Evaluating the potential causal relationship between carnitine-related metabolites and breast cancer: a Mendelian randomization and transcriptomic analysis

Feifeng Ran[1,2], Lilin Que[1,2], Lan Luo[2], Mei Gan[2], Rensheng Wang[1]*, Leifeng Liang[2]*

[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

Submitted: 16 October 2025; **Accepted:** 29 December 2025
Online publication: 3 April 2026

Arch Med Sci 2026; 22 (3): 1688–1705
DOI: https://doi.org/10.5114/aoms/216186

***Corresponding authors:**
Leifeng Liang
Department of Oncology
The Sixth Affiliated
Hospital of Guangxi
Medical University
The First People's
Hospital of Yulin
No. 495, Jiaoyu Middle Road
Yuzhou District, Yulin
Guangxi, 537000, China
Phone: 13657757927
E-mail: liangleifeng@stu.gxmu.edu.cn

Rensheng Wang
Department of
Radiation Oncology
The First Affiliated
Hospital of Guangxi
Medical University
No. 6, Shuangyong Road
Qingxiu District
Nanning, Guangxi
530000, China
Phone: 13807806008
E-mail: 13807806008@163.com

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 used 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 potential 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 potential 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 suggested that genetically predicted elevated levels of octanoylcarnitine and decanoylcarnitine were associated with 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 provides evidence supporting potential causal associations between octanoylcarnitine and decanoylcarnitine and BC risk. Integrative single-cell transcriptomics revealed six CRM-associated biomarkers and their dynamic expression within tumor microenvironments. Additional experimental and clinical studies are needed to validate these observations and clarify their biological and translational relevance.

Key words: female malignant tumors, fatty acid metabolic intermediates, genetic causal inference, tumor microenvironment regulation.

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 progression 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, transports 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 potential causal relationship between CRMs and BC, and characterize associated molecular pathways and cellular expression patterns. MR, which uses genetic variants as instrumental variables (IV), can reduce susceptibility to confounding and reverse causation, providing a useful approach for assessing potential causal relationships between exposures and outcomes [22, 23]. Based on this, we utilized large-scale genomic and transcriptomic data to investigate potential causal associations between CRMs and BC risk. Furthermore, by integrating single-cell RNA sequencing (scRNA-seq), cell communication analysis, and regulatory network construction, we characterized the cellular expression patterns and molecular features of CRM-associated genes within the TME and identified candidate genes for further investigation. This work provides genetic and transcriptomic evidence supporting the involvement of carnitine-related metabolic pathways in BC and highlights candidate genes for further mechanistic and translational investigation.

Material and methods

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. The representative histopatho-

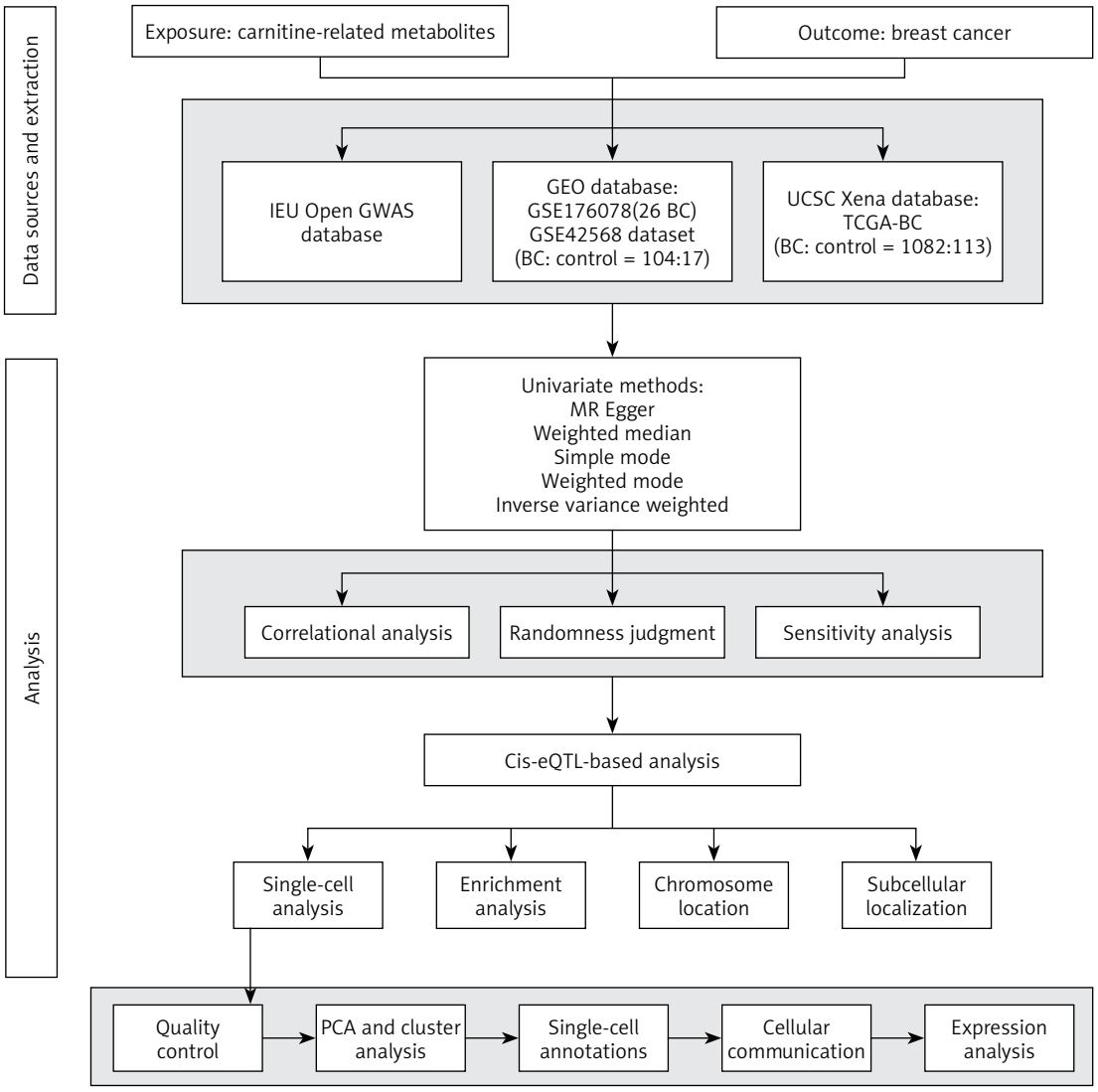

Figure 1. Flowchart of the study design

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.

logical slide images of breast cancer from the TCGA are provided in Supplementary Figure S1.

MR analysis

To investigate the potential 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 = (sample-size.exposure-2) × ((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 = (β^2)/(β^2 + SE^2 × N)). The following thresholds were used: $p < 0.05$ indicated a significant causal relationship between an expo-

sure factor and patient outcome; OR > 1 indicated risk factors; and OR < 1 indicated safety factors. MR results were visually represented using scatter plots, forest plots, and funnel plots.

Several sensitivity analyses were performed to evaluate the validity and stability of the causal relationships. Heterogeneity among IVs was assessed using the *mr_heterogeneity* function, with homogeneity indicated by *p*-values > 0.05. Horizontal pleiotropy was tested using the mr_pleiotropy_test function, where *p*-values > 0.05 suggested no directional pleiotropic effects. Additionally, leave-one-out analysis was conducted by systematically excluding individual SNPs to assess their impact on overall causal effect estimates. These analyses were implemented within the "TwoSampleMR" package [32].

Enrichment analysis

Genes associated with IVs corresponding to metabolites that showed evidence of associations with BC risk in the MR analysis were retrieved from the eQTLGen database and considered putative cis-eQTLs target genes. To investigate the potential biological functions and pathways associated with these genes in BC, an enrichment analysis was performed using the Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases via the "clusterProfiler" package [33] ($p < 0.05$).

Single-cell analysis

To characterize the cell-type-specific expression patterns and potential functional associations of CRM-related genes within the BC microenvironment, this study performed single-cell RNA sequencing analysis on BC tissues. Quality control for the GSE176078 dataset was conducted using the "Seurat" package [34], with criteria set to exclude cells containing fewer than 200 genes, genes detected in fewer than three cells, and cells with mitochondrial 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].

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

To investigate the signaling interaction networks among different cell subpopulations within the TME of BC, the "cellchat" software package was employed to analyze 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.

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.

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.

Statistical analysis

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

Results

MR analysis suggested that octanoylcarnitine and decanoylcarnitine are potential risk factors for BC

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

Sensitivity analyses supported the consistency of the MR results. No evidence of significant heterogeneity or horizontal pleiotropy was detected ($p > 0.05$) (Table II). Leave-one-out analysis indicated that no single SNP unduly influenced the overall estimates (Figures 2 G, H). In summary, the MR analysis suggested that higher genetically predicted levels of octanoylcarnitine and decanoylcarnitine are associated with increased BC risk. To explore shared instrumental variables between the two significant exposure factors, three overlapping SNPs were identified. 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.

Table I. Mendelian randomization analysis of associations between carnitine metabolites and breast cancer risk

Outcome	Exposure	nSNP	β	OR (95%CI)	*P*-value
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)	0.680
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.

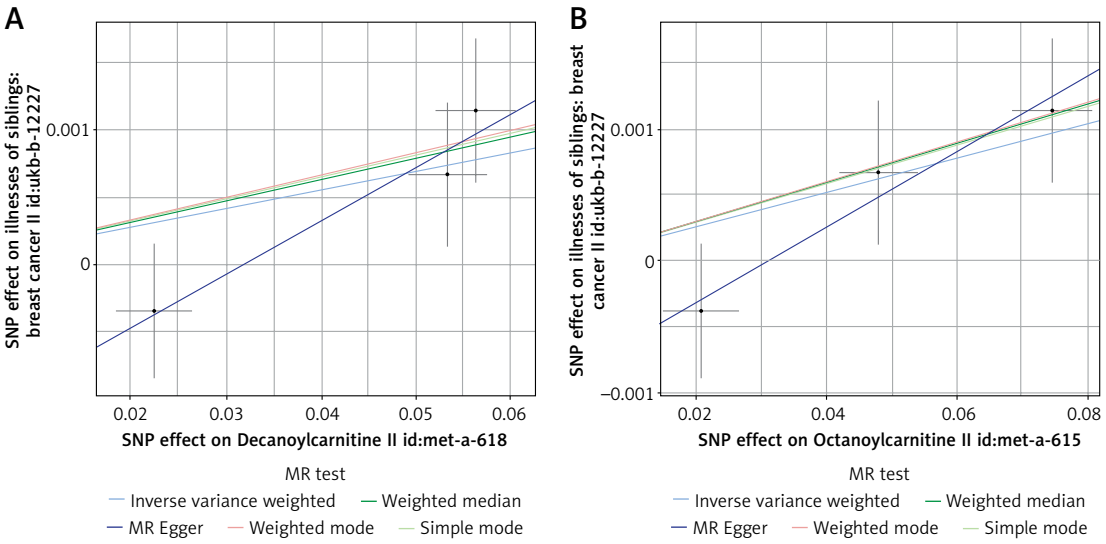

Figure 2. Scatter plots of the potential causal association between CRM and BC risk. **A** – decanoylcarnitine and BC risk, **B** – octanoylcarnitine and BC risk

SNP – single nucleotide polymorphism, MR – Mendelian randomization, OR – odds ratio.

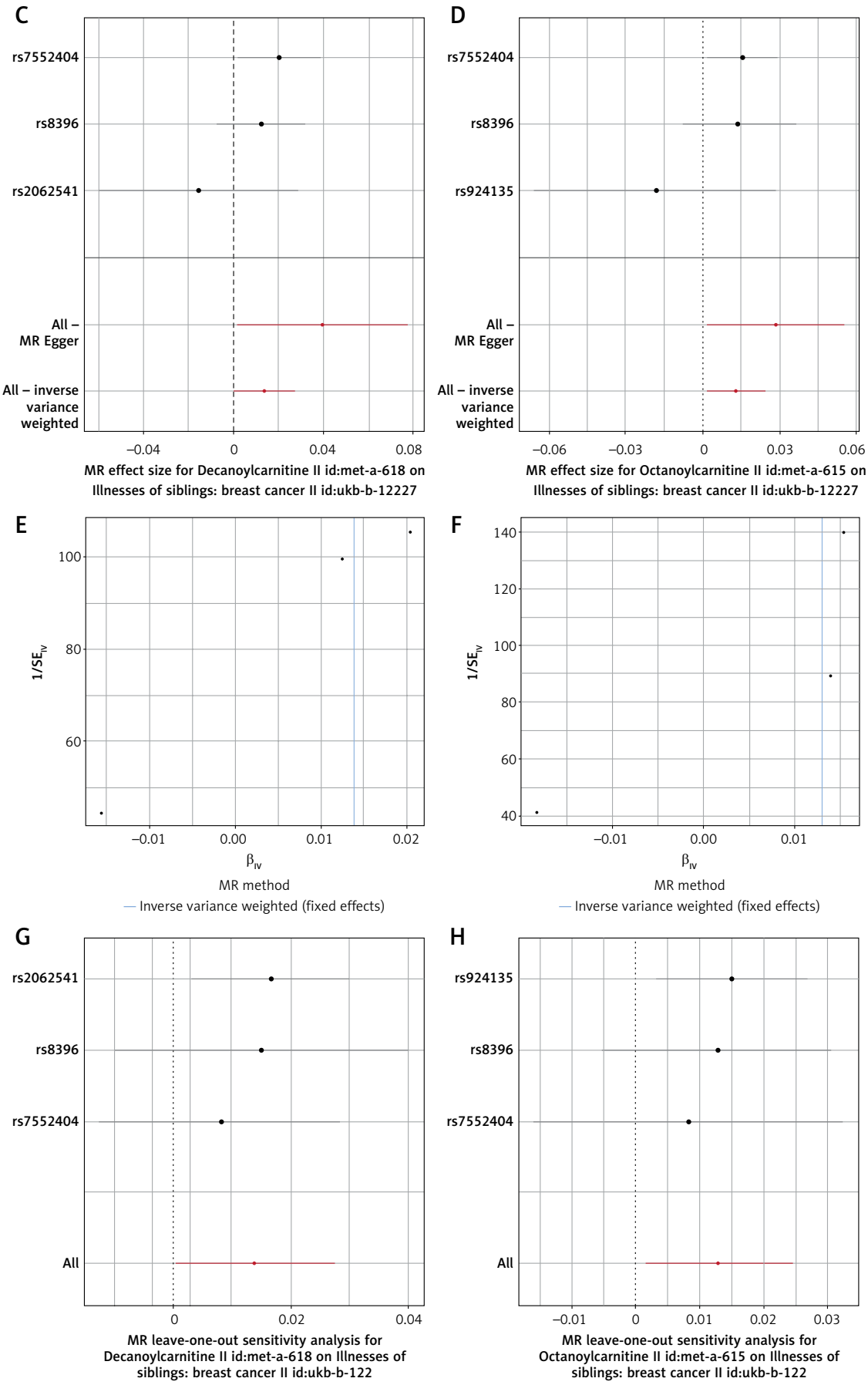

Figure 2. Cont. Forest plots for diagnostic efficacy of decanoylcarnitine (**C**) and octanoylcarnitine (**D**). **E**, **F** – Funnel plot for randomness judgment. MR leave-one-out sensitivity test for decanoylcarnitine (**G**) and octanoylcarnitine (**H**)

SNP – single nucleotide polymorphism, MR – Mendelian randomization, OR – odds ratio.

Table II. Results of sensitivity analysis

Exposure	Outcome	Heterogeneity test (IVW)			Horizontal pleiotropy test (MR-Egger)		
		Q	*P*-value	I^2	Intercept	SE	*P*-value
Octanoylcarnitine	BC	2.204126153	0.332	0	–0.001256772	0.000893921	0.393595639
Decanoylcarnitine	BC	1.790034097	0.049	9.26	–0.000893961	0.000706888	0.425941195

IVW – inverse variance weighted, BC – breast cancer.

Functional analysis of 12 candidate genes

To explore their potential 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) (Figure 3 A), while KEGG analysis

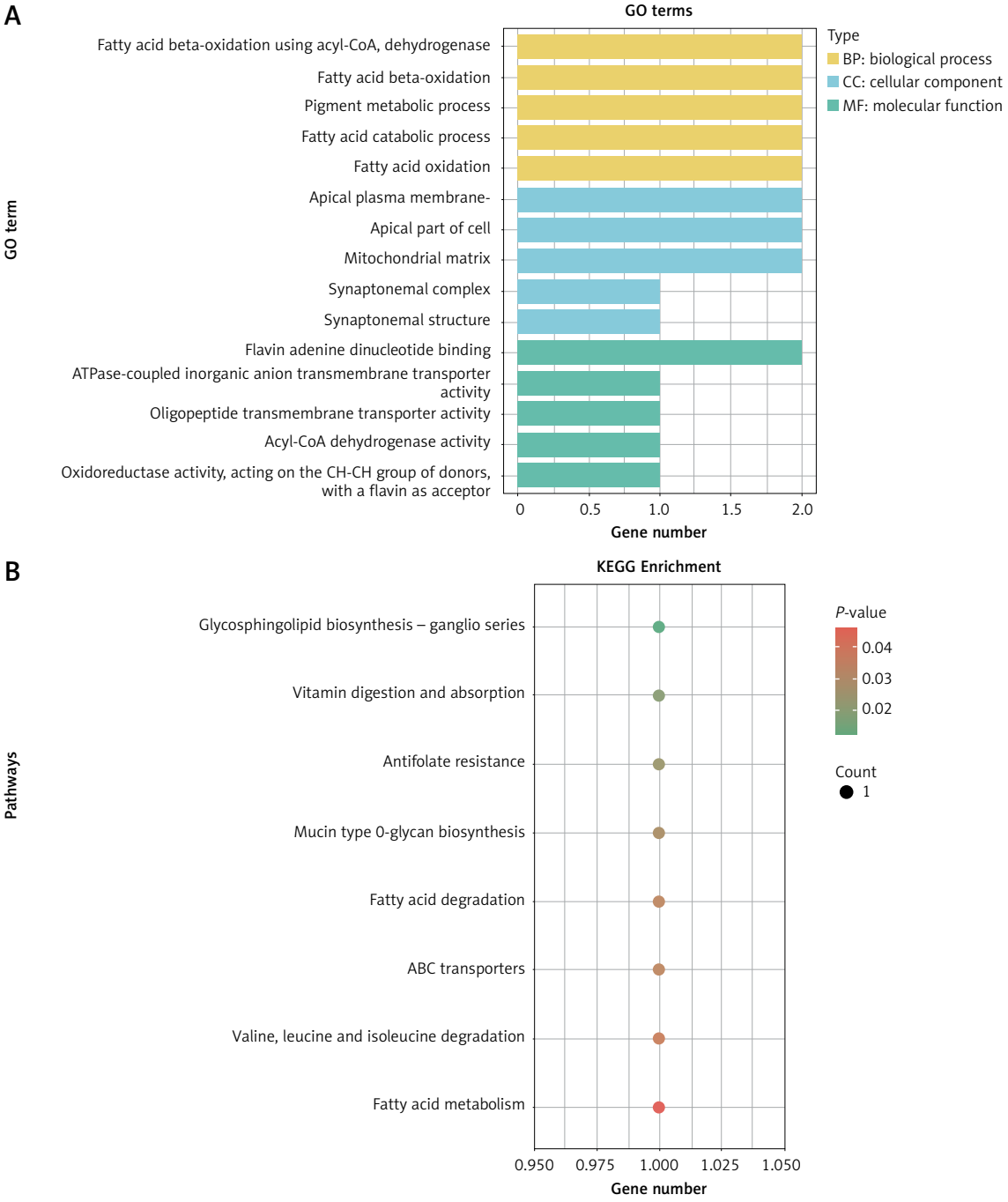

Figure 3. Functional enrichment analysis of 12 candidate genes. **A** – Results of Gene Ontology (GO) analysis. **B** – Bubble plot for Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis

indicated involvement in pathways such as glycosphingolipid biosynthesis-ganglio series, among others (Figure 3 B).

Eight cell subpopulations were identified and annotated

Quality control outcomes are shown in Supplementary Figure S2. Following standard preprocessing, 2000 genes with high variability were identified (Figure 4 A). PCA was performed, and the first 30 principal components were selected for further analysis (Figures 4 B–D). UMAP clustering analysis revealed 32 distinct cell clusters (Figure 4 E). Eight cell subpopulations were annotated, including epithelial cells, fibroblasts, and CD8+ T-cells (Figure 4 F). Cellular communication analysis showed enhanced communication between macrophages and macrophages and between macrophages and epithelial cells (Figures 4 G, H).

Expression and signaling pathways involving biomarkers in scRNA-seq

Macrophages exhibited the highest number and intensity of interactions with other annotated

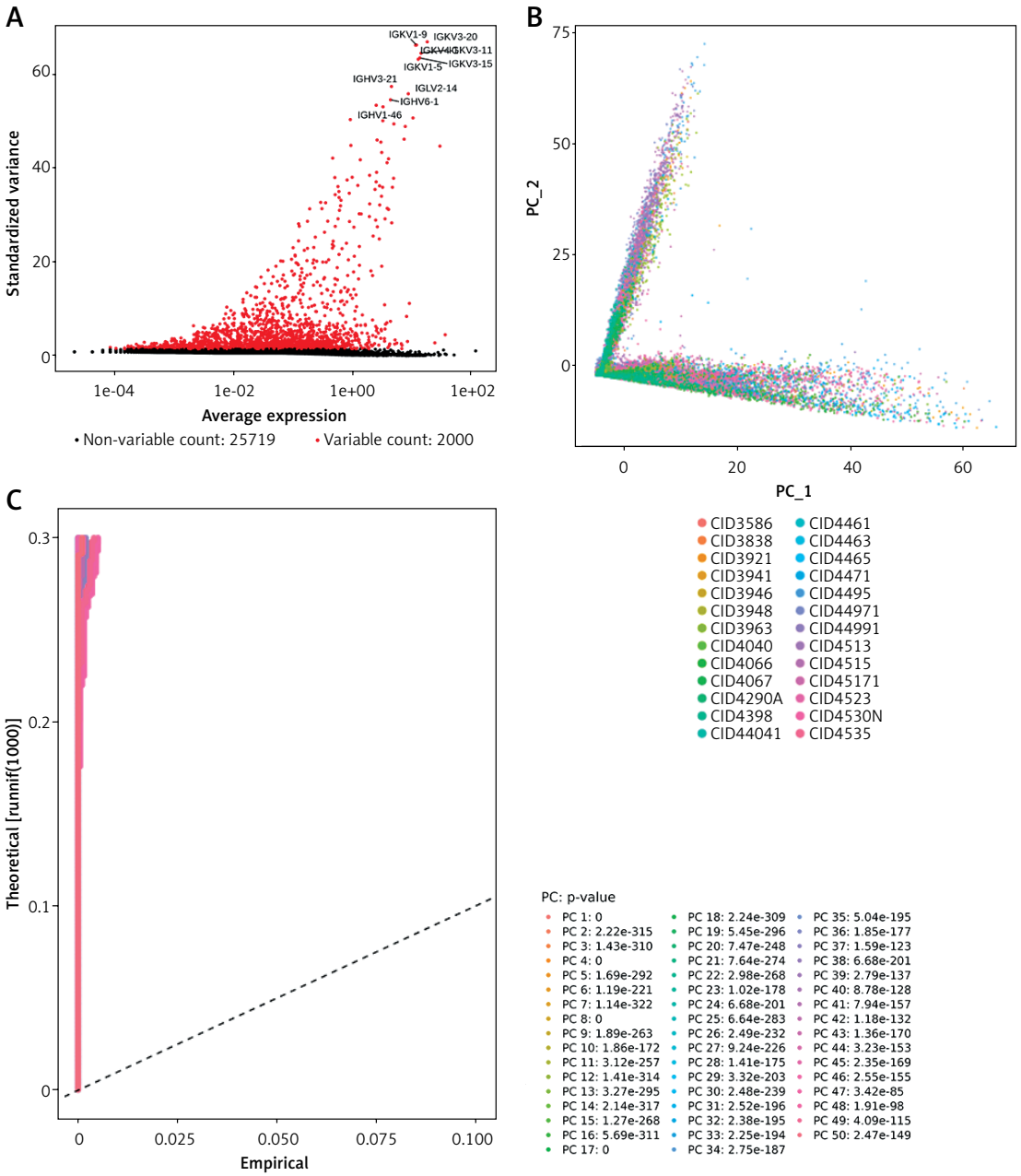

Figure 4. Single-cell RNA-seq of BC tissues. **A** – Results of screening highly variable genes. **B** – Dimension reduction by principal component analysis (PCA). **C** – Scatter plot of principal components

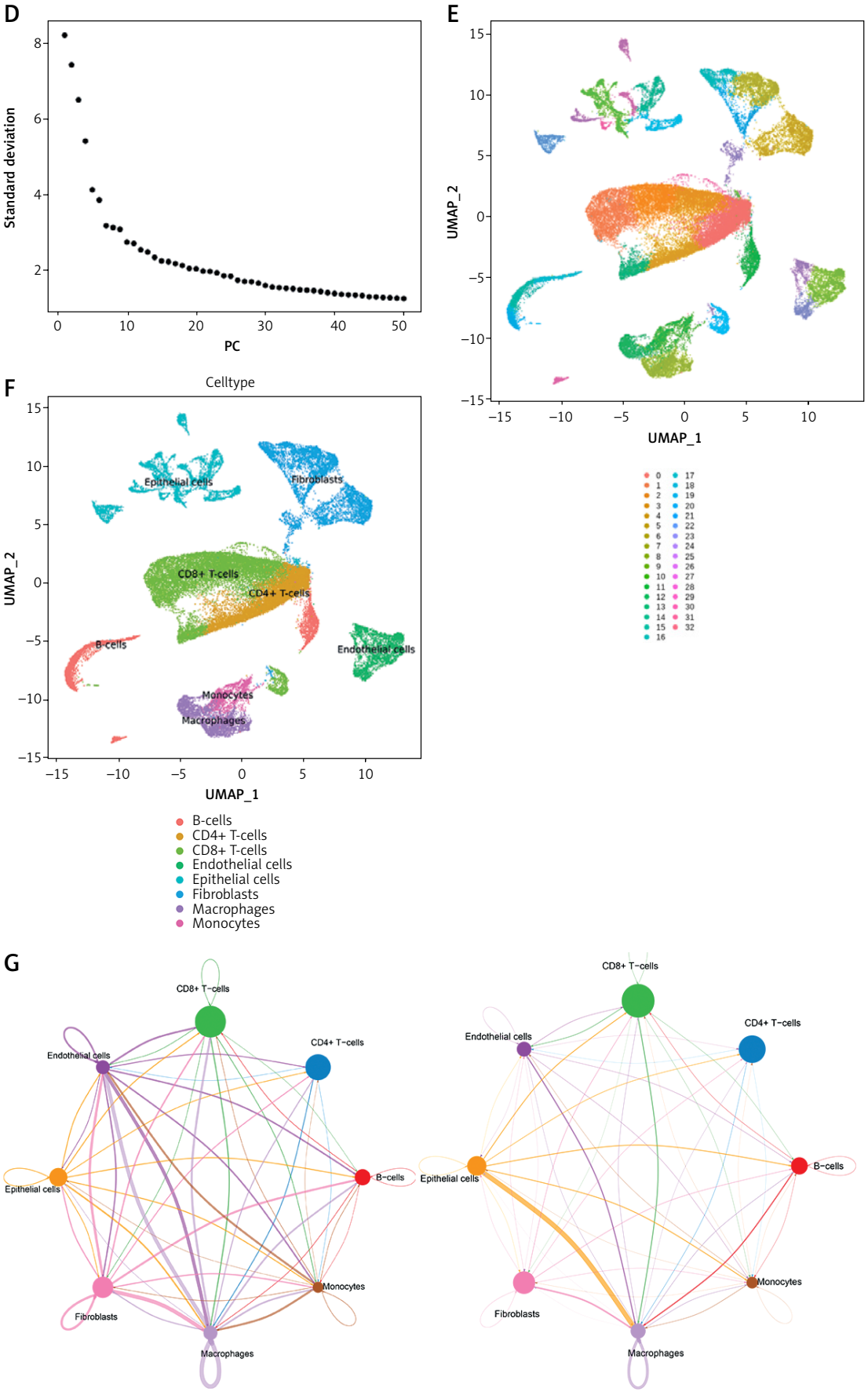

Figure 4. Cont. **E** – Uniform manifold approximation and projection (UMAP) cluster analysis. **F** – Single-cell annotations. **G** – Cellular communication

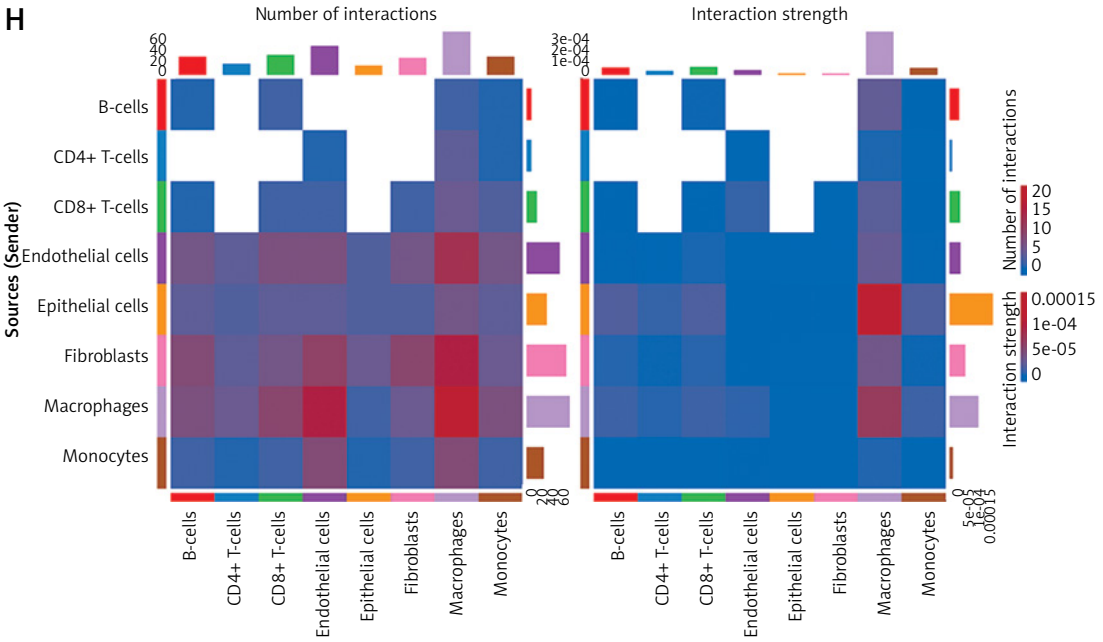

Figure 4. Cont. **H** – Heat map of cellular communication

cells during cell communication, thus designating them as key cells. A total of 7,890 scDEGs across various cell clusters were identified (Figure 5 A). KEGG pathway and GO biological process analyses of annotated cell subpopulations 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 (Figure 5 B). The overlap between the 7,890 scDEGs and the 12 SNP target genes identified six biomarkers: ACADM, FNIP2, RAPGEF2, RABGGTB, PPID, and ST6GALNAC3 (Figure 5 C). Following biomarker identification, their pseudo-temporal trajectories in macrophages were prioritized for further study. As shown in Figures 5 D, 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 (Figure 5 F).

Localization information and expression levels of biomarkers

The chromosomal locations of the biomarkers in human DNA are shown in Figure 6 A. 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 (Figure 6 B). Expression analysis of the TCGA-BC and GSE42568 datasets demonstrated significant downregulation of ACADM, RAPGEF2, and ST6GALNAC3 in BC samples (Figures 6 C, D). As presented in Figure 6 E, 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.

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 regulating three biomarkers (Figure 7 A).

Additionally, the mRNA-miRNA-lncRNA regulatory network revealed that 157 miRNAs (such as hsa-miR-105-5p, hsa-miR-106b-5p, hsa-miR-1252-5p, and hsa-miR-126-5p) and 68 lncRNAs (such as NEAT1 and XIST) interacted with the biomarkers (Supplementary Table SI, Figure 7 B). Drug prediction analysis identified 31 targeted drugs for ACADM (e.g., irbesartan, angiotensin A, and trv-120027) and 3 drugs for PPID (SCY-635, rencofilstat,

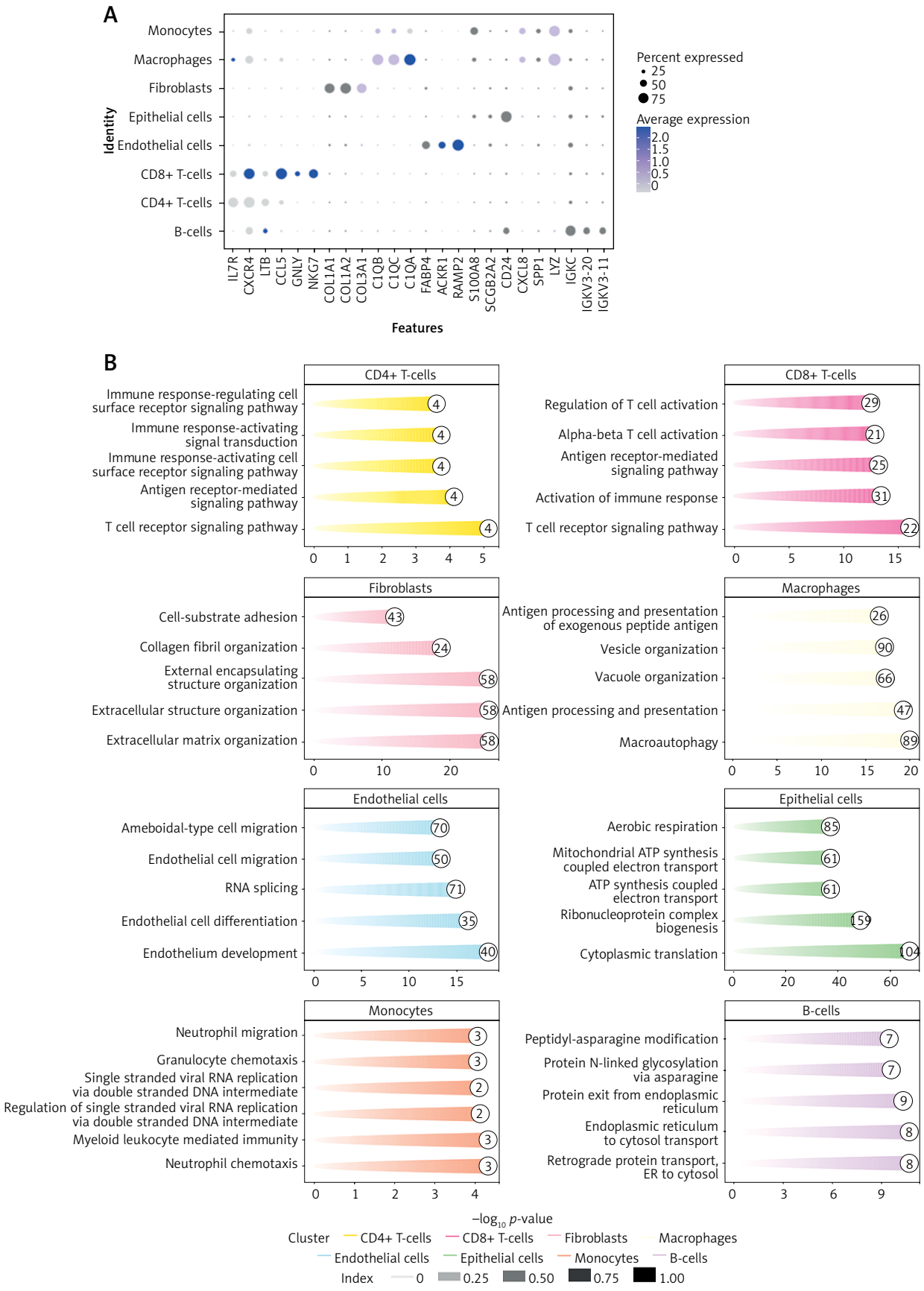

Figure 5. Expression and signaling pathways in differentially expressed genes (scRNA)-seq. **A** – Identification of single-cell scDEGs. **B** – Functional annotation of single-cell clusters

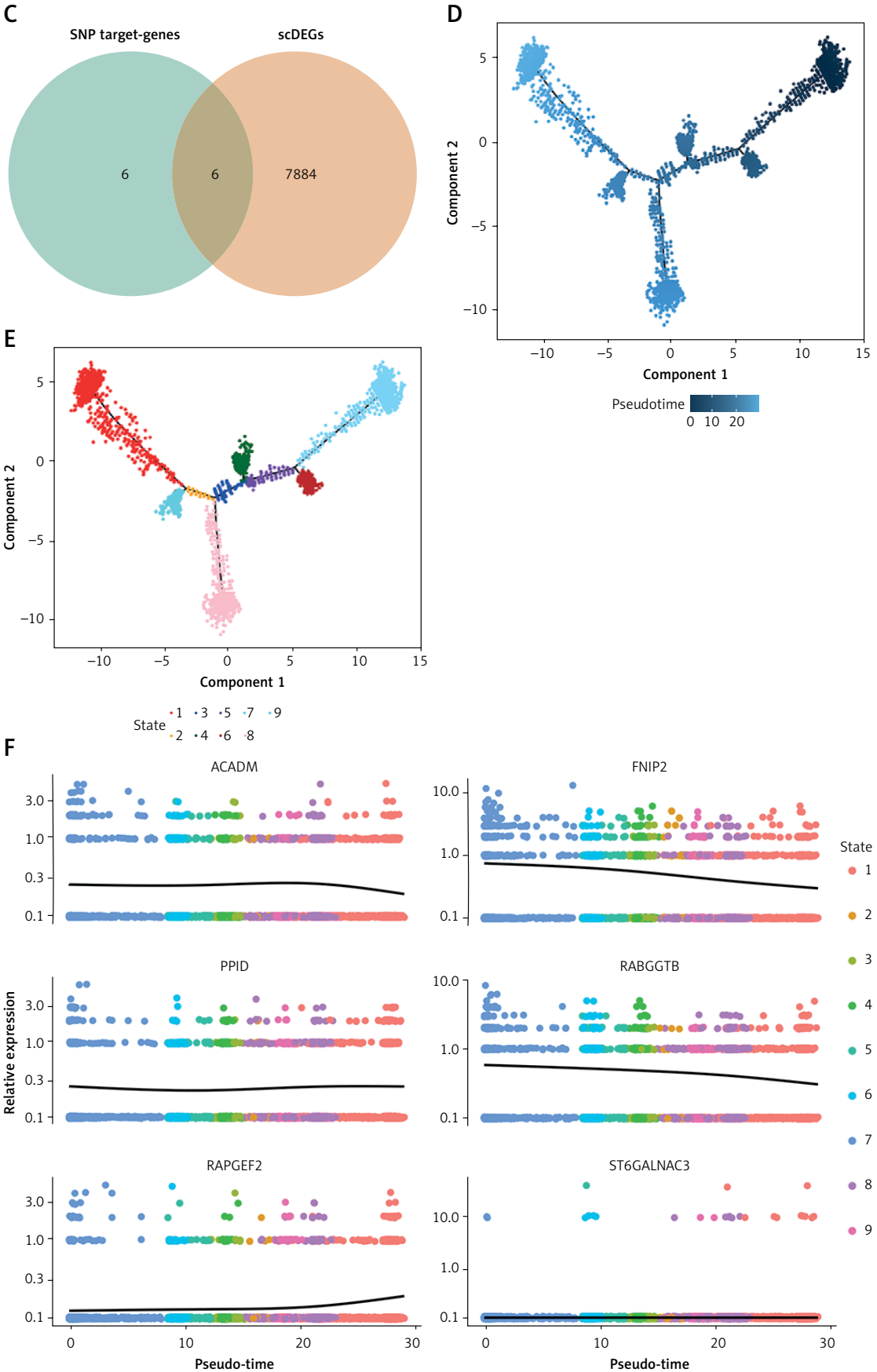

Figure 5. Cont. **C** – Identification of candidate biomarkers. **D** – Pseudotime analysis of macrophages. **E** – Differentiation state of macrophages. **F** – Expression levels of genes in development of macrophages

and cyclosporine) in BC (Figure 7 C, Supplementary Table SII). In summary, this integrative analysis provides insights into potential molecular pathways that may be associated with BC. The identification of candidate drug targets offers directions for future experimental and therapeutic investigation.

Discussion

BC exhibits significant diversity at both the cellular and molecular levels, which greatly impacts prognosis, particularly in subtypes such as TNBC. TNBC accounts for 15% of all BC cases and is associated with a relatively poor prognosis due to its high metastatic potential [42, 43]. A recent study highlighted that patients with hormone receptor-positive (HR+)/human epidermal growth factor receptor 2-negative (HER2-) BC exhibit distinct biological and clinical characteristics, suggesting the need for tailored therapeutic approaches for this group [44]. Therefore, identifying new biomarkers specific to these subtypes is critical. CRM have been proposed as potential prognostic biomarkers in multiple cancers [45, 46]. In the present study, MR analyses suggested potential associations between elevated levels of these metabolites and increased BC risk, supporting the need for further investigation of their potential clinical and biological relevance.

This study identified octanoylcarnitine and decanoylcarnitine as potential risk factors for BC. Previous studies have shown that octanoylcarnitine levels vary across BC stages and hormone receptor subtypes, with lower levels observed in T3-stage tumors and higher levels in ER/PR-positive tumors [47]. We hypothesize that elevated octanoylcarnitine may be associated with metabolic changes during BC development and progression.

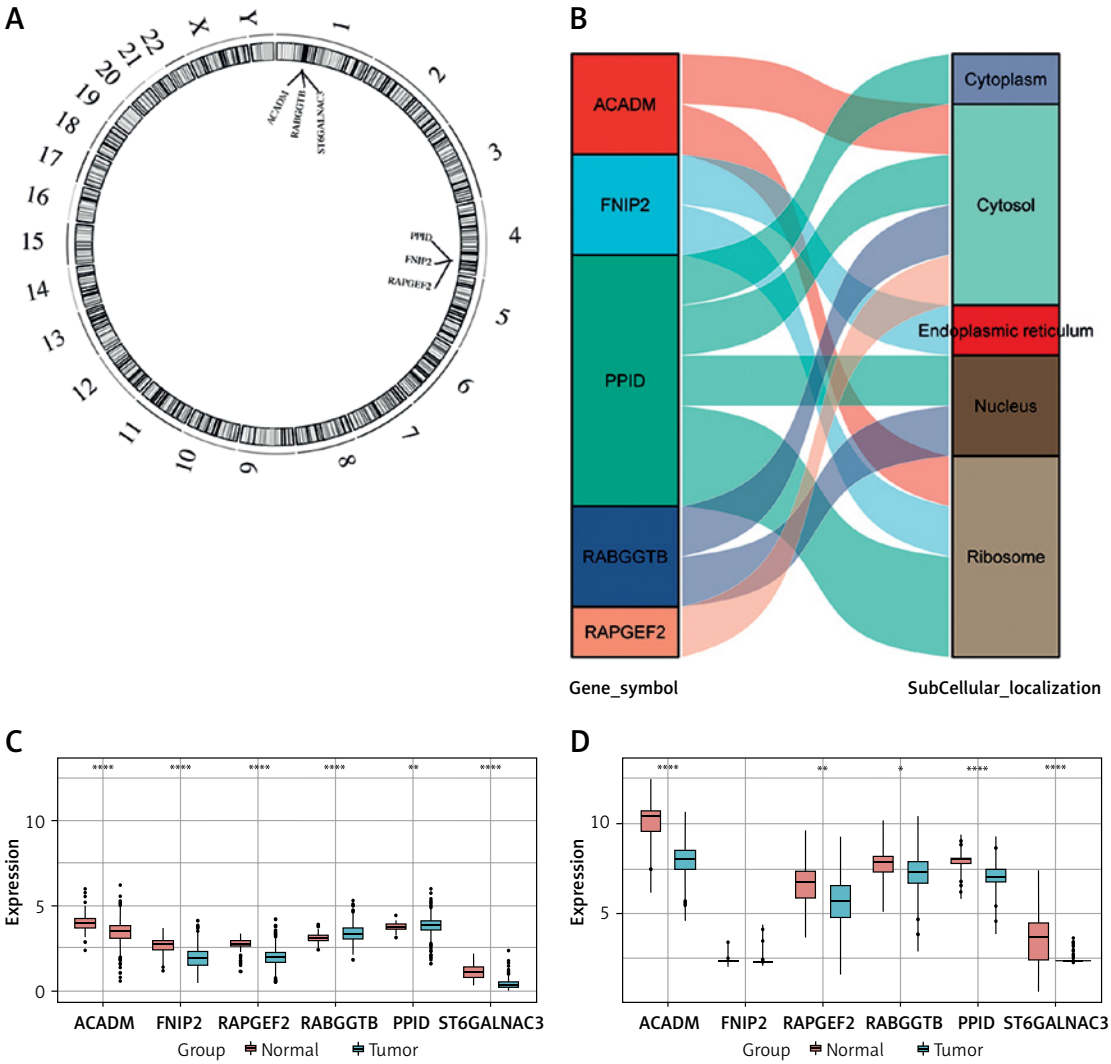

Figure 6. Localization information and expression levels of biomarkers. **A** – Location of biomarker genes on human chromosomes. **B** – Subcellular localization analysis. **C**, **D** – Expression analysis in TCGA-BRCA and GSE42568 datasets

SNP – single nucleotide polymorphism, DEG – differentially expressed gene.

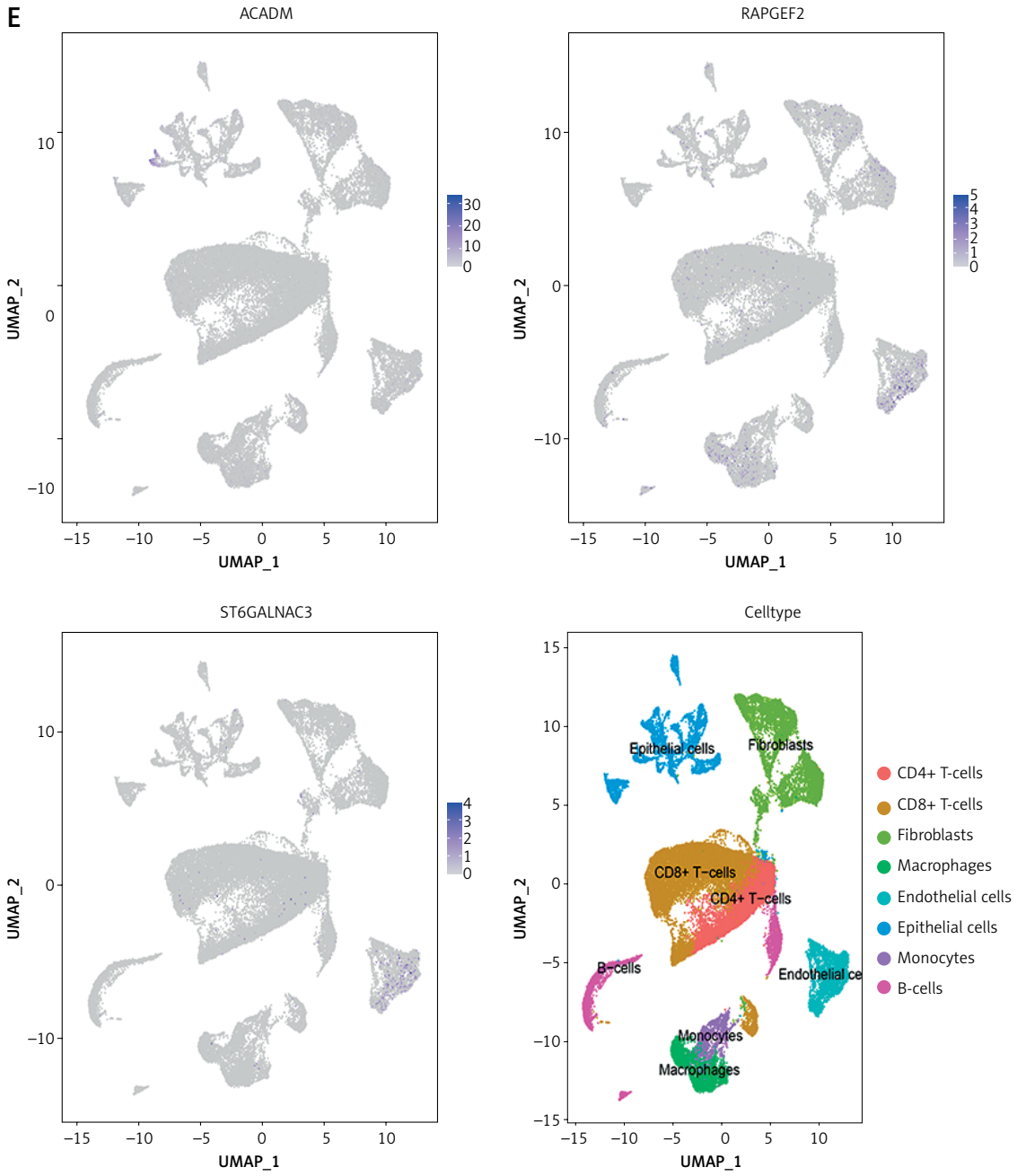

Figure 6. Cont. **E** – Expression analysis of highly variable genes

SNP – single nucleotide polymorphism, DEG – differentially expressed gene.

In advanced stages, its levels may be depleted and decline, possibly due to metabolic reprogramming, such as an 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 further characterize their dynamic

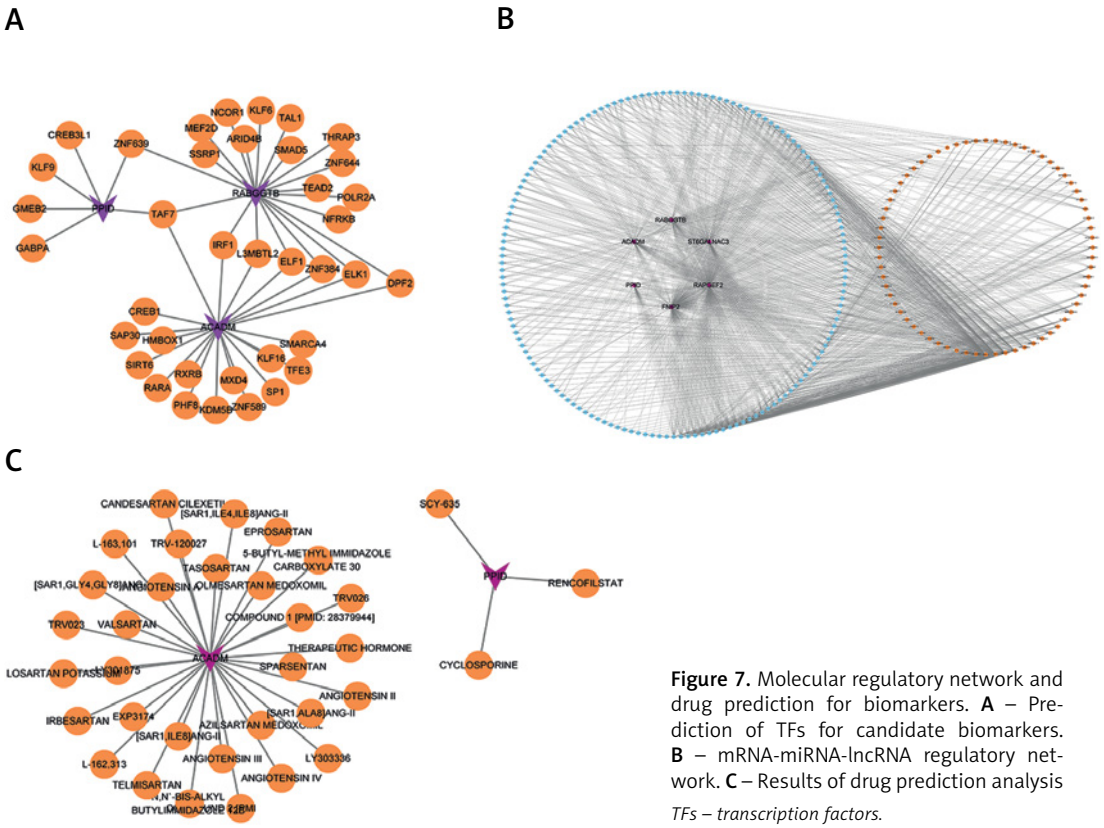

Figure 7. Molecular regulatory network and drug prediction for biomarkers. **A** – Prediction of TFs for candidate biomarkers. **B** – mRNA-miRNA-lncRNA regulatory network. **C** – Results of drug prediction analysis

TFs – transcription factors.

changes and mechanisms of action across different stages and subtypes of BC, and to evaluate 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. Previous 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 the cellular state while suppressing metastatic potential [54]. Collectively, these findings suggest that carnitine-related metabolic pathways may contribute to metabolic alterations associated with BC progression. Further research is needed to explore the mechanisms linking carnitine-related metabolites 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 [59–61]. 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 SIII). 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 [62–64]. 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.

In conclusion, this study provides evidence supporting potential associations between elevated levels of octanoylcarnitine and decanoylcarnitine and increased risk of BC. The genes annotated by the associated SNPs were enriched in FAO-related pathways, suggesting a potential association between carnitine-related metabolites and fatty acid metabolic processes in BC. 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, suggesting potential links between CRM-related metabolic pathways and cellular communication within the BC microenvironment. These findings identify candidate genes and pathways that may contribute to future studies of BC biology and metabolism-related therapeutic strategies. However, further functional validation and investigation in more diverse populations are warranted to elucidate the molecular and cellular processes through which these metabolites and their corresponding genes may influence BC progression.

Data statement

The datasets generated and analyzed 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/].

Acknowledgments

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

Funding

This study was supported by the National Natural Science Foundation of China (No. 82260627).

Ethical approval

Not applicable.

Conflict of interest

The authors declare no conflict of interest.

References

1. Giaquinto A, Sung H, Miller K, et al. Breast Cancer Statistics, 2022. CA Cancer J Clin 2022; 72: 524-41.
2. Elbasheer MMA, Dodwell D, Gathani T. Understanding global variation in breast cancer mortality. Br J Radiol 2025; 98: 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: 7537-54.
6. Engmann NJ, Golmakani MK, Miglioretti DL, Sprague BL, Kerlikowske K. Population-attributable risk proportion of clinical risk factors for breast cancer. JAMA Oncol 2017; 3: 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 EH, Kaushik NK. Functional impact of non-coding RNAs in high-grade breast carcinoma: Moving from resistance to clinical applications: a comprehensive review. Biochim Biophys Acta Rev Cancer 2023; 1878: 188915.

9. Miller K, Nogueira L, Devasia T, et al. Cancer treatment and survivorship statistics, 2022. CA Cancer J Clin 2022; 72: 409-36.
10. Yu J, Mu Q, Fung M, Xu X, Zhu L, Ho RJY. Challenges and opportunities in metastatic breast cancer treatments: Nano-drug combinations delivered preferentially to metastatic cells may enhance therapeutic response. Pharmacol Ther 2022; 236: 108108.
11. An J, Peng C, Tang H, Liu X, Peng F. New advances in the research of resistance to neoadjuvant chemotherapy in breast cancer. Int J Mol Sci 2021; 22: 9644.
12. Ward PS, Thompson CB. Metabolic reprogramming: a cancer hallmark even warburg did not anticipate. Cancer Cell 2012; 21: 297-308.
13. Thakur C, Qiu Y, Pawar A, Chen F. Epigenetic regulation of breast cancer metastasis. Cancer Metastasis Rev 2024; 43: 597-619.
14. Becker LM, O'Connell JT, Vo AP, Cain MP, Tampe D, Bizarro L, et al. Epigenetic reprogramming of cancer-associated fibroblasts deregulates glucose metabolism and facilitates progression of breast cancer. Cell Rep 2020; 31: 107701.
15. Wang Y, Dang H, Wang Q, et al. Mitochondrial pyruvate dehydrogenase phosphatase metabolism disorder in malignant tumors. Oncol Res 2025; 33: 1861-74.
16. Li S, Peng M, Tan S, et al. The roles and molecular mechanisms of non-coding RNA in cancer metabolic reprogramming. Cancer Cell Int 2024; 24: 37.
17. Chakraborty S, Kumar AS, Banerjee S. Lipids: driving forces in the underlying biology of carcinogenesis. ACS Pharmacol Transl Sci 2025; 8: 1891-918.
18. Castelli S, De Falco P, Ciccarone F, Desideri E, Ciriolo MR. Lipid catabolism and ROS in cancer: a bidirectional liaison. Cancers 2021; 13: 5484.
19. Andolino C, Cotul EK, Xianyu Z, et al. Fatty acid synthase-derived lipid stores support breast cancer metastasis. Cancer Metab 2025; 13: 35.
20. Longo N, Frigeni M, Pasquali M. Carnitine transport and fatty acid oxidation. Biochim Biophys Acta 2016; 1863: 2422-35.
21. Zhang J, Wu G, Zhu H, et al. Circulating carnitine levels and breast cancer: a matched retrospective case-control study. Front Oncol 2022; 12: 891619.
22. Smith GD, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? Int J Epidemiol 2003; 32: 1-22.
23. Ference BA. Mendelian randomization studies: using naturally randomized genetic data to fill evidence gaps. Curr Opin Lipidol 2015; 26: 566-71.
24. Wu SZ, Al-Eryani G, Roden DL, et al. A single-cell and spatially resolved atlas of human breast cancers. Nat Genet 2021; 53: 1334-47.
25. Clarke C, Madden SF, Doolan P, et al. Correlating transcriptional networks to breast cancer survival: a large-scale coexpression analysis. Carcinogenesis 2013; 34: 2300-8.
26. The Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. Nature 2012; 490: 61-70.
27. Hemani G, Zheng J, Elsworth B, et al. The MR-Base platform supports systematic causal inference across the human phenome. Elife 2018; 7: e34408
28. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. Int J Epidemiol 2015; 44: 512-25.
29. Bowden J, Davey Smith G, Haycock P, Burgess S. Consistent estimation in mendelian randomization with some invalid instruments using a weighted median estimator. Genet Epidemiol 2016; 40: 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: 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: 1985-98.
32. Bou Sleiman M, Jha P, Houtkooper R, Williams RW, Wang X, Auwerx J. The gene-regulatory footprint of aging highlights conserved central regulators. Cell Rep 2020; 32: 108203.
33. Rao S, Mondragón L, Pranjic B, et al. AIF-regulated oxidative phosphorylation supports lung cancer development. Cell Res 2019; 29: 579-91.
34. Hao Y, Hao S, Andersen-Nissen E, et al. Integrated analysis of multimodal single-cell data. Cell 2021; 184: 3573-87.e29.
35. Satija R, Farrell J, Gennert D, Schier A, Regev A. Spatial reconstruction of single-cell gene expression data. Nat Biotechnol 2015; 33: 495-502.
36. Aran D, Looney A, Liu L, et al. Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. Nat Immunol 2019; 20: 163-72.
37. Jin S, Guerrero-Juarez C, Zhang L, et al. Inference and analysis of cell-cell communication using CellChat. Nat Commun 2021; 12: 1088.
38. Xiaoning W, Guixin W, Qiaoqiao X, et al. Potential tumor-specific antigens and immune landscapes identification for mRNA vaccine in thyroid cancer. Front Oncol 2024; 14: 1480028.
39. Shukla V, Rao M, Zhang H, et al. ASXL3 is a novel pluripotency factor in human respiratory epithelial cells and a potential therapeutic target in small cell lung cancer. Cancer Res 2017; 77: 6267-81.
40. Yuanbin R, Katerina J K, Boris T, 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: e133.
41. Wilcoxon F. Individual comparisons of grouped data by ranking methods. J Econ Entomol 1946; 39: 269-70.
42. Allemani C, Matsuda T, Di Carlo V, 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: 1023-75.
43. Howard FM, Olopade OI. Epidemiology of triple-negative breast cancer: a review. Cancer J 2021; 27: 8-16.
44. Jin X, Zhou Y, Ma D, et al. Molecular classification of hormone receptor-positive HER2-negative breast cancer. Nat Genet 2023; 55: 1696-708.
45. Beloribi-Djefaflia S, Vasseur S, Guillaumond F. Lipid metabolic reprogramming in cancer cells. Oncogenesis 2016; 5: e189.
46. Lou S, Balluff B, Cleven AHG, Bovée J, McDonnell LA. Prognostic metabolite biomarkers for soft tissue sarcomas discovered by mass spectrometry imaging. J Am Soc Mass Spectrom 2017; 28: 376-83.
47. Park J, Shin Y, Kim TH, Kim DH, Lee A. Plasma metabolites as possible biomarkers for diagnosis of breast cancer. PLoS One 2019; 14: e0225129.

48. Wang Z, Wang C, Gao X. Exploring metabolic pathways in gastrointestinal cancer: new evidence of causality. Discov Oncol 2025; 16: 1777.
49. Huang Y, Lin W, Zheng X. Causal association between 637 human metabolites and ovarian cancer: a mendelian randomization study. BMC Genom 2024; 25: 97.
50. Breeur M, Ferrari P, Dossus L, et al. Pan-cancer analysis of pre-diagnostic blood metabolite concentrations in the European Prospective Investigation into Cancer and Nutrition. BMC Med 2022; 20: 351.
51. Gunchick V, Brown E, Liu J, et al. Morphomics, survival, and metabolites in patients with metastatic pancreatic cancer. JAMA Netw Open 2024; 7: e2440047.
52. Ye J, Wu S, Quan Q, et al. Fibroblast growth factor receptor 4 promotes triple-negative breast cancer progression via regulating fatty acid metabolism through the AKT/RYR2 signaling. Cancer Med 2024; 13: e70439.
53. Zeng F, Yao M, Wang Y, et al. Fatty acid β-oxidation promotes breast cancer stemness and metastasis via the miRNA-328-3p-CPT1A pathway. Cancer Gene Ther 2022; 29: 383-95.
54. Loo SY, Toh LP, Xie WH, et al. Fatty acid oxidation is a druggable gateway regulating cellular plasticity for driving metastasis in breast cancer. Sci Adv 2021; 7: eabh2443.
55. Potter S. Single-cell RNA sequencing for the study of development, physiology and disease. Nat Rev Nephrol 2018; 14: 479-92.
56. Lee AH, Happerfield LC, Bobrow LG, Millis RR. Angiogenesis and inflammation in invasive carcinoma of the breast. J Clin Pathol 1997; 50: 669-73.
57. Bingle L, Brown NJ, Lewis CE. The role of tumour-associated macrophages in tumour progression: implications for new anticancer therapies. J Pathol 2002; 196: 254-65.
58. Huang X, Cao J, Zu X. Tumor-associated macrophages: an important player in breast cancer progression. Thorac Cancer 2022; 13: 269-76.
59. Zhang T, Yu J, Cheng S, et al. Research Progress on the anticancer molecular mechanism of targets regulating cell autophagy. Pharmacology 2023; 108: 224-37.
60. Hu M, Fan JX, He ZY, Zeng J. The regulatory role of autophagy between TAMs and tumor cells. Cell Biochem Funct 2024; 42: e3984.
61. Hao Y, Duan F, Dong X, et al. Gold nanoparticle inhibits the tumor-associated macrophage M2 polarization by inhibiting m(6)A methylation-dependent ATG5/autophagy in prostate cancer. Anal Cell Pathol 2025; 2025: 6648632.
62. Zheqiong T, Yaru Z, Man Z, et al. Carnitine palmitoyl transferase 1A is a novel diagnostic and predictive biomarker for breast cancer. BMC Cancer 2021; 21: 409.
63. Huang D, Li T, Li X, et al. HIF-1-mediated suppression of acyl-CoA dehydrogenases and fatty acid oxidation is critical for cancer progression. Cell Rep 2014; 8: 1930-42.
64. Seok S, Kim Y, Byun S, et al. Fasting-induced JMJD3 histone demethylase epigenetically activates mitochondrial fatty acid β-oxidation. J Clin Invest 2018; 128: 3144-59.