Research

Integrated bulk and single-cell transcriptomic analysis unveiled a novel cuproptosis-related lipid metabolism gene molecular pattern and a risk index for predicting prognosis and antitumor drug sensitivity in breast cancer

Cheng Zeng[1] · Chang Xu[2] · Shuning Liu[1] · Yuanyi Wang[1] · Yuhan Wei[1] · Yalong Qi[1] · Yue Wang[3,4] · Jiani Wang[1] · Fei Ma[1]

Abstract

Breast cancer is the second most prevalent malignant tumor worldwide and is highly heterogeneous. Cuproptosis, a newly identified form of cell death, is intimately connected to lipid metabolism. This study investigated breast cancer heterogeneity through the lens of cuproptosis-related lipid metabolism genes (CLMGs), with the goal of predicting patient prognosis, immunotherapy efficacy, and sensitivity to anticancer drugs. By utilizing transcriptomic data from The Cancer Genome Atlas (TCGA) for breast cancer, we identified 682 CLMGs and applied the nonnegative matrix factorization (NMF) method to categorize breast cancer patients into four distinct clusters: cluster 1, "immune-cold and stroma-poor"; cluster 2, "immune-infiltrated"; cluster 3, "stroma-rich"; and cluster 4, "moderate infiltration". We subsequently developed a risk model based on CLMGs that incorporates ACSL1, ATP2B4, ATP7B, ENPP6, HSPH1, PIP4K2C, SRD5A3, and ULBP1. This model demonstrated excellent prognostic predictive performance in both the internal (testing and entire sets) and external (GSE20685 and Kaplan–Meier Plotter sets) validation sets. High-risk patients presented lower expression levels of immune checkpoint-related genes and lower immunophenoscores (IPSs), whereas low-risk patients presented higher CD8[+] T-cell infiltration levels and IPSs. Furthermore, the risk index was positively correlated with tumor cell stemness and could predict sensitivity to anticancer drugs. We also confirmed that SRD5A3 was highly expressed in breast cancer and participated in promoting the proliferation and migration of breast cancer cells. In conclusion, the results of this study provide new insights and strategies for assessing prognosis and implementing precision treatment for breast cancer through the lens of CLMGs.

Cheng Zeng and Chang Xu have contributed equally to this work.

✉ Jiani Wang, ncc_wangjiani@126.com;✉ Fei Ma, nccmafei@163.com; Cheng Zeng, zengchengstar@163.com; Chang Xu, xchang199@163.com; Shuning Liu, liushuninggg@163.com; Yuanyi Wang, wangyuanyi0709@163.com; Yuhan Wei, yuhan_wei@126.com; Yalong Qi, doctorlongxy@163.com; Yue Wang, wangyue0512@163.com | [1]Department of Medical Oncology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100021, China. [2]Department of Radiation Oncology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100021, China. [3]Department of Oncology, Wujin Hospital Affiliated With Jiangsu University, Changzhou 213000, Jiangsu Province, China. [4]Department of Oncology, Wujin Clinical College of Xuzhou Medical University, Changzhou 213000, Jiangsu Province, China.

1 Introduction

According to the epidemiological information from GLOBOCAN 2024, breast cancer became the 2nd most common malignancy worldwide in 2022 [1]. In clinical practice, factors such as patient age, cancer stage, histological grade, lymph node involvement, and molecular subtype of breast cancer remain pivotal for predicting patient prognosis and guiding treatment decisions [2, 3]. However, because of the significant diversity of breast cancer types, conventional clinico-pathological elements are insufficient to satisfy the requirements of the exact diagnosis and treatment of tumors. With the advancement and widespread use of sequencing innovations, more patients can obtain tumor expression profiles at a lower cost. Researchers can use these tumor expression profiles to identify prognosis-related genes, predict patient prognosis and guide treatment through specific gene signatures.

Cuproptosis, a recently identified type of programmed cell death, operates through copper binding to lipoacylated components of the tricarboxylic acid cycle. This interaction results in the aggregation of lipoacylated proteins and the suppression of iron−sulfur cluster proteins, culminating in cellular death [4, 5]. Research has revealed that cuproptosis-related genes, including SLC31A1, DLAT, PDHA1, and ATP7A, are strongly associated with an unfavorable prognosis in breast cancer patients, whereas elevated expression of ATP7B and LIPT1 indicates a more favorable prognosis [6–9]. Researchers have utilized cuproptosis-related gene signatures to predict OS, immune cell infiltration, and drug responsiveness in breast cancer patients [8, 10, 11]. However, the number of cuproptosis-related genes identified to date is limited, and their internal regulatory mechanisms remain poorly understood [6].

Breast cancer cells often exhibit abnormal lipid metabolism activities [12]. Metabolic changes not only provide essential energy and structural components for tumor cells but also promote tumor progression by affecting signaling pathways and gene expression [13]. Additionally, dysregulated lipid metabolism is intricately linked to immune evasion, drug resistance, and unfavorable prognosis in patients with breast cancer [14]. In tumors, lipid metabolism and copper metabolism abnormalities are often intertwined [4, 15]. Reactive oxygen species (ROS) produced by lipid metabolism can further increase the toxicity of copper ions, exacerbating oxidative damage to intracellular proteins and lipids and inducing cuproptosis [16, 17]. Copper ions can impact the activities of crucial enzymes such as fatty acid synthase and acetyl-CoA carboxylase [18, 19]. However, the intricate connection between cuproptosis and lipid metabolism remains ambiguous. Therefore, it is essential to systematically comprehend the interplay between cuproptosis genes and lipid metabolism genes, uncover cancer-specific vulnerabilities associated with cuproptosis and lipid metabolism, and enhance prognostic prediction and precision treatment for patients.

In this study, an interaction network between cuproptosis genes and lipid metabolism genes was constructed and breast cancer patients were clustered into four clusters on the basis of cuproptosis-related lipid metabolism genes (CLMGs). Furthermore, this study constructed a risk model using CLMGs to predict the prognosis of breast cancer patients and pinpoint those most likely to benefit from chemotherapy and immunotherapy. Finally, in vitro experiments confirmed that disrupting CLMG SRD5A3 could impede the advancement of breast cancer cells.

2 Materials and methods

2.1 Data collection and processing

The study design is illustrated in Fig. 1. Pan-cancer transcriptomic data from The Cancer Genome Atlas (TCGA) were obtained from the University of California, Santa Cruz (UCSC) Xena Browser (https://xenabrowser.net/datapages/) in December 2023. The dataset includes both normal and tumor samples across 24 cancer types: Bladder Urothelial Carcinoma (BLCA, normal = 19, tumor = 406), Breast Invasive Carcinoma (BRCA, normal = 111, tumor = 1039), Cervical Squamous Cell Carcinoma and Endocervical Adenocarcinoma (CESC, normal = 3, tumor = 304), Cholangiocarcinoma (CHOL, normal = 9, tumor = 35), Colon Adenocarcinoma (COAD, normal = 41, tumor = 456), Esophageal Carcinoma (ESCA, normal = 11, tumor = 162), Glioblastoma Multiforme (GBM, normal = 5, tumor = 161), Head and Neck Squamous Cell Carcinoma (HNSC, normal = 44, tumor = 502), Kidney Chromophobe (KICH, normal = 25, tumor = 65), Kidney Renal Clear Cell Carcinoma (KIRC, normal = 72, tumor = 532), Kidney Renal Papillary Cell Carcinoma (KIRP, normal = 32, tumor = 290), Liver

Fig. 1 Research flowchart
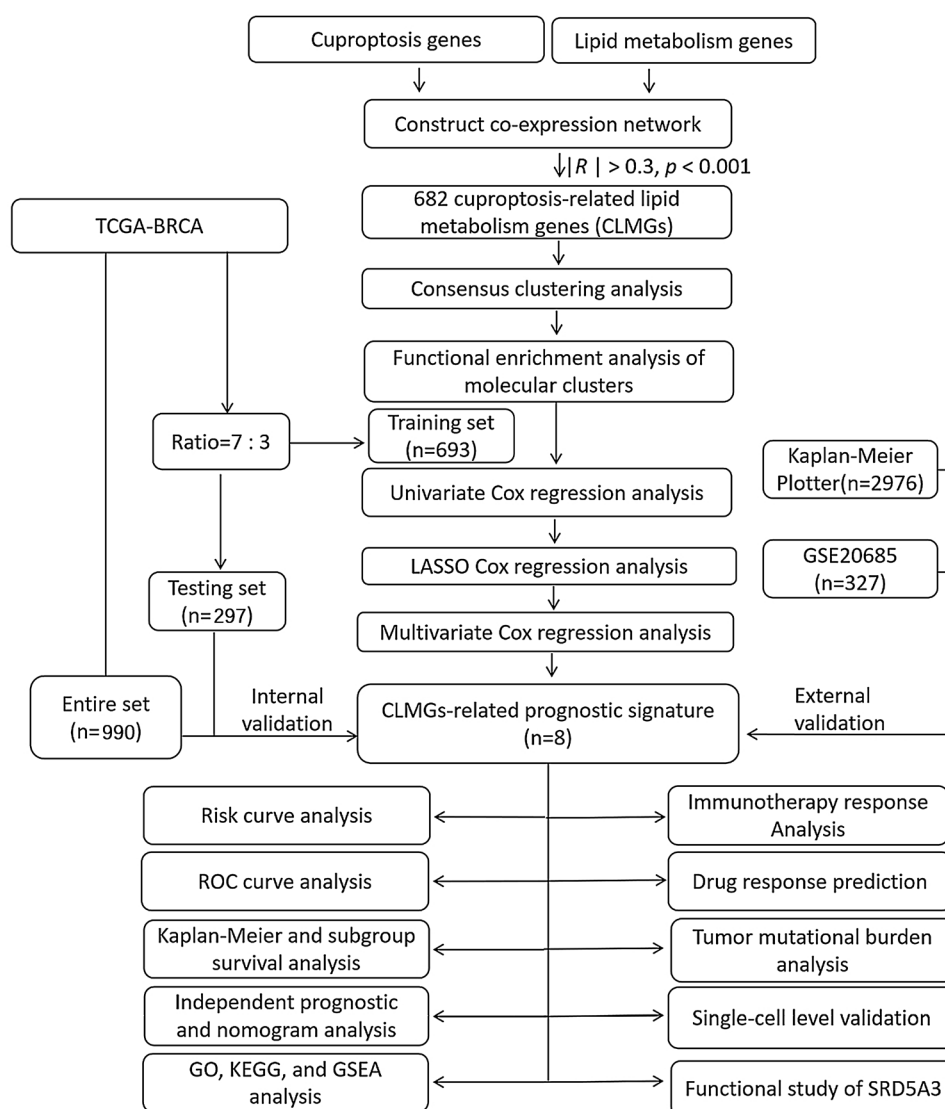
Hepatocellular Carcinoma (LIHC, normal = 50, tumor = 371), Lung Adenocarcinoma (LUAD, normal = 59, tumor = 516), Lung Squamous Cell Carcinoma (LUSC, normal = 49, tumor = 501), Pancreatic Adenocarcinoma (PAAD, normal = 4, tumor = 178), Pheochromocytoma and Paraganglioma (PCPG, normal = 3, tumor = 179), Prostate Adenocarcinoma (PRAD, normal = 52, tumor = 497), Rectum Adenocarcinoma (READ, normal = 10, tumor = 166), Sarcoma (SARC, normal = 2, tumor = 259), Skin Cutaneous Melanoma (SKCM, normal = 1, tumor = 469), Stomach Adenocarcinoma (STAD; normal = 32, tumor = 375), Thyroid Carcinoma (THCA, normal = 59, tumor = 504), Thymoma (THYM, normal = 2, tumor = 120), and Uterine Corpus Endometrial Carcinoma (UCEC, normal = 35, tumor = 545). Clinical information for the BRCA cohort was also downloaded from the UCSC Xena Browser. We excluded transcriptomic samples with missing clinical data, incomplete survival status, or survival durations of fewer than 30 days. After filtering, a total of 990 breast cancer samples with complete survival and clinical data were retained for the subsequent analyses. For external validation, we utilized the GSE20685 dataset, comprising transcriptomic data and clinical features of 327 breast cancer patients. This dataset was downloaded from the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/). Additionally, the Kaplan–Meier Plotter database (https://kmplot.com/analysis/index.php?p=service&cancer=breast_rnaseq_gse96058) was used to access survival data for 2976 breast cancer patients, forming another independent external validation cohort. A list of 27 cuproptosis-related genes were obtained from the FerrDb database in December 2023 (http://www.zhounan.org/ferrdb/current/) (Table S1). For lipid metabolism-related genes, we sourced 918 genes from the Harmonizome database in December 2023 (https://maayanlab.cloud/Harmonizome/gene_set/lipid/GeneRIF+Biological+Term+Annotations)

and 1034 genes from a previously published study [20]. After removing 212 overlapping genes, we finalized a set of 1740 unique lipid metabolism-related genes for further analysis (Table S2).

To address the differences between the microarray-based GSE20685 dataset and the RNA-seq-based TCGA-BRCA dataset, we first performed a $\log_2$ transformation ($\log_2(tcga + 1)$) on the TCGA-BRCA gene expression matrix to normalize the data and reduce its dynamic range. We then standardized the GSE20685 dataset using the normalizeBetweenArrays function to ensure consistent data distribution across samples and minimize technical biases. Finally, we applied the Combat function from the sva package (version 3.54.0) to adjust for batch effects using the Empirical Bayes method [21].

2.2 Consistent clustering analysis of CLMGs

This study first extracted transcriptome-level data for genes related to cuproptosis and lipid metabolism from the TCGA-BRCA dataset. Pearson correlation analysis was then carried out on these genes. The CLMGs were determined based on |R| > 0.3 and P < 0.001 [22, 23]. Breast cancer clusters associated with CLMGs were identified through consistent clustering analysis of mRNA expression levels using the nonnegative matrix factorization (NMF) method, implemented with the NMF package (version 0.28) [24].

2.3 Development of a risk model using CLMGs

The TCGA-BRCA cohort (n = 990) samples were randomly divided into a training set (n = 693) and a testing set (n = 297) in a 7:3 ratio using the createDataPartition function from the caret package (version 6.0.94). Initially, the CLMGs in the training set were subjected to univariate Cox regression analysis. To minimize overfitting, genes with a P-value < 0.05 from the univariate Cox regression analysis were further analyzed using least absolute shrinkage and selection operator (LASSO) regression. The analysis was performed with the glmnet package (version 4.1.8), utilizing default parameters and tenfold cross-validation. Finally, to refine the risk model, candidate genes identified by LASSO regression were subjected to multivariate Cox regression analysis to reduce their number. The risk index for each patient was calculated with reference to published papers [25, 26].

2.4 Validation of the risk model associated with CLMGs

The training set was utilized to assess the performance of the model, whereas the two internal validation sets (testing set and entire set) and the external validation set (GSE20685 and Kaplan−Meier Plotter sets) were employed to verify the model's effectiveness. First, a heatmap depicting the gene expression of the model genes and a scatter plot illustrating the survival status of breast cancer patients were generated. Kaplan−Meier survival analysis was then performed to compare survival between the high-risk and low-risk groups using the survminer package (version 0.5.0) and the survival package (version 3.7.0). Additionally, receiver operating characteristic (ROC) curve analysis was conducted to assess the sensitivity and specificity of the risk index in predicting 1-year, 3-year, and 5 year survival outcomes in breast cancer patients using the timeROC package (version 0.4).

2.5 Survival analysis of different pathological feature subgroups

To assess the performance of the risk model across various clinical subgroups, subgroup Kaplan−Meier survival analysis was performed. The subgroups analyzed included age > 65 years, age ≤ 65 years, T1-2 stage, T3-4 stage, N0 stage, N1-3 stage, M0 stage, M1 stage, stage I-II, stage III-IV, basal type, human epidermal growth factor receptor 2 (Her2)-positive type, luminal A type, and luminal B type.

2.6 Prognostic analysis and development of a nomogram

To identify prognostic factors in breast cancer patients, univariate and multivariate Cox regression analyses were conducted using the survival package (version 3.7.0), and a nomogram was constructed with the regplot package (version 1.1). Detailed analytical methods can be found in our published article [25].

2.7 Analysis of biological functional enrichment

To investigate potential differences in biological function between high-risk and low-risk groups of breast cancer patients, differential gene expression analysis was initially conducted using the limma package (version 3.62.1), with criteria of $|\log_2$-fold change (FC)$| > 1$ and adjusted $P < 0.05$ [27]. Gene Ontology (GO) analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis were subsequently performed on the differentially expressed genes (DEGs) using the clusterProfiler package (version 4.14.3) [28]. Furthermore, GSEA was conducted using h.all.v7.5.1.symbols.gmt as the annotated gene set. A total of 1000 permutations were performed, and the normalized enrichment score (NES) was computed on the Affymetrix chip platform [29]. A significance threshold of $P < 0.05$ and a false discovery rate (FDR)-adjusted $P < 0.25$ were applied for statistical significance [23].

2.8 Immune landscape analysis

To investigate the impact of the CLMG-related risk index on the immune microenvironment, Spearman correlation analysis was initially performed to assess the relationship between the risk index and immune cell infiltration levels using the CIBERSORT method [30]. Spearman correlation analysis was subsequently performed between the risk index and 47 immune checkpoint genes. To delve deeper into the disparities in the tumor microenvironment (TME) between high-risk and low-risk groups of breast cancer patients, the ESTIMATE method was used to analyze tumor purity between the two groups using the estimate package (version 1.0.13)[31]. Subsequently, single-sample GSEA (ssGSEA) methods were employed to assess variations in immune cell infiltration abundance and immune-related functions between the two subgroups using GSEABase package (version 1.68.0) [32]. A higher IPS indicates a greater response rate to anti-CTLA4 and anti-PD-1 antibodies [33]. Further analysis was conducted on the differences in IPS between the two subgroups.

2.9 Analysis of tumor mutational burden and drug sensitivity

To explore the role of the risk index in breast cancer treatment, correlations between the risk index and tumor mutational burden (TMB) and drug sensitivity were further explored. First, somatic mutation data in MAF format for breast cancer patients were downloaded from the TCGA database, followed by annotation analysis via the maftools package (version 2.22.0) [34]. Differences in TMB between the two subgroups were subsequently analyzed. Additionally, grouped Kaplan−Meier survival analysis was conducted, combining the risk index and TMB. To further explore the role of the risk index in breast cancer treatment, Spearman correlation analysis was performed between the RNA stemness index, the DNA stemness index, and the risk index. Next, differences in drug sensitivity between the two subgroups were analyzed on the basis of half maximal inhibitory concentration (IC50) values and the gene expression matrix of 198 drugs from the Genomics of Drug Sensitivity in Cancer (GDSC) database (https://www.cancerrxgene.org/).

2.10 Single-cell RNA sequencing analysis

TISCH2 database (http://tisch.comp-genomics.org/home/) includes 190 tumor single-cell datasets from 50 cancer types, comprising over 6 million cells [35]. For this study, the BRCA_GSE176078 dataset was utilized to analyze the expression levels of CLMG-related risk model genes from a single-cell perspective. The GSE176078 dataset comprises 26 samples and 89,471 cells. Breast cancer tissue samples were subjected to dimensionality reduction and clustering analysis via the uniform manifold approximation and projection (UMAP) method, followed by cell type annotation. The expression levels of the 8 CLMG-related risk model genes (ACSL1, ATP2B4, ATP7B, ENPP6, HSPH1, PIP4K2C, SRD5A3, and ULBP1) were subsequently analyzed at the single-cell level.

2.11 Cell lines, cell culture, and siRNA

The breast cancer cell lines MCF7 and MDA-MB-231 were purchased from the American Type Culture Collection (Manassas, VA, USA) [36]. MCF7 and MDA-MB-231 cells were cultured in high-glucose medium supplemented with 10% fetal bovine serum (FBS, Gibco, USA) and 1% penicillin–streptomycin (HyClone). All the cells were incubated at 37 °C in a 5% CO_2 incubator. SRD5A3 small interfering RNA (siRNA) was purchased from Beijing Tsingke Biotech (Beijing, China).

Fig. 2 Four CLMG-related clusters in breast cancer. **A** Pearson correlation analysis of 27 cuproptosis-related genes and 1,740 lipid metab- ▶
olism-related genes in breast cancer. **B** Determination of the optimal rank r value (r=2–10) for NMF. **C** Consistency clustering heatmap of
CLMGs based on the NMF method. **D** Heatmap showing the expression of prognostic-related genes in the consistency clustering analysis.
E Kaplan–Meier OS analysis of four CLMG-associated clusters in the TCGA-BRCA dataset of 990 breast cancer patients. **F** Tumor purity analy-
sis among the four CLMG-related clusters. **G** Immune cell infiltration abundance analysis among the four CLMG-related clusters based on
the MCPcounter algorithm. *CLMGs* cuproptosis-related lipid metabolism genes, *NMF* non-negative matrix factorization, *OS* overall survival;
*$P < 0.05$; **$P < 0.01$; ***$P < 0.001$; *ns* no significance

SRD5A3 siRNA and negative control (NC) siRNA were transfected into cells with Lipofectamine 3000 (Invitrogen, USA).
The sequences of the si-NC and si-SRD5A3 were as follows: si-NC (5′ > 3′)–UUCUCCGAACGUGUCACGUTT (sense) and ACG
UGACACGUUCGGAGAATT (antisense) [37]; si-SRD5A3 #1 (5′ > 3′)–GGCUAGUGGUGACAAAUGU (sense) and ACAUUUGUC
ACCACUAGCC (antisense); and si-SRD5A3 #2 (5′ > 3′)–GCUUUGGUGCCUUACUCAA (sense) and UUGAGUAAGGCACCA
AAGC (antisense).

2.12 Total RNA extraction and RT−qPCR

Total RNA was extracted from MCF7 and MDA-MB-231 cells using TRIzol reagent (Invitrogen™, USA) following the manu-
facturer's protocol. cDNA was synthesized from 2 μg of RNA using a cDNA reverse transcription kit (Thermo Fisher, USA).
The remaining steps for RT−qPCR was described in our published paper [36, 38]. The forward primer (5′ > 3′) for SRD5A3
was TGGCTGCACAGCTTACGAAG, and the reverse primer (5′ > 3′) was TCAGCACAGTTAGGCCAACAA. The forward primer
(5′ > 3′) for GAPDH was GTCAAGGCTGAGAACGGGAA, and the reverse primer (5′ > 3′) was AAATGAGCCCCAGCCTTCTC.
The relative expression levels were calculated using the $2^{-\Delta\Delta Ct}$ method.

2.13 MTS and cell scratch assays

The MTS assay was used to assess the cell proliferation capacity. Cells in the logarithmic growth phase from each group
were seeded into 96-well plates (Corning, USA), with 3000 cells in 100 μL of complete medium per well. At 0 h, 24 h, 48 h,
and 72 h, 10 μL of MTS (Promega, USA) was added to the 96-well plates. After incubation for 1 h in a 5% CO_2, 37 ℃ incuba-
tor, the optical density (OD) value was measured at a wavelength of 490 nm using a microplate reader. A scratch assay
was used to assess the cell migration capacity. Cells in the logarithmic growth phase from each group were seeded into
6-well plates (Corning, USA). The next day, when the cell confluence reached 90–100%, a sterile 100 μL pipette tip was
used to create a vertical scratch in the cell layer across the plane of the 6-well plate (Corning, USA). After being washed
with phosphate-buffered saline (PBS; Sangon Biotech, China), the cells were cultured in serum-free medium. The scratch
area was photographed under a microscope at 0 h and 24 h.

2.14 Statistical analysis

All statistical analyses and visualizations were performed using R software (version 4.1.2). To assess differences across
categories, the Wilcoxon test was used. Overall survival (OS) was compared across different categories via Kaplan−Meier
curves. Univariate and multivariate Cox regression analyses were employed to identify independent prognostic variables.
The predictive ability of the CLMG-related model was assessed using ROC curves and nomograms. All the experiments
were independently repeated three times, and the results are presented as the means ± standard deviations. Statistical
significance was defined as $P < 0.05$. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$.

3 Results

3.1 Four CLMG-related clusters in breast cancer

To explore the potential roles of CLMGs in breast cancer, we performed Pearson correlation analysis in the TCGA-BRCA
dataset, using pre-selected lists of 27 cuproptosis-related genes and 1740 lipid metabolism genes, and identified 682
CLMGs based on the criteria of |R| > 0.3 and $P < 0.001$ (Fig. 2A and Table S3). We subsequently conducted a consistent
clustering analysis of these CLMGs using the NMF method. The optimal number of clusters was identified at the point
where the cophenetic curve showed the most significant decline, revealing four clusters of breast cancer patients on the
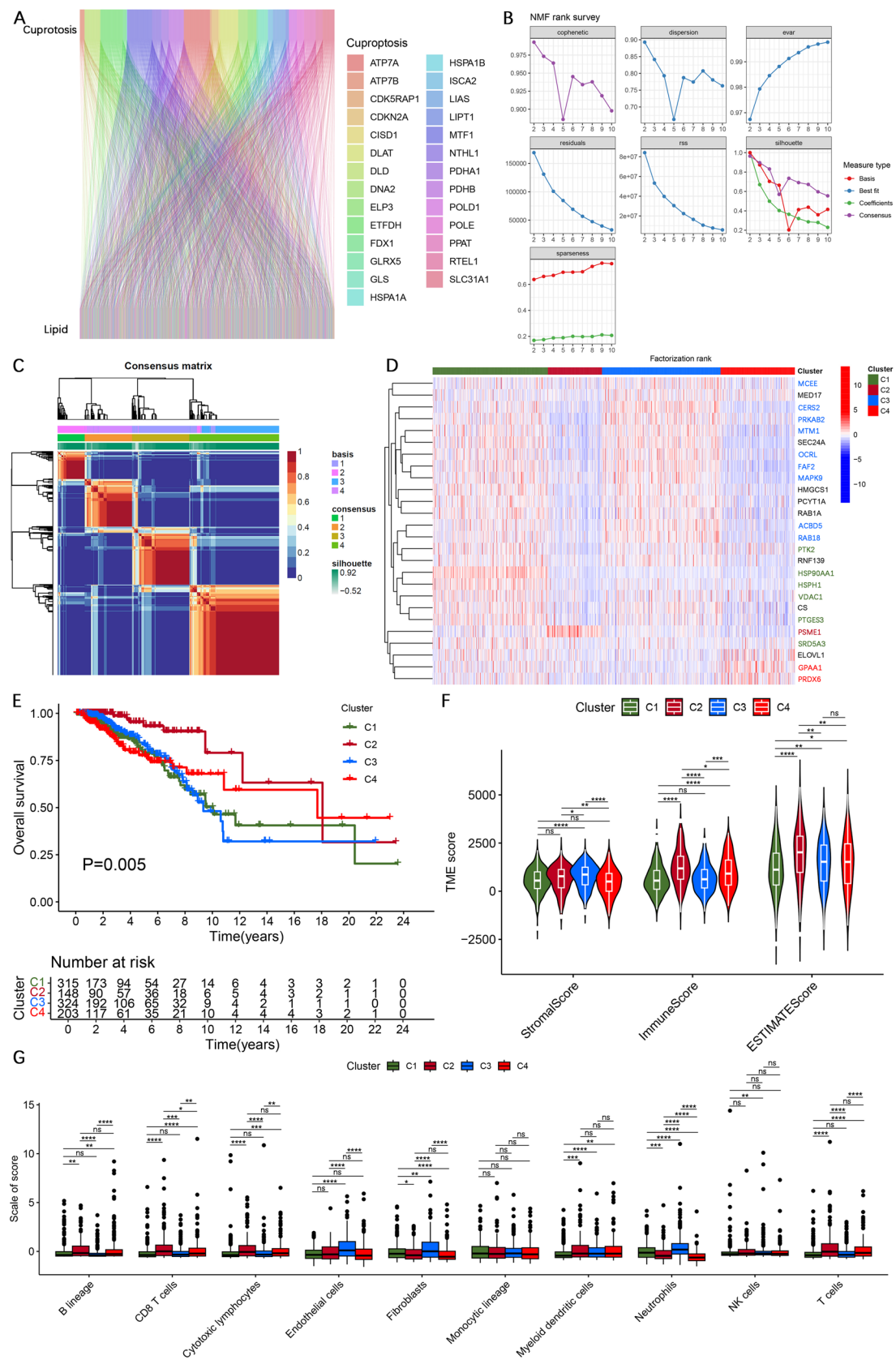
basis of CLMGs (Fig. 2B, C). The heatmap of gene expression revealed that genes associated with cluster 1 significantly overexpressed genes such as PTK2, HSP90AA1, HSPH1, VDAC1, PTGES3, and SRD5A3. Genes associated with cluster 2 showed significant overexpression of PSME1. Cluster 3 was characterized by the significant overexpression of genes including MCEE, CERS2, PRKAB2, MTM1, OCRL, FAF2, MAPK9, ACBD5, and RAB18. Lastly, genes associated with cluster 4 significantly overexpressed GPAA1 and PRDX6 (Fig. 2D and Fig. S1). The survival analysis of 990 TCGA-BRCA breast cancer patients revealed significant differences in overall survival (OS) across the four clusters, with patients in cluster 2 showing the best prognosis (Fig. 2E).

ESTIMATE was utilized to investigate variations in stromal and immune cell infiltration levels, as well as tumor purity, across the four clusters of breast cancer [31]. The results revealed that cluster 1 had the low ESTIMATE score and immune score, indicating the high tumor purity. Cluster 2 had the highest ESTIMATE score and immune score, suggesting high levels of immune cell infiltration. Cluster 3 exhibited the highest stromal score, indicating significant stromal cell infiltration in the tumor. Compared with clusters 2 and 1, cluster 4 had intermediate immune and ESTIMATE scores, reflecting moderate levels of immune cell infiltration and tumor purity (Fig. 2F). MCPcounter can be used to assess the abundance of 8 immune cell types and 2 stromal cell types in tissue samples [39]. We analyzed the abundance of immune cells and stromal cells in four CLMG-related clusters using MCPcounter. The results showed that cluster 2 had a high abundance of B lineage cells, natural killer (NK) cells, cytotoxic lymphocytes, T cells, CD8$^+$ T cells, and myeloid dendritic cells infiltration. Cluster 3 had a high abundance of neutrophils, endothelial cells and fibroblasts infiltration. Cluster 1 had a low abundance of immune and stromal cell infiltration, whereas cluster 4 had an intermediate level of immune cell infiltration between clusters 2 and 3 (Fig. 2G). The MCPcounter analysis results were consistent with the results of the ESTIMATE analysis. These findings suggest that cluster 1 is termed the "immune-cold and stroma-poor" type. Cluster 2 is considered the "immune-infiltrated" type. Cluster 3 is categorized as the "stroma-rich" type. Cluster 4 is referred to as the "moderate infiltration" type.

3.2 Development of a risk model based on CLMGs

To assess how CLMGs predict breast cancer prognosis, we developed a CLMG-related risk model to evaluate the risk levels in breast cancer patients. Initially, univariate Cox regression analysis was conducted on the differentially expressed CLMGs in the training set to identify 12 genes associated with breast cancer prognosis (Table S4). To prevent overfitting, tenfold cross-validation was applied in the LASSO regression analysis to select 11 candidate genes (Fig. 3A, B). Multivariate Cox regression analysis further refined the model, resulting in the identification of 8 key genes (ACSL1, ATP2B4, ATP7B, ENPP6, HSPH1, PIP4K2C, SRD5A3, and ULBP1) for construction of the risk model (Fig. 3C). The correlation heatmap revealed that the 8 model genes were significantly associated with most cuproptosis genes, with ATP7B being both a cuproptosis-related gene and a lipid metabolism-related gene (Fig. 3D). Survival analysis of the 8 core genes revealed that high expression levels of ACSL1, HSPH1, PIP4K2C, SRD5A3, and ULBP1 were associated with shorter OS in breast cancer patients, whereas elevated expression of ATP2B4, ATP7B, and ENPP6 was associated with longer OS (Fig. 3E–L). These findings suggest that ACSL1, HSPH1, PIP4K2C, SRD5A3, and ULBP1 may act as oncogenes and that ATP2B4, ATP7B, and ENPP6 may function as tumor suppressors.

3.3 Validation of the CLMG-related risk model

To assess the reliability of the risk model, the training set and entire set were used for internal validation, and the GSE20685 dataset was used for external validation. First, a heatmap of core gene expression was generated for the training set, which revealed that ACSL1, HSPH1, PIP4K2C, SRD5A3, and ULBP1 were predominantly expressed in the high-risk group and that ATP2B4, ATP7B, and ENPP6 were primarily expressed in the low-risk group (Fig. 4A). Similar findings were observed in both the internal validation set and the external validation set (Fig. 4E, I, M). The scatter plot revealed a greater proportion of deceased patients in the high-risk group (Fig. 4B, F, J, N). Kaplan – Meier survival analysis revealed that high-risk patients had shorter OS than did low-risk patients (Fig. 4C, G, K, O). ROC curve analysis revealed that the risk index had good accuracy in predicting the 1-year, 3-year, and 5-year survival of breast cancer patients (Fig. 4D, H, L, P). Additionally, in the Kaplan–Meier Plotter, high-risk breast cancer patients had shorter OS, progression-free survival (PFS), and distant metastasis free survival (DMFS) compared to those in the low-risk group (Fig. 4Q). Overall, the CLMG-related risk index effectively quantified the risk level of patients with breast cancer and predicted their prognosis.
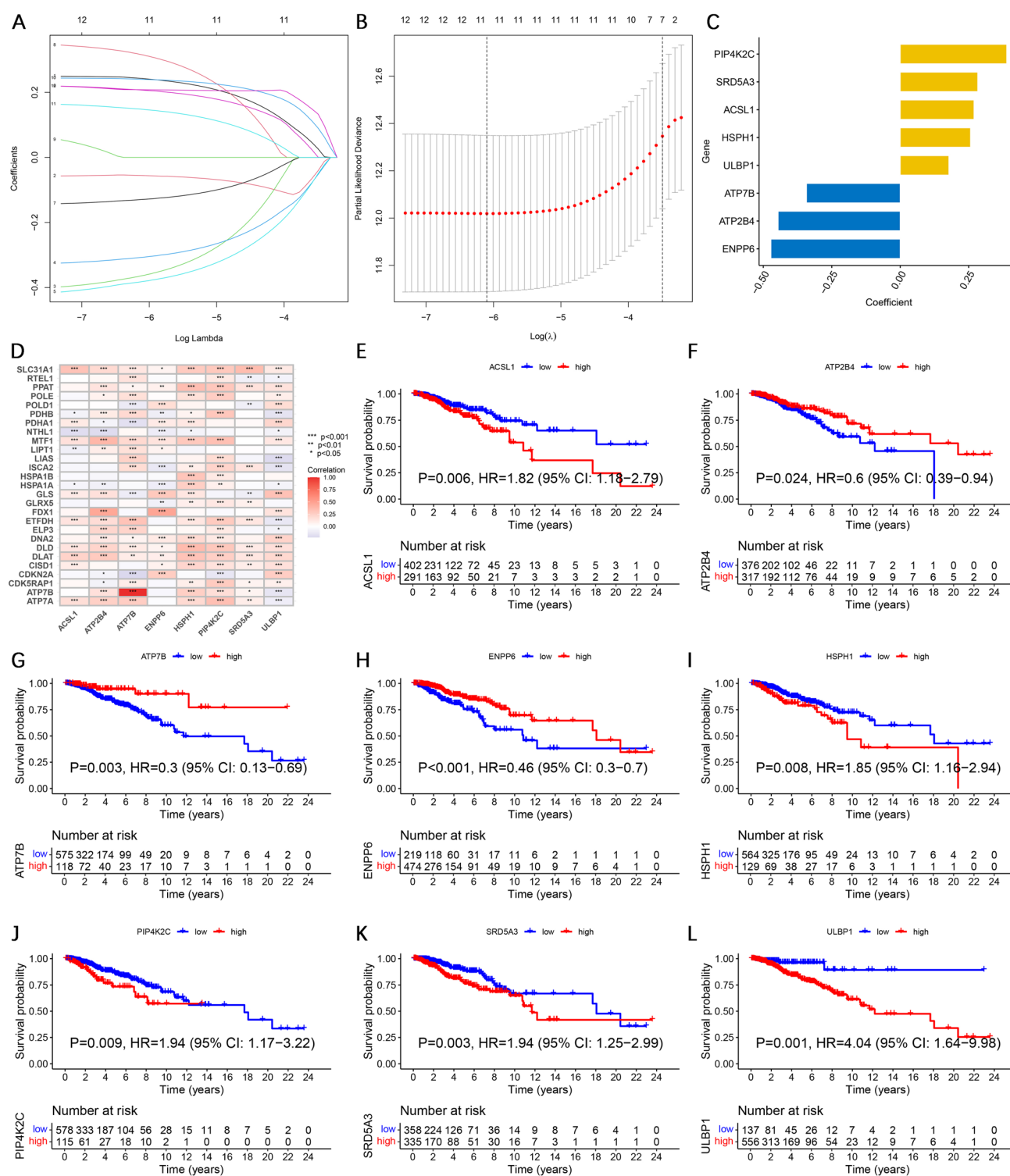
Fig. 3 Construction of the CLMG-related risk model. **A** LASSO regression analysis variable trajectory plot. Each line represents a gene. **B** The likelihood deviation values corresponding to each lambda value in the LASSO regression analysis. **C** Selection of 8 CLMGs through multivariate Cox regression analysis. **D** Heatmap showing the correlation between the 8 risk model genes and cuproptosis genes. **E–L** Survival prognosis analysis of the risk model genes in the training set, with groups stratified based on the optimal cutoff value. *CLMGs* cuproptosis-related lipid metabolism genes, *LASSO* least absolute shrinkage and selection operator; *P < 0.05; **P < 0.01; ***P < 0.001
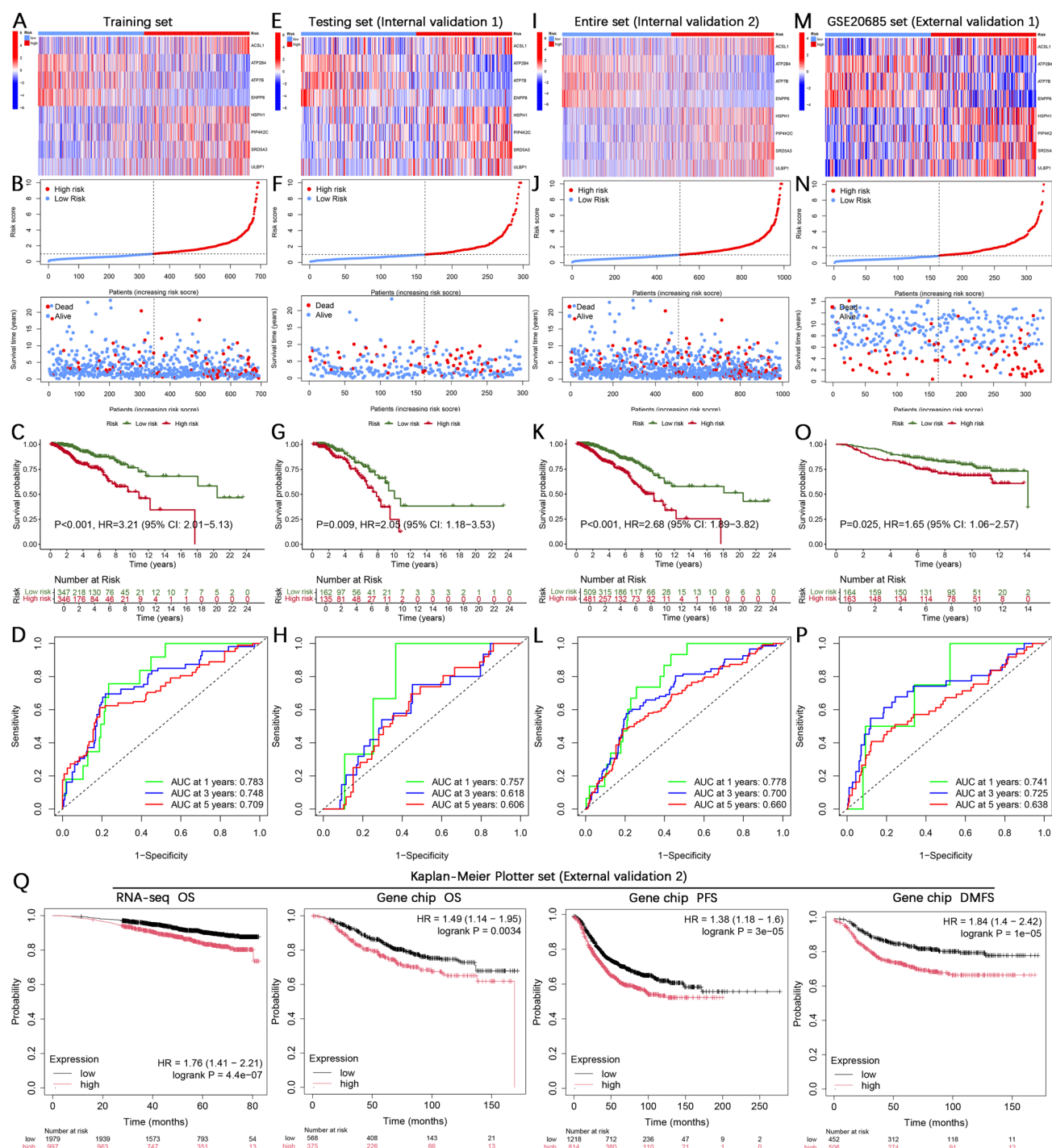
Fig. 4 Validation of the CLMG-related risk model. **A–D** Heatmaps of model gene expression levels, scatter plots of risk scores and survival status, Kaplan–Meier OS analyses between high- and low-risk groups, and time-dependent ROC curves of risk indices at 1, 3, and 5 years in the training set. The testing set and entire set were used as internal validation sets, while the GSE20685 set served as an external validation set. **E–P** The same analyses were performed across two internal validation sets (testing and entire sets) and the GSE20685 external validation set. **Q** The prognostic efficacy of CLMG-related model was validated in Kaplan–Meier plotter. *OS* overall survival, *PFS* progression-free survival, *DMFS* distant metastasis free survival, *ROC* receiver operating characteristic

3.4 Subgroup survival analysis on the basis of clinical and pathological characteristics

To further assess the reliability of the risk model across various clinical and pathological subgroups, a subgroup survival analysis was conducted in breast cancer patients. The results revealed that, except for those in the basal and Her2 subgroups, breast cancer patients in the high-risk group had shorter OS than did those in the low-risk group (Fig. S2).

3.5 Prognostic analysis and nomogram construction

Univariate and multivariate Cox regression analyses revealed that age and the risk index were independent prognostic factors for breast cancer patients (Fig. 5A, B). To more accurately predict the survival time of breast cancer patients, we constructed a nomogram incorporating clinical and pathological features along with the risk index and performed a calibration curve analysis, which demonstrated good performance (Fig. 5C, D). Additionally, multivariate ROC curve analysis revealed that the comprehensive nomogram score had the highest area under the curve (AUC) values for predicting the 1-year, 3-year, and 5 year survival status of breast cancer patients, with AUCs of 0.868, 0.773, and 0.739, respectively, further improving the efficacy of the risk index in predicting patient survival (Fig. 5E–G).

3.6 Biological function analysis

To further investigate the biological function differences between the high- and low-risk groups identified by the CLMG-related risk model, we conducted differential expression analysis in breast cancer patients. Genes were considered differentially expressed if they met the criteria of $|\log_2 FC| > 1$ and a corrected $P < 0.05$, resulting in the identification of 1,782 differentially expressed genes (Table S5). GO analysis of these DEGs revealed associations with receptor ligand activity, the T-cell receptor complex, the regulation of angiogenesis, and the regulation of epithelial cell proliferation (Fig. 6A), and KEGG analysis revealed associations with the PI3K-AKT pathway, the MAPK pathway, the cell cycle, the regulation of lipolysis in adipocytes, and the B-cell receptor pathway (Fig. 6B). Additionally, GSEA suggested that the high-risk group might be associated with E2F targets, the G2M checkpoint, MTORC1 signaling, MYC targets, and the oxidative phosphorylation pathway (Fig. 6C).

3.7 Immune landscape analysis

Cuproptosis and lipid metabolism reprogramming are not limited to tumor cells; they are also closely related to immune cell infiltration and TME function. Therefore, the impact of the risk index on the tumor immune microenvironment was explored. First, the correlations among the risk index, model genes, and the abundance of infiltrating immune cells were investigated. The results revealed a negative correlation between the risk index and CD8[+] T cells, CD4[+] memory resting T cells, plasma cells, monocytes, resting mast cells, resting dendritic cells, and naïve B cells, whereas a positive correlation was observed with resting NK cells and macrophages (Fig. 7A). Figure 7B shows that the risk index was negatively correlated with most immune checkpoints. Among the model genes, ACSL1, ATP2B4, ENPP6, and ULBP1 were positively correlated with most immune checkpoint genes, whereas PIP4K2C and SRD5A3 were negatively correlated with these genes. Therefore, the differences in the immune microenvironment between the two subgroups were further analyzed. ESTIMATE analysis revealed that low-risk patients had higher scores than did high-risk patients (Fig. 7C). CIBERSORT analysis revealed that naïve B cells, memory B cells, CD8[+] T cells, resting CD4[+] memory T cells, resting dendritic cells, and resting mast cells were more abundant in the low-risk group, whereas M0 macrophages and M2 macrophages were more abundant in the high-risk group (Fig. 7D). ssGSEA revealed that the infiltration levels of B cells, CD8[+] T cells, immature dendritic cells (iDCs), mast cells, neutrophils, plasmacytoid dendritic cells (pDCs), T helper cells, T helper 2 (Th2) cells, tumor-infiltrating lymphocytes (TILs), and regulatory T cells (Tregs) were greater in the low-risk group than in the high-risk group, whereas macrophage infiltration was lower in the low-risk group (Fig. 7E). Additionally, the immune-related functions of chemokine receptor (CCR) activity, checkpoints, cytolytic activity, HLA, parainflammation, T-cell costimulation, and the type II interferon (IFN) response were more active in the low-risk group, whereas the type I IFN response was more active in the high-risk group (Fig. 7F). Higher IPS scores indicate a greater response rate to anti-CTLA4 and PD-1 antibodies, and an analysis of IPS differences revealed that low-risk breast cancer patients had higher IPS scores than did high-risk patients (Fig. 7G). These
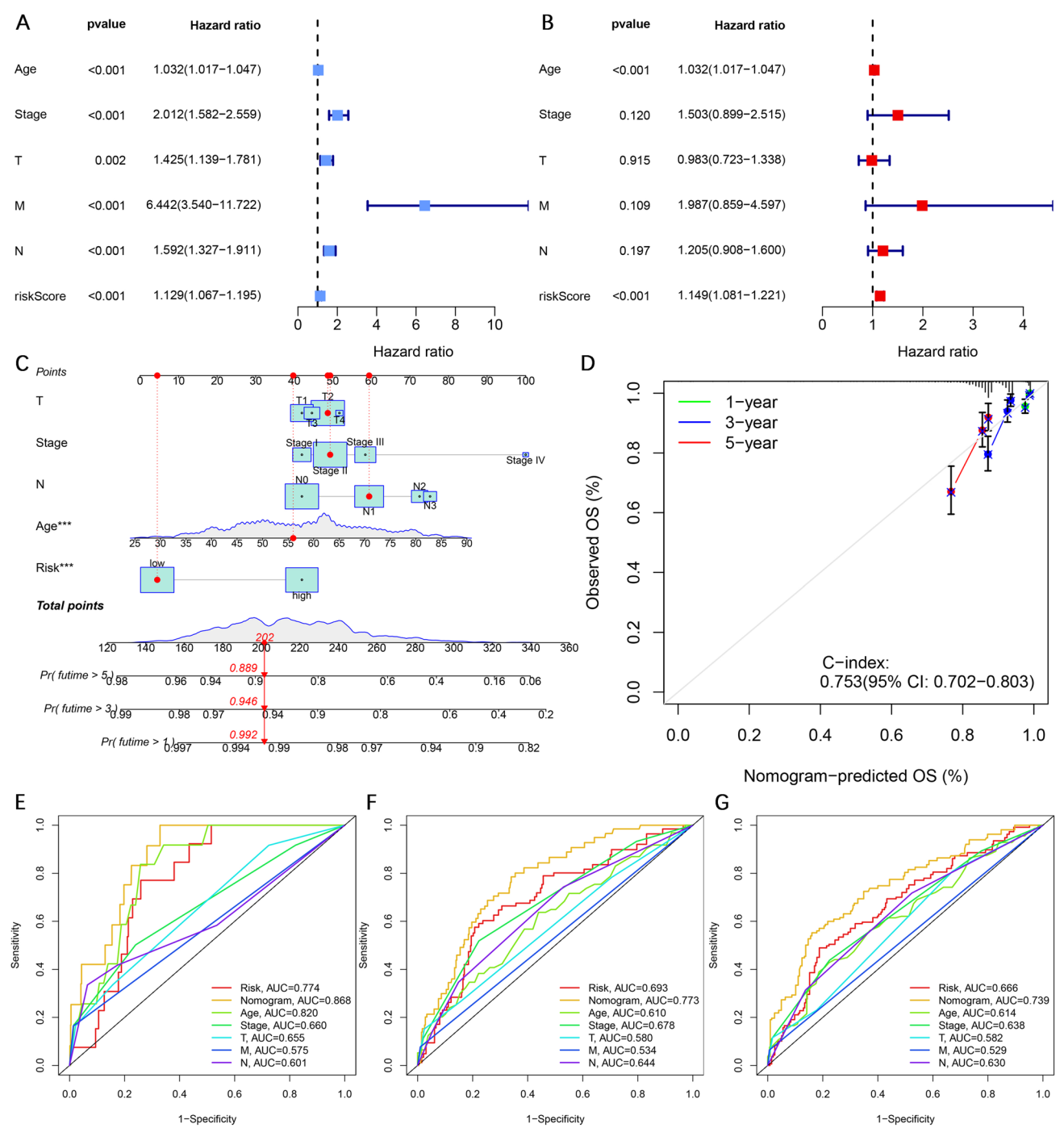
Fig. 5 Prognostic analysis and construction of the nomogram. **A–B** Univariate and multivariate Cox regression analyses. **C** Construction of the nomogram combining risk index and clinical pathological features. **D** Calibration curve of the nomogram. **E–G** Multi-parameter ROC curve analysis for 1-year, 3 year, and 5 year survival. *ROC* receiver operating characteristic
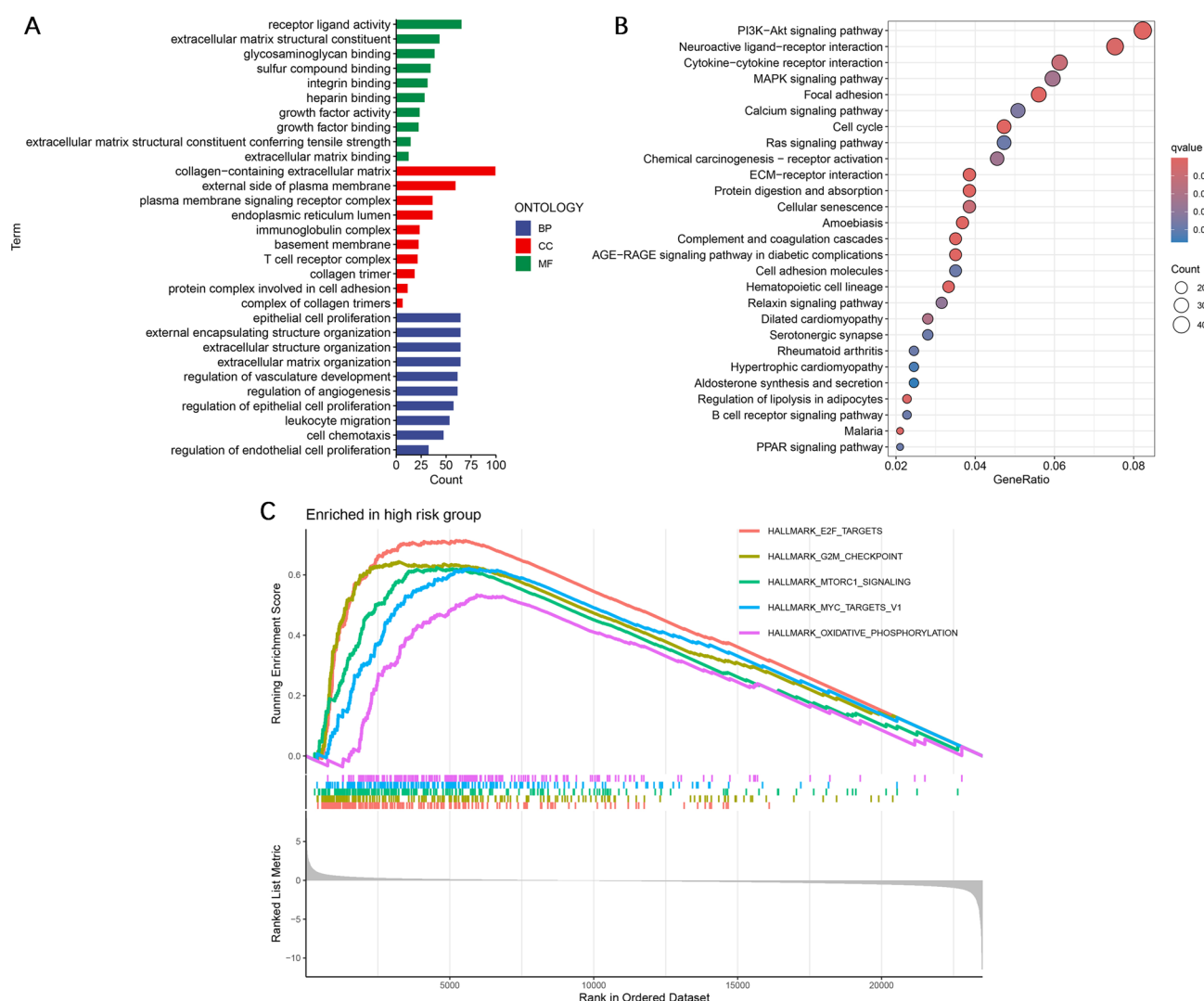
Fig. 6 Biological function analysis between high-risk and low-risk groups. **A–B** GO functional enrichment analysis and KEGG pathway enrichment analysis of differentially expressed genes between high-risk and low-risk groups. **C** GSEA analysis based on HALLMARK gene sets. *GO* gene ontology, *KEGG* kyoto encyclopedia of genes and genomes, *BP* biological process, *CC* cellular component, *MF* molecular function, *GSEA* gene set enrichment analysis

results indicate that low-risk breast cancer patients have greater immune cell infiltration and lower tumor purity, which may render them more sensitive to immunotherapy.

3.8 Analysis of TMB and drug sensitivity

Patients with a high TMB who have not received treatment tend to have a poor prognosis. Therefore, we explored the somatic TMB in patients with high- and low-risk breast cancer. The mutation frequencies of PIK3CA and CDH1 were significantly lower in the high-risk group than in the low-risk group, whereas the mutation frequencies of TP53 and TTN1 were significantly greater in the high-risk group (Fig. 8A, B). Compared with low-risk patients, high-risk patients had a

Fig. 7 Immune landscape analysis of breast cancer. **A** Spearman correlation analysis between the risk index and immune cell infiltration ▶ levels based on CIBERSORT algorithms. **B** Spearman correlation analysis between the 8 model genes and 47 immune checkpoint genes. **C** Tumor purity analysis for high-risk and low-risk breast cancer patients based on the ESTIMATE algorithm. **D** Immune cell infiltration abundance in high-risk and low-risk breast cancer patients analyzed using the CIBERSORT method. **E–F** ssGSEA analysis of immune cell infiltration and immune-related functions in high-risk and low-risk breast cancer patients. **G** IPS analysis for high-risk and low-risk breast cancer patients. *IPS* immunophenoscores; *P < 0.05; **P < 0.01; ***P < 0.001

greater TMB and worse OS (Fig. 8C, D). Multivariate survival analysis revealed that among breast cancer patients with a high TMB, those in the high-risk group had a worse prognosis, suggesting that the risk index could predict patient survival independently of TMB (Fig. 8E).

To further explore the impact of the risk index on breast cancer treatment efficacy, we first analyzed the Spearman correlation among the RNA stemness index, the DNA stemness index, and the risk index in breast cancer patients. The results revealed that the risk index was positively correlated with the stemness index (Fig. 8F, G). Studies have shown that stemness is closely related to resistance to antitumor drugs. Therefore, we analyzed the differences in antitumor drug sensitivity between high- and low-risk groups of breast cancer patients. The results revealed that the IC50 values for epirubicin, cyclophosphamide, docetaxel, paclitaxel, gemcitabine, fulvestrant, palbociclib, ribociclib, alpelisib, and rapamycin were lower in the low-risk group than in the high-risk group (Fig. 8H–Q), indicating that patients in the low-risk group were more sensitive to antitumor drugs. These findings provide a reference for the clinical individualization of antitumor drug dosage intensity.

3.9 Validation of model gene expression at the single-cell level

To validate the expression of model genes in tumor tissues, expression analysis at the single-cell level was performed using the TISCH2 database. GSE161529 includes 52 patients, with a total of 332,168 cells. These cells were dimensionally reduced and clustered into 63 groups using the UMAP method and annotated into 11 cell types (Fig. 9A, B). Among them, malignant cells, epithelial cells, and fibroblasts were the most common cell types (Fig. 9C). Single-cell analysis revealed that ACSL1 was most highly expressed in monocytes/macrophages; ATP2B4, ATP7B, PIP4K2C, SRD5A3, and ULBP1 were predominantly expressed in malignant cells; ENPP6 was most highly expressed in fibroblasts; and HSPH1 was highly expressed across various cell types (Fig. 9D–K and Fig. S3).

3.10 Interference with SRD5A3 inhibits the proliferation and migration of breast cancer cells

The multivariate Cox regression analysis used to construct the CLMG-related risk model revealed that PIP4K2C and SRD5A3 had large coefficient values (Fig. 3C), indicating their significant contribution to the risk index. Additionally, single-cell expression analysis demonstrated high expression levels of both HSPH1 and SRD5A3 in tumor cells (Fig. 9H, J), suggesting their potential roles in cancer. Among these, SRD5A3 was selected for further investigation due to its prominent coefficient value in the multivariate Cox regression analysis and its elevated expression specifically in tumor cells, which supports its potential as a key cancer-related gene. Pan-cancer analysis revealed that SRD5A3 mRNA was highly expressed in various tumor tissues, including breast cancer (Fig. 10A). In 113 paired samples of normal breast tissue and breast cancer tissue, SRD5A3 mRNA was significantly overexpressed in tumor tissues (Fig. 10B). An analysis of the Clinical Proteomic Tumor Analysis Consortium (CPTAC) database revealed that SRD5A3 protein levels were significantly higher in breast cancer tissues than in normal breast tissues (Fig. 10C). The AUC value for the ROC curve of the SRD5A3 expression level in diagnosing breast cancer was 0.810, indicating high accuracy (Fig. 10D). Survival analysis revealed that breast cancer patients with high SRD5A3 expression had shorter OS and RFS than did those with low SRD5A3 expression (Fig. 10E, F). Furthermore, in vitro cell experiments demonstrated that interference with SRD5A3 in the MCF7 and MDA-MB-231 breast cancer cell lines inhibited cell proliferation and migration **(**Fig. 10G-I). These results indicate that SRD5A3 is highly expressed in breast cancer and is associated with a poor prognosis in breast cancer patients.
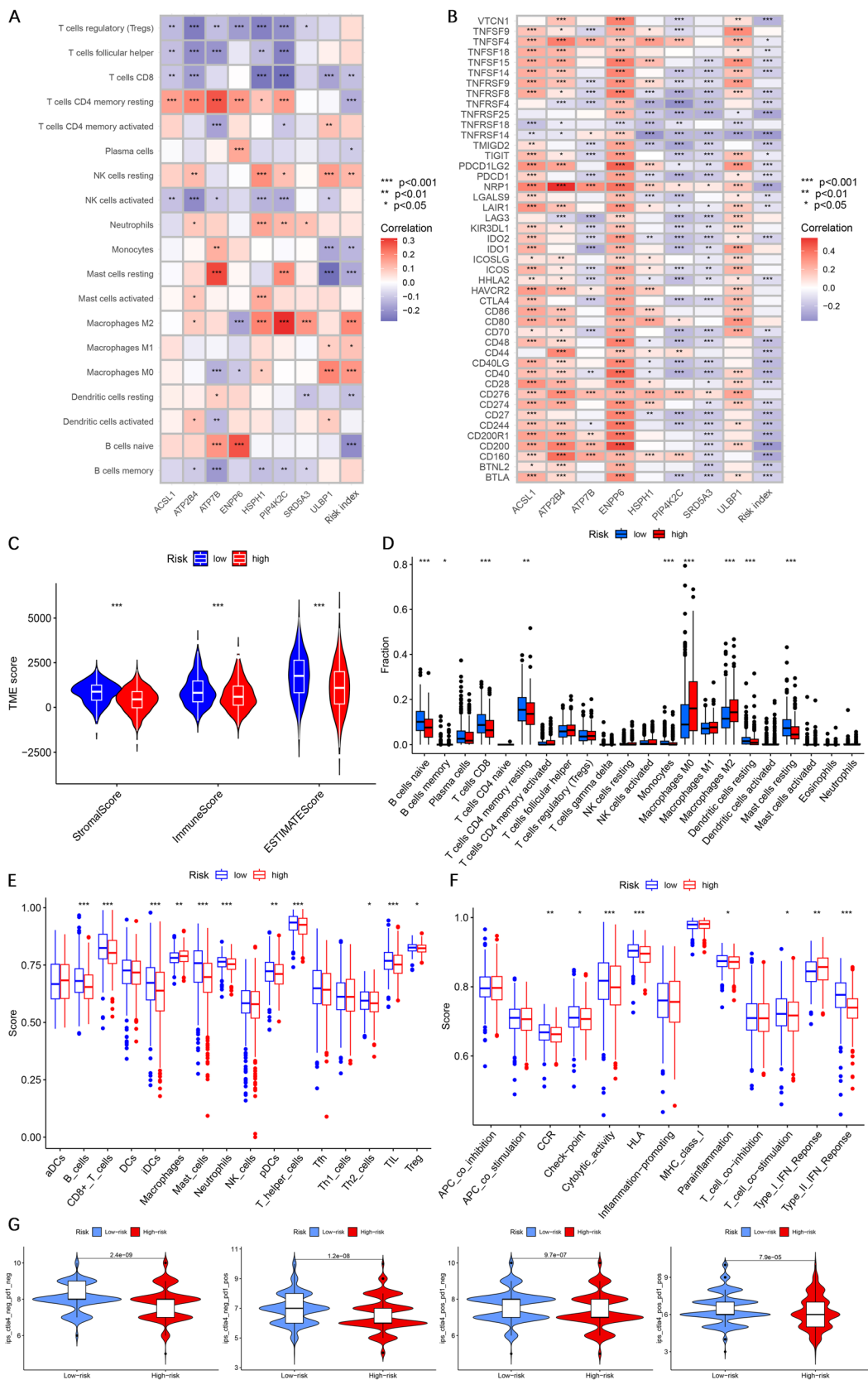
Fig. 8 TMB and drug sensitivity analysis. **A–B** Waterfall plots showing somatic gene mutations in high-risk and low-risk breast cancer ▶ patients. **C** Analysis of the differences in tumor mutational burden between high-risk and low-risk breast cancer patients. **D–E** Kaplan–Meier survival analysis of breast cancer patients based on TMB and the combination of TMB and risk index, with groups stratified using the optimal cutoff value. **F–G** Spearman correlation analysis between RNA and DNA stemness indices and the risk index. **H–Q** Drug sensitivity analysis for high-risk and low-risk breast cancer patients. *TMB* tumor mutational burden, *OS* overall survival

4 Discussion

In this study, 682 CLMGs were initially identified, and breast cancer patients were classified into four clusters using the NMF method. These four clusters presented notable differences in OS, immune cell infiltration, and stromal cell infiltration. Among them, patients in cluster 2 had the best prognosis, with high levels of CD8$^+$ T cells and cytotoxic lymphocyte infiltration; cluster 2 was classified as the "immune infiltration" type. Multiple studies have demonstrated that increased T-cell infiltration is correlated with improved prognosis and an enhanced response to immunotherapy [40, 41]. The favorable prognosis observed in patients with cluster 2 may be attributed to elevated levels of T-cell infiltration. PSME1 is a regulatory subunit of the 20S proteasome and influences antigen processing and presentation by regulating proteasome activity [42]. PSME1 is highly expressed in patients with cluster 2. PSME1 may augment antigen presentation, enhance the effectiveness of immune cell infiltration into the tumor site, and consequently enhance antitumor activity [43]. Patients in cluster 3 had a poor prognosis with abundant infiltration of stromal cells, such as endothelial cells and fibroblasts, and were classified as the "stroma-rich" type. High stromal cell infiltration is usually associated with more aggressive tumor behavior. Cancer-associated fibroblasts (CAFs) promote tumor progression through the secretion of extracellular matrix components and diverse growth factors [44]. Stromal cells in the TME can promote angiogenesis, providing the necessary oxygen and nutrients to tumor cells and thereby supporting tumor growth and expansion [45]. Moreover, stromal cells can secrete immunosuppressive factors, hindering the infiltration and activity of immune cells and allowing tumor cells to evade immune surveillance [46]. Patients in cluster 1 presented a poor prognosis with low infiltration of both immune cells and stromal cells and were classified into the "stroma-poor and immune cold" type. These patients had a high tumor purity, lacked effective immune and stromal inhibition, and were often associated with increased invasiveness, increased risk of metastasis, and a poor prognosis. Patients in cluster 4 had an intermediate prognosis with moderate infiltration of immune and stromal cells and were classified as the "moderately infiltrated" type. In our analysis, the variations in the tumor immune microenvironment of breast cancer are delineated through the lens of CLMGs.

A risk model on the basis of CLMGs exhibited robust prognostic performance and was validated across internal validation sets (testing set and entire set) and external validation set (GSE20685 and Kaplan−Meier Plotter sets). The risk model included ACSL1, ATP2B4, ATP7B, ENPP6, HSPH1, PIP4K2C, SRD5A3, and ULBP1. Breast cancer patients with elevated levels of ATP2B4, ATP7B, and ENPP6 exhibited favorable prognoses and were prominently expressed in the low-risk group, suggesting their protective role. Conversely, increased expression of ACSL1, HSPH1, PIP4K2C, SRD5A3, and ULBP1 was correlated with a poorer prognosis and showed significant elevations in the high-risk group, supporting their role as risk-associated molecules. SRD5A3, a member of the steroid 5α-reductase family, plays a key role in cholesterol metabolism and steroid hormone synthesis. Elevated levels of SRD5A3 are associated with poorer OS in liver cancer patients [47]. The loss of SRD5A3 inhibits liver cancer growth [47]. Our study revealed that SRD5A3 was highly expressed in breast cancer, which correlated with shorter OS and RFS. Interference with SRD5A3 expression can inhibit the proliferation and migration of breast cancer cells. ROC curve analysis suggested that SRD5A3 could serve as a potential diagnostic biomarker for breast cancer. However, further investigation into its molecular mechanisms in breast cancer is needed.

Cuproptosis is an emerging type of programmed cell death, and studies have shown that the expression of genes related to cuproptosis is closely associated with the immune microenvironment of patients with tumors, including those with breast cancer [8–10]. Lipids not only provide bioenergy for tumor cell proliferation but also act as signaling molecules in various cellular activities. Lipid metabolism influences not only tumor cells but also other cells within the TME, particularly immune cells. Changes in lipid metabolism within the TME can produce immunosuppressive lipid metabolites that inhibit the antitumor immunity of immune cells. Additionally, the reorganization of lipid metabolism within immune cells can suppress their function, leading to immune evasion by tumor cells. Thus, we conducted an in-depth analysis of how the CLMG-related risk index impacts the TME. CIBERSORT analysis revealed that this index was
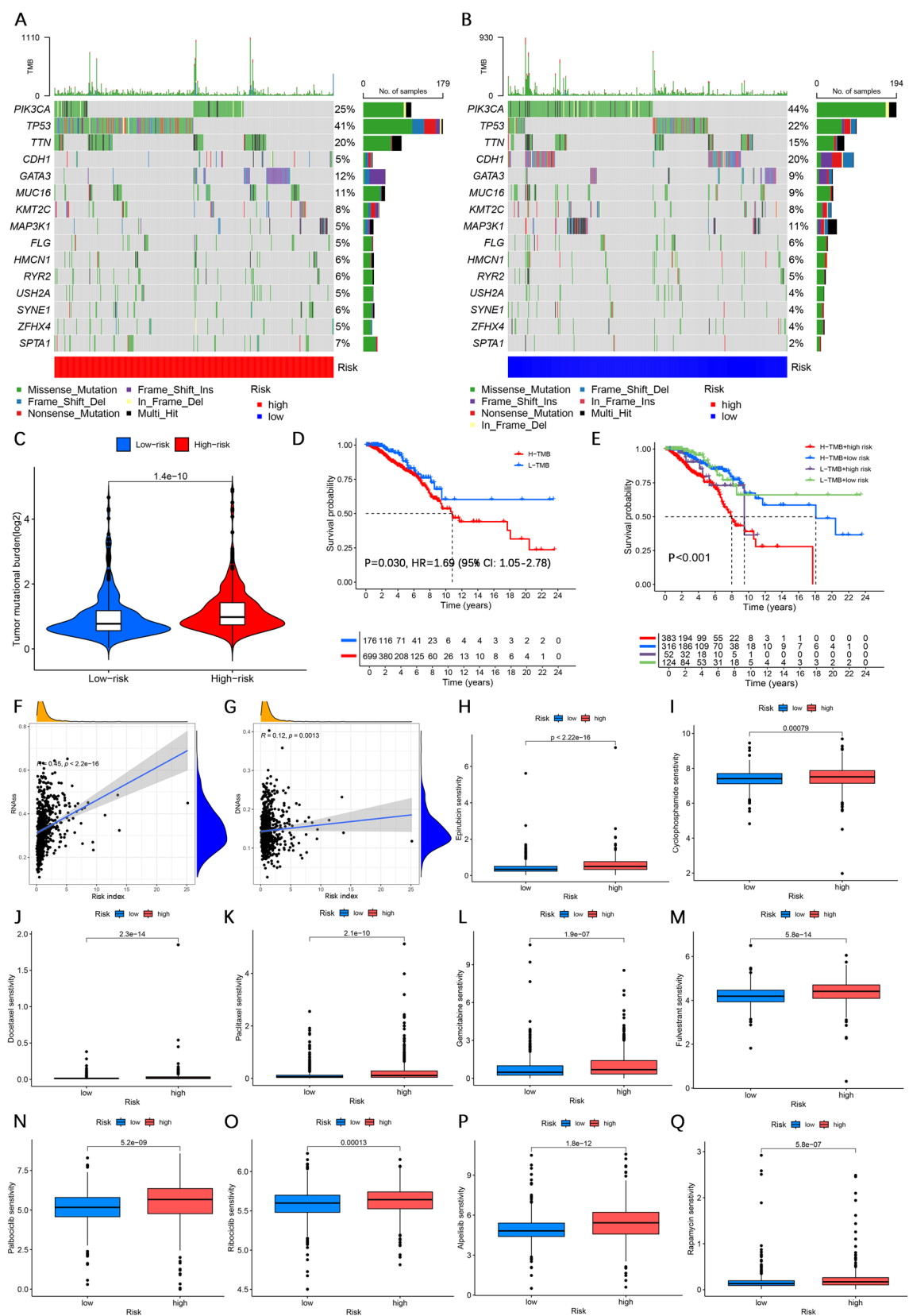
Fig. 9 Single-cell transcriptomic analysis. **A** Dimensionality reduction and clustering of breast cancer tissue using the UMAP method from ▶ the TISCH2 database. **B** Annotation of 38 clusters into 11 different cell types. **C** Pie chart showing the proportions of the 11 cell types. **D–K** Validation of the expression levels of 8 model genes (ACSL1, ATP2B4, ATP7B, ENPP6, HSPH1, PIP4K2C, SRD5A3, and ULBP1) in single-cell transcriptomic data. *UMAP* uniform manifold approximation and projection

negatively correlated with the infiltration of CD8$^+$ T cells and B cells but positively correlated with the infiltration of M2 macrophages and myeloid dendritic cells. Correlation analyses revealed negative correlations between the risk index and the expression of most immune checkpoints. These results suggest that a higher CLMG-related risk index is associated with a more immunosuppressive microenvironment. We further compared immune cell infiltration between the high- and low-CLMG-related risk groups. ESTIMATE analysis revealed that the low-risk group had greater immune and stromal cell infiltration and lower tumor purity than did the high-risk group. Both CIBERSORT and ssGSEA analyses revealed that high-risk patients had increased infiltration of M2 macrophages and reduced infiltration of CD8$^+$ T cells. Consistent findings across multiple algorithms further underscored the role of the CLMG-related risk index in predicting immunological features in breast cancer patients. Immune-related functional analyses revealed that the high-risk group exhibited more active type I IFN responses, whereas the low-risk group showed heightened type II IFN responses. Research indicates that a robust preoperative type I IFN response is linked to a reduced efficacy of PD-1 antibody therapy and shorter survival [48]. Additionally, the type I IFN response promotes the brain metastasis of tumor cells [49]. The type II IFN response enhances antitumor immune reactions by upregulating MHC expression, increasing the expression of IFN-γ and TNF, and enhancing helper T-cell responses [50]. We also explored the response to immunotherapy in patients with different risk indices. The results revealed that patients in the low-risk group had a greater IPS than did those in the high-risk group, suggesting that patients in the low-risk group had a stronger immunogenic profile and a higher response rate to immune checkpoint inhibitors, which is consistent with the aforementioned results. Moreover, our study demonstrated that the CLMG-related risk index was associated with tumor stemness, thereby predicting the sensitivity of breast cancer patients to antitumor drugs. Our research provides insights into the tumor immune microenvironment, immunotherapy response, and antitumor drug sensitivity from the perspective of CLMGs in breast cancer patients.

This study has several limitations. First, although we constructed an interaction network for cuproptosis-related and lipid metabolism-related genes at the transcriptome level, our exploration was limited to correlation analyses. The causal relationships and specific regulatory mechanisms between cuproptosis-related and lipid metabolism-related genes still need to be validated. Second, we developed a CLMG-related risk model using a TCGA-BRCA dataset and validated the model with a GEO dataset. However, the reliability of the model still lacks validation with real-world data or prospective cohort data. Third, we only preliminarily explored the biological function of the CLMG SRD5A3 in breast cancer; further investigations are needed to elucidate the specific mechanisms of this molecule.

5 Conclusion

In summary, in this study, we identified 682 CLMGs, providing insights into the expression patterns and roles of these genes in breast cancer. The CLMG-related risk index can predict the overall prognosis of patients with breast cancer, patient response to immunotherapy, and patient sensitivity to anticancer drugs. Additionally, SRD5A3 is a potential therapeutic target for breast cancer patients.
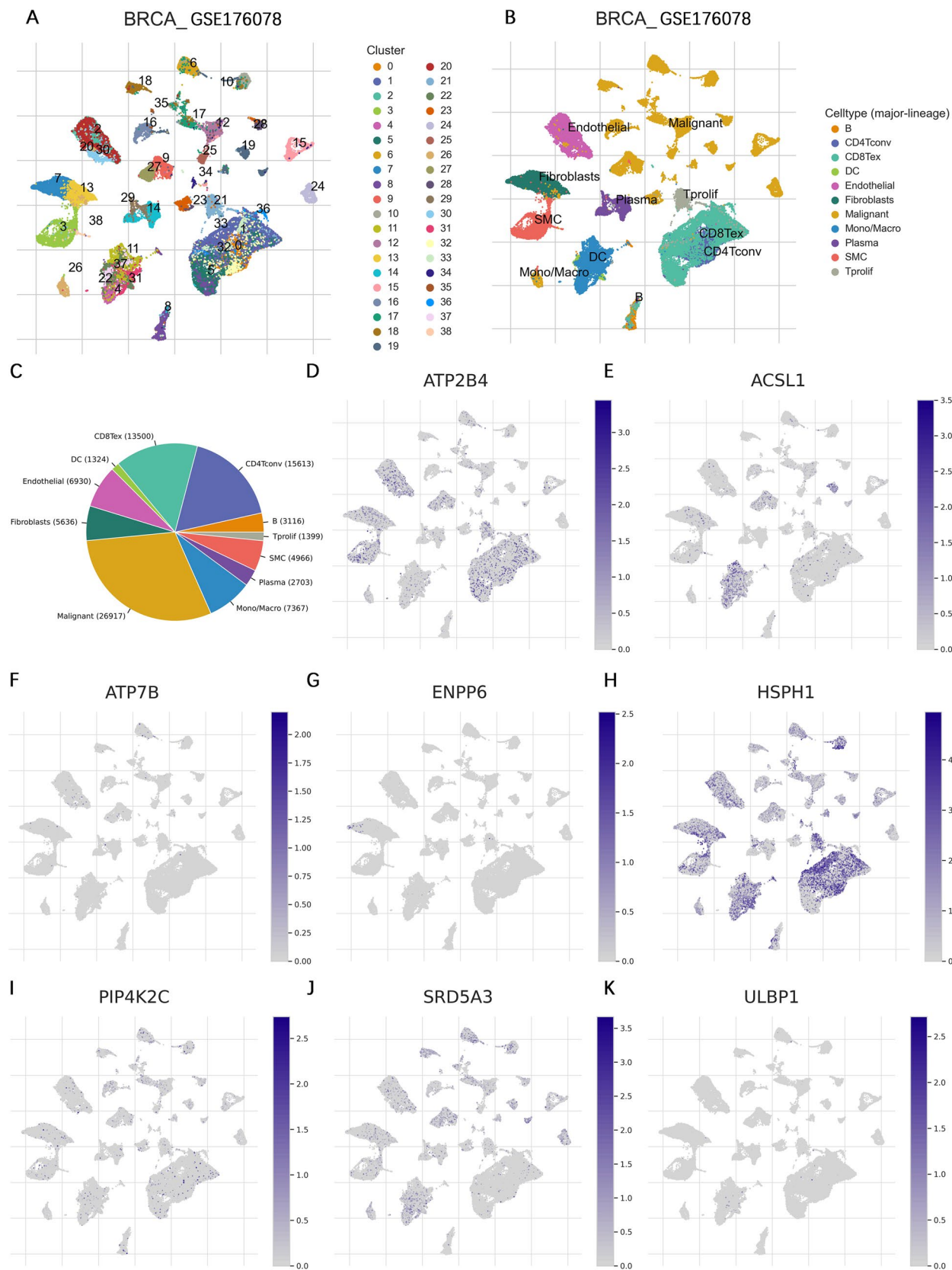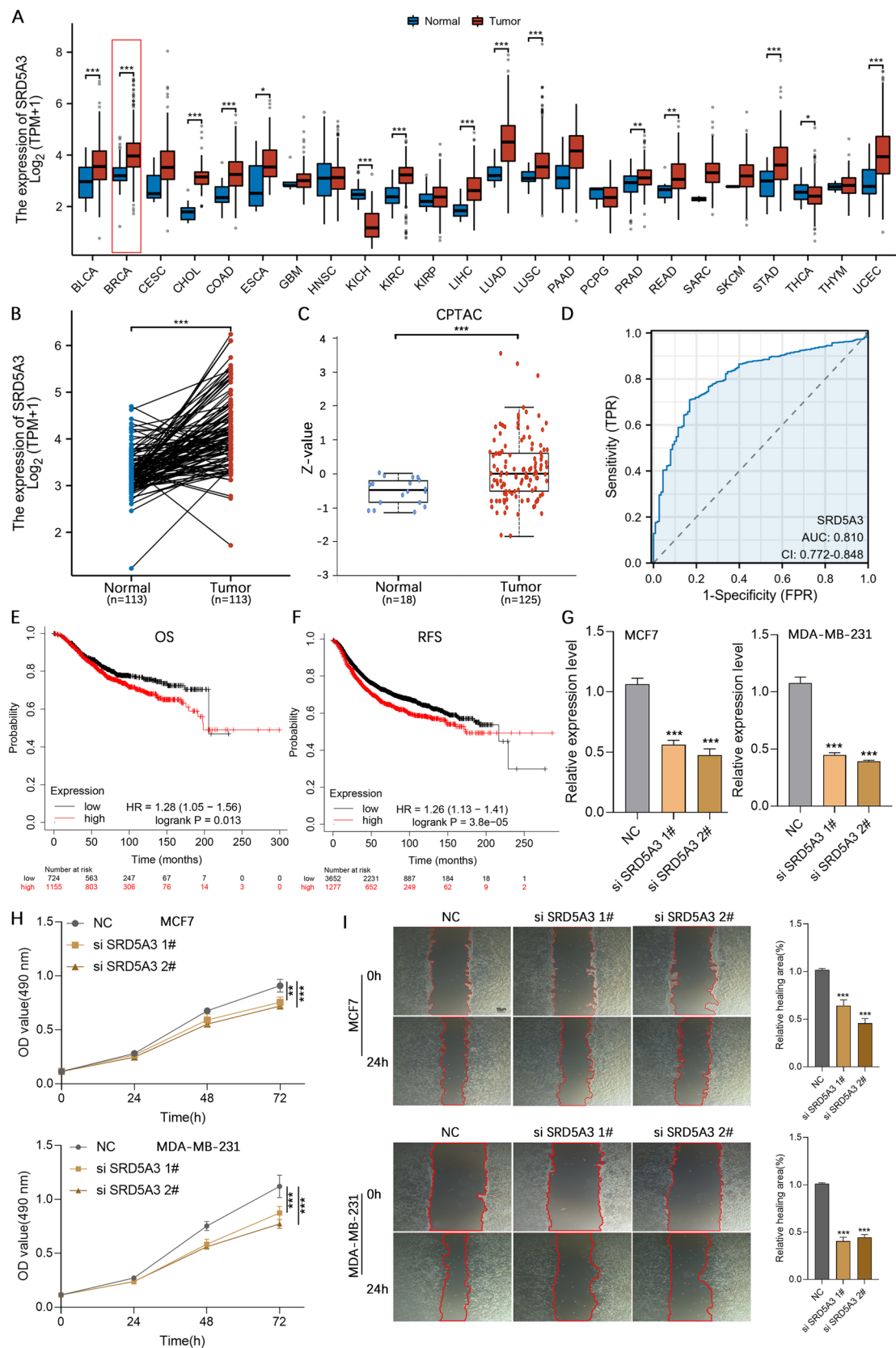
Fig. 10 Interference with SRD5A3 inhibits cell proliferation and migration in breast cancer cells. **A** Differential expression analysis of SRD5A3 ▶ between normal and tumor tissues across various cancer types in the TCGA database. **B** Comparison of SRD5A3 expression between 113 pairs of normal breast tissues and breast cancer tissues in the TCGA database. **C** Comparison of SRD5A3 protein expression levels in normal breast tissues and breast cancer tissues from the CPTAC database. **D** ROC curve analysis of SRD5A3 expression for predicting breast cancer. **E–F** Kaplan-Meier survival analysis for OS and RFS in breast cancer patients with high and low SRD5A3 expression, stratified based on the optimal cutoff value from the Kaplan-Meier Plotter database. **G** RT-qPCR analysis of SRD5A3 interference levels in MCF7 and MDA-MB-231 breast cancer cell lines. **H** MTS assay showing the effect of SRD5A3 interference on cell proliferation at 0 h, 24 h, 48 h, and 72 h in MCF7 and MDA-MB-231 cell lines. **I** Scratch assay analysis of cell migration in MCF7 and MDA-MB-231 breast cancer cell lines after SRD5A3 interference. *TCGA* the cancer genome atlas, *GTEx* genotype-tissue expression, *CPTAC* clinical proteomic tumor analysis consortium, *ROC* receiver operating characteristic, *OS* overall survival, *RFS* relapse-free survival, *NC* negative control; The data are shown as the mean ± SD. *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$

Declarations

Ethics approval and consent to participate The data for this study were sourced from publicly available databases, and ethical approval was not needed.

Competing interests The authors declare no competing interests.

References

1. Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2024;74:229–63. https://doi.org/10.3322/caac.21834.
2. Waks AG, Winer EP. Breast cancer treatment: a review. JAMA. 2019;321:288–300. https://doi.org/10.1001/jama.2018.19323.
3. Zhai J, Wu Y, Ma F, Kaklamani V, Xu B. Advances in medical treatment of breast cancer in 2022. Cancer Innov. 2023. https://doi.org/10.1002/cai2.46.
4. Tsvetkov P, Coy S, Petrova B, Dreishpoon M, Verma A, Abdusamad M, et al. Copper induces cell death by targeting lipoylated TCA cycle proteins. Science. 2022;375:1254–61. https://doi.org/10.1126/science.abf0529.
5. Tang D, Chen X, Kroemer G. Cuproptosis: a copper-triggered modality of mitochondrial cell death. Cell Res. 2022;32:417–8. https://doi.org/10.1038/s41422-022-00653-7.
6. Liu X, Luo B, Wu X, Tang Z. Cuproptosis and cuproptosis-related genes: Emerging potential therapeutic targets in breast cancer. Biochim Biophys Acta Rev Cancer. 2023;1878:189013. https://doi.org/10.1016/j.bbcan.2023.189013.
7. Li L, Li L, Sun Q. High expression of cuproptosis-related SLC31A1 gene in relation to unfavorable outcome and deregulated immune cell infiltration in breast cancer: an analysis based on public databases. BMC Bioinform. 2022;23:350. https://doi.org/10.1186/s12859-022-04894-6.
8. Huang T, Liu Y, Li J, Shi B, Shan Z, Shi Z, et al. Insights into prognosis and immune infiltration of cuproptosis-related genes in breast cancer. Front Immunol. 2022;13:1054305. https://doi.org/10.3389/fimmu.2022.1054305.
9. Liu J, Lu Y, Dai Y, Shen Y, Zeng C, Liu X, et al. A comprehensive analysis and validation of cuproptosis-associated genes across cancers: Overall survival, the tumor microenvironment, stemness scores, and drug sensitivity. Front Genet. 2022;13:939956. https://doi.org/10.3389/fgene.2022.939956.
10. Li J, Wu F, Li C, Sun S, Feng C, Wu H, et al. The cuproptosis-related signature predicts prognosis and indicates immune microenvironment in breast cancer. Front Genet. 2022;13:977322. https://doi.org/10.3389/fgene.2022.977322.
11. Song S, Zhang M, Xie P, Wang S, Wang Y. Comprehensive analysis of cuproptosis-related genes and tumor microenvironment infiltration characterization in breast cancer. Front Immunol. 2022;13:978909. https://doi.org/10.3389/fimmu.2022.978909.
12. Xiao Q, Xia M, Tang W, Zhao H, Chen Y, Zhong J. The lipid metabolism remodeling: a hurdle in breast cancer therapy. Cancer Lett. 2024;582:216512. https://doi.org/10.1016/j.canlet.2023.216512.
13. Luo X, Cheng C, Tan Z, Li N, Tang M, Yang L, et al. Emerging roles of lipid metabolism in cancer metastasis. Mol Cancer. 2017;16:76. https://doi.org/10.1186/s12943-017-0646-3.
14. Wang J, Yang Y, Shao F, Meng Y, Guo D, He J, et al. Acetate reprogrammes tumour metabolism and promotes PD-L1 expression and immune evasion by upregulating c-Myc. Nat Metab. 2024;6:914–32. https://doi.org/10.1038/s42255-024-01037-4.

15. Zhong Y, Zeng W, Chen Y, Zhu X. The effect of lipid metabolism on cuproptosis-inducing cancer therapy. Biomed Pharmacother. 2024;172:116247. https://doi.org/10.1016/j.biopha.2024.116247.

16. Lesiów MK, Komarnicka UK, Kyzioł A, Bieńko A, Pietrzyk P. ROS-mediated lipid peroxidation as a result of Cu(ii) interaction with FomA protein fragments of *F* nucleatum: relevance to colorectal carcinogenesis. Metallomics. 2019;11:2066–77. https://doi.org/10.1039/c9mt00179d.

17. Aishajiang R, Liu Z, Wang T, Zhou L, Yu D. Recent advances in cancer therapeutic copper-based nanomaterials for antitumor therapy. Molecules. 2023. https://doi.org/10.3390/molecules28052303.

18. Tang Z, Gasperkova D, Xu J, Baillie R, Lee JH, Clarke SD. Copper deficiency induces hepatic fatty acid synthase gene transcription in rats by increasing the nuclear content of mature sterol regulatory element binding protein 1. J Nutr. 2000;130:2915–21. https://doi.org/10.1093/jn/130.12.2915.

19. Sivaramakrishnan R, Incharoensakdi A. Overexpression of fatty acid synthesis genes in *Synechocystis* sp PCC 6803 with disrupted glycogen synthesis increases lipid production with further enhancement under copper induced oxidative stress. Chemosphere. 2022;291:132755. https://doi.org/10.1016/j.chemosphere.2021.132755.

20. Shen L, Huang H, Li J, Chen W, Yao Y, Hu J, et al. Exploration of prognosis and immunometabolism landscapes in ER+ breast cancer based on a novel lipid metabolism-related signature. Front Immunol. 2023;14:1199465. https://doi.org/10.3389/fimmu.2023.1199465.

21. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics. 2012;28:882–3. https://doi.org/10.1093/bioinformatics/bts034.

22. Zeng C, Yu H, Liu X, Liu Q, Jin J. Identification and validation of a novel necroptosis-related long noncoding rna prognostic signature for lung adenocarcinoma. BioMed Res Int. 2022;2022:9710540. https://doi.org/10.1155/2022/9710540.

23. Zeng C, Liu Y, He R, Lu X, Dai Y, Qi G, et al. Identification and validation of a novel cellular senescence-related lncRNA prognostic signature for predicting immunotherapy response in stomach adenocarcinoma. Front Genet. 2022;13:935056. https://doi.org/10.3389/fgene.2022.935056.

24. Lin X, Boutros PC. Optimization and expansion of non-negative matrix factorization. BMC Bioinform. 2020;21:7. https://doi.org/10.1186/s12859-019-3312-5.

25. Zeng C, He R, Dai Y, Lu X, Deng L, Zhu Q, et al. Identification of TGF-β signaling-related molecular patterns, construction of a prognostic model, and prediction of immunotherapy response in gastric cancer. Front Pharmacol. 2022;13:1069204. https://doi.org/10.3389/fphar.2022.1069204.

26. Zeng C, Xu C, Wei Y, Ma F, Wang Y. Training and experimental validation a novel anoikis- and epithelial-mesenchymal transition-related signature for evaluating prognosis and predicting immunotherapy efficacy in gastric cancer. J Cancer. 2025;16:1078–100. https://doi.org/10.7150/jca.106029.

27. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015;43:e47. https://doi.org/10.1093/nar/gkv007.

28. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, et al. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. Innovation (Camb). 2021;2:100141. https://doi.org/10.1016/j.xinn.2021.100141.

29. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA. 2005;102:15545–50. https://doi.org/10.1073/pnas.0506580102.

30. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration of cell subsets from tissue expression profiles. Nat Methods. 2015;12:453–7. https://doi.org/10.1038/nmeth.3337.

31. Yoshihara K, Shahmoradgoli M, Martínez E, Vegesna R, Kim H, Torres-Garcia W, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. Nat Commun. 2013;4:2612. https://doi.org/10.1038/ncomms3612.

32. Barbie DA, Tamayo P, Boehm JS, Kim SY, Moody SE, Dunn IF, et al. Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. Nature. 2009;462:108–12. https://doi.org/10.1038/nature08460.

33. Charoentong P, Finotello F, Angelova M, Mayer C, Efremova M, Rieder D, et al. Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade. Cell Rep. 2017;18:248–62. https://doi.org/10.1016/j.celrep.2016.12.019.

34. Mayakonda A, Lin D-C, Assenov Y, Plass C, Koeffler HP. Maftools: efficient and comprehensive analysis of somatic variants in cancer. Genome Res. 2018;28:1747–56. https://doi.org/10.1101/gr.239244.118.

35. Han Y, Wang Y, Dong X, Sun D, Liu Z, Yue J, et al. TISCH2: expanded datasets and new tools for single-cell transcriptome analyses of the tumor microenvironment. Nucleic Acids Res. 2023;51:D1425–31. https://doi.org/10.1093/nar/gkac959.

36. Guo F, Ma J, Li C, Liu S, Wu W, Li C, et al. PRR15 deficiency facilitates malignant progression by mediating PI3K/Akt signaling and predicts clinical prognosis in triple-negative rather than non-triple-negative breast cancer. Cell Death Dis. 2023;14:272. https://doi.org/10.1038/s41419-023-05746-8.

37. Li Y-Q, Sun F-Z, Li C-X, Mo H-N, Zhou Y-T, Lv D, et al. RARRES2 regulates lipid metabolic reprogramming to mediate the development of brain metastasis in triple negative breast cancer. Mil Med Res. 2023;10:34. https://doi.org/10.1186/s40779-023-00470-y.

38. Zeng C, Qi G, Shen Y, Li W, Zhu Q, Yang C, et al. DPEP1 promotes drug resistance in colon cancer cells by forming a positive feedback loop with ASCL2. Cancer Med. 2023;12:412–24. https://doi.org/10.1002/cam4.4926.

39. Becht E, Giraldo NA, Lacroix L, Buttard B, Elarouci N, Petitprez F, et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. Genome Biol. 2016;17:218. https://doi.org/10.1186/s13059-016-1070-5.

40. van der Leun AM, Thommen DS, Schumacher TN. CD8+ T cell states in human cancer: insights from single-cell analysis. Nat Rev Cancer. 2020;20:218–32. https://doi.org/10.1038/s41568-019-0235-4.

41. Schenkel JM, Pauken KE. Localization, tissue biology and T cell state - implications for cancer immunotherapy. Nat Rev Immunol. 2023;23:807–23. https://doi.org/10.1038/s41577-023-00884-8.

42. Thompson JC, Davis C, Deshpande C, Hwang W-T, Jeffries S, Huang A, et al. Gene signature of antigen processing and presentation machinery predicts response to checkpoint blockade in non-small cell lung cancer (NSCLC) and melanoma. J Immunother Cancer. 2020. https://doi.org/10.1136/jitc-2020-000974.

43. Guo Y, Dong X, Jin J, He Y. The expression patterns and prognostic value of the proteasome activator subunit gene family in gastric cancer based on integrated analysis. Front Cell Dev Biol. 2021;9:663001. https://doi.org/10.3389/fcell.2021.663001.

44. Sahai E, Astsaturov I, Cukierman E, DeNardo DG, Egeblad M, Evans RM, et al. A framework for advancing our understanding of cancer-associated fibroblasts. Nat Rev Cancer. 2020;20:174–86. https://doi.org/10.1038/s41568-019-0238-1.

45. Steenbrugge J, De Jaeghere EA, Meyer E, Denys H, De Wever O. Splenic hematopoietic and stromal cells in cancer progression. Cancer Res. 2021;81:27–34. https://doi.org/10.1158/0008-5472.CAN-20-2339.

46. de Castro LL, Lopes-Pacheco M, Weiss DJ, Cruz FF, Rocco PRM. Current understanding of the immunosuppressive properties of mesenchymal stromal cells. J Mol Med. 2019;97:605–18. https://doi.org/10.1007/s00109-019-01776-y.

47. Mai Q, Sheng D, Chen C, Gou Q, Chen M, Huang X, et al. Steroid 5 alpha-reductase 3 (SRD5A3) promotes tumor growth and predicts poor survival of human hepatocellular carcinoma (HCC). Aging. 2020;12:25395–411. https://doi.org/10.18632/aging.104142.

48. Boukhaled GM, Gadalla R, Elsaesser HJ, Abd-Rabbo D, Quevedo R, Yang SYC, et al. Pre-encoded responsiveness to type I interferon in the peripheral immune system defines outcome of PD1 blockade therapy. Nat Immunol. 2022;23:1273–83. https://doi.org/10.1038/s41590-022-01262-7.

49. Ma W, Oliveira-Nunes MC, Xu K, Kossenkov A, Reiner BC, Crist RC, et al. Type I interferon response in astrocytes promotes brain metastasis by enhancing monocytic myeloid cell recruitment. Nat Commun. 2023;14:2632. https://doi.org/10.1038/s41467-023-38252-8.

50. Boehm U, Klamp T, Groot M, Howard JC. Cellular responses to interferon-gamma. Annu Rev Immunol. 1997;15:749–95. https://doi.org/10.1146/annurev.immunol.15.1.749.