, such as blood metabolites (Wei et al., 2023), body mass index (BMI) (Wang et al., 2023), insomnia (Thorkildsen et al., 2023), lifetime smoking (Zhu et al., 2023), the role of trace elements in the pathogenesis of sepsis remains unclear. Simultaneously, understanding the dysregulation of trace element metabolism in the pathogenesis of sepsis is not comprehensive (Guo et al., 2023). Numerous micronutrients have been reported to play a crucial role in the immune system, and their deficiency may severely impair host immunity, increasing the risk of infection (Gombart et al., 2020). Some studies emphasize vitamin C as a biological and theoretical basis for sepsis treatment (Spoelstra-de Man et al., 2018); however, a randomized controlled trial found no significant improvement in sepsis-related inflammation and vascular damage with vitamin C (Fowler et al., 2019). The disparity between these two study results may be influenced by factors such as sample size, follow-up time, and confounding variables. Due to the cost and practical difficulties, conducting sufficiently large randomized controlled trials is challenging, and there is limited research providing substantial support for the relationship between micronutrients and sepsis.

Mendelian randomization (MR) is an approach used to evaluate the relationship between risk factors and diseases in terms of causality. When there are no randomized controlled trials (RCTs) or new RCTs being conducted, MR becomes a valuable alternative approach that can provide dependable evidence on the causal connection between exposure and the risk of disease (Zuccolo and Holmes, 2017). In observational studies, MR utilizes genetic variation as an instrumental variable (IV) to successfully mitigate the influence of confounding factors that are difficult to control and reduces the likelihood of reverse causation.

In this study, we employ MR method to estimate the relationship between genetically predicted blood micronutrient levels and the risk of sepsis-related outcomes. We have selected nine

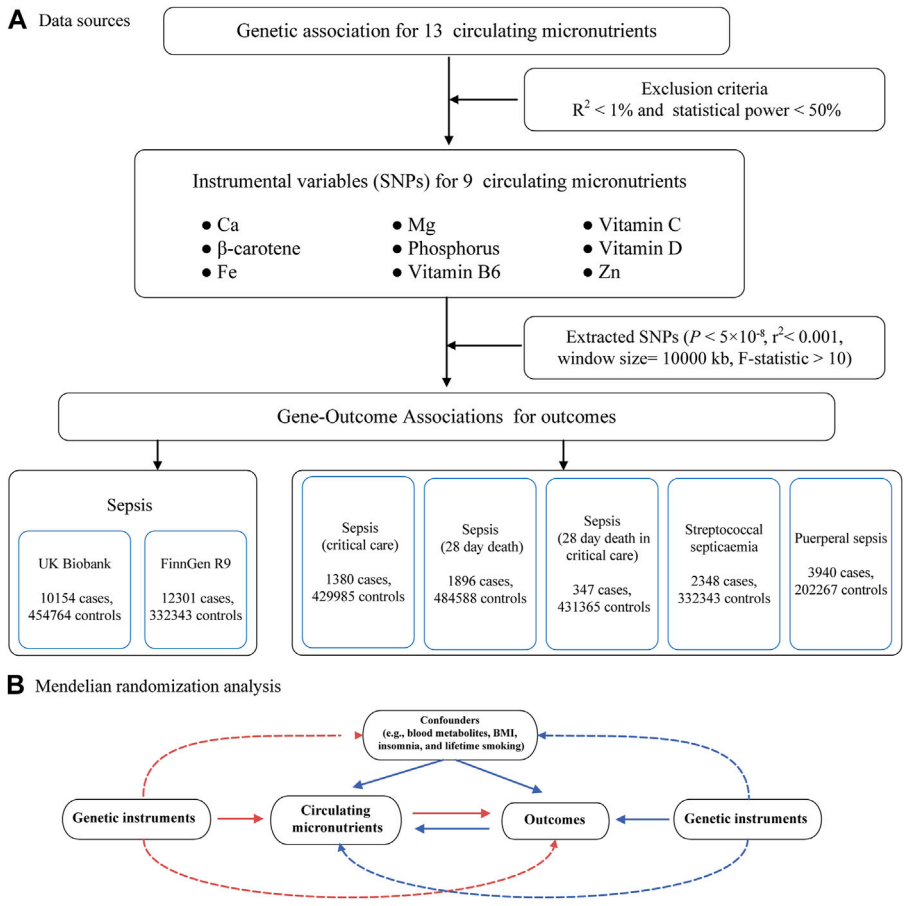


FIGURE 1
Flow chart for the Mendelian randomized analysis. **(A)** Data sources. **(B)** Mendelian randomization analysis. Abbreviations: IVW, inverse-variance weighted; Ca, Calcium; Fe, Iron; Mg, Magnesium; Zn, Zinc.

micronutrients associated with infection [calcium (Ca), β-carotene, iron (Fe), magnesium (Mg), phosphorus, vitamin C, vitamin B6, vitamin D, zinc (Zn)] and assessed the infection risk of sepsis and its susceptibility and severity.

Materials and methods

Research design

The inquiry follows the principles specified in the STROBE-MR guidelines (Supplementary Table S1). Refer to Figure 1 for an illustrative representation of the study design. In brief, we performed a comprehensive investigation of MR using data from 16 publicly accessible genome-wide association studies (GWAS) to obtain summary statistics. The objective was to elucidate the relationship between circulating micronutrient levels and sepsis. Of the 16 studies, 9 contributed exposure data, while 7 furnished outcome data. To reduce potential biases caused by population stratification, only individuals of European ancestry were included in the study for both the exposure and outcome data. For the primary MR analysis, sepsis data were procured from two independent GWAS consortia, utilized for preliminary and replicative analyses, culminating in a meta-analysis for result amalgamation. In secondary MR analyses, we scrutinized the causal nexus between micronutrients and the severity, as well as subtypes, of sepsis. Our analytic approach was bidirectional, initially scrutinizing the influence of circulating micronutrients on the susceptibility to sepsis and its associated maladies, followed by an exploration of reverse causality. A schematic overview of the study methodology is delineated in Figure 1.

The data used in this investigation were obtained from studies with the explicit consent and ethical endorsement of participants, eliminating the need for ethical approval from an institutional review board for the present study.

Data sources for circulating micronutrient

By searching on pubmed website (https//www.ncbi.nlm.nih.gov/pubmed) (last accessed on 1 March 2023), we obtained GWAS data related to micronutrients in the European population. To prevent sample overlap between the exposures and outcomes in our research, we excluded micronutrients sourced from these two databases during our micronutrient search. No genomic studies were found for vitamin B1, B2, B3, B5, B7, fluoride, chloride, sulfur, and iodine. Exclusion of global genomic investigations on vitamin K, cobalt, chromium, sodium, potassium, and molybdenum was based on the lack of significant findings across the entire genome (Meyer et al., 2010; Dashti et al., 2014; Ng et al., 2015). A preliminary identification of 15 potential micronutrients was established: Ca (O'Seaghdha et al., 2013), copper (Evans et al., 2013), Fe (Benyamin et al., 2014), Mg (Meyer et al., 2010), selenium (Cornelis et al., 2015), Zn (Evans et al., 2013), phosphorus (Kestenbaum et al., 2010), beta-carotene (Ferrucci et al., 2009), folate (Grarup et al., 2013), vitamin A (Mondul et al., 2011), vitamin B6 (Hazra et al., 2009), vitamin B12 (Grarup et al., 2013), vitamin C (Zheng et al., 2021), vitamin D (Jiang et al., 2018), and vitamin E (Major et al., 2011). However, Vitamin A and vitamin E were not considered in these GWAS because they were controlled for body mass index (BMI), which could lead to biased genetic effects caused by BMI adjustments (Aschard et al., 2015). Detailed information about the GWAS for the 13 candidate exposures was provided in Supplementary Table S2.

Data sources for sepsis-related outcomes

We used ICD-coded linked secondary care data to identify sepsis and sepsis-related outcomes. In the UK Biobank (Bycroft et al., 2018), sepsis and the severity were identified using ICD-10 codes A02, A39, A40, and A41. In the FinnGen database (Kurki et al., 2023), codes A40.9, A41, and O85 were used to identify sepsis and its subtypes, in line with recent literature (Hamilton et al., 2023). Cases were included if the code appeared in either the primary or secondary diagnostic position in Hospital Episode Statistics (HES) data or similar datasets in the devolved nations, as provided by the UK Biobank.

For our primary MR analysis, we chose sepsis as the primary outcome. For our study, we used summary statistics data from two separate cohorts of European ancestry, namely the UK Biobank and the FinnGen Release 9, which were employed as the outcomes. To ascertain the association of genetic variations with sepsis, we initially employed the latest version of sepsis GWAS summary data (10,154 cases and 454,764 controls) from the UK Biobank. To validate through replication and meta-analysis, we employed an additional collection of sepsis summary data (12,301 cases and 332,343 controls) obtained from the FinnGen consortium.

For secondary analyses, we opted for five sepsis-related outcomes, which encompassed three data points on sepsis severity obtained from the UK Biobank, severe sepsis (1,380 cases and 429,985 controls), and sepsis-related death within 28 days (1896 cases and 484,588 controls) and severe sepsis-related death within 28 days (347 cases and 431,365 controls). Moreover, the FinnGen cohort provided two variations of sepsis information, namely streptococcal septicaemia (2,348 cases and 332,343 controls) and puerperal sepsis (3,940 cases and 2202267 controls), which were obtained from https//r9.finngen.fi/pheno/. Refer to Table 1 for detailed information on data sources.

MR analysis

In the selection of genomically significant SNPs, we applied stringent thresholds, specifically $p < 5 \times 10^{-8}$, to obtain top independent SNPs strongly correlated with each micronutrient. Within 10,000-kb windows, we eliminated single nucleotide polymorphisms (SNPs) that were in linkage disequilibrium with parameters $r^2 < 0.001$. At the same time, in order to guarantee that the impact of SNPs on exposure aligns with their impact on outcomes for the identical allelic gene, we eliminated palindromic SNPs with moderate frequencies of alleles. To evaluate the statistical power, we calculated the F-statistic for every SNP. All F-statistics for the SNPs exceeded 10, indicating a minimal likelihood of weak instrumentality.

To establish the causal connection between sepsis and micronutrients, we utilized the inverse variance weighting (IVW)

TABLE 1 Source of outcome genome-wide association study summary data.

Outcome	Source	Cases	Control	Trait/Phenocode	Population
Sepsis	UK Biobank	10,154	454,764	Sepsis	European
Sepsis	FinnGen R9	12,301	332,343	AB1_other_sepsis	European
severe sepsis	UK Biobank	1,380	429,985	Sepsis (critical care)	European
sepsis-related death within 28 days	UK Biobank	1896	484,588	Sepsis (28 day death)	European
severe sepsis-related death within 28 days	UK Biobank	347	431,365	Sepsis (28 day death in critical care)	European
Streptococcal septicaemia	FinnGen R9	2,348	332,343	AB1_strepto_sepsis	European
Puerperal sepsis	FinnGen R9	3,940	202,267	O15_puerp_sepsis	European

technique as our primary analytical method. IVW is a commonly used method in MR studies, combining the wald ratios of each SNP to derive a summary estimate (Pierce and Burgess, 2013). In order to guarantee the dependability of our findings, we performed several sensitivity analyses to confirm if diversity and pleiotropy in genetic instruments could potentially cause bias in MR outcomes. The methods used in these analyses were the weighted median, MR-Egger, MR pleiotropy residual sum, and MR-PRESSO methods. Egger intercept (Bowden et al., 2015) was utilized to evaluate horizontal pleiotropy, while the MR-PRESSO test was employed for outlier identification (Verbanck et al., 2018). Cochran's Q test was performed to assess heterogeneity in the genetic instruments used across the two cohorts (Greco et al., 2015), with $p < 0.05$ indicating significant heterogeneity. Finally, for the outcomes of sepsis infection, an MR Steiger test was conducted to examine the directionality of the associations (Hemani et al., 2017).

To examine if the potential genetic tools related to micronutrients were linked to other traits like blood metabolites (Wei et al., 2023), BMI (Wang et al., 2023), insomnia (Thorkildsen et al., 2023), and lifetime smoking (Zhu et al., 2023), we employed PhenoScanner V2. The website http//www.phenoscanner.medschl.cam.ac.uk/ was accessed on 1 November 2023. If necessary, we assessed the correlation between exposure and outcome after excluding these SNPS from the MR analysis to mitigate potential pleiotropic effects.

Multivariable MR was employed to assess whether there was bias in any phenotype due to pleiotropy as identified on PhenoScanner. IEU OpenGWAS project (https://gwas.mrcieu.ac.uk/.) provided genetic diversity for potential pleiotropic traits. We performed a multivariable MR analysis to investigate the effect of Zn on the likelihood of severe sepsis-related death within 28 days. This analysis included mean corpuscular hemoglobin concentration (GWAS identifier: ebi-a-GCST90002328), reticulocyte count (GWAS identifier: ebi-a-GCST90025972), high light scatter reticulocyte count (GWAS identifier: ebi-a-GCST90025970), reticulocyte fraction of red cells (GWAS identifier: ebi-a-GCST90002406), and mean corpuscular volume (GWAS identifier: ebi-a-GCST90025963).

Power statistics

We conducted power calculations using the online platform (https://shiny.cnsgenomics.com/mRnd/) (Brion et al., 2013). Based on the sample sizes used in the meta-analysis, we computed the statistical power for each analysis under a type I error of 5%, and the results are summarized in Table 2. In order to guarantee the strength of our conclusions, we exclusively took into account micronutrients that had an R^2 value higher than 1% and/or a statistical power exceeding 50% for at least one sepsis-related outcome (Flatby et al., 2023). This criterion led to the exclusion of copper, folate, selenium, and vitamin B12 from further analysis, as detailed in Supplementary Tables S3, S4.

Replication and reverse MR analysis

To perform the primary MR analysis, we carried out a replication analysis by utilizing supplementary sepsis summary data obtained from the FinnGen consortium. The findings from the two groups (UK Biobank and FinnGen) were combined and analyzed using a f random-effects model in METAL (version 2011-03-25) (Willer et al., 2010). Additionally, to further assess whether our MR study was affected by reverse causation, we performed a reverse MR analysis on the association between genetically predicted sepsis and candidate micronutrients. In this reverse MR analysis, susceptibility and severity of sepsis were treated as exposures, and candidate micronutrients were considered as outcomes. We applied the identical rigorous standards for selecting instrumental variables, requiring a significance level of $p < 5 \times 10^{-8}$, and ensuring linkage disequilibrium with $r^2 < 0.001$ within windows of 10,000-kb.

Statistical analysis

The TwoSampleMR package (version 0.5.6) and the R package "MRPRESSO" (version 4.0.3) were utilized for conducting all MR analyses. METAL (version 2011-03-25) (Willer et al., 2010) was utilized for meta-analysis of results. A significance threshold of $p < 0.05$ was deemed to be of nominal importance, whereas the Bonferroni-adjusted statistical significance threshold (for 9 exposures) was established at $p = 0.05/9 = 5.56 \times 10^{-3}$.

Results

Instrumental variable selection

The number of instrumental variables for circulating micronutrients ranged from 2 to 11. The F-statistics for these

TABLE 2 Source of exposure genome-wide association study summary data.

Exposure	Number of SNPs	% Of variance explained	Population ancestry	Pubmed ID
Ca	7	0.841	European	24,068,962
Cu	2	4.6	European	23,720,494
Fe	3	3.04	European	25,352,340
Mg	5	1.49	European	20,700,443
P	5	1.2	European	20,558,539
Se	7	3.65	European	25,343,990
Folate	2	0.41	European	23,754,956
β-carotene	4	8.36	European	19,185,284
Vitamin B6	2	3.07	European	19,744,961
Vitamin B12	10	4.78	European	23,754,956
Vitamin C	11	1.79	European	33,203,707
Vitamin D	6	2.67	European	29,343,764
Zn	2	4.59	European	23,720,494

Abbreviations: Ca, Calcium; Cu, Copper; Fe, Iron; Mg, Magnesium; P, phosphorus; Se, Selenium; Zn, Zinc.

SNPs ranged from 25 to 1,497, with a median of 50, surpassing the conventional threshold of 10, suggesting a minimal likelihood of weak instrumentality (see Supplementary Table S4). However, it is noteworthy that one SNP for calcium (rs1550532), two SNPs for magnesium (rs7965584, rs7197653), and one SNP for vitamin D (rs8018720) were unavailable in the outcome datasets for streptococcal septicaemia and puerperal sepsis. Furthermore, the outcome dataset for severe sepsis-related death within 28 days did not include one SNP related to Fe (rs1525892), and two SNPs associated with phosphorus (rs1697421, rs9469578) were excluded due to incompatible alleles.

Primary MR analyses

Using the chosen instrumental variables, we performed an initial evaluation on the association between 9 circulating micronutrients and sepsis likelihood in the UK Biobank discovery set. We identified two independent associations with sepsis for β-carotene and Fe ($p < 5.56 \times 10^{-3}$) (Supplementary Table S5). No heterogeneity was found in the sensitivity analyses, which included Cochran's Q test and I^2 values (β-carotene $p = 0.997$, Fe $p = 0.264$). Evaluation of horizontal pleiotropy using MR-Egger suggested insufficient evidence for horizontal pleiotropy (Supplementary Table S6). In the replication set, no significant associations between micronutrients and sepsis were detected, with the relationship between β-carotene and sepsis in the opposite direction compared to the discovery set ($p = 0.119$, IVW). Following the meta-analysis, a nominal significant correlation was found solely between Fe and the susceptibility to sepsis infection (odds ratio [OR] = 1.083; 95% confidence interval [CI]: 1.00, 1.17; $p = 0.048$) (Figure 2; Supplementary Table S5).

The risk of sepsis showed no significant correlation with the levels of calcium, β-carotene, magnesium, phosphorus, vitamin B6, vitamin C, vitamin D, and Zn in the bloodstream (Figure 2; Supplementary Table S5).

Secondary MR analyses

In subgroup analyses, we observed associations between three micronutrients and three sepsis-related outcomes. As shown in Figure 3; Supplementary Table S1, we found a nominal significant negative relationship between β-carotene and the likelihood of sepsis death within 28 days (OR = 0.781; 95% CI: 0.611, 0.997; $p = 0.047$, IVW) and severe sepsis-related death within 28 days (OR = 0.449; 95% CI: 0.253, 0.799; $p = 6.48 \times 10^{-3}$, IVW). The sensitivity analyses (Supplementary Tables S7, S8), which involved the use of Cochran's Q test and I^2 values, indicated the absence of heterogeneity. Additionally, the MR-Egger analysis, with a small intercept, showed minimal influence of horizontal pleiotropy. Likewise, the MR-PRESSO examination did not detect any unusual SNPs or horizontal pleiotropy impacts on sepsis death within 28 days ($p = 0.958$) or severe sepsis-related death within 28 days ($p = 0.64$) (Supplementary Table S9). However, the MR-Egger method shows a direction opposite to IVW, and it did not pass our stringent significance threshold. Simultaneously, we also observed a nominal significant negative association between genetically predicted vitamin C and a reduced risk of puerperal sepsis (OR = 0.702; 95% CI: 0.507, 0.971; $p = 0.032$, IVW). Although sensitivity analysis found no evidence of heterogeneity or pleiotropy, it still did not meet our stringent statistical threshold.

In contrast, there was a strong correlation between the level of Zn and a decreased likelihood of severe sepsis-related death within 28 days (OR = 0.450; 95% CI: 0.263, 0.770; $p = 3.58 \times 10^{-3}$, IVW). Supplementary Table S8 presents the outcomes of sensitivity analyses. Cochran's Q test indicated no heterogeneity ($p = 0.339$); however, the restricted number of accessible SNPs (only 2) prevented the execution of MR-PRESSO and Egger regression analyses. Moreover, we employed the Steiger test to validate whether the identified causal relationships were influenced by reverse causation. The results of Steiger do not support the existence of reverse causal effects between candidate

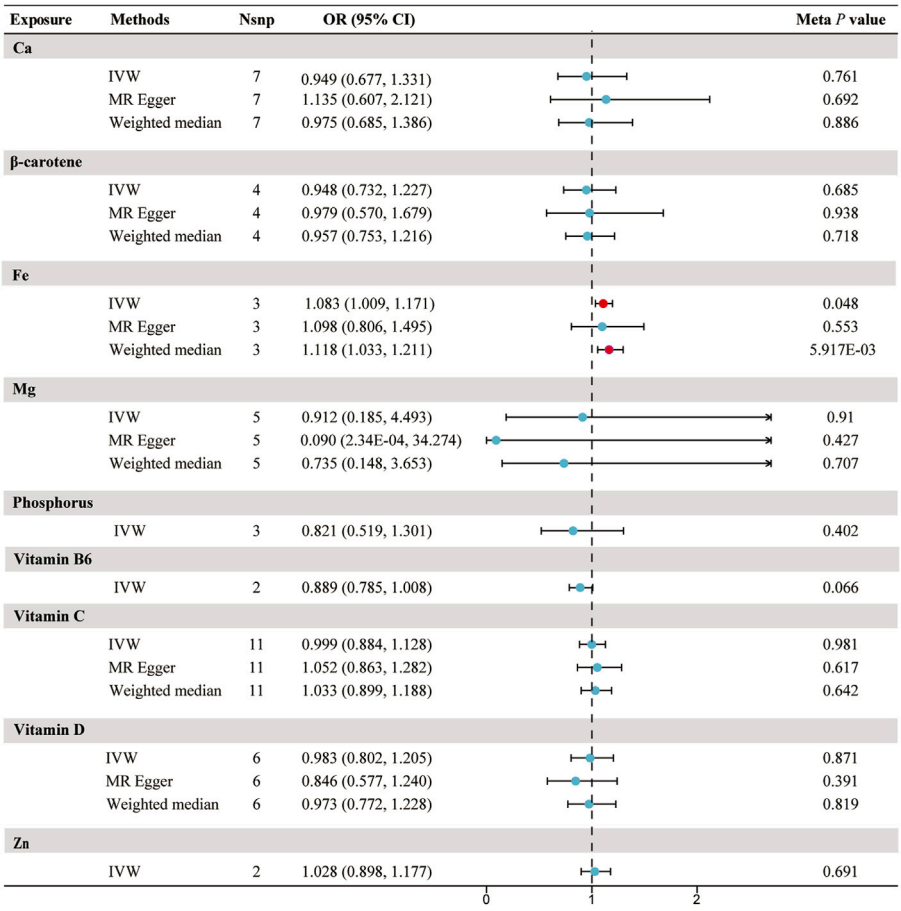


FIGURE 2
Forest plot for the meta-analysis of circulating micronutrients levels on the risk of sepsis. Abbreviations: IVW, inverse-variance weighted; Nsnp, number of SNP; OR, odds ratio; CI, confidence interval; Ca, Calcium; Fe, iron, Mg, Magnesium; Zn, zinc.

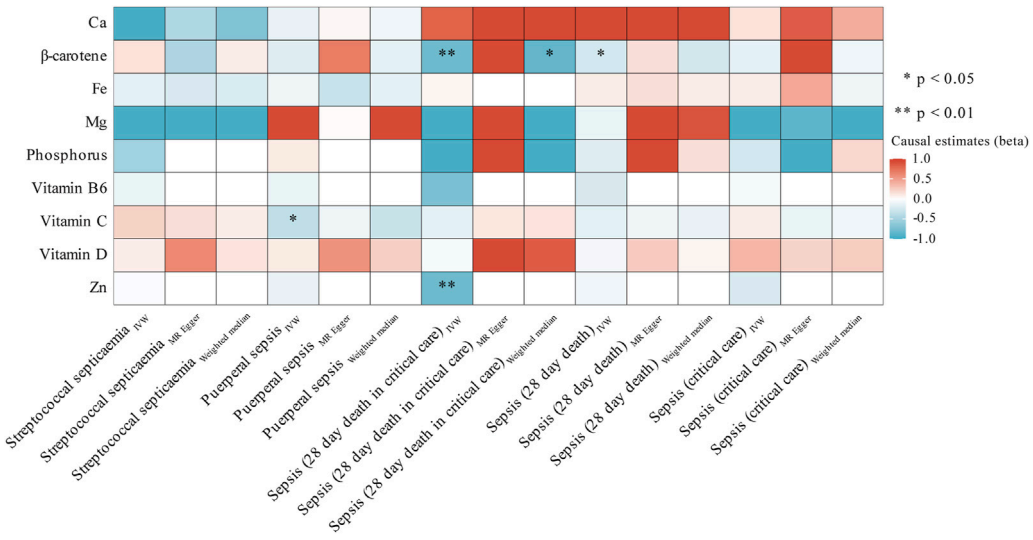


FIGURE 3
Heatmap showing the causal effects of circulating micronutrients levels on the risk of sepsis-related outcomes by using three method (IVW, MR egger, and weighted median). Abbreviations: Ca, Calcium; Fe, iron, Mg, Magnesium; Zn, zinc. IVW, inverse-variance weighted.

TABLE 3 Estimated causal effects of zinc on Sepsis (28 day death in critical care) by the multivariable Mendelian randomization analysis.

Exposure	Nsnp	Multivariable MR	
		OR (95% CI)	*P*
Mean corpuscular hemoglobin concentration	227	1.537 (0.601, 3.932)	0.370
Reticulocyte fraction of red cells	227	0.335 (0.010, 11.218)	0.541
Mean corpuscular volume	227	0.921 (0.514, 1.652)	0.783
High light scatter reticulocyte count	227	6.106 (0.839, 44.454)	0.074
Reticulocyte count	227	0.462 (0.011, 19.751)	0.687
Zinc	227	0.675 (0.508, 0.895)	6.421E-03

Abbreviations: OR, odds ratio; CI, confidence interval, Nsnp, number of SNP.

micronutrient (β-carotene, vitamin C, and Zn) and sepsis-related outcomes (Supplementary Table S9).

Given the limited availability of only two SNPs for the micronutrient Zn within the threshold of $p < 5 \times 10^{-8}$, which precluded heterogeneity and pleiotropy tests, we conducted a *post hoc* MR analysis. In this analysis, we included variants at a more liberal threshold ($p < 5 \times 10^{-7}$). As shown in Supplementary Tables S2, S10, the results aligned closely with our initial findings (OR = 0.483; 95% CI: 0.298, 0.783; $p = 3.16 \times 10^{-3}$, IVW), and the *p*-values for heterogeneity and pleiotropy were both above 0.05, suggesting a minimal likelihood of heterogeneity and pleiotropy.

Confounder and multivariable MR analyses

Although sensitivity analyses did not uncover any indications of bias that would make the MR estimates unreliable, we proceeded to examine the second characteristic linked to the leading SNP for Zn by utilizing the PhenoScanner tool [blood metabolites (Wei et al., 2023), BMI (Wang et al., 2023), insomnia (Thorkildsen et al., 2023), and lifetime smoking (Zhu et al., 2023)]. However, no connections were observed between Zn-related instrumental variables and reported risk factors (Supplementary Table S11). Nonetheless, it is noteworthy that Zn's rs1532423 was closely associated with mean corpuscular hemoglobin concentration (GWAS identifier: ebi-a-GCST90002328), reticulocyte count (GWAS identifier: ebi-a-GCST90025972), high light scatter reticulocyte count (GWAS identifier: ebi-a-GCST90025970), reticulocyte fraction of red cells (GWAS identifier: ebi-a-GCST90002406), and mean corpuscular volume (GWAS identifier: ebi-a-GCST90025963). Hence, we performed a multivariable MR analysis to examine the association between Zn and potential pleiotropic traits and the likelihood of severe sepsis-related death within 28 days.

After accounting for the impacts of mean corpuscular hemoglobin concentration, reticulocyte count, high light scatter reticulocyte count, reticulocyte fraction of red cells, and mean corpuscular volume in the multivariable MR analysis, we found comparable effects to the primary analysis, suggesting that Zn continued to exhibit a defensive influence on severe sepsis-related death within 28 days (OR = 0.675; 95% CI: 0.508, 0.895; $p = 6.421 \times 10^{-3}$, IVW) (Table 3).

Reverse MR analyses

In order to further explore the causal efficacy connection between potential micronutrients and outcomes related to sepsis, we performed reverse causal analyses by employing instrumental variables for sepsis-related outcomes. Our objective was to use IVW-MR estimates and select significant independent SNPs with a $p < 5 \times 10^{-8}$ as instrumental variables. This analysis aimed to investigate whether there was any indication of a reverse causal association between the identified Zn and the outcome of severe sepsis-related death within 28 days. As there were no significant independent SNPs identified when considering severe sepsis-related death within 28 days as the variable at a significance level of $p < 5 \times 10^{-8}$ or $p < 5 \times 10^{-7}$, we adjusted the criteria to $p < 5 \times 10^{-6}$ and included linkage disequilibrium with $r^2 < 0.001$ within 10,000-kb windows. Nevertheless, our examination revealed restricted backing for this inverse causal connection (beta = −0.014; 95% CI: −0.115, 0.087; $p = 0.783$, IVW), as specified in Supplementary Table S12.

Discussion

Drawing on our current knowledge, this study represents the initial comprehensive examination of causal connections between various circulating micronutrients in the blood and the susceptibility, severity, and subtype-specific risks of sepsis. Our research findings indicate one strong correlation and four suggestive associations among four micronutrients and sepsis-related outcomes. In particular, our main finding suggests a strong causal effect connection between genetically forecasted Zn levels in the bloods and a decreased risk of severe sepsis-related death within 28 days. Additionally, four suggestive associations were identified: elevated blood Fe levels indicating a potential link to increased susceptibility to sepsis, higher blood β-carotene levels suggestively associated with decreased risk of severe sepsis-related death within 28 days and sepsis-related death within 28 days, and a suggestively correlation between vitamin C and decreased risk of postpartum sepsis. There is no apparent association between the other five circulating micronutrients and sepsis or related outcomes.

Sepsis is an illness resulting from an infection, which causes dysfunction of organs and ultimately leads to death. Sepsis is a significant worldwide contributor to death, causing approximately 6 million fatalities each year (Evans et al., 2021). Timely

identification and medical intervention are vital for individuals who might be susceptible to septicemia. In recent years, the role of circulating micronutrients in blood in disease has garnered increasing attention. Zn, a micronutrient, has been demonstrated to be a vital metal ion for a well-functioning immune system. In the human body, it has a vital function in cellular differentiation, proliferation, and apoptosis mechanisms. According to Gammoh's research (Gammoh and Rink, 2017), Zn has the ability to control the discharge of inflammatory substances, the production of coenzymes, and the operation of T helper cells, B cells, neutrophils, natural killer cells, and macrophages.

Previous studies have found an association between Zn deficiency and compromised immune function, as well as adverse disease outcomes (Overbeck et al., 2008). Low-dose Zn supplementation has been shown to effectively treat respiratory infections and childhood diarrhea (Dhingra et al., 2020). However, the connection between Zn levels in blood and the risk of human sepsis infection has not been clearly established. Research has observed significantly lower serum Zn concentrations in ICU sepsis patients compared to healthy controls (Hoeger et al., 2017). However, a randomized controlled trial found no notable distinctions between the group that received Zn supplementation and the control group among sepsis patients (Mehta et al., 2013), and it even indicated potential adverse consequences (Braunschweig et al., 1997). The uncertain outcomes could be impacted by methodological deficiencies like limited sample sizes or remaining confounding factors. From a genetic standpoint, our MR study presents proof that genetically anticipated Zn concentrations in the bloodstream offer a safeguarding influence on the severe sepsis-related death within 28 days, despite the fact that the influence of Zn on sepsis and severe sepsis is limited. An animal experiment has identified a potential mechanism behind these findings: Zn can modulate host immune defense by blocking the IKK complex and inducing inhibition of the NF-κB pathway downstream of MAPK (Liu et al., 2013). However, supplementing Zn during infection needs to consider the risk of creating a Zn microenvironment favorable to pathogen growth while interfering with the innate immune system's ability to chelate free Zn. Our study may suggest that, although Zn is not associated with the risk of severe sepsis, it supports clinical practices of Zn supplementation in critically ill sepsis patients at high risk of mortality. Nevertheless, additional medical investigation is required to authenticate these discoveries.

We additionally discovered a slight positive correlation between genetically anticipated blood Fe levels and the susceptibility of sepsis. Consistent with previous observational studies on sepsis, septic patients had higher serum iron levels compared to healthy volunteers (Akkas et al., 2020). Fe is an essential element in various physiological processes, and deviations in Fe status (such as Fe deficiency or Fe overload) can significantly impact health. Fe status deviation exhibits noticeable gender differences, with females being more prone to Fe deficiency. Prior observational studies have suggested associations between both iron deficiency (Mohus et al., 2018) and high iron status (Brandtner et al., 2020) with an increased risk of infection. According to a recent study using magnetic resonance imaging (MRI), it was found that the addition of Fe is not likely to greatly raise the chances of infection (Butler-Laporte et al., 2023). Conversely, a separate study using the same method indicated a positive association between the predicted increase in serum Fe levels and an elevated risk of sepsis (Mohus et al., 2022). Unfortunately, the latter study only set the significance threshold at $p < 0.05$ without rigorous correction for multiple testing. It is worth noting that the observed associations do not imply a strong causal relationship, and the impact of Fe on sepsis appears to be relatively mild. Future research endeavors should explore this relationship further, conducting rigorous analyses to validate these findings.

Regarding β-carotene, long considered potent antioxidants within the organism, prior epidemiological studies have proposed a negative correlation between carotenoid intake and cancer incidence (Koklesova et al., 2020). Moreover, cancer patients exhibit a significant increase in carotenoid concentrations after anti-tumor treatment (McMillan et al., 2000). A recent MR study found a protective effect of blood β-carotene against type 2 diabetes (Chen et al., 2023).These findings collectively support the beneficial role of β-carotene in disease occurrence. In our research, we noticed a slight adverse correlation between β-carotene and the likelihood of severe sepsis as well as the mortality rate within 28 days for individuals with severe sepsis. Nevertheless, this discovery could be fortuitous as a result of conducting numerous tests and did not meet our rigorous statistical significance criteria.

Likewise, we observed a suggestive causal effect of higher circulating levels of vitamin C in reducing the risk of puerperal sepsis. Previous observational studies have noted significantly decreased average vitamin C levels in sepsis patients (Carr et al., 2017), prompting discussions on the potential therapeutic role of vitamin C as a crucial antioxidant in sepsis management. A recent review, considering findings from conducted randomized controlled trials, reported positive effects of vitamin C on reducing sepsis mortality in only 2 out of 11 projects (Ammar et al., 2021). Some studies suggest that vitamin C, compared to a placebo, may contribute to mitigating inflammation induced by severe sepsis (Fowler et al., 2014). However, a recent randomized controlled trial yielded inconsistent results, showing that vitamin C did not significantly improve sepsis-related inflammation and vascular damage (Fowler et al., 2019). The heterogeneity in vitamin C treatment regimens, initiation times, and duration of therapy has led to significant variability in results across observational studies. Our MR study did not find causal effects of vitamin C on susceptibility and severity of sepsis at genetic level. However, it revealed a mild protective effect of vitamin C specifically in one subtype of sepsis—postpartum sepsis. Nevertheless, this finding did not meet our stringent statistical thresholds, and given the limited observational studies on postpartum sepsis to date, larger-scale research is needed to further explore this relationship in the future.

Surprisingly, there were no connections discovered between genetically anticipated levels of calcium, magnesium, phosphorus, vitamin D, and vitamin B6 in the bloodstream and the likelihood of sepsis-related consequences. Contrary to a meta-analysis suggesting that vitamin D deficiency increases susceptibility and mortality in sepsis (de Haan et al., 2014). Our study does not support this viewpoint. Additionally, a prospective cohort study found insufficient evidence for vitamin D in predicting sepsis and mortality rates (Ratzinger et al., 2017). A review concluded that there is no clear evidence that selenium supplementation can prevent infection and new infection rates (Zhao et al., 2019). This may suggest that these micronutrients are not crucial risk factors for the development of sepsis and its related outcomes.

Our study has several strengths. This is the initial comprehensive study of MR that examines the association between the levels of 9 micronutrients and the risk of sepsis and its related outcomes. This helps reduce the influence of confounding factors present in observational studies. Furthermore, our examination was limited to people of European origin, reducing the occurrence of population stratification errors. Thirdly, the meta-analysis of summary data from multiple sepsis cohorts reduced random errors, enhancing credibility. However, the study also has limitations. First, some micronutrients' instrumental variables exhibited varying degrees of low statistical power. Despite all instrumental variables having F-values greater than 10, suggesting a low probability of weak instrument bias, there is still a possibility of some bias remaining. To enhance statistical power, it is imperative to conduct future GWAS on traits related to micronutrients at a larger scale. Second, the study population limited to individuals of European descent may hinder generalization to a broader population. Third, due to sparse subtype data, we could not replicate the associations between Zn and susceptibility or severity subtypes of sepsis. However, our findings suggest a protective trend of Zn against all five sepsis-related outcomes. Larger-scale clinical studies are needed to further confirm these findings and explore underlying mechanisms.

Conclusion

To summarize, our research indicates that Zn might have a safeguarding effect in decreasing the likelihood of death within 28 days for patients with severe sepsis, endorsing the medical recommendation of providing Zn supplements to patients who face a high risk of mortality due to severe sepsis. This provides new insights for further research into the role of micronutrients in the prevention and treatment of sepsis.

Scope statement

Previous research has paid limited attention to the role of trace micronutrients in the pathogenesis of sepsis. Therefore, we employed Mendelian randomization analysis to comprehensively investigate the causal relationship between the levels of nine micronutrients (including calcium, β-carotene, iron, magnesium, phosphorus, vitamin C, vitamin B6, vitamin D, and zinc) and susceptibility, severity, and subtypes of sepsis. Our research indicates that zinc might have a safeguarding effect in decreasing the likelihood of death within 28 days for patients with severe sepsis, endorsing the medical recommendation of providing zinc supplements to patients who face a high risk of mortality due to severe sepsis. This provides new insights for further research into the role of micronutrients in the prevention and treatment of sepsis.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

Author contributions

ZW: Conceptualization, Methodology, Writing–original draft, Writing–review and editing. YL: Investigation, Writing–review and editing. XM: Investigation, Validation, Writing–review and editing. JZ: Software, Writing–review and editing. FH: Writing–review and editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Acknowledgments

The authors sincerely thank all the participants and investigators of the UK Biobank and FinnGen study.

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2024.1353118/full#supplementary-material

References

Akkas, I., Ince, N., and Sungur, M. A. (2020). Serum trace element and heavy metal levels in patients with sepsis. *Aging Male* 23, 222–226. doi:10.1080/13685538.2020.1740200

Ammar, M. A., Ammar, A. A., Condeni, M. S., and Bell, C. M. (2021). Vitamin C for sepsis and septic shock. *Am. J. Ther.* 28, e649–e679. doi:10.1097/MJT.0000000000001423

Aschard, H., Vilhjalmsson, B. J., Joshi, A. D., Price, A. L., and Kraft, P. (2015). Adjusting for heritable covariates can bias effect estimates in genome-wide association studies. *Am. J. Hum. Genet.* 96, 329–339. doi:10.1016/j.ajhg.2014.12.021

Benyamin, B., Esko, T., Ried, J. S., Radhakrishnan, A., Vermeulen, S. H., Traglia, M., et al. (2014). Novel loci affecting iron homeostasis and their effects in

individuals at risk for hemochromatosis. *Nat. Commun.* 5, 4926. doi:10.1038/ncomms5926

Bowden, J., Davey Smith, G., and Burgess, S. (2015). Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int. J. Epidemiol.* 44, 512–525. doi:10.1093/ije/dyv080

Brandtner, A., Tymoszuk, P., Nairz, M., Lehner, G. F., Fritsche, G., Vales, A., et al. (2020). Linkage of alterations in systemic iron homeostasis to patients' outcome in sepsis: a prospective study. *J. Intensive Care* 8, 76. doi:10.1186/s40560-020-00495-8

Braunschweig, C. L., Sowers, M., Kovacevich, D. S., Hill, G. M., and August, D. A. (1997). Parenteral zinc supplementation in adult humans during the acute phase response increases the febrile response. *J. Nutr.* 127, 70–74. doi:10.1093/jn/127.1.70

Brion, M. J., Shakhbazov, K., and Visscher, P. M. (2013). Calculating statistical power in Mendelian randomization studies. *Int. J. Epidemiol.* 42, 1497–1501. doi:10.1093/ije/dyt179

Butler-Laporte, G., Farjoun, Y., Chen, Y., Hultstrom, M., Liang, K. Y. H., Nakanishi, T., et al. (2023). Increasing serum iron levels and their role in the risk of infectious diseases: a Mendelian randomization approach. *Int. J. Epidemiol.* 52, 1163–1174. doi:10.1093/ije/dyad010

Bycroft, C., Freeman, C., Petkova, D., Band, G., Elliott, L. T., Sharp, K., et al. (2018). The UK Biobank resource with deep phenotyping and genomic data. *Nature* 562, 203–209. doi:10.1038/s41586-018-0579-z

Carr, A. C., Rosengrave, P. C., Bayer, S., Chambers, S., Mehrtens, J., and Shaw, G. M. (2017). Hypovitaminosis C and vitamin C deficiency in critically ill patients despite recommended enteral and parenteral intakes. *Crit. Care* 21, 300. doi:10.1186/s13054-017-1891-y

Cecconi, M., Evans, L., Levy, M., and Rhodes, A. (2018). Sepsis and septic shock. *Lancet* 392, 75–87. doi:10.1016/S0140-6736(18)30696-2

Chen, L., Fan, Z., Sun, X., Qiu, W., Mu, W., Chai, K., et al. (2023). Diet-derived antioxidants and nonalcoholic fatty liver disease: a Mendelian randomization study. *Hepatol. Int.* 17, 326–338. doi:10.1007/s12072-022-10443-3

Cornelis, M. C., Fornage, M., Foy, M., Xun, P., Gladyshev, V. N., Morris, S., et al. (2015). Genome-wide association study of selenium concentrations. *Hum. Mol. Genet.* 24, 1469–1477. doi:10.1093/hmg/ddu546

Dashti, H. S., Shea, M. K., Smith, C. E., Tanaka, T., Hruby, A., Richardson, K., et al. (2014). Meta-analysis of genome-wide association studies for circulating phylloquinone concentrations. *Am. J. Clin. Nutr.* 100, 1462–1469. doi:10.3945/ajcn.114.093146

De Haan, K., Groeneveld, A. B., De, GEUS, Egal, M., and Struijs, A. (2014). Vitamin D deficiency as a risk factor for infection, sepsis and mortality in the critically ill: systematic review and meta-analysis. *Crit. Care* 18, 660. doi:10.1186/s13054-014-0660-4

Dhingra, U., Kisenge, R., Sudfeld, C. R., Dhingra, P., Somji, S., Dutta, A., et al. (2020). Lower-dose zinc for childhood diarrhea - a randomized, multicenter trial. *N. Engl. J. Med.* 383, 1231–1241. doi:10.1056/NEJMoa1915905

Evans, D. M., Zhu, G., Dy, V., Heath, A. C., Madden, P. A., Kemp, J. P., et al. (2013). Genome-wide association study identifies loci affecting blood copper, selenium and zinc. *Hum. Mol. Genet.* 22, 3998–4006. doi:10.1093/hmg/ddt239

Evans, L., Rhodes, A., Alhazzani, W., Antonelli, M., Coopersmith, C. M., French, C., et al. (2021). Surviving sepsis campaign: international guidelines for management of sepsis and septic shock 2021. *Crit. Care Med.* 49, e1063–e1143. doi:10.1097/CCM.0000000000005337

Ferrucci, L., Perry, J. R., Matteini, A., Perola, M., Tanaka, T., Silander, K., et al. (2009). Common variation in the beta-carotene 15,15'-monooxygenase 1 gene affects circulating levels of carotenoids: a genome-wide association study. *Am. J. Hum. Genet.* 84, 123–133. doi:10.1016/j.ajhg.2008.12.019

Flatby, H. M., Ravi, A., Damas, J. K., Solligard, E., and Rogne, T. (2023). Circulating levels of micronutrients and risk of infections: a Mendelian randomization study. *BMC Med.* 21, 84. doi:10.1186/s12916-023-02780-3

Fowler, A. A., Syed, A. A., Knowlson, S., Sculthorpe, R., Farthing, D., Dewilde, C., et al. (2014). Phase I safety trial of intravenous ascorbic acid in patients with severe sepsis. *J. Transl. Med.* 12, 32. doi:10.1186/1479-5876-12-32

Fowler, A. A., Truwit, J. D., Hite, R. D., Morris, P. E., Dewilde, C., Priday, A., et al. (2019). Effect of vitamin C infusion on organ failure and biomarkers of inflammation and vascular injury in patients with sepsis and severe acute respiratory failure: the CITRIS-ALI randomized clinical trial. *JAMA* 322, 1261–1270. doi:10.1001/jama.2019.11825

Gammoh, N. Z., and Rink, L. (2017). Zinc in infection and inflammation. *Nutrients* 9, 624. doi:10.3390/nu9060624

Gombart, A. F., Pierre, A., and Maggini, S. (2020). A review of micronutrients and the immune system-working in harmony to reduce the risk of infection. *Nutrients* 12, 236. doi:10.3390/nu12010236

Grarup, N., Sulem, P., Sandholt, C. H., Thorleifsson, G., Ahluwalia, T. S., Steinthorsdottir, V., et al. (2013). Genetic architecture of vitamin B12 and folate levels uncovered applying deeply sequenced large datasets. *PLoS Genet.* 9, e1003530. doi:10.1371/journal.pgen.1003530

Greco, M. F., Minelli, C., Sheehan, N. A., and Thompson, J. R. (2015). Detecting pleiotropy in Mendelian randomisation studies with summary data and a continuous outcome. *Stat. Med.* 34, 2926–2940. doi:10.1002/sim.6522

Guo, R., Duan, J., Pan, S., Cheng, F., Qiao, Y., Feng, Q., et al. (2023). The road from AKI to CKD: molecular mechanisms and therapeutic targets of ferroptosis. *Cell Death Dis.* 14, 426. doi:10.1038/s41419-023-05969-9

Hamilton, F. W., Thomas, M., Arnold, D., Palmer, T., Moran, E., Mentzer, A. J., et al. (2023). Therapeutic potential of IL6R blockade for the treatment of sepsis and sepsis-related death: a Mendelian randomisation study. *PLoS Med.* 20, e1004174. doi:10.1371/journal.pmed.1004174

Hazra, A., Kraft, P., Lazarus, R., Chen, C., Chanock, S. J., Jacques, P., et al. (2009). Genome-wide significant predictors of metabolites in the one-carbon metabolism pathway. *Hum. Mol. Genet.* 18, 4677–4687. doi:10.1093/hmg/ddp428

Hemani, G., Tilling, K., and Davey Smith, G. (2017). Orienting the causal relationship between imprecisely measured traits using GWAS summary data. *PLoS Genet.* 13, e1007081. doi:10.1371/journal.pgen.1007081

Hoeger, J., Simon, T. P., Beeker, T., Marx, G., Haase, H., and Schuerholz, T. (2017). Persistent low serum zinc is associated with recurrent sepsis in critically ill patients - a pilot study. *PLoS One* 12, e0176069. doi:10.1371/journal.pone.0176069

Jiang, X., O'Reilly, P. F., Aschard, H., Hsu, Y. H., Richards, J. B., Dupuis, J., et al. (2018). Genome-wide association study in 79,366 European-ancestry individuals informs the genetic architecture of 25-hydroxyvitamin D levels. *Nat. Commun.* 9, 260. doi:10.1038/s41467-017-02662-2

Kestenbaum, B., Glazer, N. L., Kottgen, A., Felix, J. F., Hwang, S. J., Liu, Y., et al. (2010). Common genetic variants associate with serum phosphorus concentration. *J. Am. Soc. Nephrol.* 21, 1223–1232. doi:10.1681/ASN.2009111104

Koklesova, L., Liskova, A., Samec, M., Buhrmann, C., Samuel, S. M., Varghese, E., et al. (2020). Carotenoids in cancer apoptosis-the road from bench to bedside and back. *Basel Cancers* 12. doi:10.3390/cancers12092425

Kurki, M. I., Karjalainen, J., Palta, P., Sipila, T. P., Kristiansson, K., Donner, K. M., et al. (2023). FinnGen provides genetic insights from a well-phenotyped isolated population. *Nature* 613, 508–518. doi:10.1038/s41586-022-05473-8

Liu, M. J., Bao, S., Galvez-Peralta, M., Pyle, C. J., Rudawsky, A. C., Pavlovicz, R. E., et al. (2013). ZIP8 regulates host defense through zinc-mediated inhibition of NF-κB. *Cell Rep.* 3, 386–400. doi:10.1016/j.celrep.2013.01.009

Major, J. M., Yu, K., Wheeler, W., Zhang, H., Cornelis, M. C., Wright, M. E., et al. (2011). Genome-wide association study identifies common variants associated with circulating vitamin E levels. *Hum. Mol. Genet.* 20, 3876–3883. doi:10.1093/hmg/ddr296

Mcmillan, D. C., Sattar, N., Talwar, D., O'Reilly, D. S., and Mcardle, C. S. (2000). Changes in micronutrient concentrations following anti-inflammatory treatment in patients with gastrointestinal cancer. *Nutrition* 16, 425–428. doi:10.1016/s0899-9007(00)00270-7

Mehta, K., Bhatta, N. K., Majhi, S., Shrivastava, M. K., and Singh, R. R. (2013). Oral zinc supplementation for reducing mortality in probable neonatal sepsis: a double blind randomized placebo controlled trial. *Indian Pediatr.* 50, 390–393. doi:10.1007/s13312-013-0120-2

Meyer, T. E., Verwoert, G. C., Hwang, S. J., Glazer, N. L., Smith, A. V., Van Rooij, F. J., et al. (2010). Genome-wide association studies of serum magnesium, potassium, and sodium concentrations identify six Loci influencing serum magnesium levels. *PLoS Genet.* 6, e1001045. doi:10.1371/journal.pgen.1001045

Mohus, R. M., Flatby, H., Liyanarachi, K. V., Dewan, A. T., Solligard, E., Damas, J. K., et al. (2022). Iron status and the risk of sepsis and severe COVID-19: a two-sample Mendelian randomization study. *Sci. Rep.* 12, 16157. doi:10.1038/s41598-022-20679-6

Mohus, R. M., Paulsen, J., Gustad, L., Askim, A., Mehl, A., Dewan, A. T., et al. (2018). Association of iron status with the risk of bloodstream infections: results from the prospective population-based HUNT Study in Norway. *Intensive Care Med.* 44, 1276–1283. doi:10.1007/s00134-018-5320-8

Mondul, A. M., Yu, K., Wheeler, W., Zhang, H., Weinstein, S. J., Major, J. M., et al. (2011). Genome-wide association study of circulating retinol levels. *Hum. Mol. Genet.* 20, 4724–4731. doi:10.1093/hmg/ddr387

Ng, E., Lind, P. M., Lindgren, C., Ingelsson, E., Mahajan, A., Morris, A., et al. (2015). Genome-wide association study of toxic metals and trace elements reveals novel associations. *Hum. Mol. Genet.* 24, 4739–4745. doi:10.1093/hmg/ddv190

O'Seaghdha, C. M., Wu, H., Yang, Q., Kapur, K., Guessous, I., Zuber, A. M., et al. (2013). Meta-analysis of genome-wide association studies identifies six new Loci for serum calcium concentrations. *PLoS Genet.* 9, e1003796. doi:10.1371/journal.pgen.1003796

Overbeck, S., Rink, L., and Haase, H. (2008). Modulating the immune response by oral zinc supplementation: a single approach for multiple diseases. *Arch. Immunol. Ther. Exp. Warsz.* 56, 15–30. doi:10.1007/s00005-008-0003-8

Pierce, B. L., and Burgess, S. (2013). Efficient design for Mendelian randomization studies: subsample and 2-sample instrumental variable estimators. *Am. J. Epidemiol.* 178, 1177–1184. doi:10.1093/aje/kwt084

Ratzinger, F., Haslacher, H., Stadlberger, M., Schmidt, R. L., Obermuller, M., Schmetterer, K. G., et al. (2017). 25(OH)D and 1,25(OH)D vitamin D fails to predict sepsis and mortality in a prospective cohort study. *Sci. Rep.* 7, 40646. doi:10.1038/srep40646

Rhee, C., Dantes, R., Epstein, L., Murphy, D. J., Seymour, C. W., Iwashyna, T. J., et al. (2017). Incidence and trends of sepsis in US hospitals using clinical vs claims data, 2009-2014. *JAMA* 318, 1241–1249. doi:10.1001/jama.2017.13836

Spoelstra-De Man, A. M. E., Elbers, P. W. G., and Oudemans-Van Straaten, H. M. (2018). Vitamin C: should we supplement? *Curr. Opin. Crit. Care* 24, 248–255. doi:10.1097/MCC.0000000000000510

Thorkildsen, M. S., Gustad, L. T., Mohus, R. M., Burgess, S., Nilsen, T. I. L., Damas, J. K., et al. (2023). Association of genetically predicted insomnia with risk of sepsis: a mendelian randomization study. *JAMA Psychiatry* 80, 1061–1065. doi:10.1001/jamapsychiatry.2023.2717

Verbanck, M., Chen, C. Y., Neale, B., and Do, R. (2018). Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. *Nat. Genet.* 50, 693–698. doi:10.1038/s41588-018-0099-7

Wang, J., Hu, Y., Zeng, J., Li, Q., He, L., Hao, W., et al. (2023). Exploring the causality between body mass index and sepsis: a two-sample mendelian randomization study. *Int. J. Public Health* 68, 1605548. doi:10.3389/ijph.2023.1605548

Wei, Z., Xiong, Q., Huang, D., Wu, Z., and Chen, Z. (2023). Causal relationship between blood metabolites and risk of five infections: a Mendelian randomization study. *BMC Infect. Dis.* 23, 663. doi:10.1186/s12879-023-08662-6

Willer, C. J., Li, Y., and Abecasis, G. R. (2010). METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* 26, 2190–2191. doi:10.1093/bioinformatics/btq340

Zhao, Y., Yang, M., Mao, Z., Yuan, R., Wang, L., Hu, X., et al. (2019). The clinical outcomes of selenium supplementation on critically ill patients: a meta-analysis of randomized controlled trials. *Med. Baltim.* 98, e15473. doi:10.1097/MD.0000000000015473

Zheng, J. S., Luan, J., Sofianopoulou, E., Imamura, F., Stewart, I. D., Day, F. R., et al. (2021). Plasma vitamin C and type 2 diabetes: genome-wide association study and mendelian randomization analysis in European populations. *Diabetes Care* 44, 98–106. doi:10.2337/dc20-1328

Zhu, H., Zhan, X., Wang, C., Deng, Y., Li, X., Song, L., et al. (2023). Causal associations between tobacco, alcohol use and risk of infectious diseases: a mendelian randomization study. *Infect. Dis. Ther.* 12, 965–977. doi:10.1007/s40121-023-00775-4

Zuccolo, L., and Holmes, M. V. (2017). Commentary: Mendelian randomization-inspired causal inference in the absence of genetic data. *Int. J. Epidemiol.* 46, 962–965. doi:10.1093/ije/dyw327

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 26 March 2024
DOI 10.3389/fgene.2024.1335223

OPEN ACCESS

EDITED BY
Xuanye Cao,
University of Texas MD Anderson Cancer Center, United States

REVIEWED BY
Xiuyu Cai,
Sun Yat-sen University Cancer Center (SYSUCC), China
Wei Gong,
Shanghai Jiao Tong University, China

*CORRESPONDENCE
Jun Gu,
jungu@ntu.edu.cn

†These authors have contributed equally to this work and share first authorship

RECEIVED 08 November 2023
ACCEPTED 14 March 2024
PUBLISHED 26 March 2024

CITATION
Ji T, Lv Y, Liu M, Han Y, Yuan B and Gu J (2024), Causal relationships between mitochondrial proteins and different pathological types of lung cancer: a bidirectional mendelian randomization study.
Front. Genet. 15:1335223.
doi: 10.3389/fgene.2024.1335223

Causal relationships between mitochondrial proteins and different pathological types of lung cancer: a bidirectional mendelian randomization study

Tanao Ji[1†], Yue Lv[2†], Meiqun Liu[3], Yujie Han[4], Baochang Yuan[4] and Jun Gu[4*]

[1]Department of General Practice, Affiliated Hospital of Nantong University, Medical School of Nantong University, Nantong, China, [2]Department of Hematology, Affiliated Hospital of Nantong University, Medical School of Nantong University, Nantong, China, [3]Department of Electrocardioeraphy, Qidong People's Hospital, Qidong Liver Cancer Institute, Affiliated Qidong Hospital of Nantong University, Nantong, China, [4]Department of Pulmonary and Critical Care Medicine, Affiliated Hospital of Nantong University, Medical School of Nantong University, Nantong Key Laboratory of Respiratory, Nantong, China

An increasing number of studies point to an association between mitochondrial proteins (MPs) and lung cancer (LC). However, the causal relationship between MPs and LC remains unclear. Consequently, our study employed a bidirectional Mendelian randomization (MR) analysis to explore the causal association between MPs and different pathological types of LC. A two-sample MR study was performed using the genome-wide association study (GWAS) data publicly available. We applied the primary inverse variance weighted (IVW) method along with additional MR methods to validate the causality between MPs and different pathological types of LC. To ensure the robustness of our findings, sensitivity analyses were employed. Moreover, we performed a bi-directional MR analysis to determine the direction of the causal association. We identified a total of seven MPs had significant causal relationships on overall LC, lung squamous cell carcinoma (LUSC), and small cell lung carcinoma (SCLC). We found two MPs had significant associations with overall LC, four MPs had significant associations with LUSC, and four MPs had significant associations with SCLC. Additionally, an MP was found to have a nominal relationship with LUSC. Moreover, no causality was found between MPs and lung adenocarcinoma (LUAD). Bidirectional MR showed no reverse effect between identified MPs and different pathological types of LC. In general, our findings of this MR study suggest causal associations of specific MPs with overall LC, LUSC, and SCLC. However, no such causality was found in LUAD.

KEYWORDS

lung cancer, mitochondrial protein, Mendelian randomization, causal relationship, european

1 Introduction

Cancer stands as one of the primary causes of global mortality. On a worldwide scale, in 2020, it is estimated that approximately 2.2 million new cases of lung cancer (LC) and nearly 1.8 million LC deaths occurred. Meanwhile, LC ranked as the most commonly diagnosed cancer among males and the third most frequently diagnosed cancer among females (Sung et al., 2021). The incidence of new LC cases is projected to rise until 2035 in most countries, which causes a substantial global public health challenge (Luo et al., 2023). The majority of LC patients are diagnosed at an advanced stage of the disease, resulting in a 5-year survival rate of less than 20% (Osuoha et al., 2018; Bade and Dela Cruz, 2020). Hence, it is crucial to identify modifiable protective or risk factors to prevent the occurrence and progression of LC. Smoking is the most established and well-acknowledged risk factor for LC (Leiter et al., 2023). However, as smoking prevalence decreases and the number of LC cases in nonsmokers rises, it becomes increasingly important to investigate a better understanding of LC development (Bade and Dela Cruz, 2020). As a result, further research is gradually focusing on the other risk factors of LC, encompassing environmental exposures, lifestyle, gender, and genetics (Schabath and Cote, 2019).

Mitochondria serve as the central command for cellular metabolism, maintaining equilibrium and stress responses, playing a pivotal role in regulating processes like cell growth, division, differentiation, and apoptosis (Anderson et al., 2019). Previous studies have unveiled an unforeseen complexity and versatility in mitochondrial activities, combining mitochondrial energetics with protein biogenesis, metabolic pathways, and apoptosis (Pfanner et al., 2019). Moreover, recent studies based on proteomics indicated the remarkable importance of retaining mitochondrial proteostasis in guaranteeing the correct function of mitochondria (Wachoski-Dark et al., 2022). Encoded by both nuclear and mitochondrial DNA, mitochondrial proteins (MPs) are susceptible to errors during folding and assembly on account of oxidative stress and post-translational modifications (Stefani, 2004; Santo-Domingo and Demaurex, 2012). This may result in mitochondrial dysfunction, leading to an increase in reactive oxygen species (ROS) with tumor-promoting effect (Bandy and Davison, 1990). Mitochondrial protein quality control (MPQC) employs various pathways and regulators to maintain the quality and quantity of MPs. Dysregulated MPQC results in proteotoxicity and malfunctioning mitochondria, contributing to a range of human diseases, including cancer. Numerous studies have connected the dysfunction of MPQC in the etiology and pathogenesis of multiple types of cancer, including LC (Wallace, 2012; Friedlander et al., 2021). However, due to various objective factors, including technological and methodological constraints, the majority of existing research findings about MPQC rely on the animal or cellular experiments which can be influenced by multiple variables (Friedlander et al., 2021). In summary, the causality of the relationships between MPs and LC, as well as the direction of these causal connections, remains unclear. Therefore, it is crucial to investigate if MPs contribute to the onset of LC or just outcomes of shared risk factors.

Mendelian randomization (MR) analysis is a widely used method for establishing the causal relationship between exposure factors and outcomes, with the fundamental principle of employing genetic variations as instrumental variables (IVs) to model and evaluate the causality (Sanderson, 2021). The MR approach parallels the design of a randomized controlled trial (RCT) on account of parental alleles being randomly distributed to offspring during gamete formation in Mendel's law (Emdin et al., 2017). Furthermore, the results of MR studies are more robust against residual confoundings and the bias of reverse causal effects because the genetic variations are randomly assigned during meiosis and are not linked to environmental factors (Boehm and Zhou, 2022).

In our study, we aimed to apply a comprehensive two-sample MR analysis to determine the causal effect between MPs and LC and its various pathological types. By means of employing a bidirectional MR analysis, we could investigate the causality of MPs on LC risk and also determine if LC had a causal effect on MPs. From this foundation, we aimed to elucidate the influence between MPs and different pathological types of LC, ultimately aiding in developing innovative treatment options for LC.

2 Methods

2.1 Study design

Figure 1 illustrates an overview of the bidirectional MR analyses employed in our study. All IVs selected were guided by three principal assumptions of MR studies. Namely, IVs must demonstrate a strong association with the exposure; IVs impact the outcome solely through the exposure; IVs should not exhibit any association with confounding factors in the relationship between exposure and outcome.

2.2 Genome-wide association study (GWAS) sources

The GWAS data for MPs were sourced from a GWAS study involving a total sample size of 3,301 healthy participants of European descent (Sun et al., 2018). A total of 66 mitochondrial proteins (due to limited data availability) were enrolled in the subsequent MR analysis. The GWAS data for LC were derived from a large-scale GWAS study involving 85,716 individuals with 29,266 cases and 56,450 controls, while the GWAS study ulteriorly categorized LC into specific pathological types as lung adenocarcinoma (LUAD), lung squamous cell carcinoma (LUSC), and small cell lung carcinoma (SCLC) (McKay et al., 2017). All detailed information on GWAS data for MR analyses is presented in Figure 1. The original GWAS obtained approval from their respective institutions, and all data used for this study are publicly available. Therefore, no additional ethical approval was required.

2.3 Acquisition of IVs

Due to the restricted pool of accessible SNPs, we opted for SNPs with a cutoff of $p < 1e\text{-}5$. Then genetic instruments were excluded on a linkage disequilibrium (LD) threshold of $r2 < 0.001$ and a window size = 10,000 kb. To assess the statistical strength of each SNP, the F

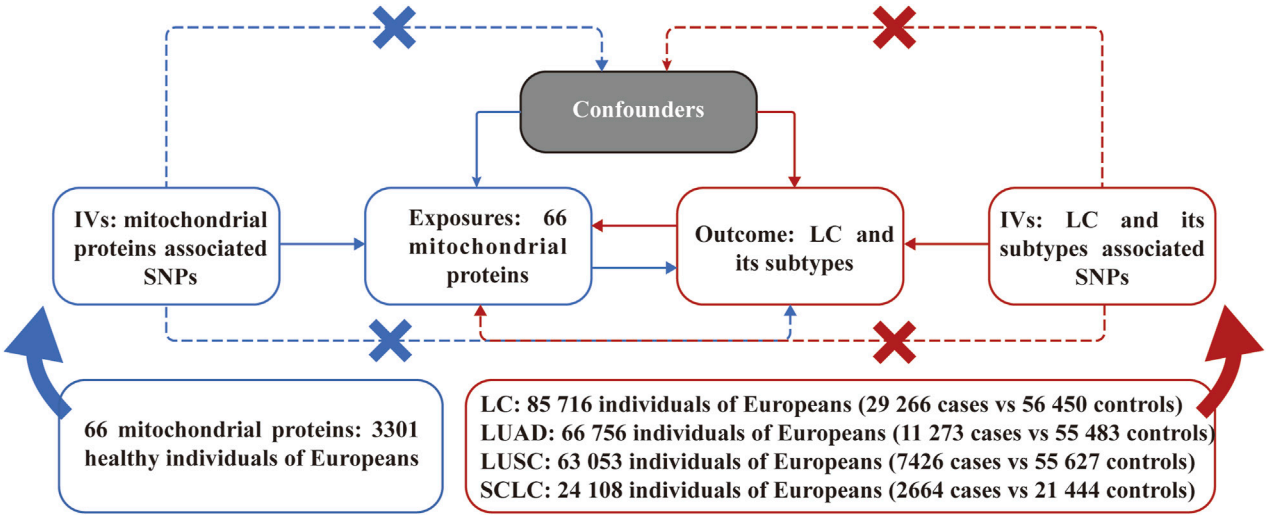


FIGURE 1
Assumptions and study design of the bidirectional Mendelian randomization study of the causal relationships between 66 mitochondrial proteins and different pathological types of lung cancer. LC, lung cancer; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; SCLC, small cell lung carcinoma; SNPs, single nucleotide polymorphisms.

statistics were also calculated, and the SNPs with F statistic <10 were eliminated for weak strength (Brion et al., 2013). We further excluded IVs that exhibited associations with potential confounding traits according to PhenoScanner (http://www.phenoscanner.medschl.cam.ac.uk/).

2.4 MR analysis

As for the two-sample analyses, we conducted the inverse variance-weighted (IVW) method as the primary approach for examining the bidirectional causal relationships between MPs and different pathological types of LC. Additionally, three complementary MR approaches were employed, including MR-Egger, weighted median (WM), and MR-Pleiotropy residual sum and outlier (MR-PRESSO), to sustain the findings derived from the IVM method. *p*-values were adjusted for false discovery rate (FDR) method, and Adjusted *p*-values (adj. *P*) < 0.05 were considered statistically significant. Also, *p*-values <0.05 were considered nominally significant.

2.5 Sensitivity analysis

Given that the IVW method could be biased by pleiotropic IVs, sensitivity analyses were employed to address the pleiotropic effects in the causal estimates. To assess potential heterogeneity, Cochrane's Q test was applied. In cases where heterogeneity was detected $p < 0.05$, a random-effects IVW analysis was performed to account for the measured heterogeneity. Additionally, the intercept of MR-Egger and MR-PRESSO global test were adopted to estimate the presence of horizontal pleio8tropy in the genetic variants ($p < 0.05$ indicated potential horizontal pleiotropy) while MR-PRESSO global test demonstrated a greater level of accuracy and assistance compared to MR-Egger in identifying horizontal pleiotropy. Furthermore, a leave-one-out analysis was conducted to determine whether the results were actuated by individual variants. We conducted all our MR analyses using the R software (version 4.3.1).

3 Results

3.1 Acquisition of IVs

After filtering for SNPs with LD, significantly linked to potential confounders (lung function and chronic obstructive pulmonary disease), and other LC-associated traits, a total of 125 SNPs were enrolled as IVs for the ensuing MR analyses, with the F statistics for each SNP being >10, demonstrating the absence of instrument bias (Supplementary Table S1).

3.2 Causal effects of MPs on LC

The results reported that both mitochondrial NADH dehydrogenase [ubiquinone]iron-sulfur protein 4 (Ndufs4) (IVW: OR = 0.971, 95% CI: 0.949–0.994, $p = 0.015$, adj. $p = 0.015$) and mitochondrial import inner membrane translocase subunit TIM14 (TIMM14/DNAJC19) (IVW: OR = 0.935, 95% CI: 0.887–0.985, $p = 0.012$, adj. $p = 0.023$) had a protective causal effect on overall LC (Figure 2). However, we observed no genetic predisposition to MPs demonstrated a causal relationship on LUAD (Supplementary Table S2).

As to LUSC, the findings suggested that mitochondrial steroidogenic acute regulatory protein (StAR) (IVW: OR = 0.878, 95% CI 0.790–0.977, $p = 0.017$, adj. $p = 0.028$), mitochondrial Ndufs4 (IVW: OR = 0.955, 95% CI 0.920–0.991, $p = 0.015$; WM: OR = 0.960, 95% CI 0.920–1.002, $p = 0.063$, adj. $p = 0.037$) and mitochondrial DNAJC19 (IVW: OR = 0.908, 95% CI 0.830–0.992,

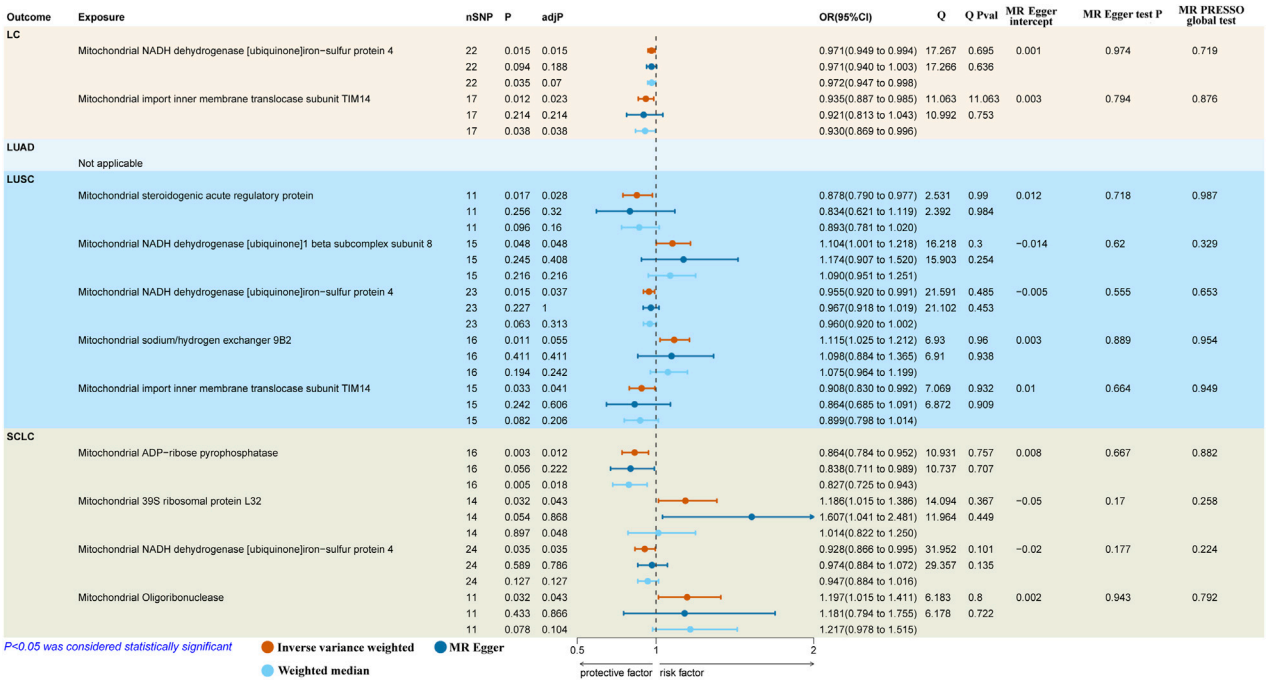


FIGURE 2
Causal effects of mitochondrial proteins on different pathological types of LC. LC, lung cancer; LUSC, lung squamous cell carcinoma; SCLC, small cell lung carcinoma; SNPs, single nucleotide polymorphisms; MR, Mendelian randomization; MR-PRESSO, MR-pleiotropy residual sum and outlier.

p = 0.033, adj. p = 0.041) had protective causal effects on LUSC. In contrast, mitochondrial NADH dehydrogenase [ubiquinone]1 beta subcomplex subunit 8 (NDUFB8) (IVW: OR = 1.104, 95% CI 1.001–1.218, p = 0.048, adj. p = 0.048) indicated pathogenic causal impacts on LUSC. Besides, mitochondrial sodium/hydrogen exchanger 9B2 (SLC9B2) (IVW: OR = 1.115, 95% CI 1.025–1.212, p = 0.011, adj. p = 0.055) showed a suggestive casual association with the higher risk of LUSC (Figure 2).

In terms of SCLC, we discovered two protective factors, including mitochondrial ADP-ribose pyrophosphatase (NUDT9) (IVW: OR = 0.864, 95% CI 0.784–0.952, p = 0.003, adj. p = 0.012) and mitochondrial Ndufs4 (IVW: OR = 0.928, 95% CI 0.866–0.995, p = 0.035, adj. p = 0.048) and two risk factors, including mitochondrial 39S ribosomal protein L32 (MRPL32) (IVW: OR = 1.186, 95% CI 1.015–1.386, p = 0.032, adj. p = 0.043) and mitochondrial oligoribonuclease (REXO2) (IVW: OR = 1.197, 95% CI 1.015–1.411, p = 0.032, adj. p = 0.043) were causal associated with SCLC (Figure 2).

3.3 Sensitivity analysis

The scatter plot showed that the causal estimates derived by the MR-Egger regression and weighted median approach were consistent in both dimension and direction with IVW method (Supplementary Figure S1). The findings of Cochrane's Q test indicated no significant heterogeneity (p > 0.05). The results revealed that the MR-Egger regression did not identify any pleiotropic effects for MPs (all p > 0.05). Additionally, the MR-PRESSO global test detected neither horizontal pleiotropic effects nor outlier SNPs (all p > 0.05). Moreover, the leave-one-out analysis validated that no individual SNP solely drove the causality between MPs and different pathological types of LC (Supplementary Figure S2).

3.4 Bidirectional causal associations between identified MPs and LC

To assess any reverse causality between identified MPs and different pathological types of LC, we considered overall LC and its subtypes as the exposure and identified MPs as the outcome. After screening, we employed 119 SNPs associated with different pathological types of LC as IVs (Supplementary Table S3). Finally, the results indicated no evidence for a reverse causal association between identified MPs and different pathological types of LC (Table 1).

4 Discussion

To our knowledge, this is the inaugural investigation of the causality between mitochondrial proteins and LC using open-access genetic databases. We employed bidirectional MR analyses to determine the causality between 66 MPs and different pathological types of LC, which enabled us to evaluate the upstream and downstream in the disease progression while avoiding reverse causation. We further ensured the robustness of our MR analyses against pleiotropic influences by implementing a variety of MR approaches, including MR-Egger and MR-PRESSO, to

10.3389/fgene.2024.1335223

TABLE 1 Causal effects of different pathological types of LC on identified mitochondrial proteins.

Exposure	Outcome	Method	nSNP	OR (95%CI)	*p*-value	adj. *p*-value
LC						
	Mitochondrial NADH dehydrogenase [ubiquinone]iron-sulfur protein 4	Inverse variance weighted	50	1.048 (0.933–1.176)	0.430	0.860
		MR Egger	50	1.083 (0.824–1.424)	0.570	0.570
		Weighted median	50	0.981 (0.831–1.157)	0.816	1
	Mitochondrial import inner membrane translocase subunit TIM14	Inverse variance weighted	50	1.034 (0.897–1.191)	0.649	0.649
		MR Egger	50	1.193 (0.856–1.664)	0.303	0.606
		Weighted median	50	1.012 (0.851–1.204)	0.894	0.894
LUSC						
	Mitochondrial steroidogenic acute regulatory protein	Inverse variance weighted	39	0.952 (0.879–1.031)	0.226	0.377
		MR Egger	39	1.121 (0.949–1.326)	0.188	0.470
		Weighted median	39	0.991 (0.879–1.117)	0.881	0.881
	Mitochondrial NADH dehydrogenase [ubiquinone]1 beta subcomplex subunit 8	Inverse variance weighted	39	0.995 (0.916–1.081)	0.911	0.911
		MR Egger	39	0.971 (0.814–1.159)	0.747	0.747
		Weighted median	39	1.030 (0.914–1.161)	0.626	1.000
	Mitochondrial NADH dehydrogenase [ubiquinone]iron-sulfur protein 4	Inverse variance weighted	39	0.993 (0.907–1.086)	0.872	1.000
		MR Egger	39	0.889 (0.736–1.072)	0.226	0.377
		Weighted median	39	1.023 (0.909–1.151)	0.706	0.883
	Mitochondrial sodium/hydrogen exchanger 9B2	Inverse variance weighted	39	0.948 (0.876–1.027)	0.191	0.478
		MR Egger	39	0.834 (0.706–0.986)	0.041	0.205
		Weighted median	39	0.967 (0.858–1.089)	0.579	1.000
	Mitochondrial import inner membrane translocase subunit TIM14	Inverse variance weighted	39	1.073 (0.987–1.167)	0.098	0.490
		MR Egger	39	1.093 (0.915–1.306)	0.333	0.416
		Weighted median	39	1.082 (0.964–1.213)	0.180	0.900
SCLC						
	Mitochondrial ADP-ribose pyrophosphatase	Inverse variance weighted	30	1.004 (0.945–1.067)	0.902	0.902

(Continued on following page)

TABLE 1 (*Continued*) Causal effects of different pathological types of LC on identified mitochondrial proteins.

Exposure	Outcome	Method	nSNP	OR (95%CI)	*p*-value	adj. *p*-value
		MR Egger	30	1.026 (0.904–1.165)	0.693	0.924
		Weighted median	30	0.987 (0.911–1.070)	0.757	1.000
	Mitochondrial 39S ribosomal protein L32	Inverse variance weighted	30	0.968 (0.911–1.029)	0.298	0.596
		MR Egger	30	0.905 (0.797–1.027)	0.134	0.536
		Weighted median	30	0.940 (0.869–1.017)	0.123	0.492
	Mitochondrial NADH dehydrogenase [ubiquinone]iron-sulfur protein 4	Inverse variance weighted	30	1.057 (0.995–1.123)	0.074	0.296
		MR Egger	30	1.033 (0.910–1.173)	0.619	1.000
		Weighted median	30	1.043 (0.957–1.137)	0.334	0.668
	Mitochondrial Oligoribonuclease	Inverse variance weighted	30	0.995 (0.936–1.057)	0.869	1.000
		MR Egger	30	0.975 (0.858–1.109)	0.708	0.708
		Weighted median	30	1.005 (0.922–1.096)	0.901	0.901

LC, lung cancer. LUAD, lung adenocarcinoma. LUSC, lung squamous cell carcinoma. SCLC, small cell lung carcinoma. SNPs, single nucleotide polymorphisms. MR, Mendelian randomization. adj. *p*-values, Adjusted *p*-values.

validate our findings. Our results indicated that a total of eight MPs may be potential protective contributors or potential risk factors to the development of overall LC, LUSC, and SCLC, whereas no such causal effect was observed in the case of LUAD. Furthermore, we found no evidence for a reverse causal effect between identified MPs and different pathological types of LC.

The majority of reactive oxygen species (ROS) within cells are produced from the mitochondrial respiratory chain. An overabundance of ROS can result in oxidative stress, causing oxidative harm to proteins and alterations in MP expression (Rabilloud et al., 2001; Poyton et al., 2009). Numerous studies have demonstrated that increased levels of ROS are correlated with the formation and advancement of LC (Weinberg et al., 2010; Jiang et al., 2022). Additionally, examining variations in the mitochondrial proteome is deemed to be an effective method to gauge the degree of mitochondrial damage under oxidative stress conditions (Gibson, 2005). Due to the absence of protective histones and a restricted range of DNA repair mechanisms, mitochondrial DNA (mtDNA) is highly susceptible to oxidative damage (Richter et al., 1988). Instability in mtDNA has been observed in several cancers, including LC (Chatterjee et al., 2006). An observational study suggested that compared to patients without LC, mutation rates in mtDNA were significantly increased in exhaled breath condensate in patients with LC (Yang Ai et al., 2013). After smoking, exposure to radon is the second leading cause of LC (Lorenzo-González et al., 2019). A study found that in radon-induced LC patients, the concentration of cell free mtDNA was significantly increased compared to other participants in the study (Bulgakova et al., 2022). Although these studies suggested a relationship between MPs and LC to some degree, the causal links remain unclear, and the direct association between them still lacks substantial research backing. Our research offered evidence supporting causal associations of MPs with LC and its subtypes by deducing the causality through genetic prediction using MR, which could also mitigate confounders effectively.

The result showed two protective factors in overall LC. Ndufs4 encodes mitochondrial complex I protein (Karamanlidis et al., 2013), while a study revealed deficiency in complex I led to elevated levels of mitochondrial ROS in macrophages in mouse models with myeloid-specific deletion of Ndufs4 (Cai et al., 2023). A similar result of Ndufs4 was also found in LUSC and SCLC, which indicated that Ndufs4 may be a vital protective factor in the development of LC. DNAJC19 plays a crucial role in preserving mitochondrial integrity, and the mutation in DNAJC19 could induce the occurrence of dilated cardiomyopathy and ataxia syndrome (Davey et al., 2006). Paradoxically, the expression of DNAJC19 was increased in NSCLC tissues compared to noncancerous adjacent tissues (Zhou et al., 2021). This conflictive result could be attributed to pathogenic variants in DNAJC19, which can lead to damage to mitochondrial function (Wachoski-Dark et al., 2022). Our finding also indicated the protective causal effect of DNAJC19 on LUSC, which further substantiates that DNAJC19 has a pivotal protective effect against LC.

In terms of LUSC, our study revealed four probable and one possible MPs with causal links, including protective factors of StAR, Ndufs4, and DNAJC19 and risk factors of NDUFB8 and SLC9B2.

StAR governs the crucial step that limits the rate of steroid biosynthesis, playing a vital role in the regulation of steroid hormones (Manna et al., 2009). ROS impairs mitochondria, leading to reduced StAR expression and steroidogenesis across various steroid-producing cells. At the same time, hormone deficiencies are considered a primary driver of human aging, which is related to the onset of various tumors (Manna et al., 2016; Jackaman et al., 2017). The selective nitration of NDUFB8 results in the disintegration of mitochondrial supercomplexes, causing the impairment of complex I activity and mitochondrial function. The activity of Complex I is recognized as a crucial factor in controlling mitochondrial respiration. Besides, nitration of NDUFB8 may represent a crucial mechanism in inflammatory conditions, which is a crucial component in the advancement of tumors (Coussens and Werb, 2002; Quintero et al., 2006; Davis et al., 2010). SLC9B2 is a sodium/hydrogen antiporter (Chintapalli et al., 2015). However, our understanding of the precise molecular functions of SLC9B2 remains limited. A previous study suggested the increased expression of SLC9B2 had a positive relationship with Autosomal-dominant polycystic kidney disease (Chapman et al., 2015). The expression level of SLC9B2 was identified significantly upregulated in Crohn's disease (Ye et al., 2022). We speculate that the inflammation may be one of the reasons for its role as a risk factor for LUSC.

For SCLC, the results identified that NUDT9 and Ndufs4 presented protective causal effects, and MRPL32 and REXO2 showed pathogenic causal effects. Adenosine diphosphate ribose (ADPR) interacts with NUDT9 homology to activate transient receptor potential melastatin 2 (TRPM2) channel (Miller and Cheung, 2016), while the decreased level of TRPM2 was considered to enhance tumor potential metastasis (Gershkovitz et al., 2018). SCLC is well known as a highly aggressive disease, thus this mechanism may explain the association between the protective factor NUDT9 and SCLC. As to risk factors, the current understanding of MRPL32 and REXO2 is limited. Prior studies demonstrated that suppressing MRPL32 could reduce oxygen-glucose deprivation/reperfusion damage (Guan et al., 2020) and that REXO2 was associated with a poorer prognosis in glioma (Wang et al., 2021).

Nevertheless, our study had several constraints. Firstly, increasing the sample size is pivotal for a more accurate determination of the causal relationship between MPs and different pathological types of LC due to the potential biases from the current fairly small MP sample size. Secondly, the participants in GWAS data were predominantly of European populations, which constrained the applicability of our results to other ethnicities and could result in biased conclusions. Finally, our study merely identified causal associations of MPs with LC and its subtypes, further in-depth research is required to clarify the exact mechanisms of the causality.

5 Conclusion

In general, we systematically assessed the causality between MPs and different pathological types of LC by performing bidirectional MR analyses. Our study identified a total of seven MPs had significant causal relationships on overall LC, LUSC, and SCLC. Our findings suggested that there were two protective causal associations with LC; two protective causal associations, two causal pathogenic associations, and a nominally protective causal association with LUSC; two protective causal associations and two causal pathogenic associations with SCLC. Additionally, the results demonstrated no MP had a causality link with LUAD, and no evidence supported the reverse causality for identified MPs with LC or its subtypes. This research underscores the causal effects of MPs on the occurrence of LC, suggesting that MPs might be a viable strategy for LC prevention.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: https://www.ebi.ac.uk/gwas/publications/29875488. https://www.ebi.ac.uk/gwas/publications/28604730.

Author contributions

TJ: Software, Writing–original draft. YL: Formal Analysis, Writing–original draft. ML: Project administration, Writing–original draft. YH: Resources, Writing–original draft. BY: Validation, Writing–original draft, Validation. JG: Conceptualization, Funding acquisition, Writing–review and editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This study was supported by grants from Jiangsu Provincial Research Hospital: YJXYY202204-XKA02.

Acknowledgments

This is a short text to acknowledge the contributions of specific colleagues, institutions, or agencies that aided the efforts of the authors.

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/fgene.2024.1335223/full#supplementary-material

References

Anderson, A. J., Jackson, T. D., Stroud, D. A., and Stojanovski, D. (2019). Mitochondria-hubs for regulating cellular biochemistry: emerging concepts and networks. *Open Biol.* 9, 190126. doi:10.1098/rsob.190126

Bade, B. C., and Dela Cruz, C. S. (2020). Lung cancer 2020: Epidemiology, etiology, and prevention. *Clin. Chest Med.* 41, 1–24. doi:10.1016/j.ccm.2019.10.001

Bandy, B., and Davison, A. J. (1990). Mitochondrial mutations may increase oxidative stress: implications for carcinogenesis and aging? *Free Radic. Biol. Med.* 8, 523–539. doi:10.1016/0891-5849(90)90152-9

Boehm, F. J., and Zhou, X. (2022). Statistical methods for Mendelian randomization in genome-wide association studies: a review. *Comput. Struct. Biotechnol. J.* 20, 2338–2351. doi:10.1016/j.csbj.2022.05.015

Brion, M.-J. A., Shakhbazov, K., and Visscher, P. M. (2013). Calculating statistical power in Mendelian randomization studies. *Int. J. Epidemiol.* 42, 1497–1501. doi:10.1093/ije/dyt179

Bulgakova, O., Kussainova, A., Kakabayev, A., Aripova, A., Baikenova, G., Izzotti, A., et al. (2022). The level of free-circulating mtDNA in patients with radon-induced lung cancer. *Environ. Res.* 207, 112215. doi:10.1016/j.envres.2021.112215

Cai, S., Zhao, M., Zhou, B., Yoshii, A., Bugg, D., Villet, O., et al. (2023). Mitochondrial dysfunction in macrophages promotes inflammation and suppresses repair after myocardial infarction. *J. Clin. Invest.* 133, e159498. doi:10.1172/JCI159498

Chapman, A. B., Devuyst, O., Eckardt, K.-U., Gansevoort, R. T., Harris, T., Horie, S., et al. (2015). Autosomal-dominant polycystic kidney disease (ADPKD): executive summary from a kidney disease: improving global outcomes (KDIGO) controversies conference. *Kidney Int.* 88, 17–27. doi:10.1038/ki.2015.59

Chatterjee, A., Mambo, E., and Sidransky, D. (2006). Mitochondrial DNA mutations in human cancer. *Oncogene* 25, 4663–4674. doi:10.1038/sj.onc.1209604

Chintapalli, V. R., Kato, A., Henderson, L., Hirata, T., Woods, D. J., Overend, G., et al. (2015). Transport proteins NHA1 and NHA2 are essential for survival, but have distinct transport modalities. *Proc. Natl. Acad. Sci. U. S. A.* 112, 11720–11725. doi:10.1073/pnas.1508031112

Coussens, L. M., and Werb, Z. (2002). Inflammation and cancer. *Nature* 420, 860–867. doi:10.1038/nature01322

Davey, K. M., Parboosingh, J. S., McLeod, D. R., Chan, A., Casey, R., Ferreira, P., et al. (2006). Mutation of DNAJC19, a human homologue of yeast inner mitochondrial membrane co-chaperones, causes DCMA syndrome, a novel autosomal recessive Barth syndrome-like condition. *J. Med. Genet.* 43, 385–393. doi:10.1136/jmg.2005.036657

Davis, C. W., Hawkins, B. J., Ramasamy, S., Irrinki, K. M., Cameron, B. A., Islam, K., et al. (2010). Nitration of the mitochondrial complex I subunit NDUFB8 elicits RIP1- and RIP3-mediated necrosis. *Free Radic. Biol. Med.* 48, 306–317. doi:10.1016/j.freeradbiomed.2009.11.001

Emdin, C. A., Khera, A. V., and Kathiresan, S. (2017). Mendelian randomization. *JAMA* 318, 1925–1926. doi:10.1001/jama.2017.17219

Friedlander, J. E., Shen, N., Zeng, A., Korm, S., and Feng, H. (2021). Failure to guard: mitochondrial protein quality control in cancer. *Int. J. Mol. Sci.* 22, 8306. doi:10.3390/ijms22158306

Gershkovitz, M., Caspi, Y., Fainsod-Levi, T., Katz, B., Michaeli, J., Khawaled, S., et al. (2018). TRPM2 mediates neutrophil killing of disseminated tumor cells. *Cancer Res.* 78, 2680–2690. doi:10.1158/0008-5472.CAN-17-3614

Gibson, B. W. (2005). The human mitochondrial proteome: oxidative stress, protein modifications and oxidative phosphorylation. *Int. J. Biochem. Cell Biol.* 37, 927–934. doi:10.1016/j.biocel.2004.11.013

Guan, X., Zhang, H., Qin, H., Chen, C., Hu, Z., Tan, J., et al. (2020). CRISPR/Cas9-mediated whole genomic wide knockout screening identifies mitochondrial ribosomal proteins involving in oxygen-glucose deprivation/reperfusion resistance. *J. Cell Mol. Med.* 24, 9313–9322. doi:10.1111/jcmm.15580

Jackaman, C., Tomay, F., Duong, L., Abdol Razak, N. B., Pixley, F. J., Metharom, P., et al. (2017). Aging and cancer: the role of macrophages and neutrophils. *Ageing Res. Rev.* 36, 105–116. doi:10.1016/j.arr.2017.03.008

Jiang, L., Zhang, J., Xu, Y., Xu, H., and Wang, M. (2022). Treating non-small cell lung cancer by targeting the PI3K signaling pathway. *Chin. Med. J. Engl.* 135, 1272–1284. doi:10.1097/CM9.0000000000002195

Karamanlidis, G., Lee, C. F., Garcia-Menendez, L., Kolwicz, S. C., Suthammarak, W., Gong, G., et al. (2013). Mitochondrial complex I deficiency increases protein acetylation and accelerates heart failure. *Cell Metab.* 18, 239–250. doi:10.1016/j.cmet.2013.07.002

Leiter, A., Veluswamy, R. R., and Wisnivesky, J. P. (2023). The global burden of lung cancer: current status and future trends. *Nat. Rev. Clin. Oncol.* 20, 624–639. doi:10.1038/s41571-023-00798-3

Lorenzo-González, M., Torres-Durán, M., Barbosa-Lorenzo, R., Provencio-Pulla, M., Barros-Dios, J. M., and Ruano-Ravina, A. (2019). Radon exposure: a major cause of lung cancer. *Expert Rev. Respir. Med.* 13, 839–850. doi:10.1080/17476348.2019.1645599

Luo, G., Zhang, Y., Etxeberria, J., Arnold, M., Cai, X., Hao, Y., et al. (2023). Projections of lung cancer incidence by 2035 in 40 countries worldwide: population-based study. *JMIR Public Health Surveill.* 9, e43651. doi:10.2196/43651

Manna, P. R., Dyson, M. T., and Stocco, D. M. (2009). Regulation of the steroidogenic acute regulatory protein gene expression: present and future perspectives. *Mol. Hum. Reprod.* 15, 321–333. doi:10.1093/molehr/gap025

Manna, P. R., Stetson, C. L., Slominski, A. T., and Pruitt, K. (2016). Role of the steroidogenic acute regulatory protein in health and disease. *Endocrine* 51, 7–21. doi:10.1007/s12020-015-0715-6

McKay, J. D., Hung, R. J., Han, Y., Zong, X., Carreras-Torres, R., Christiani, D. C., et al. (2017). Large-scale association analysis identifies new lung cancer susceptibility loci and heterogeneity in genetic susceptibility across histological subtypes. *Nat. Genet.* 49, 1126–1132. doi:10.1038/ng.3892

Miller, B. A., and Cheung, J. Y. (2016). TRPM2 protects against tissue damage following oxidative stress and ischaemia–reperfusion. *J. Physiol.* 594, 4181–4191. doi:10.1113/JP270934

Osuoha, C. A., Callahan, K. E., Ponce, C. P., and Pinheiro, P. S. (2018). Disparities in lung cancer survival and receipt of surgical treatment. *Lung Cancer* 122, 54–59. doi:10.1016/j.lungcan.2018.05.022

Pfanner, N., Warscheid, B., and Wiedemann, N. (2019). Mitochondrial proteins: from biogenesis to functional networks. *Nat. Rev. Mol. Cell Biol.* 20, 267–284. doi:10.1038/s41580-018-0092-0

Poyton, R. O., Ball, K. A., and Castello, P. R. (2009). Mitochondrial generation of free radicals and hypoxic signaling. *Trends Endocrinol. Metab.* 20, 332–340. doi:10.1016/j.tem.2009.04.001

Quintero, M., Colombo, S. L., Godfrey, A., and Moncada, S. (2006). Mitochondria as signaling organelles in the vascular endothelium. *Proc. Natl. Acad. Sci. U. S. A.* 103, 5379–5384. doi:10.1073/pnas.0601026103

Rabilloud, T., Heller, M., Rigobello, M. P., Bindoli, A., Aebersold, R., and Lunardi, J. (2001). The mitochondrial antioxidant defence system and its response to oxidative stress. *Proteomics* 1, 1105–1110. doi:10.1002/1615-9861(200109)1:9<1105::AID-PROT1105>3.0.CO;2-M

Richter, C., Park, J. W., and Ames, B. N. (1988). Normal oxidative damage to mitochondrial and nuclear DNA is extensive. *Proc. Natl. Acad. Sci. U. S. A.* 85, 6465–6467. doi:10.1073/pnas.85.17.6465

Sanderson, E. (2021). Multivariable mendelian randomization and mediation. *Cold Spring Harb. Perspect. Med.* 11, a038984. doi:10.1101/cshperspect.a038984

Santo-Domingo, J., and Demaurex, N. (2012). Perspectives on: SGP symposium on mitochondrial physiology and medicine: the renaissance of mitochondrial pH. *J. Gen. Physiol.* 139, 415–423. doi:10.1085/jgp.201110767

Schabath, M. B., and Cote, M. L. (2019). Cancer progress and priorities: lung cancer. *Cancer Epidemiol. Biomarkers Prev.* 28, 1563–1579. doi:10.1158/1055-9965.EPI-19-0221

Stefani, M. (2004). Protein misfolding and aggregation: new examples in medicine and biology of the dark side of the protein world. *Biochim. Biophys. Acta* 1739, 5–25. doi:10.1016/j.bbadis.2004.08.004

Sun, B. B., Maranville, J. C., Peters, J. E., Stacey, D., Staley, J. R., Blackshaw, J., et al. (2018). Genomic atlas of the human plasma proteome. *Nature* 558, 73–79. doi:10.1038/s41586-018-0175-2

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., et al. (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* 71, 209–249. doi:10.3322/caac.21660

Wachoski-Dark, E., Zhao, T., Khan, A., Shutt, T. E., and Greenway, S. C. (2022). Mitochondrial protein homeostasis and cardiomyopathy. *Int. J. Mol. Sci.* 23, 3353. doi:10.3390/ijms23063353

Wallace, D. C. (2012). Mitochondria and cancer. *Nat. Rev. Cancer* 12, 685–698. doi:10.1038/nrc3365

Wang, H., Wang, X., Xu, L., Zhang, J., and Cao, H. (2021). RUNX1 and REXO2 are associated with the heterogeneity and prognosis of IDH wild type lower grade glioma. *Sci. Rep.* 11, 11836. doi:10.1038/s41598-021-91382-1

Weinberg, F., Hamanaka, R., Wheaton, W. W., Weinberg, S., Joseph, J., Lopez, M., et al. (2010). Mitochondrial metabolism and ROS generation are essential for Kras-mediated tumorigenicity. *Proc. Natl. Acad. Sci. U. S. A.* 107, 8788–8793. doi:10.1073/pnas.1003428107

Yang Ai, S. S., Hsu, K., Herbert, C., Cheng, Z., Hunt, J., Lewis, C. R., et al. (2013). Mitochondrial DNA mutations in exhaled breath condensate of patients with lung cancer. *Respir. Med.* 107, 911–918. doi:10.1016/j.rmed.2013.02.007

Ye, C., Zhu, S., Gao, Y., and Huang, Y. (2022). Landscape of sialylation patterns identify biomarkers for diagnosis and prediction of response to anti-TNF therapy in crohn's disease. *Front. Genet.* 13, 1065297. doi:10.3389/fgene.2022.1065297

Zhou, J., Peng, Y., Gao, Y.-C., Chen, T.-Y., Li, P.-C., Xu, K., et al. (2021). Targeting DNAJC19 overcomes tumor growth and lung metastasis in NSCLC by regulating PI3K/AKT signaling. *Cancer Cell Int.* 21, 338. doi:10.1186/s12935-021-02054-z

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 22 April 2024
DOI 10.3389/fgene.2024.1357704

OPEN ACCESS

EDITED BY
Xuanye Cao,
University of Texas MD Anderson Cancer Center, United States

REVIEWED BY
Bingrui Li,
University of Texas MD Anderson Cancer Center, United States
Wen Jiang,
AstraZeneca, United States

*CORRESPONDENCE
Yueyang Lai,
laiyy@njucm.edu.cn
Haibo Cheng,
haibocheng@njucm.edu.cn

†These authors have contributed equally to this work

RECEIVED 18 December 2023
ACCEPTED 22 March 2024
PUBLISHED 22 April 2024

CITATION
Shang Z, Xi S, Lai Y and Cheng H (2024), Single-cell transcriptomics and Mendelian randomization reveal LUCAT1's role in right-sided colorectal cancer risk.
Front. Genet. 15:1357704.
doi: 10.3389/fgene.2024.1357704

Single-cell transcriptomics and Mendelian randomization reveal LUCAT1's role in right-sided colorectal cancer risk

Zhihao Shang[1,2†], Songyang Xi[1,2,3†], Yueyang Lai[1,2*] and Haibo Cheng[1,2*]

[1]Jiangsu Collaborative Innovation Center of Traditional Chinese Medicine Prevention and Treatment of Tumor, Nanjing University of Chinese Medicine, Nanjing, China, [2]The First School of Clinical Medicine, Nanjing University of Chinese Medicine, Nanjing, China, [3]Zhenjiang Hospital of Chinese Traditional and Western Medicine, Zhenjiang, Jiangsu, China

Background: Colorectal cancer (CRC) is a malignancy with high incidence and mortality rates globally, categorized into left-sided and right-sided CRC, each exhibiting significant differences in molecular characteristics, clinical manifestations, and prognosis.

Methods: This study employed single-cell transcriptomic data and various bioinformatics approaches, such as two-sample Mendelian randomization, reverse Mendelian randomization, colocalization analysis, directed filtering, pseudotime analysis, and intercellular communication analysis. It analyzed cellular-level disparities between left-sided and right-sided CRC, identifying distinct subpopulations with characteristic variations. For these cells, two-sample Mendelian randomization was utilized to explore gene-to-one-sided CRC causality.

Results: LUCAT1 was enriched in high-abundance monocyte subpopulations in right-sided CRC and demonstrated potential risk factor status through Mendelian randomization analysis. The specific single-nucleotide polymorphism (SNP) rs10774624 was associated with an increased risk of CRC. Moreover, metabolic pathway analysis revealed that LUCAT1^{+} monocytes exhibit lower communication activity in the tumor microenvironment and heightened activity in metabolic functions like glycosaminoglycan degradation. Its biological functions are related to the positive regulation of interleukin-6 production and NF-kappa B signaling, among others.

Conclusion: This study confirmed a potential causal relationship between LUCAT1 and right-sided CRC risk through Mendelian randomization analysis. These findings provide novel insights into the pathogenesis of right-sided CRC and may aid in developing early detection and treatment strategies for right-sided CRC.

KEYWORDS

colorectal cancer, single-cell sequencing, Mendelian randomized, LUCAT1, bioinformatics

1 Introduction

Colorectal cancer (CRC) ranks among the most prevalent and lethal malignancies worldwide (White and Sears, 2023). This neoplasm is categorized into right-sided and left-sided colorectal cancer based on the tumor's location within the colon. This classification extends beyond mere anatomical delineation, as these subtypes exhibit significant disparities in molecular characteristics, clinical manifestations, treatment responses, and prognoses. Typically, right-sided CRC is more common in older patients and is associated with a poorer prognosis, whereas left-sided CRC often responds better to treatment and shows higher survival rates (Boeckx et al., 2017; Nawa et al., 2008). These differences underscore the need for personalized approaches in the diagnosis, treatment, and management of CRC, highlighting the importance of understanding these variations to optimize treatment strategies, improve patient outcomes, and develop novel therapeutic targets.

Molecularly, right-sided and left-sided CRC demonstrate distinct characteristics. Right-sided CRC is frequently linked with microsatellite instability (MSI) and BRAF mutations, associated with immune evasion and chemotherapy resistance (Takahashi et al., 2016; van der Post and Hansson, 2014; Weiss et al., 2011). Conversely, left-sided CRC often exhibits mutations in the KRAS and p53 genes, aligning with the typical adenoma-carcinoma sequence (Brooks et al., 2001). Additionally, right-sided CRC shows higher genomic and epigenetic heterogeneity, while left-sided CRC is characterized by chromosomal instability (Kajiwara et al., 2023). These molecular features not only influence the tumor's biological behavior but also critically impact the response to various treatment modalities, playing a pivotal role in clinical decision-making. Thus, a deeper understanding of these molecular differences is crucial for developing more precise and effective treatment approaches.

Lung cancer-associated transcript 1 (LUCAT1), a long non-coding RNA (lncRNA), has garnered attention for its expression and function in various tumors. Initially identified in lung cancer, LUCAT1 has been found to regulate tumor progression in other cancer types as well (Cao et al., 2023). Its roles include promoting tumor cell proliferation, inhibiting apoptosis, enhancing cancer cell migration, and invasion, and participating in epigenetic regulation. LUCAT1 also modulates the activity of microRNAs (miRNAs) by acting as an "miRNA sponge," indirectly influencing the expression of numerous genes. A more comprehensive understanding of LUCAT1's role in tumorigenesis may pave the way for developing novel cancer treatment strategies (Xiao et al., 2021).

Single-cell technologies enable researchers to analyze gene expression, molecular characteristics, and cell states at an individual cell level. This is significant for revealing tumor heterogeneity, identifying distinct cell subpopulations, and understanding cell interactions within the tumor microenvironment. Mendelian randomization, an epidemiological method using genetic variants as instrumental variables, assesses causal relationships between exposures and disease outcomes. This approach helps mitigate confounding and reverse causation issues common in traditional observational studies (Lee et al., 2023; Shigemura et al., 2023; Wang H. et al., 2023; Wang M. et al., 2023).

Our study, leveraging single-cell transcriptomic data from right-sided and left-sided CRC obtained from the Gene Expression Omnibus (GEO) database, identifies a subgroup of mononuclear cells that promote tumor development. Through analyzing expression differences with other cell subpopulations, we identified differentially expressed genes in this subgroup. Using these genes as exposure factors and employing bioinformatics methods like Mendelian randomization, colocalization, and directional filtering, we discovered that genetic variations in LUCAT1, a long-chain non-coding RNA, are risk factors for right-sided CRC, potentially linked to the single-nucleotide polymorphism (SNP) rs10774624. These findings offer insights into the molecular mechanisms of right-sided and left-sided CRC and contribute to developing potential targeted therapeutic strategies.

2 Materials and methods

2.1 Data acquisition and preprocessing

The data for this study were sourced from the publicly accessible GEO database, specifically the GSE188711 dataset, which comprises single-cell RNA sequencing data from six colorectal cancer samples, including three from the left side and three from the right side of the colon. We used the Read10X function from the Seurat package to import data in the 10x Genomics format for data preprocessing. Each sample's data were read from the specified directory and immediately encapsulated into a Seurat object using the CreateSeuratObject function, with parameters set to default except for the project argument, which was uniquely assigned to each sample based on its origin (e.g., L1, L2, and L3 for left-sided samples and R1, R2, and R3 for right-sided samples).

2.2 Quality control and dimension reduction

The Seurat objects corresponding to individual samples were merged into a single dataset using the merge function with the default parameters to facilitate collective analysis. During the quality control (QC) phase, metrics such as the number of gene expression features (nFeature_RNA), the proportion of mitochondrial gene expression (percent.mt), and the proportion of hemoglobin gene expression (percent.HB) were computed for each cell. Cells were then filtered based on the following criteria: cells with gene counts over 200 and under 4,000 and mitochondrial gene expression below 10% to eliminate both dead or dying cells and potential doublets or multiplets.

For dimensionality reduction, we utilized principal component analysis (PCA) using the RunPCA function with features = VariableFeatures (object) to focus on highly variable genes, followed by uniform manifold approximation and projection (UMAP) for visualization purposes, employing the RunUMAP function with the default parameters except for dims = 1:10 to use the first ten principal components (Qi and Zhang, 2023).

2.3 Clustering analysis and single-cell type annotation

Clustering analysis was performed using the FindNeighbors function with default parameters but specifying dims = 1:10 to

use the data from the first ten principal components for neighborhood calculation. The FindClusters function was then applied for actual clustering, with the resolution parameter adjusted based on preliminary analyses to optimize cluster granularity. We utilized the SingleR package or cell-type annotation, leveraging the SingleR function with the reference dataset obtained from the HumanPrimaryCellAtlasData function for human samples. Annotation was performed by matching our dataset's expression profiles with those in the reference, providing a predicted cell type for each cell. In particular, for the detailed analysis of monocyte subpopulations, we isolated monocyte cells and conducted further dimension reduction, clustering, and annotation steps. The FindVariableFeatures function was used to select 2000 highly variable genes, followed by PCA and UMAP for visualization. Harmony was used for batch correction using the RunHarmony function with group.by.vars = "orig.ident," ensuring the integration of data from different samples without batch effects. The FindAllMarkers function was employed to identify subgroup-specific marker genes, with the parameters set to only.pos = TRUE, min.pct = 0.25, and logfc.threshold = 0.25 to focus on genes that were positively expressed in at least 25% of cells within any given cluster with a minimum fold change of 0.25.

This detailed methodology ensures a comprehensive and reproducible approach to analyzing single-cell RNA sequencing data, facilitating the identification of cellular subpopulations and their respective marker genes within the complex landscape of colorectal cancer (Feng et al., 2023).

2.4 Pseudotime and intercellular communication analysis

Leveraging the slingshot package, we conducted trajectory analysis on cells to delineate the developmental pathways of monocyte subtypes. Constructing SingleCellExperiment objects allowed us to convert Seurat objects for analysis with slingshot. Within the scope of the slingshot analysis, we designated "celltype" as the clustering label and employed UMAP for dimensionality reduction. Specific starting clusters [start.clus = c (3,5)] were selected, alongside setting the trajectory shrinkage parameter (shrink = 0.2) to facilitate the refined formation of trajectories. Additionally, trajectory visualization was performed, utilizing color and layout options provided by the RColorBrewer and igraph packages, graphically depicting cell state transitions and developmental trajectories.

In exploring intercellular communication, we initially refined our scRNA-seq dataset through a further filtration process, excluding non-monocyte cell types to concentrate on their communication within the colorectal cancer context. Subsequently, we engaged in a quantitative analysis of cell-to-cell communication using the CellChat package. This step involved the creation of a CellChat object and its integration with the Human Cell Communication Database (CellChatDB.human). Our focus was directed toward the "Secreted Signaling" category of cell communication, selecting supported ligand–receptor pairs from the database. By identifying overexpressed genes and ligand–receptor pairs and projecting these elements onto the PPI network, we could construct and quantify the probabilities of intercellular communication. Furthermore, we filtered the cell communication network, eliminating communications within specific cell groups that had fewer cells, adopting min.cells = 10 as a threshold for analysis. Finally, we depicted the network and bubble plots of cell communication through visualization tools, offering an intuitive presentation of the quantity and patterns of interactions among cell groups.

2.5 Metabolic pathway analysis

We employed the scMetabolism package to assess the metabolic activity within specific subpopulations of colorectal cancer cells. Initially, our focus was directed toward the "monocyte_CO2" subpopulation, which was further stratified into "LUCAT1^{+}M" and "LUCAT1^{-}M" subgroups based on the expression of the LUCAT1 gene. Additionally, subpopulations other than "monocyte_CO2" were isolated from the aggregate dataset for subsequent analysis. The scMetabolism package facilitated a quantitative assessment of metabolic activity in these cell subgroups. The specific steps of analysis included the following:

1) Utilizing the sc.metabolism.Seurat function to score metabolic pathways via the AUCell method, without imputation (imputation = F), and setting the number of parallel cores to 2 (ncores = 2). "KEGG" was selected as the metabolism type (metabolism.type = "KEGG").
2) Certain metabolic pathways (for instance, input.pathway < -rownames (scRNA_metab @assays [["METABOLISM"]] [["score"]])[61:90]) were chosen for dot plot visualization to illustrate the variability in pathway activity across different genotypes. Moreover, differential gene expression analysis was conducted using the FindAllMarkers function, identifying genes with significant expression differences between the "LUCAT1^{+}M" and "LUCAT1^{-}IM" subgroups (employing positive markers, with a log fold change threshold set at 0.5). The list of differential genes obtained was then utilized for subsequent functional enrichment analysis.

2.6 Gene conversion and Mendelian randomization preparation

Upon identifying key genes, we used the clusterProfiler and org.Hs.eg.db packages to convert these gene symbols to ENSEMBL IDs. This facilitated subsequent Mendelian randomization analysis, allowing us to extract SNP information related to these genes from publicly available Genome-Wide Association Studies (GWAS databases), serving as instrumental variables for the analysis.

2.7 Two-sample Mendelian randomization and bidirectional Mendelian randomization

In this study, differential genes in the second group of monocytes compared to other cells were used as exposure factors (Table 1), with colorectal cancer (CRC) as the outcome, for causal inference analysis (MR) using "TwoSampleMR" (https://github.

TABLE 1 Results of Mendelian randomization and colocalization with directional filtering.

Symbol	SNP	Mendelian randomization (Wald ratio)	Steiger filtering	Colocalization PPH4 (coloc.abf/coloc.susie)
LUCAT1	rs10774624	2.8365 (1.6750,4.8035)	Passed (3.20151×10^{-5})	9.297435e−01/0.00025

com/MRCIEU/TwoSampleMR). The Wald ratio was used for genes with only one eQTL available, while inverse-variance weighted MR (MR-IVW) was applied when two or more genetic tools were available. The odds ratio (OR) of increased colorectal cancer risk was expressed for carriers of specific gene polymorphisms (e.g., a particular SNP) compared to non-carriers. For the primary analysis, the Bonferroni correction was applied for multiple testing adjustments, prioritizing results for further analysis with a threshold of 0.05/888 ($p < 5.63 \times 10^{-5}$). Moreover, colorectal cancer was also used as the exposure variable, with high expression of LUCAT1 in right-sided colorectal cancer monocytes as the outcome, for a bidirectional Mendelian randomization analysis, revealing potential bidirectional causal relationships between colorectal cancer onset and genes. Two colorectal cancer datasets from the GWAS database were selected, one as a test set (ebi-a-GCST90018808) and another as a validation set (ebi-a-GCST012879).

2.8 Colocalization analysis

Bayesian colocalization analysis was employed to assess the likelihood of two traits sharing the same causal variant, using the "Coloc" package (https://github.com/chr1sw) with default settings. As previously described, Bayesian colocalization provides posterior probabilities for five hypotheses regarding whether two traits share a single variant. In this study, we tested hypothesis 4 (PPH4), where LUCAT1 and colorectal cancer were associated with the region through a shared variant. Using the Coloc.abf and Coloc.susie algorithms, a gene was defined as having evidence of colocalization if it met the criterion of a gene-based PPH4 > 80% with at least one algorithm. Following the extraction and organization of relevant SNP data, the locusComparer package was used to create regional association plots, showcasing the association degree of specific gene regions with the occurrence of colorectal cancer. The specific steps of analysis included the following:

1) Employing the vcfR::read.vcfR function to read the "./eqtl-a-ENSG00000119917.vcf" file, obtaining eQTL information for the gene of interest.
2) Utilizing the separate function from the tidyverse package to process genotype data, splitting columns containing multiple pieces of information into separate variables, including effect size (beta), standard error (se), logarithmic *p*-value (logpvalue), allele frequency (eaf), and sample size (samplesize).
3) Calculating the minor allele frequency (MAF) and filtering data based on chromosomal position to focus on SNPs within specific regions of the gene.
4) Generating regional association plots with the locuscomparer package to visually depict the degree of association between the LUCAT1 gene eQTLs and the colorectal cancer GWAS results.

2.9 Directional filtering

Directional filtering analysis was conducted to evaluate the relative impact of SNPs on LUCAT1 gene expression and colorectal cancer occurrence. This step was achieved by comparing the association strength of SNPs with the exposure and outcome, aiming to ensure that the SNPs used in the analysis were correctly aligned on the causal pathway with the exposure variable. The steiger_filtering and directionality_test functions within the TwoSampleMR package were utilized for this analysis.

2.10 Gene ontology enrichment analysis and Kyoto Encyclopedia of Genes and Genomes pathway enrichment analysis

Gene ontology (GO) enrichment analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis are pivotal methodologies in bioinformatics for deciphering the biological processes and pathways underlying gene expression data. To investigate the potential biological functions of monocyte_CO2 cells and LUCAT1$^+$ monocytes, we selected the top 100 highly variable genes from these two cellular subpopulations for GO and KEGG enrichment analyses. The org.Hs.eg.db package was utilized for ID conversion, while the clusterProfiler package facilitated the enrichment analyses. All analyses and visualizations were conducted within the R version 4.2.1 environment.

3 Results

3.1 Study design

Our investigation begins with the analysis of cell subpopulations in left- and right-sided colorectal cancer, identifying a notable monocyte_CO2 subpopulation predominance in right-sided cases. This observation is followed by differential gene expression analysis, coupled with cell communication and pseudotime analysis, to delineate gene expression patterns and cellular interactions. We then integrate two-sample Mendelian randomization, reverse Mendelian randomization, and eQTL mapping, refined by colocalization and directional filtering techniques. The study culminates with a focused downstream analysis of LUCAT1$^+$

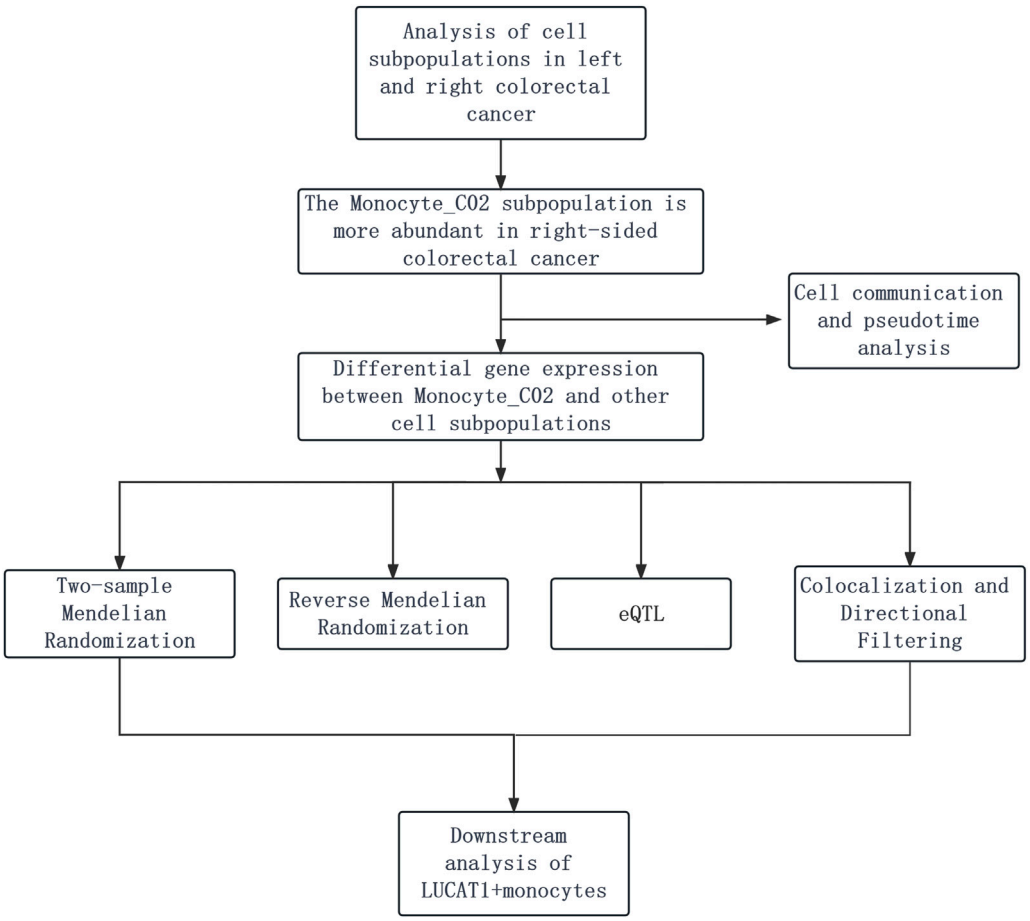


FIGURE 1
Depicts the analysis of colorectal cancer cell subpopulations, revealing a larger number of monocyte_CO2 cells in right-sided cases. It is followed by gene expression and cell interaction studies using Mendelian randomization and eQTL mapping, refined by colocalization and filtering, and ending with a focus on the role of LUCAT1⁺ monocytes in cancer progression.

monocytes to elucidate their potential role in colorectal cancer progression (Figure 1).

3.2 Monocytes, particularly monocyte_CO2, exhibit greater abundancein right-sided colorectal cancer, potentially related to the characteristics of right-sided colorectal cancer

In this investigation, single-cell sequencing data from left- and right-sided colorectal cancer were retrieved from the GEO database, identified as dataset GSE188711. Following stringent quality control (Supplementary Figures S1A, B) and batch-effect normalization (Supplementary Figures S1C–E), we procured high-quality single-cell transcriptomic data. We initially performed dimensionality reduction clustering on individual samples, followed by the presentation of marker expression for each cell cluster post-merging (Supplementary Figures S2, S3). Utilizing UMAP for clustering analysis, we segregated cells within tumor tissues into multiple subgroups, including B cells, T cells, dendritic cells (DCs), endothelial cells, epithelial cells, monocytes, neutrophils, smooth muscle cells, macrophages, and natural killer (NK) cells, each displaying distinct distribution patterns in left- and right-sided colorectal cancer samples (Figure 2A). Cell ratio diagrams further disclosed a notably higher proportion of monocytes in right-sided than left-sided colorectal cancer (Figure 2B). An independent analysis of monocytes indicated a significantly elevated proportion of the monocyte_CO2 subpopulation in right-sided colorectal cancer, signifying that this distinct subpopulation might play an enhanced role in this cancer variant (Figures 2C, D).

Further investigation into these monocyte subgroups revealed expression patterns of highly variable genes (Figure 2E). A suite of such genes, including CXCL8, TNFAIP6, CXCL3, and SPP1, were specifically upregulated in the monocyte_CO2 cells. The expression patterns of these genes might correlate with the functional dynamics of monocytes in right-sided colorectal cancer and their contribution to oncogenic processes within the tumor microenvironment. We further analyzed the biological functions enriched by the top 100 highly variable genes in the monocyte_CO2 cells, uncovering potential associations with biological processes such as leukocyte proliferation, mononuclear cell proliferation, lymphocyte proliferation, regulation of leukocyte proliferation, antigen processing and presentation, and MHC protein complex binding (Supplementary Figure S4A).

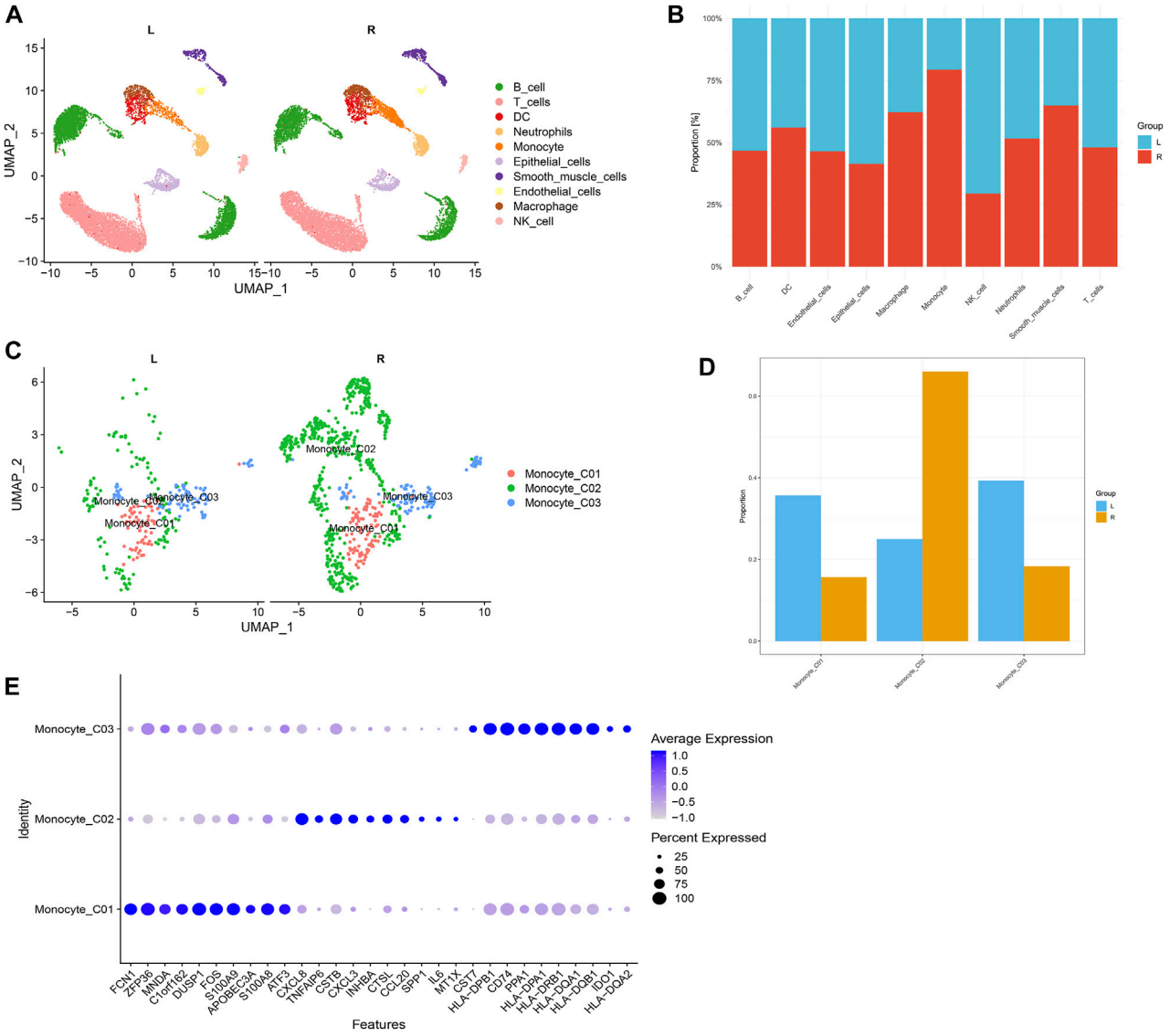


FIGURE 2
Single-cell transcriptomics reveal cellular disparities in left- and right-sided colorectal cancers. **(A)** Uniform manifold approximation and projection (UMAP) dimensionality reduction clustering identifies multiple cell subpopulations within colorectal cancer tissues, including B cells, T cells, dendritic cells, endothelial cells, epithelial cells, monocytes, neutrophils, smooth muscle cells, macrophages, and natural killer cells. **(B)** A stacked bar chart illustrates the comparative frequency of occurrence of each cell subpopulation between the two groups. **(C)** UMAP dimensionality reduction clustering displays the distribution of monocyte subpopulations. **(D)** Comparative cell proportion plots highlight the variance in monocyte subpopulations between left- and right-sided colorectal cancers. **(E)** Dotplot diagrams demonstrate the highly variable genes among three monocyte groups.

3.3 Temporal and communicative features of monocyte_CO2 cells

Figure 3A charts the evolutionary trajectory of monocyte subpopulations, transitioning from monocyte_C01 to monocyte_CO2 and monocyte_CO3 over time. Given the dynamic shifts in the tumor microenvironment, monocyte_C01 is hypothesized to represent an initial state from which they differentiate temporally. Subsequent analyses of cellular communication within the right-sided colorectal cancer (R group), in contrast to the left-sided (L group), demonstrated a marked diminution in the communication activity of monocyte_CO2 cells (Figures 3B, C). This phenomenon could reflect the biological heterogeneity between left- and right-sided colorectal cancers, potentially involving changes in cellular communication patterns and immunomodulatory processes pertinent to tumor evolution. Detailed assessments using Dotplot charts (Figures 3D, E) revealed that, against a backdrop of reduced overall cellular communication, interactions between monocyte_CO2 cells and macrophages were relatively intensified. This finding implies the formation of a more cohesive communication network among specific cell subpopulations within the tumor microenvironment, likely an adaptive response to environmental perturbations.

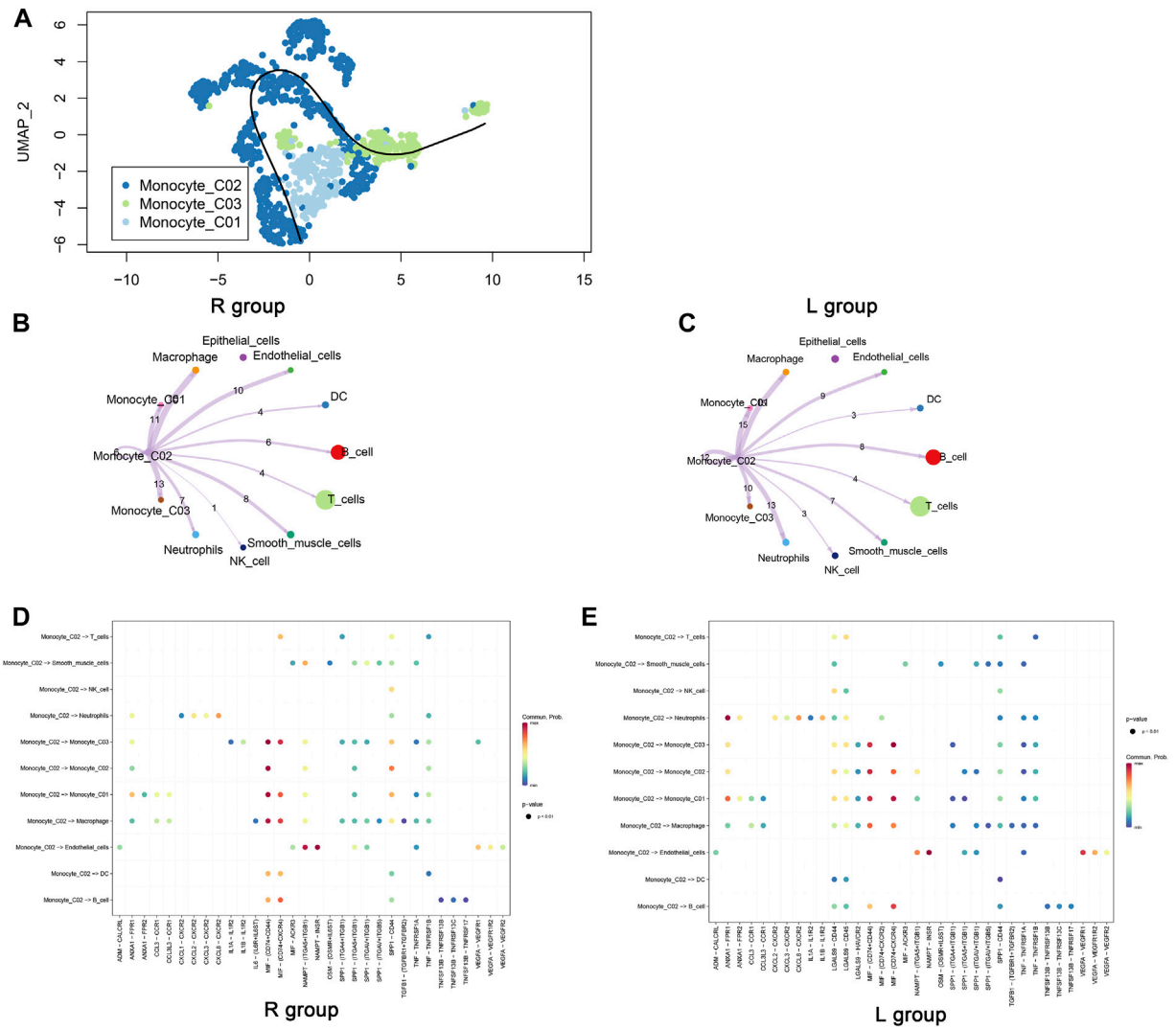


FIGURE 3
Pseudotime trajectory and intercellular communication of monocytes. **(A)** The cellular differentiation trajectory plot illustrates the differentiation pathway of monocytes, initiating from C01 and diverging toward C02 and C03. **(B, C)** The cellular communication network diagrams exhibit the communication dynamics between cells in left- and right-sided colorectal cancers originating from the Monocyte_C02 cells. **(D, E)** Dotplot diagrams display the expression of ligand–receptor pairs between monocyte_C02 cells, as the point of communication initiation, and other cells.

3.4 eQTL, colocalization, and directional filtering

The intersection of differential genes from monocyte_CO2 cells with those from other major cell classes and monocyte subpopulations was analyzed for Mendelian randomization with colorectal cancer. eQTL analysis identified numerous significant loci with a substantial impact on gene expression and associated colorectal cancer risk. Key gene loci, including LUCAT1, SDC2, and GRAMD1A, exhibited highly significant eQTL effects. Variations at these loci influenced gene expression substantially, with the variation at the LUCAT1 locus particularly demonstrating a robust association with increased colorectal cancer risk (Figure 4A).

Further colocalization analysis in tandem with GWAS findings pinpointed a specific SNP, rs653178, showing significant correlations in both eQTL analysis for the LUCAT1 gene and GWAS for colorectal cancer risk. The right-sided Manhattan plot reinforced the significant association signal of this SNP with colorectal cancer risk at a specific chromosomal location (Figure 4B). Steiger directional filtering lent additional support to the association directionality between SNP rs10774624 and colorectal cancer risk, suggesting the mutation precedes the cancer's onset (Table 1).

3.5 Bidirectional and two-sample mendelian randomization study results indicate the relationship between LUCAT1 and the risk of colorectal cancer

In Figure 5A, using the ebi-a-GCST90018808 dataset from the GWAS database for the forward MR analysis indicated a high positive correlation between genetic variation in LUCAT1 and

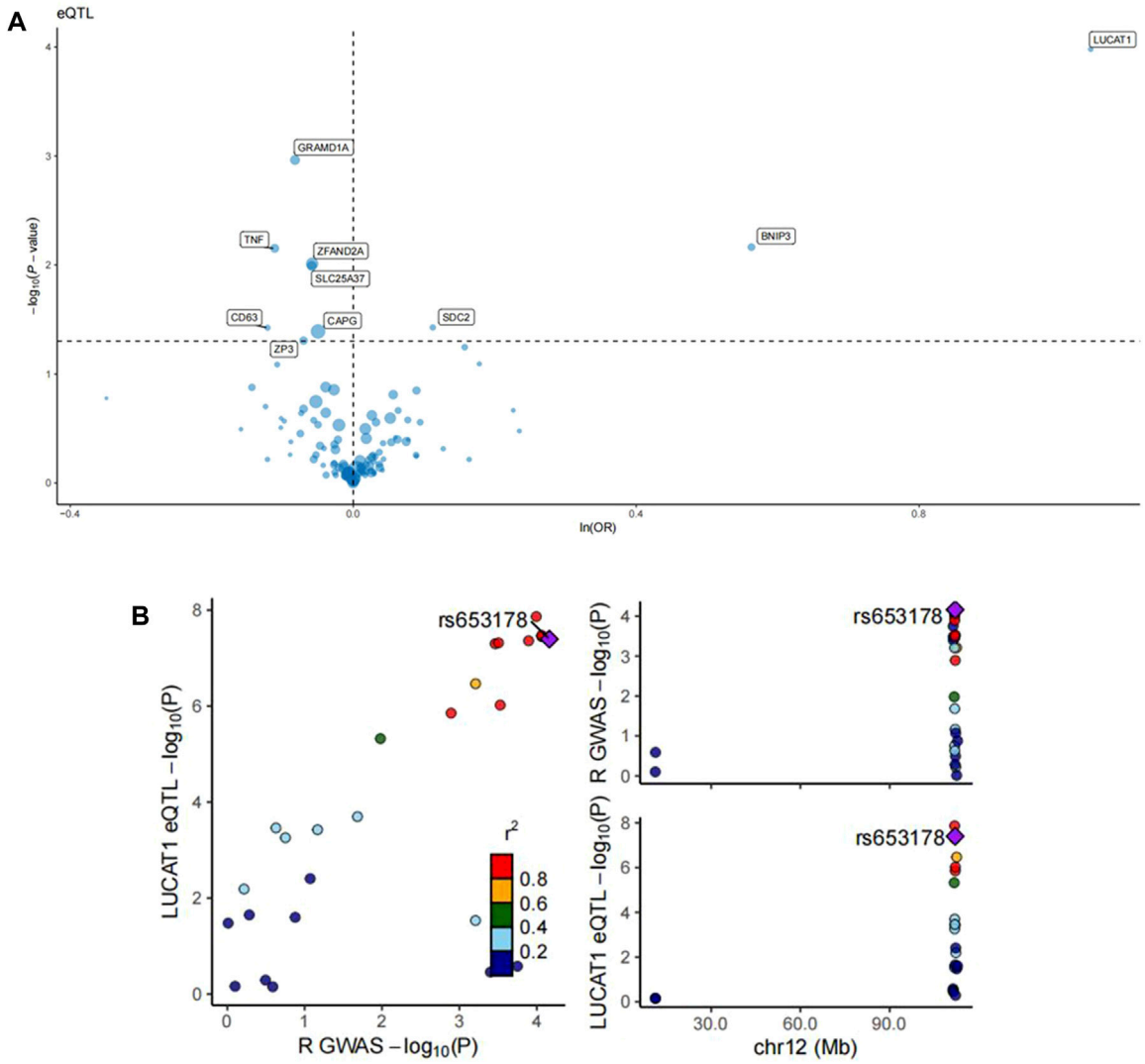


FIGURE 4
eQTL analysis of monocyte_C02 characteristic genes and regional association map of LUCAT1. **(A)** The volcano plots illustrate the Mendelian randomization (MR) results for the characteristic genes of monocyte_C02 cells in relation to colorectal cancer risk. The MR analysis utilized the Wald ratio or inverse-variance weighted method to evaluate the impact of genetic variations in the monocyte_C02 characteristic genes on colorectal cancer risk. The odds ratio (OR) for an increased risk of colorectal cancer was quantified based on per standard deviation increase in gene levels. "ln" refers to the natural logarithm, and "PVE" represents the proportion of variance explained. **(B)** The left panel depicts the relationship between the −log10(P) values of eQTLs and those of GWAS. Each point represents a single-nucleotide polymorphism (SNP), with the color indicating different r^2 values, reflecting the degree of correlation of the SNP in both eQTL and GWAS studies. The right panel is a regional association plot for SNPs within a specific region on chromosome 12 associated with a particular phenotype. Here, the *X*-axis represents the chromosomal position, while the *Y*-axis shows the −log10(P) values of the SNP's association with the specific phenotype. The color coding is consistent with the left panel, representing the r^2 values.

the risk of developing colorectal cancer (ORwr = 2.8365 [1.6750–4.8035], p = 0.0001), signifying a notable elevation in cancer risk. The reverse MR analysis (Figure 5B) assessed the impact of colorectal cancer risk on the LUCAT1 gene. Employing various MR methodologies, like MR Egger and the inverse-variance weighted method, the correlation between colorectal cancer risk and the LUCAT1 gene expression was found to be non-significant, suggesting colorectal cancer is not a direct causative factor for LUCAT1 genetic variation.

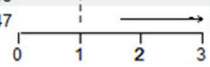
For validation, a two-sample MR analysis with dataset ebi-a-GCST012879 (Figure 6) corroborated the initial findings, indicating a positive correlation between increased LUCAT1 expression levels and a higher colorectal cancer risk.

3.6 Biological traits of LUCAT1$^+$ monocytes

Further exploration into the role of LUCAT1 within the tumor microenvironment revealed that LUCAT1 is predominantly expressed in monocytes and neutrophils (Figure 7A) and is highly expressed in right-sided colorectal cancer (Figure 7B). Echoing previous results, LUCAT1$^+$ monocytes in right-sided colorectal cancer continued to exhibit weaker communication strength (Figure 7C). Dotplot charts of LUCAT1$^+$ monocytes as both communicative sources and target cells showed widespread overexpression of the receptor–ligand pair CD74 and CD44 in LUCAT1$^+$ monocytes, other monocytes, and macrophages, with CD74 and CXCR4 also playing significant roles (Figure 7D). A

A

exposure	method	nsnp	pval	OR (95% CI)
CAPG	MR Egger	4	0.2085422141	0.9249 (0.8508 - 1.0055)
CAPG	Weighted median	4	0.0464603931	0.9502 (0.9036 - 0.9992)
CAPG	Inverse variance weighted	4	0.0407067891	0.9515 (0.9072 - 0.9979)
CAPG	Simple mode	4	0.3495607077	0.9599 (0.8928 - 1.0321)
CAPG	Weighted mode	4	0.2225670247	0.9575 (0.9058 - 1.0121)
GRAMD1A	MR Egger	6	0.7358389054	0.9786 (0.8702 - 1.1004)
GRAMD1A	Weighted median	6	0.0057163322	0.9269 (0.8783 - 0.9782)
GRAMD1A	Inverse variance weighted	6	0.0010874321	0.9210 (0.8767 - 0.9676)
GRAMD1A	Simple mode	6	0.0700576739	0.8850 (0.7974 - 0.9822)
GRAMD1A	Weighted mode	6	0.0525626453	0.9318 (0.8822 - 0.9842)
CD63	Inverse variance weighted	2	0.0376349667	0.8859 (0.7903 - 0.9931)
SLC25A37	MR Egger	7	0.0744574007	0.9185 (0.8530 - 0.9892)
SLC25A37	Weighted median	7	0.0092123773	0.9369 (0.8920 - 0.9840)
SLC25A37	Inverse variance weighted	7	0.0102056274	0.9430 (0.9017 - 0.9862)
SLC25A37	Simple mode	7	0.2796943137	0.9395 (0.8475 - 1.0414)
SLC25A37	Weighted mode	7	0.0412254013	0.9342 (0.8872 - 0.9836)
SDC2	MR Egger	5	0.1948829486	1.4178 (0.9396 - 2.1394)
SDC2	Weighted median	5	0.0908933346	1.1146 (0.9829 - 1.2639)
SDC2	Inverse variance weighted	5	0.0374125411	1.1190 (1.0066 - 1.2441)
SDC2	Simple mode	5	0.3027694913	1.0998 (0.9393 - 1.2878)
SDC2	Weighted mode	5	0.2280178554	1.1077 (0.9621 - 1.2753)
BNIP3	Wald ratio	1	0.0068538611	1.7558 (1.1675 - 2.6405)
ZFAND2A	MR Egger	4	0.2261228987	0.8289 (0.6700 - 1.0255)
ZFAND2A	Weighted median	4	0.0065779161	0.9416 (0.9016 - 0.9834)
ZFAND2A	Inverse variance weighted	4	0.0097234167	0.9436 (0.9031 - 0.9861)
ZFAND2A	Simple mode	4	0.1083975302	0.9223 (0.8599 - 0.9892)
ZFAND2A	Weighted mode	4	0.0953559660	0.9420 (0.8973 - 0.9890)
ZP3	MR Egger	3	0.6863577463	1.0411 (0.8988 - 1.2059)
ZP3	Weighted median	3	0.0250387538	0.9329 (0.8780 - 0.9913)
ZP3	Inverse variance weighted	3	0.0494272474	0.9322 (0.8692 - 0.9998)
ZP3	Simple mode	3	0.1831454787	0.8121 (0.6624 - 0.9956)
ZP3	Weighted mode	3	0.2333459864	0.9498 (0.8946 - 1.0083)
TNF	MR Egger	3	0.6666067628	0.9393 (0.7593 - 1.1618)
TNF	Weighted median	3	0.0074836239	0.8954 (0.8258 - 0.9709)
TNF	Inverse variance weighted	3	0.0070295311	0.8949 (0.8255 - 0.9702)
TNF	Simple mode	3	0.1627698698	0.8218 (0.6881 - 0.9816)
TNF	Weighted mode	3	0.1203115415	0.8992 (0.8304 - 0.9737)
LUCAT1	Wald ratio	1	0.0001048347	2.8365 (1.6750 - 4.8035)



B

exposure	method	nsnp	pval	outcome	OR (95% CI)
Colorectal cancer || id:ebi-a-GCST90018808	MR Egger	10	0.6420933	LUCAT1	1.0885 (0.7714 − 1.5360)
Colorectal cancer || id:ebi-a-GCST90018808	Weighted median	10	0.5103282	LUCAT1	1.0357 (0.9331 − 1.1495)
Colorectal cancer || id:ebi-a-GCST90018808	Inverse variance weighted	10	0.8295357	LUCAT1	1.0096 (0.9252 − 1.1017)
Colorectal cancer || id:ebi-a-GCST90018808	Simple mode	10	0.8238720	LUCAT1	1.0211 (0.8541 − 1.2208)
Colorectal cancer || id:ebi-a-GCST90018808	Weighted mode	10	0.5507487	LUCAT1	1.0405 (0.9178 − 1.1795)

FIGURE 5
Mendelian randomization (MR) analysis of genetic variations in monocyte_C02 characteristic genes and their association with colorectal cancer risk, illustrated through forest plots. **(A)** This panel employs the Mendelian randomization approach to investigate the causal relationship between genetic variations in monocyte_C02 characteristic genes and the risk of developing colorectal cancer. Analytical methods include MR Egger, weighted median, and inverse-variance weighted approaches. **(B)** This panel utilizes MR to explore the causal relationship of genetic variations in colorectal cancer on the monocyte_C02 characteristic genes. Each point estimate represents the odds ratio (OR) associated with specific genetic variations and disease risk, accompanied by a 95% confidence interval (CI). The diagram also includes the sample size (nsnp), *p*-value (pval), and overall effect size (OR and CI) for each method. Statistical significance is typically determined by a *p*-value <0.05.

pseudo-timeline depicted the association between gene expression and timing, with LUCAT1 primarily expressed early in the timeline (Figure 7E). Figure 7F presents metabolic function differences between LUCAT1$^+$ and LUCAT1$^-$ monocytes and two other monocyte groups, revealing heightened activity in metabolic functions such as glycosaminoglycan degradation, ubiquinone and other terpenoid-quinone biosynthesis, and thiamine metabolism. We further investigated the biological functions enriched by the top 100 highly variable genes in LUCAT1$^+$ monocytes, revealing that these cells may be implicated in biological processes such as positive regulation of interleukin-6 production, the NF-kappa B signaling pathway, pattern recognition receptor activity, cellular response to lipopolysaccharide, and the integrin complex (Supplementary Figure S4B).

exposure	method	nsnp	pval	OR (95% CI)
CAPG	MR Egger	4	7.211976e-01	0.9804 (0.8921 - 1.0775)
CAPG	Weighted median	4	5.057107e-01	0.9810 (0.9273 - 1.0379)
CAPG	Inverse variance weighted	4	2.888012e-01	0.9723 (0.9230 - 1.0241)
CAPG	Simple mode	4	6.340871e-01	0.9756 (0.8899 - 1.0694)
CAPG	Weighted mode	4	7.562997e-01	0.9900 (0.9344 - 1.0490)
GRAMD1A	MR Egger	6	5.577038e-01	0.9572 (0.8369 - 1.0947)
GRAMD1A	Weighted median	6	8.819534e-01	0.9946 (0.9258 - 1.0684)
GRAMD1A	Inverse variance weighted	6	6.607395e-01	0.9858 (0.9247 - 1.0509)
GRAMD1A	Simple mode	6	8.290761e-01	0.9865 (0.8774 - 1.1092)
GRAMD1A	Weighted mode	6	9.174188e-01	0.9959 (0.9244 - 1.0729)
CD63	Inverse variance weighted	2	1.392964e-01	0.8741 (0.7313 - 1.0448)
SLC25A37	MR Egger	7	4.002287e-01	1.0495 (0.9468 - 1.1633)
SLC25A37	Weighted median	7	2.088645e-01	1.0437 (0.9764 - 1.1156)
SLC25A37	Inverse variance weighted	7	1.915511e-01	1.0428 (0.9792 - 1.1105)
SLC25A37	Simple mode	7	3.579388e-01	1.0633 (0.9423 - 1.1998)
SLC25A37	Weighted mode	7	2.704637e-01	1.0452 (0.9732 - 1.1225)
SDC2	MR Egger	5	5.452674e-01	1.1844 (0.7272 - 1.9289)
SDC2	Weighted median	5	9.720532e-01	0.9970 (0.8451 - 1.1763)
SDC2	Inverse variance weighted	5	6.988289e-01	1.0273 (0.8963 - 1.1774)
SDC2	Simple mode	5	9.035911e-01	0.9835 (0.7633 - 1.2672)
SDC2	Weighted mode	5	8.707017e-01	0.9795 (0.7753 - 1.2375)
BNIP3	Wald ratio	1	9.479946e-01	1.0116 (0.7157 - 1.4299)
ZFAND2A	MR Egger	4	6.842564e-01	1.0519 (0.8521 - 1.2986)
ZFAND2A	Weighted median	4	3.604403e-01	0.9768 (0.9288 - 1.0272)
ZFAND2A	Inverse variance weighted	4	2.880928e-01	0.9736 (0.9267 - 1.0229)
ZFAND2A	Simple mode	4	7.813947e-01	0.9884 (0.9170 - 1.0655)
ZFAND2A	Weighted mode	4	4.849496e-01	0.9772 (0.9233 - 1.0343)
ZP3	Inverse variance weighted	2	8.516077e-01	1.0172 (0.8507 - 1.2163)
TNF	MR Egger	3	3.144835e-01	0.7748 (0.5918 - 1.0143)
TNF	Weighted median	3	1.344339e-02	0.8724 (0.7829 - 0.9721)
TNF	Inverse variance weighted	3	1.622076e-02	0.8741 (0.7832 - 0.9755)
TNF	Simple mode	3	9.004642e-01	1.0182 (0.7930 - 1.3073)
TNF	Weighted mode	3	1.244841e-01	0.8595 (0.7655 - 0.9650)
LUCAT1	Wald ratio	1	3.991901e-06	3.4434 (2.0360 - 5.8237)

FIGURE 6
Forest plot of Mendelian randomization (MR) analysis in a validation dataset. To enhance the robustness of the MR analysis, we conducted a two-sample MR analysis using an additional dataset. Each point estimate represents the odds ratio (OR) associated with specific genetic variations and disease risk, including a 95% confidence interval (CI). The diagram also details the sample size (nsnp), *p*-value (pval), and overall effect size (OR and CI) for each method used. Statistical significance is generally determined by a *p*-value <0.05.

4 Discussion

Single-cell transcriptomic sequencing, a transformative technology widely employed in biological and medical research in recent years, enables precise analysis of gene expression patterns at the individual cell level (Wang H. et al., 2023; Xie et al., 2023). This approach is particularly vital for revealing cellular heterogeneity in disease states, especially in complex and heterogeneous cancers such as colorectal cancer. Colorectal cancer, a prevalent malignancy, is categorized into left-sided and right-sided types, each differing significantly in molecular characteristics, clinical manifestations, and prognosis (Choi and Kim, 2023; Sustic et al., 2023; Talhouni et al., 2023). For instance, left-sided colorectal cancers are commonly associated with KRAS and BRAF mutations, whereas right-sided cancers are more linked to microsatellite instability and CpG island methylation anomalies (Bond et al., 2023; Kajiwara et al., 2023). Mendelian randomization analysis, a statistical tool, utilizes these genetic variations as natural experiments to establish causative links between specific genetic markers and cancer risk, providing crucial insights into the molecular mechanisms of colorectal cancer and the development of new therapeutic strategies.

Monocytes in tumor tissues, particularly in colorectal cancer, play a complex and multifaceted role. As integral components of the

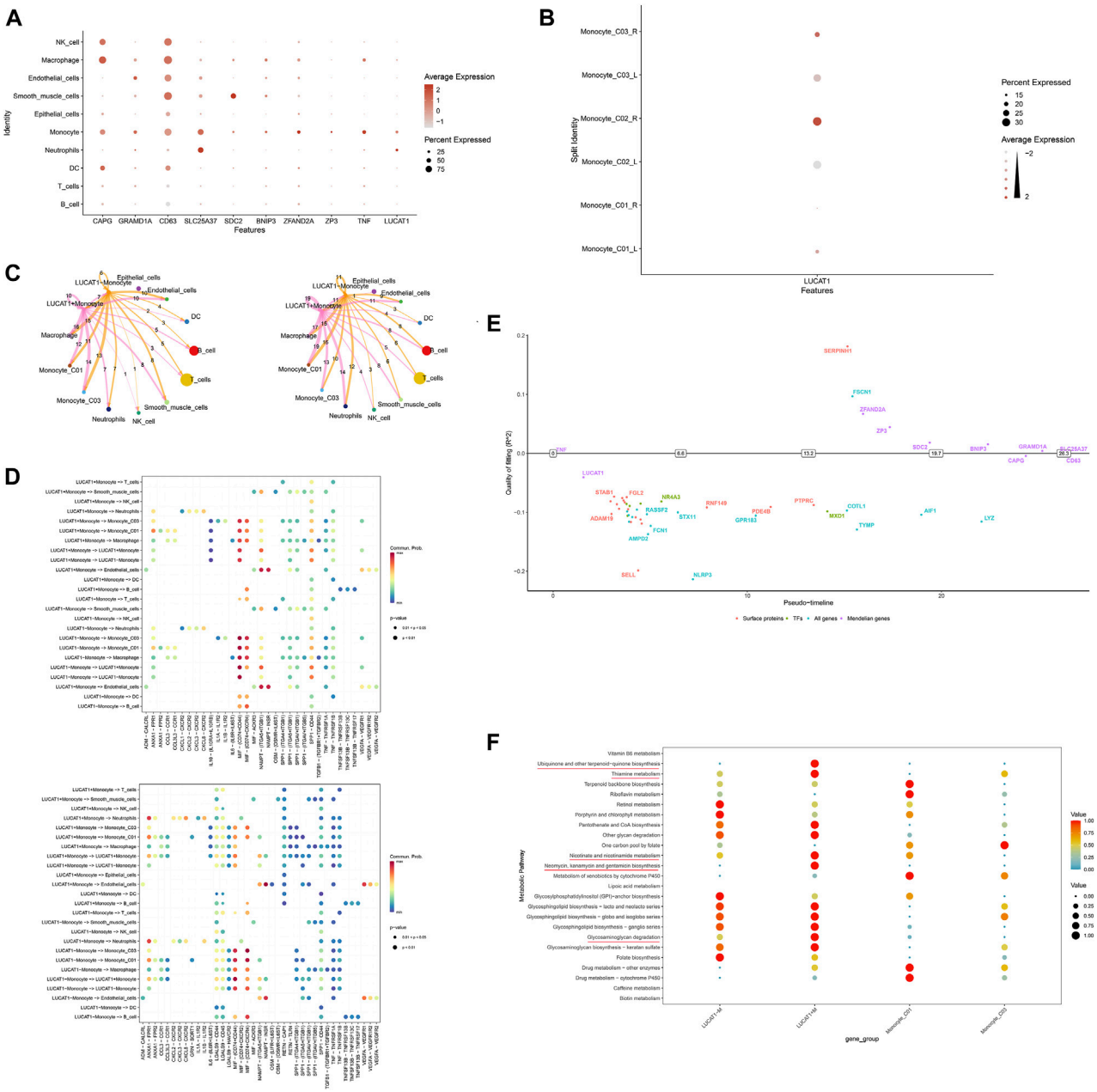


FIGURE 7
Downstream mechanisms of LUCAT1^{+} monocytes. **(A)** Expression of LUCAT1 across various cell subpopulations. Variations in size and color represent the expression level and percentage of LUCAT1 in different cell types. **(B)** LUCAT1 is highly expressed in monocyte_C02 in right-sided colorectal cancer. **(C)** Cellular communication network diagram shows the communication intensity of LUCAT1^{+} monocytes in right-sided colorectal cancer. The nodes represent different cell types, and the line width indicates the strength of communication. **(D)** Dotplot diagram displays the expression of receptor−ligand pairs involved in communication, where each dot's size and color signify the expression level and percentage of different genes across various cell types. **(E)** The pseudo-timeline graph illustrates gene expression changes over time, with the horizontal axis representing time and the vertical axis indicating gene expression levels. **(F)** The dotplot diagram highlights the metabolic functional differences between LUCAT1^{+} and LUCAT1^{-} monocytes and other monocytes. Dots of different colors represent the activity of various metabolic pathways.

immune system, they can differentiate into various cell types, including macrophages and dendritic cells, each playing intricate roles in the tumor microenvironment (Gan et al., 2023; Wu et al., 2023). On the one hand, certain differentiated monocytes, such as M2-type macrophages, may facilitate tumor growth and metastasis by secreting growth factors, pro-inflammatory cytokines, and angiogenic factors, supporting tumor proliferation and spread (Li et al., 2023; Shen et al., 2023). They may also secrete immunosuppressive molecules, aiding tumors in evading immune surveillance (Batalha et al., 2023; Gawinski et al., 2023). Conversely, some monocytes can differentiate into immunologically active cells like M1-type macrophages and certain dendritic cells, enhancing the immune response against tumor cells (Andreuzzi et al., 2022; Ito et al., 2023).

Our analysis of single-cell transcriptomic data from left- and right-sided colorectal cancers identified a higher prevalence of the monocyte_C02 subgroup in right-sided colorectal cancer. High variability genes in monocyte_C02 cells, including CXCL8, TNFAIP6, CXCL3, and SPP1, indicate a potential role in tumor promotion. Studies suggest CXCL8 can accelerate tumor cell proliferation, epithelial–mesenchymal transition (EMT), angiogenesis, and impede anti-tumor immunity (Ravi et al., 2023). CXCL3, a bioactive protein of low molecular weight, primarily recruits and activates various cells expressing the CXC chemokine receptor (CXCR)12, which is involved in cell migration, invasion, and angiogenesis and plays a crucial role in the development of cardiovascular and pulmonary diseases (Chang et al., 2023; Ji et al., 2023). Research also reveals that SPP1 impacts the tumor microenvironment by promoting inflammatory responses, immune suppression, and regulating extracellular matrix (ECM) remodeling to support tumor growth (Lv et al., 2023; Yu et al., 2023). These findings imply that the monocyte_C02 subgroup may contribute to tumor progression. The study discovered that in right-sided colorectal cancer, the reduction in communication among monocyte_CO2 cells, as well as between monocyte_CO2 cells and neutrophils, was significant. Monocytes and neutrophils can interact and influence each other's functions within the tumor microenvironment. For instance, they can regulate each other through secreted cytokines, or in certain cases, the activation of one cell type may promote the recruitment of another cell type to the tumor microenvironment. However, this interaction is complex and may vary depending on the type of tumor, specific conditions of the microenvironment, and the host's immune status. Monocytes can differentiate into tumor-associated macrophages (TAMs), which typically exhibit an M2 polarization state that promotes tumor growth, invasion, and metastasis by secreting growth factors, pro-inflammatory cytokines, and angiogenic factors, thus providing a favorable environment for the tumor. The reduction in communication may slow the progression of monocytes to M2-type macrophages, thereby steering right-sided colorectal cancer in a direction more favorable for immune response (Sawa-Wejksza et al., 2018). Figures 3D, E display the expression of different receptor–ligand pairs across various cell subpopulations, providing rich information on the differences in cellular communication between left- and right-sided colorectal cancers. Notably, the expression of ANXA1-FPR1 between monocytes and between monocytes and neutrophils is reduced in right-sided colorectal cancer. ANXA1 can regulate the function of T cells (especially regulatory T cells, or Tregs) by binding to FPR1, which is crucial for immune evasion in the tumor microenvironment. Specifically, the ANXA1–FPR1 complex may enhance the immunosuppressive function of Treg cells, thereby weakening the immune system's attack on the tumor (Sawa-Wejksza et al., 2018; Tian et al., 2023). The expression of CD74-CXCR4 is decreased in right-sided colorectal cancer. In some cancers, the expression levels of CD74 and CXCR4 may influence the behavior of tumor cells, including metastasis and invasion. This may indicate a better prognosis for right-sided colorectal cancer (Bian et al., 2023). The biological functions of monocyte_CO2 cells are primarily enriched in processes such as leukocyte proliferation, mononuclear cell proliferation, lymphocyte proliferation, regulation of leukocyte proliferation, antigen processing and presentation, and MHC protein complex binding. Monocytes, including monocytes and certain types of lymphocytes, can further differentiate into macrophages and dendritic cells, which play pivotal roles in tumor immune responses. Macrophages can have either tumor-promoting or tumor-suppressing effects depending on their states within the tumor microenvironment. The proliferation and activation of monocytes are crucial in the tumor immune editing process, as they influence immune responses through cytokine production and antigen presentation. The coordinated action of these immune functions is essential for eliciting effective tumor-specific immune responses. The coordination of immune functions such as MHC protein complex binding, MHC class II protein complex binding, and antigen processing and presentation is crucial for triggering effective tumor-specific immune responses. The ability of immune cells to efficiently recognize and respond to tumor cells largely depends on the effective processing and presentation of tumor antigens, as well as the correct expression and function of MHC molecules. This suggests that the increase in monocyte_CO2 cells is likely related to the immune environment of right-sided colorectal cancer. The reason for this could be that right-sided colorectal cancers often have a higher microsatellite instability (MSI) and BRAF mutation rate, while left-sided cancers are more associated with chromosomal instability pathways. This could also explain why right-sided colorectal cancers tend to have higher immune response rates.

Our two-sample Mendelian randomization and reverse Mendelian randomization analyses identified genetic variations in LUCAT1 as risk factors for colorectal cancer, supported by directional filtering and colocalization analyses. The mutation at the rs10774624 locus was deemed causally related to colorectal cancer onset, and additional colocalization analysis revealed another potential disease risk-related locus sharing genetic variation, rs653178. Although research on these loci is scarce, known biological functions of LUCAT1 and eQTL results suggest a strong promotive effect of its SNPs on colorectal cancer development.

To further explore the potential role of LUCAT1, we conducted separate analyses of the functions of LUCAT1$^+$ and LUCAT1$^-$ monocytes in tumor tissues. LUCAT1 expression was found to be higher in right-sided colorectal cancer, supporting the hypothesis that LUCAT1 is a risk factor that may contribute to the development of right-sided colorectal cancer. Generally, LUCAT1$^+$ monocytes demonstrated weaker communication strength than LUCAT1$^-$ monocytes, which could be attributed to the typically better prognosis of right-sided colorectal cancer, leading to lower levels of malignancy cell communication. Moreover, monocyte_C02 cells showed the most robust communication with other monocytes, likely linked to monocyte differentiation processes, especially involving MIF interactions with CD74 and CD44. In tumorigenesis, MIF-CD74 interactions are known to promote tumor cell growth and survival, with the expression levels of MIF being closely tied to tumor aggressiveness and prognosis in certain types of cancer. MIF-CD44 interactions are also crucial in affecting cell–ECM interactions, particularly important in the migration and metastasis of tumor cells. The pronounced activity of

LUCAT1 in the early stages of monocyte differentiation may be associated with its role in facilitating malignant differentiation and the onset of tumors. An analysis of differential metabolic pathways revealed increased activity in glycosaminoglycan degradation in LUCAT1$^+$ monocytes. Glycosaminoglycans (GAGs), crucial biomolecules containing uronic acid and amino sugar residues, disrupt cell–cell adhesion during tumor cell dissociation and invasion. The modification of cadherin (E-cadherin) by β1,6 N-acetylglucosaminyltransferase V (GnT-V), which adds β1,6-N-acetylglucosamine (β1,6GlcNAc)-branched N-glycans, impairs cell adhesion and aids in tumor cell invasion. Additionally, the α2,6-sialylation terminal structure interferes with tumor cell adhesion (Ahn et al., 2009; Shen et al., 2022). The biological functions of LUCAT1$^+$ monocytes are markedly enriched in pathways, including positive regulation of interleukin-6 production, NF-kappa B signaling, pattern recognition receptor activity, cellular response to lipopolysaccharide, and integrin complexes. Among these, IL-6 is a multifunctional cytokine that plays a pivotal role in the initiation and progression of tumors. Elevated levels of IL-6 are associated with various types of cancers, contributing to the proliferation, survival, and migration of tumor cells. IL-6 facilitates inflammatory reactions and the formation of the tumor microenvironment by activating the JAK/STAT3 signaling pathway, which aids in tumor growth and immune evasion (Johnson et al., 2018). The NF-κB signaling pathway is crucial in the regulation of inflammatory and immune responses. This pathway is often activated in cancer, enhancing tumor cell proliferation, survival, and invasion while inhibiting apoptosis. Activation of NF-κB is also intimately connected with the recruitment of inflammatory cells within the tumor microenvironment and the generation of tumor-promoting inflammation (Lee et al., 2007). Pattern recognition receptors (PRRs) play an essential role in innate immunity by recognizing pathogen-associated molecular patterns (PAMPs) and damage-associated molecular patterns (DAMPs). Within the tumor microenvironment, activation of PRRs can promote inflammatory responses that may favor the proliferation and survival of tumor cells. Integrins, as receptors on the cell surface, mediate interactions between cells and the extracellular matrix (ECM) that are crucial for cell migration, proliferation, and survival (Amarante-Mendes et al., 2018). In cancer, integrin activation can enhance the invasiveness and metastatic potential of tumor cells, remodeling the tumor microenvironment through interactions with the ECM. This further suggests that the expression of LUCAT1 may facilitate the progression of tumors.

5 Conclusion

Through single-cell transcriptomic sequencing, we discovered that the monocyte_C02 subpopulation is more prevalent in right-sided colorectal cancer, and its highly expressed genes, such as CXCL8, TNFAIP6, CXCL3, and SPP1, may be implicated in tumor growth promotion. Notably, we observed that the expression of LUCAT1 in these monocytes could be associated with the occurrence of colorectal cancer, with Mendelian randomization analysis further indicating a direct link between genetic variations in LUCAT1 and colorectal cancer risk. Additionally, LUCAT1$^+$ monocytes were found to be more active in right-sided colorectal cancer, potentially influencing the glycosaminoglycan degradation pathway, thereby disrupting cell adhesion and facilitating tumor cell invasion. In summary, LUCAT1 is not only a risk factor for the development of colorectal cancer but may also participate in the progression of the disease by modulating the functions of monocytes and the interactions within the tumor microenvironment.

This provides a new perspective in uncovering and understanding the complex biological mechanisms of right-sided colorectal cancer and offers valuable insights for developing potential therapeutic strategies targeting LUCAT1. Future advancements in these findings may contribute to promoting more personalized cancer treatment approaches, particularly tailored therapies for different subtypes of left- and right-sided colorectal cancer, ultimately improving patient prognosis.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding authors.

Author contributions

ZS: Conceptualization, Investigation, Software, Writing–original draft, Writing–review and editing. SX: Data curation, Formal Analysis, Methodology, Conceptualization, Funding acquisition, Investigation, Project administration, Resources, Software, Validation, Visualization, Writing–original draft, Writing–review and editing. YL: Writing–review and editing, Writing–original draft. HC: Funding acquisition, Project administration, Resources, Validation, Visualization, Writing–review and editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by the National Key R&D Program of China (2022YFC3500200; 2022YFC3500202), the Innovation Team and Talents Cultivation Program of National Administration of Traditional Chinese Medicine (No: ZYYCXTD-C-202208), a project funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions (N/A), the Natural Science Foundation of China (No:81902330), the Social Development Guiding Science and Technology Plan Project of Zhenjiang (No. FZ2021065), the Science and Technology Project of Jiangsu Provincial Administration of Traditional Chinese Medicine (No. YB201843), and the Talent and Technology Construction Fund Project of Zhenjiang Hospital of Chinese Traditional and Western Medicine (No. ZXY202301).

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/fgene.2024.1357704/full#supplementary-material

References

Ahn, Y. H., Lee, J. Y., Lee, J. Y., Kim, Y. S., Ko, J. H., and Yoo, J. S. (2009). Quantitative analysis of an aberrant glycoform of TIMP1 from colon cancer serum by L-PHA-enrichment and SISCAPA with MRM mass spectrometry. *J. Proteome Res.* 8 (9), 4216–4224. doi:10.1021/pr900269s

Amarante-Mendes, G. P., Adjemian, S., Branco, L. M., Zanetti, L. C., Weinlich, R., and Bortoluci, K. R. (2018). Pattern recognition receptors and the host cell death molecular machinery. *Front. Immunol.* 9, 2379. doi:10.3389/fimmu.2018.02379

Andreuzzi, E., Fejza, A., Polano, M., Poletto, E., Camicia, L., Carobolante, G., et al. (2022). Colorectal cancer development is affected by the ECM molecule EMILIN-2 hinging on macrophage polarization via the TLR-4/MyD88 pathway. *J. Exp. Clin. Cancer Res.* 41 (1), 60. [Journal Article]. doi:10.1186/s13046-022-02271-y

Batalha, S., Gomes, C. M., and Brito, C. (2023). Immune microenvironment dynamics of HER2 overexpressing breast cancer under dual anti-HER2 blockade. *Front. Immunol.* 14, 1267621. [Journal Article]. doi:10.3389/fimmu.2023.1267621

Bian, Y., Bi, G., Shan, G., Liang, J., Yao, G., Sui, Q., et al. (2023). Identification of the relationship between single-cell N6-methyladenosine regulators and the infiltrating immune cells in esophageal carcinoma. *Heliyon* 9 (8), e18132. doi:10.1016/j.heliyon.2023.e18132

Boeckx, N., Koukakis, R., Op, D. B. K., Rolfo, C., Van Camp, G., Siena, S., et al. (2017). Primary tumor sidedness has an impact on prognosis and treatment outcome in metastatic colorectal cancer: results from two randomized first-line panitumumab studies. *Ann. Oncol.* 28 (8), 1862–1868. [Journal Article]. doi:10.1093/annonc/mdx119

Bond, M., Bolhuis, K., Loosveld, O., de Groot, J., Droogendijk, H., Helgason, H. H., et al. (2023). First-line systemic treatment strategies in patients with initially unresectable colorectal cancer liver metastases (CAIRO5): an open-label, multicentre, randomised, controlled, phase 3 study from the Dutch Colorectal Cancer Group. *Lancet Oncol.* 24 (7), 757–771. doi:10.1016/S1470-2045(23)00219-X

Brooks, D. G., James, R. M., Patek, C. E., Williamson, J., and Arends, M. J. (2001). Mutant K-ras enhances apoptosis in embryonic stem cells in combination with DNA damage and is associated with increased levels of p19(ARF). *Oncogene* 20 (17), 2144–2152. doi:10.1038/sj.onc.1204309

Cao, Y., Zou, Z., Wu, X., Li, W., Lu, Z., Hu, J., et al. (2023). LUCAT1 inhibits ferroptosis in bladder cancer by regulating the mRNA stability of STAT3. *Gene* 894, 147974. [Journal Article]. doi:10.1016/j.gene.2023.147974

Chang, S., Zhang, G., Li, L., Li, H., Jin, X., Wang, Y., et al. (2023). Sirt4 deficiency promotes the development of atherosclerosis by activating the NF-κB/IκB/CXCL2/3 pathway. *Atherosclerosis* 373, 29–37. doi:10.1016/j.atherosclerosis.2023.04.006

Choi, Y., and Kim, N. (2023). Sex difference of colon adenoma pathway and colorectal carcinogenesis. *World J. Mens. Health* 42, 256–282. [Journal Article; Review]. doi:10.5534/wjmh.230085

Feng, X., Chen, Q., Zhong, J., Yu, S., Wang, Y., Jiang, Y., et al. (2023). Molecular characteristics of circulating B cells and kidney cells at the single-cell level in special types of primary membranous nephropathy. *Clin. Kidney J.* 16 (12), 2639–2651. doi:10.1093/ckj/sfad215

Gan, L., Ye, D., Feng, Y., Pan, H., Lu, X., Wan, J., et al. (2023). Immune cells and hypertension. *Immunol. Res.* 72, 1–13. [Journal Article; Review]. doi:10.1007/s12026-023-09414-z

Gawinski, C., Mroz, A., Roszkowska-Purska, K., Sosnowska, I., Derezinska-Wolek, E., Michalski, W., et al. (2023). A prospective study on the roles of the Lymphocyte-to-Monocyte ratio (LMR), Neutrophil-to-Lymphocyte ratio (NLR), and Platelet-to-Lymphocyte ratio (PLR) in patients with locally advanced rectal cancer. *Biomedicines* 11 (11), 3048. [Journal Article]. doi:10.3390/biomedicines11113048

Ito, M., Mimura, K., Nakajima, S., Okayama, H., Saito, K., Nakajima, T., et al. (2023). M2 tumor-associated macrophages resist to oxidative stress through heme oxygenase-1 in the colorectal cancer tumor microenvironment. *Cancer Immunol. Immunother.* 72 (7), 2233–2244. [Journal Article]. doi:10.1007/s00262-023-03406-6

Ji, H. Z., Chen, L., Ren, M., Li, S., Liu, T. Y., Chen, H. J., et al. (2023). CXCL8 promotes endothelial-to-mesenchymal transition of endothelial cells and protects cells from erastin-induced ferroptosis via CXCR2-mediated activation of the NF-κB signaling pathway. *Pharm. (Basel)* 16 (9), 1210. doi:10.3390/ph16091210

Johnson, D. E., O'Keefe, R. A., and Grandis, J. R. (2018). Targeting the IL-6/JAK/STAT3 signalling axis in cancer. *Nat. Rev. Clin. Oncol.* 15 (4), 234–248. doi:10.1038/nrclinonc.2018.8

Kajiwara, T., Nishina, T., Yamashita, R., Nakamura, Y., Shiozawa, M., Yuki, S., et al. (2023). Sidedness-Dependent prognostic impact of gene alterations in metastatic colorectal cancer in the nationwide cancer genome screening project in Japan (SCRUM-Japan GI-SCREEN). *Cancers (Basel)* 15 (21), 5172. [Journal Article]. doi:10.3390/cancers15215172

Lee, C. H., Jeon, Y. T., Kim, S. H., and Song, Y. S. (2007). NF-kappaB as a potential molecular target for cancer therapy. *Biofactors* 29 (1), 19–35. doi:10.1002/biof.5520290103

Lee, H. J., Zhao, Y., Fleming, I., Mehta, S., Wang, X., Wyk, B. V., et al. (2023). Early cellular and molecular signatures correlate with severity of West Nile virus infection. *iScience* 26 (12), 108387. [Journal Article]. doi:10.1016/j.isci.2023.108387

Li, C., Guan, R., Li, W., Wei, D., Cao, S., Xu, C., et al. (2023). Single-cell RNA sequencing reveals tumor immune microenvironment in human hypopharygeal squamous cell carcinoma and lymphatic metastasis. *Front. Immunol.* 14, 1168191. doi:10.3389/fimmu.2023.1168191

Lv, J., Gao, H., Ma, J., Liu, J., Tian, Y., Yang, C., et al. (2023). Dynamic atlas of immune cells reveals multiple functional features of macrophages associated with progression of pulmonary fibrosis. *Front. Immunol.* 14, 1230266. doi:10.3389/fimmu.2023.1230266

Nawa, T., Kato, J., Kawamoto, H., Okada, H., Yamamoto, H., Kohno, H., et al. (2008). Differences between right- and left-sided colon cancer in patient characteristics, cancer morphology and histology. *J. Gastroenterol. Hepatol.* 23 (3), 418–423. doi:10.1111/j.1440-1746.2007.04923.x

Qi, W., and Zhang, Q. (2023). Insights on epithelial cells at the single-cell level in hepatocellular carcinoma prognosis and response to chemotherapy. *Front. Pharmacol.* 14, 1292831. [Journal Article]. doi:10.3389/fphar.2023.1292831

Ravi, D., Ntinopoulou, E., Guetta, N., Weier, M., Vogel, V., Spellerberg, B., et al. (2023). Dysregulated monocyte-derived macrophage response to Group B Streptococcus in newborns. *Front. Immunol.* 14, 1268804. [Journal Article]. doi:10.3389/fimmu.2023.1268804

Sawa-Wejksza, K., Dudek, A., Lemieszek, M., Kalawaj, K., and Kandefer-Szerszen, M. (2018). Colon cancer-derived conditioned medium induces differentiation of THP-1 monocytes into a mixed population of M1/M2 cells. *Tumour Biol.* 40 (9), 1010428318797880. [Journal Article]. doi:10.1177/1010428318797880

Shen, S., Zhang, Z., Huang, H., Yang, J., Tao, X., Meng, Z., et al. (2023). Copper-induced injectable hydrogel with nitric oxide for enhanced immunotherapy by amplifying immunogenic cell death and regulating cancer associated fibroblasts. *Biomater. Res.* 27 (1), 44. [Journal Article]. doi:10.1186/s40824-023-00389-4

Shen, W., Yu, Q., Pu, Y., and Xing, C. (2022). Upregulation of long noncoding RNA MALAT1 in colorectal cancer promotes radioresistance and aggressive malignance. *Int. J. Gen. Med.* 15, 8365–8380. [Journal Article]. doi:10.2147/IJGM.S393270

Shigemura, K., Kuribayashi-Shigetomi, K., Tanaka, R., Yamasaki, H., and Okajima, T. (2023). Mechanical properties of epithelial cells in domes investigated using atomic force microscopy. *Front. Cell Dev. Biol.* 11, 1245296. [Journal Article]. doi:10.3389/fcell.2023.1245296

Sustic, I., Racetin, A., Vukojevic, K., Benzon, B., Tonkic, A., Sundov, Z., et al. (2023). Expression pattern of DAB adaptor protein 2 in left- and right-side colorectal carcinoma. *Genes (Basel)* 14 (7), 1306. doi:10.3390/genes14071306

Takahashi, Y., Sugai, T., Habano, W., Ishida, K., Eizuka, M., Otsuka, K., et al. (2016). Molecular differences in the microsatellite stable phenotype between left-sided and right-sided colorectal cancer. *Int. J. Cancer.* 139 (11), 2493–2501. [Journal Article]. doi:10.1002/ijc.30377

Talhouni, S., Fadhil, W., Mongan, N. P., Field, L., Hunter, K., Makhsous, S., et al. (2023). Activated tissue resident memory T-cells (CD8+CD103+CD39+) uniquely predict survival in left sided "immune-hot" colorectal cancers. *Front. Immunol.* 14, 1057292. doi:10.3389/fimmu.2023.1057292

Tian, Z., Zhang, P., Li, X., and Jiang, D. (2023). Analysis of immunogenic cell death in ascending thoracic aortic aneurysms based on single-cell sequencing data. *Front. Immunol.* 14, 1087978. [Journal Article]. doi:10.3389/fimmu.2023.1087978

van der Post, S., and Hansson, G. C. (2014). Membrane protein profiling of human colon reveals distinct regional differences. *Mol. Cell. Proteomics.* 13 (9), 2277–2287. doi:10.1074/mcp.M114.040204

Wang, H., Shemesh-Mayer, E., Zhang, J., Gao, S., Zeng, Z., Yang, Z., et al. (2023a). Genome resequencing reveals the evolutionary history of garlic reproduction traits. *Hortic. Res.* 10 (11), d208. [Journal Article]. doi:10.1093/hr/uhad208

Wang, M., Deng, C., Yang, C., Yan, M., Lu, H., Zhang, Y., et al. (2023b). Unraveling temporal and spatial biomarkers of epithelial-mesenchymal transition in colorectal cancer: insights into the crucial role of immunosuppressive cells. *J. Transl. Med.* 21 (1), 794. doi:10.1186/s12967-023-04600-x

Weiss, J. M., Pfau, P. R., O'Connor, E. S., King, J., LoConte, N., Kennedy, G., et al. (2011). Mortality by stage for right-versus left-sided colon cancer: analysis of surveillance, epidemiology, and end results--Medicare data. *J. Clin. Oncol.* 29 (33), 4401–4409. doi:10.1200/JCO.2011.36.4414

White, M. T., and Sears, C. L. (2023). The microbial landscape of colorectal cancer. *Nat. Rev. Microbiol.* 22, 240–254. [Journal Article; Review]. doi:10.1038/s41579-023-00973-4

Wu, L., Xia, W., Hua, Y., Fan, K., Lu, Y., Wang, M., et al. (2023). Cellular crosstalk of macrophages and therapeutic implications in non-small cell lung cancer revealed by integrative inference of single-cell transcriptomics. *Front. Pharmacol.* 14, 1295442. [Journal Article]. doi:10.3389/fphar.2023.1295442

Xiao, S. H., Wang, Y., Cao, X., and Su, Z. (2021). Long non-coding RNA LUCAT1 inhibits myocardial oxidative stress and apoptosis after myocardial infarction via targeting microRNA-181a-5p. *Bioengineered* 12 (1), 4546–4555. [Journal Article]. doi:10.1080/21655979.2021.1966351

Xie, Z., Niu, L., Zheng, G., Du, K., Dai, S., Li, R., et al. (2023). Single-cell analysis unveils activation of mast cells in colorectal cancer microenvironment. *Cell Biosci.* 13 (1), 217. [Journal Article]. doi:10.1186/s13578-023-01144-x

Yu, S., Chen, M., Xu, L., Mao, E., and Sun, S. (2023). A senescence-based prognostic gene signature for colorectal cancer and identification of the role of SPP1-positive macrophages in tumor senescence. *Front. Immunol.* 14, 1175490. [Journal Article]. doi:10.3389/fimmu.2023.1175490

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 02 May 2024
DOI 10.3389/fgene.2024.1377434

OPEN ACCESS

EDITED BY
Ziheng Wang,
University of Macau, China

REVIEWED BY
Ren Jun,
Yangzhou University, China
Jianfeng Chen,
Wuxi Traditional Chinese Medicine Hospital, China
Yidan Sun,
First Teaching Hospital of Tianjin University of Traditional Chinese Medicine, China

*CORRESPONDENCE
Zekuan Xu,
xuzekuan@njmu.edu.cn

†These authors have contributed equally to this work

RECEIVED 27 January 2024
ACCEPTED 08 April 2024
PUBLISHED 02 May 2024

CITATION
Zhang Y, Wang S, Li Q, Liu H, Xuan Z, Li F, Li Z, Xia Y, Jiang T, Xu P, Fang L, Wang L, Zhang D, Xu H, Yang L and Xu Z (2024), Associations of dietary factors with gastric cancer risk: insights from NHANES 2003–2016 and mendelian randomization analyses. *Front. Genet.* 15:1377434. doi: 10.3389/fgene.2024.1377434

Associations of dietary factors with gastric cancer risk: insights from NHANES 2003–2016 and mendelian randomization analyses

Yigang Zhang[1†], Sen Wang[1†], Qingya Li[1†], Hongda Liu[1], Zhe Xuan[1], Fengyuan Li[1], Zheng Li[1], Yiwen Xia[1], Tianlu Jiang[1], Penghui Xu[1], Lang Fang[1], Linjun Wang[1], Diancai Zhang[1], Hao Xu[1], Li Yang[1] and Zekuan Xu[1,2,3]*

[1]Department of General Surgery, The First Affiliated Hospital of Nanjing Medical University, Nanjing, China, [2]Collaborative Innovation Center for Cancer Personalized Medicine, Nanjing Medical University, Nanjing, China, [3]The Institute of Gastric Cancer, Nanjing Medical University, Nanjing, China

Background: Gastric cancer (GC) continues to be one of the leading causes of cancer-related deaths globally. Diet significantly influences the incidence and progression of GC. However, the relationship between dietary intake and GC is inconsistent.

Methods: A study was conducted with adults who participated in the National Health and Nutrition Examination Survey (NHANES) from 2003 to 2016 to investigate possible associations between 32 dietary factors and GC. To further detect potential causal relationships between these dietary factors and the risk of GC, a two-sample Mendelian randomization (MR) analysis was conducted. The primary method employed was the inverse variance weighted (IVW) analysis, and its results were further validated by four other methods.

Results: Of the 35,098 participants surveyed, 20 had a history of GC. Based on the results of weighted logistic multivariate analysis, it was observed that there was a positive correlation between total fat intake [odds ratio (OR) = 1.09, 95% confidence interval (CI): (1.01–1.17), p = 0.03] and GC as well as negative association of dietary monounsaturated fatty acids (MUFAs) intake [OR = 0.83, 95% CI: (0.76–0.92), $p < 0.001$]. Further evaluations of the odds of GC across the quartiles of dietary MUFAs showed that the top quartile of total MUFA intake was associated with a lower likelihood of GC in three different models [model1: OR = 0.03, 95% CI: (0.00–0.25), $p < 0.01$; model2: OR = 0.04, 95% CI: (0.00–0.38), p = 0.01; model3: OR = 0.04, 95% CI: (0.00–0.40), p = 0.01]. For the MR analyses, genetic instruments were selected from the IEU Open GWAS project; IVW analysis showed that GC risk was not associated with MUFAs [OR = 0.82, 95% CI: (0.59–1.14), p = 0.23] or the ratio of MUFAs to total fatty acids [OR = 1.00, 95% CI: (0.75–1.35), p = 0.98]. Similar results were observed when using the other MR methods.

Conclusion: The NHANES study revealed that consuming MUFAs was linked to a lower risk of GC, although the results of MR analyses do not provide evidence of a causal relationship. Additional research is therefore necessary to clarify these findings.

KEYWORDS

gastric cancer, dietary factors, NHANES, MUFA, Mendelian randomization

Introduction

The rapidly growing global incidence of gastric cancer (GC) presents a significant public health challenge as it remains one of the leading cause of cancer-related mortality (Siegel et al., 2023). Despite advancements in early screening and therapeutic approaches, patients with advanced GC still have poor prognosis (Thrift et al., 2023). The development of GC is multifactorial and involves influences from factors, such as diet, environment, and genetics, with the dietary factors being of particular significance (Bouras et al., 2022). Based on reflection of an old Chinese proverb that "illness comes from the mouth," it is imperative to look into the associations between dietary factors and GC. By gaining a deeper understanding of their relationship, efforts can be made to modify dietary patterns to potentially reduce the incidence of GC.

Recent studies have identified several dietary factors that may be associated with GC; of these, high glucose levels in the body are believed to be linked to greater incidence of malignancies, including GC (Tay et al., 2021). Similarly, increased fat intake has been identified as another important dietary habit that is carcinogenic and potentially related to GC (Kyrgiou et al., 2017). Protein is a fundamental component necessary for body composition and is regarded as a pivotal nutrient for GC patients (Ouyang et al., 2018; Kubota et al., 2020). Furthermore, multiple studies have highlighted the strong positive association between high salt consumption and GC, particularly with respect to salt-preserved foods (Kurosawa et al., 2006; D'Elia et al., 2012). For instance, a recent study reported that high intake of salted fish was linked to an elevated risk of GC (Bouras et al., 2022). On the other hand, high consumption of vitamin C, carotenoids, and other antioxidants, which have the potential to mitigate oxidative damage, has been reported to confer protective effects against the incidence of GC (Kong et al., 2014; Kim et al., 2018; Chen et al., 2021). However, a recent clinical trial found no significant interactions between vitamin supplements and GC incidence (Guo et al., 2020). Nevertheless, it is important to note that most current studies concentrate on a single dietary factor while neglecting the complexity, diversity, and interactions of different dietary intakes. As a result, these reports on the associations between dietary factors and GC may be one-sided. Therefore, it is imperative to shift the focus toward examining food groups or dietary patterns by taking into account multiple dietary factors and conducting comprehensive studies to gain a more holistic understanding of the associations between diet and GC.

Mendelian randomization (MR) is an approach that utilizes genetic variants as instrumental variables (IVs) and offers several advantages over observational studies; it has the potential to circumvent residual confounding and reverse causality, thereby providing a more reliable approach for evaluating causal relationships (Boyko, 2013; Sekula et al., 2016). Therefore, MR was employed in this study to further investigate the causal relationships between some dietary factors of interest and the risk of GC.

Thus, this study aims to investigate the links between dietary factors and the risk of GC by integrating an observational study and two-sample MR analyses. The main goal of this work was to establish a theoretical basis for the prevention and treatment of GC through the improvement of dietary habits.

Methods

Study design and population in NHANES

The data for this cross-sectional study were extracted from the National Health and Nutrition Examination Survey (NHANES), a multistage stratified composite design survey on the health and nutritional information of a representative selection of the non-institutionalized U.S. population, conducted by the National Centers for Health Statistics (NCHS) of the Centers for Disease Control and Prevention (CDC). The observations from seven consecutive NHANES surveys (2003–2004, 2005–2006, 2007–2008, 2009–2010, 2011–2012, 2013–2014, and 2015–2016) were combined into a single analytic sample; thus, a total of 35,098 eligible participants above the age of 18 years, who were interviewed regarding their medical conditions and dietary intakes, were included in this study. The participants who had incomplete information were excluded (n = 6,809).

Variable selection in NHANES

The diagnoses of GC were defined using two items on the Medical Status Questionnaire: "Have you ever been told by a doctor or other health professional that you had cancer or malignancy?" and "What kind of cancer was it?" Answers that indicated only "stomach cancer" were classified as the outcome variables. Some demographic covariates, including age, sex, race, education, smoking status, weight, and body mass index (BMI), were also assessed.

The study participants were asked by trained interviewers to recall two consecutive 24-h dietary periods (day 1 and day 2) to assess the total dietary intakes through comprehensive reference to the NHANES. The present study only included dietary recalls for day 1 as those for day 2 had more missing values. A total of 32 dietary factors from the dietary questionnaire in the NHANES database were included in this study as follows: energy (kcal), protein (g), carbohydrate (g), total sugars (g), dietary fibers (g), total fat (g), saturated fatty acids (SFAs, g), monounsaturated fatty

acids (MUFAs, g), polyunsaturated fatty acids (PUFAs, g), cholesterol (mg), vitamin A (μg), retinol (μg), alpha-carotene (μg), beta-carotene (μg), vitamin B1 (thiamin, mg), vitamin B2 (riboflavin, mg), vitamin B3 (niacin, mg), vitamin B6 (mg), folate (μg), vitamin B12 (μg), vitamin C (μg), vitamin E (mg), vitamin K (μg), calcium (mg), phosphorus (mg), magnesium (mg), iron (mg), zinc (mg), copper (mg), sodium (mg), potassium (mg), and selenium (μg).

Data sources for genetic instruments

The genome-wide association study (GWAS) data analyzed in the present study was obtained from the IEU open GWAS project supported by the MRC Integrative Epidemiology Unit (IEU) at the University of Bristol, collated and analyzed GWAS data from the UK Biobank, FinnGen biobank, and published articles. The single-nucleotide polymorphisms (SNPs) at the genome-wide significance level ($p < 5 \times 10^{-8}$) used in this study included MUFAs (GWAS ID: met-d-MUFA, sample size: 114,999, number of SNPs: 12,321,875, population: European, gender: both) and ratio of MUFAs to total fatty acids (GWAS ID: met-d-MUFA_pct, sample size: 114,999, number of SNPs: 12,321,875, population: European, gender: both). The data on SNPs associated with GC (GWAS ID: finn-b-C3_STOMACH, sample size: 218,792, samples with GC: 633, number of SNPs: 16,380,466, population: European, gender: both) were also extracted from the IEU open GWAS project (https://gwas.mrcieu.ac.uk/).

Genetic instrument selection

The SNPs in linkage disequilibrium (defined as $r^2 > 0.001$ or clump distance <10,000 kb) and those having weaker associations with exposure were excluded, leaving 66 independent SNPs as the IVs for MUFAs and 66 for the ratio of MUFAs to total fatty acids. The F-statistic was used to ensure strong association between the SNPs and exposure. The detailed information on the selected SNPs is presented in Supplementary Tables S1,S2.

Statistical analysis

The data in the current study were obtained and statistically evaluated using R 4.1.1 (R Foundation, Vienna, Austria). The NHANES study population was divided into two groups in accordance with the presence or absence of a history of GC, and characteristics were determined for comparison between the groups. Continuous variables were expressed in terms of the median and interquartile range (IQR) as they did not obey a normal distribution. Significance differences between the two groups were evaluated using the Wilcoxon rank-sum test. Frequency and percent were used to describe the categorical variables, and the distribution of the categorical variables was appropriately compared using the Pearson chi-squared test.

Considering the stratified multistage probabilistic sampling approach of the NHANES, the "survey" package was used to adjust the complex sampling weights in the analyses. The two-year cycle weights were divided by seven to reflect the 14 survey years. Weighted logistic multivariate analysis was used to explore the associations between the dietary factors and GC. Three different models were used to decrease the influences of the confounders, where the first model was the crude model; the second model was adjusted for age, sex, and race, and the third model was adjusted for age, sex, race, education, smoking status, and BMI. The odds ratio (OR) and 95% confidence interval (CI) were used to assess the associations.

For the MR analyses, the "TwoSampleMR" package was used to conduct the inverse variance weighted (IVW) analysis as the primary method of assessing the causal effect between MUFAs and GC risk. The IVW model is considered to have the strongest ability to detect causation in the two-sample MR analysis (Hartwig et al., 2017). MR–Egger, weighted-median, simple mode, and weighted mode were also implemented to validate the results from the IVW analysis. The possible heterogeneity and directional pleiotropy were assessed through the Cochrane Q test and intercept from MR–Egger (Qian et al., 2020). The leave-one-out sensitivity analysis was also conducted, and a p-value <0.05 (two-sided) was considered to be statistically significant in this study.

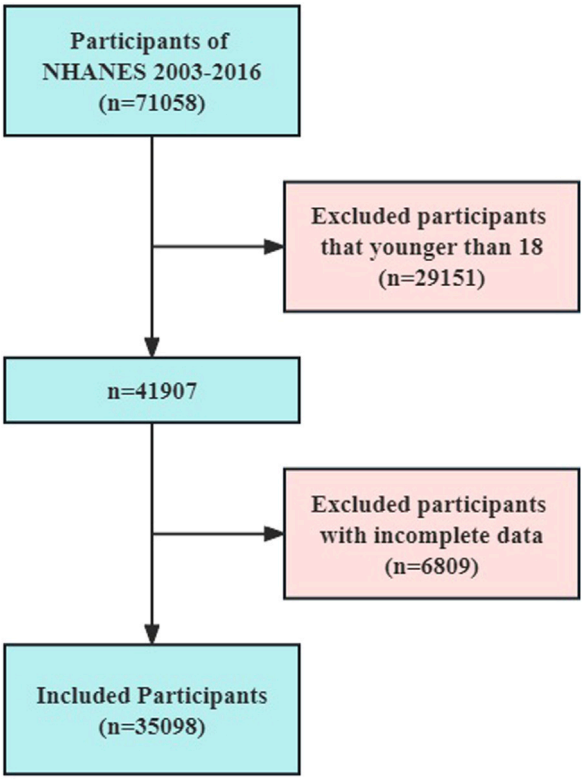


FIGURE 1
Flowchart of the NHANES study participants.

Results

Characteristics of included participants

A total of 35,098 individuals (weighted n = 219,465,579) over 18 years of age were selected for this study through the NHANES database. The flowchart illustrating the selection process of the

TABLE 1 General characteristics of the adults included in this study stratified by the presence or absence of a history of GC.

Variable	Total (unweighted n = 35,098, weighted n = 219,465,579)	GC (unweighted n = 20, weighted n = 105,634)	No GC (unweighted n = 35,078, weighted n = 219,359,945)	*p*-value
Age (median (IQR))	46 (32.59)	53 (48.67)	46 (32.59)	< **0.0001**
Sex (%)				0.06
Male	48.16	22.06	48.17	
Female	51.84	77.94	51.83	
Race (%)				0.51
Non-Hispanic Black	11.34	20.95	11.34	
Non-Hispanic White	68.44	71.97	68.44	
Mexican American	8.48	3.53	8.49	
Other Hispanic	4.9	0	4.9	
Other race	6.84	3.56	6.84	
Education (%)				**0.03**
Less than high school	16.97	27.5	16.96	
High school graduate	23.64	6.51	23.65	
Some college	31.6	58.87	31.59	
College graduate or above	27.79	7.12	27.8	
Smoking status (%)				0.9
No	78.36	76.9	78.36	
Yes	21.64	23.1	21.64	
Weight (median (IQR))	79.20 (67.00.93.70)	67.40 (67.40.81.40)	79.30 (67.00.93.70)	**0.02**
BMI (median (IQR))	27.61 (24.03.32.20)	25.90 (24.40.28.20)	27.61 (24.02.32.20)	0.1

Data are expressed as median (IQR) for skewed variables and percentage (%) for categorical variables. *p*-value for skewed variables was assessed by the Wilcoxon rank-sum test, and *p*-value for the categorical variables was determined using the Pearson chi-squared test.
The bold values mean $p < 0.05$.

participants is depicted in Figure 1. Among these participants, 20 people representing 105,634 individuals reported having a history of GC. The characteristics of these individuals were then stratified based on the presence or absence of GC, as shown in Table 1. The analysis revealed that individuals with GC were older (53 vs 46 years, $p < 0.0001$), had lower weights (67.40 vs 79.30 kg, $p = 0.02$), and had lower educational attainment ($p = 0.03$) compared to those without GC.

Dietary intakes and risk of GC

Table 2 represents the dietary intakes of the participants with and without GC. Individuals with GC consumed less energy ($p < 0.001$), carbohydrates ($p = 0.01$), dietary fibers ($p < 0.0001$), total fats ($p = 0.03$), MUFAs ($p < 0.01$), PUFAs ($p = 0.04$), vitamin B1 ($p < 0.0001$), vitamin B3 ($p < 0.0001$), vitamin B6 ($p < 0.0001$), folate ($p < 0.0001$), vitamin E ($p < 0.0001$), vitamin K ($p < 0.0001$), phosphorus ($p = 0.01$), magnesium (<0.0001), iron ($p < 0.0001$), zinc ($p < 0.01$), copper ($p = 0.01$), sodium ($p = 0.02$), and selenium ($p = 0.03$).

The correlations between the aforementioned dietary intakes and risk of GC via logistic regression analysis after adjustment for multiple potential confounders are described in Table 3. The dietary total fat intake [OR = 1.09, 95% CI: (1.01–1.17), $p = 0.03$] was positively associated with GC; meanwhile, dietary MUFA intake [OR = 0.83, 95% CI: (0.76–0.92), $p < 0.001$] was negatively associated with GC. After adjusting for age, sex, and race

TABLE 2 Comparison of dietary intakes of persons with and without self-reported GC.

Variable	Total	GC	No GC	*p*-value
Energy (kcal)	1,988.00 (1,479.00–2,649.00)	1,615.00 (1,038.00–1615.00)	1,989.00 (1,479.00–2,649.00)	< **0.001**
Proteins (g)	75.81 (54.17–103.46)	51.95 (44.83–97.60)	75.82 (54.19–103.48)	0.15
Carbohydrates (g)	236.50 (171.86–319.67)	192.76 (94.95,218.33)	236.56 (171.88–319.74)	**0.01**
Total sugars (g)	99.26 (61.94–150.26)	111.15 (38.91,111.15)	99.26 (61.98–150.26)	0.41
Dietary fibers (g)	14.70 (9.70–21.30)	9.10 (5.60–9.30)	14.70 (9.70–21.30)	< **0.0001**
Total fats (g)	74.42 (50.20–105.17)	56.02 (29.94–71.58)	74.44 (50.20–105.17)	**0.03**
SFAs (g)	23.69 (15.21–35.20)	17.39 (15.33–26.14)	23.69 (15.21–35.20)	0.19
MUFAs (g)	26.49 (17.38–38.30)	21.63 (9.29–22.84)	26.50 (17.38–38.30)	**0.002**
PUFAs (g)	15.82 (10.02–23.87)	10.90 (2.47–15.98)	15.82 (10.02–23.87)	**0.04**
Cholesterol (mg)	223.00 (131.00–380.00)	272.00 (185.00–556.00)	223.00 (131.00–380.00)	0.47
Vitamin A (μg)	494.00 (274.00–812.00)	464.00 (464.00–857.00)	494.00 (274.00–812.00)	0.43
Retinol (μg)	334.00 (171.00–566.00)	443.00 (321.00–453.00)	334.00 (171.00–566.00)	0.08
Alpha carotene (μg)	47.00 (10.00–244.00)	13.00 (0.00–158.00)	47.00 (11.00–244.00)	0.48
Beta carotene (μg)	779.00 (300.00–2,312.00)	555.00 (214.00–4,768.00)	779.00 (300.00–2,312.00)	0.78
Vitamin B1 (mg)	1.46 (1.02–2.03)	0.95 (0.63–1.13)	1.46 (1.02–2.03)	< **0.0001**
Vitamin B2 (mg)	1.97 (1.36–2.73)	2.04 (1.88–2.06)	1.97 (1.36–2.73)	0.51
Vitamin B3 (mg)	22.73 (15.89–31.80)	13.24 (13.24–18.81)	22.74 (15.90–31.81)	< **0.0001**
Vitamin B6 (mg)	1.77 (1.18–2.54)	0.91 (0.65–1.45)	1.77 (1.18–2.54)	< **0.0001**
Folate (μg)	354.00 (240.00–511.00)	247.00 (170.00–287.00)	354.00 (240.00–512.00)	< **0.0001**
Vitamin B12 (μg)	3.96 (2.25–6.47)	2.79 (2.77–4.28)	3.96 (2.25–6.47)	0.15
Vitamin C (mg)	53.40 (21.90–115.60)	26.90 (9.30–110.50)	53.50 (21.90–115.60)	0.38
Vitamin E (mg)	6.72 (4.30–10.27)	4.18 (1.81–4.69)	6.72 (4.31–10.27)	< **0.0001**
Vitamin K (μg)	63.80 (35.80–119.30)	34.10 (25.00–49.30)	63.80 (35.80–119.40)	< **0.0001**
Ca (mg)	834.00 (546.00–1,223.00)	810.00 (571.00–1,009.00)	834.00 (546.00–1,223.00)	0.55
P (mg)	1,263.00 (915.00–1,717.00)	966.00 (943.00–1,286.00)	1,263.00 (915.00–1,717.00)	**0.01**
Mg (mg)	275.00 (198.00–370.00)	179.00 (179.00–264.00)	275.00 (198.00–370.00)	< **0.0001**
Fe (mg)	13.27 (9.33–18.92)	8.80 (5.47–10.69)	13.28 (9.33–18.92)	< **0.0001**
Zn (mg)	10.17 (6.97–14.71)	6.47 (5.14–9.87)	10.17 (6.98–14.71)	**0.002**
Cu (mg)	1.14 (0.82–1.58)	0.55 (0.54–1.00)	1.14 (0.82–1.58)	**0.01**
Na (mg)	3,217.00 (2,287.00–4,406.00)	2,856.00 (1,902.00–3,101.00)	3,218.00 (2,287.00–4,408.00)	**0.02**
K (mg)	2,535.00 (1,820.00–3,365.00)	1,492.00 (1,350.00–2,531.00)	2,536.00 (1,821.00–3,365.00)	0.06
Se (μg)	101.90 (70.40–142.10)	80.40 (63.30–103.80)	101.90 (70.40–142.20)	**0.03**

p-value was determined by the Wilcoxon rank-sum test, median (IQR).
The bold values mean $p < 0.05$.

(model2) as well as age, sex, race, education, smoking status, and BMI (model3), the dietary total fat intake [model2: OR = 1.08, 95% CI: (1.01–1.16), $p = 0.02$; model3: OR = 1.08, 95% CI: (1.01–1.16), $p = 0.03$] was still associated with higher odds of GC, and the dietary MUFA intake [model2: OR = 0.83, 95% CI: (0.75–0.91), $p < 0.001$; model3: OR = 0.83, 95% CI: (0.76–0.92), $p < 0.001$] was still associated with a lower risk of GC. The average total fat intake of the participants with GC in this study was still within the 25%–35% range recommended by the 2020–2025 Dietary Guidelines for Americans (Phillips, 2021). Despite the positive association between total fat intake and GC, it is important to note that the lower daily energy intakes of those with GC than without GC may have contributed to this relationship. Therefore, it is plausible that this association

TABLE 3 Associations between dietary intakes and GC.

Variable	Model1		Model2		Model3	
	OR (95% CI)	*p*-value	OR (95% CI)	*p*-value	OR (95% CI)	*p*-value
Energy (kcal)	0.9972 (0.9918–1.0027)	0.3180	0.9978 (0.9928–1.0027)	0.3724	0.9976 (0.9922–1.0031)	0.3838
Carbohydrates (g)	1.0073 (0.9834–1.0318)	0.5471	1.0072 (0.9850–1.0299)	0.5227	1.0071 (0.9831–1.0317)	0.5596
Dietary fibers (g)	0.9574 (0.8970–1.0220)	0.1885	0.9495 (0.8813–1.0230)	0.1708	0.9605 (0.8836–1.0441)	0.3393
Total fats (g)	1.0866 (1.0099–1.1691)	**0.0266**	1.0818 (1.0114–1.1572)	**0.0226**	1.0820 (1.0069–1.1627)	**0.0321**
MUFAs (g)	0.8343 (0.7572–0.9192)	**<0.001**	0.8302 (0.7546–0.9135)	**<0.001**	0.8338 (0.7562–0.9194)	**<0.001**
PUFAs (g)	0.9688 (0.8677–1.0817)	0.5695	0.9701 (0.8634–1.0898)	0.6049	0.9727 (0.8636–1.0955)	0.6439
Vitamin B1 (mg)	0.8809 (0.1363–5.6913)	0.8929	0.8134 (0.1107–5.9757)	0.8373	0.8181 (0.1002–6.6778)	0.8495
Vitamin B3 (mg)	0.9955 (0.8870–1.1172)	0.9378	1.0059 (0.9016–1.1222)	0.9151	1.0069 (0.9113–1.1125)	0.8914
Vitamin B6 (mg)	0.6444 (0.3578–1.1605)	0.1413	0.6016 (0.3416–1.0592)	0.0776	0.5989 (0.3137–1.1432)	0.1185
Folate (μg)	0.9995 (0.9957–1.0034)	0.8137	1.0000 (0.9962–1.0037)	0.9899	0.9999 (0.9957–1.0041)	0.9707
Vitamin E (mg)	0.9831 (0.8378–1.1535)	0.8325	0.9804 (0.8212–1.1705)	0.8249	0.9847 (0.8317–1.1659)	0.8565
Vitamin K (μg)	0.9927 (0.9790–1.0065)	0.2951	0.9916 (0.9754–1.0080)	0.3073	0.9917 (0.9765–1.0072)	0.2897
P (mg)	1.0010 (0.9982–1.0038)	0.4776	1.0011 (0.9984–1.0037)	0.4293	1.0011 (0.9984–1.0039)	0.4174
Mg (mg)	1.0026 (0.9943–1.0109)	0.5358	1.0023 (0.9947–1.0100)	0.5479	1.0018 (0.9948–1.0089)	0.6067
Fe (mg)	0.9477 (0.8548–1.0507)	0.3037	0.9304 (0.8349–1.0367)	0.1885	0.9317 (0.8549–1.0153)	0.1053
Zn (mg)	1.0073 (0.9242–1.0979)	0.8666	1.0141 (0.9226–1.1146)	0.7695	1.0142 (0.9312–1.1047)	0.743
Cu (mg)	0.8428 (0.0795–8.9384)	0.8859	0.7847 (0.0777–7.9216)	0.8353	0.7996 (0.1129–5.6627)	0.8207
Na (mg)	1.0002 (0.9995–1.0008)	0.5906	1.0003 (0.9996–1.0009)	0.4116	1.0003 (0.9997–1.0009)	0.3561
Se (μg)	1.0046 (0.9786–1.0313)	0.7304	1.0043 (0.9785–1.0308)	0.7433	1.0042 (0.9792–1.0298)	0.7414

OR, odds ratio; CI, confidence interval.
Model1: crude model.
Model2: adjusted for age, sex, and race.
Model3: adjusted for age, sex, race, education, smoking status, and BMI.
The bold values mean $p < 0.05$.

may be influenced by the differences in daily energy intakes between the two groups.

MUFAs intake and GC

To further study the associations between MUFAs and risk of GC, the above three models were used to evaluate the odds of GC across the quartiles (Q1: <16.512 g, Q2: 16.512–25.429 g, Q3: 25.429–37.203 g, Q4: >37.203 g) of total intake of MUFAs (Table 4). In all three models, the top quartile of total MUFA intake had over 90% lower likelihood of GC [model1: OR = 0.03, 95% CI: (0.00–0.25), $p < 0.01$; model2: OR = 0.04, 95% CI: (0.00–0.38), $p = 0.01$; model3: OR = 0.04, 95% CI: (0.00–0.40), $p = 0.01$]. These results indicate that a diet rich in MUFAs might play a protective role against GC.

The stability of the correlation between MUFAs and GC risk was further confirmed in different populations (Table 5). Analyses stratified by race show that MUFA intake was associated with lower GC risk in black participants [OR = 0.96, 95% CI: (0.92–0.99), $p = 0.02$] and white participants [OR = 0.92, 95% CI: (0.84–1.00), $p = 0.05$]. In stratified analyses based on smoking status, MUFAs were significantly correlated with GC risk in non-smokers [OR = 0.94, 95% CI: (0.89–1.00), $p = 0.04$]. Stratification by BMI showed that MUFAs were significantly associated with lower GC risk only in people with healthy weights [BMI: 18.5–25, OR = 0.89, 95% CI: (0.87–0.91), $p < 0.0001$]. Overall, the findings of this study indicate that a high dietary intake of MUFAs decreases the risk of GC.

Causal relationship between MUFAs and GC risk

The cross-sectional study design of NHANES prevented the establishment of a causal relationship between the dietary factors and risk of GC. To avoid this limitation, MR analyses were conducted, and details of these SNPs are given in Supplementary Tables S1,S2. The F-statistic for each SNP was above 10. From the results of the IVW analyses, there were no genetic instruments associated with MUFAs or ratio of MUFAs to total fatty acids having a causal relationship with GC risk. The pooled ORs for GC risk in

TABLE 4 Odds ratios and 95% confidence intervals for GC according to the daily dietary MUFA intake level.

Variable	Model1		Model2		Model3	
	OR (95% CI)	*p*-value	OR (95% CI)	*p*-value	OR (95% CI)	*p*-value
Q1	ref		ref		ref	
Q2	0.9168 (0.1380–6.0891)	0.9277	0.9990 (0.1497–6.6683)	0.9992	1.0181 (0.1508–6.8718)	0.9852
Q3	0.2512 (0.0583–1.0819)	0.0634	0.3106 (0.0727–1.3281)	0.1135	0.3393 (0.0797–1.4449)	0.1419
Q4	0.0260 (0.0028–0.2451)	**0.0017**	0.0400 (0.0043–0.3762)	**0.0053**	0.0426 (0.0045–0.4048)	**0.0065**
P for trend		**0.0023**		**0.0119**		**0.0156**

OR: odds ratio; CI: confidence interval.
Model1: crude model.
Model2: adjusted for age, sex, and race.
Model3: adjusted for age, sex, race, education, smoking status, and BMI.
The bold values mean $p < 0.05$.

genetically predicted per unit change were 0.82 (95% CI: 0.59–1.14; $p = 0.23$) and 1.00 (95% CI: 0.75–1.35; $p = 0.98$) for MUFAs and ratio of MUFAs to total fatty acids, respectively (Table 6; Figure 2). There was no evidence of heterogeneity or pleiotropy of the aforementioned IVW analysis (Supplementary Table S3).

Four other MR methods (MR–Egger, weighted-median, simple mode, and weighted mode) were conducted, and similar results were observed to those of the IVW analyses (Table 6; Figure 2). For both genetic instruments, MR-PRESSO was conducted, and no outliers were found in this study. The leave-one-out sensitivity analysis showed that the overall result could change upon removing rs964184 for both MUFAs and ratio of MUFAs to total fatty acids as well as rs174564 for MUFAs (Supplementary Figure S1).

TABLE 5 Subgroup analysis of the associations between MUFAs and GC.

Variable	OR (95% CI)	*p*-value
Age		
<60	0.9493 (0.8969–1.0049)	0.0726
>60	0.9071 (0.8156–1.0088)	0.0718
Race		
Non-Hispanic Black	0.9572 (0.9220–0.9938)	**0.0227**
Non-Hispanic White	0.9169 (0.8411–0.9995)	**0.0488**
Mexican American	1.0251 (0.9768–1.0758)	0.3103
Other Hispanic	0.9987 (0.9956–1.0019)	0.4225
Other race	0.9823 (0.9059–1.0653)	0.6637
Smoking status		
No	0.9416 (0.8894–0.9969)	**0.0388**
Yes	0.9325 (0.8330–1.0440)	0.2228
BMI		
Underweight	0.8893 (0.8702–0.9088)	**<0.0001**
Healthy weight	0.9733 (0.9350–1.0133)	0.1855
Overweight	0.9304 (0.8538–1.0138)	0.0988
Obese	0.9479 (0.8383–1.0719)	0.3904

The bold values mean $p < 0.05$.

Discussion

Numerous studies have unequivocally demonstrated the strong relationships between dietary intake and risk of developing various types of cancers, including lung cancer (Sun et al., 2016), breast cancer (De Cicco et al., 2019), and colorectal cancer (Thanikachalam and Khan, 2019). While some studies have underscored the significant roles of specific nutrients, such as cholesterol (Pih et al., 2021), nitrates (Poorolajal et al., 2020), salt (Wu et al., 2021), and alcohol (Laszkowska et al., 2021), in the incidence and progression of GC, a common tendency in these investigations is the exclusive focus on individual dietary factors that neglects the intricate interplay and complexities of different dietary intakes. It is important to recognize that modifications to a single dietary factor can invariably lead to compensatory changes in other dietary characteristics. Thus, the present cross-sectional study sought to address this deficiency by examining the NHANES database to elucidate the relationships between 32 dietary factors and risk of GC. The findings from a weighted logistic multivariate analysis reveal a noteworthy association between the intake of MUFAs and reduced risk of GC [OR = 0.8343, 95% CI: (0.7572–0.9192), $p < 0.001$]. This underscores the significance of considering the comprehensive dietary landscape for understanding the multiple factors at play in the development of GC.

Extensive research has been conducted on MUFAs owing to their potential health benefits (Snaebjornsson et al., 2020). Foods such as olive oil, avocados, nuts, and seeds, which are common components of the Mediterranean diet, have abundant quantities of MUFAs, and this diet is renowned for its benefits against cardiovascular diseases, obesity, and malignancies (Davis et al., 2015; Schwingshackl et al., 2017; Morze et al., 2021). Current studies advocate the restriction of SFAs and incorporation of higher proportions of MUFAs and PUFAs into a healthy diet (Sacks et al., 2017). Moreover, a recent comprehensive meta-analysis involving 3,202,496 participants revealed an inverse association between the Mediterranean diet and mortality rates of several cancers, including GC, highlighting the potential role of

TABLE 6 Causal relationship between MUFAs and GC risk based on different MR methods.

Exposure	MR method	OR	95% CI	*p*-value
MUFAs	IVW	0.8214	0.5939–1.1360	0.2344
	MR–Egger	1.0091	0.5932–1.7165	0.9735
	Weighted-median	1.0375	0.6203–1.7354	0.8883
	Simple mode	0.7523	0.2980–1.8992	0.5493
	Weighted mode	1.1843	0.6340–2.2122	0.5978
Ratio of MUFAs to total fatty acids	IVW	1.0039	0.7475–1.3482	0.9794
	MR–Egger	1.2926	0.8252–2.0246	0.2668
	Weighted-median	1.3905	0.8750–2.2095	0.1630
	Simple mode	1.1772	0.5234–2.6480	0.6946
	Weighted mode	1.3437	0.8680–2.0803	0.1902

OR, odds ratio; CI, confidence interval.

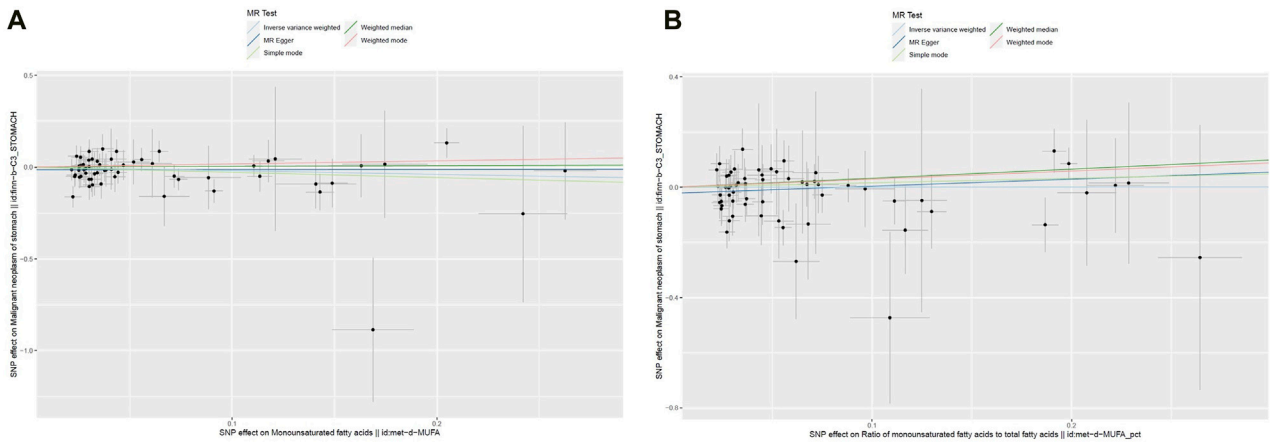


FIGURE 2
Scatter plots of the genetic associations between **(A)** MUFAs and **(B)** ratio of MUFAs to total fatty acids and GC.

MUFAs in protecting against GC (Morze et al., 2021). Additionally, two separate studies found that dietary MUFAs were linked to a reduced risk of pancreatic cancer (Nkondjock et al., 2005; Banim et al., 2018). The potential antitumor effects of MUFAs may be attributed to their antioxidant properties, capacity to reduce chronic inflammation, and cholesterol-lowering properties (Farag and Gad, 2022; Wan et al., 2022; Guo et al., 2023). Furthermore, in the present NHANES observational study, after adjusting for potential confounders, the authors observed a protective effect of a high-MUFAs diet against GC, consistent with the findings of the aforementioned studies. These findings collectively suggest the potential of MUFAs in mitigating the risk of certain types of cancer while emphasizing their value as a component of a health-promoting diet.

However, some studies have reported conflicting results regarding the association between MUFAs and cancer risk. A large-scale case-controlled study revealed that increased MUFA intake was linked to higher odds of breast cancer (Sasanfar et al., 2022). Similarly, another study reported a positive association between dietary MUFA intake and pancreatic cancer (Gong et al., 2010). Therefore, evidence regarding the relationship between dietary MUFAs and cancer risk remains inconclusive. It is worth noting that there have been no cohort studies to investigate the association between dietary MUFAs and GC to date. This indicates the need for further research to clarify the impacts of MUFA consumption on GC risk.

The primary limitation of an observational study is the challenge of establishing a causal relationship. To address this limitation, two-sample MR analyses were conducted to investigate any potential causal relationships between MUFAs or ratio of MUFAs to total fatty acids and the risk of GC. The initial findings from the IVW analysis do not support a causal relationship between MUFAs or ratio of MUFAs to total fatty acids and GC. Furthermore, this study incorporated four additional MR analyses, all of which consistently aligned with the findings of the IVW analysis, thereby enhancing the robustness of the findings. Despite the seemingly contradictory results between the NHANES observational

study and MR analyses, it would be premature to conclusively label MUFAs as "ineffective" in mitigating GC risk. The complexity of dietary factor interactions and metabolism in the human body suggests that compensatory mechanisms may occur when the intake of a specific dietary factor is altered over a short period of time. In light of this intricate interplay, it is more prudent to view each dietary factor as an integral component, akin to individual bricks contributing to the construction of a "great wall" that safeguards against malignancies such as GC.

To the best of the authors' knowledge, this study is an initial attempt to examine the correlations between dietary MUFAs intake and risk of GC by integrating an observational study with two-sample MR analyses, thereby enhancing the reliability of the findings. Although the results derived from the MR analyses do not substantiate a causal role, it remains imperative to further investigate whether increased consumption of MUFA-rich foods exerts a protective effect against GC.

Despite the findings of this study, several limitations should be considered. First, the relatively lower incidence of GC in the USA compared to East Asia implies a need for further analyses using databases of Asian participants and GWAS to bolster the findings. Second, the use of self-reported 24-h dietary recall data in NHANES may not be fully representative of the participants' long-term dietary intakes. Third, the absence of information on GC staging, histological findings, surgical history, and fatalities among the study participants precludes subgroup analyses based on the cancer stages, potentially affecting the results. Lastly, this study did not delineate the MUFAs based on their derivation from animal or plant sources, introducing the possibility of biases.

Conclusion

In conclusion, the results of the present study show no evidence to support a causal link between MUFA intake and gastric cancer risk. Larger studies are therefore required to explore the potential associations between GC risk and animal- or plant-derived MUFAs.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material; further inquiries can be directed to the corresponding authors.

Author contributions

YZ: conceptualization and writing–original draft. SW: investigation and writing–original draft. QL: software and writing–original draft. HL: data curation and writing–review and editing. ZX: methodology and writing–review and editing. FL: investigation and writing–review and editing. ZL: investigation and writing–review and editing. YX: data curation and writing–review and editing. TJ: validation and writing–review and editing. PX: project administration and writing–review and editing. LF: formal analysis and writing–review and editing. LW: supervision and writing–review and editing. DZ: supervision and writing–review and editing. HX: supervision and writing–review and editing. LY: supervision and writing–review and editing. XZ: conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, resources, software, supervision, validation, visualization, writing–original draft, and writing–review and editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This work was partially supported by the National Natural Science Foundation of China (grant nos 81871946, 82072708, and 82202849); the Special Foundation for National Science and Technology Basic Research Program of China (grant no. 2019FY101104); the Primary Research & Development Plan of Jiangsu Province (grant no. BE2016786); the Program for Development of Innovative Research Team in the First Affiliated Hospital of NJMU; the Priority Academic Program Development of Jiangsu Higher Education Institutions (grant nos PAPD, JX10231, and 801); the Jiangsu Key Medical Discipline (General Surgery) (grant no. ZDXKA2016005); the Jiangsu Key Lab of Cancer Biomarkers, Prevention and Treatment, Collaborative Innovation Centre for Cancer Personalized Medicine, Nanjing Medical University; the China Postdoctoral Science Foundation (grant no. 2021TQ0132); and The Youth Fund Program for the National Natural Science Foundation of China from the First Affiliated Hospital of Nanjing Medical University (grant no. PY2021032).

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2024.1377434/full#supplementary-material

References

Banim, P. J., Luben, R., Khaw, K. T., and Hart, A. R. (2018). Dietary oleic acid is inversely associated with pancreatic cancer - data from food diaries in a cohort study. *Pancreatol.* 18 (6), 655–660. doi:10.1016/j.pan.2018.07.004

Bouras, E., Tsilidis, K. K., Triggi, M., Siargkas, A., Chourdakis, M., and Haidich, A. B. (2022). Diet and risk of gastric cancer: an umbrella review. *Nutrients* 14 (9), 1764. doi:10.3390/nu14091764

Boyko, E. J. (2013). Observational research--opportunities and limitations. *Journal of diabetes and its complications* 27 (6), 642–648. doi:10.1016/j.jdiacomp.2013.07.007

Casanova, N., Beaulieu, K., Finlayson, G., and Hopkins, M. (2019). Metabolic adaptations during negative energy balance and their potential impact on appetite and food intake. *The Proceedings of the Nutrition Society* 78 (3), 279–289. doi:10.1017/S0029665118002811

Chen, Z., Huang, Y., Cao, D., Qiu, S., Chen, B., Li, J., et al. (2021). Vitamin C intake and cancers: an umbrella review. *Frontiers in nutrition* 8, 812394. doi:10.3389/fnut.2021.812394

Davis, C., Bryan, J., Hodgson, J., and Murphy, K. (2015). Definition of the mediterranean diet; a literature review. *Nutrients* 7 (11), 9139–9153. doi:10.3390/nu7115459

De Cicco, P., Catani, M. V., Gasperi, V., Sibilano, M., Quaglietta, M., and Savini, I. (2019). Nutrition and breast cancer: a literature review on prevention, treatment and recurrence. *Nutrients* 11 (7), 1514. doi:10.3390/nu11071514

D'Elia, L., Rossi, G., Ippolito, R., Cappuccio, F. P., and Strazzullo, P. (2012). Habitual salt intake and risk of gastric cancer: a meta-analysis of prospective studies. *Clinical nutrition (Edinburgh, Scotland)* 31 (4), 489–498. doi:10.1016/j.clnu.2012.01.003

Farag, M. A., and Gad, M. Z. (2022). Omega-9 fatty acids: potential roles in inflammation and cancer management. *Journal, genetic engineering and biotechnology* 20 (1), 48. doi:10.1186/s43141-022-00329-0

Gong, Z., Holly, E. A., Wang, F., Chan, J. M., and Bracci, P. M. (2010). Intake of fatty acids and antioxidants and pancreatic cancer in a large population-based case-control study in the San Francisco Bay Area. *International journal of cancer* 127 (8), 1893–1904. doi:10.1002/ijc.25208

Guo, Y., Li, Z. X., Zhang, J. Y., Ma, J. L., Zhang, L., Zhang, Y., et al. (2020). Association between lifestyle factors, vitamin and garlic supplementation, and gastric cancer outcomes: a secondary analysis of a randomized clinical trial. *JAMA network open* 3 (6), e206628. doi:10.1001/jamanetworkopen.2020.6628

Guo, Z., Bergeron, K. F., Lingrand, M., and Mounier, C. (2023). Unveiling the MUFA-cancer connection: insights from endogenous and exogenous perspectives. *International journal of molecular sciences* 24 (12), 9921. doi:10.3390/ijms24129921

Hartwig, F. P., Davey Smith, G., and Bowden, J. (2017). Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. *International journal of epidemiology* 46 (6), 1985–1998. doi:10.1093/ije/dyx102

Kim, J. H., Lee, J., Choi, I. J., Kim, Y. I., Kwon, O., Kim, H., et al. (2018). Dietary carotenoids intake and the risk of gastric cancer: a case-control study in korea. *Nutrients* 10 (8), 1031. doi:10.3390/nu10081031

Kong, P., Cai, Q., Geng, Q., Wang, J., Lan, Y., Zhan, Y., et al. (2014). Vitamin intake reduce the risk of gastric cancer: meta-analysis and systematic review of randomized and observational studies. *PloS one* 9 (12), e116060. doi:10.1371/journal.pone.0116060

Kubota, T., Shoda, K., Konishi, H., Okamoto, K., and Otsuji, E. (2020). Nutrition update in gastric cancer surgery. *Annals of gastroenterological surgery* 4 (4), 360–368. doi:10.1002/ags3.12351

Kurosawa, M., Kikuchi, S., Xu, J., and Inaba, Y. (2006). Highly salted food and mountain herbs elevate the risk for stomach cancer death in a rural area of Japan. *Journal of gastroenterology and hepatology* 21 (11), 1681–1686. doi:10.1111/j.1440-1746.2006.04290.x

Kyrgiou, M., Kalliala, I., Markozannes, G., Gunter, M. J., Paraskevaidis, E., Gabra, H., et al. (2017). Adiposity and cancer at major anatomical sites: umbrella review of the literature. *BMJ (Clinical research ed)* 356, j477. doi:10.1136/bmj.j477

Laszkowska, M., Rodriguez, S., Kim, J., and Hur, C. (2021). Heavy alcohol use is associated with gastric cancer: analysis of the national health and nutrition examination survey from 1999 to 2010. *The American journal of gastroenterology* 116 (5), 1083–1086. doi:10.14309/ajg.0000000000001166

Morze, J., Danielewicz, A., Przybyłowicz, K., Zeng, H., Hoffmann, G., and Schwingshackl, L. (2021). An updated systematic review and meta-analysis on adherence to mediterranean diet and risk of cancer. *European journal of nutrition* 60 (3), 1561–1586. doi:10.1007/s00394-020-02346-6

Nkondjock, A., Krewski, D., Johnson, K. C., Ghadirian, P., and Canadian Cancer Registries Epidemiology Research Group (2005). Specific fatty acid intake and the risk of pancreatic cancer in Canada. *British journal of cancer* 92 (5), 971–977. doi:10.1038/sj.bjc.6602380

Ouyang, X., Dang, Y., Zhang, F., and Huang, Q. (2018). Low serum albumin correlates with poor survival in gastric cancer patients. *Clinical laboratory* 64 (3), 239–245. doi:10.7754/Clin.Lab.2017.170804

Phillips, J. A. (2021). Dietary Guidelines for Americans, 2020-2025. *Workplace health and safety* 69 (8), 395. doi:10.1177/21650799211026980

Pih, G. Y., Gong, E. J., Choi, J. Y., Kim, M. J., Ahn, J. Y., Choe, J., et al. (2021). Associations of serum lipid level with gastric cancer risk, pathology, and prognosis. *Cancer research and treatment* 53 (2), 445–456. doi:10.4143/crt.2020.599

Poorolajal, J., Moradi, L., Mohammadi, Y., Cheraghi, Z., and Gohari-Ensaf, F. (2020). Risk factors for stomach cancer: a systematic review and meta-analysis. *Epidemiology and health* 42, e2020004. doi:10.4178/epih.e2020004

Qian, Y., Ye, D., Huang, H., Wu, D. J. H., Zhuang, Y., Jiang, X., et al. (2020). Coffee consumption and risk of stroke: a mendelian randomization study. *Annals of neurology* 87 (4), 525–532. doi:10.1002/ana.25693

Sacks, F. M., Lichtenstein, A. H., Wu, J. H. Y., Appel, L. J., Creager, M. A., Kris-Etherton, P. M., et al. (2017). Dietary fats and cardiovascular disease: a presidential advisory from the American heart association. *Circulation* 136 (3), e1–e23. doi:10.1161/CIR.0000000000000510

Sasanfar, B., Toorang, F., Zendehdel, K., and Salehi-Abargouei, A. (2022). Substitution of dietary macronutrients and their sources in association with breast cancer: results from a large-scale case-control study. *European journal of nutrition* 61 (5), 2687–2695. doi:10.1007/s00394-022-02811-4

Schwingshackl, L., Schwedhelm, C., Galbete, C., and Hoffmann, G. (2017). Adherence to mediterranean diet and risk of cancer: an updated systematic review and meta-analysis. *Nutrients* 9 (10), 1063. doi:10.3390/nu9101063

Sekula, P., Del Greco, M. F., Pattaro, C., and Köttgen, A. (2016). Mendelian randomization as an approach to assess causality using observational data. *Journal of the American Society of Nephrology JASN* 27 (11), 3253–3265. doi:10.1681/ASN.2016010098

Siegel, R. L., Miller, K. D., Wagle, N. S., and Jemal, A. (2023). Cancer statistics, 2023. *CA a cancer journal for clinicians* 73 (1), 17–48. doi:10.3322/caac.21763

Snaebjornsson, M. T., Janaki-Raman, S., and Schulze, A. (2020). Greasing the wheels of the cancer machine: the role of lipid metabolism in cancer. *Cell metabolism* 31 (1), 62–76. doi:10.1016/j.cmet.2019.11.010

Sun, Y., Li, Z., Li, J., Li, Z., and Han, J. (2016). A healthy dietary pattern reduces lung cancer risk: a systematic review and meta-analysis. *Nutrients* 8 (3), 134. doi:10.3390/nu8030134

Tay, S. W., Li, J. W., and Fock, K. M. (2021). Diet and cancer of the esophagus and stomach. *Current opinion in gastroenterology* 37 (2), 158–163. doi:10.1097/MOG.0000000000000700

Thanikachalam, K., and Khan, G. (2019). Colorectal cancer and nutrition. *Nutrients* 11 (1), 164. doi:10.3390/nu11010164

Thrift, A. P., Wenker, T. N., and El-Serag, H. B. (2023). Global burden of gastric cancer: epidemiological trends, risk factors, screening and prevention. *Nature reviews Clinical oncology* 20 (5), 338–349. doi:10.1038/s41571-023-00747-0

Wan, Y., Wu, K., Wang, L., Yin, K., Song, M., Giovannucci, E. L., et al. (2022). Dietary fat and fatty acids in relation to risk of colorectal cancer. *European journal of nutrition* 61 (4), 1863–1873. doi:10.1007/s00394-021-02777-9

Wu, B., Yang, D., Yang, S., and Zhang, G. (2021). Dietary salt intake and gastric cancer risk: a systematic review and meta-analysis. *Frontiers in nutrition* 8, 801228. doi:10.3389/fnut.2021.801228

frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 05 August 2024
DOI 10.3389/fimmu.2024.1278327

OPEN ACCESS

EDITED BY
Xuanye Cao,
University of Texas MD Anderson Cancer Center, United States

REVIEWED BY
Debora Decote-Ricardo,
Federal Rural University of Rio de Janeiro, Brazil
Bingrui Li,
University of Texas MD Anderson Cancer Center, United States

*CORRESPONDENCE
Hong-chao Feng
hcfeng@gzu.edu.cn
Ling-ling Fu
fu.lingling@hotmail.com

RECEIVED 16 August 2023
ACCEPTED 29 April 2024
PUBLISHED 05 August 2024

CITATION
Yan M, Xiao L-y, Gosau M, Smeets R, Feng H-c, Burg S, Fu L-l and Friedrich RE (2024) The role of herpes simplex virus infection in the etiology of head and neck cancer–a Mendelian randomization study. *Front. Immunol.* 15:1278327. doi: 10.3389/fimmu.2024.1278327

The role of herpes simplex virus infection in the etiology of head and neck cancer–a Mendelian randomization study

Ming Yan[1,2], Li-yuan Xiao[1], Martin Gosau[2], Ralf Smeets[2,3], Hong-chao Feng[1*], Simon Burg[2], Ling-ling Fu[1,2*] and Reinhard E. Friedrich[2]

[1]Department of Oral and Maxillofacial Surgery, Guiyang Hospital of Stomatology, Guiyang, China, [2]Department of Oral and Maxillofacial Surgery, University Medical Center Hamburg-Eppendorf, Hamburg, Germany, [3]Department of Oral and Maxillofacial Surgery, Division of Regenerative Orofacial Medicine, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

Introduction: Head and neck cancer (HNC) is a complex disease, and multiple risk factors can lead to its progression. Observational studies indicated that herpes simplex virus (HSV) may be correlated with the risk of HNC. However, the causal effects and direction between them were still unclear.

Methods: This study utilized a Mendelian randomization (MR) approach for causality assessment between HSV infection and Head and neck cancer based on the latest public health data and Genome-Wide Association Study (GWAS) data. The causal effects were estimated using IVW, weighted median, and MR-Egger. A reverse MR analysis was subsequently performed. Cochrans *Q* test, MR-Egger intercept test, leave one out analysis, and the funnel plot were all used in sensitivity analyses.

Results: Genetically predicted higher level of HSV-1 IgG was causally related to HNC (OR=1.0019, 95%CI=1.0003–1.0036, p=0.0186, IVW) and oral and oropharyngeal cancer (OR=1.0018, 95%CI=1.0004–1.0033, p=0.0105, IVW). The reverse MR analysis did not demonstrate a reverse causal relationship between HSV and HNC. However, HSV-2 infection was not causally related to HNC data and oropharyngeal cancer data. Sensitivity analysis was performed and revealed no heterogeneity and horizontal pleiotropy.

Conclusion: Collectively, a significant association was noted between HSV infection and increased risk of HNC, providing valuable insights into the etiology of this malignancy. Further in-depth study is needed to validate these findings and elucidate the underpinning mechanisms.

KEYWORDS

head and neck cancer, herpes simplex virus, Mendelian randomization, causal effect, hsv, oral and oropharyngeal cancer

1 Introduction

Head and neck cancer (HNC) is a complicated and multifactorial disease that consists of a heterogeneous group of malignant tumors in the upper respiratory tract, covering the oral cavity, pharynx, throat, and nasal cavity (1). It is an important global health burden and is responsible for a considerable proportion of morbidity and mortality relevant to cancers on global scale. Despite advances in treatment modalities, the prognosis of this malignancy is still poor, which emphasizes the demand for a deeper understanding of its etiology and identification of new risk factors (2).

Herpes simplex virus (HSV) infection is triggered by two distinct serotypes, HSV-1 and HSV-2, showing a high prevalence in the general population. HSV-1 mainly infects the lip and mouth areas, resulting in recurrent oral lesions, commonly referred to as cold sores, while HSV-2 primarily causes genital herpes (3). In addition to the well-known manifestations, HSV infection is also linked to multiple diseases, including cancer. Several studies have discussed the potential link between HSV infection and the progression of HNC and have proposed direct and indirect mechanisms (4).

Previous epidemiological investigations have reported the relationships between HSV infection and HNC, especially oropharyngeal cancer. However, the nature of these relationships and potential causal associations remain undefined (5). Observational studies have inherent limitations, such as confounding factors and reverse causality, which hinders their ability to definitively establish causality. Rigorous and innovative research designs are required to overcome these challenges and clarify the causal role of HSV infection in HNC (6).

Mendelian randomization (MR) analysis, an instrumental variable approach with genetic variants serving as instrumental variables (IVs), represents a powerful tool for assessing causality in epidemiological studies (7). Though random assignment of genetic variants during the gamete formation process and their correlations with relevant exposures, the MR analysis can provide strong evidence for causality (8). In terms of HSV infection and HNC, the MR analysis offers a unique opportunity to overcome the limitations of observational studies and clarify the potential causal mechanism of their associations (9).

Therefore, in this study, a comprehensive MR analysis was carried out to explore the causality between HSV infection and the development of HNC, particularly oropharyngeal cancer. By utilizing large-scale genomic data and HSV infection-related genetic tools, we probed into whether HSV-1 and HSV-2 infections were causally relevant to the risk of HNC (10).

The results of this study were of great significance for understanding the etiology of HNC and may pave the way for targeted interventions to attenuate the burden of this malignancy. By elucidating the causal implication of HSV infection in HNC, we could identify prevention strategies and treatments specifically targeting HSV-related pathways (11). Ultimately, these insights may contribute to the improvement of patient prognosis, early detection, and personalized management of HNC (12).

2 Materials and methods

To study the causal relation between HSV and HNC, we conducted a bidirectional two-sample Mendelian randomization (TSMR) study in accordance with the latest STROBE-MR (Strengthening the Reporting of Observational Studies in Epidemiology Using Mendelian Randomization) guidelines (13). MR is a powerful analytical method that assesses the causality in observational studies using genetic variants as IVs.

The TSMR analysis consisted of two major procedures: estimating the genetic association with exposure (HSV infection) and estimating the genetic association with outcome (HNC). These estimated values were then combined for assessing the causal impact of the exposure on the outcome (14).

Three key assumptions must be met to ensure the validity of MR analysis:

Strong IV association: the selected IVs should be closely linked to the exposure variable (HSV infection). We identified genetic variants that had previously been validated and demonstrated to be strongly associated with HSV infection on the ground of large-scale genome-wide association studies (GWAS) or other credible sources (15).

Independence of IVs: the IVs adopted in the analysis should be independent of any confounding factors that might affect the outcome (HNC). We carefully selected IVs that had been proven to be independent of known confounding factors through extensive literature review and consultation with experts in the field (16).

Exclusion restriction assumption: IVs should affect the outcomes only *via* the association with the exposure variable (HSV infection). This assumption guaranteed that IVs would not have a direct impact on outcomes independent of their impact on HSV infection (17).

To evaluate the strength of IVs and prevent the impact of weak instruments on causality, we calculated statistical values using the formula: $F=\beta^2\_exposure/SE^2\_exposure$. Weak IVs were defined as $F<10$, indicating limited statistical power to reliably estimate the causal effect (18).

The data adopted in this article were publicly available to researchers worldwide, so no additional ethical approval and informed consent were required. We gained the summary statistics of necessary genetic associations between HSV infection and HNC from publicly available GWAS datasets and consortia (19).

According to the latest STROBE-MR guidelines, this paper conducted a bidirectional TSMR study to observe the causal relation between HSV and HNC. MR study must meet three principal assumptions: IVs should be strongly linked to exposure; (2) IVs should be independent of any possible confounders; (3) IVs affected the outcomes only *via* the exposure (Figure 1A). To avoid the impact of weak IVs on causality, the statistical values of IVs were calculated based on the formula $F=\beta^2 exposure/SE^2 exposure$. A weak IV was defined if $F<10$. The data utilized in the present study were publicly available to global researchers (20). Hence, no

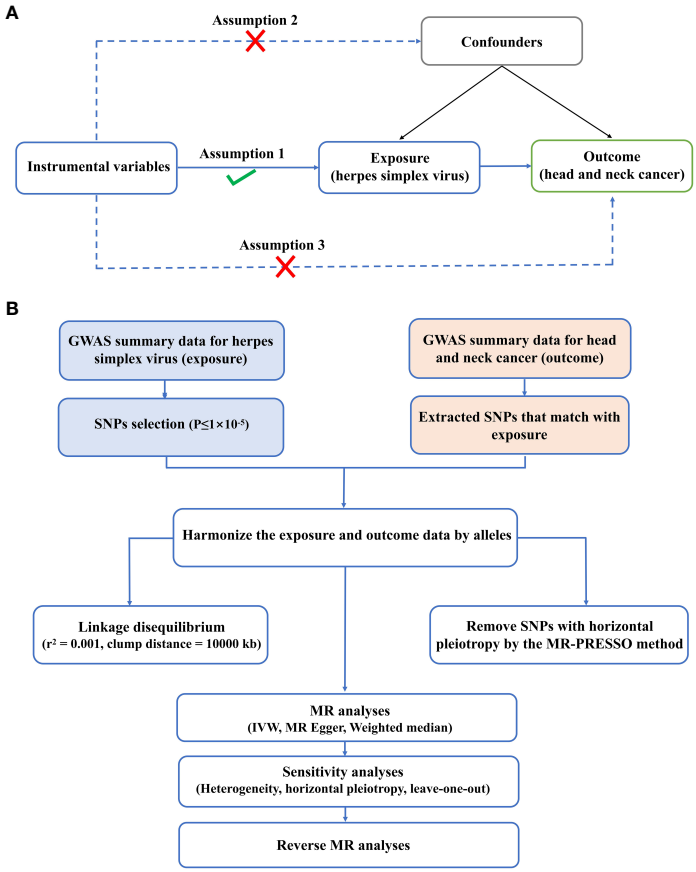


FIGURE 1
Experimental design and assumptions. **(A)** Three key assumptions of MR analysis. **(B)** Flow chart of experimental design.

additional ethical approval and informed consent were required. The procedures of the experimental design are presented in Figure 1B.

2.1 HSV infection data

Butler et al. conducted a GWAS analysis on infectious pathogens in 2020, which enrolled 8735 individuals, and serum samples were provided for the detection of antibody levels against a variety of antigens, including HSV IgG1 antibody and IgG2 antibody. Antibody detection was done using a Luminex 100 platform (Luminex Corporation, Austin, TX, USA) at a dilution of 1:1000 using a fluorescent bead-based multiplex serology technique. This approach provided the median fluorescence intensity (MFI), which allowed standardized quantification of antibodies in the samples obtained by detecting the fluorescence signal that was emitted by the analyte-trap complex. This approach and the selection of seropositive threshold had been validated for multiple infectious pathogens. The MFI seropositive threshold of the HSV IgG1 antibody and IgG2 antibody was 150. There were 6199 cases diagnosed as HSV-1 positive and 1382 cases diagnosed as HSV2 positive. The study of Butler et al. was currently the largest GWAS study on HSV serological tests (21).

The database used to obtain HSV information was from the study conducted by Butler et al. in 2020. In their study, Butler et al. carried out a GWAS analysis of infectious pathogens, including HSV, using a dataset consisting of 8735 individuals who offered serum samples for antibody detection against various antigens (22).

The fluorescent bead-based multiplex serology technology was adopted for antibody detection on the Luminex 100 platform manufactured by Luminex company (Austin, TX, USA). Serum samples were diluted at a ratio of 1:1000, and antibody levels were measured using MFI. MFI provided standardized quantification of the antibody concentration in the samples *via* the detection of the fluorescence emitted by the analyte-trap complex (23).

To determine the seropositivity of HSV, a specific threshold was established according to the effective criteria of multiple infectious diseases. In the present study, the seropositive threshold of the HSV IgG1 antibody and IgG2 antibody was set at 150 MFI. Therefore, 6199 individuals were diagnosed as HSV-1 positive, and 1382 were diagnosed as HSV-2 positive. It was worth noting that this GWAS analysis on HSV serological detection conducted by Butler et al. represented the largest such study to date.

Using the comprehensive dataset provided by Butler et al., this study adopted the information on HSV serology positivity to explore the causal relation between HSV infection and the

development of HNC. The large sample size and validated serological detection methods used in the study of Butler et al. contributed to the robustness and reliability of our analysis and enhanced the validity of the research results presented in this paper.

2.2 HNC data

UK Biobank is currently the largest GWAS database in the world, and the research population involves volunteers across the UK. The datasets, including HNC, Laryngeal cancer, Oral and oropharyngeal cancer, Oral cavity cancer, and Oropharyngeal cancer, were downloaded from UKB at https://biobank.ndph.ox.ac.uk/ukb/search.cgi (24), Details and data sources are given in Table 1.

2.3 Selection of IVs

Linkage disequilibrium (LD) thresholds were adopted for the extracted SNP ($r^2 < 0.001$, 10000 kb) to avoid the effect of LD so as to ensure independence between IVs at each exposure. Palindromic alleles were eliminated. Additionally, the F statistic was utilized to assess the strength of the IV-exposure correlation. A value of F statistic > 10 was deemed to be strong enough to avoid weak IV-induced bias. In order to satisfy the second assumption of MR, we further searched these SNPs in the PhenoScanner database (http://www.phenoscanner.medschl.cam.ac.uk/) and excluded SNPs related to other putative confounding factors (smoking, drinking frequency, etc.).

2.4 TSMR analysis

TSMR analysis was employed to analyze the causality between HSV infection on head and neck squamous carcinoma. MR methods included inverse variance weighted (IVW), MR-Egger, and weighted median (WM). As the most common MR method that could estimate the causal effect by integrating the ratio estimate of each SNP, the IVW method was the major analysis method used in this study. The MR-Egger intercept test could evaluate the horizontal pleiotropy in the MR analysis through the intercept of MR-Egger regression (horizontal pleiotropy was defined as $p<0.05$). After sequentially eliminating the SNP locus, the leave-one-out analysis used the remaining SNP loci for MR analysis to test whether there was bias caused by a specific SNP locus, and it adopted the IVW method for calculation. In MR analysis, the symmetry of the funnel plot was able to evaluate the reliability of associations. We applied the IVW method to assess the influence of all genetic variables on the outcomes. The Cochran Q test of IVW was employed to evaluate the heterogeneity between SNPs, and $p>0.1$ suggested no heterogeneity among genetic tools. Mendelian randomization-pleiotropy residual sum and outlier (MR-PRESSO) consisted of three parts: i) detection of horizontal pleiotropy; ii) correction of pleiotropy by eliminating detection outliers (genetic variants with horizontal pleiotropy); iii) comparison of the differences in causal relations before and after correction. Eventually, reverse TSMR analysis was performed with HNC as exposure and HSV infection as an outcome. MR analyses were accomplished by the "TwoSampleMR" and "MR-PRESSO" packages (R version 4.1.2). Power analysis was performed using mRnd (https://shiny.cnsgenomics.com/mRnd/). All analyses were based upon public data with no need for additional ethical approval and informed consent of participants since these had been obtained at the initial release (25).

3 Results

3.1 HSV-1 infections might related to HNC

Through the aforementioned screening conditions, 44 SNPs were found to be significantly associated with HSV infection, including 22 in HSV-1 and 22 in HSV-2, with F statistic values >10. No confounding factors of HNC were found after searching at Phenoscannerv2. The details for SNPs are described in Supplementary Table 1. As revealed by the positive MR analysis, HSV-1 IgG was causally related to HNC (OR=1.0019, 95% CI=1.0003–1.0036, p=0.0186, IVW), and oral and oropharyngeal cancer (OR=1.0018, 95%CI=1.0004–1.0033, p=0.0105, IVW) (Table 2). All causal effects of HSV on HNC assessed by the three MR methods were visualized in the scatter plot, wherein a slope greater than zero indicated a positive correlation (Figure 2). However, HSV-2 infection was not causally related to HNC data and oropharyngeal cancer data. The results of *post-hoc* power calculations were shown in Supplementary Table 2.

3.2 Sensitivity analysis revealed no heterogeneity and horizontal pleiotropy

The robustness of the aforementioned causal associations was validated based on the data from the sensitivity analysis. The

TABLE 1 General description of data sources involved in the MR analysis.

Year	Trait	Consortium	Sample size	Case	Number of SNPs	Population
2021	Head and neck cancer	UK Biobank	373122	1,106	9655080	European
2021	Oral and oropharyngeal cancer	UK Biobank	372855	839	9185233	European
2020	HSV-1 IgG	Butler-Laporte G	9724	8735	9170312	European
2020	HSV-2 IgG	Butler-Laporte G	9724	8535	9170312	European

MR, Mendelian randomization.

TABLE 2 TSMR analysis of the causal relation between HSV and head and neck cancer.

Exposures	Outcomes	SNPs	Methods	OR	95% CI	*p*
HSV-1 (IgG)	Head and neck cancer	21	MR-Egger	1.0034	0.9996–1.0073	0.0909
			Weighted median	1.0022	0.9998–1.0045	0.0651
			IVW	1.0019	1.0003–1.0036	0.0186
	Oral and oropharyngeal cancer	21	MR-Egger	1.0033	1.0008–1.0066	0.0584
			Weighted median	1.0016	0.9995–1.0037	0.1248
			IVW	1.0018	1.0004–1.0033	0.0105
HSV-2 (IgG)	Head and neck cancer	22	MR-Egger	0.9996	0.9949–1.0042	0.8724
			Weighted median	1.0007	0.9977–1.0037	0.6417
			IVW	1.0006	0.9985–1.0027	0.5621
	Oral and oropharyngeal cancer	22	MR-Egger	0.9997	0.9956–1.0037	0.8951
			Weighted median	1.0008	0.9984–1.0031	0.4997
			IVW	1.0006	0.9988–1.0024	0.5025

HSV, herpes simplex virus; IVW, inverse-variance weighted; OR, Odds ratio; CI, confidence interval.

heterogeneity test revealed no heterogeneity in the MR analysis (Cochran's Q statistic, p>0.05). The MR-Egger regression analysis failed to provide evidence for horizontal pleiotropy (MR-Egger intercept<0.01, p>0.05). The MR-PRESSO global test suggested that no noticeable outliers were able to drive the causal effect (p>0.05) (Table 3). The leave-one-out analysis further displayed no single SNP driving the causal effect (Figure 3), and the symmetry data of the funnel plot exhibited no significant heterogeneity (Figure 4).

In light of the results of the reverse MR analysis, no significant causal relations were noted between HNC and oropharyngeal cancer and HSV infection.

4 Discussion

According to the IVW genetically predicted HSV-1 was found to be positively associated with HNC risk, especially oral and oropharyngeal cancer. The methods of Inverse Variance Weighting (IVW) are deemed dependable in instances where Mendelian randomization analyses are unaffected by pleiotropy and heterogeneity. Complementarily, the Weighted Median (WM) approach is frequently employed alongside IVW. This WM technique prioritizes the estimation of causal effects by weighing and ranking the effect estimates from all instrumental variables,

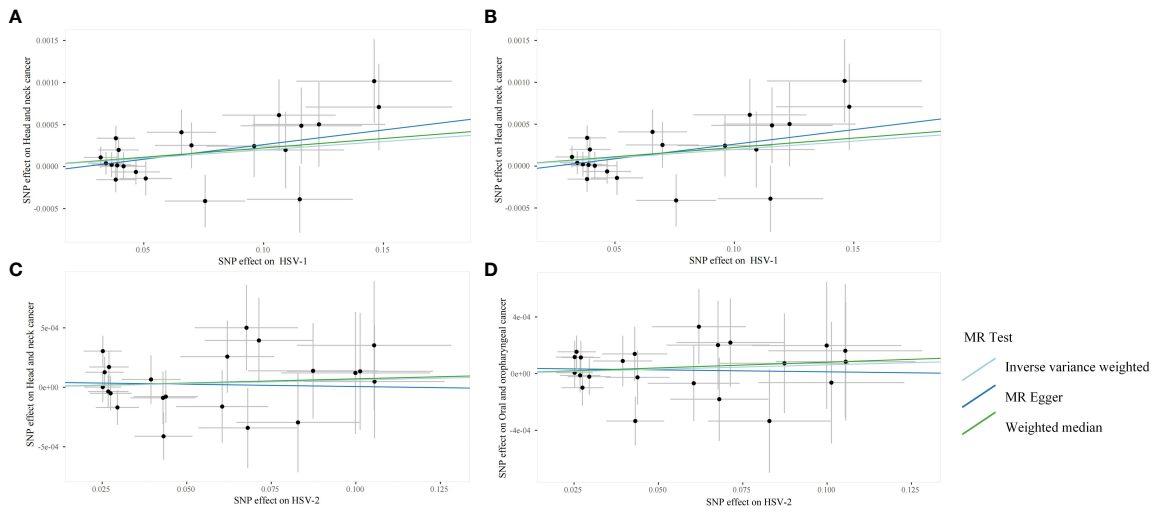


FIGURE 2
Scanner plots for the two-sample Mendelian randomization analyses. **(A)** HSV-1 and head and neck cancer; **(B)** HSV-1 and Oral and oropharyngeal cancer; **(C)** HSV-2 and head and neck cancer; **(D)** HSV-2 and Oral and oropharyngeal cancer.

TABLE 3 Results of the sensitivity analysis.

Exposure	Outcome	Heterogeneit			MR-Egger regression		MR-PRESSO
		Method	Q	Q-Pvalue	Intercept	p_intercept	Global test P
HSV-1	Head and neck cancer (UKB)	MR-Egger	19.69031	0.413425	8.67618E-05	0.407185	0.95
		IVW	20.43493	0.431035			
	Oral and oropharyngeal cancer (UKB)	MR-Egger	16.94983	0.593267	8.76034E-05	0.330346	0.99
		IVW	17.94784	0.590844			
HSV-2	Head and neck cancer (UKB)	MR-Egger	20.25135	0.442309	4.50223E-05	0.642206	0.09
		IVW	20.4767	0.491272			
	Oral and oropharyngeal cancer (UKB)	MR-Egger	12.59259	0.894172	4.02205E-05	0.632071	0.46
		IVW	12.82903	0.914494			

ultimately determining the causal effect based on the median value. In large sample sizes, the stability of each instrumental variable's estimate enhances the reliability of the median estimate. Conversely, in smaller samples, the median may exhibit greater variability due to the more volatile nature of the estimates. This paper, which investigates HSV-1 and HSV-2, operates with smaller sample sizes, potentially leading to more fluctuating results. Consequently, the primary focus of this study is on the outcomes derived from the IVW method.

Our MR analysis offered convincing evidence supporting the role of HSV-1 as a hazardous factor for HNC, especially oral and oropharyngeal cancer. However, no causal association was observed between HSV-2 infection and HNC, including oral and oropharyngeal cancer. These findings provided valuable insights into the etiology of these malignancies and were of great significance for clinical practice and future research (26).

The association between HSV-1 and HNC was consistent with previous epidemiological studies, which reported a higher rate of HSV-1 infection in patients with oral and oropharyngeal cancer compared with controls (27). HSV-1 is a common virus that mainly infects oral and oropharyngeal mucosa, resulting in recurrent oral ulcers or cold sores. The virus establishes latency in the trigeminal ganglion and can reactivate periodically, leading to virus shedding and potential transmission to others (28).

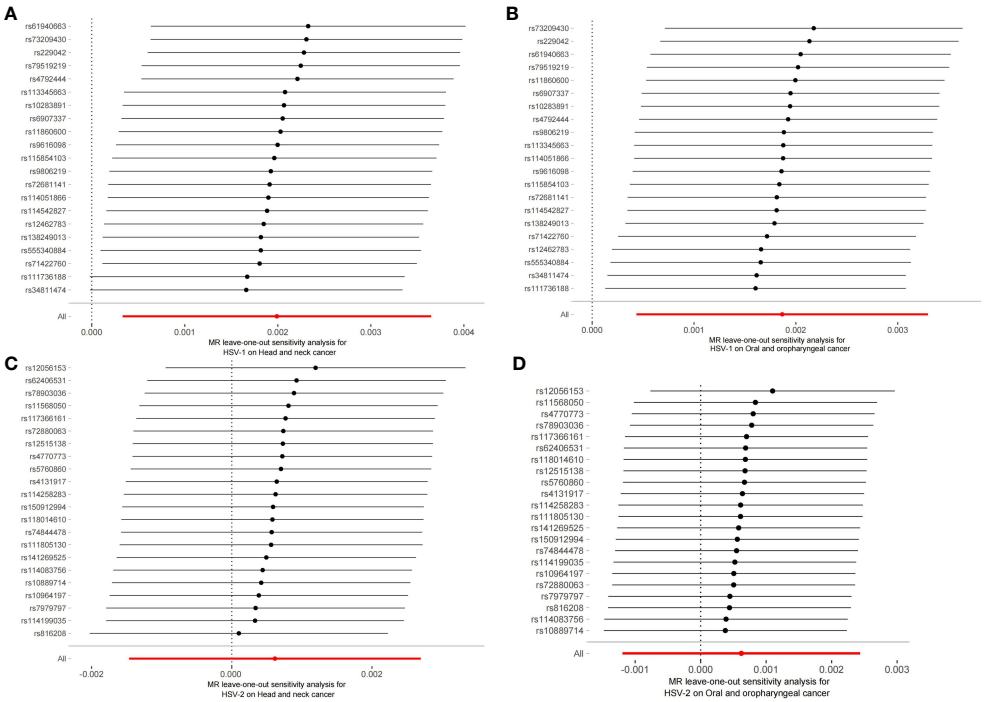


FIGURE 3
Plots of leave-one-out analyses for the two-sample Mendelian randomization analyses. **(A)** HSV-1 and head and neck cancer; **(B)** HSV-1 and Oral and oropharyngeal cancer; **(C)** HSV-2 and head and neck cancer; **(D)** HSV-2 and Oral and oropharyngeal cancer.

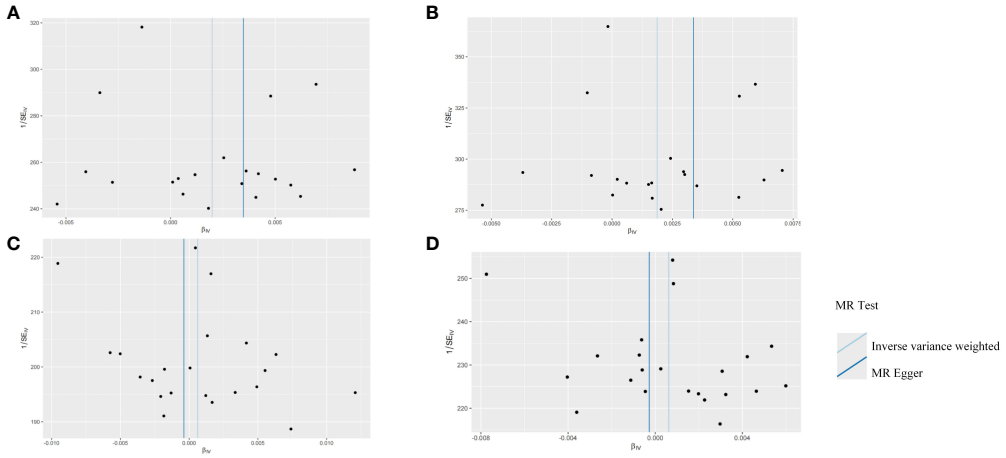


FIGURE 4
Funnel plots for the two-sample Mendelian randomization analyses. **(A)** HSV-1 and head and neck cancer; **(B)** HSV-1 and Oral and oropharyngeal cancer; **(C)** HSV-2 and head and neck cancer; **(D)** HSV-2 and Oral and oropharyngeal cancer.

The mechanism by which HSV-1 leads to the progression of HNC is multifactorial and complex. HSV-1 infection will trigger a series of immune responses, resulting in the activation of various inflammatory mediators. Sustained or repeated viral replication and shedding can lead to chronic inflammation in the oral and oropharyngeal mucosa, which in turn will promote tissue damage and genetic changes, key events in the initiation and development of cancers. Inflammatory mediators, encompassing cytokines, chemokines, and growth factors, are released in the immune response to HSV-1 infection, creating an environment conducive to cell transformation. These molecules can induce DNA damage, disrupt cell signaling pathways, and accelerate abnormal cell proliferation and survival (29). Additionally, chronic inflammation induces the release of reactive oxygen species (ROS) and reactive nitrogen species (RNs), possibly resulting in DNA damage and genomic instability, which is a hallmark of cancer development. HSV-1 has evolved several strategies to evade and modulate the host immune response, which may have a significant impact on cancer development (30). Through multiple immune evasion mechanisms, such as interfering with antigen expression, this virus can down-regulate the major histocompatibility complex (MHC) molecules and inhibit the activation and function of immune cells, including T cells and natural killer (NK) cells. Dysregulation of immune checkpoints, such as PD-1 and CTLA-4, is another key mechanism by which HSV-1 may facilitate carcinogenesis (31). HSV-1 infection can up-regulate immune checkpoint molecules on T cells, which causes cell dysfunction and impaired anti-tumor immune response. Immune checkpoint ligands, such as PD-L1 expressed by infected or malignant cells, can interact with immune checkpoint receptors on T cells, which further inhibits the immune response and stimulates the immune evasion of virus or tumor cells (32). HSV-1 infection disrupts multiple pathways engaged in cell proliferation, apoptosis, and immune response. Multiple virus-encoded proteins can manipulate the cellular signaling network, creating a favorable environment for the replication and persistence of the virus. For example, HSV-1 proteins, such as ICP0, ICP4, and ICP27, can affect host gene expression and block cellular signaling pathways, including p53, NF-κB, and MAPK-mediated signaling pathways. These alterations in cell signals can lead to dysregulation of cell proliferation, inhibition of apoptosis, and evasion from immune surveillance, contributing to the survival and growth of viruses and potentially transformed cells (33).

On the contrary, our study found no significant causal association between HSV-2 infection and HNC, including oral and oropharyngeal cancer. This finding was in agreement with several previous investigations (34). Thompson et al. conducted a systematic review and meta-analysis of the existing literature and believed that there was insufficient evidence to support the direct link between HSV-2 infection and HNC. Furthermore, Chen et al. failed to unravel a significant association between HSV-2 seropositivity and the risk of oropharyngeal cancer in a large prospective cohort study. Overall, combined with these studies, our study demonstrated that different from HSV-1, HSV-2 infection might not be an important dangerous factor for HNC (35).

It was worth noting that HSV-1 and HSV-2 exhibited different associations with HNC, which might be attributed to their different biological characteristics and modes of transmission. HSV-1 mainly infects oral and oropharyngeal mucosa, while HSV-2 mainly affects genital and anal regions. Different anatomic regions of infection may lead to different carcinogenic potentials of these two viruses. Additionally, differences in viral gene expression, immune response, and cytotaxis may also contribute to different associations (36).

However, several studies have reported conflicting results on the relationship between HSV-1 and HNC. For instance, a case-control study by Roberts et al. reported no significant association between HSV-1 seropositivity and the risk of oropharyngeal cancer. Likewise, Brown et al. failed to observe a significant relationship between HSV-1 infection and the risk of HNC in a population-based cohort study. These contradictory results might be attributed to diverse factors, including study design, sample size, population characteristics, and different HSV-1 detection methods (37).

Taken together, our research result was basically consistent with prior studies, that was, HSV-1 infection was a risk factor for HNC, especially oral and oropharyngeal cancer. Our study did not show a causal association between HSV-2 infection and HNC, which was supported by the existing literature, highlighting the importance of distinguishing these two HSV types in evaluating their potential roles in carcinogenesis.

According to the literature review, this is the first Mendelian randomization study on HSV infection and the risk of head and neck cancer. This article analyzes the association between the two at the genetic level. However, there are still some limitations in this study. First, the GWAS data on HSV are limited, and it is impossible to use multiple data sets to verify our results. Second, the study population is all Europeans, so it is impossible to predict the relationship between HSV and head and neck cancer in other populations.

5 Conclusions

In this study, MR analysis was adopted to assess whether HSV infection was causally linked to the development of HNC. We noted a significant causal relation between HSV-1 infection and the progression of HNC, particularly oral and oropharyngeal cancer, but no such causal relation was found between HSV-2 infection and HNC.

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.

Ethics statement

The manuscript presents research on animals that do not require ethical approval for their study.

Author contributions

MY: Data curation, Methodology, Software, Visualization, Writing – original draft. LX: Methodology, Project administration, Writing – review & editing. MG: Formal analysis, Investigation, Data curation, Writing – review & editing. RF: Data curation, Formal analysis, Writing – review & editing. RS: Data curation, Formal analysis, Methodology, Writing – review & editing. H-CF: Conceptualization, Formal analysis, Methodology, Project administration, Writing – review & editing. L-LF: Formal analysis, Funding acquisition, Investigation, Validation, Writing – original draft. SB: Data curation, Formal analysis, Validation, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by a grant from the Foundation of Guiyang Science and Technology Bureau (No. [2018]1-56). MY was supported by the Merit Scholarship of Hamburg university for International Students (No.7238065). L-LF was supported by the Chinese Government Scholarship (CSC Scholarship) (No.202208520014).

Acknowledgments

Authors thank the MiBioGen and FinnGen consortium for sharing the summary-level GWAS data.

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2024.1278327/full#supplementary-material

References

1. Chow LQM. Head and neck cancer. *N Engl J Med*. (2020) 382:60–72. doi: 10.1056/NEJMra1715715

2. Grigoletto I, Cavalheri V, Gobbo LA, Pozo K, Maia Filho ER, Ribeiro DG, et al. Effects of semisupervised exercise training on health outcomes in people with lung or

head and neck cancer: protocol for a randomized controlled trial. *JMIR Res Protoc*. (2023) 12:e43547. doi: 10.2196/43547

3. Menotti L, Avitabile E. Herpes simplex virus oncolytic immunovirotherapy: the blossoming branch of multimodal therapy. *Int J Mol Sci*. (2020) 21. doi: 10.3390/ijms21218310

4. Jahan N, Ghouse SM, Martuza RL, Rabkin SD. *In situ* cancer vaccination and immunovirotherapy using oncolytic HSV. *Viruses*. (2021) 13. doi: 10.3390/v13091740

5. Hong B, Sahu U, Mullarkey MP, Kaur B. Replication and spread of oncolytic herpes simplex virus in solid tumors. *Viruses*. (2022) 14. doi: 10.3390/v14010118

6. Kalke K, Orpana J, Lasanen T, Esparta O, Lund LM, Frejborg F, et al. The *in vitro* replication, spread, and oncolytic potential of finnish circulating strains of herpes simplex virus type 1. *Viruses*. (2022) 14. doi: 10.3390/v14061290

7. Bowden J, Holmes MV. Meta-analysis and Mendelian randomization: A review. *Res Synth Methods*. (2019) 10:486–96. doi: 10.1002/jrsm.1346

8. Grover S, Del Greco MF, König IR. Evaluating the current state of Mendelian randomization studies: a protocol for a systematic review on methodological and clinical aspects using neurodegenerative disorders as outcome. *Syst Rev*. (2018) 7:145. doi: 10.1186/s13643–018-0809–3

9. Woolf B, Di Cara N, Moreno-Stokoe C, Skrivankova V, Drax K, Higgins JPT, et al. Investigating the transparency of reporting in two-sample summary data Mendelian randomization studies using the MR-Base platform. *Int J Epidemiol*. (2022) 51:1943–56. doi: 10.1093/ije/dyac074

10. Spiga F, Gibson M, Dawson S, Tilling K, Davey Smith G, Munafò MR, et al. Tools for assessing quality and risk of bias in Mendelian randomization studies: a systematic review. *Int J Epidemiol*. (2023) 52:227–49. doi: 10.1093/ije/dyac149

11. Crider K, Williams J, Qi YP, Gutman J, Yeung L, Mai C, et al. Folic acid supplementation and malaria susceptibility and severity among people taking antifolate antimalarial drugs in endemic areas. *Cochrane Database Syst Rev*. (2022) 2. doi: 10.1002/14651858.Cd014217

12. Gibson MJ, Spiga F, Campbell A, Khouja JN, Richmond RC, Munafò MR. Reporting and methodological quality of studies that use Mendelian randomization in UK Biobank: a meta-epidemiological study. *BMJ Evid Based Med*. (2023) 28:103–10. doi: 10.1136/bmjebm-2022–112006

13. Skrivankova VW, Richmond RC, Woolf BAR, Yarmolinsky J, Davies NM, Swanson SA, et al. Strengthening the reporting of observational studies in epidemiology using mendelian randomization: the STROBE-MR statement. *Jama*. (2021) 326:1614–21. doi: 10.1001/jama.2021.18236

14. Wang L, Ding C. Major depression disorder may causally associate with the increased atrial fibrillation risk: evidence from two-sample mendelian randomization analyses. *BMC Med Genomics*. (2023) 16:144. doi: 10.1186/s12920–023-01565–0

15. Kwok MK, Schooling CM. Herpes simplex virus and Alzheimer's disease: a Mendelian randomization study. *Neurobiol Aging*. (2021) 99:101.e11–101.e13. doi: 10.1016/j.neurobiolaging.2020.09.025

16. Zhu M, Ma Z, Zhang X, Hang D, Yin R, Feng J, et al. C-reactive protein and cancer risk: a pan-cancer study of prospective cohort and Mendelian randomization analysis. *BMC Med*. (2022) 20:301. doi: 10.1186/s12916-022-02506-x

17. Hyppönen E, Mulugeta A, Zhou A, Santhanakrishnan VK. A data-driven approach for studying the role of body mass in multiple diseases: a phenome-wide registry-based case-control study in the UK Biobank. *Lancet Digit Health*. (2019) 1:e116–26. doi: 10.1016/s2589–7500(19)30028–7

18. Li X, Meng X, He Y, Spiliopoulou A, Timofeeva M, Wei WQ, et al. Genetically determined serum urate levels and cardiovascular and other diseases in UK Biobank cohort: A phenome-wide mendelian randomization study. *PloS Med*. (2019) 16:e1002937. doi: 10.1371/journal.pmed.1002937

19. Li X, Meng X, Spiliopoulou A, Timofeeva M, Wei WQ, Gifford A, et al. MR-PheWAS: exploring the causal effect of SUA level on multiple disease outcomes by using genetic instruments in UK Biobank. *Ann Rheum Dis*. (2018) 77:1039–47. doi: 10.1136/annrheumdis-2017–212534

20. Lai B, Yu HP, Chang YJ, Wang LC, Chen CK, Zhang W, et al. Assessing the causal relationships between gout and hypertension: a bidirectional Mendelian randomization study with coarsened exposures. *Arthritis Res Ther*. (2022) 24:243. doi: 10.1186/s13075–022-02933–4

21. Huang SY, Yang YX, Kuo K, Li HQ, Shen XN, Chen SD, et al. Herpesvirus infections and Alzheimer's disease: a Mendelian randomization study. *Alzheimers Res Ther*. (2021) 13:158. doi: 10.1186/s13195–021-00905–5

22. Chang MJ, Liu MT, Chen MR, Li N, Zhao YH, Zhang SX, et al. Mendelian randomization analysis suggests no associations of herpes simplex virus infections with systemic lupus erythematosus. *J Med Virol*. (2023) 95:e28649. doi: 10.1002/jmv.28649

23. Zhang Y, Qu J, Luo L, Xu Z, Zou X. Multigenomics reveals the causal effect of herpes simplex virus in Alzheimer's disease: A two-sample mendelian randomization study. *Front Genet*. (2021) 12:773725. doi: 10.3389/fgene.2021.773725

24. Haycock PC, Burgess S, Nounu A, Zheng J, Okoli GN, Bowden J, et al. Association between telomere length and risk of cancer and non-neoplastic diseases: A mendelian randomization study. *JAMA Oncol*. (2017) 3:636–51. doi: 10.1001/jamaoncol.2016.5945

25. Li J, Tian A, Zhu H, Chen L, Wen J, Liu W, et al. Mendelian randomization analysis reveals no causal relationship between nonalcoholic fatty liver disease and severe COVID-19. *Clin Gastroenterol Hepatol*. (2022) 20:1553–1560.e78. doi: 10.1016/j.cgh.2022.01.045

26. Gavrielatou N, Doumas S, Economopoulou P, Foukas PG, Psyrri A. Biomarkers for immunotherapy response in head and neck cancer. *Cancer Treat Rev*. (2020) 84:101977. doi: 10.1016/j.ctrv.2020.101977

27. Cohen EEW, Bell RB, Bifulco CB, Burtness B, Gillison ML, Harrington KJ, et al. The Society for Immunotherapy of Cancer consensus statement on immunotherapy for the treatment of squamous cell carcinoma of the head and neck (HNSCC). *J Immunother Cancer*. (2019) 7:184. doi: 10.1186/s40425–019-0662–5

28. Widener RW, Whitley RJ. Herpes simplex virus. *Handb Clin Neurol*. (2014) 123:251–63. doi: 10.1016/b978–0-444–53488-0.00011–0

29. Obradovic A, Graves D, Korrer M, Wang Y, Roy S, Naveed A, et al. Immunostimulatory cancer-associated fibroblast subpopulations can predict immunotherapy response in head and neck cancer. *Clin Cancer Res*. (2022) 28:2094–109. doi: 10.1158/1078–0432.Ccr-21–3570

30. Trivedi S, Sun L, Aggarwal C. Immunotherapy for head and neck cancer. *Hematol Oncol Clin North Am*. (2021) 35:1021–37. doi: 10.1016/j.hoc.2021.05.010

31. Moskovitz J, Moy J, Ferris RL. Immunotherapy for head and neck squamous cell carcinoma. *Curr Oncol Rep*. (2018) 20:22. doi: 10.1007/s11912–018-0654–5

32. Cillo AR, Kürten CHL, Tabib T, Qi Z, Onkar S, Wang T, et al. Immune landscape of viral- and carcinogen-driven head and neck cancer. *Immunity*. (2020) 52:183–199.e9. doi: 10.1016/j.immuni.2019.11.014

33. Cramer JD, Burtness B, Ferris RL. Immunotherapy for head and neck cancer: Recent advances and future directions. *Oral Oncol*. (2019) 99:104460. doi: 10.1016/j.oraloncology.2019.104460

34. Poulose JV, Kainickal CT. Immune checkpoint inhibitors in head and neck squamous cell carcinoma: A systematic review of phase-3 clinical trials. *World J Clin Oncol*. (2022) 13:388–411. doi: 10.5306/wjco.v13.i5.388

35. Botticelli A, Cirillo A, Strigari L, Valentini F, Cerbelli B, Scagnoli S, et al. Anti-PD-1 and anti-PD-L1 in head and neck cancer: A network meta-analysis. *Front Immunol*. (2021) 12:705096. doi: 10.3389/fimmu.2021.705096

36. Xu Y, Zhu G, Maroun CA, Wu IXY, Huang D, Seiwert TY, et al. Programmed death-1/programmed death-ligand 1-axis blockade in recurrent or metastatic head and neck squamous cell carcinoma stratified by human papillomavirus status: A systematic review and meta-analysis. *Front Immunol*. (2021) 12:645170. doi: 10.3389/fimmu.2021.645170

37. Wang H, Zhao Q, Zhang Y, Zhang Q, Zheng Z, Liu S, et al. Immunotherapy advances in locally advanced and recurrent/metastatic head and neck squamous cell carcinoma and its relationship with human papillomavirus. *Front Immunol*. (2021) 12:652054. doi: 10.3389/fimmu.2021.652054

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 03 March 2025
DOI 10.3389/fgene.2025.1366743

OPEN ACCESS

EDITED BY
Simon Charles Heath,
National Center for Genomic Analysis, Center for Genomic Regulation (CRG), Spain

REVIEWED BY
Xin Hu,
Sichuan University, China
Jianqiu Kong,
Sun Yat-Sen Memorial Hospital, China

*CORRESPONDENCE
Zhen Liu,
rachelly520@163.com

RECEIVED 07 January 2024
ACCEPTED 07 February 2025
PUBLISHED 03 March 2025

CITATION
Du Y, Xiao X, Liu F, Zhu W, Mo J and Liu Z (2025) Causal effects of metabolites on malignant neoplasm of bone and articular cartilage: a mendelian randomization study.
Front. Genet. 16:1366743.
doi: 10.3389/fgene.2025.1366743

Causal effects of metabolites on malignant neoplasm of bone and articular cartilage: a mendelian randomization study

Yongwei Du[1], Xiqiu Xiao[2], Fuping Liu[3], Wenqing Zhu[1], Jianwen Mo[1] and Zhen Liu[4]*

[1]Department of Orthopedics, First Affiliated Hospital of Gannan Medical University, Ganzhou, China, [2]Department of Orthopedics, 8th People Hospital of Nankang, Ganzhou, China, [3]Department of Emergency, Shangyou Hospital of Traditional Chinese Medicine, Ganzhou, China, [4]Department of Rehabilitation, First Affiliated Hospital of Gannan Medical University, Ganzhou, China

Objective: Previous research has demonstrated that metabolites play a significant role in modulating disease phenotypes; nevertheless, the causal association between metabolites and malignant malignancies of bones and joint cartilage (MNBAC)has not been fully elucidated.

Methods: This study used two-sample Mendelian randomization (MR) to explore the causal correlation between 1,400 metabolites and MNBAC. Data from recent genome-wide association studies (GWAS) involving 8,299 individuals were summarized. The GWAS summary data for metabolites were acquired from the IEU Open GWAS database, while those for MNBAC were contributed by the Finnish Consortium. We employed eight distinct MR methodologies: simple mode, maximum likelihood estimator, MR robust adjusted profile score, MR-Egger, weighted mode, weighted median, MR-PRESSO and inverse variance weighted to scrutinize the causal association between metabolites engendered by each gene and MNBAC. Consequently, we evaluated outliers, horizontal pleiotropy, heterogeneity, the impact of single nucleotide polymorphisms (SNPs), and adherence to the normal distribution assumption in the MR analysis.

Results: Our findings suggested a plausible causative relationship between N-Formylmethionine (FMet) levels, lignoceroylcarnitine (C24) levels, and MNBAC. We observed a nearly significant causal association between FMet levels and MNBAC within the cohort of 1,400 metabolites (P = 0.024, odds ratio (OR) = 3.22; 95% CI [1.16–8.92]). Moreover, we ascertained a significant causal link between levels of C24 and MNBAC (P = 0.0009; OR = 0.420; 95%CI [0.25–0.70]). These results indicate a potential causative relationship between FMet, C24 level and MNBAC.

Conclusion: The occurrence of MNBAC may be causally related to metabolites. This might unveil new possibilities for investigating early detection and treatment of MNBAC.

KEYWORDS

mendelian randomization, causality, metabolites, neoplasm, bone, articular, cartilage

1 Introduction

Malignant malignancies of bones and joint cartilage (MNBAC) is a rare but severe type of tumor. The phrase "bone tumors" refers to all cancers, including primary, secondary, and metastatic tumors, originating from skeletal or other bone tissue components (Yang et al., 2023). Primary MNBAC include osteosarcoma, chondrosarcoma,malignant, lymphoma,osteofibrosarcoma, myeloma,Ewing'ssarcoma,and chordoma (Choi and Ro, 2021). MNBAC predominantly occurs in the mobile segments of the long bones, referred to as metaphysis, encompassing the proximal tibia, proximal humerus, and distal femur (Chou et al., 2014). The major clinical symptoms of MNBAC are pain, swelling, and functional impairment (Xia et al., 2018). Osteosarcoma is the most common primary MNBAC, accounting for approximately 1% of all malignancies in the United States (Suehara et al., 2019). Osteosarcoma frequently exhibits aggressive growth and metastasizes to adjacent tissues and other locations. Ewing sarcoma (ES), the second most frequent bone tumor in teenagers, flourishes in a mechanically active microenvironment (Marturano-Kruik et al., 2018), It typically occurs in children and adolescents and originates in the bone marrow or soft tissues. Conventional methods for treating bone tumors include surgical resection, radiotherapy, and chemotherapy (Beane et al., 2017). Reconstruction of the affected area post-resection is a crucial phase that significantly impacts the overall outcome and patient wellbeing (Hu et al., 2023a; Hu et al., 2023b; Hu et al., 2022). Radiotherapy may be used to reduce tumor size preoperatively, prevent recurrence after surgery, and control metastases (Jones et al., 2018). Chemotherapy is often combined with surgery and radiation therapy to eliminate potential micrometastatic lesions (Wang et al., 2019). Although malignant bone tumors are relatively rare, they pose a significant threat to the patient's life and physical function. Therefore, exploring new targets for screening, prevention, and treatment of MNBAC is essential.

Metabolites are tiny compounds that act as intermediates and products of metabolic reactions. Multiple factors affect the levels of these metabolites, including genetics, dietary patterns, lifestyle choices, gut microbiota composition and pathological conditions (Noronha et al., 2019). Metabolites could influence the risk of maladies and be the focus of therapeutic intervention (Noronha et al., 2019). A better understanding of the causative function of metabolites in disease etiology can lead to more controllable therapeutic targets. Common genetic metabolites serve as discriminating agents in the pathogenesis of various complicated illnesses. These metabolites interact with environmental variables such as lifestyle choices, potentially influencing an individual's susceptibility to specific disease phenotypes (Li et al., 2019). To date, GWAS has identified several metabolite-related loci in human urine and blood specimens (Cai et al., 2021). Moreover, these loci correlated with the progression and prognosis of respiratory disorders (Chang et al., 2023), gastrointestinal maladies (Kim et al., 2019), cardiovascular conditions (Mihuta et al., 2023), endocrine dysregulation (Tan et al., 2023), as well as tumor diseases (Gubser and Kallies, 2020). However, limited studies have investigated the association between 1,400 metabolites and MNBAC.

Using GWAS, we can scrutinize genetic variations within extensive populations and juxtapose them with diverse metabolite concentrations, disease ramifications, and other pertinent attributes to elucidate the involvement of metabolites in disease consequences (Tang et al., 2019). Numerous metabolite levels have shown high heritability, providing the opportunity to perform Mendelian randomization (MR) (Civelek and Lusis, 2014). MR is an instrumental variable analysis approach utilizing genetic variations as tools to evaluate causal connections between potentially modifiable exposures, such as single nucleotide polymorphisms (SNPs), and clinically significant outcomes; it has been extensively employed to investigate causal inference in epidemiological studies (Liu et al., 2023a; Liu et al., 2023b; Xiang et al., 2021).

This study explored the causal relationship between 1,400 metabolites and MNBAC employing MR analysis coupled with metabolomics using GWAS data of MNBAC as the outcome file and GWAS data of 1,400 metabolites as the exposure file. Furthermore, this study identified relevant metabolites, providing novel insights into early detection and therapeutic strategies for MNBAC.

2 Methods

2.1 The flowchart and assumption of MR

The causal links between 1,400 metabolites and MNBAC were examined using a two-sample MR analysis. Summary-level GWASs data were used for the metabolites and MNBAC. The flowchart of this study is displayed in Figure 1. Furthermore, to ensure the accuracy of the findings, the MR analysis must adhere to three fundamental hypotheses: (1) The instrumental variables (IVs) employed exhibited a robust association with metabolites. (2) The selected IVs and confounding factors that influenced both the metabolites and MNBAC were mutually independent. (3) The absence of horizontal pleiotropy was ensured: IVs solely influenced MNBAC through metabolites (Davey Smith and Hemani, 2014) (Figure 1). Moreover, the results obtained were reported following the MR-STROBE protocol (Choi et al., 2022).

2.2 Exposure sources of 1,400 metabolites

Metabolic data were derived from the extensive GWAS analysis conducted by Chen et al. in the esteemed journal "Nature Genetics" (Chen et al., 2023). This investigation amalgamated 309 metabolite ratios and 1,091 individual metabolites from a cohort of 8,299 participants within the esteemed Canadian Longitudinal Study of Aging (CLSA). The CLSA cohort comprised nearly 2.1 million SNPs and 452 blood metabolites. Comprehensive GWAS summary statistics are accessible for direct retrieval from the European GWAS (GWAS ID: met-a) under the accession number GCST90199621-902010209, encompassing data for 1,400 metabolites.

2.3 Outcome sources of MNBAC

The GWAS summary data for MNBAC were obtained from the FinnGen studies, which are available through their website (https://

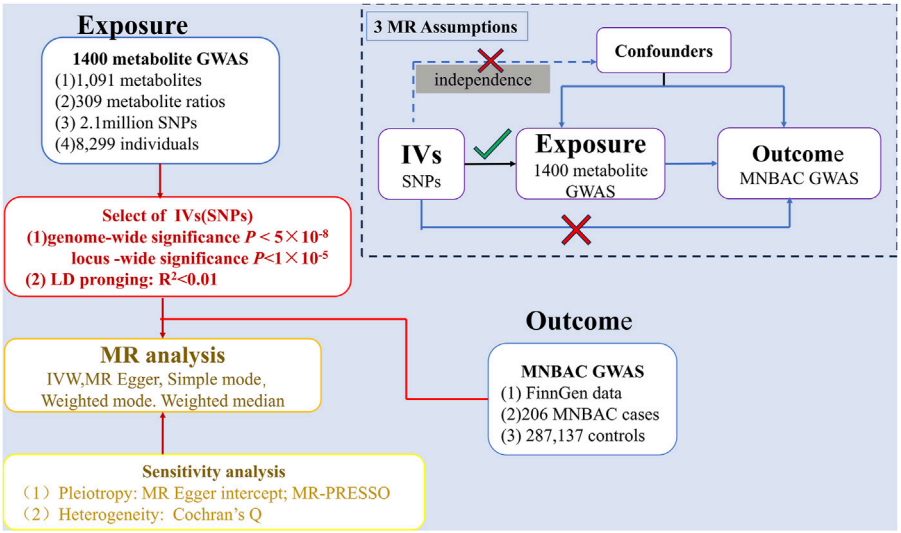


FIGURE 1
MR analyses process and major assumptions.

r9.finngen.fi/) and included individuals of European ancestry, both men and women. SAIGE (https://github.com/weizhouUMICH/SAIGE) was utilized to conduct the GWAS analysis, incorporating 20175454 variable SNPs across a cohort of 377,277 participants. After adjustments for factors such as age, gender, high genotypic individual deletions (>5%), excessive heterozygosity (4SD), and non-Finnish lineage, a subset of 206 MNBAC cases and 287,137 controls were selected for scrutiny. MNBAC was defined using the ICD-10 code M13. Further information on the data can be found on the FinnGen website.

2.4 Statistical analysis

Statistical analysis was executed utilizing R software (version 4.3.1). The "TwoSampleMR" software was employed to perform MR analysis of the causal relationship between metabolites and MNBAC. $P < 0.05$ ordinarily signifies the statistically significance of the findings, thus indicating that such a correlation may be regarded as evidence of causality (Xiang et al., 2021).

2.4.1 IVs selection

Meticulous selection of the approved IVs was imperative for enhancing the robustness of MR analysis. Initially, we pursued stringent criteria characterized by formidable values of 1×10^{-5} and 5×10^{-8}. The SNPs used in the MR test adhered to the principles of Mendelian inheritance: parental alleles were randomly allocated to offspring, impervious to acquired traits. Therefore, these alleles exhibited a high degree of independence and were potentially unrelated to confounding factors. The universal standards for SNP screening encompassed two thresholds: $P < 1 \times 10^{-5}$ and $P < 5 \times 10^{-8}$, signifying their statistically significant inclusion in the research. Lastly, we used Steiger filtration to eliminate any IVs that may lead to causal inversion.

2.4.2 Statistical analyses for MR

We examined the two cohorts using 1,400 metabolites as the exposure and MNBAC as the outcome in this study. MR analyses were executed using the "Two Sample MR" software package, with IVW analysis employed to synthesize the effects of multiple loci and evaluate numerous SNPs (Yan et al., 2023). Without horizontal pleiotropy, the IVW test was used as the principal method for assessing causal effects to obtain unbiased estimates (Du et al., 2023). The presence or absence of heterogeneity determined the existence of fixed or random effects. The effect estimates were presented as odds ratios (ORs) and 95% confidence intervals (CI).

In addition to MR analysis, the maximum likelihood estimator (MLE), MR robust adjusted profile score (MR-RAPS), MR-Egger test (Bowden et al., 2015) and the weighted median (WM) approach (Bowden et al., 2016) were employed. WM data were utilized to determine substantial causation. The absence of horizontal pleiotropy was established if $P > 0.05$. The basic model and MR-PRESSO analyses were used as part of the sensitivity analyses (Liu K. et al., 2022). The F statistic was calculated using aggregated data levels to ascertain IV exposure correlations. If $F > 10$, the correlations were considered sufficiently robust to mitigate the weak IVs bias. Within the IVW framework, the Cochrane Q statistic was utilized to evaluate heterogeneity among SNP estimates. Additionally,we validated the robustness of the data using the simple mode and the leave-one-out method (Liu K. et al., 2022).

3 Result

3.1 The study design of MR

The causal links between MNBAC and 1,400 metabolites were unveiled through a two-sample MR analysis. The categorization of metabolites and MNBAC conformed to the aggregated data

10.3389/fgene.2025.1366743

TABLE 1 Causal results of MR analysis between metabolites and MNBAC with threshold of $P < 1 \times 10^{-5}$.

Exposure	method	nsnp	pval	Or (95%CI)
N-formylmethionine levels	MR Egger	20	13.46×10^{-2}	3.19 (0.74, 13.60)
N-formylmethionine levels	Weighted median	20	0.14×10^{-2}	3.79 (1.67, 8.59)
N-formylmethionine levels	Inverse variance weighted	20	2.36×10^{-2}	2.04 (1.10, 3.78)
N-formylmethionine levels	Simple mode	20	63.71×10^{-2}	0.66 (0.11, 3.64)
N-formylmethionine levels	Weighted mode	20	1.60×10^{-2}	3.65 (1.39, 9.53)
Isoursodeoxycholate levels	MR Egger	17	1.09×10^{-2}	0.18 (0.05, 0.57)
Isoursodeoxycholate levels	Weighted median	17	6.90×10^{-2}	0.31 (0.13, 0.72)
Isoursodeoxycholate levels	Inverse variance weighted	17	4.43×10^{-2}	0.49 (0.24, 0.98)
Isoursodeoxycholate levels	Simple mode	17	15.02×10^{-2}	0.27 (0.04, 1.47)
Isoursodeoxycholate levels	Weighted mode	17	3.67×10^{-2}	0.22 (0.06, 0.81)
Methionine sulfone levels	MR Egger	29	4.69×10^{-2}	1.97 (1.04, 3.73)
Methionine sulfone levels	Weighted median	29	3.24×10^{-2}	1.74 (1.04, 2.92)
Methionine sulfone levels	Inverse variance weighted	29	4.80×10^{-2}	1.42 (1.00, 2.03)
Methionine sulfone levels	Simple mode	29	30.17×10^{-2}	1.75 (0.62, 5.01)
Methionine sulfone levels	Weighted mode	29	3.92×10^{-2}	1.78 (1.05, 3.02)
Methyl glucopyranoside (alpha + beta) levels	MR Egger	22	5.72×10^{-2}	1.41 (1.01, 1.96)
Methyl glucopyranoside (alpha + beta) levels	Weighted median	22	4.41×10^{-2}	1.45 (1.01, 2.09)
Methyl glucopyranoside (alpha + beta) levels	Inverse variance weighted	22	0.30×10^{-2}	1.49 (1.14, 1.93)
Methyl glucopyranoside (alpha + beta) levels	Simple mode	22	8.12×10^{-2}	2.15 (0.94, 4.89)
Methyl glucopyranoside (alpha + beta) levels	Weighted mode	22	2.71×10^{-2}	1.40 (1.06, 1.86)
Lignoceroylcarnitine (C24) levels	MR Egger	21	0.66×10^{-2}	0.27 (0.12, 0.63)
Lignoceroylcarnitine (C24) levels	Weighted median	21	0.77×10^{-2}	0.48 (0.28, 0.82)
Lignoceroylcarnitine (C24) levels	Inverse variance weighted	21	0.06×10^{-2}	0.50 (0.34, 0.74)
Lignoceroylcarnitine (C24) levels	Simple mode	21	93.49×10^{-2}	0.96 (0.37, 2.47)
Lignoceroylcarnitine (C24) levels	Weighted mode	21	3.26×10^{-2}	0.52 (0.29, 0.90)

acquired from the GWASs. Figure 1 depicts the flowchart outlining the MR investigation involving the metabolites and MNBAC.

3.2 Selection of IVs related to MNBAC

We meticulously selected IVs linked to MNBAC from a pool of 2.1 million SNPs associated with 1,400 metabolites. Subsequent to a quality control procedure integrating the Linkage Disequilibrium (LD) effect and retrogression method, we utilized a $P < 1 \times 10^{-5}$ for the calculations, resulting in the identification of 30,276 SNPs, 2295 SNP-metabolites, and IVs for MNBAC (threshold 1×10^{-5}). Each SNP demonstrated adequate validity (F-values ranging from 19.51 to 2,298.39, all F > 10) (Table 1). The most significant information of the IVs is presented in Supplementary Table S1 ($P < 1 \times 10^{-5}$). Additionally, to establish the robustness of the results, we adopted a more stringent threshold of 5×10^{-8} for the analysis, which identified 2295 SNP metabolites and IVs for MNBAC (F-values ranging between 29.71 and 2,298.39, all F > 10). Significant data regarding IVs were provided within the specifics. Supplementary Table S2 outlines the primary information of the IVs ($P < 5 \times 10^{-8}$).

3.3 MR analyses results ($P < 1 \times 10^{-5}$)

We evaluated the influence of 1,400 metabolites on bone tumor risk at a threshold of 1×10^{-5}, and found suggestive evidence of causality ($P < 0.05$) for five metabolites. These included N-formylmethionine (FMet) levels ($P = 0.001$; OR = 3.789; 95% CI [1.670–8.593]),isoursodeoxycholate levels ($P = 0.010$; OR = 0.183; 95% CI [0.058–0.576]), methionine sulfone levels ($P = 0.032$; OR = 1.749; 95%CI [1.048–2.921]), methyl glucopyranoside (alpha + beta) levels ($P = 0.044$; OR = 1.451; 95% CI [1.009–2.085]), and lignoceroylcarnitine (C24) levels ($P = 0.006$, OR = 0.268; 95% CI [0.115–0.625]). Table 1 and Figure 2 presented these findings. Notably, three of these metabolites, namely, methyl glucopyranoside (alpha + beta) levels, fMet levels, and

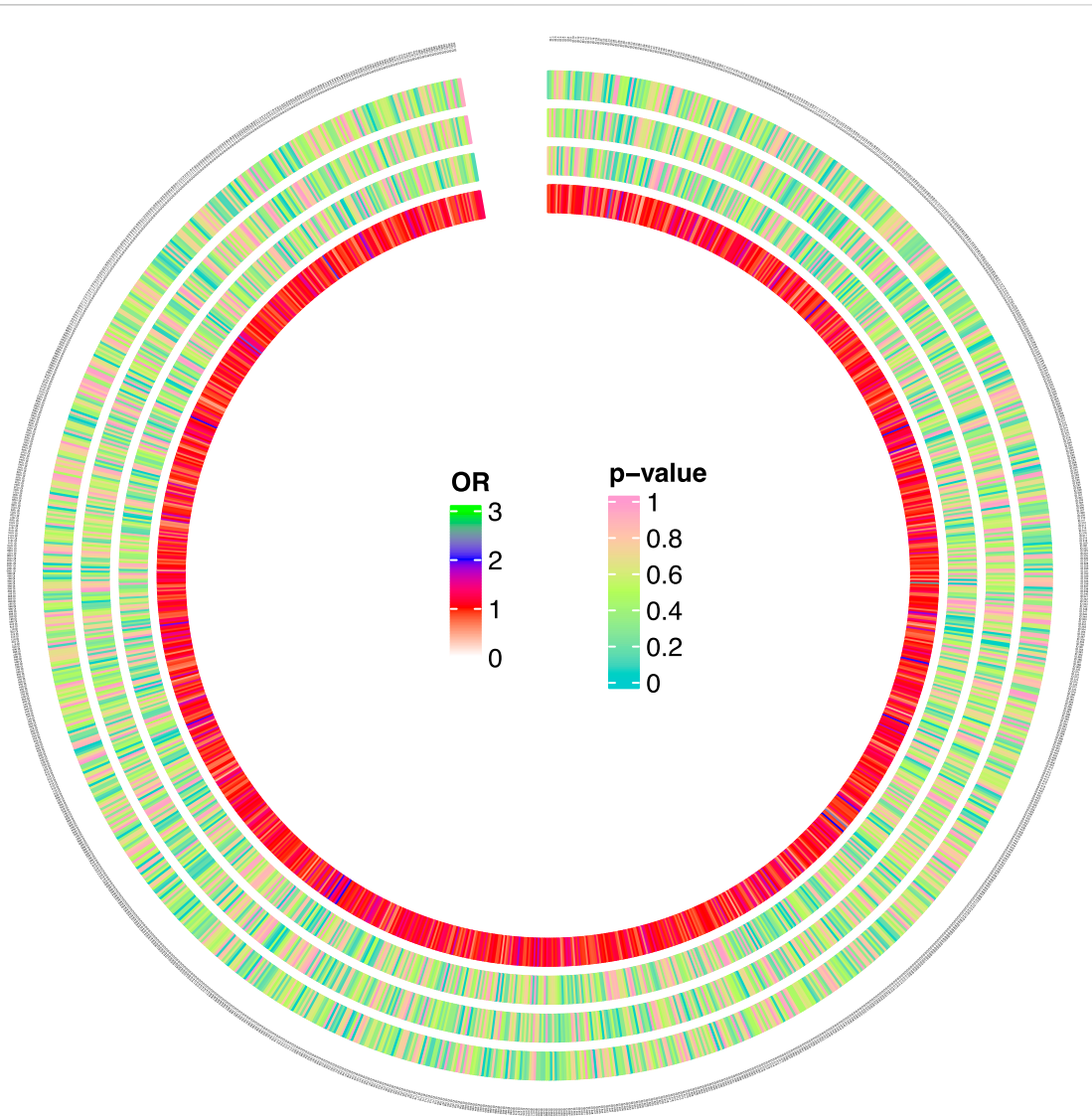


FIGURE 2
Causal analysis 562 results of 1,400 metabolites and MNBAC (locus-wide significance, $P < 1 \times 10^{-5}$). The color corresponding to the *P* value is based on the RGB color (*P* = 0, #66CCCC; *P* = 0.5, #CCFF66; *P* = 1, #FF99CC). The color corresponding to the OR value is based on the RGB color (OR = 0, white; OR = 1, red; OR = 2, blue; OR = 3, green).

methionine sulfone levels, were particularly associated with high-risk factors for bone and joint cancer. Additionally, isoursodeoxycholate concentrations and C24 levels, might be linked to low-risk bone tumor. These findings were validated using five different methods (Supplementary Table S3).

3.4 Heterogeneity analysis ($P < 1 \times 10^{-5}$)

Supplementary Table S5 lists the results of multiplicity and heterogeneity assessments for all metabolites. Through sensitivity analyses,we verified the effect of accurate MR results on metabolites of MNBAC. Notably, FMet levels ($P = 0.51$), Isoursodeoxycholate levels ($P = 0.06$), Methionine sulfone levels ($P = 0.24$), Methyl glucopyranoside (alpha + beta) levels ($P = 0.58$), and C24 levels showed no evidence of horizontal pleiotropy in relation to bone tumors (P = 0.11) (Table 2). Meanwhile, no heterogeneity was observed in FMet levels (MR-Egger: $P = 0.26$; IVW: P = 0.29), Isoursodeoxycholate levels (MR-Egger: $P = 0.15$; IVW: $P = 0.05$), Methionine sulfone levels (MR-Egger: $P = 0.55$; IVW: $P = 0.53$), and Methyl glucopyranoside (alpha + beta) levels (MR-Egger: $P = 0.12$; IVW: $P = 0.13$) (Table 2). Furthermore, the leave-one-out analysis showed no meaningful difference. in casual estimation of FMet levels. Methionine sulfone levels Isoursodeoxycholate levels. Methyl glucopyranoside (alpha + beta) levels and C24 levels on MNBAC (Figure 3).

To validate the accuracy of MR Egger regression, we further validated the significant MR results using MLE, MR-PRESSO, MR-RAPS. We found no evidence of heterogeneity in FMet levels ($P = 0.295$), isoursodeoxycholate levels ($P = 0.074$), methionine sulfone levels ($P = 0.622$), methyl glucopyranoside (alpha + beta) levels ($P = 0.238$), and C24 levels ($P = 0.519$), indicating the lack of horizontal

TABLE 2 MR results of sensitivity analysis with threshold of $P < 1 \times 10^{-5}$.

Exposure	Method	Q	Q-pval	Method	Q	Q_pval	egger_intercept	pval	MR-PRESSO
N-formylmethionine levels	IVW	21.85	0.29	MR Egger	21.32	0.26	−0.06	0.51	0.29
Isoursodeoxycholate levels	IVW	25.92	0.05	MR Egger	20.34	0.16	0.143	0.06	0.07
Methionine sulfone levels	IVW	26.70	0.53	MR Egger	25.29	0.55	−0.05	0.24	0.62
Methyl glucopyranoside (alpha + beta) levels	IVW	28.05	0.13	MR Egger	27.63	0.11	0.02	0.58	0.23
LC (C24) levels	IVW	18.32	0.56	MR Egger	15.61	0.68	0.11	0.11	0.51

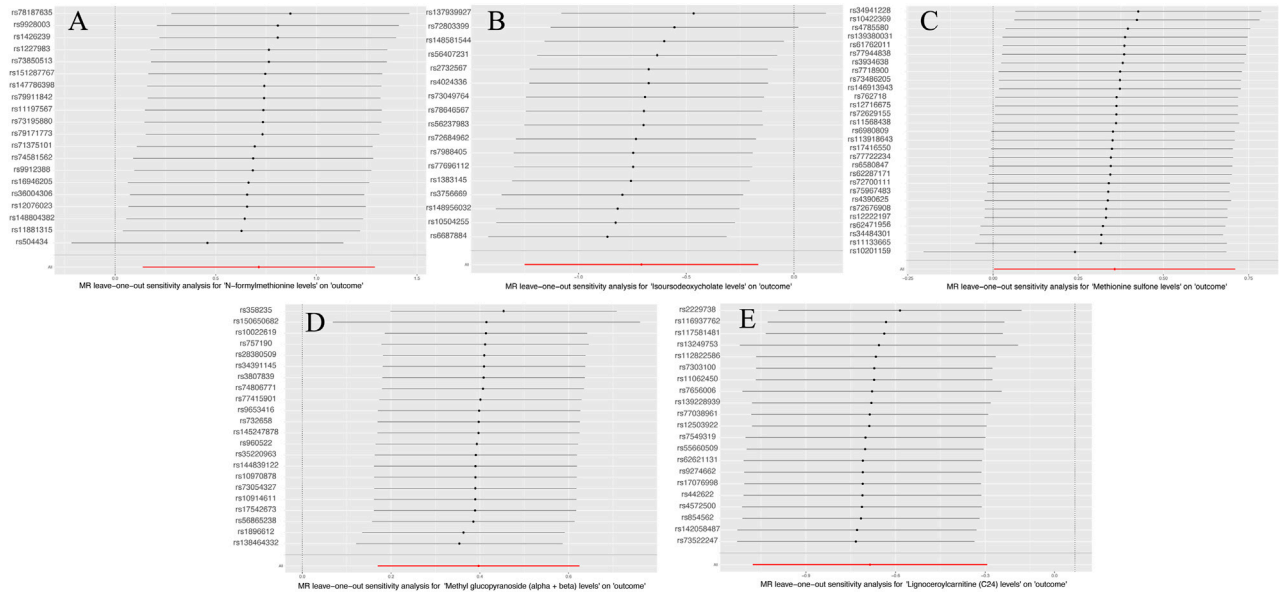


FIGURE 3
The leave-one-out results of 1,400 metabolites and MNBAC (locus-wide significance, $P < 1 \times 10^{-5}$). **(A)** N-formylmethionine levels **(B)**. Isoursodeoxycholate levels **(C)**. Methionine sulfone levels **(D)**. Methyl glucopyranoside (alpha + beta) levels **(E)**. Lignoceroylcarnitine (C24) levels.

pleiotropy (Table 2). Moreover, data robustness was reinforced through sample-by-sample exclusion analysis, which demonstrated consistent IVW results for lack of heterogeneity and pleiotropy. Based on these findings, there appeared to be a suggestive causal correlation between FMet levels, isoursodeoxycholate levels, methionine sulfone levels, methyl glucopyranoside (alpha + beta) levels, and C24 levels with MNBAC.

3.5 Results of MR analysis ($P < 5 \times 10^{-8}$)

Supplementary Table S4 presents the results pertaining FMet levels and MNBAC, illustrating a notable causal significance for FMet levels in MR analyses (IVW: OR = 3.12, 95%CI [1.16–8.92], $P = 0.05$; WM: OR = 3.22, 95%CI [1.16–8.92], $P = 0.02$; MR Egger: OR = 55.32, 95%CI [1.74–1750.53], $P = 0.26$). Furthermore, there was a significant causal relationship between the metabolite levels of C24 and MNBAC (IVW: OR = 0.42; 95%CI [0.25–0.70]; $P = 0.0009$; WM: OR = 0.47; 95% CI [0.26–0.85]; $P = 0.01$; MR Egger: OR = 0.42; 95%CI [0.26–0.85]; $P = 0.01$) (Table 3; Figure 4).

3.6 Heterogeneity analysis ($P < 5 \times 10^{-8}$)

Supplementary Table S6 displays the pleiotropy and heterogeneity test results for metabolisms. Heterogeneity analysis results of C24 levels (MR Egger: $P = 0.24$; IVW: $P = 0.41$) and multiplicity analysis (MR Egger: $P = 0.34$; MR-PRESSO: $P = 0.31$) verified reliability of the results. Likewise, the scrutiny of heterogeneity in FMet levels (MR Egger: $P = 0.62$; IVW: $P = 0.21$) and the multiplicity analysis (MR-PRESSO: NA; MR Egger: $P = 0.34$) verified the accuracy of the data (Table 4). Concurrently, the findings of sample-by-sample exclusion further validated the robustness of the data (Figure 4). Unfortunately, due to the overly stringent 5×10^{-8} threshold, only FMet levels and C24 levels were obtained with fewer instrumental variables. Notably, C24 levels

TABLE 3 Causal Results of MR analysis between metabolites and MNBAC with threshold of $P < 5 \times 10^{-8}$.

Exposure	Method	nsnp	pval	Or (95%CI)
N-formylmethionine levels levels	MR Egger	3	26.34×10^{-2}	55.32 (1.74, 1750.53)
N-formylmethionine levels levels	Weighted median	3	2.40×10^{-2}	3.22 (1.16, 8.92)
N-formylmethionine levels levels	IVW	3	5.18×10^{-2}	3.12 (0.99, 9.87)
N-formylmethionine levels levels	Simple mode	3	17.29×10^{-2}	5.55 (1.10, 27.95)
N-formylmethionine levels levels	Weighted mode	3	16.06×10^{-2}	3.90 (1.14, 13.27)
Lignoceroylcarnitine (C24) levels	MR Egger	4	42.81×10^{-2}	0.42 (0.07, 2.31)
Lignoceroylcarnitine (C24) levels	Weighted median	4	13.57×10^{-2}	0.47 (0.26, 0.85)
Lignoceroylcarnitine (C24) levels	IVW	4	0.09×10^{-2}	0.42 (0.25, 0.70)
Lignoceroylcarnitine (C24) levels	Simple mode	4	19.35×10^{-2}	0.49 (0.22, 1.12)
Lignoceroylcarnitine (C24) levels	Weighted mode	4	10.33×10^{-2}	0.47 (0.25, 0.89)

TABLE 4 MR results of sensitivity analysis, with threshold of $P < 5 \times 10^{-8}$.

Exposure	Method	Q	Q_pval	Method	Q	Q_pval	Egger intercept	pval	MR-presso
N-fet levels	IVW	3.09	0.21	MR Egger	0.23	0.62	−0.46	0.33	NA
LC (C24) levels	IVW	2.82	0.41	MR Egger	2.82	0.24	−0.004	0.98	0.30

exhibited a significant correlation with MNBAC, while FMet levels approached significance in terms of causal inference. In addition,the leave-one-out analysis showed some difference. in casual estimation of FMet levels and C24 levels on MNBAC (Figure 5; Supplementary Figure S2). Furthermore, due to an insufficiency of IVs, the multiplicity assessment for FMet levels was unattainable via MR-PRESSO.

3.7 Further validation of the MR results

To further ascertain the causal relationship between metabolites and MNBAC, we employed additional methods to validate the results. Under the threshold of $P < 1 \times 10^{-5}$, the outcomes of MR-PRESSO, MR-RAPS, and MLE provided additional substantiation of the causal nexus between FMet levels, Methionine sulfone levels, Isoursodeoxycholate levels, C24 levels, and Methyl glucopyranoside (alpha + beta) levels with bone tumors (Table 5). The results from MR-RAPS and MLE confirmed the causal relationship between FMet levels and MNBAC, albeit not verified by MR-PRESSO. Meanwhile, C24 levels were further validated by MR-PRESSO, MR-RAPS, and MLE at a threshold of $P < 5 \times 10^{-8}$ (Table 6).

3.8 Ethics statement

This summary-level data utilized in this study are de-identified public data and are accessible to download. Each GWAS in this study received ethical approval from their respective universities.

4 Discussion

This research conducted an MR analysis to investigate the potential causal relationship between 1,400 metabolites and MNBAC. By investigating the association from a host genetic perspective, we aimed to validate the role of these metabolites in altering susceptibility to MNBAC. Five MR methods were employed for the analysis. Although some of the results from various analytical approaches were inconsistent, these differences did not significantly influence our findings. The random effects IVW technique exhibited superior statistical power compared to the other approaches, hence it was selected as the major analytical approach in this work. While there was a potential causal relationship, and multiple corrections are too strict, they were also close to being corrected. The results of this study suggested that the two metabolites may be linked to a lower risk of MNBAC, while the three metabolites are related to a higher risk of MNBAC. Our findings open up possibilities for identifying novel biomarkers that can be utilized in future MNBAC studies. Moreover, our results indicated potential avenues for MNBAC prevention and treatment, including the targeted manipulation of specific metabolite levels. Notably,the cross-sectional aspect of this study made it difficult to find a definitive connection between metabolites and MNBAC. However, using MR analysis, we provided valuable insights into the potential causative association and highlight the significance of these metabolites in influencing susceptibility to MNBAC.

As the most prevalent type of MNBAC, osteosarcoma generates severe symptoms and poses a threat to individuals of all ages due to malignant neoplasia (Quintero Escobar et al., 2020). Several studies on osteosarcoma have examined aberrant metabolisms. The

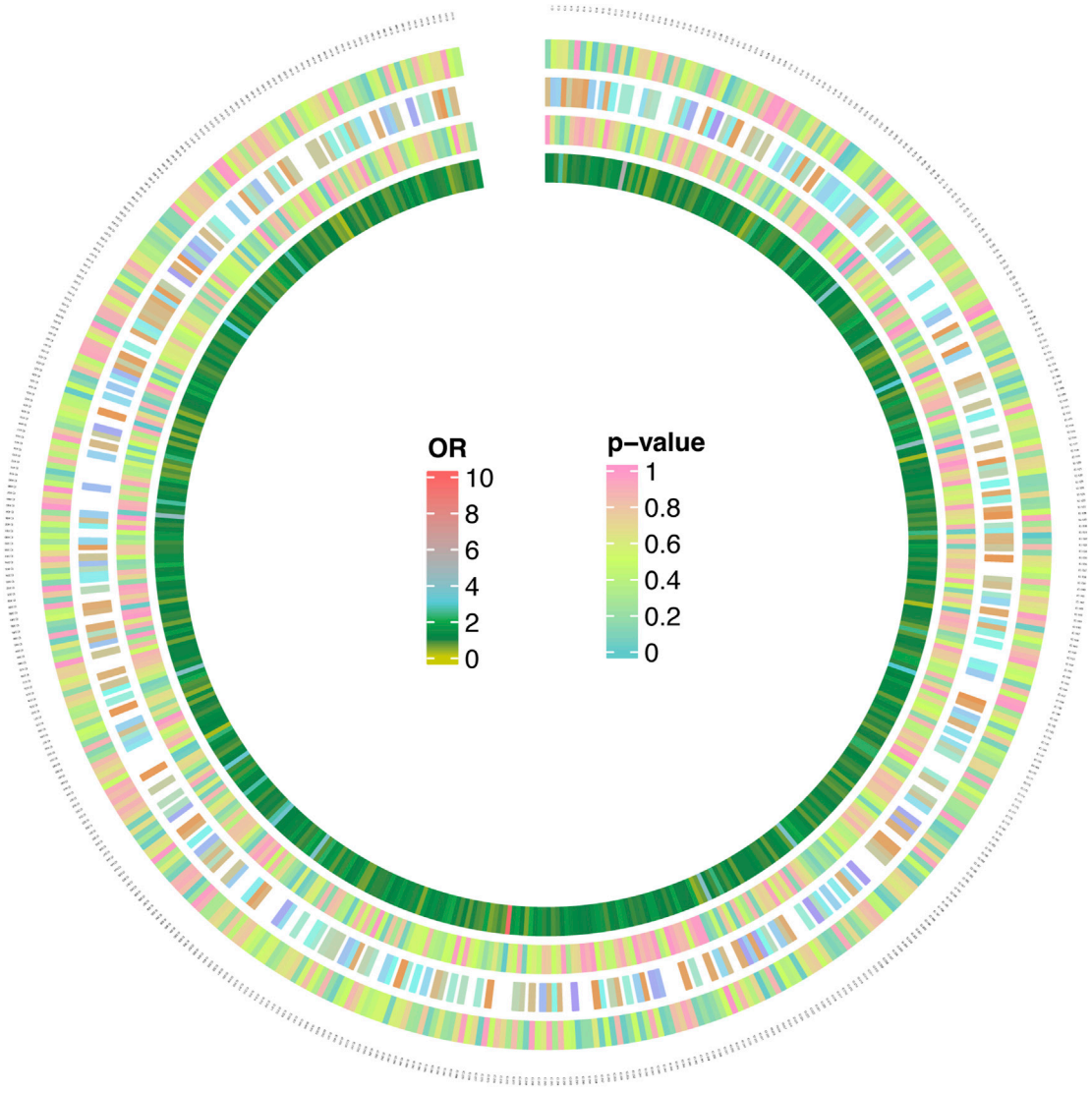


FIGURE 4
MR analysis 566 results of 1,400 metabolites and MNBAC (genome-wide statistical significance, $P < 5 \times 10^{-8}$). The color corresponding to the *P* value is based on the RGB color (*P* = 0, #66CCCC; *P* = 0.5, #CCFF66; *P* = 1, #FF99CC). The color corresponding to the OR value is based on the RGB color (OR = 0, #CCCC00; OR = 1, #088247; OR = 2, #11AA4D; OR = 3, #58CDD9; OR = 10, #FF6666).

development and progression of osteosarcoma is closely related to cellular metabolites (Velayutham et al., 2023). A promising natural metabolite, stylolite, has been discovered to activate vascular endothelial growth factor receptor 2 (VEGFR2)and trigger its downstream signaling pathways. This activation promotes endothelial cell proliferation and angiogenesis while hindering the growth and invasion of osteosarcoma cells, simultaneously enhancing the sensitivity to chemotherapy drugs. Moreover, particular metabolites such as zoledronic acid induce iron-induced death in osteosarcoma cells by reducing coenzyme Q levels and stimulating heme oxygenase 1(HMOX1)expression (Ren et al., 2022). *In vitro* experiments involving osteosarcoma stem cells have revealed comparable declines in metabolites associated with the tricarboxylic acid (TCA) cycle (Zhong et al., 2019). These reductions stem from impaired mitochondrial function and are accompanied by diminished glutamine, aspartate, and glutathione levels (Ren et al., 2020; Zhong et al., 2019). Metabolite-based biomarkers for osteosarcoma exhibit potential for diagnosis and monitoring disease progression (Fan et al., 2021). These substances have been linked to developing and regulating glucose metabolism and cellular regulatory mechanisms in osteosarcoma. In conclusion, metabolites play a vital role in MNBAC research. Here, we evaluated the effects of 1,400 metabolites on MNBAC risk and identified five metabolites that showed suggestive causal relationships with MNBAC, which were validated using more than five methods. These metabolites included the FMet levels, methionine sulfone levels, methyl glucopyranoside (alpha + beta) levels, isoursodeoxycholic acid levels, and C24 levels. Among them, our finding of significant causality between C24 levels and MNBAC is novel. The FMet levels showed near-significant results. This study is the first to discover the correlation between 1,400 metabolites and MNBAC.

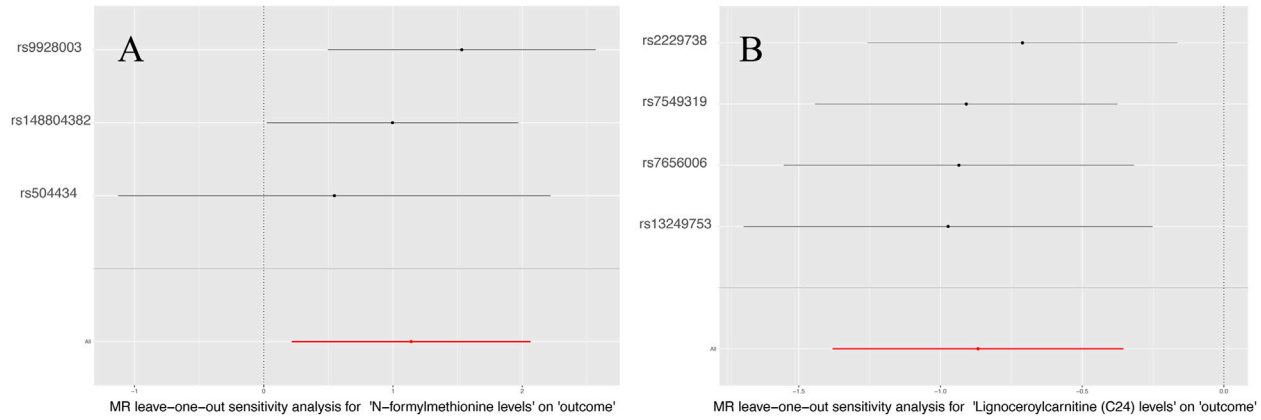


FIGURE 5
The leave-one-out results of 1,400 metabolites and MNBAC ($P < 5 \times 10^{-8}$); **(A)** N-formylmethionine levels **(B)**. Lignoceroylcarnitine (C24) levels.

TABLE 5 MR Results of sensitivity analysis with threshold of $P < 1 \times 10^{-5}$.

Exposure	Method	Or (95%CI)	p-val
N-formylmethionine levels	MR-PRESSO	1.93 (1.43, 2.63)	4.31×10^{-2}
N-formylmethionine levels	MR-RAPs	1.98 (1.47, 2.67)	2.19×10^{-2}
N-formylmethionine levels	MLE	1.97 (1.47, 2.65)	2.07×10^{-2}
Methionine sulfone levels	MR-PRESSO	1.46 (1.24, 1.72)	2.87×10^{-2}
Methionine sulfone levels	MR-RAPs	1.47 (1.23, 1.76)	3.08×10^{-2}
Methionine sulfone levels	MLE	1.47 (1.23, 1.75)	2.89×10^{-2}
Isoursodeoxycholate levels	MR-PRESSO	0.51 (0.36, 0.71)	5.93×10^{-2}
Isoursodeoxycholate levels	MR-RAPs	0.49 (0.37, 0.65)	1.08×10^{-2}
Isoursodeoxycholate levels	MLE	0.52 (0.39, 0.68)	1.84×10^{-2}
Lignoceroylcarnitine (C24) levels	MR-PRESSO	0.52 (0.43, 0.62)	1.03×10^{-3}
Lignoceroylcarnitine (C24) levels	MR-RAPs	0.51 (0.42, 0.61)	2.81×10^{-4}
Lignoceroylcarnitine (C24) levels	MLE	0.52 (0.43, 0.62)	3.16×10^{-4}
Methyl glucopyranoside (alpha + beta) levels	MR-PRESSO	1.48 (1.30, 1.69)	6.31×10^{-3}
Methyl glucopyranoside (alpha + beta) levels	MR-RAPs	1.49 (1.31, 1.70)	1.93×10^{-3}
Methyl glucopyranoside (alpha + beta) levels	MLE	1.50 (1.32, 1.70)	1.33×10^{-3}

TABLE 6 MR results of sensitivity analysis with threshold of $P < 5 \times 10^{-8}$.

Exposure	Method	Or (95%CI)	P-val
N-formylmethionine levels	MR-PRESSO	Not enough intrumental variables	NA
N-formylmethionine levels	MR-RAPs	3.19 (1.96, 5.19)	1.69×10^{-2}
N-formylmethionine levels	MLE	3.19 (1.97, 5.19)	1.66×10^{-2}
Lignoceroylcarnitine (C24) levels	MR-PRESSO	0.48 (0.35, 0.66)	8.26×10^{-2}
Lignoceroylcarnitine (C24) levels	MR-RAPs	0.48 (0.37, 0.62)	3.70×10^{-3}
Lignoceroylcarnitine (C24) levels	MLE	0.48 (0.37, 0.62)	3.70×10^{-3}

C24 is a long-chain fatty acid derivative that participates in the metabolism of fatty acids, particularly in the beta-oxidation process within the mitochondria (Abd-Allah et al., 2009). Abnormal fatty acid metabolism may be associated with tumor growth and survival, thus, changes in C24 levels may reflect alterations in tumor metabolic status in osteosarcoma cells, Researchers have found a link between the concentration of C24 in serum and the risk of MNBAC development, with higher concentrations associated with lower risks (Liu T. et al., 2022), Similarly, a study on the testing dose of C24 reported that it can reverse fatigue symptoms in MNBAC patients with C24 deficiency (Farahzadi et al., 2023). Our research has identified a noteworthy inverse relationship between C24 levels and MNBAC, indicating a possible function of C24 in inhibiting tumor growth. Assessing C24 levels could potentially assist in identifying patients at high risk of MNBAC and serve as a biomarker for tracking disease progression and evaluating treatment efficacy.

FMet, an amino acid that typically corresponds to the start codon, signifies the initiation of polypeptide chain synthesis. The role of FMet in protein degradation processes is also significant, particularly in the activity of peptidyl deformalize (PDF) (Silver, 2011). In osteosarcoma cells, PDF activity may be upregulated, thereby affecting protein stability and intracellular signal transduction (Lee et al., 2004; Pietzke et al., 2020). In our study, the significant positive correlation between the FMet levels and MNBAC indicated a potential role of FMet in promoting tumor growth. Similarly, in another 11,966 individuals, FMet levels may resulted in all-cause mortality and the risk of human cancer. including MNBAC. suggesting a substantial connection between FMet and the risk of MNBAC (Cai et al., 2021), In another cancer study, FMet was utilized as a drug precursor, converted into formic acid through the activity of PDF enzyme (Yu et al., 2015).

This study had several limitations. First, like other MR researches on metabolites, although our study satisfies the MR assumptions (IVs is closely related to the metabolite), there may be other mechanisms or factors in some cases that result in a correlation between IVs and the target variable, rather than a causal relationship. Second, our study's sample sizes was modest, which may alter the dependability of our results. Given that GWAS only included European ancestry participants, our findings may not be applicable to other racial populations. Third, the multiple statistical correction employed was overly strict and conservative, potentially overlooking the metabolites that may have a causal relationship with MNBAC. While there is potential for a causal relationship, and the multiple corrections were overly strict, they approach correction. Therefore, we considered biological plausibility and did not rely solely on the results of the multiple-hypothesis testing. Finally, due to the insufficient availability of an ample number of IVs in this study, the implementation of reverse MR analysis and multivariable Mendelian randomization analyses (MVMR) was precluded. In future research, we plan to undertake GWAS investigations specifically targeting FMet levels and C24 levels to secure a robust set of IVs. This will facilitate a more thorough validation of the causal relationship between metabolites and MNBAC through the application of reverse MR analysis and MVMR.

Concludingly, this study confirmed the causal link between metabolites and MNBAC species, including and FMet levels. These metabolites have the potential to serve as new biomarkers or treatment targets for MNBAC and novel strategies for its treatment and prevention.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

Author contributions

YD: Writing–original draft, Writing–review and editing, Conceptualization, Data curation, Formal Analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization. XX: Data curation, Formal Analysis, Conceptualization, Funding acquisition, Investigation, Software, Writing–review and editing. FL: Data curation, Funding acquisition, Methodology, Resources, Supervision, Validation, Visualization, Writing–review and editing. WZ: Data curation, Funding acquisition, Methodology, Resources, Software, Supervision, Validation, Visualization, Writing–review and editing. JM: Conceptualization, Formal Analysis, Investigation, Project administration, Software, Supervision, Validation, Visualization, Writing–review and editing. ZL: Writing–original draft, Writing–review and editing, Conceptualization, Data curation, Formal Analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This work was supported by grants from the Jiangxi Key R&D Plan Project (20203BBGL73145), the Jiangxi Provincial Department of Education Science and Technology Plan (GJJ190809), and the Ganzhou Guiding Science and Technology Plan (GZ2021ZSF001) to YD. Jiangxi Provincial Health Commission General Program (SKJP220211236) to WZ.

Acknowledgments

We want to acknowledge the participants and investigators of the FinnGen study (https://r9.finngen.fi/) (MNBAC datasource), the largest genome-wide association study (mGWAS) analysis of metabolites (metabolites datasource) published by chen et al. (Chen et al., 2023) in ⟪nature genetics⟫ for sharing the genetic data. We thank Home for Researchers editorial team (www.home-for-researchers.com) for language editing service.

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/fgene.2025.1366743/full#supplementary-material

References

Abd-Allah, A. R., Helal, G. K., Al-Yahya, A. A., Aleisa, A. M., Al-Rejaie, S. S., and Al-Bakheet, S. A. (2009). Pro-inflammatory and oxidative stress pathways which compromise sperm motility and survival may be altered by L-carnitine. *Oxidative Med. Cell. Longev.* 2, 73–81. doi:10.4161/oxim.2.2.8177

Beane, J., Campbell, J. D., Lel, J., Vick, J., and Spira, A. (2017). Genomic approaches to accelerate cancer interception. *Lancet Oncol.* 18, e494–e502. doi:10.1016/S1470-2045(17)30373-X

Bowden, J., Davey Smith, G., and Burgess, S. (2015). Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int. J. Epidemiol.* 44, 512–525. doi:10.1093/ije/dyv080

Bowden, J., Davey Smith, G., Haycock, P. C., and Burgess, S. (2016). Consistent estimation in mendelian randomization with some invalid instruments using a weighted median estimator. *Genet. Epidemiol.* 40, 304–314. doi:10.1002/gepi.21965

Cai, N., Gomez-Duran, A., Yonova-Doing, E., Kundu, K., Burgess, A. I., Golder, Z. J., et al. (2021). Mitochondrial DNA variants modulate N-formylmethionine, proteostasis and risk of late-onset human diseases. *Nat. Med.* 27, 1564–1575. doi:10.1038/s41591-021-01441-3

Chang, Q. H., Zhang, Y. C., Zhang, D. Y., Mao, T., Chang, R., Wang, N., et al. (2023). A novel methionine metabolism-related signature predicts prognosis and immunotherapy response in lung adenocarcinoma. *Aging* 15, 3498–3523. doi:10.18632/aging.204687

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

Choi, J. H., and Ro, J. Y. (2021). The 2020 WHO classification of tumors of bone: an updated review. *Adv. anatomic pathology* 28, 119–138. doi:10.1097/PAP.0000000000000293

Choi, S. S., Kim, H., Choung, Y. J., Jeong, S. J., and Lee, C. H. (2022). Risk factors for new vertebral compression fracture after kyphoplasty and efficacy of osteoporosis treatment: a STROBE-compliant retrospective study. *Medicine* 101, e32018. doi:10.1097/MD.0000000000032018

Chou, C. W., Cheng, Y. W., and Tsai, C. H. (2014). Phyllostachys edulis extract induces apoptosis signaling in osteosarcoma cells, associated with AMPK activation. *Drug Des. Dev. Ther.* 8, 1577–1584. doi:10.2147/DDDT.S69342

Civelek, M., and Lusis, A. J. (2014). Systems genetics approaches to understand complex traits. *Nat. Rev. Genet.* 15, 34–48. doi:10.1038/nrg3575

Davey Smith, G., and Hemani, G. (2014). Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum. Mol. Genet.* 23, R89–R98. doi:10.1093/hmg/ddu328

Du, Y., Xie, B., Wang, M., Zhong, Y., Lv, Z., Luo, Y., et al. (2023). Roles of sex hormones in mediating the causal effect of vitamin D on osteoporosis: a two-step Mendelian randomization study. *Front. Endocrinol.* 14, 1159241. doi:10.3389/fendo.2023.1159241

Fan, J., Jahed, V., and Klavins, K. (2021). Metabolomics in bone research. *Metabolites* 11 (7), 434. doi:10.3390/metabo11070434

Farahzadi, R., Hejazi, M. S., Molavi, O., Pishgahzadeh, E., Montazersaheb, S., and Jafari, S. (2023). Clinical significance of carnitine in the treatment of cancer: from traffic to the regulation. *Oxidative Med. Cell. Longev.* 2023, 9328344. doi:10.1155/2023/9328344

Gubser, P. M., and Kallies, A. (2020). Methio "mine"! Cancer cells steal methionine and impair CD8 T-cell function. *Immunol. cell Biol.* 98, 623–625. doi:10.1111/imcb.12385

Hu, X., Lu, M., Zhang, Y., Li, Z., Wang, J., Wang, Y., et al. (2023a). Pelvic-girdle reconstruction with three-dimensional-printed endoprostheses after limb-salvage surgery for pelvic sarcomas: current landscape. *Br. J. Surg.* 110, 1712–1722. doi:10.1093/bjs/znad310

Hu, X., Wang, Y., Li, J., Qing, P., Yang, X., Zeng, J., et al. (2023b). Three-Dimensional multimodality image reconstruction as teaching tool for case-based learning among medical postgraduates: a focus on primary pelvic bone tumour education. *BMC Med. Educ.* 23, 944. doi:10.1186/s12909-023-04916-8

Hu, X., Lu, M., He, X., Li, L., Lin, J., and Zhou, Y. (2022). Hip reconstruction using a customized intercalary prosthesis with the rhino horn-designed uncemented stem for ultrashort proximal femur segments following tumor resection: a combined biomechanical and clinical study. *BMC Musculoskelet. Disord.* 23 (1), 852.

Jones, D. T. W., Kieran, M. W., Bouffet, E., Alexandrescu, S., Bandopadhayay, P., Bornhorst, M., et al. (2018). Pediatric low-grade gliomas: next biologically driven steps. *Neuro-oncology* 20, 160–173. doi:10.1093/neuonc/nox141

Kim, Y. H., Kim, J. H., Kim, B. G., Lee, K. L., Kim, J. W., and Koh, S. J. (2019). Tauroursodeoxycholic acid attenuates colitis-associated colon cancer by inhibiting nuclear factor kappaB signaling. *J. gastroenterology hepatology* 34, 544–551. doi:10.1111/jgh.14526

Lee, M., She, Y., Soskis, M., Borella, C., Gardner, J., Hayes, P., et al. (2004). Human mitochondrial peptide deformylase, a new anticancer target of actinonin-based antibiotics. *J. Clin. Invest.* 114, 1107–1116. doi:10.1172/JCI22269

Li, J., Li, X., Zhang, S., and Snyder, M. (2019). Gene-environment interaction in the era of precision medicine. *Cell* 177, 38–44. doi:10.1016/j.cell.2019.03.004

Liu, K., Cai, Y., Song, K., Yuan, R., and Zou, J. (2023a). Clarifying the effect of gut microbiota on allergic conjunctivitis risk is instrumental for predictive, preventive, and personalized medicine: a Mendelian randomization analysis. *EPMA J.* 14, 235–248. doi:10.1007/s13167-023-00321-9

Liu, K., Zou, J., Fan, H., Hu, H., and You, Z. (2022a). Causal effects of gut microbiota on diabetic retinopathy: a Mendelian randomization study. *Front. Immunol.* 13, 930318. doi:10.3389/fimmu.2022.930318

Liu, K., Zou, J., Yuan, R., Fan, H., Hu, H., Cheng, Y., et al. (2023b). Exploring the effect of the gut microbiome on the risk of age-related macular degeneration from the perspective of causality. *Investigative Ophthalmol. and Vis. Sci.* 64, 22. doi:10.1167/iovs.64.7.22

Liu, T., Liu, C., Wang, X., Wei, Y., Li, S., Song, Y., et al. (2022b). The association of serum L-carnitine concentrations with the risk of cancer in Chinese adults with hypertension. *Nutrients* 14, 4999. doi:10.3390/nu14234999

Marturano-Kruik, A., Villasante, A., Yaeger, K., Ambati, S. R., Chramiec, A., Raimondi, M. T., et al. (2018). Biomechanical regulation of drug sensitivity in an engineered model of human tumor. *Biomaterials* 150, 150–161. doi:10.1016/j.biomaterials.2017.10.020

Mihuta, M. S., Paul, C., Borlea, A., Roi, C. M., Pescari, D., Velea-Barta, O. A., et al. (2023). Connections between serum Trimethylamine N-Oxide (TMAO), a gut-derived metabolite, and vascular biomarkers evaluating arterial stiffness and subclinical atherosclerosis in children with obesity. *Front. Endocrinol.* 14, 1253584. doi:10.3389/fendo.2023.1253584

Noronha, A., Modamio, J., Jarosz, Y., Guerard, E., Sompairac, N., Preciat, G., et al. (2019). The Virtual Metabolic Human database: integrating human and gut microbiome metabolism with nutrition and disease. *Nucleic acids Res.* 47, D614-D624–d624. doi:10.1093/nar/gky992

Pietzke, M., Meiser, J., and Vazquez, A. J. M. m. (2020). Formate metabolism in health and disease, *Mol. Metab.* 33, 23–37. doi:10.1016/j.molmet.2019.05.012

Quintero Escobar, M., Costa, T., Martins, L. G., Costa, S. S., vanHelvoort Lengert, A., Boldrini, É., et al. (2020). Insights in osteosarcoma by proton nuclear magnetic resonance serum metabonomics. *Front. Oncol.* 10, 506959. doi:10.3389/fonc.2020.506959

Ren, L., Ruiz-Rodado, V., Dowdy, T., Huang, S., Issaq, S. H., Beck, J., et al. (2020). Glutaminase-1 (GLS1) inhibition limits metastatic progression in osteosarcoma. *Cancer and metabolism* 8, 4. doi:10.1186/s40170-020-0209-8

Ren, T., Huang, J., Sun, W., Wang, G., Wu, Y., Jiang, Z., et al. (2022). Zoledronic acid induces ferroptosis by reducing ubiquinone and promoting HMOX1 expression in osteosarcoma cells. *Front. Pharmacol.* 13, 1071946. doi:10.3389/fphar.2022.1071946

Silver, L. J. C. (2011). Challenges of antibacterial discovery, *Clin. Microbiol. Rev.*, 24, 71–109. doi:10.1128/CMR.00030-10

Suehara, Y., Alex, D., Bowman, A., Middha, S., Zehir, A., Chakravarty, D., et al. (2019). Clinical genomic sequencing of pediatric and adult osteosarcoma reveals distinct molecular subsets with potentially targetable alterations. *Clin. cancer Res. official J. Am. Assoc. Cancer Res.* 25, 6346–6356. doi:10.1158/1078-0432.CCR-18-4032

Tan, H., Shi, Y., Yue, T., Zheng, D., Luo, S., Weng, J., et al. (2023). Machine learning approach reveals microbiome, metabolome, and lipidome profiles in type 1 diabetes. *J. Adv. Res.* 64, 213–221. doi:10.1016/j.jare.2023.11.025

Tang, W., Putluri, V., Ambati, C. R., Dorsey, T. H., Putluri, N., and Ambs, S. (2019). Liver- and microbiome-derived bile acids accumulate in human breast tumors and inhibit growth and improve patient survival. *Clin. cancer Res. official J. Am. Assoc. Cancer Res.* 25, 5972–5983. doi:10.1158/1078-0432.CCR-19-0094

Velayutham, N. K., Thamaraikani, T., Wahab, S., Khalid, M., Ramachawolran, G., Abullais, S. S., et al. (2023). Stylopine: a potential natural metabolite to block vascular endothelial growth factor receptor 2 (VEGFR2) in osteosarcoma therapy. *Front. Pharmacol.* 14, 1150270. doi:10.3389/fphar.2023.1150270

Wang, S., Hwang, E. E., Guha, R., O'Neill, A. F., Melong, N., Veinotte, C. J., et al. (2019). High-throughput chemical screening identifies focal adhesion kinase and aurora kinase B inhibition as a synergistic treatment combination in ewing sarcoma. *Clin. cancer Res. official J. Am. Assoc. Cancer Res.* 25, 4552–4566. doi:10.1158/1078-0432.CCR-17-0375

Xia, B., Wang, L., Feng, L., Tian, B., Tan, Y., and Du, B. (2018). Knockdown of long noncoding RNA CAT104 inhibits the proliferation, migration, and invasion of human osteosarcoma cells by regulating MicroRNA-381. *Oncol. Res.* 27, 89–98. doi:10.3727/096504018X15199511344806

Xiang, K., Wang, P., Xu, Z., Hu, Y. Q., He, Y. S., Chen, Y., et al. (2021). Causal effects of gut microbiome on systemic lupus erythematosus: a two-sample mendelian randomization study. *Front. Immunol.* 12, 667097. doi:10.3389/fimmu.2021.667097

Yan, Z., Xu, Y., Li, K., and Liu, L. (2023). Increased fruit intake is associated with reduced risk of depression: evidence from cross-sectional and Mendelian randomization analyses. *Front. public health* 11, 1276326. doi:10.3389/fpubh.2023.1276326

Yang, M., Su, Y., Xu, K., Xie, J., Wen, P., Liu, L., et al. (2023). Rheumatoid arthritis increases the risk of malignant neoplasm of bone and articular cartilage: a two-sample bidirectional mendelian randomization study. *Arthritis Res. Ther.* 25, 219. doi:10.1186/s13075-023-03205-5

Yu, Z., Han, M., and Cowan, J. J. A. C. (2015). Toward the design of a catalytic metallodrug: selective cleavage of G-quadruplex telomeric DNA by an anticancer copper-acridine-ATCUN complex. *Angew. Chem. Int. Ed. Engl.* 54, 1901–1905. doi:10.1002/anie.201410434

Zhong, Z., Mao, S., Lin, H., Li, H., Lin, J., and Lin, J. M. (2019). Alteration of intracellular metabolome in osteosarcoma stem cells revealed by liquid chromatography-tandem mass spectrometry. *Talanta* 204, 6–12. doi:10.1016/j.talanta.2019.05.088

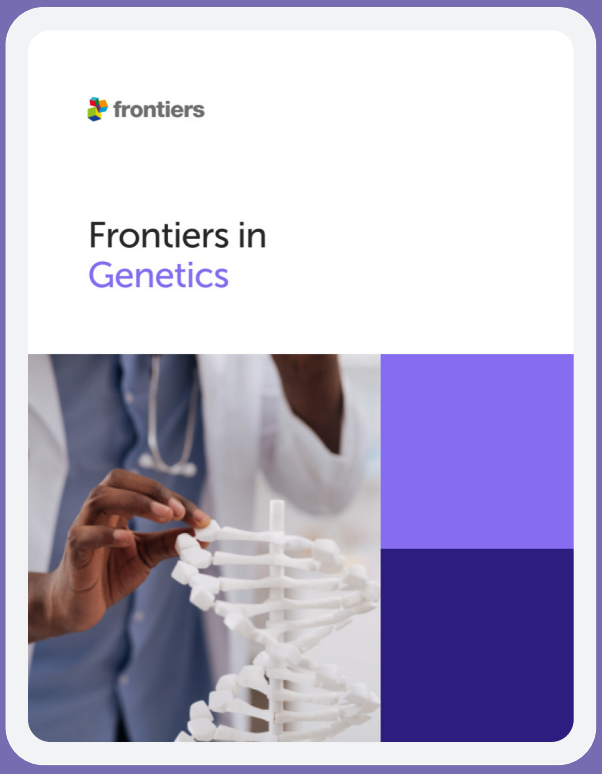
Frontiers in Genetics

Highlights genetic and genomic inquiry relating to all domains of life

The most cited genetics and heredity journal, which advances our understanding of genes from humans to plants and other model organisms. It highlights developments in the function and variability of the genome, and the use of genomic tools.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact