



Prognostic genes in hepatocellular carcinoma and their function: an analysis integrating high-throughput-based spatial transcriptomics and single-cell RNA-sequencing

Yumei Zhang^{1#}, Yu Yao^{2#}, Zongcai Yan¹, Meiling He¹, Yingjie Tang², Zhiming Zhang²

¹Department of Medical Oncology, Guangxi Medical University Cancer Hospital, Nanning, China; ²Department of Hepatobiliary Surgery, Guangxi Medical University Cancer Hospital, Nanning, China

Contributions: (I) Conception and design: Y Zhang; (II) Administrative support: Z Zhang; (III) Provision of study materials or patients: Y Yao, Z Yan; (IV) Collection and assembly of data: M He, Y Tang; (V) Data analysis and interpretation: Y Yao, M He; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work as co-first authors.

Correspondence to: Zhiming Zhang, MD. Department of Hepatobiliary Surgery, Guangxi Medical University Cancer Hospital, No. 71 Hedi Road, Qingxiu District, Nanning 530021, China. Email: zhzm1977@yeah.net.

Background: Identifying novel prognostic indicators for patients with hepatocellular carcinoma (HCC) is critical to improving patient outcomes. Mapping the spatial composition of HCC tumors remains challenging. Therefore, this study aimed to identify prognostic genes and assess their role in HCC by integrating spatial transcriptome sequencing (ST-seq) and single-cell RNA sequencing (scRNA-seq).

Methods: HCC-related datasets, including The Cancer Genome Atlas (TCGA)-HCC, International Cancer Genome Consortium (ICGC)-HCC, GSE149614, and GSE203612, were examined in this study. Single-cell and spatial transcriptome analyses were first conducted and followed by cell communication analysis. Target cell types with differential proportions and higher enrichment were identified. Least absolute shrinkage and selection operator (LASSO) Cox regression analysis was implemented to select prognostic genes and construct a risk model. Subsequently, a nomogram incorporating independent prognostic factors was developed, and immune microenvironment and immunotherapy response were analyzed. Finally, pseudotime analysis was performed to determine the developmental trajectory of the target cell types.

Results: Among hepatocytes, T/natural killer (NK) cells, B cells, myeloid cells, fibroblasts, and endothelial cells, hepatocytes and T/NK cells were identified as target cell types. ST analysis delineated seven cell subclusters (C0–C6). Cell communication analysis indicated that hepatocytes interact more strongly with other cell types compared to T/NK cells. The number of interactions between C4 and other cell subclusters was higher than the interaction intensity. Of the seven prognostic genes, *ADH4*, *IGFBP3*, and *LCAT* were downregulated in HCC, while *AKR1B10*, *GAGE2A*, *MAGEA6*, and *UCLH1* were upregulated. The patients deemed to have a higher risk according to the model had shorter survival. The nomogram that was created by combining the seven prognostic genes had a good ability to forecast the patients' probability of survival. In the immunoinfiltration analysis, *UCLH1* expression had the highest positive correlation with T/NK cell abundance. Moreover, in silico analysis suggested a potential association between high-risk status and reduced sensitivity to immunotherapy. Pseudotime analysis indicated that the expression of *UCLH1*, *GAGE2A*, *MAGEA6*, *ADH4*, and *LCAT* displayed an inverted U shape during the differentiation of hepatocytes, with the expression of *UCLH1* was elevated at the end of T/NK cell differentiation.

Conclusions: This study identified hepatocytes, T/NK cells, and seven prognostic genes in HCC, constructed a risk model and nomogram capable of predicting patient prognosis, and revealed through computational prediction that patients in the high-risk group might exhibit a poor response to immunotherapy, thereby providing new potential tools for the clinical treatment of HCC.

Keywords: Hepatocellular carcinoma (HCC); single-cell RNA sequencing (scRNA-seq); spatial transcriptomics; tumor microenvironment

Submitted Jan 06, 2026. Accepted for publication Mar 02, 2026. Published online Mar 25, 2026.

doi: 10.21037/jgo-2026-1-0014

View this article at: <https://dx.doi.org/10.21037/jgo-2026-1-0014>

Introduction

Hepatocellular carcinoma (HCC) is the most common primary liver cancer, accounting for about 90% of these case types. Globally, HCC is one of the most prevalent malignant tumors and a leading cause of cancer-related death. Patients with HCC are often asymptomatic in the early stages, which leads to late diagnosis and poor treatment outcomes, posing a serious threat to patient health (1). The exact mechanism underlying the development of

HCC remains unclear, and thus, in-depth investigation into the molecular pathogenesis and the discovery of novel intervention targets and treatments have become a key research focus, with significant implications for the clinical management of HCC (2-4). HCC's inconspicuous early symptoms and frequent metastasis at diagnosis limit the efficacy of systemic therapies, resulting in poor prognosis, high recurrence, and low 5-year survival. These clinical characteristics severely impact patients' quality of life and health. Therefore, identifying the biomarkers for HCC and its pathogenic mechanisms is critical to improving clinical management.

Sequencing technology plays a crucial role in liver disease research (5). Its advantages of high-throughput ability, high sensitivity, and high resolution allow for the detection of subtle changes in disease progression and provide crucial support for personalized medicine and precision therapy. Sequencing technology has been widely applied in HCC research. For instance, Chen *et al.* used transcriptome data from HCC samples and adjacent nontumor tissues to identify genes related to anoikis and determined that anoikis-related genes have the potential to evaluate the prognosis of patients with HCC and the efficacy of immunotherapy (5). Another study reported that the expression level of the *EXO1* gene in HCC tissues is significantly correlated with the overall survival (OS), clinical stage, and tumor status of patients with HCC, thereby demonstrating the diagnostic and prognostic value of *EXO1* in HCC (6). As can be seen, sequencing technology-based approaches can effectively evaluate the early diagnosis and prognosis of patients with HCC, thus providing novel perspectives in the search for therapeutic targets and interventions in HCC. However, traditional bulk sequencing methods can only provide the average information of all cells in the sample (7), which limits the study of the transcriptional activity and its changes in individual cells within the sample to a certain extent. Single-cell RNA sequencing (scRNA-seq) technology is a high-throughput sequencing technique that can reveal the gene expression profile of individual cells. The principle

Highlight box

Key findings

- Six cell types and seven cell subclusters were identified in hepatocellular carcinoma (HCC) along with seven prognostic genes. An accurate risk model and a nomogram were constructed. According to computational prediction, high-risk patients may exhibit reduced sensitivity to immunotherapy.

What is known, and what is new?

- Methods based on sequencing technologies can facilitate the early diagnosis and prognosis assessment of patients with HCC, offering novel insights into the research on therapeutic targets and intervention strategies for patients with HCC.
- By integrating spatial transcriptome sequencing with single-cell RNA sequencing, we were able to examine the cellular heterogeneity and spatial distribution of HCC, characterize the communication characteristics between key target cells [hepatocytes and T/natural killer (NK) cells] and the dynamic expression of prognostic genes during differentiation, assess the impact of the immune microenvironment on HCC progression, and identify potential associations between immune cells such as T/NK cells and HCC progression.

What is the implication, and what should change now?

- The seven prognostic genes can serve as candidate therapeutic targets for targeted therapy in HCC. The risk model and nomogram can aid in prognostic stratification in the clinic and guide personalized treatment for high-risk patients.
- The risk model and nomogram devised in this study can be applied clinically as an auxiliary tool for prognostic evaluation. To complement the findings on molecular mechanisms produced by this study, further cellular experiments and animal model studies should be conducted.

of scRNA-seq involves extracting RNA from single cells, reverse transcribing it into complementary DNA, and then conducting sequencing analysis to comprehensively study the gene expression profile of individual cells. Overall, elucidating the biomarkers and molecular mechanisms underlying the pathogenesis of HCC is of paramount importance for its clinical management.

Spatial transcriptomics technology has been developed and applied in recent years. The advantages of spatial transcriptomics technology include the ability to simultaneously detect the expression of a large number of genes at the tissue or cell level and revealing the spatial distribution patterns of genes. It can provide information on tissue structure and cell-cell interactions, contributing to a deeper understanding of the complex cellular organization within organisms. This technology offers more comprehensive biological information, providing additional perspectives for disease research and treatment (8,9). Through this technology, the spatial distribution of gene expression within HCC tissues at the tissue level can be determined, facilitating the discovery of characteristic tumor microenvironment features and targets. This spatially resolved information can serve as a crucial reference for studying the pathogenesis of HCC, identifying diagnostic markers, and developing personalized treatment strategies, thereby driving further progress and in-depth research in the field of HCC (10,11).

Although progress has been made in the development of transcriptome-based prognostic models for HCC (12), most studies still rely on bulk transcriptomic or scRNA-seq data, which makes it difficult to systematically dissect the spatial heterogeneity of gene expression and the spatial patterns of cell-cell interactions in the tumor microenvironment. To overcome this limitation, the present study integrated spatial transcriptomics and scRNA-seq to achieve the following objectives: (I) identify key cell types at single-cell resolution and map their spatial locations in tissues; (II) dissect the communication patterns between spatially adjacent cells; and (III) establish associations between spatial gradients of gene expression and patient prognosis. On this basis, we further systematically compared differences in clinical characteristics, immune microenvironment, and drug responsiveness between high- and low-risk patients through immune infiltration analysis, immunotherapy response evaluation, and drug sensitivity prediction. Pseudotime analysis was also applied to clarify the dynamic expression trajectories of prognostic genes during the differentiation of key cell populations.

Collectively, this study not only provides a spatial multi-omics perspective for understanding cellular heterogeneity and tumor evolution in HCC, but also offers an important foundation for identifying candidate therapeutic targets and intervention strategies. We present this article in accordance with the TRIPOD reporting checklist (available at <https://jgo.amegroups.com/article/view/10.21037/jgo-2026-1-0014/rc>).

Methods

Data collection

RNA-seq data for HCC from The Cancer Genome Atlas (TCGA)-HCC were obtained from UCSC Xena (<https://xenabrowser.net/datapages/>), including complete survival and gene expression profiles from 368 patients with HCC and 50 healthy controls. Additionally, the International Cancer Genome Consortium (ICGC) database (<https://dcc.icgc.org/>) was accessed to obtain data from the ICGC-HCC dataset, consisting of 45 HCC samples. scRNA-seq data (GSE149614) and spatial transcriptome sequencing (ST-seq) data (GSE203612) for HCC were acquired from the Gene Expression Omnibus (GEO) (<http://www.ncbi.nlm.nih.gov/geo/>). The GSE149614 dataset, containing 10 primary tumors and 8 paracancer samples, was used for single-cell analysis, while ST-seq data from a primary HCC sample (GSM6177612) in GSE203612 were utilized for spatial transcriptome analysis. The study was conducted in accordance with the Declaration of Helsinki and its subsequent amendments.

scRNA-seq analysis

The raw gene expression data from the GSE149614 dataset were processed into a Seurat object via the “Seurat” package in R version 5.0.1 (The R Foundation of Statistical Computing, Vienna, Austria) (13). Cells with low-quality metrics, including fewer than 200–6,000 genes detected (nFeature), fewer than 20,000 total gene counts (nCount), or a mitochondrial gene proportion (percent.mt) below 10%, were excluded. Principal component analysis (PCA) was performed for dimensionality reduction, with the focus on the first 2,000 highly variable genes (HVGs) as identified by the FindVariableFeatures function. The ElbowPlot function was used to determine the optimal number of principal components (PCs). Clustering was executed with the FindClusters function, yielding distinct cell clusters at

an optimum resolution of 0.4. Cell types were then classified based on marker genes referenced in a previous study (14), and the results were visualized with two-dimensional Uniform Manifold Approximation and Projection (UMAP).

Spatial transcriptome analysis and scRNA-seq-derived differentially expressed genes (scRNA-DEGs)

In order to characterize the different spatial distributions of spatial transcriptome data, the ST-seq data were analyzed with the R package “Seurat”. The nFeature and nCount of the GSM6177612 sample in the GSE203612 dataset were visualized through the SpatialFeaturePlot function. The SCTransform function was applied to standardize the ST-seq data of the GSM6177612 sample. Cell subclusters were obtained via UMAP according to PCs ascertained through PCA dimensionality reduction (optimum latitude =20). The spatial positions of cell subclusters were visualized via the SpatialFeaturePlot function, and the FindAllMarkers function was used to identify and spatially visualize the differentially expressed clustering genes between cell subclusters with each other. Feature genes extracted from cell types in the GSE149614 dataset [average $\log_2$ fold change ($\log_2$ FC) >0.3; adjusted P (adj. P) value <0.05] and feature genes obtained from cell types in the GSE203612 dataset (average $\log_2$ FC >0.3; adj. P<0.05) were selected for multimodal intersection analysis (MIA) so that the enrichment of these feature genes could be assessed. Under the condition that the expression site was greater than 10, the cell types with higher enrichment were selected as candidate cell types for subsequent analyses. We then sought to identify the scRNA-DEGs that were differentially expressed between the HCC and control samples in the GSE149614 dataset. Initially, the Wilcoxon test was used to assess cell types with differential proportions, and a box plot was generated through the R package “ggplot2” v3.4.4 (15) to represent these findings. Cell types with differential proportions were intersected with candidate cell types in order to obtain the target cell types. Using the R package “ReactomeGSA” v3.4.4 (16), we conducted functional enrichment analysis on the GSE149614 dataset in order to clarify the primary functions connected to the target cell types and identify the specific biological processes and pathways that these cell types are predominantly involved in. Additionally, the FindAllMarkers function was employed to identify the highly expressed genes specific to the target cell types. Subsequently, these genes underwent enrichment analysis via the Kyoto Encyclopedia of Genes

and Genomes (KEGG) Orthology-Based Annotation System for integrated analysis (KOBAS-i) database (<http://kobas.cbi.pku.edu.cn/kobas3>). Moreover, the 10 genes with the highest expression were selected based on degree scores in Cytoscape software v3.9.1 (17), and a protein-protein interaction (PPI) network was mapped to illustrate interactions between proteins encoded by these 10 highly expressed genes.

Cell communication

Cell communication analysis was performed to evaluate the interactions between individual cell types within the GSE149614 dataset and cell subclusters in the GSE203612 dataset. Specifically, the R package “CellChat” v1.6.1 (18) was used to infer the expression and pairing of receptors and ligands between cell types and cell subclusters. The CellChatDB.human dataset, sourced from the CellPhoneDB database, served as the reference for this analysis. The netVisual_circle function was used to display the number and intensity of interactions between any two cell subclusters, and the netVisual_heatmap function was applied to create heatmaps displaying the number and intensity of interactions.

Selection of candidate genes and functional analysis

The scRNA-DEGs between HCC and control samples in the target cell types were identified via the FindMarkers function (average $\log_2$ FC >0.5; P<0.05). The Manhattan plot was created via the “scRNAtoolVis” package v0.0.7 in R. The DEGs between HCC and controls in the TCGA-HCC dataset were acquired via the “DESeq2” package v1.36.0 (19) ($|\log_2$ FC| >2; adj. P<0.05). A volcano plot was created with the “ggplot2” package to visualize the DEGs. Moreover, the “ComplexHeatmap” package v2.14.0 (20) was responsible for displaying the expression heatmap of the DEGs. The intersection of scRNA-DEGs and DEGs was identified as candidate genes through the R package “VennDiagram” v1.7.3 (21). The “circlize” R package v0.4.15 (22) was employed to characterize the relationships between the key genes. In order to examine the functions related to the candidate genes (P<0.05), the R package “clusterProfiler” v4.7.1.003 (23) was applied for Gene Ontology (GO) items and the KEGG pathways for these genes [false discovery rate (FDR) <0.05]. Moreover, a PPI network was constructed via the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING;

<https://www.string-db.org/>) database, and the network was visualized via Cytoscape software (interaction score >0.7). Candidate genes with interacting relationships at the protein level were selected as key genes for prognostic gene screening.

Construction of a risk model

To screen for prognostic genes from pivotal genes, univariate Cox regression analysis, proportional hazards (PH) assumption test, and least absolute shrinkage and selection operator (LASSO) regression analysis were implemented. Univariate Cox regression analysis was supported by the R package “survival” v3.5-3 (24) to select genes related to survival [hazard ratio (HR) $\neq 1$; $P < 0.05$]. Subsequently, genes that passed the PH assumption test ($P > 0.05$) were analyzed by LASSO regression analysis to screen for prognostic genes maintaining high importance in HCC. The LASSO model was built via the R package “glmnet” v4.1-4 (25) (n-fold = 10). Based on the regression coefficient (coef) acquired from the LASSO model and the relative expression level (expr) of prognostic genes in every sample of the TCGA-HCC dataset, the equation of risk score was ascertained as follows: $\text{risk score} = \sum_{i=1}^n [\text{coef}(\text{gene}_i) \times \text{expr}(\text{gene}_i)]$. Moreover, the patients with HCC in the TCGA-HCC and ICGC-HCC datasets were categorized as high- and low-risk groups according to the optimal cutoff value of risk score. In the TCGA-HCC dataset, a risk curve was created to display the OS of high- and low-risk patients. The R package “survminer” v0.4.9 (26) was used to generate a Kaplan-Meier survival curve to evaluate the differences in OS between two groups of risk patients. In addition, the ability of the risk model to predict patients’ 1-, 3-, and 5-year OS was assessed via the receiver operating characteristic (ROC) curve, which was created via the “survivalROC” package v1.0.3.1 (27) [area under the curve (AUC) ≥ 0.6]. Identical analyses were performed for the ICGC-HCC dataset to confirm the risk model’s applicability and validity.

Independent prognostic analysis and stratified analysis of clinical features

Age, gender, stage, and pathological staging [tumor, node, and metastasis (T, N, and M) stages], in conjunction with risk score, were incorporated into univariate Cox regression analysis ($P < 0.05$), PH assumption test ($P > 0.05$), and multivariate Cox regression analysis ($P < 0.05$) to collect

the independent prognostic factors. A nomogram with independent prognostic factors was then developed with the R package “rms” v6.5-0 (24) to forecast survival probability. Moreover, a calibration curve was created by the R package “rms” to evaluate the nomogram’s predictive value. To assess the nomogram’s clinical utility, a decision curve was created. Finally, the Wilcoxon rank-sum test was used to estimate the difference in the risk score between clinical subgroups and to examine the relationships between these clinicopathological factors with risk score and with the two groups of risk patients ($P < 0.05$).

Assessment of immune microenvironment and immunotherapy response

In the TCGA-HCC dataset, the differences in immunoinfiltration of 28 immune cells and two stromal cells were determined according to the single-set gene set enrichment analysis (ssGSEA) scores between the high- and low-risk groups via the Wilcoxon test (FDR < 0.05) (28) through use of the ssGSEA algorithm in the R package “GSVA” v1.46.0 (29). Subsequently, the relationships between differentially infiltrated immune cells and prognostic genes were clarified via Spearman correlation analysis [$|\text{correlation coefficient}(\text{cor})| > 0.3$; FDR < 0.05]. In order to assess the differences in response to immunotherapy between the two groups of risk patients, the dysfunction, exclusion, and tumor immune dysfunction and exclusion (TIDE) scores were assessed via the TIDE algorithm in the TCGA-HCC dataset, and difference in the three corresponding scores between the low- and high-risk patients were compared through the Wilcoxon test ($P < 0.05$). Additionally, the ability of the risk score to forecast immunotherapy results was determined via the application of Spearman correlation analysis of the TIDE score and risk score.

Drug sensitivity analysis

The half-maximal inhibitory concentration (IC_{50}) of 138 different drug types was collected from the Genomics of Drug Sensitivity in Cancer (GDSC; <http://www.cancerrxgene.org/>) database. Subsequently, Spearman correlation analysis was carried out with the R package “psych” v2.14.0 (27) to screen for drugs that met the criteria of $|\text{cor}| > 0.3$ and FDR < 0.05 , with the correlation between the IC_{50} values and the risk score being evaluated. Drugs with significantly different (FDR < 0.05) IC_{50} values

between the high- and low-risk patients were identified via the Wilcoxon test. The 15 drugs with the highest sensitivity were visualized.

Characterization of the mutational landscape

The Wilcoxon test in the R package “maftools” v2.14.0 (30) was used to determine the difference in tumor mutational burden (TMB) ($P < 0.05$) between the two risk groups in the TCGA-HCC dataset. The mutation rates and mutation types of the top 20 most frequently mutated genes in the two risk groups were examined.

Secondary clustering and pseudotime analysis of target cell types

To examine the heterogeneity between different cell types, secondary clustering was performed, and target cell types were annotated. With the aim of determining the developmental trajectory of target cell types, pseudotime analysis was implemented via the “Monocle” package v2.26.0 (31).

Expression analysis of the prognostic genes

In order to clarify the expression of prognostic genes in single-cell, common transcriptome, and spatial transcriptome datasets, the Wilcoxon test was used to compare the differences in prognostic gene expression between patients with HCC and controls in the TCGA-HCC dataset ($P < 0.05$).

Statistical analysis

Bioinformatic analyses were supported by R v4.2.2. The Wilcoxon rank-sum test was used to estimate intergroup differences. P values less than 0.05 were considered to indicate statistical significance.

Results

Identification of six cell types in the GSE149614 dataset and seven cell subclusters in the GSE203612 dataset

After quality control was conducted, the GSE149614 dataset included 25,479 genes and 52,471 cells (Figure 1A,1B). The top 2,000 HVGs and 30 PCs were used for cell clustering (Figure 1C-1E). UMAP analysis identified 38

cell clusters, annotated into six cell types: myeloid cells, fibroblasts, endothelial cells, hepatocytes, T/natural killer (NK) cells, and B cells (Figure 1F,1G). In HCC samples, the proportion of hepatocytes increased to 29.58%, while that of T/NK cells was more abundant in controls (64.12%) (Figure 1H). The nFeature and nCount of GSM6177612 in the GSE203612 dataset are shown in Figure 1I. The first 20 PCs were used for clustering, yielding seven cell subclusters (C0–C6) (Figure 1J,1K). The spatial distribution of these subclusters within the analyzed sample was visualized (Figure 1L). Differential expression analysis identified clustering genes, with *CYP2A6* being highly expressed in C1, C3, C4, and C5 (Figure 1M).

Hepatocytes interacted more strongly with other cell types than with T/NK cells

Hepatocytes and T/NK cells were identified as target cell types based on their higher enrichment in spatial transcriptomics and differential abundance in single-cell data (Figure 2A,2B). Functional enrichment analysis further supported their distinct roles in HCC: hepatocytes were enriched in metabolic pathways (e.g., retinol metabolism), while T/NK cells were associated with immune cytotoxicity pathways (Figure 2C,2D). These cell types were therefore selected as the cellular context for subsequent prognostic gene screening and model construction. The PPI network showed stronger protein interactions among the top 10 most highly expressed genes in T/NK cells than in hepatocytes (Figure 2E). Hepatocytes had stronger and more numerous interactions with other cell types than did T/NK cells (Figure S1A,S1B). Fibroblasts interacted with T/NK cells and B cells mainly via MIF-CD74 and MIF-CXCR4 (Figure S1C). Analysis of this single spatial sample suggested that C4 exhibited more interaction events (higher interaction count) with other subclusters, but the interaction strength (intensity) was relatively low (Figure S1D,S1E). C6 interacted via FN1-SDC1 and FN1-SDC4, while C5 interacted with C4 via VTN-(ITGAV + ITGB5) and VTN-(ITGAV + ITGB1) (Figure S1F). The VTN-mediated pathway was used to visualize the interactions between cell subclusters (Figure S1G).

The functions of 130 candidate genes were related to cell cycle regulation and maintenance of chromosome structure and function

A total of 3,428 scRNA-DEGs were identified in

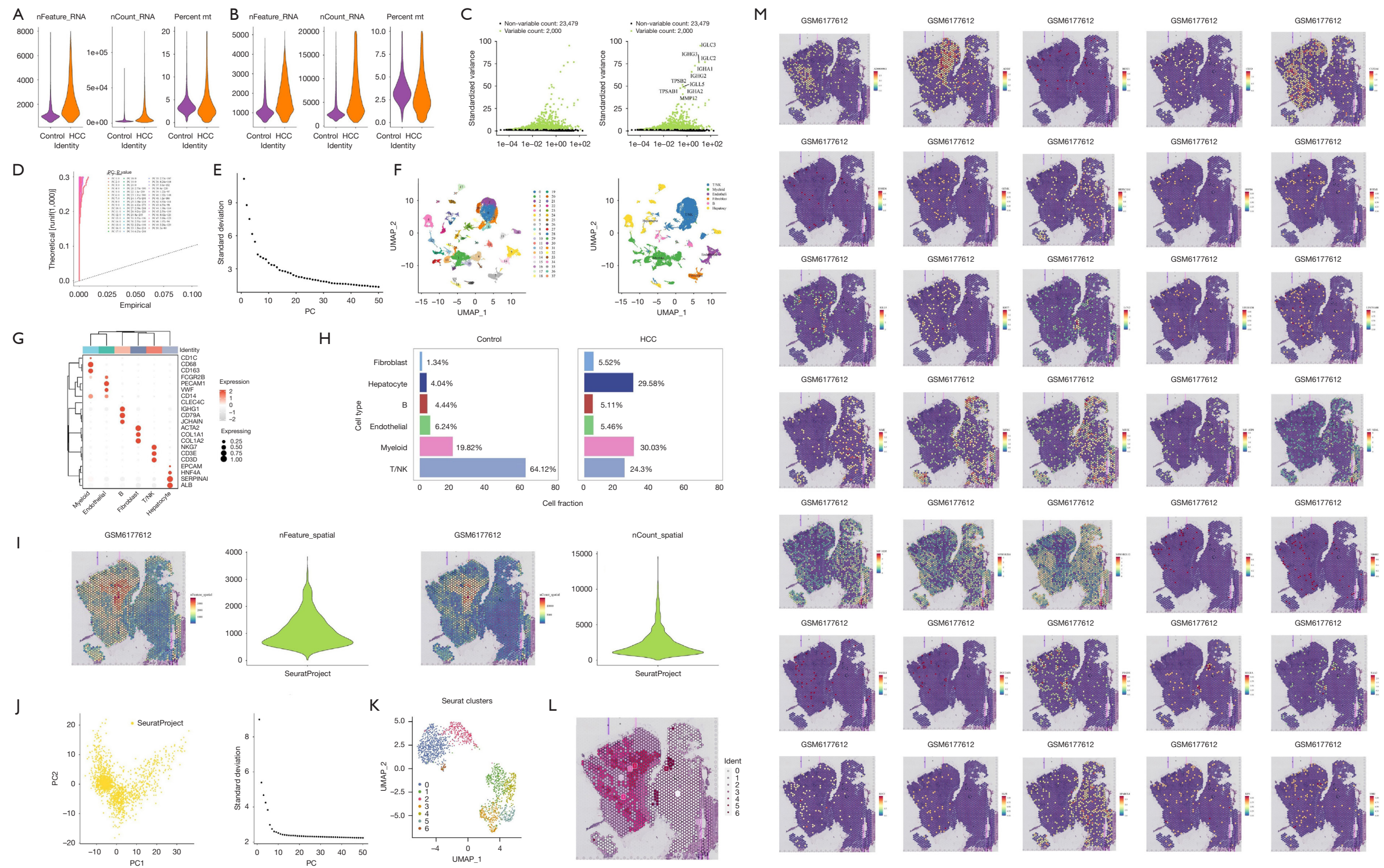


Figure 1 Single-cell quality control and spatial transcriptome analysis. (A,B) After quality control was conducted, the GSE149614 dataset included 25,479 genes and 52,471 cells. (C-E) The top 2,000 HVGs and 30 PCs were used for cell clustering. (F,G) UMAP analysis identified 38 cell clusters, annotated into six cell types: myeloid cells, fibroblasts, endothelial cells, hepatocytes, T/NK cells, and B cells. (H) In HCC samples, the proportion of hepatocytes increased to 29.58%, while that of T/NK cells was more abundant in controls (64.12%). (I) The nFeature and nCount of GSM6177612 in the GSE203612 dataset. (J,K) The first 20 PCs were used for clustering, yielding seven cell subclusters (C0–C6). (L) The spatial distribution of these subclusters was visualized. (M) Differential expression analysis identified clustering genes, with CYP2A6 being highly expressed in C1, C3, C4, and C5. HCC, hepatocellular carcinoma; HVGs, highly variable genes; NK, natural killer; PC, principal component; UMAP, Uniform Manifold Approximation and Projection.

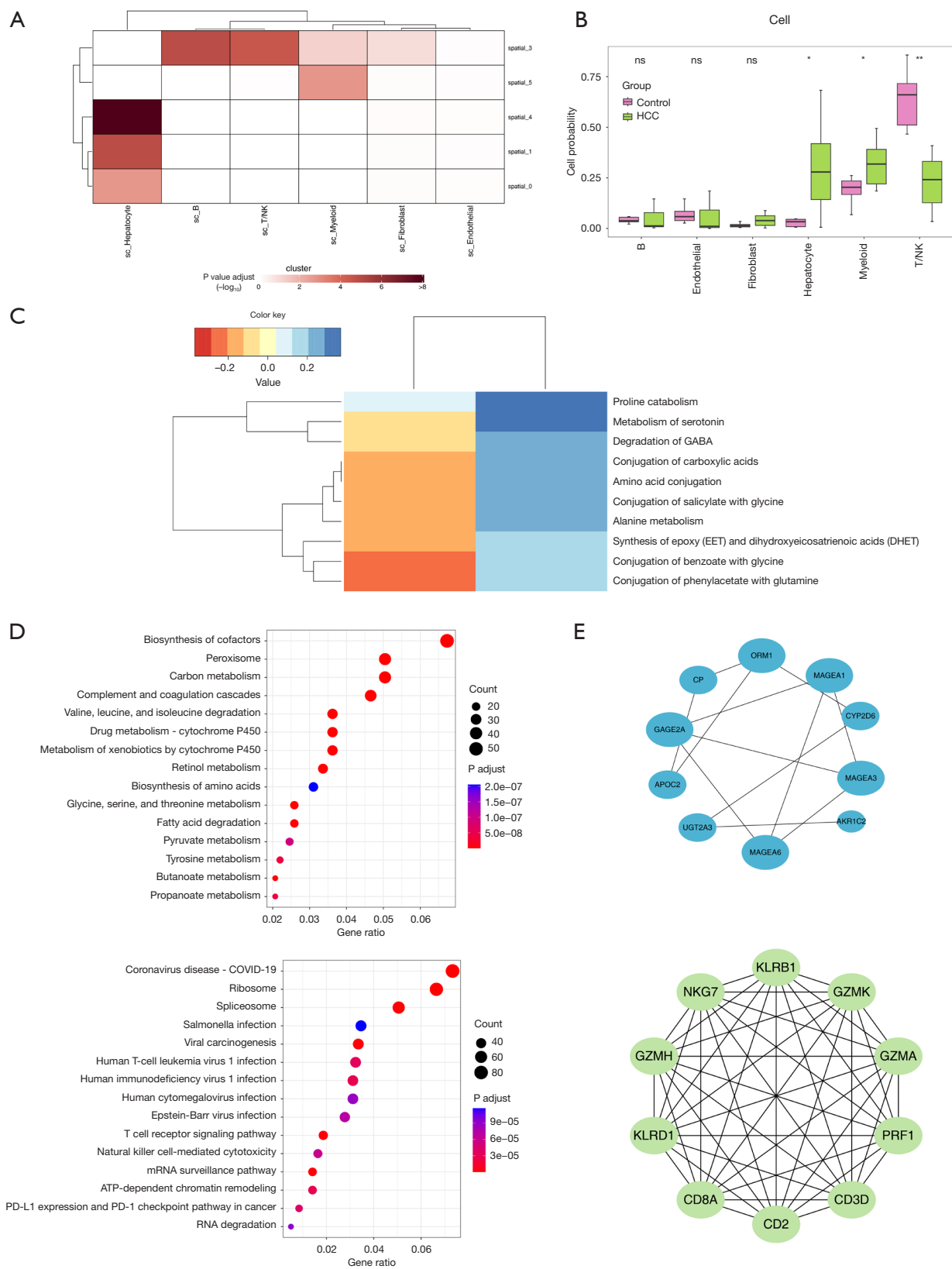


Figure 2 The characteristic genes of each cell type in individual cells subjected to MIA. (A) Enrichment analysis identified hepatocytes, T/

NK cells, and B cells as candidate cell types due to higher enrichment. (B) Among six cell types, hepatocytes, T/NK cells, and myeloid cells showed differential proportions between HCC and control samples in the GSE149614 dataset. (C) Functional enrichment analysis revealed co-enriched pathways like proline catabolism, with most pathways activated in hepatocytes and suppressed in T/NK cells. (D) Hepatocytes were involved in retinol metabolism and amino acid biosynthesis, while T/NK cells were linked to NK cell-mediated cytotoxicity. (E) The PPI network showed stronger protein interactions among the top 10 most highly expressed genes in T/NK cells than in hepatocytes. HCC, hepatocellular carcinoma; MIA, multimodal intersection analysis; NK, natural killer; PPI, protein-protein interaction.

hepatocytes and T/NK cells (Figure 3A), while 1,608 DEGs were found in the TCGA-HCC dataset (Figure 3B,3C). The intersection of these two sets yielded 130 candidate genes (Figure 3D). These genes were associated with 131 GO terms, including cellular amino acid metabolic process and peptidase regulator activity (Figure 3E), and were enriched in 6 KEGG pathways, such as the p53 signaling pathway (Figure 3F). A PPI network with 133 interaction pairs and formed by 53 pivotal genes was identified, with an average node degree of 2.090 (Figure 3G).

The risk model accurately predicted the outcomes of patients with HCC

Sixteen genes associated with the survival of patients with HCC (HR $\neq$ 1; $P < 0.05$) were identified by univariate Cox regression (Figure 4A). These genes passed the PH assumption test ($P > 0.05$) (Table 2). Seven genes (*IGFBP3*, *UCHL1*, *GAGE2A*, *MAGEA6*, *AKR1B10*, *ADH4*, and *LCAT*) were selected as prognostic genes, with a minimum lambda (0.02357759) (Figure 4B). The risk score equation is shown in Table 1. In the TCGA-HCC dataset, higher risk scores were correlated with increased death rates (Figure 4C) and upregulated prognostic genes in the high-risk patients (Figure 4D). High-risk patients had shorter survival ($P < 0.001$) (Figure 4E). The model's accuracy was validated by ROC curves, with AUCs of 0.750, 0.700, and 0.680 for 1-, 3-, and 5-year projections, respectively (Figure 4F). The risk model was further validated in the ICGC-HCC dataset, which yielded consistent results (Figure 4G-4J).

The predictive nomogram demonstrated reliability

The risk score was validated as an independent prognostic factor (Figure 5A,5B, Table 3). A nomogram integrating seven prognostic genes was constructed, with higher total points correlating with a higher probability of prolonged OS (Figure 5C). Calibration (Figure 5D) and decision curves (Figure 5E) confirmed the nomogram's high accuracy and clinical validity for predicting survival in patients

with HCC. Risk scores increased with HCC progression, particularly from T1 to T2 and T1 to T3/4 stages and from stage 1 to 2 and 1 to 3/4 (Figure 5F). High-risk patients were more likely to have stage 2 (26.5%) and stage 3–4 disease (33.3%), while low-risk patients primarily had stage 1 disease (58.6%) (Figure 5G).

Poor immunotherapy response in the high-risk group

Significant differences in ssGSEA scores were noted between the risk groups for 11 immune cells (FDR < 0.05) (Figure 6A). *UCHL1* expression was positively correlated with T/NK cell abundance (cor = 0.54; FDR < 0.05), while *ADH4* was negatively correlated with activated CD4 T cell abundance (cor = -0.33; FDR < 0.05) (Figure 6B). The high-risk group exhibited higher exclusion and TIDE scores and a lower dysfunction score (Figure 6C), suggesting a potentially suppressed immune microenvironment and reduced efficacy of immunotherapy in these patients. A positive correlation between TIDE and risk scores further suggested a potential reduced sensitivity to immunotherapy in high-risk patients (Figure 6D). Spearman correlation analysis identified 33 drugs significantly linked to risk scores, with 13 and 20 being positively and negatively correlated, respectively (Figure 6E). High-risk patients showed lower IC₅₀ values for negatively correlated drugs such as A.443654 and rapamycin but higher IC₅₀ values for positively correlated drugs such as axitinib and erlotinib (Figure 6F).

Missense mutations of the seven prognostic genes

Mutation analysis revealed distinct patterns in the high- and low-risk groups. The most common mutations in the high-risk group were in *TP53* (48%), *TTN* (31%), *MUC16* (20%), *CTNNB1* (18%), and *LRP1B* (13%), while those in the low-risk group were in *CTNNB1* (35%), *TTN* (25%), *ALB* (14%), *MUC16* (14%), and *TP53* (13%) (Figure S2A). Common mutation types in HCC were splice site, frame shift insertion, in-frame deletion, missense, and nonsense (Figure S2B). All seven prognostic genes showed missense

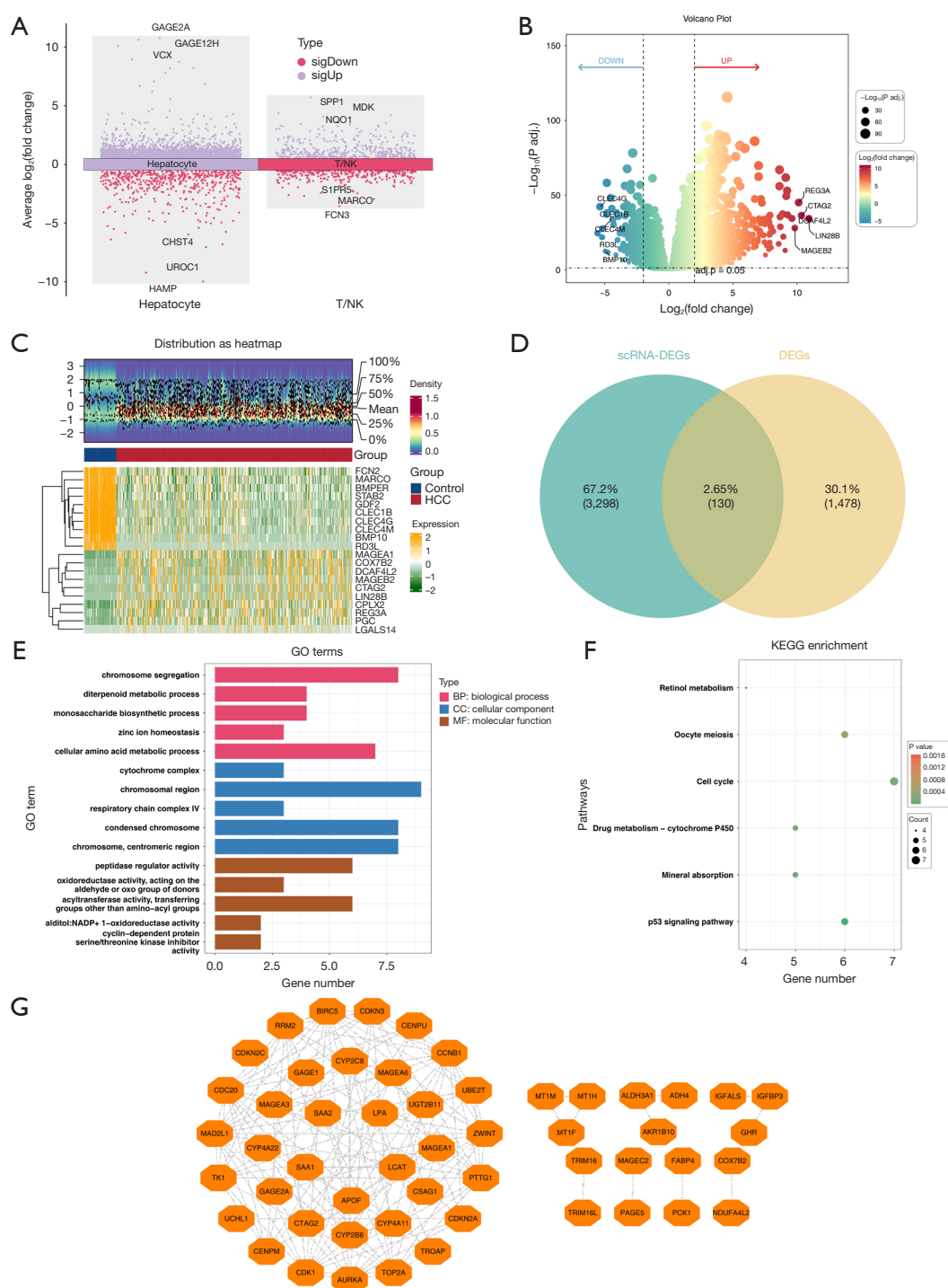


Figure 3 The functions of candidate genes were related to cell cycle regulation and maintenance of chromosome structure and function. (A) A total of 3,428 scRNA-DEGs were identified in hepatocytes and T/NK cells. (B,C) 1,608 DEGs were found in the TCGA-HCC dataset. (D) The intersection of these two sets yielded 130 candidate genes. (E) These genes were associated with 131 GO terms, including cellular amino acid metabolic process and peptidase regulator activity. (F) were enriched in 24 KEGG pathways, such as p53 signaling pathway. (G) A PPI network with 133 interaction pairs and formed by 53 pivotal genes was identified, with an average node degree of 2.090. DEGs, differentially expressed genes; GO, Gene Ontology; HCC, hepatocellular carcinoma; KEGG, Kyoto Encyclopedia of Genes and Genomes; NK, natural killer; PPI, protein-protein interaction; scRNA-DEGs, scRNA-seq-derived differentially expressed genes; TCGA, The Cancer Genome Atlas.

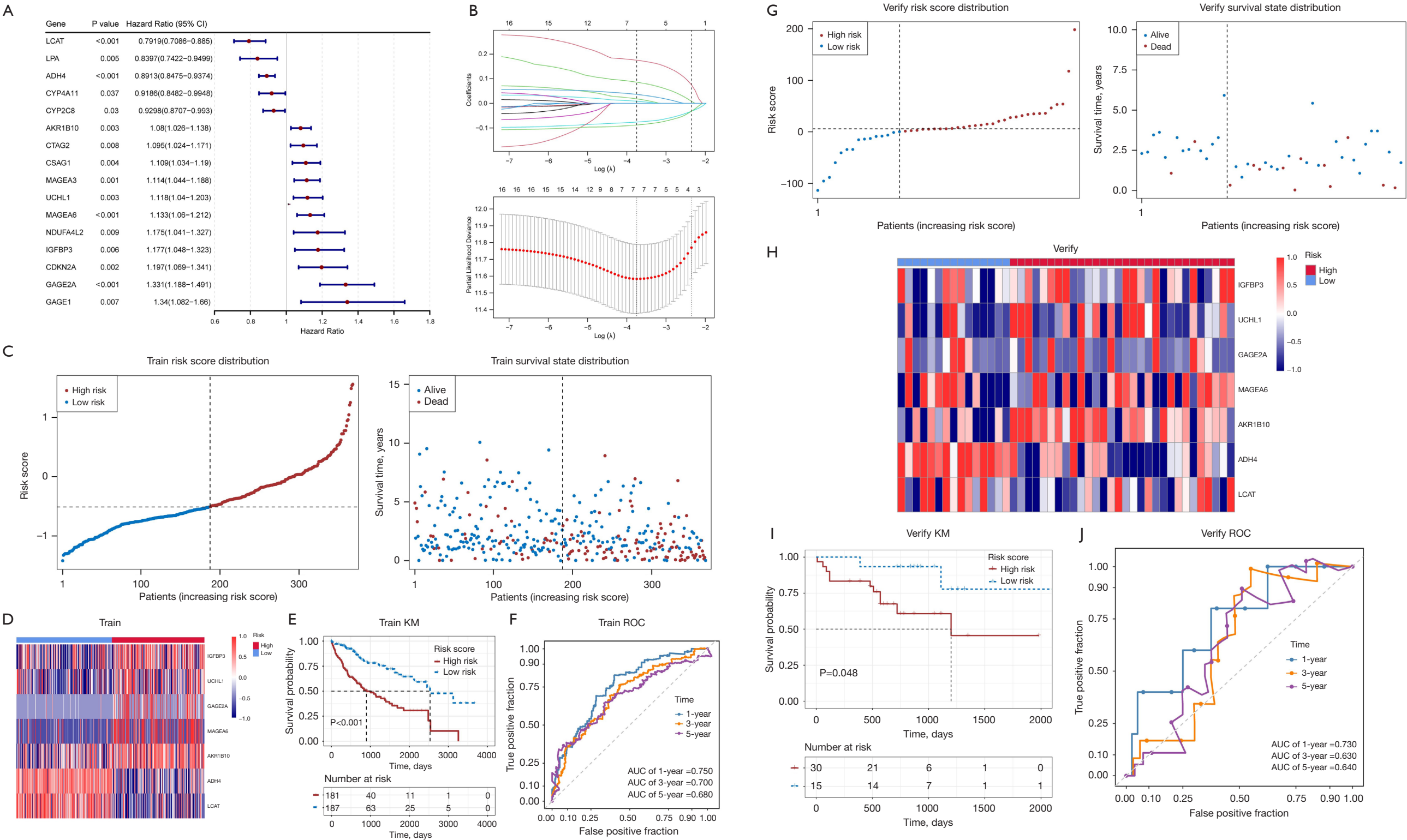


Figure 4 Development of risk models for predicting patient outcomes. (A) Sixteen genes associated with HCC survival (HR $\neq$ 1, $P < 0.05$) were identified by univariate Cox regression. (B) Seven genes (*IGFBP3*, *UCLH1*, *GAGE2A*, *MAGEA6*, *AKR1B10*, *ADH4*, and *LCAT*) were selected as prognostic genes at minimum lambda (0.02357759). (C) The risk score equation is shown in Table 1. In the TCGA-HCC dataset, higher risk scores correlated with increased death rates. (D) Upregulated prognostic genes in high-risk patients. (E) High-risk patients had shorter survival ($P < 0.001$). (F) The model's accuracy was validated by ROC curves with AUCs of 0.750, 0.700, and 0.680 for 1-, 3-, and 5-year projections, respectively. (G-J) The risk model was further validated in the ICGC-HCC dataset, showing consistent results. AUC, area under the curve; CI, confidence interval; HCC, hepatocellular carcinoma; HR, hazard ratio; ICGC, International Cancer Genome Consortium; KM, Kaplan-Meier; ROC, receiver operating characteristic; TCGA, The Cancer Genome Atlas.

Table 1 List of LASSO regression coefficients of prognostic genes

Gene	Coef
<i>IGFBP</i>	0.01822912
<i>UCLH</i>	0.00862560
<i>GAGE2A</i>	0.17519568
<i>MAGEA</i>	0.08592270
<i>AKR1B10</i>	0.03617247
<i>ADH4</i>	−0.07607377
<i>LCAT</i>	−0.08799587

Coef, coefficient; LASSO, least absolute shrinkage and selection operator.

Table 2 PH assumption test for 16 survival-related genes

Gene	Chi-sq	df	P
<i>CTAG2</i>	0.39679006	1	0.53
<i>CYP2C8</i>	1.344156429	1	0.25
<i>IGFBP3</i>	0.822528049	1	0.36
<i>CDKN2A</i>	0.01546734	1	0.90
<i>UCLH1</i>	0.627224597	1	0.43
<i>NDUFA4L2</i>	0.206339117	1	0.65
<i>CYP4A11</i>	1.928875075	1	0.16
<i>GAGE2A</i>	0.00471945	1	0.95
<i>MAGEA6</i>	0.138415711	1	0.71
<i>AKR1B10</i>	1.999748033	1	0.16
<i>ADH4</i>	1.776967504	1	0.18
<i>LPA</i>	1.128258245	1	0.29
<i>CSAG1</i>	0.850373304	1	0.36
<i>GAGE1</i>	0.826081767	1	0.36
<i>LCAT</i>	0.785499808	1	0.38
<i>MAGEA3</i>	0.489657345	1	0.48

Chi-sq, Chi-squared; df, degrees of freedom; PH, proportional hazards.

mutations, and *LCAT* also had frame shift insertions (Figure S2C).

The dynamic expression of prognostic genes during the differentiation of hepatocytes and T/NK cells

The top 30 PCs were used for clustering hepatocytes into 11

subtypes and T/NK cells into 13 subtypes (Figure 7A–7C). Pseudotime analysis identified nine differentiation states for hepatocytes and T/NK cells, with stage 1 being the initial stage and stages 7 and 8 being the terminal stages for hepatocytes and T/NK cells, respectively (Figure 7D,7E). Hepatocyte subtypes 1, 2, 5, and 7 were prevalent in the early stages, while subtype 0 was more common in the later stages. For T/NK cells, subtypes 3, 5, 9, and 10 were more common in the early stages, and subtype 7 was more abundant in the later stages (Figure 7F). The expression patterns of *UCLH1*, *GAGE2A*, *MAGEA6*, *ADH4*, and *LCAT* showed an inverted U shape during hepatocyte differentiation, while the expressions of *AKR1B10* and *IGFBP3* were higher at later stages. At the end of T/NK cell differentiation, *UCLH1* was upregulated, but *ADH4*, *GAGE2A*, and *MAGEA6* were downregulated (Figure 7G).

ADH4, *IGFBP3*, and *LCAT* were downregulated in HCC, while *AKR1B10*, *GAGE2A*, *MAGEA6*, and *UCLH1* were upregulated

In the TCGA-HCC dataset, *ADH4*, *IGFBP3*, and *LCAT* were downregulated in HC, while *AKR1B10*, *GAGE2A*, *MAGEA6*, and *UCLH1* were upregulated (P<0.001) (Figure 8A). In the GSE149614 dataset, *IGFBP3* was expressed in all six cell types, while the other five prognostic genes were more abundant in hepatocytes (Figure 8B). All prognostic genes showed significant expression differences between HCC and controls (P<0.05) (Figure 8C). Using the single spatial transcriptomics sample from the GSE203612 dataset, the spatial expression of these genes across HCC cell subclusters was visualized (Figure 8D).

Discussion

Primary liver cancer is a malignant neoplasm originating from hepatocytes or intrahepatic bile duct epithelial cells, and it is one of the most prevalent malignancies globally, with HCC being the predominant form. The early HCC manifestations are often subtle, leading diagnoses being at an advanced stage with metastases being present, which limits the efficacy of systemic therapies. This culminates in poor prognosis, high recurrence, and a low 5-year survival, significantly compromising patients’ quality of life and posing a threat to their health (1,32).

In our study, the single-cell dataset GSE149614 was subjected to quality controlled. Dimensionality was reduced, 2,000 variable genes were selected, and UMAP/

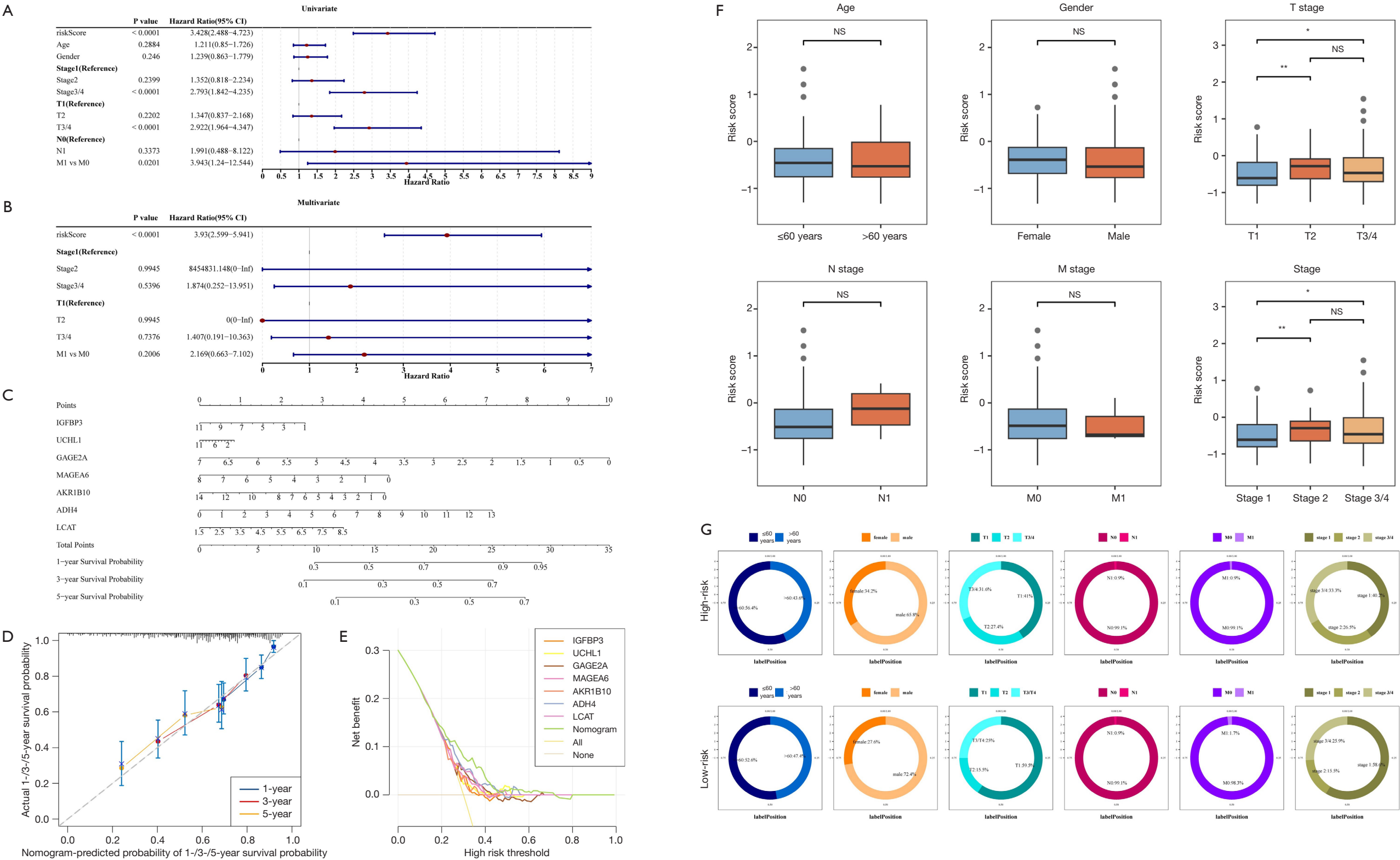


Figure 5 Survival differences between the low- and high-risk groups by subgroups of clinical characteristics. (A,B) The risk score was validated as an independent prognostic factor. (C) A nomogram integrating seven prognostic genes was constructed, with higher total points correlating with increased OS probabilities. (D) Calibration curves. (E) Decision curves. (F) Risk scores increased with HCC progression, particularly from T1 to T2 and T1 to T3/4 stages, and from stage 1 to 2 and 1 to 3/4. (G) High-risk patients were more prevalent in stage 2 (26.5%) and stage 3/4 (33.3%), while low-risk patients were mainly in stage 1 (58.6%). *, P<0.05; **, P<0.01; NS, not significant (P>0.05). CI, confidence interval; HCC, hepatocellular carcinoma; M, metastasis; OS, overall survival; T, tumor.

Table 3 PH assumption test for clinical features and risk scores

Features	Chi-sq	df	P
Risk score	0.229066758	1	0.63
Stage	1.926165258	2	0.38
T stage	1.448623984	2	0.48
M stage	2.760174707	1	0.10

Chi-sq, Chi-squared; df, degrees of freedom; M, metastasis; PH, proportional hazards; T, tumor.

marker genes used to classify cells into six types. Differential analysis identified hepatocytes, T/NK, and myeloid cells as differential types. The spatial transcriptomics data were standardized, dimensionality-reduced, and clustered into seven groups. Spatial locations were visualized, and MIA showed high enrichment of hepatocytes, T/NK, and B cells. Hepatocytes and T/NK cells were selected as target cell types. DEGs were screened between target cells in the disease and control groups. Common differential genes were identified through intergroup analysis. After enrichment and PPI screening, regression was used to find prognosis-related genes such as *IGFBP3*. A risk model was constructed and evaluated. Previous studies have found that cellular plasticity and metabolism regulation can improve cancer cell proliferation (1,30,32). Hepatocytes undergo metabolic reprogramming during HCC, and hepatocyte secretome changes are both consequence and cause of tumor development. The study results suggest that *UCHL1*, *GAGE2A*, *MAGEA6*, *ADH4*, *LCAT*, *AKR1B10*, and *IGFBP3* may be associated with the hepatocyte secretome changes in HCC progression.

Of note, *UCHL1* expression was positively correlated with T/NK cell infiltration. Although high T/NK cell infiltration generally predicts a favorable prognosis (33), *UCHL1* itself is a well-established poor prognostic marker (34), suggesting that their relationship is more complex than expected. *UCHL1* is a deubiquitinating enzyme that exerts oncogenic roles in multiple cancers, including HCC, by regulating protein stability (35). We therefore hypothesize that high *UCHL1* expression in the tumor microenvironment may not only be accompanied by increased lymphocyte counts, but also lead to impaired anti-tumor function of T/NK cells.

This hypothesis is supported by previous studies: *UCHL1* may promote T-cell dysfunction or exhaustion by stabilizing inhibitory receptors or interfering with key signaling pathways (36). In the present study, the high-risk group (with

high *UCHL1* expression) exhibited higher TIDE exclusion scores, further indicating that the microenvironment tends to be immunosuppressive and immune-evasive.

Collectively, this seemingly contradictory positive correlation may reveal a paradoxical phenomenon in *UCHL1*-high tumors: the increased T/NK cell population may be accompanied by functional impairment or exhaustion rather than effective anti-tumor immunity. Future studies confirming that *UCHL1* directly regulates T/NK cell function in HCC will provide novel strategic insights for combining immunotherapy with *UCHL1*-targeted inhibition.

To date, a variety of multi-gene signatures have been proposed for prognostic evaluation of HCC, including signatures associated with immunity, metabolism, or ferroptosis (37).

In contrast, the 7-gene signature (*IGFBP3*, *UCHL1*, *GAGE2A*, *MAGEA6*, *AKR1B10*, *ADH4*, and *LCAT*) established in this study by integrating spatial and single-cell transcriptomic data not only enables dissection of cell-type-specific expression, but also reveals spatial distribution patterns within the tumor microenvironment.

This signature covers multiple biological processes, including alcohol metabolism (*ADH4*), cholesterol transport (*LCAT*), and ubiquitination (*UCHL1*) (38-40), forming a prognostic model with both multifunctionality and spatial awareness.

Furthermore, this risk model showed favorable predictive performance in both the TCGA and ICGC cohorts, and its independent validation results suggest that it possesses prognostic potential comparable to previously reported signatures.

Therefore, the core novelty of this study lies not only in the inclusion of a novel gene panel covering multiple biological processes, but also in the integration of spatial and single-cell analyses, which endows the HCC prognostic signature with a new dimension of spatial resolution. This distinguishes our work conceptually and methodologically from previous studies primarily based on bulk transcriptome data.

Immune cell infiltration identified 10 types with differential abundance. TIDE scores differed between groups, and IC₅₀ analysis indicated correlations between drug sensitivity and risk score. The prognostic gene expression was visualized in single-cell and spatial transcriptomics datasets. Heterogeneity analysis was performed on target cell types, which were further divided into subclusters. Pseudotime analysis suggested the presence

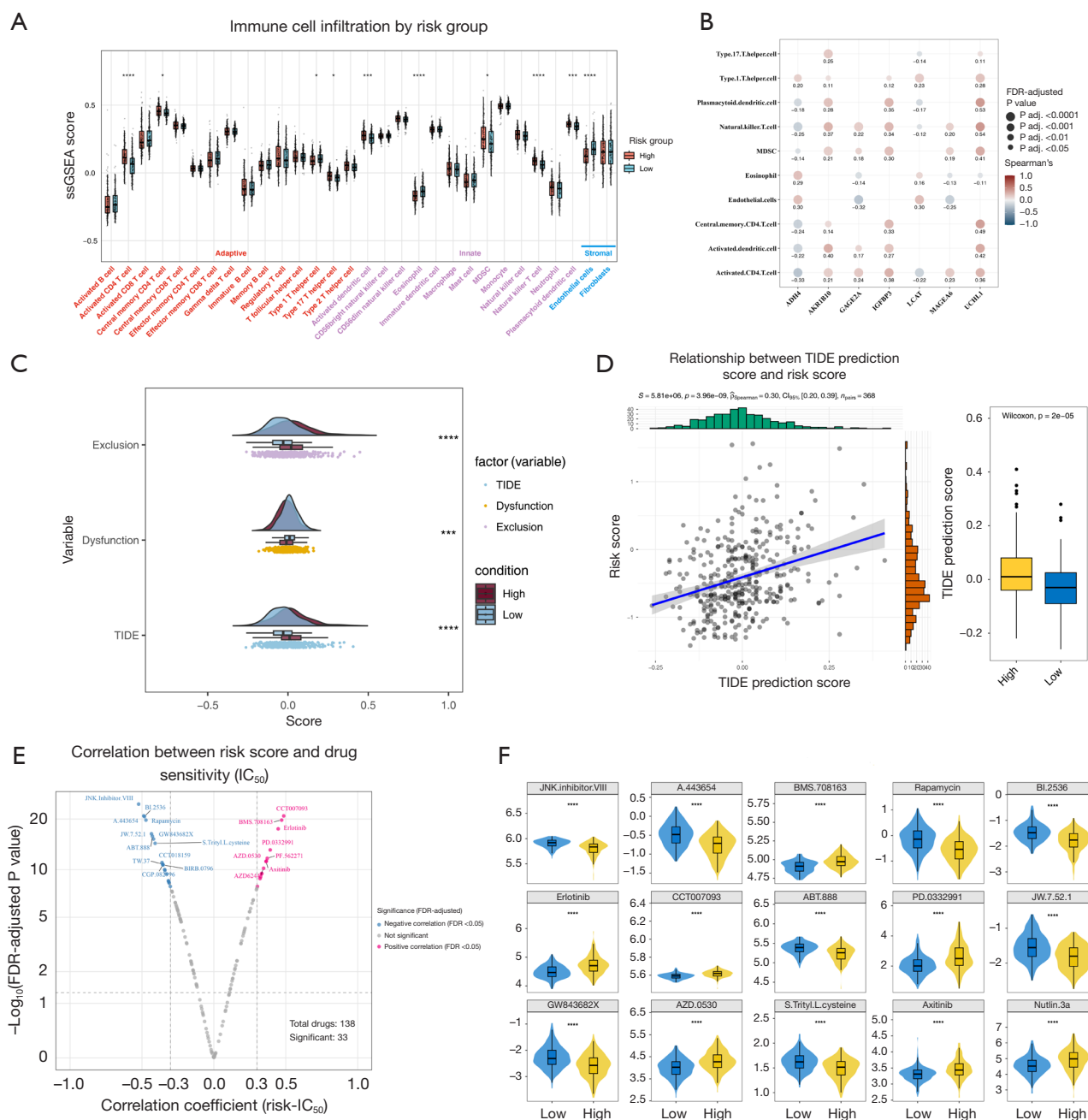


Figure 6 Characteristics of immunological infiltration that differed between the low- and high-risk groups. (A) Significant differences in ssGSEA scores were noted between risk groups for 10 immune cells (FDR < 0.05). (B) UCHL1 positively correlated with NK T cells (cor = 0.54, FDR < 0.05), while ADH4 negatively correlated with activated CD4 T cells (cor = -0.33, FDR < 0.05). (C) The high-risk group exhibited higher exclusion and TIDE scores, with a lower dysfunction score. (D) A positive correlation between TIDE and risk scores further suggests a potential reduced sensitivity to immunotherapy in high-risk patients. (E) Spearman's correlation analysis identified 33 drugs significantly linked to risk scores, with 13 positively and 20 negatively correlated. (F) High-risk patients showed lower IC_{50} values for negatively correlated drugs like A.443654 and rapamycin, and higher IC_{50} values for positively correlated drugs such as axitinib and erlotinib. *, P < 0.05; ***, P < 0.001; ****, P < 0.0001. Cor, correlation coefficient; FDR, false discovery rate; IC_{50} , half-maximal inhibitory concentration; MDSC, myeloid-derived suppressor cell; NK, natural killer; ns, not significant; ssGSEA, single-set gene set enrichment analysis; TIDE, tumor immune dysfunction and exclusion.

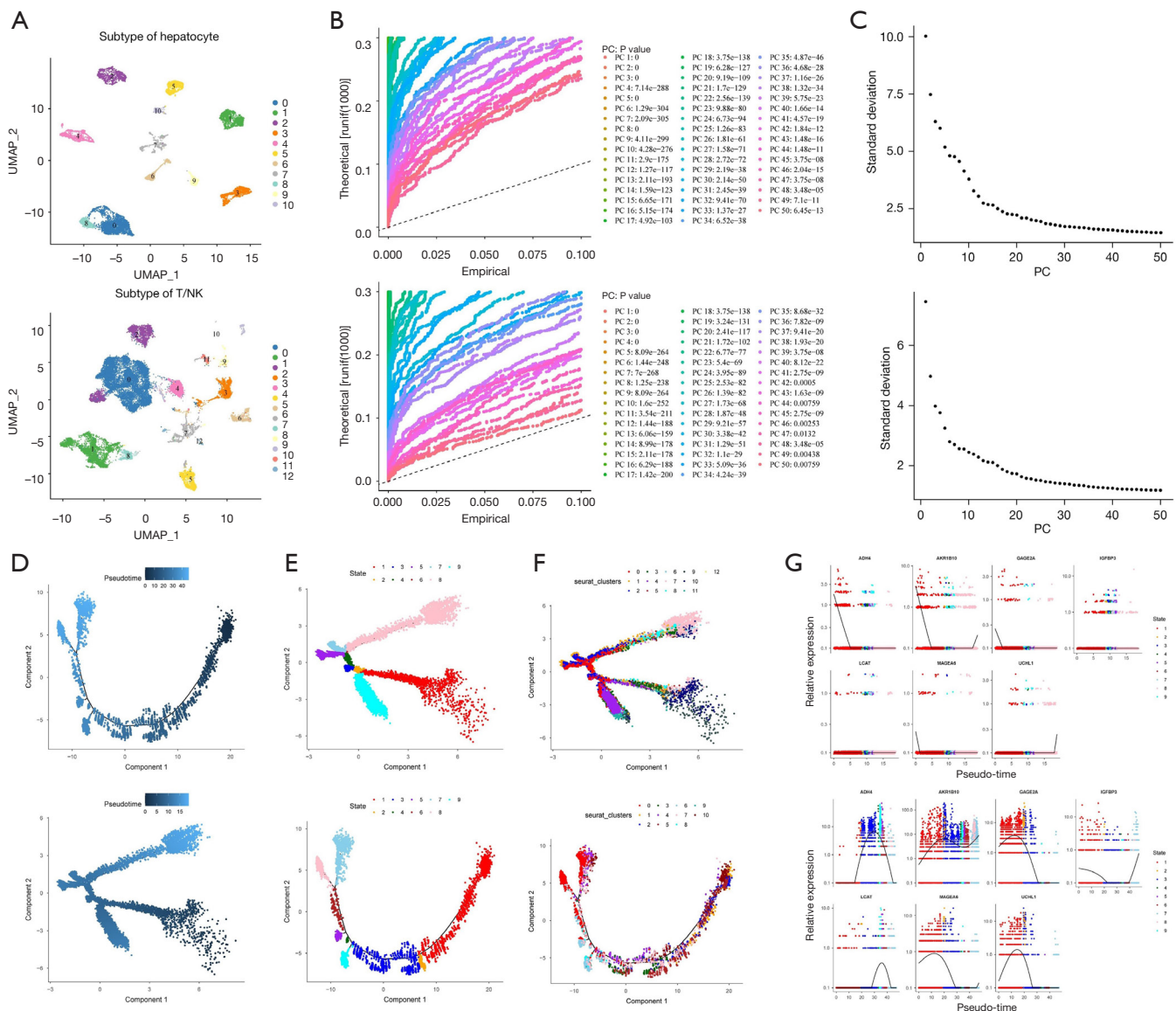
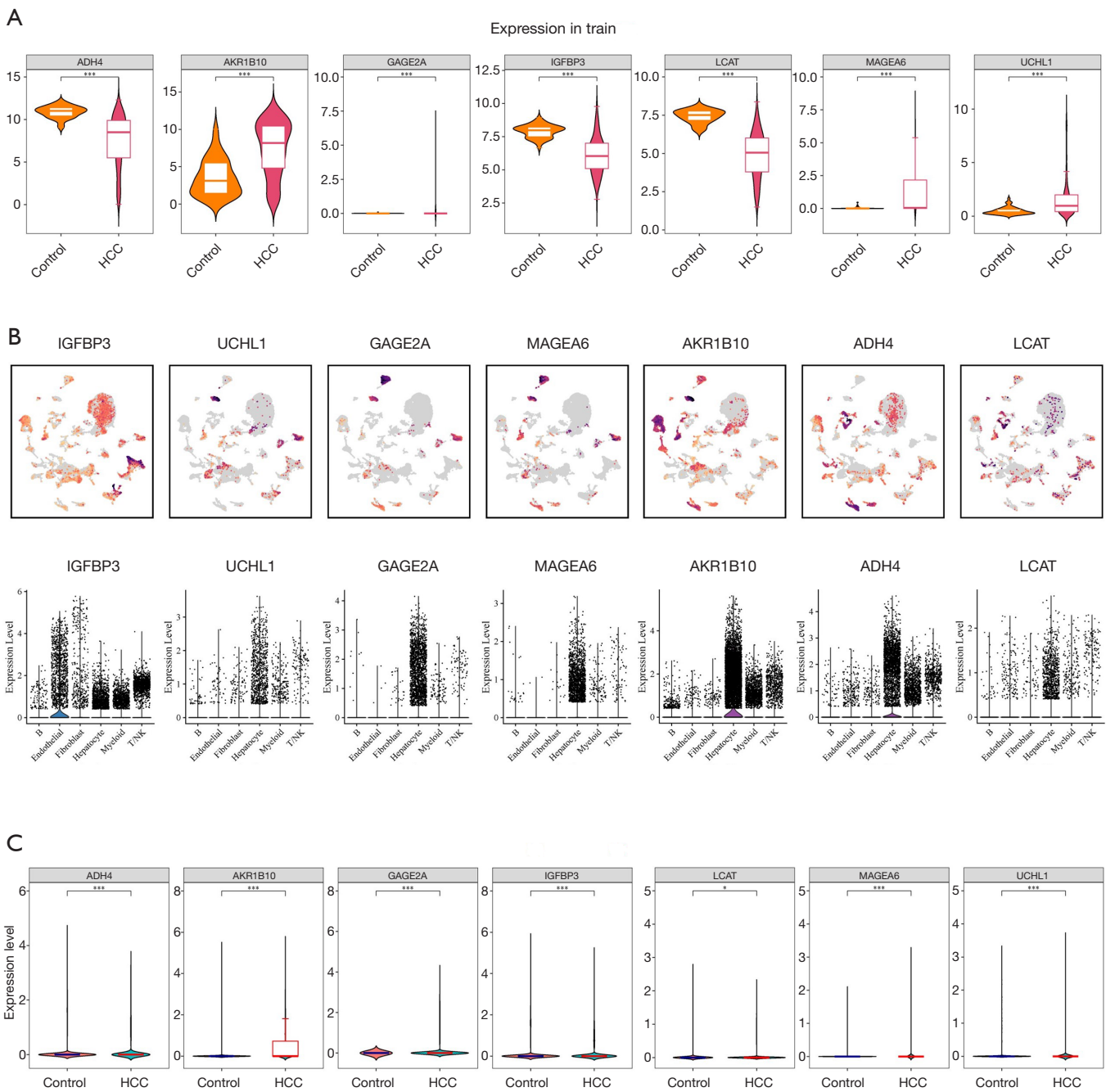


Figure 7 The pseudotime analysis of the differentiation of hepatocytes and T/NK cells. (A-C) The top 30 PCs were used for clustering hepatocytes into 11 subtypes and T/NK cells into 13 subtypes. (D,E) Pseudotime analysis identified nine differentiation states for hepatocytes and T/NK cells, with stage 1 as the initial stage and stages 7/8 and 8 as the terminal stages for hepatocytes and T/NK cells, respectively. (F) Hepatocyte subtypes 1, 2, 5, and 7 were prevalent in early stages, while subtype 0 was more common in later stages. For T/NK cells, subtypes 3, 5, 9, and 10 were more common in early stages, and subtype 7 was more abundant in later stages. (G) At the end of T/NK cell differentiation, *UCHL1* was upregulated, but *ADH4*, *GAGE2A*, and *MAGEA6* were downregulated. NK, natural killer; PC, principal component; UMAP, Uniform Manifold Approximation and Projection.

of dynamic changes in prognostic gene expression. Cell-cell communication analysis demonstrated intercellular relationships at the single-cell and spatial levels.

The results of this study provide new insights into the clinical management of HCC and also highlight the

challenges in clinical translation. The risk model constructed based on seven prognostic genes can serve as a prognostic stratification tool beyond traditional staging systems to help identify high-risk patients. Our analyses indicate that the high-risk group (associated with high *UCHL1*



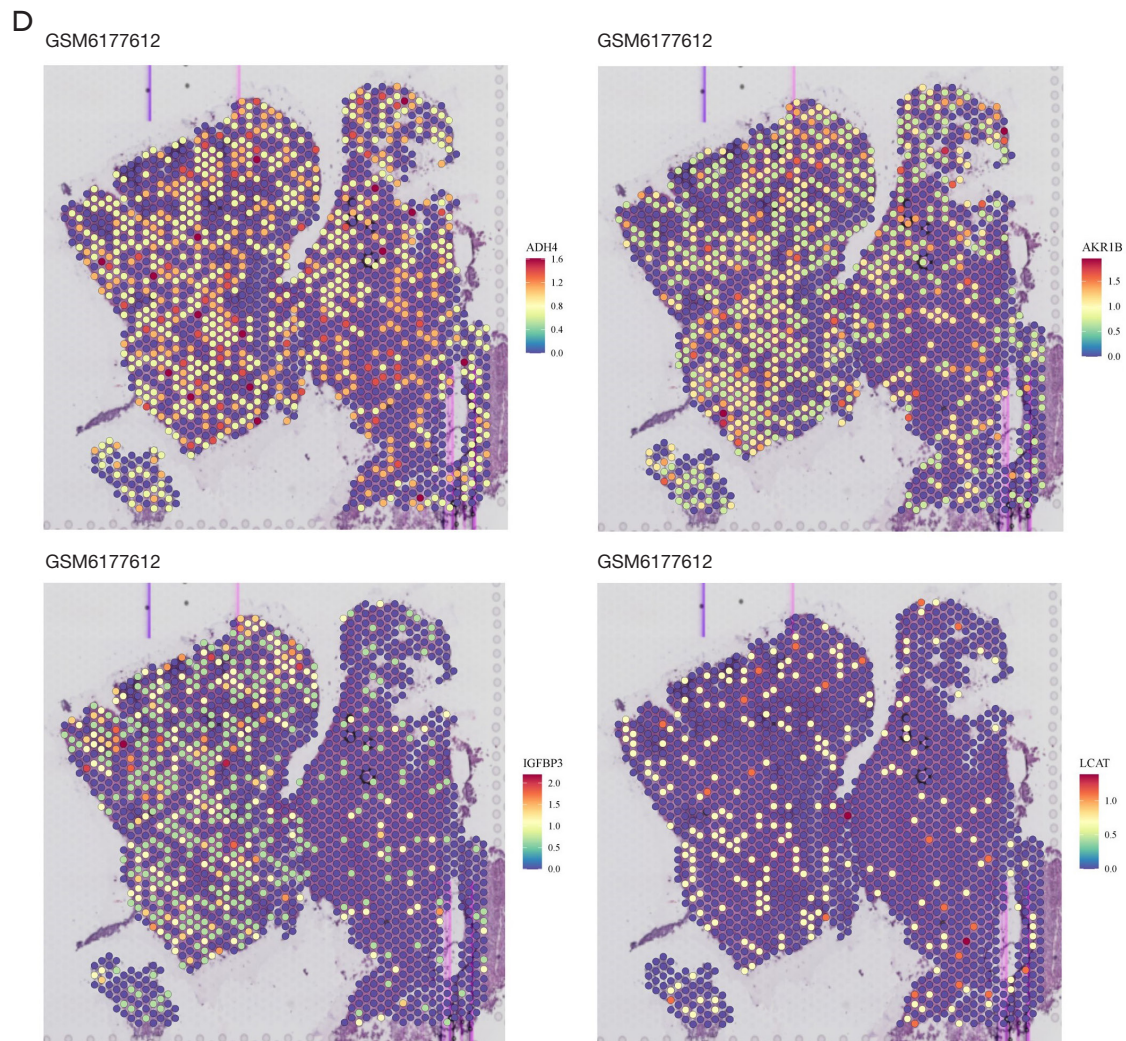


Figure 8 The expression of prognostic genes in common transcriptome, single cell, and spatial transcriptome datasets. (A) In the TCGA-HCC dataset, *ADH4*, *IGFBP3*, and *LCAT* were down-regulated, while *AKR1B10*, *GAGE2A*, *MAGEA6*, and *UCHL1* were up-regulated in HCC ($P < 0.001$). (B) In the GSE149614 dataset, *IGFBP3* was expressed in all six cell types, while the other five prognostic genes were more abundant in hepatocytes. (C) All prognostic genes showed significant expression differences between HCC and controls ($P < 0.05$). (D) Spatial expression of these genes in HCC cell subclusters was visualized in the GSE203612 dataset. *, $P < 0.05$; ***, $P < 0.001$. HCC, hepatocellular carcinoma; NK, natural killer; TCGA, The Cancer Genome Atlas.

expression) exhibits stronger immune exclusion features, suggesting that this model or its key genes (e.g., *UCHL1*) may be used to predict sensitivity to immunotherapy and thus guide individualized medication (41). In addition, the functions of the relevant genes (e.g., *UCHL1*, *ADH4*) imply novel therapeutic directions; for instance, the strategy of combining *UCHL1* inhibitors with immunotherapy may be explored in the future (35).

Nevertheless, critical bottlenecks must be overcome for

clinical translation: first, the model requires independent validation in prospective, multicenter clinical cohorts, and a standardized detection pipeline suitable for clinical samples needs to be established. Second, the current conclusions are mainly derived from computational analyses and must be verified by cellular and animal experiments to clarify the specific mechanisms of these genes (e.g., *UCHL1*) in HCC progression and immune regulation. Finally, drug development targeting novel biomarkers is inherently

characterized by long cycles and high risks. Therefore, future work should focus on promoting the translation of these findings into preclinical and clinical studies to gradually realize their clinical application value.

The insights from this study remain preliminary, and so further in-depth investigation based on the specific areas of interest is needed. This study has the following limitations. First, all findings were based on bioinformatics analyses of public datasets and lacked experimental validation. Therefore, future studies are required to validate the expression of key genes in independent patient cohorts using techniques such as quantitative real-time polymerase chain reaction and immunohistochemistry, and to perform *in vitro* and *in vivo* functional verification via gene knockout/overexpression, cellular functional assays, and animal models.

Second, although we constructed a predictive model and revealed gene expression patterns, the specific molecular mechanisms underlying their roles in HCC remain unclear. Further research should focus on constructing gene regulatory networks, screening potential targets combined with molecular docking, and thoroughly exploring downstream signaling pathways, PPIs, and their effects on the immune microenvironment.

Finally, spatial transcriptomic analysis was performed on only a single tumor sample; the results were limited by the high spatial heterogeneity of tumors and may not represent the overall characteristics of HCC. Future studies should integrate spatial transcriptomic data from multiple HCC patients to systematically validate the reliability and clinical significance of the spatial distribution patterns of cell subsets and their interaction networks observed in this study.

Conclusions

Based on a series of bioinformatics analysis, prognostic genes (*IGFBP3*, *UCHL1*, *GAGE2A*, *MAGEA6*, *AKR1B10*, *ADH4*, and *LCAT*) related to HCC prognosis were identified. Further research was conducted on the mechanisms of prognostic gene changes in the pathological process of HCC for the purpose of generating new reference data for the development of targeted drugs in HCC. Furthermore, the study examined, in depth, the impact of the immune microenvironment on the progression of HCC revealed that immune cells, including T/NK cells, may be associated with the progression of HCC, providing new perspectives for the development of candidate therapeutic targets. Finally, based on the expression patterns of the prognostic

genes during key cell differentiation processes, we further analyzed the potential connections between the prognostic genes and changes in hepatocyte secretome. This may constitute a novel approach and future research direction for understanding the pathogenic mechanisms of HCC and improving the prognosis of patients with HCC.

Acknowledgments

None.

Footnote

Reporting Checklist: The authors have completed the TRIPOD reporting checklist. Available at <https://jgo.amegroups.com/article/view/10.21037/jgo-2026-1-0014/rc>

Peer Review File: Available at <https://jgo.amegroups.com/article/view/10.21037/jgo-2026-1-0014/prf>

Funding: This study was supported by the Guangxi Medical and Health Key Discipline Construction Project.

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at <https://jgo.amegroups.com/article/view/10.21037/jgo-2026-1-0014/coif>). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki and its subsequent amendments.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

1. Zhou H, Song T. Conversion therapy and maintenance

- therapy for primary hepatocellular carcinoma. *Biosci Trends* 2021;15:155-60.
2. Song F, Wang CG, Mao JZ, et al. PANOptosis-based molecular subtyping and HPAN-index predicts therapeutic response and survival in hepatocellular carcinoma. *Front Immunol* 2023;14:1197152.
 3. Liu Z, Ma H, Lai Z. The Role of Ferroptosis and Cuproptosis in Curcumin against Hepatocellular Carcinoma. *Molecules* 2023;28:1623.
 4. Shi ZR, Duan YX, Cui F, et al. Integrated proteogenomic characterization reveals an imbalanced hepatocellular carcinoma microenvironment after incomplete radiofrequency ablation. *J Exp Clin Cancer Res* 2023;42:133.
 5. Chen Y, Huang W, Ouyang J, et al. Identification of Anoikis-Related Subgroups and Prognosis Model in Liver Hepatocellular Carcinoma. *Int J Mol Sci* 2023;24:2862.
 6. Ma J, Jin J, Lu H, et al. Exonuclease 1 is a Potential Diagnostic and Prognostic Biomarker in Hepatocellular Carcinoma. *Front Mol Biosci* 2022;9:889414.
 7. Lei Y, Tang R, Xu J, et al. Applications of single-cell sequencing in cancer research: progress and perspectives. *J Hematol Oncol* 2021;14:91.
 8. Saviano A, Henderson NC, Baumert TF. Single-cell genomics and spatial transcriptomics: Discovery of novel cell states and cellular interactions in liver physiology and disease biology. *J Hepatol* 2020;73:1219-30.
 9. Wang G, Zhou Z, Jin W, et al. Single-cell transcriptome sequencing reveals spatial distribution of IL34(+) cancer-associated fibroblasts in hepatocellular carcinoma tumor microenvironment. *NPJ Precis Oncol* 2023;7:133.
 10. Saqib J, Park B, Jin Y, et al. Identification of Niche-Specific Gene Signatures between Malignant Tumor Microenvironments by Integrating Single Cell and Spatial Transcriptomics Data. *Genes (Basel)* 2023;14:2033.
 11. Zhang S, Yuan L, Danilova L, et al. Spatial transcriptomics analysis of neoadjuvant cabozantinib and nivolumab in advanced hepatocellular carcinoma identifies independent mechanisms of resistance and recurrence. *Genome Med* 2023;15:72.
 12. Huang H, Wu F, Yu Y, et al. Multi-transcriptomics analysis of microvascular invasion-related malignant cells and development of a machine learning-based prognostic model in hepatocellular carcinoma. *Front Immunol* 2024;15:1436131.
 13. Yu L, Shen N, Shi Y, et al. Characterization of cancer-related fibroblasts (CAF) in hepatocellular carcinoma and construction of CAF-based risk signature based on single-cell RNA-seq and bulk RNA-seq data. *Front Immunol* 2022;13:1009789.
 14. Gui M, Huang S, Li S, et al. Integrative single-cell transcriptomic analyses reveal the cellular ontological and functional heterogeneities of primary and metastatic liver tumors. *J Transl Med* 2024;22:206.
 15. Liu S, Xie X, Lei H, et al. Identification of Key circRNAs/lncRNAs/miRNAs/mRNAs and Pathways in Preeclampsia Using Bioinformatics Analysis. *Med Sci Monit* 2019;25:1679-93.
 16. Griss J, Viteri G, Sidiropoulos K, et al. ReactomeGSA - Efficient Multi-Omics Comparative Pathway Analysis. *Mol Cell Proteomics* 2020;19:2115-25.
 17. Chin CH, Chen SH, Wu HH, et al. cytoHubba: identifying hub objects and sub-networks from complex interactome. *BMC Syst Biol* 2014;8 Suppl 4:S11.
 18. Fang Z, Li J, Cao F, et al. Integration of scRNA-Seq and Bulk RNA-Seq Reveals Molecular Characterization of the Immune Microenvironment in Acute Pancreatitis. *Biomolecules* 2022;13:78.
 19. Yan C, Niu Y, Li F, et al. System analysis based on the pyroptosis-related genes identifies GSDMC as a novel therapy target for pancreatic adenocarcinoma. *J Transl Med* 2022;20:455.
 20. Li J, Wang H, Cao F, et al. A bioinformatics analysis for diagnostic roles of the E2F family in esophageal cancer. *J Gastrointest Oncol* 2022;13:2115-31.
 21. Zhang W, Shi G, Wang H, et al. Molecular Mechanism of Xingnao Kaiqiao Pill for Perioperative Neurocognitive Disorder and Its Correlation With Immune and Inflammatory Signaling Pathways Based on Network Pharmacology and Molecular Docking. *Front Aging Neurosci* 2022;14:925072.
 22. Shen G, Wang WL. Circlize package in R and Analytic Hierarchy Process (AHP): Contribution values of ABCDE and AGL6 genes in the context of floral organ development. *PLoS One* 2022;17:e0261232.
 23. Yu G, Wang LG, Han Y, et al. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 2012;16:284-7.
 24. Liu TT, Li R, Huo C, et al. Identification of CDK2-Related Immune Forecast Model and ceRNA in Lung Adenocarcinoma, a Pan-Cancer Analysis. *Front Cell Dev Biol* 2021;9:682002.

25. Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *J Stat Softw* 2010;33:1-22.
26. Wang S, Su W, Zhong C, et al. An Eight-CircRNA Assessment Model for Predicting Biochemical Recurrence in Prostate Cancer. *Front Cell Dev Biol* 2020;8:599494.
27. Yang H, Wang K, Li B, et al. The Prognostic Role of Blood Inflammatory Biomarkers and EGFR Mutation Status in Stage IIIA/N2 Non-Small Cell Lung Cancer Patients Treated With Trimodality Therapy. *Front Oncol* 2021;11:707041.
28. Chen J, Sun M, Chen C, et al. Identification of hub genes and their correlation with infiltration of immune cells in MYCN positive neuroblastoma based on WGCNA and LASSO algorithm. *Front Immunol* 2022;13:1016683.
29. Hänzelmann S, Castelo R, Guinney J. GSEA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics* 2013;14:7.
30. Xu Q, Chen S, Hu Y, et al. Landscape of Immune Microenvironment Under Immune Cell Infiltration Pattern in Breast Cancer. *Front Immunol* 2021;12:711433.
31. Ni L, Yang H, Wu X, et al. The expression and prognostic value of disulfidptosis progress in lung adenocarcinoma. *Aging (Albany NY)* 2023;15:7741-59.
32. Wang Z, Qin H, Liu S, et al. Precision diagnosis of hepatocellular carcinoma. *Chin Med J (Engl)* 2023;136:1155-65.
33. Song M, Huang Y, Hong Y, et al. PD-L1-expressing natural killer cells predict favorable prognosis and response to PD-1/PD-L1 blockade in neuroblastoma. *Oncoimmunology* 2024;13:2289738.
34. Yao J, Reyimu A, Sun A, et al. UCHL1 acts as a potential oncogene and affects sensitivity of common anti-tumor drugs in lung adenocarcinoma. *World J Surg Oncol* 2022;20:153.
35. Mitra S, Epshtein Y, Sammani S, et al. UCHL1, a deubiquitinating enzyme, regulates lung endothelial cell permeability in vitro and in vivo. *Am J Physiol Lung Cell Mol Physiol* 2021;320:L497-507.
36. Xia X, Xu Q, Liu M, et al. Deubiquitination of CD36 by UCHL1 promotes foam cell formation. *Cell Death Dis* 2020;11:636.
37. Chen ZD, Luo JY, Ye YP, et al. Identification of immune-related genes and patient selection for hepatocellular carcinoma immunotherapy. *Transl Cancer Res* 2023;12:1210-31.
38. Li L, Huang YT, Wang LT, et al. ADH4-a potential prognostic marker for hepatocellular carcinoma with possible immune-related implications. *BMC Cancer* 2024;24:927.
39. Xu M, Xie P, Liu S, et al. LCAT deficiency promotes hepatocellular carcinoma progression and lenvatinib resistance by promoting triglyceride catabolism and fatty acid oxidation. *Cancer Lett* 2025;612:217469.
40. Yang G, Fan G, Zhang T, et al. Upregulation of Ubiquitin Carboxyl-Terminal Hydrolase L1 (UCHL1) Mediates the Reversal Effect of Verapamil on Chemo-Resistance to Adriamycin of Hepatocellular Carcinoma. *Med Sci Monit* 2018;24:2072-82.
41. Lu H, Zheng LY, Wu LY, et al. The immune escape signature predicts the prognosis and immunotherapy sensitivity for pancreatic ductal adenocarcinoma. *Front Oncol* 2022;12:978921.

(English Language Editor: J. Gray)

Cite this article as: Zhang Y, Yao Y, Yan Z, He M, Tang Y, Zhang Z. Prognostic genes in hepatocellular carcinoma and their function: an analysis integrating high-throughput-based spatial transcriptomics and single-cell RNA-sequencing. *J Gastrointest Oncol* 2026;17(2):75. doi: 10.21037/jgo-2026-1-0014