Evaluating the Causal Relationship Between Carnitine-Related Metabolites and Breast Cancer: A Mendelian Randomization and Transcriptomic Analysis

Keywords
female malignant tumours, fatty acid metabolic intermediates, genetic causal inference, tumour microenvironment regulation

Abstract

Introduction

Breast cancer (BC) is the most common malignancy among women worldwide. Although observational studies have linked carnitine-related metabolites (CRMs) to BC, causal inference has been limited by confounding and reverse causality. This study utilized Mendelian randomization (MR) analysis to investigate the potential causal link between CRMs and BC.

Material and methods

MR analysis was conducted using the Cancer Genome Atlas-BC and CRM-related datasets to explore the causal relationship between CRMs and BC. Variants were screened based on criteria encompassing genome-wide significance ($p < 5\times10^{-8}$), independence ($r^2 < 0.001$, kb=10000), and sufficient strength (F-statistic > 10). Various MR techniques, including Inverse-Variance Weighted, MR-Egger, simple mode, weighted mode, and weighted median approaches, were applied to assess these causal associations. Single-cell RNA sequencing (scRNA-seq) was employed to investigate the expression and biological functions of biomarkers linked to metabolites.

Results

The MR analysis revealed a significant causal relationship between elevated levels of octanoylcarnitine and decanoylcarnitine and an increased risk of BC. Functional enrichment analysis identified 12 candidate genes associated with these metabolites, which are involved in fatty acid β-oxidation (FAO) pathways. ScRNA-seq analysis revealed eight distinct cell subpopulations, with macrophages exhibiting the highest intercellular communication. Six biomarkers were identified as potential contributors to BC development: ACADM, FNIP2, RAPGEF2, RABGGTB, PPID, and ST6GALNAC3.

Conclusions

This study identifies octanoylcarnitine and decanoylcarnitine as significant causal risk factors for BC, linked to dysregulated FAO. Integration of single-cell transcriptomics revealed six CRM-associated biomarkers and their dynamic expression within tumor microenvironments. These findings provide novel candidates for BC diagnosis and personalized therapy, warranting further validation.

Evaluating the Causal Relationship Between Carnitine-Related Metabolites and Breast Cancer: A Mendelian Randomization and Transcriptomic Analysis

Feifeng Ran[a,b#], Lilin Que [a,b#], Lan Luo[b], Mei Gan[b], Rensheng Wang[a]*, Leifeng Liang[b]*

1 Department of Radiation Oncology, The First Affiliated Hospital of Guangxi Medical University, Nanning, Guangxi, China

2 Department of Oncology, The Sixth Affiliated Hospital of Guangxi Medical University, The First People's Hospital of Yulin, Yulin, Guangxi, China

#Feifeng Ran and Lilin Que contributed equally and considered as co-first authors.

*** Corresponding Author:**

Leifeng Liang

Mailing address: No. 495, Jiaoyu Middle road, Yuzhou district, Yulin, Guangxi, 537000, China

E-mail: liangleifeng@stu.gxmu.edu.cn

Tel:13657757927

Rensheng Wang

Mailing address: No. 6, Shuangyong road, Qingxiu district, Nanning, Guangxi , 530000, China

Email: 13807806008@163.com

Tel:13807806008

Abstract

Background: Breast cancer (BC) is the most common malignancy among women worldwide. Although observational studies have linked carnitine-related metabolites (CRMs) to BC, causal inference has been limited by confounding and reverse causality. This study utilized Mendelian randomization (MR) analysis to investigate the potential causal link between CRMs and BC.

Methods: MR analysis was conducted using the Cancer Genome Atlas-BC and CRM-related datasets to explore the causal relationship between CRMs and BC. Variants were screened based on criteria encompassing genome-wide significance ($p < 5\times10^{-8}$), independence ($r^2 < 0.001$, kb=10000), and sufficient strength (F-statistic > 10). Various MR techniques, including Inverse-Variance Weighted, MR-Egger, simple mode, weighted mode, and weighted median approaches, were applied to assess these causal associations. Single-cell RNA sequencing (scRNA-seq) was employed to investigate the expression and biological functions of biomarkers linked to metabolites.

Results: The MR analysis revealed a significant causal relationship between elevated levels of octanoylcarnitine and decanoylcarnitine and an increased risk of BC. Functional enrichment analysis identified 12 candidate genes associated with these metabolites, which are involved in fatty acid β-oxidation (FAO) pathways. ScRNA-seq analysis revealed eight distinct cell subpopulations, with macrophages exhibiting the highest intercellular communication. Six biomarkers were identified as potential contributors to BC development: ACADM, FNIP2, RAPGEF2, RABGGTB, PPID, and ST6GALNAC3.

Conclusions: This study identifies octanoylcarnitine and decanoylcarnitine as significant causal risk factors for BC, linked to dysregulated FAO. Integration of

single-cell transcriptomics revealed six CRM-associated biomarkers and their dynamic expression within tumor microenvironments. These findings provide novel candidates for BC diagnosis and personalized therapy, warranting further validation.

Keywords: Female malignant tumours; fatty acid metabolic intermediates; genetic causal inference; tumour microenvironment regulation

Highlights

1.This study represents the first application of Mendelian randomisation to establish a significant causal association between elevated levels of octanoylcarnitine and decanoylcarnitine and breast cancer risk, providing robust genetic evidence for the role of metabolic reprogramming in breast cancer pathogenesis.

2.Through integrating multi-omics data, six key genes closely associated with carnitine metabolism were identified. Their expression characteristics within the tumour microenvironment were elucidated at the single-cell level, providing novel molecular biomarkers for early breast cancer diagnosis and targeted therapy.

3.Furthermore, a multi-level regulatory network for these biomarkers was constructed, and potential therapeutic targets were predicted. This suggests that ACADM and PPID may serve as drug targets, opening new avenues for breast cancer treatment strategies based on metabolic intervention.

Introduction

Breast cancer (BC) is the most common malignancy among women worldwide [1]. In 2022, the number of new BC cases globally approached 2.3 million and is projected to rise to approximately 3 million by 2050 [2, 3]. Advanced BC can lead to multi-organ dysfunction due to distant metastasis, directly threatening life [4, 5]. Age, genetics, environment, and lifestyle factors are significant risk factors for BC [6]. From a pathological perspective, BC is primarily classified into non-invasive and invasive types. Rare subtypes such as mucinous carcinoma typically have a lower histological grade and higher hormone receptor expression, resulting in a relatively favorable prognosis [7]. Molecular subtyping further categorizes BC into intrinsic subtypes with distinct biological behaviors and therapeutic targets, such as Luminal A, Luminal B, HER2-positive, and triple-negative BC, based on the expression of key biomarkers (ER, PR, HER2, and Ki-67) [8]. In terms of treatment, the development of novel drugs and the application of comprehensive approaches including surgery, radiotherapy, chemotherapy, endocrine therapy, and targeted therapy have increased the 5-year survival rate for early-stage patients to nearly 90% [9]. However, significant challenges persist, including the heterogeneity of triple-negative BC (TNBC) and the development of treatment resistance across subtypes [10, 11]. Therefore, enhancing early detection and targeted anticancer therapy for BC is crucial for further improving patient survival rates.

Metabolic reprogramming, established as one of the core hallmarks of cancer, functions as a pivotal mechanism driving tumorigenesis and progression [12]. Its

intricate crosstalk with epigenetic regulation is postulated to cooperatively facilitate the advancement of BC [13, 14]. This reprogramming is primarily characterized by enhanced glucose metabolism, increased fatty acid synthesis, and an elevated rate of glutamine metabolism [15], which collectively enable cancer cells to flexibly adjust their energy metabolism and biosynthetic pathways. These adaptations provide indispensable support for rapid proliferation, invasion, metastasis, and adaptation to the complex tumor microenvironment (TME) [16]. Among these alterations, dysregulated lipid metabolism stands out as one of the most prominent metabolic disturbances in cancer, significantly influencing cell membrane architecture, diverse energy processes, and intercellular communication [17]. Specifically, the catabolism of fatty acids serves as a crucial energy source for ATP generation within cancer cells [18]. Growing evidence indicates that increased fatty acid synthesis and catabolism synergistically support the migratory capacity of metastatic BC cells [19]. Carnitine, specifically acylcarnitines, transport fatty acids into mitochondria for β-oxidation [20]. Consequently, the levels and functional status of carnitine-related metabolites (CRMs) directly govern the flux of fatty acid oxidation, thereby positioning CRMs as a critical nexus linking lipid metabolic reprogramming to cellular energy supply and the malignant phenotype in BC. While epidemiological studies have identified associations between circulating carnitine levels and BC risk [21], inherent limitations in study design and methodology have precluded deeper exploration of causality and underlying mechanisms.

Therefore, this study is the first to systematically integrate MR with multi-omics

analysis, aiming to overcome the limitations of previous correlational research, clarify the causal relationship between CRMs and BC, and explore their underlying mechanisms. MR, which uses genetic variants as instrumental variables (IV), can effectively control for confounding biases and reverse causation, providing a robust method for inferring causality between exposures and outcomes [22, 23]. Based on this, we utilized large-scale genomic and transcriptomic data to screen for CRMs with causal effects on BC risk. Furthermore, by integrating single-cell RNA sequencing (scRNA-seq), cell communication analysis, and regulatory network construction, we systematically untangle the functional roles of the CRMs within the TME and identify their downstream key biomarkers. This work provides novel causal evidence for the etiological mechanisms of BC and offers molecular targets with translational potential for its early detection and targeted therapy.

2. Materials and methods

2.1 Data sources

Figure 1 presents the comprehensive study framework. The primary data for MR analysis were obtained from the Integrative Epidemiology Unit (IEU) Open GWAS database (https://gwas.mrcieu.ac.uk/), including the BC dataset (ukb-b-12227) and CRM datasets (met-a-742, met-a-618, met-a-615, met-a-463, met-a-479, met-a-699, met-a-467, met-a-573, met-a-476, and met-a-652). Single-cell datasets GSE176078 (26 BC) [24] and GSE42568 (BC: control = 104:17) [25] were retrieved from the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/gds). The Cancer Genome Atlas (TCGA)-BC (BC: control = 1,082:113) dataset was sourced

from the University of California Santa Cruz (UCSC) Xena database (https://xenabrowser.net/datapages/) [26]. The data were accessed on September 1st, 2023.

2.2 MR analysis

To investigate the causal relationship between carnitine metabolites and BC risk, this study employed MR for analysis. The MR analysis was based on three core assumptions: (1) The IVs should be significantly correlated with CRM, (2) The IVs must not be influenced by any other confounding factors, and (3) BC must be affected by CRM solely through the IVs [22]. The *extract_instruments* function was used to assess exposure factors and screen IVs ($p < 5 \times 10^{-8}$) [27]. IVs significantly correlated with the exposure factors were identified. The "clump" variable was employed to detect closely related IVs, which were then assessed for linkage disequilibrium (LD) with criteria of $r^2 = 0.001$ and kb = 10,000. IVs for exposure factors and outcomes were harmonized, excluding SNPs with F-statistics < 10 ($F = (samplesize.exposure-2) \times ((R^2) / (1-R^2))$). The *extract_outcome_data* function was used to filter SNPs not associated with BC. MR analysis was performed using MR Egger regression [28], weighted median [29], inverse variance weighted (IVW) [30], and simple mode and weighted mode methods [31]. P-values and odds ratios (OR) for the IVW method were used, with the IVW approach treating the reciprocal of outcome variance (the square of SE) as the weight ($R^2 = (\beta^2) / (\beta^2 + SE^2 \times N)$). The following thresholds were used: $p < 0.05$ indicated a significant causal relationship between an exposure factor and patient outcome; OR > 1 indicated risk factors; and OR < 1 indicated safety

155 factors. MR results were visually represented using scatter plots, forest plots, and
156 funnel plots.
157 Several sensitivity analyses were performed to evaluate the validity and stability of
158 the causal relationships. Heterogeneity among IVs was assessed using the
159 *mr_heterogeneity* function, with homogeneity indicated by p-values > 0.05.
160 Horizontal pleiotropy was tested using the mr_pleiotropy_test function, where p-
161 values > 0.05 suggested no directional pleiotropic effects. Additionally, leave-one-out
162 analysis was conducted by systematically excluding individual SNPs to assess their
163 impact on overall causal effect estimates. These analyses were implemented within
164 the "TwoSampleMR" package [32].

165 **2.3 Enrichment analysis**

166 Genes associated with IVs showing significant causal relationships with the outcome
167 were retrieved from the eQTLGen database and considered as SNP target genes
168 related to cis-eQTLs. To elucidate the functions and processes of these SNP target
169 genes in BC, an enrichment analysis was performed using the Gene Ontology (GO)
170 and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases *via* the
171 “clusterProfiler” package [33] ($p < 0.05$).

172 **2.4 Single-cell analysis**

173 To decipher the cell-specific expression and functions of CRM-associated genes
174 within the BC microenvironment, this study performed single-cell RNA sequencing
175 analysis on BC tissues. Quality control for the GSE176078 dataset was conducted
176 using the "Seurat" package [34], with criteria set to exclude cells containing fewer
177 than 200 genes, genes detected in fewer than three cells, and cells with mitochondrial
178 gene proportions exceeding 5% [35]. Highly variable genes were identified using the

FindVariableFeatures function based on mean-variance relationships. Data normalization was performed using the *ScaleData* function. Following principal component analysis (PCA), uniform manifold approximation and projection (UMAP) was applied for cell cluster identification at a resolution of 0.25. Cell type annotation was performed using the "singleR" package [36].

2.5 Cell communication, pathway analysis, and single-cell trajectory in biological processes

To investigate the signalling interaction networks among different cell subpopulations within the TME of BC, the “cellchat” software package was employed to analyse cell communication between annotated cell types [37]. Cells exhibiting the highest number and intensity of intercellular interactions were selected as key cells. The FindALLMarkers function was applied to perform differential analysis on each sub-population, using the parameters only.pos = TRUE, min.pct = 0.2, and return.thresh = 0.01, to obtain single-cell differentially expressed genes (scDEGs). Subsequently, functional enrichment analyses of KEGG pathways and GO biological processes were performed using the org.Hs.eg.db and clusterProfiler packages (pvalueCutoff = 1 and qvalueCutoff = 1). To identify candidate biomarkers possessing both genetic association and cell-specificity, biomarkers were derived from the intersection of SNP target genes and scDEGs. Additionally, the “Monocle” package [38] was employed to visualize the developmental trajectories of key cells and analyze changes in the expression levels of biomarkers during cell differentiation.

2.6 Identification and expression analysis of biomarkers

Chromosomal localization of biomarkers was performed using the “RCircos” package [39], and subcellular localization predictions were made using the RNALocate database (http://rnalocate.org) to investigate the expression distribution of biomarkers

in cells. Expression analysis was performed using the TCGA-BC and GSE42568 datasets to compare the expression of biomarkers in BC and control groups. UMAP clustering heatmaps were generated to visualize the correlation between biomarkers and annotated cell types.

2.7 Molecular Regulatory Network and Drug Prediction

Transcription factors (TFs) regulating biomarker expression were identified through the ENCODE database chip-seq data *via* the NetworkAnalyst online platform (https://www.networkanalyst.ca/). The "multiMiR" package [40] was used to identify overlapping target microRNAs (miRNAs) from miRDB (https://www.mirdb.org/) and miRanda (http://www.microrna.org/microrna/home.do) databases. Overlapping miRNAs were considered as target miRNAs for further investigation. LncRNA analysis of biomarkers was conducted using the starBase database (http://starbase.sysu.edu.cn). To explore potential therapeutic targets for BC treatment, biomarkers were queried against the DGIdb database (https://www.dgidb.org) to identify candidate drugs.

2.8 Statistical analysis

All bioinformatics analyses were conducted using R (v 4.2.2) , with comparisons between groups performed using the Wilcoxon test [41].

3. Results

3.1 MR analysis indicated that octanoylcarnitine and decanoylcarnitine were risk factors for BC

The IVW analysis revealed a significant causal relationship ($p < 0.05$) between decanoylcarnitine (OR = 1.0140, 95%CI = 1.0003 - 1.0278, $p < 0.05$), octanoylcarnitine (OR = 1.0131, 95%CI = 1.0015 -1.0248, $p < 0.05$), and BC,

indicating that higher levels of both metabolites are associated with an increased risk of BC (Table (1)). The scatter plot showed a positive slope, further supporting their role as risk factors (Fig.(2A-B)). Additionally, forest plots corroborated these findings, demonstrating an elevated BC risk linked to increased decanoylcarnitine and octanoylcarnitine levels in the IVW analysis (Fig.(2C-D)). The funnel plot displayed approximate symmetry of IVs on both sides of the IVW line, suggesting that the MR analysis adhered to randomness (Fig.(2E-F)).

Sensitivity analyses confirmed the robustness of the MR results. Homogeneity testing ($p > 0.05$) and non-significant horizontal pleiotropy ($p > 0.05$) further validated the findings (Table (2)). Leave-one-out analysis showed consistent SNP effects on the outcomes, reinforcing the reliability of the analysis (Fig.(2G-H)). In summary, the MR analysis indicated that elevated levels of octanoylcarnitine and decanoylcarnitine significantly increase BC risk. To explore the role of IVs in the causal regulation between exposure factors and outcomes, two significant exposure factors shared three SNPs. A total of 12 SNP target genes were identified as being cis-regulated, including ACADM, ETFDH, FNIP2, RAPGEF2, RABGGTB, PPID, ABCC1, MSH4, TMEM144, ST6GALNAC3, CRYZ, and SLC44A5.

3.2 Functional analysis of 12 candidate genes

To understand their biological functions, an enrichment analysis was performed on these 12 SNP target genes. GO analysis revealed that these genes are associated with fatty acid β-oxidation（FAO） (Fig.(3A)), while KEGG analysis indicated involvement in pathways such as glycosphingolipid biosynthesis-ganglio series,

among others (Fig.(3B)).

3.3 Eight cell subpopulations are annotated

Quality control outcomes are shown in Supplementary Figure 1. Following standard preprocessing, 2000 genes with high variability were identified (Fig.(4A)). PCA was performed, and the first 30 principal components were selected for further analysis (Fig.(4B-D)). UMAP clustering analysis revealed 32 distinct cell clusters (Fig.(4E)). Eight cell subpopulations were annotated, including epithelial cells, fibroblasts, and CD8+ T-cells (Fig.(4F)). Cellular communication analysis showed enhanced communication between macrophages and macrophages and between macrophages and epithelial cells (Fig.(4G – H)).

3.4 Expression and signaling pathways involving biomarkers in scRNA-seq

Macrophages exhibited the highest number and intensity of interactions with other annotated cells during cell communication, thus designating them as key cells. A total of 7,890 scDEGs across various cell clusters were identified (Fig.(5A)). KEGG pathway and GO biological process analyses of annotated cell sub-populations revealed that macrophages were primarily enriched in processes such as antigen processing and presentation of exogenous peptide antigens, vesicle organization, vacuole organization, and macroautophagy. In contrast, epithelial cells were predominantly enriched in aerobic respiration, mitochondrial ATP synthesis coupled with electron transport, ribonucleoprotein complex biogenesis, and cytoplasmic translation (Fig.(5B)). The overlap between the 7,890 scDEGs and the 12 SNP target genes identified six biomarkers: ACADM, FNIP2, RAPGEF2, RABGGTB, PPID, and ST6GALNAC3 (Fig.(5C)). Following biomarker identification, their pseudo-temporal trajectories in macrophages were prioritized for further study. As shown in

Fig.(5D-E), macrophage differentiation proceeded in three directions. Notably, state 7 differentiated earlier, while state 1 differentiated later. The expression levels of ACADM, RAPGEF2, and FNIP2 exhibited smooth changes, gradually decreasing throughout macrophage development. PPID expression remained stable throughout differentiation. In contrast, RABGGTB expression remained stable initially and then increased, while ST6GALNAC3 expression remained consistently low during macrophage development (Fig.(5F)).

3.5 Localization information and expression levels of biomarkers

The chromosomal locations of the biomarkers in human DNA are shown in Fig.(6A). Specifically, ACADM, RABGGTB, and ST6GALNAC3 are located on chromosome 1, while FNIP2, RAPGEF2, and PPID are located on chromosome 4. Subcellular localization analysis revealed that the biomarkers were primarily distributed in the cytoplasm, ribosomes, endoplasmic reticulum, ribosome-free cytoplasm, and nucleus, with the majority of the proteins localized in the cytoplasm (Fig.(6B)). Expression analysis of the TCGA-BC and GSE42568 datasets demonstrated significant downregulation of ACADM, RAPGEF2, and ST6GALNAC3 in BC samples (Fig.(6C-D)). As presented in Fig. 6E, ACADM was among the highly variable genes in epithelial cells, whereas RAPGEF2 and ST6GALNAC3 were associated with highly variable genes in endothelial cells, and RAPGEF2 was also linked to highly variable genes in macrophages.

3.6 Results of molecular regulatory network and drug prediction for biomarkers

For ACADM, 21 TFs were predicted; for RABGGTB, 20 TFs; and for PPID, 6 TFs. No corresponding TF data were predicted for the remaining three biomarkers. These TFs were used to establish a TF-mRNA network, with the TF TAF7 co-targeting and

298 regulating three biomarkers (Fig.(7A)).

299 Additionally, the mRNA-miRNA-lncRNA regulatory network revealed that 157
300 miRNAs (such as hsa-miR-105-5p, hsa-miR-106b-5p, hsa-miR-1252-5p, and hsa-
301 miR-126-5p) and 68 lncRNAs (such as NEAT1 and XIST) interacted with the
302 biomarkers (Supplementary Table (1), Fig.(7B)). Drug prediction analysis identified
303 31 targeted drugs for ACADM (e.g., irbesartan, angiotensin A, and trv-120027) and 3
304 drugs for PPID (SCY-635, rencofilstat, and cyclosporine) in BC (Fig.(7C),
305 Supplementary Table (2)). In summary, this comprehensive analysis provides an in-
306 depth exploration of the molecular regulatory mechanisms that may contribute to BC.
307 The identification of potential drug targets offers promising avenues for therapeutic
308 intervention.

309 ## Discussion

310 BC exhibits significant diversity at both the cellular and molecular levels, which
311 greatly impacts prognosis, particularly in subtypes like TNBC. TNBC accounts for
312 15% of all BC cases and is associated with a relatively poor prognosis due to its high
313 metastatic potential [42, 43]. A recent study highlighted that patients with hormone
314 receptor-positive (HR+)/human epidermal growth factor receptor 2-negative (HER2-)
315 BC exhibit distinct biological and clinical characteristics, suggesting the need for
316 tailored therapeutic approaches for this group [44]. Therefore, identifying new
317 biomarkers specific to these subtypes is critical. CRM have been proposed as
318 potential universal prognostic biomarkers in various cancers [45, 46]. Our findings
319 reveal a significant causal relationship between elevated levels of these metabolites

and increased BC risk, emphasizing their potential as biomarkers for early detection and risk assessment in clinical settings.

This study identifies octanoylcarnitine and decanoylcarnitine as risk factors for BC. Within BC, octanoylcarnitine levels exhibit dynamic alterations. It was found that its levels are reduced in T3-stage tumors, yet elevated in ER/PR-positive tumors [47]. We hypothesize that elevated octanoylcarnitine may drive malignant transformation and proliferation in tumorigenesis and early progression by providing energy substrates and biosynthetic precursors. In advanced stages, its levels may be depleted and decline, possibly due to metabolic reprogramming, such as enhanced Warburg effect, and microenvironmental changes [15]. However, the biological roles of these acylcarnitines display notable heterogeneity and context-dependency. For instance, elevated octanoylcarnitine is associated with increased risk of gastrointestinal cancer [48], while it exhibits inhibitory effects in ovarian cancer [49]. Similarly, decanoylcarnitine has been reported to show a positive association with the risk of most cancers [50], but in metastatic pancreatic cancer patients, its higher serum levels are linked to favorable body composition characteristics (e.g., greater subcutaneous fat) and longer survival [51]. These findings suggest that the biological effects of octanoylcarnitine and decanoylcarnitine are highly dependent on cancer type, disease stage, molecular subtype, and the local metabolic microenvironment. In summary, the roles of these two metabolites in cancer are complex and dynamic. Future work will involve functional experiments and validation in clinical cohorts to systematically elucidate their dynamic changes and mechanisms of action across different stages and

subtypes of BC, and to explore their potential as stage-specific biomarkers or therapeutic targets.

The 12 SNP-targeted genes identified in this study were significantly enriched in FAO and related metabolic pathways, which represent the core processes through which carnitine regulates fatty acid metabolism. Further studies have shown that modulating fatty acid metabolism via the AKT/RYR2 signaling pathway promotes the progression of TNBC[52]. Meanwhile, FAO can enhance BC invasion and metastasis through the miRNA-328-3p/CPT1A axis [53]. Furthermore, fatty acid oxidation is a key metabolic pathway that sustains the malignant mesenchymal phenotype of tumor cells and drives metastasis. Disruption of this pathway, through pharmacological (e.g., retinoic acid) or genetic intervention, can alter lipid metabolic flux, influence epigenetic modifications, and subsequently reverse cellular state while suppressing metastatic potential [54]. These findings support the view that alterations in carnitine levels play a key role in the metabolic reprogramming of BC, and further research is needed to explore the mechanisms by which carnitine contributes to the development and progression of BC.

The complex TME of BC consists of tumor cells, immune cells, fibroblasts, vascular cells, and the extracellular matrix, all of which contribute to tumor growth, progression, and metastasis [55]. Cellular communication analysis revealed that Macrophage-Macrophage interactions and Macrophage-Epithelial cell communications were upregulated. Macrophages, as immune effector cells, play a critical role in the immune microenvironment of BC. Numerous studies have

established the protumor effects of alternatively activated (M2) tumor-associated macrophages (TAMs), correlating their presence with poor prognosis in BC and other cancers [56, 57]. TAMs contribute to tumor progression by promoting angiogenesis, suppressing T cell-mediated immunity, and modulating metabolic activities and associated metabolites, thereby shaping the overall metabolic landscape of the TME [58]. Moreover, this study revealed that macrophage subpopulations are predominantly enriched in macroautophagy. As a central regulatory process governing cellular metabolism and homeostasis, the activation of autophagy is closely associated with functional polarization, survival, and pro-tumor activity of macrophages [59, 60, 61]. Targeting autophagic processes in macrophages may thus represent a promising therapeutic strategy to reverse their pro-tumor functions.

To explore therapeutic targets related to CRM in BC tumor cells, six biomarker genes were identified by overlapping DEGs and candidate genes: ACADM, FNIP2, RAPGEF2, RABGGTB, PPID, and ST6GALNAC3 (Supplementary Table (3)). Notably, ST6GALNAC3, ACADM, and RAPGEF2 were significantly downregulated in the BC group, suggesting their involvement in BC development and progression. The identified SNP target genes, including ACADM, are integral to the carnitine shuttle mechanism and are associated with promoting FAO, a pathway that supports cancer cell proliferation, migration, and invasion. These biomarkers may hinder tumor development by inhibiting FAO, leading to energy deprivation and abnormal lipid accumulation [62, 63, 64]. The expression of these genes correlates with tumor

progression and metastasis, indicating that targeting these metabolic pathways could provide therapeutic strategies for mitigating BC.

However, this study has several limitations. First, the inclusion of participants of European ancestry may limit the applicability of the findings to diverse global populations. Future research should validate these results in more diverse cohorts and further confirm the conclusions by integrating clinical or epidemiological data. Moreover, although the MR approach enhances the reliability of causal inference and transcriptomic analysis provides mechanistic insights, the specific biological functions of the identified biomarkers and related metabolites in the development and progression of BC remain incompletely understood. These preliminary associations still require validation through in vitro and in vivo functional experiments to further assess their effects on cell proliferation, invasion, and lipid metabolism.

Conclusion

In summary, this study demonstrates that elevated levels of octanoylcarnitine and decanoylcarnitine serve as significant risk factors for BC. The genes annotated by the associated SNPs were predominantly enriched in the FAO pathway, suggesting that these metabolites may play a critical role in BC progression by modulating this metabolic process. Further integration of single-cell transcriptomic analysis enabled the identification of six CRM-related potential biomarkers: ACADM, FNIP2, RAPGEF2, RABGGTB, PPID, and ST6GALNAC3. Notably, single-cell profiling revealed dynamic expression patterns of these biomarkers within macrophage subpopulations in the TME, indicating that CRM-mediated metabolic reprogramming

may be closely associated with immune-stromal crosstalk. These findings provide novel candidate biomarkers and potential therapeutic targets for early detection, risk stratification, and personalized treatment of BC. However, further functional validations in more diverse populations are warranted to elucidate the specific molecular mechanisms through which these metabolites and their corresponding genes contribute to BC progression.

Disclosure Statement

Data statement: The datasets generated and analysed within this study are available from the Gene Expression Omnibus (GEO) database repository, [https://www.ncbi.nlm.nih.gov/geo/]; the University of California Santa Cruz (UCSC) Xena databases repository, [https://xenabrowser.net/datapages/] and the Integrative Epidemiology Unit (IEU) Open genome-wide association study (GWAS) database repository, [https://gwas.mrcieu.ac.uk/].

Funding: This study was supported the National Natural Science Foundation of China (82260627).

Conflict of Interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Ethics statement: We used publicly available summary-level data. Ethical approval and informed consent for each included study can be determined in each original publication. Therefore, no additional ethical approval was needed. Registry and the registration No. of the study/trial: N/A. Animal Studies: N/A.

Author contributions: Feifeng Ran: data curation, formal analysis, Writing - original draft. Lilin Que: Software, validation. Lan Luo: Visualization, methodology. Mei Gan: Investigation. Rensheng Wang: Conceptualization, resources, supervision, project administration. Leifeng Liang: Conceptualization, supervision, funding acquisition, project administration, writing - review & editing.

Figure Legends

Fig. (1) Flowchart of the study design. IEU Open GWAS, Integrative Epidemiology Unit Open genome-wide association study; BC, breast cancer; GEO, Gene Expression Omnibus; UCSC, University of California Santa Cruz; TCGA, The Cancer Genome Atlas; MR, Mendelian randomization; PCA, principal component analysis.

Fig. (2) Scatter plots of the causal association between CRM and BC risk. (A) decanoylcarnitine and BC risk, (B) octanoylcarnitine and BC risk. Forest plots for diagnostic efficacy of (C) decanoylcarnitine and (D) octanoylcarnitine. (E-F) Funnel plot for randomness judgment. MR leave-one-out sensitivity test for (G) decanoylcarnitine and (H) octanoylcarnitine. SNP, single nucleotide polymorphism. MR, Mendelian randomization. OR, odds ratio.

Fig. (3) Functional enrichment analysis of 12 candidate genes. (A) The results of GO analysis. (B) Bubble plot for KEGG pathway enrichment analysis. GO, Gene Ontology. KEGG, Kyoto Encyclopedia of Genes and Genomes.

Fig. (4) Single-cell RNA-seq of BC tissues. (A) Results of screening highly variable genes. (B) Dimension reduction by PCA. (C) Scatter plot of principal components. (D) Elbow plot of principal components. (E) UMAP cluster analysis. (F) Single-cell annotations. (G) Cellular communication. (H) Heat map of cellular communication. PCA, principal component analysis. UMAP, uniform manifold approximation and projection.

Fig. (5) Expression and signaling pathways in scRNA-seq. (A) Single-cell scDEGs Identification. (B) Functional annotation of single-cell clusters. (C) Identification of

core biomarkers. (D) Pseudotime analysis of macrophages. (E) The differentiation state of macrophages. (F) The expression levels of genes in development of macrophages. scDEGs, differential expressed genes.

Fig. (6) Localization information and expression levels of biomarkers. (A) Location of biomarker genes on human chromosomes. (B) Subcellular localization analysis. (C-D) Expression analysis in TCGA-BRCA and GSE42568 datasets. (E) Expression analysis of highly variable genes. SNP, single nucleotide polymorphism; DEG, differentially expressed gene.

Fig. (7) The molecular regulatory network and drug prediction for biomarkers. (A) The TFs prediction for biomarkers. (B) The mRNA-miRNA-lncRNA regulatory network. (C) The results of drug prediction analysis. TFs, transcription factors.

Table Legends

Table (1) Mendelian randomization analysis of carnitine metabolites on BC

Table (2) The results of sensitivity analysis

References

1. Giaquinto A, Sung H, Miller K, Kramer J, Newman L, Minihan A, *et al*. Breast Cancer Statistics, 2022. *CA Cancer J Clin*. 2022;72(6):524-41.

2. Elbasheer M M A, Dodwell D, Gathani T. Understanding global variation in breast cancer mortality. *Br J Radiol*. 2025;98(1173):1369-72.

3. Freihat O, Sipos D, Kovacs A. Global burden and projections of breast cancer incidence and mortality to 2050: a comprehensive analysis of GLOBOCAN data. *Front Public Health*. 2025;13:1622954.

4. Duan H, Zhang Y, Qiu H, Fu X, Liu C, Zang X, *et al*. Machine learning-based prediction model for distant metastasis of breast cancer. *Comput Biol Med*. 2024;169:107943.

5. Huang J, Lin Q. Correlation between multimodality imaging features and molecular subtypes in breast cancer: a comparative study between young (≤30 years) and middle-aged (45-55 years) women. *Quant Imaging Med Surg*. 2025;15(8):7537-54.

6. Engmann N J, Golmakani M K, Miglioretti D L, Sprague B L, Kerlikowske K. Population-Attributable Risk Proportion of Clinical Risk Factors for Breast Cancer. *JAMA Oncol*. 2017;3(9):1228-36.

7. Hu T, Huang J, Fang K. Overall Survival in Patients with Mucinous Carcinoma of Breast: A Population-Based Study. *Int J Gen Med*. 2021;14:9991-10001.

8. Jaiswal A, Kaushik N, Choi E H, Kaushik N K. Functional impact of non-coding RNAs in high-grade breast carcinoma: Moving from resistance to clinical

498 applications: A comprehensive review. *Biochim Biophys Acta Rev Cancer*.
499 2023;1878(4):188915.
500 9. Miller K, Nogueira L, Devasia T, Mariotto A, Yabroff K, Jemal A, et al. Cancer
501 treatment and survivorship statistics, 2022. *CA Cancer J Clin*. 2022;72(5):409-36.
502 10. Yu J, Mu Q, Fung M, Xu X, Zhu L, Ho R J Y. Challenges and opportunities in
503 metastatic breast cancer treatments: Nano-drug combinations delivered preferentially
504 to metastatic cells may enhance therapeutic response. *Pharmacol Ther*.
505 2022;236:108108.
506 11. An J, Peng C, Tang H, Liu X, Peng F. New Advances in the Research of
507 Resistance to Neoadjuvant Chemotherapy in Breast Cancer. *Int J Mol Sci*.
508 2021;22(17).
509 12. Ward P S, Thompson C B. Metabolic reprogramming: a cancer hallmark even
510 warburg did not anticipate. *Cancer Cell*. 2012;21(3):297-308.
511 13. Thakur C, Qiu Y, Pawar A, Chen F. Epigenetic regulation of breast cancer
512 metastasis. *Cancer Metastasis Rev*. 2024;43(2):597-619.
513 14. Becker L M, O'Connell J T, Vo A P, Cain M P, Tampe D, Bizarro L, *et al*.
514 Epigenetic Reprogramming of Cancer-Associated Fibroblasts Deregulates Glucose
515 Metabolism and Facilitates Progression of Breast Cancer. *Cell Rep*.
516 2020;31(9):107701.
517 15. Wang Y, Dang H, Wang Q, Wu S, Han L, Luo X U, *et al*. Mitochondrial pyruvate
518 dehydrogenase phosphatase metabolism disorder in malignant tumors. *Oncol Res*.
519 2025;33(8):1861-74.

16. Li S, Peng M, Tan S, Oyang L, Lin J, Xia L, *et al*. The roles and molecular mechanisms of non-coding RNA in cancer metabolic reprogramming. *Cancer Cell Int*. 2024;24(1):37.

17. Chakraborty S, Kumar A S, Banerjee S. Lipids: Driving Forces in the Underlying Biology of Carcinogenesis. *ACS Pharmacol Transl Sci*. 2025;8(7):1891-918.

18. Castelli S, De Falco P, Ciccarone F, Desideri E, Ciriolo M R. Lipid Catabolism and ROS in Cancer: A Bidirectional Liaison. *Cancers (Basel)*. 2021;13(21).

19. Andolino C, Cotul E K, Xianyu Z, Li Y, Bhat D, Ayers M, *et al*. Fatty acid synthase-derived lipid stores support breast cancer metastasis. *Cancer Metab*. 2025;13(1):35.

20. Longo N, Frigeni M, Pasquali M. Carnitine transport and fatty acid oxidation. *Biochim Biophys Acta*. 2016;1863(10):2422-35.

21. Zhang J, Wu G, Zhu H, Yang F, Yang S, Vuong A M, *et al*. Circulating Carnitine Levels and Breast Cancer: A Matched Retrospective Case-Control Study. *Front Oncol.* 2022;12:891619.

22. Smith G D, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol*. 2003;32(1):1-22.

23. Ference B A. Mendelian randomization studies: using naturally randomized genetic data to fill evidence gaps. *Curr Opin Lipidol*. 2015;26(6):566-71.

24. Wu S Z, Al-Eryani G, Roden D L, Junankar S, Harvey K, Andersson A, et al. A single-cell and spatially resolved atlas of human breast cancers. *Nat Genet*.

2021;53(9):1334-47.

25. Clarke C, Madden S F, Doolan P, Aherne S T, Joyce H, O'Driscoll L, *et al*. Correlating transcriptional networks to breast cancer survival: a large-scale coexpression analysis. *Carcinogenesis*. 2013;34(10):2300-8.

26. Comprehensive molecular portraits of human breast tumours. *Nature*. 2012;490(7418):61-70.

27. Hemani G, Zheng J, Elsworth B, Wade K, Haberland V, Baird D, et al. The MR-Base platform supports systematic causal inference across the human phenome. *Elife* 2018;7.

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(2):512-25.

29. Bowden J, Davey Smith G, Haycock P, Burgess S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genet Epidemiol* 2016;40(4):304-14.

30. Burgess S, Scott R, Timpson N, Davey Smith G, Thompson S. Using published data in Mendelian randomization: a blueprint for efficient identification of causal risk factors. *Eur J Epidemiol* 2015;30(7):543-52.

31. Hartwig F, Davey Smith G, Bowden J. Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. *Int J Epidemiol* 2017;46(6):1985-98.

32. Bou Sleiman M, Jha P, Houtkooper R, Williams R W, Wang X, Auwerx J. The

564 Gene-Regulatory Footprint of Aging Highlights Conserved Central Regulators. *Cell*
565 *Rep*. 2020;32(13):108203.
566 33. Rao S, Mondragón L, Pranjic B, Hanada T, Stoll G, Köcher T, *et al*. AIF-
567 regulated oxidative phosphorylation supports lung cancer development. *Cell Res*.
568 2019;29(7):579-91.
569 34. Hao Y, Hao S, Andersen-Nissen E, Mauck W, Zheng S, Butler A, *et al*. Integrated
570 analysis of multimodal single-cell data. *Cel*l. 2021;184(13):3573-87.e29.
571 35. Satija R, Farrell J, Gennert D, Schier A, Regev A. Spatial reconstruction of
572 single-cell gene expression data. *Nat Biotechnol* 2015;33(5):495-502.
573 36. Aran D, Looney A, Liu L, Wu E, Fong V, Hsu A, *et al*. Reference-based analysis
574 of lung single-cell sequencing reveals a transitional profibrotic macrophage. *Nat*
575 *Immunol*. 2019;20(2):163-72.
576 37. Jin S, Guerrero-Juarez C, Zhang L, Chang I, Ramos R, Kuan C, *et al*. Inference
577 and analysis of cell-cell communication using CellChat. *Nat Commun*.
578 2021;12(1):1088.
579 38. Xiaoning W, Guixin W, Qiaoqiao X, Yingxi L, Wenbin S, Zhaoyi L, *et al*.
580 Potential tumor-specific antigens and immune landscapes identification for mRNA
581 vaccine in thyroid cancer. *Front Oncol*. 2024;14(0).
582 39. Shukla V, Rao M, Zhang H, Beers J, Wangsa D, Wangsa D, *et al*. ASXL3 Is a
583 Novel Pluripotency Factor in Human Respiratory Epithelial Cells and a Potential
584 Therapeutic Target in Small Cell Lung Cancer. *Cancer Res*. 2017;77(22):6267-81.
585 40. Yuanbin R, Katerina J K, Boris T, Paula H, Richard A R, Russell B, *et al*. The

multiMiR R package and database: integration of microRNA-target interactions along with their disease and drug associations. *Nucleic Acids Res*. 2014;42(17).

41. Wilcoxon F. Individual comparisons of grouped data by ranking methods. *J Econ Entomol.* 1946;39:269.

42. Allemani C, Matsuda T, Di Carlo V, Harewood R, Matz M, Nikšić M, *et al*. Global surveillance of trends in cancer survival 2000-14 (CONCORD-3): analysis of individual records for 37 513 025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries. *Lancet*. 2018;391(10125):1023-75.

43. Howard F M, Olopade O I. Epidemiology of Triple-Negative Breast Cancer: A Review. *Cancer J*. 2021;27(1):8-16.

44. Jin X, Zhou Y, Ma D, Zhao S, Lin C, Xiao Y, *et al*. Molecular classification of hormone receptor-positive HER2-negative breast cancer. *Nat Genet*. 2023;55(10):1696-708.

45. Beloribi-Djefaflia S, Vasseur S, Guillaumond F. Lipid metabolic reprogramming in cancer cells. *Oncogenesis*. 2016;5(1):e189.

46. Lou S, Balluff B, Cleven A H G, Bovée J, McDonnell L A. Prognostic Metabolite Biomarkers for Soft Tissue Sarcomas Discovered by Mass Spectrometry Imaging. *J Am Soc Mass Spectrom*. 2017;28(2):376-83.

47. Park J, Shin Y, Kim T H, Kim D H, Lee A. Plasma metabolites as possible biomarkers for diagnosis of breast cancer. *PLoS One*. 2019;14(12):e0225129.

48. Wang Z, Wang C, Gao X. Exploring metabolic pathways in gastrointestinal cancer: new evidence of causality. *Discov Oncol*. 2025;16(1):1777.

608 49. Huang Y, Lin W, Zheng X. Causal association between 637 human metabolites
609 and ovarian cancer: a mendelian randomization study. *BMC Genomics*.
610 2024;25(1):97.
611 50. Breeur M, Ferrari P, Dossus L, Jenab M, Johansson M, Rinaldi S, *et al*. Pan-
612 cancer analysis of pre-diagnostic blood metabolite concentrations in the European
613 Prospective Investigation into Cancer and Nutrition. *BMC Med*. 2022;20(1):351.
614 51. Gunchick V, Brown E, Liu J, Locasale J W, Philip P A, Wang S C, *et al*.
615 Morphomics, Survival, and Metabolites in Patients With Metastatic Pancreatic
616 Cancer. *JAMA Netw Open*. 2024;7(10):e2440047.
617 52. Ye J, Wu S, Quan Q, Ye F, Zhang J, Song C, *et al*. Fibroblast Growth Factor
618 Receptor 4 Promotes Triple-Negative Breast Cancer Progression via Regulating Fatty
619 Acid Metabolism Through the AKT/RYR2 Signaling. *Cancer Med*.
620 2024;13(23):e70439.
621 53. Zeng F, Yao M, Wang Y, Zheng W, Liu S, Hou Z, *et al*. Fatty acid β-oxidation
622 promotes breast cancer stemness and metastasis via the miRNA-328-3p-CPT1A
623 pathway. *Cancer Gene Ther*. 2022;29(3-4):383-95.
624 54. Loo S Y, Toh L P, Xie W H, Pathak E, Tan W, Ma S, *et al*. Fatty acid oxidation is
625 a druggable gateway regulating cellular plasticity for driving metastasis in breast
626 cancer. *Sci Adv*. 2021;7(41):eabh2443.
627 55. Potter S. Single-cell RNA sequencing for the study of development, physiology
628 and disease. *Nature Reviews Nephrology*. 2018.
629 56. Lee A H, Happerfield L C, Bobrow L G, Millis R R. Angiogenesis and

630 inflammation in invasive carcinoma of the breast. *J Clin Pathol*. 1997;50(8):669-73.
631 57. Bingle L, Brown N J, Lewis C E. The role of tumour-associated macrophages in
632 tumour progression: implications for new anticancer therapies. *J Pathol*.
633 2002;196(3):254-65.
634 58. Huang X, Cao J, Zu X. Tumor-associated macrophages: An important player in
635 breast cancer progression. *Thorac Cancer*. 2022;13(3):269-76.
636 59. Zhang T, Yu J, Cheng S, Zhang Y, Zhou C H, Qin J, et al. Research Progress on
637 the Anticancer Molecular Mechanism of Targets Regulating Cell Autophagy.
638 *Pharmacology*. 2023;108(3):224-37.
639 60. Hu M, Fan J X, He Z Y, Zeng J. The regulatory role of autophagy between TAMs
640 and tumor cells. *Cell Biochem Funct*. 2024;42(2):e3984.
641 61. Hao Y, Duan F, Dong X, Bi R, Wang Y, Zhu S, *et al*. Gold Nanoparticle Inhibits
642 the Tumor-Associated Macrophage M2 Polarization by Inhibiting m(6)A Methylation-
643 Dependent ATG5/Autophagy in Prostate Cancer. *Anal Cell Pathol (Amst)*.
644 2025;2025:6648632.
645 62. Zheqiong T, Yaru Z, Man Z, Zhenzhao L, Tangwei W, Chao Z, *et al*. Carnitine
646 palmitoyl transferase 1A is a novel diagnostic and predictive biomarker for breast
647 cancer. *BMC Cancer*. 2021;21(1):409.
648 63. Huang D, Li T, Li X, Zhang L, Sun L, He X, *et al*. HIF-1-mediated suppression
649 of acyl-CoA dehydrogenases and fatty acid oxidation is critical for cancer
650 progression. *Cell Rep* 2014;8(6):1930-42.
651 64. Seok S, Kim Y, Byun S, Choi S, Xiao Z, Iwamori N, *et al*. Fasting-induced

652 JMJD3 histone demethylase epigenetically activates mitochondrial fatty acid β-
653 oxidation. *J Clin Invest* 2018;128(7):3144-59.

654

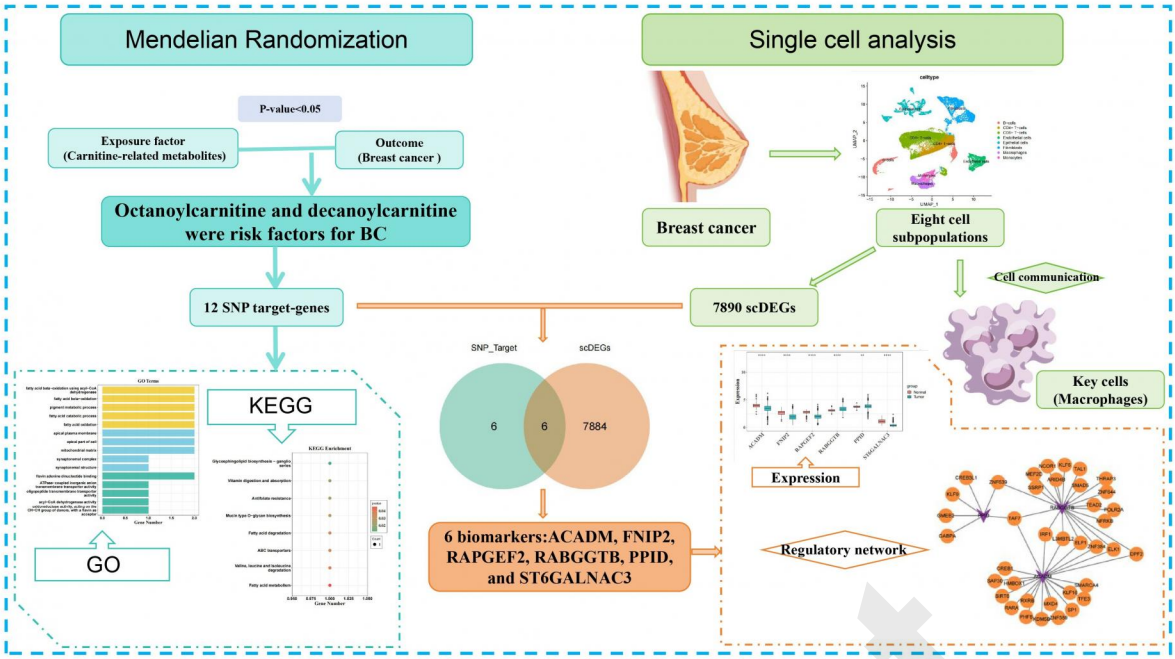
Mendelian Randomization
Single cell analysis
P-value<0.05
Exposure factor (Carnitine-related metabolites)
Outcome (Breast cancer)
Octanoylcarnitine and decanoylcarnitine were risk factors for BC
Breast cancer
Eight cell subpopulations
Cell communication
12 SNP target-genes
7890 scDEGs
Key cells (Macrophages)
SNP_Target
scDEGs
6
6
7884
KEGG
GO
Expression
Regulatory network
6 biomarkers:ACADM, FNIP2, RAPGEF2, RABGGTB, PPID, and ST6GALNAC3

Table 1 Mendelian randomization analysis of carnitine metabolites on breast cancer

Outcome	Exposure	nSNP	b	OR (95%CI)	p
ukb-b-12227	Succinylcarnitine	7	-0.001	0.9985 (0.9791 - 1.0183)	0.885
ukb-b-12227	Decanoylcarnitine	3	0.0139	1.0140 (1.0003 - 1.0278)	0.045
ukb-b-12227	Octanoylcarnitine	3	0.0130	1.0131 (1.0015 -1.0248)	0.026
ukb-b-12227	Propionylcarnitine	4	0.0126	1.0127 (0.9847 - 1.0414)	0.378
ukb-b-12227	Isobutyrylcarnitine	3	-0.0029	0.9971 (0.9833 - 1.0110)	0.68
ukb-b-12227	Acetylcarnitine	2	-0.0088	0.9912 (0.9583 - 1.0252)	0.608
ukb-b-12227	Glutaroyl carnitine	8	-0.0069	0.9931 (0.9739 - 1.0127)	0.488
ukb-b-12227	Butyrylcarnitine	5	0.0004	1.0004 (0.9905 - 1.0104)	0.939
ukb-b-12227	Isovalerylcarnitine	3	-0.0162	0.9839 (0.9487 - 1.0204)	0.382
ukb-b-12227	Hexanoylcarnitine	4	0.0136	1.0137 (0.9984 - 1.0291)	0.079

SNP, single nucleotide polymorphisms.

Table 1 The results of sensitivity analysis

Exposure	Outcome	Heterogeneity test (IVW)			Horizontal pleiotropy test (MR-Egger)		
		Q	p	I squared	Intercept	SE	p
Octanoylcarnitine	BC	2.204126153	0.332	0	-0.001256772	0.00089392 1	0.393595639
Decanoylcarnitine	BC	1.790034097	0.049	9.26	-0.000893961	0.000706888	0.425941195

IVW, inverse variance weighted; BC, breast cancer.

Fig. (1) Flowchart of the study design. IEU Open GWAS, Integrative Epidemiology Unit Open genome-wide association study; BC, breast cancer; GEO, Gene Expression Omnibus; UCSC, University of California Santa Cruz; TCGA, The Cancer Genome Atlas; MR, Mendelian randomization; PCA, principal component analysis.

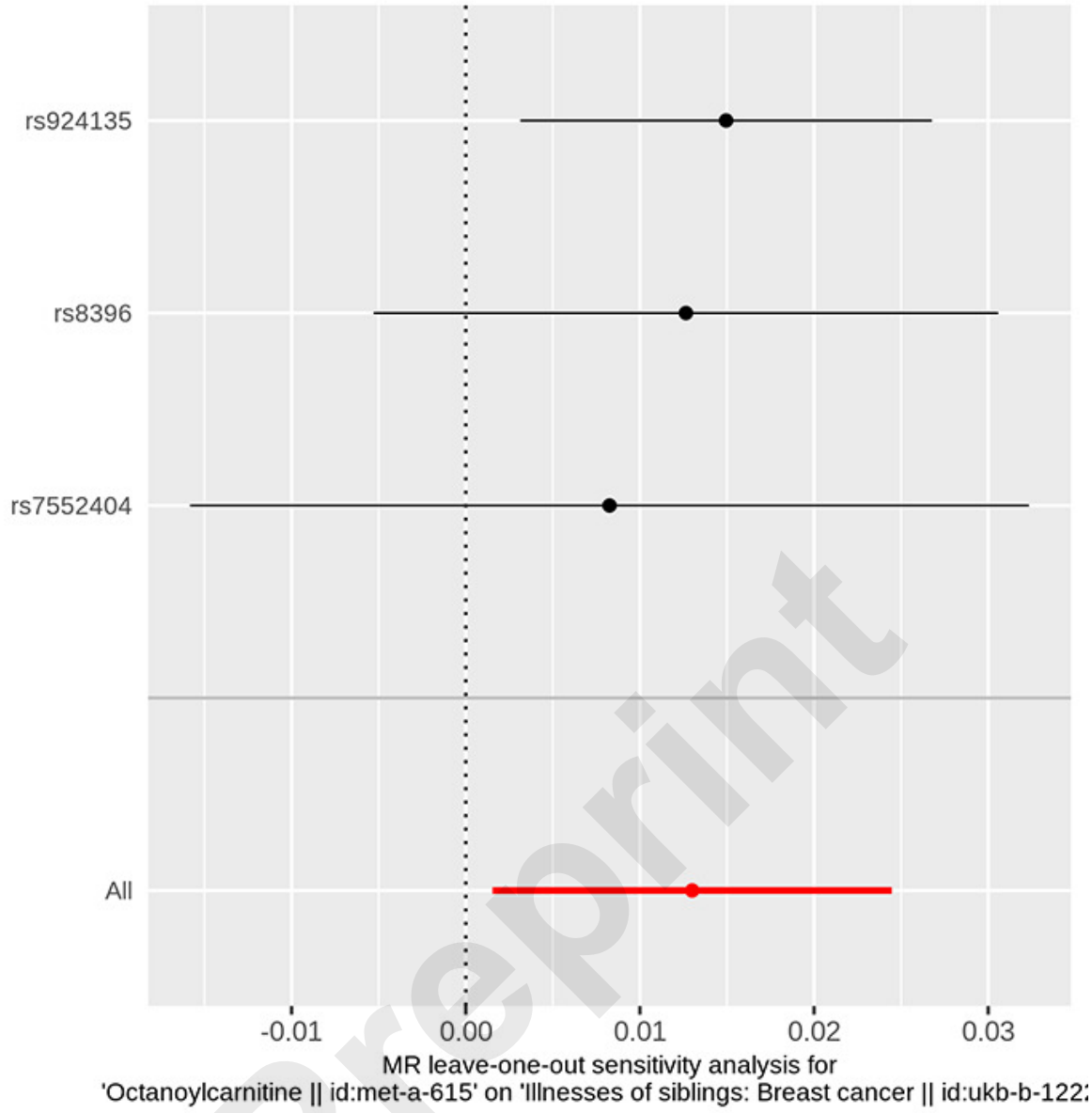

Fig. (2) Scatter plots of the causal association between CRM and BC risk. (A) decanoylcarnitine and BC risk, (B) octanoylcarnitine and BC risk. Forest plots for diagnostic efficacy of (C) decanoylcarnitine and (D) octanoylcarnitine. (E-F) Funnel plot for randomness judgment. MR leave-one-out sensitivity test for (G) decanoylcarnitine and (H) octanoylcarnitine. SNP, single nucleotide polymorphism. MR, Mendelian randomization. OR, odds ratio.

Fig. (3) Functional enrichment analysis of 12 candidate genes. (A) The results of GO analysis. (B) Bubble plot for KEGG pathway enrichment analysis. GO, Gene Ontology. KEGG, Kyoto Encyclopedia of Genes and Genomes.

Fig. (4) Single-cell RNA-seq of BC tissues. (A) Results of screening highly variable genes. (B) Dimension reduction by PCA. (C) Scatter plot of principal components. (D) Elbow plot of principal components. (E) UMAP cluster analysis. (F) Single-cell annotations. (G) Cellular communication. (H) Heat map of cellular communication. PCA, principal component analysis. UMAP, uniform manifold approximation and projection.

Fig. (5) Expression and signaling pathways in scRNA-seq. (A) Single-cell scDEGs Identification. (B) Functional annotation of single-cell clusters. (C) Identification of core biomarkers. (D) Pseudotime analysis of macrophages. (E) The differentiation state of macrophages. (F) The expression levels of genes in development of macrophages. scDEGs , differential expressed Genes .

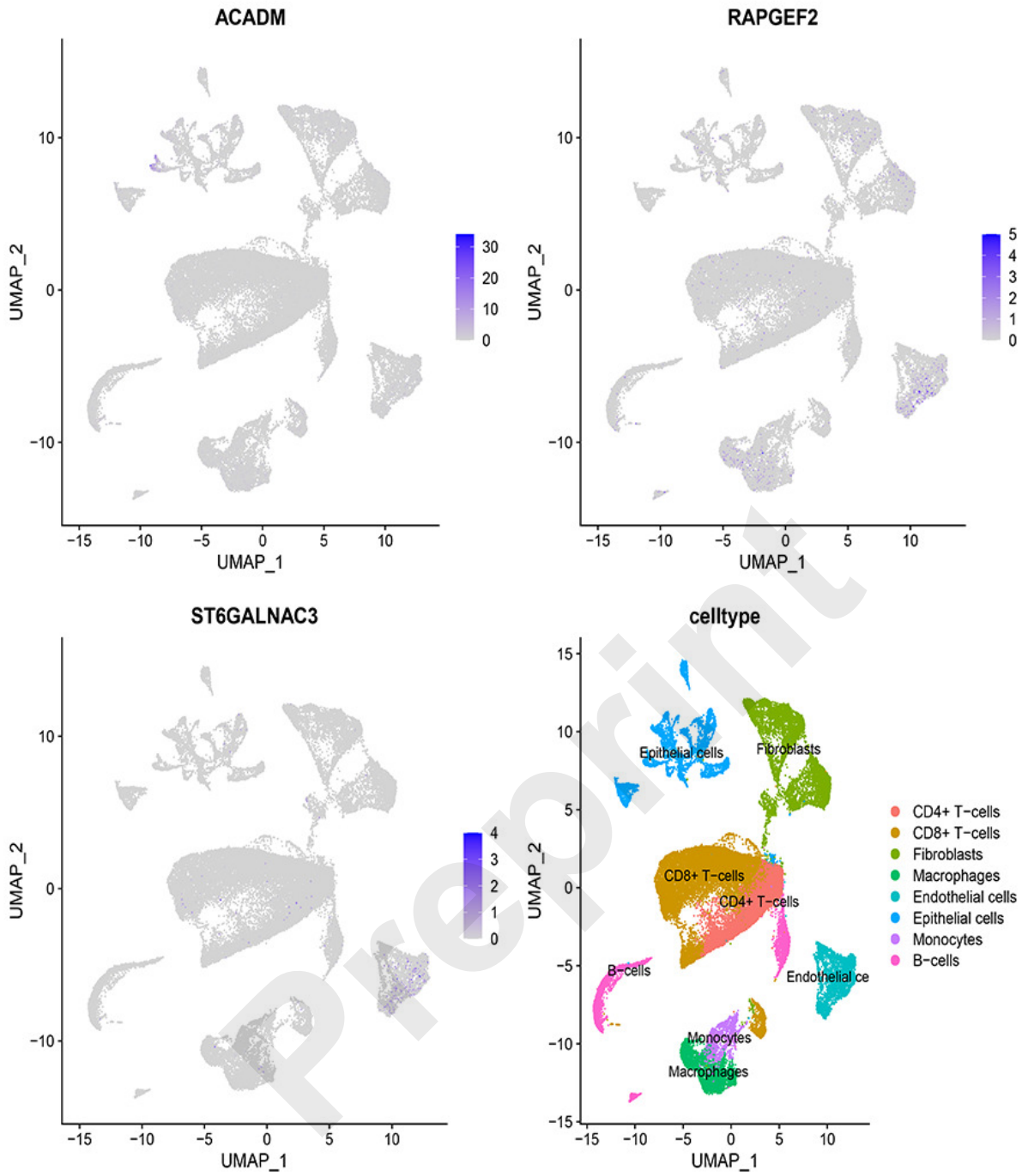

Fig. (6) Localization information and expression levels of biomarkers. (A) Location of biomarker genes on human chromosomes. (B) Subcellular localization analysis. (C-D) Expression analysis in TCGA-BRCA and GSE42568 datasets. (E) Expression analysis of highly variable genes. SNP, single nucleotide polymorphism; DEG, differentially expressed gene.

Fig. (7) The molecular regulatory network and drug prediction for biomarkers. (A) The TFs prediction for biomarkers. (B) The mRNA-miRNA-lncRNA regulatory network. (C) The results of drug prediction analysis. TFs, transcription factors.

Table Legends