



A Novel RNA Binding Protein-Related Prognostic Signature for Hepatocellular Carcinoma

Yongbiao Huang^{1†}, Sheng Chen^{2†}, Wan Qin¹, Yali Wang¹, Long Li¹, Qianxia Li^{1*} and Xianglin Yuan^{1*}

¹ Department of Oncology, Tongji Hospital, Huazhong University of Science and Technology, Wuhan, China, ² Department of Hepatobiliary Surgery, Affiliated Hospital of Hebei University, Baoding, China

OPEN ACCESS

Edited by:

You Zhou,
First people's Hospital of Changzhou,
China

Reviewed by:

Yingjie Shao,
Soochow University, China
Meiling Jin,
City of Hope National Medical Center,
United States

*Correspondence:

Qianxia Li
liqianx110@163.com
Xianglin Yuan
yuanxianglin@hust.edu.cn

[†]These authors have contributed
equally to this work

Specialty section:

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

Received: 06 July 2020

Accepted: 24 August 2020

Published: 28 October 2020

Citation:

Huang Y, Chen S, Qin W, Wang Y, Li L,
Li Q and Yuan X (2020) A Novel RNA
Binding Protein-Related
Prognostic Signature for
Hepatocellular Carcinoma.
Front. Oncol. 10:580513.
doi: 10.3389/fonc.2020.580513

Hepatocellular carcinoma (HCC) is a highly malignant and aggressive cancer with high recurrence rates and mortality. Some studies have illustrated that RNA binding proteins (RBPs) were involved in the carcinogenesis and development of multiple cancers, but the roles in HCC were still unclear. We downloaded the RNA-seq and corresponding clinical information of HCC from The Cancer Genome Atlas (TCGA) database, and 330 differentially expressed RBPs were identified between normal and HCC tissues. Through series of the univariate, the least absolute shrinkage selection operator (LASSO), and the stepwise multivariate Cox regression analyses, six prognosis-related key RBPs (CNOT6, UPF3B, MRPL54, ZC3H13, IFIT5, and PPARGC1A) were screened out from DE RBPs, and a six-RBP gene risk score signature was constructed in training set. Survival analysis indicated that HCC patients with high-risk scores had significantly worse overall survival than low-risk patients, and furthermore, the signature can be used as an independent prognostic indicator. The good accuracy of this prognostic signature was confirmed by the ROC curve analysis and was further validated in the International Cancer Genome Consortium (ICGC) HCC cohort. Besides, a nomogram based on six RBP genes was established and internally validated in the TCGA cohort. Gene set enrichment analysis demonstrated some cancer-related phenotypes were significantly gathered in the high-risk group. Overall, our study first identified an RBP-related six-gene prognostic signature, which could serve as a promising prognostic biomarker and provide some potential therapeutic targets for HCC.

Keywords: hepatocellular carcinoma, RNA binding proteins, prognostic signature, risk score, nomogram

INTRODUCTION

Liver cancer, the fifth most frequent type of malignancy with high morbidity and mortality, has become the second leading cause of cancer death globally. It was estimated that 841,080 new liver cancer cases were diagnosed worldwide and 781,631 deaths occurred in 2018 (1, 2). Hepatocellular carcinoma (HCC), the main pathological type of primary liver cancer, represents approximately 80–90% of all liver cancer

cases (3, 4). At present, the major treatments include systemic pharmacological treatment, surgical resection, transplantation, ablation therapies, transcatheter arterial chemoembolization, and radiotherapy (3, 5). In recent decades, the incidence and mortality of HCC has been increasing globally (1). In spite of the significant progress made in diagnosis and treatment, the prognosis for patients with HCC still remains poor due to the high complexity and heterogeneity of hepatocarcinogenesis (3). Therefore, it is critical to identify prognostic biomarkers and develop novel accurate prediction models for predicting prognosis of patients with HCC and guiding clinical therapy.

RNA binding proteins (RBPs) play a crucial role in post-transcriptional gene regulation (6). RBPs can bind various types of RNAs include coding RNAs (mRNAs) and non-coding RNAs (rRNAs, ncRNAs, snRNAs, miRNAs, tRNAs, snoRNAs) through an RNA-binding domain directly (7, 8). So far, more than 1,500 human RBPs (7.5% of the proteome) have been identified that contain 600 structurally distinct RNA-binding domains (7). They form ribonucleoprotein complexes by binding their target RNAs and regulate RNA metabolism, include RNA maturation, splicing, transport, localization, polyadenylation, stability, degradation, and translation (8, 9). Most RBPs are evolutionarily conserved and ubiquitously expressed to maintain cellular homeostasis (7, 10). Due to the extremely significant biological function of RBPs, its dysfunction can lead to the occurrence of multiple diseases, including cardiovascular system diseases (11), blood diseases (10), neurodegenerative disorders (12), and cancers (11–14).

Previous published studies have indicated that aberrant expression of some RBPs can affect cell growth and proliferation and promote tumor occurrence and progression (15). In addition, its aberrant expression is also significantly related to malignant degree and clinical prognosis of patients with cancer (16). For instance, the RNA binding proteins Musashi-1 and Musashi-2 were found to be overexpressed in colorectal cancer, and they regulate the mRNA stability and translation in essential oncogenic signaling pathways (17). Negative elongation factor E (NELFE) promotes metastasis of pancreatic cancer through activating the Wnt/ β -catenin signaling pathway and decreasing the NDRG2 mRNA stabilization (18). Human ribosomal protein S3 (RPS3) is upregulated in HCC and is closely relevant to the prognosis of patients with HCC. RPS3 stabilized SIRT1 mRNA through binding with the 3' UTR of SIRT1 mRNA to sustaining HCC progression and the somatic copy-number alterations of NELFE enhanced MYC signaling and promote cell proliferation in HCC (19, 20). The molecular mechanism by which RBPs promote carcinogenesis and development is still not clear.

Abbreviations: HCC, Hepatocellular carcinoma; RBPs, RNA binding proteins; TCGA, The Cancer Genome Atlas; ICGC, International Cancer Genome Consortium; OS, Overall survival; ROC, Receiver operating characteristic; DE RBPs, Differentially expressed RBPs; GO, Gene ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; FC, Fold change; FDR, false discovery rate; PPI, Protein-protein interaction; MCODE, Molecular Complex Detection; LASSO, Least absolute shrinkage selection operator; AUC, Area under the curve; GSEA, Gene set enrichment analysis; HPA, Human Protein Atlas; BP, Biological process; CC, Cell component; MF, Molecular function; NES, Normalized enrichment score.

Consequently, we considered that RBPs were potential prognostic biomarkers for HCC patients. In our study, the RNA-seq data and corresponding clinical information of HCC cases were obtained from The Cancer Genome Atlas (TCGA) database, and then we identified differentially expressed RBPs between tumor and normal liver tissue. Based on differentially expressed RBPs, survival related RBPs were screened out and an RBP-associated prognostic model was constructed to predict the clinical outcome of HCC patients. The prognostic value of this model was validated in another HCC cohort from the International Cancer Genome Consortium (ICGC) database.

MATERIALS AND METHODS

TCGA HCC Dataset and Difference Analysis

The normalized RNA-seq data (Fragments Per Kilobase Million, FPKM) and corresponding clinical data, which contained 374 HCC samples and 50 normal liver tissue samples, were downloaded from TCGA database as training set. Wilcoxon Test was utilized to perform difference analysis and identify the differentially expressed RBPs (DE RBPs) between the HCC and normal tissue. RBPs with $|\log_2$ fold change (FC)| ≥ 0.5 and adj P-value < 0.05 were used for subsequent analysis.

GO and KEGG Functional Enrichment Analyses

To explore main biological functions and signaling pathways of the differently expressed RBPs, the R package “clusterProfiler” was used to carry out Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway and Gene Ontology (GO) enrichment analyses (21), and the results were visualized via “GOplot” R package. The false discovery rate (FDR) < 0.05 was thought to be statistically significant.

PPI Network

Protein-protein interaction (PPI) networks in differently expressed RBPs were constructed by using the STRING database and visualized via Cytoscape software (22). The Cytoscape plugin Molecular Complex Detection (MCODE) was used to detect the important modules in PPI network (23), and GO and KEGG analyses were conducted to further investigate their molecular function in HCC.

Prognosis-Related Key RBPs Screening

The univariate Cox regression analysis was carried out to find the prognosis-related RBPs among the differentially expressed RBPs via “survival” R package, and P-value < 0.01 were considered for subsequent analysis, using the least absolute shrinkage selection operator (LASSO) regression analysis to further screen prognostic-related RBP genes with “glmnet” R package. Finally, the stepwise multivariate regression analysis was performed to screen out optimal key prognostic-related RBP genes and obtain their standardized regression coefficients.

Survival, Expression, and Genetic Alteration Analyses of Key Prognosis-Related RBP Genes

The Kaplan–Meier curves survival was utilized to evaluate the prognostic value of each key RBP gene in TCGA cohort, and P-value < 0.05 was considered to have statistical difference. The copy-number alterations and mutations were detected with the online database cBioPortal (24), and the protein expression was detected by Human Protein Atlas (HPA) database (25).

Construction of an RBP-Genes Prognostic Signature

A risk score signature was constructed by using multivariate Cox regression based on the previously obtained RBPs using the survival R package in TCGA. The risk score was calculated by the following formula: Risk score = Expression of gene1 × Coefficient of gene1 + Expression of gene2 × Coefficient of gene2 + ... Expression of geneN × Coefficient of geneN (26, 27). By the median value of the risk score, all HCC patients were assigned into low-risk groups and high-risk groups, and the Kaplan–Meier curve analysis and log-rank test were used to assess the survival difference between two subgroups by “Survival” R package. The receiver operating characteristic (ROC) curves were plotted and the area under the curve (AUC) values were calculated with “SurvivalROC” R package, which was used to evaluate the predictive power (28). Then, the LIRI-JP project in ICGC dataset contained 229 HCC patient cases with complete clinical information and follow-time more than 1 month was used as testing set to validate the predictive capacity of this model (29). In addition, the univariate and multivariate Cox regression analyses were utilized to determine the correlation between RBP signature and clinical characteristics and OS in the TCGA and ICGC cohort, respectively. The statistical difference of risk scores between the stratified clinicopathologic features was calculated by using the Kruskal–Wallis test.

After that, a prognostic nomogram based on key prognosis-related RBP genes was generated by using “rms” R package to predict OS of HCC patients at 1-, 3-, and 5-years in the TCGA cohort. Meanwhile, the calibration curves were plotted to appraise the prognostic performance of the nomogram.

Gene Set Enrichment Analysis

Gene set enrichment analysis (GSEA) was conducted with GSEA v4.0.3 software to identify different signaling pathways between two subgroups. Hallmark gene sets (h.all.v6.0.symbol.gmt) were downloaded from Molecular Signatures Database as the reference gene set (30). Nominal p-value < 0.05 and FDR q-value < 0.05 were set as the cut-off.

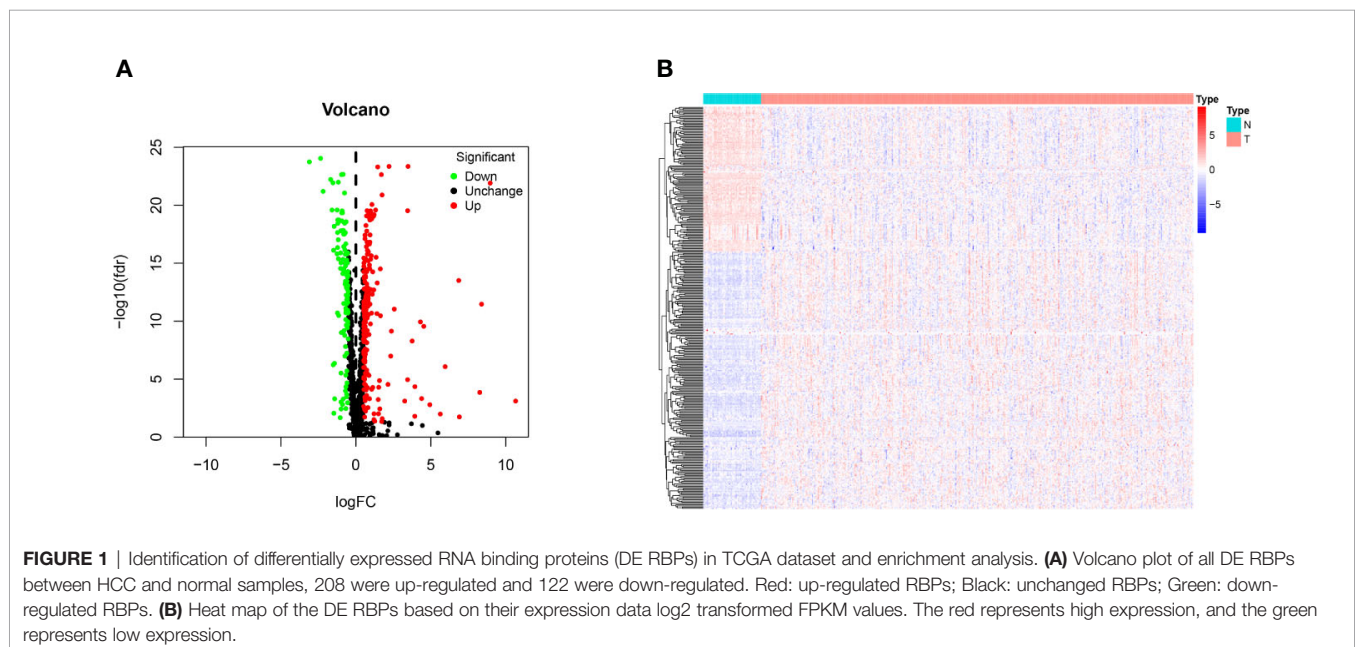
RESULTS

Identification of the Differentially Expressed RBPs

The 1,542 human RBPs found so far were included in our study (7), a total of 330 DE RBPs were identified by Wilcoxon Test between 374 HCC tissues and 50 normal liver tissues, including 208 upregulated and 122 downregulated RBPs, according to the adj P-value < 0.05, $|\log_2\text{FC}| \geq 0.5$ (Figure 1).

Enrichment Analysis of DE RBPs

We carried out the GO and KEGG pathway enrichment analyses of the DE RBPs in HCC by using the R package “clusterProfiler.” GO analysis consists of biological process (BP), cellular component (CC), and molecular function (MF). The DE RBPs were significantly gathered in ncRNA processing, RNA splicing, regulation of translation, RNA catabolic process, and RNA phosphodiester bond hydrolysis of the BP category (Figure 2A); cytoplasmic ribonucleoprotein granule, ribonucleoprotein



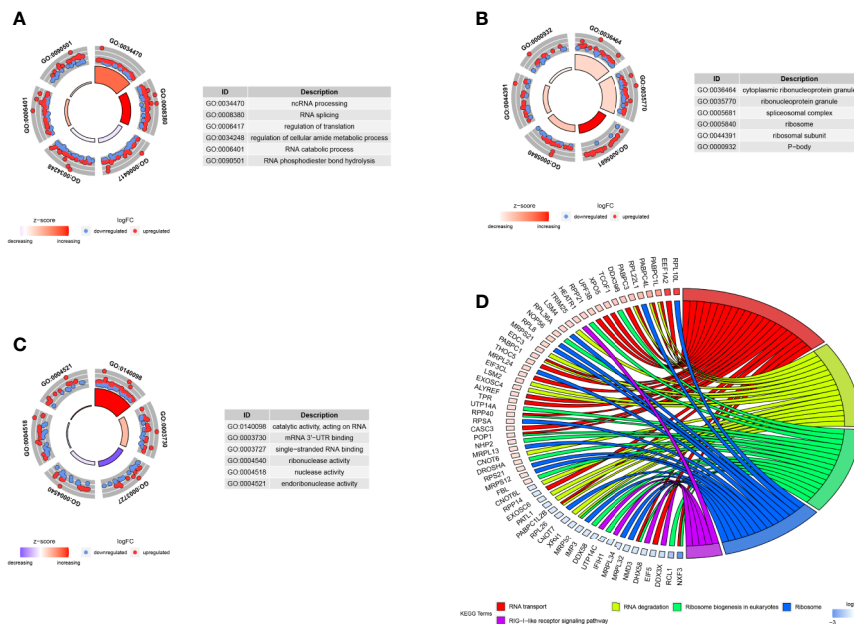


FIGURE 2 | GO and KEGG enrichment analyses of DE RBPs. **(A–C)** Top six enriched GO terms respectively enriched in Biological processes (BP), Cellular components (CC), and Molecular functions (MF); **(D)** Five significantly enriched KEGG pathways for DE RBPs. The outer circle shows a scatter plot for each term or pathway of the logFC of the assigned genes, red circles represent up-regulation and blue represent down-regulation.

granule, spliceosomal complex, ribosome, ribosomal subunit, and P-body of the CC analysis (**Figure 2B**); RNA catalytic activity, mRNA 3'-UTR binding, single-stranded RNA binding, ribonuclease activity, nuclease activity, and endoribonuclease activity of the MF analysis (**Figure 2C**). The KEGG analysis result indicated that the DE RBPs were significantly gathered in RNA transport and degradation, mRNA surveillance pathway, spliceosome, ribosome, ribosome biogenesis in eukaryotes, and RIG-I-like receptor signaling pathway (**Figure 2D**). Many emerging studies have suggested that RBPs participate in RNA metabolism and formation of mRNA spliceosomal complex, and mediate post-transcriptional gene regulation. The ribosome is a kind of ribonucleoprotein granule and is considered as a molecular machine for protein synthesis. Some RBPs are closely related to ribonucleoprotein formation, they can assemble specific RNAs to form ribonucleoprotein granules in eukaryotic cells, like P-bodies and stress granules. P-bodies are conserved cytoplasmic ribonucleoprotein granules in eukaryotic organisms, which involved in translational repression and mRNA decay and degradation (31–34). These results suggested that RBPs play an essential role in RNA processing and protein synthesis, and their aberrant expression could promote carcinogenesis and progression of a variety of tumors.

PPI Network Analysis

For further understanding the function of DE RBPs in HCC procession, we constructed a PPI network that consists of 163 nodes and 1,047 edges by using STRING database and Cytoscape software (**Figure 3A**). Moreover, pivotal modules were identified

from the PPI network using MODE plug in Cytoscape. Module 1 included 23 upregulated DE RBPs and 3 downregulated DE RBPs (**Figure 3B**), and enrichment analysis indicated they were correlated with RNA splicing, RNA 3'-end processing, and mRNA surveillance. Module 2 included 14 upregulated DE RBPs and 7 downregulated DE RBPs (**Figure 3C**), significantly enriched in ncRNA processing, rRNA processing, and ribosome biogenesis. Module 3 included 7 upregulated DE RBPs and 7 downregulated DE RBPs (**Figure 3D**), related to mitochondrial gene expression, mitochondrial translational termination, and elongation.

Selection of Prognosis-Related RBPs

We download the RNA-seq data and corresponding clinical information of HCC patients from the TCGA and ICGC databases, the TCGA HCC cohort as training set, the ICGC HCC cohort as testing set, and the clinical detailed characteristics were collated in **Table 1**. A total of 343 HCC cases with follow-up time more than 30 days in training set were included in the next series of analyses. The univariate Cox regression analysis was utilized to screen prognosis-related RBPs on DE RBPs by survival R package, and 37 survival-related RBPs among the DE RBPs were identified ($p < 0.01$) (**Figure 4A**). Thereafter, the LASSO regression analysis was conducted for further decreasing the number of survival-related RBPs using 10-fold cross validation via “glmnet” R package (**Figures 4B, C**). Finally, we obtained six key prognosis-related RBP genes: CNOT6, UPF3B, MRPL54, ZC3H13, IFIT5, and PPARGC1A by stepwise multivariate regression analysis (**Figure 4D** and **Table 1**).

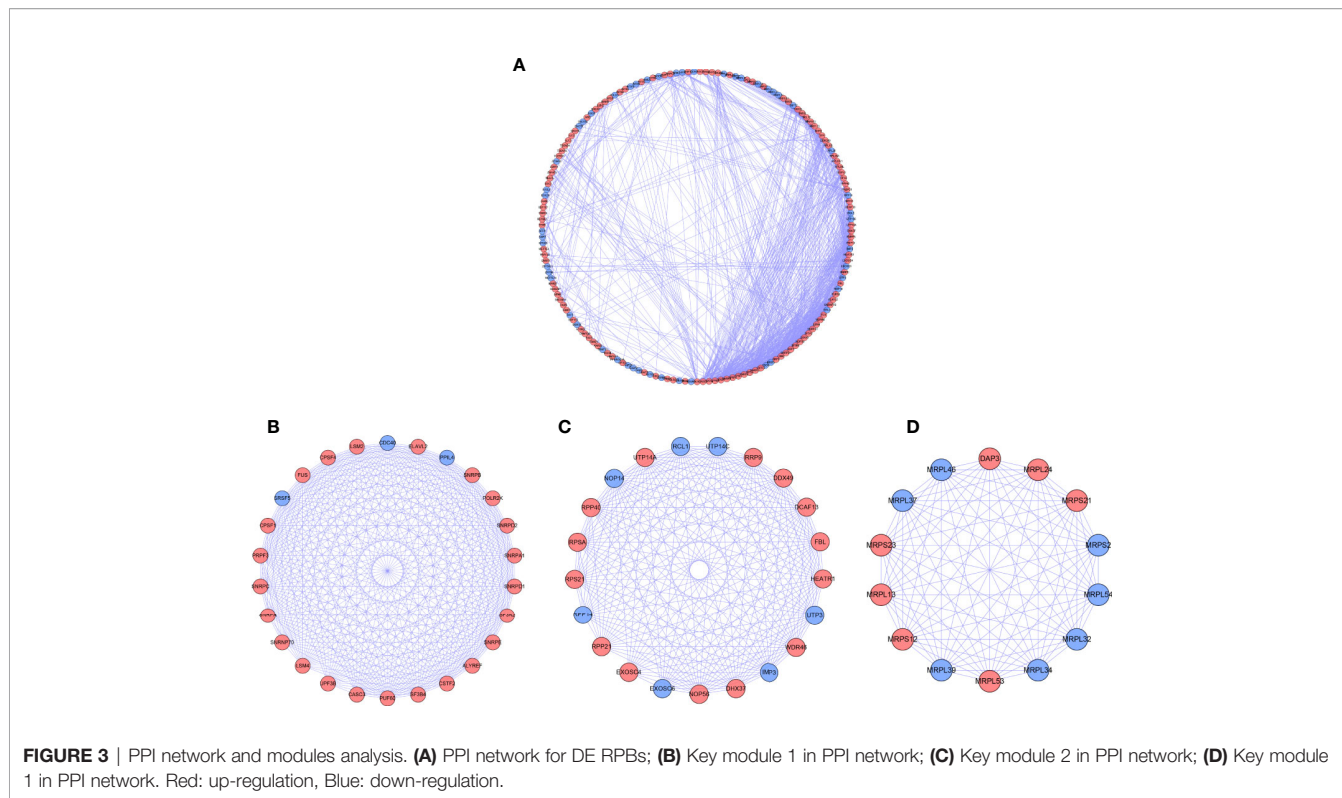


TABLE 1 | The clinical Characteristics of HCC patients from TCGA and ICGC database.

Characteristics	Detailed data		
	TCGA cohort (n=343)	ICGC cohort (n=229)	
Status	Dead	117 (34.11%)	40 (17.47%)
	Survival	226 (65.89%)	189 (82.53%)
Age at diagnosis (years)	≤65	219 (63.85%)	89 (38.86%)
	>65	124 (36.15%)	140 (61.14%)
Gender	Female	107 (31.20%)	61 (26.64%)
	Male	236 (68.80%)	168 (73.36%)
Histological grade	G1	53 (15.45%)	NA
	G2	164 (47.81%)	NA
	G3	113 (32.94%)	NA
	G4	13 (3.79%)	NA
TNM stage	I	164 (47.81%)	36 (15.72%)
	II	77 (22.45%)	106 (46.29%)
	III	81 (23.62%)	68 (29.69%)
	IV	3 (0.87%)	19 (8.31%)
	NA	18 (5.25%)	0 (0.00%)

NA, not available.

Expression, Alteration and Survival Analyses of the Six Prognosis-Related RBP Genes

We further analyzed the expression of these RBPs via HPA database, and the immunohistochemistry results of five key RBPs

in HCC and normal tissues were presented in **Figure 5A**, with PPARGC1A not included in the database. By using the cBioPortal online database, we found that 39 out of the 366 HCC patients (11%) have genetic alterations (mutations and copy-number alterations) in the six RBP genes, and ZC3H13 with the highest alteration frequency (**Figure 5B**). The six key RBP genes were analyzed by using Kaplan–Meier curve analysis to further verify their prognostic value on the TCGA cohort, and the results demonstrated that HCC patients with UPF3B and CNOT6 low-expressions had longer OS, while patients with IFIT5, MRPL54, PPARGC1A, and ZC3H13 high-expression had better survival rate (**Figure 5C**).

Construction Validation of the RBP-Related Risk Score Signature

Based on the previously obtained six key prognosis-related RBP genes, we established a risk core model, and the risk score of each HCC patients was calculated using the formula: Risk score = (0.34900 × CNOT6 Exp) + (0.50277 × UPF3B Exp) + (-0.43143 × MRPL54 Exp) + (-0.21809 × ZC3H13 Exp) + (-0.46413 × IFIT5 Exp) + (-0.19919 × PPARGC1A Exp). Among these six prognosis-related RBPs, CNOT6 and UPF3B were high-risk factors (HR>1); MRPL54, ZC3H13, IFIT5, and PPARGC1A were protective factors (HR<1) (**Table 2**).

All of the 343 HCC patients were assigned into high-risk (n = 171) and low-risk groups (n = 172) using the median risk score in the testing cohort. Low-risk patients had a significantly longer OS compared with the patients in high-risk group (p=7.588e-07) (**Figure 6A**). The AUC value for this six-RBP

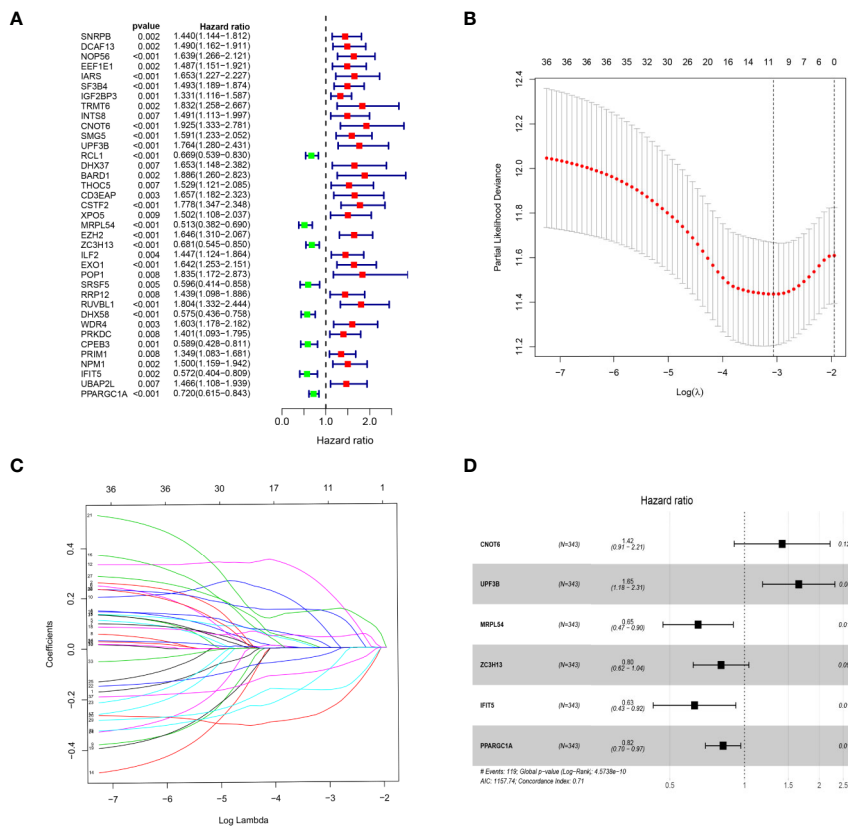


FIGURE 4 | Selection of prognosis-related RBPs in the training cohort. **(A)** Univariate Cox regression analysis; **(B, C)** LASSO regression analysis; **(D)** Multivariate Cox regression analysis to screen out the key RBPs most relevant to prognosis.

gene risk score signature was 0.762 in the 1-year ROC curve, 0.737 in the 3-year ROC curve, and 0.692 in the 5-year ROC curve (**Figure 6B**). The risk scores and survival status distribution of HCC patients between two subgroups were presented in **Figures 6C, D**. We found that as the risk score increased, the number of HCC deaths also increased in the training set. The heatmap of six RBP genes expression level was shown in **Figure 6E**.

Next, to further verify the prognostic performance of this model, we collected 229 HCC cases with follow-up time >30d as the testing set from the ICGC database, and we used the same formula to calculate their risk score. Same as TCGA cohort, according to the cut-off value of TCGA cohort, the results showed that patients with high-risk scores (n=141) had a worse OS than those in low-risk group (n=88) (p=2.55e-2), the AUC value of 1-, 3-, 5-year was 0.822, 0.738 and 0.631, respectively (**Figure 7**). These results indicated that our prognostic signature had considerable robustness in predicting OS for HCC patients.

Association Between Clinical Characteristics and the Six-RBP Gene Signature

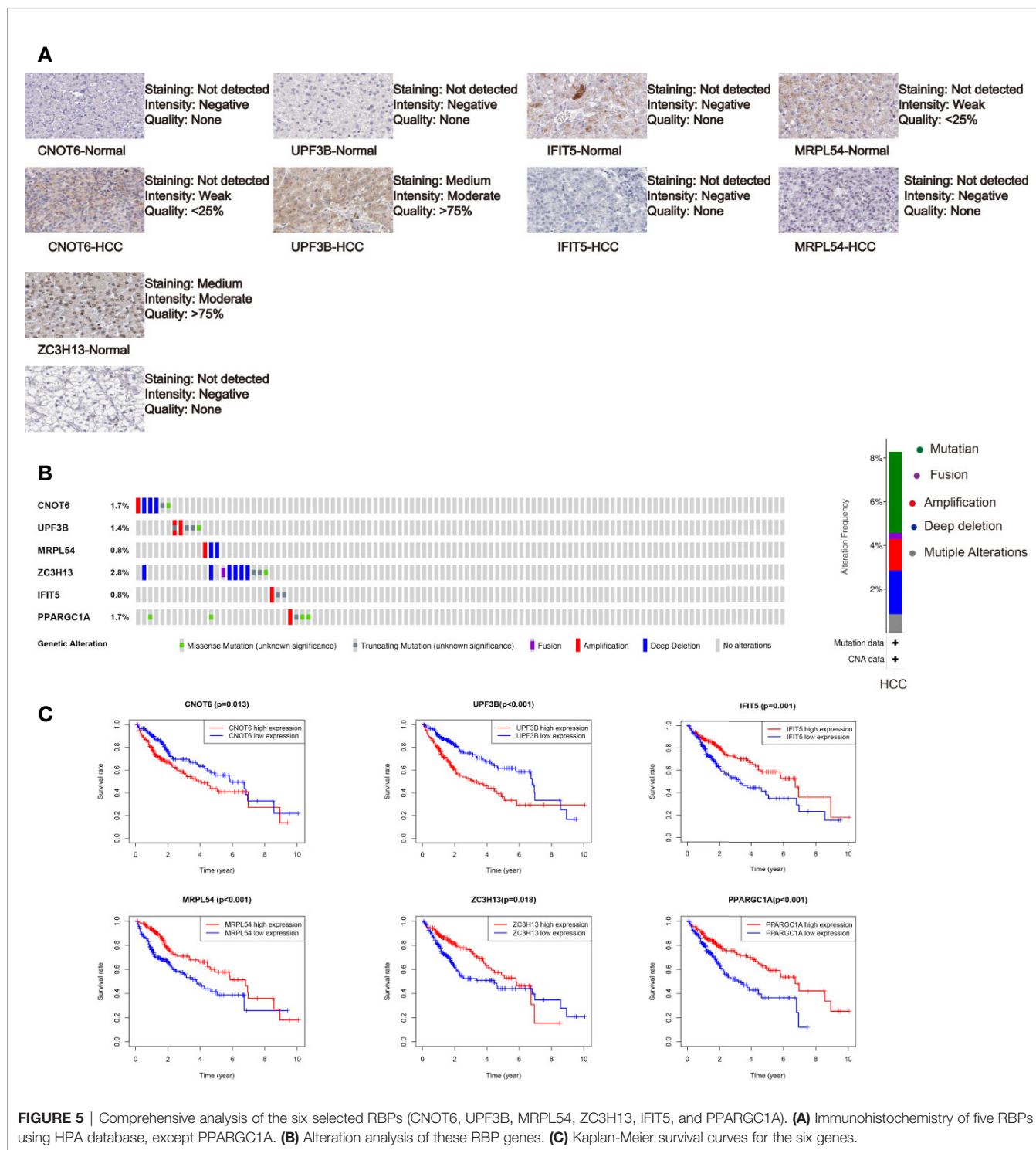
Univariate and multivariate Cox regression analyses were performed for clinical features: age, gender, grade, stage, and

risk score in training and testing set respectively. The results demonstrated the stage (P<0.001) and risk score (P<0.001) were independent prognostic indicators in the TCGA cohort (**Figures 8A, B** and **Table 3**), whereas in the ICGC cohort, the gender (P=0.014352), stage (P<0.001), and risk score (P<0.001) were independent prognostic indicators (**Figures 8C, D** and **Table 3**).

As shown in **Figure 9A**, we found that most dead patients had higher risk scores, which suggested that high-risk patients usually had worse clinical outcomes. Moreover, HCC patients with advanced tumor clinicopathological parameters (stage II and stage III, G3 and G4, pT3 and pT4) were more likely to have higher risk scores than patients with early-stage HCC. We conducted further survival analyses that were stratified by clinical characteristics, and patients in the low-risk group had greater OS than high-risk in all clinical characteristics for stratification survival analyses, including age, gender, grade, and stage (**Figure 9B**).

A Nomogram Establishment on the Six Key Prognosis-Related RBP Genes

The selected six key prognosis-related RBP genes were used to establish a prognostic nomogram through the multivariate Cox regression analysis. We can plot a perpendicular line between the total points axis and each prognostic axis, and estimated the survival probability of HCC patients at 1-, 3-, and 5-year (**Figure 10A**). We



also drew the calibration curves, which indicated that nomogram had good prediction performance in HCC patients (Figure 10B).

GSEA Analysis

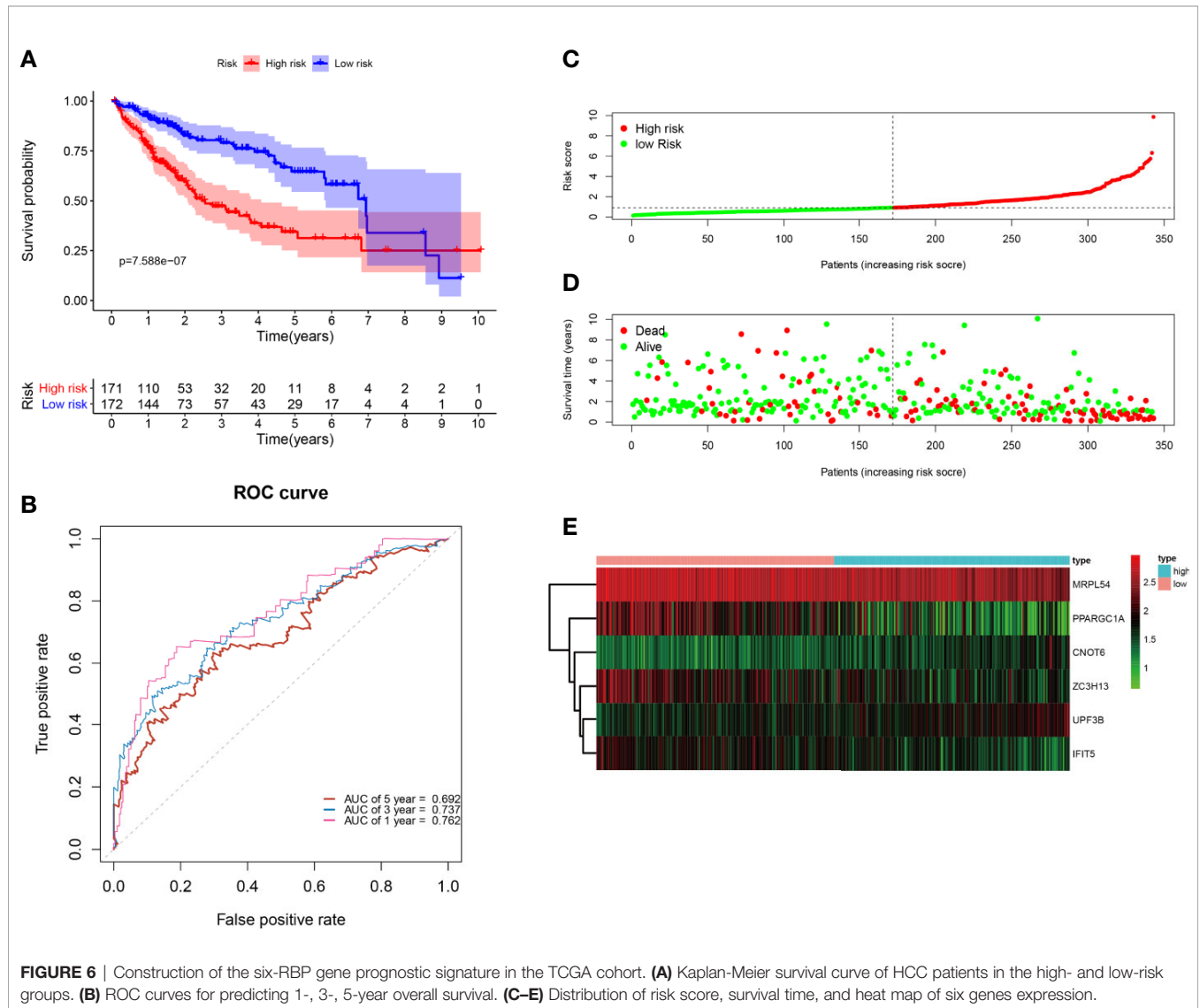
To further explore biological functions and pathways correlated with the risk score signature, GSEA was carried out between high-

and low-risk groups in the TCGA HCC cohort. Some cancer-related gene sets were significantly gathered in HCC patients with high risk score, including “DNA repair,” “MYC targets V1,” “mTORC1 signaling,” “PI3K-AKT-mTOR signaling,” “glycolysis,” “G2M checkpoint,” “E2F targets,” “Wnt/beta-catenin Signaling,” “P53 pathway,” shown in Figure 11.

TABLE 2 | The six prognosis-associated key RBPs identified by multivariate Cox regression analysis.

RBP name	Coefficient	HR	Lower 95% CI	Upper 95% CI	P-value
CNOT6	0.349007	1.417659	0.908826	2.21138	0.123921
UPF3B	0.502772	1.653298	1.18189	2.312733	0.003327
MRPL54	-0.431433	0.649577	0.468813	0.90004	0.009517
ZC3H13	-0.218092	0.804051	0.62116	1.040793	0.097658
IFIT5	-0.464131	0.628681	0.428497	0.922387	0.017643
PPARGC1A	-0.199195	0.81939	0.695361	0.965543	0.017375

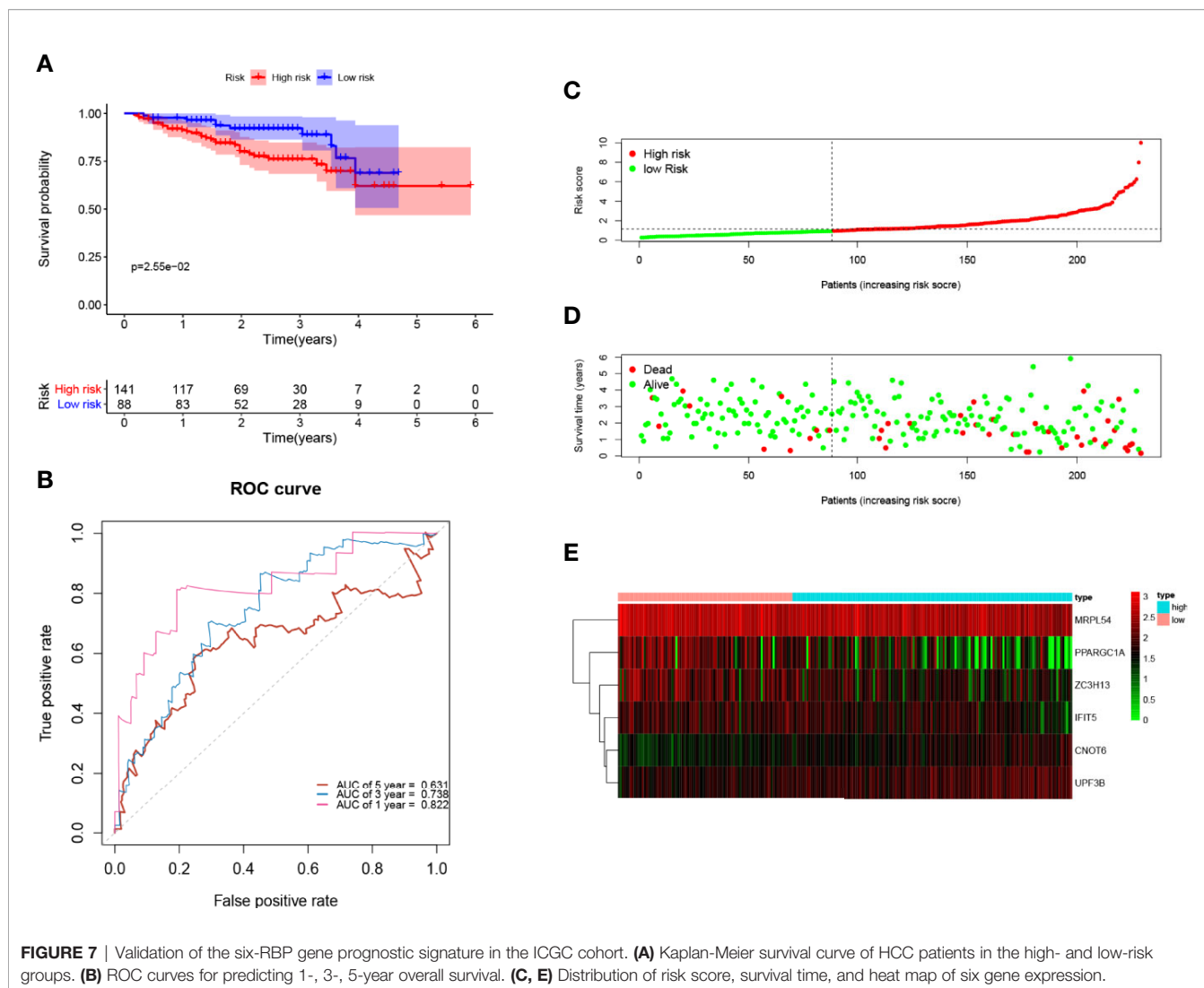
HR, hazard ratio; CI, confidence interval.



DISCUSSION

HCC has become a severe health concern in China; its incidence and mortality rate are still gradually rising due to hepatitis virus infection (1–3). Some research has shown that RBPs were closely related to the tumorigenesis and development

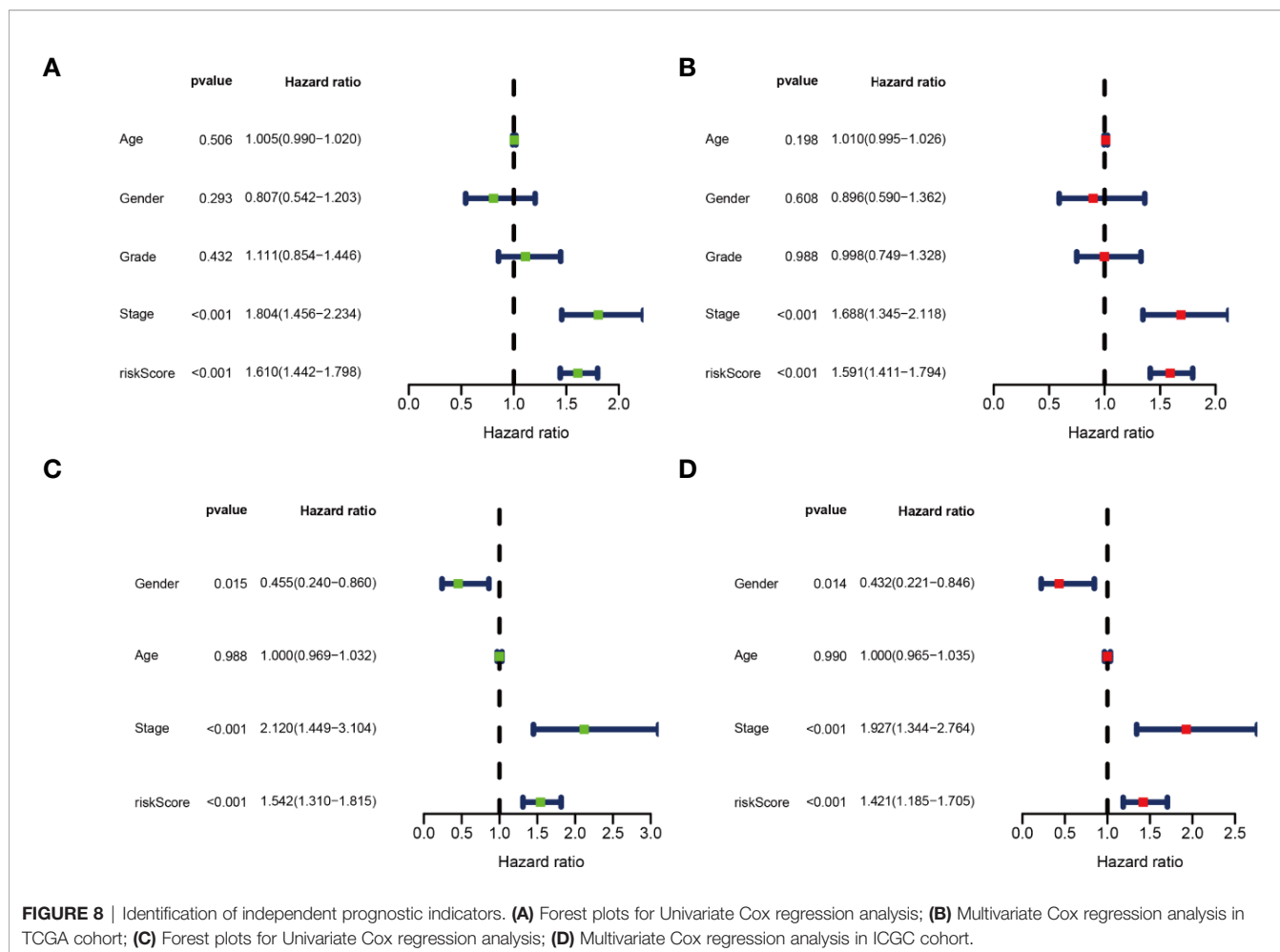
of multiple cancers, but what role RBPs play in HCC were still unclear (11–14). In this study, we integrated the RNA-seq data of HCC from TCGA database, then identified 330 DE RBPs. The PPI network and functional enrichment analysis were conducted to explore the biological function and signaling pathways of DE RBPs in HCC. Next, we filtered out six key RBP genes (CNOT6,



UPF3B, MRPL54, ZC3H13, IFIT5, and PPARGC1A) that were most relevant to prognosis by using the univariate, LASSO, and multivariate Cox regression analyses. Based on the six prognosis-related RBP genes, we established a promising six-RBP gene signature and nomogram to predict OS of HCC patients and validated its robustness in the ICGC cohort. The HCC patients were assigned into two subgroups, high- and low-risk groups, and patients in the high-risk group had poorer outcomes. Next, the GSEA analysis was utilized to investigate the differences in some critical signaling pathways between two subgroups in the TCGA HCC cohort.

Among the six key RBP genes, CNOT6 and UPF3B were highly expressed in the HCC tissues relative to the normal tissues and were considered as unfavorable factors that may lead to worse overall survival. Four genes (MRPL54, ZC3H13, IFIT5, and PPARGC1A) were downregulated and may function as tumor suppressor genes in HCC, and showed a positive correlation with prognosis. CNOT6 encodes Ccr4a protein that was a deadenylase subunit of the CCR4-Not complex (5, 35), and

the CNOT6 rs2453176 C>T polymorphism was related to an increased risk of lung cancer (36). Previous research found that CNOT6 was overexpressed in non-metastatic lung squamous cell carcinoma, and it may be associated with low invasiveness (37). Moreover, the CNOT6 expression level was significantly lower in acute leukemia patients than healthy controls (38). UPF3B encodes a protein that participated in nonsense-mediated mRNA decay, and the mutation of UPF3B was associated with mental retardation (39). However, the role of UPF3b in cancer has not yet been reported and needs further study. López et al. used the machine-learning classification model to recognize that MRPL54 may be strongly connected to breast cancer (40). In our study, we found for the first time that UPF3b and MRPL54 were independent prognostic indicators in HCC. Liu et al. suggested ZC3H13 was downregulated in HCC, and patients with lower ZC3H13 expression had poorer overall survival, consistent with our findings (41). ZC3H13 also has been demonstrated to have prognostic value in other cancer types, such as lung adenocarcinoma, clear cell renal carcinoma, and anal squamous



cell carcinoma (42–44). ZC3H13 expression was higher in lung adenocarcinoma, and its expression pattern was the same as that in HCC. However, ZC3H13 had a lower expression level in clear cell renal carcinoma. Trilla et al. suggested that the genetic variant

of ZC3H13 was associated with poor disease-free survival. IFIT5 belongs to the interferon-induced tetratricopeptide repeat (IFIT) protein family (44). Some previous studies have indicated that IFIT5 was high-expressed and negatively correlated with the

TABLE 3 | Univariate and multivariate analyses of different clinical characteristics in TCGA and ICGC cohorts.

TCGA cohort	Univariate Cox analysis			Multivariate Cox regression		
	HR	95%CI	P-value	HR	95%CI	P-value
Age	1.005	0.990–1.020	0.50645	1.01	0.994–1.026	0.198323
Gender	0.807	0.542–1.203	0.293326	0.896	0.590–1.361	0.608359
Histological grade	1.111	0.854–1.446	0.43213	0.998	0.749–1.328	0.988103
TNM stage	1.804	1.456–2.234	<0.001	1.687	1.345–2.117	<0.001
Risk score	1.61	1.442–1.798	<0.001	1.591	1.411–1.794	<0.001
ICGC cohort	Univariate Cox analysis			Multivariate Cox regression		
	HR	95%CI	P-value	HR	95%CI	P-value
Age	1	0.969–1.032	0.987595	1	0.965–1.035	0.990104
Gender	0.454	0.240–0.860	0.015353	0.432	0.221–0.846	0.014352
Histological grade	NA	NA	NA	NA	NA	NA
TNM stage	2.12	1.448–3.104	<0.001	1.927	1.344–2.764	<0.001
Risk score	1.542	1.310–1.815	<0.001	1.421	1.185–1.705	<0.001

HR, hazard ratio; CI, confidence interval; NA, not available.

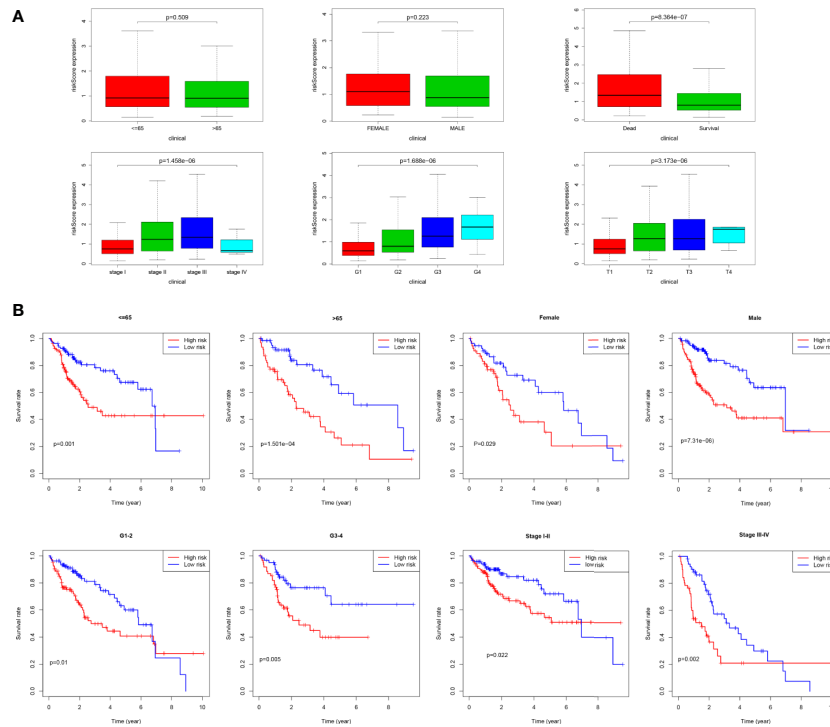


FIGURE 9 | Correlation of risk score and clinical characteristics. **(A)** Risk score distribution between different clinical characteristics; **(B)** Kaplan-Meier survival analysis of the signature stratified by clinical characteristics.

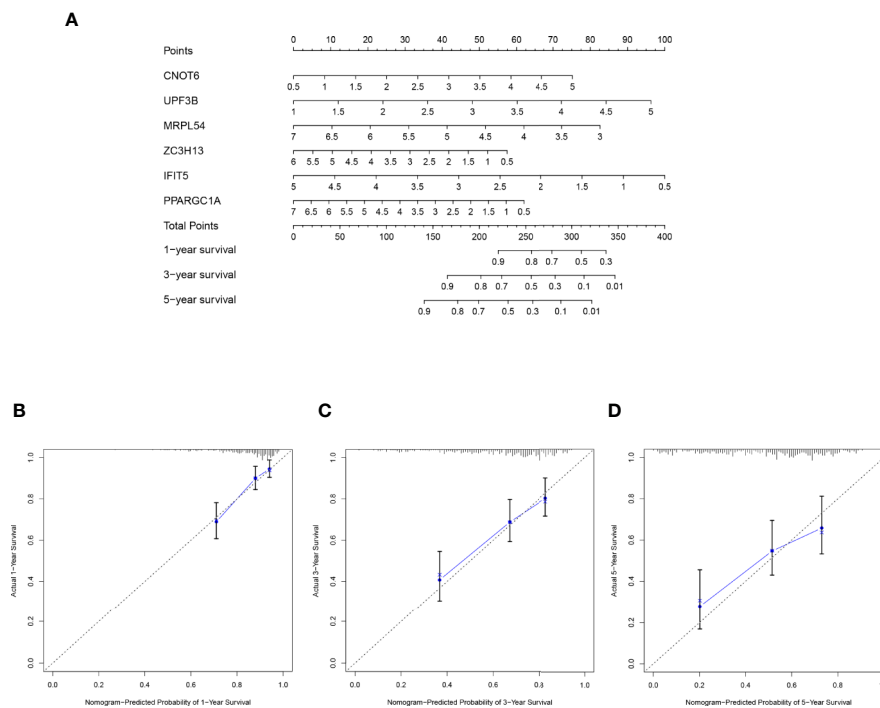
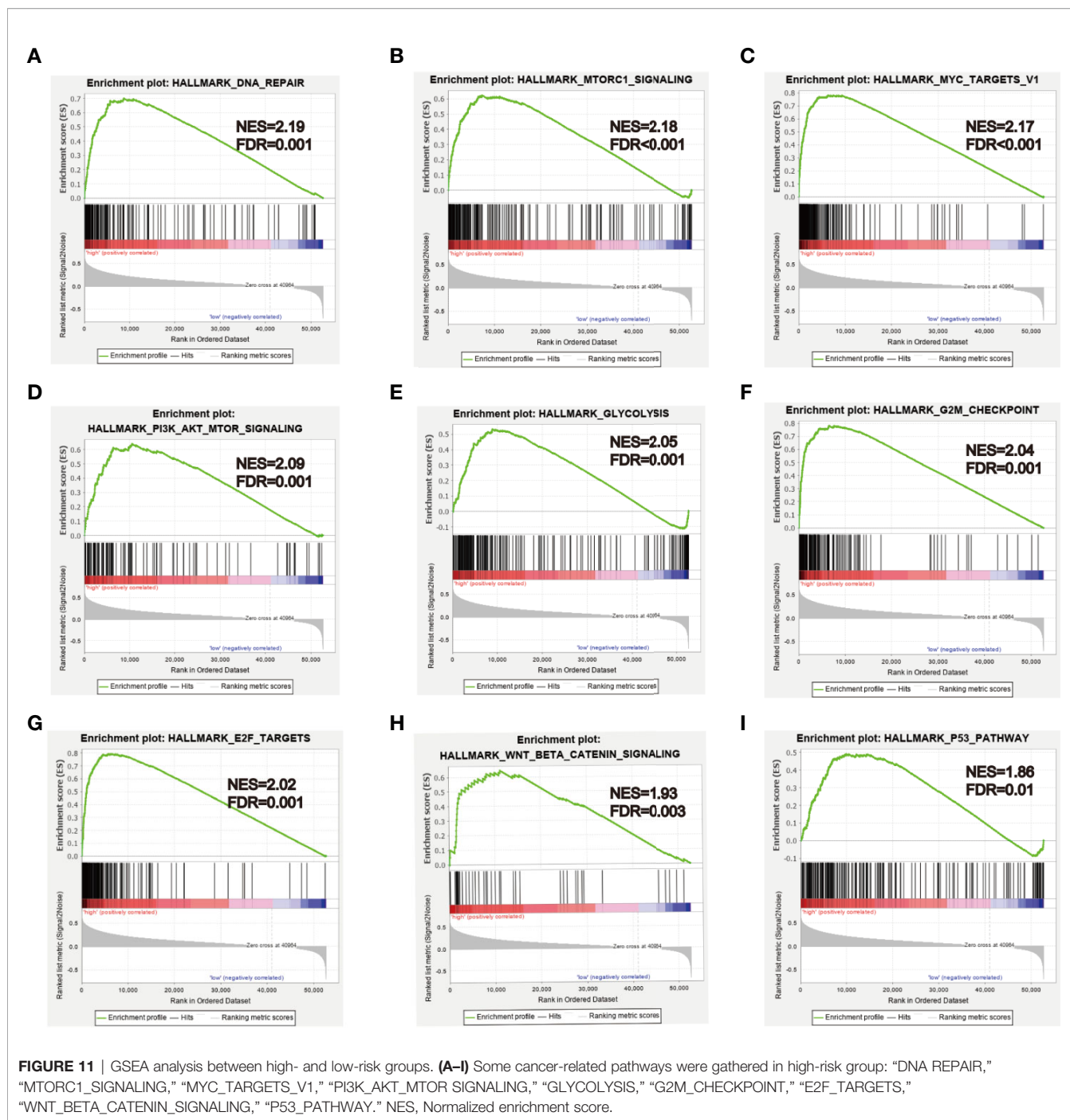


FIGURE 10 | A nomogram in TCGA HCC dataset. **(A)** The nomogram was built based on this six-RBP gene signature in the training cohort. **(B–D)** The calibration plots showed good predictive performance for OS at 1-, 3-, 5-year.



prognosis in renal cell carcinoma, bladder cancer, and prostate cancer patients. IFIT5 may function as an oncogene, promote cancer invasion, metastasis, and progression by inducing epithelial–mesenchymal transition (EMT) via modulating turnover of tumor suppressive microRNAs, including miR-363, miR-99a, and miR-128 (45–48). PPARGC1A, also known as PGC-1 α , functions as a master regulator of mitochondrial biogenesis and oxidative phosphorylation and plays a pivotal role in cancer cell metabolism and metastasis (49, 50). Some published studies have demonstrated that PPARGC1A

was upregulated in lung cancer and invasive breast cancer, and facilitated cancer metastasis and invasion. Moreover, PPARGC1A high expression was correlated to poor prognosis in patients with lung cancer and breast cancer (49, 51). However, the opposite results have been observed in some other studies, and PPARGC1A has been suggested as a tumor suppressor that suppresses prostate cancer and melanoma cell proliferation, migration, and metastasis (52, 53). In addition, Zhang et al. observed that PPARGC1A rs2970847 C>T polymorphisms associated with HCC risk (54). Given the importance of the six RBP genes

in kinds of cancer types, these genes might be potential prognostic biomarkers for patients with HCC, but detailed molecular mechanism during hepatocarcinogenesis needs further in-depth exploration.

GSEA analysis showed that some cancer-related pathways were enriched in high-risk HCC patients, such as Wnt/ β -catenin signaling, P53 pathway, PI3K-AKT-mTOR signaling, and MYC signaling. These molecular pathways have been confirmed to be implicated in HCC carcinogenesis. Autophagy can activate Wnt/ β -catenin signaling and promote HCC cells metastasis and glycolysis (55). Alpha-fetoprotein (AFP) inhibited autophagy in HCC cells by activating of PI3K/Akt/mTOR signaling, thereby promoting proliferation, migration, and invasion (56). The c-Myc was a transcription factor that plays an important role in hepatocarcinogenesis, NELFE promoted HCC progression via enhancing MYC signaling (20). P53 as a tumor suppressor protein, inhibiting the p53 pathway, may promote HCC cells proliferation and inhibit apoptosis (57).

Thanks to the great progress in microarray and next-generation sequencing technologies, a number of multigene prognostic models have been developed to predict survival for HCC patients, such as Wang et al. developed an immune-related prognostic model in HCC (58), and Li et al. developed a CIMP-associated prognostic model for HCC (59). However, RBPs-associated prognostic model for HCC has not been reported yet; this is the first study about a prognostic model in HCC patients constructed using multiple RBP genes, to our knowledge. According to our risk score signature, survival analysis displayed significant difference of OS between high- and low-risk subgroups, and usually low-risk patients had better survival than patients with high risk score. The ROC curves suggested that our prognostic model had a good accuracy, and the AUC values of 1-, 3-year were greater than 0.75 both in training and testing set. In addition, whether in the training set or testing set, TNM stage and risk score were independent prognostic indicators in HCC. Although our model has good prediction performance, there are still some limitations that need to be discussed. First, the six-RBP gene signature was built based on TCGA HCC dataset and was only validated in the ICGC HCC dataset, which has not been validated in our own clinical HCC cases cohort. Second, most HCC patients in TCGA database were Caucasian, and it is not clear whether it has the same predictive effect in non-Caucasian races. Finally, our study was retrospective and needs further validation by a larger prospective study.

In conclusion, we identified differently expressed RBP genes and constructed a promising six-RBP gene prognostic signature

to predict clinical outcomes for HCC patients. This risk score signature was proven to have good predictive ability and function as an independent prognostic indicator for HCC patients, contributing to guided clinical decision making and personalized treatment. Moreover, this study would further help us understand the prognostic value and biological function of RBPs in HCC.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: The Cancer Genome Atlas (<https://portal.gdc.cancer.gov/>).

AUTHOR CONTRIBUTIONS

YH and SC designed the study. YH, WQ, and LL carried out data extraction and statistical analysis. YH, XY, and QL drafted and modified the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by grants from the National Natural Science Foundation of China (2018, 81773360).

ACKNOWLEDGMENTS

Authors would like to thank TCGA (<https://portal.gdc.cancer.gov/>) and ICGC (<https://dcc.icgc.org/releases/current/Projects/LIRI-JP>) databases for providing the data of HCC.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2020.580513/full#supplementary-material>

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) 68:394–424. doi: 10.3322/caac.21492
2. Akinyemiju T, Abera S, Ahmed M, Alam N, Alemayohu MA, Allen C, et al. The Burden of Primary Liver Cancer and Underlying Etiologies From 1990 to 2015 at the Global, Regional, and National Level: Results From the Global Burden of Disease Study 2015. *JAMA Oncol* (2017) 3:1683–91. doi: 10.1001/jamaoncol.2017.3055
3. 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) 16:589–604. doi: 10.1038/s41575-019-0186-y
4. EASL Clinical Practice Guidelines. Management of hepatocellular carcinoma. *J Hepatol* (2018) 69:182–236. doi: 10.1016/j.jhep.2018.03.019
 5. Forner A, Reig M, Bruix J. Hepatocellular carcinoma. *Lancet* (2018) 391:1301–14. doi: 10.1016/S0140-6736(18)30010-2
 6. Keene JD. RNA regulons: coordination of post-transcriptional events. *Nat Rev Genet* (2007) 8:533–43. doi: 10.1038/nrg2111
 7. Gerstberger S, Hafner M, Tuschl T. A census of human RNA-binding proteins. *Nat Rev Genet* (2014) 15:829–45. doi: 10.1038/nrg3813
 8. Mohibi S, Chen X, Zhang J. Cancer the RBP^eutics-RNA-binding proteins as therapeutic targets for cancer. *Pharmacol Ther* (2019) 203:107390. doi: 10.1016/j.pharmthera.2019.07.001
 9. Glisovic T, Bachorik JL, Yong J, Dreyfuss G. RNA-binding proteins and post-transcriptional gene regulation. *FEBS Lett* (2008) 582:1977–86. doi: 10.1016/j.febslet.2008.03.004
 10. Hodson DJ, Screen M, Turner M. RNA-binding proteins in hematopoiesis and hematological malignancy. *Blood* (2019) 133:2365–73. doi: 10.1182/blood-2018-10-839985
 11. de Bruin RG, Rabelink TJ, van Zonneveld AJ, van der Veer EP. Emerging roles for RNA-binding proteins as effectors and regulators of cardiovascular disease. *Eur Heart J* (2017) 38:1380–8. doi: 10.1093/eurheartj/ehw567
 12. Nussbacher JK, Batra R, Lagier-Tourenne C, Yeo GW. RNA-binding proteins in neurodegeneration: Seq and you shall receive. *Trends Neurosci* (2015) 38:226–36. doi: 10.1016/j.tins.2015.02.003
 13. Pereira B, Billaud M, Almeida R. RNA-Binding Proteins in Cancer: Old Players and New Actors. *Trends Cancer* (2017) 3:506–28. doi: 10.1016/j.trecan.2017.05.003
 14. Wang ZL, Li B, Luo YX, Lin Q, Liu SR, Zhang XQ, et al. Comprehensive Genomic Characterization of RNA-Binding Proteins across Human Cancers. *Cell Rep* (2018) 22:286–98. doi: 10.1016/j.celrep.2017.12.035
 15. Lukong KE, Chang KW, Khandjian EW, Richard S. RNA-binding proteins in human genetic disease. *Trends Genet* (2008) 24:416–25. doi: 10.1016/j.tig.2008.05.004
 16. Masuda K, Kuwano Y. Diverse roles of RNA-binding proteins in cancer traits and their implications in gastrointestinal cancers. *Wiley Interdiscip Rev RNA* (2019) 10:e1520. doi: 10.1002/wrna.1520
 17. García-Cárdenas JM, Guerrero S, López-Cortés A, Armendáriz-Castillo I, Guevara-Ramírez P, Pérez-Villa A, et al. Post-transcriptional Regulation of Colorectal Cancer: A Focus on RNA-Binding Proteins. *Front Mol Biosci* (2019) 6:65. doi: 10.3389/fmolb.2019.00065
 18. Han L, Zan Y, Huang C, Zhang S. NELFE promoted pancreatic cancer metastasis and the epithelial-to-mesenchymal transition by decreasing the stabilization of NDRG2 mRNA. *Int J Oncol* (2019) 55:1313–23. doi: 10.3892/ijo.2019.4890
 19. Zhao L, Cao J, Hu K, Wang P, Li G, He X, et al. RNA-binding protein RPS3 contributes to hepatocarcinogenesis by post-transcriptionally up-regulating SIRT1. *Nucleic Acids Res* (2019) 47:2011–28. doi: 10.1093/nar/gky1209
 20. Dang H, Takai A, Forgues M, Pomyen Y, Mou H, Xue W, et al. Oncogenic Activation of the RNA Binding Protein NELFE and MYC Signaling in Hepatocellular Carcinoma. *Cancer Cell* (2017) 32:101–114.e8. doi: 10.1016/j.ccell.2017.06.002
 21. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* (2012) 16:284–7. doi: 10.1089/omi.2011.0118
 22. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res* (2019) 47:D607–13. doi: 10.1093/nar/gky1131
 23. Bader GD, Hogue CW. An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinf* (2003) 4:2. doi: 10.1186/1471-2105-4-2
 24. 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:pl1. doi: 10.1126/scisignal.2004088
 25. Uhlen M, Zhang C, Lee S, Sjöstedt E, Fagerberg L, Bidkhori G, et al. A pathology atlas of the human cancer transcriptome. *Science* (2017) 357(6352): eaan2507. doi: 10.1126/science.aan2507
 26. Chen HY, Yu SL, Chen CH, Chang GC, Chen CY, Yuan A, et al. A five-gene signature and clinical outcome in non-small-cell lung cancer. *N Engl J Med* (2007) 356:11–20. doi: 10.1056/NEJMoa060096
 27. Wu M, Li X, Zhang T, Liu Z, Zhao Y. Identification of a Nine-Gene Signature and Establishment of a Prognostic Nomogram Predicting Overall Survival of Pancreatic Cancer. *Front Oncol* (2019) 9:996. doi: 10.3389/fonc.2019.00996
 28. Heagerty PJ, Lumley T, Pepe MS. Time-dependent ROC curves for censored survival data and a diagnostic marker. *Biometrics* (2000) 56:337–44. doi: 10.1111/j.0006-341x.2000.00337.x
 29. Shiraishi Y, Fujimoto A, Furuta M, Tanaka H, Chiba K, Boroevich KA, et al. Integrated analysis of whole genome and transcriptome sequencing reveals diverse transcriptomic aberrations driven by somatic genomic changes in liver cancers. *PLoS One* (2014) 9:e114263. doi: 10.1371/journal.pone.0114263
 30. 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 USA* (2005) 102:15545–50. doi: 10.1073/pnas.0506580102
 31. Van Treecb B, Parker R. Emerging Roles for Intermolecular RNA-RNA Interactions in RNP Assemblies. *Cell* (2018) 174:791–802. doi: 10.1016/j.cell.2018.07.023
 32. Tauber D, Tauber G, Parker R. Mechanisms and Regulation of RNA Condensation in RNP Granule Formation. *Trends Biochem Sci* (2020) 45(9):764–78. doi: 10.1016/j.tibs.2020.05.002
 33. Standart N, Weil D. P-Bodies: Cytosolic Droplets for Coordinated mRNA Storage. *Trends Genet* (2018) 34:612–26. doi: 10.1016/j.tig.2018.05.005
 34. Luo Y, Na Z, Slavoff SA. P-Bodies: Composition, Properties, and Functions. *Biochemistry-US* (2018) 57:2424–31. doi: 10.1021/acs.biochem.7b01162
 35. Mittal S, Aslam A, Doidge R, Medica R, Winkler GS. The Ccr4a (CNOT6) and Ccr4b (CNOT6L) deadenylase subunits of the human Ccr4-Not complex contribute to the prevention of cell death and senescence. *Mol Biol Cell* (2011) 22:748–58. doi: 10.1091/mbc.E10-11-0898
 36. Zhou F, Wang Y, Liu H, Ready N, Han Y, Hung RJ, et al. Susceptibility loci of CNOT6 in the general mRNA degradation pathway and lung cancer risk-A re-analysis of eight GWASs. *Mol Carcinog* (2017) 56:1227–38. doi: 10.1002/mc.22585
 37. Maragozidis P, Papanastasi E, Scutelnic D, Totomi A, Kokkori I, Zarogiannis SG, et al. Poly(A)-specific ribonuclease and Nocturnin in squamous cell lung cancer: prognostic value and impact on gene expression. *Mol Cancer* (2015) 14:187. doi: 10.1186/s12943-015-0457-3
 38. Maragozidis P, Karangeli M, Labrou M, Dimoulou G, Papaspyrou K, Salataj E, et al. Alterations of deadenylase expression in acute leukemias: evidence for poly(a)-specific ribonuclease as a potential biomarker. *Acta Haematol* (2012) 128:39–46. doi: 10.1159/000337418
 39. Laumonnier F, Shoubridge C, Antar C, Nguyen LS, Van Esch H, Kleefstra T, et al. Mutations of the UPF3B gene, which encodes a protein widely expressed in neurons, are associated with nonspecific mental retardation with or without autism. *Mol Psychiatry* (2010) 15:767–76. doi: 10.1038/mp.2009.14
 40. López-Cortés A, Cabrera-Andrade A, Vázquez-Naya JM, Pazos A, González-Díaz H, Paz-Y-Miño C, et al. Prediction of breast cancer proteins involved in immunotherapy, metastasis, and RNA-binding using molecular descriptors and artificial neural networks. *Sci Rep* (2020) 10:8515. doi: 10.1038/s41598-020-65584-y
 41. Liu GM, Zeng HD, Zhang CY, Xu JW. Identification of METTL3 as an Adverse Prognostic Biomarker in Hepatocellular Carcinoma. *Dig Dis Sci* (2020) 10.1007/s10620-020-06260-z. doi: 10.1007/s10620-020-06260-z
 42. Zhuang Z, Chen L, Mao Y, Zheng Q, Li H, Huang Y, et al. Diagnostic, progressive and prognostic performance of m(6A) methylation RNA regulators in lung adenocarcinoma. *Int J Biol Sci* (2020) 16:1785–97. doi: 10.7150/ijbs.39046
 43. Chen J, Yu K, Zhong G, Shen W. Identification of a m(6A) RNA methylation regulators-based signature for predicting the prognosis of clear cell renal carcinoma. *Cancer Cell Int* (2020) 20:157. doi: 10.1186/s12935-020-01238-3
 44. Trilla-Fuertes L, Ghanem I, Maurel J, G-Pastríán L, Mendiola M, Peña C, et al. Comprehensive Characterization of the Mutational Landscape in Localized Anal Squamous Cell Carcinoma. *Transl Oncol* (2020) 13:100778. doi: 10.1016/j.tranon.2020.100778
 45. Katibah GE, Lee HJ, Huizar JP, Vogan JM, Alber T, Collins K. tRNA binding, structure, and localization of the human interferon-induced protein IFIT5. *Mol Cell* (2013) 49:743–50. doi: 10.1016/j.molcel.2012.12.015

46. Lo UG, Bao J, Cen J, Yeh HC, Luo J, Tan W, et al. Interferon-induced IFIT5 promotes epithelial-to-mesenchymal transition leading to renal cancer invasion. *Am J Clin Exp Urol* (2019) 7:31–45.
47. Huang J, Lo UG, Wu S, Wang B, Pong RC, Lai CH, et al. The roles and mechanism of IFIT5 in bladder cancer epithelial-mesenchymal transition and progression. *Cell Death Dis* (2019) 10:437. doi: 10.1038/s41419-019-1669-z
48. Lo UG, Pong RC, Yang D, Gandee L, Hernandez E, Dang A, et al. IFN γ -Induced IFIT5 Promotes Epithelial-to-Mesenchymal Transition in Prostate Cancer via miRNA Processing. *Cancer Res* (2019) 79:1098–112. doi: 10.1158/0008-5472.CAN-18-2207
49. LeBleu VS, O'Connell JT, Gonzalez HK, Wikman H, Pantel K, Haigis MC, et al. PGC-1 α mediates mitochondrial biogenesis and oxidative phosphorylation in cancer cells to promote metastasis. *Nat Cell Biol* (2014) 16:992–1003, 1–15. doi: 10.1038/ncb3039
50. Kaminski L, Torino S, Dufies M, Djabari Z, Haider R, Roustan FR, et al. PGC1 α Inhibits Polyamine Synthesis to Suppress Prostate Cancer Aggressiveness. *Cancer Res* (2019) 79:3268–80. doi: 10.1158/0008-5472.CAN-18-2043
51. Li JD, Feng QC, Qi Y, Cui G, Zhao S. PPARGC1A is upregulated and facilitates lung cancer metastasis. *Exp Cell Res* (2017) 359:356–60. doi: 10.1016/j.yexcr.2017.08.017
52. Torrano V, Valcarcel-Jimenez L, Cortazar AR, Liu X, Urošević J, Castillo-Martin M, et al. The metabolic co-regulator PGC1 α suppresses prostate cancer metastasis. *Nat Cell Biol* (2016) 18:645–56. doi: 10.1038/ncb3357
53. Luo C, Lim JH, Lee Y, Granter SR, Thomas A, Vazquez F, et al. A PGC1 α -mediated transcriptional axis suppresses melanoma metastasis. *Nature* (2016) 537:422–6. doi: 10.1038/nature19347
54. Zhang S, Jiang J, Chen Z, Wang Y, Tang W, Chen Y, et al. Relationship of PPARG, PPARGC1A, and PPARGC1B polymorphisms with susceptibility to hepatocellular carcinoma in an eastern Chinese Han population. *Oncotargets Ther* (2018) 11:4651–60. doi: 10.2147/OTT.S168274
55. Fan Q, Yang L, Zhang X, Ma Y, Li Y, Dong L, et al. Autophagy promotes metastasis and glycolysis by upregulating MCT1 expression and Wnt/ β -catenin signaling pathway activation in hepatocellular carcinoma cells. *J Exp Clin Cancer Res* (2018) 37:9. doi: 10.1186/s13046-018-0673-y
56. Wang S, Zhu M, Wang Q, Hou Y, Li L, Weng H, et al. Alpha-fetoprotein inhibits autophagy to promote malignant behaviour in hepatocellular carcinoma cells by activating PI3K/AKT/mTOR signalling. *Cell Death Dis* (2018) 9:1027. doi: 10.1038/s41419-018-1036-5
57. Kong Y, Zhang L, Huang Y, He T, Zhang L, Zhao X, et al. Pseudogene PDIA3P1 promotes cell proliferation, migration and invasion, and suppresses apoptosis in hepatocellular carcinoma by regulating the p53 pathway. *Cancer Lett* (2017) 407:76–83. doi: 10.1016/j.canlet.2017.07.031
58. Wang Z, Zhu J, Liu Y, Liu C, Wang W, Chen F, et al. Development and validation of a novel immune-related prognostic model in hepatocellular carcinoma. *J Transl Med* (2020) 18:67. doi: 10.1186/s12967-020-02255-6
59. Li G, Xu W, Zhang L, Liu T, Jin G, Song J, et al. Development and validation of a CIMP-associated prognostic model for hepatocellular carcinoma. *Ebiomedicine* (2019) 47:128–41. doi: 10.1016/j.ebiom.2019.08.064

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Huang, Chen, Qin, Wang, Li, Li and Yuan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.