



OPEN ACCESS

EDITED BY

Zhijie Xu,
Central South University, China

REVIEWED BY

Chen Xue,
Zhejiang University, China
Dingyuan Zeng,
Guangxi University, China

*CORRESPONDENCE

Qian Zhang,
✉ zqian@jlu.edu.cn

RECEIVED 12 September 2023

ACCEPTED 06 November 2023

PUBLISHED 17 November 2023

CITATION

Qi W and Zhang Q (2023), Insights on epithelial cells at the single-cell level in hepatocellular carcinoma prognosis and response to chemotherapy.
Front. Pharmacol. 14:1292831.
doi: 10.3389/fphar.2023.1292831

COPYRIGHT

© 2023 Qi and Zhang. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). 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.

Insights on epithelial cells at the single-cell level in hepatocellular carcinoma prognosis and response to chemotherapy

Wenqian Qi and Qian Zhang*

Department of digestive, China-Japan Union Hospital of Jilin University, Changchun, China

Background: Hepatocellular carcinoma (HCC) originates from Epithelial cells, and epithelial lineage plasticity has become a promising research direction for advancing HCC treatment. This study aims to focus on Epithelial cells to provide target insights for detecting HCC prognosis and response to drug therapy.

Methods: Single-cell RNA sequencing (scRNA-seq) data from GSE149614 were clustered using Seurat, and the differentiation and evolution of epithelial cells were analyzed by Monocle 2. Scissor+ and Scissor- Epithelial cells associated with the prognostic phenotypes of bulk RNA-seq of HCC were screened using the Scissor algorithm for differential analysis to screen candidate genes. Candidate genes were overlapped with prognostic related genes screened by univariate Cox regression, and the Least Absolute Shrinkage and Selection Operator (LASSO) sparse penalty was imposed on the intersection genes to construct a risk assessment system.

Results: Eight major cell subpopulations of HCC were identified, among which the proportion of epithelial cells in non-tumor liver tissues and HCC tissues was significantly different, and its proportion increased with advanced clinical stage. During the progression of HCC, the whole direction of epithelial cells differentiation trajectory was towards enhanced cell proliferation. Differential analysis between Scissor+ and Scissor- epithelial cells screened 1,265 upregulated and 191 downregulated prognostic candidate genes. Wherein, the upregulated genes were enriched in Cell processes, Genetic information processing, Metabolism and Human disease with Infection. Nevertheless, immune system related pathways took the main proportions in downregulated genes enriched pathways. There were 17 common genes between upregulated candidate genes and prognostic risk genes, of which CDC20, G6PD and PLOD2 were selected as components for constructing the risk assessment system. Risk score showed a significant correlation with tumor stage, epithelial-mesenchymal transition (EMT) related pathways and 22 therapeutic drugs, and was an independent prognostic factor for HCC.

Abbreviations: HCC, Hepatocellular carcinoma; scRNA-seq, Single-cell RNA sequencing; LASSO, Least Absolute Shrinkage and Selection Operator; GEO, Gene Expression Omnibus; PVTT, Portal vein tumor thrombus; TCGA, The Cancer Genome Atlas; HCCDB, Hepatocellular Carcinoma Database; UMAP, Uniform Manifold Approximation and Projection; RGE, Reversed graph embedding; DEG, Differentially expressed gene; BEAM, Branched Expression Analysis Modeling; FDR, False discovery rate; ROC, Receiver Operating Characteristic; AUC, The area under the ROC curve; EMT, Epithelial-mesenchymal transition; CDC20, Cell division cycle 20 homologue; G6PD, Glucose-6-phosphate dehydrogenase; STAT3, Signal transducer and activator of transcription 3; PLOD2, Procollagen-lysine,2-oxoglutarate 5-dioxygenase 2.

Conclusion: This study revealed the cellular composition of HCC, the differentiation evolution and functional landscape of epithelial cells in the further deterioration of HCC, and established a 3-gene risk model, which was closely related to clinical features, EMT, and drug sensitivity prediction. These findings provided insights in patient prognosis and drug therapy detection for HCC.

KEYWORDS

hepatocellular carcinoma, epithelial cells, differentiation, scissor algorithm, risk assessment system, prognosis, drug therapy

Introduction

Hepatocellular carcinoma (HCC) is a major liver tumor that threatens the lives of many people with poor prognosis around the world, accounting for 5% of all incident cases and 8.4% of all deaths due to cancer worldwide (Moon et al., 2020; Shu et al., 2022; Alqurashi et al., 2023; Romualdo et al., 2023). The little improvement in prognosis of HCC patients over the past 15 years is mainly attributed to inadequate monitoring of risk individuals and advanced tumor manifestations, with no effective treatment to achieve long-term survival of patients (Ogunwobi et al., 2019; Ahn et al., 2021; Huang et al., 2023). New biomarkers are needed to improve HCC detection, prognosis, treatment response prediction, and disease monitoring during therapy.

A healthy liver is static on mitosis, but after toxic injury or resection, cells can quickly enter the cell cycle to restore liver quality and function (Campana et al., 2021). Indeed, there is a high degree of plasticity among epithelial cells, hepatocytes and biliary epithelial cells in mammalian liver, which is related to the mechanism of tissue repair mechanisms and liver lesions such as cancer (Johnson, 2019). During HCC progression, the transition of epithelial cells to a mesenchymal phenotype exacerbates the motility and invasiveness of various epithelial cell types (Jayachandran et al., 2016). In this context, the scientific community believes that advances in the field of epithelial cell plasticity may hold great promise for the development of therapies for HCC patients (Ko et al., 2020). It is also worth mentioning that the concept of numerous markers of epithelial cells as therapeutic targets for liver cancer has been demonstrated in preclinical studies, such as CXCL5 combining with CXCR2 promoted epithelial-mesenchymal transition (EMT) via PI3K/Akt/GSK-3 β /Snail signaling (Xia et al., 2015; Zhou et al., 2015), Keratin 1 (Ogunnigbagbe et al., 2022), p63, the isoforms of which were related to tumor recurrence and reduced survival (Gonzalez et al., 2017), E-cadherin (Nakag et al., 2014), EpCAM, a latent marker for cancer stem cells, was reported to be intensely correlated with unfavorable clinical prognosis in HCC (Gerlach et al., 2018; Park et al., 2020). However, the cellular and molecular mechanisms of epithelial involvement in HCC remain unclear.

Single cell RNA sequencing (scRNA-seq) technology based on high-throughput sequencing at a single-cell resolution has been widely recognized as feasible and advantageous in dissecting tumor heterogeneity and analyzing the cellular mechanism and differentiation of important cell subsets (Ho et al., 2019). For example, a recent study in bladder cancer correlated scRNA-seq data disclosed the relationship between molecular and clinical features, guiding molecular diagnosis and targeted therapy

(Robertson et al., 2017). In the meantime, scRNA-seq is also applied to assess the effectiveness and safety of new drugs in clinical trials (Srivatsan et al., 2020). In this study, we aimed to explore the heterogeneity of HCC at the single-cell level, select epithelial cells to explore their differentiation evolution, and combine the analysis of bulk RNA-seq data of HCC to identify novel biomarkers of HCC to predict prognosis and response to drug therapy.

Materials and methods

Study omics of human liver cancer tissues

The omics data downloaded in this study included scRNA-seq omics data and bulk RNA-seq omics data. ScRNA-seq omics data were obtained from the Gene Expression Omnibus (GEO) database under accession number GSE149614, which gives 70,000 single-cell transcriptomes for 10 HCC patients from four relevant sites: primary tumor, portal vein tumor thrombus (PVTT), metastatic lymph node and non-tumor liver. The bulk RNA-seq omics data were obtained from three different data sources, one was the TCGA-LIHC dataset (n = 342) in The Cancer Genome Atlas (TCGA) database (<https://portal.gdc.cancer.gov/>) (Yan et al., 2022), one was the ICGC-LIRI-JP dataset (n = 212) from the Hepatocellular Carcinoma Database (HCCDB) database (<http://lifeome.net/database/hccdb/>), another one was the GSE14520 dataset (n = 221) from the GEO database.

Standard processing and clustering of scRNA-seq

ScRNA-seq data were read with the use of the Seurat package (Butler et al., 2018), and only cells with a gene number between 200 and 8,000 and a mitochondrial gene proportion of less than 10% were retained. SCTransform function was utilized to normalize the single-cell UMI count data based on a negative binomial regression model with regular parameters to remove the influence of deep sequencing. The RunHarmony function was used to set default parameters to integrate the transformed data and correct for batch effects. Then, with the aid of RunUMAP function, the dimension of Uniform Manifold Approximation and Projection (UMAP) was reduced, and FindClusters was used to perform graph-based clustering on the neighbor graph constructed by the FindNeighbors function call. The parameters of the two functions were set to $\text{dims} = 1:30$ and $\text{resolution} = 0.05$, respectively. The cell

identity of each cluster was obtained by referring to the cell marker resources provided by the CellMarker database.

Construction of pseudotime trajectory

Monocle 2 applies recently developed machine learning strategy that was described as reversed graph embedding (RGE) to accurately reconstruct complex single-cell trajectories (Qiu et al., 2017). We used monocle2 to read the count data of epithelial cells in the expression matrix, merged the phenotype information of the cells, used newCellDataSet function to construct cds objects, and filtered out the genes expressed in less than 10 cells. Then, the differentialGeneTest function (fullModelFormulaStr = "~ site") was adopted to calculate the differentially expressed genes (DEGs) between the primary tumor and metastatic samples of HCC patients. Using the reduceDimension function and setting the parameters max_components = 2, method = "DDRTree" to reduce the dimension, and under the control of the orderCells function, the cells were sorted and the trajectory was constructed. In particular, the branch of the primary tumor was set here as the starting point of the trajectory.

Differential expression analysis between trajectories branches

Monocle 2 provides Branched Expression Analysis Modeling (BEAM) method to analyze the branching of cell data and designated nodes to identify DEGs related to branching, and heatmap was generated using plot_genes_branched_heatmap to identify DEGs associated with clades. The plot_genes_branched_pseudotime function was also applied to visualize the expression of the interested genes at the branch point.

Capture of prognostic associated cells

Scissor is a method for quantifying the similarity between single-cell and batch data by measures such as Pearson's correlation for each pair of cells and batch samples to identify cell subsets from single-cell data associated with a given phenotype (Kearnes et al., 2014; Sun et al., 2022). The novelty of Scissor lies in its use of phenotypic information from bulk data to identify highly correlated cell subpopulations with diseases, thereby revealing disease mechanisms and improving disease diagnosis and treatment. The bulk expression matrix of TCGA-LIHC, ICGC and GSE14520 and the single-cell expression matrix of GSE149614 were input in Scissor to screen out the cell subsets related to prognostic phenotypes in HCC patients, and further classified them into Scissor+ and Scissor-cell subgroups.

Differential expression analysis of Scissor groups based on subgroups associated with prognostic phenotypes

The FindMarkers function in Seurat was used to identify DEGs between Scissor+ and Scissor- cell subgroups. The R

package RobustRankAggreg (Ogunwobi et al., 2019) then integrated the rankings of logFC values to obtain a comprehensive ranking list for each bulk RNA-seq cohort. Genes that showed the same trend of differential expression in the three cohorts and had an RRA score less than 0.05 were considered as candidate genes.

Construction of regression model

Univariate Cox regression of clinical survival data was performed in the three bulk RNA-seq cohorts to screen genes with $p < 0.05$ as prognostic related genes for HCC. The common prognostic risk genes (HR > 1) in the three bulk RNA-seq cohorts were used for overlap analysis with the candidate genes. LASSO sparse penalty was imposed on the overlapping genes, with high confidence to select the most important genes for the prognosis of HCC to build a risk regression model.

Drug sensitivity analysis of the risk regression model

The expression matrix of HCC in the TCGA-LIHC cohort was used as training data to predict drug response using the R package oncoPredict. The correlation between the half maximal inhibitory concentration (IC50) values of the resulting drugs and the risk regression model was analyzed by Spearman correlation test, in which the false discovery rate (FDR) was adjusted by Benjamini and Hochberg, and $FDR < 0.05$ and $|cor| > 0.5$ was defined as a significant correlation.

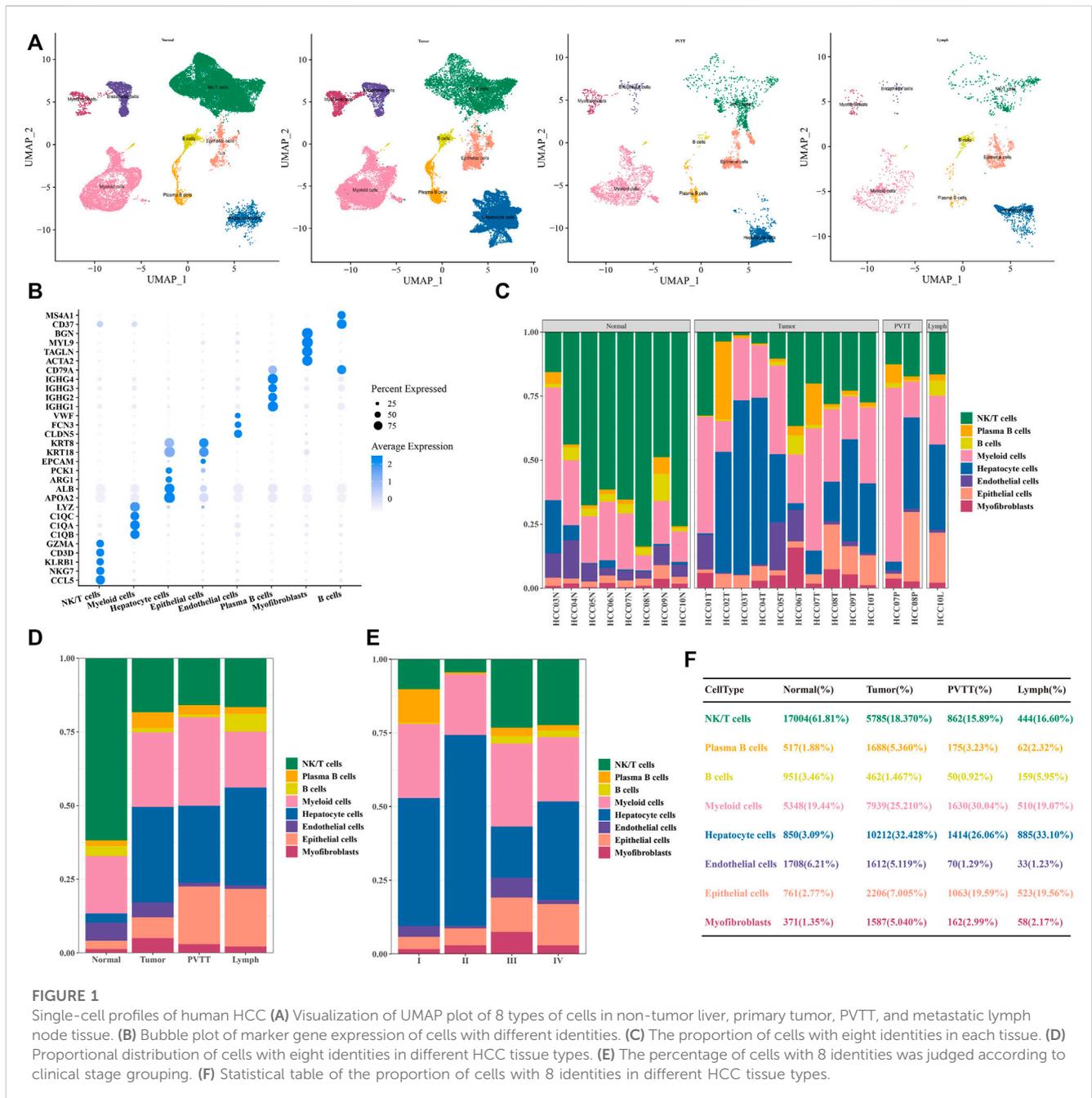
Statistical analysis

All statistical analyses were processed using R software. Functional enrichment analyses were processed in clusterProfiler. The analysis to evaluate the regression model included Kaplan-Meier survival analysis and Receiver Operating Characteristic (ROC) analysis, which were performed by the "survival" package and "pROC," respectively. Differences between two groups of continuous variables were assessed using the Wilcoxon rank-sum test, and differences between more than two groups of variables were compared using the Kruskal-Wallis test. The p -value for significance in this study was set at < 0.05 .

Results

Single-cell landscape of human HCC

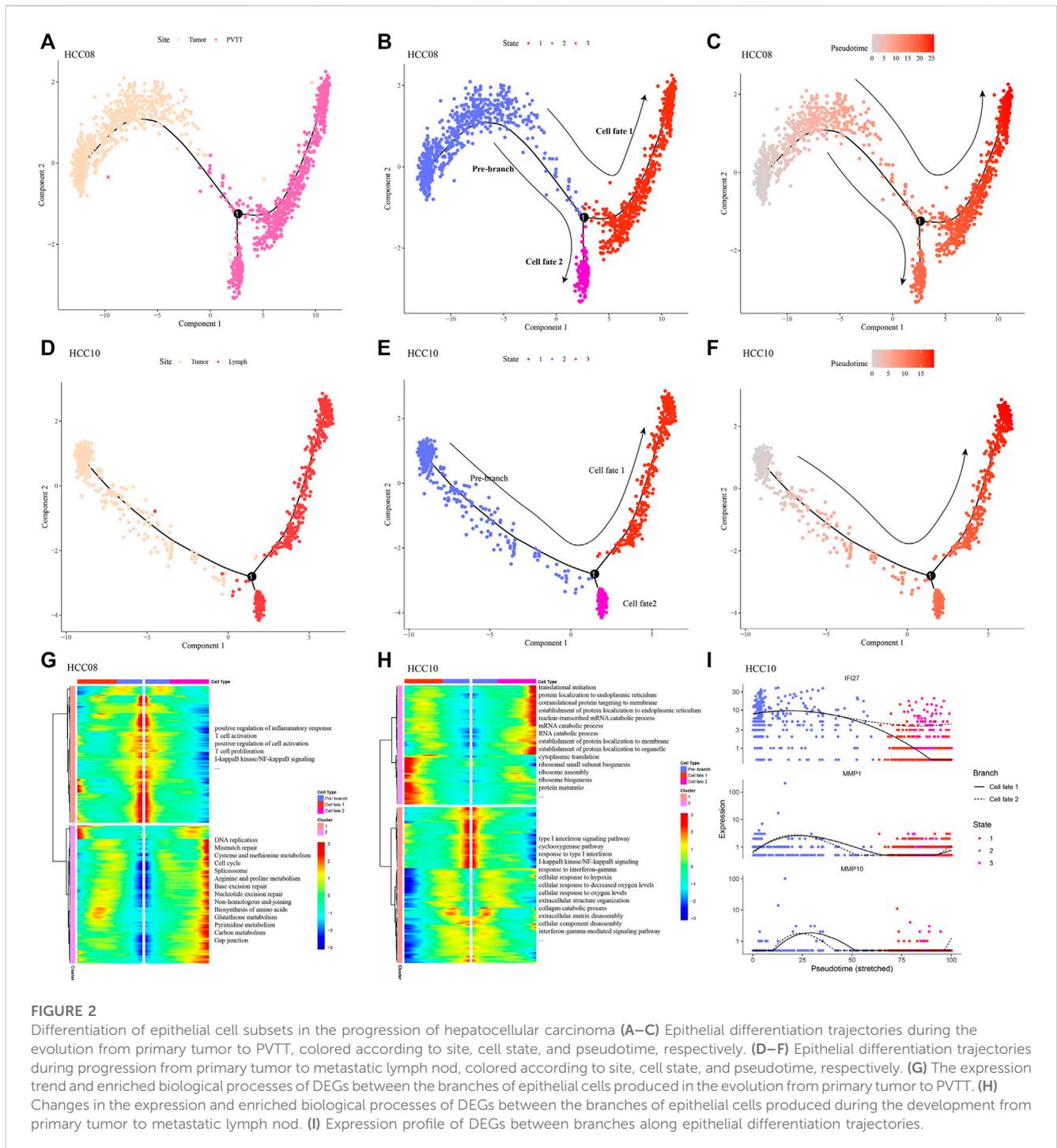
After the standard processing pipeline of scRNA-seq, cell clustering and annotation, a total of eight types of cell clusters were identified in the HCC of GSE149614. In terms of the distribution of 8 types of cells in different tissues of different origins, non-tumor liver distributed 27,510 cells, primary tumor distributed 31,491 cells, portal vein tumor thrombus (PVTT)



distributed 5,426 cells, and portal vein tumor thrombus (PVT) distributed 5,426 cells, the metastatic lymph nodes distributed 2,674 cells (Figure 1A). Each has its own specifically expressed genes, shown in Figure 1B. After our statistical analysis, we found that the proportion of NK/T cells in primary tumor, PVT and metastatic lymph node tissues was greatly reduced compared with non-tumor liver tissues. However, the proportion of hepatocyte cells and epithelial cells increased significantly (Figures 1C, D, F). Differences in the distribution of cell types were also observed among stages I-IV, and the proportion of epithelial cells increased with the advance of clinical stage (Figure 1E). Through the single-cell landscape analysis we revealed epithelial cells seemed to occupy an important position in the development of HCC.

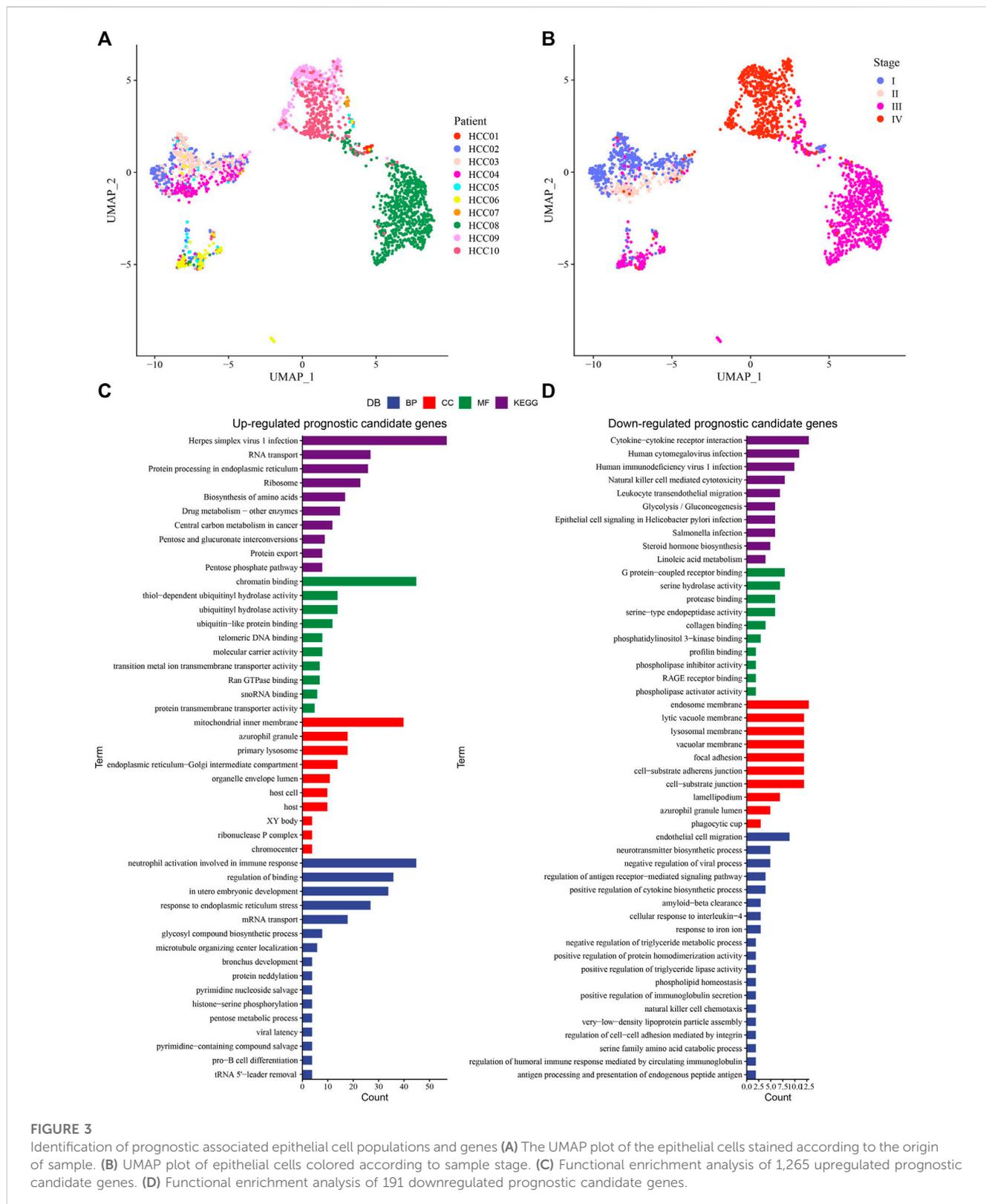
Differentiations of epithelial cell subsets in the progression of hepatocellular carcinoma

Given the difference in the proportion of epithelial cells between non-tumor liver tissues and HCC tissues, and the tendency for the proportion to increase with clinical stage, we extracted epithelial cells from the two HCC samples with the highest percentage of epithelial cells, primary tumor of HCC08, primary tumor and metastatic lymph nodes of PVT and HCC10 were included. Monocle2 was used to explore the dynamic changes of gene expression in epithelial cells during the development and progression of HCC, it was found that there were two different fates of epithelial cells in the evolution from primary tumor to PVT (Figures 2A–C): One of them



showed a gradual weakening of inflammatory response, T cell activation and proliferation, and the other showed a gradual strengthening of cell proliferation and metabolism (Figure 2G). Epithelial cells in the process of developing from primary tumor to metastatic lymph node also had two differentiation trajectories (Figures 2D–F): one was the gradual improvement of RNA catabolic, protein localization and protein translation ability, the other one was a tendency to weaken the strength of the immune response and the cellular response to hypoxia (Figure 2H). We also found that the

expression of IFI27, a differential gene between branches, was significantly higher in the differentiation pathway along cell fate 2 than in the differentiation pathway along cell fate 1. Two others branched DEGs, MMP1 and MMP10, were also presented during the differentiation of epithelial cells from primary tumor to metastatic lymph node (Figure 2I). These findings confirmed the action of Epithelial cells on cell proliferation, and forced us to propose a hypothesis that Epithelial cells may have an inhibitory effect on inflammatory response targeting HCC.



Identification of prognostic associated epithelial cell populations and genes

The epithelial cells in each primary tumor were selected and labeled according to their stage, and we found that the samples in stage I mainly

included HCC01, HCC02 and HCC03. HCC04 was at stage II, HCC05, HCC06, HCC07 and HCC08 were at stage III, and HCC09 and HCC10 were at stage IV (Figures 3A, B). Using the Scissor algorithm, we identified prognostic phenotypes-related Scissor+ and Scissor-epithelial cells, including 447 Scissor + cells and 184 Scissor-cells

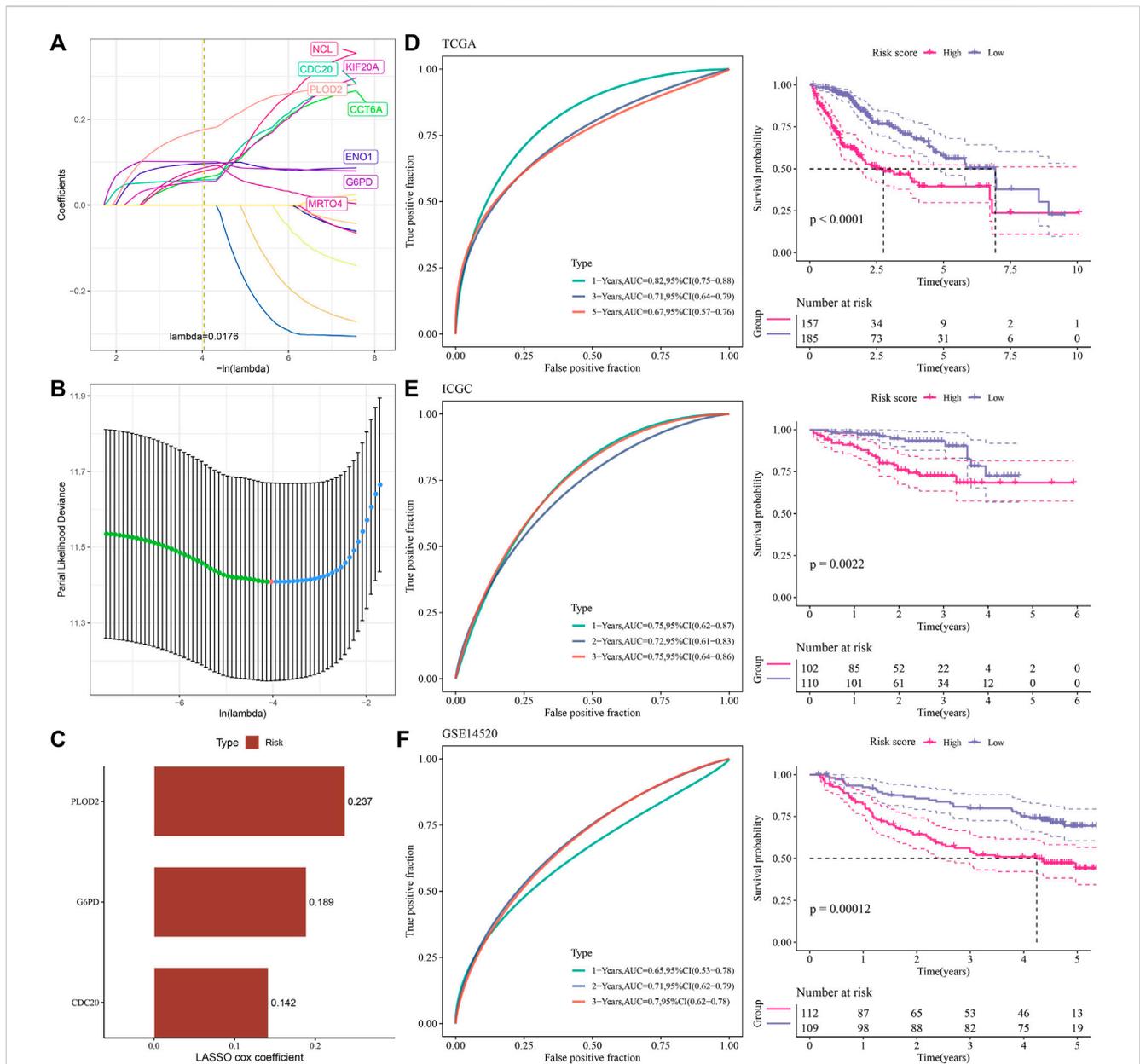


FIGURE 4

Derivation and validation of a prognostic assessment formula (A,B) LASSO sparse penalty was imposed on 17 gene. (C) The LASSO Cox coefficients of 3 genes selected by stepwise multifactor regression analysis. (D) ROC analysis and Kaplan-Meier survival analysis of the prognosis evaluation system in the TCGA-LIHC cohort. (E) ROC analysis and Kaplan-Meier survival analysis of the prognostic evaluation system in the ICGC cohort. (F) ROC curve and Kaplan-Meier curve of the prognostic evaluation system for the prognosis of patients in the GSE14520 cohort.

associated with the prognostic phenotypes of TCGA-LIHC, 673 Scissor+ cells and 131 Scissor-cells associated with prognostic phenotypes of ICGC, 877 Scissor+ cells and 10 Scissor-cells associated with prognostic phenotypes of GSE14520. Differential analysis between Scissor+ and Scissor- epithelial cells screened 1,456 DEGs, including 1,265 upregulated prognostic candidates and 191 downregulated prognostic candidates. We found that the upregulated genes were mainly enriched in Cell processes (cell cycle), Genetic information processing (RNA transport, Ribosome, etc.), Metabolism (Biosynthesis of amino acids) and Human disease with Infection (Herpes simplex virus 1 infection). As for enriched pathways with downregulated genes, immune system related pathways (Cytokine-cytokine receptor

interaction, Leukocyte transendothelial migration) and Human disease with Infection (Human cytomegalovirus infection and Human immunodeficiency virus 1 infection, etc.) took the main proportions. Collectively, the most significantly correlated pathways were those favoring cancer progression (Figures 3C, D).

Derivation and validation of a prognostic assessment formula

Using univariate Cox regression analysis, we obtained prognostic risk genes in the TCGA-LIHC, ICGC and

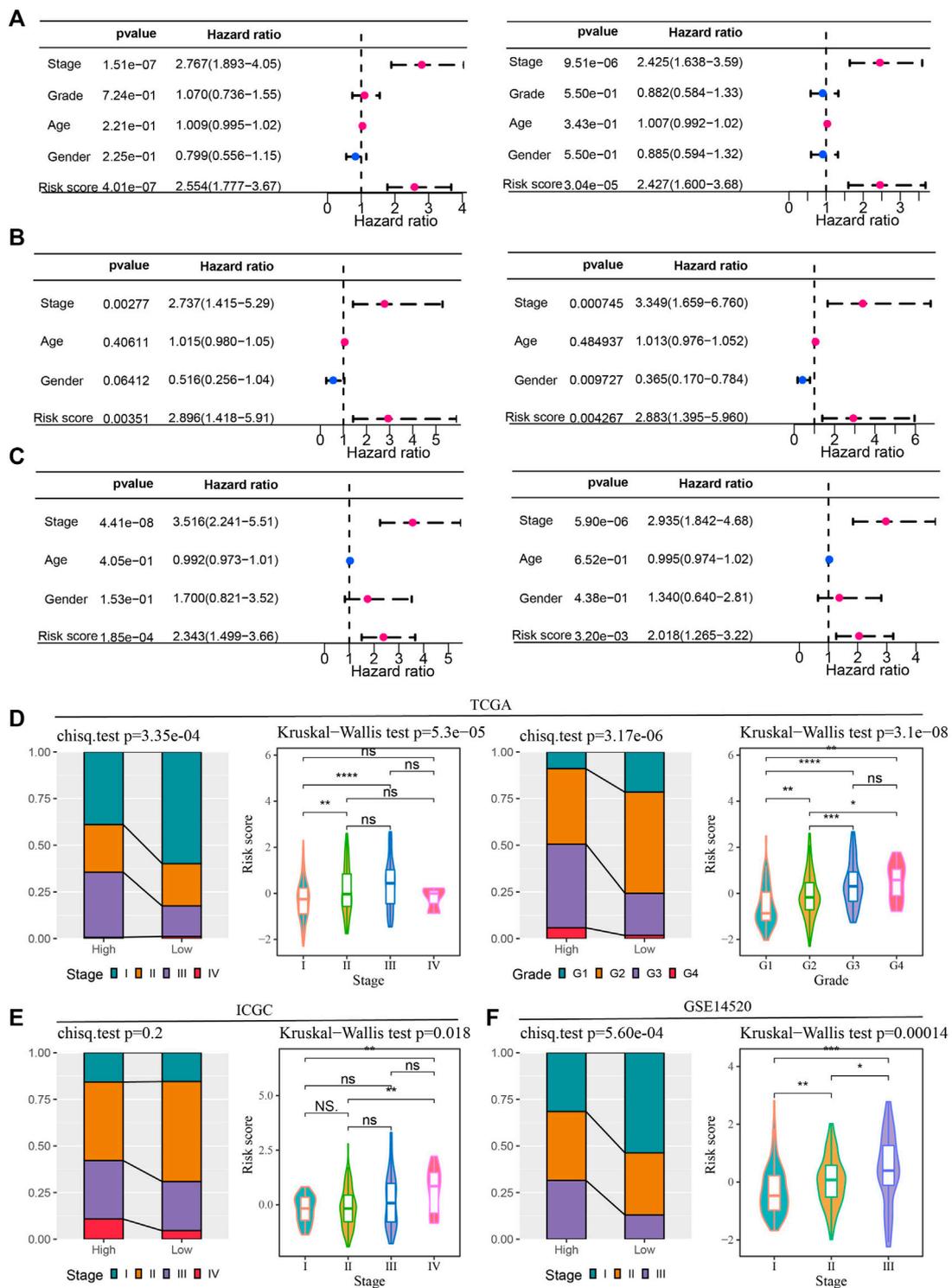


FIGURE 5

Risk score and clinical characteristics in predicting the prognosis of HCC and the relationship between them (A–C) Univariate and multivariate Cox regression analyses of clinical characteristics and risk scores in the TCGA-LIHC dataset, ICGC and GSE14520 cohorts. (D) Bar and violin plots to explore the association of risk score with stage and grade in TCGA-LIHC. (E) Stage distribution of high-risk and low-risk groups in the ICGC dataset and risk scores of samples in each stage. (F) Stage of samples grouped by risk score and risk score of samples grouped by stage in the GSE14520 dataset.

GSE14520 cohorts, which were overlapped with 1,265 upregulated prognostic candidate genes, and 17 genes were found in the overlapping part, including NCAPG, NCL, DBF4, ENO1,

KIF20A, RAN, MRT04, CDC20, CDK1, UBE2C, CCNB1, STMN1, CCT6A, DLGAP5, G6PD, SSB, PLOD2. LASSO sparse penalty was imposed on them, and eight genes had nonzero

coefficients (Figures 4A, B). In order to obtain the most parsimonious model with adequate fitting degree, stepwise multivariate regression analysis was performed, and three genes (CDC20, G6PD and PLOD2) were selected as components for the construction of risk assessment formula (Figure 4C). LASSO gave the coefficient for each gene, and the risk formula was: Risk Score = $+0.1424668 \times \text{CDC20} + 0.1886310 \times \text{G6PD} + 0.2367155 \times \text{PLOD2}$. The risk score was calculated in the bulk RNA-seq cohort of each HCC and prognosis was predicted. The risk score showed a significant inverse correlation with overall survival (OS). The accuracy of risk score in predicting prognosis was also tested by ROC curve. For 3-year OS, the area under the ROC curve (AUC) of the TCGA-LIHC, ICGC, and GSE14520 cohorts were 0.71, 0.75, and 0.7, respectively, indicating high prognostic accuracy of the model (Figures 4D–F).

Risk score and clinical characteristics in predicting the prognosis of HCC and the relationship between them

Univariate and multivariate Cox regression analyses of risk score and clinical factors were performed in 3 bulk RNA-seq cohorts of HCC to determine independent variables predicting HCC prognosis. Among the four clinical variables and risk scores provided by the TCGA-LIHC dataset, stage and risk score were independent variables that could independently predict HCC prognosis (Figure 5A). Univariate and multivariate Cox regression analysis based on clinical data and risk score in ICGA identified stage, gender and risk score as independent prognostic factors for HCC (Figure 5B). The conclusion obtained in the GSE14520 cohort was the same as that in the TCGA-LIHC dataset, that stage and risk score were independent prognostic predictors of HCC (Figure 5C). By correlating these independent prognostic variables, risk score and stage in TCGA-LIHC were found influence each other, and the proportion of stage III samples in the high-risk group was significantly higher than that in the low-risk group. The risk score of samples in stage III was significantly higher than that in stage I. Grade also showed a significant correlation with risk score. The proportion of G3-G3 patients composing the high-risk group was significantly higher than that of G3-G4 patients composing the low-risk group. Risk score were analyzed in each grade, and it was found that the risk scores of samples in G3 and G4 groups were significantly higher than those in G1 and G2 groups (Figure 5D). The high-risk and low-risk groups in the ICGC dataset did not show significant differences in stage distribution. However, the risk score of stage IV samples was significantly higher than that of stage I-stage II samples (Figure 5E). In the verification set GSE14520, stage was very significantly associated with risk score. Specifically, the proportion of stage III in the high-risk group was significantly higher than that in the low-risk group, and the risk score increased with the increase of risk score (Figure 5F).

Association of risk score with key biological processes occurring in cancer progression

The association between the risk score and pathophysiological events or carcinogenic factors occurring during cancer progression, including epithelial-mesenchymal transition (EMT), vascular stability, and hypoxia, was investigated. We found that EMT

related VEGF signaling pathway displayed significant enrichment score between high and low risk groups (Figure 6A). Then, the expression levels of VEGF-related genes were compared between high-risk and low-risk samples in the three datasets. VEGFA and VEGFB were more expressed in high-risk samples, and FLT4 was more expressed in low-risk samples (Figure 6C). CLDN5, JAM2 and TIE1 in vascular stability-related genes also showed significant negative correlation with risk score in TCGA-LIHC and GSE14520 datasets. Paradoxically, PCDH12 showed opposite trends in ICGC datasets and GSE14520 datasets, positively correlated with risk score in the former dataset, and negatively correlated with risk score in the latter dataset (Figure 6B). As a close connection between VEGF signaling pathway and Hypoxia in cancer development (Florentin et al., 2022), Hypoxia-related genes were also compared between high-risk and low-risk samples, and it is certain that HIF1A expression was significantly higher in high-risk samples than in low-risk samples (Figure 6D). Based on the above results, it was speculated that the risk score may partly affect the cancer progression.

The relationship between risk score and drug therapy

The IC50 value of each drug in the samples in TCGA-LIHC was calculated, and a total of 22 drugs were found to be significantly correlated with risk score. The IC50 values of 21 drugs were significantly negatively correlated with risk score, which may be more suitable for the treatment of HCC patients with high-risk score such as Vincristine, Vinblastine, Paclitaxel, Daporinad, Bortezomib and Axitinib, which are commonly used chemotherapy drugs for HCC. The IC50 value of SB505124, an inhibitor for TGF β receptor was more suitable for the treatment of patients with low-risk score (Figure 7).

Discussion

HCC originates from Epithelial cell population and consists of tumor cells, basement membrane, and surrounding stroma (Holczbauer et al., 2022; Yao et al., 2023). The phenotype appears to be closely associated with specific gene mutations, tumor subsets, and/or oncogenic pathways (Calderaro et al., 2019). In this study, we first focused on the composition landscape of cells in HCC under its tissue background, and emphasized the differentiation trajectory and functional influence of Epithelial cells, which was driven by the great regenerative potential of the liver, and its repair ability was mainly attributed to the ability of differentiated Epithelial cells, hepatocytes and biliary Epithelial cells to proliferate after injury (Ko et al., 2020).

The discovery of the hepatocyte profile will help decipher the molecular and cellular mechanisms that drive the healthy liver into disease states, provide insights into the detection of novel therapeutic targets, and pave the way for effective disease interventions (Ma et al., 2021). We deciphered the heterogeneity of HCC at the single cell level and identified 8 cell types. The proportion of Epithelial cells in HCC tissues increased with the increase of clinical stage. We also detected two distinct fates of

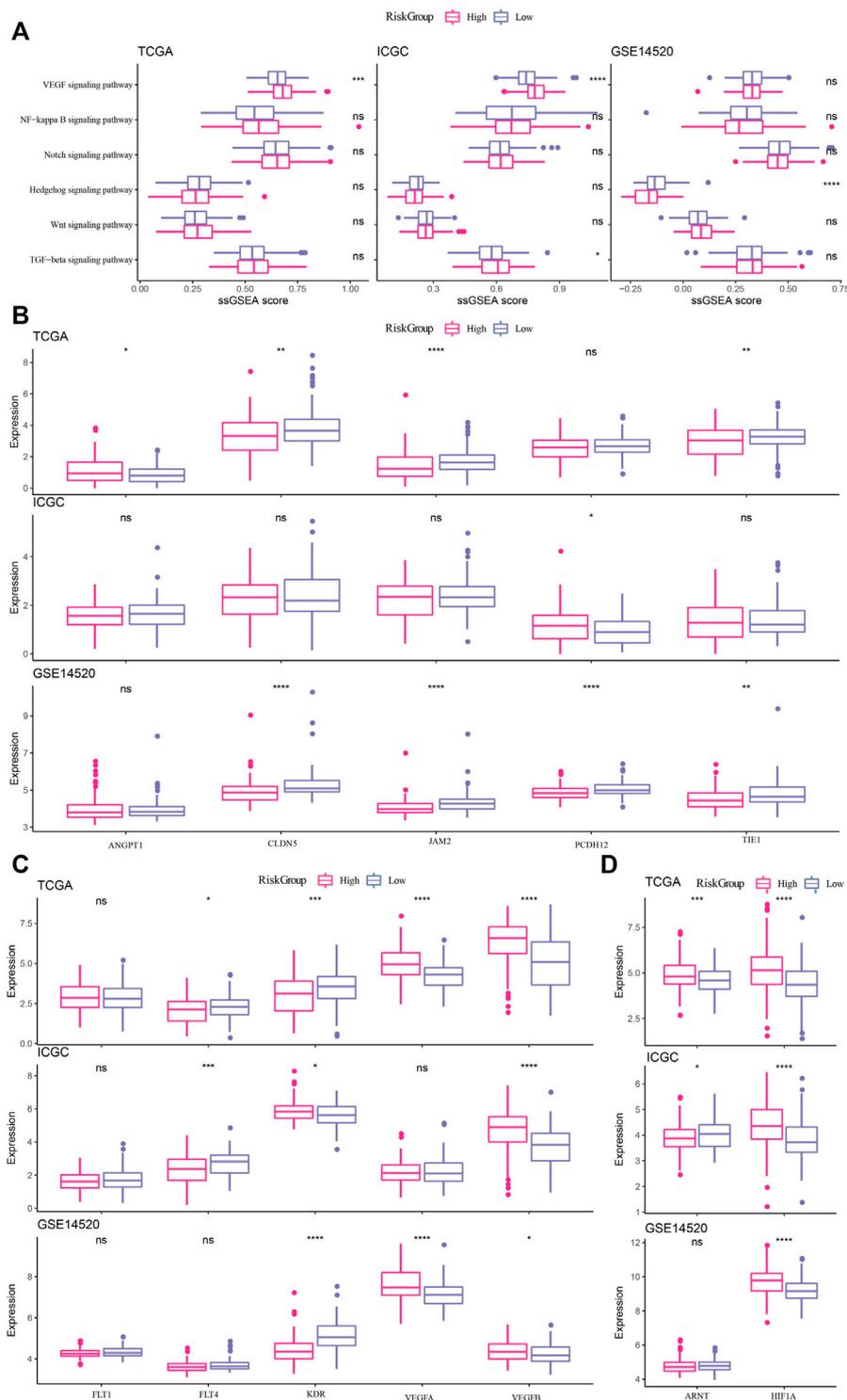


FIGURE 6

Association of risk score with key biological processes occurring in cancer progression (A) Relationship between EMT-related pathways and risk score in TCGA-LIHC, ICGC and GSE14520 datasets. (B) Relationship between vascular stability-related genes and risk score in TCGA-LIHC, ICGC and GSE14520 datasets. (C) Differential expression of VEGF-related genes between high-risk and low-risk samples in the three datasets. (D) Differential expression analysis of Hypoxia-related genes between high-risk and low-risk samples in the three datasets. * and ns denote *p*-value, * < 0.05, ** *p* < 0.001, *** *p* < 0.0001, the ns was no significant difference.

epithelial cells during the evolution from primary tumor to PVTT, one showing a trend toward progressively diminished immune activity and the other showing a trend toward increased

proliferative and metabolic capacity of the cells similar to Heparanase/Syndecan-1 Axis (Yu et al., 2022). Epithelial cells during the progression from primary tumor to metastatic lymph

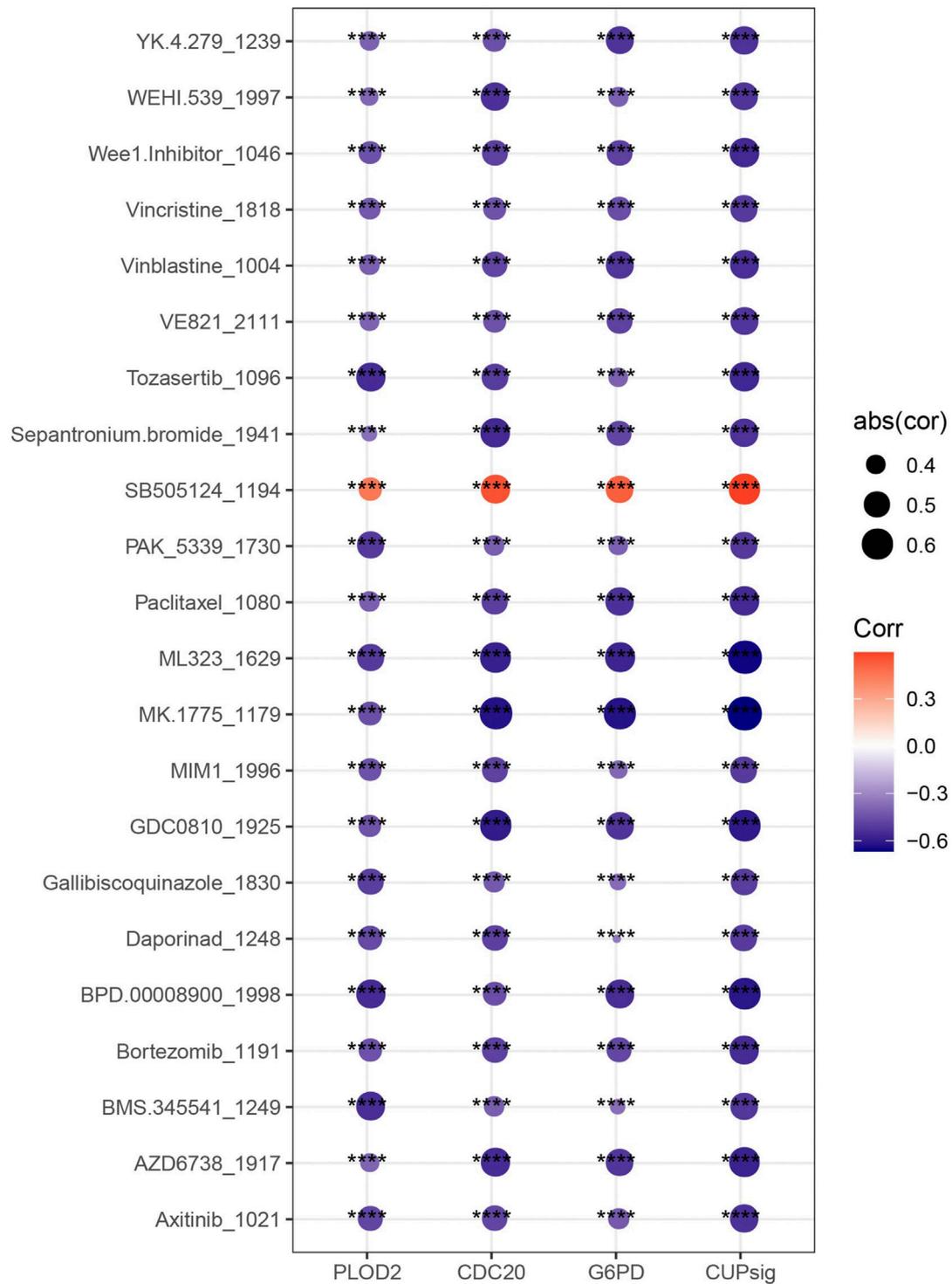


FIGURE 7
Correlation between risk score and IC50 value of drugs.

nod also had two differentiation trajectories, one phenomenon was the gradual improvement of RNA catabolic, protein localization, and protein translation, the other was the attenuation of immune response and cellular response to hypoxia, which reflects the dynamic pathophysiological regulation of Epithelial differentiation in the deterioration of HCC.

Techniques such as single-cell omics in combination with molecular and functional studies will help reveal the remaining unknowns in this field (Gadd et al., 2020). This study linked scRNA-seq analysis with the molecular and functional studies of bulk RNA-seq to explore more effects of epithelial cells on HCC using an algorithm called Scissor, which has not been widely used in combined scRNA-seq and bulk RNA-seq studies. We

found the prognostic candidate genes in each bulk RNA-seq dataset, and finally screened 3 upregulated prognostic candidate genes to develop the risk assessment system. Cell division cycle 20 homologue (CDC20) is a cell cycle regulator that controls the correct segregation of chromosomes during mitosis (Bruno et al., 2022). CDC2 has oncogenic properties and also regulates anticancer drug responses, and is considered an emerging target for cancer therapeutic intervention (Jeong et al., 2022; Wavelet-Vermuse et al., 2022). Additionally, CDC20 was reported to be related to immune infiltration in cancer. For example, a positive relationship of CDC20 with the infiltration of CD8⁺ T cells, CD4⁺ T cells, as well as natural killer cells in HCC was observed (Lai et al., 2021; Xiong et al., 2021). These findings indicated that CDC20 could be a promising target for HCC therapy in terms of cell cycle inhibition or studying tumor immune vaccine (Wang et al., 2022). Glucose-6-phosphate dehydrogenase (G6PD) is a rate-limiting enzyme in pentose phosphate pathway (PPP), which induces EMT by activating signal transducer and activator of transcription 3 (STAT3) pathway, thus promoting the metastasis of HCC cells (Lu et al., 2018). G6PD also has the properties of vascular regulation and participates in the metabolic adaptation of hypoxia (Baptista et al., 2022). More importantly, G6PD related inhibitors are being studied for cancer treatment. We believed that this newly found gene in HCC may also expand the treating method for HCC. PLOD2 expression was identified as a significant, independent factor of poor prognosis by the study of Noda et al. (2012). PLOD2 is induced by hypoxia and affects chemotherapy resistance in biliary tract cancer patients through EMT (Okumura et al., 2018). Moreover, several pharmacological inhibitors of PLOD2 have been proved to have anti-metastatic effects (Du et al., 2017). Collectively, these three genes are prospective drug research targets for HCC treatment. In the future, the expression levels of these three genes on mRNA and protein should be validated first combined with exploring their functions on cell cycle, metastasis or metabolism under hypoxia environment. In the present study, the risk assessment system considering these 3 genes simultaneously also did show significant correlations with VEGFA and VEGFB, some vascular stability-related genes, and hypoxia-related genes. The risk assessment system also coordinates the treatment response of the 22 drugs, which may also aid in the screening of patient treatments.

Conclusion

In conclusion, this study used the analysis strategy of scRNA-seq and bulk RNA-seq data to reveal the cellular composition of HCC, differentiation, evolution and functional landscape of epithelial cells

References

- Ahn, J. C., Teng, P. C., Chen, P. J., Posadas, E., Tseng, H. R., Lu, S. C., et al. (2021). Detection of circulating tumor cells and their implications as a biomarker for diagnosis, prognostication, and therapeutic monitoring in hepatocellular carcinoma. *Hepatology* 73 (1), 422–436. doi:10.1002/hep.31165
- Alqurashi, Y. E., Al-Hetty, H., Ramaiah, P., Fazaa, A. H., Jalil, A. T., Alsaikhan, F., et al. (2023). Harnessing function of EMT in hepatocellular carcinoma: from biological view to nanotechnological standpoint. *Environ. Res.* 227, 115683. doi:10.1016/j.envres.2023.115683
- Baptista, I., Karakitsou, E., Cazier, J. B., Gunther, U. L., Marin, S., and Cascante, M. (2022). TKTL1 knockdown impairs hypoxia-induced glucose-6-phosphate dehydrogenase and glyceraldehyde-3-phosphate dehydrogenase overexpression. *Int. J. Mol. Sci.* 23 (7), 3574. doi:10.3390/ijms23073574
- Bruno, S., Luserna di Rora, A., Napolitano, R., Soverini, S., Martinelli, G., and Simonetti, G. (2022). CDC20 in and out of mitosis: a prognostic factor and therapeutic target in hematological malignancies. *J. Exp. Clin. Cancer Res.* 41 (1), 159. doi:10.1186/s13046-022-02363-9
- Butler, A., Hoffman, P., Smibert, P., Papalexi, E., and Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* 36 (5), 411–420. doi:10.1038/nbt.4096
- Calderaro, J., Ziol, M., Paradis, V., and Zucman-Rossi, J. (2019). Molecular and histological correlations in liver cancer. *J. Hepatol.* 71 (3), 616–630. doi:10.1016/j.jhep.2019.06.001
- Campana, L., Esser, H., Huch, M., and Forbes, S. (2021). Liver regeneration and inflammation: from fundamental science to clinical applications. *Nat. Rev. Mol. Cell Biol.* 22 (9), 608–624. doi:10.1038/s41580-021-00373-7

in HCC, provide a prognostic model of HCC, and give our insights in the detection of drug therapy for patients.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

Author contributions

WQ: Conceptualization, Data curation, Funding acquisition, Investigation, Project administration, Resources, Supervision, Validation, Writing—original draft. QZ: Data curation, Formal Analysis, Funding acquisition, Methodology, Project administration, Software, Writing—original draft, Writing—review and editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This study was supported by the Jilin Provincial Department of Finance (SY2021SCZ2) and Jilin Provincial Department of Science and Technology Project (20230203185).

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.

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.

- Du, H., Pang, M., Hou, X., Yuan, S., and Sun, L. (2017). PLOD2 in cancer research. *Bioméd. Pharmacother.* 90, 670–676. doi:10.1016/j.biopha.2017.04.023
- Florentin, J., O'Neil, S. P., Ohayon, L. L., Uddin, A., Vasamsetti, S. B., Arunkumar, A., et al. (2022). VEGF receptor 1 promotes hypoxia-induced hematopoietic progenitor proliferation and differentiation. *Front. Immunol.* 13, 882484. doi:10.3389/fimmu.2022.882484
- Gadd, V. L., Aleksieva, N., and Forbes, S. J. (2020). Epithelial plasticity during liver injury and regeneration. *Cell Stem Cell* 27 (4), 557–573. doi:10.1016/j.stem.2020.08.016
- Gerlach, J. C., Foka, H. G., Thompson, R. L., Gridelli, B., and Schmelzer, E. (2018). Epithelial cell adhesion molecule fragments and signaling in primary human liver cells. *J. Cell Physiol.* 233 (6), 4841–4851. doi:10.1002/jcp.26286
- Gonzalez, R., De la Rosa, A. J., Ruffini, A., Rodriguez-Hernandez, M. A., Navarro-Villaran, E., Marchal, T., et al. (2017). Role of p63 and p73 isoforms on the cell death in patients with hepatocellular carcinoma submitted to orthotopic liver transplantation. *PLoS One* 12 (3), e0174326. doi:10.1371/journal.pone.0174326
- Ho, D. W., Tsui, Y. M., Sze, K. M., Chan, L. K., Cheung, T. T., Lee, E., et al. (2019). Single-cell transcriptomics reveals the landscape of intra-tumoral heterogeneity and stemness-related subpopulations in liver cancer. *Cancer Lett.* 459, 176–185. doi:10.1016/j.canlet.2019.06.002
- Holzcbauer, A., Wangenstein, K. J., and Shin, S. (2022). Cellular origins of regenerating liver and hepatocellular carcinoma. *JHEP Rep.* 4 (4), 100416. doi:10.1016/j.jhepr.2021.100416
- Huang, D. Q., Singal, A. G., Kanwal, F., Lampertico, P., Buti, M., Sirlin, C. B., et al. (2023). Hepatocellular carcinoma surveillance - utilization, barriers and the impact of changing aetiology. *Nat. Rev. Gastroenterol. Hepatol.* doi:10.1038/s41575-023-00818-8
- Jayachandran, A., Dhungel, B., and Steel, J. C. (2016). Epithelial-to-mesenchymal plasticity of cancer stem cells: therapeutic targets in hepatocellular carcinoma. *J. Hematol. Oncol.* 9 (1), 74. doi:10.1186/s13045-016-0307-9
- Jeong, S. M., Bui, Q. T., Kwak, M., Lee, J. Y., and Lee, P. C. (2022). Targeting Cdc20 for cancer therapy. *Biochim. Biophys. Acta Rev. Cancer* 1877 (6), 188824. doi:10.1016/j.bbcan.2022.188824
- Johnson, R. L. (2019). Hippo signaling and epithelial cell plasticity in mammalian liver development, homeostasis, injury and disease. *Sci. China Life Sci.* 62 (12), 1609–1616. doi:10.1007/s11427-018-9510-3
- Kearnes, S. M., Haque, I. S., and Pande, V. S. (2014). SCISSORS: practical considerations. *J. Chem. Inf. Model* 54 (1), 5–15. doi:10.1021/ci400264f
- Ko, S., Russell, J. O., Molina, L. M., and Monga, S. P. (2020). Liver progenitors and adult cell plasticity in hepatic injury and repair: knowns and unknowns. *Annu. Rev. Pathol.* 15, 23–50. doi:10.1146/annurev-pathmechdis-012419-032824
- Lai, E., Tai, Y., Jiang, J., Zhao, C., Xiao, Y., Quan, X., et al. (2021). Prognostic evaluation and immune infiltration analysis of five bioinformatic selected genes in hepatocellular carcinoma. *J. Cell Mol. Med.* 25 (24), 11128–11141. doi:10.1111/jcmm.17035
- Lu, M., Lu, L., Dong, Q., Yu, G., Chen, J., Qin, L., et al. (2018). Elevated G6PD expression contributes to migration and invasion of hepatocellular carcinoma cells by inducing epithelial-mesenchymal transition. *Acta Biochim. Biophys. Sin. (Shanghai)* 50 (4), 370–380. doi:10.1093/abbs/gmy009
- Ma, L., Khatib, S., Craig, A. J., and Wang, X. W. (2021). Toward a liver cell Atlas: understanding liver biology in health and disease at single-cell resolution. *Semin. Liver Dis.* 41 (3), 321–330. doi:10.1055/s-0041-1729970
- Moon, A. M., Singal, A. G., and Tapper, E. B. (2020). Contemporary epidemiology of chronic liver disease and cirrhosis. *Clin. Gastroenterol. Hepatol.* 18 (12), 2650–2666. doi:10.1016/j.cgh.2019.07.060
- Nakagawa, H., Hikiba, Y., Hirata, Y., Font-Burgada, J., Sakamoto, K., Hayakawa, Y., et al. (2014). Loss of liver E-cadherin induces sclerosing cholangitis and promotes carcinogenesis. *Proc. Natl. Acad. Sci. U. S. A.* 111 (3), 1090–1095. doi:10.1073/pnas.1322731111
- Noda, T., Yamamoto, H., Takemasa, I., Yamada, D., Uemura, M., Wada, H., et al. (2012). PLOD2 induced under hypoxia is a novel prognostic factor for hepatocellular carcinoma after curative resection. *Liver Int.* 32 (1), 110–118. doi:10.1111/j.1478-3231.2011.02619.x
- Ogunnigbagbe, O., Bunick, C. G., and Kaur, K. (2022). Keratin 1 as a cell-surface receptor in cancer. *Biochim. Biophys. Acta Rev. Cancer* 1877 (1), 188664. doi:10.1016/j.bbcan.2021.188664
- Ogunwobi, O. O., Harricharran, T., Huaman, J., Galuza, A., Odumuwaun, O., Tan, Y., et al. (2019). Mechanisms of hepatocellular carcinoma progression. *World J. Gastroenterol.* 25 (19), 2279–2293. doi:10.3748/wjg.v25.i19.2279
- Okumura, Y., Noda, T., Eguchi, H., Sakamoto, T., Iwagami, Y., Yamada, D., et al. (2018). Hypoxia-induced PLOD2 is a key regulator in epithelial-mesenchymal transition and chemoresistance in biliary tract cancer. *Ann. Surg. Oncol.* 25 (12), 3728–3737. doi:10.1245/s10434-018-6670-8
- Park, D. J., Sung, P. S., Kim, J. H., Lee, G. W., Jang, J. W., Jung, E. S., et al. (2020). EpCAM-high liver cancer stem cells resist natural killer cell-mediated cytotoxicity by upregulating CEACAM1. *J. Immunother. Cancer* 8 (1), e000301. doi:10.1136/jitc-2019-000301
- Qiu, X., Mao, Q., Tang, Y., Wang, L., Chawla, R., Pliner, H. A., et al. (2017). Reversed graph embedding resolves complex single-cell trajectories. *Nat. Methods* 14 (10), 979–982. doi:10.1038/nmeth.4402
- Robertson, A. G., Kim, J., Al-Ahmadie, H., Bellmunt, J., Guo, G., Cherniack, A. D., et al. (2017). Comprehensive molecular characterization of muscle-invasive bladder cancer. *Cell.* 171 (3), 540–556.e25. doi:10.1016/j.cell.2017.09.007
- Romualdo, G. R., Heidor, R., Bacil, G. P., Moreno, F. S., and Barbisan, L. F. (2023). Past, present, and future of chemically induced hepatocarcinogenesis rodent models: perspectives concerning classic and new cancer hallmarks. *Life Sci.* 330, 121994. doi:10.1016/j.lfs.2023.121994
- Shu, X., Wang, Q., and Wu, Q. (2022). The eph/ephrin system in hepatocellular carcinoma: functional roles and potential therapeutic targets. *Oncologie* 24 (3), 427–439. doi:10.32604/oncologie.2022.023248
- Srivatsan, S. R., McFaline-Figueroa, J. L., Ramani, V., Saunders, L., Cao, J., Packer, J., et al. (2020). Massively multiplex chemical transcriptomics at single-cell resolution. *Science* 367 (6473), 45–51. doi:10.1126/science.aax6234
- Sun, D., Guan, X., Moran, A. E., Wu, L. Y., Qian, D. Z., Schedin, P., et al. (2022). Identifying phenotype-associated subpopulations by integrating bulk and single-cell sequencing data. *Nat. Biotechnol.* 40 (4), 527–538. doi:10.1038/s41587-021-01091-3
- Wang, D., Chen, H., and Hu, Y. (2022). Polarized autologous macrophages (PAM) can be a tumor vaccine. *Oncologie* 24 (3), 441–449. doi:10.32604/oncologie.2022.024898
- Wavelet-Vermuse, C., Odnokoz, O., Xue, Y., Lu, X., Cristofanilli, M., and Wan, Y. (2022). CDC20-Mediated hnRNPU ubiquitination regulates chromatin condensation and anti-cancer drug response. *Cancers (Basel)* 14 (15), 3732. doi:10.3390/cancers14153732
- Xia, J., Xu, X., Huang, P., He, M., and Wang, X. (2015). The potential of CXCL5 as a target for liver cancer - what do we know so far? *Expert Opin. Ther. Targets* 19 (2), 141–146. doi:10.1517/14728222.2014.993317
- Xiong, C., Wang, Z., Wang, G., Zhang, C., Jin, S., Jiang, G., et al. (2021). Identification of CDC20 as an immune infiltration-correlated prognostic biomarker in hepatocellular carcinoma. *Invest. New Drugs* 39 (5), 1439–1453. doi:10.1007/s10637-021-01126-1
- Yan, D., Li, C., Zhou, Y., Yan, X., Zhi, W., Qian, H., et al. (2022). Exploration of combinational therapeutic strategies for HCC based on TCGA HCC database. *Oncologie* 24 (1), 101–111. doi:10.32604/oncologie.2022.020357
- Yao, C., Wu, S., Kong, J., Sun, Y., Bai, Y., Zhu, R., et al. (2023). Angiogenesis in hepatocellular carcinoma: mechanisms and anti-angiogenic therapies. *Cancer Biol. Med.* 20 (1), 25–43. doi:10.20892/j.issn.2095-3941.2022.0449
- Yu, S., Lv, H., Zhang, J., Zhang, H., Ju, W., Jiang, Y., et al. (2022). Heparanase/Syndecan-1 Axis regulates the grade of liver cancer and proliferative ability of hepatocellular carcinoma cells. *Oncologie* 24 (3), 539–551. doi:10.32604/oncologie.2022.024882
- Zhou, S. L., Zhou, Z. J., Hu, Z. Q., Li, X., Huang, X. W., Wang, Z., et al. (2015). CXCR2/CXCL5 axis contributes to epithelial-mesenchymal transition of HCC cells through activating PI3K/Akt/GSK-3 β /Snail signaling. *Cancer Lett.* 358 (2), 124–135. doi:10.1016/j.canlet.2014.11.044