Am J Cancer Res 2023;13(6):2269-2284
www.ajcr.us /ISSN:2156-6976/ajcr0149790

Original Article

SFPQ promotes the proliferation, migration and invasion of hepatocellular carcinoma cells and is associated with poor prognosis

Youyu Lin[1,2*], Wenting Zhong[2*], Qili Lin[5*], Yan Ye[2*], Shengzhao Li[2], Hui Chen[2], Hongxia Liu[2], Luyun Xu[2], Wei Zhuang[3], Shaoqing Chen[6], Yao Lin[1,4], Qingshui Wang[1,2]

[1]The Second Affiliated Hospital of Fujian Traditional Chinese Medical University, Fujian-Macao Science and Technology Cooperation Base of Traditional Chinese Medicine-Oriented Chronic Disease Prevention and Treatment, Innovation and Transformation Center, Fujian University of Traditional Chinese Medicine, Fuzhou, Fujian, China; [2]College of Life Sciences, Fujian Normal University, Fuzhou, Fujian, China; [3]Department of Urology, The Second Affiliated Hospital of Fujian Medical University, Quanzhou, Fujian, China; [4]Collaborative Innovation Center for Rehabilitation Technology, Fujian University of Traditional Chinese Medicine, Fuzhou, Fujian, China; [5]Department of Pathology, Fujian Medical University Union Hospital, Fuzhou, Fujian, China; [6]Fujian University of Traditional Chinese Medicine, Fuzhou, Fujian, China. [*]Equal contributors.

Received February 21, 2023; Accepted May 15, 2023; Epub June 15, 2023; Published June 30, 2023

Abstract: Liver cancer is a prevalent type of tumor worldwide. CRISPR-Cas9 technology can be utilized to identify therapeutic targets for novel therapeutic approaches. In this study, our goal was to identify key genes related to the survival of hepatocellular carcinoma (HCC) cells by analyzing the DepMap database based on CRISPR-Cas9. We screened candidate genes associated with HCC cell survival and proliferation from DepMap and identified their expression levels in HCC from the TCGA database. To develop a prognostic risk model based on these candidate genes, we performed WGCNA, functional pathway enrichment analysis, protein interaction network construction, and LASSO analysis. Our findings show that 692 genes were critical for HCC cell proliferation and survival, and among them, 571 DEGs were identified in HCC tissues. WGCNA categorized these 584 genes into three modules, and the blue module consisting of 135 genes was positively linked to the tumor stage. Using the MCODE approach in Cytoscape, we identified ten hub genes in the PPI network, and through Cox univariate analysis and Lasso analysis, we developed a prognostic model consisting of three genes (SFPQ, SSRP1, and KPNB1). Furthermore, knocking down SFPQ inhibited HCC cell proliferation, migration, and invasion. In conclusion, we identified three core genes (SFPQ, SSRP1, and KPNB1) that are essential for the proliferation and survival of HCC cells. These genes were used to develop a prognostic risk model, and knockdown of SFPQ was found to inhibit the proliferation, migration, and invasion of HCC cells.

Keywords: HCC, DepMap, proliferation, SFPQ, SSRP1, KPNB1

Introduction

Hepatocellular carcinoma (HCC) is a prevalent malignant tumor of the digestive system, accounting for more than 90% of all liver cancer cases [1]. Despite significant progress in clinical and experimental treatment of HCC due to advances in surgical techniques, most patients only exhibit clinical symptoms in advanced stages, and the complex pathophysiological mechanisms result in a poor prognosis [2]. Combining biomarkers can enhance diagnostic accuracy, representing a new direction in the clinical diagnosis of cancer [3, 4]. Therefore, identifying reliable biodiagnostic markers of HCC and developing novel therapeutic approaches remain necessary prerequisites for effective treatment of HCC patients.

As more mutated genes associated with the development of human cancer are being identified, there is a rapid development of therapeutics targeting these genes [5]. However, this discovery alone is not enough, and it is necessary to link these dependent genes with the genetic or molecular characteristics of tumors

to discover new therapeutic targets. The Cancer Dependency Map (DepMap) project (https://depmap.org/portal/) provides histological data derived from various sources such as the Cancer Cell Lineage Encyclopedia (CCLE), small molecule sensitivity data, Broad Achilles Project genetic dependent data, and Sanger's GDSC [6]. The project uses CRISPR-Cas9 and RNAi technology to silence or eliminate individual genes and identify those critical to the survival of cells [7, 8]. The CERES algorithm used for CRISPR screening is also obtained from DepMap, where the CERES score represents the effect of gene knockdown or knockout. Lower scores indicate that the gene may be more important for cell growth and survival [9].

In this study, we identified core essential genes that are differentially expressed in HCC and critical for HCC cell survival using the Depmap website and TCGA database. We then obtained the most relevant module for HCC progression through WGCNA analysis and functionally annotated the module genes using KEGG and GO analysis. Next, we developed and verified a risk predictive model. Finally, we used CCK8, transwell, and migration assays to detect the effect of the screened core gene-SFPQ (Splicing Factor Proline and Glutamine Rich) on the function of HCC cells. SFPQ is a DNA and RNA-binding nuclear protein that plays a crucial role in RNA translocation, apoptosis, and DNA repair [10-12]. It can bind to the dsDNA of many gene promoters and affect the transcriptional regulation of genes [13]. SFPQ can also bind to long-stranded non-coding RNAs (LncRNAs) of certain genes and participate in tumor formation and metastasis [11]. In recent years, further research has found that SFPQ is associated with many diseases, such as Alzheimer's disease [13-15]. Moreover, abnormal SFPQ function is associated with many tumor progressions. A high level of SFPQ promotes proliferation and development in colorectal cancer [16], while up-regulated SFPQ in melanoma is negatively associated with patient survival [17]. This study is the first to investigate the expression and function of SFPQ in HCC.

Methods

Identification of crucial HCC genes

The DepMap uses RNAi and CRISPR-Cas9 knockdown screens to obtain genomic information and sensitivity to genetic and small molecule perturbations in individual cell lines to distinguish genes that are critical for the proliferation and survival (https://depmap.org/portal/). The DepMap contains 483 cell lines screened with a library of approximately 74,000 sgRNAs targeting approximately 17,000 genes, and the resulting CERES scores are taken to conclude the gene-level dependence [18]. The positive CERES score showed that a gene knockdown enhances the survival and proliferation of cell. The native CERES score indicates that a gene knockdown inhibits the survival and proliferation of cell [19]. In the study, we classified candidate genes as essential genes with CERES scores < -1 in more than 75% of HCC cell lines.

Clinical data collection and extraction

Transcriptome data and clinical information of HCC were obtained from the TCGA (https://cancergenome.nih.gov/). Information on LIRI-JP data was obtained from the International Cancer Genome Consortium (ICGC) (https://icgc.org/) and was used as the validation cohort. Additionally, mRNA expression data of 27 melanoma patients who received anti-PD-1 checkpoint inhibitor therapy was downloaded from the GSE78220 dataset in the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/).

Screening for differentially expressed genes (DEGs)

The Limma R package was utilized to identify differentially expressed genes (DEGs) between HCC and normal liver tissues. Genes with a fold-change (FC) greater than 1.5 or less than 0.67 (which corresponds to $|\log2FC| \geq 0.585$) and a *p*-value less than 0.05 were deemed as significant DEGs [20].

Weighted gene co-expression network analysis (WGCNA) analysis

WGCNA analysis was conducted on the transcriptional expression profiles of DEGs in the TCGA database using the WGCNA R package [21]. A soft-threshold parameter of β=5 to the power and a scale-free R^2=0.9 were selected for this analysis. A signed network was utilized, and a minimum cluster size was applied for dendrogram clustering detection. In total, three modules were created.

Construction of prognostic risk model

The least absolute shrinkage and selection operator (LASSO) is a contemporary statistical technique that enables active selection from a large set of potentially multicollinear variables in regression analysis, resulting in a more correlated and explainable set of predictors [22]. By analyzing the TCGA-HCC transcriptomic data, LASSO was developed to establish a risk score formula: risk score = (0.0288 × expression of SFPQ) + (0.0168 × expression of SSRP1) + (0.0005 × expression of KPNB1).

GSEA analysis

We obtained the GSEA software from the official website (https://software.broadinstitute.org/gsea/index.jsp) and utilized biosignaling gene sets from MSigDB (https://software.broadinstitute.org/gsea/downloads.jsp) as reference gene sets to divide HCC patients into high-risk and low-risk groups. Each analysis was conducted with 1000 iterations using the default weighted enrichment statistical method.

Nomogram model

Clinical factors, including risk score, age, gender, stage, and grade, were incorporated to construct a nomogram to evaluate the probability of 1-, 3-, and 5-OS of HCC in the entire set. The nomogram's performance was evaluated using the 'rms' package by generating a calibration plot [23]. This plot compared the predicted probabilities of the nomogram with the actual observed rates.

Cell cultivation and transfection

HCC Hep3B cells and Huh7 cells were obtained from ATCC (American Type Culture Collection, Manassas, VA, USA). The cells were cultured in DMEM medium supplemented with 10% fetal bovine serum (FBS) and 1% penicillin/streptomycin sulfate. The cells were maintained in a humidified sterile cell culture incubator at 37°C with 5% CO_2. To transfect the encoded vectors, approximately 2 × 10^6 cells were inoculated in 6 cm culture dishes and Lipofectamine 2000 (Invitrogen, Carlsbad, CA, USA) was used for transfection of both Huh7 and Hep3B cells.

CCK-8 assay

The level of cell proliferation was evaluated using the CCK-8 assay. Transfected cells were seeded into four 96-well plates, with 2000 cells per well and six replicate wells per plate. After cell attachment, the CCK-8 reagent purchased from Beyotime (Shanghai, China) was added following the manufacturer's instructions, and the OD450 was measured using an enzyme marker after 1.5 hours. The absorbance of the cells was measured separately at 0, 24, 48, and 72 hours of incubation.

Invasion assay

The invasion capacity of Huh7 and Hep3B cells was assessed using the Trans-well assay. Cells were added to 24-well plates with an invasion chamber containing serum-free medium. The lower chamber was filled with complete medium and cultured at 37°C, 5% CO_2 for 48 hours. The invaded cells passing through the membrane were fixed with a methanol: PBS=7:3 mixture and stained with 500 μL of 0.1% crystal violet for 10-20 minutes. Cells on the upper chamber surface were wiped off with cotton swabs, and the number of invading tumor cells was randomly photographed at 6 different areas.

Migration assay

The migration ability of Huh7 and Hep3B cells was assessed using a wound healing assay. Firstly, the cells were cultured in 6-well plates until they reached approximately 90% confluency. Next, any dislodged cellular debris was washed away with PBS at least three times. A 10 μL pipette tip was then used to create a scratch wound, and any remaining debris was washed away with PBS. Finally, serum-free medium was added to the plates and the cells were incubated at 37°C, 5% CO_2 for 48 hours before being observed microscopically.

HCC cDNA microarray

The HCC cDNA microarrays were purchased from Shanghai Xinchao Biotechnology Co., Ltd., with the chip lot number: HLivH090Su01 and array number: 96*R100-M-20170915-**. The cDNA microarray contained 64 cancer tissue samples from patients with HCC and 24 adjacent non-cancerous tissue samples.

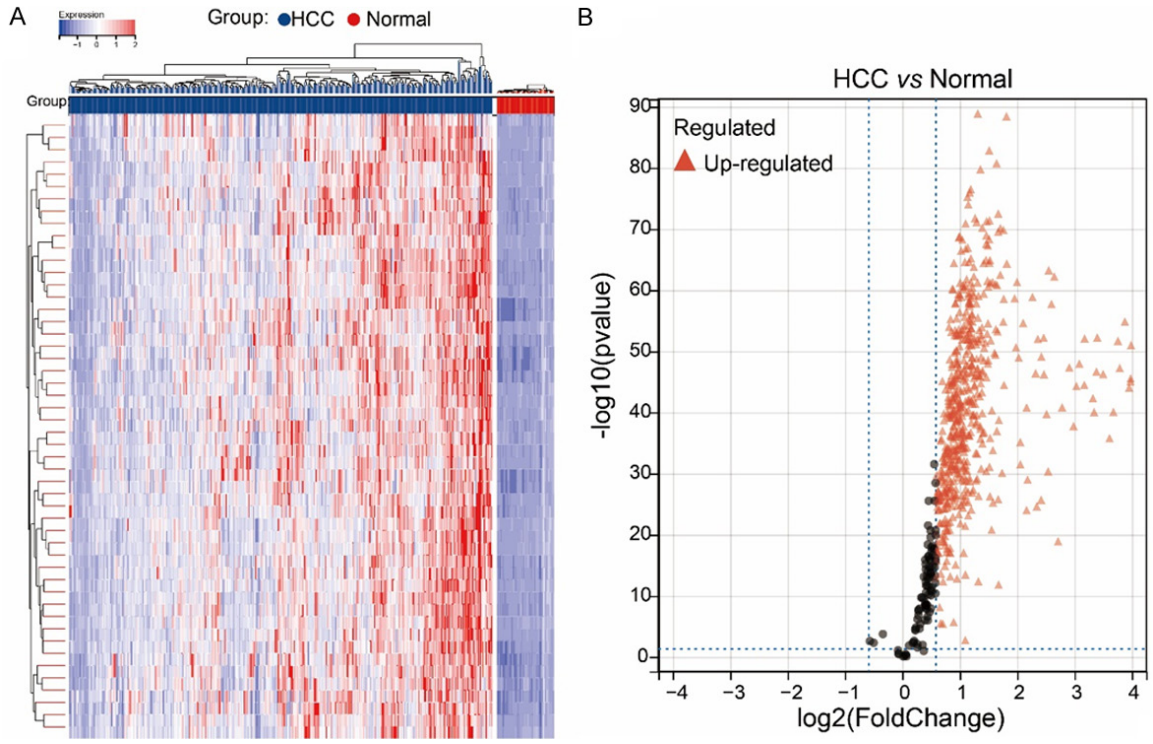


Figure 1. Analysis the expression of 692 candidate genes in HCC based on TCGA database. A. Heat map. Red color represents up-regulated DEGs and blue color represents down-regulated DEG. B. Volcano plot. Black plot indicates non-DEG, red plot indicates up-regulated DEG.

Real-time quantitative PCR

Real-time quantitative PCR was carried out using SYBR Green PCR kit (Takara, Dalian, China). The reaction conditions involved an initial denaturation at 95°C for 10 minutes followed by 40 cycles of 95°C for 1 minute and 58°C for 30 seconds. GAPDH was utilized as an internal control. Primer sequences were as follows: SFPQ forward 5'-GGGCTGT TGTAATAGTGGATGA-3'; SFPQ reverse 5'-CCAAAGACTCCATCGCTGA-3'. GAPDH forward 5'-CCTGCACC ACCAACTGCTTAG-3'; GAPDH reverse: 5'-GTGGATGCAGGGATGATGTTC-3'.

Statistical analysis

In this study, we plotted Kaplan-Meier curves to analyze overall survival (OS) using survival data from HCC patients in the TCGA and LIRI-JP databases. Additionally, we drew a receiver operating characteristic (ROC) curve to evaluate the prediction accuracy, with the area under the curve (AUC) representing the level of accuracy. A statistical significance of $P < 0.05$ was deemed significant.

Results

Identifying the essential HCC genes

We obtained CERES dependency scores for 24 HCC cell lines from DepMap and identified essential HCC genes with a CERES score of less than -1 in over 75% of the cell lines. A total of 692 genes were found to be crucial for sustaining the survival of HCC cell lines and were designated as candidate genes (Supplementary Table 1). We then aimed to identify differentially expressed genes (DEGs) in HCC among these 692 candidates. Our analysis of the TCGA database revealed that 571 candidate genes were significantly upregulated in HCC tissues, while the expression levels of the remaining genes did not show significant downregulation (**Figure 1**, Supplementary Table 2). This could be due to the fact that these candidate genes play important roles in the survival of HCC cell lines.

Next, we constructed gene co-expression modules using WGCNA and filtered for phase-related modules. We utilized mRNA expression of 571 DEGs for WGCNA analysis and selected the

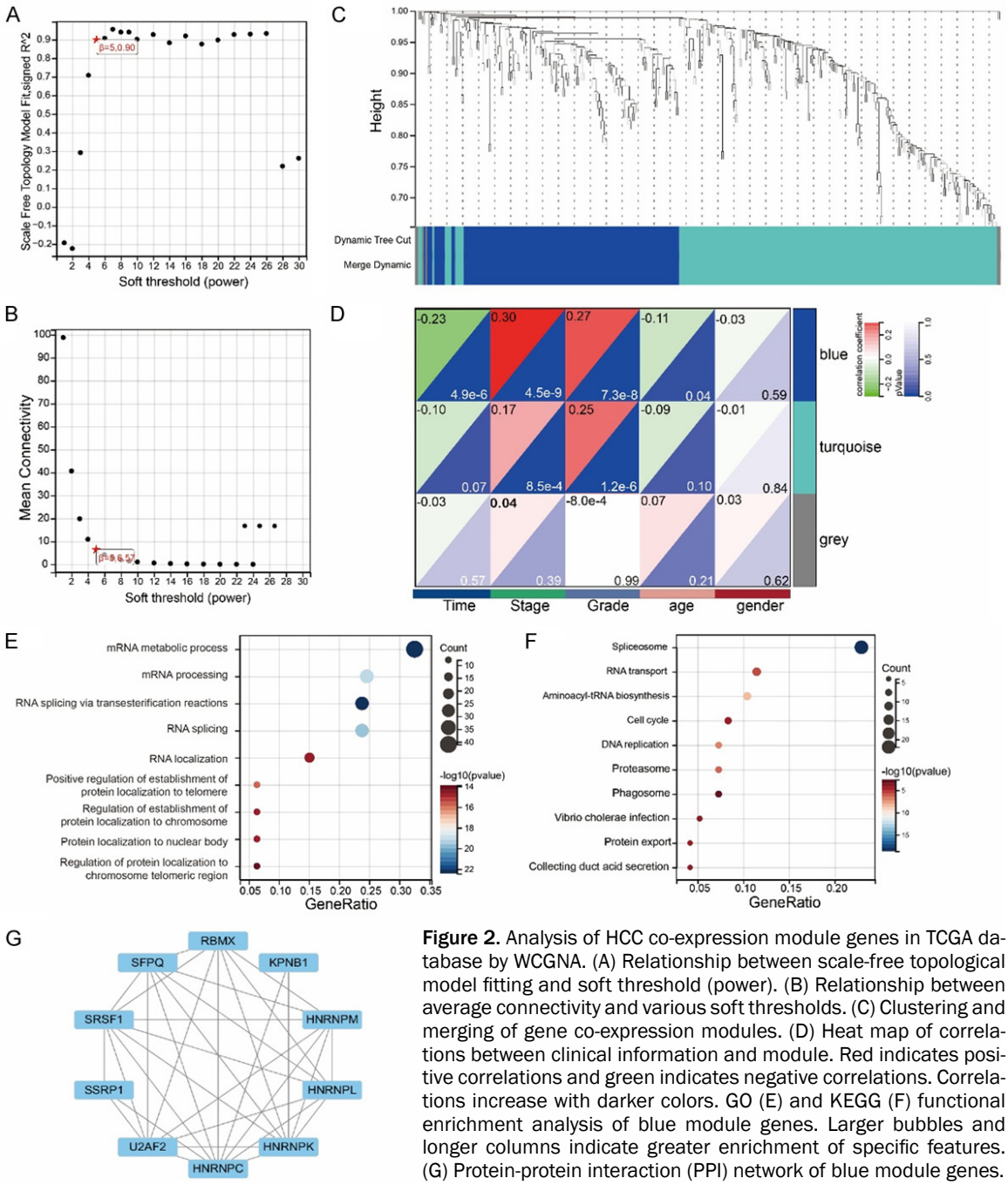


Figure 2. Analysis of HCC co-expression module genes in TCGA database by WCGNA. (A) Relationship between scale-free topological model fitting and soft threshold (power). (B) Relationship between average connectivity and various soft thresholds. (C) Clustering and merging of gene co-expression modules. (D) Heat map of correlations between clinical information and module. Red indicates positive correlations and green indicates negative correlations. Correlations increase with darker colors. GO (E) and KEGG (F) functional enrichment analysis of blue module genes. Larger bubbles and longer columns indicate greater enrichment of specific features. (G) Protein-protein interaction (PPI) network of blue module genes.

optimal β=5 as the soft threshold power to establish a scale-free network (**Figure 2A**, **2B**). The WGCNA analysis divided the 571 DEGs into three modules based on their co-expression patterns. The blue module consisted of 135 genes, the turquoise module consisted of 194 genes, and the grey module consisted of 6 genes (**Figure 2C**, Supplementary Table 3). Based on the WGCNA results, the blue module was found to be positively correlated with tumor stage (R=0.3, P < 0.001; **Figure 2D**). To further investigate the 135 genes in the blue module, we performed GO and KEGG pathway analyses. The top 10 GO enrichment terms included mRNA metabolic process, RNA splicing via transesterification reactions, mRNA processing, RNA splicing, RNA localization, and others (**Figure 2E**). The KEGG pathway analysis revealed that these 135 genes were enriched in the spliceosome, RNA transport, aminoacyl-

tRNA biosynthesis, cell cycle, DNA replication, proteasome, phagosome, vibrio cholerae infection, protein export, and collecting duct acid secretion (**Figure 2F**). Subsequently, we mapped these 135 genes as a PPI network and used Cytoscape software to construct a PPI network and identify 10 hub genes through the MCODE application. The 10 hub genes included RNA binding motif protein X-linked (RBMX), karyopherin subunit Beta 1 (KPNB1), heterogeneous nuclear ribonucleoprotein M (HNRNPM), heterogeneous nuclear ribonucleoprotein C (HNRNPC), heterogeneous nuclear ribonucleoprotein L (HNRNPL), U2 small nuclear RNA auxiliary factor 2 (U2AF2), heterogeneous nuclear ribonucleoprotein K (HNRNPK), structure-specific recognition protein 1 (SSRP1), serine and arginine-rich splicing factor 1 (SRSF1), and splicing factor proline and glutamine rich (SFPQ) (**Figure 2G**).

Construction of a prognostic model according to essential HCC genes

We further evaluated the prognostic value of these 10 genes for HCC patients through a univariate Cox regression analysis of the TCGA dataset. The results revealed that high expression of SFPQ, HNRNPK, HNRNPL, HNRNPM, KPNB1, SRSF1, SSRP1, U2AF2, and RBMX were significantly negatively associated with the overall survival of patients with HCC, while HNRNPC was not significantly correlated with the overall survival of patients with HCC (**Figure 3A**). Next, we used LASSO analysis of the 10-fold cross-validation to analyze these 10 genes (**Figure 3B**), and identified three genes that constructed the risk score model, which included SFPQ, SSRP1, and KPNB1 (**Figure 3C**). The calculation of the risk scores is as follows: Risk score = (0.0288 × expression of SFPQ) + (0.0168 × expression of SSRP1) + (0.0005 × expression of KPNB1). Based on the risk score, each HCC patient was grouped into a low-risk group and a high-risk group. Kaplan-Meier survival rates showed an association between the low-risk group and better prognosis for HCC patients (**Figure 3D**). The results of the ROC analysis demonstrated the powerful predictive potential of risk scores for the prognosis of HCC patients at 1 year (AUC=0.78), 3 years (AUC=0.72), and 5 years (AUC=0.77) (**Figure 3E**). The distribution of mRNA levels of the three genes, risk scores, and survival of HCC patients based on the TCGA dataset were shown in **Figure 3F**.

To investigate the underlying mechanism between the low- and high-risk score samples, we conducted GSEA of the differential expression of the low- and high-risk score samples. The GSEA enrichment analysis revealed that the high-risk score samples were mainly enriched in the cell cycle, oocyte meiosis, ubiquitin-mediated proteolysis, spliceosome, base excision repair, and DNA replication (Supplementary Figure 1A-F).

The relationship between the risk scores and immune characteristics

Furthermore, we investigated the association of risk scores with immune characteristics by analyzing the correlation between risk scores and 24 immune indicators (**Figure 4**). The results demonstrated a positive correlation between the risk score and macrophage M0 cells.

Correlation of risk signature with anti-PD-1 immunotherapy

Immune therapy, represented by immune checkpoint blockade (PD-1/L1 and CTLA-4), has shown excellent clinical efficacy as a major breakthrough in cancer treatment. Previous studies have demonstrated that anti-PD-1 antibody therapy exhibits stronger anti-tumor effects, but only a few patients achieve durable responses. Therefore, in this study, we investigated whether risk score can predict patient response to PD-1 immune checkpoint blockade therapy (GSE78220 cohort). Encouragingly, patients with lower risk scores had better OS than those with higher risk scores (**Figure 5A**). Partial responders had lower risk scores compared to non-responders. Moreover, although not statistically significant, patients with complete immune therapy responses had lower risk scores compared to non-responders, which may be due to the limited number of patients in the cohort (**Figure 5B**). These results strongly suggest that risk features are significantly associated with response to anti-PD-1 immune therapy and may serve as new biomarkers for predicting response to PD-1/L1 immune therapy.

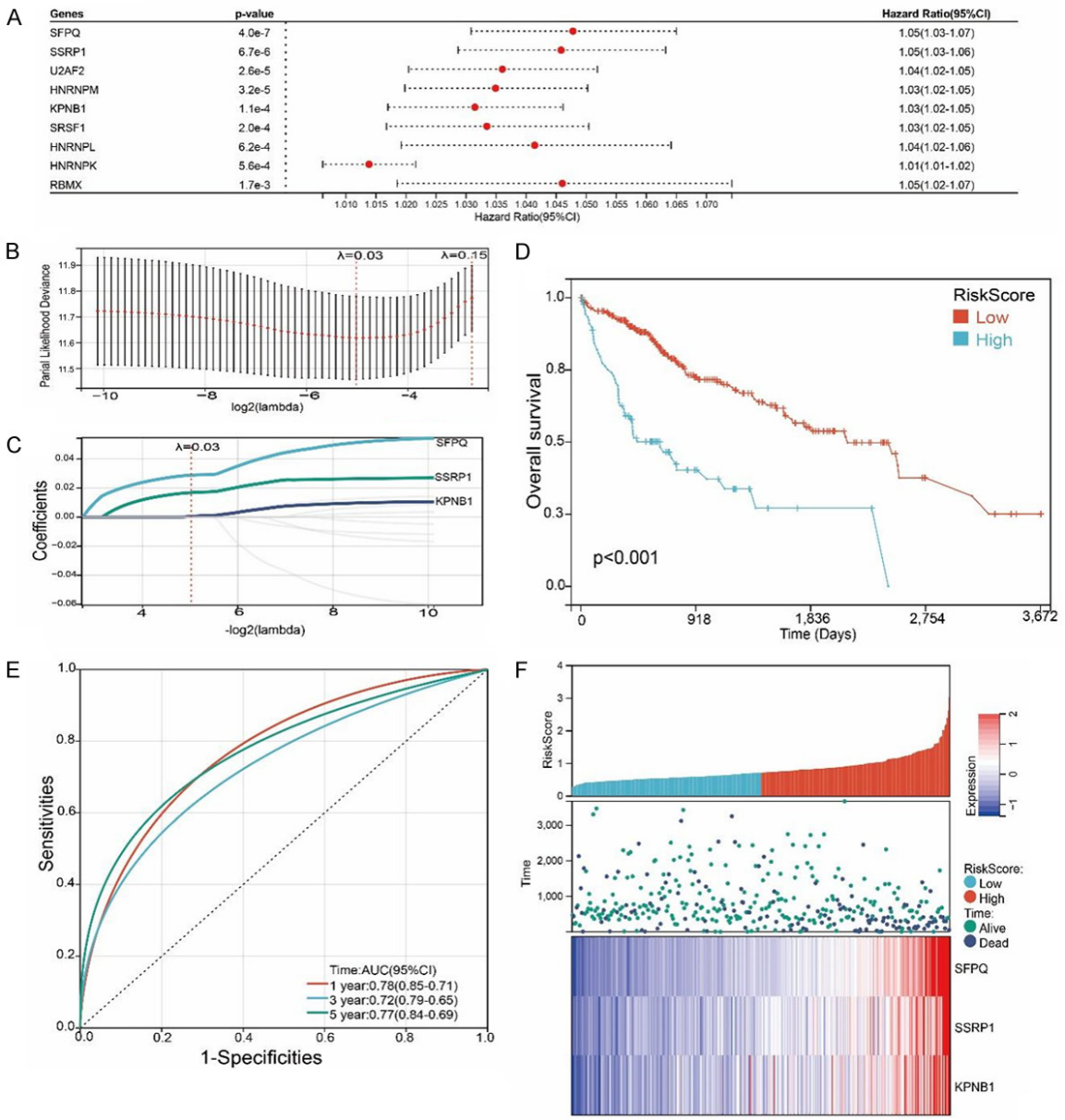


Figure 3. Construction of classifiers to predict prognosis of HCC patients based on 10 genes. A. Univariate Cox survival analysis on the 10 genes. B. Lasso coefficient profile of overall survival with partial likelihood deviation. C. Distribution of lasso coefficients for 10 OS genes. D. Kaplan-Meier comparing the OS of the HCC patients within low-risk and high-risk. E. ROC was implemented to assess the ability of risk scores to differentiate between different prognostic times. F. Distribution of survival status, risk scores, and mRNA expression of SFPQ, SSRP1 and KPNB1 in HCC patients based on the TCGA database.

Validation of the prognostic model

We further divided HCC patients into high-risk and low-risk groups based on different clinical factors, including age, gender, tumor stage, grade, and TNM. The survival analysis showed significantly shorter survival times for HCC patients in the high-risk group compared to those in the low-risk group (Supplementary Figure 2A-N). These findings demonstrate the robust stratification ability of the prognostic risk model. Additionally, we utilized information from the LIRI-JP database and computed each patient's prognostic risk score, classifying them into high and low-risk groups. The Kaplan-Meier curves analysis revealed that low-risk scored

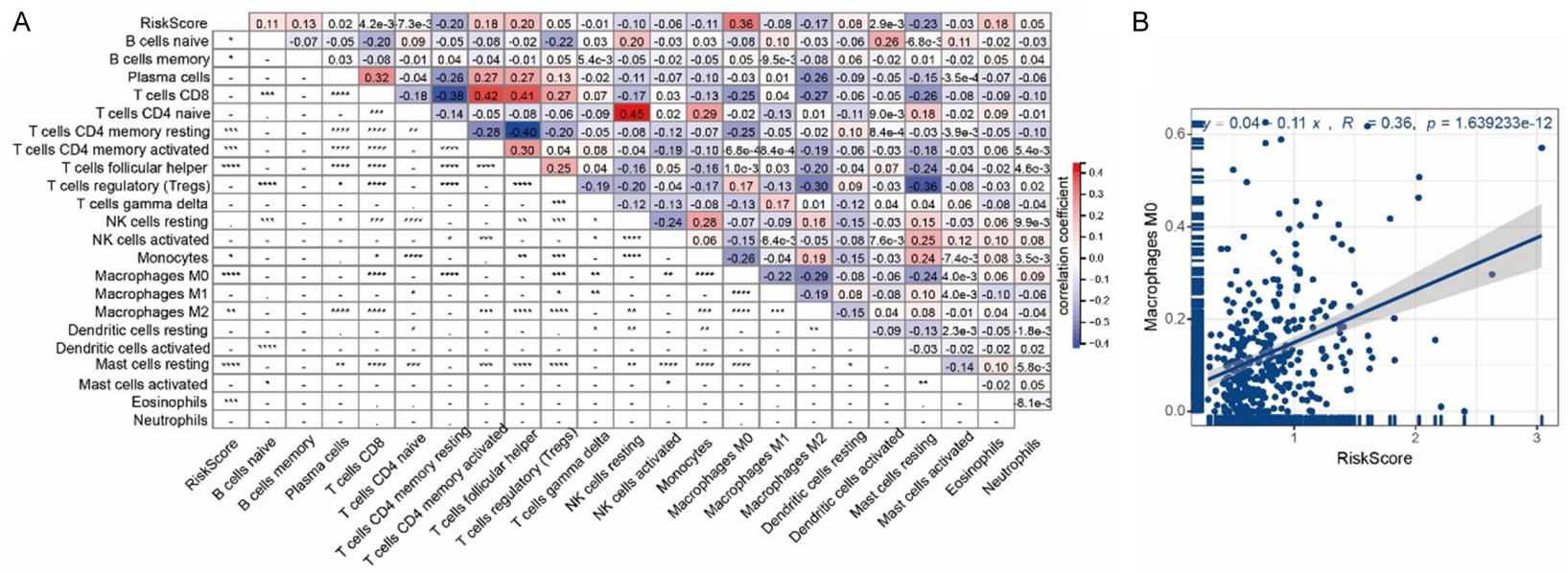


Figure 4. Correlation between the immune indicators and risk scores. A. Heat map of correlation between risk score and 24 immune indicators generated by analyzing HCC patient data from TCGA database. B. Correlation analysis between macrophage M0 and risk score.

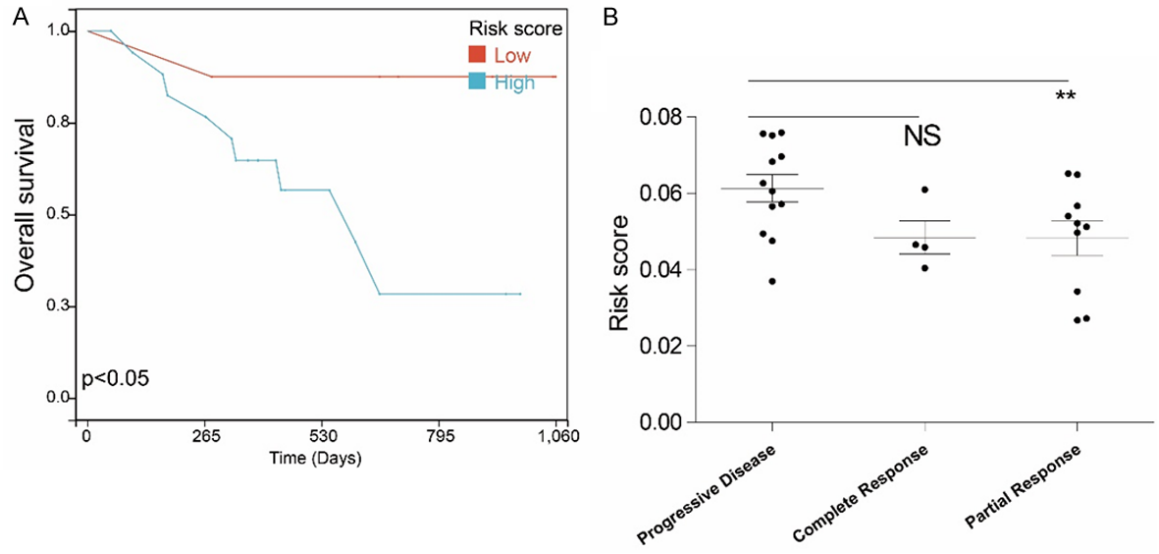


Figure 5. Association of risk score with anti-PD-1 immunotherapy treatment response of GSE78220 cohort. A. Kaplan-Meier analysis of OS in patients treated with anti-PD-1 immunotherapy with different risk scores. B. Comparison of risk scores between patients with complete, partial, and no response to anti-PD-1 immunotherapy. NS, $P > 0.05$; **, $P < 0.01$.

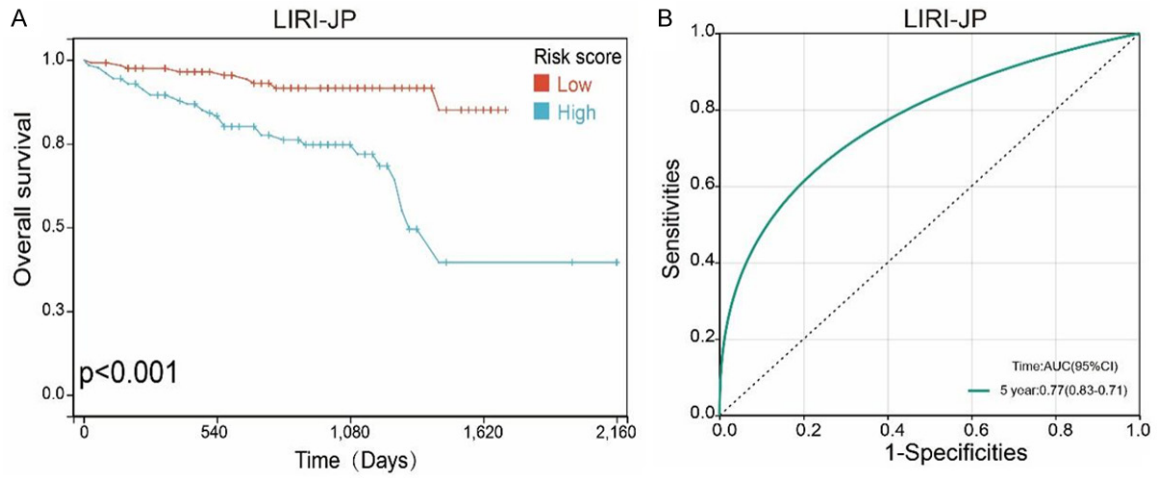


Figure 6. Validation of the prognostic model based on LIRL-JP database. A. Kaplan-Meier analysis of the prognostic model by using the LIRL-JP database. B. Plotting ROC curves to assess the accuracy of the prognostic model based on LIRL-JP database.

patients had a better overall survival than high-risk scored patients. The ROC curve for 5-year overall survival showed excellent predictive value (AUC=0.77), demonstrating the prognostic model's accuracy (**Figure 6A**, **6B**).

Construction and validation of a nomogram model

Subsequently, we developed a nomogram model that included independent predictors of overall survival, such as tumor stage, tumor grade, age, gender, and risk score. The nomogram model provides a visual statistical predictive tool for HCC patients' survival, enabling the risk-scoring model to be applied to clinical practice (**Figure 7A**). Furthermore, the calibration curves demonstrated good consistency between the predicted and observed values of the nomogram (**Figure 7B**).

SFPQ is highly expressed in HCC and is associated with poor prognosis

Previous studies have demonstrated that SSRP1 and KPNB1 genes promote HCC cell

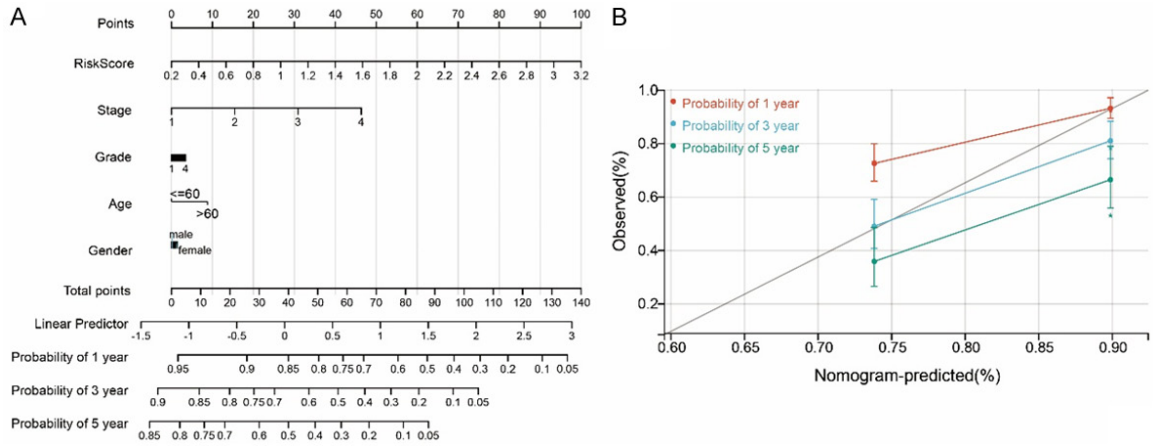


Figure 7. A constructed nomogram for prognostic prediction of the HCC patients. A. Nomogram was used to assess the probability of survival of HCC patients based on risk score. B. Calibration curves for 1-, 3- and 5-year overall survival of HCC patients.

proliferation. However, the expression and effect of SFPQ on HCC are unknown. **Figures 1** and **3** have shown the high expression of SFPQ in liver cancer and its correlation with poor prognosis in liver cancer patients. Based on the TCGA database, we further analyzed the expression of SFPQ in different stages and in HCC with or without metastasis. The results revealed that the expression of SFPQ was lowest in Stage I and highest in Stage III&IV (**Figure 8A**). There was no significant difference in SFPQ expression between HCC tissues with and without metastasis (**Figure 8B**). Additionally, RT-PCR was used to detect the expression of SFPQ in HCC tissues and adjacent non-cancerous tissues from 24 patients. The results demonstrated that the expression of SFPQ was significantly upregulated in HCC tissues (**Figure 8C**). Furthermore, survival analysis revealed that high expression of SFPQ was a poor prognostic factor in 66 HCC patients (**Figure 8D**).

Knockdown SFPQ inhibits the HCC cells proliferation, migration and invasion

To investigate the functional roles of SFPQ in HCC cell proliferation, we knocked down SFPQ using SFPQ-shRNAs (**Figure 9A**, **9B**). Subsequently, CCK-8 experiments were performed to evaluate the effects of SFPQ inhibition on the proliferation, migration, and invasion of Huh-7 and Hep3B cells (**Figure 9C-H**). The results demonstrated that SFPQ inhibition significantly reduced the proliferation, migration, and invasion of Huh-7 and Hep3B cells (**Figure 9C**, **9D**).

Discussion

In this study, we identified genes critical for HCC cell proliferation and survival from the DepMap project using the CERES algorithm for CRISPR screening to calculate dependency scores. Among these 692 essential genes, we identified 584 DEGs in HCC tissues using the TCGA database. Using WGCNA, we categorized these 584 genes into three modules. The blue module consisted of 135 genes and was positively associated with tumor stage. These 135 genes were mainly enriched in spliceosome, RNA transport, and cell cycle pathways.

We utilized the MCODE approach in Cytoscape to identify 10 hub genes, including SFPQ, HNRNPK, HNRNPL, HNRNPC, HNRNPM, KPNB1, SRSF1, SSRP1, U2AF2, and RBMX, in the PPI network. Using Cox univariate analysis and Lasso, we developed a prognostic model that included three genes (SFPQ, SSRP1, and KPNB1). To further assess the reliability of the prognostic risk model, we performed subgroup analysis and external validation. The AUC of the ROC curves of the prognostic model was greater than 0.7, indicating that the prognostic model had good performance in predicting the prognosis of patients with HCC.

We next explored the potential mechanisms underlying the impact of the risk score on prognosis. GSEA analysis revealed significant enrichment of signaling pathways, such as cell cycle, oocyte meiosis, ubiquitin-mediated proteolysis, and spliceosome, in the high-risk

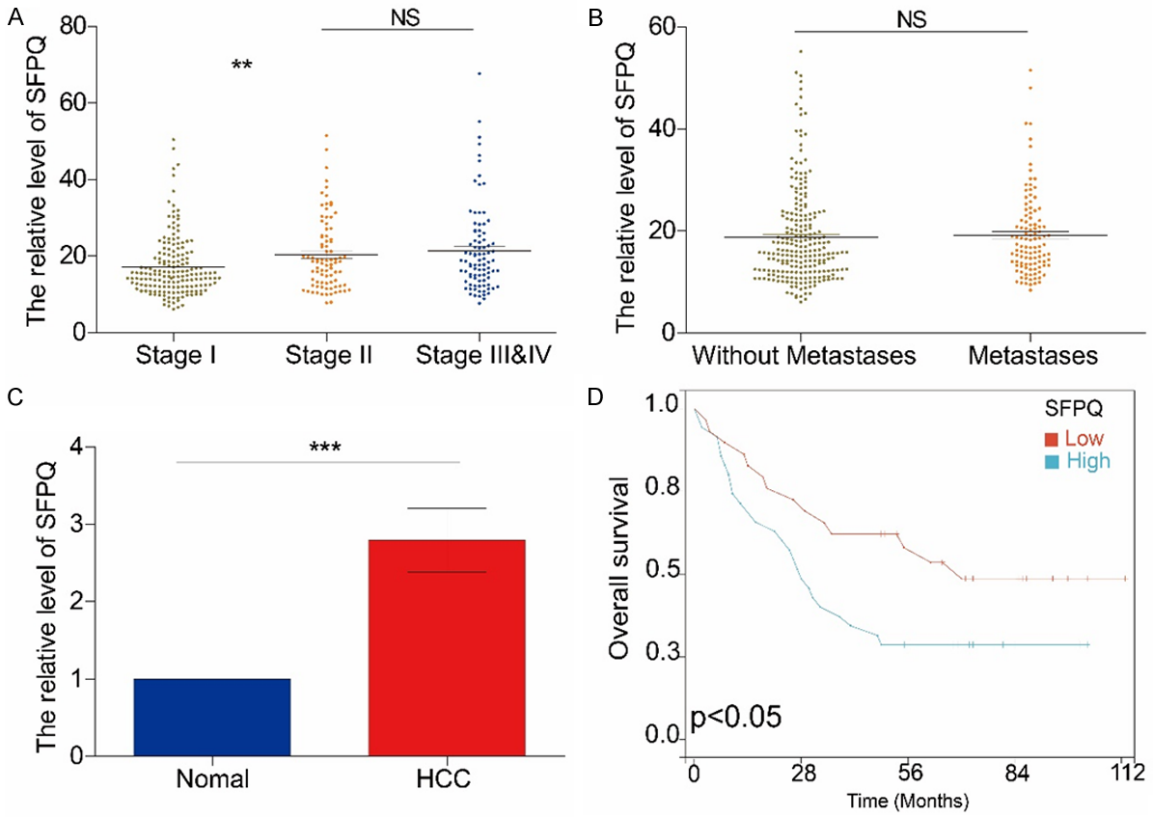


Figure 8. Expression and prognostic analysis of SFPQ in HCC. A. Expression analysis of SFPQ in different stages of HCC based on TCGA database. B. Expression analysis of SFPQ in HCC with/without metastasis based on TCGA database. C. Differential expression analysis of SFPQ in HCC tissue and adjacent normal tissue based on HCC cDNA microarray. D. Prognostic analysis of SFPQ in HCC based on HCC cDNA microarray. NS, $P > 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

group. All of these biological processes have been found to be involved in the progression of HCC [24], providing evidence for the rationality of risk score stratification in HCC patients and molecular hypotheses.

Additionally, we further analyzed the tumor-infiltrating immune cell landscape in patients with different risk score groups. We found that the risk score had a positive association with the abundance of macrophages M0 cells. Macrophages M0 are commonly referred to as non-activated macrophages, which together with the M1 and M2 phenotypes constitute tumor-associated macrophages. The altered tumor microenvironment, such as hypoxia, inflammation, chemicals released by tumor cells, and exacerbation of inflammation, may promote the accumulation of macrophages [25]. Macrophages have been reported to be recruited to tumor tissues and act as pro-angiogenic cells, which is significantly associated with poor overall survival and disease-free survival in HCC [25]. Therefore, the varied types of tumor-infiltrating immune cells present in patients with varying risk score groups likely play a role in determining their distinct prognoses. Nonetheless, additional research is necessary to uncover the specific underlying mechanisms involved.

As a significant breakthrough in cancer treatment, immunotherapy represented by immune checkpoint blockade (PD-1/L1 and CTLA-4) has shown promising clinical efficacy. Early studies have demonstrated that anti-PD-1 antibodies display good clinical efficacy. However, only a small number of patients are able to achieve a durable response, so identifying suitable HCC patients for PD-1 immune therapy may be urgent and clinically meaningful [32]. Encouragingly, in this study, we found that the risk score is significantly associated with the response to PD-1 immune therapy. A decrease

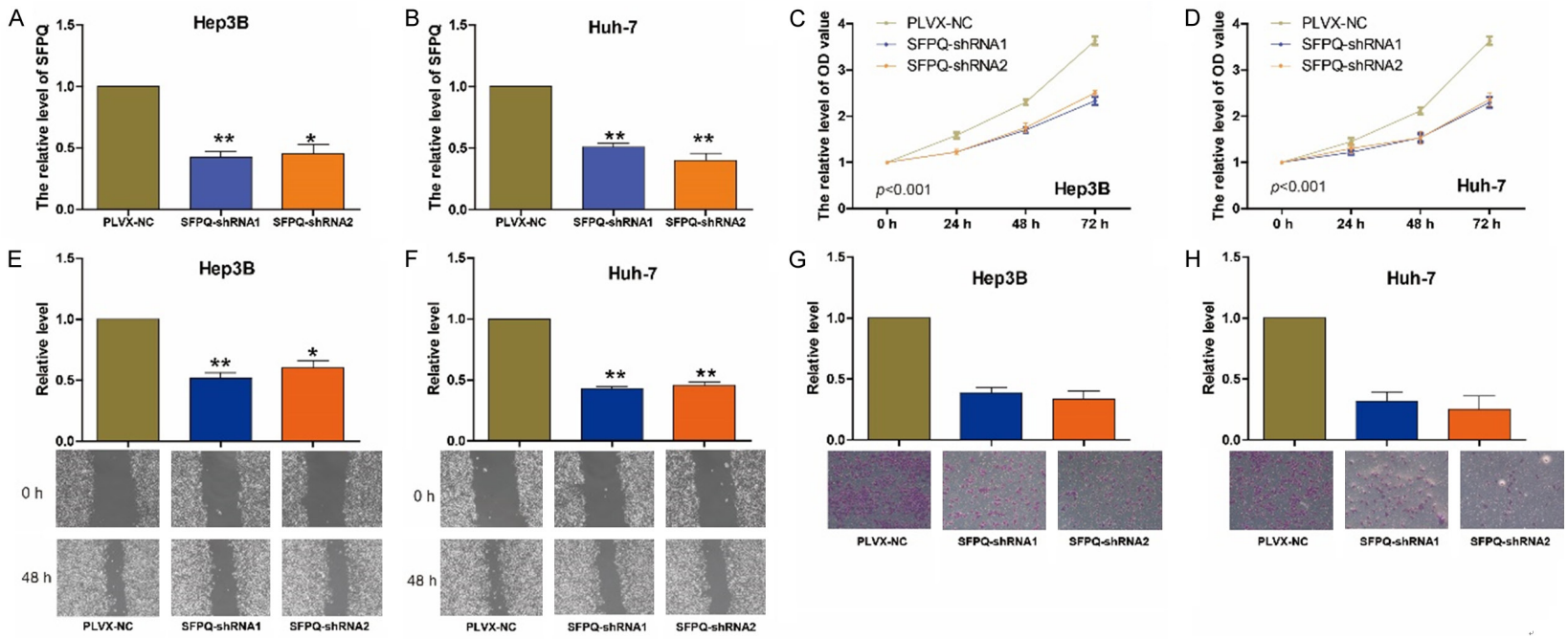


Figure 9. SFPQ knockdown inhibits the proliferation of HCC cells. (A, B) The SFPQ expression in the HCC (A) Hep3B and (B) Huh-7 cells transfected with SFPQ-shRNAs. (C, D) The proliferation of Hep3B (C) and Huh-7 (D) cells transfected with SFPQ-shRNAs was detected by CCK8. (E, F) The migrative of Hep3B (E) and Huh-7 (F) cells transfected with SFPQ-shRNAs was detected by wound healing assay. (G, H) The invasive of Hep3B (G) and Huh-7 (H) cells transfected with SFPQ-shRNAs was detected by transwell assay. *, $P < 0.05$; **, $P < 0.01$.

in the risk score was found in responders to PD-1 immune therapy, and patients with a lower risk score had better OS than those with a higher risk score. This strongly suggests that the risk score may be used as a new biomarker for predicting the response to PD-1 immune therapy.

The prognostic signature was developed using three genes, namely SFPQ, SSRP1, and KPNB1. Among these, SSRP1 and KPNB1 have been previously found to be related to HCC [26, 27].

SSRP1 is one of the subunits of the facilitated chromatin transcription complex and is involved in processes such as transcriptional regulation, DNA replication, cell cycyle and DNA repair [28-30]. Studies have indicated that up-regulated SSRP1 is related to poor pathological features and poor prognosis in a variety of tumor, including bladder cancer [31], breast cancer, colorectal cancer [32], and HCC. In HCC, downregulation of SSRP1 was found to cause cell cycle arrest in G0/G1 phase and promote apoptosis, and elevated SSRP1 decreased cell sensitivity to chemotherapeutic drugs [27, 33]. MiR-497, a post-transcriptional regulator of SSRP1, was found to effectively suppress SSRP1 expression [27]. Silencing of SSRP1 in colorectal cancer activated AKT pathway and inhibited proliferation, migration, and invasion [34]. In CRC, SSRP1 was upregulated and related to Dukes stage of tumors, and miR-28-5p was found to be another upstream target of SSRP1 [32]. Altered expression of SSRP1 in tumor leads to changes in EMT-related signaling molecules that contribute to tumor progression [35, 36].

KPNB1 is a soluble nuclear transport receptor that transports proteins or RNA out or into of the nucleus by binding directly to proteins via the non-classical transport pathway or forming KPNA/importin β1/protein trimers via the classical transport pathway [37-39]. This nuclear transport system has been found to be related to cell cycle and mitosis, and therefore its dysregulation is associate with tumor genesis and tumor progression [40, 41]. Elevated level of KPNB1 has been found in several types of tumors, including liver cancer [42], lung cancer [43, 44], gastric cancer [45], cervical cancer [38, 46], and prostate cancer [47]. In HCC patients, the expression of KPNB1 was related to tumor histological differentiation, and KPNB1 was found to be involved in regulating the HCC cells proliferation of HCC [42, 48]. Knockdown of KPNβ1 expression caused IκBα to be downregulated in the cytoplasm, and NF-κB, which originally formed a complex, was translocated into the nucleus and induced apoptosis [48]. Similarly in prostate cancer, targeting KPNB1 was found to significantly inhibit the NF-κB' nuclear translocation, inducing reduction in apoptosis and tumorsphere formation in cancer cells [47]. In neuropathic pain, miR-101 inhibited the activation of NF-κB by targeting KPNB1 mRNA to control the nuclear translocation of p65 [49]. In colorectal cancer, KPNB1 knockdown significantly reduced the proliferation and invasion of CRC cells by EMT related signal expression [50]. At present, there is no study report on the effect of SFPQ on the HCC cells proliferation. In the study, knockdown SFPQ could inhibit the HCC cells proliferation, migration, and invasion.

Although risk score has shown good performance in predicting the prognosis of HCC, there are still some limitations that need to be addressed. Firstly, although the prognostic value of risk score has been validated in external cohorts, an independent cohort consisting of more HCC patients is needed to further validate the model. Secondly, we did not directly analyze the potential mechanisms of risk score in the development and progression of HCC. In vitro studies are needed to further validate the potential mechanisms revealed by bioinformatics.

Conclusion

The current study focused on the identification core essential genes that are important in the proliferation and survival in HCC using DepMap database. Three core genes including SFPQ, SSRP1 and KPNB1 were identified through WGCNA, univariate Cox analysis and LASSO analysis were to construct a prognostic model. In addition, konckdown SFPQ could inhibit the HCC cells proliferation, migration, and invasion.

Acknowledgements

This research was funded by the fund of National Science Foundation for Young Scientists of China (82003095), Natural Science Foundation of Fujian Province (2020J01311402

& 2022J01273), and National Science Foundation of China (81973924).

Disclosure of conflict of interest

None.

Address correspondence to: Qingshui Wang and Yao Lin, The Second Affiliated Hospital of Fujian Traditional Chinese Medical University, Fujian-Macao Science and Technology Cooperation Base of Traditional Chinese Medicine-Oriented Chronic Disease Prevention and Treatment, Innovation and Transformation Center, Fujian University of Traditional Chinese Medicine, Fuzhou, Fujian, China. E-mail: wangqingshui@fjnu.edu.cn (QSW); yaolin@fjtcm.edu.cn (YL); Shaoqing Chen, Fujian University of Traditional Chinese Medicine, Fuzhou, Fujian, China. E-mail: chensq@fjtcm.edu.cn; Wei Zhuang, Department of Urology, The Second Affiliated Hospital of Fujian Medical University, Quanzhou, Fujian, China. E-mail: 612032@fjmu.edu.cn

References

[1] Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A and Bray F. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2021; 71: 209-249.

[2] Kim DW, Talati C and Kim R. Hepatocellular carcinoma (HCC): beyond sorafenib-chemotherapy. J Gastrointest Oncol 2017; 8: 256-265.

[3] Nia A and Dhanasekaran R. Genomic landscape of HCC. Curr Hepatol Rep 2020; 19: 448-461.

[4] Samant H, Amiri HS and Zibari GB. Addressing the worldwide hepatocellular carcinoma: epidemiology, prevention and management. J Gastrointest Oncol 2021; 12 Suppl 2: S361-S373.

[5] Friedman AA, Letai A, Fisher DE and Flaherty KT. Precision medicine for cancer with next-generation functional diagnostics. Nat Rev Cancer 2015; 15: 747-756.

[6] Ghandi M, Huang FW, Jané-Valbuena J, Kryukov GV, Lo CC, McDonald ER 3rd, Barretina J, Gelfand ET, Bielski CM, Li H, Hu K, Andreev-Drakhlin AY, Kim J, Hess JM, Haas BJ, Aguet F, Weir BA, Rothberg MV, Paolella BR, Lawrence MS, Akbani R, Lu Y, Tiv HL, Gokhale PC, de Weck A, Mansour AA, Oh C, Shih J, Hadi K, Rosen Y, Bistline J, Venkatesan K, Reddy A, Sonkin D, Liu M, Lehar J, Korn JM, Porter DA, Jones MD, Golji J, Caponigro G, Taylor JE, Dunning CM, Creech AL, Warren AC, McFarland JM, Zamanighomi M, Kauffmann A, Stransky N, Imielinski M, Maruvka YE, Cherniack AD, Tsherniak A, Vazquez F, Jaffe JD, Lane AA, Weinstock DM, Johannessen CM, Morrissey MP, Stegmeier F, Schlegel R, Hahn WC, Getz G, Mills GB, Boehm JS, Golub TR, Garraway LA and Sellers WR. Next-generation characterization of the cancer cell line encyclopedia. Nature 2019; 569: 503-508.

[7] Shimada K, Bachman JA, Muhlich JL and Mitchison TJ. shinyDepMap, a tool to identify targetable cancer genes and their functional connections from cancer dependency map data. Elife 2021; 10: e57116.

[8] Tsherniak A, Vazquez F, Montgomery PG, Weir BA, Kryukov G, Cowley GS, Gill S, Harrington WF, Pantel S, Krill-Burger JM, Meyers RM, Ali L, Goodale A, Lee Y, Jiang G, Hsiao J, Gerath WFJ, Howell S, Merkel E, Ghandi M, Garraway LA, Root DE, Golub TR, Boehm JS and Hahn WC. Defining a cancer dependency map. Cell 2017; 170: 564-576, e16.

[9] Meyers RM, Bryan JG, McFarland JM, Weir BA, Sizemore AE, Xu H, Dharia NV, Montgomery PG, Cowley GS, Pantel S, Goodale A, Lee Y, Ali LD, Jiang G, Lubonja R, Harrington WF, Strickland M, Wu T, Hawes DC, Zhivich VA, Wyatt MR, Kalani Z, Chang JJ, Okamoto M, Stegmaier K, Golub TR, Boehm JS, Vazquez F, Root DE, Hahn WC and Tsherniak A. Computational correction of copy number effect improves specificity of CRISPR-Cas9 essentiality screens in cancer cells. Nat Genet 2017; 49: 1779-1784.

[10] Kumar P, Soni DK, Sen C, Larsen MB, Mazan-Mamczarz K, Piao Y, De S, Gorospe M, Frizzell RA and Biswas R. SFPQ rescues F508del-CFTR expression and function in cystic fibrosis bronchial epithelial cells. Sci Rep 2021; 11: 16645.

[11] Imamura K, Imamachi N, Akizuki G, Kumakura M, Kawaguchi A, Nagata K, Kato A, Kawaguchi Y, Sato H, Yoneda M, Kai C, Yada T, Suzuki Y, Yamada T, Ozawa T, Kaneki K, Inoue T, Kobayashi M, Kodama T, Wada Y, Sekimizu K and Akimitsu N. Long noncoding RNA NEAT1-dependent SFPQ relocation from promoter region to paraspeckle mediates IL8 expression upon immune stimuli. Mol Cell 2014; 53: 393-406.

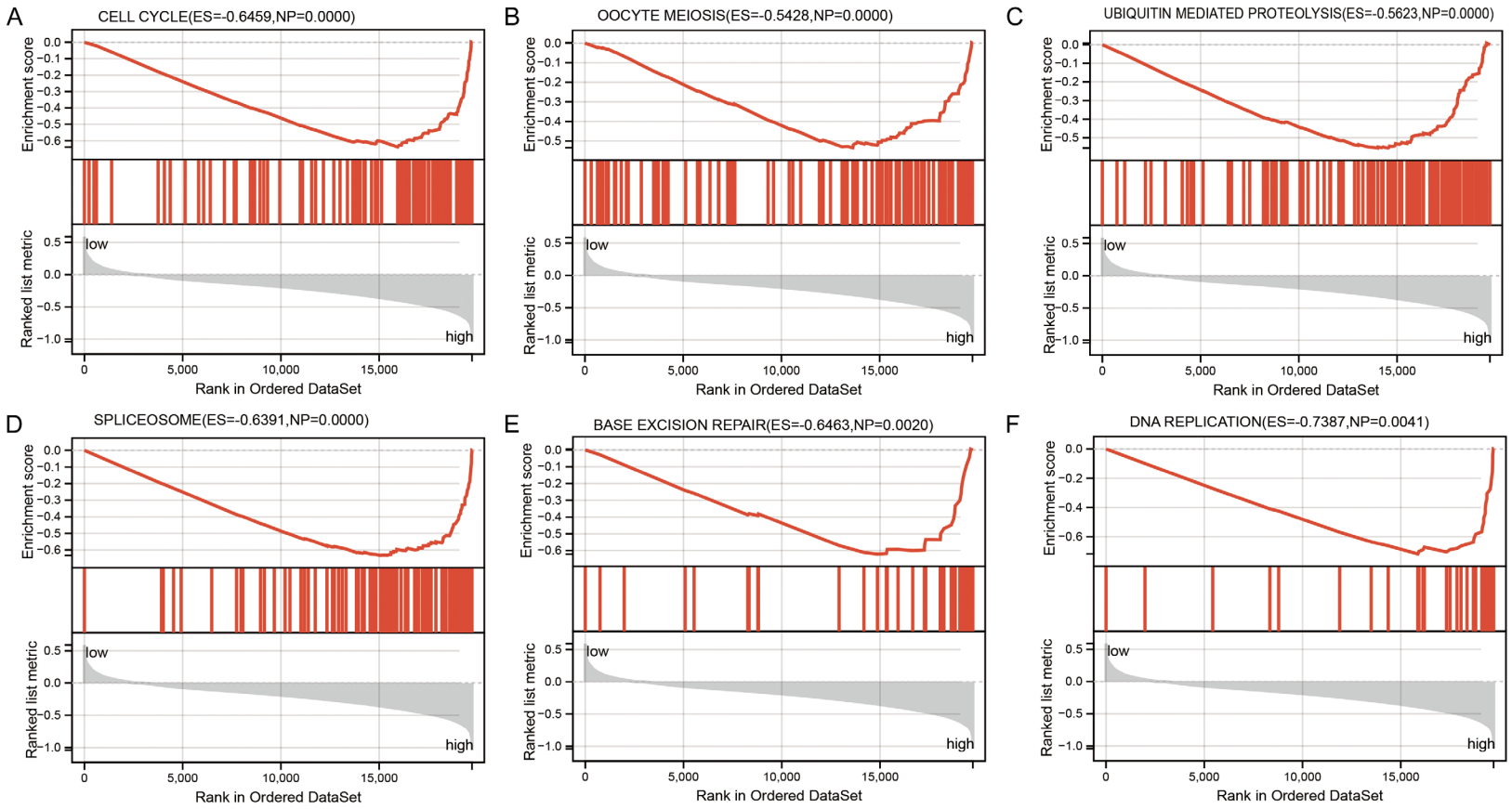
[12] Stagsted LVW, O'Leary ET, Ebbesen KK and Hansen TB. The RNA-binding protein SFPQ preserves long-intron splicing and regulates circRNA biogenesis in mammals. Elife 2021; 10: e63088.

[13] Lim YW, James D, Huang J and Lee M. The emerging role of the RNA-binding protein SFPQ in neuronal function and neurodegeneration. Int J Mol Sci 2020; 21: 7151.

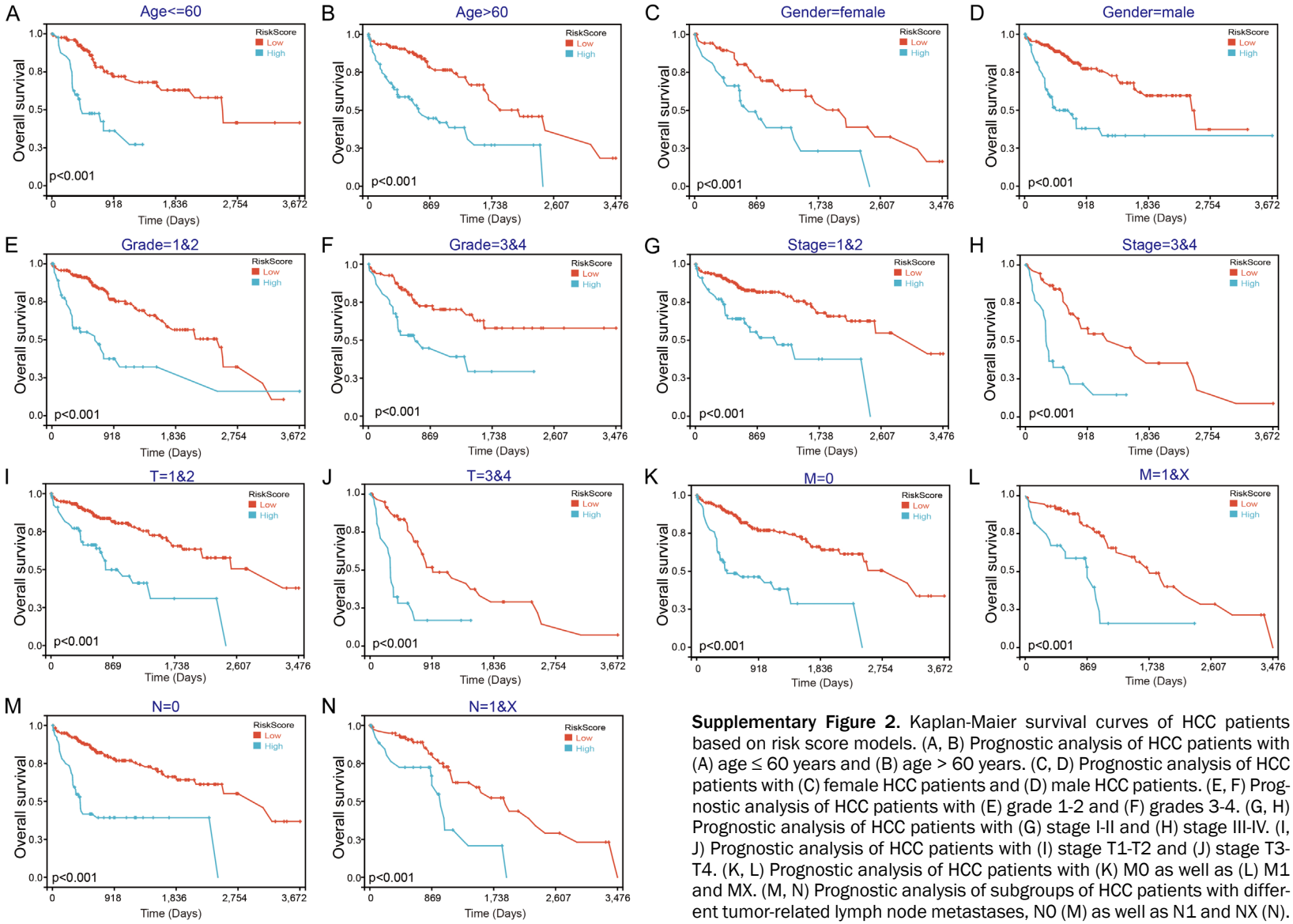
[14] Fukuda Y, Pazyra-Murphy MF, Silagi ES, Tasdemir-Yilmaz OE, Li Y, Rose L, Yeoh ZC, Vangos NE, Geffken EA, Seo HS, Adelmant G, Bird GH,

Walensky LD, Marto JA, Dhe-Paganon S and Segal RA. Binding and transport of SFPQ-RNA granules by KIF5A/KLC1 motors promotes axon survival. J Cell Biol 2021; 220: e202005051.
[15] Younas N, Zafar S, Shafiq M, Noor A, Siegert A, Arora AS, Galkin A, Zafar A, Schmitz M, Stadelmann C, Andreoletti O, Ferrer I and Zerr I. SFPQ and Tau: critical factors contributing to rapid progression of Alzheimer's disease. Acta Neuropathol 2020; 140: 317-339.
[16] Meng Y, Li S, Zhang Q, Ben S, Zhu Q, Du M and Gu D. LncRNA-422 suppresses the proliferation and growth of colorectal cancer cells by targeting SFPQ. Clin Transl Med 2022; 12: e664.
[17] Bi O, Anene CA, Nsengimana J, Shelton M, Roberts W, Newton-Bishop J and Boyne JR. SFPQ promotes an oncogenic transcriptomic state in melanoma. Oncogene 2021; 40: 5192-5203.
[18] Barretina J, Caponigro G, Stransky N, Venkatesan K, Margolin AA, Kim S, Wilson CJ, Lehár J, Kryukov GV, Sonkin D, Reddy A, Liu M, Murray L, Berger MF, Monahan JE, Morais P, Meltzer J, Korejwa A, Jané-Valbuena J, Mapa FA, Thibault J, Bric-Furlong E, Raman P, Shipway A, Engels IH, Cheng J, Yu GK, Yu J, Aspesi P Jr, de Silva M, Jagtap K, Jones MD, Wang L, Hatton C, Palescandolo E, Gupta S, Mahan S, Sougnez C, Onofrio RC, Liefeld T, MacConaill L, Winckler W, Reich M, Li N, Mesirov JP, Gabriel SB, Getz G, Ardlie K, Chan V, Myer VE, Weber BL, Porter J, Warmuth M, Finan P, Harris JL, Meyerson M, Golub TR, Morrissey MP, Sellers WR, Schlegel R and Garraway LA. The cancer cell line encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature 2012; 483: 603-607.
[19] Klann TS, Black JB, Chellappan M, Safi A, Song L, Hilton IB, Crawford GE, Reddy TE and Gersbach CA. CRISPR-Cas9 epigenome editing enables high-throughput screening for functional regulatory elements in the human genome. Nat Biotechnol 2017; 35: 561-568.
[20] Ma X, Liu J, Wang H, Jiang Y, Wan Y, Xia Y and Cheng W. Identification of crucial aberrantly methylated and differentially expressed genes related to cervical cancer using an integrated bioinformatics analysis. Biosci Rep 2020; 40: BSR20194365.
[21] Cheng Y, Liu C, Liu Y, Su Y, Wang S, Jin L, Wan Q, Liu Y, Li C, Sang X, Yang L, Liu C, Wang X and Wang Z. Immune microenvironment related competitive endogenous RNA network as powerful predictors for melanoma prognosis based on WGCNA analysis. Front Oncol 2020; 10: 577072.
[22] Kim SM, Kim Y, Jeong K, Jeong H and Kim J. Logistic LASSO regression for the diagnosis of breast cancer using clinical demographic data and the BI-RADS lexicon for ultrasonography. Ultrasonography 2018; 37: 36-42.
[23] Wang S, Yang L, Ci B, Maclean M, Gerber DE, Xiao G and Xie Y. Development and validation of a nomogram prognostic model for SCLC patients. J Thorac Oncol 2018; 13: 1338-1348.
[24] Wu T and Dai Y. Tumor microenvironment and therapeutic response. Cancer Lett 2017; 387: 61-68.
[25] Zhang Z, Wang Z and Huang Y. Comprehensive analyses of the infiltrating immune cell landscape and its clinical significance in hepatocellular carcinoma. Int J Gen Med 2021; 14: 4695-4704.
[26] Kojima Y, Nishina T, Nakano H, Okumura K and Takeda K. Inhibition of importin β1 augments the anticancer effect of agonistic anti-death receptor 5 antibody in TRAIL-resistant tumor cells. Mol Cancer Ther 2020; 19: 1123-1133.
[27] Ding Q, He K, Luo T, Deng Y, Wang H, Liu H, Zhang J, Chen K, Xiao J, Duan X, Huang R, Xia Z, Zhou W, He J, Yu H, Jiao X and Xiang G. SSRP1 contributes to the malignancy of hepatocellular carcinoma and is negatively regulated by miR-497. Mol Ther 2016; 24: 903-914.
[28] Gao Y, Li C, Wei L, Teng Y, Nakajima S, Chen X, Xu J, Leger B, Ma H, Spagnol ST, Wan Y, Dahl KN, Liu Y, Levine AS and Lan L. SSRP1 cooperates with PARP and XRCC1 to facilitate single-strand DNA break repair by chromatin priming. Cancer Res 2017; 77: 2674-2685.
[29] Marcianò G, Da Vela S, Tria G, Svergun DI, Byron O and Huang DT. Structure-specific recognition protein-1 (SSRP1) is an elongated homodimer that binds histones. J Biol Chem 2018; 293: 10071-10083.
[30] Falbo L, Raspelli E, Romeo F, Fiorani S, Pezzimenti F, Casagrande F, Costa I, Parazzoli D and Costanzo V. SSRP1-mediated histone H1 eviction promotes replication origin assembly and accelerated development. Nat Commun 2020; 11: 1345.
[31] Song H, Zeng J, Lele S, LaGrange CA and Bhakat KK. APE1 and SSRP1 is overexpressed in muscle invasive bladder cancer and associated with poor survival. Heliyon 2021; 7: e06756.
[32] Wu W, He K, Guo Q, Chen J, Zhang M, Huang K, Yang D, Wu L, Deng Y, Luo X, Yu H, Ding Q and Xiang G. SSRP1 promotes colorectal cancer progression and is negatively regulated by miR-28-5p. J Cell Mol Med 2019; 23: 3118-3129.
[33] Luo G, Xu J, Xia Z, Liu S, Liu H, He K and Xiang G. SSRP1 is a prognostic biomarker correlated with CD8(+) T cell infiltration in hepatocellular carcinoma (HCC). Biomed Res Int 2021; 2021: 9409836.

[34] Wang Q, Jia S, Jiao Y, Xu L, Wang D, Chen X, Hu X, Liang H, Wen N, Zhang S, Guo B and Zhang L. SSRP1 influences colorectal cancer cell growth and apoptosis via the AKT pathway. Int J Med Sci 2019; 16: 1573-1582.
[35] Gao L and Xiong X. MiR-223 inhibits the proliferation, invasion and EMT of nasopharyngeal carcinoma cells by targeting SSRP1. Int J Clin Exp Pathol 2018; 11: 4374-4384.
[36] Yu Y, Gao Y and Yu Y. SSRP1 worsens malignant melanoma progression by activating MAPKs pathway. Ann Clin Lab Sci 2021; 51: 783-789.
[37] Ha S, Jeong J, Oh J, Rhee S and Ham SW. A small organic molecule blocks EGFR transport into the nucleus by the nonclassical pathway resulting in repression of cancer invasion. Chembiochem 2018; 19: 131-135.
[38] Stelma T and Leaner VD. KPNB1-mediated nuclear import is required for motility and inflammatory transcription factor activity in cervical cancer cells. Oncotarget 2017; 8: 32833-32847.
[39] Wang T, Huang Z, Huang N, Peng Y, Gao M, Wang X and Feng W. Inhibition of KPNB1 inhibits proliferation and promotes apoptosis of chronic myeloid leukemia cells through regulation of E2F1. Onco Targets Ther 2019; 12: 10455-10467.
[40] Chi RA, van der Watt P, Wei W, Birrer MJ and Leaner VD. Inhibition of Kpnβ1 mediated nuclear import enhances cisplatin chemosensitivity in cervical cancer. BMC Cancer 2021; 21: 106.
[41] Ajayi-Smith A, van der Watt P, Mkwanazi N, Carden S, Trent JO and Leaner VD. Novel small molecule inhibitor of Kpnβ1 induces cell cycle arrest and apoptosis in cancer cells. Exp Cell Res 2021; 404: 112637.
[42] Zhuang W, Sun H, Zhang S, Zhou Y, Weng W, Wu B, Ye T, Huang W, Lin Z, Shi L and Shi K. An immunogenomic signature for molecular classification in hepatocellular carcinoma. Mol Ther Nucleic Acids 2021; 25: 105-115.
[43] Du W, Zhu J, Zeng Y, Liu T, Zhang Y, Cai T, Fu Y, Zhang W, Zhang R, Liu Z and Huang JA. KPNB1-mediated nuclear translocation of PD-L1 promotes non-small cell lung cancer cell proliferation via the Gas6/MerTK signaling pathway. Cell Death Differ 2021; 28: 1284-1300.
[44] Sekimoto N, Suzuki Y and Sugano S. Decreased KPNB1 expression is induced by PLK1 inhibition and leads to apoptosis in lung adenocarcinoma. J Cancer 2017; 8: 4125-4140.
[45] Zhu J, Wang Y, Huang H, Yang Q, Cai J, Wang Q, Gu X, Xu P, Zhang S, Li M, Ding H and Yang L. Upregulation of KPNβ1 in gastric cancer cell promotes tumor cell proliferation and predicts poor prognosis. Tumour Biol 2016; 37: 661-672.
[46] Kodama M, Kodama T, Newberg JY, Katayama H, Kobayashi M, Hanash SM, Yoshihara K, Wei Z, Tien JC, Rangel R, Hashimoto K, Mabuchi S, Sawada K, Kimura T, Copeland NG and Jenkins NA. In vivo loss-of-function screens identify KPNB1 as a new druggable oncogene in epithelial ovarian cancer. Proc Natl Acad Sci U S A 2017; 114: E7301-E7310.
[47] Yang J, Guo Y, Lu C, Zhang R, Wang Y, Luo L, Zhang Y, Chu CH, Wang KJ, Obbad S, Yan W and Li X. Inhibition of Karyopherin beta 1 suppresses prostate cancer growth. Oncogene 2019; 38: 4700-4714.
[48] Yang L, Hu B, Zhang Y, Qiang S, Cai J, Huang W, Gong C, Zhang T, Zhang S, Xu P, Wu X and Liu J. Suppression of the nuclear transporter-KPNβ1 expression inhibits tumor proliferation in hepatocellular carcinoma. Med Oncol 2015; 32: 128.
[49] Liu JC, Xue DF, Wang XQ, Ai DB and Qin PJ. MiR-101 relates to chronic peripheral neuropathic pain through targeting KPNB1 and regulating NF-κB signaling. Kaohsiung J Med Sci 2019; 35: 139-145.
[50] Zhang Y and Li KF. Karyopherin β1 deletion suppresses tumor growth and metastasis in colorectal cancer (CRC) by reducing MET expression. Biomed Pharmacother 2019; 120: 109127.



Supplementary Figure 1. GSEA analysis. Patients with HCC were divided into two groups based on the risk score using mean value as a cutoff. (A-F) GSEA was carried out using the TCGA database, high-risk score group enriched in cell cycle (A), oocyte meiosis (B), ubiquitin-mediated proteolysis (C), spliceosome (D), base excision repair (E), and DNA replication (F) pathways.



Supplementary Figure 2. Kaplan-Maier survival curves of HCC patients based on risk score models. (A, B) Prognostic analysis of HCC patients with (A) age ≤ 60 years and (B) age > 60 years. (C, D) Prognostic analysis of HCC patients with (C) female HCC patients and (D) male HCC patients. (E, F) Prognostic analysis of HCC patients with (E) grade 1-2 and (F) grades 3-4. (G, H) Prognostic analysis of HCC patients with (G) stage I-II and (H) stage III-IV. (I, J) Prognostic analysis of HCC patients with (I) stage T1-T2 and (J) stage T3-T4. (K, L) Prognostic analysis of HCC patients with (K) M0 as well as (L) M1 and MX. (M, N) Prognostic analysis of subgroups of HCC patients with different tumor-related lymph node metastases, N0 (M) as well as N1 and NX (N).