



## OPEN ACCESS

## EDITED BY

Fu Wang,  
Xi'an Jiaotong University, China

## REVIEWED BY

Yang Xiao,  
Southeast University, China  
Liao Yuanmei,  
Andalusian Center of Molecular Biology  
and Regenerative Medicine (CABIMER),  
Spain

## \*CORRESPONDENCE

Cuifang Han  
✉ hancuifang@gdmu.edu.cn  
Hongbing Yu  
✉ hbyu@gdmu.edu.cn  
Zhiwei He  
✉ hezhiwei@gdmu.edu.cn

<sup>†</sup>These authors have contributed equally to this work

## SPECIALTY SECTION

This article was submitted to  
Molecular and Cellular Oncology,  
a section of the journal  
Frontiers in Oncology

RECEIVED 21 November 2022

ACCEPTED 02 January 2023

PUBLISHED 31 January 2023

## CITATION

Han C, Chen J, Huang J, Zhu R, Zeng J,  
Yu H and He Z (2023) Single-cell  
transcriptome analysis reveals the  
metabolic changes and the prognostic  
value of malignant hepatocyte  
subpopulations and predict new  
therapeutic agents for hepatocellular  
carcinoma.  
*Front. Oncol.* 13:1104262.  
doi: 10.3389/fonc.2023.1104262

## COPYRIGHT

© 2023 Han, Chen, Huang, Zhu, Zeng, Yu  
and He. This is an open-access article  
distributed under the terms of the [Creative  
Commons Attribution License \(CC BY\)](#). The  
use, distribution or reproduction in other  
forums is permitted, provided the original  
author(s) and the copyright owner(s) are  
credited and that the original publication in  
this journal is cited, in accordance with  
accepted academic practice. No use,  
distribution or reproduction is permitted  
which does not comply with these terms.

# Single-cell transcriptome analysis reveals the metabolic changes and the prognostic value of malignant hepatocyte subpopulations and predict new therapeutic agents for hepatocellular carcinoma

Cuifang Han<sup>1\*</sup>, Jiaru Chen<sup>1,2†</sup>, Jing Huang<sup>1</sup>, Riting Zhu<sup>1,2</sup>,  
Jincheng Zeng<sup>3</sup>, Hongbing Yu<sup>1\*</sup> and Zhiwei He<sup>1\*</sup>

<sup>1</sup>Guangdong Provincial Key Laboratory of Medical Molecular Diagnostics, The First Dongguan Affiliated Hospital, Guangdong Medical University, Dongguan, China, <sup>2</sup>School of Pharmacy, Guangdong Medical University, Dongguan, China, <sup>3</sup>Dongguan Key Laboratory of Medical Bioactive Molecular Developmental and Translational Research, Guangdong Provincial Key Laboratory of Medical Molecular Diagnostics, Guangdong Medical University, Dongguan, China

**Background:** The development of HCC is often associated with extensive metabolic disturbances. Single cell RNA sequencing (scRNA-seq) provides a better understanding of cellular behavior in the context of complex tumor microenvironments by analyzing individual cell populations.

**Methods:** The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) data was employed to investigate the metabolic pathways in HCC. Principal component analysis (PCA) and uniform manifold approximation and projection (UMAP) analysis were applied to identify six cell subpopulations, namely, T/NK cells, hepatocytes, macrophages, endothelial cells, fibroblasts, and B cells. The gene set enrichment analysis (GSEA) was performed to explore the existence of pathway heterogeneity across different cell subpopulations. Univariate Cox analysis was used to screen genes differentially related to The Overall Survival in TCGA-LIHC patients based on scRNA-seq and bulk RNA-seq datasets, and LASSO analysis was used to select significant predictors for incorporation into multivariate Cox regression. Connectivity Map (CMap) was applied to analysis drug sensitivity of risk models and targeting of potential compounds in high risk groups.

**Results:** Analysis of TCGA-LIHC survival data revealed the molecular markers associated with HCC prognosis, including MARCKSL1, SPP1, BSG, CCT3, LAGE3, KPNA2, SF3B4, GTPBP4, PON1, CFHR3, and CYP2C9. The RNA expression of 11 prognosis-related differentially expressed genes (DEGs) in normal human hepatocyte cell line MIHA and HCC cell lines HCC-LM3 and HepG2 were

compared by qPCR. Higher KPNA2, LAGE3, SF3B4, CCT3 and GTPBP4 protein expression and lower CYP2C9 and PON1 protein expression in HCC tissues from Gene Expression Profiling Interactive Analysis (GEPIA) and Human Protein Atlas (HPA) databases. The results of target compound screening of risk model showed that mercaptopurine is a potential anti-HCC drug.

**Conclusion:** The prognostic genes associated with glucose and lipid metabolic changes in a hepatocyte subpopulation and comparison of liver malignancy cells to normal liver cells may provide insight into the metabolic characteristics of HCC and the potential prognostic biomarkers of tumor-related genes and contribute to developing new treatment strategies for individuals.

#### KEYWORDS

cancer metabolism, hepatocellular carcinoma, malignant hepatocytes, prognostic biomarker, single-cell RNA sequencing

## 1 Introduction

The mortality rate for liver cancer is the third highest among all cancers, and it is the sixth most frequent cancer overall (1). Hepatocellular carcinoma (HCC) is a tumour of hepatocellular origin. HCC is the predominant pathological type of primary liver cancer (PLC), as it represents 75-85% of all instances of PLC (2). A vast majority of HCCs are caused by chronic disease, and most of these cases reportedly evolve from chronic liver disease. This is primarily because of viral infections, including hepatitis B virus (HBV) and hepatitis C virus (HCV), and alcohol misuse (3). It is recommended that patients diagnosed with HCC in the early stages receive surgical resection, liver transplantation, and local resection (radiofrequency ablation) according to the Barcelona Clinic Liver Cancer (BCLC) staging system. Those in the intermediate stage are widely treated with trans-arterial chemoembolization (TACE), whereas systemic therapies are mainly considered for advanced-stage patients. Advanced-stage patients are often symptomatic, although they exhibit some degree of impaired liver function (4, 5). Notably, few or no treatments are available to improve survival rates for patients in advanced stages.

The development of treatment modalities for advanced HCC has dramatically expanded recently. To date, the FDA has approved several oral tyrosine kinase inhibitors (lenvatinib, regorafenib and cabozantinib), immune checkpoint inhibitors (nivolumab and pembrolizumab) and

immunotherapies, such as monoclonal antibodies (6–8). These therapies have steadily improved the overall survival (OS) of HCC patients. However, the prognosis for HCC patients continues to be poor because of recurrence and elevated metastasis rates (9). HCC features have been attributed to a small subpopulation of tumour cells that carry more aggressive genetic or phenotypic alterations that allow them to escape conventional detection methods (10).

Although conventional bulk RNA sequencing (bulk RNA-seq) can provide sufficient gene expression profiles of large blocks of tissue, it does not effectively distinguish between different cell lineages and cellular interactions (11). Recently, the emergence of single-cell sequencing technology has bridged the gap between traditional high-throughput sequencing technologies and microarray data to provide genomic, transcriptomic, and epigenetic information from individual cells (12). Tumours consist of three major cell types, namely, malignant, immune and stromal cells, whose spatiotemporal interactions constitute a complex ecosystem (13). Unravelling the interactions between these types involves understanding tumour development and prognosis and therapeutic options. Since the advent of single-cell sequencing, various researchers have produced a relatively complete picture of human cell atlas, which has subsequently provided a great reference for understanding the complex composition of the organs of the body (14). Additionally, single-cell sequencing has been extensively employed to reveal the molecular mechanisms underlying HCC. For instance, studies have mapped the single-cell landscape of the early recurrent HCC ecosystem by relying on the high recurrence and low survival rates of HCC patients to advance the immunotherapy guidelines for HCC (13). Numerous studies have utilized single-cell sequencing techniques to elucidate the heterogeneity of malignant tumour cells, stromal cells, and immune cells. The large scale single-cell omics study targeting tumor-associated T cells published by Zhang et al. sketched the tumor immune landscape and laid the groundwork for a multifaceted understanding of T-cell characteristics associated with liver cancer (15). Single-cell technology can also identify rare subpopulations that were previously undetected by bulk RNA sequencing techniques, and these cell types are pivotal in determining tumor characteristics, including stemness-associated malignant cells and cancer-associated fibroblasts (16–18).

**Abbreviations:** HCC, hepatocellular carcinoma; PCA, Principal component analysis; UMAP, Uniform Manifold Approximation and Projection; KEGG, Kyoto Encyclopedia of Genes and Genomes; PON1, paraoxonase 1; TNM, tumor-node-metastasis; GSEA, gene set enrichment analysis; CNV, Copy number variation; DEDs, Differentially Expressed Genes; OS, Overall Survival; LASSO, least absolute shrinkage and selection operator; MARCKSL1, MARCKS Like 1; SPP1, Secreted Phosphoprotein 1; BSG, Basigin, also called CD147 or EMMPRIN; CCT3, chaperonin containing TCP1 subunit 3; LAGE3, L antigen family member 3; KPNA2, karyopherin subunit alpha 2; SF3B4, Splicing Factor 3b Subunit 4; GTPBP4, GTP Binding Protein 4; PON1, Paraoxonase 1; CFHR3, Complement factor H-related 3; CYP2C9, cytochrome P450 family 2 subfamily C member 9.

The reprogramming of energy metabolism characterizes tumour cells and causes rapid cell growth and proliferation. Thus, it is one of the hallmarks of cancer. Tumour cells actively take up glucose through the uncommon process of anaerobic glycolysis (Warburg effect). Studies have shown that this process provides energy to tumour cells, permitting intermediates to enter the anabolic bypass to maintain the *de novo* synthesis of nucleotides, lipids, and amino acids needed for cell proliferation (19). HCC is closely linked to metabolic abnormalities, as the liver is the primary metabolic organ. Most previous studies concerned with liver cancer have focused on sequencing at the tissue level to reveal the overall metabolic alterations. Single-cell sequencing technology can compensate for the shortcomings of bulk sequencing, thereby allowing one to pinpoint the cell groups most significantly associated with metabolic alterations from a large number of cell types. This also allows researchers to comprehensively describe the overall changes in gene expression patterns and reveal changes across specific cell groups. Therefore, scRNA-seq and bulk RNA-seq integration are important techniques for studying tumour development and heterogeneity. We analysed published single-cell transcriptome sequencing data to identify metabolically relevant HCC subpopulations, namely, hepatic epithelial cells. We then used the identified differentially expressed genes to designate a prognostic model for HCC patients.

## 2 Materials and methods

### 2.1 Data collection

The scRNA-seq data for HCC patients were acquired from GEO (<https://www.ncbi.nlm.nih.gov/geo/>, accession number GSE149614) and TCGA (<https://portal.gdc.cancer.gov/>) databases, respectively. TCGA-LIHC samples with complete clinical information were utilized as the model training set, and HCC samples from the GEO database (GSE76427) were utilized as the external validation set.

We first constructed a human liver cell atlas by performing cell classification and marker gene identification relying on Seurat. There were 17 samples in total from 10 HCC patients. These included 8 tumour samples (PT), 8 normal paraneoplastic samples (NTL), and one metastatic lymph node sample (MLN). The data for these samples were obtained from the GSE149614 project.

### 2.2 Identification of HCC cell subtypes

The scRNA-seq data were assessed by the Seurat package implemented in R software (4.1.1), with the exclusion of samples with more than 30% mitochondrial genes. The data were normalized using the Normalize Data function, and 2,000 genes with high intercellular coefficients of variation were subsequently extracted. Principal component analysis (PCA) was then performed, with 15 PCs selected for subsequent uniform manifold approximation and projection (UMAP) analysis. Cell types within the obtained clusters

were annotated by the reported cell marker genes, and the expression matrix was generated for further analysis.

### 2.3 Analysis of intercellular communications

To investigate the potential interactions between tumor and paracancerous normal HCC samples, we employed the CellChat (1.5.0) package to analyse intercellular communication. We performed CellChat analysis of the annotated cellular gene expression profile data according to the official workflow. This package mimics intercellular communication by assessing the binding ligands and receptors along with their cofactors (20). Depending on receptor expression in one cell type and ligand expression in the other, enriched receptor–ligand interactions between the two cell types were inferred. Signaling pathways were visualized using the “netVisual\_aggregate” function, where ligands were defined as efferent signals and receptors were defined as afferent signals.

### 2.4 Identification of important metabolic pathways at the single-cell level

Next, we employed the ‘scMetabolism’ package (0.2.1) to calculate the metabolic state between different cell types in the HCC dataset. This package combines published gene sets from the Kyoto Encyclopedia of Genes and Genomes (KEGG) database and the Reactome database to easily quantify single-cell metabolic activity. (21). Here, we used the authors’ integrated list of metabolism-related gene sets from the Reactome database to explore metabolic pathway changes among six cell subpopulations and further looked at metabolic changes in epithelial cell subpopulations between tumor and paracancerous normal HCC samples.

### 2.5 Copy number variation analysis

To identify malignant cells in HCC patients, we compared patterns of chromosomal gene expression across cancer cells to those of their putative noncancerous counterparts using the infercnv package (version 1.12.0). First, we downloaded the human genome annotation file from the gencode database (<https://www.gencodegenes.org/human/>), converting it into a genomic location file. We used paracancerous epithelial tissue expression profiles from HCC patients as a reference group. Because our data were 10x scRNA-seq data, we set 0.1 as the cut-off value, and the denoise = T. Referring to the two indicators used by Itay Tirosh et al. to determine benign versus malignant cells, here we used the overall copy number variant (CNV) and the correlation with the average CNV of the top 5% of cells from the same tumor to estimate the malignancy or non-malignancy of the cells (22). The following correlation reference thresholds for determining the malignancy or not of cells were given: malignant cells: overall CNV > 0.2 & CNV correlation of the top 5% of tumors > 0.2; non-malignant cells: overall CNV < 0.2 & CNV correlation of the top 5% of tumors < 0.2.

## 2.6 Identification of significantly related pathways across different epithelial cell types

After scoring individual cells using a variety of enrichment methods, we derived multiple gene set enrichment score matrixes using the 'irGSEA' package (<https://github.com/chuiqin/irGSEA/>). Next, we calculated the differentially expressed gene sets for every single cell subpopulation within the enrichment score matrix for every gene set using the Wilcoxon test. Employing heat maps, certain specific enrichment pathways were labelled and then visualized.

## 2.7 Generation and validation of prognostic features

Univariate Cox analysis was used to screen genes associated with OS in TCGA-LIHC patients based on scRNA-seq and bulk RNA-seq datasets, and then, LASSO analysis was used to select significant predictors for incorporation into multivariate Cox regression. Next, we selected and used prognostic characteristics to generate polygenic risk scores and stratify TCGA-LIHC samples into either low- or high-risk groups. We also generated time-dependent receiver operating characteristic (ROC) curves to assess the predictive power of the prognostic features. The GSE76427 dataset was used to validate the prognostic value of the prognostic features. The entire analysis and visualization processes were performed by the survival, survminer, rms, and time ROC packages in R.

## 2.8 Gene expression of prognostic genes

Total RNA from cells was extracted with TRIzol reagent (Thermo Fisher Scientific, 15596026) following the manufacturer's instructions. Complementary DNA (cDNA) was synthesized and PCRs with cDNA as template were performed using a real-time detector (The Applied Biosystems QuantStudio 5 Real-Time PCR System) using Hieff qPCR SYBR Green Master Mix. The primer sequences are shown in [Supplementary Table S1](#). Transcript levels were normalized against beta-actin levels as an internal reference and were evaluated using the  $2^{-\Delta\Delta Ct}$  method. All experiments were repeated three times.

The Human Protein Atlas (HPA) tool was used to visually display the protein expression of prognostic genes in the form of immunohistochemical staining. The Gene Expression Profiling Interaction Analysis (GEPIA) database was applied to further demonstrate the credibility of the results.

## 2.9 Cell culture and western blot

An immortalized nontumorigenic normal human hepatocyte cell line MIHA and HCC cell lines HCC-LM3 and HepG2 were purchased from the Fenghui Biotech Co., Ltd. (Hunan, China) with STR report.

The MIHA cells were cultured in RPMI-1640 and HCC-LM3 and HepG2 were cultured in Dulbecco's modified Eagle medium (Gibco, Gaithersburg, MD, USA) with 10% fetal bovine serum (FBS, Sigma), 100 µg/mL penicillin and 100 µg/mL streptomycin (Solarbio, Shanghai, China) at 37°C and 5% CO<sub>2</sub>.

Total protein was extracted by using Takara kit. The Protein concentration was detected by BCA assay. The primary antibodies used in this study were anti-CYP2C9 (1:1000, Abcam), anti-PON1 (1:1000, Abcam) anti-beta-Actin (1:1000, Cell Signaling Technology).

## 2.10 Drug sensitivity analysis

Genomics of Drug Sensitivity in Cancer (GDSC, <https://www.cancerrxgene.org/>) is the largest pharmacogenomic database that is freely accessible for predicting responses to anticancer drugs. GDSC comprises 2 databases, namely, GDSC1, which contains 958 cell lines and 367 drugs, and GDSC2, which contains 805 cell lines and 198 drugs (23). To explore the differences in drug treatment effects among HCC patients, drug inferred sensitivity scores were assessed in GDSC2 by the 'oncoPredict' package.

## 2.11 Connective map analysis

The Connective Map (CMap) database stores a large-scale resource of expression profile data of cell lines under different drug treatments, which allows rapid targeting of drug candidates for the treatment of target diseases based on aberrant transcriptomic features in tumor cells (24). These drugs have an inverse relationship with tumor-promoting factors and may regulate aberrantly expressed genes in the opposite direction.

Recently, Yang et al. used the Library of Integrated Network-based Cellular Signatures (LINCS) database to demonstrate that using the eXtreme Sum (XSum) algorithm is most likely to yield optimal results in matching compounds and disease features, demonstrating better drug retrieval performance than the other five available methods, and obtaining practical targets with desirable results in liver cancer (25). In addition, the parameters for achieving the best prediction performance in this study were set at a number of disease molecular features of 100. Considering the significant difference in dimensionality between CMap data and LINCS, we incorporate more query signatures using top300 genes for XSum analysis for potential drug prediction.

## 2.12 Statistical analysis

All statistical analyses were carried out using packages implemented in R version 4.2.0 (<https://www.r-project.org/>). Student's t test was used to perform comparisons of continuous variables between two groups, and the Wilcoxon rank sum test was used to compare more than two groups. Kaplan–Meier curves with log-rank statistics were used to compare differences in OS between the two groups. Statistical significance was represented by  $p < 0.05$ .

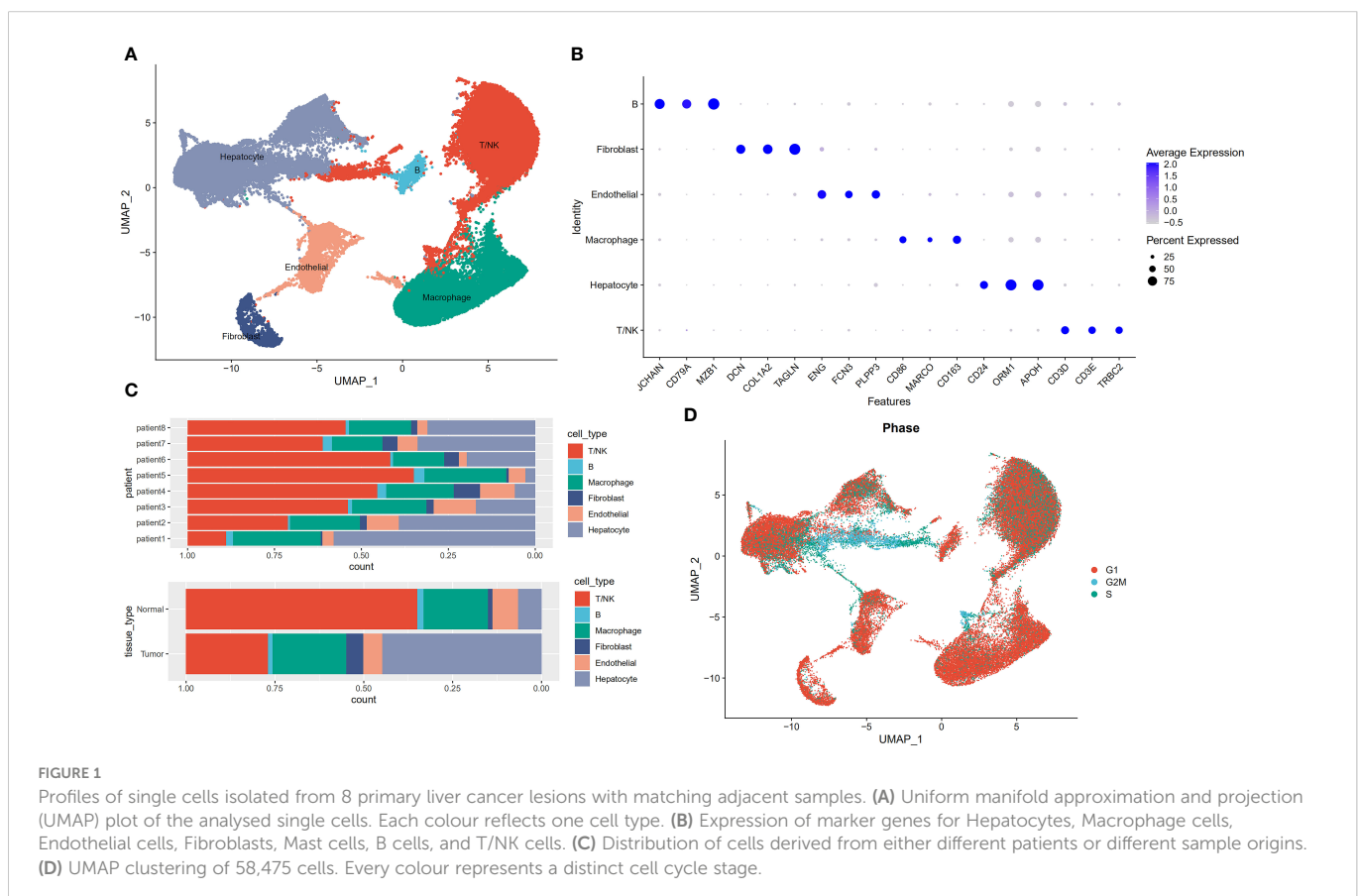
## 3 Results

### 3.1 Single-cell gene expression profiles reveal six major cell types in the TME of primary HCC tumours

We performed descending and unsupervised cell clustering to recognize cell types based on their expression profiles. The raw dataset was read using the Seurat package. Then, an initial screening of genes and cells was performed using the following criteria: a gene had to be expressed in at least 3 cells, and at least 200 genes were measured in this cell. This was followed by further quality control to extract cells with >200 and <8000 expressed genes and <30% of mitochondrial genes. Next, the data were normalized to obtain 2000 highly variable genes for subsequent downscaling. Removal of the cell cycle effect resulted in an expression matrix comprising 58,475 cells and 24,746 genes. Next, we employed known marker genes to define broad cell categories and obtained the following six major cell subpopulations: T/NK cells, hepatocytes, macrophages, endothelial cells, fibroblasts, and B cells (Figures 1A, B). Cells from tumours and normal paracancerous tissues from different patients were classified into six categories (Figure 1C). Because proliferation is a hallmark of tumour cells, we employed the cell cycle scoring method to analyse the cell cycle. This image shows the results indicated that most of the cells were in the G1 phase, and a small number of cells were in the G2/M and S phases (Figure 1D).

### 3.2 Genes associated with the glucose and lipid metabolic pathway are upregulated in hepatocytes

To explore the existence of pathway heterogeneity across different cell subpopulations, we performed pathway activity and GSEA using signature genomes. Numerous pathways associated with cancer were upregulated in the hepatocyte subpopulation; these pathways included oxidative phosphorylation, glycolysis, and the metabolism of fatty acids, bile acids, and xenobiotics (Figure 2A). Next, we used the scMetabolism package to calculate scores for each metabolic pathway in each cell. We found that the epithelial cell subpopulation was enriched in most metabolic pathways, mainly those regulating pyruvate metabolism, the citrate tricarboxylic acid cycle, and the metabolism of triglycerides, pyruvates, lipids, carbohydrates, amino acids and their derivatives, ketone bodies, glucose, and fatty acids, and FoxO-mediated oxidative stress (Figure 2B). The genes of glucose metabolism and lipid metabolism pathways were also upregulated in epithelial cells (Figures 2C, D). To determine the differences in metabolic pathways of hepatic epithelial cells between tumor and paracancerous tissues, we extracted a separate subpopulation of hepatocytes and analysed the enrichment of metabolic pathways. Strikingly, we found an opposite trend between the glucose metabolism and lipid metabolism pathways in tumour and paracancerous cells (Figure 2E). Consequently, we subjected this cell subpopulation to more in-depth analysis.





macrophages, and fibroblasts in tumor samples, but not in normal samples adjacent to cancer, further supporting the critical role of SPP1 in the tumor ecosystem.

### 3.4 Transcriptome heterogeneity of hepatocytes in HCC

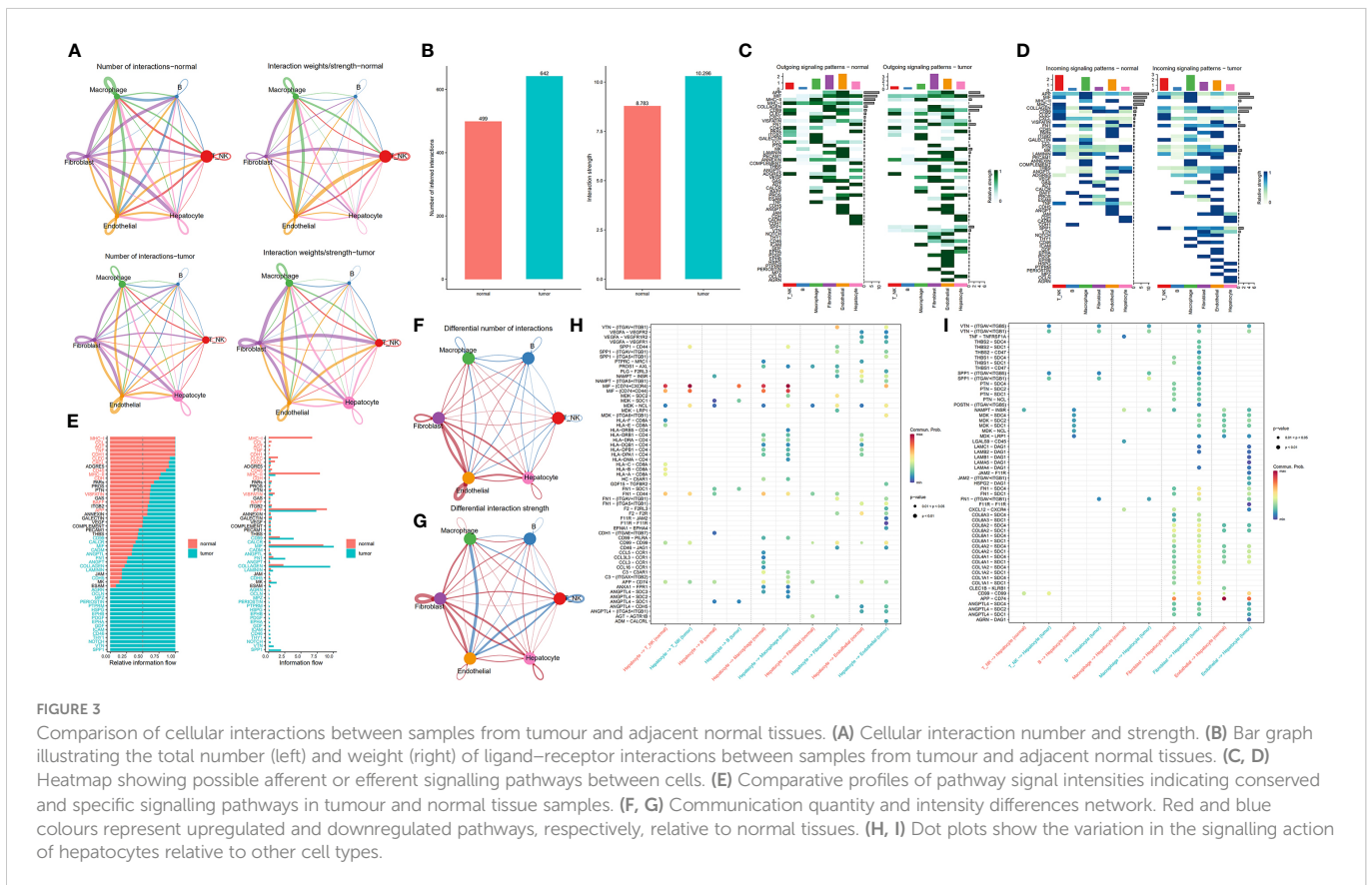
Despite previous batch effects, tumour cells continued to show patient-specific expression patterns. This suggests a high degree of heterogeneity, which could possibly be caused by CNVs. Six major cell subpopulations were identified after the entire malignant and normal hepatocytes reclustering (Figure 4A). In addition, UMAP plots revealed distinct clusters of malignant cells that corresponded to the sample origin (Figure 4B). Figure 4C illustrates the marker genes for each cell subpopulation. Next, the irGSEA package was employed to perform scRNA-seq gene set enrichment analysis and found that these subpopulations have unique activation signals. These signals include the Hedgehog signalling pathway (subpopulation 0), the early oestrogen response (subpopulation 1), the IL6/STAT3 and TNF signalling pathways (subpopulation 2), the xenobiotic metabolism and reactive oxygen species signalling pathways (subpopulation 3), and the KRAS signalling pathway (subpopulation 4). Moreover, multiple cell proliferation-related pathways were upregulated in subpopulation 5; these pathways included those involving the MYC targets V1 and V2, G2M checkpoints, E2F targets, WNT signalling, and P53 targets (Figure 4D). Activated KRAS is a major driver of cancer stem cell (CSC) proliferation and tumour metastasis (31). The results of the present study revealed that the KRAS signalling pathway

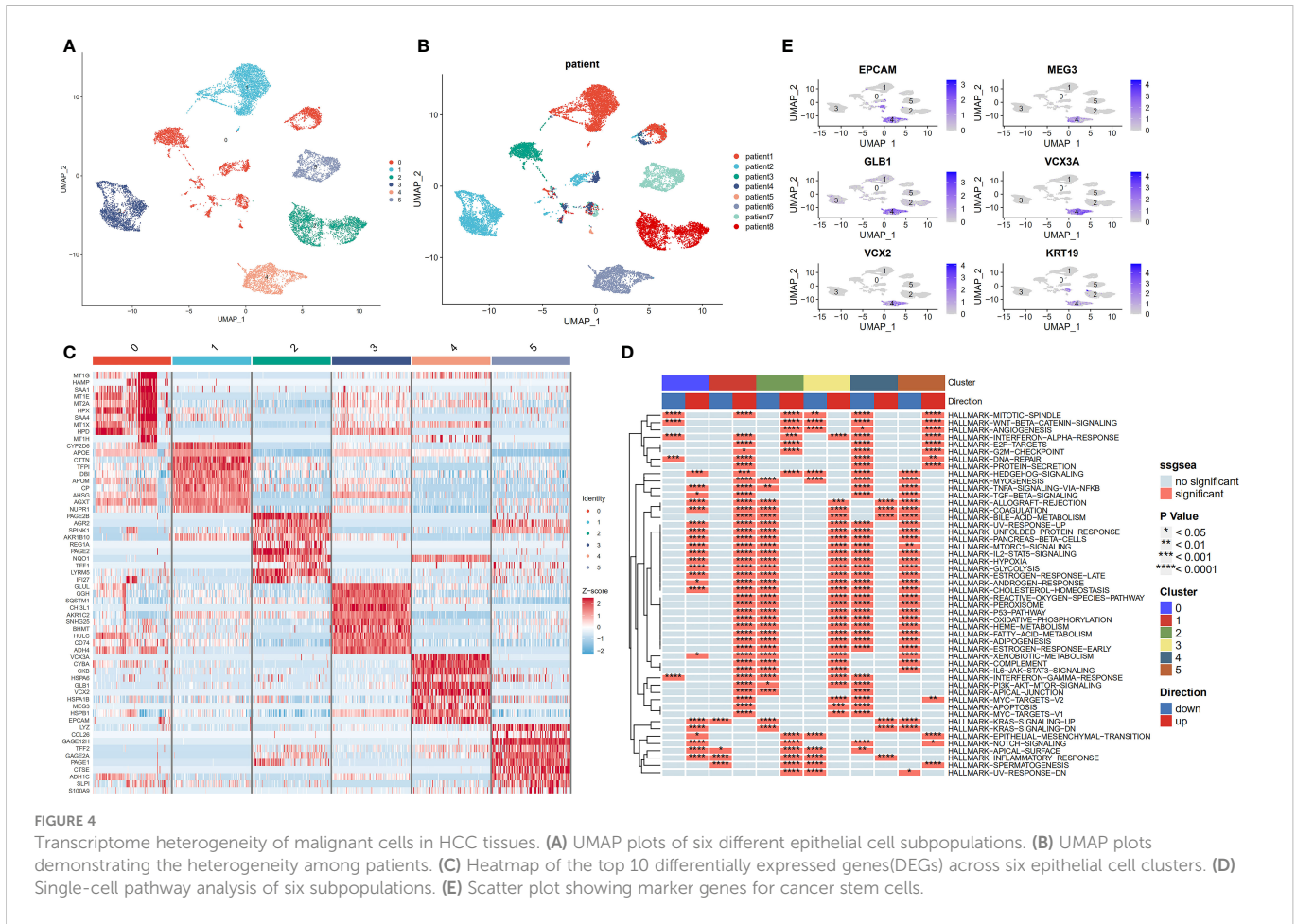
was significantly upregulated in subpopulation 4, and the marker genes for CSCs were also distributed in this subpopulation (Figure 4E).

### 3.5 Profiles of chromosomal CNV in hepatocytes subpopulations

Next, we determined the chromosomal CNVs in each sample based on transcriptomic data to understand the malignancy of the epithelial cell subpopulation. This image shows the results revealed low and high CNV in adjacent normal epithelial cell subpopulations (control samples) and tumour epithelial cells, respectively. Chromosome amplification primarily occurs within chromosomes 1, 3, 5, 6, 7, 8, 12, 15, 17, 20, 21, and 22, with deletions most commonly observed in chromosomes 4, 9, 10, 11, 13, 14, 16 and 18 (Figure 5A). First, the copy value (CNV value) was calculated based on the sum of squares for all genes in each sample. Next, we ranked the CNV values of the tumour cells, using the top 5% as a reference, and then calculated correlation coefficients between other epithelial and reference cells. The determination of tumour cells was achieved at a threshold CNV >0.2 and a correlation coefficient >0.2. With CNV value as the horizontal coordinate and correlation coefficient as the vertical coordinate, black dots represent tumor cells and blue dots represent normal cells (Figure 5B). Finally, 13,502 tumor cells and 1,718 normal cells were identified and projected on the UMAP map (Figure 5C).

Thereafter, we employed the FindAllMarkers function and set the screening conditions logfc = 0.25 (difference multiplicity), min. pct =





0.25 (minimum differential gene expression ratio) and pct. diff >0.1 (pct.1-pct.2) to identify marker genes in the hepatic malignant and normal epithelial cell subsets. The results revealed a total of 564 marker genes (Supplementary Table S3). We hypothesize that their function in HCC differs from that in normal epithelial cells, although further research exploration is needed.

### 3.6 Malignant hepatocyte subpopulations are associated with HCC prognosis

Next, we explored the prognostic role of hepatocyte subpopulations in HCC patients. Analysis of the mRNA expression data from HCC samples across the TCGA database yielded 2,900 differentially expressed genes (Figure 6A). Marker genes from malignant and nonmalignant cells of hepatocyte subpopulations intersected with DEGs related to HCC development in the TCGA database. Notably, 2,900 DEGs overlapped with 564 marker genes, resulting in 203 differentially expressed marker genes in HCC. These were subsequently named hepatocyte differential genes (HDGs) (Figure 6B). Univariate Cox regression analysis revealed 101 differentially expressed marker genes that were significantly related to the prognosis of HCC patients. To obtain a more robust prognostic profile, we employed the LASSO regression algorithm at 10-fold cross-validation with a lambda-min of 0.06321515 to designate a prognostic model consisting of 11 genes, namely, MARCKSL1

(MARCKS Like 1), SPP1 (Secreted Phosphoprotein 1), BSG (Basigin, also called CD147 or EMMPRIN), CCT3 (chaperonin containing TCP1 subunit 3), LAGE3 (L antigen family member 3), KPNA2 (karyopherin subunit alpha 2), SF3B4 (Splicing Factor 3b Subunit 4), GTPBP4 (GTP Binding Protein 4), PON1 (Paraoxonase 1), CFHR3 (Complement factor H-related 3) and CYP2C9 (cytochrome P450 family 2 subfamily C member 9) (Figures 6C, D).

Next, the median risk score was used to stratify the patients into high- and low-risk groups. Patients in the low-risk group showed significantly higher OS rates than their counterparts in the high-risk group ( $p < 0.001$ ) (Figure 6E). Application of the 11-gene signature in the validation cohort also indicated that patients in the low-risk group had longer OS rates than their counterparts in the high-risk group ( $p < 0.001$ ) (Figure 6F). To test the prognostic performance of the 11-gene signature, time-dependent ROC curves were generated targeting TCGA-LIHC samples. The results revealed area under the curve (AUC) values of 0.8, 0.7, and 0.7 for 1-, 3- and 5-year survival, respectively, in the testing cohort (Figure 6G) and 0.8, 0.8, and 0.87, respectively, in the validation cohort (GSE76427) (Figure 6H). These findings suggest that the 11-gene signature had good prognostic value in both cohorts. For the association analysis between the clinicopathological characteristics and the prognostic model, we analysed gender, TNM, stage, and risk scores in the TCGA-LIHC sample. The multivariate Cox regression analysis results revealed that the risk score was a significant independent prognostic factor for patients with LIHC ( $p < 0.001$ ) (Figure 6I). Moreover, we generated a



nomogram encompassing gender, stage, age, grade, risk score and 1-, 2- and 3-year survival. Next, we employed a one-sample GSEA approach to calculate scores for each sample across 175 pathways based on the risk score to identify relevant regulatory pathways. Thereafter, the correlation between each pathway and the risk score was no less than 0.3 for the evaluation. The results revealed 39 positive and 50 negative correlations with the sample risk score. Pathways that were positively correlated with the risk score included those related to cancer development, whereas the negatively correlated pathways included those regulating glycolysis/glycogenesis, glycine, and the metabolism of fatty acids, serine, threonine, glyoxylate and dicarboxylate (Figures 6K, L).

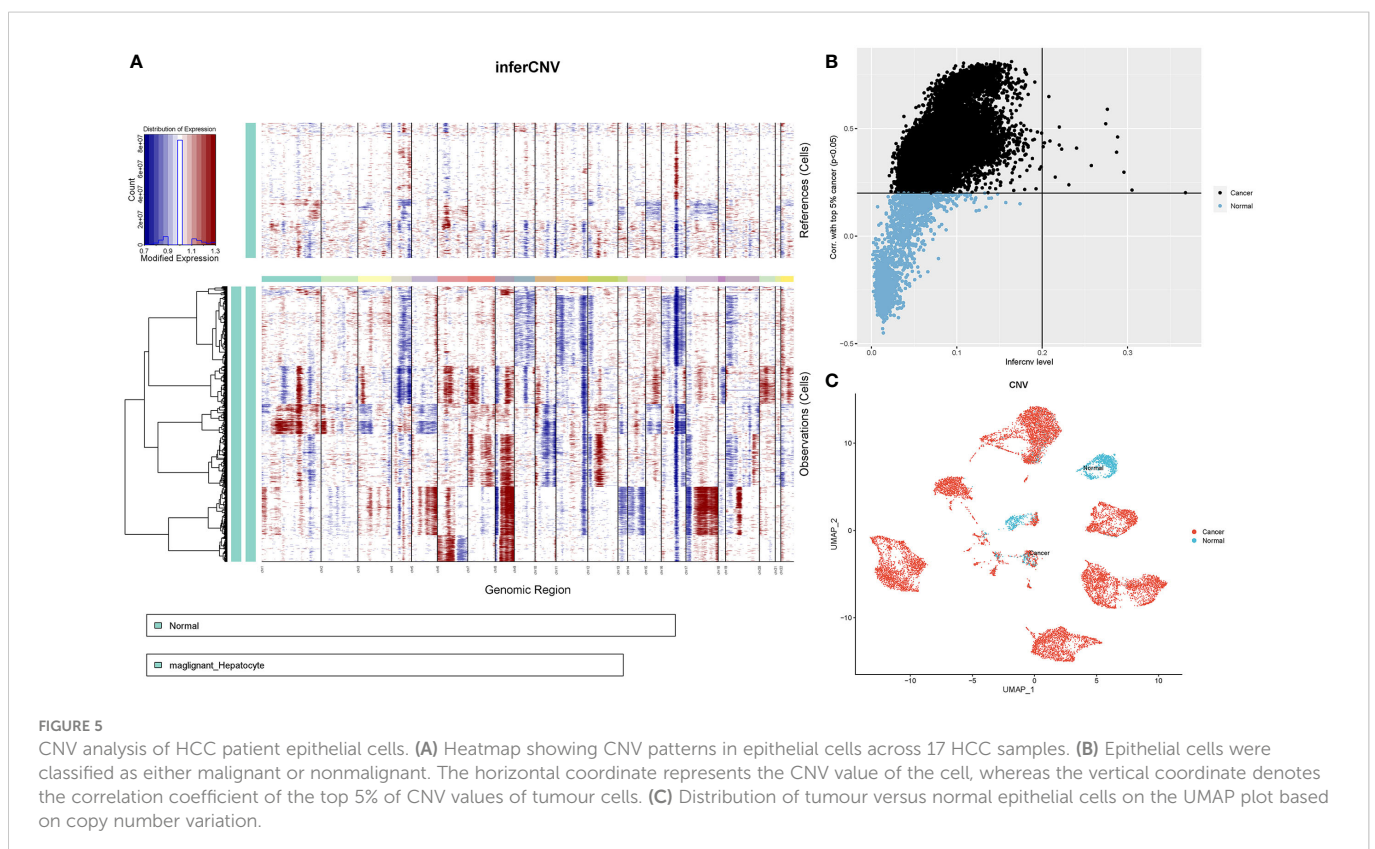
### 3.7 The relative RNA expression level and protein expression level of MARCKSL1, SPP1, BSG, CCT3, LAGE3, KPNA2, SF3B4, GTPBP4, PON1, CFHR3 and CYP2C9

Based on the initial trend of differentially up- and down-regulated genes (Supplementary Table S4), To further investigate the gene expression characteristics of 11 prognosis-related differentially expressed genes (MARCKSL1, SPP1, BSG, CCT3, LAGE3, KPNA2, SF3B4, GTPBP4, PON1, CFHR3 and CYP2C9) in the high-risk and low-risk groups of HCC patients, we performed a correlation analysis between gene expression levels and risk scores. The results showed that all eight genes were positively correlated with risk scores, except for PON1, CFHR3 and CYP2C9, whose mRNA expression levels were

significantly negatively correlated with risk scores (Supplementary Figure S3A). Meanwhile, in order to classify the high and low risk genes, we could see from the forest plot of 11 prognostic genes that the hazard ratio of MARCKSL1, SPP1, BSG, CCT3, LAGE3, KPNA2, SF3B4 and GTPBP4 were all greater than 1, suggesting that these 8 genes might be poor prognostic factors and belong to high risk genes, while PON1, CFHR3 and CYP2C9 were all less than 1, suggesting that these three genes may be factors with a better prognosis (Supplementary Figure S3B).

The RNA expression of MARCKSL1, SPP1, BSG, CCT3, LAGE3, KPNA2, SF3B4, GTPBP4, PON1, CFHR3 and CYP2C9 in normal human hepatocyte cell line MIHA and HCC cell lines HCC-LM3 (high metastatic HCC cells) and HepG2 (low metastatic HCC cells) were compared by qPCR. It was found that CYP2C9, PON1 and CFHR3 were low expressed and MARCKSL1, SPP1, BSG, CCT3, LAGE3, KPNA2, SF3B4, GTPBP4 were over expressed in human hepatoma cells compared with normal human hepatocyte cells (Unpaired t-test,  $p < 0.01$ ) (Figure 7A). Figure 7B shows the results of the protein expression levels of CYP2C9 and PON1 were down regulated in HepG2 and HCC-LM3 compared to MIHA. At the same time, CPTAC database analysis results showed that the protein expression of PON1, CFHR3 and CYP2C9 were low expressed in tumor tissues compared with paracancer normal tissue, while other genes were over expressed (Figure 7C).

Furthermore, immunohistochemical analysis from HPA database confirmed higher KPNA2, LAGE3, SF3B4, CCT3 and GTPBP4 protein expression and lower CYP2C9 and PON1 protein expression in HCC tissues (Figures 7D–J).



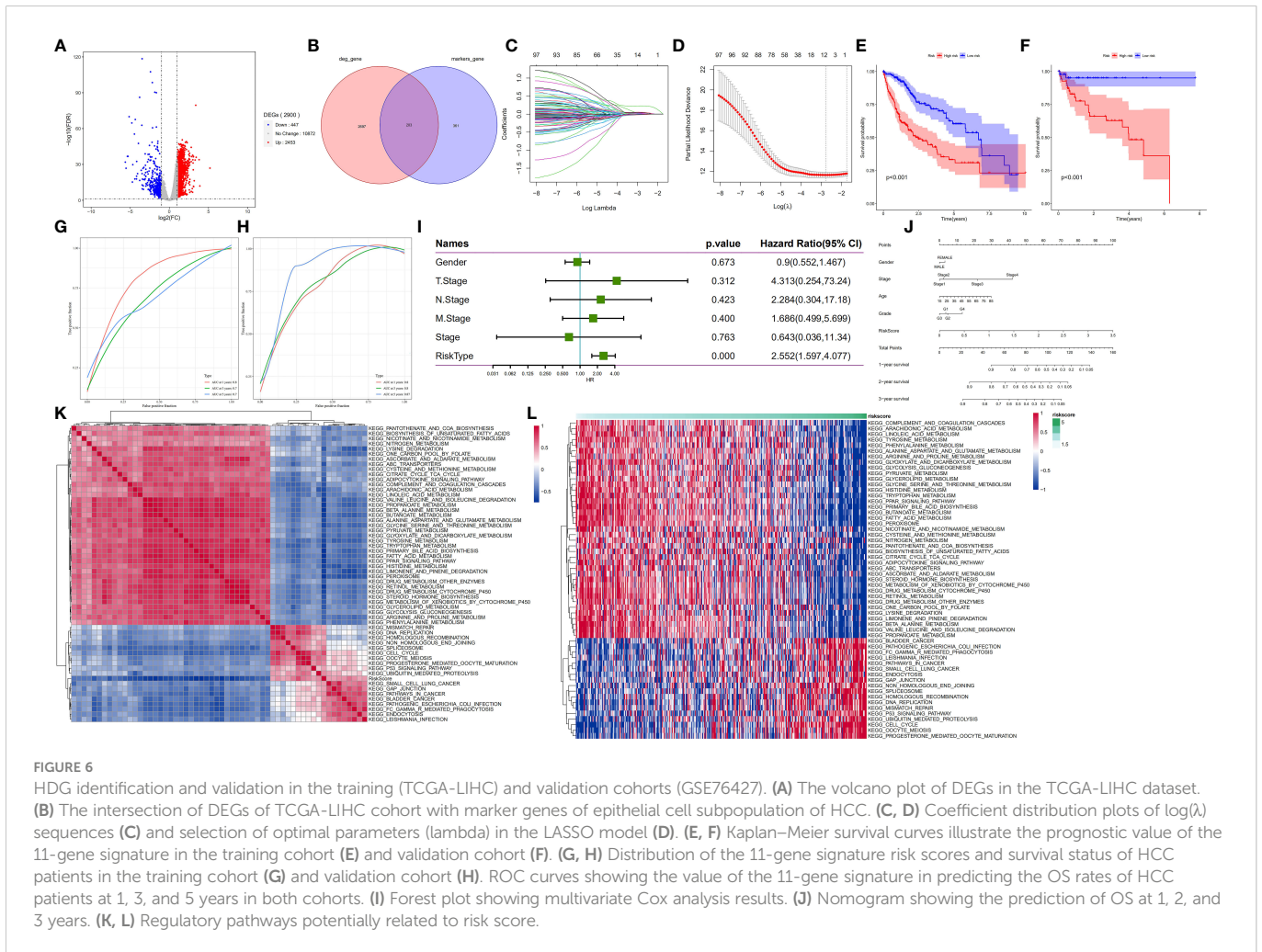
### 3.8 Drug sensitivity analysis of risk models and targeting of potential compounds in high risk groups using connectivity map (CMap)

To determine the impact of risks on clinical practice, we evaluated the IC50 values of several chemotherapeutic agents in the high- and low-risk groups using the “oncoPredict” package. This analysis identified 123 drugs that were statistically significant ( $p < 0.01$ ) (Supplementary Table S5). The results showed that afatinib, dasatinib, 5-fluorouracil, lapatinib, SCH772984, and cediranib had lower IC50 values in the high-risk group than in the low-risk group, suggesting that patients in the high-risk group may benefit more from these drugs. In contrast, JQ1, AT13148, axitinib, AZ960, AZD1208, and irinotecan had lower IC50 values in the low-risk group, suggesting that low-risk patients may benefit more from the above chemotherapeutic agents (Figure 8A).

While single-cell sequencing strategies are powerful tools for constructing disease signatures specific to individual cell types, CMap provides unprecedented convenience for researchers to tightly link the triad of drug, gene and disease in a context where

deep understanding is lacking, as this method does not require the detailed mechanism of action or drug target to be provided in advance to predict therapeutic potential. Therefore, by combining a high-resolution single-cell sequencing strategy with CMap, we have been able to directly target effective therapeutic agents based on individual cell-level expression signatures and thus provide a more accurate prediction for screening potential drugs for disease.

We used a computational drug discovery strategy based on “signature reversion” (32) to identify drugs with a high risk of reversion using the large amount of data in the CMap database (Figure 8B). The top 300 genes with the highest fold change in the high- and low-risk groups were extracted for XSum analysis (Supplementary Table S6). The results of the CMap analysis revealed several compounds with gene expression patterns opposite to those specific to the high-risk group, with lower CMap scores indicating a higher perturbation ability. PHA.00816795, mercaptopurine, W.13, NU.1025 and arachidonyltrifluoromethane were the five potentially valuable small molecule drug candidates, as they were ranked as the top 5 candidates (Figure 8C). Among the top three of these candidates is mercaptopurine, which is a common chemotherapeutic drug that produces anticancer effects by interfering



**FIGURE 6** HDG identification and validation in the training (TCGA-LIHC) and validation cohorts (GSE76427). (A) The volcano plot of DEGs in the TCGA-LIHC dataset. (B) The intersection of DEGs of TCGA-LIHC cohort with marker genes of epithelial cell subpopulation of HCC. (C, D) Coefficient distribution plots of  $\log(\lambda)$  sequences (C) and selection of optimal parameters (lambda) in the LASSO model (D). (E, F) Kaplan–Meier survival curves illustrate the prognostic value of the 11-gene signature in the training cohort (E) and validation cohort (F). (G, H) Distribution of the 11-gene signature risk scores and survival status of HCC patients in the training cohort (G) and validation cohort (H). ROC curves showing the value of the 11-gene signature in predicting the OS rates of HCC patients at 1, 3, and 5 years in both cohorts. (I) Forest plot showing multivariate Cox analysis results. (J) Nomogram showing the prediction of OS at 1, 2, and 3 years. (K, L) Regulatory pathways potentially related to risk score.

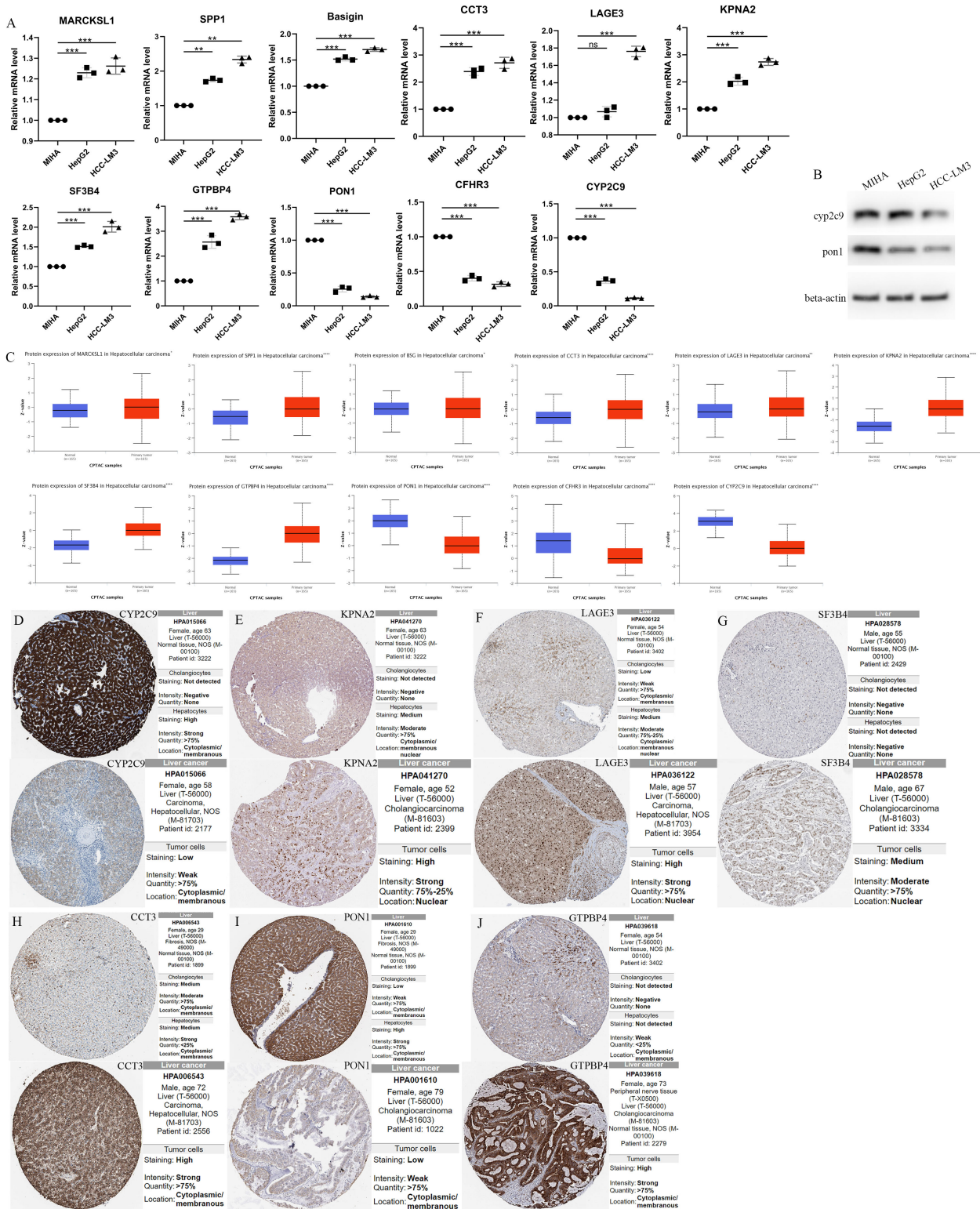
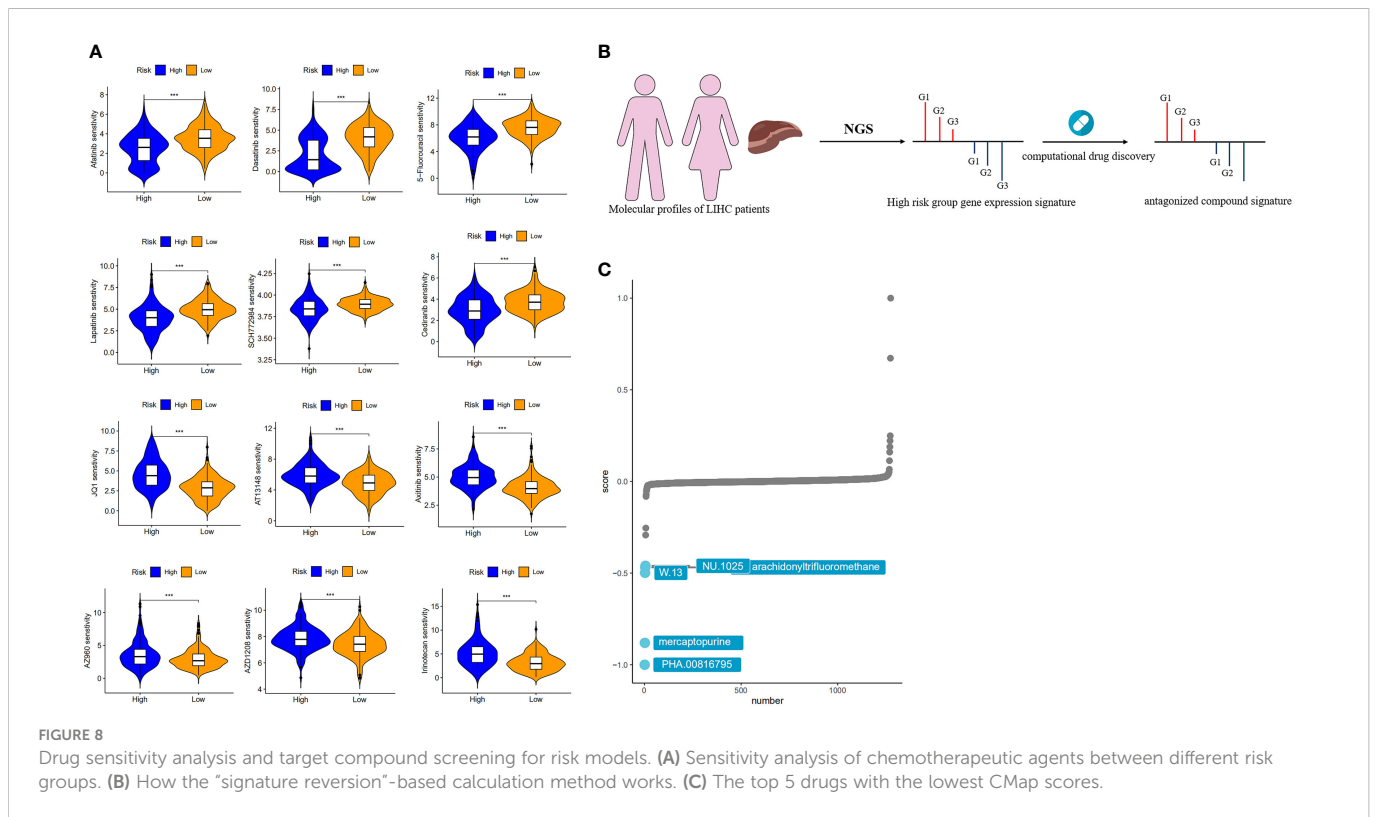


FIGURE 7

The Relative RNA Expression Level and Protein Expression Level of prognosis-related differentially expressed genes. (A) The Relative RNA Expression Level of MARCKSL1, SPP1, BSG, CCT3, LAGE3, KPNA2, SF3B4, GTPBP4, PON1, CFHR3 and CYP2C9. (B) Expression of CYP2C9 and PON1 in normal human hepatocyte cell line MIHA and HCC cell lines HCC-LM3 and HepG2 through western blot analysis. (C) Box plots showed the differential protein expression of 11 hub genes in the CPTAC dataset in HCC tumor tissue and adjacent normal. (D–J) Immunohistochemical analysis of the CYP2C9, KPNA2, LAGE3, SF3B4, CCT3, PON1 and GTPBP4 in HCC and liver tissues from the HPA database. HCC, hepatocellular carcinoma; CPTAC, The National Cancer Institute’s Clinical Proteomic Tumor Analysis Consortium. HPA, Human Protein Atlas. (Unpaired t-test, \*\*P < 0.01, \*\*\*P < 0.001, \*\*\*\*P < 0.0001 and ns, no significance).



with cell division or DNA synthesis (33). Yu et al. obtained five drugs associated with HCC by integrating multiple data to define the types of genes, considering the effect of genetic changes on HCC and the positive and negative relationships between drugs and HCC (34). Among these drugs, mercaptopurine is a potential anti-HCC drug.

## 4 Discussion

The liver is a major site for many metabolic processes, and metabolic dysregulation is vital for HCC progression and development (35). Evidence from numerous studies has shown that HCC originates from adult hepatocytes (36, 37). In this study, we found that HCC occurs in adult hepatocytes. Moreover, there were metabolic changes in the hepatic epithelial cells. While normal hepatocytes produce energy primarily through oxidative phosphorylation, malignant hepatocytes convert glucose into lactate through glycolysis to generate energy, a phenomenon known as the Warburg effect (38). The dysregulation of oxidative phosphorylation is related to elevated HCC tumorigenicity (39, 40). The liver synthesizes lactic acid and can store and breakdown lipids. Therefore, in HCC, aberrant lipid metabolism generates the lipids required for membrane formation and energy production, and posttranslational modifications support tumorigenesis (41). In our study, the gluconeogenic pathway (aerobic gluconeogenesis) was found to be enhanced in normal hepatic epithelial cells adjacent to cancerous cells, whereas the lipid metabolism pathway was enriched in malignant hepatocytes (Figure 2E).

HCC is a heterogeneous disease influenced by multiple factors, which makes it difficult to diagnose and perform individualized treatment. HCC patients are often diagnosed after curative surgical

approaches are no longer possible because these patients are at an advanced stage of the disease. Traditional sequencing methods often mask the underlying heterogeneity in phenotypically defined cell subpopulations. In contrast, scRNA-seq allows the in-depth exploration of tumour heterogeneity and the analysis of tumour development, drug resistance, intercellular communication and immune infiltration patterns (12). Thus, this technique was employed to comprehensively analyse the HCC landscape at single-cell resolution.

To understand the interactions among hepatocytes and stromal cells and immune cells, we conducted intercellular communication analysis. This revealed enhanced interactions between hepatic epithelial cells and fibroblasts and reduced contact with immune cells, macrophages and endothelial cells in tumour samples compared to normal adjacent samples (Supplementary Figure S2). Cancer-associated fibroblasts (CAFs) are a major part of the tumour stroma and contribute to HCC progression. Furthermore, CAFs interact with tumour cells, immune cells, or vascular endothelial cells in the TME through direct intercellular contacts or indirect paracrine interactions to promote HCC (42–44). Similarly, Wang et al. performed a single-cell level analysis of samples from normal and malignant livers and found that in HCC, the most significant alteration was the expansion of ACTA2+ fibroblast populations and malignant cells. This suggests that the transition of hepatocytes from normal to malignant is accompanied by alterations in intercellular contact with other cells in the tumour microenvironment, which produce the complex intra- and intertumoral heterogeneity of HCC (45).

Differentially expressed genes between malignant and normal hepatocytes were identified by analysing copy number variations in single-cell transcriptome data and isolating malignant and nonmalignant cells from hepatocytes. In addition, analysis of

TCGA-LIHC survival data revealed the molecular markers associated with HCC prognosis, including MARCKSL1, SPP1, BSG, CCT3, LAGE3, KPNA2, SF3B4, GTPBP4, PON1, CFHR3, and CYP2C9. The identified prognostic risk factors showed good prediction performance in both HCC cohorts. Based on this, we also constructed a nomogram risk assessment model, which combines risk scores with clinical characteristics to facilitate the clinical application of HCC. It has been reported that CYP2C9 is involved in the metabolism of many carcinogens and drugs, and is down-regulated in HCC (46). Wang et al. used time serial transcriptome to reveal that *Cyp2c29* is a key gene in the development of hepatocellular carcinoma in the mouse model, and its overexpression enhances the production of 14,15-EET and inhibits inflammation induced hepatocellular proliferation by inhibiting the IKK-NF- $\kappa$ B pathway during liver injury (47). Meanwhile, the expression of the human homologous of *Cyp2c29* gene in mice was positively correlated with the survival time of HCC patients, further suggesting that CYP2C epoxygenases may be a potential therapeutic target for liver disease. Chen and others have revealed *IncZic2*/depletion/MARCKS/MARCKSL1 pathways can eliminate the liver tumor-initiating cells (TICs) (48). The overexpression of myristoylated alanine-rich protein kinase C substrate (MARCKS) and MARCKS like 1 (MARCKSL1) can drive the self-renewal of TICs. Yang et al. demonstrated that BSG may be a tumor-promoting factor in HCC (49). The potential diagnostic role of BSG in differentiating HCC specimens from non-tumor specimens was demonstrated by analysis of multiple cohorts. BSG mRNA expression levels were significantly upregulated in both HCC specimens and HCC cell lines, and significantly shorter Overall Survival (OS) ( $P = 0.0014$ ) and Disease Free Survival (DFS) ( $P = 0.0097$ ) were observed in patients with high BSG expression relative to those with low BSG expression. Han et al. revealed that CCT3 is a new complementary biomarker for HCC screening and diagnosis (50). Several studies have shown that CCT3 is overexpressed in HCC patients by quantitative RT-PCR and western blotting. CCT3 can influence the progression of HCC by affecting phosphorylation signaling and translocation of STAT3/STAT3 into the nucleus of HCC cells (51, 52). The study of Li et al. showed that LAGE3 has prognostic value in HCC, which may affect the progression path of HCC tumor by promoting the proliferation, survival, migration, invasion and anti-apoptosis of HCC cells through the PI3K/AKT/mTOR and Ras/RAF/MAPK pathways (53). Guo et al. identified KPNA2 as a potential diagnostic and prognostic biomarker for HCC, which may affect HCC cell proliferation and migration by regulating cell cycle and DNA replication (54). Splicing factor 3b subunit 4 (SF3B4) has been revealed to be associated with the diagnosis and prognosis of HCC (55, 56). Liu et al. further demonstrated that SF3B4 drives cell proliferation and metastasis in HCC (57). Deng et al. further studied the mechanism and revealed the interaction between SF3B4 and ENAH in HCC, that is, SF3B4-regulated ENAH promotes the development of HCC by activating Notch signaling (58). It has been reported that Guanosine triphosphate binding protein 4 (GTPBP4) is associated with poor prognosis in HCC patients (59). Additional reports have explored the role of GTPBP4 in metabolic regulation and the potential mechanisms involved in HCC development and metastasis (60). GTPBP4 induces the dimer conformation of PKM2 through the SUMOylation to promote the

aerobic glycolysis of HCC, thus promoting the progression and metastasis of HCC (61). Serum Paraoxonase 1 (PON1) has been reported as a biomarker for evaluating microvascular infiltration in hepatocellular carcinoma. Complement factor H related 3 (CFHR3) can be used to predict the prognosis of HCC. Overexpression of CFHR3 can affect the proliferation and apoptosis of hepatocellular carcinoma (62). Recent reports suggest that overexpression of CFHR3 may be a potential strategy for overcoming hypoxia and treating HCC (63). These studies confirm the significance and plausibility of these prognostic signatures.

Currently, liver transplantation and resection are efficient treatment options for early-stage disease; however, these treatments are appropriate for only 20-30% of HCC patients (64). Chemotherapy is another viable treatment option for advanced HCC. Recently, there has been significant progress in the development of molecularly targeted treatments for liver cancer (65). These include sorafenib, levatinib, and regorafenib, which have been approved as first- and second-line treatments for HCC. In this study, the sensitivities of HCC to various treatments were predicted. Low-risk patients showed higher sensitivity to afatinib, dasatinib, 5-fluorouracil, lapatinib, SCH72984, and cediranib than high-risk patients, which may be attributed to their higher metabolic activity. Various drugs were suggested for low-risk group patients, such as JQ1, AT13148, axitinib, AZ960, AZD1208, and irinotecan. Cancerous cells have the potential to evade the immune system (66). Immune escape can be achieved through a variety of mechanisms. Thus, therapeutic strategies that block checkpoint inhibitors of the PD-1/PD-L1 and CTLA-4 pathways can promote tumour-reactive T-cell aggregation, thereby improving the antitumour response (67, 68).

To the best of our knowledge, the 11-gene signature is the first to explore the overall molecular prognostic feature of subpopulations associated with metabolic disorders from single-cell sequencing data. This risk model exhibited excellent ability to predict the prognosis of HCC patients, and the AUC values at year 1, year 3 and year 5 were all greater than 0.7, with the optimal value of 0.8. Meanwhile, a novel XSum algorithm was used to predict potential drugs targeting high-risk groups from the Cmap database, and 5 drugs were finally obtained, including PHA.00816795, mercaptopurine, W.13, NU. 1025 and arachidonyl trifluoromethane. Small molecule drugs, serving as candidates, embrace potential value conducive to providing medication strategies for accurate treatment of HCC patients.

This research also has certain drawbacks. First, more perspective data with larger sample size should be collected to validate the accuracy of our established prognostic model. Second, the characteristics of different fractionated epithelial cells have not been generated and validated. Further in-depth analysis from specific epithelial cell subtypes closely related to metabolic changes will be conducive to obtaining more accurate and valid prognostic characteristics. Nevertheless, scRNA-seq analysis sheds new light on the metabolic characteristics of individual cell subsets in HCC, and anchors the survival and prognosis of relevant cell subsets with the most significant metabolic changes, which is a key step forward in clinical practice.

In conclusion, the present study identified prognostic genes significantly associated with metabolic changes in a hepatocyte

subpopulation at the single-cell level, and explored the heterogeneity of this subpopulation and its interrelationships with other cells in the tumor microenvironment. A prognostic model for OS prediction in HCC patients was established and validated and the results demonstrated favourable predictive ability. Additionally, differences in chemosensitivity between high-risk and low-risk groups were evaluated, and five potential drugs that might reverse the risk score were forecasted. These results provided an in-depth understanding of the metabolic characteristics of HCC. Furthermore, the characteristics of potential prognostic biomarker can be clarified through the comparison of tumor-related genes constructed by liver malignant cells and normal hepatocytes. The above may be conducive to new strategies of individualized therapy.

## Data availability statement

The original contributions presented in the study are included in the article/[Supplementary Material](#). Further inquiries can be directed to the corresponding authors.

## Ethics statement

Ethical approval was not provided for this study on human participants because TCGA and GEO belong to public databases. The patients involved in the database have obtained ethical approval. Users can download relevant data for free for research and publish relevant articles. Our study is based on open source data, so there are no ethical issues and other conflicts of interest. The Cancer Genome Atlas (TCGA), a landmark cancer genomics program, molecularly characterized over 20,000 primary cancer and matched normal samples spanning 33 cancer types. This joint effort between NCI and the National Human Genome Research Institute began in 2006, bringing together researchers from diverse disciplines and multiple institutions. Over the next dozen years, TCGA generated over 2.5 petabytes of genomic, epigenomic, transcriptomic, and proteomic data. The data, which has already led to improvements in our ability to diagnose, treat, and prevent cancer, will remain publicly available for anyone in the research community to use. (<https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga>). Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

## References

- Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* (2021) 71(3):209–49. doi: 10.3322/caac.21660
- Singal AG, Lampertico P, Nahon P. Epidemiology and surveillance for hepatocellular carcinoma: New trends. *J Hepatol* (2020) 72(2):250–61. doi: 10.1016/j.jhep.2019.08.025
- Li X-Y, Shen Y, Zhang L, Guo X, Wu J. Understanding initiation and progression of hepatocellular carcinoma through single cell sequencing. *Biochim Biophys Acta (BBA)-Reviews Cancer* (2022) p:188720. doi: 10.1016/j.bbcan.2022.188720
- Llovet J, Kelley RK, Villanueva A, Singal AG, Pikarsky E, Roayaie S, et al. Hepatocellular carcinoma. *Nat Rev Dis Primers* (2021) 7(1):6. doi: 10.1038/s41572-020-00240-3
- Balogh J, Victor D 3rd, Asham EH, Burroughs SG, Bektour M, Saharia A, et al. Hepatocellular carcinoma: a review. *J hepatocellular carcinoma* (2016) 3:41. doi: 10.2147/JHC.S61146
- Zhang T, Merle P, Wang H, Zhao H, Kudo M. Combination therapy for advanced hepatocellular carcinoma: do we see the light at the end of the tunnel? *Hepatobiliary Surg Nutr* (2021) 10(2):180. doi: 10.21037/hbsn-2021-7

## Author contributions

CH and JC contributed conception and design of the study; ZH and HY refined the research design idea. JC and JH collected the data; CH and JC performed the statistical analysis; JH and RZ accomplished the RT-qPCR assay and Western blot; CH and JZ further supplemented and validated the data and assisted with the interpretation of the results; CH and JC drafted the manuscript; ZH and HY revised the logic of the manuscript and polished the language. All authors contributed to manuscript and approved the submitted version.

## Funding

The present study was supported by the National Natural Science Foundation of China (project no. 81902855), Natural Science Foundation of Guangdong Province (project no. 2020A1515010026), Funds for PhD Researchers of Guangdong Medical University in 2019 (project no. B2019019), and Discipline Construction Project of Guangdong Medical University (project no. 4SG22005G).

## 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.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2023.1104262/full#supplementary-material>

7. Todesca P, Marzi L, Critelli RM, Cuffari B, Caporali C, Turco L, et al. Angiopoietin-2/Tie2 inhibition by regorafenib associates with striking response in a patient with aggressive hepatocellular carcinoma. *Hepatology* (2019) 70(2):745–7. doi: 10.1002/hep.30588
8. Gao X, Chen H, Huang X, Li H, Liu Z, Bo X. ARQ-197 enhances the antitumor effect of sorafenib in hepatocellular carcinoma cells via decelerating its intracellular clearance. *OncoTargets Ther* (2019) 12:1629. doi: 10.2147/OTT.S196713
9. Li D, Sedano S, Allen R, Gong J, Cho M, Sharma S, et al. Current treatment landscape for advanced hepatocellular carcinoma: patient outcomes and the impact on quality of life. *Cancers* (2019) 11(6):841. doi: 10.3390/cancers11060841
10. Vande Lune P, Abdel Aal AK, Klimkowski S, Zorzour JG, Gunn AJ. Hepatocellular carcinoma: Diagnosis, treatment algorithms, and imaging appearance after transarterial chemoembolization. *J Clin Transl Hepatol* (2018) 6(2):175–88. doi: 10.14218/JCTH.2017.00045
11. Newman AM, Steen CB, Liu CL, Gentles AJ, Chaudhuri AA, Scherer F, et al. Determining cell type abundance and expression from bulk tissues with digital cytometry. *Nat Biotechnol* (2019) 37(7):773–82. doi: 10.1038/s41587-019-0114-2
12. González-Silva L, Quevedo L, Varela I. Tumor functional heterogeneity unraveled by scRNA-seq technologies. *Trends Cancer* (2020) 6(1):13–9. doi: 10.1016/j.trecan.2019.11.010
13. Sun Y, Wu L, Zhong Y, Zhou K, Hou Y, Wang Z, et al. Single-cell landscape of the ecosystem in early-relapse hepatocellular carcinoma. *Cell* (2021) 184(2):404–421.e16. doi: 10.1016/j.cell.2020.11.041
14. Han X, Zhou Z, Fei L, Sun H, Wang R, Chen Y, et al. Construction of a human cell landscape at single-cell level. *Nature* (2020) 581(7808):303–9. doi: 10.1038/s41586-020-2157-4
15. Zheng C, Zheng L, Yoo JK, Guo H, Zhang Y, Guo X, et al. Landscape of infiltrating T cells in liver cancer revealed by single-cell sequencing. *Cell* (2017) 169(7):1342–1356.e16. doi: 10.1016/j.cell.2017.05.035
16. Zheng H, Pomyen Y, Hernandez MO, Li C, Livak F, Tang W, et al. Single-cell analysis reveals cancer stem cell heterogeneity in hepatocellular carcinoma. *Hepatology* (2018) 68(1):127–40. doi: 10.1002/hep.29778
17. Ho D, Tsui YM, Sze KM, Chan LK, Cheung TT, Lee E, et al. Single-cell transcriptomics reveals the landscape of intra-tumoral heterogeneity and stemness-related subpopulations in liver cancer. *Cancer Lett* (2019) 459:176–85. doi: 10.1016/j.canlet.2019.06.002
18. Yu L, Shen N, Shi Y, Shi X, Fu X, Li S, 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. doi: 10.3389/fimmu.2022.1009789
19. Satriano L, Lewinska M, Rodrigues PM, Banales JM, Andersen JB. Metabolic rearrangements in primary liver cancers: cause and consequences. *Nat Rev Gastroenterol Hepatol* (2019) 16(12):748–66. doi: 10.1038/s41575-019-0217-8
20. Jin S, Guerrero-Juarez CF, Zhang L, Chang I, Ramos R, Kuan CH, et al. Inference and analysis of cell-cell communication using CellChat. *Nat Commun* (2021) 12(1):1–20. doi: 10.1038/s41467-021-21246-9
21. Wu Y, Yang S, Ma J, Chen Z, Song G, Rao D, et al. Spatiotemporal immune landscape of colorectal cancer liver metastasis at single-cell level. *Cancer Discovery* (2022) 12(1):134–53. doi: 10.1158/2159-8290.CD-21-0316
22. Tirosh I, Izar B, Prakadan SM, Wadsworth MH2nd, Treacy D, Trombetta JJ, et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* (2016) 352(6282):189–96. doi: 10.1126/science.aad0501
23. Maeser D, Gruener RF, Huang RS. oncoPredict: an R package for predicting in vivo or cancer patient drug response and biomarkers from cell line screening data. *Briefings Bioinform* (2021) 22(6):bbab260. doi: 10.1093/bib/bbab260
24. Al Mahi N, Zhang EY, Sherman S, Yu JJ, Medvedovic M. Connectivity map analysis of a single-cell RNA-sequencing-derived transcriptional signature of mTOR signaling. *Int J Mol Sci* (2021) 22(9):4371. doi: 10.3390/ijms22094371
25. Yang C, Zhang H, Chen M, Wang S, Qian R, Zhang L, et al. A survey of optimal strategy for signature-based drug repositioning and an application to liver cancer. *Elife* (2022) 11:e71880. doi: 10.7554/eLife.71880
26. Tanese K, Hashimoto Y, Berkova Z, Wang Y, Samaniego F, Lee JE, et al. Cell surface CD74-MIF interactions drive melanoma survival in response to interferon- $\gamma$ . *J Invest Dermatol* (2015) 135(11):2901. doi: 10.1038/jid.2015.259
27. Jeong H, Lee SY, Seo H, Kim BJ. Recombinant mycobacterium smegmatis delivering a fusion protein of human macrophage migration inhibitory factor (MIF) and IL-7 exerts an anticancer effect by inducing an immune response against MIF in a tumor-bearing mouse model. *J Immunother Cancer* (2021) 9(8). doi: 10.1136/jitc-2021-003180
28. Wirtz T, Saal A, Bergmann I, Fischer P, Heinrichs D, Brandt EF, et al. Macrophage migration inhibitory factor exerts pro-proliferative and anti-apoptotic effects via CD74 in murine hepatocellular carcinoma. *Br J Pharmacol* (2021) 178(22):4452–67. doi: 10.1111/bph.15622
29. Ma L, Du W, Chen Z, Xiang C. Single-cell atlas of tumor clonal evolution in liver cancer. *bioRxiv* (2020). doi: 10.1101/2020.08.18.254748
30. Song Z, Chen W, Athavale D, Ge X, Desert R, Das S, et al. Osteopontin takes center stage in chronic liver disease. *Hepatology (Baltimore Md.)* (2021) 73(4):1594–608. doi: 10.1002/hep.31582
31. Moon B, Jeong WJ, Park J, Kim TI, Min Do S, Choi KY. Role of oncogenic K-ras in cancer stem cell activation by aberrant wnt/ $\beta$ -catenin signaling. *J Natl Cancer Institute* (2014) 106(2):dj373. doi: 10.1093/jnci/dj373
32. Li J, Zheng S, Chen B, Butte AJ, Swamidass SJ, Lu Z. A survey of current trends in computational drug repositioning. *Briefings Bioinform* (2016) 17(1):2–12. doi: 10.1093/bib/bbv020
33. Sadia M. A review on CRISPR-cas system based applications in oncology. *Brac University* (2022).
34. Yu L, Xu F, Gao L. Predict new therapeutic drugs for hepatocellular carcinoma based on gene mutation and expression. *Front Bioengineering Biotechnol* (2020) 8:8. doi: 10.3389/fbioe.2020.00008
35. Du D, Liu C, Qin M, Zhang X, Xi T, Yuan S, et al. Metabolic dysregulation and emerging therapeutical targets for hepatocellular carcinoma. *Acta Pharm Sin B* (2021) 2(2):558–80. doi: 10.1016/j.apsb.2021.09.019
36. Font-Burgada J, Shalapur S, Ramaswamy S, Hsueh B, Rossell D, Umemura A, et al. Hybrid periportal hepatocytes regenerate the injured liver without giving rise to cancer. *Cell* (2015) 162(4):766–79. doi: 10.1016/j.cell.2015.07.026
37. Marquardt JU. *Deconvolution of the cellular origin in hepatocellular carcinoma: Hepatocytes take the center stage*. Wiley Online Library (2016) p. 1020–3.
38. Vander Heiden MG, Cantley LC, Thompson CB. Understanding the warburg effect: the metabolic requirements of cell proliferation. *science* (2009) 324(5930):1029–33. doi: 10.1126/science.1160809
39. Lee Y-K, Lim JJ, Jeoun UW, Min S, Lee EB, Kwon SM, et al. Lactate-mediated mitochondrial defects impair mitochondrial oxidative phosphorylation and promote hepatoma cell invasiveness. *J Biol Chem* (2017) 292(49):20208–17. doi: 10.1074/jbc.M117.809012
40. Soukupova J, Malfettone A, Hyrossova P, Hernandez-Alvarez MI, Penuelas-Haro I, Bertran E, et al. Role of the transforming growth factor- $\beta$  in regulating hepatocellular carcinoma oxidative metabolism. *Sci Rep* (2017) 7(1):1–15. doi: 10.1038/s41598-017-12837-y
41. Zaidi N, Lupien L, Kuemmerle NB, Kinlaw WB, Swinnen JV, Smans K. Lipogenesis and lipolysis: the pathways exploited by the cancer cells to acquire fatty acids. *Prog Lipid Res* (2013) 52(4):585–9. doi: 10.1016/j.plipres.2013.08.005
42. Peng H, Zhu E, Zhang Y. Advances of cancer-associated fibroblasts in liver cancer. *biomark Res* (2022) 10(1):1–18. doi: 10.1186/s40364-022-00406-z
43. Liu C, Liu L, Chen X, Cheng J, Zhang H, Zhang C, et al. LSD1 stimulates cancer-associated fibroblasts to drive Notch3-dependent self-renewal of liver cancer stem-like CellsLSD1 regulates liver CSC self-renewal via Notch3 signaling. *Cancer Res* (2018) 78(4):938–49. doi: 10.1158/0008-5472.CAN-17-1236
44. Zhang M, Yang H, Wan L, Wang Z, Wang H, Ge C, et al. Single-cell transcriptomic architecture and intercellular crosstalk of human intrahepatic cholangiocarcinoma. *J Hepatol* (2020) 73(5):1118–30. doi: 10.1016/j.jhep.2020.05.039
45. Wang H, Feng C, Lu M, Zhang B, Xu Y, Zeng Q, et al. Integrative single-cell transcriptome analysis reveals a subpopulation of fibroblasts associated with favorable prognosis of liver cancer patients. *Trans Oncol* (2021) 14(1):100981. doi: 10.1016/j.tranon.2020.100981
46. Yu D, Green B, Marrone A, Guo Y, Kadlubar S, Lin D, et al. Suppression of CYP2C9 by microRNA hsa-miR-128-3p in human liver cells and association with hepatocellular carcinoma. *Sci Rep* (2015) 5(1):1–9. doi: 10.1038/srep08534
47. Wang Q, Tang Q, Zhao L, Zhang Q, Wu Y, Hu H, et al. Time serial transcriptome reveals Cyp2c29 as a key gene in hepatocellular carcinoma development. *Cancer Biol Med* (2020) 17(2):401. doi: 10.20892/j.issn.2095-3941.2019.0335
48. Chen Z, Liu Y, Yao L, Guo S, Gao Y, Zhu P. The long noncoding RNA lncZic2 drives the self-renewal of liver tumor-initiating cells via the protein kinase c substrates MARCKS and MARCKSL1. *J Biol Chem* (2018) 293(21):7982–92. doi: 10.1074/jbc.RA117.001321
49. Yang L, Deng WL, Zhao BG, Xu Y, Wang XW, Fang Y, et al. FOXO3-induced lncRNA LOC554202 contributes to hepatocellular carcinoma progression via the miR-485-5p/BSG axis. *Cancer Gene Ther* (2022) 29(3):326–40. doi: 10.1038/s41417-021-00312-w
50. Qian E-N, Han SY, Ding SZ, Lv X. Expression and diagnostic value of CCT3 and IQGAP3 in hepatocellular carcinoma. *Cancer Cell Int* (2016) 16(1):1–8. doi: 10.1186/s12935-016-0332-3
51. Zhang Y, Wang Y, Wei Y, Wu J, Zhang P, Shen S, et al. Molecular chaperone CCT3 supports proper mitotic progression and cell proliferation in hepatocellular carcinoma cells. *Cancer Lett* (2016) 372(1):101–9. doi: 10.1016/j.canlet.2015.12.029
52. Cui X, Hu ZP, Li Z, Gao PJ, Zhu JY. Overexpression of chaperonin containing TCP1, subunit 3 predicts poor prognosis in hepatocellular carcinoma. *World J Gastroenterology: WJG* (2015) 21(28):8588. doi: 10.3748/wjg.v21.i28.8588
53. Li Y, Xiong H. Correlation of LAGE3 with unfavorable prognosis and promoting tumor development in HCC via PI3K/AKT/mTOR and Ras/RAF/MAPK pathways. *BMC Cancer* (2022) 22(1):1–13. doi: 10.1186/s12885-022-09398-3
54. Guo X, Wang Z, Zhang J, Xu Q, Hou G, Yang Y, et al. Upregulated KPNA2 promotes hepatocellular carcinoma progression and indicates prognostic significance across human cancer types. *Acta Biochim Biophys Sin* (2019) 51(3):285–92. doi: 10.1093/abbs/gmz003
55. Shen Q, Nam SW. SF3B4 as an early-stage diagnostic marker and driver of hepatocellular carcinoma. *BMB Rep* (2018) 51(2):57. doi: 10.5483/BMBRep.2018.51.2.021

56. Iguchi T, Komatsu H, Masuda T, Nambara S, Kidogami S, Ogawa Y, et al. Increased copy number of the gene encoding SF3B4 indicates poor prognosis in hepatocellular carcinoma. *Anticancer Res* (2016) 36(5):2139–44.
57. Liu Z, Li W, Pang Y, Zhou Z, Liu S, Cheng K, et al. SF3B4 is regulated by microRNA-133b and promotes cell proliferation and metastasis in hepatocellular carcinoma. *EBioMedicine* (2018) 38:57–68. doi: 10.1016/j.ebiom.2018.10.067
58. Deng G, Luo Y, Zhang Y, Zhang J, He Z. Enabled homolog (ENAH) regulated by RNA binding protein splicing factor 3b subunit 4 (SF3B4) exacerbates the proliferation, invasion and migration of hepatocellular carcinoma cells via notch signaling pathway. *Bioengineered* (2022) 13(2):2194–206. doi: 10.1080/21655979.2021.2023983
59. Liu W-B, Jia WD, Ma JL, Xu GL, Zhou HC, Peng Y, et al. Knockdown of GTPBP4 inhibits cell growth and survival in human hepatocellular carcinoma and its prognostic significance. *Oncotarget* (2017) 8(55):93984. doi: 10.18632/oncotarget.21500
60. Zhou Q, Yin Y, Yu M, Gao D, Sun J, Yang Z, et al. GTPBP4 promotes hepatocellular carcinoma progression and metastasis via the PKM2 dependent glucose metabolism. *Redox Biol* (2022) 56:102458. doi: 10.1016/j.redox.2022.102458
61. Ding G-Y, Zhu XD, Ji Y, Shi GM, Shen YH, Zhou J, et al. Serum PON1 as a biomarker for the estimation of microvascular invasion in hepatocellular carcinoma. *Ann Trans Med* (2020) 8(5):204. doi: 10.21037/atm.2020.01.44
62. Liu H, Zhang L, Wang P. Complement factor h-related 3 overexpression affects hepatocellular carcinoma proliferation and apoptosis. *Mol Med Rep* (2019) 20(3):2694–702. doi: 10.3892/mmr.2019.10514
63. Zeng Z, Lei S, Wang J, Yang Y, Lan J, Tian Q, et al. A novel hypoxia-driven gene signature that can predict the prognosis of hepatocellular carcinoma. *Bioengineered* (2022) 13(5):12193–210. doi: 10.1080/21655979.2022.2073943
64. Wang S, Xie J, Zou X, Pan T, Yu Q, Zhuang Z, et al. Single-cell multiomics reveals heterogeneous cell states linked to metastatic potential in liver cancer cell lines. *Science* (2022) 25(3):103857. doi: 10.1016/j.isci.2022.103857
65. Yan T, Yu L, Zhang N, Peng C, Su G, Jing Y, et al. The advanced development of molecular targeted therapy for hepatocellular carcinoma. *Cancer Biol Med* (2022) 19(6):802. doi: 10.20892/j.issn.2095-3941.2021.0661
66. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell* (2011) 144(5):646–74. doi: 10.1016/j.cell.2011.02.013
67. Marin-Acevedo JA, Kimbrough EO, Lou Y. Next generation of immune checkpoint inhibitors and beyond. *J Hematol Oncol* (2021) 14(1):1–29. doi: 10.1186/s13045-021-01056-8
68. Faivre S, Rimassa L, Finn RS. Molecular therapies for HCC: Looking outside the box. *J Hepatol* (2020) 72(2):342–52. doi: 10.1016/j.jhep.2019.09.010