

A Novel Ferroptosis-Related Gene Signature For Clinically Predicting Recurrence After Hepatectomy of Hepatocellular Carcinoma Patients.

Huaxiang Wang

Fuzong clinical medical college of Fujian medical university

Fengfeng Xu

Fuzong clinical medical college of Fujian Medical University

Ruisheng Ke

The First Affiliated Hospital of Xiamen University

Huanzhang Hu

Fuzong clinical medical college of Fujian Medical University

Jian Fang

The third People's hospital of Fujian university of traditional medicine

Yi Jiang (✉ jiangyi1963@126.com)

Fuzong clinical medical college of Fujian Medical university <https://orcid.org/0000-0003-0825-6409>

Research Article

Keywords: Hepatocellular carcinoma, Ferroptosis, Gene signature, Recurrence-free survival, Nomogram, Decision curve analysis, Prognosis, TCGA, GEO, Genetic alteration

Posted Date: October 26th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-982158/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: The high HCC recurrence rate was the main reason for the poor prognosis after hepatectomy. In this present study, we identified a novel ferroptosis-related gene signature for clinically predicting HCC recurrence after hepatectomy.

Methods: Ferroptosis-related genes were obtained from the FerrDb database. We identified the ferroptosis-related differentially expressed genes (FDEGs) between HCC tissues and normal tissues from the GSE14520 dataset. These FDEGs were used to perform a univariate and multivariate regression analysis to construct HCC recurrence models. An independent HCC cohort from TCGA database was used to validate the reliability of the multi-gene HCC recurrence model. The predictive nomogram and DCA was built to estimate the recurrence predicting capacity of the multi-gene signature. GO, KEGG and GSEA were used to further investigate the potential mechanism of these FDEGs.

Results: A total of 39 FDEGs were identified. A seven-gene signature (MAPK9, SLC1A4, PCK2, ACSL3, STMN1, CDO1, and CXCL2) was constructed for HCC recurrence prediction. Patients in high-risk groups exhibited a significantly poor prognosis compared with low-risk patients in both the training set (GSE14520 cohort) and the validation set (TCGA cohort). Multivariate cox regression analysis demonstrated that the 7-gene signature were independent risk factors for RFS in HCC patients. The nomograms incorporating 7-gene signature and clinical prognostic risk factors were able to effectively predict RFS. KEGG analysis showed that FDEGs were mainly enriched in Ferroptosis, Hepatocellular carcinoma pathway, MAPK signaling pathway, and so on. GSEA analysis revealed that the high-risk group was enriched with multiple oncology characteristics and invasive-related pathways.

Conclusion: Our study constructed a seven ferroptosis-related gene signature and established a prognostic nomogram for clinically predicting recurrence after hepatectomy and offered novel research directions for personalized treatment in HCC patients.

Introduction

Primary liver cancer is one of the most common malignant tumors, in which hepatocellular carcinoma (HCC) accounts for about 85-90% of cases[1]. HCC causes more than 800,000 deaths per year and imposes a huge economic and health burden worldwide[2, 3]. According to data released by the American Cancer Society in 2021, the 5-year survival rate of HCC patients for all stages was only 20%[4]. With the improvement of diagnostic capabilities, the proportion of surgical resection of HCC has increased, but those who have undergone radical resection still have a 70% recurrence rate within 5 years[5]. The high recurrence rate of HCC was mainly responsible for the death of patients. Therefore, establishing an effective model for predicting postoperative recurrence and identifying high-risk patients early, and taking the initiative to take clinical actions is of great value to improve the prognosis. The traditional recurrence prediction model integrated tumor stage, tumor size, microvascular invasion, tumor differentiation, and other relevant clinical characteristics, and supplemented by a single serum alpha-fetoprotein

expression[6–8]. But their specificity and sensitivity were not enough to distinguish patients with heterogeneity.

Ferroptosis is a newly discovered cell death form that results from severe lipid peroxidation of intracellular iron overload and differs from apoptosis, necrosis, and autophagy in terms of morphology[9, 10]. In recent years, an increasing number of studies have revealed that ferroptosis plays a non-negligible role in regulating the growth and proliferation of some types of tumors[11–13]. Specifically, ferroptosis has a pivotal role in killing tumor cells and inhibiting tumor invasion and metastasis[14, 15]. A previous study has shown that ferroptosis was an effective method to induce HCC cell death and has the role of block the cytotoxic effect of sorafenib on HCC[16]. At present, some ferroptosis-related genes such as NRF2, NQO1, HCAR1, MCT1, ZFP36, etc. have been validated playing the role of cancer-promoting or suppressor factors in HCC[17–19]. However, rarely studies focus on the predictive effects of these ferroptosis-related genes on the recurrence of HCC patients.

In this current study, we constructed an HCC recurrence model using ferroptosis-related differentially expressed genes (FDEGs) obtained from the gene expression omnibus (GEO)[20] and FerrDb database[21]. Then, we validated the reliability of the multi-gene HCC recurrence model in an independent The Cancer Genome Atlas (TCGA) cohort[22]. The predictive nomogram and decision curve analysis (DCA) was built to estimate the recurrence predictive capacity of the multi-gene signature. In addition, we investigated the correlations between the genetic alteration of the seven-gene signature and recurrence-free survival (RFS) in the cBioPortal database[23]. Finally, Gene ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG) and gene set enrichment analyses (GSEA) was used to explore the intrinsic regulating mechanisms of these ferroptosis-related genes[24, 25]. On the whole, our results showed that the ferroptosis-related seven-gene signature and nomogram might help effectively predict the RFS of patients with HCC.

Methods

Gene datasets and clinical characteristics collection.

We downloaded the mRNA expression data and corresponding clinical characteristics data in the GSE14520 dataset and LIHC dataset. The ferroptosis-related gene list was obtained from the FerrDb database and literature published in PubMed. The GSE14520 dataset from the GEO database, which contains RNA sequencing and clinical data of 242 HCC patients, was used to construct the predictive model for recurrence. The liver hepatocellular carcinoma (LIHC) cohort from the TCGA database, containing gene expression and clinical data of 372 HCC patients, was used to validate the results of the predictive model. All gene expression and clinical data were obtained from the publicly available database, hence it was not required for additional ethical approval.

Identification of ferroptosis-related differentially expressed genes (FDEGs).

The R software (version 4.0.2) and built-in limma package were utilized to perform differentially expressed genes analysis using RNA sequence data between the HCC tumor tissues and paired normal tissues, which were downloaded from the GSE14520 dataset. The screening of differentially expressed genes must meet two standards: \log_2 fold change (FC) > 1.0 or \log_2 fold change (FC) < -1.0, adjusted P-value < 0.05. Next, the overlapping gene between the ferroptosis-related genes obtained from the FerrDb database and the differentially expressed genes identified from the GSE14520 dataset were screened as FDEGs.

Establishment and validation of the ferroptosis-related gene signature.

We performed the univariate Cox regression and Multivariate Cox regression on FDEGs sequentially to obtain the ferroptosis-related genes and their regression coefficients. These genes expression was an independent risk factor for the RFS of HCC patients and was used to establish a prognostic risk signature. Next, the 242 HCC patients from the GSE14520 dataset were separated into high-risk and low-risk groups based on the median Risk score, which was calculated based on the following formula: Risk score = $\sum_{i=1}^n v_i \times \beta_i$ (v represent the expression value of the gene and β represent the corresponding regression coefficients.) In addition, the Kaplan–Meier survival analysis and time-dependent receiver operating characteristic (ROC) curves were used to evaluate the predictive performance of this gene signature for RFS. We further investigated the correlations between prognostic gene signature and relative clinical characteristics. Moreover, we explored the independent prognostic role of the gene signature for RFS by the univariate and multivariate Cox regression analyses. In this process, some clinical characteristics were integrated, such as age, gender, tumor stage, tumor size, serum Alpha-fetoprotein (AFP) level, and so on. Finally, we validated the reliability of the risk score model using an independent LIHC cohort in the TCGA database.

Establishment and validation of a predictive nomogram.

We established a nomogram that integrated all the independent risk factors identified from the multivariate Cox regression analysis to predict the RFS of 1-year, 3-year, and 5-year. We calculated the concordance index (C-index) using the “survival” R package to assess the predictive performance of the nomogram. We next plotted calibration curves of RFS probability at different years. In addition, the time-dependent ROC curve was plotted via the “timeROC” R package to assess the performance of the nomogram. Furthermore, we performed the DCA analysis via the “ggDCA” R package to select the best model that has the highest clinical net benefit.

Genetic alteration and protein expression analysis of gene signature.

To investigate the effect of gene alterations on the aberrant expression, we queried the genetic alterations and mutation hotspot of gene signature using the liver Hepatocellular Carcinoma dataset (TCGA, Firehose Legacy) in the cBioPortal database. Then, we compared the overall survival (OS) probability, disease-free probability, progression-free survival probability, and disease-specific survival probability between the alteration groups with no alteration groups. Furthermore, we investigated the differential

expression of the seven-gene signature protein between the HCC tissues and adjacent normal liver tissues in the Human Protein Atlas database.

Functional enrichment analysis via GO, KEGG, and GSEA.

We performed the GO and KEGG enrichment analysis on the FDEGs using the DAVID database to explore the potential mechanism that these genes regulating the tumorigenesis and progression of HCC[26]. The results were visualized using the “clusterProfiler,” “enrichplot,” and “ggplot2” R packages.

RNA sequence (RNA seq) data in the TCGA database was selected to perform the GSEA enrichment analysis using the GSEA software (version 4.1.0). We separated the 372 HCC patients into high-risk groups and low-risk groups taking the median Risk score value as the critical point. In the process of GSEA, the KEGG gene set (c2.cp.kegg.v7.0.symbols.gmt) was selected as the functional gene set, and the number of permutations was set as 1000. Other parameters were set to default values. The adjusted p-value<0.05 and false discovery rate (FDR) q-value<0.25 were considered statistically significant.

Statistical analysis.

The R software (version 4.0.2) was utilized to perform the statistical analysis and plot the statistical figures. The association between the risk score with the clinicopathological characters was analyzed using Pearson’s chi-square test. Univariate and multivariate cox regression analysis was used to identify the risk factors or independent risk factors of RFS. Kaplan-Meier analysis with the log-rank test was utilized to compare the RFS between the high-risk group and low-risk group. The area under the curve (AUC) of ROC was utilized to estimate the predictive performance of gene signature. P<0.05 was considered as a statistically significant difference.

Results

FDEGs identification in HCC.

We exhibited the scheme of our study in a flow chart (Fig. 1). A total of 1014 differential expressed genes, including 539 down-regulated and 375 up-regulated genes, were identified in 242 HCC tissues compared with 246 adjacent normal liver tissues from the GSE14520 dataset. Next, we obtained 254 ferroptosis-related genes from the FerrDb database. Then, the overlapping 39 genes between the 1014 differential expressed genes with 254 ferroptosis-related genes were identified as FDEGs.

Establishment of the prognostic seven-gene signature.

The univariate Cox regression and multivariate Cox regression on 39 FDEGs were performed sequentially to identify the independent prognostic genes for RFS. After the Multivariate Cox regression, seven genes were identified to establish a predictive gene signature. These seven genes were mitogen-activated protein kinase 9 (MAPK9), solute carrier family 1 member 4 (SLC1A4), phosphoenolpyruvate carboxykinase 2 (PCK2), acyl-CoA synthetase long-chain family member 3 (ACSL3), stathmin 1 (STMN1),

cysteine dioxygenase type 1 (CDO1), and chemokine ligand 2 (CXCL2). The risk score = $(-0.167) * \text{expression}_{\text{MAPK9}} + (-0.086) * \text{expression}_{\text{SLC1A4}} + (-0.167) * \text{expression}_{\text{PCK2}} + 0.203 * \text{expression}_{\text{ACSL3}} + 0.201 * \text{expression}_{\text{STMN1}} + (-0.003) * \text{expression}_{\text{CDO1}} + 0.109 * \text{expression}_{\text{CXCL2}}$.

Table 1
Multivariate Cox regression analysis of the 7-gene signature

Gene	Coef	HR	Lower 95%CI	Upper 95%CI	P-Value
MAPK9	-0.167	0.497	0.330	0.750	0.001
SLC1A4	-0.086	0.600	0.387	0.929	0.022
PCK2	-0.167	0.516	0.315	0.847	0.009
ACSL3	0.203	1.784	1.105	2.879	0.018
STMN1	0.201	1.851	1.102	3.112	0.020
CDO1	-0.003	0.593	0.379	0.926	0.022
CXCL2	0.109	1.606	1.027	2.512	0.038
HR-hazard ratio; CI-confidence interval.					

Internal validation of the prognostic gene signature.

We calculated the seven-gene-based risk score for each HCC patient in the training set (GSE14520). Next, the 242 patients were separated into high-risk and low-risk groups based on the median Risk score (Fig. 2A). The Kaplan–Meier survival analysis and time-dependent ROC curves were used to evaluate the predictive performance of this gene signature for RFS. The ROC curve revealed that AUCs of 1-year, 3-year, 5-year for RFS were 0.68, 0.64, and 0.61 (Fig. 2B). Furthermore, The Kaplan–Meier survival analysis revealed that patients in the high-risk group exhibited a poorer RFS than patients in the low-risk group (Fig. 2C). In addition, the correlation analysis demonstrated that high-risk score correlated to tumor-node-metastasis (TNM) stage ($P=0.020$), serum AFP level ($P<0.001$), alanine aminotransferase (ALT) (0.025), predicted risk metastasis signature (PRMS) ($P<0.001$), recurrence ($P=0.001$), and death ($P=0.002$) (Table 2). Moreover, we explored the independent prognostic value of the gene signature for RFS by the univariate and multivariate Cox regression analyses. We found that gender ($P=0.009$), PRMS (0.006), tumor size ($P=0.045$), TNM stage($P<0.001$), and high-risk score ($P<0.001$) were risk factors of recurrence-free survival by performing univariate cox regression analysis. The multivariate Cox regression analysis confirmed that gender (HR (95%CI): 2.092(1.081-4.049); $P=0.028$), TNM stage (HR (95%CI): 2.608(1.262-3.389); $P=0.004$), and high-risk score (HR (95%CI): 1.879(1.215-2.906); $P=0.005$) were independent risk factors for RFS (Fi. 3). These results exhibited a favorite performance of the seven-gene signature for RFS prediction.

Table 2

Correlation between risk score and clinicopathological features of HCC patients for RFS in the GSE14520 HCC cohort.

Characteristics		N	Risk score level		χ^2	P-Value
			Low	High		
Age	>55	117	64	53	2.370	0.124
	<=55	125	56	69		
Gender	Male	211	106	105	0.279	0.598
	Female	31	14	17		
Main tumor size	>5cm	88	36	52	0.131	0.064
	<=5cm	154	84	70		
TNM stage	I/II	174	97	77	5.400	0.020
	III	51	19	32		
Serum AFP level	>300ng/ml	110	41	69	12.393	<0.001
	<=300ng/ml	128	77	51		
ALT	>50U/L	100	41	59	5.027	0.025
	<=50U/L	142	79	63		
Multinodular	Yes	52	22	30	1.404	0.236
	No	190	98	92		
Cirrhosis	Yes	223	107	116	2.926	0.087
	No	19	13	6		
PRMS classification	High	121	25	96	80.997	<0.001
	Low	121	95	26		
Recurrence	Yes	136	55	81	10.389	0.001
	No	106	65	41		
Death	Yes	96	36	60	9.299	0.002
	No	146	84	62		

TNM - tumor, node, metastasis, AFP - alpha fetoprotein, ALT- alanine aminotransferase, PRMS- Predicted risk Metastasis Signature. P-Value<0.05 were considered statistically significant.

Validation of the prognostic gene signature in the TCGA database.

We download the RNA seq and corresponding clinical characteristic data in the TCGA database to validate the performance of the seven-gene signature for RFS prediction. We calculated the seven-gene-based risk score for each HCC patient in the validation set (TCGA HCC cohort). The 372 patients were separated into high-risk and low-risk groups based on the median Risk score (Fig. 4A). The ROC analysis showed that the AUCs of 1-year, 3-year, 5-year for RFS were 0.69, 0.73, and 0.74, respectively (Fig. 4B). Furthermore, consistent with the result of the GSE14520 dataset, the Kaplan–Meier survival analysis revealed that patients in the high-risk group exhibited a poorer RFS than patients in the low-risk group (Fig. 4C). In addition, the correlation analysis demonstrated that high-risk score correlated to tumor grade ($P=0.026$), preoperative pharmaceutical ($P=0.036$), T 3/4($P=0.002$), lymph node invasion ($P=0.001$), metastasis ($P=0.048$), recurrence ($P<0.001$), and death ($P=0.048$) (Table 3). Moreover, we explored the independent prognostic value of the gene signature for RFS by the univariate and multivariate Cox regression analyses. We found that preoperative pharmaceutical ($P=0.044$), pathologic stage (0.003), stage 3/4($P<0.001$), lymph node invasion ($P=0.040$), and high-risk score ($P<0.001$) were risk factors of recurrence-free survival via univariate Cox regression analysis. The multivariate Cox regression analysis confirmed that T 3/4 (HR (95%CI): 2.056(1.320-3.204); $P=0.001$), and high-risk score (HR (95%CI): 1.779(1.286-2.462); $P=0.001$) were independent risk factors for RFS (Fig. 5). Overall, these results validated the good performance of the seven-gene signature for RFS prediction.

Table 3
Correlation between risk score and clinicopathological features of HCC patients for RFS in the TCGA HCC cohort.

Characteristics		N	Risk score level		χ^2	P-Value
			Low	High		
Age	>60	180	83	97	1.971	0.160
	<=60	191	102	89		
Gender	Male	250	128	122	0.546	0.460
	Female	121	57	64		
Race	White	192	91	101	0.971	0.325
	Other	179	94	85		
Tumor grade	G1/G2	236	128	108	4.959	0.026
	G3/G4	135	57	78		
Radiation	Yes	10	3	7	1.600	0.203
	No	361	182	179		
Pharmaceutical	Yes	24	7	17	4.397	0.036
	No	247	178	169		
Pathologic stage	I/II	266	137	129	1.009	0.315
	III/ IV	105	48	57		
T	T1/T2	260	143	117	9.165	0.002
	T3/T4	111	42	69		
N	Yes	15	1	14	11.669	0.001
	No	156	184	172		
M	Yes	13	3	10	3.867	0.048
	No	358	182	176		
Recurrence	Yes	171	62	109	23.496	<0.001
	No	200	123	77		
Death	Yes	130	55	75	4.572	0.032
	No	241	130	111		

Establishment and validation of a predictive nomogram.

We next established a nomogram that integrated all the independent risk factors including gender, risk score, and tumor stages identified from the multivariate Cox regression analysis to predict the RFS of 1-year, 3-year, and 5-year (Fig. 6A). The C-index of the combined nomogram model was 0.872. The calibration curve representing the actual and combined model-predictive RFS of the training set (GSE14520) at 1-year, 3-year, and 5-year exhibited an accurate performance for predicting recurrence of the nomogram (Fig. 6B-D). Furthermore, the ROC analysis showed that the AUCs for 1-year, 3-year, 5-year RFS predictions were 0.824, 0.807, and 0.762, respectively (Fig. 6E). In addition, the DCA curve exhibited the best net benefit of the combined model for 1-year, 3-year, and 5-year RFS prediction compared with the individual predictive factors (Fig. 6F-H). Summarizing, these results demonstrated that the combined model of nomogram exhibited an excellent predictive ability for 1-year, 3-year, and 5-year RFS of HCC patients, which might help the clinical therapy decision.

Genetic alteration of gene signature correlated with poor survival probability.

We queried the genetic alteration of the seven-gene signature in the Liver Hepatocellular Carcinoma cohort (TCGA, Firehose Legacy) of the cBioPortal database. Among 352 HCC patients, 95 patients (27.0%) shown genetic alterations of gene signature (Fig. 7A). In addition, the patients with genetic alteration have poorer overall survival probability ($P=2.98e-3$, Fig. 7B), disease-free probability ($P=0.0274$, Fig. 7C), progression-free survival probability ($P=0.0474$, Fig. 7D), and disease-specific survival probability ($P=0.0118$, Fig. 7E) than the patients without genetic alterations. We further investigate the protein expression of gene signature between the HCC tissues and adjacent normal liver tissues in the Human Protein Atlas database. In HCC tissues, the expression of MAPK9, SLC1A4, ACSL3, and STMN1 proteins increased, while the expression of PCK2 and CDO1 proteins decreased (Fig. 7F). However, we didn't find the CXCL2 protein expression in the Human Protein Atlas database.

GO, KEGG enrichment analysis and protein-protein interaction (PPI) network constructions of FDEGs.

We performed the GO and KEGG enrichment analysis on the 39 FDEGs to explore the potential mechanism that these genes regulating the tumorigenesis and progression of HCC. GO analysis revealed that in the biological process, the FDEGs significantly enriched in the pathway of response to nutrient levels, response to the metal ion, response to starvation, response to oxidative stress, etc (Fig. 8A). In the cellular component, the FDEGs are mainly enriched in neuron projection cytoplasm, pigment granules, melanosome, etc (Fig. 8B). In the molecular function, the FDEGs are mainly enriched in protein serine kinase activity, decanoate-CoA ligase activity, etc. (Fig. 8C) KEGG analysis showed that the FDEGs significantly enriched in the pathway of Ferroptosis, Hepatocellular carcinoma, MAPK signaling pathway, and other cancer-related pathways (Fig. 8D). We further constructed a PPI network of these FDEGs in the STRING database and visualized it utilizing the Cytoscape software (Fig. 8E).

GSEA enrichment analysis.

We performed the GSEA analysis to further investigate the significant signaling pathway that genes of the high-risk score and low-risk score patients were enriched. We found that the high-risk score group was

significantly correlated with pathways of cancers, such as cell cycle (NES=1.5, P<0.001), notch signaling pathway (NES=1.5, P<0.001), pathway in cancer (NES=1.5, P<0.001), VEGF signaling pathway (NES=1.7, P<0.001), etc (Fig. 8F). Meanwhile, the low-risk score group was negatively correlated with the signaling pathway of lysine degradation (NES=-2.2, P<0.001), peroxisome (NES=-2.1, P<0.001), propanoate metabolism (NES=-1.9, P<0.001), etc (Fig. 8G).

Discussion

Cancer, a currently urgent public health issue that should be addressed, has brought a tremendous economic and health burden worldwide[27]. HCC, a type of highly aggressive cancer, was a primary cause of cancer-related death in many areas of the world, especially in East Asia and sub-Saharan Africa[28, 29]. According to data released by the American Cancer Society in 2021, the 5-year survival probability of HCC patients for all stages was only 20%[4]. With the heightened health-conscious and improved diagnostic capacity, an increasing number of HCC patients were diagnosed during the physical examinations, and those patients can be treated by curative surgery. However, studies revealed that those who have undergone radical resection still have a 70% recurrence rate within 5 years[5, 30–32]. The high recurrence rate was the main responsibility for the short overall survival and poor prognosis of HCC patients. Therefore, establishing an effective model for predicting postoperative recurrence and identifying high-risk patients early, and taking the initiative to take clinical actions is of great value to improve the prognosis. The conventional predictive model of recurrence integrated tumor stage, tumor size, microvascular invasion, tumor differentiation, and other relevant clinical characteristics, and some supplemented by a single serum alpha-fetoprotein expression. But their specificity and sensitivity were not enough to distinguish patients with heterogeneity. In recent years, the gene signature based on the mRNA aberrant expression has been reported to address the problems of heterogeneity and exhibited an amplified diagnostic sensitivity and specificity. Massive studies focus on establishing prognostic-related gene signature models to improve diagnostic efficiency and overall survival prognosis[33]. Wang et al. established an RNA-binding proteins-related gene signature to predict the overall survival and found that this gene signature can be an independent risk factor for HCC patients[34]. Yang and colleagues constructed a two-gene signature (HNRNPA2B1 and RBM15) to identify and treat HBV-related HCC patients and has the predictive value for OS[35]. However, rarely are studies devoted to investigating the recurrence-related gene signature of HCC[36].

In recent years, it has been confirmed that ferroptosis plays a significant role in inducing HCC cell death and inhibiting cell proliferation and metastasis, and the ferroptosis-related genes were the carriers of this function. Previous research revealed that DAZAP1 was the ferroptosis suppressor gene and significantly overexpressed in HCC cells. DAZAP1 also promoted proliferation and significantly reduced the cellular sensitivity to sorafenib[37]. Another study reported that metallothionein (MT)-1G increases the sorafenib-resistance of HCC cells by inhibiting the process of ferroptosis[38]. Furthermore, a recent study revealed that ACSL4, a positive-activating enzyme of ferroptosis, can increase the sensitivity of HCC patients to sorafenib by activating the ferroptosis[39–41]. Some studies focus on the predictive value of ferroptosis-

related gene signature for overall survival of HCC patients, but rarely research to establish the gene signature to predict the RFS probability[42–45].

In this current study, we established a novel ferroptosis-related seven-gene signature (including MAPK9, SLC1A4, PCK2, ACSL3, STMN1, CDO1, and CXCL2) for HCC RFS prediction. The seven-gene signature exhibited an excellent predictive performance in the training set (GSE14520). The Kaplan–Meier survival analysis revealed that patients in the high-risk group exhibited a poorer RFS than patients in the low-risk group. And the correlation analysis demonstrated that high-risk scores correlated to TNM stage, serum AFP level, ALT, predicted risk metastasis PRMS, recurrence, and death. Moreover, the multivariate Cox regression analyses revealed that high-risk score was independent risk factors for RFS. All these results were also verified in the validation set (TCGA HCC cohort).

We next established a nomogram that integrated all the independent risk factors to predict the RFS of 1-year, 3-year, and 5-year. The C-index of the combined nomogram model was 0.872. The calibration curve representing the actual and combined model-predictive RFS at 1-year, 3-year, and 5-year exhibited an accurate performance for predicting recurrence of the nomogram. In addition, the DCA curve exhibited the best net benefit of the combined model for 1-year, 3-year, and 5-year RFS prediction compared with the individual predictive factors. Summarizing, these results demonstrated that the combined model of nomogram exhibited an excellent predictive ability for 1-year, 3-year, and 5-year RFS of HCC patients, which might help the clinical therapy decision. We also queried the genetic alteration of the seven-gene signature in the cBioPortal database and found these genes alteration correlated with poor survival probability. Moreover, the KEGG analysis revealed that the FDEGs significantly enriched in the pathway of Ferroptosis, Hepatocellular carcinoma, MAPK signaling pathway, and other cancer-related pathways. Finally, GSEA further revealed several significantly enriched oncological signaling pathways, which might explain the underlying molecular mechanisms of the gene signature.

Conclusion

Our study constructed a seven ferroptosis-related gene signature and established a prognostic nomogram for clinically predicting recurrence after hepatectomy and offered novel research directions for personalized treatment in HCC patients.

Abbreviations

HCC: hepatocellular carcinoma; FDEGs: ferroptosis-related differentially expressed genes; TCGA: The Cancer Genome Atlas; DCA: decision curve analysis; GO: Gene Ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes; GSEA: Gene set enrichment analyses; RFS: Recurrence-free survival; LIHC: The liver hepatocellular carcinoma; ROC: Receiver operating characteristic; AFP: Alpha-fetoprotein; C-index: concordance index; OS: Overall survival; FDR: false discovery rate; AUC: Area under the curve; MAPK9: mitogen-activated protein kinase 9; SLC1A4: solute carrier family 1 member 4; PCK2: phosphoenolpyruvate carboxykinase 2; ACSL3: acyl-CoA synthetase long-chain family member 3;

STMN1: stathmin 1; CDO1: cysteine dioxygenase type 1; CXCL2: chemokine ligand 2; TNM: tumor-node-metastasis; PRMS: predicted risk metastasis signature; HR: Hazard ratio; CI: Confidence interval; PPI: Protein-protein interaction; NES: Normal enrichment score.

Declarations

Acknowledgements

The authors are grateful to the contributors of the GEO and TCGA for sharing the HCC expression profile data set on open access. In addition, we would like to acknowledge to all the people who have given us help on our article.

Authors' contributions

YJ and JF had designed the study. HW, FX and RK had collected data. HW, FX and RK had analyzed and interpreted the data. All authors were involved in writing paper and approved of the submitted and published versions. All authors read and approved the final manuscript.

Funding

This work was supported by The Key Project of Natural Science Foundation of Fujian Province, (No. 2020Y0078), the Surface Project of Natural Science Foundation of Fujian Province (No. 2020J011144), the Civil-Military Integration Project of the 900th Hospital (No: 2018J06).

Data Availability

The datasets generated and/or analyzed during the current study are available from the corresponding author on reasonable request.

Ethics approval

This article does not contain any studies with human participants or animals performed by any of the authors. All gene expression and clinical data were obtained from the publicly available database, hence it was not required for additional ethical approval.

Animal Research

Not applicable.

Consent to participate

This article does not contain any studies with human participants or animals performed by any of the authors.

Consent to Publish

All participants provided written informed consent for publication of the data and any associated images.

Clinical Trials Registration

Not applicable.

Plant Reproducibility

Not applicable.

Conflicts of interest

The authors Huaxiang Wang, Fengfeng Xu, Ruisheng Ke, Huanzhang Hu, Yi Jiang, and Jian Fang declare that there are no conflicts of interest.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2018;68(6):394–424.
2. Forner A, Reig M, Bruix J. Hepatocellular carcinoma. *Lancet.* 2018;391(10127):1301–14.
3. Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, et al. Cancer statistics in China, 2015. *CA Cancer J Clin.* 2016;66(2):115–32.
4. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer Statistics. 2021. *CA Cancer J Clin.* 2021;71(1):7-33.
5. Zhou T, Cai Z, Ma N, Xie W, Gao C, Huang M, et al. A Novel Ten-Gene Signature Predicting Prognosis in Hepatocellular Carcinoma. *Front Cell Dev Biol.* 2020;8:629.
6. Chan AWH, Zhong J, Berhane S, Toyoda H, Cucchetti A, Shi K, et al. Development of pre and post-operative models to predict early recurrence of hepatocellular carcinoma after surgical resection. *J Hepatol.* 2018;69(6):1284–93.
7. Villanueva A, Hoshida Y, Battiston C, Tovar V, Sia D, Alsinet C, et al. Combining clinical, pathology, and gene expression data to predict recurrence of hepatocellular carcinoma. *Gastroenterology.* 2011;140(5):1501–12 e1502.
8. Ren A, Li Z, Zhou X, Zhang X, Huang X, Deng R, et al. Evaluation of the Alpha-Fetoprotein Model for Predicting Recurrence and Survival in Patients With Hepatitis B Virus (HBV)-Related Cirrhosis Who Received Liver Transplantation for Hepatocellular Carcinoma. *Front Surg.* 2020;7:52.
9. Yu H, Guo P, Xie X, Wang Y, Chen G. Ferroptosis, a new form of cell death, and its relationships with tumourous diseases. *J Cell Mol Med.* 2017;21(4):648–57.
10. Stockwell BR, Jiang X, Gu W. Emerging Mechanisms and Disease Relevance of Ferroptosis. *Trends Cell Biol.* 2020;30(6):478–90.

11. Mou Y, Wang J, Wu J, He D, Zhang C, Duan C, et al. Ferroptosis, a new form of cell death: opportunities and challenges in cancer. *J Hematol Oncol*. 2019;12(1):34.
12. Wang Y, Wei Z, Pan K, Li J, Chen Q. The function and mechanism of ferroptosis in cancer. *Apoptosis*. 2020;25(11-12):786–98.
13. Liang C, Zhang X, Yang M, Dong X. Recent Progress in Ferroptosis Inducers for Cancer Therapy. *Adv Mater*. 2019;31(51):e1904197.
14. Friedmann Angeli JP, Krysko DV, Conrad M. Ferroptosis at the crossroads of cancer-acquired drug resistance and immune evasion. *Nat Rev Cancer*. 2019;19(7):405–14.
15. Hassannia B, Vandenabeele P, Vanden Berghe T. Targeting Ferroptosis to Iron Out Cancer. *Cancer Cell*. 2019;35(6):830–49.
16. Louandre C, Ezzoukhry Z, Godin C, Barbare JC, Maziere JC, Chauffert B, et al. Iron-dependent cell death of hepatocellular carcinoma cells exposed to sorafenib. *Int J Cancer*. 2013;133(7):1732–42.
17. Zhao Y, Li M, Yao X, Fei Y, Lin Z, Li Z, et al. HCAR1/MCT1 Regulates Tumor Ferroptosis through the Lactate-Mediated AMPK-SCD1 Activity and Its Therapeutic Implications. *Cell Rep*. 2020;33(10):108487.
18. Sun X, Ou Z, Chen R, Niu X, Chen D, Kang R, et al. Activation of the p62-Keap1-NRF2 pathway protects against ferroptosis in hepatocellular carcinoma cells. *Hepatology*. 2016;63(1):173–84.
19. Zhang Z, Guo M, Li Y, Shen M, Kong D, Shao J, et al. RNA-binding protein ZFP36/TTP protects against ferroptosis by regulating autophagy signaling pathway in hepatic stellate cells. *Autophagy*. 2020;16(8):1482–505.
20. Barrett T, Edgar R. Gene expression omnibus: microarray data storage, submission, retrieval, and analysis. *Methods Enzymol*. 2006;411:352–69.
21. Zhou N, Bao J. FerrDb: a manually curated resource for regulators and markers of ferroptosis and ferroptosis-disease associations. *Database (Oxford)*. 2020;2020.
22. Cancer Genome Atlas Research N, Weinstein JN, Collisson EA, Mills GB, Shaw KR, Ozenberger BA, et al. The Cancer Genome Atlas Pan-Cancer analysis project. *Nat Genet*. 2013;45(10):1113–20.
23. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal*. 2013;6(269):pl1.
24. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*. 2005;102(43):15545–50.
25. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*. 2000;28(1):27–30.
26. Dennis G Jr, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, et al. DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol*. 2003;4(5):P3.
27. Global Burden of Disease. Cancer C, Fitzmaurice C, Allen C, Barber RM, Barregard L, Bhutta ZA, et al. Global, Regional, and National Cancer Incidence, Mortality, Years of Life Lost, Years Lived With

- Disability, and Disability-Adjusted Life-years for 32 Cancer Groups, 1990 to 2015: A Systematic Analysis for the Global Burden of Disease Study. *JAMA Oncol.* 2017;3(4):524–48.
28. Yang JD, Hainaut P, Gores GJ, Amadou A, Plymoth A, Roberts LR. A global view of hepatocellular carcinoma: trends, risk, prevention and management. *Nat Rev Gastroenterol Hepatol.* 2019;16(10):589–604.
 29. Villanueva A. Hepatocellular Carcinoma. *N Engl J Med.* 2019;380(15):1450–62.
 30. Hasegawa K, Kokudo N, Makuuchi M, Izumi N, Ichida T, Kudo M, et al. Comparison of resection and ablation for hepatocellular carcinoma: a cohort study based on a Japanese nationwide survey. *J Hepatol.* 2013;58(4):724–9.
 31. Zhang X, Li C, Wen T, Yan L, Li B, Yang J, et al. Appropriate treatment strategies for intrahepatic recurrence after curative resection of hepatocellular carcinoma initially within the Milan criteria: according to the recurrence pattern. *Eur J Gastroenterol Hepatol.* 2015;27(8):933–40.
 32. Xu XF, Xing H, Han J, Li ZL, Lau WY, Zhou YH, et al. Risk Factors, Patterns, and Outcomes of Late Recurrence After Liver Resection for Hepatocellular Carcinoma: A Multicenter Study From China. *JAMA Surg.* 2019;154(3):209–17.
 33. Liu GM, Zeng HD, Zhang CY, Xu JW. Identification of a six-gene signature predicting overall survival for hepatocellular carcinoma. *Cancer Cell Int.* 2019;19:138.
 34. Wang L, Zhou N, Qu J, Jiang M, Zhang X. Identification of an RNA binding protein-related gene signature in hepatocellular carcinoma patients. *Mol Med.* 2020;26(1):125.
 35. Fang Q, Chen H. The significance of m6A RNA methylation regulators in predicting the prognosis and clinical course of HBV-related hepatocellular carcinoma. *Mol Med.* 2020;26(1):60.
 36. Long J, Chen P, Lin J, Bai Y, Yang X, Bian J, et al. DNA methylation-driven genes for constructing diagnostic, prognostic, and recurrence models for hepatocellular carcinoma. *Theranostics.* 2019;9(24):7251–67.
 37. Wang Q, Guo Y, Wang W, Liu B, Yang G, Xu Z, et al. RNA binding protein DAZAP1 promotes HCC progression and regulates ferroptosis by interacting with SLC7A11 mRNA. *Exp Cell Res.* 2021;399(1):112453.
 38. Sun X, Niu X, Chen R, He W, Chen D, Kang R, et al. Metallothionein-1G facilitates sorafenib resistance through inhibition of ferroptosis. *Hepatology.* 2016;64(2):488–500.
 39. Feng J, Lu PZ, Zhu GZ, Hooi SC, Wu Y, Huang XW, et al. ACSL4 is a predictive biomarker of sorafenib sensitivity in hepatocellular carcinoma. *Acta Pharmacol Sin.* 2021;42(1):160–70.
 40. Xia H, Lee KW, Chen J, Kong SN, Sekar K, Deivasigamani A, et al. Simultaneous silencing of ACSL4 and induction of GADD45B in hepatocellular carcinoma cells amplifies the synergistic therapeutic effect of aspirin and sorafenib. *Cell Death Discov.* 2017;3:17058.
 41. Doll S, Proneth B, Tyurina YY, Panzilius E, Kobayashi S, Ingold I, et al. ACSL4 dictates ferroptosis sensitivity by shaping cellular lipid composition. *Nat Chem Biol.* 2017;13(1):91–8.

42. Liu Y, Zhang X, Zhang J, Tan J, Li J, Song Z. Development and Validation of a Combined Ferroptosis and Immune Prognostic Classifier for Hepatocellular Carcinoma. *Front Cell Dev Biol.* 2020;8:596679.
43. Tang B, Zhu J, Li J, Fan K, Gao Y, Cheng S, et al. The ferroptosis and iron-metabolism signature robustly predicts clinical diagnosis, prognosis and immune microenvironment for hepatocellular carcinoma. *Cell Commun Signal.* 2020;18(1):174.
44. Liang JY, Wang DS, Lin HC, Chen XX, Yang H, Zheng Y, et al. A Novel Ferroptosis-related Gene Signature for Overall Survival Prediction in Patients with Hepatocellular Carcinoma. *Int J Biol Sci.* 2020;16(13):2430–41.
45. Du X, Zhang Y. Integrated Analysis of Immunity- and Ferroptosis-Related Biomarker Signatures to Improve the Prognosis Prediction of Hepatocellular Carcinoma. *Front Genet.* 2020;11:614888.

Figures

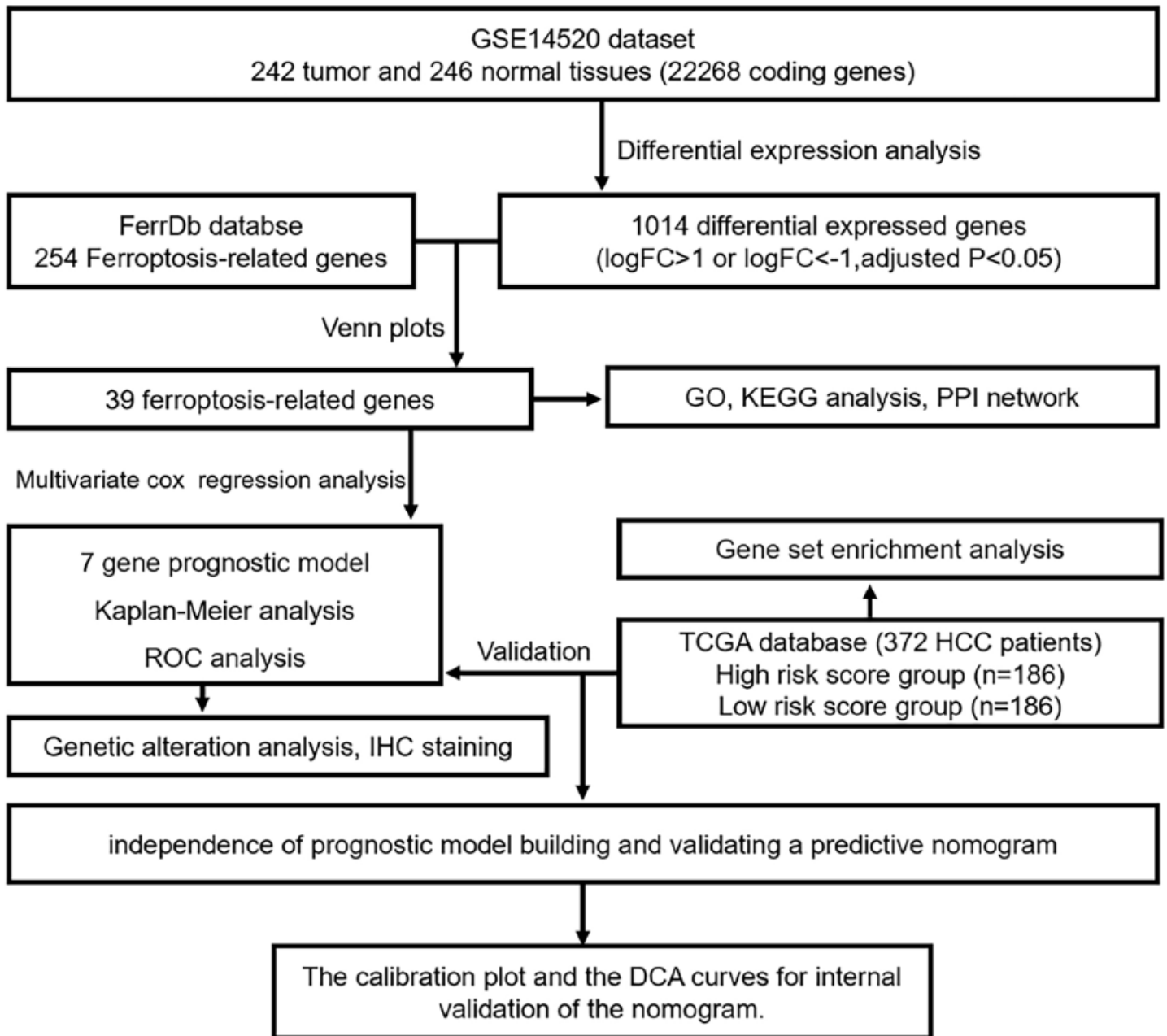


Figure 1

The flow chart showing the scheme of our study on ferroptosis-related gene signatures for clinically predicting recurrence after hepatectomy of hepatocellular carcinoma patients.

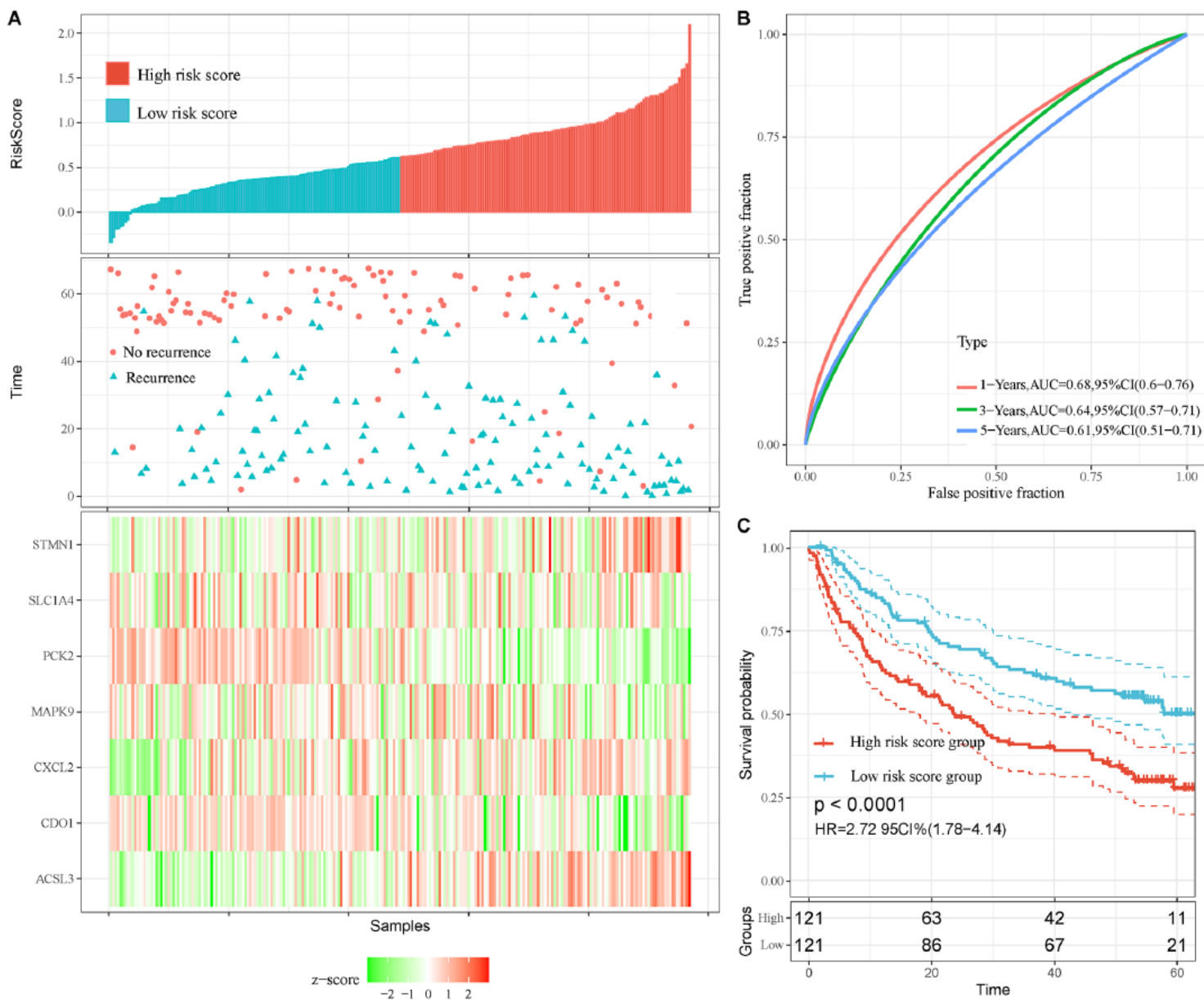


Figure 2

Risk score analysis, time dependent ROC analysis, and Kaplan–Meier analysis for the seven gene signature in HCC in the training set (GSE14520 HCC cohort). (A) Risk score, heatmap of mRNA expression of the seven gene signature in the training set of GSE14520 HCC cohort. (B) AUC of time-dependent ROC curves validated the prognostic performance of the risk score in the GSE14520 HCC cohort. (C) Kaplan–Meier curves for the RFS of HCC patients in the high-risk group and low-risk group in the GSE14520 cohort.

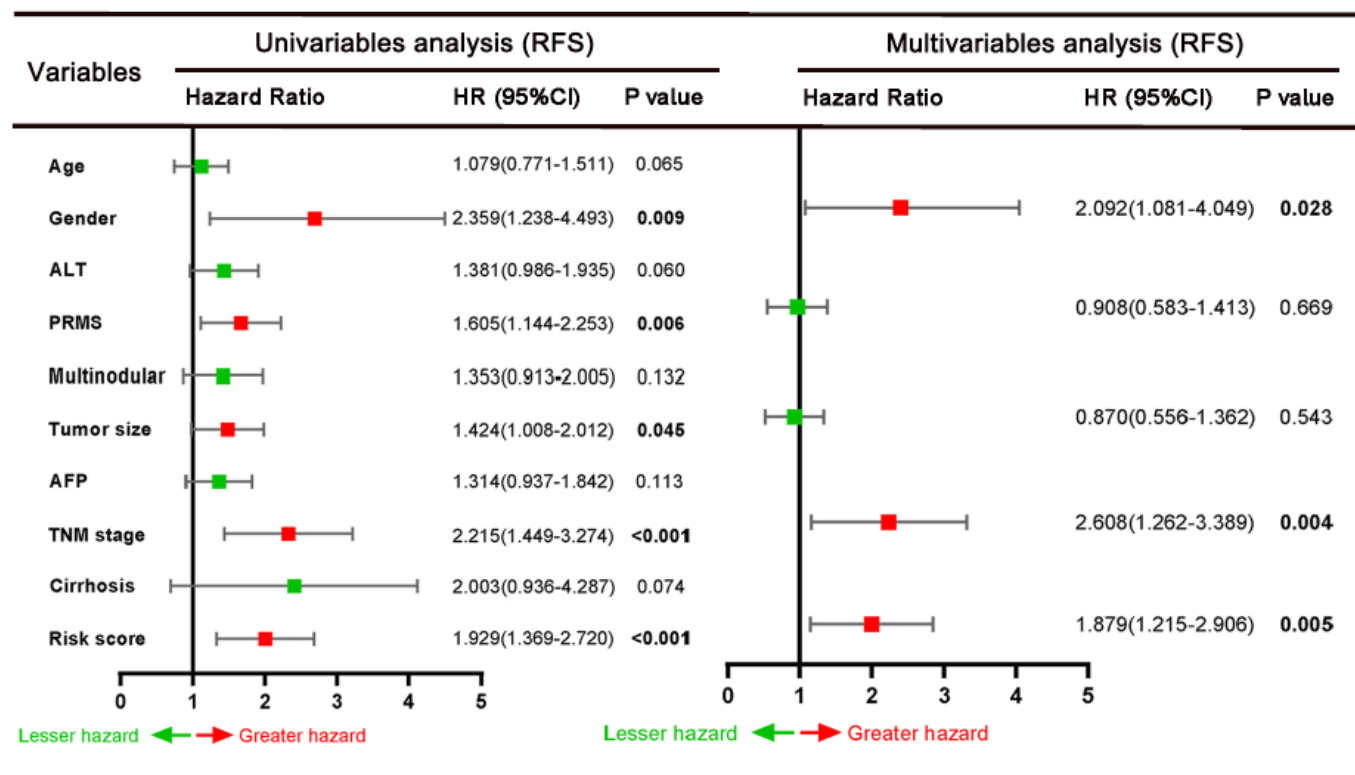


Figure 3

Forrest plot of the univariate and multivariate Cox regression analysis in GSE14520 HCC cohort.

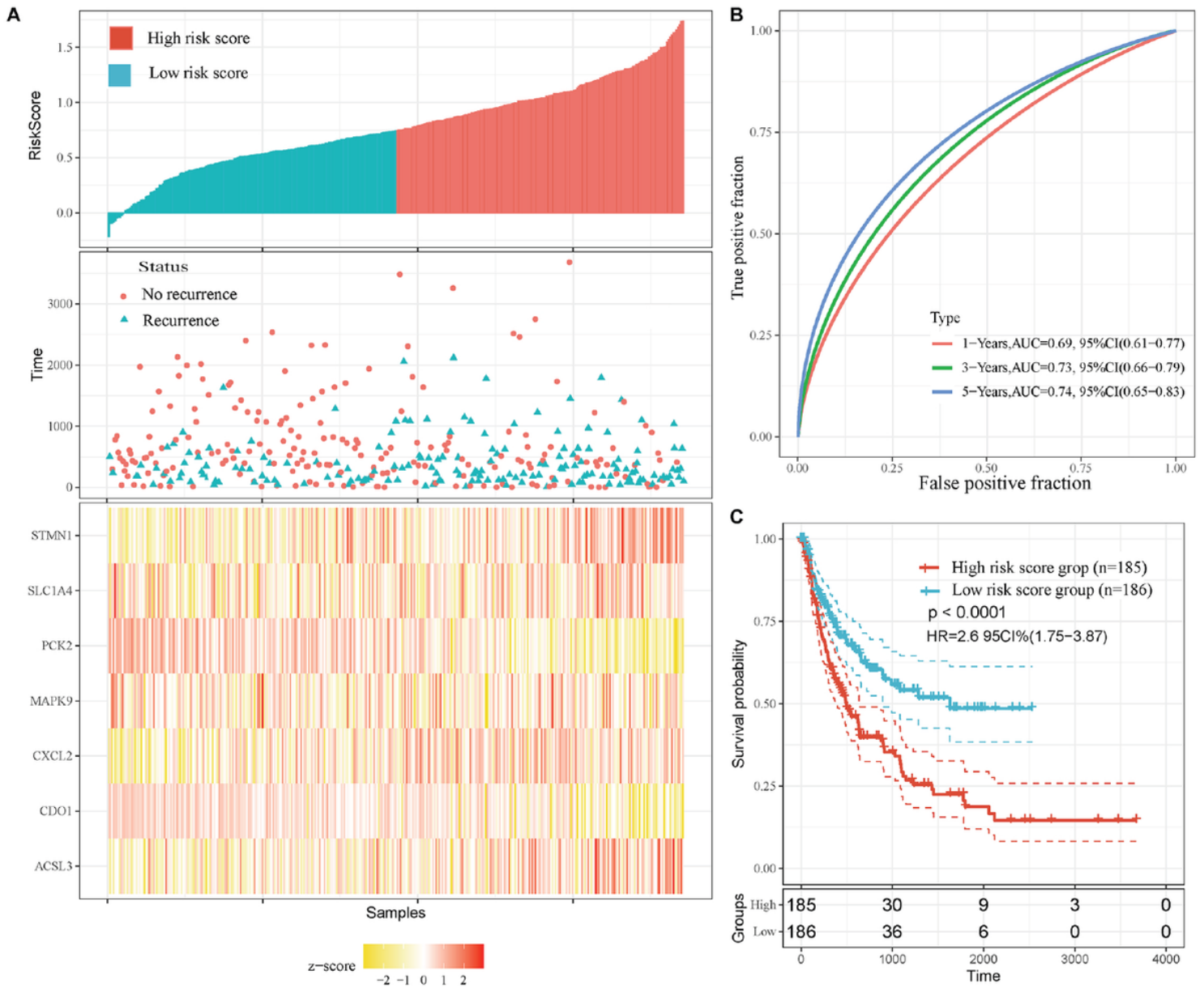


Figure 4

Risk score analysis, time dependent ROC analysis, and Kaplan–Meier analysis for the seven gene signature in HCC in the validation set (TCGA HCC cohort). (A) Risk score, heatmap of mRNA expression of the seven gene signature in the validation set of TCGA HCC cohort. (B) AUC of time-dependent ROC curves validated the prognostic performance of the risk score in the TCGA HCC cohort. (C) Kaplan-Meier curves for the RFS of HCC patients in the high-risk group and low-risk group in the TCGA cohort.

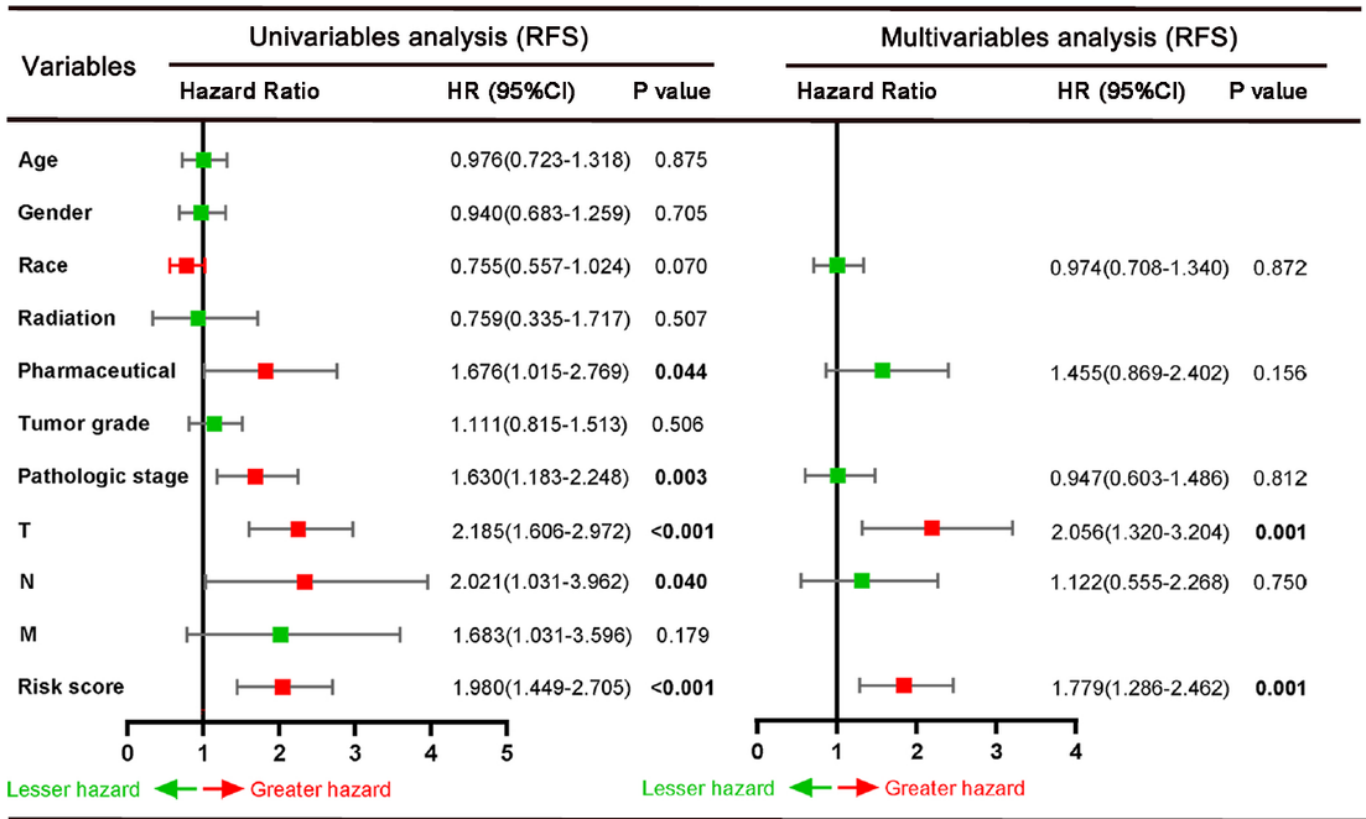


Figure 5

Forrest plot of the univariate and multivariate Cox regression analysis in TCGA HCC cohort.

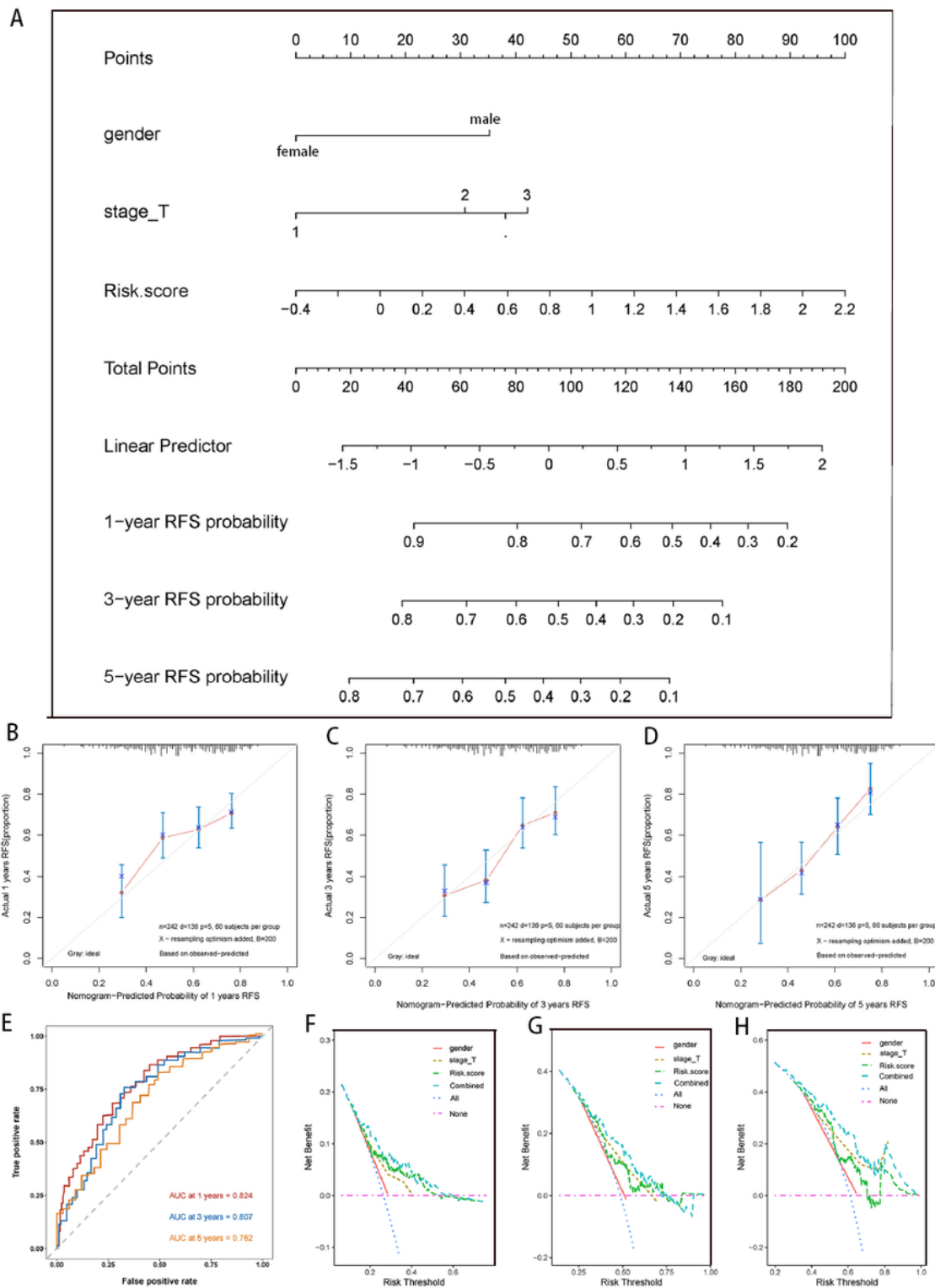


Figure 6

Nomogram predicting recurrence-free survival for HCC patients. (A) Nomogram established based on gender, TNM stage, and risk score as predictive factors to predict 1-year, 3-year, and 5-year recurrence-free survival probability. (B-D) The calibration plot for the recurrence-free survival probability at 1-year (B), 3-year (C), and 5-year (D) for internal validation of the nomogram. The Y-axis and X-axis represents actual survival and nomogram-predicted survival, respectively. (E) The time-dependent ROC curves for

1-year,3-year, and 5-year recurrence-free survival prediction of the nomogram. (F-G) DCA curves of gender, TNM stage, risk score and combined model to evaluate the clinical utility of different decision strategies. The blue line represented the combined nomogram and exhibited the best net benefit for predicting the recurrence-free survival probability at 1-year (F), 3-year (G), and 5-year (H).

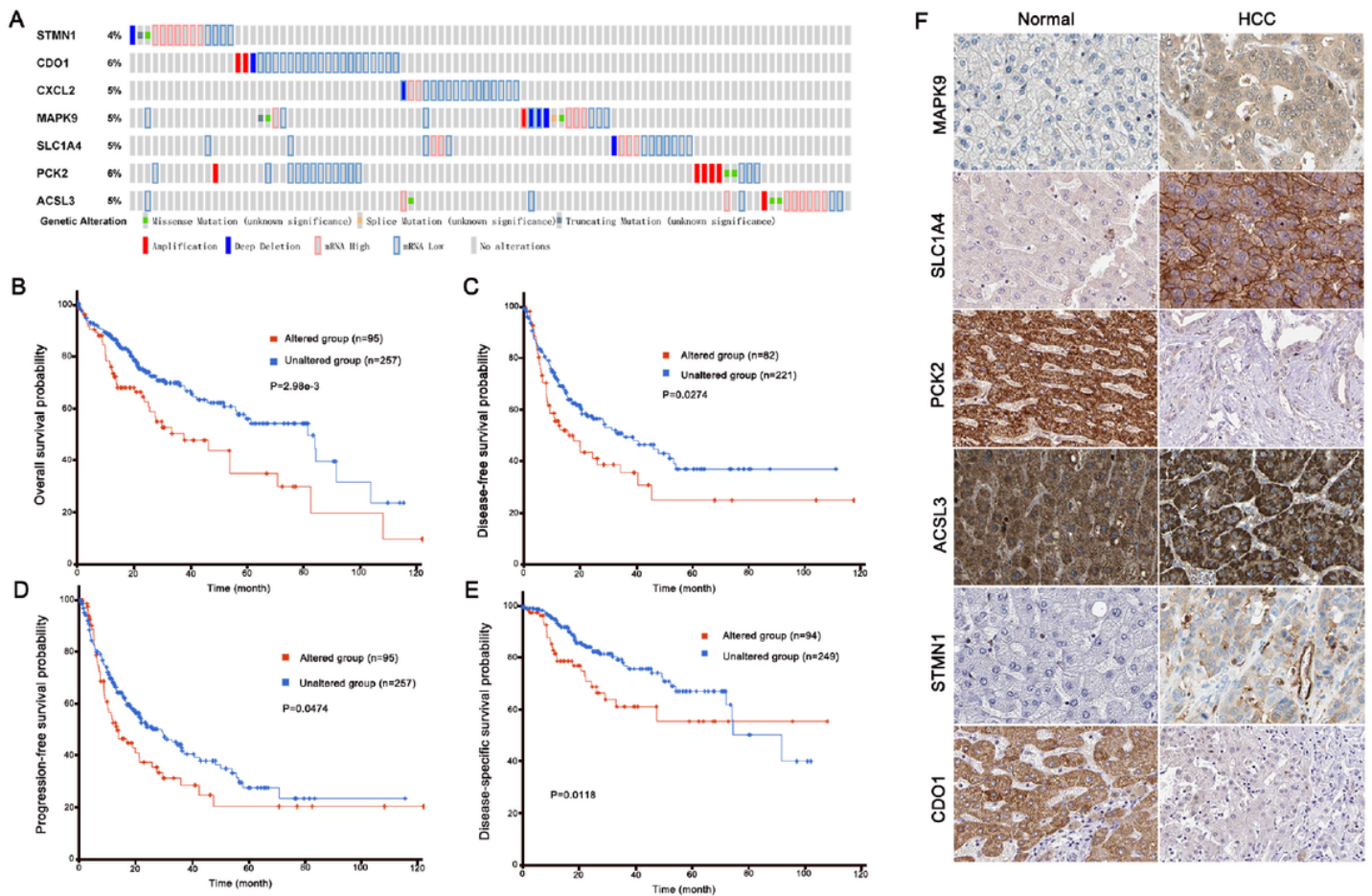


Figure 7

The genetic alterations and protein expression analysis of the seven-gene signature in HCC. (A) The genetic alteration profiles of the seven genes in the TCGA liver cancer RNA-seq (n=352) dataset from the cBioPortal database. (B-E) Patients with genetic alteration have poorer overall survival probability (B), disease-free probability (C), progression-free survival probability (D), and disease-specific survival probability (E) than the patients without genetic alterations. (F) The representative protein expression of the MAPK9, SLC1A4, PCK2, ACSL3, STMN1, and CDO1 in HCC and normal liver tissue. Data were from the Human Protein Atlas database.

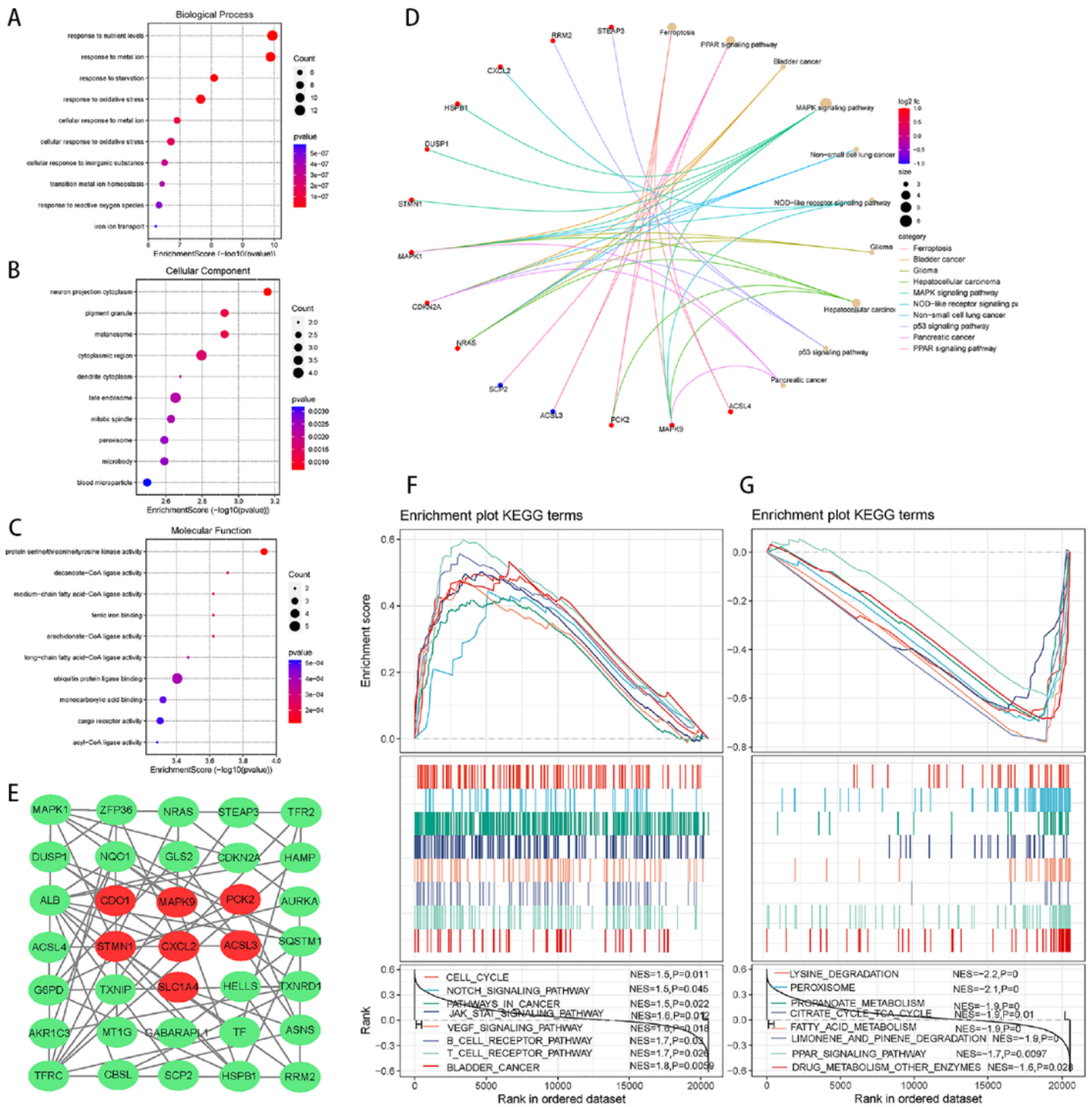


Figure 8

Representative results of functional enrichment analysis via GO, KEGG, and GSEA. (A-D) The 39 ferroptosis-related differentially expressed genes were mainly enriched pathway based on the biological process (A), cellular component (B), molecular function (C) and KEGG pathway (D). (E) Protein-protein interaction network shown the interactions between the 39 ferroptosis-related differentially expressed

genes. (F-G) The significant signaling pathway that genes of the high-risk score (F) and low-risk score patients (G) were enriched in the GSEA.