# The causal effects of 1400 genetically determined human blood metabolites and metabolite ratios on the risk of gastrointestinal tumors: a Mendelian randomization study

Qi Fu[1], Lele Zhang[2], Ji Di[1]*

[1]Qinghai University Affiliated Hospital (The Clinical Medical School), Qinghai University, Xining, China
[2]Gansu Corps Hospital of CAPF, Lanzhou, China

**Submitted:** 21 February 2025; **Accepted:** 4 May 2025
**Online publication:** 22 June 2025

Arch Med Sci
DOI: https://doi.org/10.5114/aoms/204639


***Corresponding author:**
Ji Di
Qinghai University Affiliated Hospital (The Clinical Medical School),Qinghai University, Xining, China
E-mail: luosangdj@126.com

## Abstract

**Introduction:** Recently, studies investigating the association between blood metabolites and gastrointestinal tumors have gained increased attention. A Mendelian randomization (MR) study is considered the second most persuasive research method to explore the causal relationship between exposure and outcome after RCT.

**Material and methods:** This analysis utilized the inverse variance weighted (IVW) method, the weighted median (WM) method, and MR-Egger regression. Initially, we analyzed GWAS data from the FinnGen database to identify various metabolites and their ratios. Subsequently, we repeatedly analyzed GWAS data from the Open GWAS database to filter out duplicate results.

**Results:** 5-methyluridine (FinnGen : odds ratio (OR) = 1.16, 95% confidence interval (CI) = 1.02–1.31, $p$ = 0.03, FDR-P = 0.04; Open GWAS: OR = 1.08, 95% CI = 1.01–1.17, $p$ = 0.03, FDR-P = 0.04) and 1-dihomo-linolenylglycerol (FinnGen: OR = 1.30, 95% CI = 1.02–1.65, $p$ = 0.03, FDR-P = 0.04; Open GWAS: OR = 1.16, 95% CI = 1.02–1.31, $p$ = 0.03, FDR-P = 0.04) were positively associated with the risk of gastric cancer (GC). Sphingomyelin (FinnGen: OR = 0.73, 95% CI = 0.54–0.98, $p$ = 0.04, FDR-P = 0.04; Open GWAS: OR = 0.81, 95% CI = 0.67–0.97, $p$ = 0.02, FDR-P = 0.04) was negatively correlated with GC risk. Carnitine to propionylcarnitine (C3) ratio (FinnGen: OR = 1.11, 95% CI = 1.01–1.22, $p$ = 0.03, FDR-P = 0.04; Open GWAS: OR = 1.07, 95% CI = 1.01–1.14, $p$ = 0.04, FDR-P = 0.04), arachidonate to linoleate ratio (FinnGen: OR = 1.10, 95% CI = 1.02–1.19, $p$ = 0.02, FDR-P = 0.04; Open GWAS: OR = 1.12, 95% CI = 1.06–1.18, $p = 4.44 \times 10^{-5}$, FDR-P = $3.55 \times 10^{-4}$), and androsterone sulfate (FinnGen: OR = 1.07, 95% CI = 1.01–1.14, $p$ = 0.03, FDR-P = 0.04; Open GWAS: OR = 1.05, 95% CI = 1.01–1.10, $p$ = 0.04, FDR-P = 0.04) were positively associated with the risk of colorectal cancer (CRC). 1-oleoyl-2-docosahexaenoyl-GPC (FinnGen: OR = 0.89, 95% CI = 0.81–0.98, $p$ = 0.02, FDR-P = 0.04; Open GWAS: OR = 0.93, 95% CI = 0.87–0.99, $p$ = 0.02, FDR-P = 0.04) was negatively correlated with CRC risk.

**Conclusions:** Three blood metabolites were found to be associated with the risk of GC; 4 blood metabolites and metabolite ratios were associated with the risk of CRC. These findings may provide valuable guidance for the early clinical diagnosis and treatment of gastrointestinal tumors.

**Key words:** human blood metabolites, metabolite ratios, gastrointestinal tumors, risk, Mendelian randomization.

## Introduction

Gastrointestinal tumors encompass gastric cancer (GC) and colorectal cancer (CRC) [1]. Globally, GC has the fifth highest incidence rate among all cancers and is the fourth leading cause of cancer-related deaths [2, 3]. CRC is the second leading cause of cancer mortality, with 1 million deaths attributed to CRC in 2020 [3]. This highlights that gastrointestinal tumors remain a significant public health issue worldwide. Previous research has identified several risk factors associated with gastrointestinal tumors, including smoking, alcohol consumption, high fat intake, low fiber intake, genetics, age, and certain infections [4–6]. Recently, studies investigating the connection between blood metabolites and various diseases have gained increased attention.

Metabolomics has emerged as a critical area in the medical field. It identifies changes in metabolites or metabolite pathways that help uncover the causes of disease development from a molecular perspective [7]. Metabolites can participate in human physiological activities as signaling molecules, endogenous toxins, immunomodulators, and environmental sensors, promoting or inhibiting the occurrence of diseases [8]. Determination of metabolites can help medical staff find new diagnostic or prognostic biomarkers for various diseases [9]. With the rapid advancement of metabolomics, researchers are increasingly focusing on the changes in blood metabolites associated with gastrointestinal tumors. Numerous studies have demonstrated that blood metabolites play a role in the development of these tumors [10, 11]. Specifically, some primary bile acids and short-chain fatty acids have been found to promote the growth of gastrointestinal tumors, while ursodeoxycholic acid and butyric acid have shown beneficial effects [12]. In addition, Dai *et al.* investigated the impact of 150 metabolites on the progression of GC [13]. Furthermore, Coker *et al.* examined the relationship between 20 metabolites and the progression of CRC [14]. However, the metabolites that have been studied represent only a small fraction of the vast number of metabolites, and the relationships between multitudinous metabolites and gastrointestinal tumors need to be explored. Traditional observational studies lack credibility due to potential biases, such as reverse causations and confounding factors. A randomized controlled trial (RCT) is considered the best way to confirm causality, but related RCTs are rare due to the large number of metabolites.

A Mendelian randomization (MR) study is considered the second most persuasive research method to explore the causal relationship between exposure and outcome after an RCT [15]. MR studies use single nucleotide polymorphisms (SNPs) as instrumental variables (IVs) to infer causal associations between exposures and outcomes, which can overcome the impact of potential biases on causal inferences [16]. According to Mendel's laws of inheritance, parents randomly assign alleles to their offspring. Therefore, MR studies are less affected by confounding factors [17]. In addition, genetic variation is determined earlier than exposure, so MR studies do not present a problem with reverse causation [18].

The relationships between blood metabolites and gastrointestinal tumors remain unclear, so further research is necessary. We used data from genome-wide association studies (GWAS) to conduct MR analyses to comprehensively screen 1400 blood metabolites and metabolite ratios for the causal relationships with gastrointestinal tumor risk. Identifying new specific biomarkers among a large number of blood metabolites can guide the clinical diagnoses and prognoses of gastrointestinal tumors.

## Material and methods

### Study design

We systematically assessed the potential causal relationships between 1400 human blood metabolites and metabolite ratios and gastrointestinal tumor risk using a two-sample randomization approach and designed the study based on the three core hypotheses of the MR analysis [19]: (1) Relevance assumption: The IVs selected must be closely related to exposures (GC, CRC). (2) Independence assumption: The selected IVs are only allowed to be related to exposures and should not be related to potential confounding factors. (3) Exclusion assumption: The selected IVs should influence outcomes entirely through exposures.

### GWAS data for 1400 human blood metabolites and metabolite ratios

We used the most up-to-date and comprehensive GWAS datasets currently available for the human metabolome [20]. Based on the Canadian Longitudinal Study on Aging (CLSA) cohort, researchers analyzed data on 1,091 blood metabolites and 309 metabolite ratios by examining 8,299 participants and approximately 15.4 million SNPs. The full GWAS summary statistics of the 1400 biomarkers were publicly available.

### GWAS data for GC and CRC

GWAS summary data for GC and CRC are from the FinnGen database (https://www.finngen.fi/fi).

Data on GC (phenocode: C3_STOMACH_EXALLC) include a total of 288,444 European participants (case: 1,307; control: 287,137). Data on CRC (phenocode: C3_COLORECTAL_EXALLC) include a total of 293,646 European participants (case: 6,509; control: 287,137).

To further verify the results of this study, we repeatedly analyzed another set of GWAS summary data of GC and CRC, all from the Open GWAS database (https://gwas.mrcieu.ac.uk/). The latest data on GC include a total of 476,116 European participants (cases: 1,029; controls: 475,087), with a total of 24,188,662 SNPs (GWAS ID: ebi-a-GCST90018849) [21]. The latest data on CRC include a total of 470,002 European participants (cases: 6,581; controls: 463,421), with a total of 24,182,361 SNPs (GWAS ID: ebi-a-GCST90018808) [21].

### Selection of IVs

IVs associated with exposure should meet the following requirements: (1) All IVs should have gene-wide significance. Considering the limited number of SNPs with significant genome-wide effects, we relaxed the threshold to $p < 5 \times 10^{-5}$ and obtained IVs from 1400 blood metabolites and metabolite ratios. This strategy is consistent with the approach of previous studies [22, 23]. (2) The linkage disequilibrium threshold applied was $R^2 < 0.001$, and the genomic region was within 10,000 kb. (3) The F value represents the intensity of MR, and $F > 10$ indicates that IVs of exposure factors have a strong ability to predict results. The formula for calculating the $F$ value is as follows: $F = R^2 (n - k - 1)/[k (1 - R^2)]$, $R^2 = 2 \times \text{EAF} \times (1 - \text{EAF}) \times \beta^2$ [24]. We further coordinated the SNPs of exposures and outcomes and removed the SNPs with palindromic effects and allele discordances. Then, the final results were subjected to MR analyses.

### MR analysis and sensitivity analysis

We mainly used the inverse variance weighted (IVW) method for MR analysis, which uses a meta-analysis method to integrate the Wald ratios of individual SNPs. This can assume that IVs can only affect the results through specific exposures. Therefore, the IVW method can achieve a robust result without polymorphism [25]. The IVW method uses the reciprocal of each IV variance as a weight to calculate weighted results. This process is carried out on the premise of ensuring that all IVs are valid to evaluate level pleiotropy [26]. However, when there are uncertainties in genetic associations and risk factors, such as weak IVs, the IVW method is biased to underestimate actual results [27]. Therefore, MR-Egger regression and weighted median (WM) methods are used as supplementary analytical methods. MR-Egger regression uses the reciprocal of the outcome variance as a weight to fit. It adds an intercept term to the regression to perform weighted linear regression when IVs are invalid to produce causal estimates [28]. The WM method is the median of the weighted empirical density function of the ratio estimate. It combines data from multiple genetic variants into a single causal estimate. When the proportion of invalid IVs is as high as 50%, and the accuracy of estimates varies widely between IVs, the WM method can still provide consistent effect estimates [29].

In order to further test the stability and reliability of the results, we conducted a sensitivity analysis, heterogeneity test, and pleiotropic test on the results. A leave-one-out analysis was used to analyze the sensitivity of the results. It removes each SNP in turn and then calculates the results of all remaining SNPs. When there is no statistically significant difference between the results of a single SNP and the total results, it means that the SNP will not have a non-specific impact on the results [30]. Cochran's Q test was used to quantify the heterogeneity of instrumental variables. $P > 0.05$ proves no heterogeneity, and the fixed-effect IVW method is used; $p < 0.05$ indicates significant heterogeneity, and the random-effect IVW method is used [30]. MR-Egger regression was used to detect horizontal pleiotropy. When its intercept term is close to 0 and $p > 0.05$, it means that there is no horizontal pleiotropy [31]. The MR-pleiotropy residual sum and outlier (MR-PRESSO) method was used to remove significant outliers and further reduce horizontal pleiotropy [32]. The funnel plot was used to detect publication bias. The roughly symmetrical plot illustrates that the results have no significant publication bias. The false discovery rate (FDR) method was employed to adjust the $p$-values of the final results.

### Replication analysis

To further verify the credibility of candidate blood metabolites, we performed the same MR analysis on another GC and CRC GWAS dataset. The overlapping blood metabolites of the two MR analyses were considered to have a significant causal relationship with GC or CRC.

### Metabolic pathway analysis

To further clarify the biological mechanism of the impact of screened blood metabolites on GC and CRC, we used MetaboAnalyst 5.0 (https://www.metaboanalyst.ca/) to conduct metabolic pathway analyses. The dataset for pathway analysis came from the Small Molecular Pathway Database (SMPDB) and the Kyoto Encyclopedia of Genes and Genomes (KEGG).

**Table I.** Sensitivity analysis results of overlapped metabolites with respect to gastric cancer

| Metabolites | MR-Egger | | WM | | IVW | | MR-Egger intercept | | MR-PRESSO global test | Cochran's Q test |
|---|---|---|---|---|---|---|---|---|---|---|
| | OR (95% CI) | *P*-value | OR (95% CI) | *P*-value | OR (95% CI) | *P*-value | FDR-*P* | intercept | *P*-value | *P*-value | *P*-value |
| 5-methyluridine | | | | | | | | | | | |
| FinnGen | 1.08 (–0.11–0.28) | 0.39 | 1.18 (0.002–0.33) | 0.05 | 1.16 (1.02–1.31) | 0.03 | 0.04 | 0.01 | 0.42 | 0.78 | 0.72 |
| Open GWAS | 1.12 (–0.01–0.24) | 0.09 | 1.06 (–0.05–0.16) | 0.20 | 1.08 (1.01–1.17) | 0.03 | 0.04 | –0.01 | 0.52 | 0.38 | 0.35 |
| 1-dihomo-linolenylglycerol | | | | | | | | | | | |
| FinnGen | 1.27 (–0.30–0.78) | 0.39 | 1.35 (–0.05–0.65) | 0.09 | 1.30 (1.02–1.65) | 0.03 | 0.04 | 0.003 | 0.94 | 0.68 | 0.57 |
| Open GWAS | 1.29 (–0.15–0.66) | 0.22 | 1.05 (–0.12–0.21) | 0.59 | 1.16 (1.02–1.31) | 0.02 | 0.04 | –0.01 | 0.58 | 0.11 | 0.08 |
| Sphingomyelin | | | | | | | | | | | |
| FinnGen | 0.78 (0.43–1.44) | 0.45 | 0.87 (0.58–1.29) | 0.49 | 0.73 (0.54–0.98) | 0.04 | 0.04 | –0.01 | 0.80 | 0.84 | 0.75 |
| Open GWAS | 0.53 (0.28–0.99) | 0.06 | 0.91 (0.71–1.16) | 0.43 | 0.81 (0.67–0.97) | 0.02 | 0.04 | 0.04 | 0.19 | 0.39 | 0.31 |

*OR – odds ratio, CI – confidence interval, GWAS – genome-wide association study, WM – weighted median, IVW – inverse variance weighted.*

## Results

### MR analysis of the causal relationship between 1400 human blood metabolites and metabolite ratios and the risk of GC

#### Preliminary analysis

After strict quality control of the IVs, we obtained SNPs for 62 metabolites and metabolite ratios in the FinnGen database. (Details of SNPs are listed in Supplementary Table SI). The screened IVs contained 15–39 SNPs. The *F* values of all the SNPs were greater than 10, indicating that the included IVs have strong predictive capabilities. (Details of IVs are shown in Supplementary Table SII). We used multiple sensitivity analysis to assess the heterogeneity and pleiotropy of each result. We finally strictly screened out 60 metabolites and metabolite ratios, including 38 metabolites of known chemical properties, 13 metabolites of unknown chemical properties, and 9 metabolite ratios (Supplementary Table SIII). Thirty-two metabolites or metabolite ratios were associated with a decreased risk of GC, and 28 metabolites or metabolite ratios were associated with an increased risk of GC.

#### Replication analysis

To further verify the results, we used the same method to analyze the latest data on GC in the open GWAS database. Fifty-one metabolites and metabolite ratios were screened out by the IVW method. The filtered IVs contained 17–42 SNPs, and the F value of each SNP was greater than 10 (Supplementary Table SIV). All the 51 metabolites and metabolite ratios, including 30 metabolites of known chemical properties, 11 metabolites of unknown chemical properties, and 10 metabolite ratios, passed the multiplex sensitivity analyses (Supplementary Table SV). Twenty metabolites or metabolite ratios were associated with reduced GC risk, and 31 metabolites or metabolite ratios were associated with increased GC risk.

Combined with previous MR analysis from the FinnGen database, we found 3 overlapping metabolites. 5-methyluridine and 1-dihomo-linolenylglycerol were associated with increased GC risk, and sphingomyelin was associated with decreased GC risk. The sensitivity analysis showed that none of the results had apparent heterogeneity or a pleiotropic effect (Table I, Figure 1), and the *p*-values after FDR correction were all less than 0.05. The forest plots are shown in Figure 2. In addition, as shown in Figure 3, the leave-out-one analysis showed that excluding any SNP would not significantly affect the overall results, supporting the reliability and stability of the MR analysis. The funnel plots were approximately symmetrical,

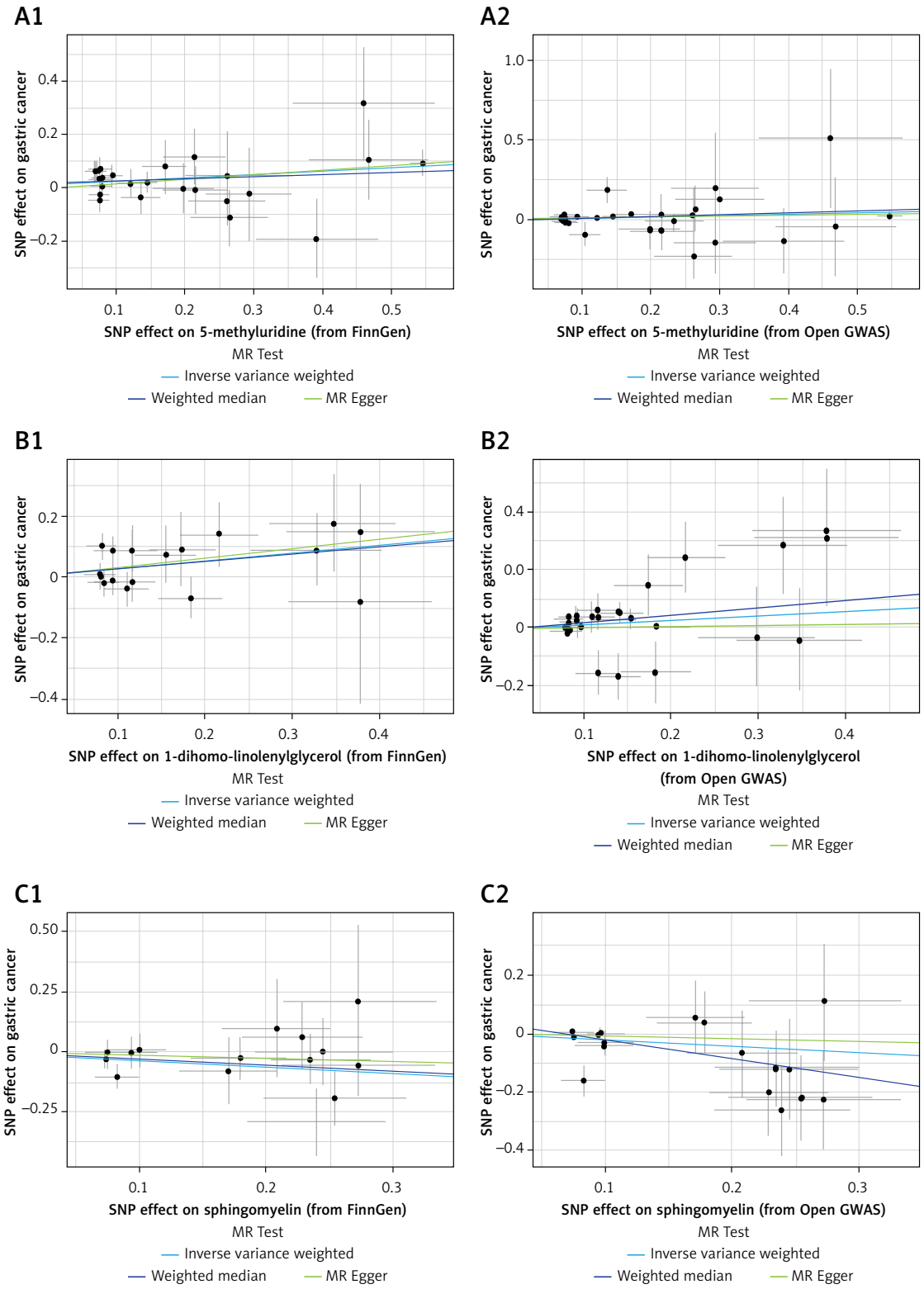


**Figure 1.** Scatter plots depicting association of overlapped metabolites and gastric cancer. **A1**, **A2** – 5-methyluridine; **B1**, **B2** – 1-dihomo-linolenylglycerol; **C1**, **C2** – sphingomyelin. Each of these points represents an instrumental variable. The vertical and horizontal lines at the center of the dot represent 95% CI. The slope of the colored line represents the size of the causal relationship. SNPs – single-nucleotide polymorphisms

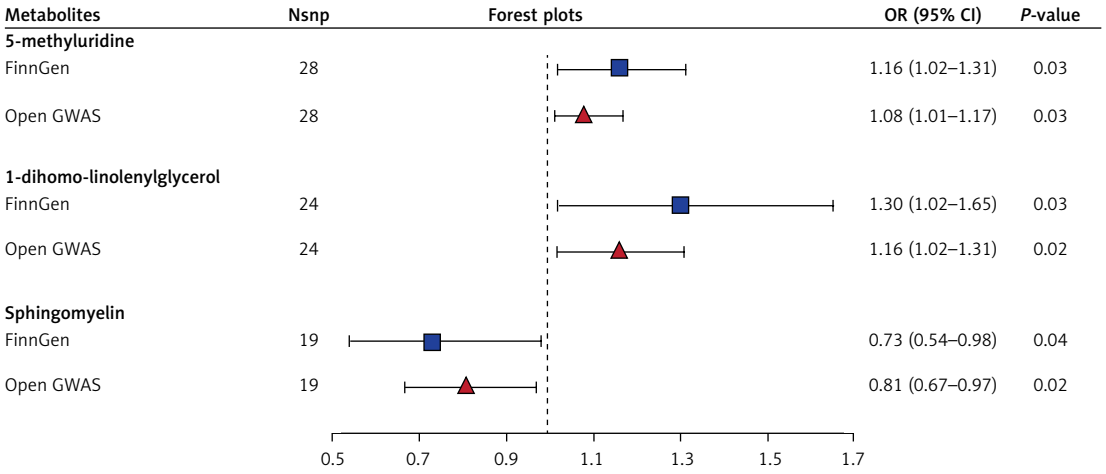


**Figure 2.** Forest plots for association of overlapped metabolites and gastric cancer risk

*SNPs – single-nucleotide polymorphisms, OR – odds ratio, CI – confidence interval, GWAS – genome-wide association study.*

indicating no publication bias in the results (Supplementary Figure S1).

### Metabolic pathway analysis

Unfortunately, because there are only 3 blood metabolites, MetaboAnalyst 5.0 cannot perform meaningful enrichment analysis.

## MR analysis of the causal relationship between 1400 human blood metabolites and metabolite ratios and the risk of CRC

### Preliminary analysis

After implementing strict quality control for each IV, we identified SNPs associated with 92 metabolites and metabolite ratios in the FinnGen database. IVs contained 12–40 SNPs. The *F* values of all SNPs were greater than 10. Detailed information about the IVs is provided in Supplementary Table SVI. We conducted the multiple sensitivity analysis to evaluate the heterogeneity and pleiotropy of each result, ultimately rigorously filtering down to 81 metabolites and metabolite ratios, including 59 metabolites with known chemical properties, 10 metabolites with unknown chemical properties, and 12 metabolite ratios (Supplementary Table SVII). Of these, 53 metabolites or metabolite ratios were negatively correlated with CRC risk, while 28 were positively correlated with CRC risk.

### Replication analyses

To further validate our results, we analyzed the latest data on CRC from the open GWAS database using the same method. We screened 110 metabolites and metabolite ratios using the IVW method. The identified IVs included 15–45 SNPs, with each SNP exhibiting an F value greater than 10 (Supplementary Table SVIII). After conducting thorough screenings, a total of 82 metabolites and metabolite ratios passed multiple sensitivity analyses. There were 48 metabolites with known chemical properties, 18 metabolites with unknown chemical properties, and 16 metabolite ratios (Supplementary Table SIX). Among these, 5 metabolites or metabolite ratios were negatively correlated with CRC risk, while 47 metabolites or metabolite ratios showed a positive correlation with CRC risk.

In conjunction with the previous MR analysis of the FinnGen database, we identified 4 overlapping metabolites and metabolite ratios associated with CRC risk. Carnitine to propionylcarnitine (C3) ratio, arachidonate to linoleate ratio, and androsterone sulfate were found to be positively correlated with CRC risk. In contrast, 1-oleoyl-2-docosahexaenoyl-GPC exhibited a negative correlation with CRC risk. The sensitivity analyses indicated that our results showed no significant heterogeneity or pleiotropy effects (Table II, Figure 4), and all *p*-values were less than 0.05 after the FDR correction. Forest plots illustrating these results can be seen in Figure 5. Additionally, the leave-one-out analysis confirmed the robustness and stability of the MR analysis (Figure 6). The funnel plots displayed a roughly symmetrical shape, indicating the absence of publication biases (Supplementary Figure S2).

### Metabolic pathway analyses

Unfortunately, because there are only 3 blood metabolites, MetaboAnalyst 5.0 cannot perform meaningful enrichment analysis.

## Discussion

We conducted a rigorous MR analysis of GWAS data from two separate databases to explore the

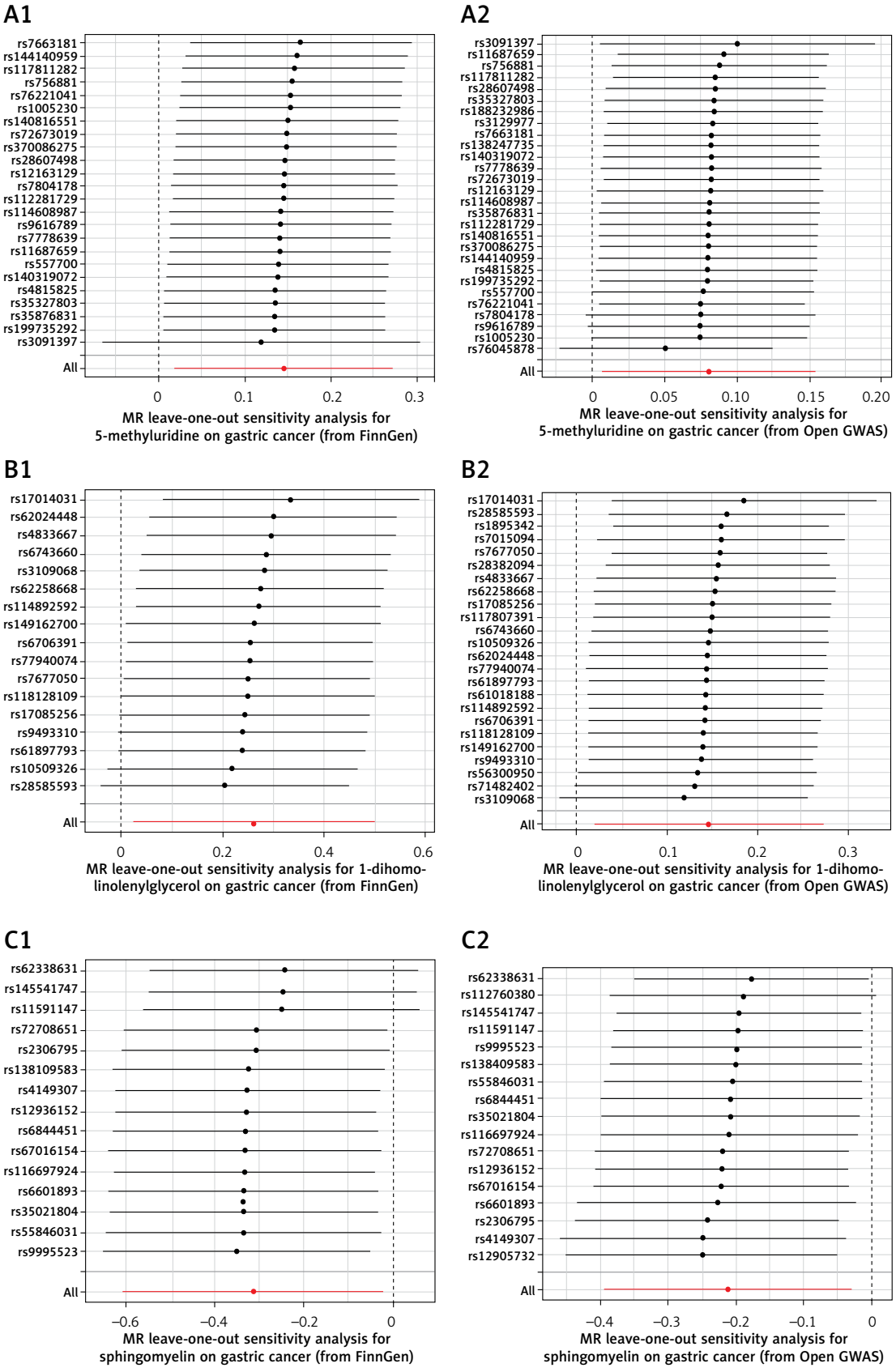


**Figure 3.** MR leave-one-out sensitivity analyses for association of overlapped metabolites and gastric cancer risk. **A1**, **A2** – 5-methyluridine; **B1**, **B2** – 1-dihomo-linolenylglycerol; **C1**, **C2** – sphingomyelin. There is no statistically significant difference between the result of a single SNP and the total result

**Table II.** Sensitivity analysis results of overlapped metabolites/ratios with respect to colorectal cancer

| Metabolites | MR-Egger | | WM | | IVW | | | MR-Egger intercept | | MR-PRESSO global test | Cochran's Q test |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | OR (95% CI) | *P*-value | OR (95% CI) | *P*-value | OR (95% CI) | *P*-value | FDR-*P* | intercept | *P*-value | *P*-value | *P*-value |
| Carnitine to propionylcarnitine (C3) ratio | | | | | | | | | | | |
| FinnGen | 1.03 (0.87–1.21) | 0.77 | 1.08 (0.94–1.25) | 0.27 | 1.11 (1.01–1.22) | 0.03 | 0.04 | 0.01 | 0.28 | 0.56 | 0.59 |
| Open GWAS | 1.00 (0.88–1.13) | 0.96 | 1.04 (0.95–1.14) | 0.41 | 1.07 (1.01–1.14) | 0.04 | 0.04 | 0.01 | 0.22 | 0.63 | 0.63 |
| 1-oleoyl-2-docosahexaenoyl-GPC | | | | | | | | | | | |
| FinnGen | 0.93 (0.69–1.25) | 0.64 | 0.85 (0.75–0.98) | 0.02 | 0.89 (0.81–0.98) | 0.02 | 0.04 | –0.01 | 0.75 | 0.78 | 0.54 |
| Open GWAS | 0.85 (0.71–1.02) | 0.10 | 0.95 (0.86–1.04) | 0.25 | 0.93 (0.87–0.99) | 0.02 | 0.04 | 0.01 | 0.34 | 0.32 | 0.29 |
| Arachidonate to linoleate ratio | | | | | | | | | | | |
| FinnGen | 1.11 (0.97–1.26) | 0.13 | 1.19 (1.09–1.30) | $1.00 \times 10^{-4}$ | 1.10 (1.02–1.19) | 0.02 | 0.04 | –0.002 | 0.88 | 0.11 | 0.12 |
| Open GWAS | 1.17 (1.08–1.27) | 0.001 | 1.16 (1.09–1.23) | $1.53 \times 10^{-6}$ | 1.12 (1.06–1.18) | $4.44 \times 10^{-5}$ | $3.55 \times 10^{-4}$ | –0.01 | 0.18 | 0.26 | 0.19 |
| Androsterone sulfate | | | | | | | | | | | |
| FinnGen | 1.05 (0.97–1.15) | 0.26 | 1.07 (0.99–1.16) | 0.05 | 1.07 (1.01–1.14) | 0.03 | 0.04 | 0.004 | 0.60 | 0.59 | 0.27 |
| Open GWAS | 1.08 (0.99–1.18) | 0.09 | 1.07 (0.99–1.16) | 0.10 | 1.05 (1.01–1.10) | 0.04 | 0.04 | 0.01 | 0.47 | 0.63 | 0.64 |

*OR – odds ratio, CI – confidence interval, GWAS – genome-wide association study, WM – weighted median, IVW – inverse variance weighted.*

causal relationships between 1400 blood metabolites and metabolite ratios and gastrointestinal tumor risk. In order to ensure the credibility of the results, we chose the intersection of the results of the two databases. The overlapping results from two authoritative databases indicated that 5-methyluridine and 1-dihomo-linolenylglycerol were positively correlated with GC risk, and sphingomyelin was negatively correlated with GC risk. Carnitine to propionylcarnitine (C3) ratio, arachidonate to linoleate ratio, and androsterone sulfate were positively correlated with CRC risk, and 1-oleoyl-2-docosahexaenoyl-GPC was negatively correlated with CRC risk.

*Helicobacter pylori* infection is recognized as a significant high-risk factor for GC [33]. Other risk factors for GC include family history, alcohol consumption, smoking, advanced age, and a high-salt diet [34]. For CRC, the main risk factors are family history, advanced age, smoking, alcohol consumption, obesity, long-term inflammatory bowel disease, a heavy intake of red meat, and an imbalanced intestinal microbiome [35, 36]. Although many risk factors associated with gastrointestinal tumors have been identified, the underlying mechanisms of their development remain complex and not fully understood. Gastrointestinal tumors often develop insidiously, with no obvious clinical symptoms in the early stages. As a result, many patients are diagnosed only at advanced stages of the disease [37]. Therefore, early diagnosis of gastrointestinal tumors is crucial in preventing further disease progression and improving patient survival. Identifying reliable serum markers can significantly facilitate the early diagnosis of gastrointestinal tumors. In recent years, the widespread use of metabolomics has prompted scientists to investigate the relationship between blood metabolites and tumor risk. However, the vast number of blood metabolites poses challenges, as many studies lack clinical validation or present conflicting results. Therefore, before more RCTs are conducted, this MR study can provide credible insights into the risk relationships between blood metabolites and gastrointestinal tumors, potentially identifying effective markers for early clinical diagnosis and treatment. Previously, Lu *et al.* examined the connection between 469 blood metabolites and the risk of gastrointestinal tumors [38], while Yun *et al.* focused on 486 blood metabolites in relation to colorectal cancer [39]. In contrast, our study investigated the relationship between 1,400 blood metabolites and metabolite ratios concerning gastrointestinal tumor risk, significantly broadening the scope of research. Additionally, unlike previous studies that relied on single databases, we integrated results from two databases, enhancing the credibility of our findings.

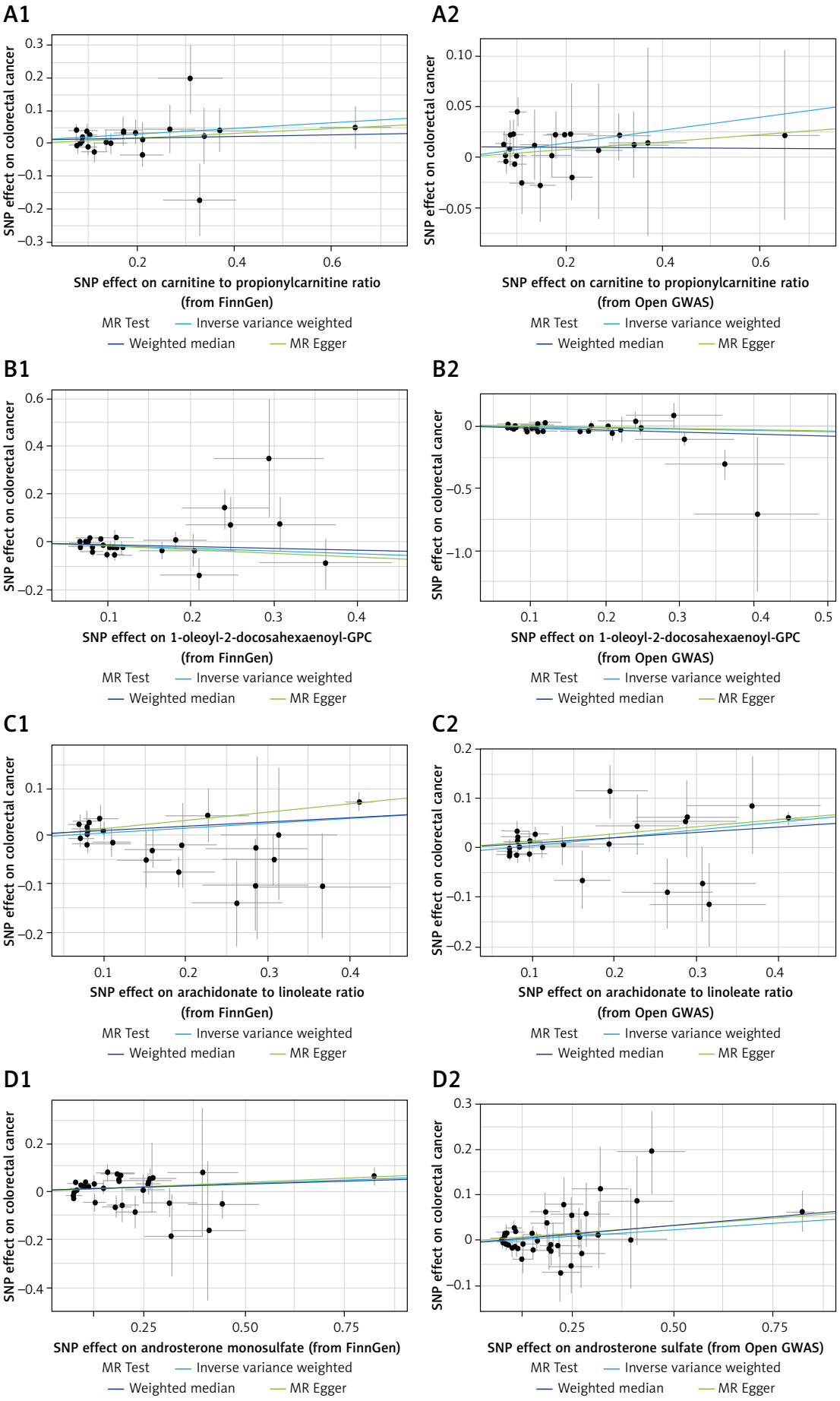


**Figure 4.** Scatter plots depicting association of overlapped metabolites/ratios and colorectal cancer risk. **A1**, **A2** – carnitine to propionylcarnitine (C3) ratio; **B1**, **B2** – 1-oleoyl-2-docosahexaenoyl-GPC; **C1**, **C2** – arachidonate to linoleate ratio; **D1**, **D2** – androsterone sulfate. Each of these points represents an instrumental variable. The vertical and horizontal lines at the center of the dot represent 95%CI. The slope of the colored line represents the size of the causal relationship

*SNPs – single-nucleotide polymorphisms.*

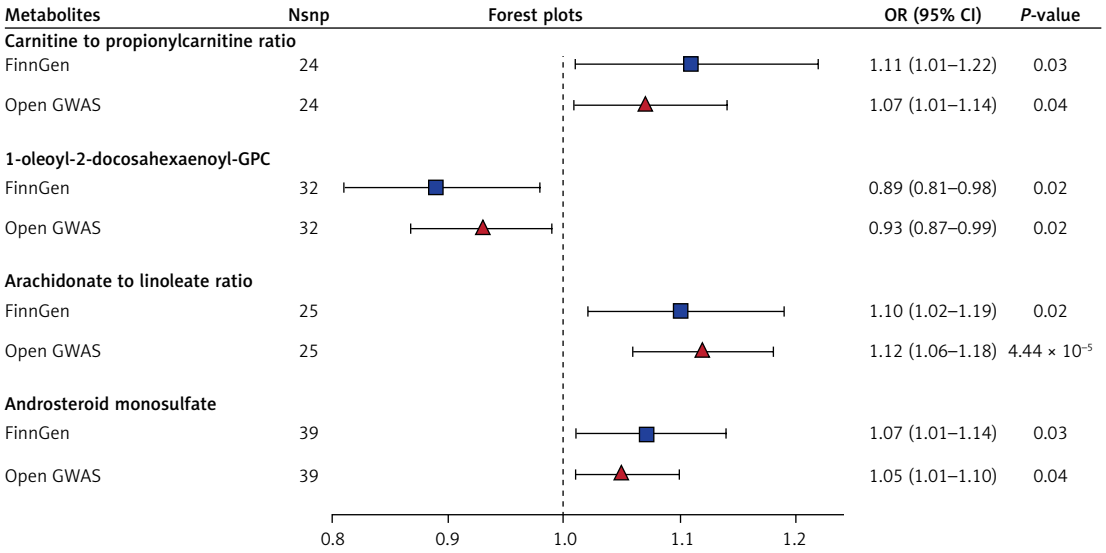


**Figure 5.** Forest plots for association of overlapped metabolites/ratios and colorectal cancer

*SNPs – single-nucleotide polymorphisms, OR – odds ratio, CI – confidence interval, GWAS – genome-wide association study.*

5-methyluridine, also known as m5U, is an integral part of mRNAs, rRNAs, tRNAs, and lncRNAs and affects their function, which plays a role in epi transcriptome variation [40, 41]. There are few studies on the relationship between m5U and tumors. Some studies have shown that m5U is related to the development of breast cancer, but the specific mechanism has not been reported [42]. m5U is the main chemical that modifies tRNA, which promotes protein synthesis [43]. The growth, invasion, and metastasis of tumor cells require a large number of proteins [44, 45]. Therefore, we speculated that the excessive modification of m5U could provide abundant essential proteins for gastrointestinal tumor cells, thus promoting the occurrence and development of tumors. Few studies have conducted in-depth research on 1-dihomo-linolenylglycerol, and its relationship with gastric cancer risk requires further experimental confirmation. Sphingomyelin is an indispensable substance for cells. It is closely related to lipid transport and affects cell proliferation [46]. Recent studies have shown that sphingomyelin can inhibit tumor development. For instance, Fhaner *et al.* found that metastatic CRC cells exhibit lower levels of sphingomyelin [47]. Similarly, Wang *et al.* reported that the most aggressive breast cancer cells have lower sphingomyelin levels compared to normal cells [48]. Additionally, sphingomyelin content in lung and esophageal cancer tissues has also been observed to decline [49, 50]. Furthermore, silencing the sphingomyelin

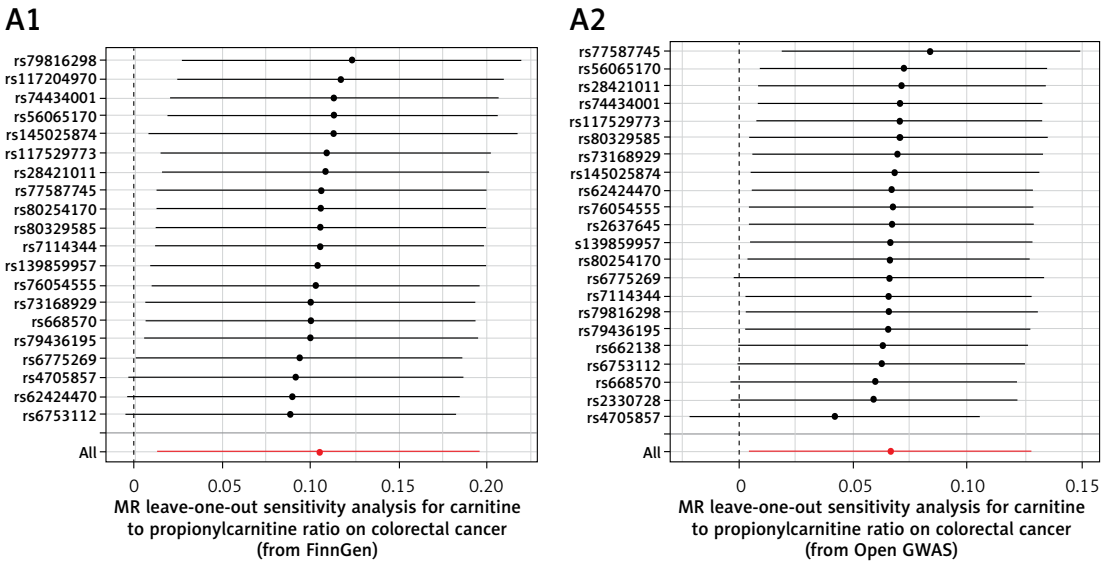


**Figure 6.** MR leave-one-out sensitivity analyses for overlapped metabolites/ratios and colorectal cancer. **A1**, **A2** – carnitine to propionylcarnitine (C3) ratio

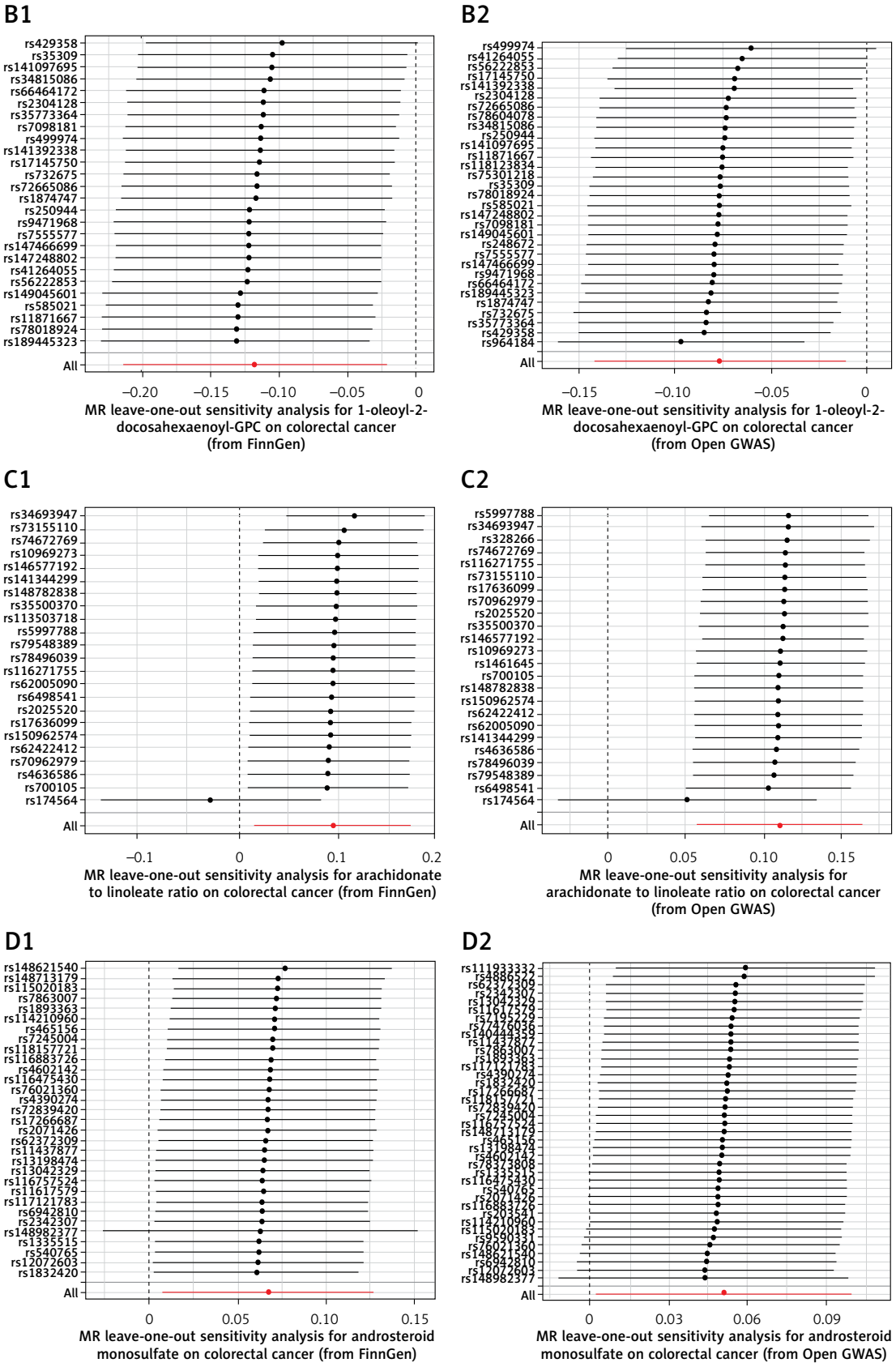


**Figure 6.** Cont. **B1**, **B2** – 1-oleoyl-2-docosahexaenoyl-GPC; **C1**, **C2** – arachidonate to linoleate ratio; **D1**, **D2** – androsterone sulfate. There is no statistically significant difference between the result of a single SNP and the total result

synthase SGMS1 disrupts adhesion and junctions between renal collecting duct cells, transforming epithelial cells into mesenchymal cells [51]. Moreover, the activation of sphingomyelin biosynthesis through 2-hydroxyoleic acid has been shown to induce cell death in glioma cell lines [52]. Sphingomyelin also has the potential to enhance the anti-tumor effects of chemotherapeutic drugs that interfere with lipid metabolism. For example, alkyl-lysophospholipid tends to accumulate in micro-domains rich in sphingomyelin. Increasing sphingomyelin levels can therefore improve the bioavailability of these chemotherapeutic agents, subsequently enhancing their anti-tumor effects [53]. Sphingomyelin influences tumor cell proliferation through various biological processes, including cell cycle arrest, endoplasmic reticulum stress, autophagy, and sphingomyelin deposition. However, the specific mechanisms underlying these effects require further investigation.

Carnitine plays a vital role in energy metabolism, primarily by transporting fatty acids to the mitochondria for β-oxidation [54]. Increasing evidence suggests that carnitine may inhibit tumor development. For instance, Chang *et al*. discovered that carnitine can prevent mitochondrial damage and impede the progression of liver cancer [55]. Furthermore, it has been reported that inhibiting carnitine metabolism can promote the stem cell-like nature of liver cancer [56]. In contrast, there is limited research on the relationship between propionylcarnitine and tumors. Current findings suggest that propionylcarnitine may contribute to cardiovascular diseases through mechanisms such as insulin resistance, energy metabolism disorders, and inflammation [57]. These same factors can also promote tumor development, indicating that propionylcarnitine might act as a metabolite that supports tumor growth. Overall, carnitine and propionylcarnitine exhibit different effects on tumor occurrence. The combined influence of these two substances on tumor promotion, especially concerning gastrointestinal tumors, requires further investigation through RCTs. Arachidonate is catalyzed by cyclooxygenase (COX)-1 and COX-2 to produce prostaglandins and leukotrienes, which have been shown to induce tumor microenvironment remodeling and immunosuppression, thereby promoting the development of various tumors, including CRC [58]. Research has confirmed that prostaglandins can enhance the proliferation of CRC cells through the RAS-ERK and β-catenin signaling pathways [59, 60]. Linoleate is an essential unsaturated fatty acid for humans; however, excessive intake of linoleate has been associated with an increased risk of CRC [61, 62]. During the production of prostaglandin E2, linoleate interacts with free radicals produced by COX, contributing to CRC development [63]. The ratio of arachidonate to linoleate is expected to serve as a potential indicator for the clinical diagnosis of CRC, although more RCTs are needed to confirm this. Reports on androsterone sulfate are limited, and its physiological role remains largely unknown. The few existing studies indicate a close relationship to cardiovascular disease [64]. There are many similarities between the mechanisms of blood vessel formation and tumor development, suggesting that androsterone sulfate may emerge as a potential tumor marker. Given the mechanistic overlaps between cardiovascular diseases and tumors, androsterone sulfate might be a promising marker in tumor biology. Lastly, 1-oleoyl-2-docosahexaenoyl-GPC is a phospholipid, with no RCTs demonstrating an association with tumors. However, an MR study has confirmed that 1-oleoyl-2-docosahexaenoyl-GPC serves as a protective factor against thoracic aortic aneurysm [65]. Our study concluded that 1-oleoyl-2-docosahexaenoyl-GPC is also protective against CRC. Further research is needed to explore how 1-oleoyl-2-docosahexaenoyl-GPC exerts its protective effects.

This MR study has several notable advantages. First, unlike previous observational studies, MR studies can simulate RCTs and are not impacted by reverse causation or confounding factors. Second, our study was conducted and tested for heterogeneity, bias, and horizontal pleiotropy in strict accordance with the three main assumptions of MR studies, and the results obtained were plausible. Third, compared with previous MR studies with several hundred metabolites and disease causality, we included 1400 metabolites and metabolite ratios, which explored a wider range and uncovered more metabolites and ratios. Fourth, unlike earlier MR analyses that relied on single databases, we combined results from the two largest and most authoritative databases, enhancing the understanding of the relationship between 1,400 metabolites and metabolite ratios and the risk of gastrointestinal tumors. Lastly, we applied the FDR method to correct *p*-values, ensuring that our results are more rigorous.

However, there are some limitations to this study. Of note, the relaxation of the significance threshold for selecting IVs is common in MR studies, but in this case, due to the limited number of SNPs with genome-wide significance, it potentially violates the relevance assumption of MR analysis. While the study ensured that the *F*-values of all SNPs were greater than 10 to exclude weak IVs, the relaxed threshold may still compromise the robustness of the causal inferences. Furthermore, genetic variations and metabolic profiles can differ significantly across populations, and the identified metabolites may not have the same associations

in non-European cohorts. Another limitation is the small number of metabolites that survived the rigorous screening process, which hindered meaningful metabolic pathway analysis using MetaboAnalyst 5.0. The inability to perform enrichment analysis limits the understanding of the biological mechanisms underlying the observed associations. Additionally, the sample size, while substantial, may still be insufficient to detect weaker associations or rare metabolic signals. Also, while MR studies are less susceptible to confounding and reverse causation compared to observational studies, they are not immune to horizontal pleiotropy, where genetic variants influence the outcome through pathways other than the exposure. Although the study employed multiple sensitivity analyses to address this issue, residual pleiotropy could still affect the results. From another viewpoint, the study's focus on blood metabolites may not capture the full complexity of metabolic interactions within the tumor microenvironment. Tissue-specific metabolites or those derived from the gut microbiome could also play significant roles in gastrointestinal tumor development but were not addressed in this analysis. Furthermore, the study did not account for potential interactions between metabolites, such as synergistic or antagonistic effects, which could influence tumor risk. Lastly, the clinical utility of these metabolites remains to be established. Translating these biomarkers into diagnostic or prognostic tools will require extensive validation in independent cohorts and experimental studies to elucidate their functional roles in tumor biology.

In conclusion, this MR study is the first to investigate the causal relationship between 1400 blood metabolites and metabolite ratios and gastrointestinal tumor risk. The overlapping results of two authoritative databases show that 3 blood metabolites are associated with the risk of GC; 4 blood metabolites and ratios are associated with the risk of CRC. The results can guide early clinical diagnosis and treatment of gastrointestinal tumors.

## Funding

No external funding.

## Ethical approval

Not applicable.

## Conflict of interest

The authors declare no conflict of interest.

## References

1. Siegel RL, Miller KD, Wagle NS, et al. Cancer statistics, 2023. CA Cancer J Clin 2023; 73: 17-48.
2. Ilic M, Ilic I. Epidemiology of stomach cancer. World J Gastroenterol 2022; 28: 1187-203.
3. Sung H, Ferlay J, Siegel RL, 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: 209-49.
4. Smyth EC, Nilsson M, Grabsch HI, et al. Gastric cancer. Lancet 2020; 396: 635-48.
5. Murphy N, Moreno V, Hughes DJ, et al. Lifestyle and dietary environmental factors in colorectal cancer susceptibility. Mol Aspects Med 2019; 69: 2-9.
6. Hughes LAE, Simons CCJM, van den Brandt PA, et al. Lifestyle, diet, and colorectal cancer risk according to (Epi)genetic instability: current evidence and future directions of molecular pathological epidemiology. Curr Colorectal Cancer Rep 2017; 13: 455-69.
7. Johnson CH, Ivanisevic J, Siuzdak G. Metabolomics: beyond biomarkers and towards mechanisms. Nat Rev Mol Cell Biol 2016; 17: 451-9.
8. Wishart DS. Metabolomics for investigating physiological and pathophysiological processes. Physiol Rev 2019; 99: 1819-75.
9. Kala P, Hnat T, Padrova K, et al. Eicosanoids in human heart failure: pilot study of plasma epoxyeicosatrienoic and dihydroxyeicosatrienoic acid levels. Arch Med Sci 2023; 19: 513-7.
10. Antonowicz S, Kumar S, Wiggins T, et al. Diagnostic metabolomic blood tests for endoluminal gastrointestinal cancer--a systematic review and assessment of quality. Cancer Epidemiol Biomarkers Prev 2016; 25: 6-15.
11. Song G, Wang L, Tang J, et al. Circulating metabolites as potential biomarkers for the early detection and prognosis surveillance of gastrointestinal cancers. Metabolomics 2023; 19: 36.
12. Qu R, Zhang Y, Ma Y, et al. Role of the gut microbiota and its metabolites in tumorigenesis or development of colorectal cancer. Adv Sci (Weinh) 2023; 10: e2205563.
13. Dai D, Yang Y, Yu J, et al. Interactions between gastric microbiota and metabolites in gastric cancer. Cell Death Dis 2021; 12: 1104.
14. Coker OO, Liu C, Wu WKK, et al. Altered gut metabolites and microbiota interactions are implicated in colorectal carcinogenesis and can be non-invasive diagnostic biomarkers. Microbiome 2022; 10: 35.
15. Zuccolo L, Holmes MV. Commentary: Mendelian randomization-inspired causal inference in the absence of genetic data. Int J Epidemiol 2017; 46: 962-5.
16. Liang G, Miao D, Du C. Causal associations between blood metabolites and breast cancer. Arch Med Sci 2024; 21: 206-14.
17. Castle WE. Mendel's law of heredity. Science 1903; 18: 396-406.
18. Zheng J, Baird D, Borges MC, et al. Recent developments in mendelian randomization studies. Curr Epidemiol Rep 2017; 4: 330-45.
19. Thomas DC, Conti DV. Commentary: the concept of 'Mendelian Randomization'. Int J Epidemiol 2004; 33: 21-5.
20. Chen Y, Lu T, Pettersson-Kymmer U, et al. Genomic atlas of the plasma metabolome prioritizes metabolites implicated in human diseases. Nat Genet 2023; 55: 44-53.
21. Sakaue S, Kanai M, Tanigawa Y, et al. A cross-population atlas of genetic associations for 220 human phenotypes. Nat Genet 2021; 53: 1415-24.
22. Cai J, Li X, Wu S, et al. Assessing the causal association between human blood metabolites and the risk of epilepsy. J Transl Med 2022; 20: 437.

23. Xiao G, He Q, Liu L, et al. Causality of genetically determined metabolites on anxiety disorders: a two-sample Mendelian randomization study. J Transl Med 2022; 20: 475.
24. Lawlor DA, Harbord RM, Sterne JA, et al. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. Stat Med 2008; 27: 1133-63.
25. Hemani G, Zheng J, Elsworth B, et al. The MR-Base platform supports systematic causal inference across the human phenome. Elife 2018; 7: e34408.
26. Burgess S, Bowden J, Fall T, et al. Sensitivity analyses for robust causal inference from mendelian randomization analyses with multiple genetic variants. Epidemiology 2017; 28: 30-42.
27. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. Genet Epidemiol 2013; 37: 658-65.
28. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. Int J Epidemiol 2015; 44: 512-25.
29. Bowden J, Davey Smith G, Haycock PC, et al. Consistent estimation in mendelian randomization with some invalid instruments using a weighted median estimator. Genet Epidemiol 2016; 40: 304-14.
30. Gronau QF, Wagenmakers EJ. Limitations of Bayesian leave-one-out cross-validation for model selection. Comput Brain Behav 2019; 2: 1-11.
31. Yuan S, Kim JH, Xu P, et al. Causal association between celiac disease and inflammatory bowel disease: a two-sample bidirectional Mendelian randomization study. Front Immunol 2023; 13: 1057253.
32. Verbanck M, Chen CY, Neale B, et al. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. Nat Genet 2018; 50: 693-8.
33. Sánchez-López JY, Díaz-Herrera LC, Rizo-de la Torre LDC. Pepsinogen I, pepsinogen II, gastrin-17, and Helicobacter pylori serological biomarkers in the diagnosis of precursor lesions of gastric cancer. Arch Med Sci 2024; 20: 1016-21.
34. Smyth EC, Nilsson M, Grabsch HI, et al. Gastric cancer. Lancet 2020; 396: 635-48.
35. Dekker E, Tanis PJ, Vleugels JLA, et al. Colorectal cancer. Lancet 2019; 394: 1467-80.
36. Xing C, Zhihao L, Ji D. Diagnostic value of fecal Fusobacterium nucleatum in colorectal cancer. Arch Med Sci 2023; 19: 1929-33.
37. Guan WL, He Y, Xu RH. Gastric cancer treatment: recent progress and future perspectives. J Hematol Oncol 2023; 16: 57.
38. Lu J, Feng Y, Guo K, et al. Association between human blood metabolome and the risk of gastrointestinal tumors. PLoS One 2024; 19: e0304574.
39. Yun Z, Guo Z, Li X, et al. Genetically predicted 486 blood metabolites in relation to risk of colorectal cancer: a Mendelian randomization study. Cancer Med 2023; 12: 13784-99.
40. Desrosiers R, Friderici K, Rottman F. Identification of methylated nucleosides in messenger RNA from Novikoff hepatoma cells. Proc Natl Acad Sci USA 1974; 71: 3971-5.
41. Shaon MSH, Karim T, Ali MM, et al. A robust deep learning approach for identification of RNA 5-methyluridine sites. Sci Rep 2024; 14: 25688.
42. Jonkhout N, Tran J, Smith MA, et al. The RNA modification landscape in human disease. RNA 2017; 23: 1754-69.
43. Witzenberger M, Burczyk S, Settele D, et al. Human TRMT2A methylates tRNA and contributes to translation fidelity. Nucleic Acids Res 2023; 51: 8691-710.
44. Cardano M, Tribioli C, Prosperi E. Targeting proliferating cell nuclear antigen (PCNA) as an effective strategy to inhibit tumor cell proliferation. Curr Cancer Drug Targets 2020; 20: 240-52.
45. Steeg PS. Tumor metastasis: mechanistic insights and clinical challenges. Nat Med 2006; 12: 895-904.
46. Shakor ABA, Taniguchi M, Kitatani K, et al. Sphingomyelin synthase 1-generated sphingomyelin plays an important role in transferrin trafficking and cell proliferation. J Biol Chem 2011; 286: 36053-62.
47. Fhaner CJ, Liu S, Ji H, et al. Comprehensive lipidome profiling of isogenic primary and metastatic colon adenocarcinoma cell lines. Anal Chem 2012; 84: 8917-26.
48. Wang S, Chen X, Luan H, et al. Matrix-assisted laser desorption/ionization mass spectrometry imaging of cell cultures for the lipidomic analysis of potential lipid markers in human breast cancer invasion. Rapid Commun Mass Spectrom 2016; 30: 533-42.
49. Marien E, Meister M, Muley T, et al. Non-small cell lung cancer is characterized by dramatic changes in phospholipid profiles. Int J Cancer 2015; 137: 1539-48.
50. Merchant TE, de Graaf PW, Minsky BD, et al. Esophageal cancer phospholipid characterization by 31P NMR. NMR Biomed 1993; 6: 187-93.
51. Brandán YR, Guaytima EDV, Favale NO, et al. The inhibition of sphingomyelin synthase 1 activity induces collecting duct cells to lose their epithelial phenotype. Biochim Biophys Acta Mol Cell Res 2018; 1865: 309-22.
52. Barceló-Coblijn G, Martin ML, de Almeida RF, et al. Sphingomyelin and sphingomyelin synthase (SMS) in the malignant transformation of glioma cells and in 2-hydroxyoleic acid therapy. Proc Natl Acad Sci USA 2011; 108: 19569-74.
53. Van der Luit AH, Budde M, Zerp S, et al. Resistance to alkyl-lysophospholipid-induced apoptosis due to downregulated sphingomyelin synthase 1 expression with consequent sphingomyelin- and cholesterol-deficiency in lipid rafts. Biochem J 2007; 401: 541-9.
54. Gnoni A, Longo S, Gnoni GV, et al. Carnitine in human muscle bioenergetics: can carnitine supplementation improve physical exercise? Molecules 2020; 25: 182.
55. Chang B, Nishikawa M, Nishiguchi S, et al. L-carnitine inhibits hepatocarcinogenesis via protection of mitochondria. Int J Cancer 2005; 113: 719-29.
56. Wang J, Zhou Y, Zhang D, et al. CRIP1 suppresses BBOX1-mediated carnitine metabolism to promote stemness in hepatocellular carcinoma. EMBO J 2022; 41: e110218.
57. Lee YH, Park S. Genetic and lifestyle-related factors influencing serum hyper-propionylcarnitine concentrations and their association with metabolic syndrome and cardiovascular disease risk. Int J Mol Sci 2023; 24: 15810.
58. Wang D, Dubois RN. Eicosanoids and cancer. Nat Rev Cancer 2010; 10: 181-93.
59. Wang D, Buchanan FG, Wang H, et al. Prostaglandin E2 enhances intestinal adenoma growth via activation of the Ras-mitogen-activated protein kinase cascade. Cancer Res 2005; 65: 1822-9.
60. Castellone MD, Teramoto H, Williams BO, et al. Prostaglandin E2 promotes colon cancer cell growth through

a Gs-axin-beta-catenin signaling axis. Science 2005; 310: 1504-10.
61. Daniel CR, McCullough ML, Patel RC, et al. Dietary intake of omega-6 and omega-3 fatty acids and risk of colorectal cancer in a prospective cohort of U.S. men and women. Cancer Epidemiol Biomarkers Prev 2009; 18: 516-25.
62. Pot GK, Geelen A, van Heijningen EM, et al. Opposing associations of serum n-3 and n-6 polyunsaturated fatty acids with colorectal adenoma risk: an endoscopy-based case-control study. Int J Cancer 2008; 123: 1974-7.
63. Karpisheh V, Nikkhoo A, Hojjat-Farsangi M, et al. Prostaglandin E2 as a potent therapeutic target for treatment of colon cancer. Prostaglandins Other Lipid Mediat 2019; 144: 106338.
64. Yu B, Heiss G, Alexander D, et al. Associations between the serum metabolome and all-cause mortality among African Americans in the Atherosclerosis Risk in Communities (ARIC) study. Am J Epidemiol 2016; 183: 650-6.
65. Tong X, Cui Y. Mendelian randomization analysis of the causal relationship between serum metabolites and thoracic aortic aneurysm. Medicine (Baltimore) 2024; 103: e39686.