Biological Procedures Online

RESEARCH

Open Access

Tumor Prognostic Risk Model Related to Monocytes/Macrophages in Hepatocellular Carcinoma Based on Machine Learning and Multi-Omics

Xinliang Wan[1†], Yongchun Zou[1†], Qichun Zhou[1], Qing Tang[1], Gangxing Zhu[1], Luyu Jia[1], Xiaoyan Yu[1], Handan Mo[1], Xiaobing Yang[1,2*] and Sumei Wang[1,2*]

Abstract

Tumor-associated macrophages (TAMs) are crucial in hepatocellular carcinoma (HCC) development and invasion. This study explores monocyte/ macrophage-associated gene expression profiles in HCC, constructs a prognostic model based on these genes, and examines its relationship with drug resistance and immune therapy responses. Single-cell RNA sequencing(scRNA-seq) data from 10 HCC tissue biopsy samples, totaling 24,597 cells, were obtained from the GEO database to identify monocyte/macrophage-associated genes. A prognostic model was constructed and validated using external datasets and Western blot. Relationships between the model, clinical correlates, drug sensitivity, and immune therapy responses were investigated. From scRNA-seq data, 2,799 monocyte/macrophage marker genes were identified. Using the TCGA dataset, a prognostic model based on the single-gene UQCRH was constructed, stratifying patients into high-risk and low-risk groups based on overall survival rates. High-risk group patients showed reduced survival rates and higher UQCRH expression in tumor tissues. Western blot analysis further confirmed the elevated expression of UQCRH in HCC cell lines. Spatial transcriptomics analysis revealed that high UQCRH expression co-localized with malignant cells in the tumor tissue. Drug sensitivity analysis revealed that the high-risk group had lower sensitivity to sorafenib and axitinib. Immune therapy response analysis indicated poorer outcomes in the high-risk group, with more pronounced APC inhibition and a weaker IFN-II response. Clinical indicator analysis showed a positive correlation between high UQCRH expression and tumor invasion. Enrichment analysis of UQCRH and associated molecules indicated involvement in oxidative phosphorylation and mitochondrial electron transport. This study introduces a prognostic model for HCC

†Xinliang Wan and Yongchun Zou contributed equally to this work.

*Correspondence:
Xiaobing Yang
yangxiaobing2002@126.com
Sumei Wang
wangsumei198708@163.com

Full list of author information is available at the end of the article

patients based on monocyte/macrophage marker genes. The single-gene model predicts HCC patient survival and treatment outcomes, identifying high-risk individuals with varying drug sensitivities and immune suppression states.

Keywords Hepatocellular carcinoma, Machine learning, Multi-omics, Monocytes/ macrophages, Prognostic model, UQCRH

Introduction

Hepatocellular carcinoma (HCC) is the primary malignancy of the liver and constitutes a significant health challenge globally, particularly in regions such as Africa and Asia, where there is a high incidence of the disease related to risk factors like chronic hepatitis B and C virus infections, cirrhosis, and non-alcoholic fatty liver disease (NAFLD) [1]. As the third leading cause of cancer-related deaths worldwide, there is an urgent need for improved treatment strategies and accurate diagnostic methods for HCC [2]. The incidence of HCC is significantly influenced by geographical differences, with varying prevalence and related risk factors across different regions, highlighting the complexity of HCC with environmental factors and genetic susceptibility. Advances in targeted therapies and the emergence of immunotherapy bring new hope to the field, yet treatment remains challenging due to high recurrence rates and the late stage of disease at diagnosis for many patients [2, 3]. Current treatments focus on innovative methods such as local region treatments, systemic treatments, and immunotherapy, aiming to utilize the human immune system to fight cancer [4]. Despite these advancements, the prognosis for HCC remains poor [5], emphasizing the urgent need for early detection methods and more effective treatment options. Research on molecular biomarkers and the tumor immune microenvironment remains crucial, with ongoing studies aiming to develop personalized treatment plans that could significantly improve patient outcomes [1]. These efforts are essential for managing the disease, which is often diagnosed at an advanced stage and difficult to treat curatively [5].

Against this backdrop, the role of monocytes/macrophages in the HCC immune microenvironment has become a key area of investigation. These immune cells play a critical role in shaping the immune landscape of the tumor and directly influence its progression [6]. The intrinsic characteristics of tumor cells, along with various external factors in the tumor microenvironment, contribute to significant changes in the metabolism, phenotype, and biological functions of macrophages [7, 8]. These alterations drive the reprogramming of macrophages within the tumor microenvironment to support tumor growth [9, 10].

In recent years, bioinformatics approaches have been used to identify gene clusters and regulatory networks related to HCC prognosis, immune microenvironment, and ferroptosis, providing valuable insights into HCC outcomes [11–13]. However, traditional RNA-seq analysis, which reflects the average gene expression across all cells in a sample, fails to capture the molecular expression changes in specific cell types within the tumor microenvironment [14].

Single-cell RNA sequencing (scRNA-seq) is an advanced technology that enables gene expression analysis at the single-cell level, revealing differences between cell types and their interactions, as well as the heterogeneity of disease gene expression [15–21]. Machine learning (ML), particularly in oncology genomics, has become a crucial tool in improving medical accuracy and patient prognosis. Its applications in risk assessment, early diagnosis, prognosis evaluation, and therapeutic drug selection have already demonstrated significant effectiveness [22].

This study integrates machine learning with multi-omics sequencing to perform single-cell RNA sequencing and transcriptomic analysis on HCC tumor samples, focusing on cellular and molecular changes. A single-gene prognostic risk model for HCC patients is developed and validated using transcriptomic and proteomic datasets through survival analysis and expression differentials. Experimental validation of gene expression and spatial transcriptomics is also performed to examine gene localization in malignant cells. Finally, the model's drug sensitivity, predicted immune therapy responses, and immune function are analyzed. The study workflow is illustrated in Fig. 1.

Materials and methods

Acquisition and Preprocessing of Single-Cell Data

The single-cell sequencing data were sourced from the article by Lu et al. titled "A single-cell atlas of the multicellular ecosystem of primary and metastatic hepatocellular carcinoma", available in the GEO database under the dataset identifier GSE149614 [23]. We extracted single-cell sequencing data from ten primary tumor samples within this dataset for analysis. The raw data for each sample were processed using the Seurat package in RStudio [24]. Quality control was performed on the analysis samples, selecting cells containing more than 300 genes and fewer than 2500 genes, totaling 24,597 cells for further bioinformatics analysis.

Data Integration and Dimension Reduction

The Seurat objects containing gene expression data for each sample were processed using the Read10x()

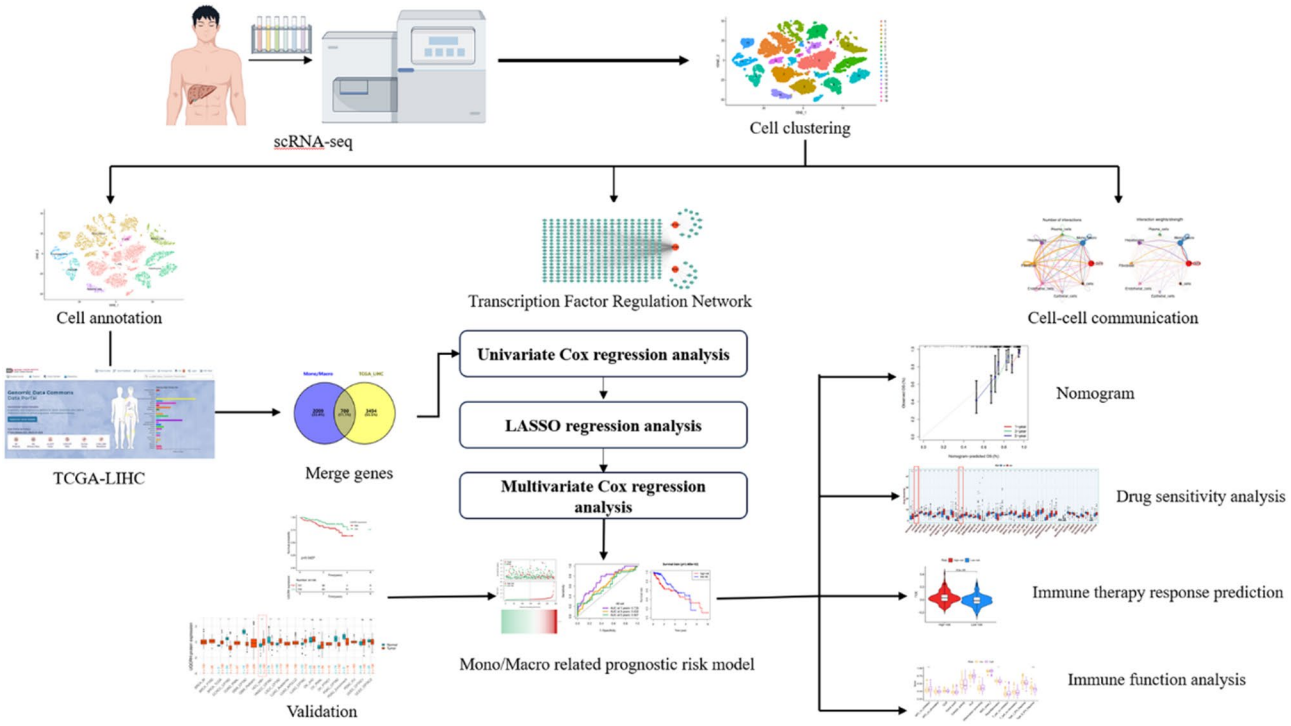

Fig. 1 The workflow of the present study. This study integrates machine learning with multi-omics sequencing to analyze HCC. A prognostic risk model is developed and validated, with analyses of drug sensitivity, immune therapy responses, and immune function

function [25]. After quality control, data were normalized using the 'NormalizeData' function. Gene expression was represented by scaling each gene's expression per sample by a factor of 10,000, followed by transformation to natural logarithm and normalization after adding one to avoid log of zero. The top 2000 highly variable genes (HVG) were identified using 'FindVariableFeatures' and the top 10 HVG were visualized.

Cell Clustering and Annotation

Cells were clustered using Seurat based on PCA scores, where each principal component (PC) represents a 'meta-feature' combining information from a related feature set. Dimensionality reduction was performed on the selected top 2000 highly variable genes using the 'RunPCA' function, and the effects of the reduction were visualized. Seurat employed spectral clustering to identify cell clusters and used 'FindNeighbors' and 'FindClusters' functions to distinguish all cell types. T-distributed Stochastic Neighbor Embedding (tSNE) was used for nonlinear dimension reduction visualization, as shown in Fig. 2(C) [26, 27]. Cell markers for hepatocellular carcinoma from CellMarker 2.0 (http://117.50.127.228/CellMarker/CellMarker_ help.html) [28] were referenced to annotate the discerned clusters of cell types. Subsequently, the proportion of each cell type in each sample was analyzed.

Marker Gene Identification

Using the 'FindAllMarkers' function in R software, marker genes for monocytes/macrophages in single-cell data were identified with a minimum percentage (min.pct) of 0.01 and a log2 fold change threshold (log2FC.threshold) of 0.01.

Cell-Cell Communication Analysis

Systematic analysis of signaling pathways between cells helps us understand the overall communication between them [29]. Using the R package CellChat, cell communication networks were constructed to identify key signaling pathways between macrophages and other cell types in the tumor microenvironment, such as T cells, inferring intercellular communication at the signaling pathway level.

Construction of Transcription Factor Gene Regulatory Networks

SCENIC is a computational method that reconstructs gene regulatory networks(GRN) and identifies cell states simultaneously from single-cell RNA-seq data [30]. Using SCENIC, stable cell states can be identified. Initially, co-expression networks of transcription factors and their target genes were analyzed using the SCENIC framework's Genie3 and AUCell on single-cell RNA data. These tools were used to derive transcriptional regulatory networks for monocyte/macrophage cell types and identify

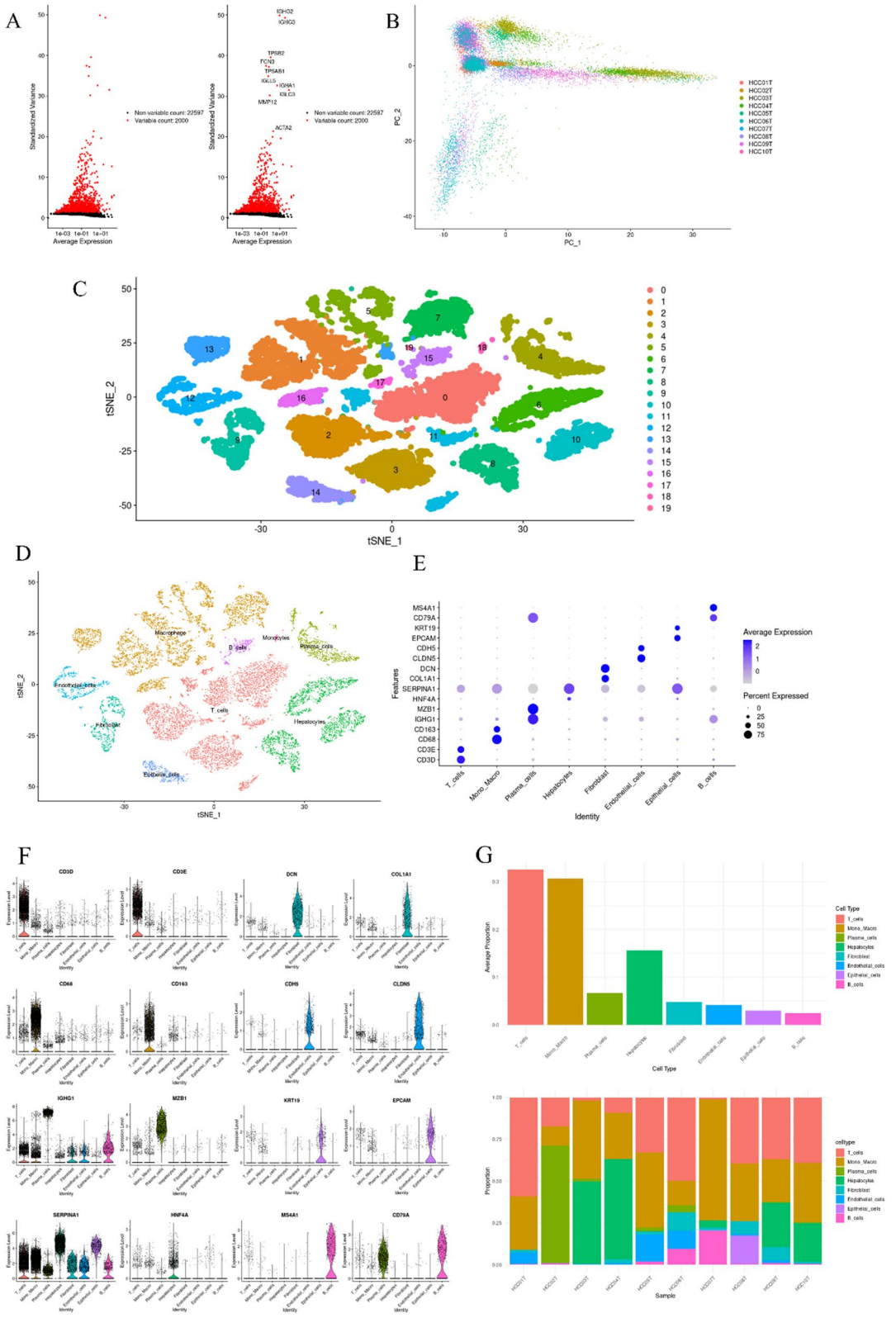

Fig. 2 Preliminary analysis of single-cell RNA sequencing samples. (**A**) Visualization of highly variable genes. (**B**) Principal component analysis (PCA) of the 10 included samples. (**C**) Clustering of cells from the included samples. (**D**) Annotation of cell types through marker genes. (**E**) Bubble chart showing the expression of marker genes across eight cell types. (**F**) Violin plots displaying the expression characteristics of marker genes in eight cell types. (**G**) Bar charts analyzing the average expression ratio of eight cell types across 10 single-cell samples and their individual proportions in each of the 10 samples

binding motifs of regulators using RcisTarget, thereby determining direct target relationships in the network. Subsequently, Regulons of single cells were scored based on each regulator and its target gene set, and cell activity states were predicted using a generated binary matrix. Finally, the transcription factor gene regulatory networks were constructed using Cytoscape. This process not only aids in understanding the functional states of cells but also reveals complex interaction networks between them.

Extraction and Analysis of Transcriptomic and Proteomic Data

Data on HCC mRNA and related clinicopathological features were downloaded from the TCGA database (https://portal.gdc.cancer.gov/), which included gene expression data and clinical features from 424 samples [31]. Additionally, transcriptomic sequencing samples from 389 HCC patients and proteomic sequencing samples from 330 patients were obtained from the ICGC database (https://dcc.icgc.org/) and the CPTAC database (https://proteomic.datacommons.cancer.gov/pdc/) respectively [32, 33]. After organizing and analyzing gene expression in LIHC samples from the TCGA database, genes in the top 70% of average gene expression in the TCGA-LIHC database were intersected with monocyte/macrophage marker genes from the single-cell data, and the extracted expression matrix was normalized.

Construction and Validation of a Single-Gene Prognostic Model

To identify significant gene expression markers affecting patient survival times, univariate Cox regression analysis was first performed on the intersected genes to identify potential differentially expressed genes [34]. Then, variables with $p<0.05$ were selected for regression analysis using the LASSO method, and the Cox proportional hazards model-based Lasso regression was constructed using the glmnet() function in the R package glmnet [35]. To evaluate the predictive performance of the model and prevent overfitting, the model's C-index was calculated and implemented using the cv.glmnet() function, selecting the C-index as the performance metric. Significant non-zero coefficient genes identified from the Lasso model were extracted, and their expression data along with patient survival information were used for constructing the prognostic model. To further evaluate the role of key genes identified from Lasso regression analysis in predicting patient survival, a multivariate Cox regression model was employed, and the prognostic model's Risk Score was calculated [36, 37]. The Risk Score is calculated using the following formula:

$$\text{Risk Score} = \sum_{i=0}^{k} \beta i * expi^{1} \quad (1)$$

Patients were then divided into high and low-risk groups based on the median Risk Score using the R packages survival and survminer, and Kaplan-Meier survival curves and risk plots were drawn and visualized along with survival status and differences for each sample. Additionally, the risk score's predictive performance for 1-year, 3-year, and 5-year patient survival was evaluated using the R packages pROC and timeROC. Integrating risk and clinical data, the function of the timeROC package was applied using the Aalen additive model to compute ROC curves and the corresponding area under the curve (AUC) at various time points without competing risks, quantifying the efficacy of risk scores and other clinical variables in predicting patient survival, with the AUC value visually demonstrating the risk model's ability to predict long-term patient survival [38]. Furthermore, HCC datasets from the ICGC database, GEO database(GSE14520) and the CPTAC dataset were selected as independent external cohorts to validate the prognostic model's effectiveness.

Correlation of Risk Model and its Gene Expression With Clinicopathological Features

To assess the correlation of clinical variables with patient prognosis, we used a nomogram for visual analysis. Integrating risk and clinical data, a Cox regression model was built using the cph function in the R package rms, and based on this model's results, a nomogram was generated using the regplot package. The nomogram intuitively shows the relative contribution of each clinical variable to the prognosis, with each variable assigned points according to its coefficient in the model. By summing the points of all variables, the predictive probability, i.e., the survival probability of patients at 1 year, 3 years, and 5 years, is obtained. Additionally, to verify the accuracy of the model, we also used the calibrate function to plot calibration curves, comparing the model-predicted survival probabilities with the actual observed survival rates. This method not only enhances the interpretability of the predictive model but also increases its practical value in clinical decision-making [39]. Furthermore, we used the R package ComplexHeatmap to draw heatmaps to visually assess the correlation between genes in the risk model and clinical features.

Drug Sensitivity Analysis, Predicting Immune Therapy Response, and Immune Function Scoring

Next, we conducted drug sensitivity analysis using the R package oncoPredict to assess the differences in drug responses between the high and low-risk groups of the

prognostic model. By combining TCGA-LIHC data and GDSC drug sensitivity data, oncoPredict was used to calculate drug sensitivities for the test samples [40]. Subsequently, the Wilcoxon test was used to compare drug sensitivities between the high and low-risk groups, and the results were visually presented using box plots drawn by the R package ggplot2. We then analyzed immune therapy responses in tumor samples using the Tumor Immune Dysfunction and Exclusion (TIDE) method [41]. By organizing and normalizing expression files from the TCGA-LIHC dataset and merging them with TIDE prediction results and clinical risk data, we compared TIDE scores between high and low-risk groups within the prognostic model to assess the potential immune therapy responses of patients with different risk levels. The analysis results were graphically displayed using the ggviolin in the R package ggpubr, further revealing the relationship between immune therapy responses and patient prognostic risk. Lastly, we conducted immune function scoring on the TCGA-LIHC dataset using the ssGSEA method and analyzed the immune-related gene sets of each sample using the R package GSVA. By evaluating the normalized scores combined with clinical risk data, we assessed the relationship between immune function scoring and patient prognosis.

Immunohistochemical Expression of Single Genes in the Prognostic Model, Paired Sample Differential Expression, and Enrichment Analysis of Directly Associated Proteins

We obtained immunohistochemical data for single genes in the prognostic model from normal liver tissue and HCC tissue from the HPA database (https://www.proteinatlas.org/) [42]. Additionally, we used the TCGA-LIHC dataset to analyze differential expression between normal and tumor samples for single genes in the prognostic model. Finally, based on the STRING database (https://string-db.org/), proteins directly associated with the single protein in the prognostic model were identified, and their protein interaction networks were visualized using Cytoscape. Subsequent KEGG and GO enrichment analyses were performed to understand the biological processes involving the single gene in the prognostic model [43].

Cell Lines and Antibodies

Human HCC cell lines, Hep-G2 and Huh-7, as well as LX2 cells, were obtained from Guangzhou Saiku Biotechnology Co., Ltd. (China). All cells were cultured at 37 °C in a humidified atmosphere with 5% CO2 and 95% air, in RPMI-1640 medium (Life Technologies, Carlsbad, CA, USA) supplemented with 10% fetal bovine serum (FBS, Gibco, USA) and 0.5% penicillin-streptomycin sulfate (Invitrogen, Carlsbad, CA, USA). Cell counts were performed using an automated cell counter (Star, Invitrogen, Carlsbad, CA, USA). GAPDH antibody was purchased from Cell Signaling Technology Inc. (Beverly, MA, USA), and anti-UQCRH (ab154803) antibody was purchased from Abcam (Cambridge, MA, USA).

Western Blot Analysis

Cells were harvested, washed, and lysed with 1× RIPA buffer. Protein concentrations were determined using the Thermo BCA Protein Assay Kit. Equal amounts of protein from the cell lysates were mixed with 5× SDS sample buffer, separated by 12% SDS-polyacrylamide gel electrophoresis, and transferred to polyvinylidene fluoride (PVDF) membranes. Membranes were blocked with 5% non-fat milk in TBST and incubated with primary antibodies against UQCRH and GAPDH at 4 °C overnight. After washing, the membranes were incubated with a secondary antibody against rabbit IgG for 1 h, followed by further washing. Detection was performed using ECL solution (Millipore, Darmstadt, Germany) and images were captured using the Bio-Rad ChemiDoc XRS + Chemiluminescence Imaging System (Bio-Rad, Hercules, CA, USA). Protein expression levels were quantified using ImageJ software.

Spatial Transcriptomics Analysis of UQCRH Expression

We downloaded HCC spatial transcriptomics samples (Sample ID: VISDS000508, VISDS000511) from the CROST database (https://ngdc.cncb.ac.cn/crost/). Using the R package spacexr, we mapped the single-cell transcriptomic data onto the tissue space and analyzed the gene expression patterns in different tissue regions based on spatial coordinates.

Statistical Analysis

The Wilcoxon signed-rank test was used to compare the expression differences of individual genes between normal and tumor samples. LASSO regression and Cox regression were employed to establish predictive models for disease prognosis, and Kaplan-Meier survival curves were used to evaluate the survival differences between different risk groups, with the log-rank test determining statistical significance. All statistical analyses were performed in Rstudio, and a two-sided p-value of less than 0.05 was considered statistically significant.

Results

Cellular Composition of Hepatocellular Carcinoma and Gene Expression in Monocytes/Macrophages

This study included scRNA-seq samples from 10 primary hepatocellular carcinoma tumors. Each sample was analyzed by integrating the top 2000 highly variable genes, followed by PCA (Principal Component Analysis) and subsequent clustering dimension reduction (Fig. 2A, B). t-SNE clustering visualization showed that 24,597 cells from hepatocellular carcinoma tissues were grouped

into 19 subgroups (Fig. 2C), and clusters were identified as different cell types based on marker genes for hepatocellular carcinoma cell types retrieved from CellMarker 2.0 (Fig. 2D). Information about the cell subgroups is provided in Table 1. To clarify the characteristics of each cell subgroup, we defined the marker genes for each cell subgroup and visualized them through bubble plots and violin plots (Fig. 2E, F). Through cell annotation of the clusters, a total of 8 cell types were identified, including T cells, monocytes/macrophages, plasma cells, hepatocytes, fibroblasts, endothelial cells, epithelial cells, and B cells, with T cells, monocytes/macrophages, and hepatocytes present in multiple cell clusters. Subsequently, we analyzed and visualized the average number of each cell type across all samples and their numbers in each individual sample (Fig. 2G), finding a higher number of monocytes/ macrophages, which suggests that monocytes/macrophages may play a significant role in the development and invasion progression of HCC. Consequently, we further identified marker genes related to monocytes/macrophages in the single-cell data, with related data presented in Supplementary Table 1.

Cell Communication Analysis Related To Monocytes/ Macrophages in scRNA-seq

Using CellChat, we constructed intercellular communication networks by calculating cell communication probabilities, while independently analyzing the communication networks between each cell type and other cell types (Fig. 3A-C). Additionally, we inferred the direct communication networks involving monocytes/macrophages with other cell types and discovered pathways such as FN1 and SPP1, which exhibited intensive communication between monocytes/ macrophages and other cell types (especially hepatocytes) (Fig. 3D, E).

Transcription Factor Regulatory Networks Related To Monocytes/ Macrophages in scRNA-seq

Based on SCENIC analysis, key regulatory networks of transcription factors within monocytes/macrophages were predicted. ZNF385A, NFKB2, and NONO (three core transcription factors) will serve as central hubs controlling HCC-Mono/Macro (Fig. 4A). Subsequently, we visualized the predicted transcription factors and their regulatory genes using Cytoscape 3.10.1 to construct the transcription factor-gene regulatory network (Fig. 4B), where red represents transcription factors, and green represents regulated genes.

Table 1 Information about the cell subpopulations

Cluster ID	Cell type	Cluster ID	Cell type
0	T_cells	10	Hepatocytes
1	Mono_Macros	11	T_cells
2	T_cells	12	Endothelial_cells
3	T_cells	13	Mono_Macros
4	Plasma_cells	14	Epithelial_cells
5	Mono_Macros	15	B_cells
6	Hepatocytes	16	Mono_Macros
7	Mono_Macros	17	T_cells
8	Hepatocytes	18	Mono_Macros
9	Fibroblast	19	B_cells

Construction and Validation of Prognostic Model and Gene Expression

In the TCGA-LIHC database, genes in the top 70% of average gene expression were intersected with 2799 monocyte/macrophage marker genes from single-cell data (Supplementary Table 2), and visualized using a Venn diagram (Fig. 5A). The extracted expression matrix was then normalized. Subsequently, univariate Cox regression analysis was performed on 700 intersecting genes, identifying 197 genes with potential prognostic significance ($p < 0.05$) (Supplementary Table 3). To prevent overfitting of the prognostic model, Lasso regression analysis was conducted, resulting in the selection of 8 genes with prognostic significance (Fig. 5B, Supplementary Table 4). Ultimately, multivariate Cox regression analysis led to the identification of one gene with significant prognostic relevance, namely, Ubiquinol-cytochrome c reductase hinge protein (UQCRH), for which a single-gene prognostic model was constructed (Supplementary Tables 5–6). First, paired sample differential expression analysis of UQCRH revealed that this gene is significantly overexpressed in tumor samples from the TCGA-LIHC dataset (Fig. 5C). Immunohistochemical results from the HPA database showed that UQCRH has higher expression in hepatocellular carcinoma tumor cells compared to normal liver tissue (Fig. 5D, E). Subsequently, based on the median Risk Score calculated from the single-gene prognostic model, patients were divided into high and low-risk groups for survival analysis and ROC analysis. Survival curve analysis indicated that the high-risk group and the single-gene high-expression group had worse overall survival (Fig. 5F-H). Additionally, in the TCGA-LIHC data, ROC analysis demonstrated that the single-gene prognostic model's risk score effectively predicts patient overall survival, with AUC values of 0.739, 0.628, and 0.605 for 1-year, 3-year, and 5-year OS, respectively (Fig. 5I). Moreover, by analyzing the GEO database, ICGC database and CPTAC database, Kaplan-Meier survival curve analysis and differential expression analysis of the UQCRH gene in HCC at the transcriptomic and proteomic levels revealed that high expression of UQCRH is associated with poor OS prognosis and is overexpressed in tumor samples, as shown in Fig. 5J, K, L.

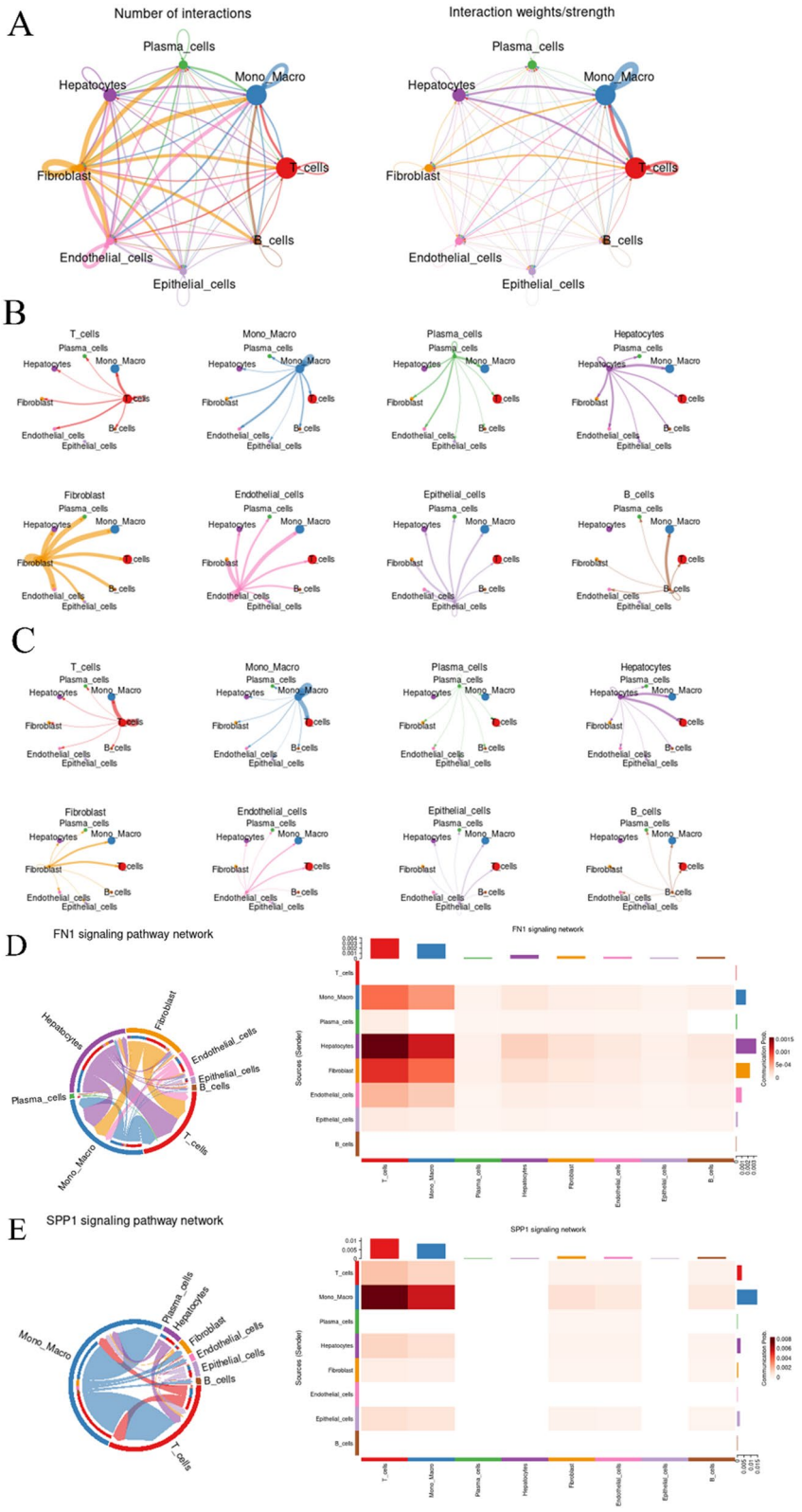

Fig. 3 Cell communication analysis. (**A**) Construction of a cell communication network for the 10 included tumor samples. (**B**) Independent display of the number of cell communication events between each cell type. (**C**) Independent display of the communication weights between each cell type. (**D**) In-depth analysis of the role of the FN1 signaling pathway in cell communication using chord diagrams and heatmaps. (**E**) Further analysis of the SPP1 signaling pathway in cell communication through chord diagrams and heatmaps

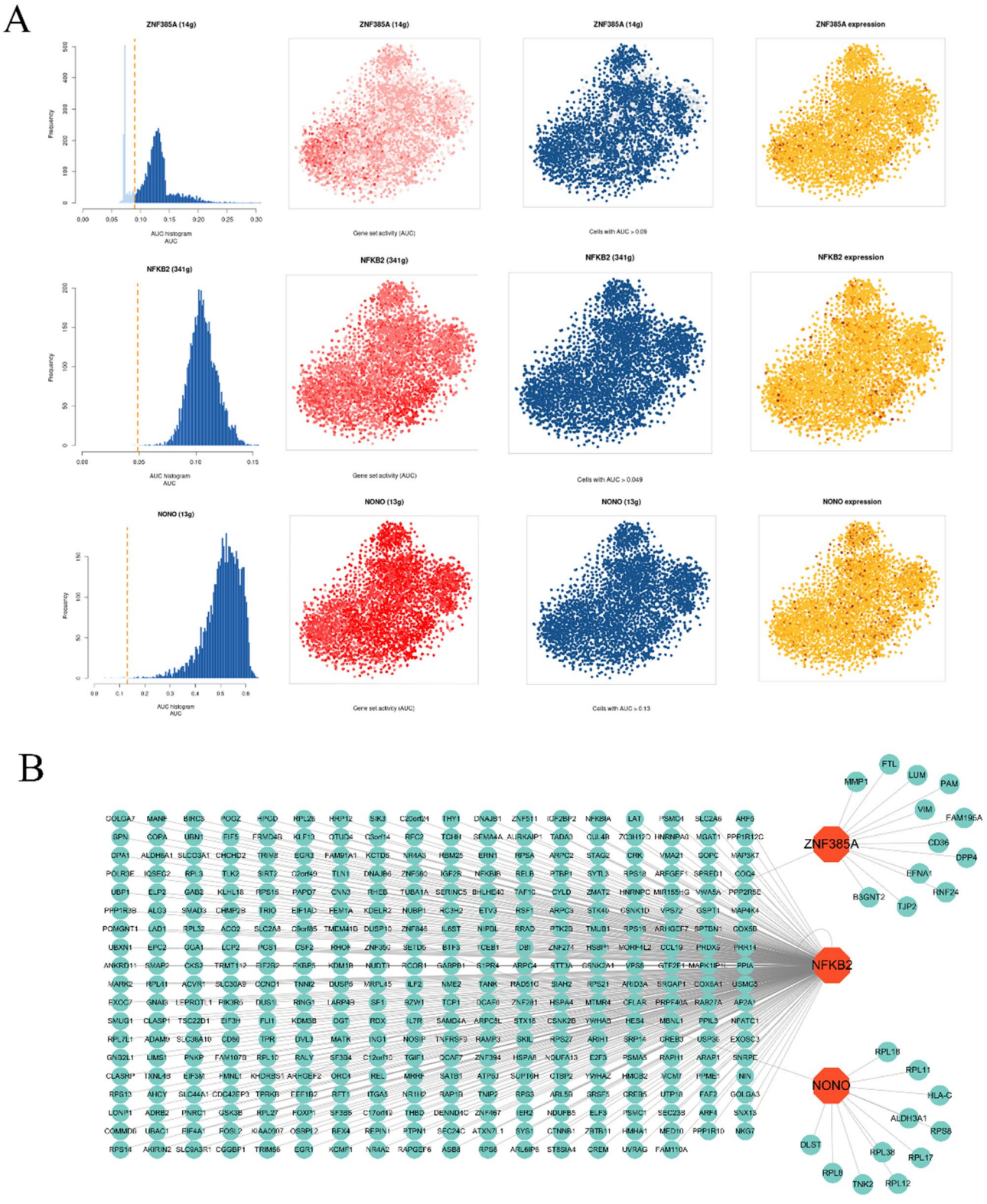

Fig. 4 Predicting transcription factor regulatory networks associated with monocytes/macrophages in scRNA-seq samples using SCENIC. (**A**) ZNF385A, NFKB2, and NONO (three core transcription factors) will become the central hubs controlling the LIHC state in monocytes/macrophages. (**B**) In-depth analysis of the transcription factor-gene regulatory network centered on ZNF385A, NFKB2, and NONO

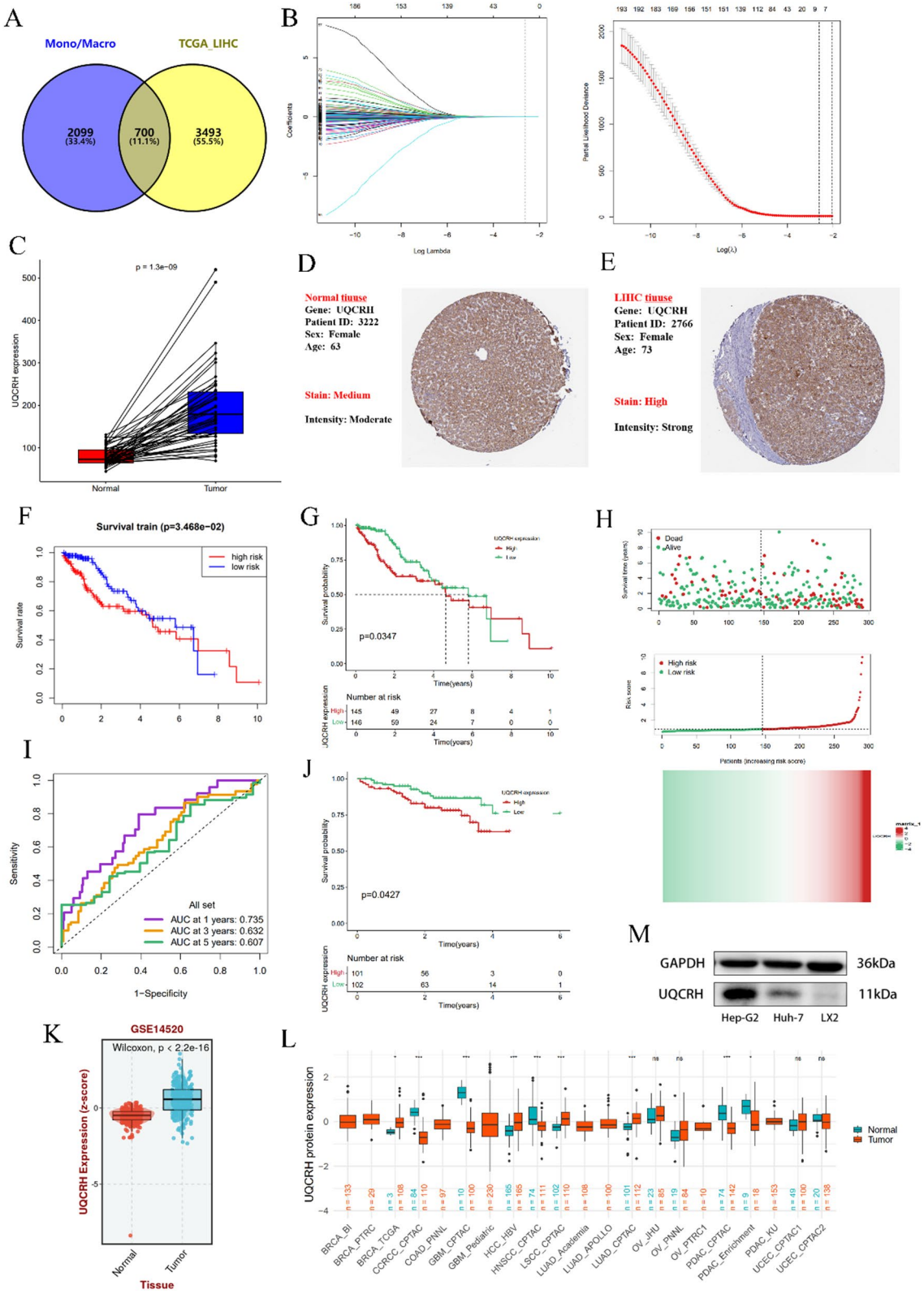

Fig. 5 (See legend on next page.)

(See figure on previous page.)

Fig. 5 Construction of a single-gene prognostic risk model and validation of UQCRH overexpression in HCC using external independent datasets and Western blot. (**A**) Intersection of monocyte/macrophage expressed genes from scRNA-seq with TCGA-LIHC expressed genes using a Venn diagram. (**B**) To prevent model overfitting, the prognostic risk model is further constructed based on LASSO regression analysis. (**C**) Differential expression analysis in paired samples from the TCGA-LIHC database revealed overexpression of UQCRH in tumor samples. (**D, E**) Analysis from the HPA database shows high expression of UQCRH in HCC samples. (**F-H**) Survival risk curves indicate that overexpression and high-risk groups in the single-gene prognostic model are associated with poor prognosis in HCC patients. (**I**) ROC analysis of the single-gene prognostic model risk shows good performance in predicting patient OS with an AUC score. (**J**) Transcriptomic analysis from the ICGC-HCC database confirms the poor prognosis associated with high UQCRH expression in patients ($p = 0.0427$). (**K**) Transcriptomic analysis from the GEO database(GSE14520) confirms overexpression of UQCRH in HCC tumor samples ($p < 2.2e\text{-}16$). (**L**) Proteomic analysis from the CPTAC database confirms overexpression of UQCRH in HCC tumor samples ($p < 0.001$). (**M**) Western blot analysis confirms the significantly elevated expression of UQCRH in HCC cell lines

UQCRH Expression Differs Significantly between Non-HCC and HCC Cells

Our Western blot results show that UQCRH is barely expressed in the non-HCC cell line LX2, but significantly upregulated in the HCC cell lines Hep-G2 and Huh-7 (Fig. 5M). The high expression in Hep-G2 and Huh-7, compared to LX2, suggests that UQCRH is specific to HCC cell lines, indicating its potential role in the development and progression of HCC.

Spatial Transcriptomic Analysis of UQCRH Expression in HCC

We performed cell annotation using scRNA-seq data of HCC from GSE166635 (Fig. 6A, B), followed by joint analysis and spatial mapping of two HCC spatial transcriptomics datasets obtained from the CROST database. Our analysis revealed that high expression of UQCRH co-localizes spatially with malignant cells within the HCC microenvironment (Fig. 6C-F). Additionally, high expression of UQCRH also co-localizes with monocyte/macrophages, particularly in HCC sample 2 (Supplementary Fig. 1).

Analysis of Prognostic Model and its Single Gene's Clinical Relevance

We first extracted clinical data from the TCGA-LIHC dataset, including patient age, gender, T, N, M, and Stage, to analyze the correlation between UQCRH expression and these clinical parameters. We found that high expression of UQCRH is positively correlated with tumor invasion and infiltration (especially T and Stage) (Fig. 7A). Next, based on the algorithm, age, gender, stage, T, N, M, and risk score were included in a nomogram (Fig. 7B). Then, the points of each variable were combined into a total score to predict the probability of OS at 1 year, 3 years, and 5 years (Fig. 7C), and the calibration curve shows that the nomogram's predictions are well aligned with the actual outcomes.

UQCRH-Related Protein Interaction and Enrichment Analysis

We used the STRING database to retrieve proteins directly associated with UQCRH and used Cytoscape to construct a protein interaction network, followed by enrichment analysis. We found that UQCRH is associated with proteins such as UQCRC1, UQCRC2, and UQCR11, which are involved in biological processes and signaling pathways including mitochondrial electron transport, cellular energy metabolism, and oxidative phosphorylation (Fig. 8A-C).

Drug Sensitivity, Immune Therapy, and Immune Function Analysis of the Prognostic Model

Using machine learning methods, we conducted drug sensitivity analysis on the single-gene prognostic model. Our analysis found that the high-risk group is less sensitive to anti-cancer drugs such as sorafenib and axitinib. (Fig. 9A). Additionally, in the immune therapy prediction analysis, we found that the high-risk group has higher TIDE scores, indicating poorer outcomes with immune therapy, whereas the low-risk group shows better effects (Fig. 9B). We then used the ssGSEA algorithm to score immune functions in the high and low-risk groups of the prognostic model, finding that the high-risk group may exhibit more active suppression of antigen presentation, while the low-risk group shows better IFN-II response (Fig. 9C).

Discussion

This study utilized a scRNA-seq dataset originally from Lu et al.'s research, with no overlap in the analysis focus compared to the original study [23]. This research analyzed scRNA-seq data from 10 primary HCC samples from a public dataset, aiming to identify clinically relevant tumor immune microenvironments and tumor characteristics. The analysis revealed a high frequency of interactions, particularly between monocytes/macrophages and other cell types such as hepatocytes and fibroblasts. These interactions are not merely incidental but indicative of the pivotal role monocytes/macrophages play in sculpting the TME. Monocytes/macrophages engage in signaling processes that contribute to tumor progression, immune evasion, and the establishment of an immunosuppressive microenvironment. This observation aligns with established knowledge in HCC pathogenesis, highlighting the prominent role of immune cells in facilitating cancer invasion and progression. Further insights were gleaned from the

interaction strength between these cell types. Strong communication signals were noted between immune cells and tumor cells, underscoring the deep involvement of monocytes/macrophages in altering the immune landscape. For instance, existing studies have found that FN1 is involved in cell adhesion and migration processes, and it is associated with promoting tumor invasion and poor prognosis [44]. Xu demonstrated through experiments that targeting the downregulation of FN1 significantly inhibits the invasion of HCC cells [45]. Liu considers SPP1 as an immune-related prognostic factor in HCC, mediated by SPP1-CD44 and SPP1-PTGER4 interactions between HCC cells and macrophages [46]. In vitro experiments have shown that SPP1 can trigger macrophage polarization towards M2-like tumor-associated macrophages, revealing its carcinogenic mechanism in HCC. Additionally, analysis of the transcription factor-gene regulatory network in monocytes/ macrophages of HCC found that ZNF385A, NFKB2, and NONO play significant gene regulatory roles. Peng's research found that Zinc finger protein 385 A (ZNF385A) is overexpressed in HCC and related to poor prognosis [47]. Hepatitis B virus infection can lead to overexpression of ZNF385A, accompanied by increased apoptosis and chronic inflammation. Moreover, they are positively correlated with immune-suppressive cells, inflammatory cytokines, immune checkpoint genes, and poor immunotherapy outcomes. Knockdown of ZNF385A expression inhibited tumor cell invasion and migration. NFKB2, a classic cancer-related transcription factor, encodes the p100 protein, which transforms into p52 in the non-canonical NF-κB signaling pathway, regulating cell survival, proliferation, and apoptosis. Although its role in HCC is not clearly defined, NFKB2 mutations and fusions are observed in other cancers like bone, skin, and cervical cancers [48]. Tang's study indicated that inhibiting the NF-κB signaling pathway, involving Akt and NFKB2, can suppress liver cancer cell proliferation, migration, and invasiveness [49]. Shen found that NONO plays a significant role in HCC, especially under hypoxic conditions, where NONO interacts with and stabilizes the HIF-1 and HIF-2 complexes, thereby activating the transcription of hypoxia-induced genes, allowing malignant proliferation of tumor cells and promoting the development of liver cancer [50]. However, it is important to note that the transcription factors identified in the SCENIC analysis primarily regulate gene expression related to monocytes/macrophages, and these factors may not directly regulate all genes involved in HCC progression. In contrast, the key prognostic markers in our study were identified through a broader multi-omics analysis, which integrated various data beyond specific networks. This broader analysis allowed us to specifically focus on the expression of monocyte/macrophage-related genes, as they play a key role in shaping the immune microenvironment of HCC.

Building on these observations, our study aims to explore the expression of monocyte/macrophage marker genes in HCC and construct a prognostic model for HCC. To identify differentially expressed genes in monocytes/macrophages in HCC, we utilized scRNA-seq data from the literature, selecting tissue biopsy samples from 10 primary HCC patients, encompassing data from 24,597 single cells. Based on these observations, our study aims to explore the expression of monocyte/macrophage marker genes in HCC and construct a prognostic model for HCC. To identify differentially expressed genes in monocytes/macrophages in HCC, we utilized scRNA-seq data from literature, selecting tissue biopsy samples from 10 primary HCC patients, encompassing data from 24,597 single cells. Based on this data, we developed a prognostic risk model for HCC targeting monocyte/macrophage marker genes on the TCGA database, further validated through data from the ICGC and CPTAC databases. Our analyses, including single-gene survival analysis, differential expression analysis, prognostic model risk scoring, and correlation analysis with clinical indicators, demonstrated that UQCRH is overexpressed in HCC tumor samples. Notably, although immunohistochemical analysis shows UQCRH expression in normal tissues, its expression is significantly higher in HCC, and its presence in normal tissues may be related to its basic function in mitochondrial metabolism. UQCRH and its prognostic model may play a role in promoting cancer progression in HCC by facilitating tumor cell invasion. Mitochondrial membrane complexes, such as Complexes I, II, III, and IV of the respiratory chain, and ATP synthase, mainly located on the inner mitochondrial membrane, are directly involved in the energy conversion process, generating ATP through oxidative phosphorylation. The function of mitochondrial membrane complexes is closely related to mitochondrial membrane dynamics and plays a central role in cell death, particularly programmed cell death, with increased permeability of the mitochondrial outer membrane allowing the release of apoptotic inducers like cytochrome c into the cytoplasm, a critical step in apoptosis [51, 52]. UQCRH, a vital component of mitochondrial respiratory chain Complex III, helps maintain the stability of the complex's structure and function, indirectly influencing the transfer of electrons from ubiquinone through Complex III to cytochrome c. This function is crucial in the energy metabolism and cellular respiration processes of normal cells but is often altered in expression and functionality in cancer cells [53–56]. Modena et al. found that UQCRH expression is reduced or even absent in some cancer cell lines, such as ovarian and breast cancer, possibly due to structural rearrangements or epigenetic

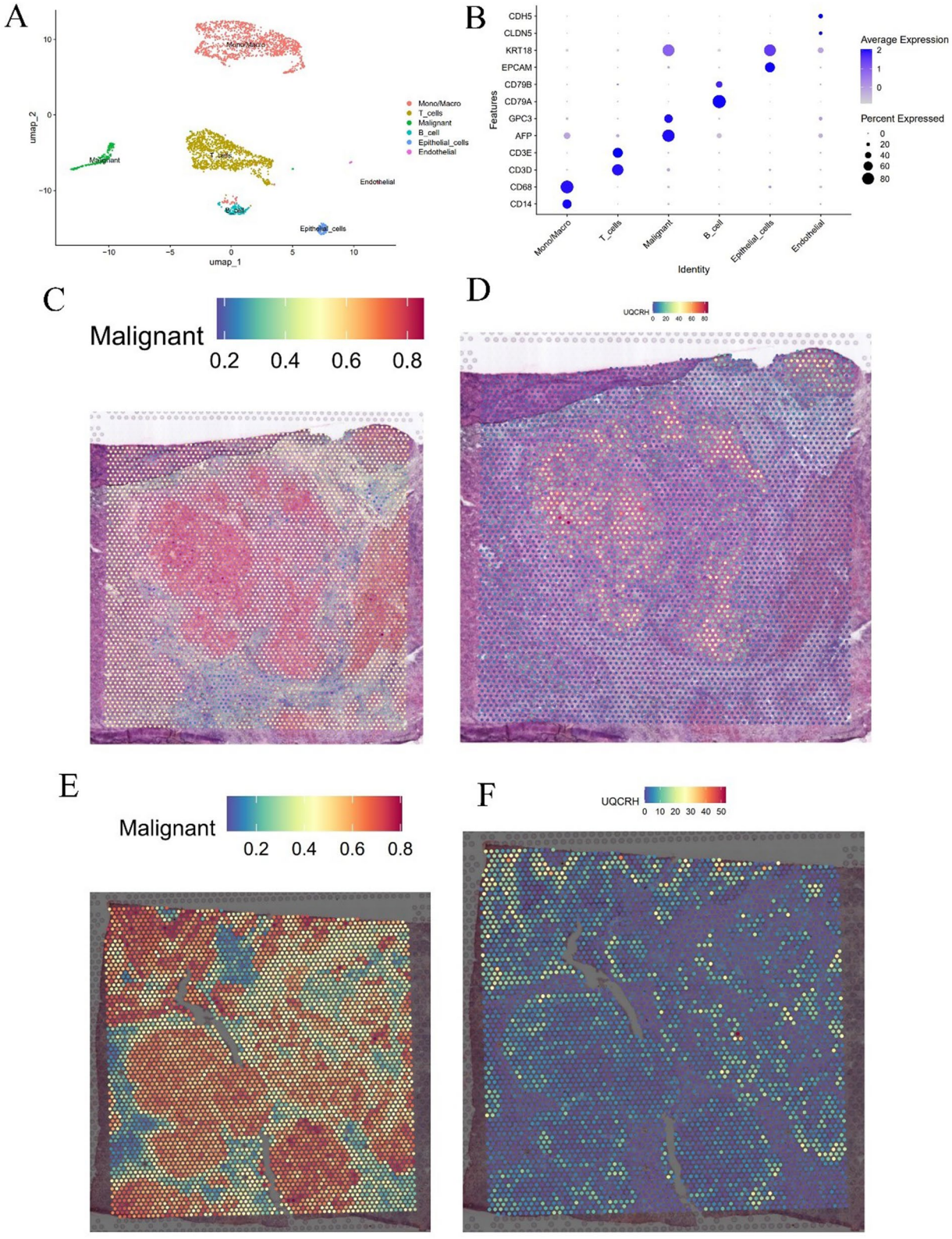

Fig. 6 Spatial analysis of UQCRH expression and malignant cell distribution in HCC samples. (**A**) UMAP clustering showing different cell types in HCC, including malignant cells, T cells, and others. (**B**) Dot plot of marker gene expression across cell clusters, highlighting markers for malignant cells. (**C**) Spatial distribution of malignant cells in HCC tissue, with redder colors indicating higher abundance. (**D**) Spatial expression of UQCRH in HCC sample 1, with redder colors representing higher expression levels. (**E**) Spatial distribution of malignant cells in HCC sample 2, similar to sample 1. (**F**) Spatial expression of UQCRH in HCC sample 2, showing co-localization with malignant cells

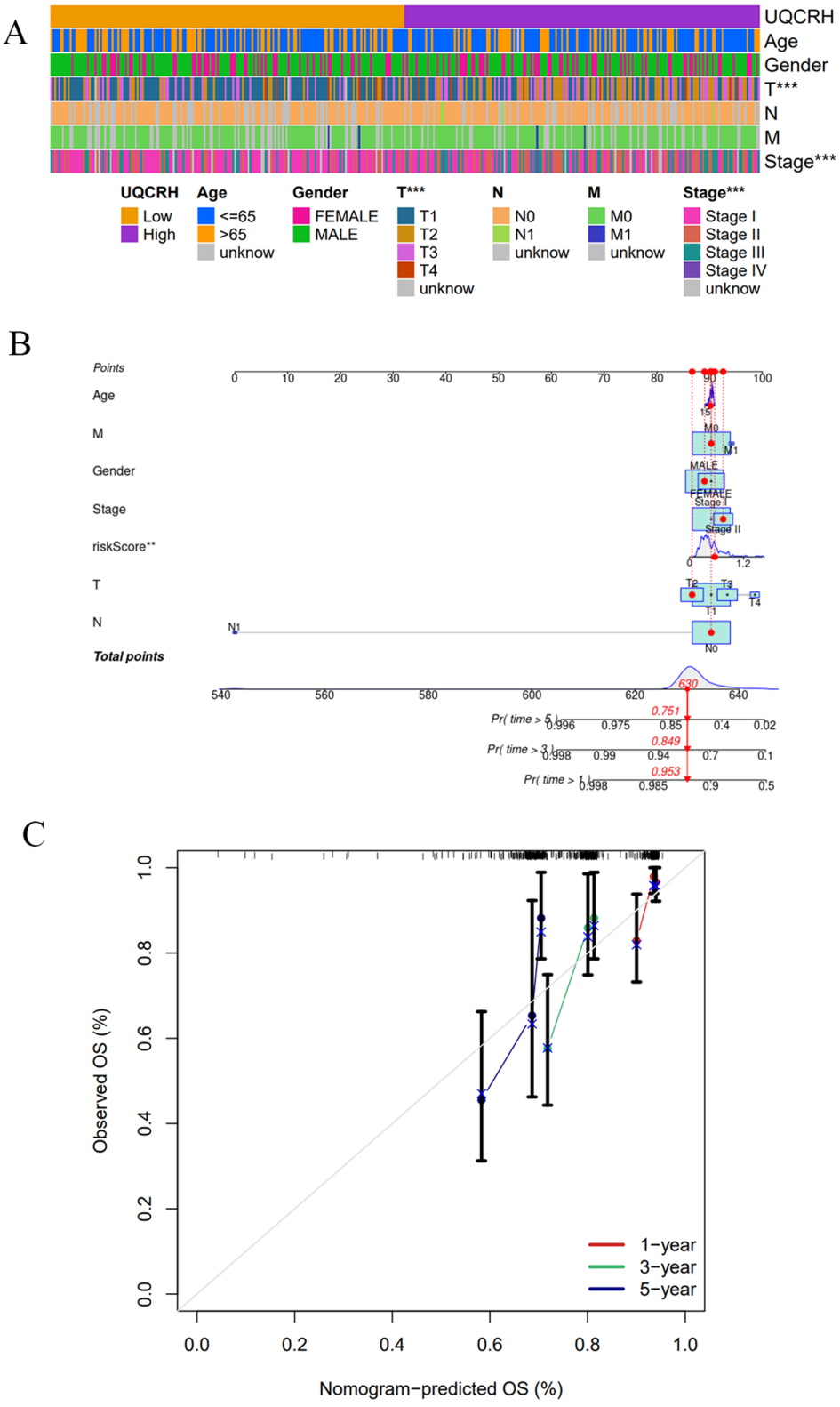

Fig. 7 Prognostic value of UQCRH expression and risk score in HCC. (**A**) In the TCGA-LIHC database, high expression of UQCRH is positively correlated with clinical indicators T and Stage ($p < 0.001$). (**B**) Nomogram analysis reveals that patient prognosis is primarily associated with Risk Score ($p < 0.01$). (**C**) Calibration curves show good agreement between the predicted values of the nomogram and actual values for 1-year, 3-year, and 5-year OS probabilities

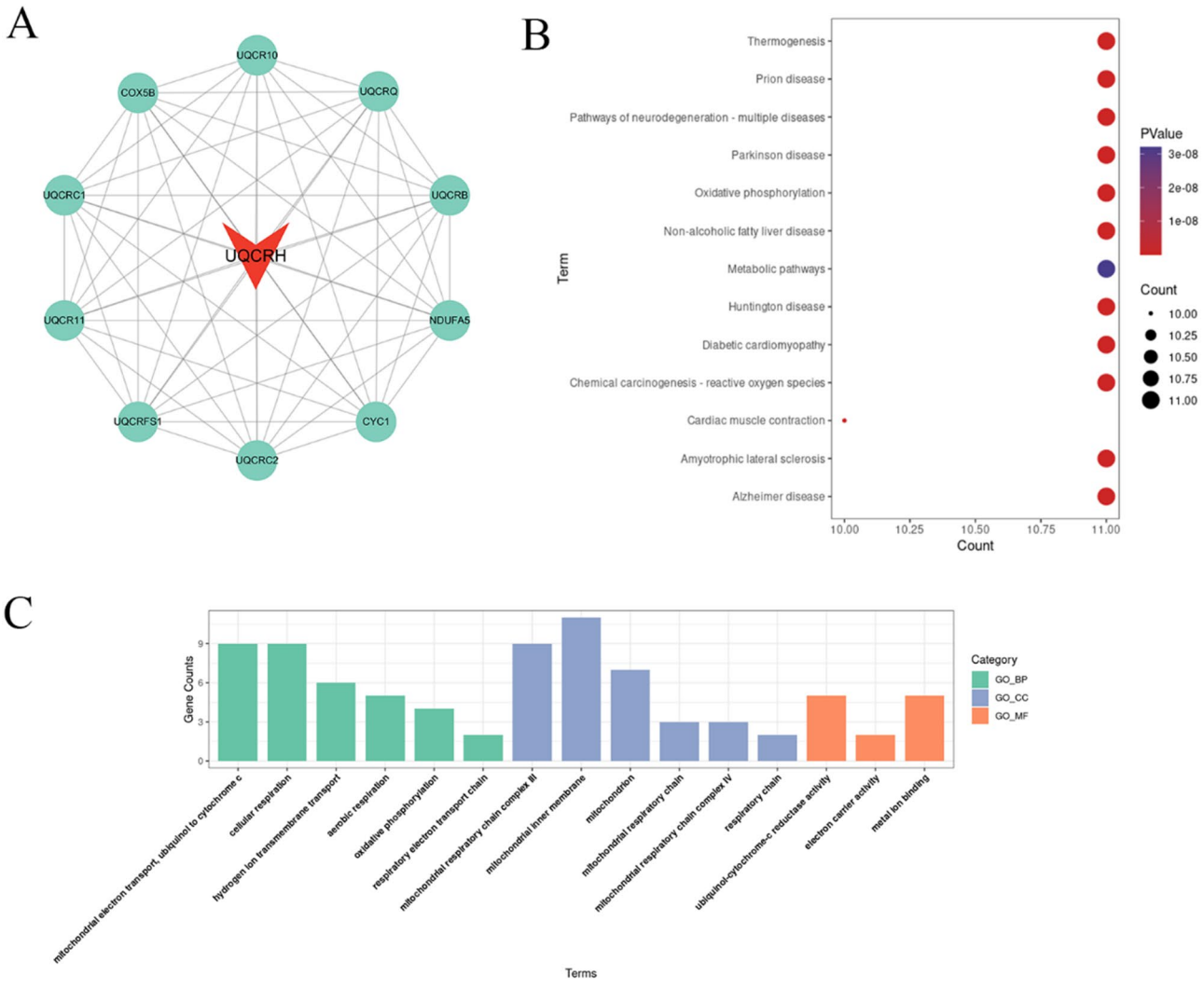

Fig. 8 Functional analysis of UQCRH and its associated proteins in HCC. (**A**) Protein-protein interaction analysis identifies direct associations between UQCRH and proteins such as UQCRB and UQCRQ. (**B, C**) KEGG and GO enrichment analyses show that UQCRH and its directly associated proteins are involved in biological processes and signaling pathways such as mitochondrial electron transport, maintenance of cellular energy metabolism, and oxidative phosphorylation

inactivation mechanisms like promoter methylation in these cancer cell lines [57]. Additionally, in clear cell renal cell carcinoma (ccRCC), UQCRH often exhibits DNA hypermethylation, leading to mRNA downregulation. This hypermethylation status is closely associated with poor clinical outcomes, indicating UQCRH's potential tumor-suppressive role in tumor progression. In ccRCC, UQCRH downregulation is associated with shorter survival times, and its overexpression can restore mitochondrial function, increase oxygen consumption, reduce the Warburg effect, and induce apoptosis in cancer cells [53, 54]. Conversely, in lung adenocarcinoma, UQCRH can induce the production of reactive oxygen species (ROS), and its expression in patients with lung adenocarcinoma is significantly higher than in patients with pneumonia and healthy controls, showing high sensitivity and specificity for diagnosing lung adenocarcinoma [55]. Eun-Ran Park found through clinical sample analysis that UQCRH is overexpressed in HCC patient samples and is associated with tumor size, poor differentiation, and vascular invasion. Overexpression of UQCRH in HCC patients significantly shortens overall and recurrence-free survival times, especially in patients with elevated alpha-fetoprotein levels, thus indicating that UQCRH could be an indicator of poor prognosis in HCC [56]. This finding aligns with the results from our bioinformatics analysis, further confirming UQCRH's oncogenic role in HCC, suggesting its potential as an important clinical prognostic marker.

Based on this understanding, in order to more comprehensively evaluate the guidance significance of UQCRH expression on therapeutic strategies, we conducted

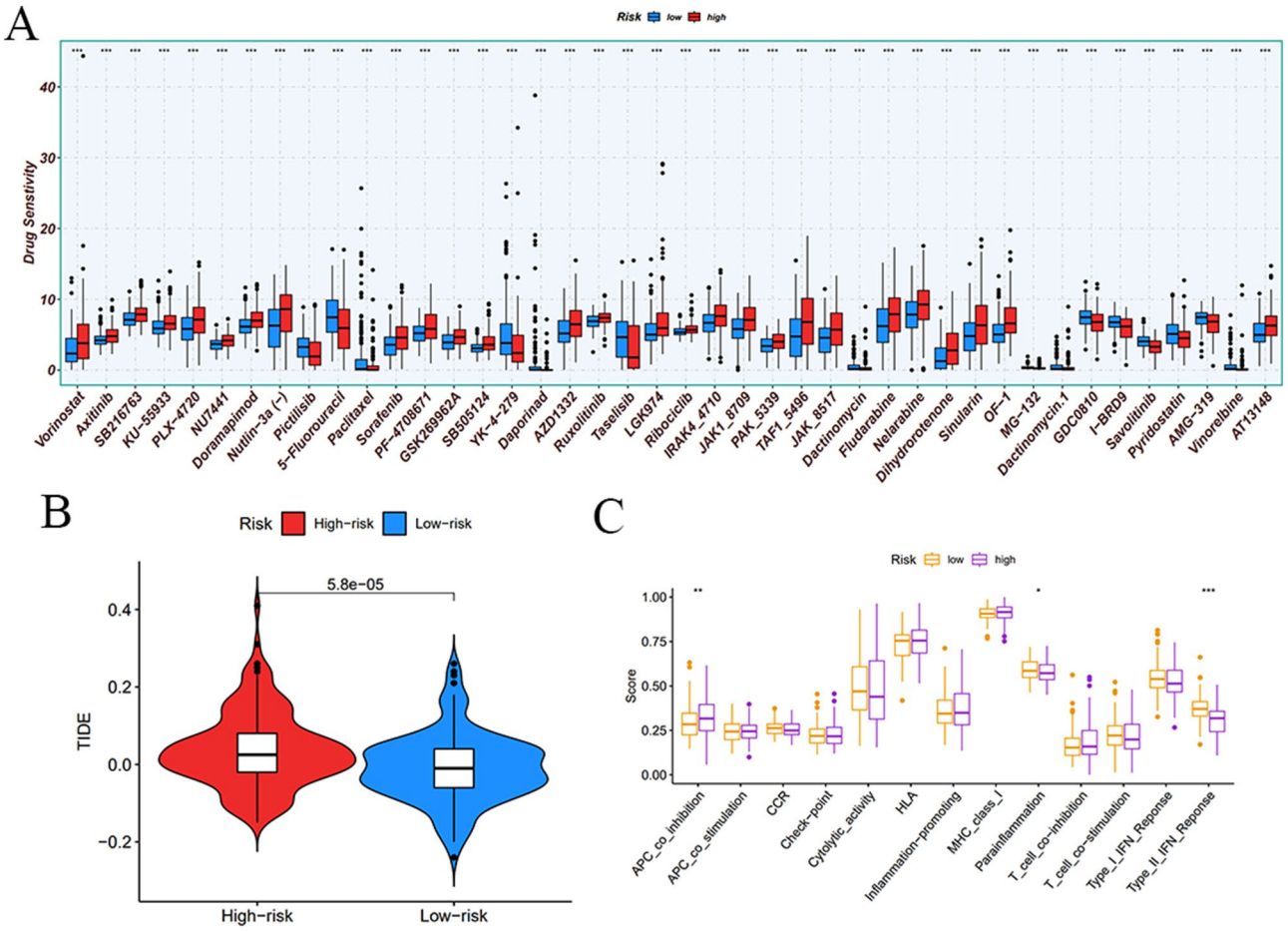

Fig. 9 Drug sensitivity, immune therapy response, and immune function analysis in high-risk and low-risk groups of HCC patients. (**A**) Drug sensitivity analysis indicates that the high-risk group in the prognostic model is less sensitive to commonly used anti-cancer drugs such as sorafenib and axitinib. (**B**) Immune therapy response prediction within the prognostic model shows that the high-risk group exhibits less pronounced responses to immune therapy. (**C**) ssGSEA immune function scoring reveals that the high-risk group in the prognostic model exhibits suppressed antigen presentation and poor response to IFN-II in terms of tumor immunity

further analyses on drug sensitivity, immune therapy responses, and immune function scoring in our prognostic model. These analyses helped us validate the biomarker potential of UQCRH from multiple perspectives and explore its value in precision medicine. In our prognostic model, the high-risk group exhibited insensitivity to various targeted therapy drugs, including sorafenib and axitinib. This may be due to mitochondrial dysfunction caused by abnormal expression of UQCRH, which affects the energy and drug metabolism of HCC tumor cells, leading to reduced intracellular drug accumulation, weakened drug efficacy, and the promotion of drug resistance in HCC tumor cells [58]. Additionally, in our prognostic model, patients in the high-risk group showed poorer efficacy in immune therapy, which is associated with suppressed antigen presentation responses and insensitivity to Type II interferon (IFN-γ) found in immune function analysis. Interestingly, existing literature indicates that IFN-γ is a critical factor determining the success of immune therapy, primarily through activating immune cells and promoting antigen presentation in tumor cells. However, tumor cells may hijack this pathway by disrupting the IFNG/IFNGR/JAK/STAT signaling pathway or preferentially expressing certain interferon-stimulated genes (ISGs), increasing resistance to IFN-γ and thus circumventing immune system attack [59]. Understanding this mechanism not only strengthens the biological foundation of our prognostic model but also highlights the key challenges that need to be overcome in immune therapy strategies. Further research into these resistance mechanisms will aid in the development of new treatment methods, potentially including inhibitors targeting specific signaling pathways or strategies to modulate ISG expression, to enhance the responsiveness of HCC patients to immune therapy.

Furthermore, we analyzed the spatial expression pattern of UQCRH in HCC using spatial transcriptomics. The results demonstrated that high expression of

UQCRH showed significant spatial co-localization with malignant cells, suggesting that this gene may play an important role in the tumor microenvironment of HCC. Through joint analysis of two different HCC samples, we found that UQCRH exhibited elevated expression in regions enriched with malignant cells, further supporting its association with tumor progression. These findings reveal the critical role of UQCRH in HCC, particularly in the spatial distribution within the malignant cell microenvironment, indicating that UQCRH may be involved in tumor initiation and progression. Future research should further explore the specific mechanisms of UQCRH in the tumor microenvironment, including its functional differences in TAMs and tumor cells, and how these differences influence HCC progression and therapeutic responses.

There are several limitations in this study. First, our study is based on data from public databases to construct the prognostic model, and further prospective clinical studies are needed to verify the authenticity and stability of this model. Second, although a prognostic model has been developed, its validation remains limited. So far, we have only performed a preliminary analysis of UQCRH expression in non-HCC and HCC cells using Western blot. Future experiments are needed to explore the mechanisms of action of this model on targeted drug resistance, immune therapy responses, and immune function. Additionally, the spatial transcriptomics analysis used publicly available datasets with a small sample size, and future studies should incorporate more spatial omics data for further validation. In constructing the prognostic model, we used Lasso regularization analysis to avoid overfitting, which is particularly suitable for high-dimensional data with small sample sizes. While Elastic Net provides a more balanced feature selection and coefficient shrinkage, Lasso was sufficient for our needs due to the limited multicollinearity in the data. Future studies may explore Elastic Net as a complementary method to further validate the robustness of the results. Moreover, although the UQCRH single-gene prognostic model demonstrated good performance for 1-year survival prediction (AUC = 0.735), its predictive performance for 3-year and 5-year survival (AUC = 0.632 and 0.607, respectively) still requires improvement. To enhance predictive performance, we integrated UQCRH risk scores with key clinical variables and constructed a comprehensive nomogram, which helps improve the predictive accuracy of the model to some extent.

Conclusion

Through scRNA-seq, we constructed a comprehensive single-cell transcriptomic atlas of HCC, revealing the heterogeneity of monocytes and macrophages in intercellular communication and transcription factor regulatory networks. By integrating scRNA-seq with transcriptomic analysis, we identified prognostic features for predicting overall survival in HCC patients. Spatial transcriptomics further revealed UQCRH's spatial expression patterns in HCC, showing significant co-localization with malignant cells, suggesting its role in the tumor microenvironment and progression. Our findings demonstrate that the UQCRH-centered prognostic risk model provides independent prognostic value for survival prediction in HCC. This research also advances our understanding of targeted drug resistance, immune therapy responses, and immune function in HCC, warranting further validation through clinical and basic research.

Abbreviations

HCC	Hepatocellular carcinoma
UQCRH	Ubiquinol-cytochrome c reductase hinge protein
scRNA-seq	Single-cell RNA sequencing
OS	Overall survival
GEO	Gene Expression Omnibus
TCGA	The Cancer Genome Atlas
ICGC	International Cancer Genome Consortium
CPTAC	Clinical Proteomic Tumor Analysis Consortium
ML	Machine learning
APC	Antigen-Presenting Cells
IFN-II	Interferon-Gamma
PCA	Principal Component Analysis
t-SNE	t-Distributed Stochastic Neighbor Embedding
HCG	Highly Variable Genes
ZNF385A	Zinc Finger Protein 385 A
NONO	Non-POU Domain-Containing Octamer-Binding Protein
NF-κB2	Nuclear Factor Kappa B Subunit 2

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12575-025-00270-9 .

Supplementary Material 1

Supplementary Material 2

Supplementary Material 3

Supplementary Material 4

Supplementary Material 5

Supplementary Material 6

Supplementary Material 7

Acknowledgements

We thank the the Gene Expression Omnibus (GEO), the Cancer Genome Atlas (TCGA) Database and the National Genomics Data Center database for sharing a large amount of data.

Author Contributions

Concept and design: SM. Manuscript writing: XL and YC. Technical support: QC, QT, GX, LY, XY and HD; Manuscript revision and editing: SM and XB. Obtained funding: SM. Supervision: SM.

Funding

This study was supported by grants from the Research Fund for Bajian Talents of Guangdong Provincial Hospital of Chinese Medicine (BJ2022KY13), the Guangzhou science and technology plan project (2024A03J0548, 202201010863), State Key Laboratory of Traditional Chinese Medicine Syndrome (QZ2023ZZ47), the Key project of State Key Laboratory of

dampness syndrome of Chinese medicine jointly built by province and Ministry (SZ2021ZZ38), the Guangdong Provincial Key Laboratory of Chinese Medicine for Prevention and Treatment of Refractory Chronic Diseases (2023KT15521), and the Foundation for basic and applied research of Guangdong Province (2017B030314166).

Data Availability
Data is provided within the manuscript and supplementary information files.

Declarations

Ethics Approval and Consent to Participate
Not applicable.

Consent for Publication
Not applicable.

Competing Interests
The authors declare no competing interests.

Author details
[1]Clinical and Basic Research Team of TCM Prevention and Treatment of NSCLC, Department of Oncology, The Second Clinical College of Guangzhou University of Chinese Medicine, Chinese Medicine Guangdong Laboratory, Guangdong Provincial Key Laboratory of Clinical Research on Traditional Chinese Medicine Syndrome, State Key Laboratory of Dampness Syndrome of Chinese Medicine, Guangzhou University of Chinese Medicine, Guangdong Provincial Hospital of Chinese Medicine, Guangzhou, Guangdong 510120, China
[2]The Second Affiliated Hospital of Guangzhou University of Chinese Medicine, Guangdong Provincial Hospital of Chinese Medicine, Guangzhou University of Chinese Medicine, Chinese Medicine Guangdong Laboratory, 111 Dade Rd, Guangzhou, Guangdong Province 510120, China

Received: 10 December 2024 / Accepted: 13 February 2025
Published online: 10 March 2025

References
1. Dhanasekaran R, Limaye A, Cabrera R. Hepatocellular carcinoma: current trends in worldwide epidemiology, risk factors, diagnosis, and therapeutics. Hepat Med. 2012;8:19–37. https://doi.org/10.2147/HMER.S16316.
2. Farasati Far B, Rabie D, Hemati P, Fooladpanjeh P, Faal Hamedanchi N, Broomand Lomer N, Karimi Rouzbahani A, Naimi-Jamal MR. Unresectable hepatocellular carcinoma: A review of new advances with focus on targeted therapy and immunotherapy. Livers. 2023;3:121–60. https://doi.org/10.3390/livers3010011.
3. Damaskos C, Garmpis N, Dimitroulis D, Garmpi A, Psilopatis I, Sarantis P, Koustas E, Kanavidis P, Prevezanos D, Kouraklis G, et al. Targeted therapies for hepatocellular carcinoma treatment: A new era Ahead—A systematic review. Int J Mol Sci. 2022;23:14117. https://doi.org/10.3390/ijms232214117.
4. Kelley C-D, Xie C, Hrones DM, Ghabra S, Greten TF, Monge C. The current landscape of therapies for hepatocellular carcinoma, carcinogenesis, 44, issue 7, July 2023, Pages 537–48, https://doi.org/10.1093/carcin/bgad052
5. Tabrizian P, Holzner ML, Zaret D, Meyerovich G, Fagenson A, Schiano T. Liver transplantation and hepatocellular carcinoma 2023: a narrative review of management and outcomes. Ann Palliat Med. 2024;13(1):126–40. https://doi.org/10.21037/apm-23-341.
6. Huang J, Wu Q, Geller DA, Yan Y. Macrophage metabolism, phenotype, function, and therapy in hepatocellular carcinoma (HCC). J Transl Med. 2023;21(1):815. https://doi.org/10.1186/s12967-023-04716-0.
7. Garcia-Alvarez A, Hernando J, Carmona-Alonso A, Capdevila J. What is the status of immunotherapy in thyroid neoplasms? Front Endocrinol. 2022;13:929091. https://doi.org/10.3389/fendo.2022.929091.
8. Russo M, Nastasi C. Targeting the tumor microenvironment: a close up of tumor-associated macrophages and neutrophils. Front Oncol. 2022;12:871513. https://doi.org/10.3389/fonc.2022.871513.
9. Wang S, Liu G, Li Y, Pan Y. Metabolic reprogramming induces macrophage polarization in the tumor microenvironment. Front Immunol. 2022;13:840029. https://doi.org/10.3389/fimmu.2022.840029.
10. Hasan MN, Capuk O, Patel SM, Sun D. The role of metabolic plasticity of tumor-associated macrophages in shaping the tumor microenvironment immunity. Cancers. 2022;14:3331. https://doi.org/10.3390/cancers14143331.
11. Li D, Li K, Zhang W, Yang KW, Mu DA, Jiang GJ, Shi RS, Ke D. The m6A/m5C/m1A regulated gene signature predicts the prognosis and correlates with the immune status of hepatocellular carcinoma. Front Immunol. 2022;13:918140. https://doi.org/10.3389/fimmu.2022.918140.
12. Liang JY, Wang DS, Lin HC, Chen XX, Yang H, Zheng Y, Li YH. A novel Ferroptosis-related gene signature for overall survival prediction in patients with hepatocellular carcinoma. Int J Biol Sci. 2020;16(13):2430–41. https://doi.org/10.7150/ijbs.45050.
13. Hu B, Ma X, Fu P, Sun Q, Tang W, Sun H, Yang Z, Yu M, Zhou J, Fan J, Xu Y. The mRNA-miRNA-lncRNA regulatory network and factors associated with prognosis prediction of hepatocellular carcinoma. Genomics Proteom Bioinf. 2021;19(6):913–25. https://doi.org/10.1016/j.gpb.2021.03.001.
14. Ziegenhain C, Vieth B, Parekh S, et al. Comparative analysis of Single-Cell RNA sequencing methods. Mol Cell. 2017;65(4):631–e6434. https://doi.org/10.1016/j.molcel.2017.01.023.
15. Kong SL, Li H, Tai JA, et al. Concurrent Single-Cell RNA and targeted DNA sequencing on an automated platform for comeasurement of genomic and transcriptomic signatures. Clin Chem. 2019;65(2):272–81. https://doi.org/10.1373/clinchem.2018.295717.
16. Li Y, Wang J, Wang F, Gao C, Cao Y, Wang J. Identification of specific cell subpopulations and marker genes in ovarian Cancer using Single-Cell RNA sequencing. Biomed Res Int. 2021. https://doi.org/10.1155/2021/1005793. 2021:1005793. Published 2021 Oct 7.
17. Papalexi E, Satija R. Single-cell RNA sequencing to explore immune cell heterogeneity. Nat Rev Immunol. 2018;18(1):35–45. https://doi.org/10.1038/nri.2017.76.
18. Qi R, Ma A, Ma Q, Zou Q. Clustering and classification methods for single-cell RNA-sequencing data. Brief Bioinform. 2020;21(4):1196–208. https://doi.org/10.1093/bib/bbz062.
19. Zhang Z, Cui F, Wang C, Zhao L, Zou Q. Goals and approaches for each processing step for single-cell RNA sequencing data. Brief Bioinform. 2021;22(4):bbaa314. https://doi.org/10.1093/bib/bbaa314.
20. Zhao M, He W, Tang J, Zou Q, Guo F. A hybrid deep learning framework for gene regulatory network inference from single-cell transcriptomic data. Brief Bioinform. 2022;23(2):bbab568. https://doi.org/10.1093/bib/bbab568.
21. Wang J, Zou Q, Lin C. Brief Bioinform. A comparison of deep learning-based pre-processing and clustering approaches for single-cell RNA sequencing data. 2022;23(1):bbab345. https://doi.org/10.1093/bib/bbab345.
22. Zhang B, Shi H, Wang H. Machine learning and AI in Cancer prognosis, prediction, and treatment selection: A critical approach. J Multidiscip Healthc. 2023;16:1779–91. https://doi.org/10.2147/JMDH.S410301. Published 2023 Jun 26.
23. Lu Y, Yang A, Quan C, et al. A single-cell atlas of the multicellular ecosystem of primary and metastatic hepatocellular carcinoma. Nat Commun. 2022;13(1):4594. https://doi.org/10.1038/s41467-022-32283-3. Published 2022 Aug 6.
24. Giorgi FM, Ceraolo C, Mercatelli D, The R, Language. Life (Basel). An Engine for Bioinformatics and Data Science. 2022;12(5):648. Published 2022 Apr 27. https://doi.org/10.3390/life12050648.
25. Stuart T, Butler A, Hoffman P, et al. Comprehensive integration of Single-Cell data. Cell. 2019;177(7):1888–902. e21.
26. Zhou B, Jin W. Visualization of single cell RNA-Seq data using t-SNE. Methods Mol Biol. 2020;2117:159–67. https://doi.org/10.1007/978-1-0716-0301-7_8.
27. Stuart T, Satija R. Integrative single-cell analysis. Nat Rev Genet. 2019;20(5):257–72. https://doi.org/10.1038/s41576-019-0093-7.
28. Hu C, Li T, Xu Y, et al. CellMarker 2.0: an updated database of manually curated cell markers in human/mouse and web tools based on scRNA-seq data. Nucleic Acids Res. 2023;51(D1):D870–6. https://doi.org/10.1093/nar/gkac947.
29. Jin S, Guerrero-Juarez CF, Zhang L, et al. Inference and analysis of cell-cell communication using cellchat. Nat Commun. 2021;12(1):1088. https://doi.org/10.1038/s41467-021-21246-9. Published 2021 Feb 17.
30. Aibar S, González-Blas CB, Moerman T, et al. SCENIC: single-cell regulatory network inference and clustering. Nat Methods. 2017;14(11):1083–6. https://doi.org/10.1038/nmeth.4463.

31. Tomczak K, Czerwińska P, Wiznerowicz M. The Cancer genome atlas (TCGA): an immeasurable source of knowledge. Contemp Oncol (Pozn). 2015;19(1A):A68–77. https://doi.org/10.5114/wo.2014.47136.
32. Zhang J, Bajari R, Andric D, et al. The international cancer genome consortium data portal. Nat Biotechnol. 2019;37(4):367–9. https://doi.org/10.1038/s41587-019-0055-9.
33. Li Y, Dou Y, Da Veiga Leprevost F, et al. Proteogenomic data and resources for pan-cancer analysis. Cancer Cell. 2023;41(8):1397–406. https://doi.org/10.1016/j.ccell.2023.06.009.
34. van Dijk PC, Jager KJ, Zwinderman AH, Zoccali C, Dekker FW. The analysis of survival data in nephrology: basic concepts and methods of Cox regression. Kidney Int. 2008;74(6):705–9. https://doi.org/10.1038/ki.2008.294.
35. Tibshirani R. The Lasso method for variable selection in the Cox model. Stat Med. 1997;16(4):385–95. https://doi.org/10.1002/(sici)1097-0258(19970228)16:4%3C385::aid-sim380%3E3.0.co;2-3.
36. Bradburn MJ, Clark TG, Love SB, Altman DG. Survival analysis part II: multivariate data analysis–an introduction to concepts and methods. Br J Cancer. 2003;89(3):431–6. https://doi.org/10.1038/sj.bjc.6601119.
37. Christensen E. Multivariate survival analysis using Cox's regression model. Hepatology. 1987;7(6):1346–58. https://doi.org/10.1002/hep.1840070628.
38. Fang Y, Huang S, Han L, Wang S, Xiong B. Comprehensive analysis of peritoneal metastasis sequencing data to identify LINC00924 as a prognostic biomarker in gastric Cancer. Cancer Manag Res. 2021;13:5599–611. https://doi.org/10.2147/CMAR.S318704. Published 2021 Jul 12.
39. Wu J, Zhang H, Li L, et al. A nomogram for predicting overall survival in patients with low-grade endometrial stromal sarcoma: A population-based analysis. Cancer Commun (Lond). 2020;40(7):301–12. https://doi.org/10.1002/cac2.12067.
40. Yang W, Soares J, Greninger P, et al. Genomics of drug sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. Nucleic Acids Res. 2013;41(Database issue):D955–61. https://doi.org/10.1093/nar/gks1111.
41. Wang Q, Li M, Yang M, et al. Analysis of immune-related signatures of lung adenocarcinoma identified two distinct subtypes: implications for immune checkpoint Blockade therapy. Aging. 2020;12(4):3312–39. https://doi.org/10.18632/aging.102814.
42. Uhlén M, Fagerberg L, Hallström BM, et al. Proteomics. Tissue-based map of the human proteome. Science. 2015;347(6220):1260419. https://doi.org/10.1126/science.1260419.
43. Szklarczyk D, Kirsch R, Koutrouli M, et al. The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. Nucleic Acids Res. 2023;51(D1):D638–46. https://doi.org/10.1093/nar/gkac1000.
44. Wang H, Zhang J, Li H, et al. FN1 is a prognostic biomarker and correlated with immune infiltrates in gastric cancers. Front Oncol. 2022;12:918719. https://doi.org/10.3389/fonc.2022.918719. Published 2022 Aug 23.
45. Xu X, Liu Z, Zhou L, Xie H, Cheng J, Ling Q, Wang J, Guo H, Wei X, Zheng S. Characterization of genome-wide TFCP2 targets in hepatocellular carcinoma: implication of targets FN1 and TJP1 in metastasis. J Exp Clin Cancer Res. 2015;34(1):6. https://doi.org/10.1186/s13046-015-0121-1.
46. Liu L, Zhang R, Deng J, Dai X, Zhu X, Fu Q, Zhang H, Tong Z, Zhao P, Fang W, Zheng Y, Bao X. Construction of TME and identification of crosstalk between malignant cells and macrophages by SPP1 in hepatocellular carcinoma. Cancer Immunol Immunother. 2022;71(1):121–36. https://doi.org/10.1007/s00262-021-02967-8.
47. Peng Q, Li J, Wu Q, et al. ZNF385A and ZNF346 serve as prognostic biomarkers associated with an inflamed immunosuppressive tumor microenvironment in hepatocellular carcinoma. Int J Mol Sci. 2023;24(4):3155. https://doi.org/10.3390/ijms24043155. Published 2023 Feb 5.
48. AACR Project GENIE Consortium. AACR Project GENIE: Powering Precision Medicine through an International Consortium. Cancer Discov. 2017;7(8):818–31. https://doi.org/10.1158/2159-8290.CD-17-0151.
49. Tang Y, Lv P, Sun Z, Han L, Zhou W. 14-3-3β promotes migration and invasion of human hepatocellular carcinoma cells by modulating expression of MMP2 and MMP9 through PI3K/Akt/NF-κB pathway. PLoS ONE. 2016;11(1):e0146070. https://doi.org/10.1371/journal.pone.0146070. Published 2016 Jan 5.
50. Shen M, Zhang R, Jia W, et al. Nuclear scaffold protein p54nrb/NONO facilitates the hypoxia-enhanced progression of hepatocellular carcinoma. Oncogene. 2021;40(24):4167–83. https://doi.org/10.1038/s41388-021-01848-9.
51. Tait SW, Green DR. Mitochondria and cell death: outer membrane permeabilization and beyond. Nat Rev Mol Cell Biol. 2010;11(9):621–32. https://doi.org/10.1038/nrm2952.
52. Giacomello M, Pyakurel A, Glytsou C, Scorrano L. The cell biology of mitochondrial membrane dynamics. Nat Rev Mol Cell Biol. 2020;21(4):204–24. https://doi.org/10.1038/s41580-020-0210-7.
53. Luo Y, Medina Bengtsson L, Wang X, et al. UQCRH downregulation promotes Warburg effect in renal cell carcinoma cells. Sci Rep. 2020;10(1):15021. https://doi.org/10.1038/s41598-020-72107-2. Published 2020 Sep 14.
54. Liu WS, Liu YD, Fu Q, et al. Prognostic significance of ubiquinol-cytochrome C reductase hinge protein expression in patients with clear cell renal cell carcinoma. Am J Cancer Res. 2016;6(4):797–805.
55. Gao F, Liu Q, Li G, et al. Identification of ubiquinol cytochrome C reductase hinge (UQCRH) as a potential diagnostic biomarker for lung adenocarcinoma. Open Biol. 2016;6(6):150256. https://doi.org/10.1098/rsob.150256.
56. Park ER, Kim SB, Lee JS, Kim YH, Lee DH, Cho EH, Park SH, Han CJ, Kim BY, Choi DW, Yoo YD, Yu A, Lee JW, Jang JJ, Park YN, Suh KS, Lee KH. The mitochondrial hinge protein, UQCRH, is a novel prognostic factor for hepatocellular carcinoma. Cancer Med. 2017;6(4):749–60. https://doi.org/10.1002/cam4.1042.
57. Modena P, Testi MA, Facchinetti F, et al. UQCRH gene encoding mitochondrial hinge protein is interrupted by a translocation in a soft-tissue sarcoma and epigenetically inactivated in some cancer cell lines. Oncogene. 2003;22(29):4586–93. https://doi.org/10.1038/sj.onc.1206472.
58. Jin P, Jiang J, Zhou L, et al. Mitochondrial adaptation in cancer drug resistance: prevalence, mechanisms, and management. J Hematol Oncol. 2022;15(1):97. https://doi.org/10.1186/s13045-022-01313-4. Published 2022 Jul 18.
59. Han J, Wu M, Liu Z. Dysregulation in IFN-γ signaling and response: the barricade to tumor immunotherapy. Front Immunol. 2023;14:1190333. https://doi.org/10.3389/fimmu.2023.1190333.

Publisher's Note