

#### **OPEN ACCESS**

EDITED BY Takaji Matsutani, Anymics Inc, Japan

REVIEWED BY
Ahmed Yaqinuddin,
Alfaisal University, Saudi Arabia
Yusheng Lin,
Jinan University, China

<sup>†</sup>These authors have contributed equally to this work and share first authorship

RECEIVED 18 June 2025 ACCEPTED 07 October 2025 PUBLISHED 22 October 2025

#### CITATION

Liu H, Sun Z, Wang X, Zhou B and Cha L (2025) Identification of immunogenic cell death signature genes in hepatocellular carcinoma: from single-cell transcriptomics to *in vitro* mechanistic validation and comprehensive prognostic modeling with hundreds of machine learning algorithms. *Front. Immunol.* 16:1649618. doi: 10.3389/fimmu.2025.1649618

#### COPYRIGHT

© 2025 Liu, Sun, Wang, Zhou and Cha. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Identification of immunogenic cell death signature genes in hepatocellular carcinoma: from single-cell transcriptomics to in vitro mechanistic validation and comprehensive prognostic modeling with hundreds of machine learning algorithms

Hongliang Liu<sup>1,2†</sup>, Zhenni Sun<sup>3†</sup>, Xi Wang<sup>4</sup>, Bin Zhou<sup>1,2</sup> and Lichao Cha<sup>1,2\*</sup>

<sup>1</sup>Department of Hepatobiliary and Pancreatic Surgery, The Affiliated Hospital of Qingdao University, Qingdao, China, <sup>2</sup>Retroperitoneal Tumor Surgery, The Affiliated Hospital of Qingdao University, Qingdao, China, <sup>3</sup>Department of Oncology, Qingdao Municipal Hospital, Qingdao, China, <sup>4</sup>Department of Oncology, Women and Children's Hospital Affiliated to Qingdao University, Qingdao, China

**Background:** Hepatocellular carcinoma (HCC) lacks reliable prognostic biomarkers for immunotherapy. Immunogenic cell death (ICD) represents a promising therapeutic target, but its comprehensive characterization in HCC remains unexplored.

**Methods:** We performed multi-omics integration of single-cell RNA sequencing data from 7 HCC samples (GSE112271, 44,461 cells) with bulk transcriptomics from three independent cohorts (TCGA-HCC [n=371], GSE14520 [n=242], ICGC [n=445]). ICD activity was quantified using ssGSEA. We identified HCC-specific ICD-related (HCC-ICDR) genes via WGCNA and optimized a prognostic model by benchmarking machine learning algorithms. Experimental validation included functional assays using CLIC1 and NAP1L1 overexpression in HepG2 cells.

**Results:** The ICD-based risk score (ICDRS) demonstrated superior prognostic accuracy (C-index=0.839), validated across cohorts. Single-cell profiling revealed macrophages exhibited the highest ICD activity. High-risk patients displayed immunosuppressive microenvironments with enriched Tregs, M0 macrophages, and neutrophils, alongside hyperactivated DNA repair and MYC signaling. Low-risk patients showed anti-tumor immunity with increased CD8+ T cells and M1 macrophages. ICDRS predicted differential therapeutic vulnerabilities: low-risk patients showed enhanced sensitivity to standard immunotherapy-compatible treatments including sorafenib and doxorubicin, while high-risk patients demonstrated preferential sensitivity to EGFR-targeted therapies. Experimental validation confirmed CLIC1 and NAP1L1 significantly promoted HCC malignant behaviors.

**Conclusions:** We established the comprehensive ICD-based prognostic framework for HCC, revealing novel tumor-immune interactions and therapeutic vulnerabilities. This model provides robust stratification for

immunotherapy selection and advances precision medicine in HCC management. Future clinical translation includes prospective validation and development of companion diagnostics, offering potential pathways for personalized HCC treatment implementation.

KEYWORDS

immunogenic cell death, hepatocellular carcinoma, multi-omics integration, precision medicine, tumor microenvironment, machine learning

### 1 Introduction

Hepatocellular carcinoma (HCC) accounts for approximately 80% of all liver cancers. Globally, there are about 700,000 new cases of liver cancer annually, with a male-to-female incidence ratio of 2-4:1. Each year, approximately 600,000 people die from liver cancer, making it the third leading cause of cancer-related deaths (1, 2).

Understanding the underlying causes of this global burden reveals significant geographic variations in HCC etiology. The etiology of hepatocellular carcinoma varies by geographic region. In Asia and sub-Saharan Africa, HCC is predominantly associated with hepatitis B and C viral infections, while in Western countries, it is linked to non-alcoholic fatty liver disease (NAFLD) and alcoholic liver disease. Overall, liver cancer development is commonly associated with chronic liver disease (3, 4). Chronic viral infections, DNA damage caused by persistent inflammation, dysregulated cellular regeneration in the context of cirrhosis, and activation of oncogenes coupled with loss of tumor suppressor gene function contribute to the development of liver cancer.

Early-stage HCC is often asymptomatic due to the liver's compensatory capacity and is typically detected incidentally during imaging. When HCC is detected early and treated with surgery or ablation, the 5-year survival rate can reach 75%. However, advanced HCC typically can only be managed with chemotherapy or local arterial embolization, with a 2-year survival rate of merely 20-25% (5, 6). Chemotherapy or targeted therapy for advanced HCC patients is challenged by tumor drug resistance, which leads to disease progression or recurrence (7). One significant cause of this resistance is immunosuppression in the tumor microenvironment, making immune enhancement particularly important for effective treatment.

Given these therapeutic limitations, novel approaches targeting immune mechanisms have gained attention. Immunogenic cell death (ICD) represents one such promising strategy. ICD, distinct from other forms of cell death, not only induces apoptosis but also activates the body's adaptive immune system against tumor cells. The underlying mechanism involves dying cells releasing or exposing a series of damage-associated molecular patterns (DAMPs), such as extracellular ATP, ecto-calreticulin (CRT), and high mobility group box 1 protein (HMGB1) (8–10). These signals promote dendritic cell uptake of tumor antigens and activate T cell

responses, thereby forming specific anti-tumor immune memory responses (11). Some research has begun to explore certain drugs such as Mecheliolide (7) and Icaritin (12), or physical effect (13)-mediated ICD enhancement mechanisms. This provides new directions for tumor treatment.

Current multi-omics studies have revealed ICD characteristic changes and potential therapeutic targets in neuroblastoma (14). clear cell renal cell carcinoma (15), and gastric cancer (16). However, comprehensive ICD characterization in hepatocellular carcinoma remains lacking. To address this gap, we systematically investigated the ICD landscape in HCC, identified key regulatory genes, and validated their functional roles through experimental approaches, providing new therapeutic targets for HCC treatment.

### 2 Materials and methods

### 2.1 Data sources

Our study integrated HCC data from three major databases:

- The Cancer Genome Atlas (TCGA) database (https://portal.gdc.cancer.gov/) (17) provided gene expression profiles and corresponding survival data from 371 HCC patients.
- 2. The Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/) contributed two key datasets. GSE112271 (18) contains single-cell RNA sequencing data from 7 HCC samples, which was utilized to provide insights at the cellular level. GSE14520 (19) encompasses tissue sequencing data and survival information from 242 HCC patients and was employed as an independent validation cohort.
- The International Cancer Genome Consortium (ICGC)
  platform (https://dcc.icgc.org/) (20) supplemented our
  study with transcriptome data and clinical follow-up
  information from 445 HCC patients.

Raw count data were converted to TPM values, log2-transformed [log2(TPM + 1)], and subjected to stringent quality control. Only samples with complete genomic and clinical data were

retained. Additionally, we corrected batch effects across cohorts using the ComBat algorithm prior to integration analysis. Specifically, the ComBat algorithm (from the sva R package, version 3.48.0) was applied to adjust for batch effects arising from the different data sources (TCGA, GEO, ICGC). The batch covariate was explicitly defined as the dataset of origin. ComBat was run using its standard parametric empirical Bayes framework (par.prior=TRUE) to model and adjust for location and scale shifts between batches. Prior to correction, genes with zero variance across samples were removed to ensure algorithm stability. The success of integration and removal of major technical artifacts were assessed by performing Principal Component Analysis (PCA) on the expression matrix before and after correction. The resulting PCA plots were visually inspected, which confirmed substantial reduction in batch-specific clustering and improved mixing of samples from different cohorts post-correction.

### 2.2 Single-cell analysis

### 2.2.1 Data preprocessing

Single-cell data preprocessing was performed using Seurat package (v5.0.0) (21). Quality control filters included: minimum 3 cells per gene, minimum 250 genes per cell, mitochondrial gene expression <15%, and total RNA counts >1000. This yielded a reliable dataset for downstream analysis.

### 2.2.2 Data normalization and dimensionality reduction

Data normalization is a critical step in single-cell transcriptomic analysis. The "LogNormalize" method (22) was adopted for data normalization, using 10000 as the scaling factor, and 2000 highly variable genes were selected through the "vst" method. To eliminate batch effects, we performed the Harmony algorithm (harmony v1.0) (23) for integration of data from multiple samples with the following optimized parameters: group.by.vars="orig.ident" for batch variable specification, assay.use="SCT" for normalized data input, and max.iter.harmony=20 for convergence optimization. For both bulk and single-cell data integration, we systematically evaluated batch correction effectiveness through: (1) Principal component analysis (PCA) visualization to assess sample clustering patterns before and after correction; (2) UMAP dimensional reduction plots to confirm elimination of sample-specific clustering while preserving biological cell type distinctions; (3) For single-cell data, successful batch integration was validated by examining the mixing of samples from different batches in the same cell type clusters, ensuring that technical variation was removed while biological heterogeneity was maintained.

Following successful integration, we proceeded with downstream analysis. PCA was performed on the integrated data, and the top 30 principal components were selected based on elbow plot analysis for subsequent dimensionality reduction. Cellular topology and heterogeneity were effectively visualized using UMAP and t-SNE methods based on the Harmony-corrected embeddings (24).

### 2.2.3 Cell type annotation

Cell type annotation was performed using a combination of automatic and manual approaches. First, we used the SingleR package (25) and the Human Cell Atlas database (https://www.humancellatlas.org/) to carry out automatic cell type annotation. The specific marker genes for each of the initial clusters are provided in Supplementary Table S2. Next, manual validation was performed by detecting reported cell type-specific markers, including ALB and SERPINA1 (Hepatocyte), GPC3 (Cancer cell), AQP1, TIE1, VWF, EDNRB, CCL14 (Endothelial cell), ACTA2, COL1A1, DCN, COL1A2 (Fibroblast cell), FOLR2, AIF1, CD68 (Macrophage), NKG7, GNLY, CCL5 (NKT cell).

### 2.2.4 Analysis of immune cell death scores

Using single-sample gene set enrichment analysis (ssGSEA) (26), we quantified ICD scores for individual cells based on established ICD gene sets (Supplementary Table S3) (27). To comprehensively evaluate ICD scores distribution patterns across cell types, Kruskal-Wallis rank sum test was used to assess overall differences. For significant differences (p<0.05), we performed pairwise comparisons with Wilcoxon rank sum test, applying Bonferroni correction for multiple testing. Box plots with significance indicators were generated comparing macrophages to other cell populations. Furthermore, we stratified cancer cells and macrophages into high and low groups based on median ICD scores, then identified DEGs using the FindAllMarkers function (minimum expression threshold=0.35). This approach aimed to reveal molecular characteristics and cellular heterogeneity of ICD at single-cell resolution in HCC.

## 2.3 Weighted gene co-expression network analysis

To explore the relationship between ICD scores and gene expression patterns, weighted gene co-expression network analysis (WGCNA) was performed on published ICD-related gene sets (27) using the WGCNA R package (28). TCGA-HCC data was preprocessed by filtering zero-variance genes and outlier samples (29). For the scale-free topology network construction, a soft threshold power of five was chosen. This was the lowest power at which the network's scale-free topology fit index (R<sup>2</sup>) first reached the standard threshold of 0.85, ensuring a balance between network properties and connectivity. Other key parameters included a minimum module size of 50 and a merge cut height of 0.15. Subsequently, gene modules were identified by applying the dynamic tree cutting algorithm, and module eigengenes (MEs) (30) were calculated. We further analyzed associations between modules and ICD scores through Pearson correlation, and identified significant modules using Student's ttest. Finally, gene significance (GS) and module membership (MM) were calculated and visualized in scatter plots to identify key genes in significant modules, termed module genes. Additionally, differential expression analysis between TCGA-HCC samples and normal samples was performed to identify TCGA-DEGs, visualized

using an enhanced volcano plot and a circular heatmap displaying the top 50 up-regulated and top 50 down-regulated DEGs. Finally, we acquired a set of genes related to ICD in HCC (HCC-ICDR genes) by intersecting TCGA-DEGs with module genes obtained from WGCNA.

## 2.4 Integration and comparison of multiple machine learning models

To identify the most robust prognostic prediction model, various machine learning algorithms and their combinations were systematically evaluated for HCC prognosis prediction performance. We first perform batch effect correction on TCGA, GSE14520, and ICGC datasets. The expression matrices underwent standardization processing through the ComBat algorithm (31), which eliminated batch differences from different data sources. Furthermore, TCGA served as the training set, and the GSE14520 and ICGC datasets served as external testing sets to ensure the robustness of the model. Basic algorithms were tested in this study included Random Survival Forest (RSF), Elastic Net (Enet), Stepwise Cox regression (StepCox), CoxBoost, Partial Least Squares Cox regression (plsRcox), Super Principal Component analysis (SuperPC), Gradient Boosting Machine (GBM), Survival Support Vector Machine (survival-SVM), Ridge regression, and Lasso regression. Algorithm-specific parameters were optimized: CoxBoost (penalty coefficient and iteration steps), GBM (interaction depth=3, minimum observations=10, optimal tree number via cross-validation). In addition, we investigated combination strategies of basic algorithms, such as RSF+GBM, RSF+Lasso, and CoxBoost+GBM, ultimately evaluating up to 114 different model combinations. The Concordance index (C-index) (32) was adopted as the primary evaluation metric, which measures the accuracy of predicted survival time rankings. We constructed each model on the training set, and tested its generalization ability on two independent test sets. Different models' C-indices across datasets were visualized through heatmaps, and models were ranked according to the average C-index values of validation sets to select the final prognostic prediction tool with optimal predictive performance and stability. We finally selected the optimal model based on internal validation performance, external validation results across two independent cohorts, and model stability across different datasets.

## 2.5 Establishment and validation of the consensus signatures

Based on the selected optimal model, namely the RSF model, this study identified the top 10 key features with the strongest prognostic predictive power from candidate features by evaluating feature importance, termed the ICD-related signatures (ICDRS), to predict overall survival (OS) of HCC patients. Specifically, key gene markers were systematically ranked and identified by analyzing each feature's contribution to model prediction accuracy. To visually demonstrate key predictors, we plotted feature importance bar charts, with features sorted and visualized according to their contribution to

model predictions. The risk score for each patient, termed ICDRS, was derived directly from the trained RSF model. Specifically, we used the predict function from the randomForestSRC package to obtain the predicted mortality risk for each sample. This output, denoted as predicted, represents the ensemble mortality estimate from all trees in the forest and serves as the continuous risk score. A higher ICDRS indicates a greater probability of experiencing the event (death). Patients were then dichotomized into high- and low-risk groups using the median ICDRS of the training cohort (TCGA) as the cutoff threshold.

To comprehensively evaluate the model's predictive performance, a multi-faceted validation strategy was adopted. First, we thoroughly explored survival differences between different risk stratifications through Kaplan-Meier survival curve analysis. Subsequently, time-dependent receiver operating characteristic (ROC) curve analysis (33) was introduced to calculate the area under the curve (AUC) for 1-year, 3-year, and 5-year predictions, fully reflecting the model's predictive accuracy at different follow-up time points. Meanwhile, we applied the model to two independent validation sets, and performed identical survival and ROC analyses to verify the model's generalization ability.

## 2.6 Clinical feature correlation and survival analysis

To thoroughly evaluate the clinical utility of ICDRS, a multidimensional analysis was conducted on the TCGA-HCC dataset, encompassing correlation studies with clinical features and survival analysis. First, we constructed a circos plot of clinical characteristics to visually demonstrate the distribution patterns of TNM staging, age, gender, and survival status among different risk groups, with chi-square tests being employed to assess the significance of intergroup differences. The distribution characteristics of risk scores across different T stages were analyzed in depth through violin plots and box plots, with statistical differences evaluated based on the Wilcoxon rank-sum test. We then created stacked bar charts to illustrate the proportion of clinical features in high and low risk groups, comprehensively elucidating the association between risk scores and tumor staging. Gene expression data were also analyzed, with heatmaps being generated to display differential gene expression, intuitively revealing the connection between risk scores and gene expression profiles. Furthermore, we developed a logistic regression model to predict M staging, with its predictive performance assessed through ROC curves. Finally, stratified Kaplan-Meier survival analyses were performed according to age and clinical staging to compare survival differences between high and low risk groups, aiming to comprehensively validate the prognostic capability of ICDRS.

### 2.7 Construction of nomographs

To further enhance the model's predictive accuracy and prognostic capability, a nomogram combining ICD and clinical

features was developed for quantifying the expected survival period of HCC patients. Key variables including age, gender, T stage, N stage, and M stage initially underwent univariate Cox regression analysis, which aimed to identify potential prognostic factors associated with overall survival. Subsequently, a multivariate Cox proportional hazards regression model was constructed, with the aforementioned clinical covariates being adjusted to determine independent prognostic factors. We then created a forest plot to visually demonstrate the prognostic impact, presenting the hazard ratios of each variable along with their 95% confidence intervals. Building upon this foundation, a comprehensive nomogram (34) integrating risk scores and key clinical parameters was developed to provide individualized predictions of 1-, 3-, and 5-year survival probabilities. We rigorously evaluated predictive accuracy and clinical utility of the nomogram at different time points through the calibration curves and decision curve analysis. The C-index was utilized to quantitatively measure the discriminative ability of the model, providing robust statistical validation for the prognostic model.

### 2.8 Functional enrichment analysis

To explore the biological significance of risk stratification based on ICDRS, multiple methods were employed for functional enrichment analysis. First, we performed the DEG analysis between high-risk and low-risk patient groups using the limma package. Subsequently, gene set enrichment analysis (GSEA) was conducted using the Hallmark gene sets from MSigDB (v2023.1) to investigate functional pathways of DEGs. For each gene set, we calculated normalized enrichment scores and significance levels after multiple testing correction (FDR q-value), and selected gene sets having FDR<0.05 for visualization. Second, gene set variation analysis (GSVA) was applied to quantify pathway activities in individual samples, and pathway enrichment results were visualized through t-value-based bar plots, highlighting riskassociated pathways. We then constructed the correlation heatmaps to intuitively demonstrate the relationships between pathway activities and risk scores. Finally, for significant pathways (log-rank p<0.05), hazard ratios (HR) and 95% confidence intervals were computed using Cox proportional hazards models.

## 2.9 Mutation analysis and heterogeneity assessment between the two risk groups

To investigate genomic heterogeneity features associated with ICD, we first calculated Mutant-Allele Tumor Heterogeneity (MATH) scores for each sample using the maftool package (35), then intuitively presented distribution characteristics through violin plots, and evaluated statistical significance using the Wilcoxon rank-sum test. Afterwards, patients were divided into high and low groups based on the median MATH score, followed by Kaplan-

Meier survival analysis to explore the association between tumor heterogeneity and prognosis. Subsequently, by combining MATH scores with risk scores, we further classified patients into four subgroups, aiming to comprehensively reveal their joint prognostic value. Meanwhile, mutation landscape analysis was performed for high and low risk groups separately, displaying the top 20 mutated genes through waterfall plots, and calculating tumor mutation burden (TMB). Using the "somaticInteractions" function, we conducted co-occurrence and mutual exclusivity analysis, which revealed interaction patterns of gene mutations in high and low risk groups.

## 2.10 Validation of risk signatures and analysis of intercellular communication based on single-cell data

To validate the biological significance of our constructed ICDRS at the single-cell level, I applied the 10 previously identified key genes to single-cell RNA sequencing dataset (GSE112271) for verification analysis. First, the expression distribution patterns of these 10 genes across different cell types in UMAP dimensionality reduction space were visualized using the "FeaturePlot" function. Based on ICDRS, we then calculated risk scores for each cell using the ssGSEA algorithm with a Poisson distribution kernel density function. The risk scores were standardized through Z-score normalization, and cells were classified into high-risk and lowrisk groups using a threshold of Z-score greater than 0. To explore functional differences between cells in different risk groups, we identified DEGs between risk groups using the "FindAllMarkers" function (logfc.threshold=0.35, min.pct=0.35) and revealed relevant biological processes through KEGG pathway enrichment analysis, with particular focus on biological pathways related to ICD. Simultaneously, GSEA was employed based on MSigDB Hallmark gene sets to identify biological pathways specifically enriched in the high-risk group.

Furthermore, to elucidate differences in communication patterns between different risk cancer cells and other cell types in the microenvironment, we constructed composite labels combining risk stratification and cell types, redefining cancer cells as "high-riskscore cancer cells" and "low-riskscore cancer cells" while maintaining original annotations for other cell types. Cell-cell communication analysis was performed using the CellChat package (36), identifying overexpressed ligand-receptor pairs using the human CellChatDB database and incorporating protein interaction network information. We then applied a minimum cell count threshold of 10 for communication filtering and calculated cell-cell communication probabilities and signaling pathway activities. Finally, for key signaling pathways such as MDK, VEGF, and MIF pathways, signaling communication heatmaps were generated to systematically compare differences in signaling intensity between high/low-riskscore cancer cells and microenvironment cells, comprehensively revealing the communication characteristics and potential biological mechanisms of cancer cells stratified by ICDRS in the TME.

## 2.11 Correlation analysis between tumor immune microenvironment characteristics and risk score model

To systematically evaluate the relationship between immune characteristics in the HCC tumor microenvironment and our constructed risk scoring model, multiple computational methods were employed for comprehensive analysis. First, we conducted the ESTIMATE score analysis (37) on HCC samples from the TCGA database using the IOBR package (38), calculating stromal score, immune score, and ESTIMATE score for each sample, and compared the differences between high and low risk groups. Stromal components and immune cell infiltration levels in tumor samples are assessed by the ESTIMATE algorithm through specific gene expression feature. We then used Wilcoxon rank-sum test to compare differences between high and low risk groups, and created the boxplots using the ggplot2 package (39) for visualization.

Thereafter, ssGSEA method was adopted for enrichment analysis of immune-related pathways. We integrated a series of immune-related pathway gene sets. Activity scores of these pathways in each sample were first calculated using the "gsva" function. We then computed the significance of pathway activity differences between high and low risk groups through the "diff\_pathway" function, while heatmaps were created using the pheatmap package (https://cran.r-project.org/web/ packages/pheatmap/index.html) to visually display differential patterns of immune pathway activity between different risk groups. To gain deeper insight into immune cell composition in the TME, CIBERSORT algorithm (40) was used to quantitatively analyze the relative abundance of 22 immune cell types in samples. This algorithm, based on linear support vector regression principles, can accurately estimate distribution proportions of various immune cell types. Then, violin plots were created using the "vioplot\_plot" function to visually compare differences in abundance of various immune cells between high and low risk groups. Additionally, we analyzed the correlations between 10 key genes in the risk model and immune cell infiltration, with correlation heatmaps visually displaying these complex association networks. A significance level of p<0.05 was used for all statistical analyses, with Benjamini-Hochberg method applied for correction during multiple comparisons.

To enhance reliability of analysis results, comprehensive evaluation was conducted in three aspects: (1) Spearman correlation analysis was performed to assess correlations between immune cell infiltration levels and risk scores, and correlation scatter plots were created using ggplot2; (2) Wilcoxon rank-sum test was used to compare differences in immune cell infiltration between high and low risk groups; (3) The relationship between immune cell infiltration levels and the patients' OS was evaluated through Kaplan-Meier survival analysis and log-rank test, generating independent survival curves for each immune cell type, and recording corresponding p-values and optimal cutoff points. We further established Venn diagrams using the VennDiagram package (https://cran.r-project.org/web/packages/ VennDiagram/index.html) to identify key immune cell types simultaneously satisfying these three conditions, which may play important roles in HCC development, progression, and prognosis.

### 2.12 Drug sensitivity analysis prediction

The Cancer Genomics Project 2016 (CGP2016) database was utilized to predict the sensitivity of high-risk and low-risk HCC patient samples to common anticancer drugs. We constructed a cell line-based ridge regression model using the pRRophetic package (41), and estimated the half-maximal inhibitory concentration (IC50) (42) values for each drug sample using the ICDRS from HCC. The specific analysis process follows: First, the CGP2016 dataset and corresponding gene expression data were loaded. Subsequently, we performed drug sensitivity predictions for each compound using the "pRRopheticPredict" function. To ensure data quality, only samples recorded in both drug sensitivity data and risk score data were included in the analysis. For each drug, we compared the differences in IC50 values between high-risk and low-risk groups using the Wilcoxon rank-sum test, with p<0.05 established as the statistical significance threshold. The median IC50 values for each risk group were calculated, and drug sensitivity differences were visualized through box plots. To facilitate interpretation of results, we sorted all analyzed drugs by P-value, and saved drug sensitivity results with statistical significance separately. This approach enables systematic evaluation of differential sensitivity patterns to anticancer drugs across different ICDRS risk groups, providing important reference for the development of individualized treatment strategies.

### 2.13 Functional verification analysis

### 2.13.1 Cell culture and vector construction

The human hepatocellular carcinoma cell line HepG2 was purchased from the Affiliated hospital of Qingdao university and maintained in DMEM medium (Gibco, USA) supplemented with 10% fetal bovine serum (FBS, Gibco, USA) and 1% penicillinstreptomycin (Invitrogen, USA) at 37 °C with 5% CO<sub>2</sub>. The complete coding sequences of human CLIC1 and NAP1L1 genes were amplified by PCR and then cloned into pcDNA3.1(+) eukaryotic expression vector (Invitrogen, USA) to construct overexpression plasmids pcDNA3.1-CLIC1 and pcDNA3.1-NAP1L1, with empty pcDNA3.1 vector serving as negative control. All recombinant plasmids were verified by DNA sequencing before being used for transfection experiments.

### 2.13.2 Establishment of stable cell lines

The complete coding sequences of human CLIC1 and NAP1L1 genes were cloned into the lentiviral vector pLVX-IRES-Puro (Clontech, USA) to construct recombinant plasmids pLVX-CLIC1 and pLVX-NAP1L1, with the empty vector pLVX-IRES-Puro serving as a negative control. We employed a three-plasmid system for lentiviral packaging: the recombinant plasmids were co-transfected with packaging plasmid psPAX2 and envelope plasmid pMD2.G into 293T cells at a mass ratio of 4:3:1, using Lipofectamine 3000 transfection reagent (Invitrogen, USA) according to the manufacturer's instructions. Virus-containing supernatants were collected at 48 and 72 hours post-transfection

and filtered through a 0.45 $\mu$ m filter to remove cellular debris. HepG2 cells were infected when they reached 70-80% confluence by adding virus-containing medium supplemented with polybrene (8 $\mu$ g/mL, Sigma, USA). We then replaced the medium with fresh culture medium 24 hours after infection. Puromycin (2 $\mu$ g/mL, Sigma, USA) was added 48 hours post-infection for positive clone selection, which continued for 10–14 days until stable expression cell lines were established. Moreover, we designated these cell lines as OE-CLIC1, OE-NAP1L1, and OE-NC (empty vector control), and verified stable expression of target proteins by Western blot.

## 2.13.3 Western blot analysis for protein expression

Total cellular proteins were extracted and quantified using standard protocols. Proteins (30 $\mu$ g) were separated by SDS-PAGE, transferred to PVDF membranes, and probed with primary antibodies against CLIC1, NAP1L1,  $\beta$ -tubulin, and GAPDH, followed by HRP-conjugated secondary antibodies. Protein bands were visualized by ECL and quantified using ImageJ software, with  $\beta$ -tubulin or GAPDH as loading controls.

### 2.13.4 CCK-8 cell proliferation assay

The stably transfected cells were seeded in 96-well plates at a density of  $3\times10^3$  cells per well, with six replicate wells for each group. After 24, 48, and 72 hours of culture, we added 10  $\mu$ L of CCK-8 reagent (Dojindo, Japan) to each well and incubated at 37 °C for 2 hours. The absorbance at 450 nm was measured using a microplate reader (BioTek, USA), and cell growth curves were plotted.

### 2.13.5 Clone formation experiment

The stably transfected cells were seeded in 6-well plates at a density of 1000 cells per well, with three replicate wells established for each group. We cultured cells under standard conditions for 14 days, during which we replaced the culture medium every 3 days. At the end of the experiment, cells were fixed with 4% paraformaldehyde for 20 minutes and stained with 0.1% crystal violet solution for 15 minutes, followed by thorough washing with PBS and air-drying. The number of colonies (defined as cell clusters containing  $\geq$ 50 cells) was counted under a microscope.

### 2.13.6 Transwell migration and invasion assays

 $5{\times}10_4$  cells were suspended in 200  $\mu L$  serum-free DMEM and added to the upper chamber of Transwell inserts (8  $\mu m$  pore size, Corning, USA), while 600  $\mu L$  of complete medium containing 10% FBS was placed in the lower chamber as a chemoattractant. After incubation at 37 °C for 24 hours, we gently removed the non-migrated cells on the upper surface with sterile cotton swabs. Cells that had migrated through the membrane were fixed with 4% paraformaldehyde for 20 minutes and stained with 0.1% crystal violet for 15 minutes. We then counted the migrated cells in five randomly selected fields (200× magnification).

For the invasion assay, the upper chamber was pre-coated with 50  $\mu L$  of diluted Matrigel (BD Biosciences, USA; 1:8 dilution) and

incubated at 37 °C for 1 hour to allow gelation. We performed the remaining experimental procedures as described for the migration assay.

### 2.13.7 Wound healing assay

The stably transfected cells were seeded in 6-well plates at a density of  $5{\times}10^5$  cells per well and cultured until cell confluence exceeded 90%. We created a straight line wound on the cell monolayer using a sterile 200  $\mu L$  pipette tip, followed by gentle washing with PBS three times to remove detached cells and debris. The medium was then replaced with serum-free DMEM for continued culture. We took the photographs at the same position at 0, 24, and 48 hours to document wound healing (100× magnification). ImageJ software was used to measure the wound area, and the healing rate was calculated as: Healing rate (%)=(Initial wound area - Wound area at detection time point)/Initial wound area  $\times$  100%.

### 2.14 Statistical analysis

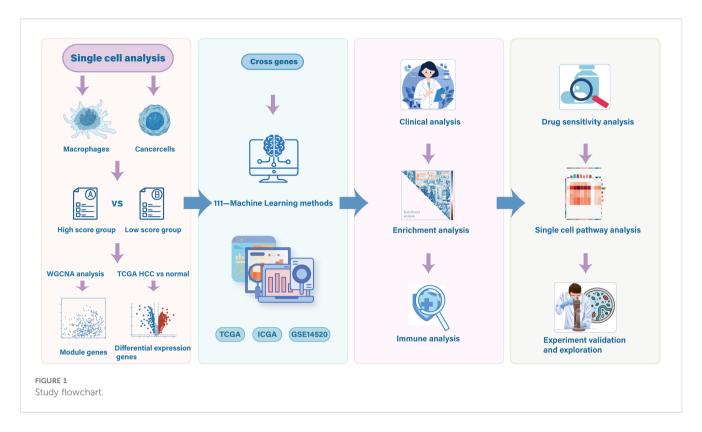
All experiments were independently repeated at least three times, and data are presented as mean  $\pm$  standard error of the mean (SEM). We performed the statistical analyses using GraphPad Prism 8.0 software. Comparisons between two groups were analyzed using Student's t-test, while comparisons among multiple groups were conducted using one-way ANOVA followed by Tukey's multiple comparison test. p<0.05 was considered statistically significant, with significance levels indicated as:  $^*p<0.05$ ,  $^{**}p<0.01$ ,  $^{***}p<0.001$ ,  $^{***}p<0.0001$ , ns means no significant difference.

### 3 Results

An overview of the study design is presented in Figure 1.

## 3.1 Single-cell transcriptome reveals ICD characteristic in HCC

Utilizing single-cell RNA sequencing, we comprehensively examined the expression landscape of ICD across different cell types. Eleven distinct cell clusters were initially identified, and their spatial distribution was visualized through UMAP (Figure 2A). We subsequently employed canonical marker gene expression profiles to identify and characterize major cell populations using UMAP dimensionality reduction. Six primary cell types were successfully delineated (Figures 2B, C), encompassing a total of 44,461 cells: (1) cancer cells expressing ALB, SERPINA1, and GPC3; (2) macrophages with high expression of FOLR2, AIF1, and CD68; (3) endothelial cells specifically expressing AQP1, TIE1, VWF, EDNRB, and CCL14; (4) fibroblasts enriched with ACTA2, COL1A1, DCN, and COL1A2; (5) hepatocytes with high expression of ALB and SERPINA1; and (6) NKT cells specifically expressing NKG7, GNLY, and CCL5. Additionally, we quantified



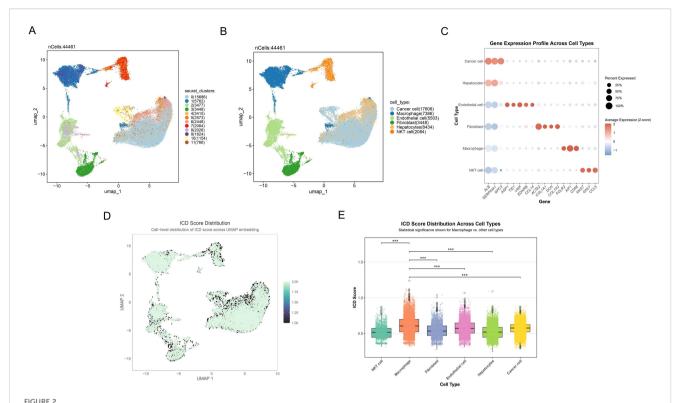
the activity of ICD in different cell types, presenting the continuous distribution of ICD scores using UMAP (Figure 2D). Statistical analysis (Figure 2E) further revealed that immune cell like macrophages exhibited significantly highest ICD scores compared to other cell types (p<0.001).

## 3.2 WGCNA network analysis identifies ICD-DEGs in bulk RNA sequencing

In the study of ICD in HCC, the TCGA-HCC dataset was analyzed using WGCNA to identify and characterize ICD-DEGs between different ICD score groups. Initially, we analyzed DEGs by comparing macrophages and cancer cells across ICD score stratified groups. A total of 317 common DEGs were then identified (Figure 3A), comprising 710 DEGs from macrophages and 964 DEGs from cancer cells, revealing shared molecular signatures between these cell types. To further explore the molecular mechanisms of ICD, we conducted WGCNA on the common DEGs. Intricate sample clustering patterns and ICD score distributions were revealed through the hierarchical clustering dendrogram (Figure 3B). Furthermore, the dynamic tree-cutting algorithm was applied to identify three distinct functional gene modules, as visualized in the cluster dendrogram (Figure 3C). Notably, module-trait relationship heatmap demonstrated that the turquoise module exhibited the most significant positive correlation with ICD traits (cor=0.4, p=2e-15), while the blue module revealed a pronounced negative association (cor=-0.37, p=2e-13, Figure 3D). Moreover, significant positive correlations between GS and MM were revealed in the turquoise (cor=0.17, p=0.047, Figure 3E) and blue (cor=0.25, p=0.01, Figure 3F) modules, suggesting functional coherence related to ICD. To narrow down the candidate gene pool, enhanced volcano plot showed the TCGA-DEGs between TCGA-HCC samples and normal sample (Figure 3G). Subsequently, the circular heatmap was conducted to further show top 100 regulated DEGs (Figure 3H). Moreover, a Venn diagram (Figure 3I) revealed 106 intersecting genes, termed HCC-ICDR genes, between the identified modules and TCGA-DEGs. These genes demonstrated significant involvement in ICD mechanisms across both whole-tissue and single-cell transcriptomic levels.

## 3.3 Prognostic feature selection and validation in HCC using machine learning

We developed a consensus signature (ICDRS) using integrated machine-learning algorithms. The RSF algorithm achieved the highest C-index (0.839) with parameters: ntree=1000, nodesize=5, and splitrule="logrank" (Figure 4A, Supplementary Table 1). This selection was further supported by consistent external validation performance (GSE14520: AUC 0.809-0.839; ICGC: AUC 0.821-0.832) and RSF's methodological advantages, including stable feature selection through ensemble mechanisms and built-in importance ranking that eliminates additional computational overhead required by other algorithms. Moreover, the top 10 features including CLIC1, NAP1L1, CBX3, RAN, APOE, CD63, CLTA, SNRPG, FTL and POMP in the RSF model were systematically identified and ranked based on their variable importance, showing the high relative importance (Figure 4B). To rigorously evaluate the prognostic potential of ICDRS, Kaplan-



Single-cell transcriptome reveals ICD characteristic in HCC. (A) Umap plot reflecting the partitioning of all cells into 11 distinct clusters. Different colors represent various cell clusters. (B) Umap plot showing the 6 distinct cell types identified with the DEGs across 11 clusters. Different colors represent distinct cell types. (C) Dot plot depicting the gene expression profile across cell types. Dot size indicates the percentage of cells expressing the genes, and dot color represents the average expression level, with red indicating positive expression and blue representing negative expression. (D) Cell-level distribution of ICD score across UMAP embedding. Darker colors represent lower ICD scores, while lighter colors indicate higher ICD scores. (E) Box plot showing the ICD score distribution across different cell types, highlighting significant differences between macrophages and other cell populations. \*\*\*p<0.001.

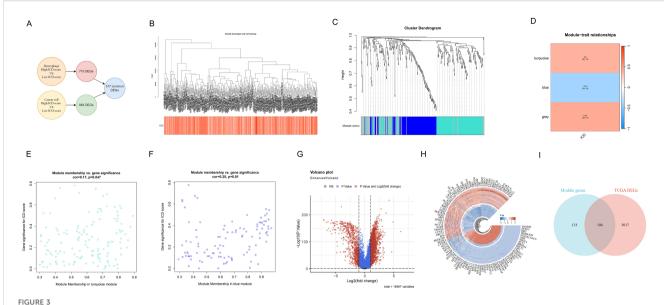
Meier survival analyses were conducted across three cohorts. Patients were categorized into distinct risk groups, which unveiled statistically significant survival differences in the TCGA training set (p<0.001) (Figure 4C), with consistent findings observed in the subsequent testing sets (all p<0.001, Figures 4E, G).

To systematically assess the time-dependent predictive performance of the prognostic model, time-dependent ROC curves were generated at 1-, 3-, and 5-year intervals. The AUC values for the TCGA dataset demonstrated robust predictive accuracy, with 0.962 (1-year), 0.985 (3-year), and 0.972 (5-year) (Figure 4D). The predictive performance was then confirmed in two validation datasets. Specifically, the GSE14520 validation dataset exhibited AUC values of 0.809 (1-year), 0.839 (3-year), and 0.819 (5-year) (Figure 4F). Similarly, the ICGC validation set showed AUC values of 0.821 (1-year), 0.832 (3-year), and 0.796 (5-year) (Figure 4H), demonstrating robust predictive consistency and significant clinical utility.

## 3.4 Performance evaluation and clinical relevance analysis of ICDRS in HCC

We comprehensively evaluated the clinical utility of the ICDRS for HCC. Initially, significant differences in clinical characteristics

were revealed through the pie chart (Figure 5A) between high-risk (n=183) and low-risk (n=182) groups, including T stage (p<0.001), gender (p<0.0267), and survival status (p<0.001). Furthermore, Figure 5B illustrated the risk score distributions across T1-T4 stages, while Figure 5C compared early (T1-2) and late (T3-4) stages. Notably, a significant upward trend in risk scores was observed as tumor staging progressed. The stacked bar plot (Figure 5D) then visually presented T stage proportions across different risk groups. A higher proportion of late-stage T classifications (T3-T4) was significantly concentrated in the highrisk group, suggesting that high-risk patients may face more severe tumor progression and adverse prognosis. Additionally, Figure 5E illustrated that the gene variables ultimately selected for the model were generally upregulated in the high-risk group, providing crucial insights into the biological underpinnings of ICDRS. We employed the ROC curve (Figure 5F) to assess the model's performance in predicting distant metastasis, with an AUC of 0.752. Additionally, Kaplan-Meier survival curves (Figures 5G-J) demonstrated significantly higher survival probabilities (all p<0.001) for the low-risk group across various clinical subgroups, including early (I-II) and late (III-IV) stages, as well as age-stratified cohorts (≤60 and >60 years). High-risk patients consistently exhibited markedly shorter survival periods across all subgroups, further validating the prognostic value of ICDRS in HCC.



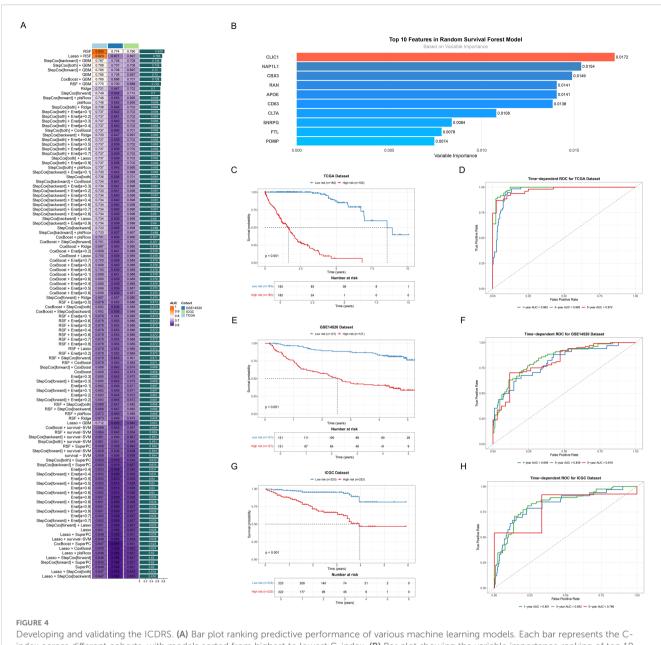
WGCNA network analysis reveals key gene modules of ICD. (A) Common differential gene expression analysis in macrophages and cancer cells stratified by ICD scores. (B) Hierarchical clustering dendrogram (upper) showing sample relationships and heatmap (lower) revealing ICD score distribution. Color intensity represents score levels. (C) Cluster dendrogram illustrating gene relationships, with colored bands at the bottom representing distinct functional gene modules. (D) Heatmap showing module-trait relationships. Each row depicts a co-expression module, with numerical values and color intensity reflecting correlational strength to ICD scores. (E, F) Scatter plots showing MM-GS correlation of the (E) turquoise and (F) blue modules. X-axis: module membership; Y-axis: gene significance for ICD score. (G) Enhanced volcano plot of DEGs between TCGA-HCC samples and normal samples. Red and blue dots represent statistically significant upregulated and downregulated genes, with total 18,367 variables analyzed. (H) Circular heatmap of the top 50 up-regulated (red) and 50 down-regulated (blue) DEGs. Inner to outer rings show gene names, and expression change direction. (I) Venn plot showing the intersecting genes between the module genes and TCGA DEGs in bulk RNA-seq.

## 3.5 Construction and validation of a prognostic nomogram model integrating ICDRS and clinical features

To evaluate the potential of ICDRS as an independent prognostic factor for HCC, we conducted comprehensive cox regression analyses that integrated clinical parameters (age, gender, TNM stage, clinical stage, grade) with the risk score. As shown in Figure 6A, T stage (p<0.001), M stage (p<0.019), clinical stage (p<0.001), and risk score (p<0.001) were identified as potential prognostic factors associated with OS. The subsequent multivariate analysis (Figure 6B) further confirmed that age (p<0.012), and risk score (p<0.001) still significantly influenced OS after adjusting for other clinical characteristics, serving as truly independent prognostic factors. A prognostic scoring nomogram was constructed based on the ICDRS and clinical characteristics (Figure 6E). Furthermore, the calibration curves showed excellent alignment between predicted and observed 1-year, 3-year, and 5-year survival rates (Figure 6C). Subsequently, the nomogram's superior net benefit within specific high-risk thresholds was demonstrated by decision curve analysis compared to individual clinical characteristics (Figure 6D). Finally, the comparative analysis of the C-index (Figure 6F) further confirmed the nomogram's enhanced predictive capability for OS, outperforming individual clinical features.

## 3.6 Transcriptomic characteristics across risk score patients based on ICDRS

To further explore the molecular mechanisms underlying the correlation between (ICDRS) and HCC prognosis, we conducted comprehensive functional enrichment analyses. The significant enrichment of five cancer-related hallmark pathways were revealed in the high-risk group through GSEA, including DNA repair, E2F targets, MYC targets V1, PI3K/AKT/MTOR signaling, and reactive oxygen species pathway (Figure 7A, FDR < 0.05). GSVA then uncovered multiple significantly upregulated pathways in the high-risk group, including: (1) cell cycle and proliferationrelated pathways: MYC targets V1/V2, G2M checkpoint, Mitotic spindle, and E2F targets; (2) stress and microenvironment-related pathways: DNA repair, reactive oxygen species pathway, hypoxia, and unfolded protein response; (3) signal transduction pathways: MTORC1 signaling and PI3K/AKT/MTOR signaling (all adjusted p<0.05, Figure 7B). Correlation analysis between risk scores and pathway activities (Figure 7C) validated these findings. Moreover, the forest plot (Figure 7D) demonstrated that pathways enriched in the high-risk group, including G2M Checkpoint, E2F Targets, Glycolysis, and DNA Repair, were associated with higher HR, suggesting these pathways may be closely linked to poor prognosis.

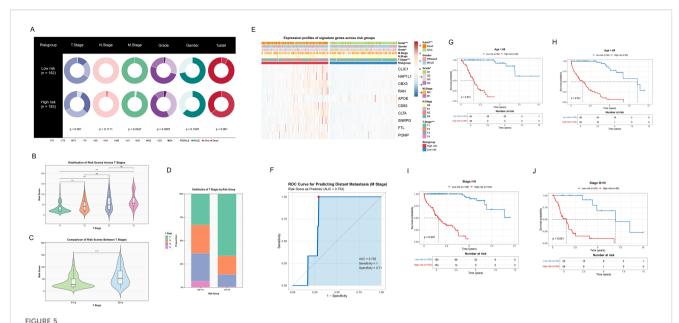


# Developing and validating the ICDRS. (A) Bar plot ranking predictive performance of various machine learning models. Each bar represents the C-index across different cohorts, with models sorted from highest to lowest C-index. (B) Bar plot showing the variable importance ranking of top 10 features in RSF model. (C, E, G) Kaplan-Meier survival curves of all datasets. Curves respectively showed survival differences between high (red line) and low risk (blue line) groups based on risk scores for (C) TCGA, (E) GSE14520 and (G) ICGC dataset. (D, F, H) Time-dependent ROC curves of all datasets. Curves demonstrated the predictive accuracy of the model at 1-, 3-, and 5-year time points for (D) TCGA, (F) GSE14520 and (H) ICGC dataset.

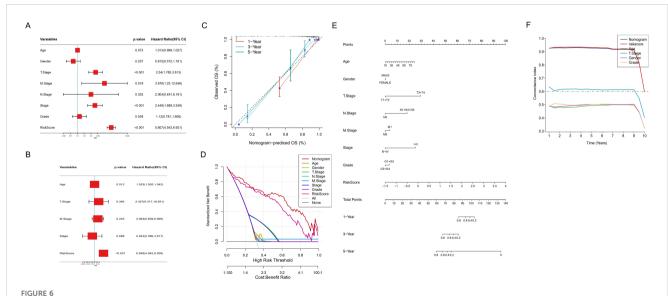
### 3.7 Mutation profiling of ICD related genes

To deeply understand the genomic characteristics of HCC patients, we systematically analyzed the mutation patterns of ICD genes. Firstly, the MATH scores in the high-risk group were significantly higher than those in the low-risk group (p=7.1e-06), indicating greater tumor heterogeneity in the high-risk group (Figure 8A). Subsequently, Kaplan-Meier survival analysis based on MATH scores revealed that patients with higher MATH scores had poorer prognosis (p=0.043, Figure 8B), further confirming the correlation between tumor heterogeneity and prognosis. Moreover, we combined MATH scores

with ICDRS risk stratification, demonstrating their interactive predictive impact on prognosis. The survival curves showed that the low-risk and low-MATH score group had the best prognosis (p<0.001, Figure 8C). Moreover, differential mutated genes between low-risk and high-risk groups were revealed by a distinct mutational landscape analysis (Figures 8D, E), with significant co-occurring mutations being observed (Figures 8F, G). Notably, TP53, a critical tumor suppressor gene, showed a mutation rate of 36% in the high-risk group compared to 21% in the low-risk group, indicating accelerated tumor proliferation and poorer prognosis in the high-risk group. Ultimately, mutation co-occurrence and exclusivity analysis revealed the complex interaction



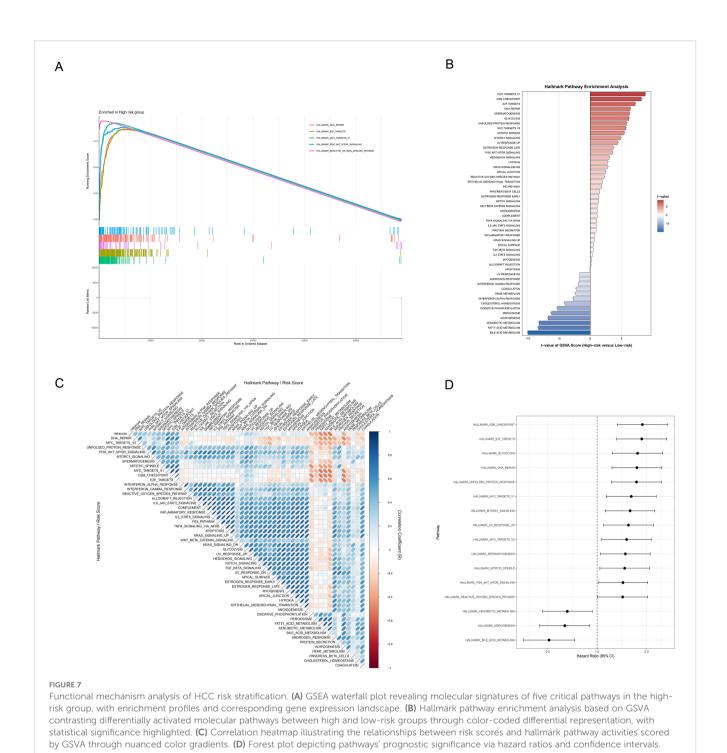
Assessing the performance and clinical utility of the ICDRS in HCC. (A) Pie charts revealing distribution of clinical features across low- and high-risk groups. (B) Violin integrated with box plots showing distribution of risk scores at different T stages, including T1, T2, T3, and T4. Non-significant differences are marked as "ns". (C) Violin integrated with box plots demonstrating the comparison of risk scores between different T stages, including T1–2 and T3–4. (D) Stacked bar plot showing the proportion of T stage distribution in risk groups. (E) Heatmap showing expression profiles of signature genes and clinical features across risk groups. (F) ROC curve for predicting distant metastasis (M stage). (G, H) Kaplan-Meier survival curves after age stratification, including age  $\leq$  60 and age>60. (I, J) Kaplan-Meier survival curves after stage stratification, including early stage (stage I+II) and later stage (stage III+IV).



Comprehensive Cox regression and nomogram analysis. (A) Forest plot of univariate cox regression analysis of clinical variables and risk score. (B) Forest plot of multivariate cox regression analysis revealing adjusted hazard ratios for clinical variables and risk score. (C) Calibration curve of the nomogram for 1, 3, and 5-year observed OS. Nomogram-predicted versus observed OS probabilities: red (1-year), blue (3-year), green (5-year) survival curves. (D) Decision curve analysis depicting prognostic model utility across risk thresholds. (E) Nomogram integrating risk score and clinical characteristics. (F) Comparison of the C-index between the nomogram and clinical characteristics.

patterns of gene mutations in high-risk and low-risk groups. In the high-risk group (Figure 8F), (1) TP53 exhibited a co-occurrence pattern with FAT3 and PCLO; (2) TTN and DOCK2 demonstrated significant co-occurrence characteristics; (3) no significant exclusivity

patterns were observed. In the low-risk group (Figure 8G), (1) CTNNB1 showed an exclusivity pattern with AXIN1 and TP53; (2) TTN and ALB displayed mutation co-occurrence features. Notably, key tumor suppressor genes like TP53 exhibited significant differences in

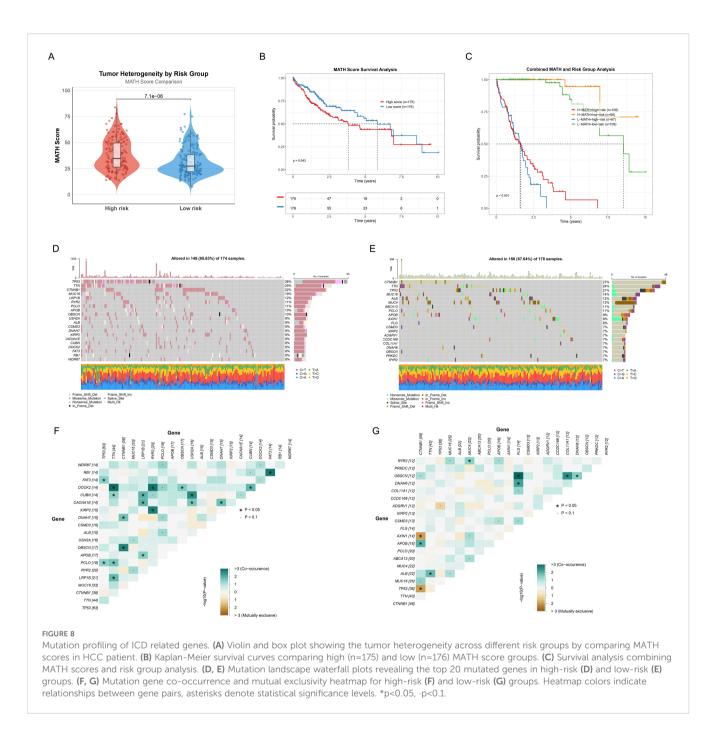


mutation patterns across different risk groups, these findings unveiling the molecular heterogeneity of ICD-related gene mutations in HCC.

## 3.8 Correlation analysis between ICDRS and single-cell features

To investigate the role of ICDRS in the tumor microenvironment (TME) at the single-cell level, we applied the established model to evaluate individual tumor cells using the top 10 genes. The top 10

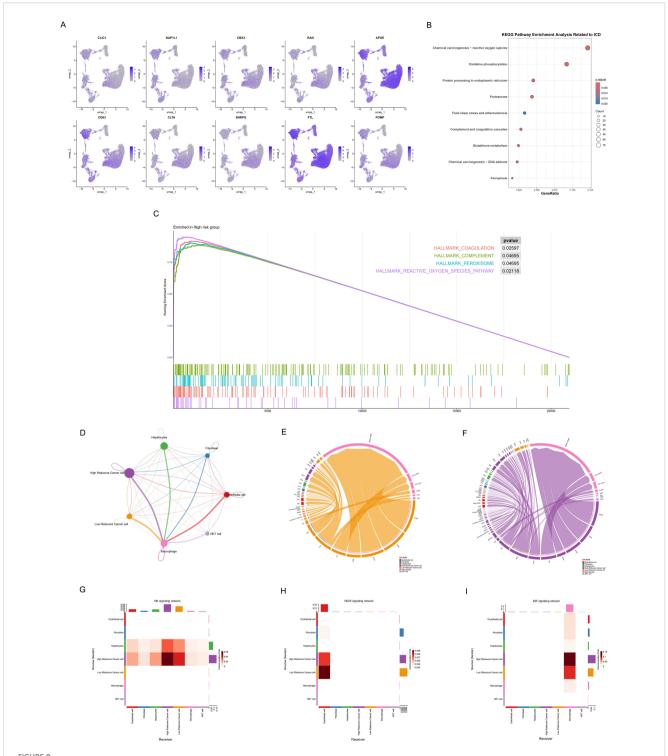
upregulated and downregulated genes from the RSF model were integrated, including CLIC1, NAP1L1, CBX3, RAN, APOE, CD63, CLTA, SNRPG, FTL, and POMP. We performed a comprehensive analysis of ICDRS expression and functional associations across different single-cell types (Figure 9). The expression patterns of these 10 genes across various cell types were determined (Figure 9A), revealing their predominant expression in cancer cells and macrophages. KEGG pathway enrichment analysis was conducted to identify the major functional pathways involving differentially expressed genes between high-risk and low-risk cells



(Figure 9B). We conducted KEGG pathway enrichment analysis to identify the major functional pathways involving DEGs between high-risk and low-risk cells (Figure 9B). Multiple important biological processes and signaling pathways related to ICD were significantly enriched, including Chemical carcinogenesis – reactive oxygen species, Oxidative phosphorylation, Protein processing in endoplasmic reticulum, Proteasome, Complement and coagulation cascades, Glutathione metabolism, and Chemical carcinogenesis – DNA adducts pathways. These pathways primarily involve oxidative stress, mitochondrial function, endoplasmic reticulum stress, protein degradation, cell death, and immune response biological processes, indicating that DEGs play important roles in ICD mechanisms. Through GSEA analysis, we further discovered multiple

significantly enriched HALLMARK pathways in the high-risk group, including coagulation (p=0.03), complement (p=0.05), peroxisome p=0.05), and reactive oxygen species pathway (p=0.02). Key biological processes that high-risk cell populations may participate in were revealed by the enrichment of these pathways (Figure 9C).

Subsequently, cancer cells were divided into high- and low-riskscore groups, and their interactions with other cell types in the TME were investigated. We observed different communication patterns in cancer cells with varying ICDRS scores (Figure 9D), with seven cell subpopulations in the high-risk group exhibiting more complex interaction networks. More active communication networks between high riskscore cancer cells and macrophages,



Association of ICDRS with single-cell characteristics. (A) Umap plots showing the expression of 10 genes of ICDRS in various cell types, as analyzed by single-cell RNA sequences. (B) Dot plot revealing the KEGG pathway enrichment analysis of DEGs between high-risk and low-risk cells related to ICD. Dot size: gene count; color: adjusted p-value from 0.02 to 0.005; x-axis: gene ratio. (C) Waterfall plot presenting GSEA analysis of hallmark gene sets in the high-risk cells. (D) Cell signaling pathway network among 7 cell types. The thickness of the line indicates the interaction strength. 7(E, F) Ligand-receptor interactions in the (E) low- and (F) high – riskscore cancer cell. Both interaction numbers and interaction strengths are shown. (G-I) Heatmaps showing the role of different cell types in (G) MDK, (H) VEGF, and (I) MIF signaling networks. The Y-axis shows the signal transmitter cells and the X-axis represents the signal receiver cells. Shades of color indicate intensity of interaction, with darker reds indicating stronger communication.

endothelial cells were demonstrated. In contrast, the intercellular communication network of low riskscore cancer cells was significantly simpler, primarily manifesting as interactions with macrophages, while communication intensity with other cell types was markedly reduced. Ligand-receptor interaction diagrams (Figures 9E, F) also showed that the high-risk group possessed denser and more complex communication networks, suggesting that elevated ICD levels may enhance signal exchange between cells in the tumor microenvironment. Notably, significant specificity in the MDK signaling pathway was exhibited by the high-risk group compared to the low-risk group. Active MDK signal transduction in the high-risk group may indicate that these tumor cells possess stronger proliferative capacity, angiogenic potential, and microenvironmental regulatory ability, thereby promoting invasive tumor growth and distant metastasis.

To further explore pathway expression and correlations in different cell types, three pathways were selected for heatmap visualization, including the MDK pathway related to tumor invasiveness and poor prognosis, the VEGF pathway associated with angiogenesis, and the MIF pathway involved in immune regulation (Figures 9G-I). The MDK signaling pathway (Figure 9G) demonstrated more complex communication patterns in high-riskscore cancer cells, establishing bidirectional signal exchange with hepatocytes. In contrast, low-riskscore cancer cells showed significantly reduced activity as MDK signal senders, indicating their limited capacity for actively regulating the microenvironment. In the VEGF signaling pathway (Figure 9H), VEGF signals secreted by cancer cells were accepted by endothelial cells, which served as the primary signal receivers. Notably, lowriskscore cancer cells exhibited stronger activity as VEGF signal senders compared to high-riskscore cancer cells, transmitting more VEGF signals to endothelial cells. This difference may be attributed to distinct angiogenic regulatory mechanisms employed by the two groups. In the MIF signaling pathway (Figure 9I), both high- and low-riskscore cancer cells regulated macrophage function and polarization states through MIF secretion. More significant MIF signal sending capacity was demonstrated by high-riskscore cancer cells, indicating the stronger immune microenvironment regulatory ability.

## 3.9 Immune landscape associated with ICDRS in HCC

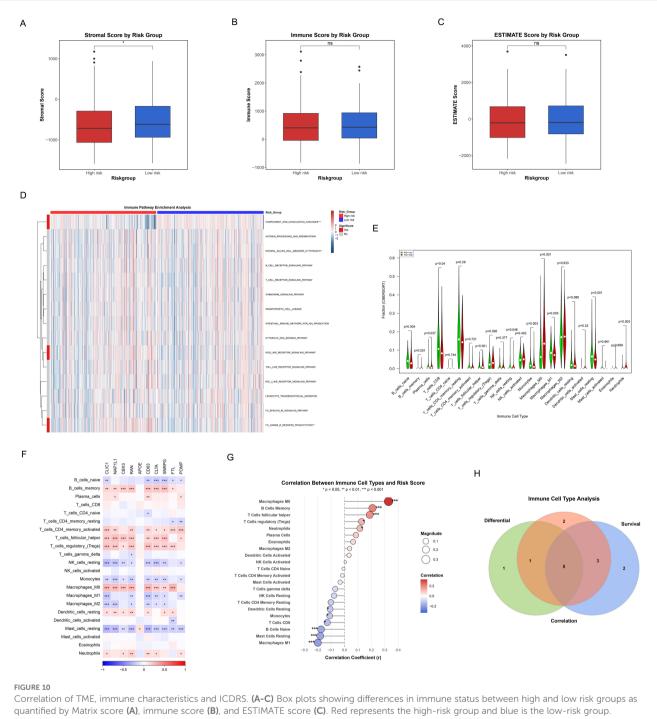
Figure 10 systematically demonstrated the significant differences in tumor microenvironment immune characteristics between high-risk and low-risk groups of HCC patients. Significantly higher Stromal Score in low-risk patients compared to high-risk patients (p<0.05) were revealed by ESTIMATE algorithm analysis (Figure 10A), suggesting that the TME in low-risk patients possesses a higher degree of stromal infiltration. However, no significant differences were observed between the two groups in terms of Immune Score (Figure 10B) and ESTIMATE Score (Figure 10C). Further immune-related pathway analysis (Figure 10D) identified three signaling pathways that

showed significant differences between high-risk and low-risk groups, mainly including: complement and coagulation cascades (p<0.001), NOD-like receptor signaling pathway (p<0.05), and fc gamma R mediated phagocytosis (p<0.01). These findings indicate that the activation status of these pathways differs across TME with varying risk levels.

Differences in immune cell infiltration patterns between the two patient groups were further elucidated by quantitative analysis of 22 immune cell types using the CIBERSORT algorithm (Figure 10E). The results demonstrated significant differences in the abundance distribution of multiple immune cell subsets between high-risk and low-risk groups: (1) Immune cell types significantly higher in the high-risk group included: plasma cells (p=0.037), T cells follicular helper (p=0.001), T cells regulatory (tregs) (p=0.026), macrophages M0 (p<0.001), and neutrophils (p=0.003); (2) Immune cell types significantly higher in the low-risk group comprised: B cells naive (p=0.004), B cells memory (p=0.001), T cells CD8 (p=0.04), NK cells resting (p=0.048), monocytes (p=0.003), macrophages M1 (p=0.003), and mast cells resting (p<0.001). More diverse and active immune cell repertoires, encompassing effector T cells, B cell subsets, monocytes, and M1 macrophages that represent anti-tumor immune cells, were demonstrated in low-risk patients. In contrast, the high-risk group was enriched with more immunosuppressive cells such as regulatory T cells and M0 macrophages. These differences in immune cell composition, particularly the enrichment of immunosuppressive cells alongside a relative deficiency in cytotoxic CD8+ T cells in the high-risk group (Figure 10E), may reflect the establishment of an immunosuppressive microenvironment prone to immune evasion mechanisms such as immune exclusion.

The complex association network between the 10 key genes embedded in ICDRS and various immune cell types was illustrated by the correlation heatmap (Figure 10F), with some genes such as CLIC1, NAP1L1, and CBX3 showing extensive positive correlations with multiple immune cell infiltration levels. A correlation scatter plot (Figure 10G) revealed that the risk score exhibited significant positive or negative correlations with specific immune cells, with macrophages M0, B cells memory, T cells follicular helper, T cells regulatory (tregs), and neutrophils showing significant positive correlations, while macrophages M1, mast cells resting, B cells naive, T cells CD8, monocytes, and dendritic cells resting displayed significant negative correlations.

Additionally, Kaplan-Meier survival analysis was performed for individual immune cells (Supplementary Figure 1), through which we identified immune cell types that were significantly associated with overall survival of HCC patients (p<0.05). Finally, by integrating the results from differential analysis (Figure 10E), correlation analysis (Figure 10G), and survival analysis (Supplementary Figure 1), we utilized a Venn diagram to identify 5 key immune cell types that simultaneously satisfied all three evaluation criteria (Figure 10H), including B cells memory, macrophages M0, macrophages M1, mast cells resting, and neutrophils. The core immune effector cells that influence HCC patient prognosis may be represented by these immune cells identified in the intersection, providing important clues for subsequent mechanistic research and therapeutic target screening.

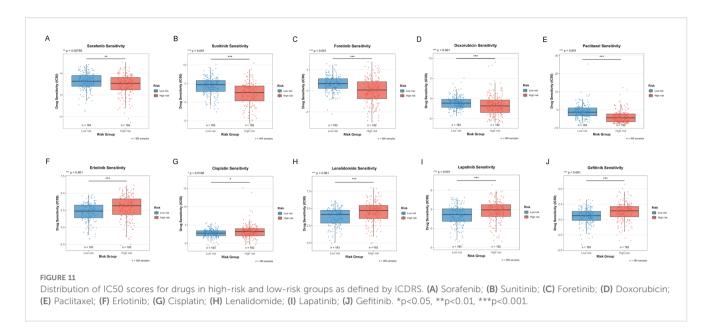


Correlation of TME, immune characteristics and ICDRS. (A-C) Box plots showing differences in immune status between high and low risk groups as quantified by Matrix score (A), immune score (B), and ESTIMATE score (C). Red represents the high-risk group and blue is the low-risk group. (D) Heatmaps reflecting differences in immune-related pathway activity between high and low risk groups. (E) Violin plots showing the level of immune cell infiltration between the high and low risk groups. Green represents the low-risk group, and red represents the high-risk group. (F) The correlation heatmap representing the correlation between the degree of immune cell infiltration and the hub genes of ICDRS. Red indicates the positive correlation, blue indicates the negative correlation, and the depth of the color represents the strength of the correlation. (G) The correlation scatter plot showing the association of the risk score with the infiltration level of key immune cells. The size of the dot represents the absolute value of the correlation coefficient, and the color indicates the direction of correlation and statistical significance. (H) Venn diagram representing the recognition of five key immune cells.

### 3.10 Differential analysis of drug sensitivity

The differences in drug sensitivity between high-risk and low-risk groups were investigated to assess the potential clinical value of the risk stratification model in personalized treatment. Drug sensitivity analysis revealed that 10 compounds exhibited significantly different responses between risk groups (Figures 11A-J).

In the treatment of HCC, increased sensitivity to targeted therapies was observed in the low-risk group, including:



- Sorafenib (first-line standard treatment for HCC, multikinase inhibitor, p=0.00785, Figure 11A)
- Sunitinib (multi-kinase inhibitor, investigational drug for HCC, p<0.001, Figure 11B)</li>
- Foretinib (c-Met/VEGFR dual inhibitor, p<0.001, Figure 11C)
- Doxorubicin (anthracycline chemotherapeutic agent, commonly used in TACE for HCC, p<0.001, Figure 11D)</li>
- Paclitaxel (microtubule stabilizer, taxane class, p<0.001, Figure 11E)</li>

By contrast, enhanced sensitivity to the following agents was demonstrated by the high-risk group:

- Erlotinib (EGFR inhibitor, p<0.001, Figure 11F)
- Cisplatin (standard agent for interventional therapy in HCC, platinum-based chemotherapy, p=0.0108, Figure 11G)
- Lenalidomide (immunomodulatory drug, p<0.001, Figure 11H)</li>
- Lapatinib (EGFR/HER2 dual inhibitor, targeted agent, p<0.001, Figure 11I)</li>
- Gefitinib (EGFR inhibitor, p<0.001, Figure 11J)

These findings not only provide a theoretical basis for risk stratification-based personalized drug strategies but also reveal that tumors with different molecular characteristics may require different therapeutic approaches.

## 3.11 Functional validation experiments of candidate genes

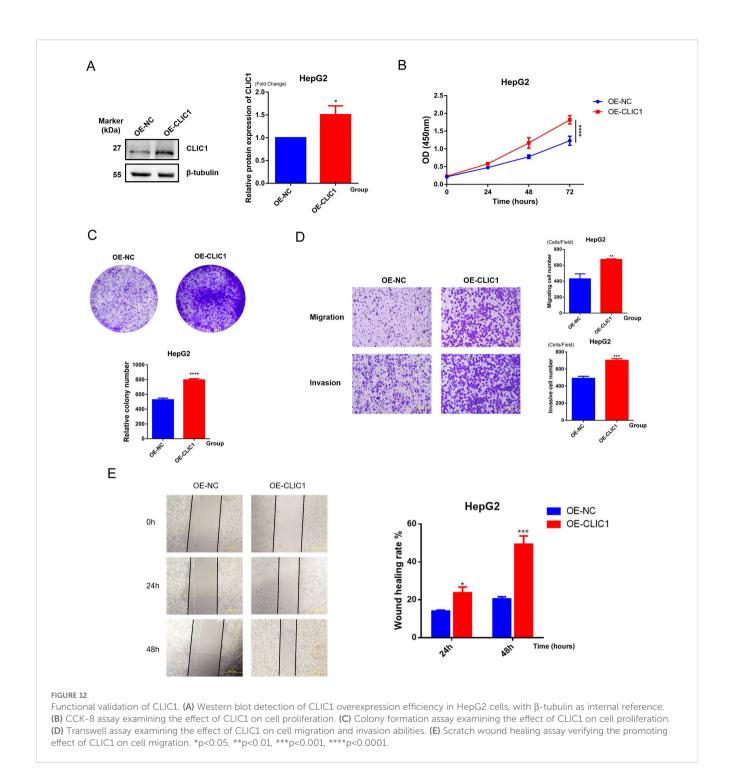
## 3.11.1 CLIC1 enhances malignant biological behaviors of HCC cells

Based on the high expression profile of CLIC1 in HCC, we constructed a CLIC1 overexpression cell line in HepG2 cells

(designated as OE-CLIC1) and established an empty vector control group (designated as OE-NC). All experiments were repeated at least three times with consistent results, and representative results are shown in Figure 12. To verify the transfection efficiency of the CLIC1 overexpression vector, we detected the CLIC1 protein expression levels by Western blot in HepG2 cells after transfection. Results demonstrated that CLIC1 protein expression was significantly upregulated by approximately 1.5-fold in OE-CLIC1 compared to OE-NC (p<0.05), while the expression level of the reference protein β-tubulin showed no significant difference between groups, indicating successful and efficient transfection of the CLIC1 overexpression vector (Figure 12A). The proliferative capacity of HepG2 cells was significantly promoted by CLIC1 overexpression as revealed by CCK-8 proliferation assay results, with particularly notable differences at the 72-hour time point (p<0.0001, Figure 12B). Colony formation assays further confirmed the proliferation-promoting effect of CLIC1, as the overexpression group formed significantly more colonies than the control group (p<0.0001, Figure 12C).

Given the potential role of CLIC1 in tumor metastasis, its effects on cell migration and invasion abilities were examined. Transwell assay results showed that CLIC1 overexpression significantly enhanced both migration (p<0.01) and invasion capabilities (p<0.001, Figure 12D) of HepG2 cells. These findings suggest that CLIC1 overexpression can significantly enhance the ability of hepatocellular carcinoma cells to traverse the extracellular matrix, indicating its potentially important role in the tumor metastasis process.

To further validate the effect of CLIC1 on cell migration ability, wound healing assays were performed. The results demonstrated that the wound closure rate of the OE-CLIC1 group was markedly faster than that of the OE-NC control group at both 24-hour and 48-hour observation time points. At 48 hours, approximately 50% wound healing rate was reached by the OE-CLIC1 group, while only about 20% was achieved by the OE-NC control group, with the difference between the two groups being highly statistically



significant (p<0.001). This result was consistent with the Transwell migration assay findings, further confirming that CLIC1 overexpression can significantly promote the migration ability of hepatocellular carcinoma cells (Figure 12E).

### 3.11.2 NAP1L1 promotes proliferation, migration, and invasion of HCC cells

NAP1L1-overexpressing cell lines (OE-NAP1L1) and control groups (OE-NC) were constructed by us, with experiments repeated three times, and representative results are shown in

Figure 13. Western blot revealed that the protein level of NAP1L1 in the OE-NAP1L1 group was elevated approximately 1.5-fold compared to the control group (p<0.01), while GAPDH expression remained stable, confirming successful transfection (Figure 13A). The biological effects of NAP1L1 were unveiled through cellular functional analyses. CCK-8 assay detected enhanced proliferative activity in the NAP1L1 high-expression group during the late culture period (72 hours) (p<0.0001, Figure 13B), and colony formation ability was also significantly enhanced (p<0.0001, Figure 13C), indicating that NAP1L1 can effectively promote HCC

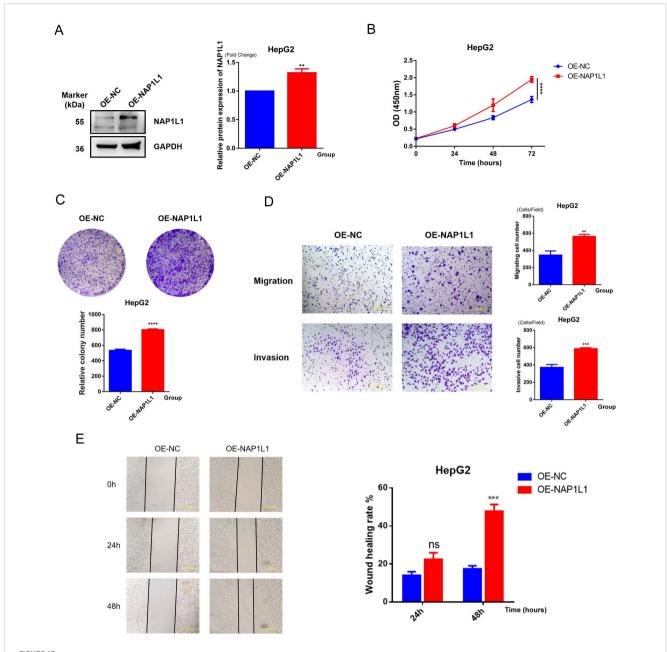


FIGURE 13
Functional verification of NAP1L1. (A) Western blot detection of NAP1L1 overexpression efficiency in HepG2 cells, with GAPDH as internal reference.
(B) CCK-8 assay to detect the effect of NAP1L1 on cell proliferation. (C) Colony formation assay to detect the effect of NAP1L1 on cell proliferation.
(D) Transwell assay to detect the effect of NAP1L1 on cell migration and invasion abilities. (E) Scratch wound healing assay to verify the promoting effect of NAP1L1 on cell migration. \*\*p<0.01, \*\*\*\*p<0.001, \*\*\*\*p<0.0001, and ns means no significant difference.

cell proliferative potential (Figure 13B). Cells with high NAP1L1 expression were demonstrated by Transwell migration and invasion analyses to possess stronger transmembrane migration capacity (p<0.01) and matrix invasion ability (p<0.001, Figure 13D). We further employed the scratch wound healing assay for verification, with results showing that the NAP1L1 overexpression group significantly exceeded the control group in wound closure efficiency within 48 hours (p<0.001, Figure 13E). These findings collectively suggest that a key role in regulating the malignant phenotype of hepatocellular carcinoma cells is played by NAP1L1.

### 4 Discussion

Through the integration of multi-omics data and advanced computational methods, we systematically explored the role and clinical significance of HCC-ICD. Cellular heterogeneity in the HCC microenvironment was revealed by single-cell RNA sequencing technology, identifying six major cell subtypes and analyzing the differences in ICD activity among these cell types. We then identified gene modules closely related to ICD through WGCNA. The HCC-related regulatory network of ICD was

visualized after integrating the corresponding differentially expressed genes, module genes, and single-cell characteristic genes. Based on these findings, we further constructed an ICDRS using these genes. Good predictive capability was demonstrated by the model in both the TCGA dataset and external independent datasets (GSE14520 and ICGC), providing strong evidence for risk stratification of HCC patients. We further elucidated key biological pathways associated with ICD through functional enrichment analyses, including GSEA and GSVA. Associations between ICD and gene mutations, the immune microenvironment, and drug sensitivity were explored by our research group, revealing the comprehensive and multifaceted role of ICD in HCC occurrence, development, and treatment response. The promoting effects of key genes CLIC1 and NAP1L1 on HCC proliferation, migration, and invasion behaviors were verified through in vitro experiments, providing important experimental evidence for the development of ICD-related therapeutic targets.

## 4.1 Research innovation and clinical comparative advantages

This study constructed a prognostic model for HCC based on ICD by integrating single-cell RNA sequencing, WGCNA, and large-scale machine learning algorithms. Extensive research has been conducted on HCC prognostic and drug resistance models, focusing primarily on molecular markers such as metabolism-related genes (43–45) and epigenetic modifications (46, 47). However, research on the role of ICD in HCC remains limited.

Our ICDRS model demonstrated superior predictive accuracy in comprehensive performance comparisons. In external validation, our RSF model achieved AUC values of 0.821, 0.832, and 0.796 for 1-year, 3-year, and 5-year OS, respectively. In comparison, Wang et al. (44) developed a mitochondrial-related transcriptome model with 3-year AUC of 0.77, while Chen et al. (45) constructed an oxidative phosphorylation-based model with AUC values of 0.690, 0.726, and 0.720 for 1-year, 2-year, and 3-year OS, respectively. These comparisons demonstrate that the ICD-based HCC prognostic model developed in this study shows superior performance compared to existing HCC prognostic models based on other molecular features. Our model enhanced predictive performance and greater stability across different survival timeframes, highlighting its clinical application value in HCC prediction.

Single-cell analysis revealed that macrophages had significantly higher ICD scores than other cell types (p<0.001), consistent with recent findings by Han et al. regarding tumor-associated macrophages' role in ICD (48). Through WGCNA analysis, we identified 106 HCC-ICDR genes that were correlated with ICD scores in the turquoise module (cor=0.4, p=2e-15) and blue module (cor=-0.37, p=2e-13), providing insights into the molecular regulatory network of ICD in hepatocellular carcinoma.

## 4.2 Molecular mechanisms: the multi-layer regulatory network of ICD and tumor progression

### 4.2.1 Composition and functional mechanisms of ICDRS

To understand the mechanistic basis of our predictive model, we analyzed the molecular mechanisms associated with ICDRS. The ICDRS comprises ten key genes with distinct functional roles. CLIC1, NAP1L1, APOE, and CD63 have been demonstrated to play crucial roles in ICD mechanisms: CLIC1, serving as a chloride ion channel protein, is involved in the regulation of apoptosis and ICD (49, 50). NAP1L1 is involved in the ICD process through its regulation of nucleosome assembly and chromatin remodeling, which affects the DNA damage response (51). APOE plays a key role in immune regulation and lipid metabolism, potentially influencing ICD effects through the modulation of macrophage polarization (52). CD63, as a marker of extracellular vesicles, is involved in intercellular signaling and immune activation (53, 54). The remaining genes CBX3, RAN, CLTA, SNRPG, FTL, and POMP primarily regulate transcription, nucleocytoplasmic transport, intracellular trafficking, RNA splicing, iron metabolism, and proteasome function, thereby indirectly modulating ICD-related cellular stress responses (55-58).

Patients in the high-risk group exhibited distinct activation of pathways associated with malignant tumor characteristics. Using functional enrichment analysis to identify these differential pathways, we found that five key cancer-related pathways were significantly enriched in the high-risk group by GSEA analysis (FDR < 0.05), including DNA repair pathway, E2F targets, MYC targets V1, PI3K/AKT/MTOR signaling, Reactive oxygen species pathway. Subsequent GSVA analysis further revealed multiple significantly upregulated pathways in the high-risk group, encompassing cell cycle and proliferation related pathways, stress and microenvironment related pathways, and signal transduction pathways.

### 4.2.2 DNA repair and genomic instability

DNA repair pathway activation plays a crucial role in maintaining genomic stability. Notably, a complex association pattern was observed between DNA repair pathway enrichment and TP53 mutations in the high-risk group. Our mutation analysis revealed that TP53 mutation rates were significantly higher in high-risk group patients compared to the low-risk group (36% vs 21%). This was accompanied by elevated MATH scores, reflecting greater tumor heterogeneity in the high-risk population.

Typically, TP53 mutations result in cell cycle checkpoint defects and reduced DNA damage repair capacity. However, expression of DNA repair-related genes was paradoxically increased in the highrisk group. Nevertheless, this compensatory repair activation is incomplete and still leads to the accumulation of genomic instability while potentially promoting the development of treatment resistance (59).

### 4.2.3 Cell cycle dysregulation and checkpoint defects

E2F targets pathway enrichment reflects cell cycle dysregulation. E2F transcription factors promote cell proliferation through S phase progression and DNA replication regulation. Hyperactivation of the E2F pathway has been recognized as an important driving factor in the malignant progression of HCC (60). Dong et al. (61) further revealed that TP53 mutations affect E2F1-mediated cell cycle progression by regulating the overexpression of histone variant H2AFZ. Specifically, H2AFZ overexpression regulates cell cycle signal transduction and DNA replication through pathways involving multiple cancer-associated kinases and E2F1, providing a molecular mechanism explanation for the association between E2F pathway activation and TP53 mutations observed in our high-risk group.

Furthermore, activation of the G2M checkpoint pathway has special significance in the context of TP53 mutations. Under normal circumstances, the TP53-mediated G1/S checkpoint serves as the primary barrier preventing the replication of damaged DNA. When this checkpoint is dysfunctional, more cells carrying DNA damage enter the S phase for replication. In this scenario, the G2M checkpoint faces greater pressure, needing to detect and respond to increased DNA damage and replication stress generated during the S phase. However, the G2M checkpoint function has inherent limitations. While the G2M checkpoint can temporarily prevent cells carrying DNA damage from entering mitosis, if DNA repair mechanisms cannot completely repair all damage, these cells may eventually bypass the checkpoint and enter division, leading to the transmission of chromosomal instability. More critically, in tumors with TP53 functional defects, the G2M checkpoint itself may also become dysfunctional, with reduced sensitivity to DNA damage, thereby allowing more genomically unstable cells to complete division, further exacerbating the accumulation of genomic instability (62).

### 4.2.4 Activation of oncogenic signaling pathways

In addition to cell cycle dysregulation, the PI3K/AKT/mTOR signaling pathway regulates cell survival, proliferation, and autophagy by integrating nutritional status, growth factor signals, and energy metabolism. Activation of this pathway in HCC is associated with tumor progression, angiogenesis, invasive metastasis, and the development of multidrug resistance (63-65). Similarly, MYC targets V1 not only regulates cell proliferation, but also is a major regulator of cell metabolism, and its abnormal activation is closely related to the aggressive phenotype and poor prognosis of HCC patients (55, 66). The activation of the reactive oxygen species pathway plays a dual role in the pathogenesis of HCC. While moderate oxidative stress promotes tumor cell survival and proliferation, excessive oxidative stress can cause extensive DNA damage, including base modifications, DNA strand breaks, and chromosomal aberrations (67). This oxidative damage is consistent with the increased genomic instability we observed and may be one of the important factors contributing to elevated MATH scores.

## 4.3 Immune microenvironment remodeling and mechanisms of therapeutic resistance

Despite the ICD signature, tumors with high ICDRS scores do not necessarily exhibit effective immune-mediated tumor control (48). Instead, poor prognosis in high ICDRS tumors appears to result from the promotion of an immunosuppressive microenvironment, a state that can be explained by mechanisms such as immune exclusion, where cytotoxic immune cells are prevented from infiltrating the tumor core, and/or T cell dysfunction, where infiltrated cells lose their effector functions (68, 69). Our immune cell infiltration analysis revealed the underlying mechanisms of this seemingly paradoxical phenomenon.

Specifically, the high-risk group demonstrated abundant immune cell infiltration, but these cells were predominantly composed of immunosuppressive components: plasma cell, follicular helper T cells, regulatory T cells, M0 macrophages, and neutrophils were significantly increased. This enrichment of immunosuppressive cells creates an immunosuppressive environment. The concomitant significant reduction in CD8+ T cell infiltration in the high-risk group (Figure 10E) strongly suggests that immune exclusion is a key mechanism underlying the ineffective anti-tumor immunity. Similar phenomena were confirmed in a clear cell renal cell carcinoma study by Wen et al. (70), who found that high-risk group patients exhibited increased Treg infiltration and decreased M1 macrophages, thereby forming an immunosuppressive environment.

In contrast, tumors with lower ICDRS scores demonstrated markedly different immune characteristics. The low-risk group was enriched with more effector immune cells, including naive B cells, memory B cells, CD8+ T cells, resting NK cells, monocytes, M1 macrophages, and resting mast cells. Extensive literature supports the central role of effector immune cells such as CD8+ T cells and M1 macrophages in anti-tumor immunity (71–74). The presence of these anti-tumor immune cells may more readily stimulate effective cytotoxic T cell responses, suggesting a more favorable immune environment and better prognosis.

Single-cell communication analysis provided a mechanistic foundation for this immune exclusion. In high ICDRS tumors, cancer cells exhibited more complex intercellular communication networks, establishing extensive signal exchanges with the microenvironment via MDK and MIF signaling pathways. Particularly, the MDK signaling pathway demonstrated stronger activity in the high-risk group, with bidirectional signal exchanges established between high-risk cancer cells and hepatocytes, potentially promoting tumor growth and microenvironment remodeling. More importantly, high-risk cancer cells were characterized by enhanced MIF signal-sending capacity, which promotes the establishment of an immunosuppressive microenvironment by regulating macrophage function and polarization states. Chen et al. (75) confirmed that MIF can induce macrophage polarization toward the M2 phenotype, thereby supporting tumor growth. This active reprogramming of the microenvironment towards an immunosuppressive and barrier-

like state offers a plausible mechanism for the observed exclusion of CD8+ T cells.

Notably, intercellular communication networks of low-risk cancer cells were significantly simplified, primarily characterized by interactions with macrophages, while communication intensity was relatively weakened. This may reflect a more balanced state of tumor-immune interaction.

## 4.4 Clinical application potential and personalized treatment guidance

The potential clinical application of ICDRS in HCC management could significantly enhance current treatment paradigms. At initial diagnosis, ICDRS could complement conventional staging systems and risk stratification approaches, potentially identifying high-risk patients who might benefit from earlier aggressive intervention or closer surveillance protocols.

These findings were further validated in a drug sensitivity analysis. Patients in the low-risk group showed higher sensitivity to multi-kinase inhibitors (sorafenib, sunitinib, fretinib) and chemotherapy drugs (doxorubicin, paclitaxel), which may be related to their more active anti-tumor immune microenvironment. Tumors are often classified into "cold" tumors that are immunosuppressed and "hot" tumors that are immunoactive and inflammatory (76), with the latter having better immunogenicity; At the same time, Smith et al. (77) also inversely confirmed that tumors with higher immune function scores showed worse sensitivity to a variety of therapeutic drugs, further supporting the close correlation between immune microenvironment status and treatment responsiveness. On the contrary, patients in the high-risk group showed relatively better response to specific targeted drugs (erlotinib, lapatinib, gefitinib) and cisplatin, suggesting that different risk groups may have different molecular target dependence and need individualized treatment strategy selection. These findings provide a theoretical basis for risk stratification-based personalized drug strategies and reveal that tumors with different molecular characteristics may require different therapeutic approaches.

From an implementation perspective, the 10-gene signature (CLIC1, NAP1L1, CBX3, RAN, APOE, CD63, CLTA, SNRPG, FTL, and POMP) could be assessed using RT-PCR or targeted RNA sequencing, making it feasible for integration into clinical testing workflows alongside established prognostic factors to guide treatment decisions throughout the HCC care continuum. Our nomogram integrating clinical characteristics provides individualized 1-, 3-, and 5-year survival probability predictions, serving as a valuable prognostic tool for clinical decision-making.

Near-term clinical translation focuses on three actionable pathways: (1) validation of the 10-gene ICDRS in prospective HCC cohorts undergoing sorafenib or immunotherapy, with initial feasibility studies planned within 12–18 months using existing clinical samples; (2) development of CLIC1 and NAP1L1 as therapeutic targets, building on our functional validation showing their roles in HCC proliferation and invasion, with potential for existing drug repurposing or novel inhibitor

development; (3) integration of ICDRS with current BCLC staging to create enhanced risk stratification algorithms, particularly valuable for intermediate-stage patients where treatment decisions are most challenging. The model's ability to predict differential drug sensitivity suggests potential clinical utility for treatment selection, though this requires further validation. While the identified immune microenvironment patterns offer biomarkers for immunotherapy response prediction.

### 4.5 Study limitations and future directions

Although this study has made significant progress in uncovering the importance of genes involved in immunogenic cell death in HCC, several limitations remain. First, our heavy reliance on public datasets introduces a potential selection bias. While these resources are invaluable, they may not fully capture the heterogeneity of HCC across different etiologies, clinical settings, and geographical regions. This limitation crucially impacts the generalizability of our ICDRS model and suggests that its performance should be cautiously evaluated in specific subpopulations not well-represented in current databases (e.g., NAFLD-driven HCC in Western cohorts). Additionally, the temporal heterogeneity inherent in these datasets, where samples were collected across different time periods with varying clinical practices and technological platforms, introduces confounding variables that our batch correction methods may not fully address. The interconnected nature of genomic databases, where similar patient populations, sample processing protocols, and analytical pipelines are often shared across studies, limits the true independence of our external validation and may create systematic biases that cannot be eliminated through conventional validation strategies. Furthermore, differences in survival endpoints, follow-up protocols, and clinical management practices across cohorts introduce outcome measurement bias that could artificially influence our model's apparent predictive accuracy, making the observed performance potentially non-generalizable to contemporary clinical settings.

Beyond dataset-derived biases, we must critically address the risk of model overfitting. Although we employed robust methodologies to mitigate this-including regularization techniques within our RSF algorithm, external validation in two independent cohorts, and systematic cross-validation during model selection—the possibility that the ICDRS is overly optimized for available dataset characteristics cannot be entirely ruled out. The exceptional performance observed in the TCGA training set compared to validation cohorts, combined with the highdimensional nature of transcriptomic data, means that some features within our 10-gene signature might inadvertently capture dataset-specific technical variances or biological redundancies rather than solely the core biological signals of immunogenic cell death. While the consistent performance across cohorts is reassuring, the model's efficacy could potentially diminish when applied to populations with substantially different genetic backgrounds or data generation protocols.

Furthermore, the transition of the ICDRS model from computational prediction to clinical utility faces significant barriers. The absence of prospective clinical validation means its real-world effectiveness remains unproven. Beyond a simple lack of validation, we critically acknowledge that practical implementation would face challenges such as the standardization of clinical assay methods, patient acceptance, cost-effectiveness, and integration into clinical workflows—factors often overlooked in bioinformatics studies but fundamental to translational impact.

A core set of limitations pertains to our experimental validation. The functional analysis of CLIC1 and NAP1L1 was conducted initially in only one cell line (HepG2), chosen for its well-established use in HCC research and stable transfection characteristics. However, we critically recognize that this approach is insufficient to account for the well-documented molecular heterogeneity of HCC, particularly given HepG2's specific genotype (e.g., HBV-positive, wild-type p53). Consequently, the oncogenic roles we observed may be cell-contextdependent. Future validation in genetically distinct cell lines (e.g., Huh7, SNU-398, PLC/PRF/5) representing diverse HCC subtypes is essential to confirm the generalizability of our findings. Furthermore, while computationally robust, our model lacks experimental validation in patient-derived samples (e.g., IHC confirmation of protein expression in tissue microarrays). This gap not only limits the future translational impact but also leaves the biological plausibility of our computational inferences less firmly established. Moreover, another central limitation of this study is the lack of direct experimental evidence linking the identified hub genes, CLIC1 and NAP1L1, to the modulation of ICD itself. While our functional assays confirmed their pro-tumorigenic roles in proliferation, migration, and invasion, these experiments did not specifically quantify hallmark ICD events. Furthermore, the precise signaling pathways downstream of CLIC1 and NAP1L1, such as AKT/mTOR, remain unexplored, limiting our mechanistic understanding. Thus, the precise functional relationship between these genes and the core immunogenic cell death process in HCC remains to be experimentally defined.

Similarly, our drug sensitivity analysis, while insightful, is derived from *in vitro* cell line data. We critically argue that these predictions likely cannot fully recapitulate the complex tumor microenvironment *in vivo* or account for patient-specific pharmacokinetics and pharmacodynamics. The lack of correlation with clinical response data means these predictions remain hypothetical and should be viewed primarily as generating testable hypotheses for future study, not as definitive therapeutic guides.

Finally, the mechanistic links between ICD and immune microenvironment remodeling, particularly their temporal dynamics, remain correlative. A critical limitation is that our study design cannot confirm causality, which is essential for developing targeted interventions that modulate ICD to improve clinical outcomes.

To address these interconnected limitations directly and bridge the gap between computational prediction and clinical application, we have formulated a targeted future research plan. We are establishing a prospective clinical cohort study at our institution. This initiative is specifically designed to: (1) expand functional experiments to include a panel of HCC cell lines with diverse genetic backgrounds (e.g., Huh7, SNU-398, PLC/PRF/5) and,

crucially, to directly investigate the role of CLIC1 and NAP1L1 in regulating ICD by assessing key hallmarks such as calreticulin exposure, ATP release, and HMGB1 translocation, combined with mechanistic investigations via Western blot analysis of key signaling pathways (e.g., AKT/mTOR) and qPCR for related genes; (2) Validate characteristic gene protein expression and its association with immune-mediated cell death phenotypes in external patient tissues; (3) Correlate the ICDRS risk score with actual treatment responses (e.g., sorafenib, EGFR inhibitors), ultimately evaluating its practical value in guiding personalized treatment decisions. This comprehensive strategy is aimed expressly at transforming the limitations identified herein into focused research objectives, thereby enhancing the translational potential of our findings.

Despite these limitations, our study establishes a solid foundation for clinical translation. In conclusion, our ICD-based prognostic framework represents a paradigm shift toward precision oncology in HCC management. The potential translational opportunities include clinical validation of the 10-gene signature, therapeutic targeting of CLIC1/NAP1L1 pathways, and implementation of risk-stratified treatment protocols. Long-term prospects encompass integration with emerging technologies such as liquid biopsies, radiomics, and AI-driven treatment optimization, potentially revolutionizing HCC patient outcomes through truly personalized medicine approaches.

### Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author/s.

### **Ethics statement**

Ethical approval was not required for the studies on humans in accordance with the local legislation and institutional requirements because only commercially available established cell lines were used.

### **Author contributions**

HL: Writing – original draft, Writing – review & editing, Project administration, Software, Methodology, Conceptualization. ZS: Software, Writing – review & editing, Conceptualization, Writing – original draft, Project administration, Methodology. XW: Writing – review & editing, Methodology. BZ: Writing – original draft, Visualization, Validation. LC: Writing – review & editing, Formal Analysis.

### **Funding**

The author(s) declare that no financial support was received for the research and/or publication of this article.

### **Acknowledgments**

We appreciated all the fellows in our department for their excellent work.

### Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

### Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial

intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

### Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

### Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2025. 1649618/full#supplementary-material

### References

- 1. Toh MR, Wong EYT, Wong SH, Ng AWT, Loo LH, Chow PK, et al. Global epidemiology and genetics of hepatocellular carcinoma. *Gastroenterology.* (2023) 164:766–82. doi: 10.1053/j.gastro.2023.01.033
- 2. Sagnelli E, Macera M, Russo A, Coppola N, Sagnelli C. Epidemiological and etiological variations in hepatocellular carcinoma. *Infection*. (2020) 48:7–17. doi: 10.1007/s15010-019-01345-y
- 3. Vogel A, Meyer T, Sapisochin G, Salem R, Saborowski A. Hepatocellular carcinoma. *Lancet*. (2022) 400:1345–62. doi: 10.1016/s0140-6736(22)01200-4
- 4. Petrick JL, Florio AA, Znaor A, Ruggieri D, Laversanne M, Alvarez CS, et al. International trends in hepatocellular carcinoma incidence, 1978-2012. *Int J Cancer*. (2020) 147:317–30. doi: 10.1002/ijc.32723
- 5. Ho SY, Liu PH, Hsu CY, Hsia CY, Huang YH, Lei HJ, et al. Evolution of etiology, presentation, management and prognostic tool in hepatocellular carcinoma. *Sci Rep.* (2020) 10:3925. doi: 10.1038/s41598-020-61028-9
- 6. European Association for the Study of the Liver. EASL Clinical Practice Guidelines on the management of hepatocellular carcinoma. *J Hepatol.* (2025) 82:315–74. doi: 10.1016/j.jhep.2024.08.028
- 7. Xu Z, Xu J, Sun S, Lin W, Li Y, Lu Q, et al. Mecheliolide elicits ROS-mediated ERS driven immunogenic cell death in hepatocellular carcinoma. *Redox Biol.* (2022) 54:102351. doi: 10.1016/j.redox.2022.102351
- 8. Galluzzi L, Guilbaud E, Schmidt D, Kroemer G, Marincola FM. Targeting immunogenic cell stress and death for cancer therapy. *Nat Rev Drug Discov.* (2024) 23:445–60. doi: 10.1038/s41573-024-00920-9
- 9. Huang C, Lin B, Chen C, Wang H, Lin X, Liu J, et al. Synergistic reinforcing of immunogenic cell death and transforming tumor-associated macrophages via a multifunctional cascade bioreactor for optimizing cancer immunotherapy. *Adv Mater.* (2022) 34e2207593. doi: 10.1002/adma.202207593
- 10. Sen Santara S, Lee DJ, Crespo Â, Hu JJ, Walker C, Ma X, et al. The NK cell receptor NKp46 recognizes ecto-calreticulin on ER-stressed cells. *Nature*. (2023) 616:348–56. doi: 10.1038/s41586-023-05912-0
- 11. Kroemer G, Galluzzi L, Kepp O, Zitvogel L. Immunogenic cell death in cancer therapy. *Annu Rev Immunol.* (2013) 31:51–72. doi: 10.1146/annurev-immunol-032712-100008
- 12. Yu Z, Guo J, Hu M, Gao Y, Huang L. Icaritin exacerbates mitophagy and synergizes with doxorubicin to induce immunogenic cell death in hepatocellular carcinoma. ACS Nano. (2020) 14:4816–28. doi: 10.1021/acsnano.0c00708
- 13. Chen G, Yang Z, Du J, He Z, Zhang Y, Zheng K, et al. Topological regulating bismuth nano-semiconductor for immunogenic cell death-mediated sonocatalytic hyperthermia therapy. *Small.* (2023) 19:e2304032. doi: 10.1002/smll.202304032
- 14. Hong P, Hu Z, Lin J, Cui K, Gao Z, Tian X, et al. Multi-omics revealed that ELAVL3 regulates MYCN in neuroblastoma via immunogenic cell death: Risk stratification and experimental research. *Int J Biol Macromol.* (2024) 282:137045. doi: 10.1016/j.ijbiomac.2024.137045

- 15. Liu J, Shi Y, Zhang Y. Multi-omics identification of an immunogenic cell death-related signature for clear cell renal cell carcinoma in the context of 3P medicine and based on a 101-combination machine learning computational framework. *EPMA J.* (2023) 14:275–305. doi: 10.1007/s13167-023-00327-3
- 16. Dai SL, Pan JQ, Su ZR. Multi-omics features of immunogenic cell death in gastric cancer identified by combining single-cell sequencing analysis and machine learning. *Sci Rep.* (2024) 14:21751. doi: 10.1038/s41598-024-73071-x
- 17. Tomczak K, Czerwińska P, Wiznerowicz M. The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. *Contemp Oncol (Pozn)*. (2015) 19:A68–77. doi: 10.5114/wo.2014.47136
- 18. Losic B, Craig AJ, Villacorta-Martin C, Martins-Filho SN, Akers N, Chen X, et al. Intratumoral heterogeneity and clonal evolution in liver cancer. *Nat Commun.* (2020) 11:291. doi: 10.1038/s41467-019-14050-z
- 19. Roessler S, Jia HL, Budhu A, Forgues M, Ye QH, Lee JS, et al. A unique metastasis gene signature enables prediction of tumor relapse in early-stage hepatocellular carcinoma patients. *Cancer Res.* (2010) 70:10202–12. doi: 10.1158/0008-5472.Can-10.2607
- 20. Zhang J, Bajari R, Andric D, Gerthoffert F, Lepsa A, Nahal-Bose H, et al. The international cancer genome consortium data portal. *Nat Biotechnol.* (2019) 37:367–9. doi: 10.1038/s41587-019-0055-9
- 21. Hao Y, Stuart T, Kowalski MH, Choudhary S, Hoffman P, Hartman A, et al. Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nat Biotechnol.* (2024) 42:293–304. doi: 10.1038/s41587-023-01767-y
- 22. Luo P, Zhang R, Ren J, Peng Z, Li J. Switchable normalization for learning-to-normalize deep representation. *IEEE Trans Pattern Anal Mach Intell.* (2021) 43:712–28. doi: 10.1109/tpami.2019.2932062
- 23. Petegrosso R, Li Z, Kuang R. Machine learning and statistical methods for clustering single-cell RNA-sequencing data. *Brief Bioinform.* (2020) 21:1209–23. doi: 10.1093/bib/bbz063
- 24. Linderman GC, Steinerberger S. Clustering with t-SNE, provably. SIAM J Math Data Sci. (2019) 1:313–32. doi: 10.1137/18m1216134
- 25. Aran D, Looney AP, Liu L, Wu E, Fong V, Hsu A, et al. Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. *Nat Immunol.* (2019) 20:163–72. doi: 10.1038/s41590-018-0276-y
- 26. Jin Y, Wang Z, He D, Zhu Y, Chen X, Cao K. Identification of novel subtypes based on ssGSEA in immune-related prognostic signature for tongue squamous cell carcinoma. Cancer Med. (2021) 10:8693-707. doi: 10.1002/cam4.4341
- 27. 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. doi: 10.1016/j.celrep.2016.12.019
- 28. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinf.* (2008) 9:559. doi: 10.1186/1471-2105-9-559

- 29. Wu H, Gong K, Qin Y, Yuan Z, Xia S, Zhang S, et al. In silico analysis of the potential mechanism of a preventive Chinese medicine formula on coronavirus disease 2019. *J Ethnopharmacol.* (2021) 275:114098. doi: 10.1016/j.jep.2021.114098
- 30. Han Y, Ye X, Cheng J, Zhang S, Feng W, Han Z, et al. Integrative analysis based on survival associated co-expression gene modules for predicting Neuroblastoma patients' survival time. *Biol Direct.* (2019) 14:4. doi: 10.1186/s13062-018-0229-2
- 31. 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. doi: 10.1093/bioinformatics/bts034
- 32. Song M, Zhang Q, Song C, Liu T, Zhang X, Ruan G, et al. The advanced lung cancer inflammation index is the optimal inflammatory biomarker of overall survival in patients with lung cancer. *J Cachexia Sarcopenia Muscle*. (2022) 13:2504–14. doi: 10.1002/jcsm.13032
- 33. Howe MD, Britton KJ, Joyce HE, Menard W, Emrani S, Kunicki ZJ, et al. Clinical application of plasma P-tau217 to assess eligibility for amyloid-lowering immunotherapy in memory clinic patients with early Alzheimer's disease. *Alzheimers Res Ther.* (2024) 16:154. doi: 10.1186/s13195-024-01521-9
- 34. Gittleman H, Sloan AE, Barnholtz-Sloan JS. An independently validated survival nomogram for lower-grade glioma. *Neuro Oncol.* (2020) 22:665–74. doi: 10.1093/neuonc/noz191
- 35. Mayakonda A, Lin DC, Assenov Y, Plass C, Koeffler HP. Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome Res.* (2018) 28:1747–56. doi: 10.1101/gr.239244.118
- 36. Jin S, Guerrero-Juarez CF, Zhang L, Chang I, Ramos R, Kuan CH, et al. Inference and analysis of cell-cell communication using CellChat. *Nat Commun.* (2021) 12:1088. doi: 10.1038/s41467-021-21246-9
- 37. Tennant PWG, Arnold KF, Ellison GTH, Gilthorpe MS. Analyses of 'change scores' do not estimate causal effects in observational data. *Int J Epidemiol.* (2022) 51:1604–15. doi: 10.1093/ije/dyab050
- 38. Zeng D, Fang Y, Qiu W, Luo P, Wang S, Shen R, et al. Enhancing immuno-oncology investigations through multidimensional decoding of tumor microenvironment with IOBR 2.0. Cell Rep Methods. (2024) 4:100910. doi: 10.1016/j.crmeth.2024.100910
- 39. Wickham H. Ggplot2: elegant graphics for data analysis: ggplot2: elegant graphics for data analysis. (2009). doi: 10.1007/978-0-387-98141-3
- 40. Chen B, Khodadoust MS, Liu CL, Newman AM, Alizadeh AA. Profiling tumor infiltrating immune cells with CIBERSORT. *Methods Mol Biol.* (2018) 1711:243–59. doi: 10.1007/978-1-4939-7493-1\_12
- 41. Geeleher P, Cox N, Huang RS. pRRophetic: an R package for prediction of clinical chemotherapeutic response from tumor gene expression levels. *PloS One.* (2014) 9:e107468. doi: 10.1371/journal.pone.0107468
- 42. Sebaugh JL. Guidelines for accurate EC50/IC50 estimation. Pharm Stat. (2011) 10:128–34. doi: 10.1002/pst.426
- 43. Feng NN, Du XY, Zhang YS, Jiao ZK, Wu XH, Yang BM. Overweight/obesity-related transcriptomic signature as a correlate of clinical outcome, immune microenvironment, and treatment response in hepatocellular carcinoma. *Front Endocrinol (Lausanne)*. (2022) 13:1061091. doi: 10.3389/fendo.2022.1061091
- 44. Wang Y, Song F, Zhang X, Yang C. Mitochondrial-related transcriptome feature correlates with prognosis, vascular invasion, tumor microenvironment, and treatment response in hepatocellular carcinoma. *Oxid Med Cell Longev.* (2022) 2022:1592905. doi: 10.1155/2022/1592905
- 45. Chen W, Yang Z, Chen Y. A novel oxidative phosphorylation-associated gene signature for prognosis prediction in patients with hepatocellular carcinoma. *Dis Markers.* (2022) 2022:3594901. doi: 10.1155/2022/3594901
- 46. Wu Y, Li L, Wang L, Zhang S, Zeng Z, Lu J, et al. m(1)A regulator-mediated methylation modification patterns correlated with autophagy to predict the prognosis of hepatocellular carcinoma. *BMC Cancer*. (2024) 24:506. doi: 10.1186/s12885-024-12235-4
- 47. Li Y, Qi D, Zhu B, Ye X. Analysis of m6A RNA methylation-related genes in liver hepatocellular carcinoma and their correlation with survival. *Int J Mol Sci.* (2021) 22:1474. doi: 10.3390/ijms22031474
- 48. Yan H, Ji X, Li B. Advancing personalized, predictive, and preventive medicine in bladder cancer: a multi-omics and machine learning approach for novel prognostic modeling, immune profiling, and therapeutic target discovery. *Front Immunol.* (2025) 16:1572034. doi: 10.3389/fimmu.2025.1572034
- 49. Hanson S, Dharan A, Jinsha PV, Pal S, Nair BG, Kar R, et al. Paraptosis: a unique cell death mode for targeting cancer. *Front Pharmacol.* (2023) 14:1159409. doi: 10.3389/fphar.2023.1159409
- 50. Li RQ, Wang W, Yan L, Song LY, Guan X, Zhang W, et al. Identification of tumor antigens and immune subtypes in breast cancer for mRNA vaccine development. *Front Oncol.* (2022) 12:973712. doi: 10.3389/fonc.2022.973712
- 51. Cruickshank B, Giacomantonio M, Marcato P, McFarland S, Pol J, Gujar S. Dying to be noticed: epigenetic regulation of immunogenic cell death for cancer immunotherapy. *Front Immunol.* (2018) 9:654. doi: 10.3389/fimmu.2018.00654
- 52. Mabrouk N, Lecoeur B, Bettaieb A, Paul C, Végran F. Impact of lipid metabolism on antitumor immune response. *Cancers (Basel)*. (2022) 14:1850. doi: 10.3390/cancers14071850
- 53. Zhang S, Liu S, Dong H, Jin X, Sun J, Sun J, et al. CD63-high macrophage-derived exosomal miR-6876-5p promotes hepatocellular carcinoma stemness via PTEN/Akt-mediated EMT pathway. *Hepatol Commun.* (2025) 9:e0616. doi: 10.1097/he9.000000000000616

- 54. Yu S, Chen J, Quan M, Li L, Li Y, Gao Y. CD63 negatively regulates hepatocellular carcinoma development through suppression of inflammatory cytokine-induced STAT3 activation. *J Cell Mol Med.* (2021) 25:1024–34. doi: 10.1111/jcmm.16167
- 55. Karaosmanoğlu O. Recurrent hepatocellular carcinoma is associated with the enrichment of MYC targets gene sets, elevated high confidence deleterious mutations and alternative splicing of DDB2 and BRCA1 transcripts. *Adv Med Sci.* (2025) 70:17–26. doi: 10.1016/j.advms.2024.10.004
- 56. Xu Y, Yao Y, Yu L, Fung HL, Tang AHN, Ng IO, et al. Clathrin light chain A facilitates small extracellular vesicle uptake to promote hepatocellular carcinoma progression. *Hepatol Int.* (2023) 17:1490–9. doi: 10.1007/s12072-023-10562-5
- 57. Li A, Li Y, Li X, Tang C, Yang Y, Li N, et al. Ferritin light chain as a potential biomarker for the prognosis of liver hepatocellular carcinoma. *Heliyon*. (2024) 10e36040. doi: 10.1016/j.heliyon.2024.e36040
- 58. Wu Y, Peng H, Chen G, Tu Y, Yu X. Bulk and single-cell RNA sequencing identify prognostic signatures related to FGFBP2(+) NK cell in hepatocellular carcinoma. *PeerJ.* (2025) 13:e19337. doi: 10.7717/peerj.19337
- 59. Gillman R, Lopes Floro K, Wankell M, Hebbard L. The role of DNA damage and repair in liver cancer. *Biochim Biophys Acta Rev Cancer.* (2021) 1875:188493. doi: 10.1016/j.bbcan.2020.188493
- 60. Zhan L, Huang C, Meng XM, Song Y, Wu XQ, Miu CG, et al. Promising roles of mammalian E2Fs in hepatocellular carcinoma. *Cell Signal.* (2014) 26:1075–81. doi: 10.1016/j.cellsig.2014.01.008
- 61. Dong M, Chen J, Deng Y, Zhang D, Dong L, Sun D. H2AFZ is a prognostic biomarker correlated to TP53 mutation and immune infiltration in hepatocellular carcinoma. *Front Oncol.* (2021) 11:701736. doi: 10.3389/fonc.2021.701736
- 62. Li Y, Lee W, Zhao ZG, Liu Y, Cui H, Wang HY. Fatty acid binding protein 5 is a novel therapeutic target for hepatocellular carcinoma. *World J Clin Oncol.* (2024) 15:130–44. doi: 10.5306/wjco.v15.i1.130
- 63. Tian LY, Smit DJ, Jücker M. The role of PI3K/AKT/mTOR signaling in hepatocellular carcinoma metabolism. Int J Mol Sci. (2023) 24:2652. doi: 10.3390/ijms24032652
- 64. Bamodu OA, Chang HL, Ong JR, Lee WH, Yeh CT, Tsai JT. Elevated PDK1 expression drives P13K/AKT/MTOR signaling promotes radiation-resistant and dedifferentiated phenotype of hepatocellular carcinoma. *Cells.* (2020) 9:746. doi: 10.3390/cells9030746
- 65. Wu Y, Zhang Y, Qin X, Geng H, Zuo D, Zhao Q. PI3K/AKT/mTOR pathway-related long non-coding RNAs: roles and mechanisms in hepatocellular carcinoma. *Pharmacol Res.* (2020) 160:105195. doi: 10.1016/j.phrs.2020.105195
- 66. Dong R, Wang T, Dong W, Zhang H, Li Y, Tao R, et al. TGM2-mediated histone serotonylation promotes HCC progression via MYC signalling pathway. *J Hepatol.* (2025) 83:105–18. doi: 10.1016/j.jhep.2024.12.038
- 67. Wang X, Li P, Ji H, Xu Z, Xing H. Single-cell transcriptomics reveals over-activated reactive oxygen species pathway in hepatocytes in the development of hepatocellular carcinoma. *Sci Rep.* (2024) 14:29809. doi: 10.1038/s41598-024-81481-0
- 68. Wherry EJ, Kurachi M. Molecular and cellular insights into T cell exhaustion. *Nat Rev Immunol.* (2015) 15:486–99. doi: 10.1038/nri3862
- 69. Joyce JA, Fearon DT. T cell exclusion, immune privilege, and the tumor microenvironment. *Science*. (2015) 348:74–80. doi: 10.1126/science.aaa6204
- 70. Wen X, Hou J, Qi T, Cheng X, Liao G, Fang S, et al. Anoikis resistance regulates immune infiltration and drug sensitivity in clear-cell renal cell carcinoma: insights from multi omics, single cell analysis and *in vitro* experiment. *Front Immunol.* (2024) 15:1427475. doi: 10.3389/fimmu.2024.1427475
- 71. Elia I, Rowe JH, Johnson S, Joshi S, Notarangelo G, Kurmi K, et al. Tumor cells dictate anti-tumor immune responses by altering pyruvate utilization and succinate signaling in CD8(+) T cells. *Cell Metab.* (2022) 34:1137–1150.e1136. doi: 10.1016/j.cmet.2022.06.008
- 72. Park J, Hsueh PC, Li Z, Ho PC. Microenvironment-driven metabolic adaptations guiding CD8(+) T cell anti-tumor immunity. *Immunity*. (2023) 56:32-42. doi: 10.1016/ j.immuni.2022.12.008
- 73. Boutilier AJ, Elsawa SF. Macrophage polarization states in the tumor microenvironment. *Int J Mol Sci.* (2021) 22:6995. doi: 10.3390/ijms22136995
- 74. Gunassekaran GR, Poongkavithai Vadevoo SM, Baek MC, Lee B. M1 macrophage exosomes engineered to foster M1 polarization and target the IL-4 receptor inhibit tumor growth by reprogramming tumor-associated macrophages into M1-like macrophages. *Biomaterials*. (2021) 278:121137. doi: 10.1016/ibiomaterials.2021.121137
- 75. Chen M, Liu H, Hong B, Xiao Y, Qian Y. MIF as a potential diagnostic and prognostic biomarker for triple-negative breast cancer that correlates with the polarization of M2 macrophages. *FASEB J.* (2024) 38:e23696. doi: 10.1096/fj.202400578R
- 76. Popel AS. Immunoactivating the tumor microenvironment enhances immunotherapy as predicted by integrative computational model. *Proc Natl Acad Sci U S A.* (2020) 117:4447–9. doi: 10.1073/pnas.2001050117
- 77. Smith MP, Sanchez-Laorden B, O'Brien K, Brunton H, Ferguson J, Young H, et al. The immune microenvironment confers resistance to MAPK pathway inhibitors through macrophage-derived TNFα. *Cancer Discov.* (2014) 4:1214–29. doi: 10.1158/2159-8290.Cd-13-1007