

Causal associations between blood metabolites and breast cancer

Keywords

breast cancer, Metabolites, Mendelian randomization

Abstract

Introduction

The associations between blood metabolites and breast cancer remain unclear. We conducted a systematic two-sample Mendelian randomization (MR) analysis to identify key human blood metabolites and uncover potential biomarkers for breast cancer development.

Material and methods

The data were extracted from large-scale genome-wide association study (GWAS) public databases. Instrumental variables were selected from a cohort study of 453 metabolic profiles from 7,824 participants. Breast cancer incidence data were obtained from a large cohort study involving 138,389 cases and 240,341 controls. Causal associations between human blood metabolites and breast cancer incidence were assessed using inverse-variance weighting, and MR-Egger regression.

Results

Five human blood metabolites were identified as biomarkers for breast cancer: serine (OR, 2.25; 95% CI: 1.18–4.27), 10-undecenoate (11:1n1) (OR, 1.38; 95% CI: 1.00–1.90), X-12696 (OR, 2.15; 95% CI: 1.14–4.08), X-14626 (OR, 1.68; 95% CI: 1.15–2.46), and succinyl carnitine (OR, 1.58; 95% CI: 1.06–2.34). The sensitivity analysis results indicate no pleiotropy between the metabolites and breast cancer risk, confirming the robustness of the findings.

Conclusions

This study in metabolomics research identified five human blood metabolites — serine, 10-undecenoate (11:1n1), X-12696, X-14626, and succinylcarnitine — as potential biomarkers for assessing breast cancer risk. Among these metabolites, serine and X-12696 showed the strongest associations with the likelihood of developing breast cancer.

Causal associations between blood metabolites and breast cancer

Abstract

Introduction: The associations between blood metabolites and breast cancer remain unclear. We conducted a systematic two-sample Mendelian randomization (MR) analysis to identify key human blood metabolites and uncover potential biomarkers for breast cancer development.

Material and methods: The data were extracted from large-scale genome-wide association study (GWAS) public databases. Instrumental variables were selected from a cohort study of 453 metabolic profiles from 7,824 participants. Breast cancer incidence data were obtained from a large cohort study involving 138,389 cases and 240,341 controls. Causal associations between human blood metabolites and breast cancer incidence were assessed using inverse-variance weighting, and MR-Egger regression.

Results: Five human blood metabolites were identified as biomarkers for breast cancer: serine (OR, 2.25; 95% CI: 1.18–4.27), 10-undecenoate (11:1n1) (OR, 1.38; 95% CI: 1.00–1.90), X-12696 (OR, 2.15; 95% CI: 1.14–4.08), X-14626 (OR, 1.68; 95% CI: 1.15–2.46), and succinyl carnitine (OR, 1.58; 95% CI: 1.06–2.34). The sensitivity analysis results indicate no pleiotropy between the metabolites and breast cancer risk, confirming the robustness of the findings.

Conclusion: This study in metabolomics research identified five human blood metabolites—serine, 10-undecenoate (11:1n1), X-12696, X-14626, and succinylcarnitine—as potential biomarkers for assessing breast cancer risk. Among

these metabolites, serine and X-12696 showed the strongest associations with the likelihood of developing breast cancer.

Keywords: Metabolites; breast cancer; Mendelian randomization

Preprint

Introduction

Breast cancer, as the most prevalent form of invasive malignancy, stands as the primary cause of cancer-related deaths among women due to its elevated incidence and mortality rates [1]. In 2020, breast cancer led to nearly 685,000 female fatalities globally and represented 30% of the anticipated cancer incidence in women for 2021 [2], underscoring the significant prevalence and mortality rates. Given the limited accessibility of breast cancer treatments and the costly nature of clinical trials [3], there is a critical need to investigate potential biomarkers linked to the development of breast cancer.

Breast cancer is influenced by a variety of both internal and external risk factors. Numerous epidemiological studies have identified mediators of breast cancer, with some Mendelian randomization (MR) studies confirming biomarkers associated with the disease. For instance, insulin-like growth factor-1 levels have been linked to a moderate increase in breast cancer risk [4]. Mitochondrial dysfunction, driven by genetic factors, has also been shown to play a causal role in breast cancer, with certain mitochondria-related genes implicated in disease development [5]. Additionally, serum C-reactive protein (CRP) has emerged as a potential biomarker for assessing overall cancer risk and risks specific to certain sites [6]. Despite these findings, research on the connection between the metabolome and breast cancer risk remains limited. Metabolomics, which focuses on the study of small molecules related to metabolic processes, can offer valuable insights when integrated with other histological platforms [7]. Understanding the causal relationships between metabolites

and breast cancer development is crucial, as it may provide genetic evidence supporting the impact of key blood metabolites on breast cancer risk.

Genome-wide association studies (GWAS) are instrumental in identifying correlations across the genome between traits and single-nucleotide polymorphisms (SNPs), shedding light on the significance and impact of various genetic variants on different traits [8]. Recent GWAS have successfully uncovered causal links between the human metabolome and diseases [9]. On the other hand, MR analysis, a more robust method for inferring causality that has emerged in recent years, leverages genetic variation as instrumental variables (IV) to evaluate the causal relationship between risk factors and disease outcomes, thereby mitigating reverse causality bias [10]. Conducting a two-sample MR analysis necessitates data from distinct sources, such as two independent GWAS, to ascertain exposure and outcomes [11]. In this study, we utilized two GWAS databases operating at different levels to perform a large-scale, two-sample MR analysis aimed at systematically examining 100 human blood metabolites and pinpointing potential causal associations with breast cancer incidence.

Materials and Methods

Research design

This study utilized the publicly available GWAS database for a two-sample MR analysis to investigate the causal relationship between human blood metabolites and breast cancer. Ethical approval for data collection and written informed consent from participants were obtained in the original GWAS. SNPs were used as instrumental variables in this study, ensuring that they met the basic assumptions required for MR

analysis.

Assumption 1: Genetic variants must be strongly associated with human blood metabolites;

Assumption 2: The genetic variants may be associated with the development of breast cancer specifically through human blood metabolites.

Assumption 3: Genetic factors must not be associated with any confounders of human blood metabolites and breast cancer.

The study design process is illustrated in Figure 1. Approval for this study was obtained from the Affiliated Cancer Hospital of Harbin Medical University. Written informed consent was obtained from all participants or their legal representatives during the recruitment process.

Data sources

The study utilized human blood metabolite exposure and breast cancer genomic data from the Integrated Epidemiology Research Center's Open Genome-Wide Association Studies (IEU OpenGWAS project) database, which includes two extensive GWAS cohorts totaling 265,554 individuals of European ancestry (detailed in Table 1). The human blood metabolite data was sourced from Shin et al.'s study, analyzing 453 metabolic profiles of 7,824 participants with approximately 3 million SNPs. The outcome data came from a cohort study by Sakaue et al. [12], involving 257,730 participants with 138,389 cases and 240,341 controls[13]. In this MR study, SNPs associated with 453 metabolites from the exposure cohort were examined to reflect blood metabolite expression at the gene level.

Data processing

Exposure data screening

A total of 104 SNPs associated with the exposure cohort were extracted from the GWAS database based on the screening criteria of $P < 5 \times 10^{-8}$ [14]. To ensure independence among individual human blood metabolites, standard parameters for linkage disequilibrium removal were applied: linkage disequilibrium coefficient $R^2 < 0.1$, with a window size of 10,000 kb. The strength of the selected SNPs was evaluated using the F-statistic with a window size of 10,000 kb, and SNPs with $F > 10$ was excluded.

Processing of outcome data

After merging the exposure data with the outcome data, the processed metabolites were aligned with the GWAS data on breast cancer incidence to pinpoint the instrumental variables linked to the outcome. Following this, the data sets were harmonized based on the statistical parameters of human blood metabolites and the GWAS data on breast cancer sharing the same loci, ensuring that the effect values of human blood metabolites and breast cancer were aligned to the same effect allele. Metabolites with less than three relevant SNPs in the genome were excluded, as at least three SNPs are required to be associated with exposure in certain MR sensitivity analyses [15]. Ultimately, 100 significant human blood metabolites were included in this study for further analysis.

Statistical analysis

The study primarily utilized the inverse variance weighting (IVW) method as

the primary MR method to investigate the causal relationship between blood metabolite concentrations and breast cancer risk [16]. Cochran's Q test, with a P-value greater than 0.05 indicating homogeneity in the results [17]. In cases of non-heterogeneous results, a fixed-effects model was employed, while a random-effects model was used for heterogeneous results to evaluate the MR effect. Sensitivity analyses were conducted to validate the reliability of IVW, including the maximum likelihood method, MR-robust adjusted profile scoring (MR-RAPS), and the MR-Egger method to detect horizontal pleiotropy [18]. Causal estimates between metabolites and breast cancer risk were presented as odds ratios (ORs) with corresponding 95% confidence intervals (CIs).

Data extraction, processing, and analysis were conducted using the TwoSampleMR (version 0.5.6) software package in R (version 4.1.3). A statistically significant difference was defined as $P < 0.05$.

Results

Results of instrumental variable screening

In this study, instrumental variables were carefully selected based on screening principles and criteria to identify 100 human blood metabolites associated with breast cancer. Cochran's Q test was used to assess heterogeneity, with the random-effects model applied in the presence of heterogeneity and the fixed-effects model used when no heterogeneity was detected. The causal relationship between human blood metabolites and breast cancer risk was then examined using the IVW method. The analysis revealed that out of the 100 metabolites, 11 were found to have a

causal link with breast cancer: proline, serine, 10-undecenoate (11:1n1), X-11440, bilirubin (E,Z or Z,E)*, X-12696, X-13431, X-13431-nonanoylcarnitine*, dihomolinolenate (20:3n3 or n6), X-14626, succinylcarnitine, and 4-androsten-3beta,17beta-diol disulfate 1*. Detailed results can be seen in Exhibit 3.

Associations of important human blood metabolites with breast cancer risk

The IVW method was utilized in the primary Mendelian randomization analysis to assess the causal relationships between 11 key human blood metabolites and the risk of breast cancer. The outcomes are comprehensively illustrated in Figure 2. The forest plot highlights five human blood metabolites that were identified as risk factors for breast cancer, specifically serine (OR, 2.25; 95% CI: 1.18–4.27), 10-undecenoate (11:1n1) (OR, 1.38; 95% CI: 1.00–1.90), X-12696 (OR, 2.15; 95% CI: 1.14–4.08), X-14626 (OR, 1.68; 95% CI: 1.15–2.46), and succinyl carnitine (OR, 1.58; 95% CI: 1.06–2.34). Notably, serine and X-12696 exhibited the most robust associations with breast cancer risk, boasting ORs of 2.25 and 2.15, respectively.

Associations of the identified important human blood metabolites with breast cancer risk

The results of the analysis in Figure 3 indicate that alleles of five metabolites, including proline, X-11440, bilirubin (E,Z or Z,E)*, X-13431-nonanoylcarnitine*, and 4-androsten-3beta,17beta-diol disulfate 1*, are negatively associated with breast cancer risk, suggesting a decrease in risk with higher allele counts. Conversely, alleles of five blood metabolites, such as serine, 10-undecenoate (11:1n1), X-12696, X-14626, and succinylcarnitine, are positively linked

to breast cancer risk, indicating an increase in risk with higher allele counts. These findings align with the primary MR analysis, which identifies serine, 10-undecenoate (11:1n1), X-12696, X-14626, and succinyl carnitine as potential biomarkers for breast cancer development.

The 11 important human metabolites identified were found to have 55 SNPs, with detailed characterization of each SNP variant provided in Exhibits 4 through 14.

Sensitivity analysis

In sensitivity analyses using both the maximum likelihood and MR-RAPS methods, genetically determined serine, 10-undecenoate (11:1n1), X-12696, X-14626, and succinyl carnitine were found to be significantly associated with an increased risk of breast cancer development. MR-Egger regression intercept results indicated no evidence of directed pleiotropy among these five human blood metabolites and breast cancer risk. Therefore, these metabolites were identified as potential biomarkers for assessing breast cancer risk (Table 1). Specifically, genetically determined serine (OR, 2.25; 95% CI: 1.18–4.27), 10-undecenoate (11:1n1) (OR, 1.38; 95% CI: 1.00–1.90), X-12696 (OR, 2.15; 95% CI: 1.14–4.08), X-14626 (OR, 1.68; 95% CI: 1.15–2.46), and succinyl carnitine (OR, 1.58; 95% CI: 1.06–2.34) showed an increased risk of breast cancer per 1 standard deviation increase.

Discussion

The combined metabolomics and genomics approach in this MR study offers novel insights into the risk of breast cancer and potential drug targets. Out of 100 human blood metabolites examined, five metabolites were found to have potential

causal links with breast cancer: serine, 10-undecenoate (11:1n1), X-12696, X-14626, and succinylcarnitine. This indicates that genetically predicted higher levels of these metabolites may be linked to a heightened risk of breast cancer.

Serine is a crucial precursor for the synthesis of various essential biomolecules such as proteins, lipids, nucleotides, and other amino acids, playing a central role in biosynthetic reactions necessary for cell division and growth [19]. The involvement of serine in cancer progression has garnered significant attention in the academic community. Research indicates that many cancer cells rely heavily on serine as a primary source of 1C units [20]. Previous studies have extensively validated the impact of serine on cancer development. For instance, oncogenes have been found to target enzymes in the serine biosynthetic pathway (SBP) [21], with the expression of these enzymes linked to inflammation in breast cancer. Additionally, heightened serine synthesis has been observed in breast cancer tissues [22]. Mechanisms through which increased serine synthesis accelerates carcinogenesis include altering glucose carbon flux, maintaining specific NAD(P)/NAD(P)H ratios, and regulating metabolite synthesis or expression [23, 24]. Building on these insights, our study further supports the role of serine as a mediator in breast cancer development, suggesting that targeting mitochondrial serine synthesis could be a promising strategy to impede breast carcinogenesis.

10-Undecenoate (11:1n1), a metabolite associated with gut microbiota, has been linked to various diseases. However, there is a lack of experimental evidence regarding its relationship with breast cancer risk or the impact of breast cancer on

blood 10-undecenoate (11:1n1) levels. Further research is required to assess the role of blood 10-undecenoate (11:1n1) concentrations in diagnosing and treating breast cancer. Studies have suggested a potential causal link between 10-undecenoate (11:1n1) and Crohn's disease, depression [25, 26], as well as low concentrations in patients with non-alcoholic fatty liver disease [27]. This neutral hydrophobic molecule remains poorly understood in the literature, but lifestyle factors such as diet and habits may influence breast cancer risk [28]. Overall, these findings offer insight into exploring the interplay of intestinal flora, metabolism, and breast cancer treatment, as well as shedding light on the connections between depression, Crohn's disease, and breast cancer development.

X-12696 and X-14626 are newly discovered blood metabolites that have not been previously documented in scientific literature. Interestingly, X-12696 has shown a strong association with breast cancer risk, ranking second only to serine. This highlights the importance of further research to investigate the role of X-12696 in the human body.

Succinyl carnitine, an acylcarnitine involved in fatty acid metabolism and mitochondrial function [29, 30], has been linked to various health conditions. Studies have shown elevated levels of succinyl carnitine in blood associated with Alzheimer's disease and maternal concentrations during pregnancy possibly contributing to coronary heart disease in offspring [31, 32]. Additionally, succinyl carnitine has been connected to total cholesterol, low-density lipoproteins, and breast cancer risk, highlighting its potential as a significant factor in disease development. Furthermore,

as a newly identified urinary biomarker for γ -hydroxybutyric acid, a substance linked to brain metabolism and recreational drug use, succinyl carnitine has also been approved for treating narcolepsy [33, 34]. These findings underscore the need for further research on the relationship between exogenous substances, such as recreational drugs and sleep disorder medications, and breast cancer risk. The study suggests potential associations between breast cancer and conditions like coronary heart disease and Alzheimer's disease, warranting further investigation for a better understanding of disease pathogenesis.

Our study has several strengths. Firstly, it is among the limited number of systematic MR studies that utilize blood metabolites as exposures to evaluate their causal impact on breast cancer risk. Secondly, this MR study utilized data from two extensive GWAS, enabling us to draw valid causal conclusions with robust statistical power. Thirdly, the study adhered to rigorous quality control measures and included a variety of sensitivity analyses and validity assessments, ensuring the stability and reliability of the results.

This study also has limitations. Firstly, the GWAS data used were solely from white European populations, thus limiting the generalizability of our findings to other racial and ethnic groups. Further research is necessary to confirm if our results are applicable to other populations. Secondly, the lack of detailed demographic information, such as age and gender, in the extracted data prevented subgroup analyses from being conducted. In addition, due to constraints in time and funding, experimental validation was not included.

Conclusion

This systematic meta-analysis identified serine, 10-undecenoate (11:1n1), X-12696, X-14626, and succinyl carnitine as potential biomarkers for predicting the risk of developing breast cancer. Specifically, serine and the previously unidentified blood metabolite X-12696 demonstrated the most significant associations with breast cancer prognosis.

Preprint

Conflict of interest: Authors declare that they have no conflict of interest for this article.

Author Contributions Statement: Li GY contributed to the conception and design of this study and drafted the manuscript; Du C made substantial contributions to the acquisition, analysis and interpretation of data, and obtaining funding; Miao DZ contributed to the revision and supervision of the manuscript. All authors and contributors have agreed to conditions noted on the Authorship Agreement form.

Acknowledgements: We thank all authors for their contributions to the article.

Funding: The research was sponsored by Medical clinical youth scientific research project (number: 2019-KYYWF-0358)

Ethics approval: This study was approved by Affiliated Cancer Hospital of Harbin Medical University.

Informed consent: All of the participants or their legal representatives provided written informed consent during recruitment.

Disclosure statement: None.

Data availability: The datasets generated and analyzed during the current study are available from the corresponding author on reasonable request.

References

1. 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: a cancer journal for clinicians*. 2021; 71:209-49.
2. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer Statistics, 2021. *CA: a cancer journal for clinicians*. 2021; 71:7-33.
3. Key TJ, Appleby PN, Reeves GK, Roddam AW. Insulin-like growth factor 1 (IGF1), IGF binding protein 3 (IGFBP3), and breast cancer risk: pooled individual data analysis of 17 prospective studies. *The Lancet Oncology*. 2010; 11:530-42.
4. Li Y, Sundquist K, Zhang N, Wang X, Sundquist J, Memon AA. Mitochondrial related genome-wide Mendelian randomization identifies putatively causal genes for multiple cancer types. *EBioMedicine*. 2023; 88:104432.
5. 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 medicine*. 2022; 20:301.
6. Ussher JR, Elmariah S, Gerszten RE, Dyck JR. The Emerging Role of Metabolomics in the Diagnosis and Prognosis of Cardiovascular Disease. *Journal of the American College of Cardiology*. 2016; 68:2850-70.
7. Swerdlow DI, Kuchenbaecker KB, Shah S, Sofat R, Holmes MV, White J, et al. Selecting instruments for Mendelian randomization in the wake of genome-wide association studies. *International journal of epidemiology*. 2016; 45:1600-16.

8. Yun Z, Guo Z, Li X, Shen Y, Nan M, Dong Q, et al. Genetically predicted 486 blood metabolites in relation to risk of colorectal cancer: A Mendelian randomization study. *Cancer medicine*. 2023; 12:13784-99.
9. Emdin CA, Khera AV, Kathiresan S. Mendelian Randomization. *Jama*. 2017; 318:1925-6.
10. Hartwig FP, Davies NM, Hemani G, Davey Smith G. Two-sample Mendelian randomization: avoiding the downsides of a powerful, widely applicable but potentially fallible technique. *International journal of epidemiology*. 2016; 45:1717-26.
11. Shin SY, Fauman EB, Petersen AK, Krumsiek J, Santos R, Huang J, et al. An atlas of genetic influences on human blood metabolites. *Nature genetics*. 2014; 46:543-50.
12. Dehghan A. Genome-Wide Association Studies. *Methods in molecular biology (Clifton, NJ)*. 2018; 1793:37-49.
13. Sakaue S, Kanai M, Tanigawa Y, Karjalainen J, Kurki M, Koshihara S, et al. A cross-population atlas of genetic associations for 220 human phenotypes. *Nature genetics*. 2021; 53:1415-24.
14. Hemani G, Zheng J, Elsworth B, Wade KH, Haberland V, Baird D, et al. The MR-Base platform supports systematic causal inference across the human phenome. *eLife*. 2018; 7.
15. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. *Genetic epidemiology*. 2013; 37:658-65.
16. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent Estimation in

Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genetic epidemiology*. 2016; 40:304-14.

17. Burgess S, Thompson SG. Interpreting findings from Mendelian randomization using the MR-Egger method. *European journal of epidemiology*. 2017; 32:377-89.

18. Ye J, Fan J, Venneti S, Wan YW, Pawel BR, Zhang J, et al. Serine catabolism regulates mitochondrial redox control during hypoxia. *Cancer discovery*. 2014; 4:1406-17.

19. Labuschagne CF, van den Broek NJ, Mackay GM, Vousden KH, Maddocks OD. Serine, but not glycine, supports one-carbon metabolism and proliferation of cancer cells. *Cell reports*. 2014; 7:1248-58.

20. Wilcz-Villega E, Carter E, Ironside A, Xu R, Mataloni I, Holdsworth J, et al. Macrophages induce malignant traits in mammary epithelium via IKK ϵ /TBK1 kinases and the serine biosynthesis pathway. *EMBO molecular medicine*. 2020; 12:e10491.

21. Kim HY, Lee KM, Kim SH, Kwon YJ, Chun YJ, Choi HK. Comparative metabolic and lipidomic profiling of human breast cancer cells with different metastatic potentials. *Oncotarget*. 2016; 7:67111-28.

22. Yang M, Vousden KH. Serine and one-carbon metabolism in cancer. *Nature reviews Cancer*. 2016; 16:650-62.

23. Fan J, Teng X, Liu L, Mattaini KR, Looper RE, Vander Heiden MG, et al. Human phosphoglycerate dehydrogenase produces the oncometabolite D-2-hydroxyglutarate. *ACS chemical biology*. 2015; 10:510-6.

24. van der Spek A, Stewart ID, Kühnel B, Pietzner M, Alshehri T, Gauß F, et al.

Circulating metabolites modulated by diet are associated with depression. *Molecular psychiatry*. 2023; 28:3874-87.

25. Puri P, Baillie RA, Wiest MM, Mirshahi F, Choudhury J, Cheung O, et al. A lipidomic analysis of nonalcoholic fatty liver disease. *Hepatology (Baltimore, Md)*. 2007; 46:1081-90.

26. De Cicco P, Catani MV, Gasperi V, Sibilano M, Quaglietta M, Savini I. Nutrition and Breast Cancer: A Literature Review on Prevention, Treatment and Recurrence. *Nutrients*. 2019; 11.

27. Mai M, Tönjes A, Kovacs P, Stumvoll M, Fiedler GM, Leichtle AB. Serum levels of acylcarnitines are altered in prediabetic conditions. *PloS one*. 2013; 8:e82459.

28. Panyard DJ, McKetney J, Deming YK, Morrow AR, Ennis GE, Jonaitis EM, et al. Large-scale proteome and metabolome analysis of CSF implicates altered glucose and carbon metabolism and succinylcarnitine in Alzheimer's disease. *Alzheimer's & dementia : the journal of the Alzheimer's Association*. 2023; 19:5447-70.

29. Taylor K, McBride N, Zhao J, Oddie S, Azad R, Wright J, et al. The Relationship of Maternal Gestational Mass Spectrometry-Derived Metabolites with Offspring Congenital Heart Disease: Results from Multivariable and Mendelian Randomization Analyses. *Journal of cardiovascular development and disease*. 2022; 9.

30. Starek-Świechowicz B, Budziszewska B, Starek A. Endogenous estrogens-breast cancer and chemoprevention. *Pharmacological reports : PR*. 2021; 73:1497-512.

31. Tao T, He T, Mao H, Wu X, Liu X. Non-Targeted Metabolomic Profiling of Coronary Heart Disease Patients With Taohong Siwu Decoction Treatment. *Frontiers*

in pharmacology. 2020; 11:651.

32. Kim S, Lee MS, Kim M, Ko BJ, Lee HS, Lee S. Derivatization-assisted LC-MS/MS method for simultaneous quantification of endogenous gamma-hydroxybutyric acid and its metabolic precursors and products in human urine. *Analytica chimica acta*. 2022; 1194:339401.

33. Steuer AE, Raeber J, Steuer C, Boxler MI, Dornbierer DA, Bosch OG, et al. Identification of new urinary gamma-hydroxybutyric acid markers applying untargeted metabolomics analysis following placebo-controlled administration to humans. *Drug testing and analysis*. 2019; 11:813-23.

Preprint

Figure Legends

Figure 1 Diagram of the mendelian randomization (MR) study design

This MR study aims to investigate the causal associations between human blood metabolites (exposure) and breast cancer (outcome). The assumption is that the instrumental variables are associated with metabolites, but not with confounders, and the instrumental variables are not associated with risk of breast cancer on metabolites and confounders. Abbreviation: SNP, single-nucleotide polymorphism; IVW, inverse-variance weighted; MR-RAPS, Mendelian randomization robust adjusted profile score

Figure 2 The associations of metabolites with the risk of breast cancer using the inverse-variance weighted mendelian randomization analysis

Abbreviations: OR, odds ratio; 95% CI, 95% confidence interval

Figure 3 Associations of genetic variants about identified metabolites with the risk of breast cancer

The line indicates the estimate of causal effect using inverse-variance weighted method. Circles indicate associations of each genetic variant related to metabolites with the risk of breast cancer. Error bars genetic indicate 95% confidence interval.

A, Proline; B, Serine; C, 10-undecenoate (11:1n1); D, X-11440; E, Bilirubin (E,Z or Z,E)*; F, X-12696; G, X-13431--nonanoylcarnitine*; H, Dihomo-linolenate (20:3n3 or n6);I, X-14626; J, Succinylcarnitine; K, 4-androsten-3beta,17beta-diol disulfate 1*

Abbreviations: SNP: single-nucleotide polymorphism

Table 1 Associations between metabolites and the risk of breast cancer in sensitivity analysis

Metabolites	Maximum likelihood		MR-RAPS		MR-Egger		MR-Egger Intercept	
	OR (95% CI)	<i>P</i> value	OR (95% CI)	<i>P</i> value	OR (95% CI)	<i>P</i> value	OR (95% CI)	<i>P</i> value
Proline	0.70 (0.50-0.97)	0.034	0.72 (0.52-1.01)	0.054	0.47 (0.23-0.99)	0.046	1.02 (0.99-1.04)	0.166
Serine	2.25 (1.17-4.34)	0.015	2.25 (1.16-4.38)	0.016	2.47 (0.00-3265.03)	0.805	1.00 (0.87-1.15)	0.979
10-undecenoate (11:1n1)	1.38 (1.00-1.90)	0.048	1.38 (1.00-1.91)	0.050	1.25 (0.62-2.54)	0.533	1.00 (0.98-1.03)	0.764
X-11440	0.82 (0.71-0.95)	0.008	0.85 (0.74-0.98)	0.027	0.78 (0.56-1.08)	0.139	1.01 (0.99-1.03)	0.503
Bilirubin (E,Z or Z,E)*	0.79 (0.63-1.00)	0.047	0.79 (0.63-1.00)	0.049	0.76 (0.42-1.37)	0.357	1.00 (0.97-1.03)	0.860
X-12696	2.16 (1.13-4.13)	0.020	2.16 (1.12-4.19)	0.022	2.05 (0.25-17.13)	0.506	1.00 (0.95-1.05)	0.963
X-13431--nonanoylcarnitine*	0.83 (0.71-0.96)	0.013	0.83 (0.71-0.96)	0.014	1.04 (0.68-1.61)	0.849	0.98 (0.96-1.01)	0.259
Dihomo-linolenate (20:3n3 or n6)	0.38 (0.20-0.74)	0.004	0.38 (0.20-0.74)	0.005	1.10 (0.08-15.21)	0.943	0.98 (0.93-1.03)	0.408
X-14626	1.69 (1.15-2.47)	0.007	1.69 (1.15-2.48)	0.008	1.81 (0.74-4.39)	0.192	1.00 (0.97-1.03)	0.864
Succinylcarnitine	1.59 (1.07-2.37)	0.022	1.47 (1.01-2.15)	0.046	1.69 (0.73-3.93)	0.219	1.00 (0.98-1.01)	0.708
4-androsten-3beta,17beta-diol disulfate 1*	0.86 (0.75-0.98)	0.024	0.86 (0.75-0.98)	0.024	0.76 (0.52-1.10)	0.141	1.01 (0.98-1.04)	0.422

Abbreviations: MR-RAPS, mendelian randomization robust adjusted profile score. OR, Odds ratios; CI, confidence intervals.

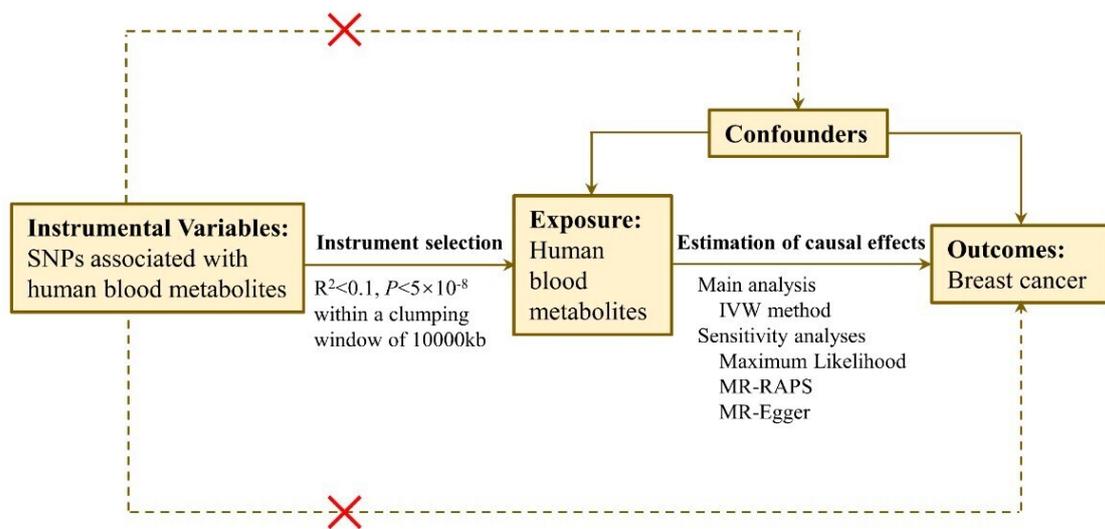


Figure 1 Diagram of the mendelian randomization (MR) study design

This MR study aims to investigate the causal associations between human blood metabolites (exposure) and breast cancer (outcome). The assumption is that the instrumental variables are associated with metabolites, but not with confounders, and the instrumental variables are not associated with risk of breast cancer on metabolites and confounders

Abbreviation: SNP, single-nucleotide polymorphism; IVW, inverse-variance weighted; MR-RAPS, Mendelian randomization robust adjusted profile score

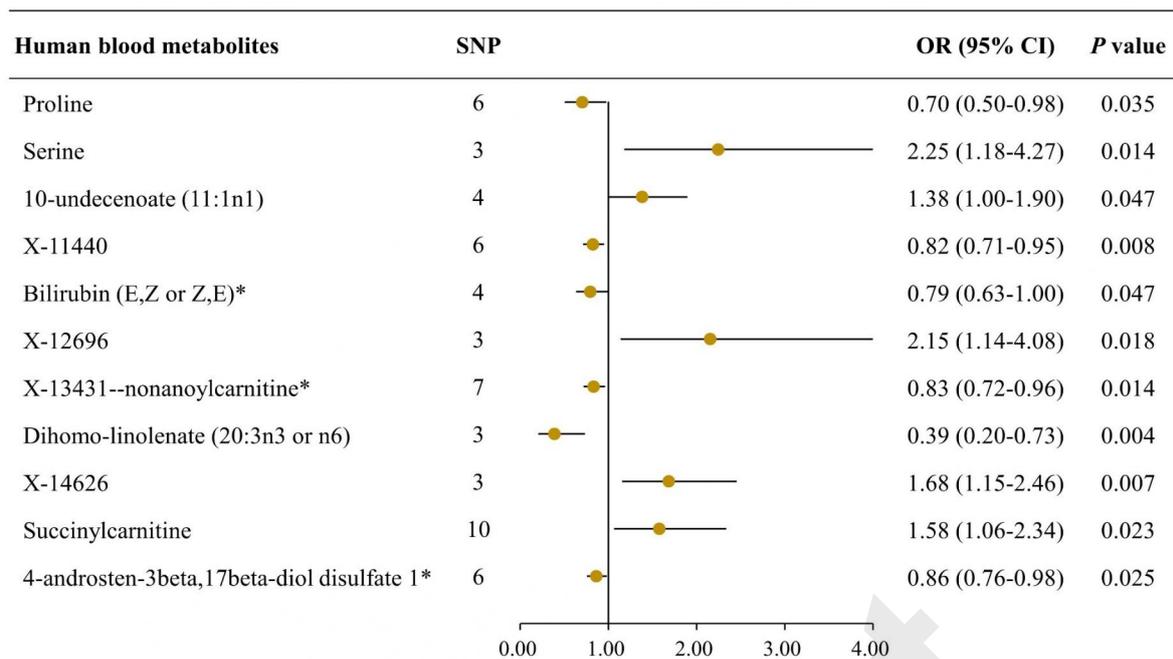


Figure 2 The associations of metabolites with the risk of breast cancer using the inverse-variance weighted mendelian randomization analysis
 Abbreviations: OR, odds ratio; 95% CI, 95% confidence interval

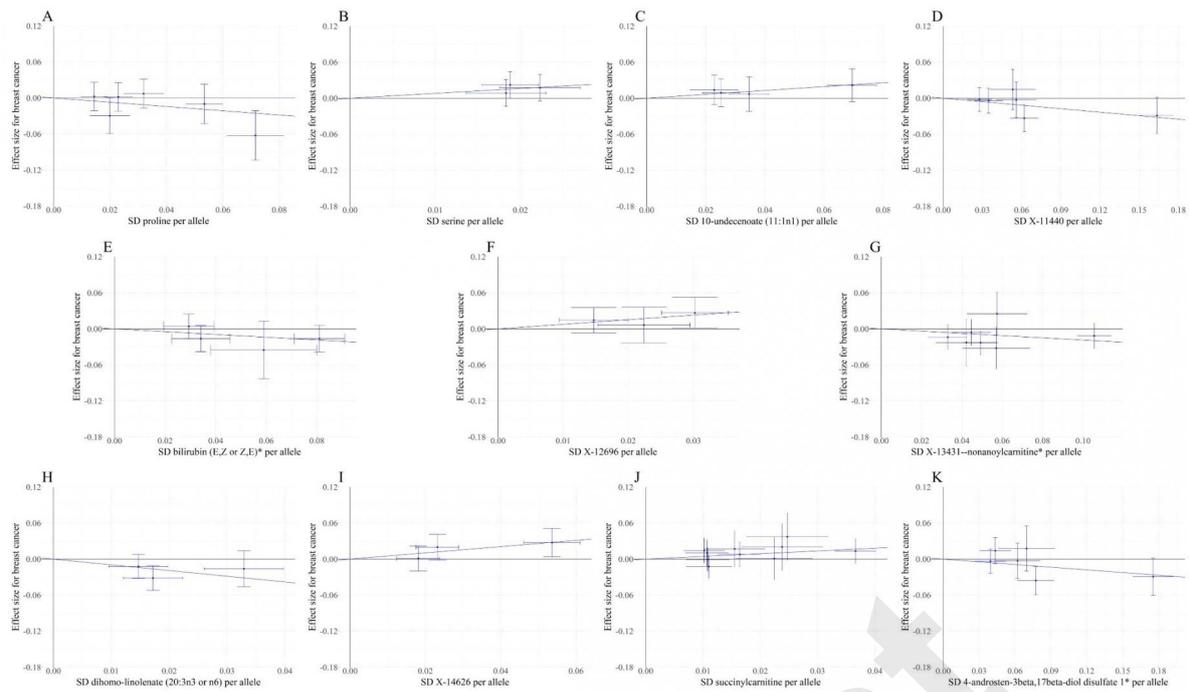


Figure 3 Associations of genetic variants about identified metabolites with the risk of breast cancer

The line indicates the estimate of causal effect using inverse-variance weighted method.

Circles indicate associations of each genetic variant related to metabolites with the risk of breast cancer. Error bars genetic indicate 95% confidence interval

A, Proline; B, Serine; C, 10-undecenoate (11:1n1); D, X-11440; E, Bilirubin (E,Z or Z,E)*; F, X-12696; G, X-13431--nonanoylcarnitine*; H, Dihomo-linolenate (20:3n3 or n6); I, X-14626; J, Succinylcarnitine; K, 4-androsten-3beta,17beta-diol disulfate 1*

Abbreviations: SNP: single-nucleotide polymorphism