



# Identification of Seven-Gene Hypoxia Signature for Predicting Overall Survival of Hepatocellular Carcinoma

Yuping Bai<sup>1,3†</sup>, Wenbo Qi<sup>2,3†</sup>, Le Liu<sup>2,3</sup>, Jing Zhang<sup>1</sup>, Lan Pang<sup>1</sup>, Tiejun Gan<sup>1</sup>, Pengfei Wang<sup>1</sup>, Chen Wang<sup>2\*</sup> and Hao Chen<sup>2,3\*</sup>

## OPEN ACCESS

### Edited by:

Shicheng Guo,  
University of Wisconsin-Madison,  
United States

### Reviewed by:

Chenhao Zhou,  
Fudan University, China  
Qin An,  
Salk Institute for Biological Studies,  
United States  
Haitao Zhao,  
Peking Union Medical College  
Hospital (CAMS), China

### \*Correspondence:

Chen Wang  
wcdjjack@163.com  
Hao Chen  
ery\_chenh@zju.edu.cn;  
chenhaodrs@163.com

† These authors have contributed  
equally to this work and share first  
authorship

### Specialty section:

This article was submitted to  
Cancer Genetics,  
a section of the journal  
Frontiers in Genetics

Received: 03 December 2020

Accepted: 15 March 2021

Published: 09 April 2021

### Citation:

Bai Y, Qi W, Liu L, Zhang J,  
Pang L, Gan T, Wang P, Wang C and  
Chen H (2021) Identification  
of Seven-Gene Hypoxia Signature for  
Predicting Overall Survival  
of Hepatocellular Carcinoma.  
Front. Genet. 12:637418.  
doi: 10.3389/fgene.2021.637418

<sup>1</sup> Department of MR, Lanzhou University Second Hospital, Lanzhou, China, <sup>2</sup> Department of Surgical Oncology, Lanzhou University Second Hospital, Lanzhou, China, <sup>3</sup> The Key Laboratory of the Digestive System Tumors of Gansu Province, Second Hospital of Lanzhou University, Lanzhou, China

**Background:** Hepatocellular carcinoma (HCC) is ranked fifth among the most common cancer worldwide. Hypoxia can induce tumor growth, but the relationship with HCC prognosis remains unclear. Our study aims to construct a hypoxia-related multigene model to predict the prognosis of HCC.

**Methods:** RNA-seq expression data and related clinical information were download from TCGA database and ICGC database, respectively. Univariate/multivariate Cox regression analysis was used to construct prognostic models. KM curve analysis, and ROC curve were used to evaluate the prognostic models, which were further verified in the clinical traits and ICGC database. GSEA analyzed pathway enrichment in high-risk groups. Nomogram was constructed to predict the personalized treatment of patients. Finally, real-time fluorescence quantitative PCR (RT-qPCR) was used to detect the expressions of KDELR3 and SCARB1 in normal hepatocytes and 4 HCC cells. The expressions of SCARB1 in hepatocellular carcinoma tissue in 46 patients were detected by immunohistochemistry, and the correlation between its expressions and disease free survival of patient was calculated.

**Results:** Through a series of analyses, seven prognostic markers related to HCC survival were constructed. HCC patients were divided into the high and low risk group, and the results of KM curve showed that there was a significant difference between the two groups. Stratified analysis, found that there were significant differences in risk values of different ages, genders, stages and grades, which could be used as independent predictors. In addition, we assessed the risk value in the clinical traits analysis and found that it could accelerate the progression of cancer, while the results of GSEA enrichment analysis showed that the high-risk group patients were mainly distributed in the cell cycle and other pathways. Then, Nomogram was constructed to predict the overall survival of patients. Finally, RT-qPCR showed that KDELR3 and SCARB1 were highly expressed in HepG2 and L02, respectively. Results of IHC staining showed that SCARB1 was highly expressed in cancer tissues compared to adjacent normal liver tissues and its

expression was related to hepatocellular carcinoma differentiation status. The Kaplan-Meier survival showed a poor percent survival in the SCARB1 high group compared to that in the SCARB1 low group.

**Conclusion:** This study provides a potential diagnostic indicator for HCC patients, and help clinicians to deepen the comprehension in HCC pathogenesis so as to make personalized medical decisions.

**Keywords:** HCC, hypoxia, gene signature, prognosis, bioinformatics

## INTRODUCTION

Hepatocellular carcinoma (HCC), characterized by high morbidity and mortality, poses a major challenge to global public health (El-Serag, 2011). At the time of diagnosis, most patients have lost the opportunity for curative treatment, including transplantation, resection or ablation. In addition, due to the high recurrence rate, patients receiving potential treatment still have a poor prognosis (Yang et al., 2019). If patients are at higher risk of recurrence, strict follow-up is required, and patients may also benefit from adjuvant therapy after cure, although no adjuvant therapy has hitherto been considered standard treatment (França et al., 2004; Wang et al., 2013). There has been no consensus on the exploitation to predict the prognosis of HCC, though a slew of attempts and efforts have been made. Previous studies have mostly used parametric prediction models constructed with clinical baseline characteristics (such as tumor size, cirrhosis, tumor number, and microvascular infiltration) and single-molecule biomarkers (such as alpha-fetoprotein [AFP] and Des- $\gamma$  carboxyl-Carboxyl of enzyme) to predict the prognosis of HCC (AlSalloom, 2016). But recently, with the development of genome sequencing technology, the integration of prognostic gene signatures and traditional parameters in the prognosis of HCC has shown great advantages. Nevertheless, it is still of great necessity to make endeavor for the application of these neoteric genetic properties in clinical practice.

Hypoxia is one of the markers of tumor microenvironment. Due to insufficient blood supply, growing tumors often occur in a hypoxia state (Lou et al., 2017). Unlike healthy cells, tumors respond to low oxygen levels by initiating multiple adaptive behaviors (for example, angiogenesis, proliferation, and invasion) that ultimately promote a more aggressive tumor phenotype. For example, glioblastoma can cause extensive tissue hypoxia, which facilitates the induction and maintenance of malignant phenotypes. For the glioma group, tumor hypoxia is associated with anti-apoptosis, tumor recurrence, resistance to chemotherapy and radiotherapy, invasion potential and reduced patient survival (D'Alessio et al., 2019). In addition, previous studies have shown that nearly 50% of locally advanced breast cancers suffer from hypoxia, leading to failed chemoradiotherapy resistance (Vaupel et al., 2002). However, despite considerable efforts on the relationship between hypoxia and tumor, the prediction of the correlation between hypoxic-related gene

expression and overall survival rate in HCC patients has not been reported.

In this study, seven hypoxia gene signatures associated with HCC prognosis were constructed using TCGA dataset and validated in the ICGC dataset. Through GSEA functional enrichment analysis, we examined the important role of the prognostic marker gene in the development and progression of HCC. The final results showed that the model had high reliability in predicting the prognosis of HCC patients, and could help clinicians better carry out individualized treatment. After these bioinformatics analyses, two genes not previously reported in HCC, KDELR3 and SCARB1, were selected to study their expression levels in normal hepatocytes L02, HCC cells SMMC-7721, HepG2, huh7, and SK-Hep-1, respectively.

**TABLE 1 |** Major demographic and clinicopathological characteristics of hepatocellular carcinoma cases ( $n = 50$ ).

Patients characteristic	Frequency (n)	Percentage (%)
<b>Age (years)</b>		
<65 years	18	36%
$\geq 65$ years	32	64%
<b>Sex</b>		
Male	36	72%
Female	14	28%
<b>Tumor size (cm)</b>		
$\leq 5$	13	26%
$> 5$	37	74%
<b>Tumor number</b>		
Single	37	74%
Multiple	13	26%
<b>AFP (ng/ml)</b>		
$\leq 200$	28	56%
$> 200$	22	44%
<b>Liver cirrhosis</b>		
No	26	52%
Yes	24	48%
<b>Grade</b>		
G1-2	29	58%
G3-4	21	42%
<b>TNM stage</b>		
Early (Stage I and II)	23	46%
Advanced (Stage III and IV)	27	54%
DFS (Median and range)	590 (0-28.0)	

Values are expressed as median (range) or n (%). TNM, tumor-nodes-metastasis; AFP,  $\alpha$ -fetoprotein; DFS, disease-free survival.

**Abbreviations:** TCGA, the cancer genome atlas; ICGC, international cancer genome consortium; KM, Kaplan-Meier; qRT-PCR quantitative reverse transcription polymerase chain reaction; HCC, hepatocellular carcinoma; GSEA, gene set enrichment analysis; ROC, receiver operating characteristic

## MATERIALS AND METHODS

### Sample Collection

RNAseq data of this sample mainly comes from TCGA cancer database<sup>1</sup> and ICGC international cancer database<sup>2</sup>, including TCGA dataset as a training set, and ICGC dataset as a verification set. In order to ensure the accuracy and reliability of the

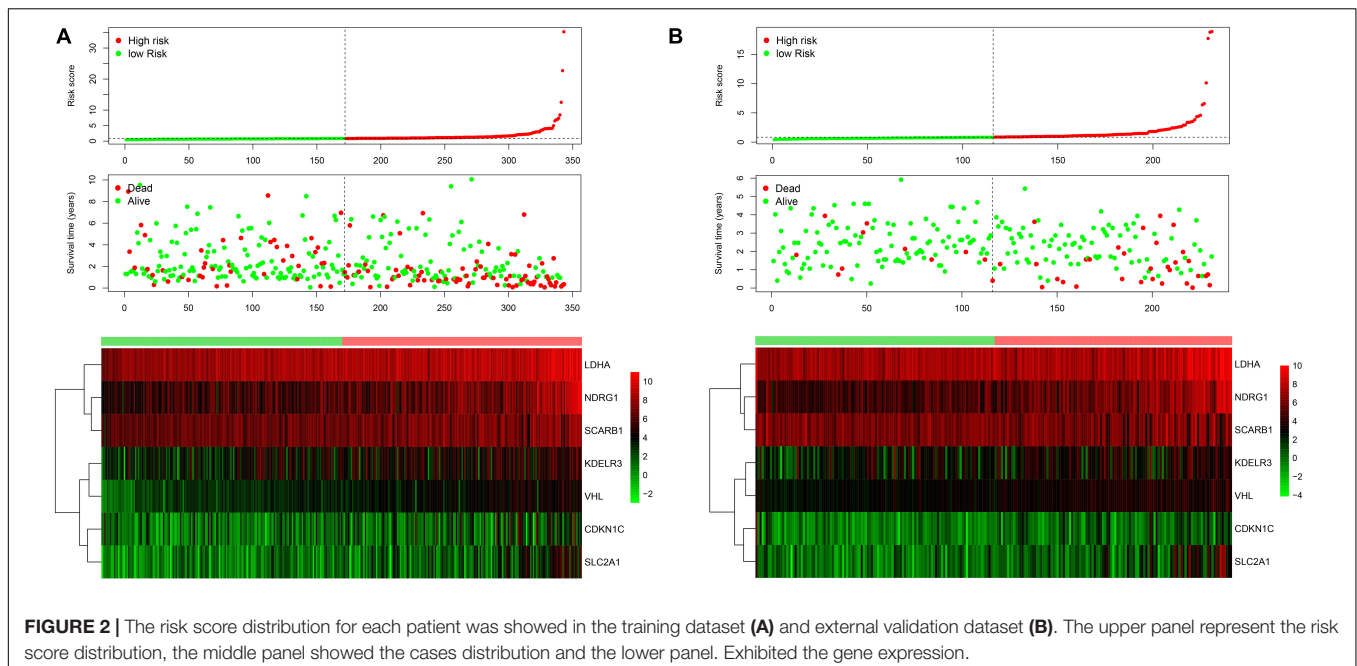
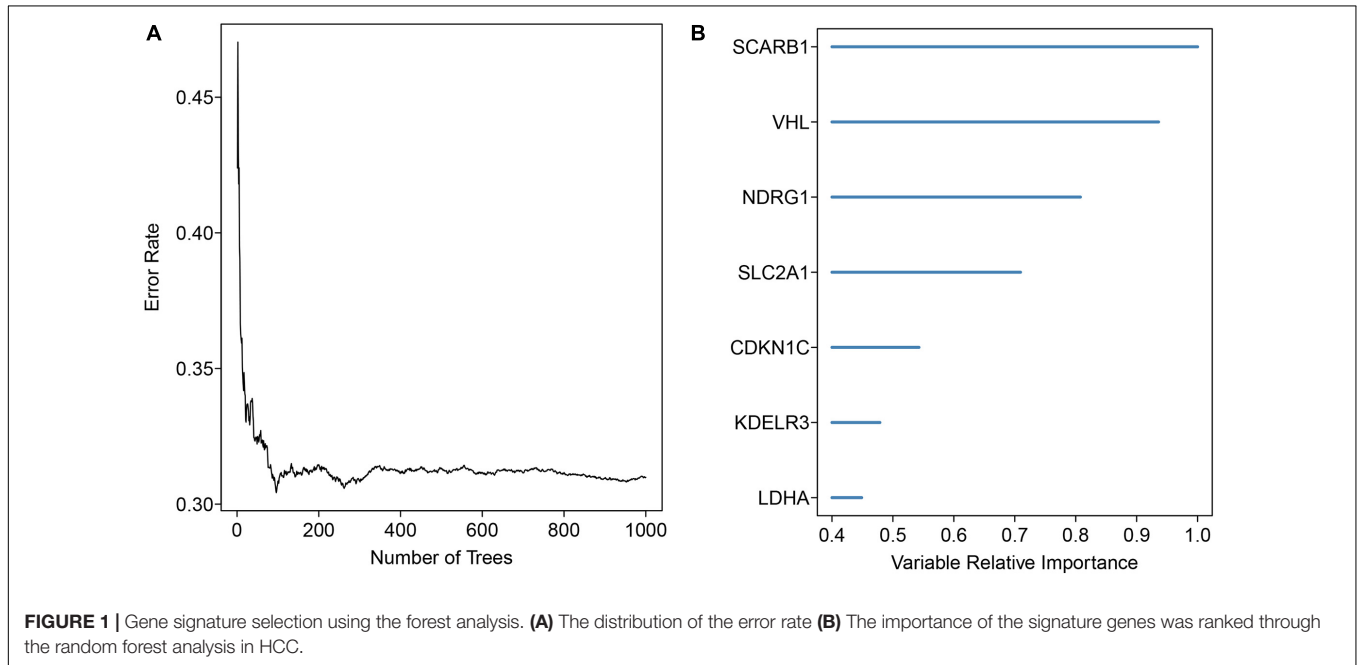
<sup>1</sup><https://portal.gdc.cancer.gov/>

<sup>2</sup><https://ICGC.org/>

data results, the samples with incomplete clinical information and TCGA samples with survival time less than 30 days were eliminated, and finally retained 343 TCGA liver cancer samples and 231 ICGC samples. The 200 hypoxia-related genes were retrieved from Molecular Signatures Database (MSiDb v7.1) and named as HALLMARK\_HYPOXIA.

### Prognostic Model Construction

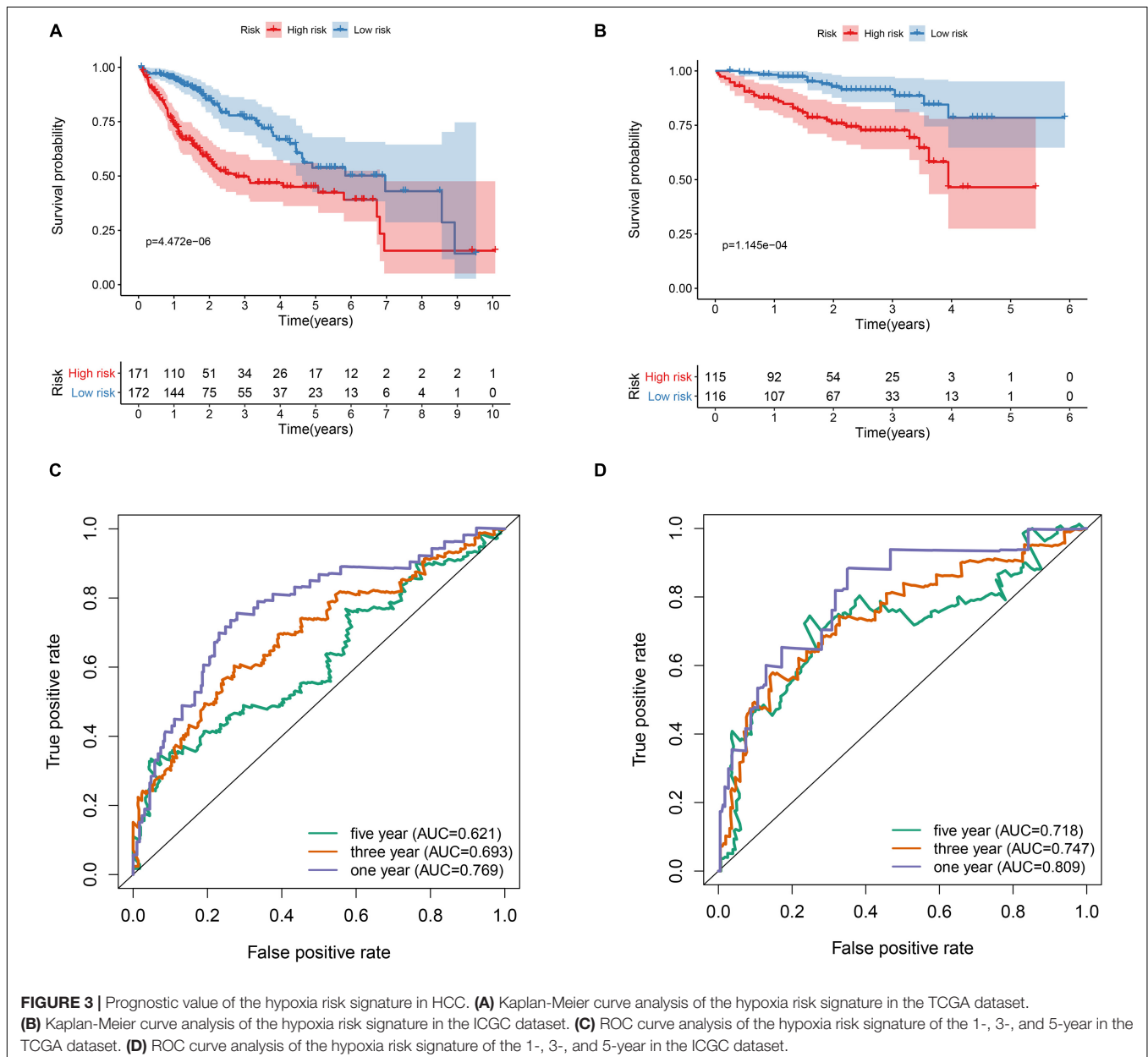
In order to establish a reliable prognosis signature, we first use Univariate Cox regression analysis to screening the prognostic



genes. We then using the random forest algorithm to make a feature selection using the “randomForestSRC” R package. We also applied the randomSurvivalForest algorithm to screening the importance of prognostic-related genes. The genes with a relative importance >0.4 as the final signature. We then using the multivariate cox regression analysis to established an prognostic model based on these genes, and the risk score for each patients was calculated according to therisk formula =  $\sum \text{Coef}_{\text{gene}} \times \text{Exp}_{\text{genes}}$ , where  $\text{Coef}_{\text{gene}}$  represent the coefficient of each prognostic gene, the  $\text{Exp}_{\text{genes}}$  represents each gene expression (Dong and Mingjun, 2019). The patients were further divided into high-risk and low-risk groups based on the median risk score in the TCGA and ICGC dataset, respectively.

### Prognostic Model Evaluation

Kaplan-Meier (KM) curve was used to analyze and compare the survival differences between the high-risk and low-risk groups in the prognostic model. Then, ROC curve was used to evaluate the specificity and sensitivity of the prognostic model by using the “survivalROC” R package. In addition, the KM curve was also used to assess the association between prognostic model risk values and clinical traits. To further assess whether the prognostic model could be used as an independent predictor, univariate and multivariate Cox regression analyses were used to evaluate clinical traits and risk values in training sets and external validation sets, and it was found that the prognostic model could be used as an independent predictor. Furthermore, we constructed a nomogram and calibration curve related to these





independent predictors for personalized independent survival prediction through the “rms” R package.

## GSEA Enrichment Analysis

In order to assess the metabolic pathways involved in the prognostic model, we used GSEA enrichment analysis to assess the enrichment pathways of patients in the high-risk and low-risk groups. We are going to use *c2.cp.kegg.v7.1.symbols.gmt* as background. The screening of significant pathways was considered to be statistically significant with  $p$ -value < 0.05 and error discovery rate FDR < 0.05.

## Cell Culture

Human normal cell line L02, HCC cell lines SMMC-7721, Hep G2, huh7 and SK-Hep-1 were purchased from Cell Resource Center, PMUC (Beijing China). All cells were maintained in Dulbecco's modified Eagle's medium (DMEM; Gibco, United States) supplemented with 10% fetal bovine serum (FBS; Gibco), 1% penicillin and streptomycin (Gibco). The cells were cultured in a 5% CO<sub>2</sub>-humidified atmosphere at 37°C.

## RNA Extraction, Reverse-Transcription RNA, and Quantitative Real-Time Polymerase Chain Reaction

The total RNA was extracted from cell lines by using the TRIzol reagent (Invitrogen), reverse transcription was performed by using the PrimeScript RT reagent Kit (Takara, Japan) and cDNA

was synthesized according to the manufacturer's instructions. The qPCR assay was performed by LightCycler480 system (Roche, Switzerland) and SYBR Green (Takara).

SCARB1: Primer name(F): GGAGATCCCATCCCCTTCTAT, Primer name(R): CTGAACTCCCTGTACACGTAG.

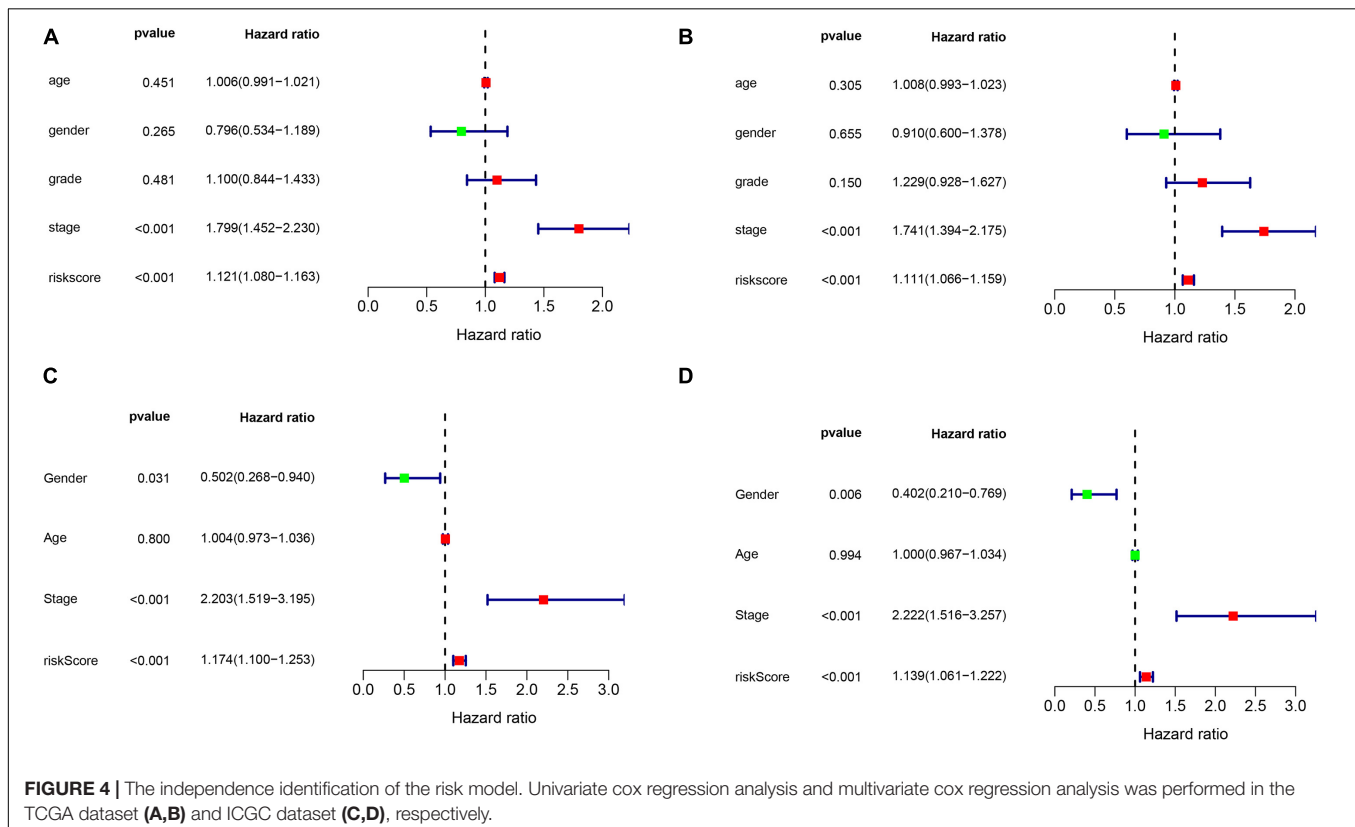
KDEL3: Primer name(F): GAGGCTGAGACCATAACTA CTC, Primer name(R):AGAAATTCTCAGTCTGGTACCG.

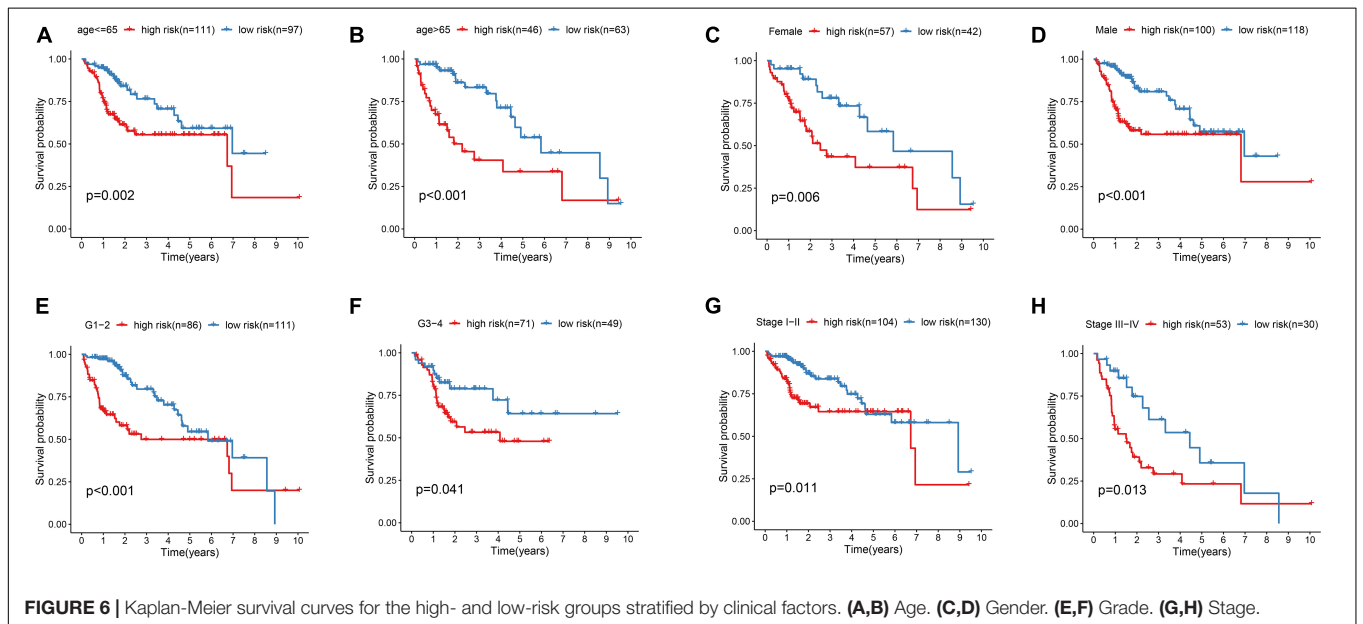
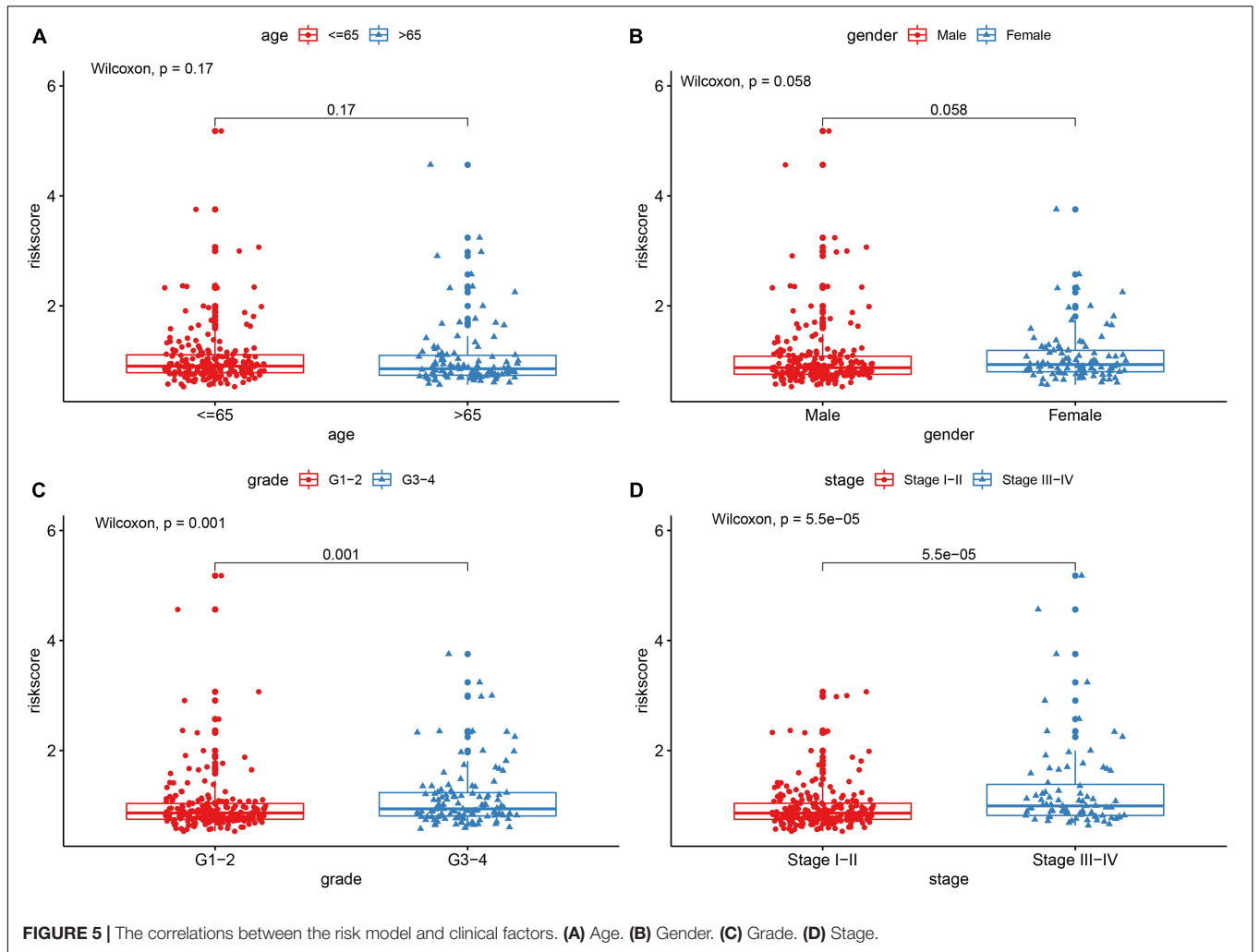
## Patients and Specimens

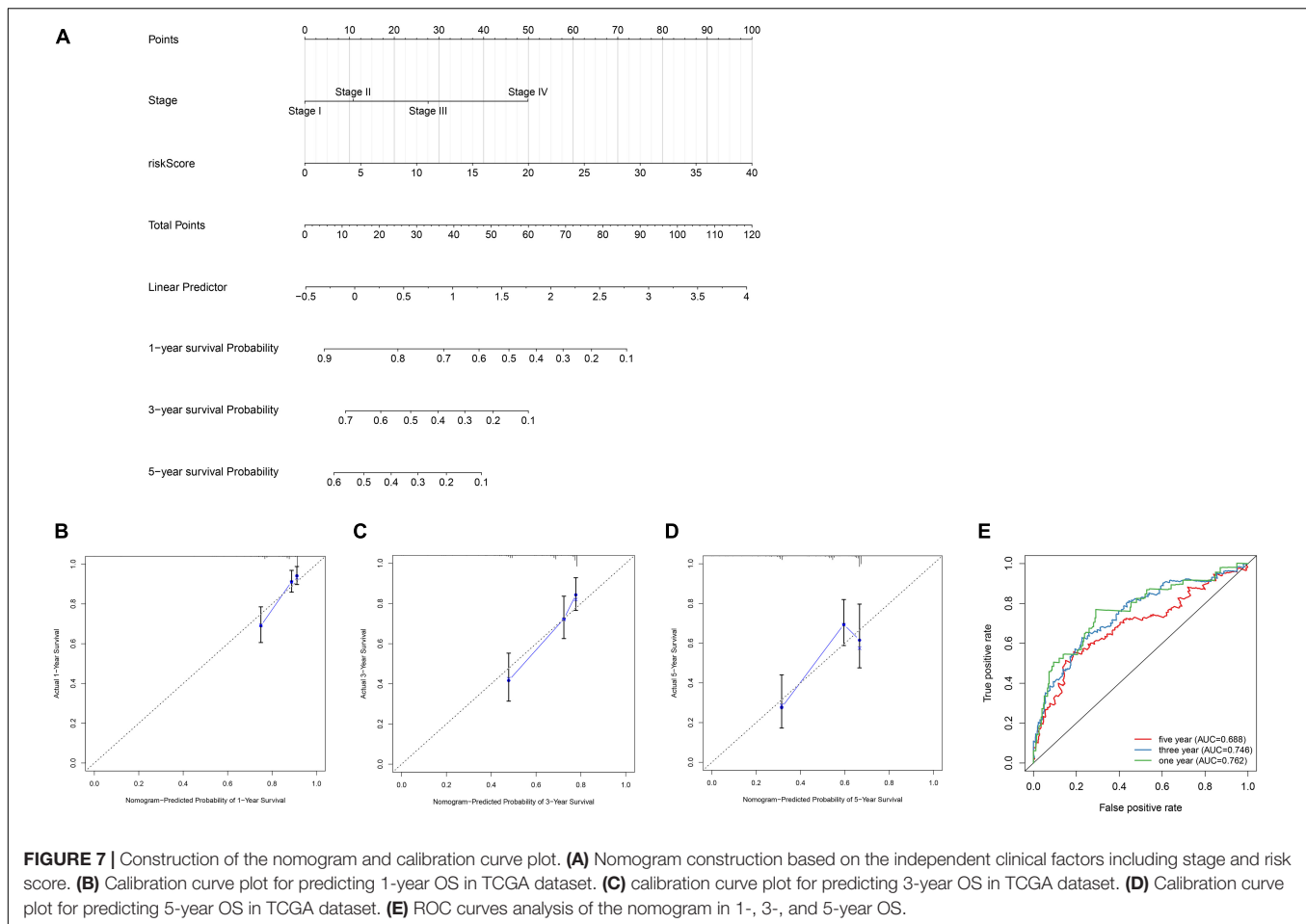
A total 46 formalin-fixed and paraffin-embedded HCC tissue samples from patients who underwent curative surgical resection at the Lanzhou University Second Hospital (Lanzhou, China), from 2016 to 2020, were included. The study protocol was approved by the Institutional Ethics Committee of Lanzhou University Second Hospital (Lanzhou, China) and the patients provided written informed consent regarding the use of their tissues. The main clinicopathological features of these patients are listed (Table 1). All patients underwent radical surgery. Median follow-up was 32 months. Disease free survival (DFS) was defined as time interval from operation date to recurrence.

## Immunohistochemical (IHC) Staining Procedure

Tissue microarray samples were cut into 4 um serial sections and then were placed in an oven at 67°C for 30 min, and further dewaxing in xylene and alcohol. Then tissue samples were treated with TE buffer (pH 9.3, 1 mM EDTA, and 10 mM Tris) at 98°C for 30 min. Next, tissue samples were immersed







in 3% H<sub>2</sub>O<sub>2</sub> in order to eliminate the endogenous peroxidase activity. Then samples were incubated with primary antibodies (SCARB1, 1:100, ABclonal, United States) in phosphate-buffered saline+Tween-20 (PBST) containing 3 mg/ml goat globulin (Sigma, St. Louis, MO, United States) for 60 min at room temperature (RT). Then anti-mouse/rabbit antibody (Envision plus, Dako) were used to incubated with tissue samples for 30 min at RT. Lastly, chromogenic agent 3, 3'-diaminobenzidine (Dako) was used to stain tissue samples.

We were examined and scored the IHC results. According to the positive cells' proportion and the staining intensity, scores were assigned: [score 0], no or less than 5% positive cells; [score 1], 6–20% positive cells; [score 2], 21–50% positive cells; [score 3], more than 50% positive cells. The staining scores of 2 or 3 were considered as high expression, the staining results were scored as high and low.

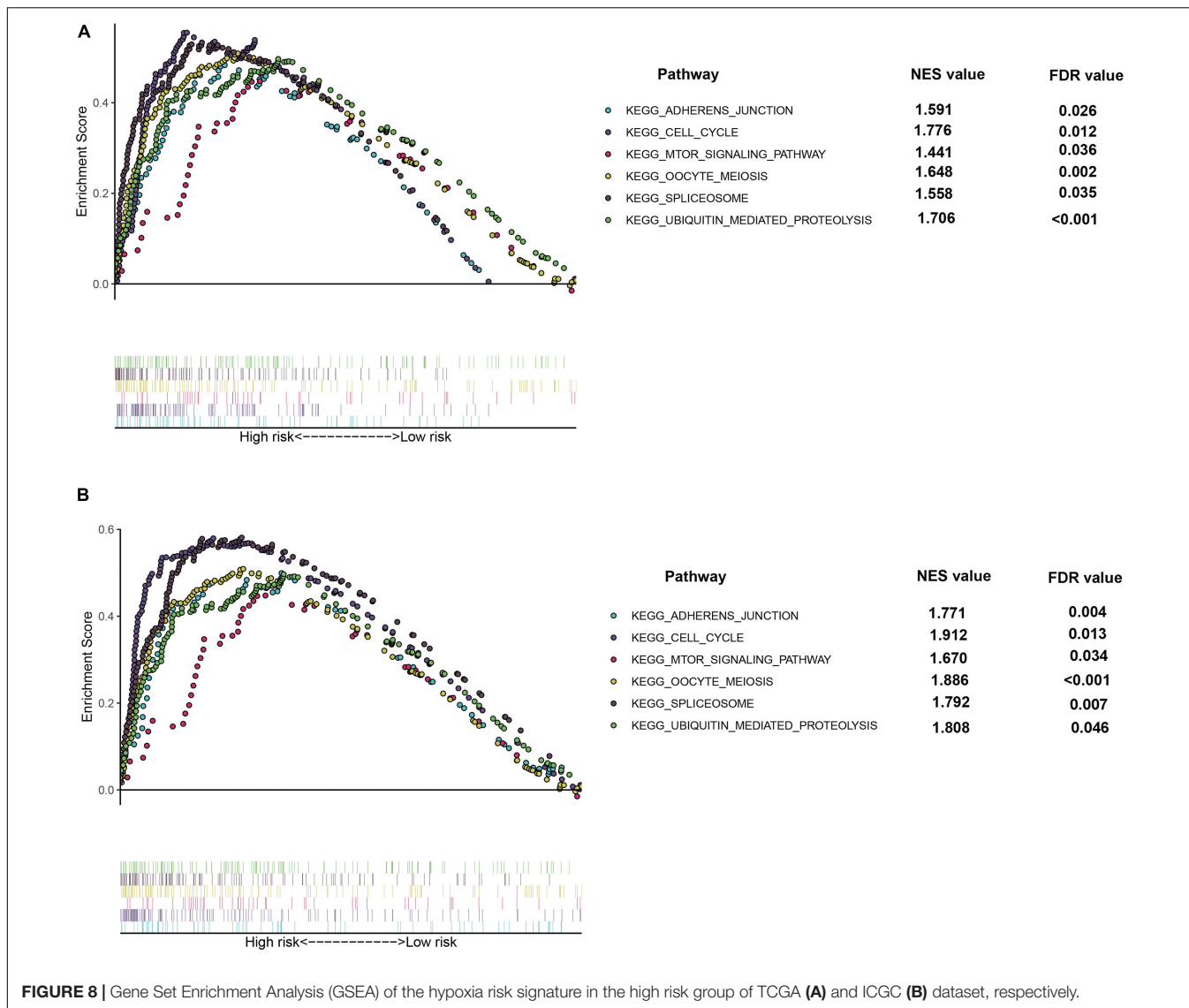
## Statistical Analysis

Statistical software SPSS 20.0 was used for data analysis. Wilcoxon rank test analysis was used to compare different groups.  $P < 0.05$  was considered statistically significant. The survival difference was analyzed using the Kaplan-Meier curve analysis and log-rank test analysis.

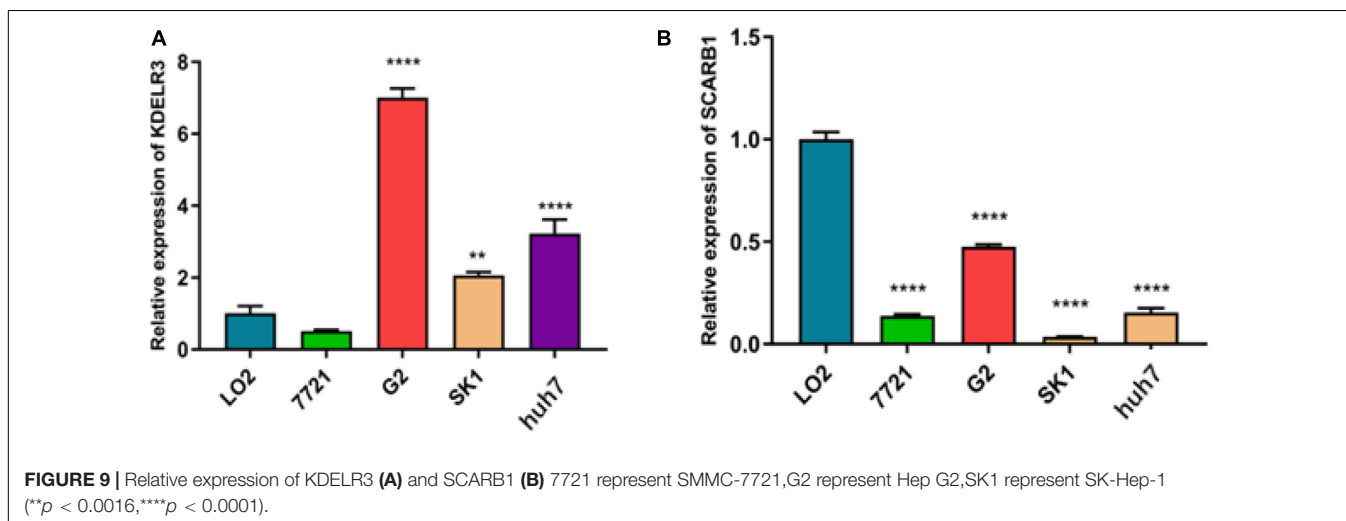
## RESULTS

### Establishment and Validation of Hypoxia-Related Prognostic Models

343 patients with liver cancer and 200 genes associated with hypoxia were used to identify the prognostic model. Using univariate cox regression analysis, 79 survival-related hypoxic genes were selected first and then the random forest algorithm was used for feature selection. We identified the genes of the relative important gene  $>0.4$  was identified as the final feature (**Figures 1A,B**). In view of these characteristics, we used multivariate Cox regression analysis to construct the prognosis model containing seven hypoxic gene-related genes. Through the correlation coefficient, the risk formula was constructed, as follows:  $r$  risk score =  $LDHA \times 0.000812695 + KDELR3 \times 0.000649537 + CDKN1C \times 0.002057653 + SLC2A1 \times 0.004190531 + NDRG1 \times 0.001235623 + VHL \times 0.023669962 + SCARB1 \times 0.00060351$ . Through the risk formula, the risk values of each patient in the training set and external validation set were calculated, and the patients were further divided into high-risk group and low-risk group on the basis of the median risk value. We found the number of death toll from the high-risk group of patients is significantly



**FIGURE 8 |** Gene Set Enrichment Analysis (GSEA) of the hypoxia risk signature in the high risk group of TCGA (A) and ICGC (B) dataset, respectively.



**FIGURE 9 |** Relative expression of KDEL3 (A) and SCARB1 (B) 7721 represent SMMC-7721, G2 represent Hep G2, SK1 represent SK-Hep-1 (\*\* $p < 0.0016$ , \*\*\*\* $p < 0.0001$ ).

higher than low-risk group (Figures 2A,B), moreover, the KM curve analysis that according to the results of the survival of high and low risk group has obvious differences, the survival rate of patients with low risk is far higher than the risk group ( $P < 0.05$ ) (Figures 3A,B), in addition, the results show that the training sample set and validation set outside the ROC curve prognosis is of high accuracy (Figures 3C,D).

## Independent Assessment of Prognostic Model

To assess the prognostic independence of this prognostic model in both the training set and the external validation set samples. We then performed the univariate cox regression analysis and multivariate cox regression analysis in both data sets. As showed in Figures 4A,B, we could discovered that the risk model can serve as an independent prognostic factor in HCC. In addition, we also observed that the risk model could act as an independent prognostic factor in the external validation dataset by performing univariate cox regression analysis (Figure 4C) and multivariate cox regression analysis (Figure 4D).

## Association Between Prognostic Models and Clinical Cause Groups

In order to explore the association between prognostic models and clinical traits, we first assessed the distribution of risk values in clinical traits. As showed in Figures 5A,B, the risk score distribution have no significantly difference in age and gender. While risk values for G3-4 were significantly higher than G1-2 ( $P < 0.05$ ) (Figure 5C). In addition, the risk values for stage III-IV were significantly higher than that of Stage I-II (Figure 5D). These results suggest that a higher risk score is associated with a higher degree of HCC malignance. Therefore, this prognostic model can accurately predict the progression of HCC. In addition, in order to study the prognostic value of the model stratified by clinicopathological variables for HCC patients, stratified analysis was conducted for HCC patients according to age (Figures 6A,B), gender (Figures 6C,D), grade (Figures 6E,F), and stage. For all the different stratifications, the Overall Survival (OS) time was significantly shorter in the high-risk group than in the low-risk group (Figure 6). These results suggest that the prognostic model can predict the prognosis of HCC patients without considering clinicopathological variables.

## The Construction and Verification of Nomogram in TCGA Data Set and ICGC Data Set

In order to establish quantitative prognostic methods for HCC, nomogram was established by independent prognostic factors and prognostic models of two data sets. Based on multivariate Cox analysis, point ratios in nomogram were used to assign points. We drew a horizontal line to determine the points for each variable, calculated the total points for each patient by adding the points for all variables, and normalized it to a distribution of 0 to 100. By drawing a vertical line between the total point axis and each pre-posterior axis, we can calculate the estimated 1-, 3-, and 5-year survival rates of HCC patients, which may be helpful for

**TABLE 2 |** The correlation between SCARB1 and clinicopathological characteristics in HCC.

Patients characteristic	SCARB1		P-value
	Low (n = 17)	High (n = 29)	
Age (years)			0.232
<65	5	9	
≥65	12	20	
Sex			0.232
Male	12	20	
Female	5	9	
Tumor size (cm)			0.265
≤5	9	14	
>5	8	15	
Tumor number			0.568
Single	13	24	
Multiple	4	5	
AFP (ng/ml)			0.075
≤200	11	14	
>200	6	15	
Liver cirrhosis			0.156
No	8	15	
Yes	9	14	
Grade			<0.001
G1-2	10	2	
G3-4	7	27	
TNM stage			<0.001
Early (Stage I and II)	13	6	
Advanced (Stage III and IV)	4	23	

Values are expressed as n (%). TNM, tumor-nodes-metastasis; AFP, alpha-fetoprotein.

practitioners to conduct clinical decisions about the prognosis of HCC patients. In addition, we evaluated the accuracy and consistency of the nomogram by performing the ROC curve and the calibration curve, respectively (Figure 7).

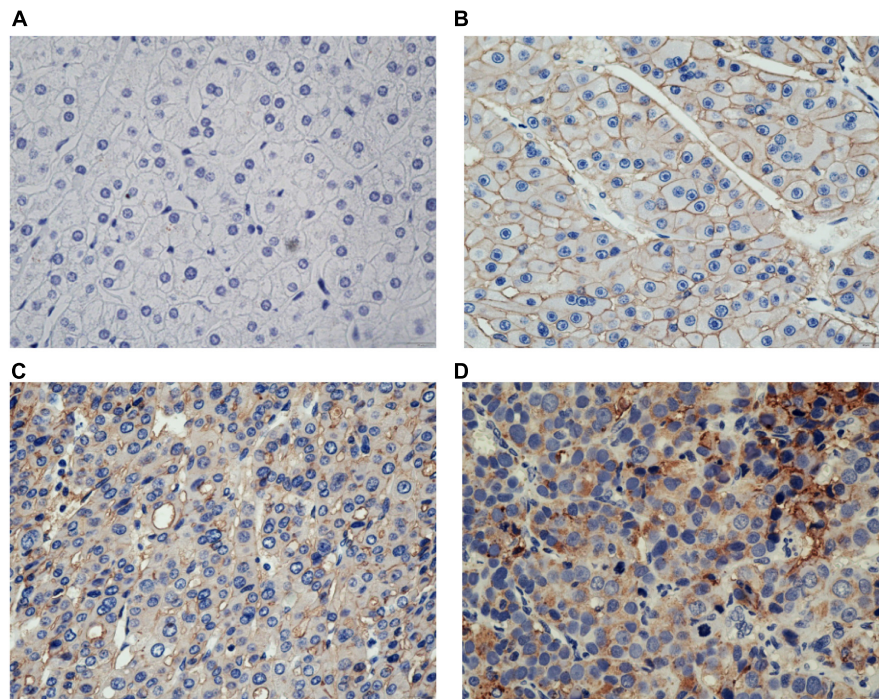
## Functional Enrichment Analysis of Prognostic Model

In order to further explore the potential function and role of prognostic models in HCC, GSEA enrichment analysis was used for enrichment analysis of high and low risk groups. The results showed the cell cycle, MTOR signaling pathways, OOCYTEMEIOSIS and UBIQUITINMEDIATESPROTEOLYSIS pathway were significantly enriched in the high-risk group (Figure 8).

## Expression Levels of kdelr3 and SCARB1

To better explain the biological function of these genes in the pathogenesis and development of HCC, we selected two genes not reported in HCC studies for RT-qPCR to study the differences in their expression levels. We found that the expression of KDELR3 was the highest in G2, while the lowest in 7721, L02 was the highest in SCARB1, and SK-Hep-1 was the lowest ( $p < 0.05$ ; Figure 9).





**FIGURE 10 |** Immunostaining of SCARB1 in HCC or the control tissue. **(A)** Negative staining of SCARB1 in normal tissue adjacent to HCC (400 magnification); **(B)** Positive staining of SCARB1 in well differentiated HCC tissue (400 magnification); **(C)** Expression of SCARB1 in moderate differentiated HCC tissue (400 magnification); and **(D)** Expression of SCARB1 in poor differentiated HCC tissue (400 magnification). SCARB1 = scavenger receptor class B member 1.

## Survival Analysis

For the 46 HCC cases, the median DFS was 9.0 months (range, 28 days-28.0 months). As expected, several clinical factors were associated with the DFS of HCC patients, including tumor Grade, TNM stage ( $p < 0.001$  for all)(Table 2). Furthermore, SCARB1 was highly expressed in cancer tissues and the positive expression rate increased with the degrade of differentiation (Figure 10). The Kaplan-Meier survival was analyzed to study the effect of SCARB1 on HCC patient survival; it revealed that a poor percent survival in the SCARB1 high group compared to that in the SCARB1 low group ( $P < 0.001$ ) (Figure 11).

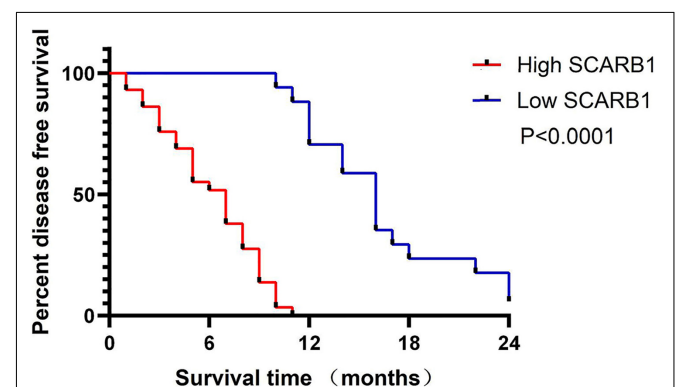
## Univariate and Multivariate Analysis

In Univariate and Multivariate analysis, traditional clinicopathological features (including tumor size, Grade and TNM stage), SCARB1 were associated with DFS. The results show that Grade, TNM stage and SCARB1 are correlated with DFS respectively ( $P = 0.001$ ,  $P = 0.036$ ,  $P < 0.001$ ) (Table 3).

## DISCUSSION

Hepatocellular carcinoma is one of the major health threats around the world, particularly in East Asia. Even after radical resection, the long-term outcome of HCC patients remains depressed (Mazzanti et al., 2016). Therefore, it is very important

to develop a prognostic model suitable for HCC patients. Recently, with the improvement of genome sequencing, biochip and high-throughput sequencing technologies, more and more studies have applied bioinformatics methods to chip dataset analysis, which provides an effective new method for the diagnosis, treatment and prognosis of HCC. In this study, a total of 574 HCC samples were downloaded from TCGA and ICGC databases, using TCGA as the training set and ICGC as the test set. Through bioinformatics analysis, a prognostic



**FIGURE 11 |** Kaplan-Meier analysis of the survival rate of patients with HCC. Disease-free survival of all patients in relation to SCARB1 expression ( $p < 0.0001$ ).

**TABLE 3** | Multivariate analyses of factors associated with DFS.

Clinicopathological factor	Univariate analysis			Multivariate analysis		
	P-value	HR	95% CI	P-value	HR	95% CI
Age (<605 vs. ≥65 years)	0.348	1.004	0.981-1.020			
Sex (female vs. male)	0.281	0.768	0.540-1.198			
Tumor size (≤5 vs. >5)	0.679	0.789	0.343-2.160	0.281	1.452	0.712-2.654
Tumor number (single vs. multiple)	0.310	0.378	0.131-1.418			
AFP (≤200 vs. >200)	0.097	1.812	0.891-3.189			
Liver cirrhosis (no vs. yes)	0.058	0.813	0.514-0.868			
Grade(G1-2 vs. G3-4)	0.041	1.177	0.850-1.530	0.036	1.102	0.618-1.890
TNM stage (early vs. advanced) <sup>a</sup>	0.038	1.755	1.465-2.437	<0.001	1.812	0.944-3.828
SCARB1 expression (low vs. high)	0.005	1.002	1.001-1.006	0.025	1.870	1.082-3.233

<sup>a</sup>Early, TNM stage I/II; advanced, TNM stage III/IV. DFS, disease-free survival; SCARB1, scavenger receptor class B member 1; TNM, tumor-nodes-metastasis; HR, hazard ratio; CI, confidence interval.

model of HCC with seven genes associated with hypoxia was constructed for the first time. Our prognostic model can effectively stratify the survival of patients. We found the efficacy of our prognostic model in both the training set and the external validation set, suggesting that the model has strong prognostic value. In addition, the prognostic model showed a significant correlation with clinicopathological factors, further supporting the robustness of the prognostic role of our model. In addition, univariate and multivariate Cox regression analysis were used to validate our prognostic model as an independent predictor. Nomogram of independent predictors (staging and prognostic models) were constructed and showed that the model performed well in predicting 1-, 3-, and 5-year OS, which may be useful for planning short-term follow-up for individual treatment. In summary, the predictive prognostic value of our signatures is greatly reflected in these results, but it is worth noting that only two databases were selected. To validate the model on a large scale, the signatures need to be validated in a more independent queue.

Among the seven genes in the prognostic model that we constructed, LDHA (lactate dehydrogenase A) is a crucial REDOX enzyme in the glycolysis pathway in organisms, which can reversely catalyze the oxidation of lactic acid to pyruvate, and this catalytic reaction is the final product of anaerobic glycolysis (Guo et al., 2019). In addition, LDHA can be used

as a possible prognostic marker for lung adenocarcinoma survival, and the high expression of LDHA is associated with poor prognosis (Yu et al., 2020). NDRG1 is a known metastasis inhibitor in a variety of cancers, participating in embryogenesis, cell growth, lipid biogenesis, stress response, and immunity. During metastasis, tumor growth and invasion require angiogenesis, and overexpression of NDRG1 is associated with a decrease in pro-angiogenic factors, resulting in a decrease of angiogenesis in pancreatic cancer (Hosoi et al., 2009). The expression levels of NDRG2 and LDHA are closely related to the prognosis of HCC patients and can be used as prognostic markers (Guo et al., 2019). KDELR3 is the third confirmed member in the KDEL family, which encodes proteins associated with the endoplasmic reticulum (ER). Reports have shown previously that KDELR3 expression in arteriosclerosis macrophages could be obviously differ from that in non-arteriosclerosis tissues, and the higher expression level in non-arteriosclerosis tissues, which can be used as a potential prognostic factor (Huang et al., 2019). CDKN1C, DKN1C, also known as p57Kip2, are cyclin-dependent kinase inhibitors. The gene encoding CDKN1C is located on chromosome 11p15.5. CDKN1C belongs to the kinase inhibitor protein /CDK interacting protein (Kip/Cip) family, which consists of three members, namely CDKN1A/p21Cip1, CDKN1B/p27Kip1, and CDKN1C/p57Kip2. Previously, CDKN1C was identified as a tumor suppressor gene with decreased expression in various cancers including HCC, colorectal cancer and ovarian cancer (Peng et al., 2015). Thus, upregulation of CDKN1C leads to inhibition of markers involved in cell growth, differentiation, cell death, and angiogenesis in malignant tumors. SLC2A1 has been extensively studied as a major glucose transporter and has been identified as a possible prognostic factor for several cancers, including HCC and NSCLC et al. (Shang et al., 2020). The absence of von Hippel-Lindau (VHL), a tumor suppressor gene, is a hallmark of clear cell renal carcinoma. In addition, VHL inactivation leads to constitutive activation of hypoxia-inducing factor (HIF) HIF-1 and HIF-2 and their downstream targets (including pro-angiogenic factors VEGF and PDGF), while the activation of HIF and its downstream targets induces tumor formation (Gossage et al., 2015). SCARB1 (Scavenger receptor class B member 1) is a protein-coding gene. The related pathways include lipoprotein metabolism and folic acid metabolism. Studies have also found that SCARB1 can be a potential target for prostate cancer (Gordon et al., 2019).

The relationship between cell cycle and malignancy is now fully established, and our GSEA results show that the associated pathways are enriched in high-risk patients. The activation of mTOR related pathways can promote the vitality and motility of HCC cell lines (Wu et al., 2018). Enrichment of The OOCYTEMESIS pathway was found in bladder cancer and cervical cancer (Gao et al., 2018; Wu et al., 2019). UBIQUITINMEDIATEDPROTEOLYSIS related pathways plays an important role in the disease progression of colorectal cancer and non-small cell lung cancer (Lu et al., 2018; Wu et al., 2020). These three pathways were enriched in our high-risk patient group.

Although we have made a lot of efforts to study the prognostic model, there are still many defects. The clinical information of external validation set and test set is not fully matched, leading to the omission of partial clinical information and the inability to fully understand the correlation between the prognostic model and clinical practice. Moreover, none of genes have been verified, because some of them have been studied in HCC. In view of this, two genes, KDELR3 and SCARB1, have not been studied in HCC, laying a foundation for further studies in the future. In the present study, we have verified the RNA expression level of KDELR3 and SCARB1, protein expression level of SCARB1 and its percent survival, which verified our prognostic model. And more research is needed.

In conclusion, a prognostic model consists of seven genes associated with hypoxia was constructed to provide potential biomarkers for the prognosis of HCC, which contributed to the understanding of the underlying HCC pathogenesis.

## DATA AVAILABILITY STATEMENT

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

## REFERENCES

- AlSallloom, A. A. (2016). An update of biochemical markers of hepatocellular carcinoma. *Int. J. Health Sci.* 10, 121–136.
- D'Alessio, A., Proietti, G., Sica, G., and Scicchitano, B. M. (2019). Pathological and molecular features of glioblastoma and its peritumoral tissue. *Cancers* 11:469. doi: 10.3390/cancers11040469
- Dong, J., and Mingjun, X. (2019). A 19 miRNA Support Vector Machine classifier and a 6 miRNA risk score system designed for ovarian cancer patients. *Oncol. Rep.* 41, 3233–3243. doi: 10.3892/or.2019.7108
- El-Serag, H. B. (2011). Hepatocellular carcinoma. *N. Engl. J. Med.* 365, 1118–1127. doi: 10.1056/NEJMra1001683
- França, A. V. C., Elias, J. Jr., Lima, B. L., Martinelli, A. L., and Carrilho, F. J. (2004). Diagnosis, staging and treatment of hepatocellular carcinoma. *Braz. J. Med. Biol. Res.* 37, 1689–1705. doi: 10.1590/s0100-879x2004001100015
- Gao, X., Chen, Y., Chen, M., Wang, S., Wen, X., and Zhang, S. (2018). Identification of key candidate genes and biological pathways in bladder cancer. *PeerJ* 6:e6036. doi: 10.7717/peerj.6036
- Gordon, J. A., Noble, J. W., Midha, A., Derakhshan, F., Wang, G., Adomat, H. H., et al. (2019). Upregulation of scavenger receptor B1 is required for steroidogenic and nonsteroidogenic cholesterol metabolism in prostate cancer. *Cancer Res.* 79, 3320–3331. doi: 10.1158/0008-5472.CAN-18-2529
- Gossage, L., Eisen, T., and Maher, E. R. (2015). VHL, the story of a tumour suppressor gene. *Nature Reviews. Cancer* 15, 55–64. doi: 10.1038/nrc3844
- Guo, Y., Li, X., Sun, X., Wang, J., Yang, X., Zhou, X., et al. (2019). Combined aberrant expression of NDRG2 and LDHA predicts hepatocellular carcinoma prognosis and mediates the anti-tumor effect of gemcitabine. *Int. J. Biol. Sci.* 15, 1771–1786. doi: 10.7150/ijbs.35094
- Hosoi, F., Izumi, H., Kawahara, A., Murakami, Y., Kinoshita, H., Kage, M., et al. (2009). N-myc downstream regulated gene 1/Cap43 suppresses tumor growth and angiogenesis of pancreatic cancer through attenuation of inhibitor of kappaB kinase beta expression. *Cancer Res.* 69, 4983–4991. doi: 10.1158/0008-5472.CAN-08-4882
- Huang, H. M., Jiang, X., Hao, M. L., Shan, M. J., Qiu, Y., Hu, G. F., et al. (2019). Identification of biomarkers in macrophages of atherosclerosis by microarray analysis. *Lipids Health Dis.* 18:107. doi: 10.1186/s12944-019-1056-x
- Lou, J., Zhang, L., Lv, S., Zhang, C., and Jiang, S. (2017). Biomarkers for hepatocellular carcinoma. *Biomark. Cancer* 9, 1–9. doi: 10.1177/1179299X16684640
- Lu, J., Li, Y., Wu, Y., Zhou, S., Duan, C., Dong, Z., et al. (2018). MICAL2 mediates p53 Ubiquitin Degradation through Oxidating p53 Methionine 40 and 160 and promotes colorectal cancer malignance. *Theranostics* 8, 5289–5306. doi: 10.7150/thno.28228
- Mazzanti, R., Arena, U., and Arena, U. (2016). Hepatocellular carcinoma: where are we? *World J. Exp. Med.* 6, 21–36. doi: 10.5493/wjem.v6.i1.21
- Peng, Y. T., Wu, W. R., Chen, L. R., Kuo, K. K., Tsai, C. H., Huang, Y. T., et al. (2015). Upregulation of cyclin-dependent kinase inhibitors CDKN1B and CDKN1C in hepatocellular carcinoma-derived cells via goniothalamin-mediated protein stabilization and epigenetic modifications. *Toxicol. Rep.* 2, 322–332. doi: 10.1016/j.toxrep.2015.01.010
- Shang, R., Wang, M., Dai, B., Du, J., Wang, J., Liu, Z., et al. (2020). Long noncoding RNA SLC2A1-AS1 regulates aerobic glycolysis and progression in hepatocellular carcinoma via inhibiting the STAT3/FOXM1/GLUT1 pathway. *Mol. Oncol.* 14, 1381–1396. doi: 10.1002/1878-0261.12666
- Vaupel, P., Briest, S., and Höckel, M. (2002). Hypoxia in breast cancer: pathogenesis, characterization and biological/therapeutic implications. *Wien. Med. Wochenschr.* 152, 334–342. doi: 10.1046/j.1563-258x.2002.02032.x
- Wang, J., He, X. D., Yao, N., Liang, W. J., and Zhang, Y. C. (2013). A meta-analysis of adjuvant therapy after potentially curative treatment for hepatocellular carcinoma. *Can. J. Gastroenterol.* 27, 351–363. doi:10.1155/2013/417894
- Wu, D. M., Deng, S. H., Zhou, J., Han, R., Liu, T., Zhang, T., et al. (2020). PLEK2 mediates metastasis and vascular invasion via the ubiquitin-dependent degradation of SHIP2 in non-small cell lung cancer. *Int. J. Cancer* 146, 2563–2575. doi: 10.1002/ijc.32675

## AUTHOR CONTRIBUTIONS

HC designed this study. YB and WQ wrote the manuscript. LL and JZ collected the data. LP and TG analyzed the data. PW revised the manuscript. All authors approved the final version for submission.

## FUNDING

This study was supported by Key Project of Science and Technology in Gansu province (19ZD2WA001) and Cuiying Scientific and Technological Innovation Program of Lanzhou University Second Hospital (No. CY2017-ZD01) and Science and technology project of Chengguan District of Lanzhou City (2019RCCX0034).

## ACKNOWLEDGMENTS

We thank HC of Lanzhou University for Project design. The code involved in the analysis of the manuscript can be found online at: <https://github.com/Bioinformatics-analysis/Hypoxia-hepatic-carcinoma>.

- Wu, T., Dong, X., Yu, D., Shen, Z., Yu, J., and Yan, S. (2018). Natural product pectolinarigenin inhibits proliferation, induces apoptosis, and causes G2/M phase arrest of HCC via PI3K/AKT/mTOR/ERK signaling pathway. *OncoTargets Ther.* 11, 8633–8642. doi: 10.2147/OTT.S186186
- Wu, X., Peng, L., Peng, L., Chen, S., Lei, Q., Li, G., et al. (2019). Identification of key genes and pathways in cervical cancer by bioinformatics analysis. *Int. J. Med. Sci.* 16, 800–812. doi: 10.7150/ijms.34172
- Yang, J. D., Hainaut, P., Gores, G. J., Amadou, A., Plymoth, A., and Roberts, L. R. (2019). A global view of hepatocellular carcinoma: trends, risk, prevention and management. *Nature reviews. Gastroenterol. Hepatol.* 16, 589–604. doi: 10.1038/s41575-019-0186-y
- Yu, X., Zhang, X., and Zhang, Y. (2020). Identification of a 5-Gene metabolic signature for predicting prognosis based on an integrated analysis of tumor microenvironment in lung adenocarcinoma. *J. Oncol.* 2020:5310793. doi: 10.1155/2020/5310793
- 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.
- Copyright © 2021 Bai, Qi, Liu, Zhang, Pang, Gan, Wang, Wang and Chen. 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.*