Development and Validation of a RNA Binding Protein-Associated Prognostic Model for Hepatocellular Carcinoma

Ming Wang, MD[1], Feng Jiang, MD[2,3] , Ke Wei, MD[4], Jimei Wang, PhD[3] ,
Guoping Zhou, PhD[2], Chuyan Wu, MD[5] , and Guoyong Yin, MD[6]

Abstract

Background: Dysregulation of RNA binding proteins (RBPs) has been identified in multiple malignant tumors correlated with tumor progression and occurrence. However, the function of RBPs is not well understood in hepatocellular carcinoma (HCC). **Methods:** The RNA sequence data of HCC was extracted out of the Cancer Genome Atlas (TCGA) database and different RBPs were calculated between regular and cancerous tissue. The study explored the expression and predictive value of the RBPs systemically with a series of bioinformatic analyzes. **Results:** A total of 330 RBPs, including 208 up-regulated and 122 down-regulated RBPs, were classified differently. Four RBPs (MRPL54, EZH2, PPARGC1A, EIF2AK4) were defined as the forecast related hub gene and used to construct a model for prediction. Further study showed that the high-risk subgroup is poor survived (OS) compared to the model-based low-risk subgroup. The area of the prognostic model under the time-dependent receiver operator characteristic (ROC) curve is 0.814 in TCGA training group and 0.729 in validation group, indicating a strong prognostic model. We also created a predictive nomogram and a web-based calculator (https://dxyjiang.shinyapps.io/RBPpredict/) based on the 4 RBPs and internal validation in the TCGA cohort, which displayed a beneficial predictive ability for HCC. **Conclusions:** Our results provide new insights into HCC pathogenesis. The 4-RBP gene signature showed a reliable HCC prediction ability with possible applications in therapeutic decision making and personalized therapy.

[1] Department of Plastic and Burn Surgery, The First Affiliated Hospital of Nanjing Medical University, Nanjing, China
[2] Department of Pediatrics, The First Affiliated Hospital of Nanjing Medical University, Nanjing, China
[3] Department of Neonatology, Obstetrics and Gynecology Hospital of Fudan University, Shanghai, China
[4] Medical Service Section, The First Affiliated Hospital of Nanjing Medical University, Nanjing, China
[5] Department of Rehabilitation Medicine, The First Affiliated Hospital of Nanjing Medical University, Nanjing, China
[6] Department of Orthopedics, The First Affiliated Hospital of Nanjing Medical University, Nanjing, China

Corresponding Authors:
Guoyong Yin, MD, Department of Orthopedics, The First Affiliated Hospital of Nanjing Medical University, Nanjing 210029, China.
Chuyan Wu, MD, Department of Rehabilitation Medicine, The First Affiliated Hospital of Nanjing Medical University, Nanjing 210029, China.
Emails: guoyong_yin@sina.com; chuyan_wu@hotmail.com

Introduction

Hepatocellular carcinoma is one of the world's most common malignant tumors.[1] It has a high degree of malignancy, a powerful invasion and metastasis, and weak prognosis.[2] While the number of tumor markers used for the diagnosis and forecast of associated cancer has increased, no one has the high recognition, sensitivity, and specific characteristics necessary for the evaluation of HCC condition, effectiveness and prognosis. Biomarkers with diagnostic and prognostic accuracy are urgently needed.[3,4]

RNA binding proteins are proteins that associate and are omnipresent in cells with various types of RNA.[5,6] The high-throughput scanning of human cells, containing 7.5 percent of all protein coding genes, detected a minimum of 1542 RBPs.[7,8] These RBP's influence cell proliferation and cell metabolism post-transcription activities and are therefore involved in many biological processes, including RNA splicing, mRNA stabilization, cytoplasm exports, localization, and protein translation.[9,10] Provided that RBPs serve multiple essential functions in post-transcriptional activities, it is shocking that variations in RBPs are strongly associated with the production and growth of numerous human diseases.[11,12] Nevertheless, RBPs remain largely unexplored throughout the origin and development of cancer.

In recent years, genome-wide studies have shown that certain RBPs are correlated with dysregulated tumor expression in adjacent tissue normal and are correlated with patient prognosis.[13] Dysregulation of RBPs is well known to be primarily the product of genomic modifications, micro-mediated regulation, epigenetic pathways and post-translation changes in cancer cells. Previous research has correlated cancer drivers with RBP dysregulation.[14,15] For, e.g. G3BP1, which has been shown that the knockdown of G3BP1 decreases the assembly of stress granules and thus provides chemical resistance and survival advantage in tumor cells.[16-18] In prokaryotic and eukaryotic cells, YB1, an RNA binding protein, plays a critical role in the control of cell growth, differentiation, and stress reaction in transcription and interpreting cells.[19,20] In addition, LARP4B is an RBP containing lanthanum and adjacent motifs for RNA recognition that enable it to participate in post-transcriptional RNA regulation and participate in the progression of several cancers.[21-24] Taken together, these findings demonstrate that RBPs are strongly connected to the development and incidence of human tumors. However, only a limited fraction of RBPs have been intensively researched and have to date been linked to cancer. Therefore, in order to analyze the possible molecular role and the therapeutic importance of RBPs in HCC, we gathered all applicable HCC data in the TCGA base and conducted the current systematic study. We have found multiple RBPs associated with HCC that provide new insights into the pathogenesis of the disease. Some of them may act as possible diagnostic and prognostic biomarkers.

Materials and Methods

Data Pre-Processing and Identification of RBPs Presented Differentially

Data on RNA sequencing of 374 HCC samples and 50 regular hepatic tissue samples have been collected from TCGA. The raw data were pre-processed with the package DESeq2. The RBPs expressed differently were defined based on a false discovery rate <0.05 and |log (fold change) >1|. All RBPs represented differentially were of an average magnitude of more than 1.

Functional Enrichment Study of GO and KEGG

GO enrichment systemically analyzed the biological processes of the various expressions of these RBPs and used 3 terms: molecular function, biological process, and cellular component. The KEGG database was used to extract possible biological pathways from RBPs presented differently. Both analyzes for GO and KEGG pathways were conducted with "ggplot2," "enrichplot" and "clusterProfiler" packages.

Network Development and Module Screening of Protein-Protein Interaction

STRING was used to detect protein-protein interactions (PPIs) between all differential expressed RBPs and their network was developed using Cytoscape 3.7.2 software. The following main modules were scanned from the PPI network by using the MCODE (Molecular Complex Detection) plug-in on Cytoscape with scores >7 and node counts >5. To choose the hub genes, the cytoHubba plug-in was used. $P < 0.05$ was considered to show a significant difference.

## Validation of RBP Expression and Efficacy Assessment

The Human Protein Atlas (HPA) database was used to detect the translation expression of 4 hub genes. We used R software to constructed the Receiver operating characteristic (ROC) curves to calculate the ability to discriminate between normal and tumor tissue.

Copy-Number Alterations and Hub RBP Mutation Analysis

The alteration of the copy number and mutation data from all PPI network hub RBPs were defined by segmentation analyzes and the GISTIC algorithm in cBioPortal. We then conducted a co-expression study of all hub RBPs. We then developed a network that included all hub genes and the 20 most frequently altered neighboring genes.

Selection of RBP Related to Prognosis

The differential expressed RBPs are evaluated using the survival package in R for a univariate Cox regression study. Subsequently, RBPs were analyzed through multiple step by step,
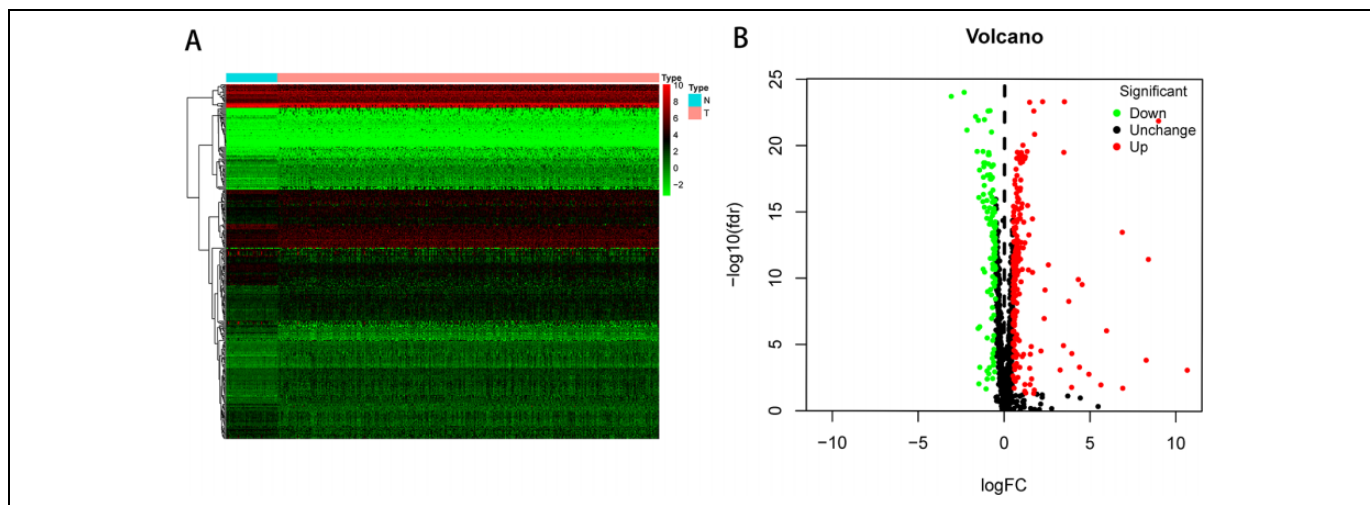
Figure 1. The RBPs expressed differently in HCC. (A) Heat map; (B) volcano plot.

Cox regression. We have used a Kaplan-Meier method to determine the predictive importance of each RBP candidate. The *P*-value with RBPs less than 0.05 is known to be valid prognostic.

Prognostic Model Construction and Evaluation

The multivariate Cox regression test for the coefficients (β_i) of the hub genes was performed. This model was developed as: Risk score $= \Sigma(\beta_i \times \text{Exp}i)(i = n)$. Patients were randomly divided into the population of training and validation group. Each category of patients was classified into subgroups of low-risk and high-risk depending on the median risk score. Then, we performed the ROC curve to test the HCC signature predictive value. A Kaplan-Meier study was performed in order to assess the mortality difference between 2 categories, where the HCC patients were divided into high- and low-groups with the median prediction value as the threshold. In addition, we have drawn a nomogram plot to estimate the OS probability using the rms R package. Validation population was used to validate the prrognostic model's predictive ability. We used the "shiny" and "DynNom" packages to generate a web-based survival rate calculator, which dynamically predicted overall survival rates (https://www.shinyapps.io/).

Results

Identification of RBPs Expressed Differently in HCC Patients

In this research, we performed a systematic analysis using many specialized techniques of key RBP functions and predictive values in HCC. The design of the analysis was shown in Figure S1. TCGA composite 374 tumor samples and 50 regular liver tissue samples are extracted from HCC database. The R software packages were used to manage the data and to Figure out the various RBPs. An analysis

included a total of 1495 RBPs and 330 RBPs met the screening standards ($P < 0.05$, |log (fold change) >1|), composed of 208 up-regulated RBPs—and 122 downregulated RBPs. Figure 1 reveals the expression distribution of these different RBPs.

Analysis of the Various Expressed RBPs by GO and KEGG Pathway Enrichment

We divided these RBPs into 2 categories to analyze the function and mechanisms of defined RBPs: up-or down-regulated expression. We then submitted these RBPs to the WebGestalt online platform for functional enrichment evaluation. As shown in Table S1 and Table S2. The results showed that downregulated RBPs expressed differently were significantly enriched in the biological process related to regulation of translation, regulation of cellular amide metabolic process, RNA phosphodiester bond hydrolysis, negative regulation of translation, and negative regulation of cellular amide metabolic process (Figure 2B). The up-regulated RBPs were significantly enriched in ncRNA processing, RNA splicing, RNA splicing via transesterification reactions with bulged adenosine as nucleophile, mRNA splicing via spliceosome and RNA splicing via transesterification reactions (Figure 2A). In Cellular component (CC) study, the decreased expressed RBPs were notably enriched in cytoplasmic ribonucleoprotein granule, ribonucleoprotein granule, *P*-body, CCR4-NOT complex and ribosome (Figure 2B), while the up-regulated RBPs expressed differently were significantly enriched in spliceosomal complex, cytoplasmic ribonucleoprotein granule, ribonucleoprotein granule, small unclear ribonucleoprotein complex and Sm-like protein family complex (Figure 2A). By way of molecular function (MF) study, the decreased RBPs expressed differently were enriched in catalytic activity acting on RNA, single-stranded RNA
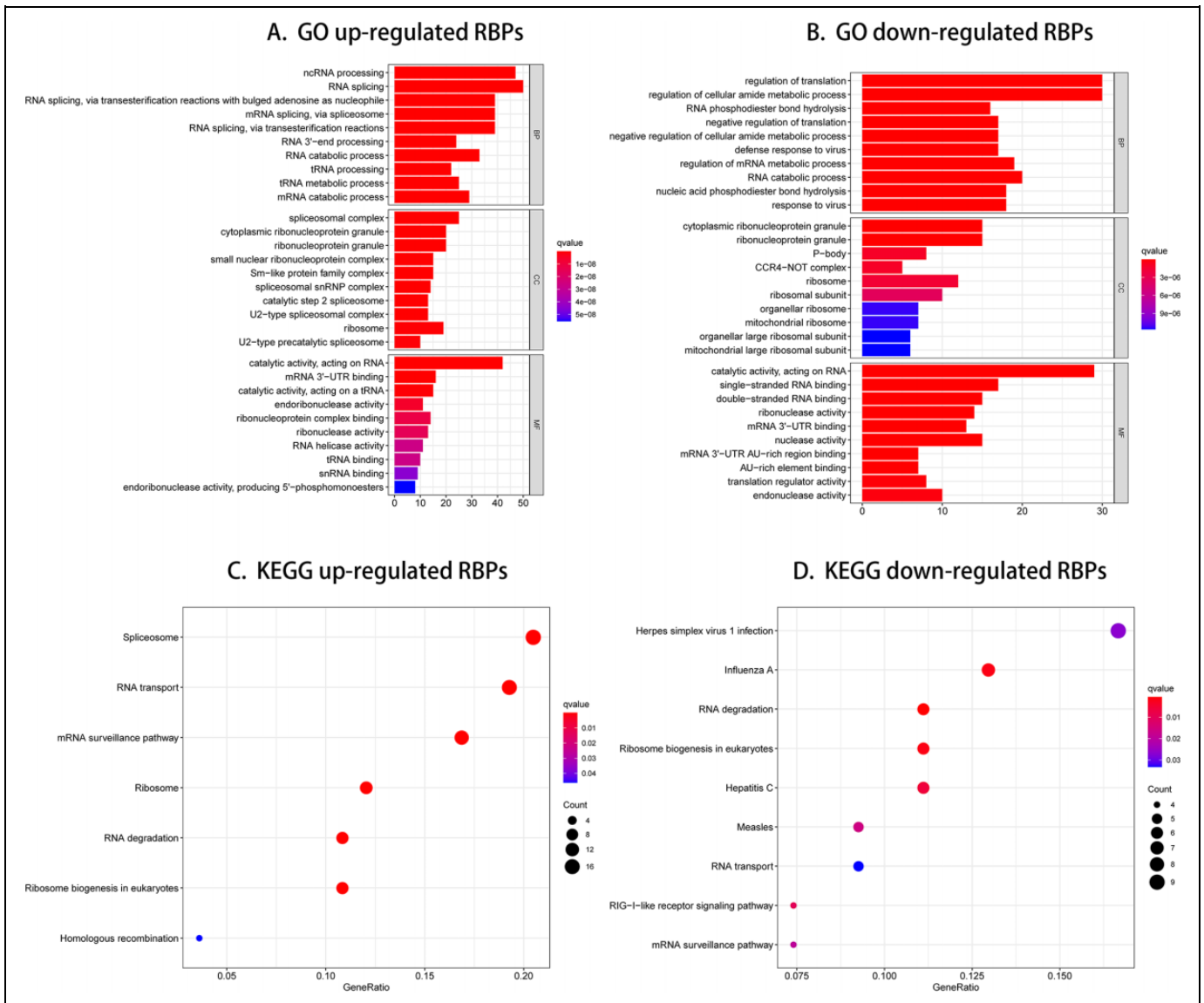
**Figure 2.** GO enrichment and KEGG pathway of aberrantly expressed RBPs.

binding, double-stranded RNA binding, ribonuclease activity and mRNA 3'-UTR binding (Figure 2B), while upregulated RBPs expressed differently were mainly enriched in catalytic activity acting on RNA, mRNA 3'-UTR binding, catalytic activity acting on a tRNA, endoribonuclease activity and ribonucleoprotein complex binding (Figure 2A). In addition, we found that KEGG downregulated RBPs expressed differently were mainly enriched in herpes simplex virus 1 infection, influenza A, RNA degradation, ribosome biogenesis in eukaryotes and hepatitis C (Figure 2D), while up-regulated RBPs were significantly enriched for spliceosome, RNA transport, mRNA surveillance pathway, ribosome and RNA degradation (Figure 2C). Overlapped genes involved in GO enrichment and KEGG pathways were shown in Figure S2.

Network Development of Protein-Protein Interaction (PPI) and Core Modules Collecting

In order to further research the functions of RNA binding proteins in HCC, we have built a PPI network utilizing Cytoscape software with 311 nodes and 2942 edges based on STRING's database (Figure 3A). The MODE tool analyzed the co-expression network to recognize potential core modules and the first relevant modules that have been acquired consisting of 53 nodes and 709 edges (Figure 3B). The RBPs in the key module 1 were greatly enriched in RNA splicing via transesterification reactions with bulged adenosine as nucleophile, mRNA splicing via spliceosome, RNA splicing via transesterification reactions, ncRNA processing and RNA splicing.
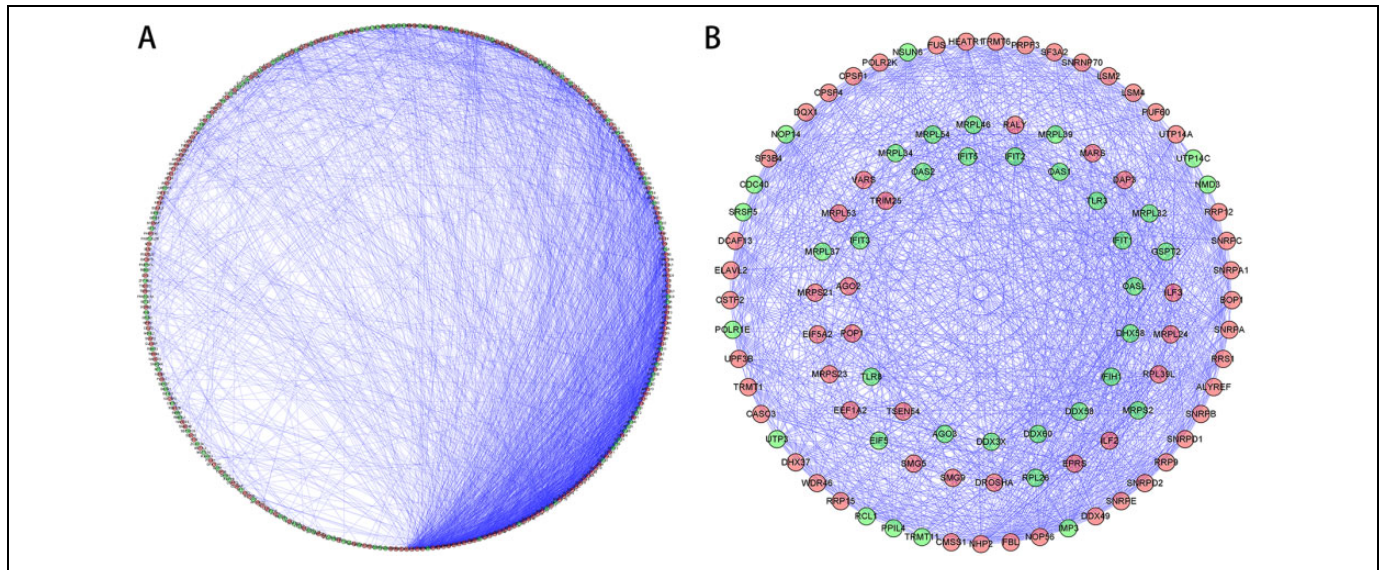
Figure 3. PPI network and modules analysis. (A) PPI network of RBPs expressed differentially; (B) the essential module in the PPI network. Green circles: down-regulation with more than 2 fold change; red circles: up-regulation of more than 2 fold change.
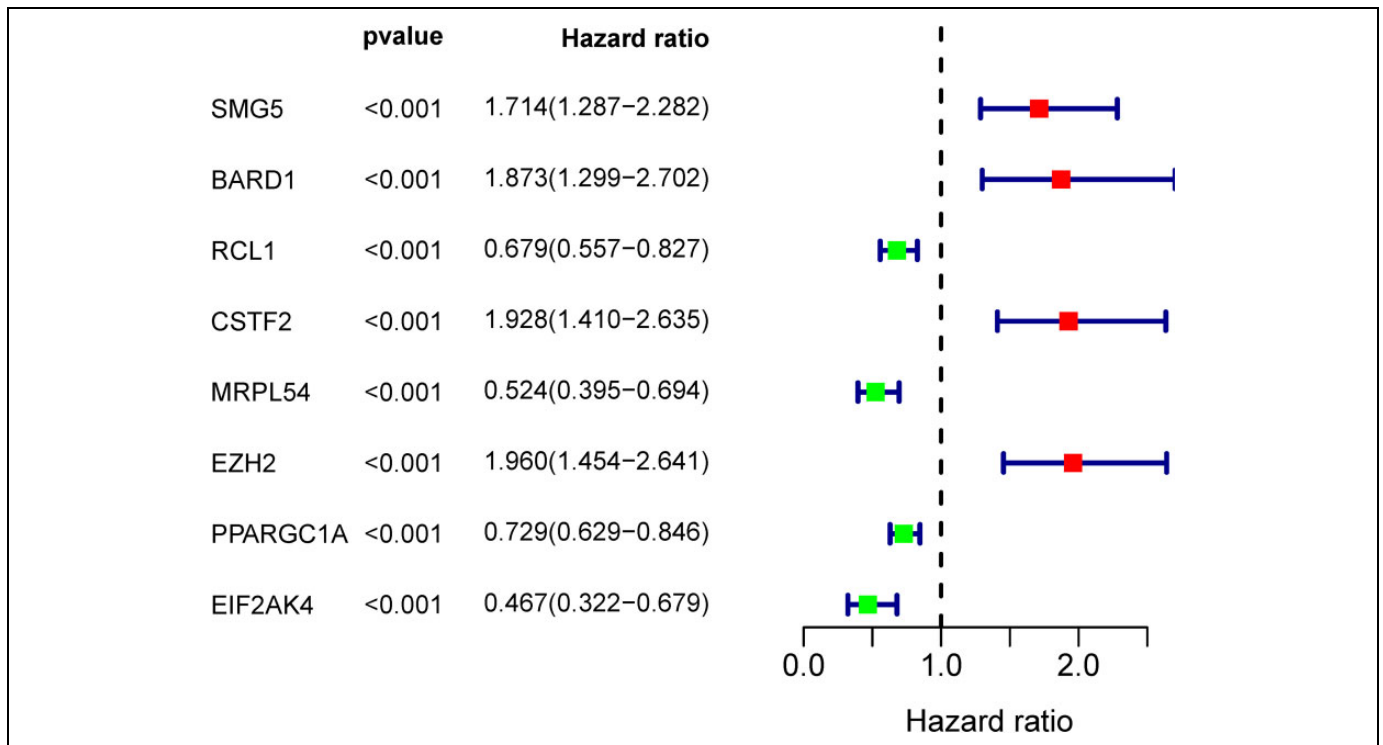
Figure 4. Univariate Cox regression study in the training group for the detection of hub RBPs.

RBPs Selecting Related to Prognosis

A total of 98 different key RBPs in the PPI network have been established. In order to study the prognostic significance of these RBPs, we conducted a univariate Cox regression analysis and obtained 8 hub RBPs associated to prognosis (Figure 4). Those 8 RBPs were analyzed in several step-by-step Cox regression for their impact on patient survival time and clinical outcomes, 4 RBPs were detected in the HCC patients as independent predictors (Figure 5).
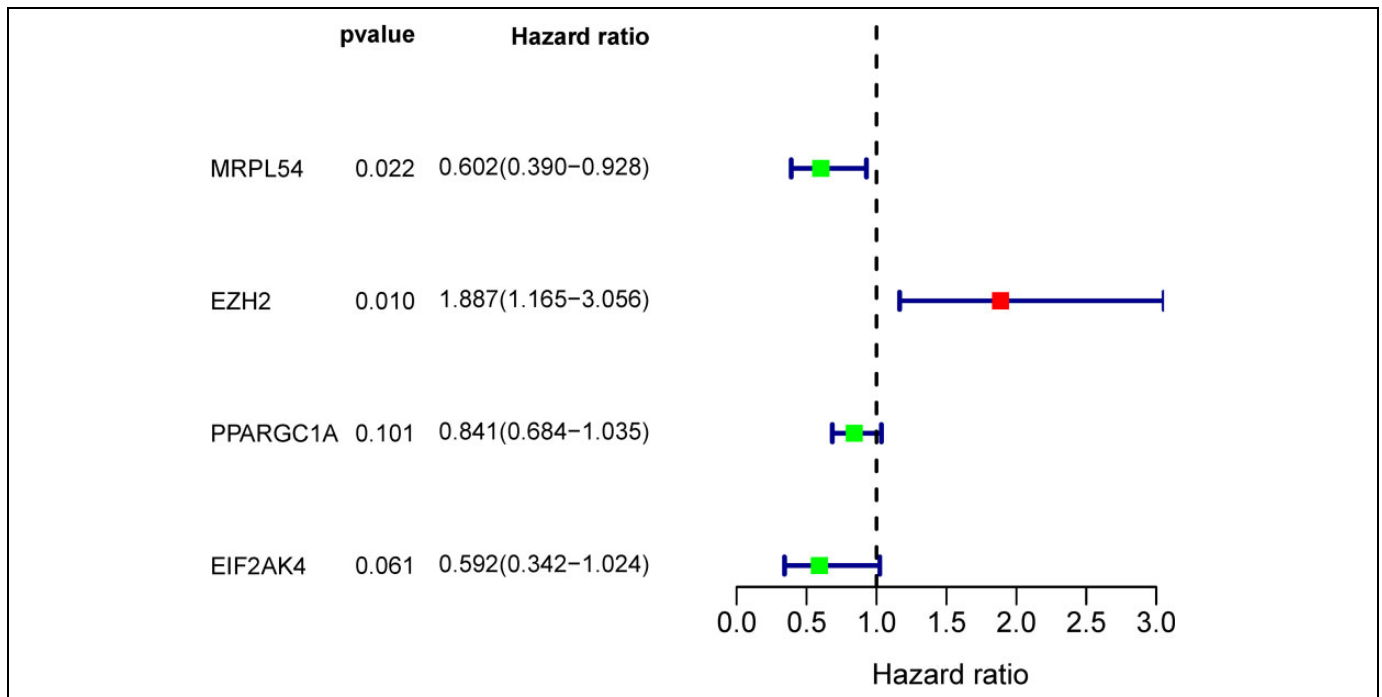
Figure 5. Multivariate Cox regression analysis to determine RBPs associated with prognosis.

Candidate Hub Genes Mutation and Copy Number Alteration Analyzes in HCC Patients

The online tool cBioportal for the HCC was used to co analyze the mutation and copy-number alteration (CNA) of the hub genes MRPL54, EZH2, PPARGC1A, and EIF2AK4. These 4 hub genes were altered in 366 samples of 442 samples of HCC (82.8%) (Figure 6A). The EZH2 amplification was the most common copy number alteration for the 4 hub genes (Figure 6B). Then we performed an interaction network with 24 nodes containing 4 hub genes and 20 neighboring genes which were most frequently altered (Figure 6C). We also found that the mitochondrial genes, including EED, SUZ12, were closely related to alterations of the 4 hub genes.

Genetic Risk Score Model Construction and Analysis Related to Prognosis

In order to construct the precious model, 4 hub RBPs identified from the multiple step-to-stop Cox regression analysis were used. The risk score was determined using the following formula for every patient:

Risk score = (-0.5078*expressionMRPL54) + (0.6350* expressionEZH2) + (-0.1731*expressionPPARGC1A) + (-0.5245*expressionEIF2AK4).

We then performed a survival analysis to evaluate the predictive capacity. A total of 370 HCC patients were randomly divided into the population of training (n = 186) and validation (n = 184). Each group of patients had

a median risk score in low and high-risk subgroups. The training population was used to assess the prediction model and the mathematical stratification of probabilities. The validation population was used to verify the prediction model. The findings showed that compared with the patients in the low-risk subgroups, the high-risk subgroup had poor OS (Figure 7A). A time-dependent ROC study was carried out to further assess the prognostic ability of the 4-RBP biomarker. We found that the area of the risk score model under the ROC curve (AUC) was 0.814 (Figure 7B), indicating moderate diagnostic efficiency. Figure 7C displayed the expression heatmap, patients' survival rate, and the signature risk score of 4 RBPs in the low- and high-risk subgroups. The findings indicated in the validation group that the prognostic model is more sensitive and more accurate (Figure 8A–C).

Construction of a Nomogram Based on the 4 Hub RBPs

To develop a quantitative HCC prognostic system, we integrated a nomogram with the signature of 4 RBPs (Figure 9A). The point scale in the nomogram for the individual variables was used to focus on the multivariate Cox analysis. A horizontal line is drawn to evaluate the point of each variable and to quantify the cumulative points for each individual by adding up the points of all variables and normalizing it to a 0 to 100 distribution. We can calculate the estimated survival rates for HCC patients at 3, and 5 years and draft a vertical line between the total point axis and each prognosis axis. This may help the practitioners make clinical
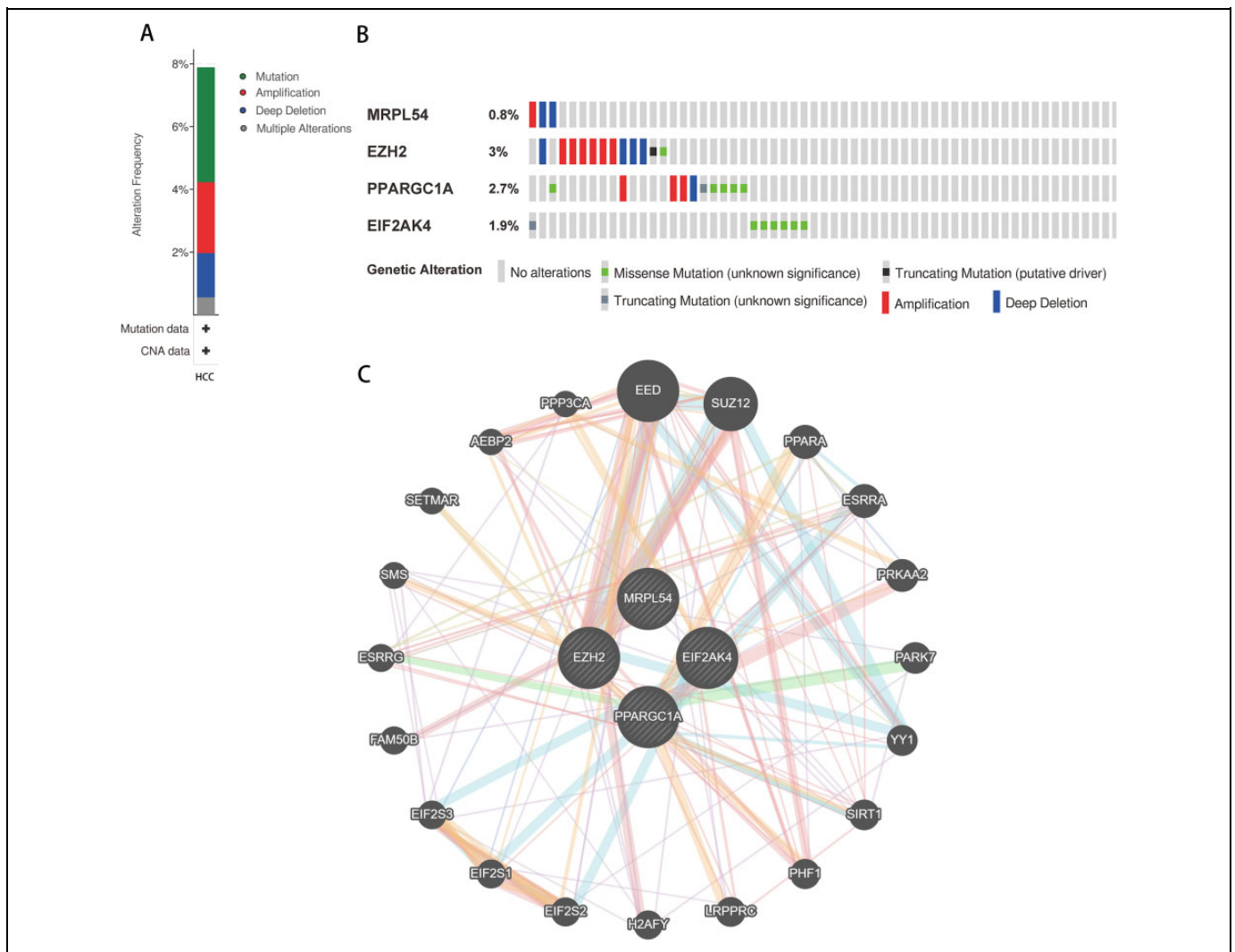
Figure 6. Hub RBP expression and HCC study of alteration. (A) Mutation frequency of hub genes; (B) mutation frequency of each gene; and (C) interaction network.

judgments for HCC patients. We developed calibration plots, which show that the expected and observed findings are well compliant (Figure 9B). In addition, by conducting COX regression analyses, we evaluated the prognostic significance of various clinical characteristics in HCC TCGA patients. The results showed a correlation between stage and risk score and the OS of patients with HCC ($P < .05$) (Table 1). However, we noticed that stage and risk score were independent OS-related predictor variables through multiple regression analyzes ($P < 0.05$) (Table 1).

According to these results, we established a dynamic web-based calculator (https://dxyjiang.shinyapps.io/RBPpredict/) to predict the OS of patients with HCC according to the nomogram (Figure 10). The calculator predicted the survival of patients according to their points of variables. For example, the 5-year OS rate was approximately 77% for a patient whose MRPL54 score was 7, EZH2 score was 3, PRARGC1A score was 4 and EIF2AK4 score was 4.

Prognostic Value and Expression of Hub RBPs Validation

The Kaplan-Meier was used to further analyze the predictive value of 4 Hub RBPs in the HCC to assess the relationship between the 4 RBP Hubs (MRPL54, EZH2, PPARGC1A, EIF2AK4) and OS. The findings of the log-rank analysis revealed that in HCC cases, the 4 RBPs were correlated with the OS (Figure 11). In addition, we have used immunohistochemistry results from HPA data to determine how these hub RBPs are exposed in HCC to show that EZH2 has been significantly increased in HCC versus normal liver tissue (Figure 12). Nevertheless, the antibody staining concentrations of EIF2AK4 and MRPL54 were relatively reduced in HCC tissue and PPARGC1A immunohistochemistry was not found in this database which must be further illustrated.

Discussion

The worldwide high HCC mortality rate was influenced by viral diseases, nutrition, environmental conditions and other
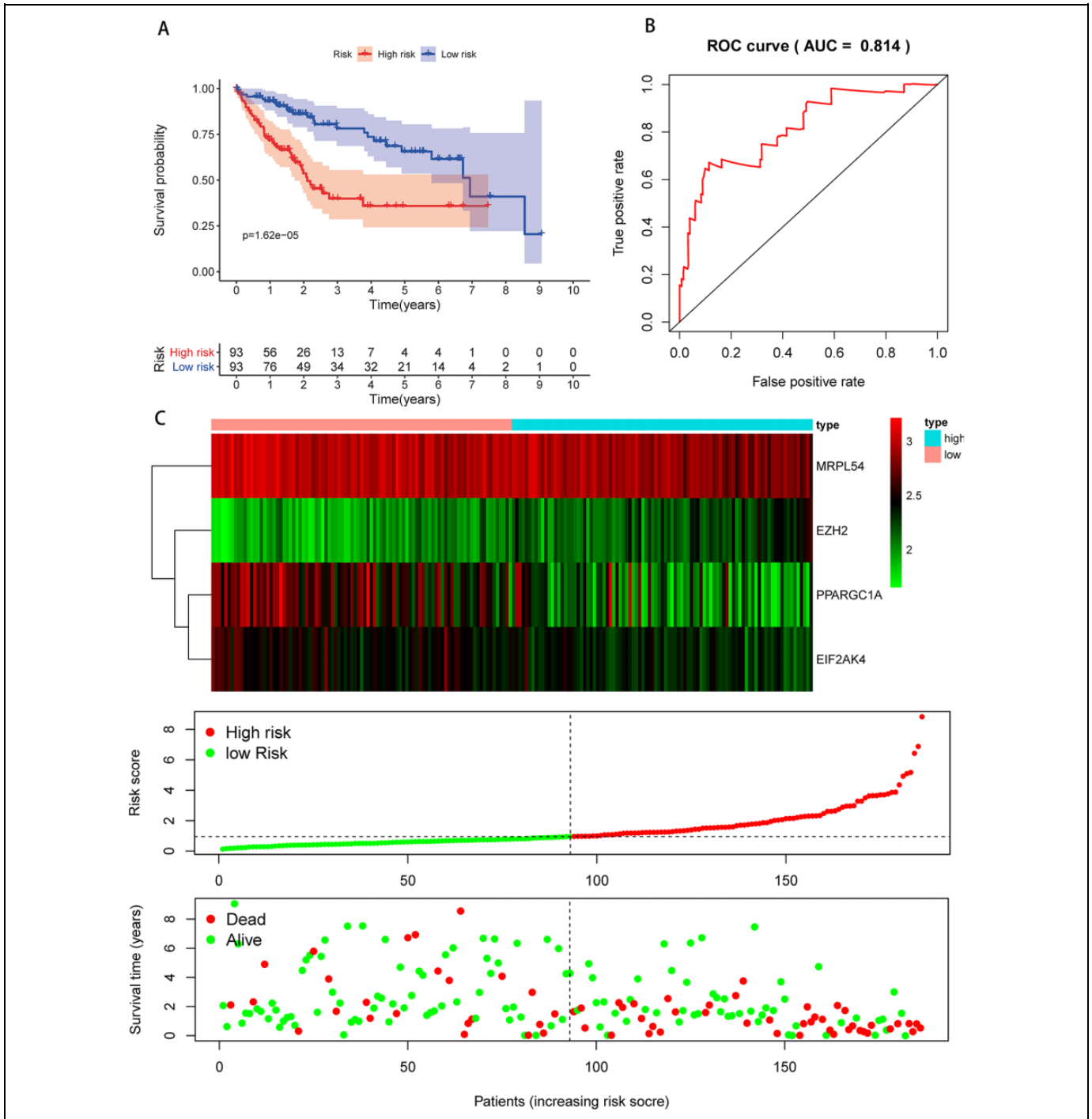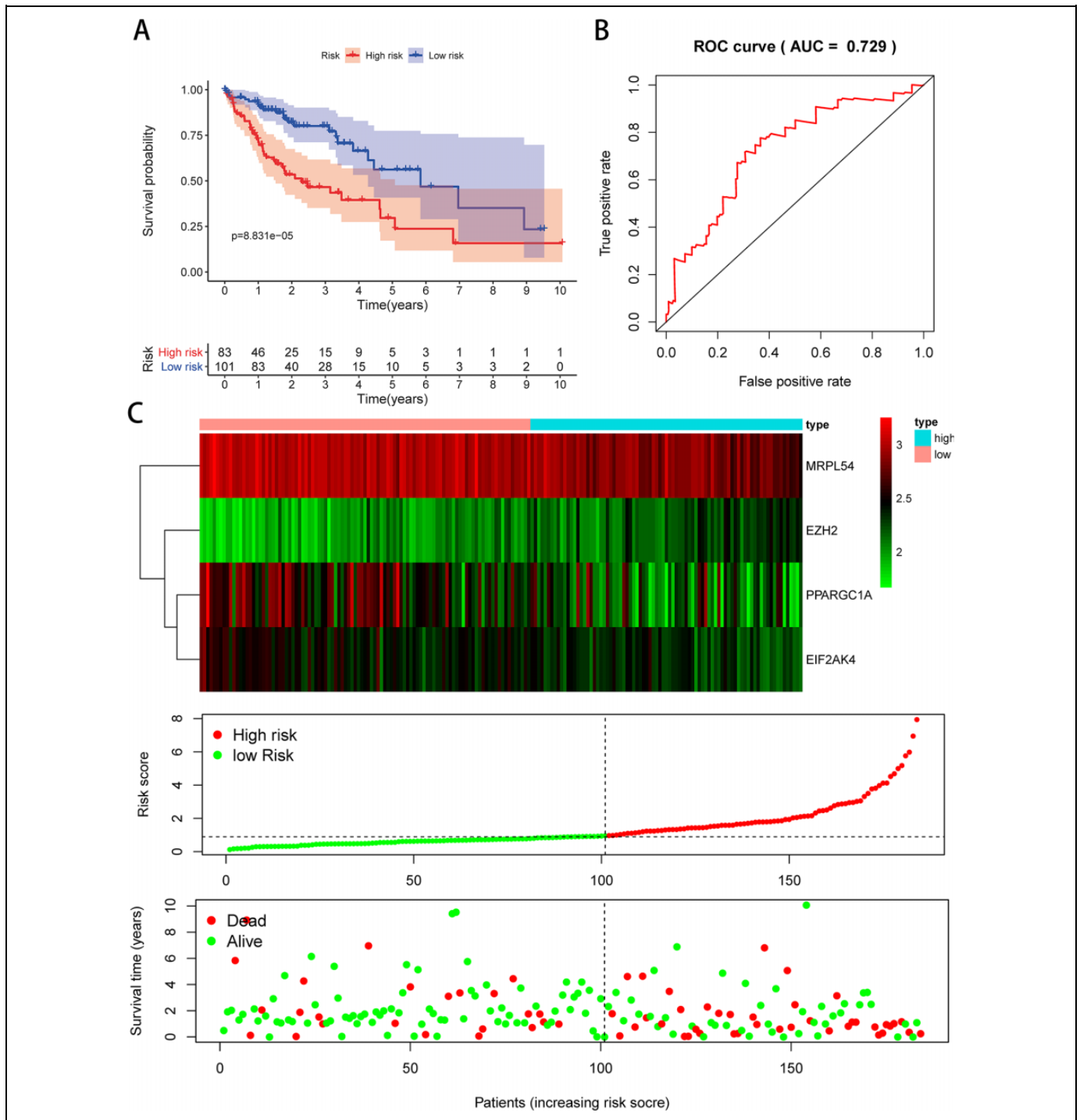
Figure 7. Risk score analysis of the TCGA training group using 4-genes prognostic model. (A) Low- and high-risk subgroup survival curve; (B) risk score-based ROC prediction curves; (C) expression heat map, risk score distribution, and survival status.

factors.[25] Constant developments in the area of surgical medicine, chemotherapy and molecular biology have improved our understanding of cancer biology, and significant advancement has been made in the treatment of HCC.[26,27] Dysregulation of RBPs has been documented in various malignant tumors.[28] Nevertheless, only a limited number of RBPs have been extensively studied and verified as contributing to cancer growth. We have established 330 RBPs

between the tumor and the normal tissue based on HCC data from TCGA in this study. We studied the biological pathways of these RBPs systematically, established co-expression network and PPI network. In addition, we conducted univariate Cox regression analysis and ROC analyzes of RBP hubs to further explore its biological roles and clinical importance. We developed a risk model for HCC forecasting based on 4 prognostic-related RBP hub genes. Such findings will help

Figure 8. Risk score study of 4-genes prognostic model in the TCGA validation group. (A) Low- and high-risk subgroup survival curve; (B) risk score-based ROC prediction curves; (C) expression heat map, risk score distribution, and survival status.

create novel biomarkers for the diagnosis and prognosis of HCC patients.

The enrichment function pathway evaluation shows that specific RBPs have been enriched considerably in ncRNA processing, RNA splicing, mRNA splicing, regulation of translation, regulation of cellular amide metabolic process, RNA 3'-end processing. Previous researches have shown that a number of human

diseases have occurred and have progressed in regulation of translation, RNA processing, and RNA splicing. RNA stability regulation after transcription is an important process in the processing of gene expression. RBPs are able to associate with RNA to build ribonucleoprotein complexes, thereby increasing the stability and gene expression of target mRNAs that play a key role in the development of various diseases. ILF3 plays an important
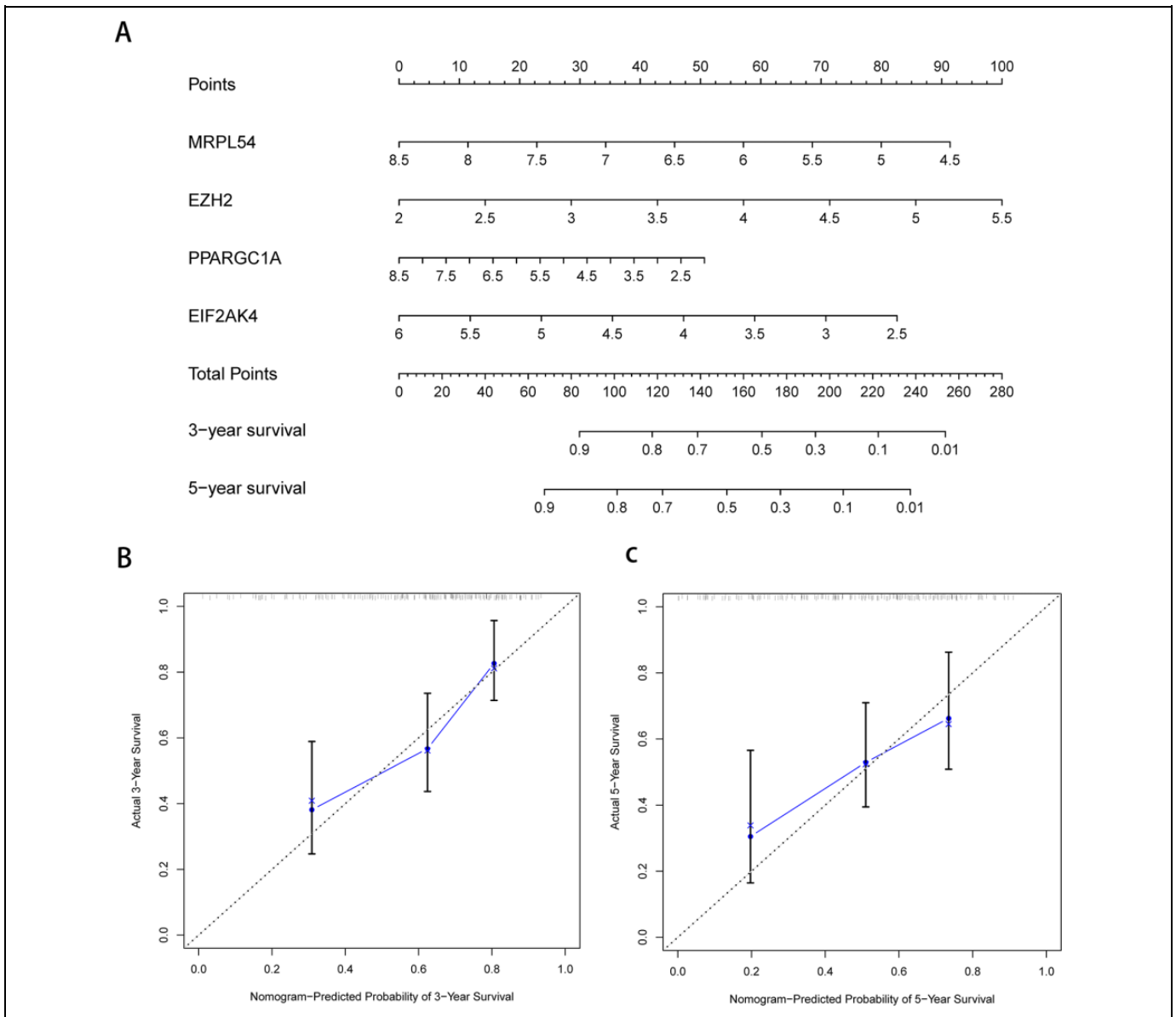
Figure 9. Nomogram and calibration plots of 4 RBPs. (A) Nomogram for the TCGA training group to estimate 3- and 5-year OS. (B, C) Calibration plots of the nomogram to estimate OS at 3 and 5 years.

Table 1. The Predictive Effect of Various Clinical Parameters.

	Univariate analysis			Multivariate analysis		
	HR	95% CI	P-value	HR	95% CI	P-value
Age	1.001	0.982-1.021	0.902	1.007	0.986-1.029	0.497
Gender	0.669	0.390-1.145	0.143	0.604	0.343-1.064	0.081
Grade	1.138	0.801-1.616	0.472	1.038	0.706-1.525	0.850
Stage	1.549	1.164-2.061	0.003	1.442	1.061-1.960	0.019
riskScore	1.453	1.215-1.738	<0.001	1.450	1.189-1.769	<0.001

role in the regulation of several aspects of the metabolism of RNA, primarily because of its dual-stranded RNA binding motives[29] and was related to many malignant tumors and progression.[30-33] HNRNPM has also been shown to facilitate the

invasion and metastasis of breast cancer by controlling the alternate splitting of CD44 and may control mesenchymal epithelial transformation in non-small-cell lung cancer and modulate apoptosis in osteosarcoma.[34-37] The KEGG pathways analysis shows that the aberrantly expressed RBPs regulate the tumorigenesis and progression of HCC through effects on spliceosome, RNA transport, mRNA surveillance pathway and ribosome.

In addition, we built a network of protein-protein interactions between these various RBPs and have a key module containing 53 main RBPs. Some of these key RBPs have been found to play a significant role in tumor growth and progression. EZH2 is an epigenetic regulator of the polycomb group of proteins with important roles in the regulation of embryonic stem cells. Varambally *et al* stated that EZH2 was overexpressed and correlated with the underexpression of miR-101 in prostate
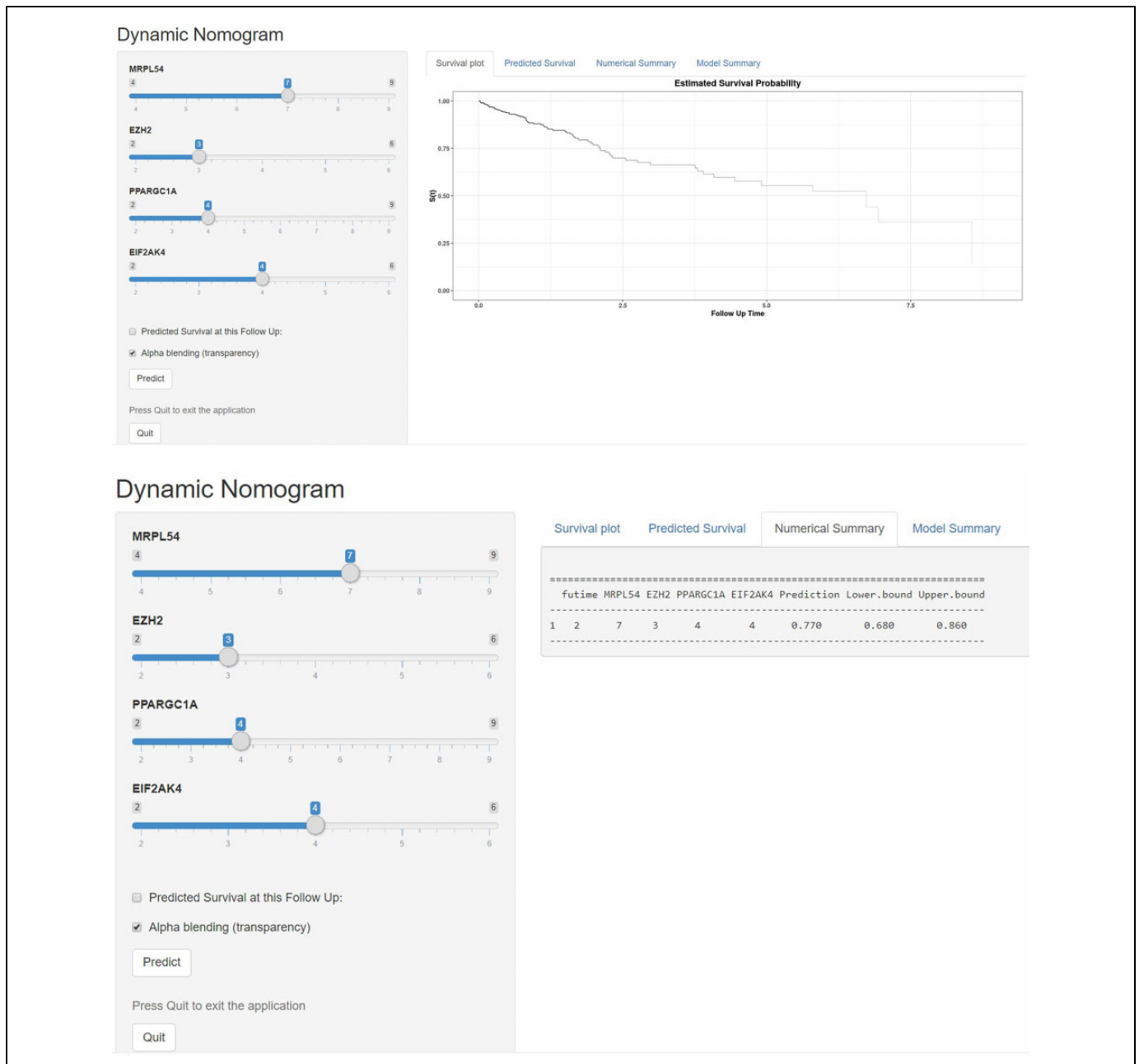
Figure 10. A 95% confidence interactive web-based calculator according to the web survival rate calculator.

cancer.[38,39] In reaction to a significant number of stress-induced signs, such as heat shock, oxidotic tension, and viral infection, the EIF2AK4 gene has been associated with the serin-51 phosphorylate eIF2α and rapidly inhibits translation initiatives.[40,41] The gene PPARGC1A encodes a transcriptional coactivator protein that plays a key role in metabolism of energy by mitochondrial biogenesis, fatty acid oxidation, hepatic gluconeogenesis, glucose absorption, and lipid metabolism.[42-45] The module study in the PPI network revealed that HCC is linked to RNA splicing with bulge adenosine as the nucleophile, mRNA splicing via spliceosome, RNA splicing via transesterification reactions, ncRNA processing and RNA splicing.

The hub RBPs were also chosen based on univariate Cox regression analysis, survival analyses, and multiple Cox regression analysis. A total of 4 RBPs, including MRPL54, EZH2, PPARGC1A, and EIF2AK4 were identified as prognosis related hub RBPs. Previous studies indicated that tumorigenesis and development of HCC patients were related to the expression MRPL54, EZH2, PPARGC1A, and EIF2AK4, which was compatible with our findings. Next, we have produced a risk model to predict HCC prognosis by using the 4 hub RBP genes coding analysis. The ROC curve analysis revealed the better diagnostic capability of these 4 genes in order to select poorly diagnosed HCC patients. However, these 4 RBPs'
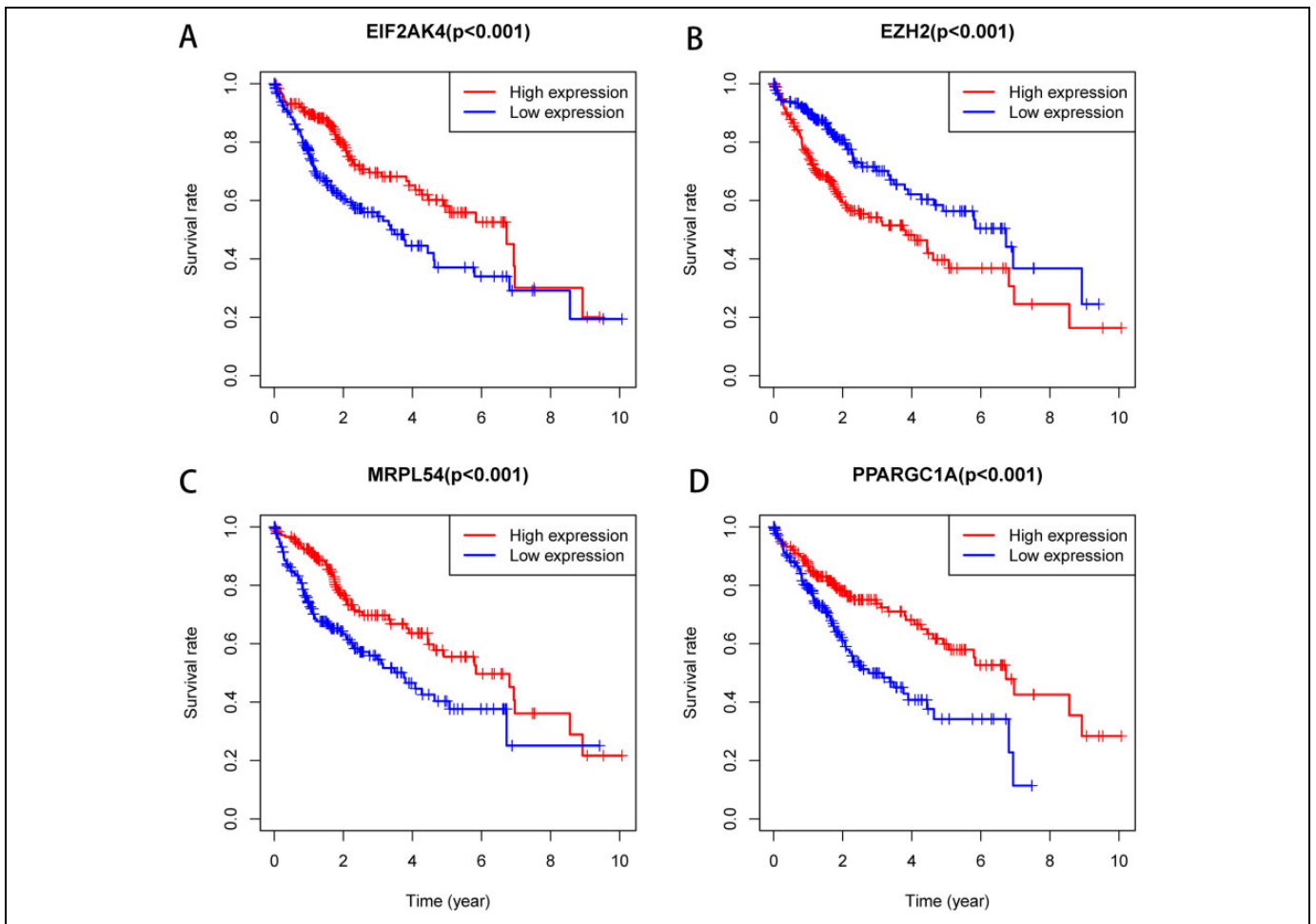
Figure 11. Prognostic value of 4 key RBPs in HCC. (A) EIF2AK4; (B) EZH2; (C) MRPL54; (D) PPARGC1A.
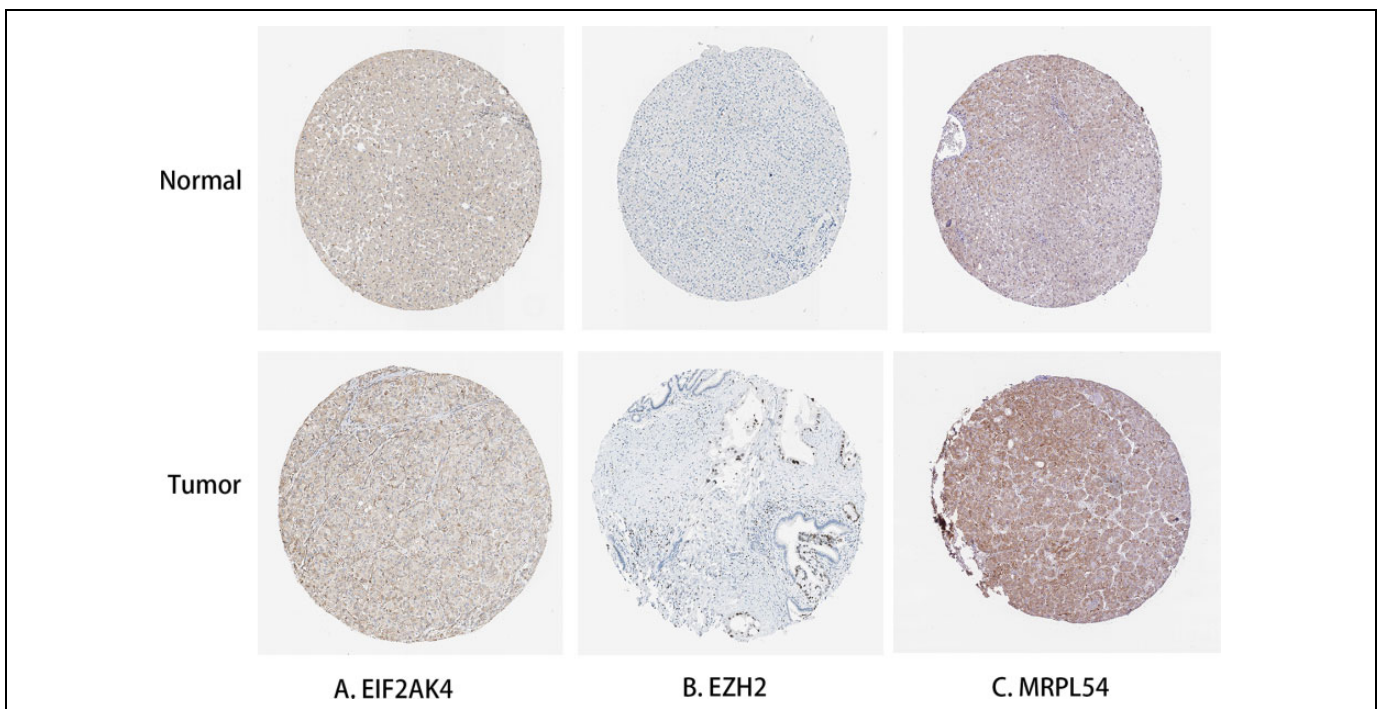
Figure 12. Protein expression validation of hub genes in normal liver tissue and HCC by HPA database. (A) EIF2AK4; (B) EZH2; (C) MRPL54.

molecular pathways leading to the HCC and further investigation of possible pathways is still not well known. A nomogram was then developed to help us more intuitively estimate 3 and 5 years OS. The Kaplan-Meier plotter to used to predict the prognostic significance of the 4 RBP genes. The findings of the predictions of the TCGA cohort were essentially accurate. Such fundings revealed that the 4-gene signature prognostic model has a certain value in adjusting HCC patients' treatment plans.

In general, our prognostic model relies on 4 gene-coding RBPs, which reduces sequencing costs significantly and is more suitable for clinical applications. Furthermore, the 4 genes predictive model has improved survival prediction efficiency in HCC patients. In addition, the RBPs related gene signature demonstrated essential biological activity, indicating it may be used for clinic assistant treatment, which in previous research was not generally always the case. Nevertheless, in this analysis, there are certain limitations. First of all, our prognostic model was based on TCGA data that is not validated in the cohort of clinical patients and other databases. Second, our research was focused on retrospective examination and further work should be carried out to validate the findings. Third, the data sets did not have clinical details that could improve multivariate Cox regression analyses' predictive validity and reliability. The data of TCGA are updated constantly, and future work could focus on further validating the model using updated TCGA datasets or more samples. In addition, more computational methods about differential and co-expressed analysis for genes should be added to improve the results.

Conclusion

In an HCC series of bioinformatic analyses, we examined the expression, potential functions and predictive values of aberrantly expressed RBPs. The RBPs were related to oncogenesis, growth, invasion and metastasis. A 4-RBP model was designed as an independent forecast signature for HCC patients. To our full knowledge, this is the first study creating a prognostic model for HCC consistent with RBP. Such results give valuable insight into HCC pathogenesis, which may assist in therapeutic decision making and individual therapy.

Authors' Note

Data curation, Feng Jiang and Chuyan Wu; Methodology, Ke Wei; Software, Feng Jiang; Verification, Ming Wang; Visualization, Chuyan Wu; Writing-original draft, Feng Jiang and Chuyan Wu; Writing-review and editing, Guoping Zhou and Jimei Wang. Feng Jiang, MD, and Chuyan Wu, MD, contributed equally to this work. The TCGA data were retrieved from GDC data portal (https://portal.gdc.cancer.gov/). The R software (https://www.r-project.org) was used for all statistical analyses. The data we utilized were drawn from TCGA dataset, which were not obtained from our own clinical samples, so our study did not require an ethical board approval.

Author Contributions

Ming Wang and Feng Jiang contributed equally to this work.

ORCID iD

Feng Jiang, MD 🆔 https://orcid.org/0000-0002-3525-8752
Chuyan Wu, MD 🆔 https://orcid.org/0000-0002-5622-6020
Jimei Wang, PhD 🆔 https://orcid.org/0000-0002-6755-8709

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2018;68(6):394.

2. Negoita S, Feuer EJ, Mariotto A, et al. Annual report to the nation on the status of cancer, part II: recent changes in prostate cancer trends and disease characteristics. *Cancer-Am Cancer SoC* 2018; 124(13):2801.

3. Zhang W, Mu D, Feng K. Hierarchical potential differentiation of liver cancer stem cells. *Adv Clin Exp Med*. 2017;26(7):1137.

4. Dawkins J, Webster RM. The hepatocellular carcinoma market. *Nat Rev Drug Discov*. 2019;18(1):13-14.

5. Gerstberger S, Hafner M, Tuschl T. A census of human RNA-binding proteins. *Nat Rev Genet*. 2014;15(12):829-845.

6. Velasco MX, Kosti A, Penalva L, Hernandez G. The diverse roles of RNA-binding proteins in glioma development. *Adv Exp Med Biol*. 2019;1157:29-39.

7. 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(3):e1520.

8. Gerstberger S, Hafner M, Tuschl T. A census of human RNA-binding proteins. *Nat Rev Genet*. 2014;15(12):829-845.

9. Duan Y, Du A, Gu Y, et al. PARylation regulates stress granule dynamics, phase separation, and neurotoxicity of disease-related RNA-binding proteins. *Cell Res*. 2019;29(3):233-247.

10. Calabretta S, Richard S. Emerging roles of disordered sequences in RNA-binding proteins. *Trends Biochem Sci* 2015;40(11): 662-672.

11. Johnson E, Dammer EB, Duong DM, et al. Deep proteomic network analysis of Alzheimer's disease brain reveals alterations in RNA binding proteins and RNA splicing associated with disease. *Mol Neurodegener*. 2018;13(1):52.

12. 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(18):1380-1388.

13. Conrad T, Albrecht AS, de Melo CV, Sauer S, Meierhofer D, Orom UA. Serial interactome capture of the human cell nucleus. *Nat Commun*. 2016;7:11212.

14. Wurth L. Versatility of RNA-binding proteins in cancer. *Comp Funct Genomics*. 2012;2012:178525.

15. Castello A, Hentze MW, Preiss T. Metabolic enzymes enjoying new partnerships as RNA-binding proteins. *Trends Endocrinol Metab*. 2015;26(12):746-757.

16. Ghisolfi L, Dutt S, McConkey ME, Ebert BL, Anderson P. Stress granules contribute to alpha-globin homeostasis in differentiating erythroid cells. *Biochem Biophys Res Commun*. 2012;420(4): 768-774.

17. El-Naggar AM, Sorensen PH. Translational control of aberrant stress responses as a hallmark of cancer. *J Pathol*. 2018;244(5): 650-666.

18. Takahashi M, Higuchi M, Matsuki H, et al. Stress granules inhibit apoptosis by reducing reactive oxygen species production. *Mol Cell Biol*. 2013;33(4):815-829.

19. Homer C, Knight DA, Hananeia L, et al. Y-box factor YB1 controls p53 apoptotic function. *Oncogene*. 2005;24(56):8314-8315.

20. Lasham A, Print CG, Woolley AG, Dunn SE, Braithwaite AW. YB-1: oncoprotein, prognostic marker and therapeutic target? *Biochem J*. 2013;449(1):11-23.

21. Bayfield MA, Yang R, Maraia RJ. Conserved and divergent features of the structure and function of La and La-related proteins (LARPs). *Biochim Biophys Acta*. 2010;1799(5-6):365-378.

22. Castello A, Fischer B, Eichelbaum K, et al. Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. *Cell*. 2012;149(6):1393-1406.

23. Stavraka C, Blagden S. The La-related proteins, a family with connections to cancer. *Biomolecules*. 2015;5(4):2701-2722.

24. Koso H, Yi H, Sheridan P, et al. Identification of RNA-binding protein LARP4B as a tumor suppressor in glioma. *Cancer Res*. 2016;76(8):2254-2264.

25. Henley SJ, Thomas CC, Lewis DR, et al. Annual report to the nation on the status of cancer, part II: Progress toward Healthy People 2020 objectives for 4 common cancers. *Cancer Am Cancer Soc*. 2020;126(10):2250-2266.

26. Wu L, Pan C, Wei X, et al. lncRNA KRAL reverses 5-fluorouracil resistance in hepatocellular carcinoma cells by acting as a ceRNA against miR-141. *Cell Commun Signal*. 2018;16(1):47.

27. Lin X, Shen J, Dan P, et al. RNA-binding protein LIN28B inhibits apoptosis through regulation of the AKT2/FOXO3A/BIM axis in ovarian cancer cells. *Signal Transduct Target Ther*. 2018;3:23.

28. Wang K, Li L, Fu L, et al. Integrated bioinformatics analysis the function of RNA binding proteins (RBPS) and their prognostic value in breast cancer. *Front Pharmacol*. 2019;10:140.

29. Castella S, Bernard R, Corno M, Fradin A, Larcher JC. Ilf3 and NF90 functions in RNA biology. *Wiley Interdiscip Rev Rna*. 2015;6(2):243-256.

30. Hu Q, Lu YY, Noh H, et al. Interleukin enhancer-binding factor 3 promotes breast tumor progression by regulating sustained urokinase-type plasminogen activator expression. *Oncogene*. 2013;32(34):3933-3943.

31. Zhang Y, Yang C, Zhang M, et al. Interleukin enhancer-binding factor 3 and HOXC8 co-activate cadherin 11 transcription to promote breast cancer cells proliferation and migration. *Oncotarget*. 2017;8(64):107477-107491.

32. Guo Y, Fu P, Zhu H, et al. Correlations among ERCC1, XPB, UBE2I, EGF, TAL2 and ILF3 revealed by gene signatures of histological subtypes of patients with epithelial ovarian cancer. *Oncol Rep*. 2012;27(1):286-292.

33. Zhu H, Yu JJ. Gene expression patterns in the histopathological classification of epithelial ovarian cancer. *Exp Ther Med*. 2010; 1(1):187-192.

34. Kedzierska H, Piekielko-Witkowska A. Splicing factors of SR and hnRNP families as regulators of apoptosis in cancer. *Cancer Lett*. 2017;396:53-65.

35. Goncalves V, Pereira J, Jordan P. Signaling pathways driving aberrant splicing in cancer cells. *Genes (Basel)*. 2017; 9(1):9.

36. Xu Y, Gao XD, Lee JH, et al. Cell type-restricted activity of hnRNPM promotes breast cancer metastasis via regulating alternative splicing. *Genes Dev*. 2014;28(11):1191-1203.

37. Yang JH, Chiou YY, Fu SL, et al. Arginine methylation of hnRNPK negatively modulates apoptosis upon DNA damage through local regulation of phosphorylation. *Nucleic Acids Res*. 2014;42(15):9908-9924.

38. Varambally S, Dhanasekaran SM, Zhou M, et al. The polycomb group protein EZH2 is involved in progression of prostate cancer. *Nature*. 2002;419(6907):624-629.

39. Varambally S, Cao Q, Mani RS, et al. Genomic loss of microRNA-101 leads to overexpression of histone methyltransferase EZH2 in cancer. *Science*. 2008;322(5908): 1695-1699.

40. Taniuchi S, Miyake M, Tsugawa K, Oyadomari M, Oyadomari S. Integrated stress response of vertebrates is regulated by four eIF2alpha kinases. *Sci Rep*. 2016;6:32886.

41. Berlanga JJ, Ventoso I, Harding HP, et al. Antiviral effect of the mammalian translation initiation factor 2alpha kinase GCN2 against RNA viruses. *Embo J*. 2006;25(8):1730-1740.

42. Puigserver P, Spiegelman BM. Peroxisome proliferator-activated receptor-gamma coactivator 1 alpha (PGC-1 alpha): transcriptional coactivator and metabolic regulator. *Endocr Rev*. 2003; 24(1):78-90.

43. Kunej T, Globocnik PM, Dovc P, Peterlin B, Petrovic D. A Gly482Ser polymorphism of the peroxisome proliferator-activated receptor-gamma coactivator-1 (PGC-1) gene is associated with type 2 diabetes in Caucasians. *Folia Biol (Praha)*. 2004; 50(5):157-158.

44. Medina-Gomez G, Gray S, Vidal-Puig A. Adipogenesis and lipotoxicity: role of peroxisome proliferator-activated receptor gamma (PPARgamma) and PPARgammacoactivator-1 (PGC1). *Public Health Nutr* 2007;10(10A):1132-1137.

45. Da FA, Ochioni AC, Martins R, et al. Adiponectin, retinoic acid receptor responder 2, and peroxisome proliferator-activated receptor-gamma coativator-1 genes and the risk for obesity. *Dis Markers* 2017;2017:5289120.