Discover Oncology

Analysis

Unraveling the significance of cuproptosis in hepatocellular carcinoma heterogeneity and tumor microenvironment through integrated single-cell sequencing and machine learning approaches

Wang Liu[1] · Liangjing Xia[2] · Yuan peng[3] · Qiang Cao[4] · Ke Xu[3] · Huiyan Luo[3] · Yongjun Peng[5] · Yanping Zhang[6]

Received: 8 February 2025 / Accepted: 13 May 2025
Published online: 24 May 2025

Abstract

Background Hepatocellular carcinoma (HCC) exhibits pronounced heterogeneity, which significantly limits the effectiveness of precision therapies. A comprehensive understanding of the biological characteristics and molecular mechanisms underlying HCC cell subpopulations is crucial for improving prognostic predictions and refining treatment strategies.

Methods Single-cell RNA sequencing data were obtained from the GEO database and processed using the Seurat R package for quality control, including data filtering, batch effect correction, and dimensionality reduction via PCA and UMAP to visualize cell distribution and identify distinct subpopulations. Cell types were annotated using established marker genes and literature references. The GSVA method was applied to evaluate the activity of 18 programmed cell death pathways. Cell developmental trajectories were reconstructed using Monocle 2 and validated with cytoTRACE to assess differentiation potential. Metabolic pathway activity was analyzed using the scMetabolism package. Bulk RNA sequencing data from the TCGA cohort were integrated to identify prognosis-associated genes through univariate Cox regression. The malignant potential of tumor subpopulations was quantified using GSVA scoring. Weighted gene co-expression network analysis (WGCNA) was employed to identify cuproptosis-related genes. A risk scoring model was constructed using LASSO regression and multivariate Cox regression based on cuproptosis-related genes and marker genes of cuproptosis-characterized tumor cells. The model's performance was validated across TCGA, GEO, and ICGC datasets. Additionally, the relationships between risk scores, clinical characteristics, key signaling pathways, and immunotherapy responses were explored. Finally, a prognostic nomogram was developed to support clinical decision-making.

Results 12 programmed cell death pathways were enriched in tumors, with cuproptosis defining HCC, particularly in the C2 subpopulation. GSVA highlighted high-risk patient enrichment in proliferation, DNA repair, and metabolism, reflecting aggressive malignancy. Developmental trajectory and metabolic analyses confirmed greater stemness and metabolic activity in C2. TCGA linked cuproptosis-related subpopulations to poor prognosis. The risk model stratified patients (validated in TCGA/GEO/ICGC), correlating with clinical grade, T-stage, survival (HR = 2.597, 95%CI 2.051–3.289, $P < 0.05$). The nomogram showed strong predictive power (C-index = 0.716), aiding clinical decisions.

Conclusion The C2 subpopulation represents the most malignant subset of HCC cells, with cuproptosis serving as a defining characteristic of this subgroup. The risk scoring and nomogram models based on cuproptosis-related genes offer

Wang Liu, Liangjing Xia, Yuan peng and Qiang Cao have contributed equally to this work.

✉ Huiyan Luo, Luohuiyan2024@163.com; ✉ Yongjun Peng, 18382971713@163.com; ✉ Yanping Zhang, ypz1024@hotmail.com | [1]Department of General Surgery, Cheng Fei Hospital, Chengdu, Sichuan 610000, People's Republic of China. [2]College of Chinese Medicine, Hong Kong Baptist University, Kowloon Tong, Hong Kong, China. [3]Department of Oncology, Chongqing General Hospital, Chongqing University, Chongqing 401147, China. [4]Department of Earth Sciences, Kunming University of Science and Technology, Kunming 650093, China. [5]Department of Orthopedics, Xichong People's Hospital, Nanchong 637200, China. [6]Department of Gastroenterology, Anqing Municipal Hospital, Anqing, Anhui 246000, People's Republic of China.

novel insights and a robust scientific foundation for prognostic prediction and personalized treatment in HCC patients. These findings highlight the potential of targeting cuproptosis and tumor microenvironment interactions to improve therapeutic outcomes in HCC.

Keywords HCC · Cuproptosis · Single-Cell Sequencing · Machine Learning · Tumor microenvironment

1 Introduction

HCC stands as one of the foremost contributors to cancer-related mortality worldwide, emerging from a complex interplay of etiological factors. These include chronic hepatitis B virus (HBV) and hepatitis C virus (HCV) infections, as well as metabolic dysfunction-associated steatohepatitis (MASH), a condition increasingly linked to hepatocellular carcinogenesis. According to data from the World Health Organization (WHO), more than 800,000 individuals worldwide are diagnosed with HCC annually, with nearly the same number succumbing to the disease [1, 2]. Surgical resection is regarded as the treatment of choice; however, due to its insidious onset, aggressive nature, and rapid progression, coupled with the challenges of early diagnosis, the majority of patients are identified at an advanced stage or after distant metastasis, rendering them ineligible for surgery. Alternative therapies, including local radiofrequency ablation, transcatheter arterial chemoembolization, and systemic chemotherapy, often fail to achieve satisfactory outcomes [3, 4]. Hence, the pursuit of novel biomarkers and therapeutic targets to enhance the clinical efficacy of HCC treatment is of profound significance.

A novel form of cell death, termed cuproptosis, has recently been identified [5]. Cuproptosis arises from the excessive accumulation of intracellular copper ions, which leads to the abnormal aggregation of lipoylated proteins, disruption of mitochondrial iron-sulfur cluster proteins involved in respiration, and subsequent induction of proteotoxic stress, culminating in cell death. This process is distinct from other known forms of cell death, such as apoptosis, ferroptosis, autophagy, and programmed necrosis [6, 7]. Specifically, after extracellular copper ions enter the cell via copper ion carriers or ATP7 A/B transporters, a portion is delivered to the mitochondrial matrix by molecular chaperones. Within the mitochondria, FDX1 reduces $Cu^{2+}$to Cu^{+}, which, on one hand, disrupts iron-sulfur protein function by inhibiting Fe-S cluster synthesis, thereby inducing apoptosis, and on the other hand, Cu^{+}binds to fatty acylated TCA cycle enzymes, leading to abnormal protein oligomerization and triggering proteotoxic stress, ultimately synergistically initiating programmed cell death. Studies have revealed significantly elevated copper ion levels in the serum and liver tissues of HCC patients. Serum copper and ceruloplasmin levels have been proposed as potential biomarkers for HCC detection [8]. Furthermore, research indicates that serum copper ion levels in HCC patients are strongly correlated with specific survival periods and overall survival, suggesting that serum copper levels may serve as an independent prognostic factor [9]. Additionally, it has been reported that copper ion treatment enhances the proliferative and migratory capacity of HCC cells [10, 11]. These findings underscore the pivotal role of copper ions in the initiation and progression of HCC, highlighting the potential of inducing cuproptosis as a novel therapeutic approach.Although some studies have established a connection between cuproptosis and HCC, the molecular characteristics and biological behaviors of tumor cells exhibiting cuproptotic traits remain inadequately elucidated. Current research lacks a detailed exploration of the metabolic properties and gene expression patterns of these cells at the single-cell level. Furthermore, the spatial distribution of cuproptosis-related cells within the tumor microenvironment and their interactions with other cell types have not been comprehensively investigated. These gaps in understanding hinder a complete appreciation of the role of cuproptosis in HCC pathogenesis, progression, and its therapeutic potential.Therefore, future studies must focus on single-cell resolution to uncover the specific gene expression profiles, metabolic pathway activities, and microenvironmental adaptation strategies of cuproptotic tumor cells. Such investigations will be instrumental in unraveling the mechanistic roles of cuproptosis in HCC and advancing the development of targeted therapeutic strategies.

This study integrates single-cell sequencing datasets from HCC patients to extract malignant epithelial cells and perform enrichment analysis of cuproptosis signaling pathways across different cellular subpopulations, identifying cell clusters exhibiting cuproptosis-related traits. Further investigation into the malignancy, metabolic characteristics, and potential roles of these cuproptosis-featured subpopulations in HCC progression was conducted. Additionally, a risk model based on cuproptosis-related genes was developed to predict patient prognosis and therapeutic response, providing novel targets and a theoretical foundation for personalized treatment of HCC.

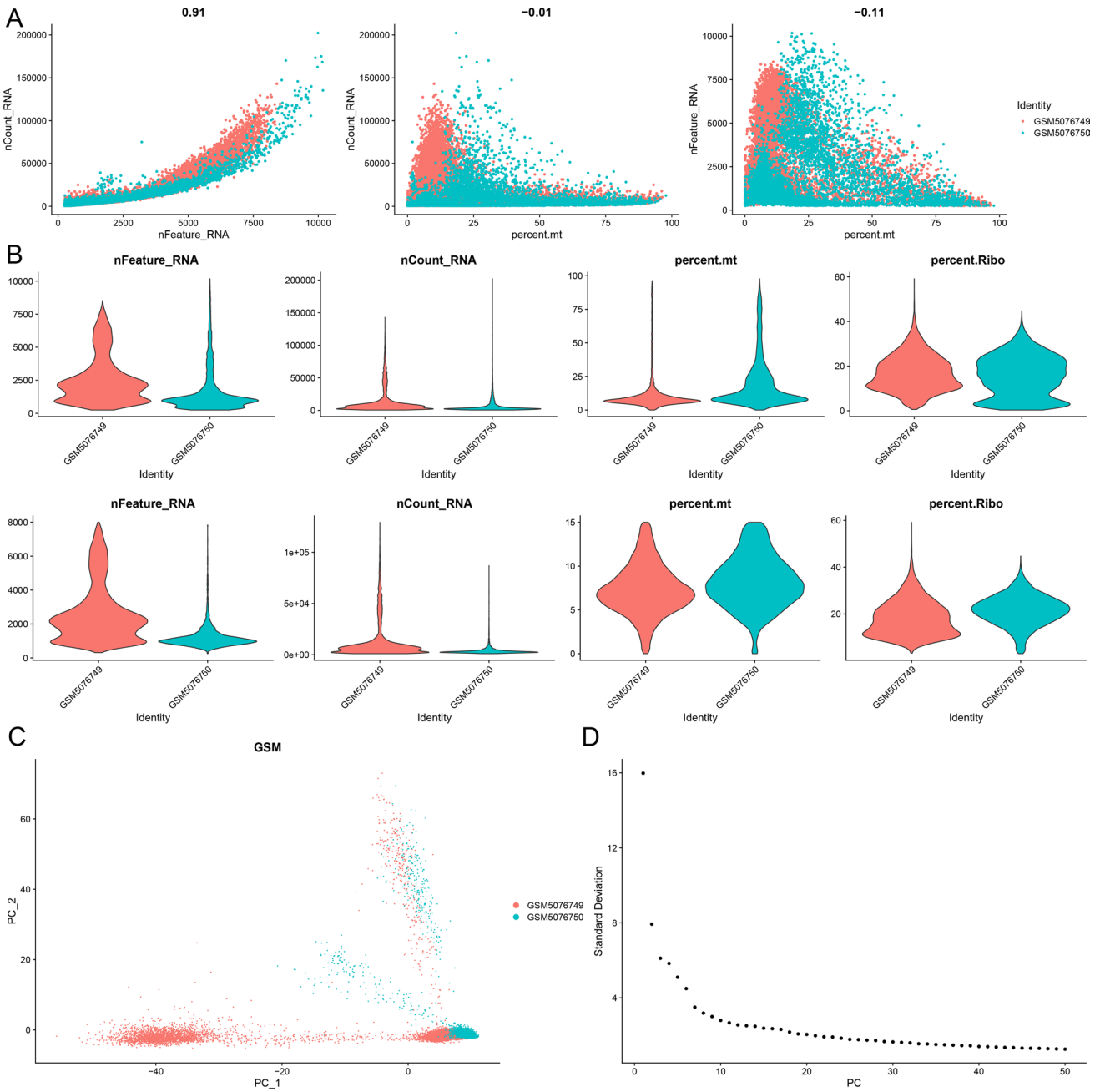


Fig. 1 **A** Correlation between basic data metrics prior to filtering. **B** Statistical comparison of the number of cells and genes before and after filtering. **C**, **D** PCA distribution of all samples and anchor plot visualization

2 Materials and methods

2.1 1Data collection and processing

The single-cell dataset GSE166635 was obtained from the GEO database and processed for quality control using the Seurat R package [12]. Cells with fewer than 200 detected genes or more than 7000 genes were excluded to prevent contamination by doublets or multiplets, and low-quality cells with mitochondrial gene proportions exceeding 20% were removed. Batch effects were corrected using the Harmony R package, with the Adjusted Rand Index (ARI) employed to evaluate the extent of batch effect elimination. Data normalization was conducted using the `NormalizeData` function,

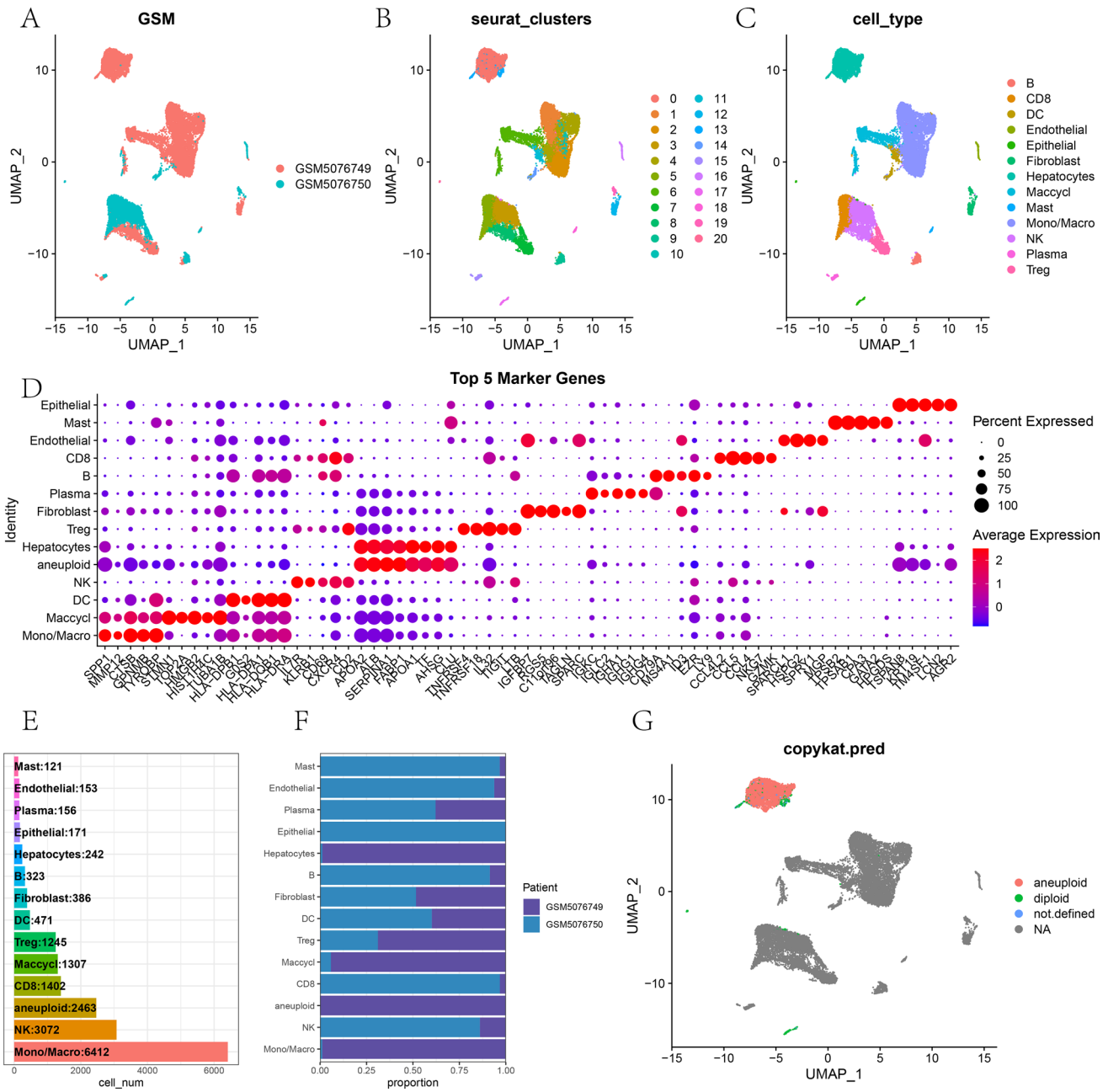


Fig. 2 **A** UMAP plot illustrating the distribution of two samples. **B** UMAP plot displaying the distribution of 21 clusters after clustering. **C** UMAP plot of annotated cell subclusters. **D** Bubble plot showing the expression of the top five marker genes in annotated subclusters. **E** Proportion and cell count of subclusters in tumor tissues versus adjacent normal tissues. **F** UMAP plot showing the distribution of malignant and non-malignant cells predicted using the copykat package

followed by downstream analysis of the quality-controlled cells. Due to the high dimensionality of the data, Principal Component Analysis (PCA) was performed using the `RunPCA` function for linear dimensionality reduction, and the optimal number of principal components was determined via the `ElbowPlot` function. The `FindNeighbors` function in Seurat was used to calculate K-nearest neighbor (KNN) relationships, constructing a k-nearest neighbor matrix. The shared nearest neighbor (SNN) clustering algorithm, implemented through the `FindClusters` function, was then applied to identify smaller partitioned clusters, forming an SNN graph.To visualize the dimensionality reduction results

Fig. 3 UMAP visualization of marker gene expression

in two-dimensional space intuitively, the UMAP method was applied to the PCA-transformed data, creating a clear mapping that revealed the intrinsic patterns and structures within the single-cell dataset. Finally, cell types were annotated based on previously published literature and known marker genes.

2.2 RCD score calculation

Single-cell gene enrichment and differential pathway analysis were conducted using the GSVA R package [13]. GSVA, a non-parametric and unsupervised method, enables the estimation of differential pathway enrichment based on gene expression datasets. For functional enrichment analysis at the single-cell level, 18 RCD scores—encompassing apoptosis, ferroptosis, disulfidoptosis, autophagy, and others—were curated from previously published literature [14]. These scores were utilized to identify functionally relevant gene categories significantly associated with differential genes across distinct cell populations, offering deeper insights into the functional characteristics and biological processes of various cell groups.

2.3 Machine learning algorithms

Following GSVA pathway scoring of single-cell data, we subsequently employed the Random Forest (RF) algorithm, executed via the "randomForest" R package, for feature selection. The default iteration count for RF was 100. A model comprising 500 decision trees was deemed sufficiently robust. Gene importance was scored based on "mean decrease in accuracy", and the gene set exhibiting the highest importance scores was selected [15].

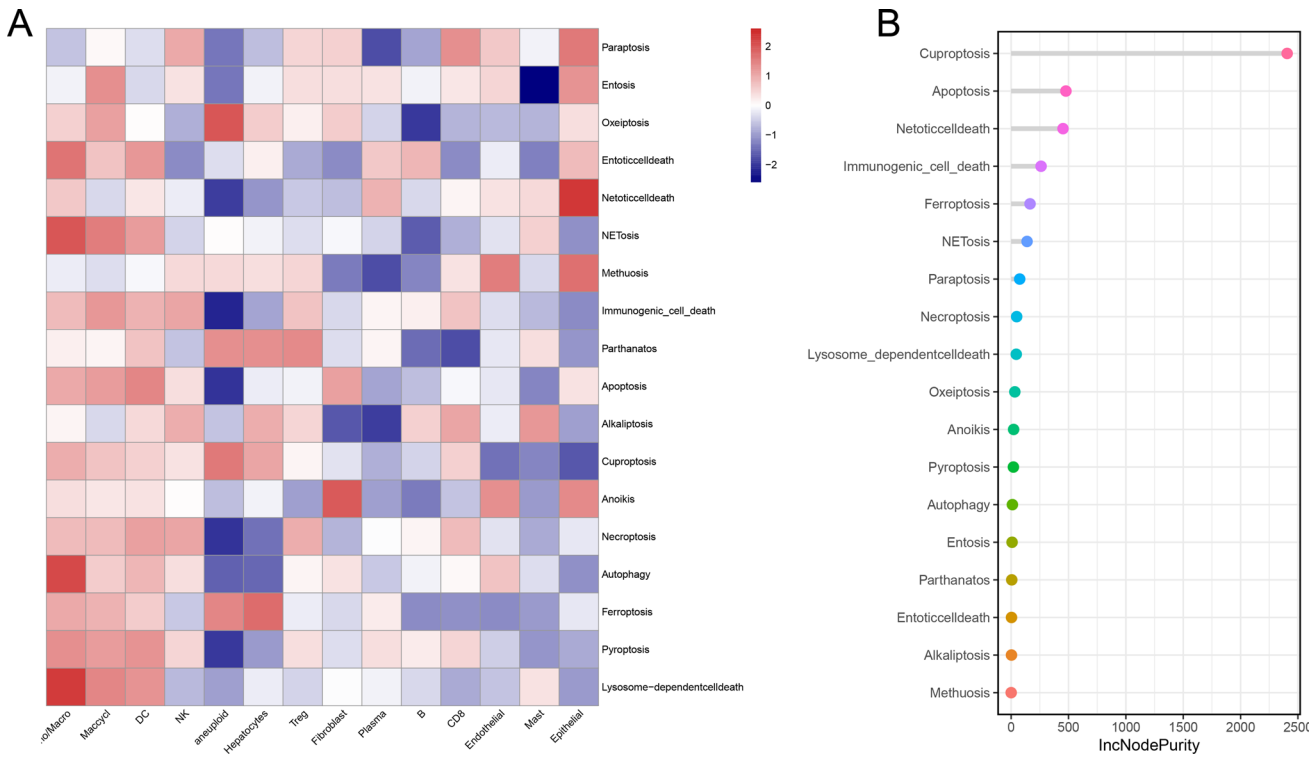


Fig. 4 **A** Heatmap of enrichment scores for various cell death pathways across different cell subclusters. **B** Heatmap showcasing key pathway enrichment scores in high- and low-risk groups, highlighting the most significant cell death pathways using a random forest algorithm

2.4 Cell developmental trajectory analysis

Developmental trajectory analysis was performed using Monocle2 on cells of the same type or those with developmental relationships [16, 17]. Monocle2 leverages key gene expression patterns and machine learning to calculate the dynamic gene expression changes that each cell undergoes during development. Cells were then ordered based on pseudotime values, simulating the dynamic changes that occur throughout the developmental process. This ordering approach projects high-dimensional data into a low-dimensional space, effectively visualizing developmental trajectories. The total length of a cell's developmental trajectory was defined by the cumulative transcriptional changes experienced from the initial state to the terminal state. Finally, cytoTRACE [18] was employed to generate a trajectory depicting the transition from low differentiation to high differentiation states.

2.5 Metabolic pathway analysis

Metabolic activity scores for specific pathways across different cell types were calculated using the"scMetabolism"package [19], a tool specifically designed for metabolic pathway analysis in single-cell sequencing data. Additionally, the metabolic states of individual cells were assessed using the AUCell method, providing a comprehensive evaluation of cellular metabolic activity.

2.6 Weighted correlation network analysis

Initially, GSVA analysis was performed on the TCGA dataset to derive activity scores for each gene set in individual samples. These activity scores reflect the functional activity levels of the gene sets within the samples. To identify potential biomarkers associated with the cuproptosis phenotype, the activity scores obtained from GSVA were used to construct a gene co-expression network, and highly correlated gene modules were identified through the WGCNA method [20]. This analysis focused on assessing the interconnectivity of gene sets and their associations with the observed phenotype. An optimal soft-threshold power (β) was first determined to calculate the adjacency between genes. Subsequently, the

adjacency matrix was transformed into a topological overlap matrix. Genes exhibiting similar expression patterns were then grouped into distinct modules. Finally, genes from the modules most strongly associated with the cuproptosis phenotype were extracted for further analysis.

2.7 Construction and evaluation of the predictive model

To explore the relationship between specific genes and patient prognosis, LASSO regression analysis [21] was performed on the gene matrix using the"glmnet"package in R. This approach identified highly correlated genes, reducing the number of genes and mitigating overfitting in the final risk model. The selected genes were then subjected to multivariate Cox regression analysis, and a risk score model was constructed based on the formula: Risk Score = $\sum ki = 0\ \beta i * expi^{-1}$,

whereβi represents the regression coefficient of each gene identified in the Cox analysis, and expi denotes the expression level of the gene. Patients were stratified into high-risk and low-risk groups based on the median risk score. Survival curves and risk plots were generated using the"survminer"and"ggrisk"packages in R, visualizing differences in survival and status among patients. Additionally, ROC curves were plotted using the"timeROC"package to evaluate the predictive performance of the risk score for 1-year, 3-year, and 5-year overall survival (OS) in patients with cholangiocarcinoma.

The relationship between risk scores and clinicopathological features in the TCGA cohort was further investigated. Cox regression analysis was performed using the"survival"package in R to determine whether the risk score could serve as an independent prognostic factor for HCC patients. Forest plots of univariate and multivariate Cox regression analyses were generated using the"forestplot"package, and a nomogram was constructed based on the TCGA cohort. Risk scores and other clinical variables were subjected to univariate and multivariate Cox regression analyses to calculate hazard ratios (HRs) and 95% confidence intervals (CIs). A nomogram was built based on the results of multivariate logistic regression analysis. Furthermore, gene set enrichment analysis was conducted using the"org.Hs.eg.db,""clusterProfiler,"and"enrichplot"packages in R to identify significantly enriched pathways between high-risk and low-risk groups. The TIDE algorithm [22] was employed to evaluate the potential response of high-risk and low-risk groups to immunotherapy, thereby assessing the effectiveness of immunotherapy strategies.

3 Results

3.1 1Data collection and processing

Initially, single-cell data were filtered by setting a threshold where each gene was required to be expressed in at least three cells, and each cell was required to express a minimum of 250 genes. The mitochondrial and rRNA content was calculated using the `PercentageFeatureSet` function, ensuring that each cell expressed between 100 and 10,000 genes, with a UMI count exceeding 100. This rigorous filtering process resulted in a dataset of 17,924 cells (Fig. 1A, B). Batch effects were subsequently eliminated to ensure compatibility among different samples (Fig. 1C).The integrated dataset was subjected to PCA for dimensionality reduction, retaining the top 30 characteristic marker genes (Fig. 1D). Cells were clustered using the `FindNeighbors` and `FindClusters` functions, yielding 21 distinct subgroups (Fig. 2B). Dimensionality reduction was further refined with the `RunUMAP` function, and UMAP plots vividly depicted the distribution of different samples (Fig. 2A).Cell subpopulations were annotated by referencing known marker genes, identifying cell types including B cells, CD8 +T cells, dendritic cells, endothelial cells, epithelial cells, fibroblasts, hepatocytes, cycling proliferative cells, mast cells, monocytes/macrophages, cycling cells, natural killer (NK) cells, plasma cells, and regulatory T cells (Tregs) (Fig. 2C). To further analyze the features of these subpopulations, the `FindAllMarkers` function was employed with a logFC threshold of 0.5 and a minimum expression fraction (Minpct) of 0.35 to identify marker genes for each subgroup. Bubble plots were utilized to display the top five marker genes expressed in each subgroup (Fig. 2D). We utilized the CopyKAT package to identify tumor cells, which are visualized in Fig. 2G using a UMAP plot. Figure 2E illustrates the number of each cell type, while Fig. 2F presents the proportion of each cell type across different groups.Finally, the distinctive characteristics of all subpopulations were examined, with specific marker gene features visualized in Fig. 3.

Fig. 5 **A** UMAP plot depicting the spatial distribution of tumor subclusters. **B** Boxplot illustrating prognostic scores across different tumor clusters. **C** Pseudotime analysis tracing the developmental trajectories of various tumor cell clusters, with cells color-coded based on tumor subcluster or pseudotime progression. **D** CytoTRACE analysis showing CytoTRACE scores for different tumor clusters, with higher scores indicating greater stemness. **E** Heatmap visualizing gene set activity differences across tumor cell clusters evaluated via ssGSEA. **F** Metabolic pathway activity of tumor cell clusters assessed using scMetabolism, visualized through heatmaps

3.2 Identification of HCC cells exhibiting cuproptosis characteristics

To investigate the specific modes of cell death most closely associated with HCC cells, we employed the ssGSEA method to calculate scores for 12 types of programmed cell death. Notably, oxeiptosis, parthanatos, cuproptosis, ferroptosis, and methuosis were significantly enriched in tumor cells (Fig. 4A). Subsequently, a random forest-based machine learning approach was utilized to rank the importance of each programmed cell death score within individual cells, with feature importance illustrated in Fig. 4B. Our findings revealed that cuproptosis emerged as the most defining feature in HCC cells. To gain deeper insights into the characteristics of cuproptosis in HCC cells, tumor cells were further stratified into four distinct subpopulations (Fig. 5A). Using ssGSEA, we calculated the cuproptosis scores for each subpopulation, which demonstrated that the C2 subpopulation exhibited the highest cuproptosis scores (Fig. 5B). Furthermore, differentiation potential across these subpopulations was assessed using Cytotrace, with the results indicating that the C2 subpopulation possessed the most robust undifferentiation potential (Fig. 5D). Combined with pseudotime trajectory analysis, it was determined that the C2 subpopulation predominantly represents an early developmental stage, with its prevalence decreasing over time, reflecting stem cell-like properties(Fig. 5C).

3.3 The crucial role of cuproptosis-driven tumor cells in metabolic activity and malignancy

Subsequently, ssGSEA was employed to perform enrichment analysis across different tumor cell subpopulations, uncovering distinct enrichment patterns of key gene sets within the four identified tumor subpopulations (Fig. 5E). Notably, cells in the C2 subpopulation were significantly enriched in pathways related to the cell cycle and DNA repair, underscoring the pivotal role of MYC in transcriptional regulation. Using scMetabolism, we further evaluated the metabolic pathway activity of different epithelial tumor cell subtypes. The results revealed that C2 cells exhibited heightened metabolic activity, with notable involvement in folate, pyrimidine, and purine metabolic pathways. This metabolic hyperactivity suggests that C2 cells likely play a critical role in promoting cellular proliferation and differentiation (Fig. 5F).

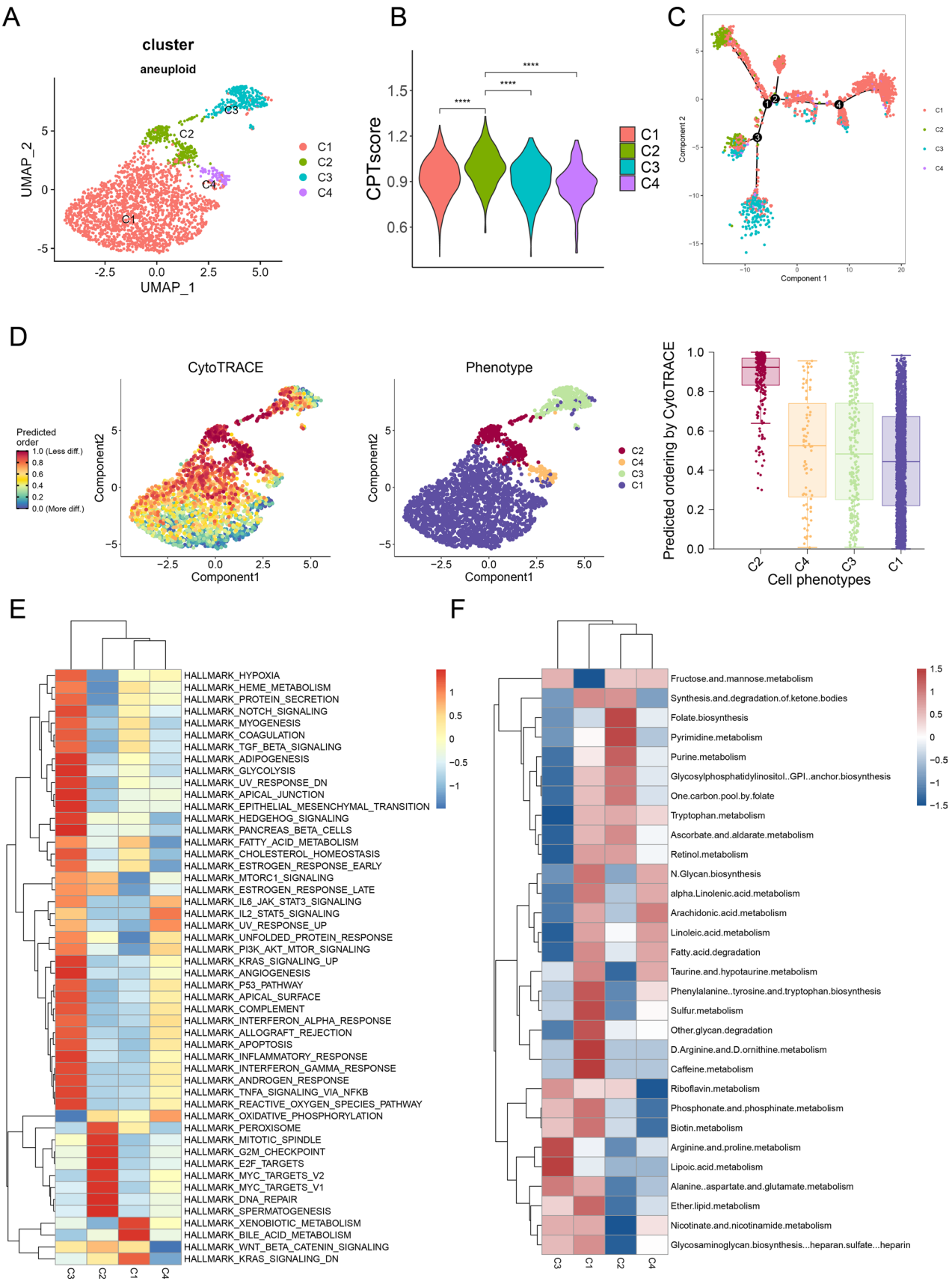
To further investigate the malignancy of the C2 subpopulation, we analyzed HCC datasets from the TCGA database. A total of 4,326 prognostic-related genes were identified using univariate Cox regression analysis. Subsequently, GSVA was utilized to assess the prognostic scores of the four tumor subpopulations, revealing that the C2 subpopulation exhibited the highest prognostic scores. This finding indicates that C2 cells possess the highest degree of malignancy among the identified subpopulations (Fig. 6A).

3.4 Identification of cuproptosis-associated genes

HCC datasets from the TCGA database were analyzed using Weighted Gene Co-Expression Network Analysis (WGCNA) to identify genes associated with the cuproptosis phenotype. Hierarchical clustering was first employed to classify the samples, and a scale-free topology standard of $R^2 > 0.85$ was used to filter the data, with $\beta = 6$ determined as the optimal soft-thresholding power to achieve scale-free independence (Fig. 6B). By setting the clustering height of module eigengenes to 0.25, ten gene modules were identified. A dendrogram was constructed to visually represent the hierarchical structure of each module (Fig. 6C). Further correlation analysis between these modules and the cuproptosis phenotype was conducted (Fig. 6D), revealing that the grey and yellow modules were significantly associated with cuproptosis. From these two modules, a total of 2,031 genes were extracted for subsequent in-depth analysis.

3.5 Construction of a risk model based on cuproptosis-associated genes

Through the intersection of marker genes from tumor cell subpopulation C2 and cuproptosis-associated genes identified via WGCNA, a total of 447 candidate genes were selected (Figs. 6E, 7A). Univariate Cox regression analysis was performed on these genes to identify potential prognostic markers for HCC within the TCGA cohort. Subsequently, LASSO regression


A
cluster
aneuploid
UMAP_2
UMAP_1
C1
C2
C3
C4
B
CPTscore
****
****
****
C1
C2
C3
C4
C
Component 2
Component 1
D
CytoTRACE
Phenotype
Predicted order
1.0 (Less diff.)
0.8
0.6
0.4
0.2
0.0 (More diff.)
Component2
Component1
Predicted ordering by CytoTRACE
Cell phenotypes
E
HALLMARK_HYPOXIA
HALLMARK_HEME_METABOLISM
HALLMARK_PROTEIN_SECRETION
HALLMARK_NOTCH_SIGNALING
HALLMARK_MYOGENESIS
HALLMARK_COAGULATION
HALLMARK_TGF_BETA_SIGNALING
HALLMARK_ADIPOGENESIS
HALLMARK_GLYCOLYSIS
HALLMARK_UV_RESPONSE_DN
HALLMARK_APICAL_JUNCTION
HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION
HALLMARK_HEDGEHOG_SIGNALING
HALLMARK_PANCREAS_BETA_CELLS
HALLMARK_FATTY_ACID_METABOLISM
HALLMARK_CHOLESTEROL_HOMEOSTASIS
HALLMARK_ESTROGEN_RESPONSE_EARLY
HALLMARK_MTORC1_SIGNALING
HALLMARK_ESTROGEN_RESPONSE_LATE
HALLMARK_IL6_JAK_STAT3_SIGNALING
HALLMARK_IL2_STAT5_SIGNALING
HALLMARK_UV_RESPONSE_UP
HALLMARK_UNFOLDED_PROTEIN_RESPONSE
HALLMARK_PI3K_AKT_MTOR_SIGNALING
HALLMARK_KRAS_SIGNALING_UP
HALLMARK_ANGIOGENESIS
HALLMARK_P53_PATHWAY
HALLMARK_APICAL_SURFACE
HALLMARK_COMPLEMENT
HALLMARK_INTERFERON_ALPHA_RESPONSE
HALLMARK_ALLOGRAFT_REJECTION
HALLMARK_APOPTOSIS
HALLMARK_INFLAMMATORY_RESPONSE
HALLMARK_INTERFERON_GAMMA_RESPONSE
HALLMARK_ANDROGEN_RESPONSE
HALLMARK_TNFA_SIGNALING_VIA_NFKB
HALLMARK_REACTIVE_OXYGEN_SPECIES_PATHWAY
HALLMARK_OXIDATIVE_PHOSPHORYLATION
HALLMARK_PEROXISOME
HALLMARK_MITOTIC_SPINDLE
HALLMARK_G2M_CHECKPOINT
HALLMARK_E2F_TARGETS
HALLMARK_MYC_TARGETS_V2
HALLMARK_MYC_TARGETS_V1
HALLMARK_DNA_REPAIR
HALLMARK_SPERMATOGENESIS
HALLMARK_XENOBIOTIC_METABOLISM
HALLMARK_BILE_ACID_METABOLISM
HALLMARK_WNT_BETA_CATENIN_SIGNALING
HALLMARK_KRAS_SIGNALING_DN
C3 C2 C1 C4
F
Fructose.and.mannose.metabolism
Synthesis.and.degradation.of.ketone.bodies
Folate.biosynthesis
Pyrimidine.metabolism
Purine.metabolism
Glycosylphosphatidylinositol..GPI..anchor.biosynthesis
One.carbon.pool.by.folate
Tryptophan.metabolism
Ascorbate.and.aldarate.metabolism
Retinol.metabolism
N.Glycan.biosynthesis
alpha.Linolenic.acid.metabolism
Arachidonic.acid.metabolism
Linoleic.acid.metabolism
Fatty.acid.degradation
Taurine.and.hypotaurine.metabolism
Phenylalanine..tyrosine.and.tryptophan.biosynthesis
Sulfur.metabolism
Other.glycan.degradation
D.Arginine.and.D.ornithine.metabolism
Caffeine.metabolism
Riboflavin.metabolism
Phosphonate.and.phosphinate.metabolism
Biotin.metabolism
Arginine.and.proline.metabolism
Lipoic.acid.metabolism
Alanine..aspartate.and.glutamate.metabolism
Ether.lipid.metabolism
Nicotinate.and.nicotinamide.metabolism
Glycosaminoglycan.biosynthesis...heparan.sulfate...heparin
C3 C1 C2 C4

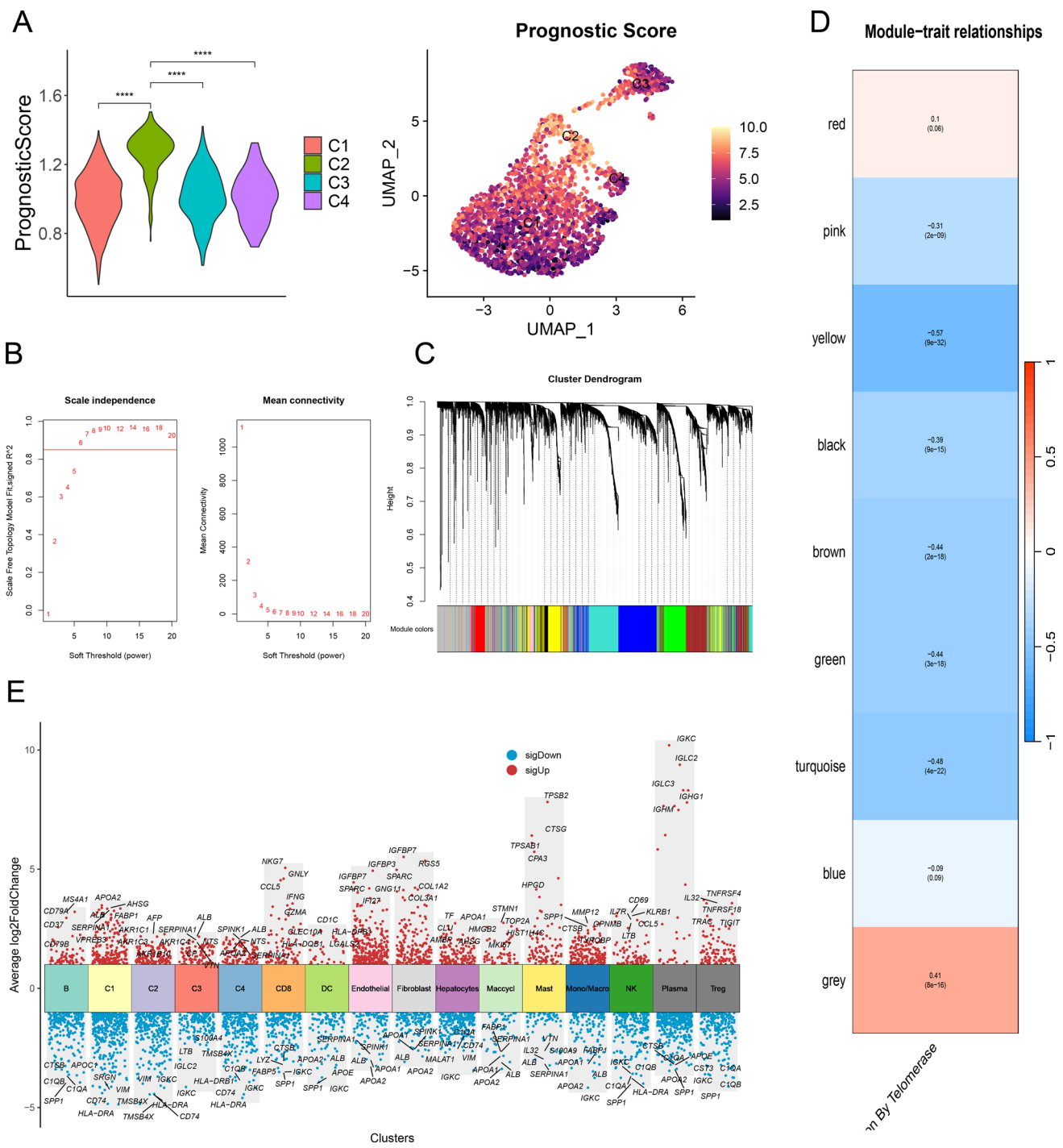


Fig. 6 **A** UMAP plot illustrating prognostic scores across different tumor clusters. **B** Violin plot depicting prognostic scores for various tumor clusters. **C** Analysis of network topology across diverse soft-thresholding powers. **D** Gene dendrogram with corresponding module colors. **E** Correlation results between 10 modules and copper death scores. **F** Top five differentially expressed genes (DEGs) with corresponding log2 FC values

analysis was employed to reduce the number of differentially expressed genes (DEGs) in the final risk model and to prevent overfitting (Fig. 7B). This process yielded 10 genes significantly associated with patient prognosis: SCGN, PPP1R1 A, TUSC1, C1S, AHCY, TFPI, MLEC, DNAJB1, TXN, and LDHA.The functional mechanisms of these ten genes in HCC involve

a multidimensional interplay encompassing metabolic reprogramming, immunomicroenvironmental modulation, and epigenetic aberrations. C1S promotes tumor microenvironment immunosuppression by regulating the complement cascade; AHCY disrupts DNA methylation by affecting S-adenosylhomocysteine hydrolase activity, driving tumor stem cell properties; the DNAJB1-PRKACA gene fusion forms an aberrant protein kinase, directly activating the PKA pathway to promote fibrolamellar HCC. In terms of metabolism, TXN enhances glycolysis and immune evasion by modulating HIF-2α signaling, while LDHA suppresses CD8 +T cell function by promoting lactate secretion. Furthermore, SPON2, a core marker of the TIMES system, inhibits recurrence by enhancing NK cell migration and activation, with its mechanism involving the activation of IFN-γ secretion and CD8 +T cell interaction. Other genes such as TFPI delay progression by inhibiting angiogenesis, and TUSC1 may participate in regulation through tumor suppression pathways, but the specific mechanisms require further validation.

Using the coefficients derived from these genes, risk scores were calculated for all samples. Patients in the TCGA cohort were stratified into high-risk and low-risk groups based on the median risk score (Fig. 7C). Kaplan–Meier survival analysis demonstrated that patients in the high-risk group had significantly shorter OS and poorer prognoses compared to those in the low-risk group (Fig. 7F). Furthermore, the risk model exhibited robust predictive accuracy for OS in the TCGA cohort, with AUCvalues of 0.76, 0.76, and 0.80 for 1-year, 3-year, and 5-year survival, respectively (Fig. 7I).

This risk model was further validated in the GEO dataset, yielding similar results (Fig. 7D). Kaplan–Meier survival curves indicated that high-risk patients had lower survival probabilities than their low-risk counterparts (Fig. 7G), and the model achieved comparable predictive accuracy, with AUC values of 0.69, 0.67, and 0.66 for 1-year, 3-year, and 5-year OS, respectively (Fig. 7J). The model was also applied to the ICGC dataset, where consistent trends were observed(Fig. 7E). Kaplan–Meier analysis again revealed worse survival outcomes for high-risk patients (Fig. 7H), and the predictive performance of the model remained robust, with AUC values of 0.74, 0.70, and 0.68 for 1-year, 3-year, and 5-year OS, respectively (Fig. 7K).

3.6 Correlation analysis between risk scores and clinical characteristics

The association between risk scores and clinical features was further examined, revealing significant correlations with pathological grade, T stage, and survival status in HCC patients ($p < 0.05$, Fig. 8A). Subsequently, univariate and multivariate Cox regression analyses were conducted in the TCGA cohort to evaluate whether the risk score could serve as an independent prognostic factor. Univariate Cox regression analysis demonstrated a positive correlation between the risk score and OS (HR:2.718, 95% CI 2.176–3.395, $p < 0.05$; Fig. 8A). Multivariate Cox analysis further confirmed that the risk score was independently associated with OS (HR: 2.597, 95% CI 2.051–3.289, $p < 0.05$); Fig. 8C).

3.7 Correlation between risk scores and key biological pathways

Through GSVA analysis, significant differences in enriched gene pathways were identified between high-risk and low-risk groups. In the high-risk group, pathways related to cell proliferation and DNA repair were prominently enriched, including the p53 signaling pathway, cell cycle, homologous recombination, DNA replication, base excision repair, nucleotide excision repair, mismatch repair, and non-homologous end joining. Collectively, these pathways underscore enhanced DNA damage repair mechanisms and genomic stability maintenance. Additionally, pathways associated with protein degradation, transcription, and metabolism were significantly enriched in the high-risk group, such as ubiquitin-mediated proteolysis, the proteasome, ribosome, RNA degradation, and spliceosome pathways. Metabolic pathways, including glycolysis (e.g., the pentose phosphate pathway), amino acid metabolism (e.g., aminoacyl-tRNA biosynthesis), and related metabolic routes, also exhibited heightened enrichment. This extensive enrichment of proliferation and DNA repair pathways suggests that high-risk group cells possess elevated proliferative activity and enhanced DNA repair capacity, thereby driving rapid tumor progression and malignant transformation (Fig. 8D).

3.8 Immunotherapy response of the risk score model

Given the absence of immunotherapy datasets specific to HCC patients, the IMvigor210 cohort[23] was utilized to predict the efficacy of chemotherapy across the two risk groups. Notably, the low-risk group exhibited a significantly higher

Fig. 7 **A** Venn diagram identifying 447 key genes intersecting marker genes from tumor cell subcluster 2 and copper death-related genes from WGCNA. **B** Trajectory of each independent variable with lambda values and coefficient distribution along the logarithmic (lambda) sequence used for parameter selection. **C**–**E** Risk plots and expression profiles of 10 genes across high- and low-risk groups in the TCGA, GEO, and ICGC cohorts. **F**–**H** Kaplan–Meier survival curves of the risk models constructed with 10 genes in TCGA, GEO, and ICGC cohorts. **I**–**K** ROC curves evaluating the performance of the risk models based on 10 genes in the TCGA, GEO, and ICGC cohorts

chemotherapy response rate compared to the high-risk group (Fig. 9A). Kaplan–Meier survival analysis further revealed that the survival probability of the high-risk group was markedly inferior to that of the low-risk group (Fig. 9B).

3.9 Construction of a nomogram for predicting patient prognosis

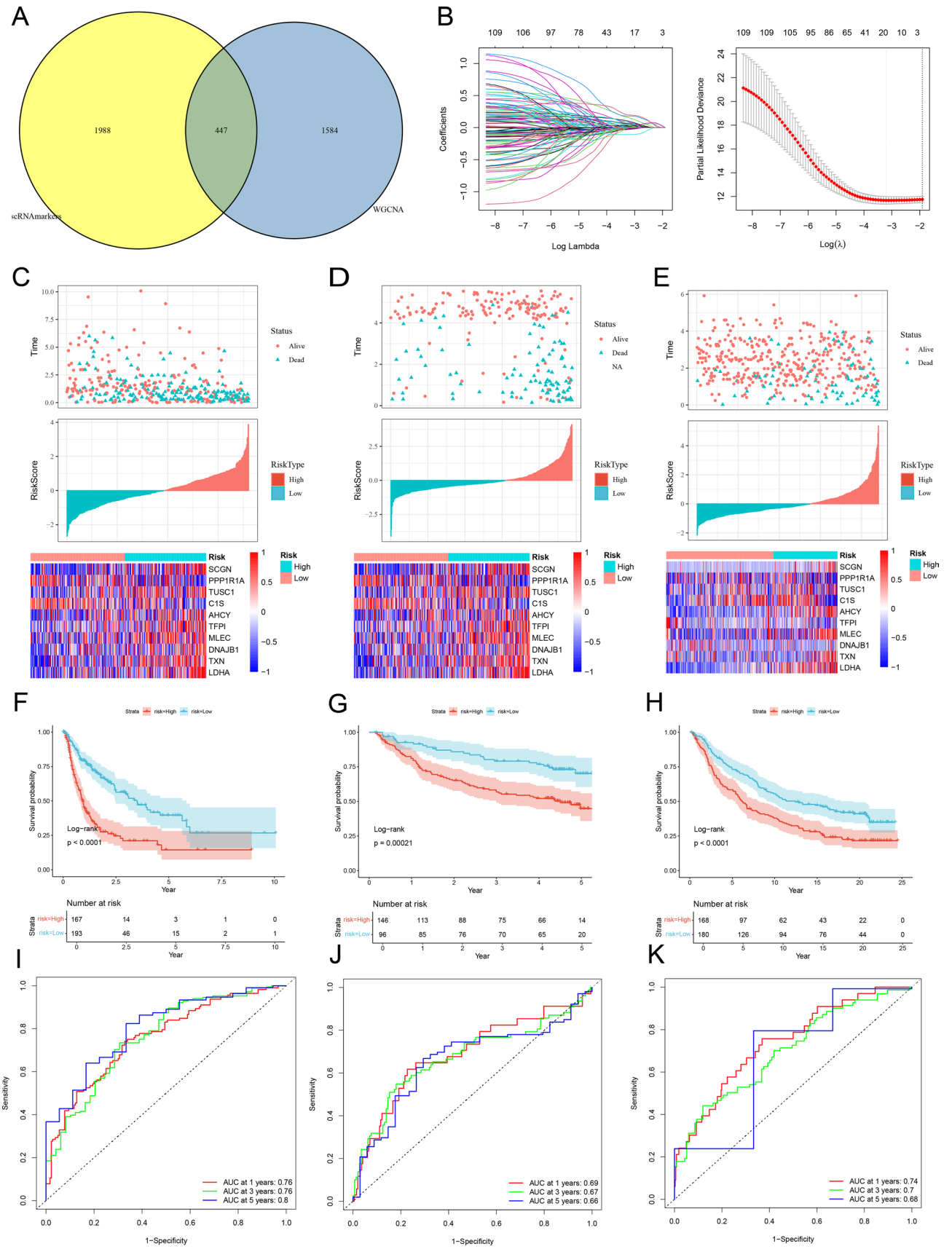
In the TCGA cohort, a nomogram was constructed by integrating variables significantly associated with patient prognosis, as determined through multivariate Cox regression analysis. This nomogram demonstrated robust predictive performance for clinical outcomes, with a concordance index (C-index) of 0.716 (95% CI 0.674–0.757; Fig. 9D). Calibration curves confirmed the alignment of the nomogram's predictions with actual clinical outcomes (Fig. 9C). Decision curve analysis (DCA) further illustrated that the nomogram provided superior clinical benefit compared to the independent use of the risk score or other clinical features alone (Fig. 9E).

4 Discussion

HCC is a prevalent malignancy of the digestive system, characterized by its insidious onset, rapid progression, and high degree of malignancy. Current clinical treatments remain largely ineffective, underscoring the urgent need to identify novel prognostic biomarkers and therapeutic targets. Cuproptosis, a recently discovered copper-dependent form of programmed cell death, involves mitochondrial dysfunction triggered by copper overload [24, 25]. However, the precise relationship between cuproptosis and HCC remains unclear. This study aims to investigate the role of cuproptosis-characteristic HCC cells in tumor progression, offering potential insights into new screening strategies, therapeutic targets, and prognostic evaluation methods for HCC.

Our findings reveal that the cuproptosis signaling pathway is significantly upregulated in HCC cells, with specific expression in a distinct subpopulation. Transcriptomic analysis further confirms the strong association between this subpopulation and HCC prognosis, suggesting that cuproptosis may play a pivotal role in the initiation and progression of HCC. Notably, FDX1 expression is significantly downregulated in HCC compared to normal liver tissue, and its reduced levels are linked to poor patient prognosis [26–28]. Prognostic risk scores based on cuproptosis-related genes (CRGs) demonstrate that patients in the high-risk group exhibit markedly shorter OS. Gao et al. constructed a cuproptosis-based prognostic model using four CRGs and quantified risk scores through three key genes—DLAT, CDKN2 A, and LIPT1. The results highlight CRG risk scores as independent prognostic factors for HCC, with lower CRG scores correlating with improved survival rates, establishing CRG scores as reliable biomarkers for HCC patients [29]. Further expression analyses using online databases and experimental validation show that CDKN2 A, DLAT, and LIPT1 are significantly upregulated in HCC tissues compared to normal tissues, with positive correlations to clinical staging and grading. Additionally, treating HCC cells with erastin (ES) resulted in reduced PLC/PRF/5 cell viability and elevated DLAT expression. Conversely, silencing DLAT restored cell viability, underscoring its critical role in HCC progression. These findings suggest that correcting the dysregulation and mutations of cuproptosis-related genes to induce cuproptosis in HCC cells could represent a promising novel therapeutic approach for treating HCC.

We identified ten cuproptosis-related genes—CGN, PPP1R1 A, TUSC1, C1S, AHCY, TFPI, MLEC, DNAJB1, TXN, and LDHA—and constructed a risk factor model based on these genes. Time-dependent ROC curves demonstrated the robust predictive capability of this model. Nomograms, a commonly used tool for evaluating tumor prognosis, offer more detailed clinical insights compared to the traditional TNM staging system, enabling more accurate survival probability estimations [30, 31]. Chen et al. utilized parameters such as TNM staging, age, and cuproptosis-related lncRNA risk scores to construct a prognostic nomogram for HCC patients, achieving high accuracy [32]. Similarly, Liu et al. developed a prognostic nomogram incorporating tumor staging and cuproptosis-related gene risk scores, yielding excellent predictive performance [33]. For the first time, this study incorporated CGN, PPP1R1 A, TUSC1, C1S, AHCY, TFPI, MLEC, DNAJB1, TXN, LDHA, T-stage, and tumor status as predictive factors to construct a prognostic nomogram for HCC patients. Calibration


A
1988
447
1584
scRNAmarkers
WGCNA
B
Coefficients
Log Lambda
Partial Likelihood Deviance
Log(λ)
C
D
E
Time
Status
Alive
Dead
NA
RiskScore
RiskType
High
Low
Risk
SCGN
PPP1R1A
TUSC1
C1S
AHCY
TFPI
MLEC
DNAJB1
TXN
LDHA
F
G
H
Survival probability
Log-rank
p < 0.0001
p = 0.00021
Year
Number at risk
Strata
risk=High
risk=Low
I
J
K
Sensitivity
1−Specificity
AUC at 1 years: 0.76
AUC at 3 years: 0.76
AUC at 5 years: 0.8
AUC at 1 years: 0.69
AUC at 3 years: 0.67
AUC at 5 years: 0.66
AUC at 1 years: 0.74
AUC at 3 years: 0.7
AUC at 5 years: 0.68

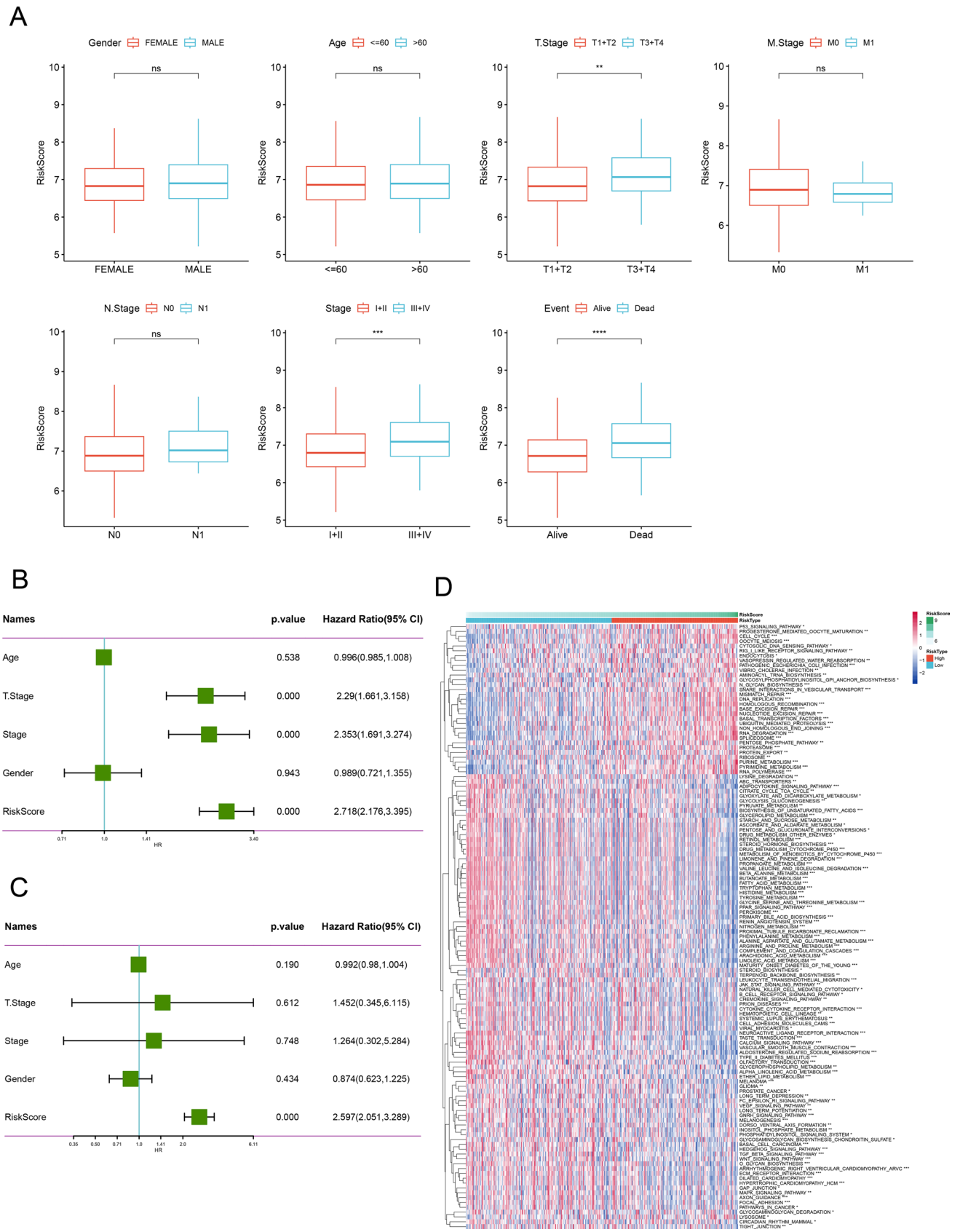


Fig. 8 **A** Association between risk scores and clinical features of HCC. **B** Univariate analysis of clinical features and risk scores. **C** Multivariate analysis of clinical features and risk scores. **D** Heatmap displaying enrichment scores of key pathways in high- and low-risk groups

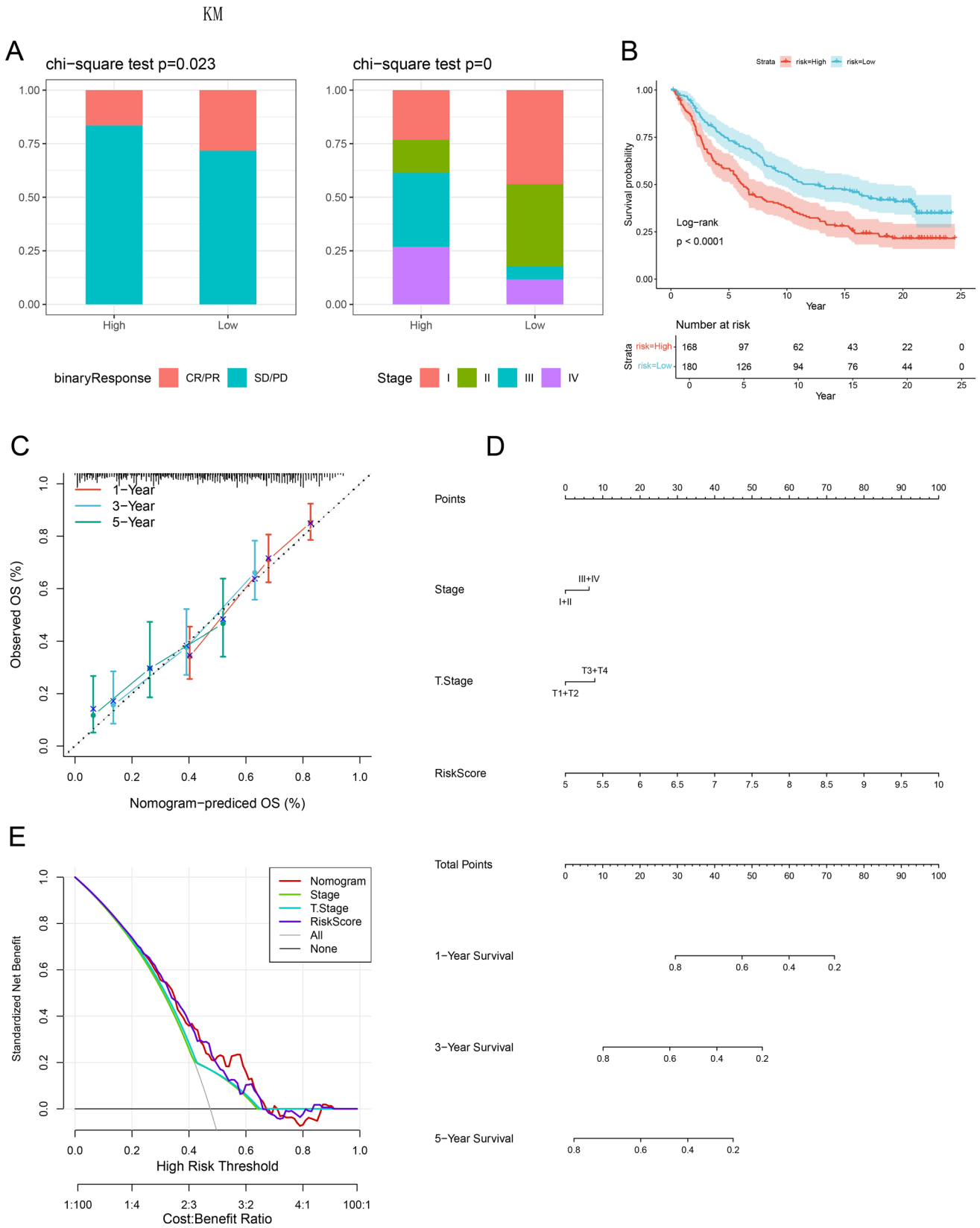


Fig. 9 **A** Differences in risk scores across immune therapy responses in the IMvigor210 cohort. **B** Prognostic differences between risk score groups in the IMvigor210 cohort. **C** Calibration curves for 1-, 3-, and 5-year predictions using the nomogram. **D** Integrated nomogram model combining risk scores and clinical features. **E** Decision curve analysis of the nomogram model

curves indicated a high degree of concordance between the nomogram's predictions and actual observations. Thus, the nomogram developed in this study provides clinicians with a novel and more effective method for assessing the prognosis of HCC patients and devising personalized treatment strategies.

Additionally, this study comprehensively analyzed the biological characteristics of tumor cells exhibiting cuproptosis features and demonstrated that the risk factor model based on cuproptosis-related genes effectively predicts the prognosis of HCC patients. The expression of these genes was significantly associated with tumor staging and the efficacy of immunotherapy, highlighting their potential as prognostic biomarkers and novel therapeutic targets for HCC.However, several limitations exist in this study. The exploration of tumor cells with cuproptosis characteristics was based solely on bioinformatics analyses, lacking experimental validation. These limitations underscore the need for future research to incorporate further data verification and experimental studies to enhance the clinical applicability and translational value of the findings.

5 Conclusion

In summary, this study, through a comprehensive evaluation of the spatial distribution of cuproptosis and the heterogeneity of HCC, highlights the pivotal role of cuproptosis in the heterogeneity of this cancer. These findings not only deepen our understanding of the mechanisms by which cuproptosis influences tumor initiation and progression but also offer fresh perspectives for the advancement of precision medicine, particularly in integrating cuproptosis inhibitors with other therapies targeting the tumor microenvironment. This research lays a theoretical foundation for exploring more effective combination treatment strategies and provides compelling preliminary evidence for the broader therapeutic application of cuproptosis inhibitors, underscoring their potential in the future landscape of cancer treatment.

Acknowledgements We gratefully acknowledge the authors and participants of all data from which we used in the study.

Author contributions WL and YZ designed this study. WL, LX, YP, QC, KX, HL, YJP and YZ conducted analyses and drafted the manuscript. YZ directed the analytical strategy and supervised the study from conception to completion. HL and YZ performed the data analyses.YJP and YZ revised the manuscript draft. All authors contributed to the interpretation of data and critically revised the manuscript.

Funding This study is supported by the funding from the Anhui Provincial Health Commission (No.AHWJ2023BAa20028), the Chengdu 2024 Medical Research Project (No. 2024530), the Medical Research Foundation of Chongqing General Hospital (No. Y2023YXYJMSXM04), and the Science and Technology Research Program of Chongqing Municipal Education Commission (No. KJQN202400116).

Data availability scRNA-seq data from CRC patients were retrieved from the GEO database (https://www.ncbi.nlm.nih.gov/geo/). Bulk RNA-seq data were obtained from TCGA database (https://portal.gdc.cancer.gov/),GEO and ICGC database.

Declarations

Ethics approval and consent to participate Not applicable.

Consent for publication Not applicable.

Competing interests The authors declare no competing interests.

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2021;71(3):209–49.

2. Fu C, Cheng C, Zhang Y. A novel signature of the ligand and receptor genes associated with disulfidoptosis for prediction of prognosis, immunologic therapy responses in hepatocellular carcinoma. Heliyon. 2023;9(9): e19502.
3. Villanueva A, Schwartz ME, Llovet JM. Liver cancer. In: Oh W, Chari A, editors. Mount Sinai expert guides: oncology. Hoboken: Wiley; 2019. p. 89–100.
4. Anwanwan D, Singh SK, Singh S, Saikam V. Singh R (2020) Challenges in liver cancer and possible treatment approaches. Biochim Biophys Acta BBA Rev Cancer. 1873;1:188314.
5. Tang D, Chen X, Kroemer G. Cuproptosis: a copper-triggered modality of mitochondrial cell death. Cell Res. 2022;32(5):417–8.
6. Xie J, Yang Y, Gao Y, He J. Cuproptosis: mechanisms and links with cancers. Mol Cancer. 2023;22(1):46.
7. Wang D, Tian Z, Zhang P, Zhen L, Meng Q, Sun B, Xu X, Jia T, Li S. The molecular mechanisms of cuproptosis and its relevance to cardiovascular disease. Biomed Pharmacother. 2023;163: 114830.
8. Baldari S, Di Rocco G, Toietta G. Current biomedical use of copper chelation therapy. Int J Mol Sci. 2020;21(3):1069.
9. Fang AP, Chen PY, Wang XY, Liu ZY, Zhang DM, Luo Y, Liao GC, Long JA, Zhong RH, Zhou ZG. Serum copper and zinc levels at diagnosis and hepatocellular carcinoma survival in the Guangdong Liver Cancer Cohort. Int J Cancer. 2019;144(11):2823–32.
10. Davis CI, Gu X, Kiefer RM, Ralle M, Gade TP, Brady DC. Altered copper homeostasis underlies sensitivity of hepatocellular carcinoma to copper chelation. Metallomics. 2020;12(12):1995–2008.
11. Fan T, Jiang L, Zhou X, Chi H, Zeng X. Deciphering the dual roles of PHD finger proteins from oncogenic drivers to tumor suppressors. Front Cell Dev Biol. 2024;12:1403396.
12. Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM, Hao Y, Stoeckius M, Smibert P, Satija R. Comprehensive integration of single-cell data. Cell. 2019;177(7):1888-1902. e1821.
13. Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinform. 2013;14:1–15.
14. Zhang W, Zhu Y, Liu H, Zhang Y, Liu H, Adegboro AA, Dang R, Dai L, Wanggou S, Li X. Pan-cancer evaluation of regulated cell death to predict overall survival and immune checkpoint inhibitor response. NPJ Prec Oncol. 2024;8(1):77.
15. He G, Jiang L, Zhou X, Gu Y, Tang J, Zhang Q, Hu Q, Huang G, Zhuang Z, Gao X, et al. Single-cell transcriptomics reveals heterogeneity and prognostic markers of myeloid precursor cells in acute myeloid leukemia. Front Immunol. 2024;15:1494106.
16. Trapnell C, Cacchiarelli D, Grimsby J, et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. Nature Biotechnology. 2014;32(4):381–386.
17. Chen H, Zuo H, Huang J, Liu J, Jiang L, Jiang C, Zhang S, Hu Q, Lai H, Yin B, et al. Unravelling infiltrating T-cell heterogeneity in kidney renal clear cell carcinoma: Integrative single-cell and spatial transcriptomic profiling. J Cell Mol Med. 2024;28(12): e18403.
18 Kang M, Armenteros JJA, Gulati GS, Gleyzer R, Avagyan S, Brown EL, Zhang W, Usmani A, Earland N, Wu Z. Mapping single-cell developmental potential in health and disease with interpretable deep learning. bioRxiv. 2024. https://doi.org/10.1101/2024.03.19.585637.
19. Wu Y, Yang S, Ma J, Chen Z, Song G, Rao D, Cheng Y, Huang S, Liu Y, Jiang S. Spatiotemporal immune landscape of colorectal cancer liver metastasis at single-cell level. Cancer Discov. 2022;12(1):134–53.
20. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008;9:1–13.
21. Tibshirani R. Regression shrinkage and selection via the lasso. J R Stat Soc Ser B Stat Methodol. 1996;58(1):267–88.
22. Jiang P, Gu S, Pan D, Fu J, Sahu A, Hu X, Li Z, Traugh N, Bu X, Li B. Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. Nat Med. 2018;24(10):1550–8.
23. Hoffman-Censits JH, Grivas P, Van Der Heijden MS, Dreicer R, Loriot Y, Retz M, Vogelzang NJ, Perez-Gracia JL, Rezazadeh A, Bracarda S. IMvigor 210, a phase II trial of atezolizumab (MPDL3280A) in platinum-treated locally advanced or metastatic urothelial carcinoma (mUC). J Clin Oncol. 2016. https://doi.org/10.1200/jco.2016.34.2_suppl.355.
24. Chen L, Min J, Wang F. Copper homeostasis and cuproptosis in health and disease. Signal Transduct Target Ther. 2022;7(1):378.
25. Wang W, Lu Z, Wang M, Liu Z, Wu B, Yang C, Huan H, Gong P. The cuproptosis-related signature associated with the tumor environment and prognosis of patients with glioma. Front Immunol. 2022;13: 998236.
26. Zhang Z, Zeng X, Wu Y, Liu Y, Zhang X, Song Z. Cuproptosis-related risk score predicts prognosis and characterizes the tumor microenvironment in hepatocellular carcinoma. Front Immunol. 2022;13: 925618.
27. Quan Y, Li W, Yan R, Cheng J, Xu H, Chen L. Tumor cuproptosis and immune infiltration improve survival of patients with hepatocellular carcinoma with a high expression of ferredoxin 1. Front Oncol. 2023;13:1168769.
28. Xiong J, Zhou X, Su L, Jiang L, Ming Z, Pang C, Fuller C, Xu K, Chi H, Zheng X. The two-sided battlefield of tumour-associated macrophages in glioblastoma: unravelling their therapeutic potential. Discov Oncol. 2024;15(1):590.
29 Gao F, Yuan Y, Ding Y, Li P-y, Chang Y, He XX. DLAT as a cuproptosis promoter and a molecular target of elesclomol in hepatocellular carcinoma. Curr Med Sci. 2023;43(3):526–38.
30. Zhang W, Ji L, Wang X, Zhu S, Luo J, Zhang Y, Tong Y, Feng F, Kang Y, Bi Q. Nomogram predicts risk and prognostic factors for bone metastasis of pancreatic cancer: a population-based analysis. Front Endocrinol. 2022;12: 752176.
31. Wang J, Wen Q, Wang X, Liu C, Zhao K, Li Y, Yang J, Liang X. Nomogram for predicting cancer-specific survival of patients with clear-cell renal cell carcinoma: a SEER-based population study. Gen Physiol Biophys. 2022;41(6):591.
32. Chen S, Liu P, Zhao L, Han P, Liu J, Yang H, Li J. A novel cuproptosis-related prognostic lncRNA signature for predicting immune and drug therapy response in hepatocellular carcinoma. Front Immunol. 2022;13: 954653.
33. Liu Z, Qi Y, Wang H, Zhang Q, Wu Z, Wu W. Risk model of hepatocellular carcinoma based on cuproptosis-related genes. Front Genet. 2022;13:1000652.