Name of Journal: World Journal of Gastrointestinal Oncology
Manuscript NO: 92204
Manuscript Type: ORIGINAL ARTICLE

Clinical and Translational Research
Casual associations between Blood Metabolites and colon cancer

A Mendelian randomization study

Keyue Hu, Yiquan Cheng, Zhilong Shi, Fupeng Ren, Gangfeng Xiao
Abstract

BACKGROUND

Limited knowledge exists regarding the casual associations linking blood metabolites and the risk of developing colorectal cancer.

AIM

To investigate causal associations between blood metabolites and colon cancer.

METHODS

The study utilized a two-sample Mendelian randomization (MR) analysis to investigate the causal impact of 486 blood metabolites on colorectal cancer. The primary method of analysis used was the inverse variance weighted (IVW) model. To further validate the results several sensitivity analyses were performed, including Cochrane's Q test, MR-Egger intercept test, and Mendelian randomization robust adjusted profile score (MR-RAPS). These additional analyses were conducted to ensure the reliability and robustness of the findings.

RESULTS

After rigorous selection for genetic variation, 486 blood metabolites were included in the MR analysis. We found Mannose (odds ratio (OR) = 2.09 (1.10-3.97), \( P = 0.024 \)), N-acetylglycine (OR = 3.14 (1.78-5.53), \( P = 7.54 \times 10^{-8} \)), X-11593--O-methylascorbate* (OR = 1.68 (1.04-2.72), \( P = 0.034 \)), 1-arachidonoylglycerophosphocholine* (OR = 4.23 (2.51-7.12), \( P = 6.35 \times 10^{-8} \)) and 1- arachidonoylglycerophosphoethanolamine*4 (OR = 3.99 (1.17-13.54), \( P = 0.027 \)) were positively causally associated with colorectal cancer, and we also found a negative causal relationship between Tyrosine (OR = 0.08 (0.01-0.63), \( P = 0.014 \)), Urate (OR = 0.25 (0.10-0.62), \( P = 0.003 \)), N-acetylglycine (OR = 0.73 (0.54-0.98), \( P = 0.033 \)), X-12092 (OR = 0.89 (0.81-0.99), \( P = 0.028 \)), Succinylcarnitine (OR = 0.48 (0.27-0.84),
P = 0.09] with colorectal cancer. A series of sensitivity analyses were performed to confirm the rigidity of the results.

CONCLUSION
This study showed a causal relationship between 10 blood metabolites and colorectal cancer, of which 5 blood metabolites were found to be causal for the development of colorectal cancer and were confirmed as risk factors. The other five blood metabolites are protective factors.

Key Words: Metabolites; Colon cancer; Mendelian randomization; GWAS; Casual

Hu K, Cheng Y, Shi Z, Ren F, Xiao G. Casual associations between Blood Metabolites and colon cancer. World J Gastrointest Oncol 2024; In press

Core Tip: Our findings demonstrate a casual relationship between Mannose, N-acetylglycine, X - 11592 - O - methylascorbate *, 1 - arachidonoylglycerophosphocholine *, and 1 - arachidonoylglycerophosphoethanolamine * and colorectal cancer.

INTRODUCTION
Colorectal cancer is globally ranked as the third most prevalent type of cancer, and it is the second leading cause of cancer-related mortality. In 2018, there were approximately 18 million new cases reported, resulting in 860,000 deaths[3]. Projections based on population data suggest that the annual burden of colorectal cancer will exceed 3 million new cases and 16 million deaths by 2040[1, 2, 3]. Disparities in colorectal cancer incidence between countries and insights obtained from international migration studies have indicated a potential correlation between diet, lifestyle factors, and the development of disease[3]. Metabolites, serving as substrates and products of metabolism. Not, are indispensable cellular components. They not only drive fundamental cellular activities but also play a
critical role as functional intermediates in predicting or influencing the onset and progression of diseases[4, 5, 6]. Observational studies have identified notable differences in blood metabolites between colorectal cancer patients and healthy individuals, primarily involving inflammation-related pathways, amino acids, and lipid metabolism[5]. For instance, Su et al conducted a study that utilized liquid chromatography mass spectrometry (LC-MS/MS) metabolomics to examine small metabolites in serum samples from colorectal cancer patients. The results indicated a significant increase in S-(3-methylbutyryl)-dihydrolipoamide-E and N-nonyl-glycine, while S-phenyl-d-cysteine demonstrated a substantial decrease in colorectal cancer patients compared to the control group[7]. However, traditional observational studies are prone to confounding factors and reverse causality, which have resulted in ongoing debates concerning the causal relationship between colorectal cancer and blood metabolites.

Mendelian Randomization (MR) is a statistical method used to assess causality in diseases of interest by utilizing single nucleotide polymorphisms (SNPs) as instrumental variables (IVs) for associated risk factors[8, 9]. This method estimates the causal relationship between an exposure and an outcome based on genetic variation[10]. Similar to randomized controlled trials[10, 11], genetic variants are randomly assigned to offspring along with gametes before the development of disease. As a result, they are less susceptible to confounding factors and reverse causality. While MR studies have been conducted on blood metabolites in various diseases, no studies have been performed on colorectal cancer.

This study utilized the metabolite database from a highly comprehensive metabolite study for exposure assessment. A systematic two-sample MR analysis was employed, using the GECCO, CORECT, CCFR, and other European cohorts’ genome-wide association studies (GWAS) data as the phenotypic data for colorectal cancer[12]. The study extensively examines the causal association between 486 blood metabolites and colorectal cancer, thereby providing insights into the etiology and pathogenesis of
metabolic-related colorectal cancer. These findings have important implications for risk prediction and treatment approaches.

MATERIALS AND METHODS

Study Design

A two-sample MR analysis was conducted to assess the causal impact of human blood metabolites on the risk of colorectal cancer, using summary statistics obtained from GWAS. Three assumptions needed to be met by the chosen IVs: (1) They must exhibit a strong association with the exposure of interest, i.e., human blood metabolites; (2) The IVs must be independent of unobserved confounders; and (3) The IVs must have a relationship with the outcome, i.e., colorectal cancer, solely through their influence on the exposure of interest, rather than via confounding factors. In this MR study, human blood metabolites were considered as the exposure, while colorectal cancer was treated as the outcome. The entire process is depicted in the flow chart shown in Figure 1. This study utilizes magnetic resonance imaging (MRI) to investigate the causal associations between blood metabolites in humans (exposure) and colorectal cancer (outcome). The study assumes that the IVs are specifically related to metabolites, and do not have any connection to confounding factors. Additionally, it assumes that the IVs are not associated with the risk of developing colorectal cancer with respect to both metabolites and confounding factors. The Ningbo No. 2 Hospital reviewed and approved this study.

Human blood metabolite samples

The metabolite database used in this study was obtained from a comprehensive metabolite study conducted by Shin et al (13). The study utilized GWAS data of human blood metabolites, which was obtained from the metabolomics GWAS server (https://metabolomics.helmholtz-muenchen.de/gwas/). The study cohort consisted of 7,824 individuals of European descent, and a total of 2.1 million SNPs were tested for 486 different metabolites. Among these metabolites, 177 were classified as unknown. In addition, 309 metabolites were classified based on their chemistry using the Kyoto
Encyclopedia of Genes and Genomes (KEGG) database. These metabolites were further assigned to eight broad metabolomes, which include amino acids, peptides, lipids, cofactors and vitamins, carbohydrates, energy, nucleotides, and xenobiotics (Table 1). Colorectal cancer samples

Summary statistics of colorectal cancer phenotypes were obtained from a GWAS conducted by Huyghe et al. The study utilized cohorts from various populations, including GECCO, CORECT, CCFR, and other European populations. The GWAS data comprised a total of 19,948 cases and 12,124 controls, yielding a final dataset of 32,072 samples. The research data sources are publicly available in the IEU GWAS database (https://gwas.mrcieu.ac.uk), and further details can be found in Table 1.

Selection of IVs

In this MRI study, we adjusted the association threshold to \( p < 1.0 \times 10^{-5} \), in accordance with the findings of Cai et al. For the investigation of metabolites, we identified eleven SNPs that exhibited no linkage disequilibrium (LD) with other SNPs (\( r^2 < 0.1 \) within a clustering window of 500 kb). These SNPs were selected as genetic tools. The F-statistics were calculated to assess genetic variation, and SNPs with F-statistics below 10 were excluded due to inadequate strength. To ensure accuracy in allele coding and strand orientation, we removed palindromic SNPs. During the alignment process, we aligned the alleles to the human genome reference sequence (build 37) and eliminated any ambiguous or duplicate SNPs. Metabolites that demonstrated significant associations with the outcome (\( p < 1.0 \times 10^{-5} \)) and those that were associated with fewer than three SNPs were excluded. The remaining SNPs, following the aforementioned procedure, were then utilized as IVs.

Mendelian randomization analysis

We conducted a MR analysis to assess the causal association between blood metabolites and colorectal cancer. For the main analysis, we employed the standard inverse variance weighted (IVW) method, which combines the Wald ratios of individual SNPs with outcomes to obtain pooled estimates of causality. This method considers potential overdispersion and is widely recognized and utilized in the field of MR analysis.
In addition, several other MR analyses have been used as complementary approaches to the IVW method. These include MR-Egger regression and the weighted median method, which aim to improve the robustness of estimates in a wider range of scenarios. MR-Egger regression is particularly useful for assessing and testing unbalanced pleiotropy and substantial heterogeneity, although it requires a larger sample size compared to IVW to achieve the same of level precision in assessing underexposure variation. On the other hand, the weighted median method can provide consistent effect estimates as long as at least half of the weighted variance is valid, even in the presence of horizontal pleiotropy.

Sensitivity analysis plays a crucial role in MR studies as it helps to identify potential genetic polymorphisms and heterogeneity in MR estimates. In order to achieve this, we conducted additional analyses using maximum likelihood, MR-RAPS, and MR-Egger intercept tests with the aim of detecting the presence of pleiotropy and evaluating the robustness of our findings. To assess the reliability of the causal estimation assumed by the IVW method, we employed MR-Egger\(^{[16]}\), maximum likelihood\(^{[17]}\), and robust adjusted Profile score (MR-RAPS)\(^{[18]}\). Moreover, to examine heterogeneity among SNPs, we utilized Cochran Q values obtained from the IVW and MR-Egger models. Consistent with common practice, we set a significance threshold of \(P < 0.05\) to indicate the presence of significant heterogeneity. Lastly, to evaluate potential horizontal pleiotropy, we utilized MR-Egger regression, where an intercept close to 0 and a P-value greater than 0.05 suggest the absence of pleiotropy in the SNPs. The identification of eligible candidate metabolites related to colorectal cancer was based on the following criteria: (1) a significant agreement among the three MR Methods in terms of direction and magnitude (all \(p < 0.05\)); (2) the absence of heterogeneity; (3) the absence of pleiotropy at any level. All analyses were performed using the R package TwoSampleMR (version 0.5.6) in R (version 4.0.0).

RESULTS
After a comprehensive quality control process, IVW identified a total of 104 IVs associated with colorectal cancer (Table S1; Table S2). The number of SNPs ranged from 5 to 19 for each metabolite, with all F values >10 considered for SNPs. After applying the Bonferroni correction, the IVW analysis revealed evidence of association between colorectal cancer and only 10 metabolites (Figure 2; Figure 3), and these results remained robust even after supplementary and sensitivity analyses. Among the known metabolites, 5 showed a positive correlation with colorectal cancer, while 5 showed a negative correlation. Specifically, genetically determined high levels of Mannose (odds ratio [OR] = 2.09 [1.10-3.97], \( P = 0.024 \)), N-acetylglucosamine (OR = 0.73 [0.54-0.98], \( P = 0.033 \)), X - 11593 - O - methylascorbate* (OR = 1.68 [1.04-2.72], \( P = 0.034 \)), 1 - arachidonoylglycerophosphocholine* (OR = 4.23 [2.51-7.12], \( P = 6.35 \times 10^{-8} \)), and 1 - arachidonoylglycerophosphoethanolamine* (OR = 0.48 [0.27-0.83], \( P = 0.027 \)) were associated with the occurrence and development of colorectal cancer. Additionally, the IVW method identified genetically determined high levels of Arachidonate (20:4n6) (OR = 3.14 [1.78-5.53], \( P = 7.54 \times 10^{-5} \)), Tyrosine (OR = 0.08 [0.01-0.61], \( P = 0.015 \)), and arachidonate (20:4n6) (OR = 3.14 [1.78-5.53], \( P = 7.54 \times 10^{-5} \)) did not have a causal relationship with colorectal cancer. Urate (OR = 0.25 [0.10-0.61], \( P = 0.003 \)), X - 12092 (OR = 0.89 [0.81-0.99], \( P = 0.028 \)), and Succinylcarntine (OR = 0.48 [0.27-0.83], \( P = 0.009 \)) also did not show a causal relationship with colorectal cancer. Notably, the strongest positive correlation was observed between 1 - arachidonoylglycerophosphocholine* (OR [95%CI]: 4.23 [2.51-7.12]) and 1 - arachidonoylglycerophosphoethanolamine* (OR [95%CI]: 3.99 [1.17-13.54]), while the strongest negative correlation was observed for Tyrosine (OR [95%CI]: 0.08 [0.01, 0.61]; Figure 2; Figure 3).

Sensitivity Analysis

To assess the robustness of the findings, we conducted several sensitivity analyses, including Maximum Likelihood, MR-RAPS, and MR-Egger Intercept. Our examination did not detect any heterogeneity in the IVs pertaining to blood metabolites that exhibited significant associations with colorectal cancer, as indicated by Cochran's IVW Q test (\( P > 0.05 \)). Furthermore, the MR-Egger regression intercept analysis did not
provide substantial evidence of directional pleiotropy. It is noteworthy that the direction of effect remained consistent across all three methods, aligning with the IVW method. Furthermore, the IVW radial MR Results demonstrated that the corrected findings were in agreement with the pre-corrected results (Table S3-S12; Table 2). The robustness of causality is substantiated by the Maximum likelihood and MR (Mendelian randomization) Estimates produced by MR-Egger, which consistently demonstrate the same direction and magnitude. The p-values derived from Cochran Q test indicate the absence of heterogeneity. Moreover, the MR-Egger intercept suggests the absence of horizontal pleiotropy (Table 2). Likewise, the MR-RAPS analysis also fails to reveal any significant level of pleiotropy.

DISCUSSION

This study represents the first investigation utilizing MR analysis to explore the causal link between human blood metabolites and colorectal cancer. The findings demonstrate that specific genes, namely Mannose, Arachidonate (20:4 n6), X - 11593 - O - methylascorbate *, 1 - arachidonoylglycerophosphocholine *, and 1 - arachidonoylglycerophosphoethanolamine *, are associated with an increased risk of colorectal cancer, suggesting their potential role in promoting its development. Conversely, genetically determined levels of Tyrosine, Urate, N-acetylglycine, X-12092, and Succinylcarnitine exhibit an inverse relationship with colorectal cancer risk, indicating a potential protective effect. Furthermore, these results were robust and verified through multiple analytical approaches. These findings are consistent with previous studies in the field of tumor metabolism. For example, lipid metabolites such as arachidonate, X-11593-O-methylascorbate, and 1-arachidonoylglycerophosphocholine have been proposed to play a crucial role in tumor development[19]. These metabolites may regulate inflammatory pathways and enhance the body's immune response to tumor progression[13]. This was demonstrated in a metabonomics study involving a cancer survival cohort of 1812 Finnish men, which identified 49 metabolites associated with cancer survival. These metabolites include...
phosphatidylcholine, glutamate, arachidonic acid (20:4n6), and glutamylamino acids such as gamma-glutamylglutamic acid, gamma-glutamylglycine, and gamma-glutamylleucine. Higher levels of these lipid molecules were found to be associated with increased cancer-specific mortality\textsuperscript{[20, 21, 22]}. The dysregulation of inflammatory pathways in cancer may be attributed to elevated levels of specific lipid molecules. Moreover, the reprogramming of lipid metabolism has been identified as a novel marker for malignant tumors\textsuperscript{[23]}. The expression profile of lipid metabolism-related genes (LMRG) in colorectal cancer has also been linked to a high tumor mutation burden \textsuperscript{[24]}. Furthermore, mannose has been identified as a marker for the malignant progression of early colorectal cancer. Alterations in sugar molecules have been implicated in the occurrence and progression of various types of cancer, including colorectal cancer \textsuperscript{[25]}.

The available physiological evidence on the relationship between X-11593--O-methylascorbate* and colorectal cancer in humans is currently inadequate. Nonetheless, an animal experiment revealed that two new synthetic derivatives of ascorbic acid, namely 3-O-ethyl ascorbic acid (EAsA) and 3-O-dodecylmethylascorbate, were observed to enhance cancer progression \textsuperscript{[26]}. Urate, which acts as an antioxidant, has the potential to reduce the risk of cancer by reducing oxidative stress. This finding is consistent with previous studies exploring the role of antioxidants in preventing cancer. These findings provide new insights into the metabolic mechanisms involved in colon cancer and may have implications for future therapeutic strategies \textsuperscript{[27, 28, 29]}.

Multiple cancer-related cellular pathways have been identified, with protein phosphorylation and dephosphorylation, particularly on tyrosine residues, being prominent regulatory mechanisms. The intricate equilibrium between these processes is tightly controlled by protein tyrosine kinase (PTK) and protein tyrosine phosphatase (PTP) \textsuperscript{[30, 31]}. An abnormal activity of oncogenic PTK has been observed in a significant proportion of human cancers. PTPs, on the other hand, have long been considered as tumor suppressors due to their ability to counterbalance the effects of phosphorylation-
based signaling activation. Activation of PTP leads to elevated tyrosine expression, indicating its inhibitory impact on cancer development \cite{30, 31}. The contribution of Succinylcarnitine to the metabolic dysregulation of cancer metastasis is evident, as elevated levels of Succinylcarnitine have been reported in metastatic rectal cancer cell lines. Additionally, X-12092, a specific blood metabolite, has received comparatively less attention, despite being highly expressed in intracranial aneurysms \cite{32}. Our study provides valuable insights into the metabolic mechanisms underlying colon cancer, which have important implications for future prevention and treatment strategies. To fully understand the relationship between these metabolites and the risk of colon cancer, more detailed investigations are needed. It is crucial to explore their role in the initiation and progression of colon cancer to gain a comprehensive understanding. Additionally, future research should focus on investigating the interplay between metabolites and their correlation with other biomarkers, in order to establish a more comprehensive metabolic profile for colon cancer. By pursuing these research endeavors, significant progress can be made towards developing more targeted cancer prevention and personalized treatment approaches.

This study has several notable strengths. Firstly, unlike previous Mendelian randomization investigations that focused on single or conventional exposure factors, this groundbreaking study integrated metabolomics with genomics to analyze the causal association between 486 human blood metabolites and colorectal cancer. Secondly, multiple techniques were employed to ensure the validity of the study findings.

However, there are several limitations that should be acknowledged in this study. Firstly, the study relied on summary statistics rather than individual data, which precluded subgroup analysis. In future MR studies, the use of individual-level data is crucial to obtain a more comprehensive perspective. Secondly, it is important to note that MR estimates are not adjusted for multiple testing. Nonetheless, we addressed this concern by conducting repeated analyses using an independent GWAS dataset, which significantly enhanced the validity of the findings. Lastly, it is crucial to consider that all
participants in this study were of European ancestry. Therefore, caution should be exercised when generalizing the conclusions to individuals from other ethnic backgrounds. Future studies should aim to incorporate diverse populations to ensure the generalizability of the findings while recognizing the importance of individualized treatment approaches.

In conclusion, this study underscores the importance of genetic factors in establishing a causal relationship between metabolites and colorectal cancer. This finding provides a new perspective on the etiology of colorectal cancer and offers potential preventive strategies through the integration of metabolomics and genomics. Nevertheless, due to the limited number of studies directly linking these metabolites to colorectal cancer, additional experimental and clinical investigations are necessary to validate these findings and elucidate the underlying mechanisms.

CONCLUSION

Our study presents the first systematic assessment of the causal relationship between plasma metabolites and colorectal cancer. To achieve this, we utilized SNPs identified through GWAS as IVs in a Mendelian randomization study design. By employing this approach, we have successfully identified a number of significant blood metabolites that are associated with colorectal cancer. These findings lay the foundation for a more comprehensive understanding of the etiology of colorectal cancer, providing insights into the interplay between colorectal cancer and plasma metabolites in the disease's development.

ARTICLE HIGHLIGHTS

Research background

Limited knowledge exists regarding the casual associations linking blood metabolites and the risk of developing colorectal cancer.

Research motivation
Colorectal cancer development is associated with the presence of five specific blood metabolites. These metabolites have been identified as causal agents and have been validated as risk factors for the disease.

**Research objectives**

To investigate causal associations between blood metabolites and colon cancer.

**Research methods**

The study utilized a two-sample Mendelian randomization (MR) analysis to investigate the causal impact of 486 blood metabolites on colorectal cancer. The primary method of analysis used was the inverse variance weighted (IVW) model. To further validate the results several sensitivity analyses were performed, including Cochran’s Q test, MR-Egger intercept test, and Mendelian randomization robust adjusted profile score (MR-RAPS). These additional analyses were conducted to ensure the reliability and robustness of the findings.

**Research results**

After rigorous selection for genetic variation, 486 blood metabolites were included in the MR analysis. We found Mannose (odds ratio (OR) = 2.09 (1.10-3.97), \( P = 0.024 \)), N-acetylglycine [OR = 3.14 (1.78-5.53), \( P = 7.54 \times 10^{-8} \)], X-11593-O-methylascorbate* [OR = 1.68 (1.04-2.72), \( P = 0.034 \)], 1-arachidonoylglycerophosphocholine* [OR = 4.23 (2.51-7.12), \( P = 6.35 \times 10^{-8} \)] and 1- arachidonoylglycerophosphoethanolamine*4 [OR = 3.99 (1.17-13.54), \( P = 0.027 \)] were positively causally associated with colorectal cancer, and we also found a negative causal relationship between Tyrosine [OR = 0.08 (0.01-0.63), \( P = 0.014 \)], Urate [OR =0.25 (0.10-0.62), \( P = 0.003 \)], N-acetylglycine [0.73 (0.54-0.98), \( P = 0.033 \)], X-12092[OR =0.89 (0.81-0.99), \( P = 0.028 \)], Succinylcarnitine [OR =0.48 (0.27-0.84), \( P = 0.09 \)] with colorectal cancer. A series of sensitivity analyses were performed to confirm the rigidity of the results.
**Research conclusions**

This study showed a causal relationship between 10 blood metabolites and colorectal cancer, of which 5 blood metabolites were found to be causal for the development of colorectal cancer and were confirmed as risk factors. The other five blood metabolites are protective factors.

**Research perspectives**

A significant inverse relationship has been observed between the remaining five blood metabolites and the development of colorectal cancer, establishing them as protective.

Crossref


Crossref