Am J Cancer Res 2022;12(5):1995-2011
www.ajcr.us /ISSN:2156-6976/ajcr0142657

*Original Article*

# A novel ferroptosis-related gene signature for clinically predicting recurrence after hepatectomy of hepatocellular carcinoma patients

Huaxiang Wang[1,2*], Chengkai Yang[2*], Yi Jiang[2,3], Huanzhang Hu[2,3], Jian Fang[4], Fang Yang[2,3]

*[1]Department of Hepatobiliary and Pancreatic Surgery, Taihe Hospital, Affiliated Hospital of Hubei University of Medicine, Shiyan, Hubei 442000, China; [2]The Fuzong Clinical Medical College of Fujian Medical University, Fuzhou 350025, Fujian, China; [3]Department of Hepatobiliary Surgery, 900 Hospital of The Joint Logistics Team, Fuzhou 350025, Fujian, China; [4]Department of Hepatobiliary Medicine, The Third People's Hospital of Fujian University of Traditional Chinese Medicine, Fuzhou 350108, Fujian, China. [*]Equal contributors.*

Received March 2, 2022; Accepted April 14, 2022; Epub May 15, 2022; Published May 30, 2022

**Abstract:** High recurrence rate in HCC is the primary cause of the poor prognosis after hepatectomy. Therefore, in this study, we aimed to construct a gene signature for predicting the recurrence rate in HCC. The mRNA expression profiles and clinical information of HCC patients from GEO and TCGA databases were used, and ferroptosis-related gene list was obtained from the FerrDb database. We identified 39 ferroptosis-related genes (FDEGs) that were differentially expressed between HCC samples and normal tissues from the GSE14520 dataset. The univariate and multivariate Cox regression analyses were employed to construct a prognostic signature. Seven FDEGs (MAPK9, SLC1A4, PCK2, ACSL3, STMN1, CDO1, and CXCL2) were included to construct a risk model, which was validated in the TCGA dataset. Patients in high-risk groups exhibited a significantly poor prognosis compared with patients in low-risk groups in both the training set (GSE14520 cohort) and the validation set (TCGA cohort). Multivariate cox regression analyses demonstrated that the 7-gene signature was an independent risk factor for RFS in HCC patients. KEGG analysis showed that FDEGs were mainly enriched in Ferroptosis, Hepatocellular carcinoma pathway, and MAPK signaling pathway. GSEA analysis suggested that the high-risk group was correlated with multiple oncogenic signatures and invasive-related pathways. These results indicated that this risk model can accurately predict recurrence after hepatectomy and offer novel research directions for personalized treatment in HCC patients.

**Keywords:** Hepatocellular carcinoma, ferroptosis, gene signature, recurrence-free survival, nomogram, decision curve analysis

## 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 the data released by the American Cancer Society in 2021, the 5-year survival rate for HCC of all stages is only 20% [4]. With the improvement in diagnosis, the proportion of HCC receiving surgical resection has increased, but those who have undergone radical resection still have a 70% recurrence rate within 5 years [5]. The high recurrence rate is the main cause of death in HCC. Therefore, establishing an effective model to predict the postoperative recurrence and identify the high-risk patients early is of great value to improve the prognosis in HCC. The traditional recurrence prediction model integrates data on tumor stage, tumor size, microvascular invasion, tumor differentiation, and other relevant clinical characteristics supplemented by a single serum alpha-fetoprotein expression [6-8]. But the specificity and sensitivity of this model is not high enough to distinguish patients with heterogeneity.

Ferroptosis is a newly discovered form of cell death 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 is crucial in regulating the initiation and progression of some 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 is an effective method to induce HCC cell death and the cytotoxic effect of sorafenib in HCC [16]. At present, some ferroptosis-related genes such as NRF2, NQO1, HCAR1, MCT1, and ZFP36 have been proven as cancer-promoting or cancer-suppressing factors in HCC [17-19]. However, few studies focus on the predictive value of these ferroptosis-related genes on the recurrence of HCC.

In this study, we constructed a 7-gene HCC recurrence model by using differentially expressed ferroptosis-related genes (FDEGs) extracted from the gene expression omnibus (GEO) [20] and FerrDb database [21]. We further validated the reliability of this 7-gene HCC recurrence model in an independent cohort of The Cancer Genome Atlas (TCGA) [22]. The predictive nomogram and decision curve analysis (DCA) were built to estimate the recurrence predictive capacity of this 7-gene signature. In addition, we investigated the correlation between the genetic alteration of this 7-gene signature and the 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) were used to explore the intrinsic regulatory mechanisms of these ferroptosis-related genes [24, 25]. Together, our results suggest that the ferroptosis-related 7-gene signature and nomogram have the potential to effectively predict the RFS for patients with HCC.

## Methods

### *Gene datasets and data collection*

We downloaded the mRNA expression data and the corresponding clinical characteristics data from the GEO and TCGA databases. The ferroptosis-related gene list was obtained from the FerrDb database and literature in PubMed. In the GEO database, we searched the keywords "hepatocellular carcinoma", "HCC". In addition, "Homo sapiens" and "Expression profiling by array" were included in the next round of screening. The datasets we would select must meet the following criteria: (1) they are human hepatocellular carcinoma samples; (2) the number of tumor and non-tumor liver control tissue samples is more than 100; (3) complete clinicopathological and survival data are available. At the end, the GSE14520 dataset containing RNA sequencing of 242 HCC samples and 246 normal liver samples was chosen to construct the predictive model for recurrence. The liver hepatocellular carcinoma (LIHC) cohort from the TCGA database, containing the gene expression and the clinical data of 372 HCC patients, was used as the validation set. All the gene expression and the clinical data were obtained from the publicly available database; hence, no additional ethical approval was required.

### *Identification of the differentially expressed ferroptosis-related genes (FDEGs)*

The R software (version 4.0.2) and built-in limma package were utilized to perform the analysis for differentially expressed genes using RNA sequence data between the HCC tumor tissues and the paired normal tissues, which were downloaded from the GSE14520 dataset. The selection of the differentially expressed genes must meet two standards: log2 fold change (FC) >1.0 or log2 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 extracted as FDEGs.

### *Establishment and validation of the ferroptosis-related gene signature*

We performed the univariate proportional hazards Cox regression on FDEGs to identify the prognostic value of the ferroptosis-related genes for RFS, and genes with a *P* value <0.05 were considered statistically significant. Subsequently, FDEGs identified by univariate Cox regression were included in the multivariate Cox's proportional hazard model with forward LR model to identify the genes which were independent risk factors for RFS of patients with HCC. In addition, the regression coefficients of FDEGs were also obtained from the

multivariate Cox regression. Finally, these genes were used to establish a prognostic risk signature. Next, the 242 HCC patients from the GSE14520 dataset were divided into high-risk and low-risk groups based on the median Risk score, which was calculated based on the following formula: Risk score = $\sum_{i=1}^{n} vi \times \beta i$ ($v$ represent the expression value of the gene and $\beta$ represent the corresponding regression coefficients). In addition, the Kaplan-Meier survival analysis and the time-dependent receiver operating characteristic (ROC) curve were used to evaluate the predictive performance of this gene signature for RFS. Moreover, to validate the independent prognostic role of this gene signature for RFS and to identify independent prognostic parameters, the univariate and multivariate Cox regression analyses were performed in the GSE14520 dataset based on this prognostic gene signature and some clinical characteristics such as age, gender, tumor stage, tumor size, serum Alpha-fetoprotein (AFP) level, TNM stage, cirrhosis, and alanine aminotransferase (ALT). P<0.05 was considered statistically significant. Those characteristics with P<0.05 from the univariate analysis were further included in the multivariate Cox's proportional hazard model with forward LR model. Finally, we validated the reliability of this 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 1-, 3-, and 5-year RFS rate. We calculated the concordance index (C-index) using the "survival" R package to assess the predictive performance of the nomogram. We next plotted the calibration curve of RFS at different year points. 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 with 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. We then compared the probability in overall survival (OS), disease-free, progression-free survival, and disease-specific survival between the alteration groups and the no alteration groups. Furthermore, we investigated the differential protein expression of the 7-gene signature between the HCC tissues and the adjacent normal liver tissues in the Human Protein Atlas database.

*Functional enrichment analysis via GO, KEGG, and GSE*

We performed the GO and KEGG enrichment analyses on the FDEGs using the DAVID database to explore the potential mechanisms by which these genes regulate 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 were selected to perform the GSEA enrichment analysis using the GSEA software (version 4.1.0). We divided the 372 HCC patients into high-risk and low-risk groups based on the median Risk score. 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 diagram. The association between the risk score and the clinicopathological characteristics was analyzed using Pearson's chi-square test. Univariate and multivariate cox regression analyses were used to identify the risk factors or independent risk factors for RFS. Kaplan-Meier analysis with the log-rank test was employed to compare the RFS between the high-risk group and the low-risk group. The area under the curve (AUC) of ROC was utilized to estimate the predictive performance of the gene signature. Two-sided P<

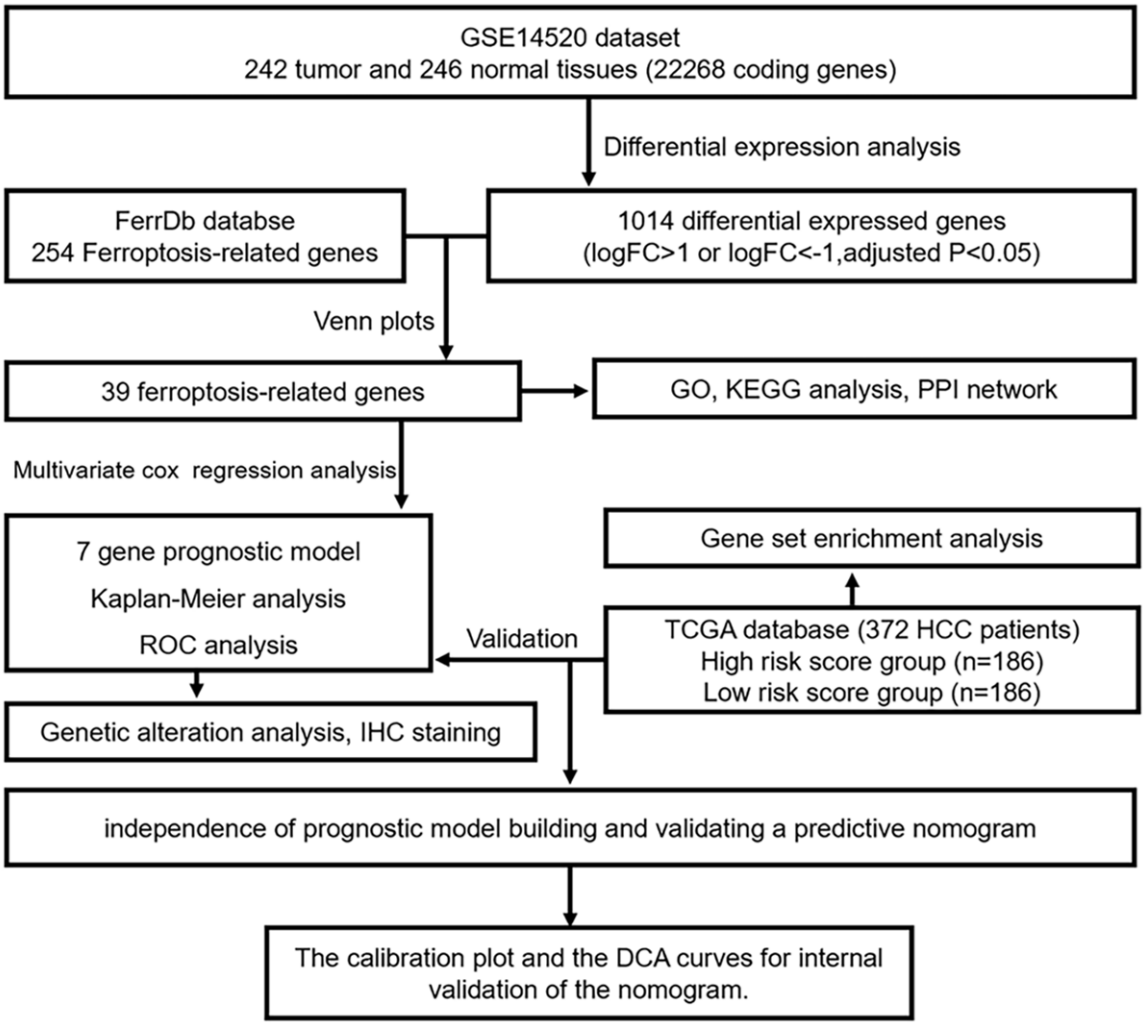



**Figure 1.** Schematic diagram showing the construction of ferroptosis-related gene signatures for clinically predicting recurrence after hepatectomy of hepatocellular carcinoma patients.

0.05 was considered as a statistically significant difference.

## Results

*FDEGs identification in HCC*

**Figure 1** showed the schematic diagram of our study (**Figure 1**). A total of 1014 differentially expressed genes, including 539 down-regulated and 375 up-regulated genes, were identified by comparing 242 HCC tissues with 246 adjacent normal liver tissues from the GSE14520 dataset. Meanwhile, we extracted 254 ferroptosis-related genes from the FerrDb database. The overlapping 39 genes between the 1014 differentially expressed genes and the 254 ferroptosis-related genes were identified as FDEGs.

*Establishment of the prognostic 7-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 analysis, seven genes were identified to construct a predictive gene signature (**Table 1**). 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) * $expression_{MAPK9}$ + (-0.086) * $expression_{SLC1A4}$ + (-0.167) * expres-

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

| Gene | Coef | aHR | 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 |

aHR-adjusted hazard ratio; CI-confidence interval.

$\text{sion}_{PCK2}$ + 0.203 * $\text{expression}_{ACSL3}$ + 0.201 * $\text{expression}_{STMN1}$ + (-0.003) * $\text{expression}_{CDO1}$ + 0.109 * $\text{expression}_{CXCL2}$.

*Internal validation of the prognostic gene signature*

We calculated the 7-gene-based risk score for each HCC patient in the training set (GSE14520). Next, the 242 patients were divided into high-risk and low-risk groups based on the median Risk score (**Figure 2A**). The Kaplan-Meier survival analysis and the time-dependent ROC curve were used to evaluate the predictive performance of this gene signature for RFS. The ROC curve revealed that AUCs of 1-, 3-, and 5-year RFS rate were 0.68, 0.64, and 0.61, respectively (**Figure 2B**). Furthermore, The Kaplan-Meier survival analysis revealed that patients in the high-risk group exhibited a worse RFS than patients in the low-risk group (**Figure 2C**). In addition, the correlation analysis demonstrated that high-risk score correlated with 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. Univariate cox regression analysis showed 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 for recurrence-free survival. The multivariate Cox regression analysis confirmed that gender (aHR (95% CI): 2.092 (1.081-4.049); P=0.028), TNM stage (aHR (95% CI): 2.608 (1.262-3.389); P=0.004), and high-risk score (aHR (95% CI): 1.879 (1.215-2.906); P=0.005) were independent risk factors for RFS (**Figure 3**). These results demonstrated the feasibility of this 7-gene signature for RFS prediction.

*Validation of the prognostic gene signature in the TCGA database*

We download the RNA seq and the corresponding clinical characteristic data in the TCGA database to validate the performance of this 7-gene signature for RFS prediction. We calculated the 7-gene-based risk score for each HCC patient in the validation set (TCGA HCC cohort). The 372 patients were divided into high-risk and low-risk groups based on the median Risk score (**Figure 4A**). The ROC analysis showed that the AUCs of 1-, 3-, and 5-year RFS rate were 0.69, 0.73, and 0.74, respectively (**Figure 4B**). Furthermore, consistent with the result from the GSE14520 dataset, the Kaplan-Meier survival analysis revealed that patients in the high-risk group exhibited a worse RFS than patients in the low-risk group (**Figure 4C**). In addition, correlation analysis demonstrated that high-risk score correlated with 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 this gene signature for RFS by the univariate and multivariate Cox regression analyses. Univariate Cox regression analysis revealed 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. The multivariate Cox regression analysis confirmed that T 3/4 (aHR (95% CI): 2.056 (1.320-3.204); P=0.001) and high-risk score (aHR (95% CI): 1.779 (1.286-2.462); P=0.001) were independent risk factors for RFS (**Figure 5**). Together, these results validated our construction of the 7-gene signature for RFS prediction.

*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 analy-

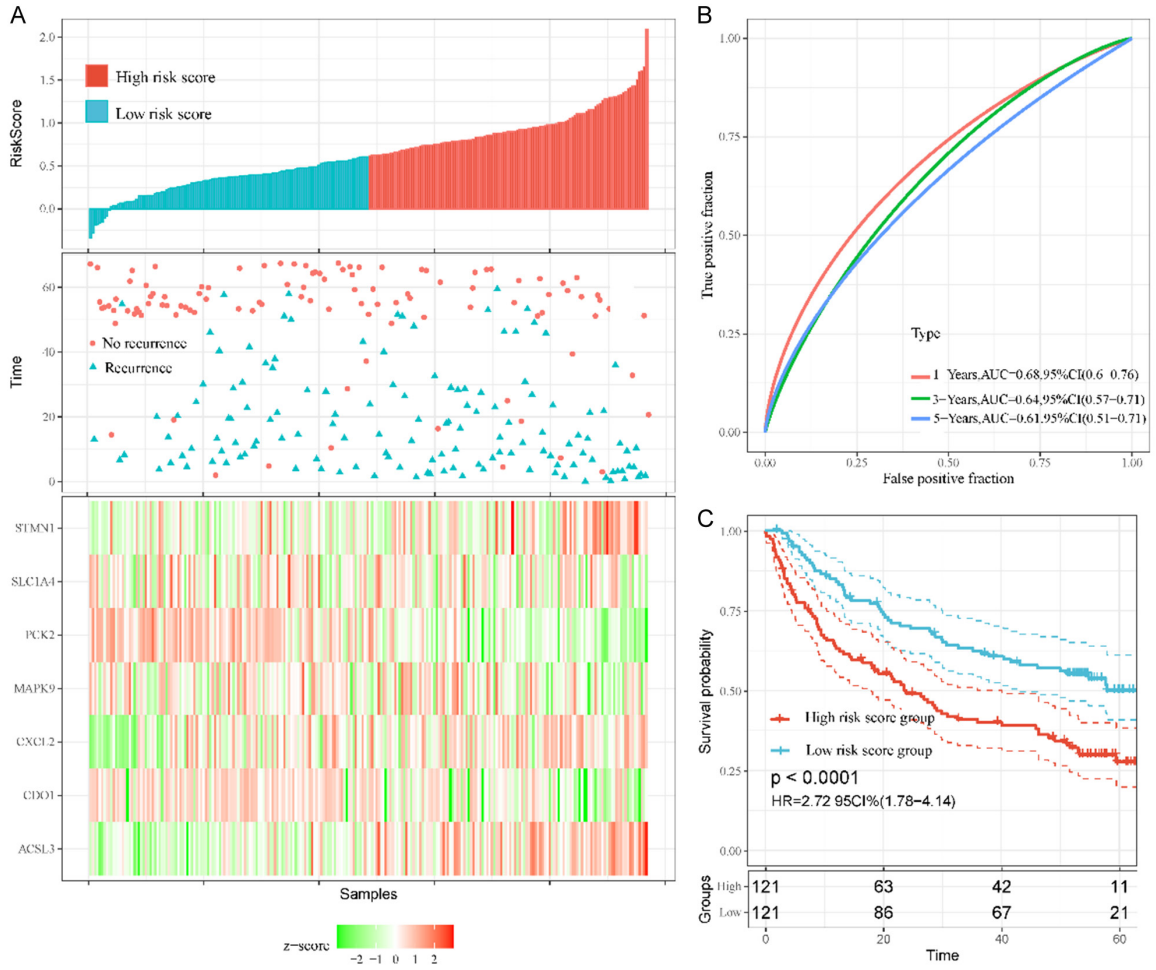



**Figure 2.** Risk score analysis, time-dependent ROC analysis, and Kaplan-Meier analysis for the 7-gene signature in HCC in the training set (GSE14520 HCC cohort). A. Risk score, heatmap of mRNA expression of the 7-gene signature in the training set of GSE14520 HCC cohort. B. AUC of the time-dependent 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.

sis to predict the 1-, 3-, and 5-year RFS (**Figure 6A**). The C-index of the combined nomogram model was 0.872. The calibration curve of the nomogram representing the actual and the combined model in the training set (GSE14-520) exhibited an accurate prediction for the 1-, 3-, and 5-year RFS (**Figure 6B-D**). Furthermore, the ROC analysis showed that the AUCs for predicting 1-, 3-, and 5-year RFS were 0.824, 0.807, and 0.762, respectively (**Figure 6E**). From the DCA curve, we could see that the combined model was superior for 1-, 3-, and 5-year RFS prediction compared with the individual predictive factors (**Figure 6F-H**). In sum, these results demonstrated that the combined model of nomogram exhibited an excellent predictive ability for 1-, 3-, and 5-year RFS of HCC patients, which might be useful in clinical practice.

*Genetic alteration of gene signature correlated with poor survival*

The genetic alteration of this 7-gene signature was explored in the Liver Hepatocellular Carcinoma cohort (TCGA, Firehose Legacy) of the cBioPortal database. Among 352 HCC patients examined, 95 patients (27.0%) showed genetic alterations in this 7-gene signature (**Figure 7A**). In addition, the patients with genetic alteration had poor rate in overall survival (P=2.98e-3, **Figure 7B**), disease-free (P=0.0274, **Figure 7C**), progression-free survival (P=0.0474, **Figure 7D**), and disease-specific

**Table 2.** Correlation between risk score and clinicopathological features of HCC patients for RFS in the GSE14520 HCC cohort

| Characteristics | | N | Risk score level | | $X^2$ | *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 | >5 cm | 88 | 36 | 52 | 0.131 | 0.064 |
| | ≤5 cm | 154 | 84 | 70 | | |
| TNM stage | I/II | 174 | 97 | 77 | 5.400 | 0.020 |
| | III | 51 | 19 | 32 | | |
| Serum AFP level | >300 ng/ml | 110 | 41 | 69 | 12.393 | <0.001 |
| | ≤300 ng/ml | 128 | 77 | 51 | | |
| ALT | >50 U/L | 100 | 41 | 59 | 5.027 | 0.025 |
| | ≤50 U/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.

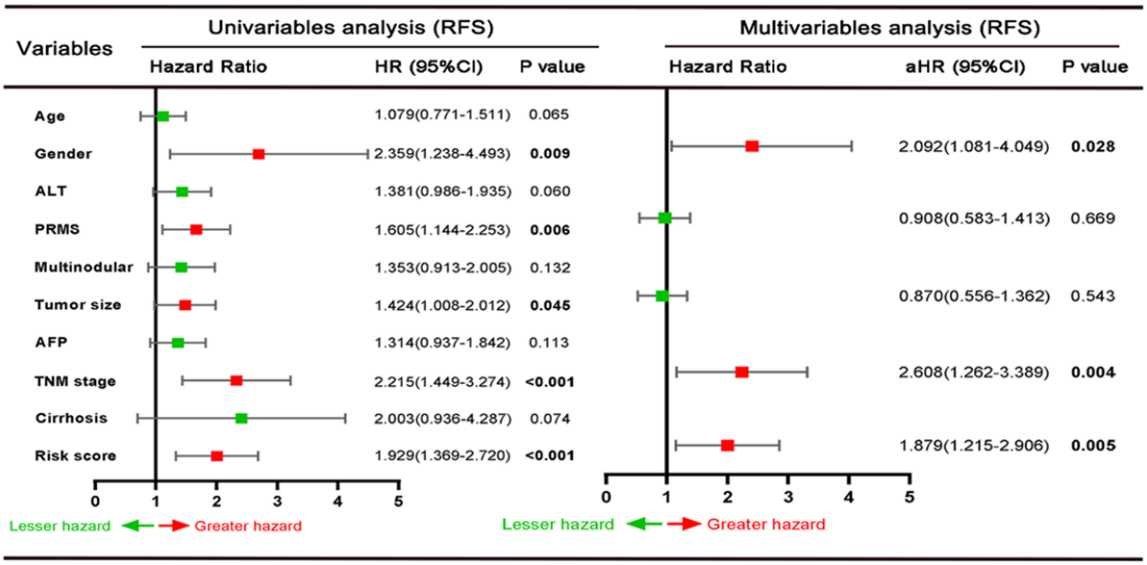



**Figure 3.** Forrest plot of the univariate and multivariate Cox regression analyses in GSE14520 HCC cohort.

survival (P=0.0118, **Figure 7E**) than the patients without genetic alterations. We further investigated the protein expression level of this 7-gene signature in HCC tissues and their adja-

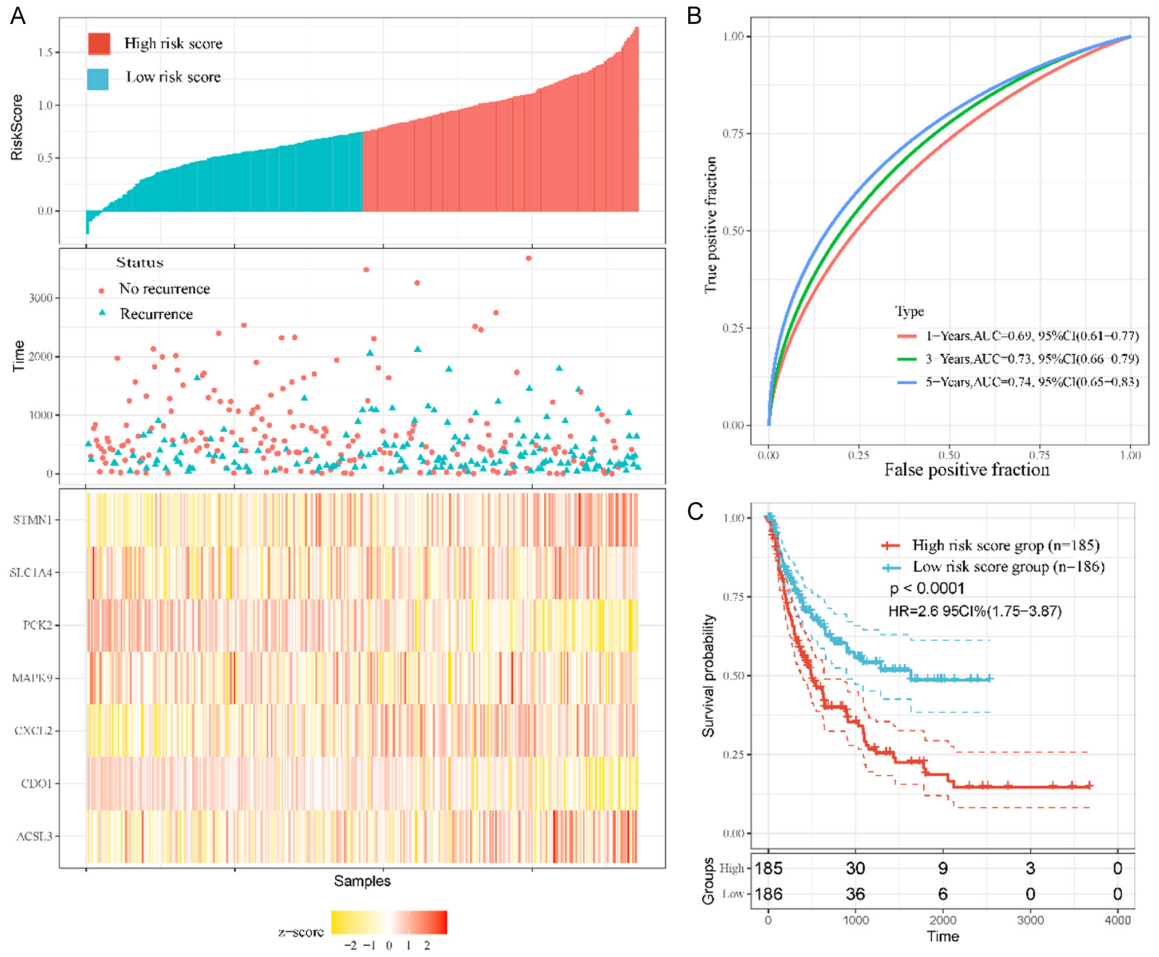



**Figure 4.** Risk score analysis, timedependent ROC analysis, and Kaplan-Meier analysis for the 7-gene signature in HCC in the validation set (TCGA HCC cohort). A. Risk score, heatmap of mRNA expression of the 7-gene signature in the validation set of TCGA HCC cohort. B. AUC of the time-dependent ROC 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.

cent normal liver tissues deposited 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 (**Figure 7F**). The information for CXCL2 protein expression was not available in the Human Protein Atlas database.

*GO and KEGG enrichment analyses and protein-protein interaction (PPI) network constructions of FDEGs*

We performed the GO and KEGG enrichment analyses on the 39 FDEGs to explore the potential mechanisms by which these genes regulate the tumorigenesis and progression of HCC. GO analysis revealed that the FDEGs were significantly enriched in the signaling pathways involved in responding to nutrient levels, the metal ion, starvation, and oxidative stress (**Figure 8A**). Interestingly, the FDEGs were also found enriched in neuron projection cytoplasm, pigment granules, and melanosome (**Figure 8B**). At the molecular level, the FDEGs were associated with protein serine kinase activity and decanoate-CoA ligase activity (**Figure 8C**). Further KEGG analysis showed that the FDEGs were significantly enriched in the pathways of Ferroptosis, Hepatocellular carcinoma, MAPK signaling pathway, and other growth-related pathways (**Figure 8D**). Additionally, we constructed a PPI network of these FDEGs in the STRING database and visualized it utilizing the Cytoscape software (**Figure 8E**).

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

| Characteristics | | N | Risk score level | | $X^2$ | *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 | | |

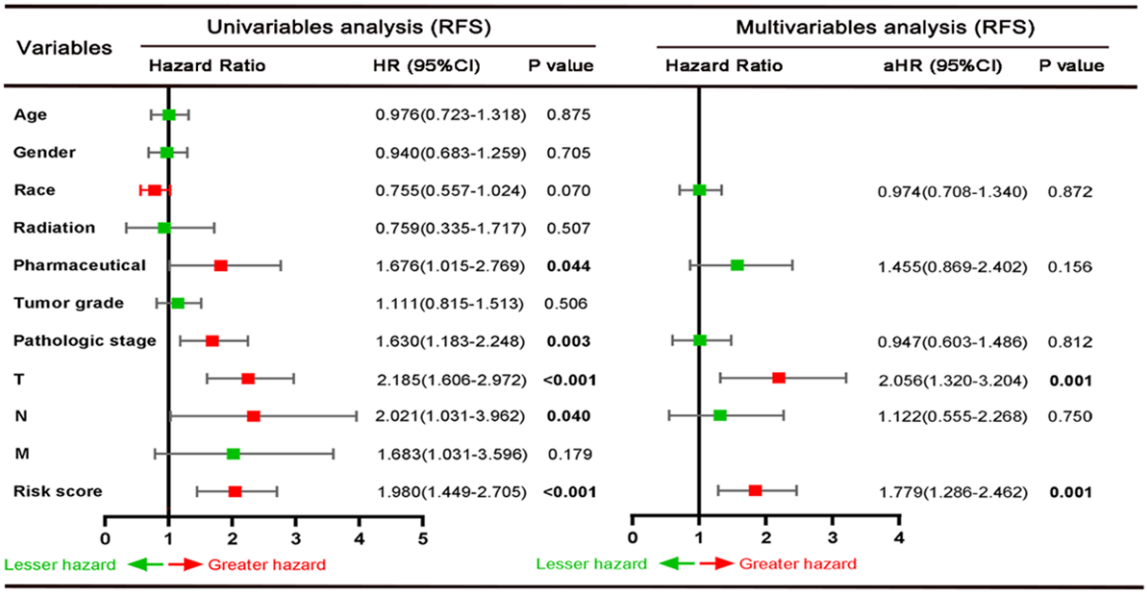



**Figure 5.** Forrest plot of the univariate and multivariate Cox regression analyses in TCGA HCC cohort.

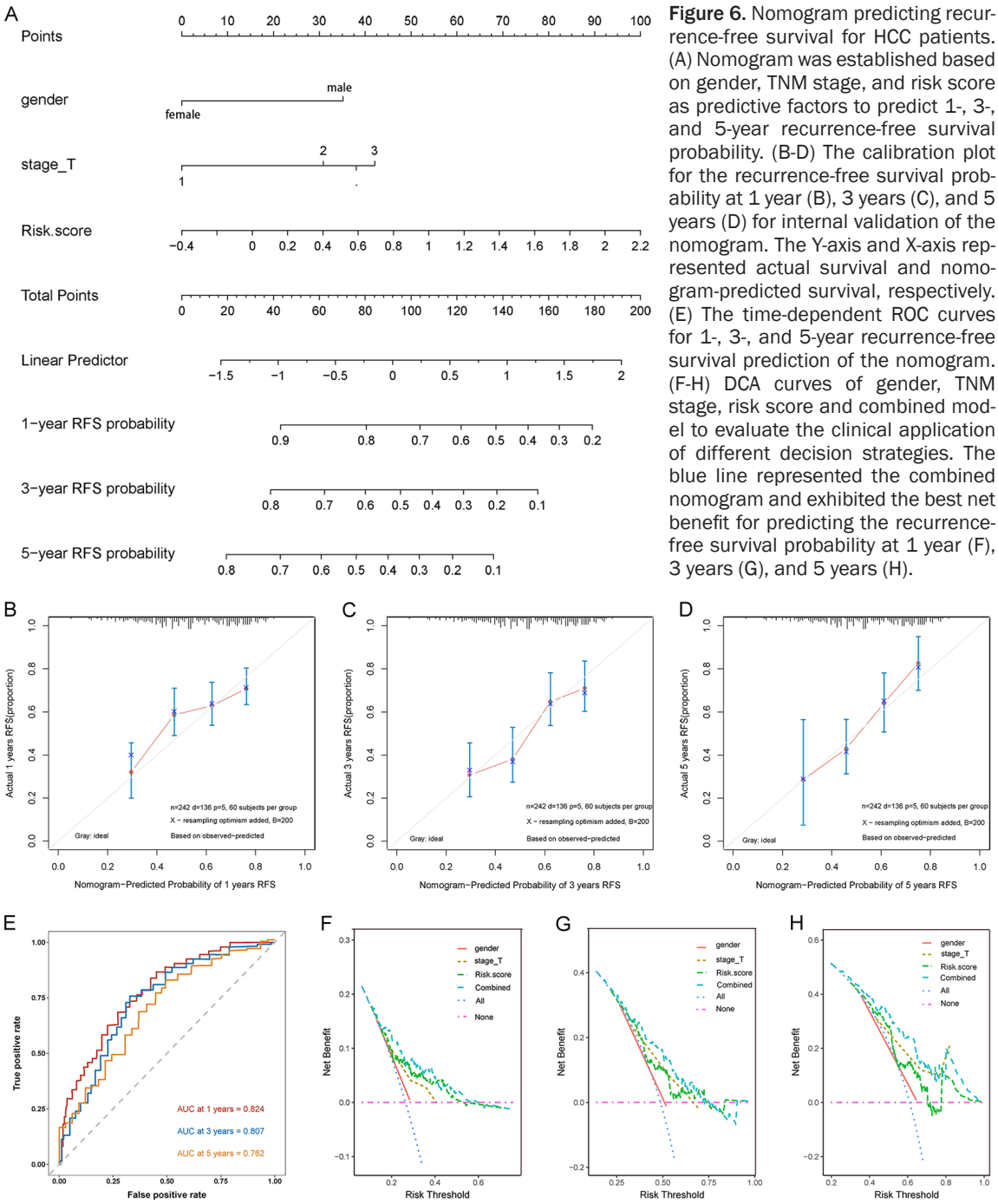



**Figure 6.** Nomogram predicting recurrence-free survival for HCC patients. (A) Nomogram was established based on gender, TNM stage, and risk score as predictive factors to predict 1-, 3-, and 5-year recurrence-free survival probability. (B-D) The calibration plot for the recurrence-free survival probability at 1 year (B), 3 years (C), and 5 years (D) for internal validation of the nomogram. The Y-axis and X-axis represented actual survival and nomogram-predicted survival, respectively. (E) The time-dependent ROC curves for 1-, 3-, and 5-year recurrence-free survival prediction of the nomogram. (F-H) DCA curves of gender, TNM stage, risk score and combined model to evaluate the clinical application 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 years (G), and 5 years (H).

*GSEA enrichment analysis*

GSEA analysis was carried out to further investigate the significant signaling pathways in which the 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 critical for cancer development including cell cycle (NES=1.5, P<0.001), notch signaling (NES=1.5, P<0.001), pathway in cancer (NES=1.5, P<0.001), and VEGF signaling pathway (NES=1.7, P<0.001) (**Figure 8F**). In contrast, the low-risk score group was negatively correlated with the signaling pathways for lysine degradation (NES=-2.2, P<0.001), peroxisome (NES=-2.1, P< 0.001) and propanoate metabolism (NES=-1.9, P<0.001) (**Figure 8G**).

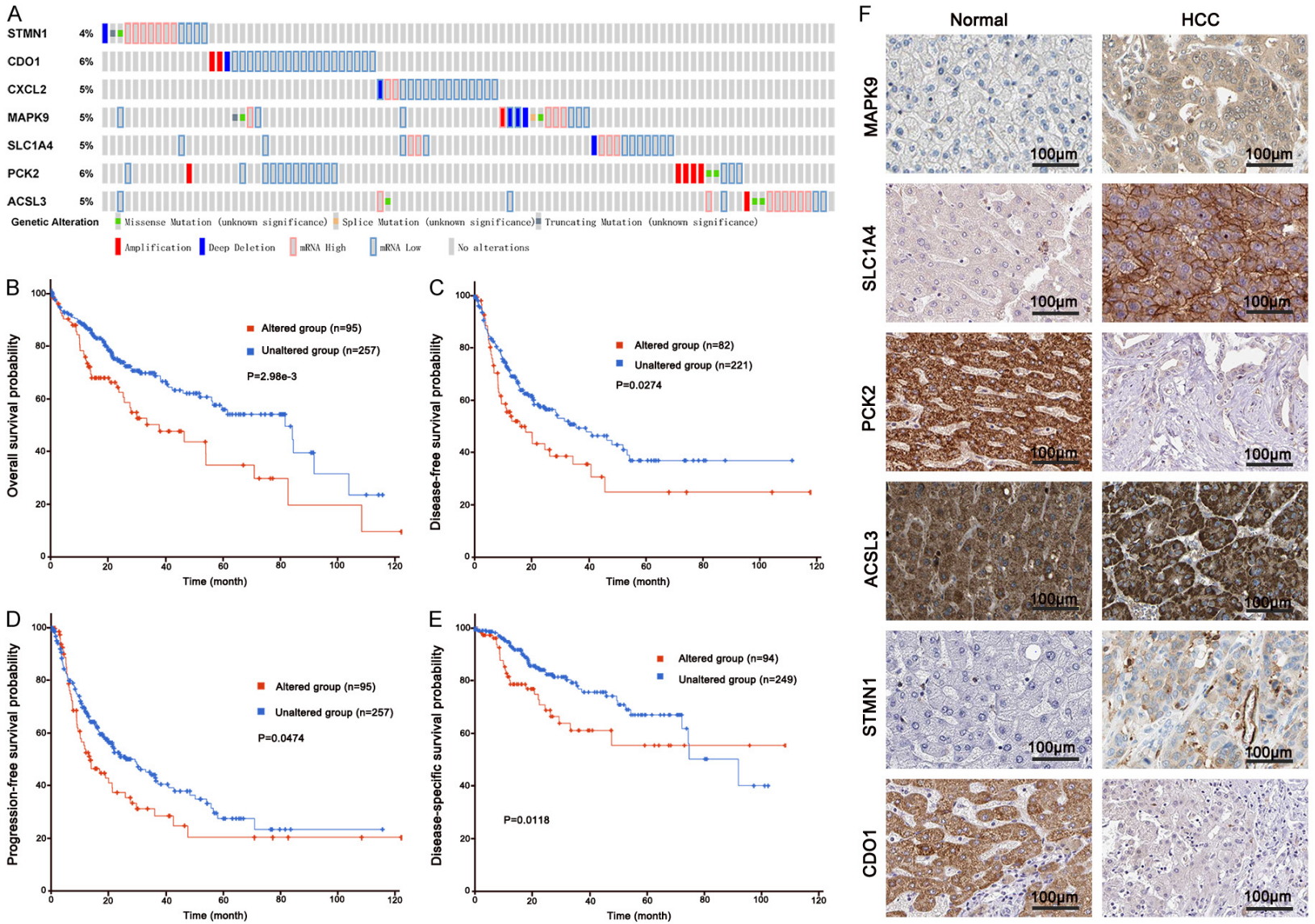
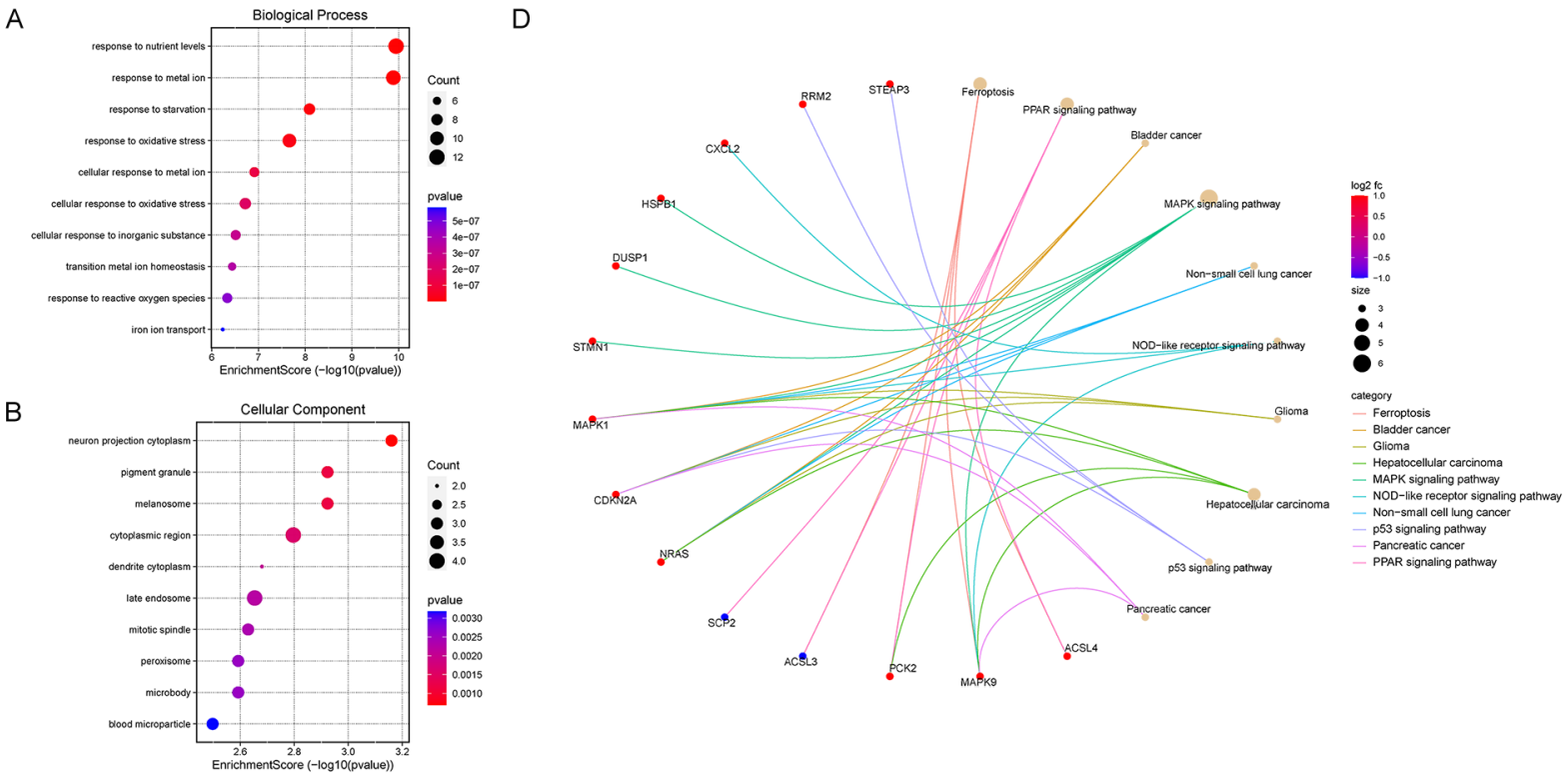



**Figure 7.** The genetic alterations and protein expression analysis of the 7-gene signature in HCC. (A) The genetic alteration profiles of the seven genes in the TCGA liver cancer RNA-seq (n=352) dataset from the cBioPortal database. (B-E) Patients with genetic alteration had poorer overall survival 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 MAPK9, SLC1A4, PCK2, ACSL3, STMN1, and CDO1 in HCC and normal liver tissue. Data were from the Human Protein Atlas database (×400 magnification).


A
Biological Process
response to nutrient levels
response to metal ion
response to starvation
response to oxidative stress
cellular response to metal ion
cellular response to oxidative stress
cellular response to inorganic substance
transition metal ion homeostasis
response to reactive oxygen species
iron ion transport
6
7
8
9
10
EnrichmentScore (−log10(pvalue))
Count
6
8
10
12
pvalue
5e−07
4e−07
3e−07
2e−07
1e−07
B
Cellular Component
neuron projection cytoplasm
pigment granule
melanosome
cytoplasmic region
dendrite cytoplasm
late endosome
mitotic spindle
peroxisome
microbody
blood microparticle
2.6
2.8
3.0
3.2
EnrichmentScore (−log10(pvalue))
Count
2.0
2.5
3.0
3.5
4.0
pvalue
0.0030
0.0025
0.0020
0.0015
0.0010
D
RRM2
STEAP3
Ferroptosis
PPAR signaling pathway
Bladder cancer
CXCL2
HSPB1
MAPK signaling pathway
DUSP1
Non−small cell lung cancer
STMN1
NOD−like receptor signaling pathway
Glioma
MAPK1
CDKN2A
Hepatocellular carcinoma
NRAS
p53 signaling pathway
Pancreatic cancer
SCP2
ACSL3
PCK2
MAPK9
ACSL4
log2 fc
1.0
0.5
0.0
−0.5
−1.0
size
3
4
5
6
category
Ferroptosis
Bladder cancer
Glioma
Hepatocellular carcinoma
MAPK signaling pathway
NOD−like receptor signaling pathway
Non−small cell lung cancer
p53 signaling pathway
Pancreatic cancer
PPAR signaling pathway

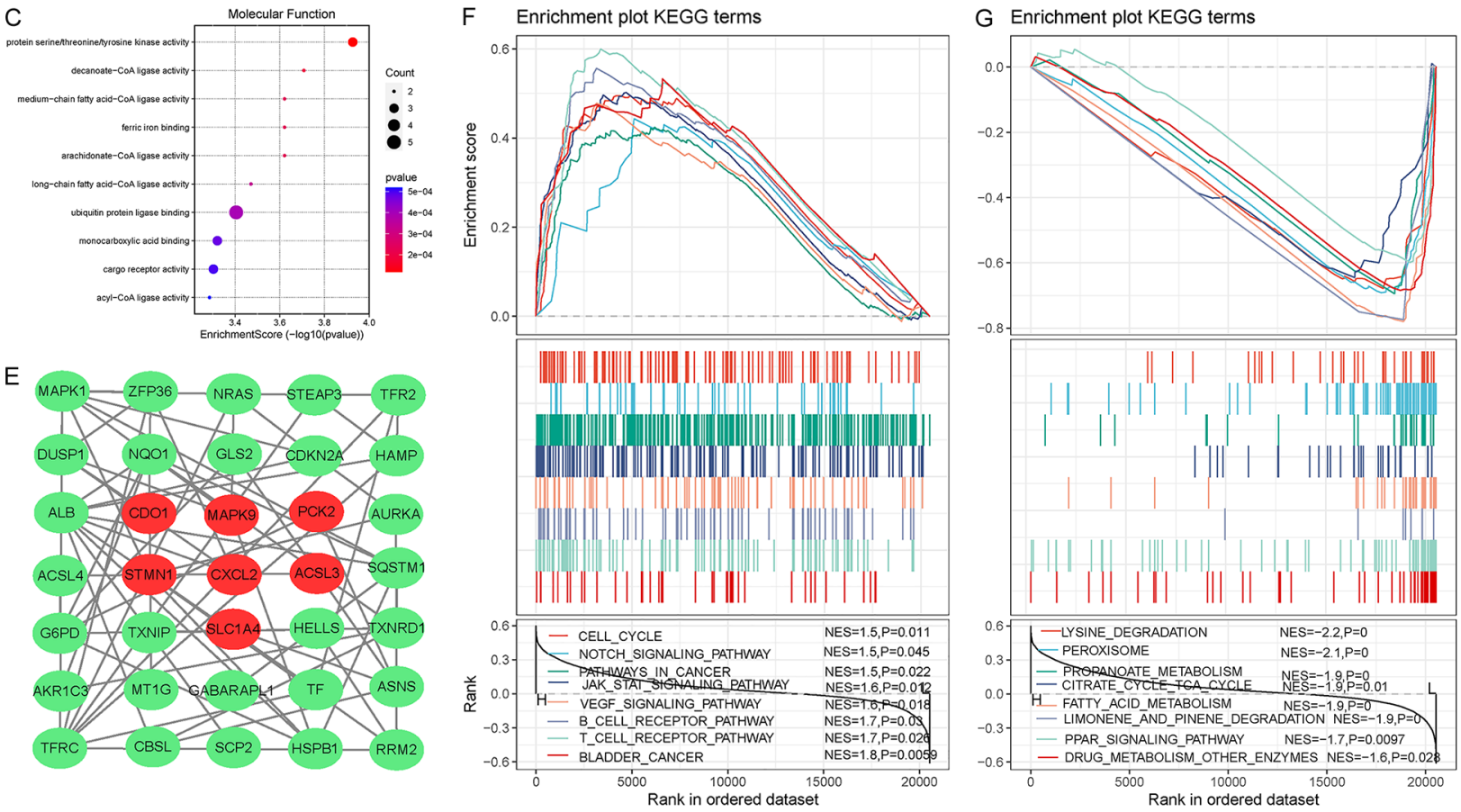



**Figure 8.** Representative results of functional enrichment analysis via GO, KEGG, and GSEA. (A-D) The 39 differentially expressed ferroptosis-related genes were mainly enriched in pathways based on the biological process (A), cellular component (B), molecular function (C) and KEGG pathway (D). (E) Protein-protein interaction network showed the interactions between these 39 ferroptosis-related genes. (F, G) The signaling pathways in which genes of the high-risk score (F) and low-risk score patients (G) were significantly enriched in the GSEA.

## Discussion

Cancer is a broad public health issue and has brought a tremendous economic and health burden worldwide. HCC, a highly aggressive cancer, was the primary cause of cancer-related death in many areas of the world, especially in East Asia and sub-Saharan Africa [27, 28]. Although the improvement in physical examination has increased the early detection and subsequent curative surgery in HCC patients, studies reveal that those who have undergone radical resection still have a 70% recurrence rate within 5 years [5, 29-31]. The high recurrence rate is the main cause for the short overall survival and poor prognosis of HCC patients. Therefore, it is urgent to establish an effective model for predicting postoperative recurrence and identifying high-risk patients early to improve the prognosis of HCC. The traditional predictive model of recurrence uses information about tumor stage, tumor size, microvascular invasion, tumor differentiation, and other relevant clinical characteristics, and supplemented by a single serum alpha-fetoprotein expression. But its specificity and sensitivity were not high enough to distinguish patients with heterogeneity. In recent years, the gene signature based on the mRNA aberrant expression has been reported to address the heterogeneity; thereby, there are many studies on establishing gene signature to improve the diagnosis and prognosis in HCC [32]. Wang et al. established an RNA-binding proteins-related gene signature to predict the overall survival and found that this gene signature could be an independent risk factor for HCC patients [33]. Yang and colleagues constructed a two-gene signature (HNRNPA2B1 and RBM15) to identify and treat HBV-related HCC patients and showed the predictive value for OS [34]. However, few studies were designed to investigate the recurrence-related gene signature of HCC [35]. In our current study, we for the first time established a novel ferroptosis-related 7-gene signature for HCC RFS prediction.

Ferroptosis plays a significant role in inducing HCC cell death and inhibiting cell proliferation and metastasis. Furthermore, many ferroptosis-related genes have been identified in regulating the activity of ferroptosis. Previous publications have reported that DAZAP1 is the ferroptosis suppressor gene and is significantly overexpressed in HCC cells. DAZAP1 also promotes cell proliferation and significantly reduces the cellular sensitivity to sorafenib [36]. Another study reports that metallothionein (MT)-1G increases the sorafenib-resistance of HCC cells by inhibiting the process of ferroptosis [37]. Furthermore, a recent study reveals that ACSL4, a positive-activating enzyme of ferroptosis, can increase the sensitivity of HCC patients to sorafenib by activating ferroptosis [38-40]. These studies focus on the predictive value of ferroptosis-related gene signature for overall survival of HCC patients; however, the significance of ferroptosis-related gene signature in predicting the RFS in HCC is largely unknown [41-44].

Here, we found that ferroptosis-related 7-gene signature (including MAPK9, SLC1A4, PCK2, ACSL3, STMN1, CDO1, and CXCL2) could predict RFS in HCC. The 7-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 worse RFS than patients in the low-risk group. And the correlation analysis demonstrated that high-risk scores correlated with TNM stage, serum AFP level, ALT, predicted risk metastasis PRMS, recurrence, and death. Moreover, the multivariate Cox regression analyses indicated that high-risk score was independent risk factors for RFS. More importantly, all these results were verified in the validation set (TCGA HCC cohort).

We further established a nomogram that integrated all the independent risk factors to predict the 1-, 3-, and 5-year RFS. The results showed an accurate performance in predicting recurrence. In addition, the combined model of nomogram performed better than the individual predictive factors, suggesting its potential in clinical application. We also examined the genetic alteration of this 7-gene signature in the cBioPortal database and found that 27% of the patients had genetic alterations in the 7-gene signature. These genetic changes correlated with poor OS and RFS. Since genetic alteration is responsible for the dysregulation of gene expression [45, 46], the genetic alteration of this 7-gene signature may have predictive value for RFS. Moreover, the KEGG analysis revealed that the FDEGs were significantly enriched in the pathways of Ferroptosis, Hepatocellular carcinoma, MAPK signaling pa-

thway, and other growth-related pathways, confirming the functional importance of these FDECs in cancer development. Finally, our GSEA analysis revealed several significantly enriched tumorigenic signaling pathways. Since the number of ferroptosis-related genes is small, GSEA identifies functions that are not related to ferroptosis but to cancer-related signaling, which might explain the underlying molecular mechanisms of this gene signature.

Nonetheless, our study has several limitations. First, although the 7-gene signature and the predictive nomogram were built and validated by different databases, they need to be further tested in clinical trials. Second, in addition to examine the expression of this gene signature in HCC and normal tissues, the specific functions and molecular mechanisms should be investigated by various in vitro and vitro approaches. In conclusion, our study constructed a 7 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.

## Acknowledgements

We would like to acknowledge all the people who have given us help with our article. This work was supported by The Key Project of Natural Science Foundation of Fujian Province (No. 2019J01533).

## Disclosure of conflict of interest

None.

## 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.

**Address correspondence to:** Fang Yang, Department of Hepatobiliary Surgery, 900 Hospital of The Joint Logistics Team, No. 156 The Second West Ring Road, Fuzhou 350025, Fujian, China. Tel: +86-18060689757; E-mail: yfang105@163.com; Jian Fang, Department of Hepatobiliary Medicine, The Third People's Hospital of Fujian University of Traditional Chinese Medicine, 363 Guobin Road, Fuzhou 350108, Fujian, China. Tel: +86-1387277-0922; E-mail: 55443025@qq.com

## References

[1] Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA and 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: 394-424.

[2] Forner A, Reig M and Bruix J. Hepatocellular carcinoma. Lancet 2018; 391: 1301-1314.

[3] Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, Jemal A, Yu XQ and He J. Cancer statistics in China, 2015. CA Cancer J Clin 2016; 66: 115-132.

[4] Siegel RL, Miller KD, Fuchs HE and Jemal A. Cancer statistics, 2021. CA Cancer J Clin 2021; 71: 7-33.

[5] Zhou T, Cai Z, Ma N, Xie W, Gao C, Huang M, Bai Y, Ni Y and Tang Y. 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, Tada T, Chong CCN, Xiang BD, Li LQ, Lai PBS, Mazzaferro V, García-Fiñana M, Kudo M, Kumada T, Roayaie S and Johnson PJ. Development of pre and post-operative models to predict early recurrence of hepatocellular carcinoma after surgical resection. J Hepatol 2018; 69: 1284-1293.

[7] Villanueva A, Hoshida Y, Battiston C, Tovar V, Sia D, Alsinet C, Cornella H, Liberzon A, Kobayashi M, Kumada H, Thung SN, Bruix J, Newell P, April C, Fan JB, Roayaie S, Mazzaferro V, Schwartz ME and Llovet JM. Combining clinical, pathology, and gene expression data to predict recurrence of hepatocellular carcinoma. Gastroenterology 2011; 140: 1501-1512.

[8] Ren A, Li Z, Zhou X, Zhang X, Huang X, Deng R and Ma Y. 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 and Chen G. Ferroptosis, a new form of cell death, and its relationships with tumourous diseases. J Cell Mol Med 2017; 21: 648-657.
[10] Stockwell BR, Jiang X and Gu W. Emerging mechanisms and disease relevance of ferroptosis. Trends Cell Biol 2020; 30: 478-490.
[11] Mou Y, Wang J, Wu J, He D, Zhang C, Duan C and Li B. Ferroptosis, a new form of cell death: opportunities and challenges in cancer. J Hematol Oncol 2019; 12: 34.
[12] Wang Y, Wei Z, Pan K, Li J and Chen Q. The function and mechanism of ferroptosis in cancer. Apoptosis 2020; 25: 786-798.
[13] Liang C, Zhang X, Yang M and Dong X. Recent progress in ferroptosis inducers for cancer therapy. Adv Mater 2019; 31: e1904197.
[14] Friedmann Angeli JP, Krysko DV and Conrad M. Ferroptosis at the crossroads of cancer-acquired drug resistance and immune evasion. Nat Rev Cancer 2019; 19: 405-414.
[15] Hassannia B, Vandenabeele P and Vanden Berghe T. Targeting ferroptosis to iron out cancer. Cancer Cell 2019; 35: 830-849.
[16] Louandre C, Ezzoukhry Z, Godin C, Barbare JC, Maziere JC, Chauffert B and Galmiche A. Iron-dependent cell death of hepatocellular carcinoma cells exposed to sorafenib. Int J Cancer 2013; 133: 1732-1742.
[17] Zhao Y, Li M, Yao X, Fei Y, Lin Z, Li Z, Cai K, Zhao Y and Luo Z. HCAR1/MCT1 regulates tumor ferroptosis through the lactate-mediated AMPK-SCD1 activity and its therapeutic implications. Cell Rep 2020; 33: 108487.
[18] Sun X, Ou Z, Chen R, Niu X, Chen D, Kang R and Tang D. Activation of the p62-Keap1-NRF2 pathway protects against ferroptosis in hepatocellular carcinoma cells. Hepatology 2016; 63: 173-184.
[19] Zhang Z, Guo M, Li Y, Shen M, Kong D, Shao J, Ding H, Tan S, Chen A, Zhang F and Zheng S. RNA-binding protein ZFP36/TTP protects against ferroptosis by regulating autophagy signaling pathway in hepatic stellate cells. Autophagy 2020; 16: 1482-1505.
[20] Barrett T and Edgar R. Gene expression omnibus: microarray data storage, submission, retrieval, and analysis. Methods Enzymol 2006; 411: 352-369.
[21] Zhou N and Bao J. FerrDb: a manually curated resource for regulators and markers of ferroptosis and ferroptosis-disease associations. Database (Oxford) 2020; 2020: baaa021.
[22] Cancer Genome Atlas Research N, Weinstein JN, Collisson EA, Mills GB, Shaw KR, Ozenberger BA, Ellrott K, Shmulevich I, Sander C and Stuart JM. The Cancer Genome Atlas Pan-Cancer analysis project. Nat Genet 2013; 45: 1113-1120.
[23] Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, Sun Y, Jacobsen A, Sinha R, Larsson E, Cerami E, Sander C and Schultz N. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal 2013; 6: pl1.
[24] Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES and Mesirov JP. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 2005; 102: 15545-15550.
[25] Kanehisa M and Goto S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res 2000; 28: 27-30.
[26] Dennis G Jr, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC and Lempicki RA. DAVID: database for annotation, visualization, and integrated discovery. Genome Biol 2003; 4: P3.
[27] Yang JD, Hainaut P, Gores GJ, Amadou A, Plymoth A and Roberts LR. A global view of hepatocellular carcinoma: trends, risk, prevention and management. Nat Rev Gastroenterol Hepatol 2019; 16: 589-604.
[28] Villanueva A. Hepatocellular carcinoma. N Engl J Med 2019; 380: 1450-1462.
[29] Hasegawa K, Kokudo N, Makuuchi M, Izumi N, Ichida T, Kudo M, Ku Y, Sakamoto M, Nakashima O, Matsui O and Matsuyama Y. Comparison of resection and ablation for hepatocellular carcinoma: a cohort study based on a Japanese nationwide survey. J Hepatol 2013; 58: 724-729.
[30] Zhang X, Li C, Wen T, Yan L, Li B, Yang J, Wang W, Xu M, Lu W and Jiang L. 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: 933-940.
[31] Xu XF, Xing H, Han J, Li ZL, Lau WY, Zhou YH, Gu WM, Wang H, Chen TH, Zeng YY, Li C, Wu MC, Shen F and Yang T. Risk factors, patterns, and outcomes of late recurrence after liver resection for hepatocellular carcinoma: a multicenter study from China. JAMA Surg 2019; 154: 209-217.
[32] Liu GM, Zeng HD, Zhang CY and Xu JW. Identification of a six-gene signature predicting overall survival for hepatocellular carcinoma. Cancer Cell Int 2019; 19: 138.
[33] Wang L, Zhou N, Qu J, Jiang M and Zhang X. Identification of an RNA binding protein-related

gene signature in hepatocellular carcinoma patients. Mol Med 2020; 26: 125.
[34] Fang Q and 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: 60.
[35] Long J, Chen P, Lin J, Bai Y, Yang X, Bian J, Lin Y, Wang D, Yang X, Zheng Y, Sang X and Zhao H. DNA methylation-driven genes for constructing diagnostic, prognostic, and recurrence models for hepatocellular carcinoma. Theranostics 2019; 9: 7251-7267.
[36] Wang Q, Guo Y, Wang W, Liu B, Yang G, Xu Z, Li J and Liu Z. RNA binding protein DAZAP1 promotes HCC progression and regulates ferroptosis by interacting with SLC7A11 mRNA. Exp Cell Res 2021; 399: 112453.
[37] Sun X, Niu X, Chen R, He W, Chen D, Kang R and Tang D. Metallothionein-1G facilitates sorafenib resistance through inhibition of ferroptosis. Hepatology 2016; 64: 488-500.
[38] Feng J, Lu PZ, Zhu GZ, Hooi SC, Wu Y, Huang XW, Dai HQ, Chen PH, Li ZJ, Su WJ, Han CY, Ye XP, Peng T, Zhou J and Lu GD. ACSL4 is a predictive biomarker of sorafenib sensitivity in hepatocellular carcinoma. Acta Pharmacol Sin 2021; 42: 160-170.
[39] Xia H, Lee KW, Chen J, Kong SN, Sekar K, Deivasigamani A, Seshachalam VP, Goh BKP, Ooi LL and Hui KM. 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.
[40] Doll S, Proneth B, Tyurina YY, Panzilius E, Kobayashi S, Ingold I, Irmler M, Beckers J, Aichler M, Walch A, Prokisch H, Trumbach D, Mao G, Qu F, Bayir H, Fullekrug J, Scheel CH, Wurst W, Schick JA, Kagan VE, Angeli JP and Conrad M. ACSL4 dictates ferroptosis sensitivity by shaping cellular lipid composition. Nat Chem Biol 2017; 13: 91-98.
[41] Liu Y, Zhang X, Zhang J, Tan J, Li J and Song Z. Development and validation of a combined ferroptosis and immune prognostic classifier for hepatocellular carcinoma. Front Cell Dev Biol 2020; 8: 596679.
[42] Tang B, Zhu J, Li J, Fan K, Gao Y, Cheng S, Kong C, Zheng L, Wu F, Weng Q, Lu C and 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.
[43] Liang JY, Wang DS, Lin HC, Chen XX, Yang H, Zheng Y and Li YH. A novel ferroptosis-related gene signature for overall survival prediction in patients with hepatocellular carcinoma. Int J Biol Sci 2020; 16: 2430-2441.
[44] Du X and Zhang Y. Integrated analysis of immunity- and ferroptosis-related biomarker signatures to improve the prognosis prediction of hepatocellular carcinoma. Front Genet 2020; 11: 614888.
[45] Jia P and Zhao Z. Impacts of somatic mutations on gene expression: an association perspective. Brief Bioinform 2017; 18: 413-425.
[46] Wang H, Xu F, Lu L, Yang F, Huang X, Lv L, Hu H and Jiang Y. The diagnostic and prognostic significance of small nuclear ribonucleoprotein Sm D1 aberrantly high expression in hepatocellular carcinoma. J Cancer 2022; 13: 184-201.