

# Prognostic value of fatty acid metabolism-related genes in patients with hepatocellular carcinoma

Dongsheng He<sup>1</sup>, Lifang Cai<sup>1</sup>, Weiming Huang<sup>1</sup>, Qingqing Weng<sup>1</sup>, Xi Lin<sup>1</sup>, Mengxing You<sup>1,\*</sup>, Shengyin Liao<sup>1,\*</sup>

<sup>1</sup>Department of Medical Oncology, The First Hospital of Putian, Teaching Hospital, Fujian Medical University, Putian 351100, Fujian Province, China

\*Equal contribution

**Correspondence to:** Mengxing You, Shengyin Liao; email: [youmengxing\\_oncol@163.com](mailto:youmengxing_oncol@163.com), <https://orcid.org/0000-0002-5579-3023>; [Shengyinliao@bbmc.edu.cn](mailto:Shengyinliao@bbmc.edu.cn)

**Keywords:** hepatocellular carcinoma, TCGA, ICGC, fatty acid metabolism, prognosis

**Received:** February 18, 2021

**Accepted:** June 2, 2021

**Published:** July 13, 2021

**Copyright:** © 2021 He et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/3.0/) (CC BY 3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

## ABSTRACT

The deregulation of fatty acid metabolism plays a crucial role in cancer. However, the prognostic value of genes involved in the metabolism in hepatocellular carcinoma (HCC) remains largely unknown. We first constructed a multi-fatty acid metabolic gene prognostic model of HCC based on The Cancer Genome Atlas (TCGA) and further validated it using the International Cancer Genome Consortium (ICGC) database. The model was integrated with the clinical parameters, and a nomogram was built and weighted. Moreover, immune cell infiltration of the tumor microenvironment was investigated. A prognostic model was constructed using 6 selected fatty acid metabolism-related genes, and HCC patients were divided into high- and low-risk groups. Receiver operating characteristic curve (ROC) analysis, principal component analysis (PCA), and t-distributed stochastic neighbor embedding (t-SNE) analysis showed the optimal performance of the model. The concordance index (C-index), ROC curve, calibration plot and decision curve analysis (DCA) all confirmed the satisfactory predictive capacity of the nomogram. The analysis of immune cell infiltration in HCC patients revealed a correlation with different risk levels. Our findings indicate that a prognostic model based on fatty acid metabolism-related genes has superior predictive capacities, which provides the possibility for further improving the individualized treatment of patients with HCC.

## INTRODUCTION

Liver cancer is the sixth most common cancer and leads to unacceptable mortality, with more than 800,000 deaths reported in 2020 [1]. As the most common primary component of liver cancer, hepatocellular carcinoma (HCC) has already become a serious global public health problem [2]. Although multiple treatments, such as surgery, chemotherapy, radiotherapy and targeted therapy, have been developed in recent years, the 5-year survival rate of patients with HCC is unsatisfactory due to recurrence in a large proportion of patients after hepatectomy and diagnosis at an advanced stage [3]. Consequently, studies exploring novel

therapeutic targets and developing novel prognostic models are urgently needed for patients with HCC.

Metabolic dysregulation is one of the ten hallmarks of cancers, and accumulating evidence indicates that metabolic reprogramming plays a crucial role in the initiation and development of cancer [4–6]. Lipid metabolic reprogramming is one of the most prominent metabolic changes observed in cancer cells and has received increasing attention. As an important component of lipid metabolism, the accumulation of fatty acids has been observed, which can be used to satisfy the requirement for lipids to synthesize signaling molecules and membranes. Fatty acids are required for

membrane synthesis, energy storage and the generation of signaling molecules, and accumulating evidence suggests that fatty acids are indispensable for the initiation and development of cancers such as breast cancer and colorectal cancer [7–9]. Moreover, deregulated fatty acids might not only disturb the curative effect of chemotherapeutic and radiation treatments on patients with cancer [10, 11] but also affect immunotherapy, which has been a breakthrough in oncotherapy in recent years. Fatty acids in the tumor microenvironment affect the function and phenotype of infiltrating immune cells, which are associated with immunosuppression [12]. Treatments targeting deregulated fatty acids in cancer might slow tumor growth and exert synergistic effects with immune checkpoint inhibitors [13]. Nevertheless, the distinct fatty acid metabolism-related genes, prognostic value and relationship with immunotherapy in HCC remain largely unknown.

In the present study, we first comprehensively analyzed the correlation between differentially expressed fatty acid metabolism-related genes and the prognosis of patients with HCC and constructed a prognostic model based on The Cancer Genome Atlas (TCGA) database. Then, the prognostic risk model was further validated using the International Cancer Genome Consortium (ICGC) database, and patients with HCC were divided into high- and low-risk groups. Moreover, the prognostic risk model was integrated with the clinical parameters of patients with HCC. Subsequently, a clinical prognostic model of patients with HCC was structured and validated from different perspectives. Finally, the difference in immune cell infiltration in high- and low-risk patients with HCC was explored. Taken together, our findings indicated that fatty acid metabolism-related genes might be potential prognostic markers and therapeutic targets in patients with HCC and further improve the efficacy of treatment in patients with HCC through personalized treatment.

## RESULTS

### Identification of fatty acid metabolism-related genes in patients with HCC

To identify genes related to fatty acid metabolism, the intersection of the three gene sets related to fatty acid metabolism was filtered, and 309 genes were extracted after the overlapping genes were removed (Figure 1A). The “Limma” R package was used to identify 12,450 differentially expressed genes (DEGs) from TCGA and 2,939 DEGs from the ICGC database (Figure 1B, 1C).

After excluding data from patients with a follow-up time of less than 90 days, the fatty acid metabolism-

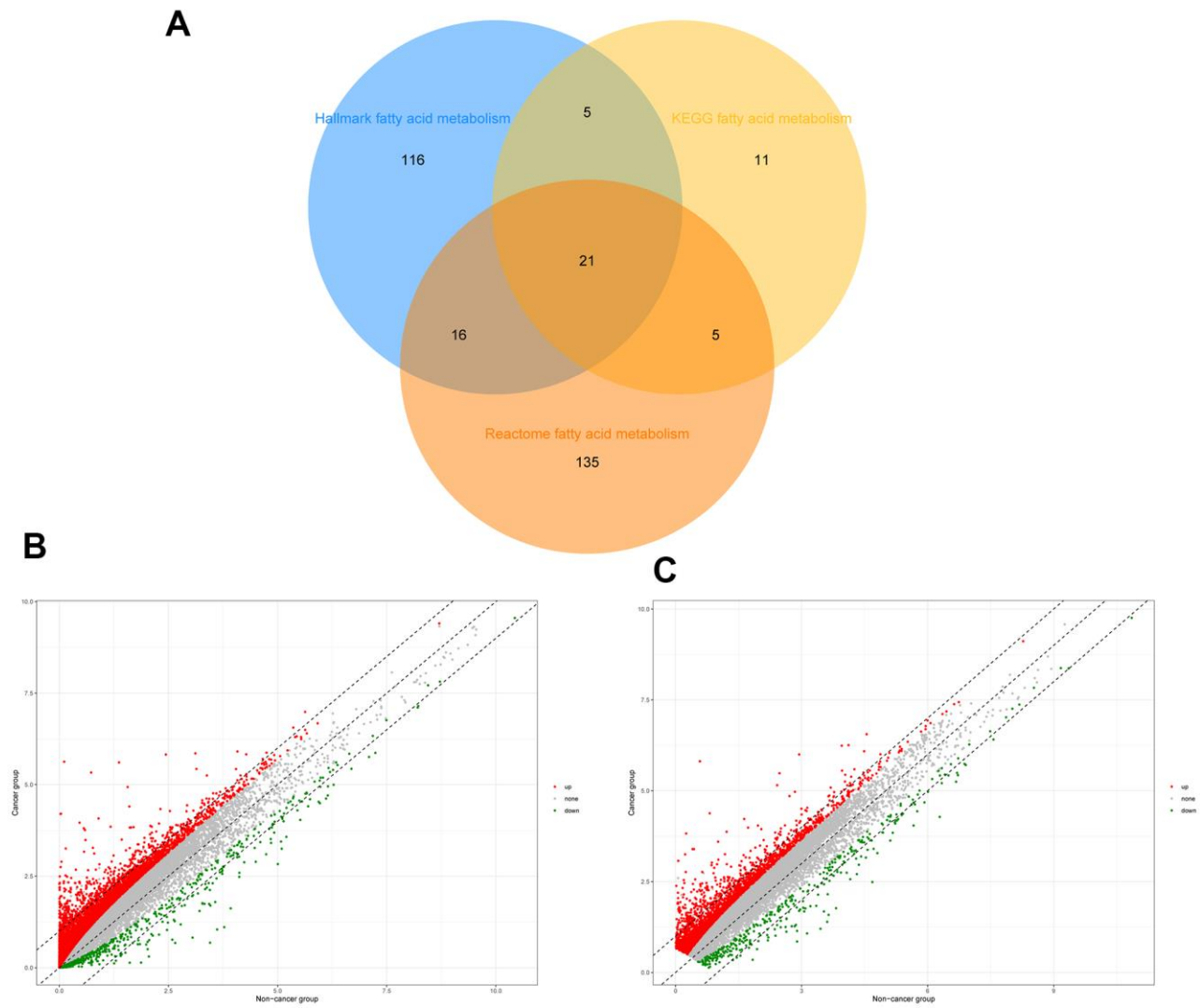
related genes and the corresponding expression profile were filtered in 329 patients with HCC in TCGA. First, 105 fatty acid metabolism-related genes correlated with overall survival (OS) were identified by performing a univariate Cox proportional hazard analysis. Next, 47 differentially expressed fatty acid metabolism-related genes correlated with OS were selected by intersecting the DEGs, and 105 fatty acid metabolism-related genes were correlated with the OS of patients in TCGA (Figure 2A). Then, the 47 selected fatty acid metabolism-related genes were validated by analyzing DEGs in the ICGC database, and 26 genes were filtered (Figure 2B).

### Construction and assessment of a prognostic model based on selected fatty acid metabolism-related genes

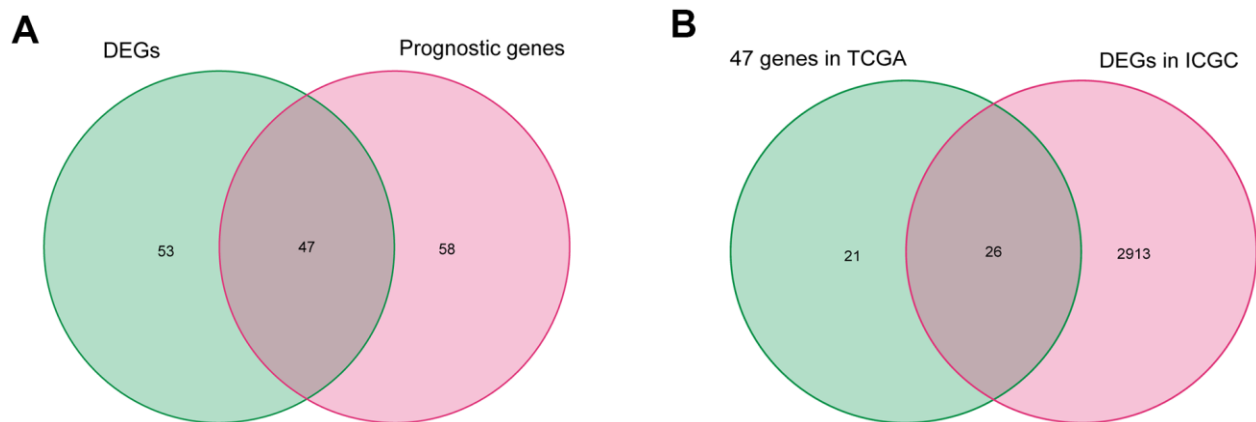
The 26 filtered genes were further included in the least absolute shrinkage and selection operator (LASSO) logistic regression algorithm (Figure 3A), cross-validation was conducted, and 11 prognostic signatures were screened to forecast an individual’s prognostic risk (Figure 3B). The heatmap of 11 genes is shown in Figure 3C, and the correlation network of the 11 genes is shown in Figure 3D. Then, 6 independent prognostic genes were further screened by performing a multivariable Cox regression analysis of the 11 prognostic signatures, and a prognostic model was constructed. Risk score =  $(-0.081 \times \text{level of the } ADH4 \text{ mRNA}) + (0.339 \times \text{level of the } ELOVL1 \text{ mRNA}) + (0.137 \times \text{level of the } ME1 \text{ mRNA}) + (0.397 \times \text{level of the } ACACA \text{ mRNA}) + (-0.330 \times \text{level of the } ACADS \text{ mRNA}) + (-0.417 \times \text{level of the } ACSL6 \text{ mRNA})$ . Three hundred twenty-nine patients with HCC were separated into high-risk (n=164) and low-risk (n=165) groups by the median risk score as the cutoff value. The OS of the two groups was significantly different in the Kaplan-Meier (K-M) analysis ( $P < 0.001$ ) (Figure 4A). A receiver operating characteristic (ROC) curve was constructed to estimate the model and assess the reliability of the risk score, and the areas under the curve (AUCs) at 1 year, 2 years and 3 years were 0.800, 0.714 and 0.697, respectively (Figure 4C–4E). Moreover, principal component analysis (PCA) and t-distributed stochastic neighbor embedding (t-SNE) analysis were performed to identify low-risk patients and high-risk patients. As shown in Figure 5A–5D, patients with HCC were divided into different risk groups with a relatively clear resolution.

### Validation of the risk model in patients with HCC

To further validate the risk model, the risk model was used to assess patients in the ICGC database according to the median risk score of the 6 genes in TCGA. The P



**Figure 1. Identification of differentially expressed fatty acid metabolic genes in HCC. (A)** Venn diagram of three fatty acid metabolic gene sets. **(B)** Differentially expressed genes of HCC patients in TCGA. **(C)** Differentially expressed genes of HCC patients in ICGC.



**Figure 2. Identification of differentially expressed fatty acid metabolic genes correlated with OS in HCC. (A)** Differentially expressed fatty acid metabolic genes correlated with OS in TCGA. **(B)** Further validation in ICGC.

value for the K-M analysis of the high-risk and low-risk groups was less than 0.001 in the ICGC database (Figure 4B). Meanwhile, the AUCs for 1 year, 2 years and 3 years were 0.732, 0.687 and 0.682, respectively (Figure 4F–4H). PCA and t-SNE analysis were also conducted to test the risk model based on the ICGC database. As shown in Figure 5E–5H, high-risk and low-risk patients were divided by different analyses.

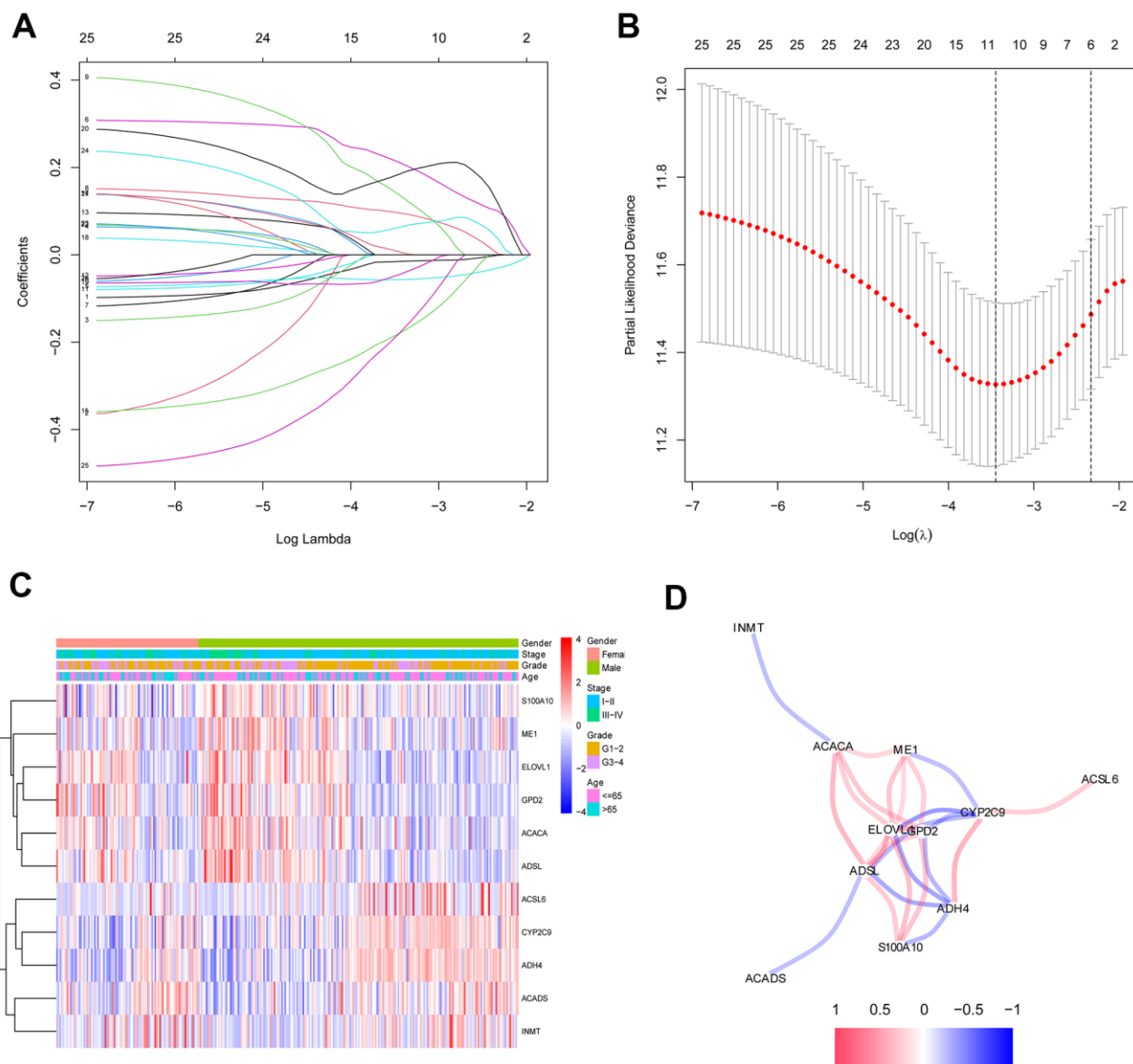
### Functional enrichment analysis of the screened genes

The most highly enriched molecular functions of the 6 screened genes were nicotinamide adenine dinucleotide

(NAD) binding, oxidoreductase activity and fatty acid activity (Figure 6A). The involvement of these molecular functions in fatty acid metabolism has been previously reported [14, 15]. Moreover, the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis suggested that the 6 genes might participate in pathways related to fatty acids (Figure 6B).

### Integrated analysis of the risk score and clinical parameters of patients with HCC

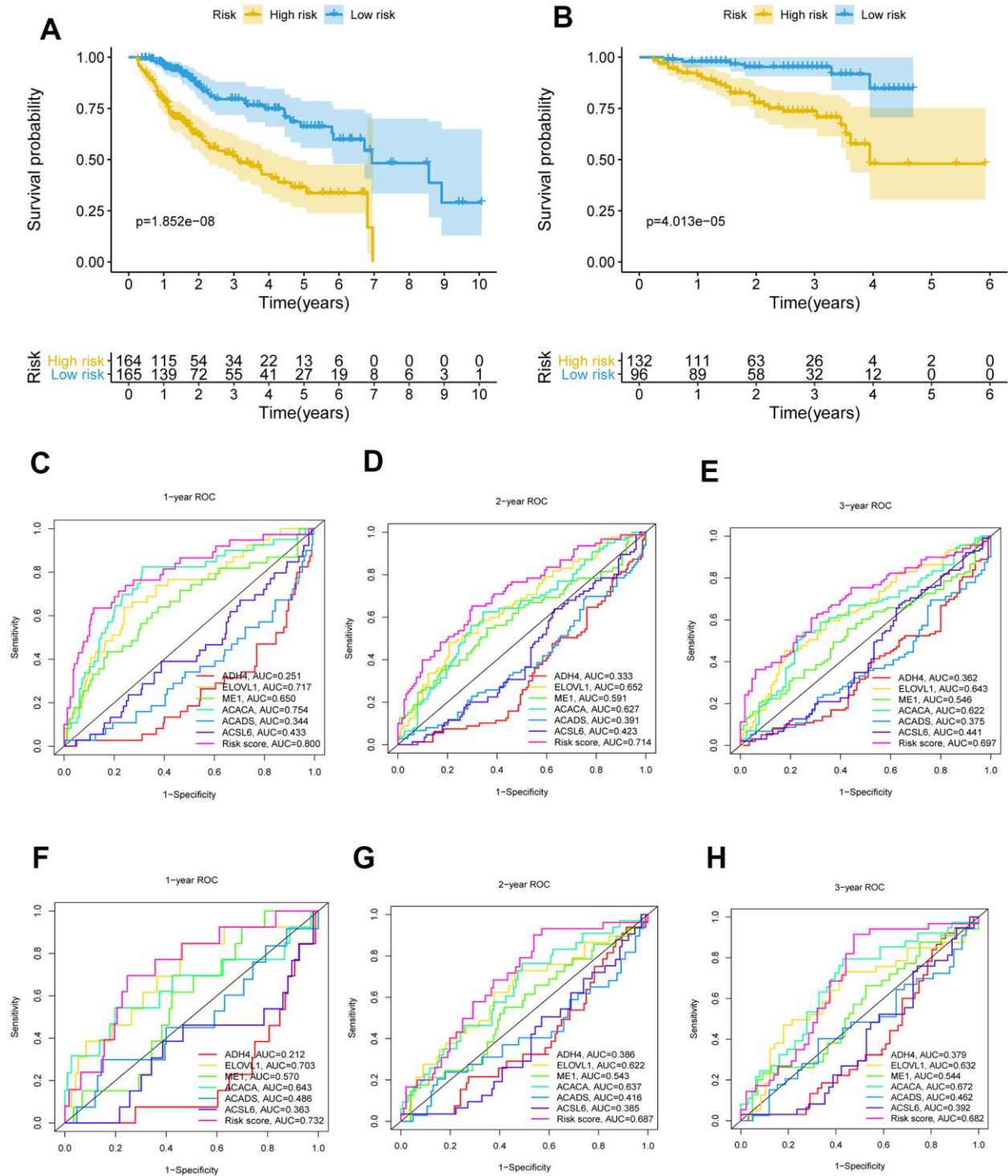
Comprehensive data were extracted for 306 patients by integrating the clinical data and risk score of patients



**Figure 3. Identification of fatty acid metabolic signatures by LASSO regression algorithm in HCC. (A)** LASSO coefficient profiles of the 26 fatty acid metabolic genes. **(B)** Cross-validation for tuning parameter selection in the proportional hazards model. **(C)** The heatmap of the 11 fatty acid metabolic genes. **(D)** The correlation of the 11 fatty acid metabolic genes.

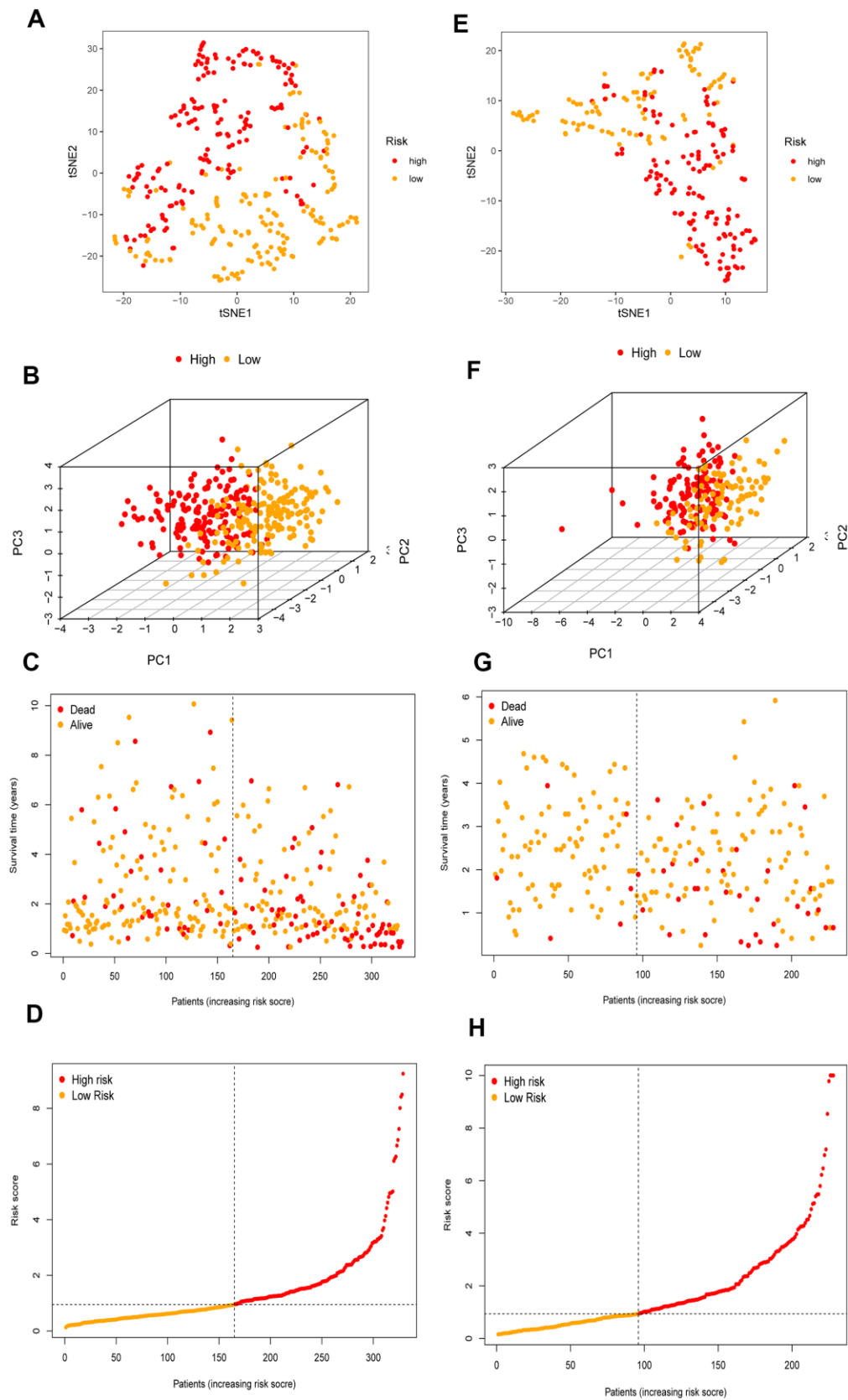
with HCC in TCGA database. First, univariate Cox regression analyses were conducted to filter the parameters (age, sex, stage, grade and risk score), and two factors (stage and risk score) were identified to be

correlated with the OS of patients with HCC ( $P < 0.001$ ) (Figure 7A). Then, the stage and risk score were suggested to be independent prognostic parameters of patients with HCC by the multivariable Cox regression



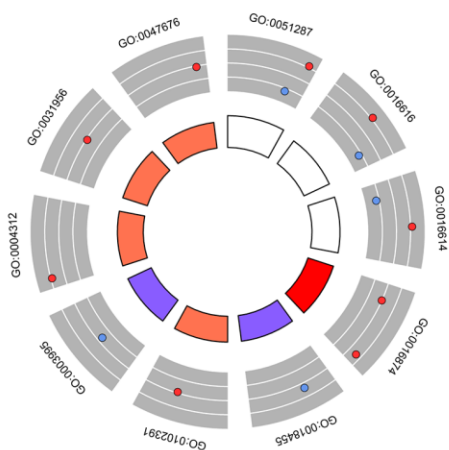
**Figure 4. K-M survival analysis and ROC curves of risk prognostic model in HCC patients. (A, B)** K-M survival analysis of risk prognostic model of HCC patients in TCGA and ICGC. **(C–E)** ROC curves analysis of risk prognostic model of HCC patients at 1 year, 2 years and 3 year in TCGA. **(F–H)** ROC curves analysis of risk prognostic model of HCC patients at 1 year, 2 years and 3 year in ICGC.





**Figure 5. The validation of the risk prognostic model in HCC patients. (A–D)** t-SNE, PCA, survival status scatter plots and risk score distribution shown the power prognostic ability of the risk prognostic model in TCGA. **(E–H)** t-SNE, PCA, survival status scatter plots and risk score distribution shown the power prognostic ability of the risk prognostic model in ICGC.

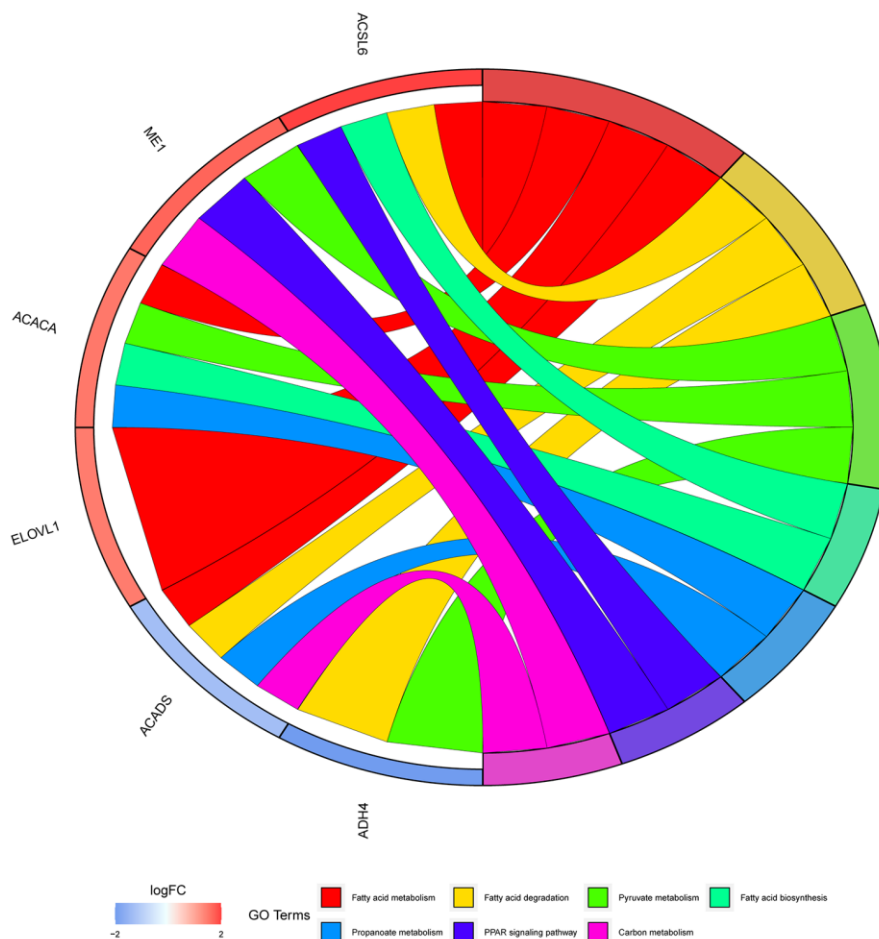
**A**



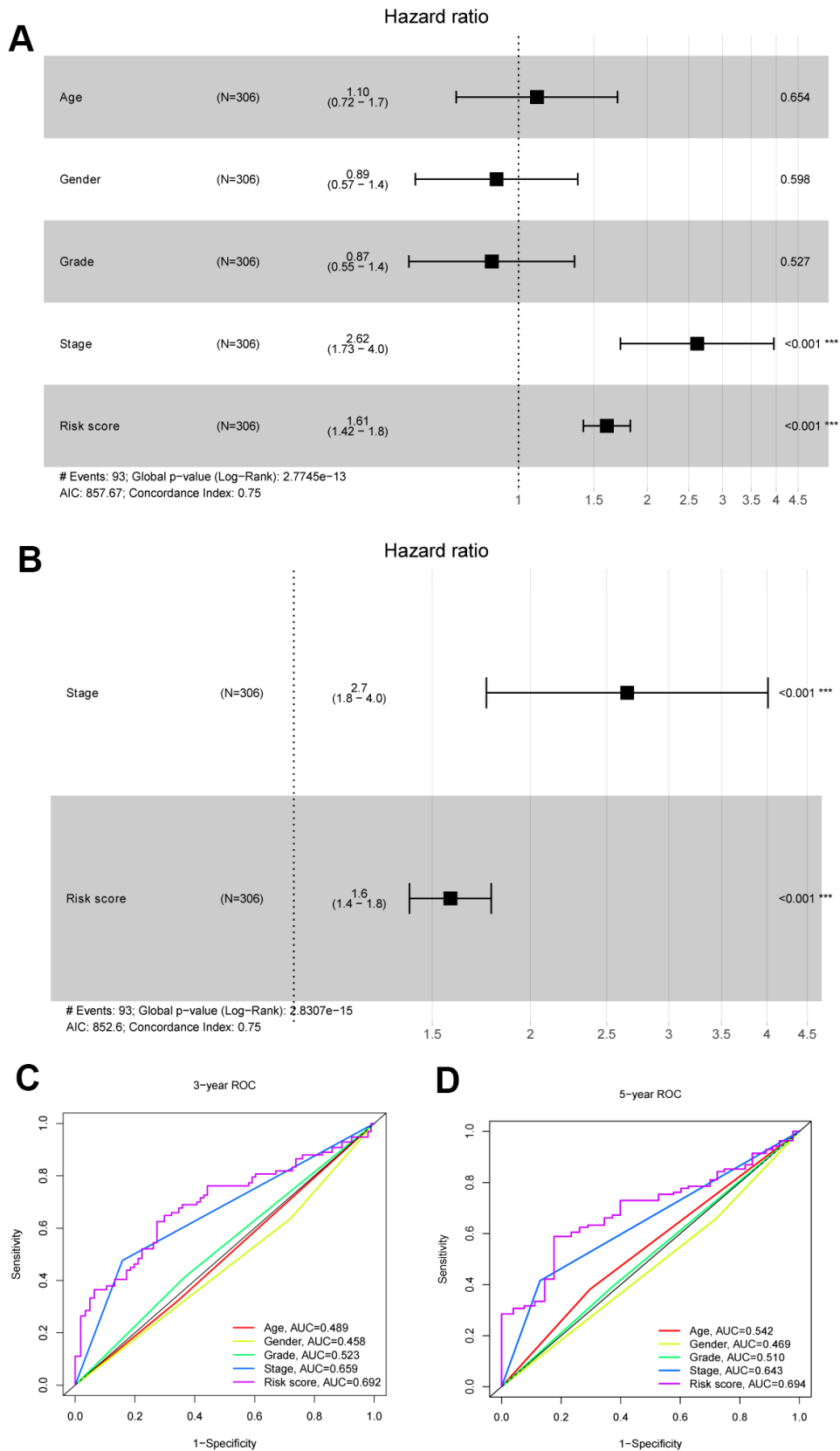
ID	Description
GO:0051287	NAD binding
GO:0016616	oxidoreductase activity, acting on the CH-OH group of donors, NAD or NADP as acceptor
GO:0016614	oxidoreductase activity, acting on CH-OH group of donors
GO:0016874	ligase activity
GO:0018455	alcohol dehydrogenase [NAD(P)+] activity
GO:0102391	decanoate-CoA ligase activity
GO:0003995	acyl-CoA dehydrogenase activity
GO:0004312	fatty acid synthase activity
GO:0031956	medium-chain fatty acid-CoA ligase activity
GO:0047676	arachidonate-CoA ligase activity



**B**



**Figure 6. Representative results of GO and KEGG analyses. (A)** The molecular functions of the 6 screened genes. **(B)** The potential biological pathways of the screened genes.



**Figure 7. Identification independent prognostic parameters in HCC patients.** (A, B) The univariate and multivariate Cox regression analysis of clinical parameters in patients with HCC. (C, D) ROC curves of risk score and clinical characteristics predicting 3- and 5-year survival in HCC patients.



algorithm ( $P < 0.001$ ) (Figure 7B). Figure 7C, 7D shows the highly sensitive and specific predictive performance of the risk score determined using the ROC curve analysis.

### **Construction and assessment of the nomogram for patients with HCC**

The risk score and stage were included in the construction of a nomogram with an optimal concordance index (C-index, 0.709) to further analyze the ability of independent prognostic factors to estimate individual OS (Figure 8A). The application of the nomogram to patients in the ICGC database also showed a satisfactory predictive effect (C-index, 0.727). In addition, the performance of the nomogram in predicting the 3- and 5-year OS rates of patients with HCC was calculated. A ROC curve analysis was performed to validate the practicability of the nomogram, and the AUCs were computed for 3- and 5-year survival (0.725 and 0.724, respectively) (Figure 8B, 8C). The calibration plot was close to the ideal curve (Figure 8D, 8E). Moreover, decision curve analysis (DCA), a novel reliable evaluation tool to quantify clinical values, showed that the nomogram obtained a better net benefit than a single independent predictive parameter (Figure 8F, 8G). Collectively, the prognostic capacity of the tumor prognostic model was verified from multiple perspectives.

### **Conjoint analysis of the tumor mutation burden and immune cell infiltration in patients with HCC**

The tumor mutation burden (TMB) has been suggested to be a marker for identifying patients with cancer who might benefit from immunotherapy and has been used to predict the curative effect of immune checkpoint inhibitors [16]. After intersecting the ID code of 364 HCC samples related to single nucleotide variation (SNV) and high- and low-risk patients with HCC, the SNV data of 156 high-risk HCC samples and 161 low-risk HCC samples were obtained. Figure 9A, 9B shows the difference in TMB in the high- and low-risk groups of patients with HCC. Moreover, the correlation analysis indicated that the TMB was positively correlated with the risk score ( $R = 0.14$ ;  $P = 0.012$ ) (Figure 9C). The 5-year survival rate of the low-risk and low-TMB group was better than that of the high-risk and high-TMB group ( $P < 0.001$ ) (Figure 9D).

The infiltration of 22 immune cell types in high- and low-risk patients with HCC is shown in Figure 10, which suggested a potential difference in the immune microenvironment in patients with HCC. Moreover, different types of immune cells infiltrated different risk

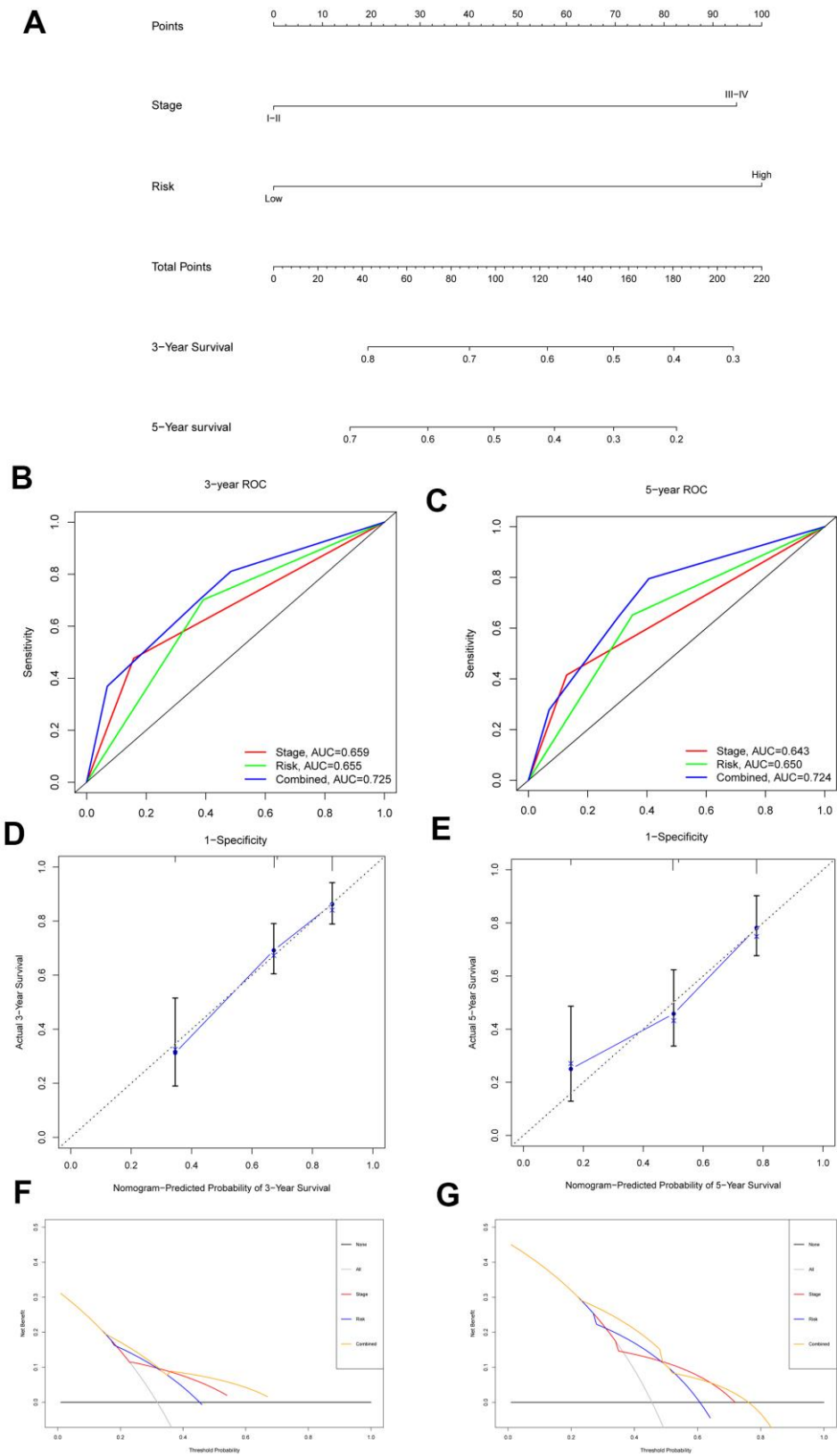
groups. In high-risk patients with HCC, higher expression of markers of M0 macrophages ( $P < 0.001$ ), follicular helper T cells ( $P = 0.003$ ), resting dendritic cells ( $P = 0.017$ ), memory B cells ( $P < 0.001$ ) and neutrophils ( $P < 0.001$ ) was observed (Figure 11A–11E), while lower expression of markers of M2 macrophages ( $P < 0.01$ ), resting mast cells ( $P < 0.01$ ), monocytes ( $P < 0.01$ ), resting NK cells ( $P < 0.01$ ), CD8 T cells ( $P = 0.027$ ), gamma delta T cells ( $P < 0.01$ ) and naive B cells ( $p = 0.017$ ) were observed than in the low-risk group (Figure 11F–11L).

## **DISCUSSION**

Convincing evidence has shown that one of the hallmarks of tumor cells, namely, metabolic dysregulation, contributes to abnormal cell biological behaviors, such as cell growth, angiogenesis, proliferation and invasion [4]. Abnormal glycolytic metabolism has already been confirmed in various tumors and has been widely used in clinical diagnosis and treatment [17]. Fatty acid metabolism, an important component of energy metabolism, has received increasing attention, and this metabolic pathway participates in energy production, membrane synthesis and signal transduction in tumor initiation and progression [7]. As a molecularly heterogeneous malignant tumor, the molecular characteristics of HCC are correlated with tumor biological behaviors [18]. Therefore, key molecular markers related to fatty acid metabolism in HCC must be identified.

Currently, gene signatures that correlate with the prediction of prognosis based on specific characteristics have become a hotspot in cancer research [19–21]. To the best of our knowledge, we have built a novel risk prediction model for HCC based on fatty acid metabolism-related genes and further constructed a clinical predictive model for patients with HCC that might be used to assess the individual's prognosis.

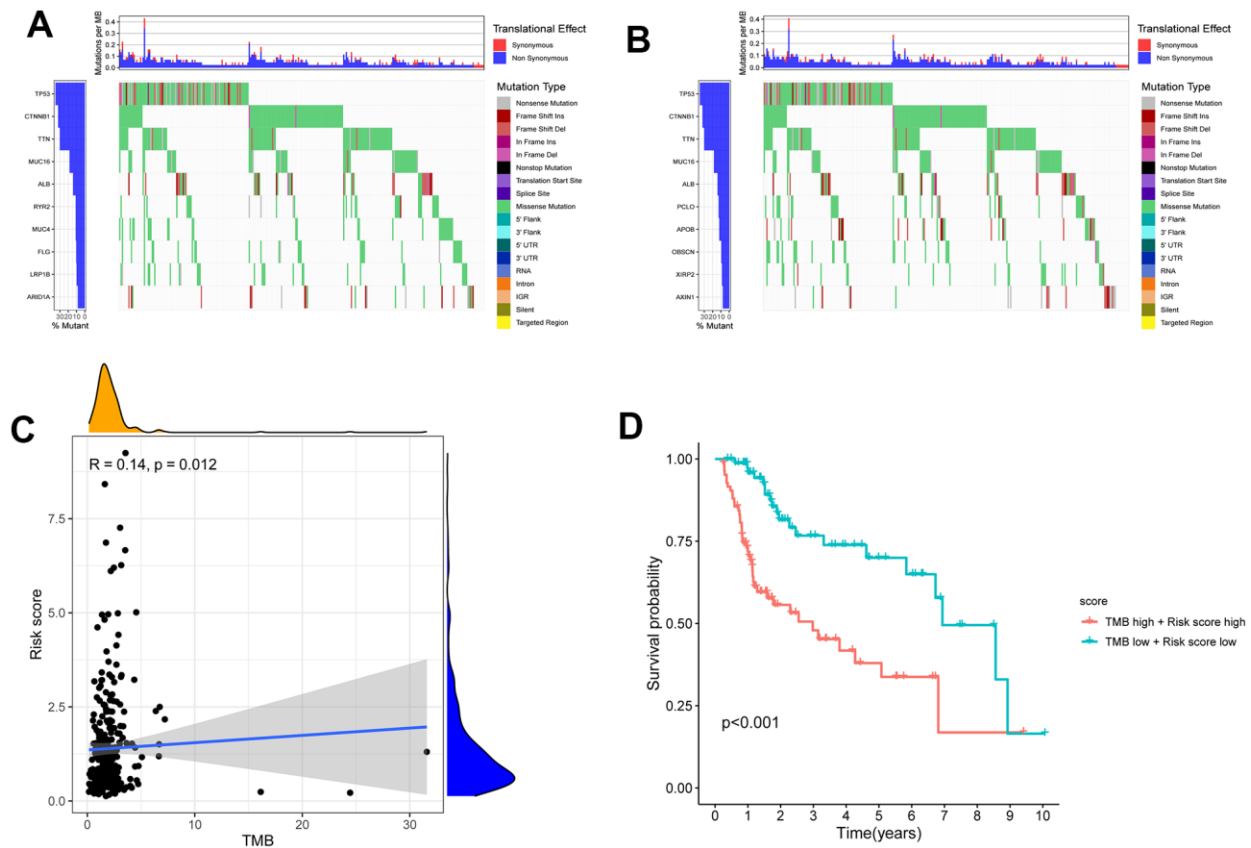
In the present study, we first systematically investigated fatty acid metabolism-related genes in patients with HCC. The 26 fatty acid metabolism DEGs related to OS in patients with HCC were identified with strict filter standards, and the TCGA and ICGC databases were used for validation. Furthermore, the LASSO regression algorithm was used to further filter the core genes related to fatty acid metabolism in patients with HCC. Multivariable Cox regression analysis was further used to identify 6 independent prognostic signatures, and a risk score model was constructed. The LASSO regression algorithm and multivariable Cox regression analysis are beneficial to avoid overfitting and improve the clinical practicability of the prognostic model.



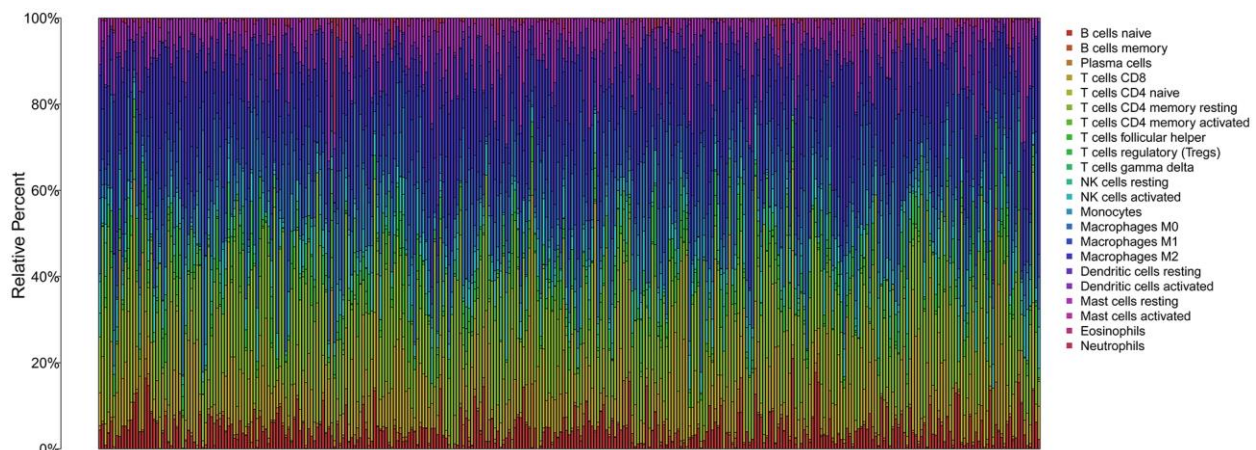
**Figure 8. Nomogram to predict 3- and 5- year OS and its validation in HCC patients. (A)** Nomogram to predict 3- and 5- year OS of HCC patients. **(B, C)** ROC curves to assess the accuracy of nomogram to predict 3- and 5- year OS in HCC patients. **(D, E)** Calibration plot analysis to assess the accuracy of nomogram to predict 3- and 5- year OS in HCC patients. **(F, G)** DCA to assess the accuracy of nomogram to predict 3- and 5- year OS in HCC patients.

Most of the 6 signature genes have been suggested to be involved in multiple cancers. However, the roles of these genes in determining the prognosis of HCC remain to be explored. *ADH4* and *ACADS* were

previously suggested to be related to HCC. Alcohol dehydrogenase 4 (*ADH4*) is mainly correlated with ethanol oxidation in the presence of intoxicating levels of alcohol and hydroxyl fatty acids [22, 23]. A few



**Figure 9.** The differences of TMB in high and low risk HCC patients. (A) The TMB in high risk HCC patients. (B) The TMB in low risk HCC patients. (C) The TMB was positively correlated with the risk score in patients with HCC. (D) The OS of the HCC patients of high risk and high TMB were lower than those in HCC patients with low risk and low TMB.



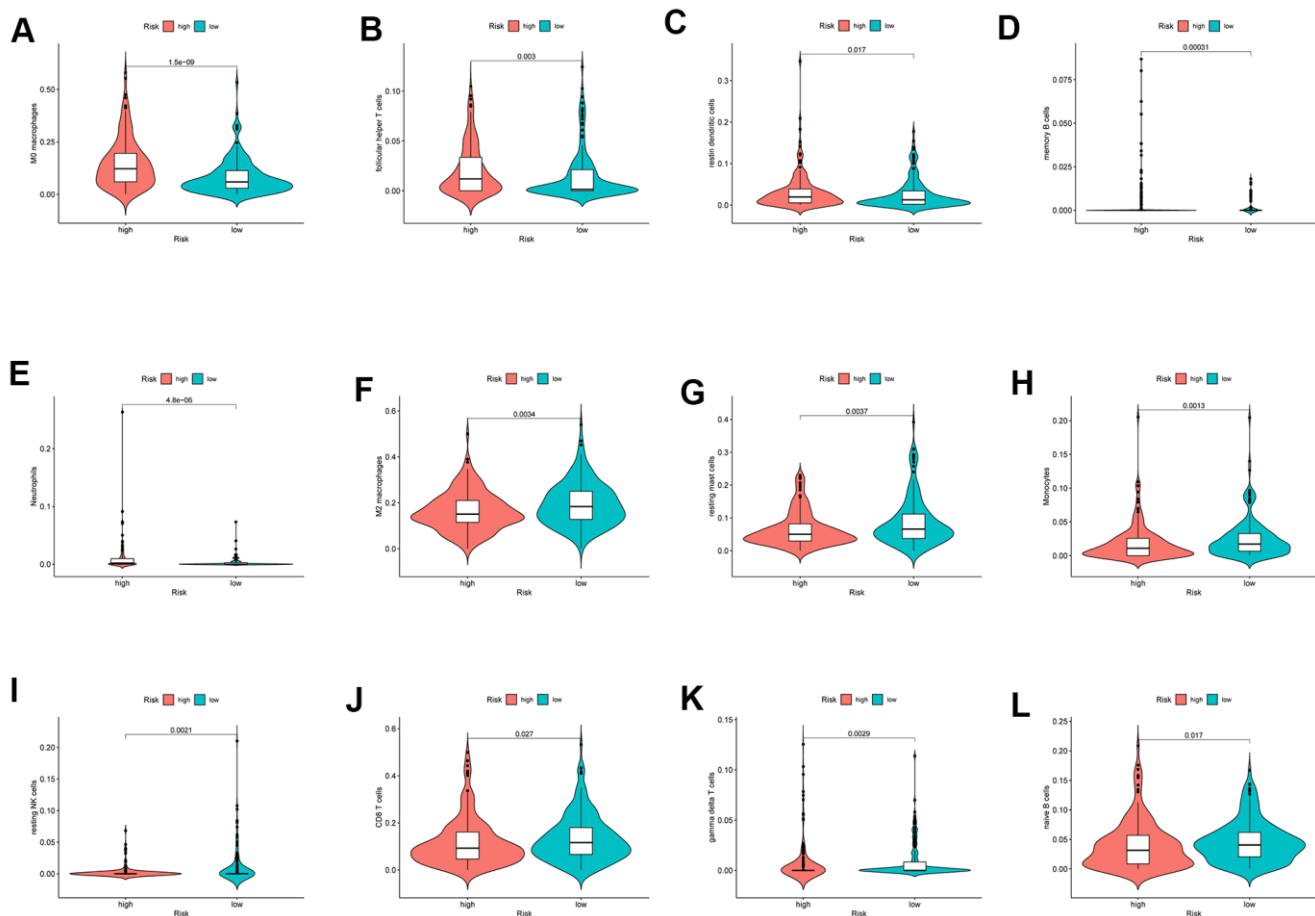
**Figure 10.** The immune infiltration of 22 immune cell types in high and low risk patients with HCC.

studies reported that *ADH4* might be a tumor suppressor and independent prognostic factor for HCC [24]. Our study again verified this speculation and further explored the ability of the gene to predict the risk score of patients with HCC. *ACADS* is a key short variant of the acyl-CoA dehydrogenase (*ACAD*) gene family and might be a methylation biomarker in HCC. The expression level of *ACADS* was negatively correlated with the DNA methylation level [25]. However, as a key metabolic enzyme, the role of *ACADS* in fatty acid metabolism in patients with HCC is unclear. The present study explored the role of *ACADS* in fatty acid metabolism in patients with HCC and included it in the construction of a risk score model.

In addition, elongation of very-long-chain fatty acids 1 (*ELOVL1*) is a single elongase catalyzing the synthesis of both saturated very-long-chain fatty acids (VLCFAs) and monounsaturated VLCFAs [26]. *ELOVL* enzymes, including *ELOVL1*, play different roles, possess characteristic substrate specificities and show differential

expression patterns in mammalian tissues. *ELOVL1* is expressed in multiple organs, including the lungs, kidney, skin and stomach, and dysfunctional *ELOVL1* proteins have been suggested to be correlated with different types of human diseases [27, 28]. *ELOVL1* has been suggested to be involved in various diseases, including cancers. However, the role of this gene related to fatty acid metabolism in HCC has not been reported in previous studies. In the present study, *ELOVL1* was upregulated. The gene correlated with OS and was included in the construction of the prognostic risk model for patients with HCC, which further suggested the importance of the roles of the gene in the origination and development of HCC. Thus, the genes related to fatty acid metabolism might be potential therapeutic and prognostic targets in HCC, and further exploration of the relationships between the genes and HCC is warranted.

After a risk predictive model was constructed based on the selected 6 genes, ROC curve, PCA and t-SNE



**Figure 11. The different immune infiltration in the high and low risk HCC patients.** (A–E) The expression of five types of immune cells is higher in high risk group compared with low risk group. (F–L) The expression of seven types of immune cells is higher in low risk group compared with high risk patients.

analyses were together used to suggest that patients with HCC could be divided into high- and low-risk groups with an optimistic definition based on the model. Then, the risk score was integrated with the clinical parameters (age, sex, stage, and grade) of patients with HCC, and the risk and stage were identified as independent prognostic factors, which further indicated the practicability of the model.

Next, we included the risk score based on the core fatty acid metabolism-related genes and stage to construct a tumor predictive model, and a nomogram was produced. The accuracy of the nomogram in calculating an individual's prognosis was assessed using the C-index, ROC curve, calibration plot analysis and DCA. DCA is a novel method for evaluating predictive models and better meets clinical requirements than a ROC curve analysis [29, 30]. The four validation methods together showed the practicability of the nomogram at different levels. The nomogram designed to predict the prognosis of patients with HCC was more accurate than the models reported in some previous studies with a better C-index (0.709 VS 0.676) and AUC for 3 years (0.725 VS 0.694) and 5 years (0.724 VS 0.667) [31, 32]. The construction of the model might help to identify patients with HCC and a poor prognosis and assist with the timely implementation of interventions.

Tumor immunotherapy is only effective for some patients with HCC, although the therapy has been shown to be a crucial therapeutic method by activating the body's own immune system. In recent years, a high TMB has been suggested to be correlated with the curative effect of immune checkpoint inhibitor treatments [33]. Moreover, the level of fatty acids in the microenvironment affects the function and phenotype of infiltrating immune cells [12, 34]. Studies have suggested that the effect of immunotherapy might be improved by regulating fatty acid metabolism in patients with cancer. In the present study, 12 types of infiltrating immune cells were identified between high- and low-risk patients with HCC, which might be beneficial for individualized immunotherapy and further improvements in the therapeutic effect.

Certainly, the potential limitations of the study should be noted. First, the follow-up time of the validation dataset from the ICGC database was insufficient, and thus the nomogram based on TCGA database must be validated using additional external datasets. Second, the in-depth molecular mechanisms of the fatty acid metabolism-related genes, including the genes included in prognostic models, must be further verified in experimental studies. In addition, the study was based

only on research data from public databases, which might contribute to selection bias. Therefore, a multicenter and large-scale study should be implemented to further validate the clinical utility of our model.

In summary, for the first time, this study established and validated a novel prognostic model based on 6 fatty acid metabolism-related genes and further constructed a clinical prognostic model for patients with HCC using strict standards. Moreover, the differences in infiltrating immune cells between high- and low-risk patients with HCC were explored, which might be helpful in providing a synergistic effect when administering treatments targeting fatty acid metabolism and immunotherapy. Therefore, our findings suggest that the novel fatty acid-related gene signature might be beneficial to the development of individualized treatments and improve the OS of patients with HCC.

## MATERIALS AND METHODS

### Patients and clinical data collection

Messenger RNA (mRNA) sequencing data from patients with HCC were obtained from TCGA database (<https://portal.gdc.cancer.gov/>) (374 HCC samples and 50 normal liver samples) and the ICGC database (<https://icgc.org/>) (243 HCC samples and 202 normal liver samples). In addition, the clinical information of patients with HCC (n=377, and 260, respectively) was downloaded from the TCGA and ICGC databases.

### Acquisition of fatty acid metabolism-related genes in patients with HCC

Three gene sets related to fatty acid metabolism (Hallmark fatty acid metabolism genes, KEGG fatty acid metabolism pathways, and Reactome fatty acid metabolism genes) were acquired from the Molecular Signature Database v7.2 (MSigDB), and fatty acid metabolism-related genes were retrieved after overlapping genes were removed [35]. The "Limma" R package was used to filter DEGs with a false discovery rate (FDR) <0.05 and fold change >2 [36].

### Construction and assessment of the risk score predictive model

The expression profile of fatty acid metabolism-related genes was extracted from the sequencing data of HCC and nontumor liver tissues in TCGA. After excluding patients with a follow-up time of less than 90 days, the survival data were integrated with sequencing data for fatty acid metabolism-related genes from patients with HCC in TCGA database. A univariate Cox regression



analysis was used to screen fatty acid metabolism-related genes associated with OS in patients with HCC ( $P < 0.05$ ), and then DEGs related to fatty acid metabolism were filtered by matching the DEGs in patients with HCC from TCGA.

To further improve the accuracy, the filtered fatty acid metabolism-related genes in patients with HCC were proven to be DEGs in the ICGC database. To avoid overfitting, the LASSO logistic regression algorithm was used to further filter core fatty acid metabolism-related genes correlated with OS.

Moreover, multivariable Cox regression analyses were subsequently conducted to identify independent prognostic genes and construct a risk score predictive model for patients with HCC based on TCGA database. Then, the model was tested in the ICGC database to further validate the feasibility of the model. Moreover, multiple test methods were conducted. Patients with HCC were separated into high-risk and low-risk groups according to the median risk score, and the K-M method was used to analyze the difference in survival. ROC, PCA, and t-SNE analyses were used to further validate the feasibility of the risk score predictive model in patients with HCC.

### **Functional enrichment analysis of the independent prognostic genes**

Gene Ontology (GO) analysis is a major bioinformatics tool for annotating genes and gene functions [37]. KEGG is a collection of databases that contain information about genomes, biological pathways, diseases, and chemical substances [38]. The GO analysis and KEGG pathway enrichment analysis of the independent prognostic genes were conducted using the “clusterProfiler” package [39]. An FDR value  $< 0.05$  was considered statistically significant.

### **Integrated analysis of the risk score predictive model and clinical parameters of patients with HCC**

The following inclusion criteria for clinical factors were used to completely identify independent pretreatment predictors in patients with HCC: 1) neoadjuvant radiotherapy or chemotherapy was not received; 2) the follow-up time was more than 90 days; and 3) an operation was performed on all patients with HCC. The clinical information (age, sex, grade and stage) and risk score of patients with HCC ( $n = 306$ ) were integrated according to the ID of patients with HCC after excluding missing data. Then, univariate and multivariable Cox regression analyses were conducted to identify independent prognostic indicators among the clinical factors ( $P < 0.05$ ).

### **Construction and assessment of the nomogram for patients with HCC**

To further investigate individual OS, the “nomogram” R package was used to build a predictive model for patients with HCC based on the independent clinical parameters. The C-index, ROC curve, calibration plot and DCA were used to weigh the prognostic ability of the nomogram from multiple perspectives.

### **Conjoint analysis of immune cell infiltration and the risk score in patients with HCC**

The SNV data for patients with HCC were downloaded from TCGA database, and the TMB was calculated. Spearman’s correlation analysis was used to explore the correlation between the risk score and TMB. K-M analysis was used to analyze the 5-year survival rate between patients with a high TMB and risk score and patients with a low TMB and risk score. Moreover, the “Cell Type Identification by Estimating Relative Subsets of RNA Transcripts (CIBERSORT)” deconvolution algorithm with 1,000 permutations was applied to quantify 22 types of tumor-infiltrating lymphocytes (TILs) in the microenvironment of high- and low-risk patients with HCC [40].

### **Abbreviations**

HCC: Hepatocellular carcinoma; TCGA: The Cancer Genome Atlas; ICGC: International Cancer Genome Consortium; MSigDB: Molecular Signature Database; DEGs: Differentially expressed genes; FDR: False discovery rate; OS: Overall survival; LASSO: Least absolute shrinkage and selection operator; ROC: Receiver operating characteristic curve; AUCs: Areas under the curve; K-M: Kaplan-Meier; PCA: Principal component analysis; t-SNE: t-distributed stochastic neighbor embedding; C-index: Concordance index; DCA: Decision curve analysis; TMB: Tumor mutation burden; SNV: Simple nucleotide variation; CIBERSORT: Cell type Identification By Estimating Relative Subsets Of RNA Transcripts; TILs: Tumor-infiltrating lymphocytes; KEGG: Kyoto Encyclopedia of Genes and Genomes; GO: Gene Ontology.

### **AUTHOR CONTRIBUTIONS**

Conceptualization, D.-S.H., M.-X.Y., and S.-Y.L.; Methodology, D.-S.H. and L.-F.C.; Validation, D.-S.H. and W.-M.H.; Formal Analysis, L.-F.C. and W.-M.H.; Investigation, Q.-Q.W. and X.L.; Writing – Original Draft Preparation, D.-S.H.; Writing – Review and Editing, S.-Y.L.; Visualization, D.-S.H. and M.-X.Y.; Supervision, M.-X.Y., and S.-Y.L.; Project Administration, M.-X.Y., and S.-Y.L.; Funding



Acquisition, M.-X.Y., S.-Y.L. and L.-F.C. All authors support publishing.

## ACKNOWLEDGMENTS

We thank the researchers who provided data for this study.

## CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

## FUNDING

This work was supported by grants from the youth scientific research project of Fujian provincial health and family planning commission (No. 2017-2-80), the project of Putian science and technology department (No. 2020SY006), and Key Clinical Specialty Construction Project of Putian city (N0.2021-2023). The funding body had no role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

## REFERENCES

1. International Agency for Research on Cancer, World Health Organization. Cancer today. 2020. <https://gco.iarc.fr/>
2. Sayiner M, Golabi P, Younossi ZM. Disease Burden of Hepatocellular Carcinoma: A Global Perspective. *Dig Dis Sci*. 2019; 64:910–17. <https://doi.org/10.1007/s10620-019-05537-2> PMID:30835028
3. Njei B, Rotman Y, Ditah I, Lim JK. Emerging trends in hepatocellular carcinoma incidence and mortality. *Hepatology*. 2015; 61:191–99. <https://doi.org/10.1002/hep.27388> PMID:25142309
4. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell*. 2011; 144:646–74. <https://doi.org/10.1016/j.cell.2011.02.013> PMID:21376230
5. Crunkhorn S. Targeting cancer cell metabolism in glioblastoma. *Nat Rev Cancer*. 2019; 19:250. <https://doi.org/10.1038/s41568-019-0139-3> PMID:30944412
6. Rosario SR, Long MD, Affronti HC, Rowsam AM, Eng KH, Smiraglia DJ. Pan-cancer analysis of transcriptional metabolic dysregulation using The Cancer Genome Atlas. *Nat Commun*. 2018; 9:5330. <https://doi.org/10.1038/s41467-018-07232-8> PMID:30552315
7. Currie E, Schulze A, Zechner R, Walther TC, Farese RV Jr. Cellular fatty acid metabolism and cancer. *Cell Metab*. 2013; 18:153–61. <https://doi.org/10.1016/j.cmet.2013.05.017> PMID:23791484
8. Madak-Erdogan Z, Band S, Zhao YC, Smith BP, Kulkoyluoglu-Cotul E, Zuo Q, Santaliz Casiano A, Wrobel K, Rossi G, Smith RL, Kim SH, Katzenellenbogen JA, Johnson ML, et al. Free Fatty Acids Rewire Cancer Metabolism in Obesity-Associated Breast Cancer via Estrogen Receptor and mTOR Signaling. *Cancer Res*. 2019; 79:2494–510. <https://doi.org/10.1158/0008-5472.CAN-18-2849> PMID:30862719
9. Wang YN, Zeng ZL, Lu J, Wang Y, Liu ZX, He MM, Zhao Q, Wang ZX, Li T, Lu YX, Wu QN, Yu K, Wang F, et al. CPT1A-mediated fatty acid oxidation promotes colorectal cancer cell metastasis by inhibiting anoikis. *Oncogene*. 2018; 37:6025–40. <https://doi.org/10.1038/s41388-018-0384-z> PMID:29995871
10. Han S, Wei R, Zhang X, Jiang N, Fan M, Huang JH, Xie B, Zhang L, Miao W, Butler AC, Coleman MA, Vaughan AT, Wang Y, et al. CPT1A/2-Mediated FAO Enhancement-A Metabolic Target in Radioresistant Breast Cancer. *Front Oncol*. 2019; 9:1201. <https://doi.org/10.3389/fonc.2019.01201> PMID:31803610
11. Wu Y, Fabritius M, Ip C. Chemotherapeutic sensitization by endoplasmic reticulum stress: increasing the efficacy of taxane against prostate cancer. *Cancer Biol Ther*. 2009; 8:146–52. <https://doi.org/10.4161/cbt.8.2.7087> PMID:19182512
12. Corn KC, Windham MA, Rafat M. Lipids in the tumor microenvironment: From cancer progression to treatment. *Prog Lipid Res*. 2020; 80:101055. <https://doi.org/10.1016/j.plipres.2020.101055> PMID:32791170
13. Veglia F, Tyurin VA, Blasi M, De Leo A, Kossenkov AV, Donthireddy L, To TK, Schug Z, Basu S, Wang F, Ricciotti E, DiRusso C, Murphy ME, et al. Fatty acid transport protein 2 reprograms neutrophils in cancer. *Nature*. 2019; 569:73–78. <https://doi.org/10.1038/s41586-019-1118-2> PMID:30996346
14. Ilbeigi D, Nourbakhsh M, Pasalar P, Meshkani R, Shokri Afra H, Panahi GH, Borji M, Sharifi R. Nicotinamide Phosphoribosyltransferase Knockdown Leads to Lipid Accumulation in HepG2 Cells through The SIRT1-AMPK Pathway. *Cell J*. 2020 (Suppl 1); 22:125–32. <https://doi.org/10.22074/cellj.2020.7013> PMID:32779442

15. Aleksenko VA, Anand D, Remeeva A, Nazarenko VV, Gordeliy V, Jaeger KE, Krauss U, Gushchin I. Phylogeny and Structure of Fatty Acid Photodecarboxylases and Glucose-Methanol-Choline Oxidoreductases. *Catalysts*. 2020; 10:1072.  
<https://doi.org/10.3390/catal10091072>
16. Chan TA, Yarchoan M, Jaffee E, Swanton C, Quezada SA, Stenzinger A, Peters S. Development of tumor mutation burden as an immunotherapy biomarker: utility for the oncology clinic. *Ann Oncol*. 2019; 30:44–56.  
<https://doi.org/10.1093/annonc/mdy495>  
PMID:30395155
17. Munkley J, Elliott DJ. Hallmarks of glycosylation in cancer. *Oncotarget*. 2016; 7:35478–89.  
<https://doi.org/10.18632/oncotarget.8155>  
PMID:27007155
18. Erstad DJ, Fuchs BC, Tanabe KK. Molecular signatures in hepatocellular carcinoma: A step toward rationally designed cancer therapy. *Cancer*. 2018; 124:3084–104.  
<https://doi.org/10.1002/cncr.31257>  
PMID:29663340
19. Cheng G, Liu D, Liang H, Yang H, Chen K, Zhang X. A cluster of long non-coding RNAs exhibit diagnostic and prognostic values in renal cell carcinoma. *Aging (Albany NY)*. 2019; 11:9597–615.  
<https://doi.org/10.18632/aging.102407>  
PMID:31727869
20. Song Q, Shang J, Yang Z, Zhang L, Zhang C, Chen J, Wu X. Identification of an immune signature predicting prognosis risk of patients in lung adenocarcinoma. *J Transl Med*. 2019; 17:70.  
<https://doi.org/10.1186/s12967-019-1824-4>  
PMID:30832680
21. Liu Y, Wu L, Ao H, Zhao M, Leng X, Liu M, Ma J, Zhu J. Prognostic implications of autophagy-associated gene signatures in non-small cell lung cancer. *Aging (Albany NY)*. 2019; 11:11440–62.  
<https://doi.org/10.18632/aging.102544>  
PMID:31811814
22. Ditlow CC, Holmquist B, Morelock MM, Vallee BL. Physical and enzymatic properties of a class II alcohol dehydrogenase isozyme of human liver: pi-ADH. *Biochemistry*. 1984; 23:6363–68.  
<https://doi.org/10.1021/bi00321a012>  
PMID:6397223
23. Höög JO, Strömberg P, Hedberg JJ, Griffiths WJ. The mammalian alcohol dehydrogenases interact in several metabolic pathways. *Chem Biol Interact*. 2003; 143–44:175–81.  
[https://doi.org/10.1016/s0009-2797\(02\)00225-9](https://doi.org/10.1016/s0009-2797(02)00225-9)  
PMID:12604202
24. Wei RR, Zhang MY, Rao HL, Pu HY, Zhang HZ, Wang HY. Identification of ADH4 as a novel and potential prognostic marker in hepatocellular carcinoma. *Med Oncol*. 2012; 29:2737–43.  
<https://doi.org/10.1007/s12032-011-0126-3>  
PMID:22147505
25. Chen D, Feng X, Lv Z, Xu X, Lu Y, Wu W, Wu H, Liu H, Cao L, Ye S, Chen J, Wu J. ACADS acts as a potential methylation biomarker associated with the proliferation and metastasis of hepatocellular carcinomas. *Aging (Albany NY)*. 2019; 11:8825–44.  
<https://doi.org/10.18632/aging.102292>  
PMID:31652420
26. Ofman R, Dijkstra IM, van Roermund CW, Burger N, Turkenburg M, van Cruchten A, van Engen CE, Wanders RJ, Kemp S. The role of ELOVL1 in very long-chain fatty acid homeostasis and X-linked adrenoleukodystrophy. *EMBO Mol Med*. 2010; 2:90–97.  
<https://doi.org/10.1002/emmm.201000061>  
PMID:20166112
27. Tvrdik P, Westerberg R, Silve S, Asadi A, Jakobsson A, Cannon B, Loison G, Jacobsson A. Role of a new mammalian gene family in the biosynthesis of very long chain fatty acids and sphingolipids. *J Cell Biol*. 2000; 149:707–18.  
<https://doi.org/10.1083/jcb.149.3.707> PMID:10791983
28. Bhandari S, Choe SK. The Role of a Fatty Acid Chain Elongase, Elov1, in Zebrafish and Its Implications for Kidney Disease in Humans. *The Molecular Nutrition of Fats*, Academic Press. 2019; 249–60.  
<https://doi.org/10.1016/B978-0-12-811297-7.00019-6>
29. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making*. 2006; 26:565–74.  
<https://doi.org/10.1177/0272989X06295361>  
PMID:17099194
30. Rousson V, Zumbo T. Decision curve analysis revisited: overall net benefit, relationships to ROC curve analysis, and application to case-control studies. *BMC Med Inform Decis Mak*. 2011; 11:45.  
<https://doi.org/10.1186/1472-6947-11-45>  
PMID:21696604
31. Wang Y, Ruan Z, Yu S, Tian T, Liang X, Jing L, Li W, Wang X, Xiang L, Claret FX, Nan K, Guo H. A four-methylated mRNA signature-based risk score system predicts survival in patients with hepatocellular carcinoma. *Aging (Albany NY)*. 2019; 11:160–73.  
<https://doi.org/10.18632/aging.101738>  
PMID:30631005
32. Tang B, Zhu J, Li J, Fan K, Gao Y, Cheng S, Kong C, Zheng L, Wu F, Weng Q, Lu C, Ji J. The ferroptosis and iron-metabolism signature robustly predicts clinical

- diagnosis, prognosis and immune microenvironment for hepatocellular carcinoma. *Cell Commun Signal.* 2020; 18:174.  
<https://doi.org/10.1186/s12964-020-00663-1>  
PMID:[33115468](https://pubmed.ncbi.nlm.nih.gov/33115468/)
33. Galuppini F, Dal Pozzo CA, Deckert J, Loupakis F, Fassan M, Baffa R. Tumor mutation burden: from comprehensive mutational screening to the clinic. *Cancer Cell Int.* 2019; 19:209.  
<https://doi.org/10.1186/s12935-019-0929-4>  
PMID:[31406485](https://pubmed.ncbi.nlm.nih.gov/31406485/)
34. Wu H, Han Y, Rodriguez Sillke Y, Deng H, Siddiqui S, Treese C, Schmidt F, Friedrich M, Keye J, Wan J, Qin Y, Kühl AA, Qin Z, et al. Lipid droplet-dependent fatty acid metabolism controls the immune suppressive phenotype of tumor-associated macrophages. *EMBO Mol Med.* 2019; 11:e10698.  
<https://doi.org/10.15252/emmm.201910698>  
PMID:[31602788](https://pubmed.ncbi.nlm.nih.gov/31602788/)
35. Zhang S, Chang W, Wu H, Wang YH, Gong YW, Zhao YL, Liu SH, Wang HZ, Svatek RS, Rodriguez R, Wang ZP. Pan-cancer analysis of iron metabolic landscape across the Cancer Genome Atlas. *J Cell Physiol.* 2020; 235:1013–24.  
<https://doi.org/10.1002/jcp.29017>  
PMID:[31240715](https://pubmed.ncbi.nlm.nih.gov/31240715/)
36. Wu D, Zhang P, Li F, Shen Y, Chen H, Feng Y, He A, Wang F. CD138<sup>+</sup> multiple myeloma cells express high level of CHK1 which correlated to overall survival in MM patient. *Aging (Albany NY).* 2020; 12:23067–81.  
<https://doi.org/10.18632/aging.104066>  
PMID:[33197893](https://pubmed.ncbi.nlm.nih.gov/33197893/)
37. Gaudet P, Škunca N, Hu JC, Dessimoz C. Primer on the Gene Ontology. *Methods Mol Biol.* 2017; 1446:25–37.  
[https://doi.org/10.1007/978-1-4939-3743-1\\_3](https://doi.org/10.1007/978-1-4939-3743-1_3)  
PMID:[27812933](https://pubmed.ncbi.nlm.nih.gov/27812933/)
38. Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* 2017; 45:D353–61.  
<https://doi.org/10.1093/nar/gkw1092> PMID:[27899662](https://pubmed.ncbi.nlm.nih.gov/27899662/)
39. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS.* 2012; 16:284–87.  
<https://doi.org/10.1089/omi.2011.0118>  
PMID:[22455463](https://pubmed.ncbi.nlm.nih.gov/22455463/)
40. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, Hoang CD, Diehn M, Alizadeh AA. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods.* 2015; 12:453–57.  
<https://doi.org/10.1038/nmeth.3337> PMID:[25822800](https://pubmed.ncbi.nlm.nih.gov/25822800/)