Am J Transl Res 2022;14(6):4207-4228
www.ajtr.org /ISSN:1943-8141/AJTR0139612

Original Article

Overexpressed SNRPB/D1/D3/E/F/G correlate with poor survival and immune infiltration in hepatocellular carcinoma

Fu-Ping Li[1], Gao-Hua Liu[2], Xue-Qin Zhang[3], Wei-Jie Kong[1], Jian Mei[4], Mao Wang[5], Yin-Hai Dai[1,5]

[1]Department of Clinical Medicine, Shaanxi University of Chinese Medicine, Xianyang 712046, Shaanxi, China; [2]Department of Oncology, Fujian Medical University Union Hospital, Fuzhou 350001, Fujian, China; [3]Jincheng Institute of Sichuan University, Chengdu 610000, Sichuan, China; [4]Department of Orthopaedic Surgery, The First Affiliated Hospital of Fujian Medical University, Fuzhou 350005, China; [5]Department of Surgical Oncology Medicine, Second Affiliated Hospital of Shaanxi University of Chinese Medicine, Xianyang 712000, Shaanxi, China

Received October 6, 2021; Accepted May 23, 2022; Epub June 15, 2022; Published June 30, 2022

Abstract: Background: Prior reports have indicated that the abnormal expression of small nuclear ribonucleoproteins (snRNPs) genes is related to malignant tumors. However, in hepatocellular carcinoma (HCC), the precise role of snRNPs is not well understood. Therefore, the purpose of this study was to evaluate the prognostic roles of SNRPB/D1/D2/D3/E/F/G and their correlation to immune infiltration in HCC. Methods: The study was carried out via the following databases, software, and experimental validation: ONCOMINE, GEPIA2, UALCAN, The Cancer Genome Atlas, Gene Expression Omnibus, ArrayExpress, Kaplan-Meier plotter, cBioPortal, STRING, DAVID 6.8, TIMER, Cytoscape software, and immunohistochemistry experiments. Results: Overexpressed SNRPB/D1/D2/D3/E/F/G proteins were found in HCC tissues. The transcription levels of 7 snRNPs genes were related to the TP53 mutation and tumor grades. SNRPB/D1/D2/D3/F/G expression was significantly correlated with cancer staging, whereas SNRPE was not. Moreover, Kaplan-Meier survival analysis showed that upregulation of SNRPB/D1/D2/E/G was relevant to worse OS in HCC patients, especially in patients with alcohol consumption and those without viral hepatitis. Multivariate Cox regression analysis indicated that expression of SNRPB/D1/D3/E/F/G were independent prognostic factors for unfavorable OS in HCC. In addition, a high mutation rate of snRNPs genes (44%) was also found in HCC. The mRNA expression levels of snRNPs were meaningfully and positively related to six types of infiltrating immune cells (B cells, $CD4^+$ T cells, $CD8^+$ T cells, neutrophil, macrophage, and dendritic cells). Also, SNRPB/D1/G genes were significantly associated with molecular markers of various immune cells in HCC. Conclusions: SNRPB/D1/D3/E/F/G are potential prognostic biomarkers for a short OS in HCC, and SNRPB/D1/G were novel immune therapy targets in HCC patients.

Keywords: Hepatocellular carcinoma, snRNPs, prognostic biomarker, ONCOMINE, TCGA, GEO

Introduction

Hepatocellular carcinoma (HCC) is one of the world's most fatal malignant tumors, which often occurs in patients with chronic hepatitis, liver cirrhosis and dysplastic nodules. Most patients are diagnosed in the middle and late stages because the progression of HCC is subtle at first. Although the treatments of HCC are continuously evolving and improving, including earlier detection and more effective therapy strategies (particularly, the emergence of immune checkpoint blockade, multi-kinase inhibitors, and antiangiogenics) over the past few decades, the 5-year survival rate is still only about 12% [1, 2]. All in all, the current focus of liver cancer research is to study the invasion and metastasis of HCC and its interaction with microenvironment and organism at the molecular level in order to find new prognostic biomarkers and therapeutic targets.

The spliceosome is a ribonucleoprotein (RNP) with a complex ring-shaped structure, which mainly consists of small nuclear ribonucleoproteins (snRNPs) encoded by seven SNRP genes (SNRPB, SNRPD1, SNRPD2, SNRPD3, SNRPE, SNRPF, and SNRPG) [3]. It is mainly responsible

for splicing the pre-RNA into mRNA [3]. Accurate splicing is essential to ensure normal cellular function like cell proliferation, apoptosis, migration, and invasion. Furthermore, previous research has proved that the aberrant expression of snRNPs genes are related to some human cancers, including cervical cancer [4], non-small-cell lung cancer (NSCLC) [5], glioblastoma [6], breast cancer [7], and hepatocellular carcinoma [8].

The human liver is a central immunological organ containing various immune cell subtypes, which play an essential role in preventing the invasion of microorganisms and tumor progression [9, 10]. A report from Marta Garnelo et al. [11] indicated that tumor-infiltrating T cells were correlated with better prognosis in HCC patients while associating with B cell infiltration. The reduction in $CD8^+$ T cell numbers was related to short survival in HCC patients [12]. Moreover, $CD74^+$ macrophages [13] and dendritic cells (DCs) [14] predicted favorable prognosis in HCC patients. Currently, a few studies have revealed spliceosomes are tightly associated with the immune microenvironment [15-17]. For instance, the deletion of some genes encoding splicing factors can cause severe defects in thymocyte development, thus a significant reduction of T cells [17].

Up to now, there have been limited studies investigating the connections between the abnormal expression of snRNPs genes and HCC. For instance, the report from Zhan Y T et al. [8] showed that overexpression of SNRPB facilitated HCC cell proliferation via indirect activation of the Akt pathway. In the study directed by Jia D et al., SNRPE was proved to be a putative oncogene in HCC through various *in vitro* and *in vivo* experiments [18]. However, the precise functional roles of snRNPs in HCC are yet unclear. The aim of our study was to present the expressions and mutations of SNRPB/D1/D2/D3/E/F/G and their associations with infiltrated immune cells in HCC via integrated bioinformatic analysis and to find promising candidate biomarkers which may be helpful for the treatment of HCC patients.

Materials and methods

ONCOMINE

ONCOMINE (http://www.oncomine.org) is a large database of tumor gene chips and a powerful data-mining platform [19]. In this paper, we obtained the data of snRNPs mRNA expression in various cancers from the Oncomine database and used Student's t test to analyze them. We defined the following threshold: *P*-value of 0.001, fold change of 1.5, gene rank of 10%, and data type of mRNA.

The cancer genome atlas (TCGA)

The Cancer Genome Atlas (https://portal.gdc.cancer.gov) is a freely available public platform containing rich cancer data [20]. In this study, the RNA-sequencing data of 374 HCC patients were obtained from TCGA, and then, we applied a Cox regression analysis to investigate the relationship between the expression levels of snRNPs and the overall survival (OS) of HCC patients further. A parameter of $P<0.1$ was included in that analysis, and *P*-value <0.05 was deemed significant.

Gene expression omnibus (GEO)

Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/geo/) is a free public database that contains sequencing data for over 4,636,000 samples [21]. In this study, the microarray data of 257 samples were obtained from the GEO database (GSE102079) to compare snRNPs expression levels in HCC and normal samples.

GEPIA 2

GEPIA 2 (http://gepia2.cancer-pku.cn/#index) is a powerful visualization website [22]. In this paper, we compared the transcription levels of snRNPs between HCC and normal samples and explored the correlation between snRNPs gene expression and tumor staging via the "Expression DIY" module of GEPIA 2. Moreover, the "Similar Genes Detection" module was used to find 50 neighboring genes related to snRNPs genes.

ArrayExpress

ArrayExpress (https://www.ebi.ac.uk/arrayexpress) is a large and comprehensive database, which includes 74,785 high-throughput functional genomics experiments data and 2,545,580 assays to date [23]. In this paper, we explored the transcription levels of snRNPs genes by different genomic experiments in the Expression Atlas. Log_2 fold change ≥1, *P*-value <0.05 was deemed significant.

UALCAN

UALCAN (http://ualcan.path.uab.edu) is an interactive online website based on the TCGA and the CPTAC databases [24]. In our study, the transcription levels of snRNPs in tumor grades and TP53 mutation of HCC patients were analyzed with the liver hepatocellular carcinoma dataset. *P*-value <0.05 was considered statistically meaningful.

Kaplan-Meier plotter

Kaplan-Meier plotter (http://kmplot.com/) is a friendly survival analysis website [25, 26]. In this study, the survival curves (OS, relapse-free survival (RFS), progression-free survival (PFS), and disease-specific survival (DDS) of snRNP expression in HCC were shown via Kaplan-Meier plotter. Meanwhile, we also explored if the upregulation of snRNPs genes had any additional impacts on the OS of HCC patients with alcohol consumption and viral hepatitis. *P*-value <0.05 was considered statistically meaningful.

CBioPortal

CBioPortal (http://www.cbioportal.org) is a robust online website based on TCGA data [27, 28]. In our study, the liver hepatocellular carcinoma dataset (TCGA, Firehose Legacy) was analyzed and visualized as the map of gene mutations, expression heatmap, and co-expression map of snRNPs using the CBioPorta database. The z-Score threshold was set to ± 1.8.

STRING

STRING (https://string-db.org/) is a convenient online tool to visualize protein interactions [29]. In this paper, the tool of STRING was used to construct the interaction networks between the 7 snRNPs proteins and their 50 frequently neighboring genes.

DAVID 6.8

DAVID 6.8 (https://david.ncifcrf.gov/) is a powerful online tool for gene functional classification [30]. This study performed gene ontology (GO) enrichment analysis and KEGG pathway enrichment analysis via DAVID 6.8. The GO enrichment analysis consisted of three parts: biological processes (BP), cellular components (CC), and molecular functions (MF).

Cytoscape

Cytoscape (https://cytoscape.org/) is a software focused on open-source web visualization and analysis [31]. In our study, the "CytoHubba" plugin was used to identify hub genes between the 7 snRNPs genes and their 50 frequently neighboring genes.

TIMER

TIMER (https://cistrome.shinyapps.io/timer/) is a web-based tool for analyzing immune cell infiltration in different tumor tissues [32, 33]. Our study explored the association between the 7 snRNPs genes and six types of immune cells and molecular markers of various immune cells in HCC via the "Gene" module.

Immunohistochemistry (IHC)

A total of 24 formalin-fixed and paraffin-embedded HCC tissues and 24 normal liver tissues were used in the IHC analysis. Firstly, 3-mm tumor sections were incubated with commercial rabbit polyclonal antibodies against SNRPB, SNRPD1, SNRPD2, SNRPD3, SNRPE, SNRPF, and SNRPG (SNRPB/D2/E/F from BioSwamp Life Science Lab; SNRPD1/D3/G from Elabscience Biotechnology) at 1/50 dilution overnight at 4°C. Then, the sections were conjugated with horse-radish peroxidase (HRP) antibody (Zhongshan Goldenbridge Biotechnology, CO., Ltd.) at room temperature for 10 minutes. All sections were sequentially treated with a biotinylated anti-rabbit immunoglobulin (Pan-Specific antibody) for 20 minutes at 37°C, then covered by 3, 3-diaminobenzidine (DAB), and slides were mounted with Vectashield mounting medium. Subsequently, all fields were observed under light microscopy. Control experiments without primary antibodies demonstrated that the signals observed were specific. All specimens were scored independently by two experienced pathologists. The staining index was calculated as follows: staining index = score of staining intensity + score of stained tumor cells. The intensity of staining was scored according to: 0 (no staining, -); 1 (weak staining, light yellow, +); 2 (moderate staining, yellow brown, ++); 3 (strong staining, brown, +++). Tumor cell proportion

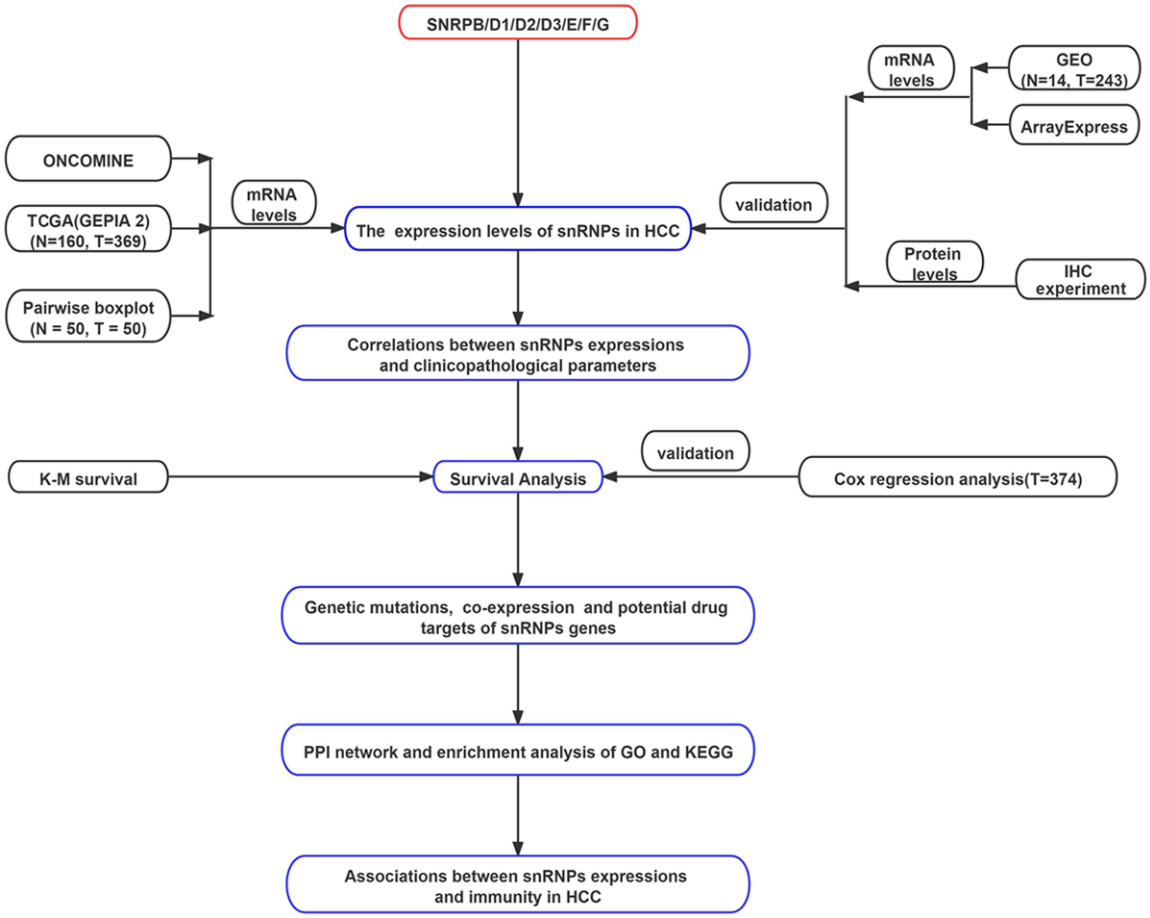


Figure 1. The present study flow chart.

was scored as follows: 0 (no positive tumor cells); 1 (<25% positive tumor cells); 2 (26%-49% positive tumor cells); 3 (50%-74% positive tumor cells); 4 (≥75% positive tumor cells). Each sample was graded by the staining index as 0 (-), 1-2 (+), 3-5 (++), and 6-8 (+++).

Statistical analysis methods

The statistical methods of the ANOVA in the GEPIA 2 database, the logarithmic rank test in Kaplan-Meier survival analysis, the Pearson coefficient of co-expression analysis, and the Cox proportional risk models in univariate and multivariate analyses, were used in statistical analysis. The TCGA HCC data were analyzed using the XIAN TAO platform (www.xiantao.love) based on R software package "survival". The GEO data were visualized via the platform of Assistant for Clinical Bioinformatics (https://www.aclbi.com) based on R software package "ggplot2". The data of staining index were performed in GraphPad Prism (v9.0.2) (San Diego, CA, United States) and presented as the mean ± SD. Wilcox rank sum test was used for statistical analyses between the data pairs where appropriate. *P*-value <0.05 was considered statistically meaningful.

Results

Characteristics of the research

This study covered many aspects, including differential expression analysis, association with clinicopathological parameters, survival analysis, gene mutations, co-expression, functional enrichment analysis, and correlation with immunity, revealing the roles of SNRPB/D1/D2/D3/E/F/G in HCC. The present study flow chart was shown in **Figure 1**.

The expression levels of snRNPs were significantly increased in HCC patients

Firstly, the Oncomine database (http://www.oncomine.org) was used to compare the mRNA

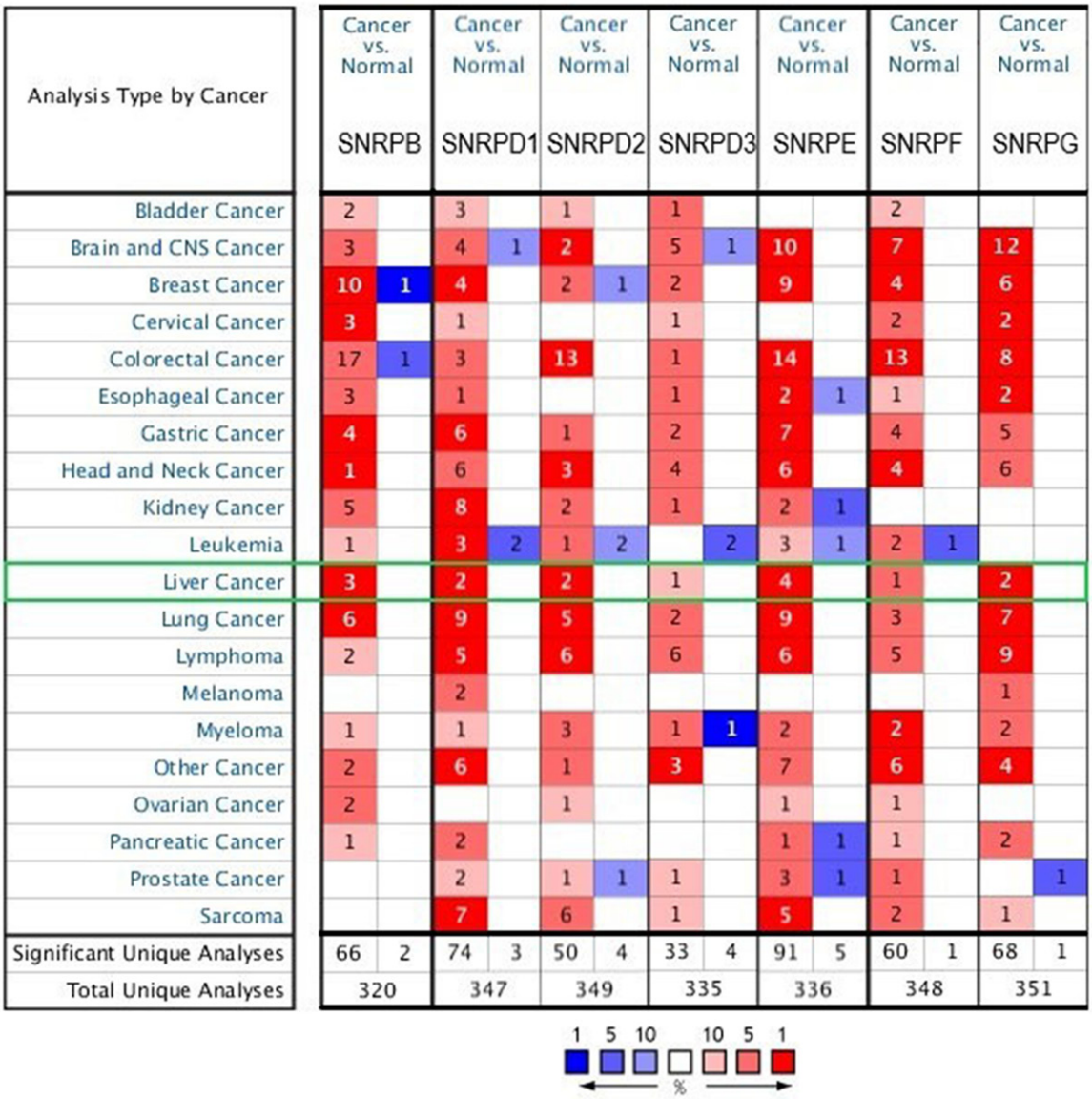


Figure 2. The transcription levels of snRNPs genes in different human cancers (ONCOMINE).

levels of snRNPs in HCC and non-cancer samples. From **Figure 2** and **Table 1**, we could see that the mRNA levels of SNRPB/D1/D2/D3/E/F/G were meaningfully higher in HCC than in non-cancer tissues. In the meantime, we used the GEPIA2 database (http://gepia2.cancer-pku.cn/#index) (N=160; T=369) to investigate the mRNA expressions of snRNPs in HCC and normal liver samples further. The results were also consistent with the mRNA levels of snRNPs which were upregulated in HCC compared to normal liver tissues (**Figure 3A**, **3B**). In addition, a pairwise boxplot (N=50; T=50) also confirmed that mRNA expression of snRNPs were overexpressed in HCC (**Figure 4A**). Moreover, the above results were validated via the ArrayExpress database (https://www.ebi.ac.uk/arrayexpress) and the GEO database (N=14; T=243). The transcription levels of snRNPs genes also significantly increased in HCC (**Table 2**; Figure S1).

Lastly, 24 paired HCC tissues and normal tissue samples were utilized to perform IHC experiments to further verify the results of bioinformatics analysis. The staining index of 24 patients was summarized in Table S1. The results of IHC also indicated that snRNPs proteins were more highly expressed in HCC tis-

Table 1. The mRNA levels of snRNPs were significantly higher in HCC than in normal liver tissues (ONCOMINE)

	Types of HCC VS normal liver tissue	Fold Change	*P*-value	T-test	Ref
SNRPB	Hepatocellular Carcinoma	2.315	4.23E-75	22.843	Roessler Liver 2 [34]
	Hepatocellular Carcinoma	1.696	7.99E-16	8.768	Chen Liver [35]
	Hepatocellular Carcinoma	2.338	1.26E-6	5.744	Roessler Liver [34]
SNRPD1	Hepatocellular Carcinoma	3.270	2.55E-97	27.765	Roessler Liver 2 [34]
	Hepatocellular Carcinoma	2.880	7.91E-9	7.287	Roessler Liver [34]
SNRPD2	Hepatocellular Carcinoma	2.160	4.05E-82	24.017	Roessler Liver 2 [34]
	Hepatocellular Carcinoma	2.052	5.00E-9	7.882	Roessler Liver 2 [34]
SNRPD3	Hepatocellular Carcinoma	1.532	6.29E-40	14.617	Roessler Liver 2 [34]
SNRPE	Hepatocellular Carcinoma	2.971	1.18E-103	28.959	Roessler Liver 2 [34]
	Hepatocellular Carcinoma	2.046	1.10E-25	12.246	Chen Liver [35]
	Hepatocellular Carcinoma	2.344	2.12E-7	6.554	Roessler Liver [34]
	Hepatocellular Carcinoma	1.906	1.75E-4	4.474	Wurmbach Liver [36]
SNRPF	Hepatocellular Carcinoma	1.675	5.18E-50	17.473	Roessler Liver 2 [34]
SNRPG	Hepatocellular Carcinoma	1.750	1.32E-69	21.226	Roessler Liver 2 [34]
	Hepatocellular Carcinoma	1.552	4.36E-6	5.191	Roessler Liver 2 [34]

sues than in normal tissues (**Figure 5A**, **5B**; all P<0.001), which was consistent with the results of bioinformatics analysis on RNA levels.

The expression of snRNPs genes were related to clinicopathological parameters of HCC patients

We explored the correlations between the expression of snRNPs genes and TP53 mutations, tumor grades, and cancer staging via the UALCAN database (http://ualcan.path.uab.edu) and the GEPIA2 database. From **Figure 4B**, we could see that the mRNA expression of snRNPs in HCC with TP53 mutation were higher than those without TP53 mutation (**Figure 4B**). Also, the mRNA expression of snRNPs was meaningfully relevant to tumor grades. The patients in grade 4 expressed the highest mRNA levels of SNRPB/D1/D2/D3/F/G, while the highest SNRPE expression was found in grade 3 (**Figure 6A**). Similarly, as can be seen in **Figure 6B**, the mRNA levels of SNRPB/D1/D2/D3/F/G varied meaningfully across different stages (P<0.05), whereas SNRPE did not (P>0.05).

The upregulation of snRNPs genes is related to poor survival outcomes in HCC patients

Using Kaplan-Meier Plotter (http://kmplot.com/), we assessed the prognostic values of snRNPs in HCC patients. From **Figure 7**, we could see that high expression of SNRPB/D1/E/G were associated with short OS (P< 0.05). In addition, high mRNA expression of SNRPB/D1/D2/G were correlated with poor RFS and PFS (P<0.05). However, highly expressed SNRPD3 and SNRPF were not significantly related to the survival of HCC patients (all P>0.05).

As is well known, chronic hepatitis virus and alcohol abuse are common risk factors leading to HCC [37]. So, we explored if the up-regulation of snRNPs genes had any additional impacts on the OS of HCC patients. From **Figure 8**, we could see that the up-regulation of SNRPD1 (HR: 1.96 VS. 2.13), SNRPD2 (HR: 1.5 VS. 1.96), SNRPE (HR: 1.55 VS. 3.08), and SNRPG (HR: 1.68 VS. 2.17) had meaningfully adverse effects on the OS of HCC patients with alcohol consumption (**Figure 8B**, **8C**, **8E** and **8G**). Furthermore, the up-regulation of SNRPB (HR: 2.16 VS. 1.71), SNRPD1 (HR: 2.23 VS. 1.76), SNRPD2 (HR: 1.83 VS. 1.61), SNRPE (HR: 2.29 VS. 1.91), SNRPF (HR: 1.88 VS. 0.65), and SNRPG (HR: 2.24 VS. 1.51) had significant negative impacts on the OS of HCC patients without viral hepatitis (**Figure 8A-C**, **8E-G**). However, these adverse impacts were not meaningfully related to the OS of HCC patients with viral hepatitis. Therefore, more research is needed to interpret potential mechanisms of the above-described events.

In a subsequent study, we attempted to evaluate the independent prognostic value of snRNPs genes for OS of HCC patients. The

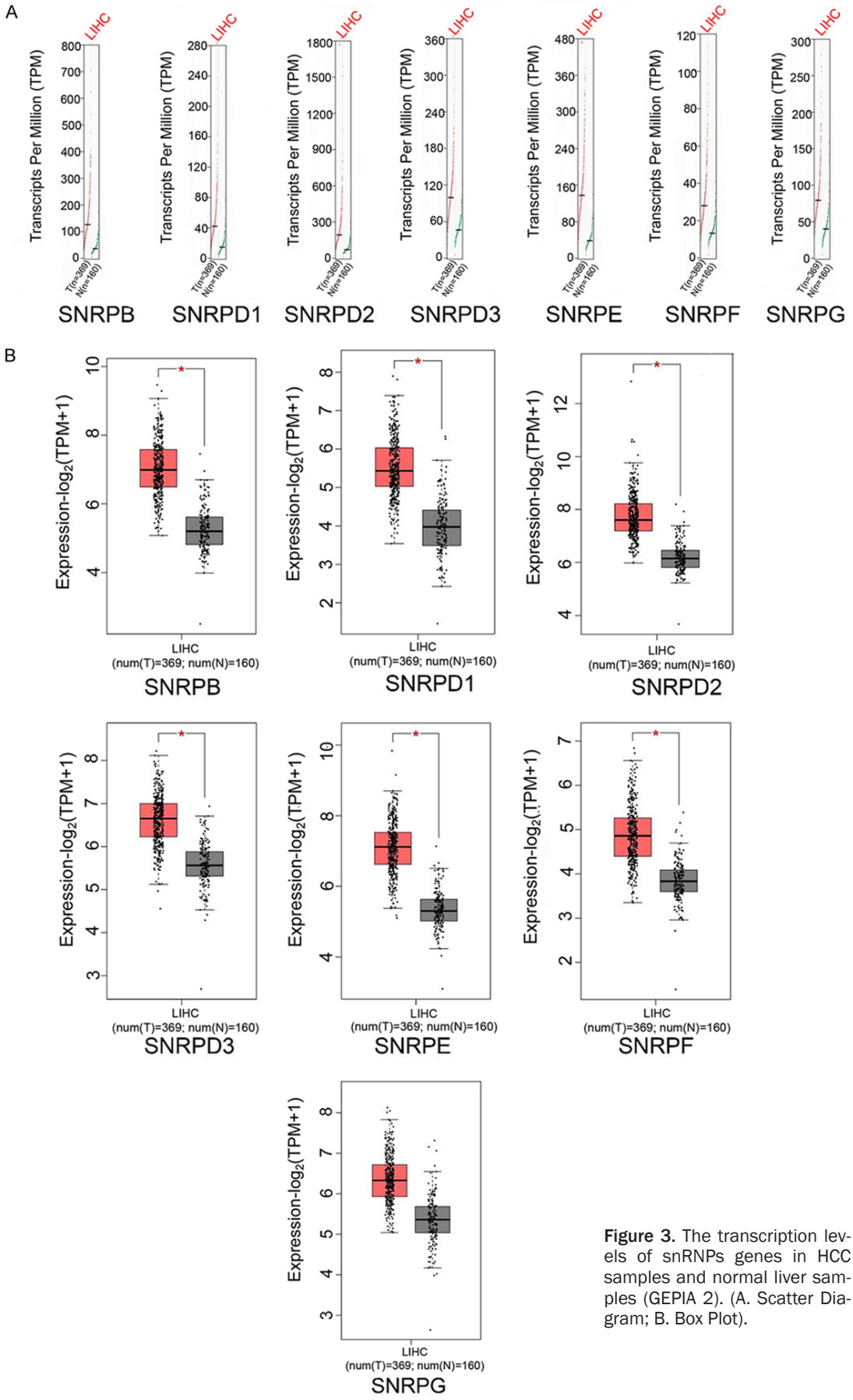


Figure 3. The transcription levels of snRNPs genes in HCC samples and normal liver samples (GEPIA 2). (A. Scatter Diagram; B. Box Plot).

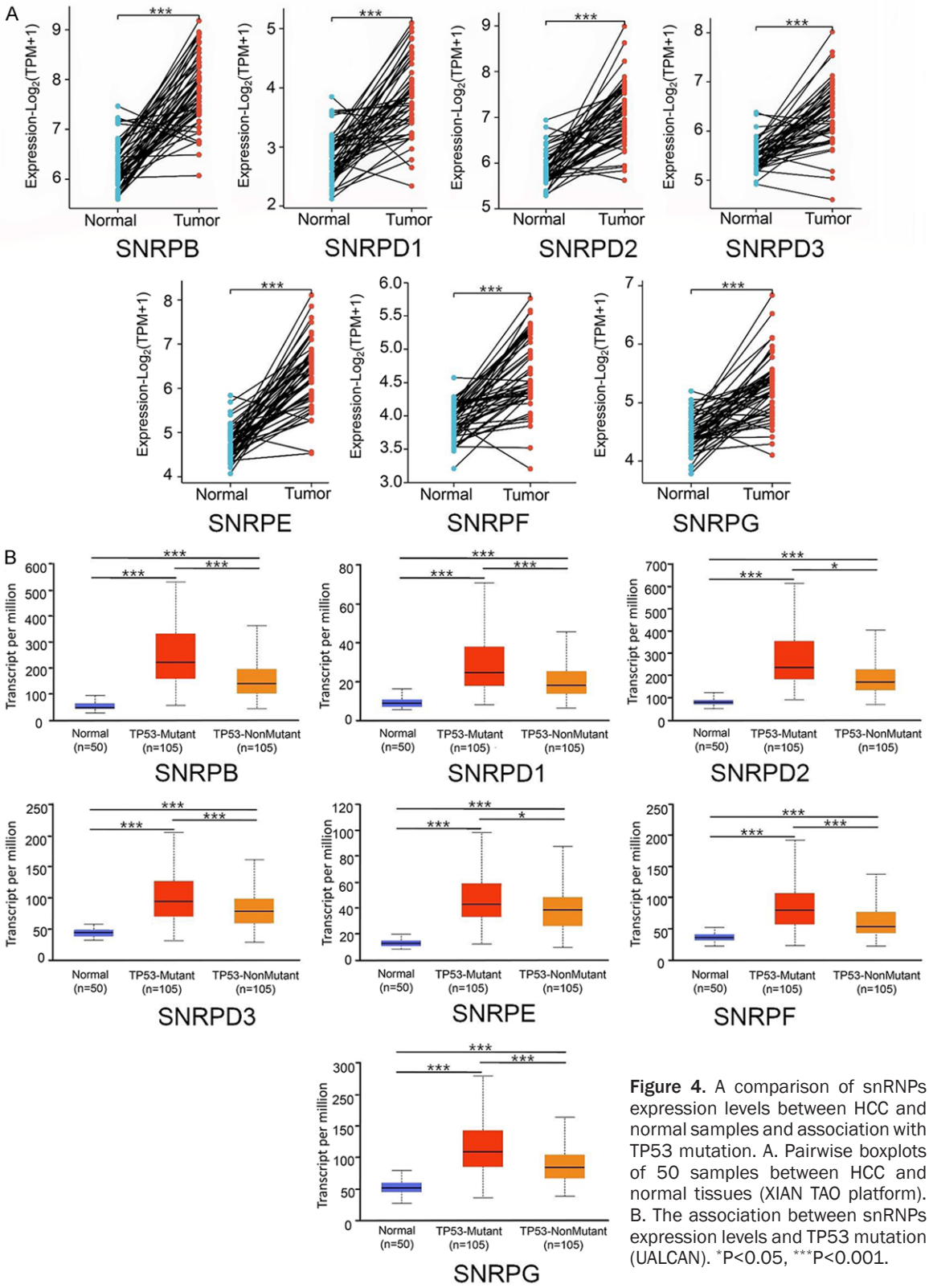


Figure 4. A comparison of snRNPs expression levels between HCC and normal samples and association with TP53 mutation. A. Pairwise boxplots of 50 samples between HCC and normal tissues (XIAN TAO platform). B. The association between snRNPs expression levels and TP53 mutation (UALCAN). *P<0.05, ***P<0.001.

RNA-sequencing data and clinical information (Table S2) of 374 HCC patients were obtained from TCGA for Cox regression analysis. Univariate Cox regression analysis indicated that

Table 2. The transcription levels of snRNPs genes were meaningfully higher in HCC than in normal tissue (ArrayExpress)

Gene	Experiment accession	Comparison	log_2 fold change	Adjusted *P*-value
SNRPB	E-GEOD-33294	Hepatocellular Carcinoma vs. Normal Tissue	1.8	7.47888E-05
SNRPD1	E-GEOD-33294	Hepatocellular Carcinoma vs. Normal Tissue	1.6	0.000186
SNRPD2	E-GEOD-55048	Hepatocellular Carcinoma vs. Normal Tissue	1.1	0.019011
SNRPD3	E-MTAB-9712	Hepatocellular Carcinoma vs. Normal Tissue	1.3	2.15E-213
SNRPE	E-GEOD-33294	Hepatocellular Carcinoma vs. Normal Tissue	1.8	0.001404
SNRPF	E-MTAB-9712	Hepatocellular Carcinoma vs. Normal Tissue	1.3	4.97E-119
SNRPG	E-MTAB-9712	Hepatocellular Carcinoma vs. Normal Tissue	1.5	3.38E-147

T stage, pathologic stage, SNRPB, SNRPD1, SNRPD2, SNRPD3, SNRPE, SNRPF, and SNRPG were all meaningfully correlated with short OS of HCC patients (all P<0.05; Figure S2). In multivariate Cox regression analysis, SNRPB, SNRPD1, SNRPD3, SNRPE, SNRPF, and SNRPG were observed to have significant associations with poor OS of HCC patients (all P<0.05; Figure S3). Taken together, SNRPB/D1/D3/E/F/G were considered as independent prognostic factors for unfavorable OS in HCC.

Gene mutations, co-expression, and potential drug targets of snRNPs genes in HCC patients

Using the cBioPortal database (http://www.cbioportal.org/), we explored the mutation rate and co-expression of snRNPs. As shown in **Figure 9A** and B, 44% (160/360) of HCC patients were found to have more than one gene mutation. SNRPB, SNRPD1, SNRPD2, SNRPD3, SNRPE, SNRPF, and SNRPG were altered in 14%, 10%, 9%, 10%, 29%, 12%, and 12% of the 360 HCC patients (**Figure 9A**). Likewise, the Expression Heatmap also displayed the degree of snRNPs mutations (**Figure 9B**). SNRPE had the highest mutation rate among snRNPs in HCC. In the meantime, we performed co-expression analysis for snRNPs. Significant and positive correlations were observed among snRNPs in **Figure 9C**, including SNRPB with SNRPD1/D2/D3/E/F/G; SNRPD1 with SNRPB/D2/D3/E/F/G; SNRPD3 with SNRPB/D1/D2/E/F/G; SNRPE with SNRPB/D1/D2/D3/F/G; SNRPF with SNRPB/D1/D2/D3/E/G; and SNRPG with SNRPB/D1/D2/D3/E/F (**Figure 9C**). Furthermore, we used the COSMIC database (https://cancer.sanger.ac.uk/cosmic) to investigate the 3D structure of the seven snRNPs proteins. Blue groups indicate predicted binding sites that are predicted to be "drug targets" (Figure S4).

Constructed protein-protein interaction (PPI) network and performed functional enrichment analysis of snRNPs in HCC patients

In order to explore the possible snRNPs protein-mediated biological pathways in hepatocellular carcinoma, we constructed a protein interaction network based on the 7 snRNPs proteins (**Figure 9D**) and another for snRNPs proteins and their 50 frequently neighboring proteins (**Figure 9E**) via the STRING database (https://string-db.org/). Meanwhile, we used Cytoscape software (https://cytoscape.org/) to find out these hub genes related to snRNPs protein-mediated biological pathways. As was shown in **Figure 9F**, eight hub genes, including CDCA3, PTTG1, NXT1, CCNB1, NOP56, CDC45, TRIM28, and H2AFX, were tightly correlated with the alterations of snRNPs.

Afterwards, snRNPs genes and their 50 neighboring genes were subjected to GO and KEGG functional enrichment analysis in HCC via the tool of DAVID 6.8 (https://david.ncifcrf.gov). The results were shown in **Figure 10** and Table S3. Biological processes (BP) included RNA splicing; RNA splicing, via transesterification reactions; mRNA splicing, via spliceosome; RNA splicing, via transesterification reactions with bulged adenosine as nucleophile; and mitotic nuclear division (**Figure 10A**). Cellular components (CC) suggested that SNRP genes existed mainly in spliceosomal complex; catalytic step 2 spliceosome; spliceosome snRNP complex; U2-type catalytic step 2 spliceosome; and U12-type spliceosome complex (**Figure 10A**). Molecular function (MF) indicated that SNRP genes were related to the structural constituents of ribonucleoprotein complex binding,

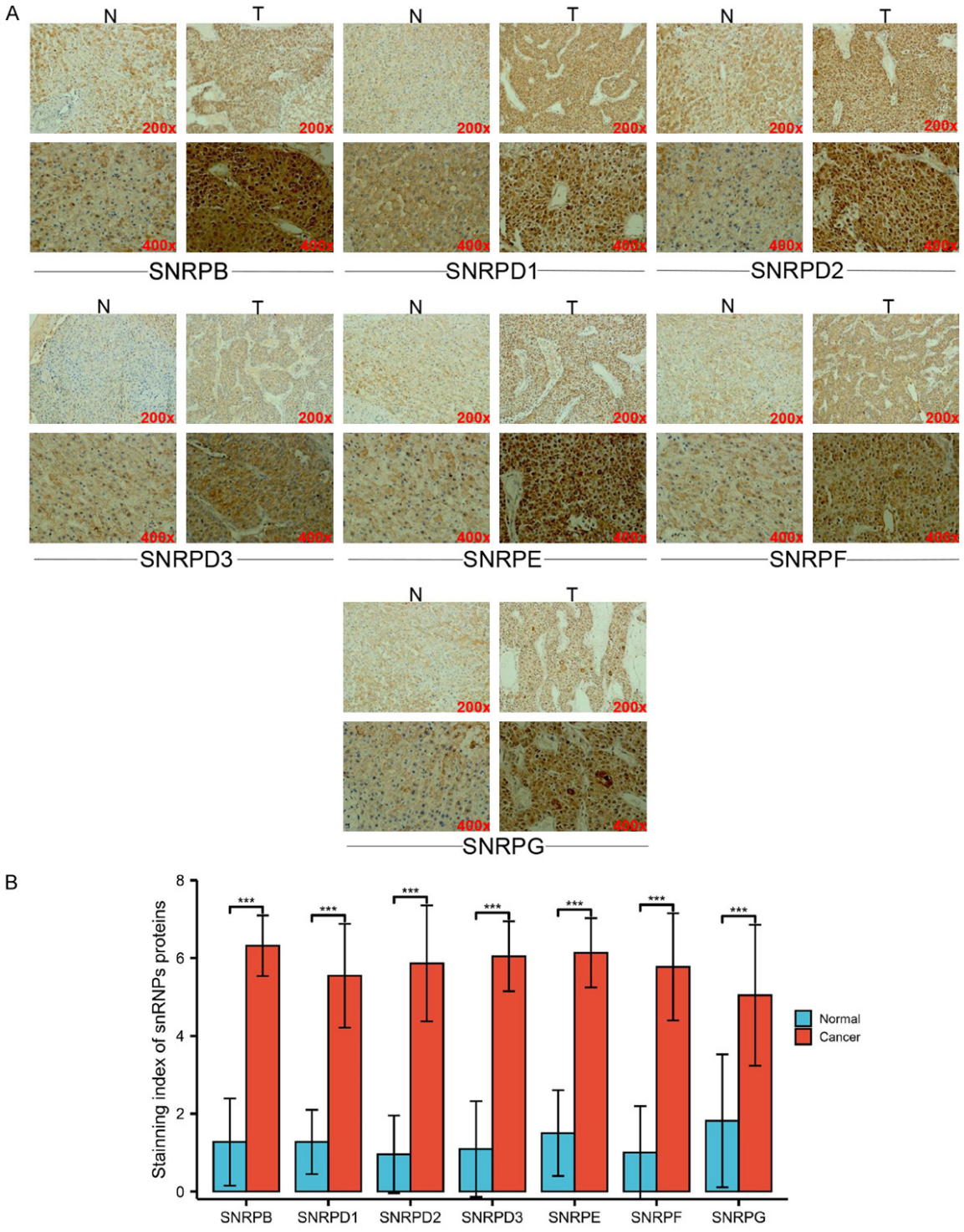


Figure 5. The protein expression levels of snRNPs in HCC (IHC). A. The IHC images of snRNPs expression in HCC tissues and normal tissues. B. Staining index of snRNPs expression in HCC tissues and normal tissues. N: normal, T: Tumor; ***P<0.001.

DNA helicase activity, and cyclin-dependent protein serine/threonine kinase regulator activity (**Figure 10A**). Moreover, KEGG pathway enrichment analysis suggested that spliceosome, cell cycle, and human T-cell leukaemia virus infection were significantly associated

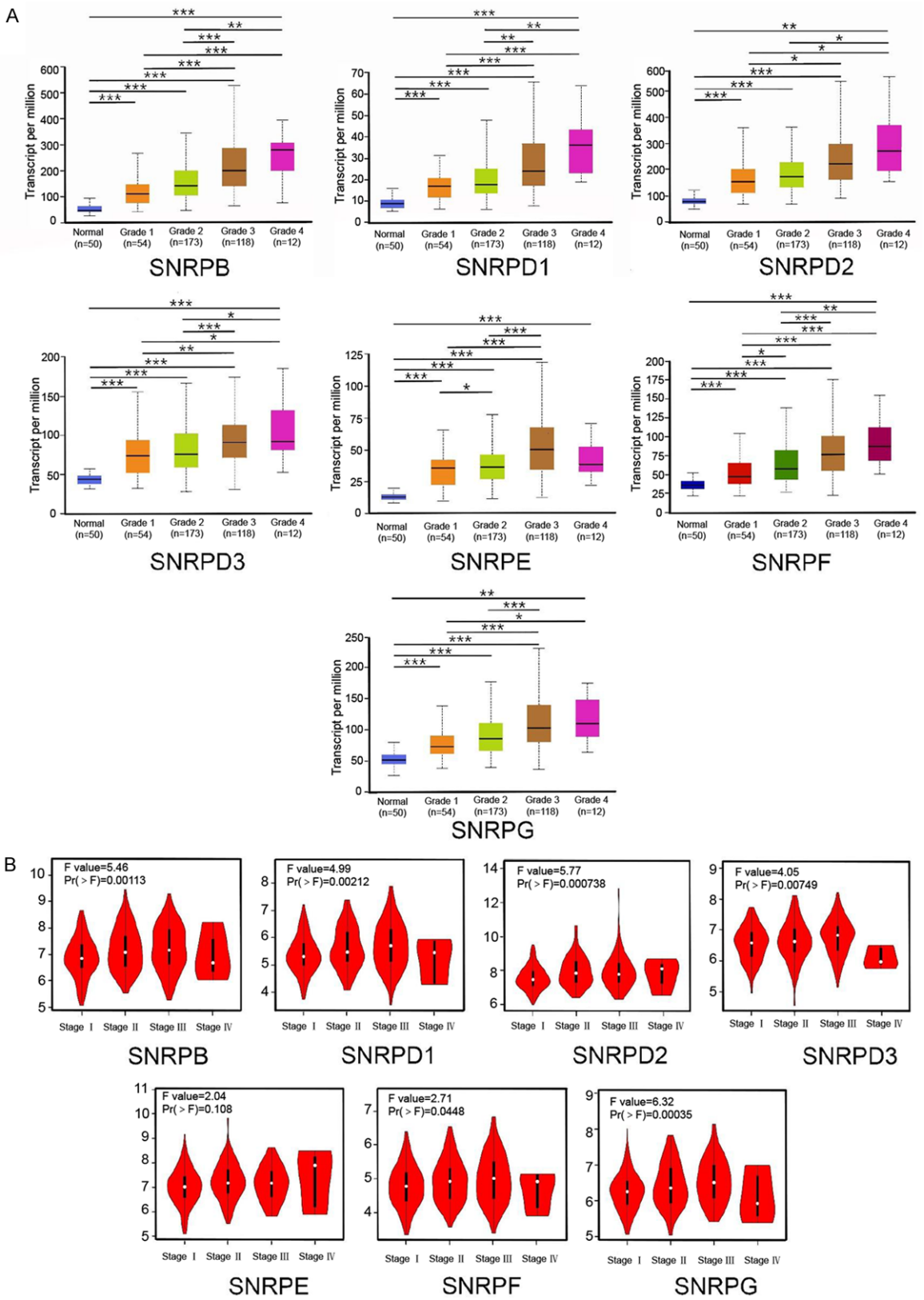
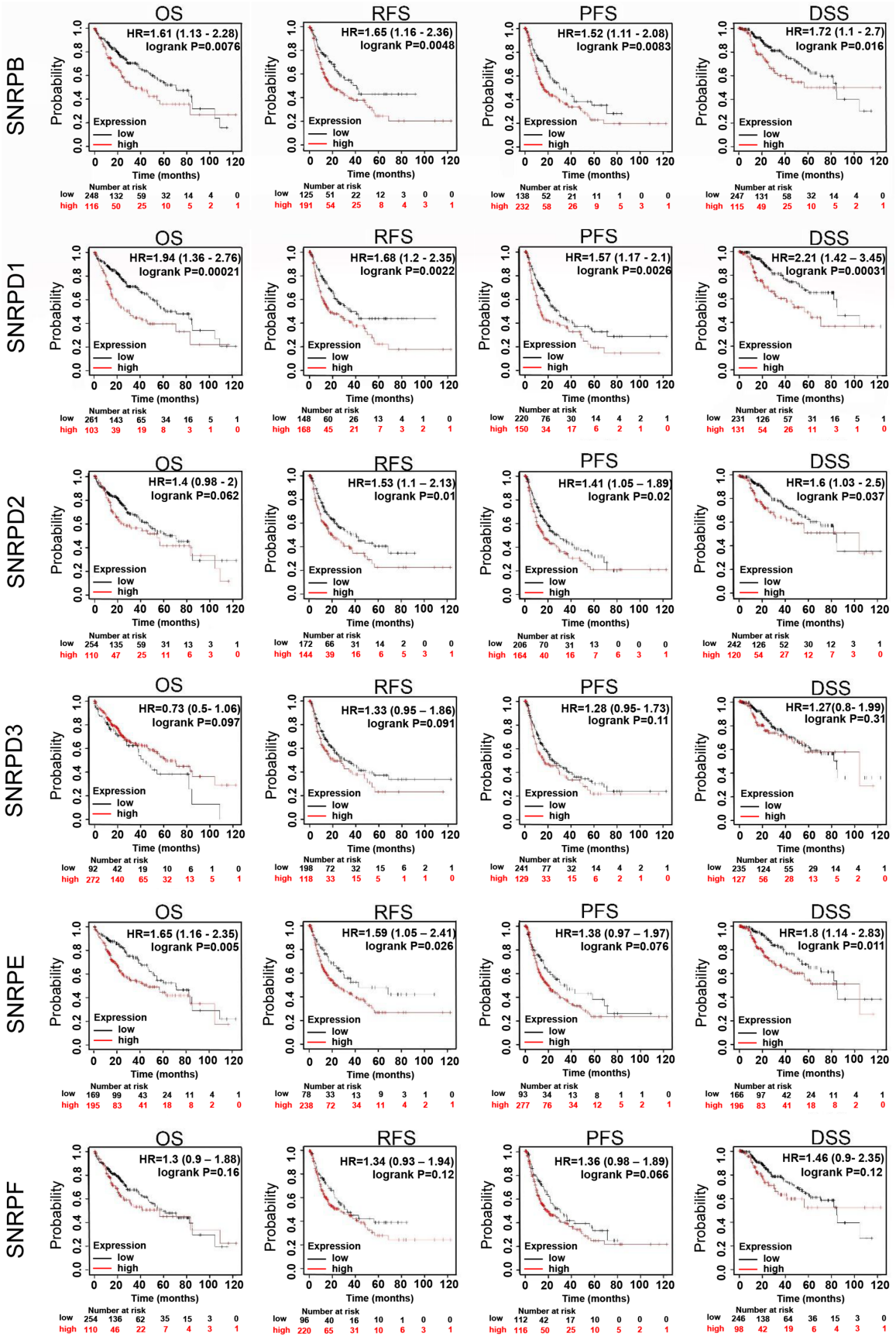


Figure 6. The association between snRNPs expression levels and clinicopathological parameters of HCC patients. A. Correlations between snRNPs expression levels and tumor grades in HCC (UALCAN). B. Association between snRNPs expression levels and cancer staging (GEPIA2). *P<0.05, **P<0.01, ***P<0.001.


SNRPB
OS
HR=1.61 (1.13 - 2.28)
logrank P=0.0076
RFS
HR=1.65 (1.16 - 2.36)
logrank P=0.0048
PFS
HR=1.52 (1.11 - 2.08)
logrank P=0.0083
DSS
HR=1.72 (1.1 - 2.7)
logrank P=0.016
SNRPD1
OS
HR=1.94 (1.36 - 2.76)
logrank P=0.00021
RFS
HR=1.68 (1.2 - 2.35)
logrank P=0.0022
PFS
HR=1.57 (1.17 - 2.1)
logrank P=0.0026
DSS
HR=2.21 (1.42 – 3.45)
logrank P=0.00031
SNRPD2
OS
HR=1.4 (0.98 - 2)
logrank P=0.062
RFS
HR=1.53 (1.1 – 2.13)
logrank P=0.01
PFS
HR=1.41 (1.05 – 1.89)
logrank P=0.02
DSS
HR=1.6 (1.03 - 2.5)
logrank P=0.037
SNRPD3
OS
HR=0.73 (0.5- 1.06)
logrank P=0.097
RFS
HR=1.33 (0.95 – 1.86)
logrank P=0.091
PFS
HR=1.28 (0.95- 1.73)
logrank P=0.11
DSS
HR=1.27(0.8- 1.99)
logrank P=0.31
SNRPE
OS
HR=1.65 (1.16 - 2.35)
logrank P=0.005
RFS
HR=1.59 (1.05 – 2.41)
logrank P=0.026
PFS
HR=1.38 (0.97 – 1.97)
logrank P=0.076
DSS
HR=1.8 (1.14 - 2.83)
logrank P=0.011
SNRPF
OS
HR=1.3 (0.9 – 1.88)
logrank P=0.16
RFS
HR=1.34 (0.93 – 1.94)
logrank P=0.12
PFS
HR=1.36 (0.98 – 1.89)
logrank P=0.066
DSS
HR=1.46 (0.9- 2.35)
logrank P=0.12
Probability
Expression
low
high
Time (months)
Number at risk

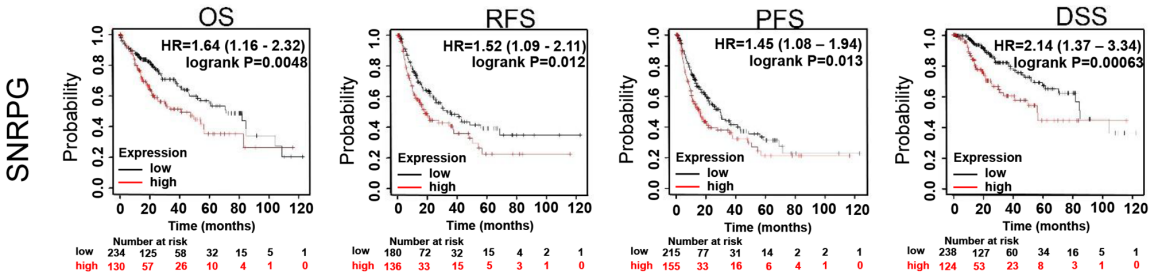
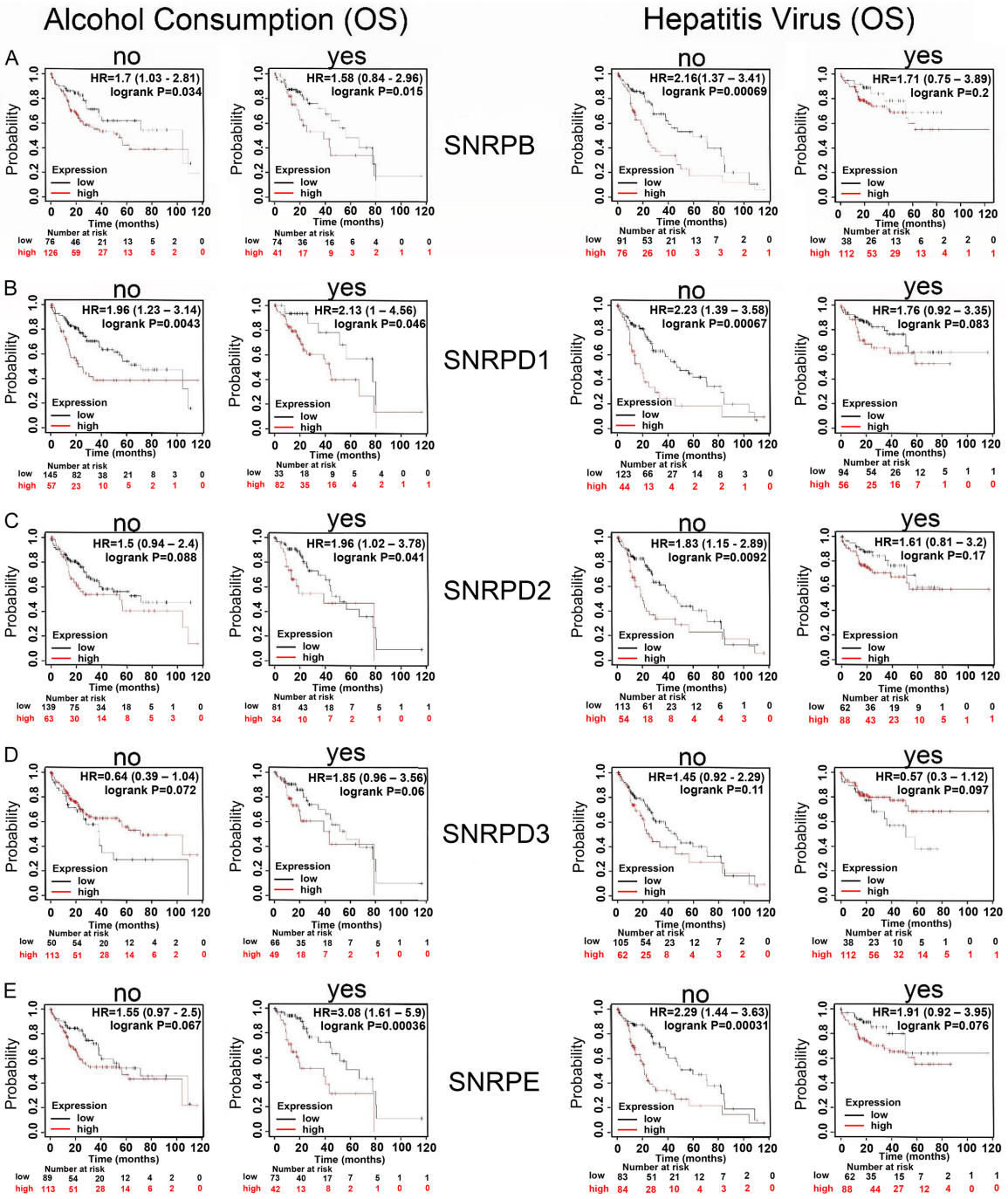


Figure 7. The Kaplan-Meier survival curves were used to evaluate the survival of HCC patients with snRNPs expressions (Kaplan-Meier Plotter).



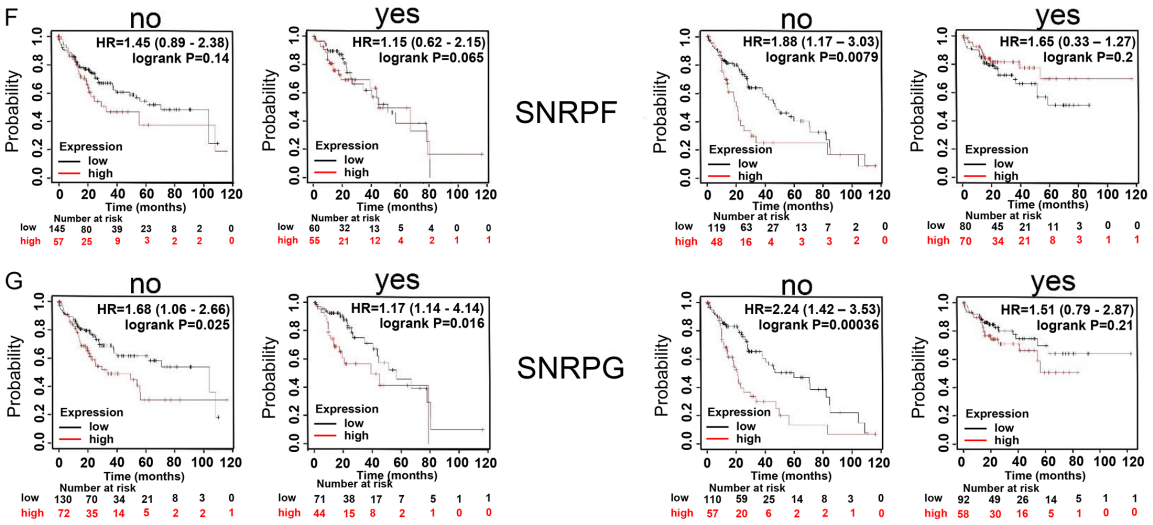
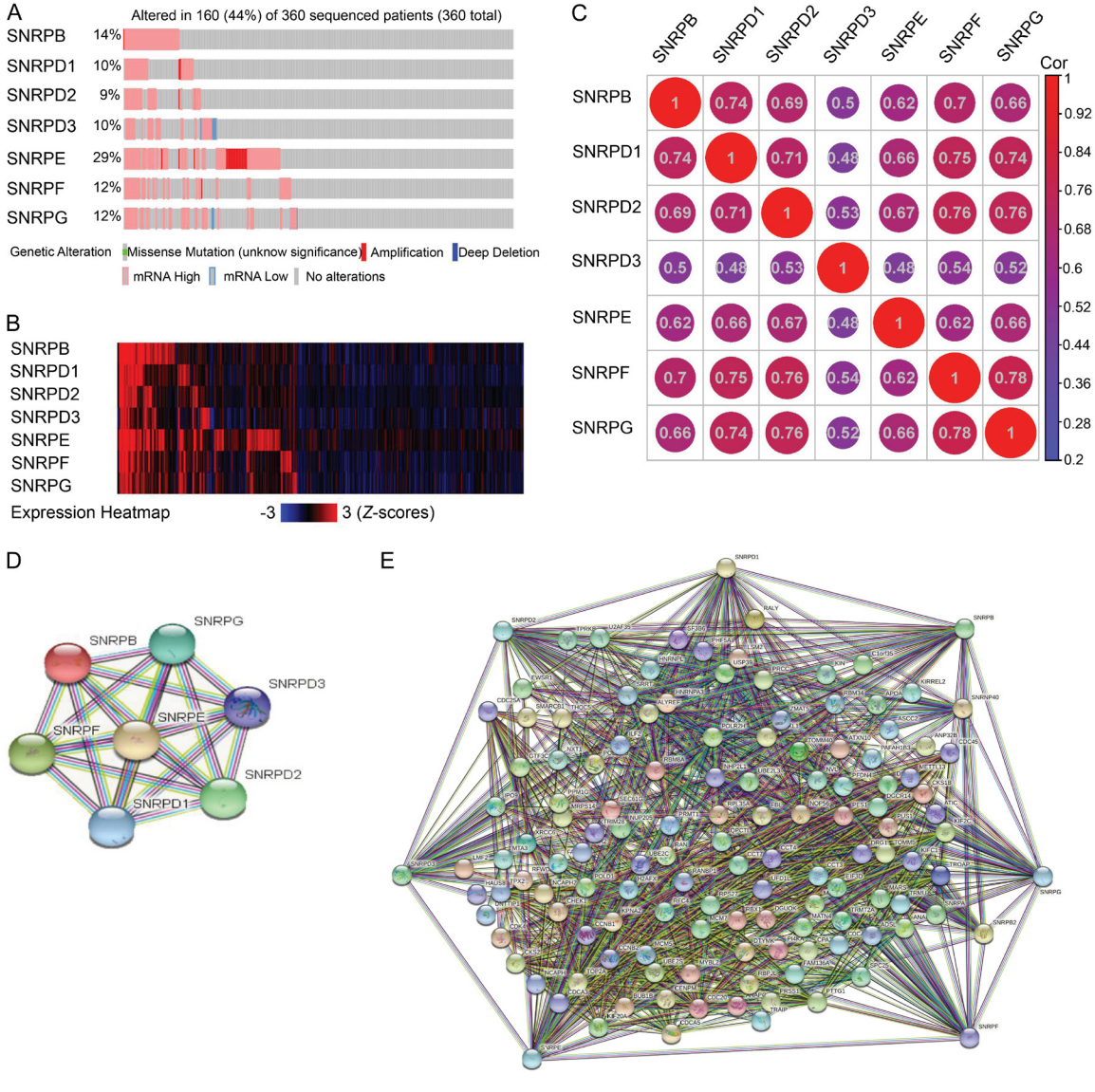


Figure 8. Up-regulated snRNPs genes are adverse survival factors in alcohol-consuming HCC patients without viral hepatitis.



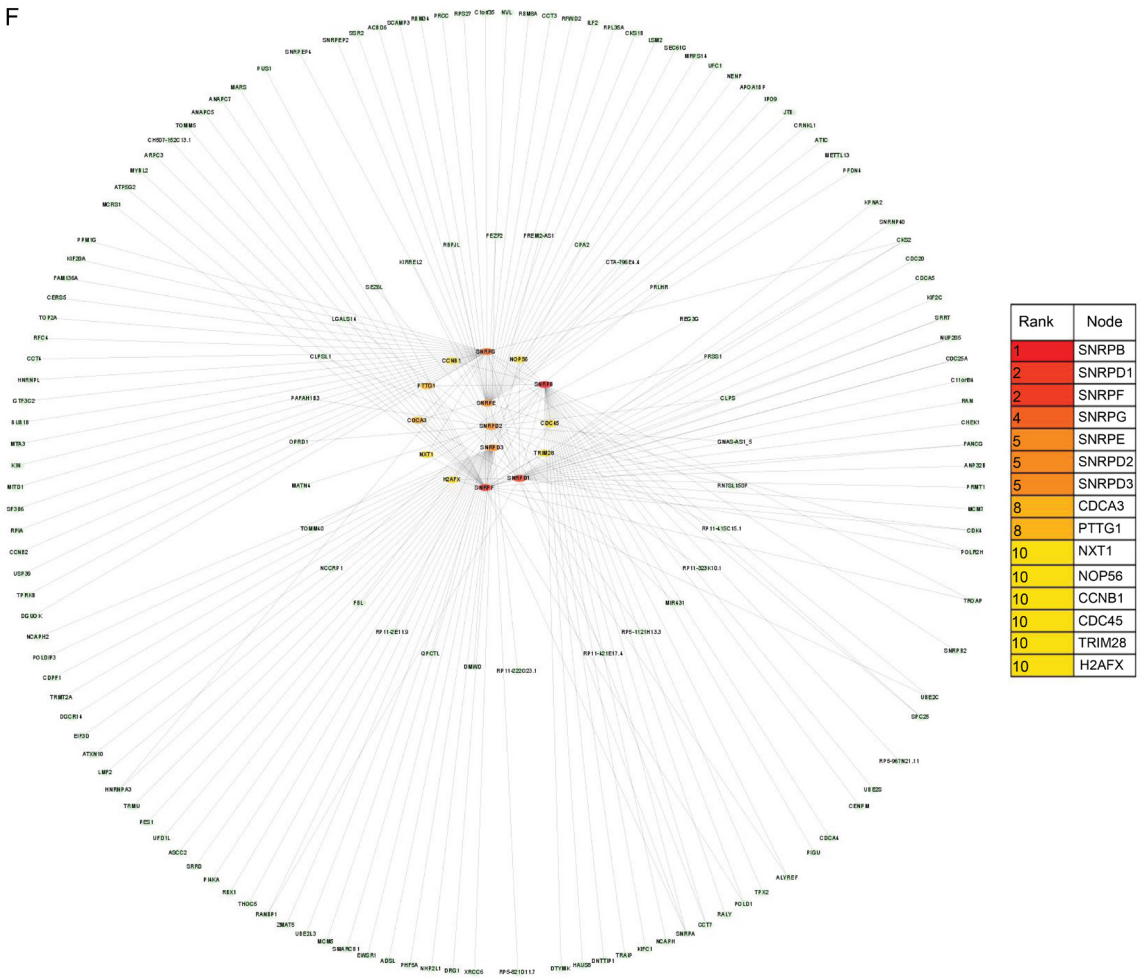


Figure 9. Gene mutations, co-expression, and PPI network for snRNPs in HCC (cBioPortal, STRING, Cytoscape). A, B. Gene mutations and expression heatmap of snRNPs in HCC patients (cBioPortal). C. The co-expression network for snRNPs (cBioPortal). D. PPI network based on the seven snRNPs proteins (STRING). E. PPI network based on snRNPs proteins and their 50 neighboring proteins (STRING). F. Eight hub genes, including CDCA3, PTTG1, NXT1, CCNB1, NOP56, CDC45, TRIM28, and H2AFX, are shown in the PPI network (Cytoscape).

with the alterations of snRNPs (**Figure 10B**). The complex process of mRNA/RNA splicing was shown in Figure S5.

The snRNPs expression was related to immune infiltration in HCC patients

Lastly, we investigated the relationships between 7 snRNPs proteins and immune infiltration via the TIMER database (https://cistrome.shinyapps.io/timer/). Seen from **Figure 11** and **Table 3**. SNRPB expression was positively relevant to these infiltrated cells, including B cells (Cor=0.435, *P*=2.50E-17), $CD8^+$ T cells (Cor=0.3, *P*=1.61E-8), $CD4^+$ T cells (Cor=0.265, *P*=6.12E-7), macrophage (Cor=0.383, *P*=2.26E-13), neutrophil (Cor=0.244, *P*=4.60E-6), and dendritic cells (Cor=0.386, *P*=1.71E-13) (**Figure 11A**). SNRPD1 expression was positively related to the infiltration of B cells (Cor=0.484, *P*=1.28E-21), $CD8^+$ T cells (Cor=0.394, *P*=3.63E-14), $CD4^+$ T cells (Cor=0.298, *P*=1.65E-8), macrophage (Cor=0.457, *P*=5.67E-19), neutrophil (Cor=0.346, *P*=3.96E-11), and dendritic cells (Cor=0.484, *P*=2.46E-21) (**Figure 11B**). SNRPD2 expression was positively correlated with these infiltrated cells, including B cells (Cor=0.308, *P*=5.13E-09), $CD8^+$ T cells (Cor=0.299, *P*=1.63E-8), macrophage (Cor=0.279, *P*=1.63E-7), and dendritic cells (Cor=0.304, *P*=1.06E-8) (**Figure 11C**). SNRPD3 expression was positively relevant to the infiltration of B cells (Cor=0.256, *P*=1.66E-6), macrophage (Cor=0.273, *P*=3.04E-7) and dendritic cells (Cor=0.247, *P*=4.10E-6) (**Figure 11D**). SNRPE expression

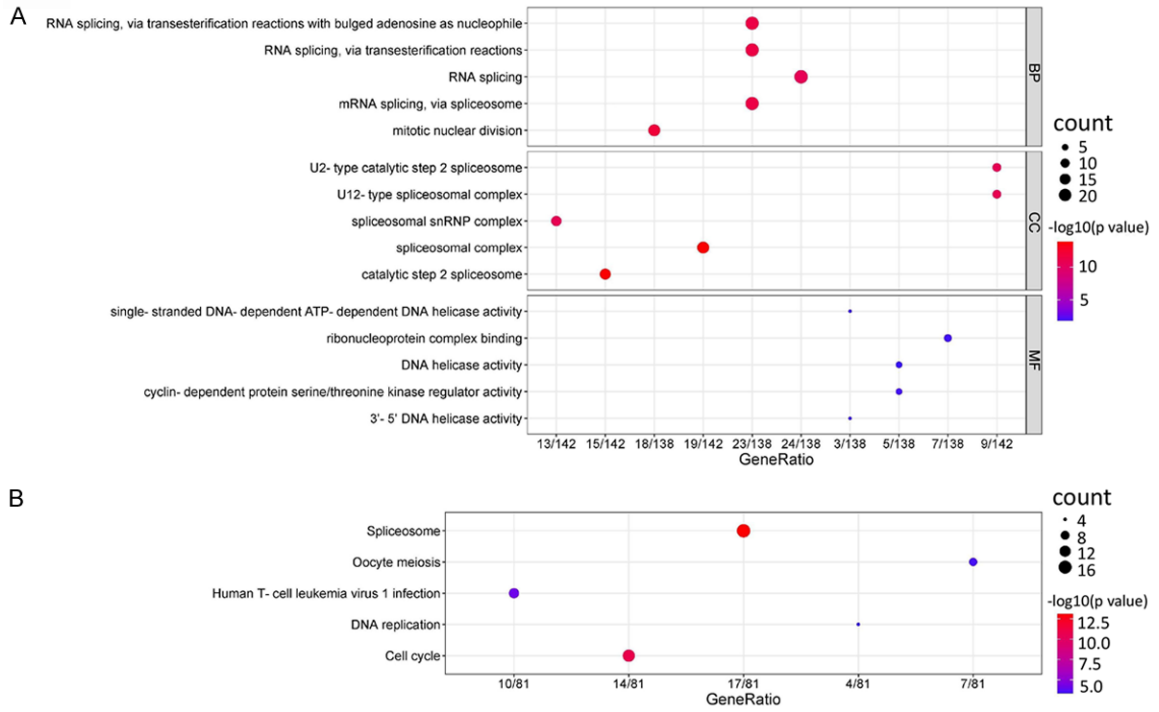


Figure 10. Functional and enrichment analysis for snRNPs and their 50 neighboring genes in HCC (DAVID 6.8). A. GO enrichment analysis for snRNPs genes and their 50 neighboring genes (DAVID 6.8). B. KEGG pathway enrichment analysis for snRNPs genes and their 50 neighboring genes (DAVID 6.8).

was positively associated with these infiltrated cells, including B cells (Cor=0.31, *P*=4.29E-9), $CD8^+$ T cells (Cor=0.218, *P*=4.71E-5), macrophage (Cor=0.274, *P*=2.83E-7), and dendritic cells (Cor=0.25, *P*=3.08E-6) (**Figure 11E**). SNRPF expression was positively related to the infiltration of B cells (Cor=0.336, *P*=1.58E-10), $CD8^+$ T cells (Cor=0.306, *P*=7.83E-9), macrophage (Cor=0.309, *P*=5.77E-9), and dendritic cells (Cor=0.288, *P*=6.28E-8) (**Figure 11F**). SNRPG expression was positively correlated with the infiltration of B cells (Cor=0.385, *P*=1.26E-13), $CD8^+$ T cells (Cor=0.334, *P*=2.27E-10), macrophage (Cor=0.396, *P*=3.22E-14), neutrophil (Cor=0.286, *P*=6.57E-8), and dendritic cells (Cor=0.369, *P*=2.08E-12) (**Figure 11G**).

Furthermore, we also explored the linkages between snRNPs genes and molecular markers of diverse immune cells using the tool of TIMER. The results indicated that SNRPB/D1/G genes were meaningfully related to molecular markers of various immune cells in HCC, including CD19, CD21, and CD70 of B cells; CD8A and CD8B of $CD8^+$ T cells; CD183, CD185, and CD278 of Tfh cells; CD212, CD191 and CD195 of Th1 cells; CD360 and CD196 of Th17 cells; TGFB1 of Treg cells; CXCR4 of Th2 cells; CD163 of TAMs; CD68 and CD11B of macrophage; CD1C, NRP1, and ITGAX of dendritic cells (**Table 3**). In a nutshell, the above findings indicated that snRNPs expression is tightly linked to the immune microenvironment in different ways in HCC patients.

Discussion

HCC is the second leading cause of death from cancer with no satisfactory cure by current therapy approaches [38]. Thus it has significant implications for finding novel prognostic biomarkers of HCC and developing effective treatment strategies. Previous studies have shown that several snRNPs genes are related to some human cancers [4-8]. However, the specific roles of snRNPs in HCC have not yet been fully elucidated. In the present study, we comprehensively explored the mRNA expression and prognostic values of SNRPB/D1/D2/D3/E/F/G and their associations with immune infiltration in HCC patients. We hope, of course, that our research results can be of useful help in the diagnosis and treatment of HCC.

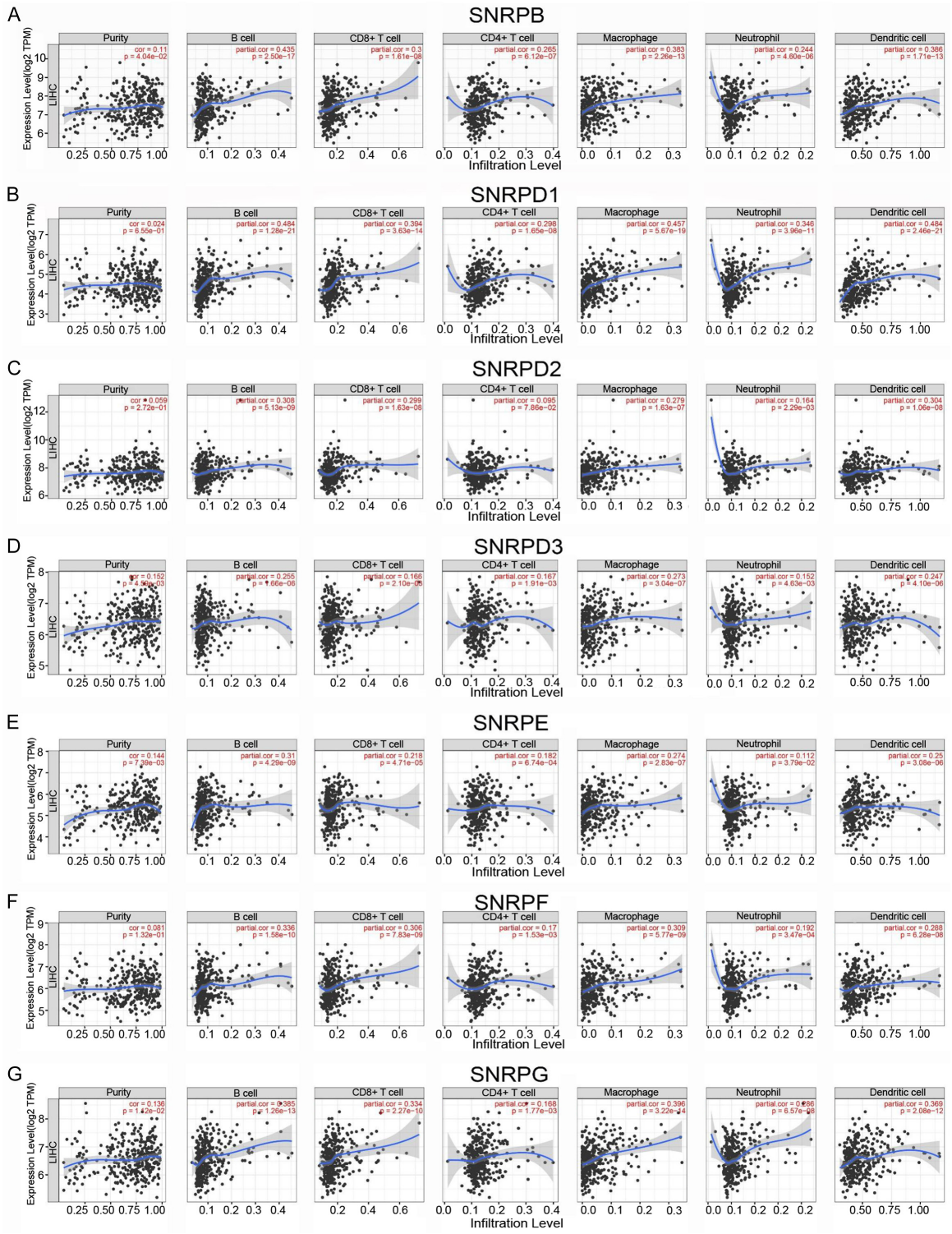


Figure 11. The association between snRNPs expressions and six types of infiltrated immune cells (TIMER).

SNRPB is one of the most studied snRNPs in human cancers. For instance, Zhu L et al. [4] indicated that down-regulated SNRPB inhibited cells proliferation and tumor growth of cervical carcinoma via *in vivo* and *in vitro* experiments. A report from Nianli L et al. [5] showed SNRPB upregulation contributed to non-small cell lung cancer tumorigenesis by regulating RAB26

Table 3. The correlations between SNRPB/D1/D2/D3/E/F/G and gene markers of immune cells

Types	Gene marker	SNRPB		SNRPD1		SNRPD2		SNRPD3		SNRPE		SNRPF		SNRPG	
		Cor	P	Cor	P	Cor	P	Cor	P	Cor	P	Cor	P	Cor	P
B cells	CD19	0.16	**	0.15	*	0.031	0.56	0.11	*	0.1	*	0.09	0.1	0.11	*
	CD21	0.11	*	0.12	*	-0.01	0.97	0.04	0.44	0.1	0.05	0.04	0.79	0.06	*
	CD70	0.33	***	0.21	***	0.055	0.29	0.14	**	0.052	0.32	0.17	***	0.3	***
$CD8^+$ T cells	CD8A	0.27	***	0.22	***	0.12	0.82	0.098	0.06	0.024	0.65	0.02	*	0.26	***
	CD8B	0.34	***	0.29	***	0.055	0.29	0.16	**	0.12	*	0.22	***	0.3	***
Tfh	CD183	0.28	***	0.22	***	0.006	0.91	0.11	*	0.066	0.21	0.12	*	0.23	**
	CD185	0.4	***	0.4	***	0.16	**	0.26	***	0.34	***	0.41	***	0.45	***
	CD278	0.22	***	0.23	***	0.001	0.99	0.073	0.16	0.062	0.23	0.4	0.05	0.22	***
Th1	CD212	0.29	***	0.27	***	0.01	0.85	0.12	*	0.046	0.38	0.13	*	0.31	***
	CD191	0.16	**	0.19	***	-0.18	0.56	0.065	0.22	-0.06	0.25	0.06	0.29	0.25	**
	CD195	0.27	***	0.27	***	0.003	0.96	0.12	*	0.34	0.52	0.12	*	0.27	***
Th17	CD360	0.29	***	0.27	***	0.025	0.64	0.12	*	0.8	0.12	0.11	*	0.31	***
	CD196	0.21	***	0.28	***	-0.03	0.78	0.16	*	0.19	*	0.18	***	0.17	***
Treg	TGFB1	0.38	***	0.4	***	0.13	**	0.21	**	0.19	**	0.23	***	0.35	***
Th2	CXCR4	0.25	***	0.28	***	0.11	0.72	0.14	**	0.11	*	0.14	**	0.26	***
TAM	CD163	0.26	***	0.23	***	0.028	0.6	0.15	**	0.018	0.73	0.15	**	0.3	***
Macrophage	CD68	0.2	***	0.23	***	-0.01	0.88	0.12	*	0.18	0.73	0.09	0.07	0.19	***
	CD11B	0.19	***	0.24	***	0.032	0.54	0.13	*	0.082	0.12	0.09	0.09	0.23	***
Dendritic cells	CD1C	0.13	*	0.2	***	-0.03	0.63	0.079	0.13	0.21	*	0.04	0.47	0.11	*
	NRP1	0.24	***	0.23	***	-0.01	0.96	0.2	***	0.11	0.42	0.07	0.21	0.28	***
	ITGAX	0.17	***	0.23	***	-0.02	0.75	0.08	0.14	-0.01	0.96	0.12	*	0.17	***

TAM: tumor-associated macrophages. *P<0.05, **P<0.01, ***P<0.001.

expression. A mechanistic study also reported that overexpressed SNRPB played a carcinogenic role in the progression of HCC and was mediated by c-Myc [39]. Likewise, in our study, the expression level of SNRPB was found to be upregulated in HCC compared to non-cancer tissues. In addition, SNRPB expression was meaningfully correlated with TP53 mutation, tumor grades, and cancer staging. Highly expressed SNRPB was associated with poor OS in HCC patients, especially in patients without viral hepatitis. Multivariate Cox regression analysis showed that SNRPB was an independent prognostic factor for poor OS of HCC patients. Given the above results, our study revealed that SNRPB promoted oncogenesis in HCC.

In most previous studies, SNRPD1 (smD1) has been generally considered to be related to systemic lupus erythematosus (SLE) [40, 41]. However, only a small number of studies have reported the role of SNRPD1 in human cancers. For instance, research from Bao M et al. [42] indicated that lower SNRPD1 expression was associated with poor survival in patients with ovarian cancer. Dai X et al. [7] found that SNRPD1 upregulation contributed to breast cancer cell proliferation. In this study, it was confirmed that SNRPD1 expression was significantly increased in HCC compared to non-cancer tissues. Furthermore, SNRPD1 expression was relevant to TP53 mutation, tumor grades, and cancer staging. SNRPD1 overexpression was correlated with unfavorable OS in HCC patients, especially in patients with alcohol consumption and those without viral hepatitis. Multivariate Cox regression analysis indicated that SNRPD1 was an independent prognostic factor for poorer OS of HCC patients. Taken together, all of the above findings agreed on the carcinogenic effect of SNRPD1 in HCC.

SNRPD2 is found closely associated with several cancers, including triple-negative breast cancer [43] and hepatocellular carcinoma [44]. Interestingly, A report from Tao Y et al. [45] showed that SNRPD2 was predicted as a pathogenic gene in Alzheimer's disease (AD) by a complex analysis of public databases. However, the specific role SNRPD2 plays in HCC is unknown. In this study, SNRPD2 was found to be overexpressed in HCC. Additionally, high SNRPD2 expression was correlated with TP53 mutation, tumor grades, and cancer staging.

SNRPD3 (smD3), like SNRPD1, is essential in diagnosing SLE [46, 47]. Moreover, several studies have shown that SNRPD3 expression is

relevant to breast cancer and NSCLC [48, 49]. For instance, in the study directed by Siebring-van O E et al. [49], they investigated how silencing SNRPD3 was able to promote p53 gene expression and kill NSCLC cells effectively. However, until now, there have been almost no reports about the functional role of SNRPD3 in HCC. In this study, higher SNRPD3 expression was found in HCC. In addition, SNRPD3 expression was associated with TP53 mutation, tumor grades, and cancer staging. Multivariate Cox regression analysis suggested a meaningful association between overexpressed SNRPD3 with shorter OS of HCC patients, which seemed congruent with the oncogenic character of SNRPD3.

The SNRPE protein is a core component of snRNPs, which has been reported to be related to some malignancies, including bladder cancer [50], non-small cell lung cancer [51], and hepatocellular carcinoma [18, 52]. As an example, Jia D et al. [18] found that SNRPE was a putative oncogene of HCC by a genome-wide analysis. Similarly, in our report, higher SNRPE expression was found in HCC than in the noncancerous tissues. Furthermore, SNRPDE expression was relevant to TP53 mutation and tumor grades. Overexpressed SNRPDE was an independent prognostic factor for poorer OS in HCC patients, especially in patients with alcohol consumption and without viral hepatitis. Taken together, our study revealed that SNRPE promoted oncogenesis in HCC.

SNRPF dysregulation has been found in some human cancers, including colorectal cancer [53], laryngeal squamous cell carcinoma cells [54], and clear cell renal cell carcinoma [55]. For example, a study from Chung F H et al. [53] indicated that SNRPF proved to be a novel biomarker of colorectal cancer via complex gene interaction networks. However, the role of SNRPF was rarely reported in HCC. In our study, highly expressed SNRPF was found in HCC compared to non-cancer tissues. Moreover, SNRPF expression was associated with TP53 mutation, tumor grades, and cancer staging. Multivariate Cox regression analysis indicated that SNRPDF was an independent prognostic factor for poor OS of HCC patients. Given the above results, our study showed that SNRPF played a carcinogenic role in HCC.

SNRPG has been found to be tightly associated with several tumors [56-58]. For instance, a mechanistic study directed by Lan Y et al. [56] showed that downregulated SNRPG could inhabit glioblastoma cell proliferation by the p53 signaling pathway. In this paper, higher mRNA expression of SNRPDG was found in HCC. Furthermore, SNRPD1 expression was relevant to TP53 mutation, tumor grades, and cancer staging. High mRNA expression of SNRPG was associated with poor OS in HCC patients, especially in patients with alcohol consumption and without viral hepatitis. Multivariate Cox regression analysis indicated that SNRPG was an independent prognostic factor for poor OS of HCC patients, suggesting that SNRPG played a carcinogenic role in HCC.

It is worth noting that in this study, we assessed the relationship between snRNPs expression in HCC and the levels of immune infiltration. snRNPs gene expression is meaningfully and positively related to the infiltration of B cells, $CD4^+$ T cells, $CD8^+$ T cells, neutrophils, macrophages, and dendritic cells. Meanwhile, we also found that snRNPs gene expression was significantly linked to molecular markers of multiple immune cells in HCC, such as CD19, CD21, and CD70 of B cells; CD8A and CD8B of $CD8^+$ T cells; and so on. Furthermore, many previous studies also suggested that the spliceosome was tightly associated with the immune microenvironment [15-17]. Nevertheless, further studies are still needed to develop immunosuppressants of individual snRNPs members and apply them to HCC patients.

In this study, there were a few limitations that need to be recognized. Firstly, the data used to evaluate the prognostic values of snRNPs in HCC patients were mainly derived from the TCGA database. Although the TCGA sequencing data was experimentally confirmed, additional large sample studies on HCC patients from other databases are necessary to validate our results. Secondly, the multivariate Cox regression analysis indicated that the expression of SNRPB/D1/D3/E/F/G was independently related to shorter OS in HCC. Consequently, SNRPB/D1/D3/E/F/G were considered as potential prognostic biomarkers for HCC, although overexpressed SNRPD3/F were not meaningfully associated with poor OS in K-M survival analysis. We speculated that these differences might be explained by the presence of confounding factors in this study. The K-M survival analysis describes only the relationship between the univariate and survival and ignores the effects

of other variables. However, in multivariate analysis, regression models are needed to adjust for potential confounders so that the actual effects of the independent variables can be found. Finally, further experiments in cells and animal models are needed to elucidate the underlying mechanisms of how snRNPs play roles in HCC.

Conclusion

In conclusion, our study indicated that SNRPB/D1/D3/E/F/G were potential prognostic biomarkers for short OS of HCC, and SNRPB/D1/G were novel immune therapy targets in HCC patients.

Acknowledgements

We thank Fang Liu and Xiang-Yong Tian of the Department of Pathology, in the Second Affiliated Hospital of Shaanxi University of Chinese Medicine for their help in IHC experiments. Moreover, we acknowledge the TCGA and GEO databases for providing their platforms and contributors for uploading their meaningful datasets. Additionally, we acknowledge the Subject Innovation Team of the Second Affiliated Hospital of Shaanxi University of Chinese Medicine (Grant No. 2020XKTDB01) and Discipline Innovation-team Project of Shaanxi University of Chinese Medicine (Grant No. 2019-SY03).

Disclosure of conflict of interest

None.

Address correspondence to: Yin-Hai Dai, Department of Surgical Oncology Medicine, Second Affiliated Hospital of Shaanxi University of Chinese Medicine, No. 831. Taiguan Road, Xixian New District, Xianyang 712000, Shaanxi, China. Tel: +86-13891488955; E-mail: 1627442702@qq.com

References

[1] Forner A, Reig M and Bruix J. Hepatocellular carcinoma. Lancet 2018; 391: 1301-1314.
[2] Siegel RL, Miller KD, Fuchs HE and Jemal A. Cancer statistics, 2021. CA Cancer J Clin 2021; 71: 7-33.
[3] Chari A, Golas MM, Klingenhager M, Neuenkirchen N, Sander B, Englbrecht C, Sickmann A, Stark H and Fischer U. An assembly chaperone collaborates with the SMN complex to generate spliceosomal SnRNPs. Cell 2008; 135: 497-509.
[4] Zhu L, Zhang XZ and Sun ZQ. SNRPB promotes cervical cancer progression through repressing p53 expression. Biomed Pharmacother 2020; 125: 109948.
[5] Liu NL, Wu ZY, Chen AX, Wang YQ, Cai DF, Zheng JN, Liu Y and Zhang LZ. SNRPB promotes the tumorigenic potential of NSCLC in part by regulating RAB26. Cell Death Dis 2019; 10: 667.
[6] Correa BR, de Araujo PR, Qiao M, Burns SC, Chen C, Schlegel R, Agarwal S, Galante PA and Penalva LO. Functional genomics analyses of RNA-binding proteins reveal the splicing regulator SNRPB as an oncogenic candidate in glioblastoma. Genome Biol 2016; 17: 125.
[7] Dai XF, Yu LH, Chen X and Zhang JY. SNRPD1 confers diagnostic and therapeutic values on breast cancers through cell cycle regulation. Cancer Cell Int 2021; 21: 229.
[8] Zhan YT, Li L, Zeng TT, Zhou NN, Guan XY and Li Y. SNRPB-mediated RNA splicing drives tumor cell proliferation and stemness in hepatocellular carcinoma. Aging (Albany NY) 2020; 13: 537-554.
[9] Aizarani N, Saviano A, Sagar, Mailly L, Durand S, Herman JS, Pessaux P, Baumert TF and Grun D. A human liver cell atlas reveals heterogeneity and epithelial progenitors. Nature 2019; 572: 199-204.
[10] Heymann F and Tacke F. Immunology in the liver--from homeostasis to disease. Nat Rev Gastroenterol Hepatol 2016; 13: 88-110.
[11] Garnelo M, Tan A, Her Z, Yeong J, Lim CJ, Chen J, Lim KH, Weber A, Chow P, Chung A, Ooi LL, Toh HC, Heikenwalder M, Ng IO, Nardin A, Chen Q, Abastado JP and Chew V. Interaction between tumour-infiltrating B cells and T cells controls the progression of hepatocellular carcinoma. Gut 2017; 66: 342-351.
[12] Fu JL, Xu DP, Liu ZW, Shi M, Zhao P, Fu BY, Zhang Z, Yang HY, Zhang H, Zhou CB, Yao JX, Jin L, Wang HF, Yang YP, Fu YX and Wang FS. Increased regulatory T cells correlate with CD8 T-cell impairment and poor survival in hepatocellular carcinoma patients. Gastroenterology 2007; 132: 2328-2339.
[13] Xiao N, Li KS, Zhu XD, Xu B, Liu XF, Lei M and Sun HC. $CD74^+$ macrophages are associated with favorable prognosis and immune contexture in hepatocellular carcinoma. Cancer Immunol Immunother 2022; 71: 57-69.
[14] Cai XY, Gao Q, Qiu SJ, Ye SL, Wu ZQ, Fan J and Tang ZY. Dendritic cell infiltration and prognosis of human hepatocellular carcinoma. J Cancer Res Clin Oncol 2006; 132: 293-301.
[15] Bowling EA, Wang JH, Gong F, Wu W, Neill NJ, Kim IS, Tyagi S, Orellana M, Kurley SJ, Dominguez-Vidana R, Chung HC, Hsu TY, Dubrulle J, Saltzman AB, Li H, Meena JK, Canlas GM, Chamakuri S, Singh S, Simon LM, Olson CM, Do-

brolecki LE, Lewis MT, Zhang B, Golding I, Rosen JM, Young DW, Malovannaya A, Stossi F, Miles G, Ellis MJ, Yu L, Buonamici S, Lin CY, Karlin KL, Zhang XH and Westbrook TF. Spliceosome-targeted therapies trigger an antiviral immune response in triple-negative breast cancer. Cell 2021; 184: 384-403, e21.

[16] Kattah NH, Kattah MG and Utz PJ. The U1-snRNP complex: structural properties relating to autoimmune pathogenesis in rheumatic diseases. Immunol Rev 2010; 233: 126-145.

[17] Yang H, Beutler B and Zhang DW. Emerging roles of spliceosome in cancer and immunity. Protein Cell 2021; [Epub ahead of print].

[18] Jia DS, Wei L, Guo WJ, Zha RP, Bao MY, Chen Z, Zhao YJ, Ge C, Zhao FY, Chen TY, Yao M, Li JJ, Wang HY, Gu JR and He XH. Genome-wide copy number analyses identified novel cancer genes in hepatocellular carcinoma. Hepatology 2011; 54: 1227-1236.

[19] Rhodes DR, Yu JJ, Shanker K, Deshpande N, Varambally R, Ghosh D, Barrette T, Pandey A and Chinnaiyan AM. ONCOMINE: a cancer microarray database and integrated data-mining platform. Neoplasia 2004; 6: 1-6.

[20] Tomczak K, Czerwinska P and Wiznerowicz M. The cancer genome atlas (TCGA): an immeasurable source of knowledge. Contemp Oncol (Pozn) 2015; 19: A68-77.

[21] Clough E and Barrett T. The gene expression omnibus database. Methods Mol Biol 2016; 1418: 93-110.

[22] Tang ZF, Li CW, Kang BX, Gao G, Li C and Zhang ZM. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. Nucleic Acids Res 2017; 45: W98-W102.

[23] Sarkans U, Fullgrabe A, Ali A, Athar A, Behrangi E, Diaz N, Fexova S, George N, Iqbal H, Kurri S, Munoz J, Rada J, Papatheodorou I and Brazma A. From ArrayExpress to BioStudies. Nucleic Acids Res 2021; 49: D1502-D1506.

[24] Chandrashekar DS, Bashel B, Balasubramanya SAH, Creighton CJ, Ponce-Rodriguez I, Chakravarthi BVSK and Varambally S. UALCAN: a portal for facilitating tumor subgroup gene expression and survival analyses. Neoplasia 2017; 19: 649-658.

[25] Gyorffy B, Surowiak P, Budczies J and Lanczky A. Online survival analysis software to assess the prognostic value of biomarkers using transcriptomic data in non-small-cell lung cancer. PLoS One 2013; 8: e82241.

[26] Nagy A, Munkacsy G and Gyorffy B. Pancancer survival analysis of cancer hallmark genes. Sci Rep 2021; 11: 6047.

[27] Cerami E, Gao JJ, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, Jacobsen A, Byrne CJ, Heuer ML, Larsson E, Antipin Y, Reva B, Goldberg AP, Sander C and Schultz N. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov 2012; 2: 401-404.

[28] Gao JJ, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, Sun YC, Jacobsen A, Sinha R, Larsson E, Cerami E, Sander C and Schultz N. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal 2013; 6: pl1.

[29] Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P, Jensen LJ and Mering CV. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res 2019; 47: D607-D613.

[30] Huang da W, Sherman BT and Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 2009; 4: 44-57.

[31] Otasek D, Morris JH, Boucas J, Pico AR and Demchak B. Cytoscape automation: empowering workflow-based network analysis. Genome Biol 2019; 20: 185.

[32] Li B, Severson E, Pignon JC, Zhao HQ, Li TW, Novak J, Jiang P, Shen H, Aster JC, Rodig S, Signoretti S, Liu JS and Liu XS. Comprehensive analyses of tumor immunity: implications for cancer immunotherapy. Genome Biol 2016; 17: 174.

[33] Li TW, Fan JY, Wang BB, Traugh N, Chen QM, Liu JS, Li B and Liu XS. TIMER: a web server for comprehensive analysis of tumor-infiltrating immune cells. Cancer Res 2017; 77: e108-e110.

[34] Roessler S, Jia HL, Budhu A, Forgues M, Ye QH, Lee JS, Thorgeirsson SS, Sun Z, Tang ZY, Qin LX and Wang XW. A unique metastasis gene signature enables prediction of tumor relapse in early-stage hepatocellular carcinoma patients. Cancer Res 2010; 70: 10202-10212.

[35] Chen X, Cheung ST, So S, Fan ST, Barry C, Higgins J, Lai KM, Ji JF, Dudoit S, Ng IO, Van De Rijn M, Botstein D and Brown PO. Gene expression patterns in human liver cancers. Mol Biol Cell 2002; 13: 1929-1939.

[36] Wurmbach E, Chen YB, Khitrov G, Zhang WJ, Roayaie S, Schwartz M, Fiel I, Thung S, Mazzaferro V, Bruix J, Bottinger E, Friedman S, Waxman S and Llovet JM. Genome-wide molecular profiles of HCV-induced dysplasia and hepatocellular carcinoma. Hepatology 2007; 45: 938-947.

[37] Candia J, Bayarsaikhan E, Tandon M, Budhu A, Forgues M, Tovuu LO, Tudev U, Lack J, Chao A, Chinburen J and Wang XW. The genomic landscape of Mongolian hepatocellular carcinoma. Nat Commun 2020; 11: 4383.

[38] Yang JD and Heimbach JK. New advances in the diagnosis and management of hepatocellular carcinoma. BMJ 2020; 371: m3544.
[39] Peng NF, Li JD, He JR, Shi XM, Huang H, Mo YS, Ye H, Wu GB, Wu FX, Xiang BD, Zhong JH, Li LQ and Zhu SL. c-Myc-mediated SNRPB upregulation functions as an oncogene in hepatocellular carcinoma. Cell Biol Int 2020; 44: 1103-1111.
[40] Riemekasten G, Weiss C, Schneider S, Thiel A, Bruns A, Schumann F, Blass S, Burmester GR and Hiepe F. T cell reactivity against the SmD1 (83-119) C terminal peptide in patients with systemic lupus erythematosus. Ann Rheum Dis 2002; 61: 779-785.
[41] Zhao ZH, Ren JL, Dai C, Kannapell CC, Wang HY, Gaskin F and Fu SM. Nature of T cell epitopes in lupus antigens and HLA-DR determines autoantibody initiation and diversification. Ann Rheum Dis 2019; 78: 380-390.
[42] Bao MY, Zhang LH and Hu YQ. Novel gene signatures for prognosis prediction in ovarian cancer. J Cell Mol Med 2020; 24: 9972-9984.
[43] Koedoot E, van Steijn E, Vermeer M, Gonzalez-Prieto R, Vertegaal ACO, Martens JWM, Le Devedec SE and van de Water B. Splicing factors control triple-negative breast cancer cell mitosis through SUN2 interaction and sororin intron retention. J Exp Clin Cancer Res 2021; 40: 82.
[44] Gu Y, Li J, Guo DL, Chen BY, Liu PP, Xiao YS, Yang K, Liu ZS and Liu QY. Identification of 13 Key genes correlated with progression and prognosis in hepatocellular carcinoma by weighted gene co-expression network analysis. Front Genet 2020; 11: 153.
[45] Tao Y, Han Y, Yu LJ, Wang Q, Leng SX and Zhang HY. The Predicted key molecules, functions, and pathways that bridge mild cognitive impairment (MCI) and Alzheimer's disease (AD). Front Neurol 2020; 11: 233.
[46] Mahler M, Fritzler MJ and Bluthner M. Identification of a SmD3 epitope with a single symmetrical dimethylation of an arginine residue as a specific target of a subpopulation of anti-Sm antibodies. Arthritis Res Ther 2005; 7: R19-29.
[47] Mahler M, Waka A, Hiepe F and Fritzler MJ. Effect of dsDNA binding to SmD-derived peptides on clinical accuracy in the diagnosis of systemic lupus erythematosus. Arthritis Res Ther 2007; 9: R68.
[48] Flodrova D, Benkovska D, Macejova D, Bialesova L, Hunakova L, Brtko J and Bobalova J. Proteomic analysis of changes in the protein composition of MCF-7 human breast cancer cells induced by all-trans retinoic acid, 9-cis retinoic acid, and their combination. Toxicol Lett 2015; 232: 226-232.
[49] Siebring-van Olst E, Blijlevens M, de Menezes RX, van der Meulen-Muileman IH, Smit EF and van Beusechem VW. A genome-wide siRNA screen for regulators of tumor suppressor p53 activity in human non-small cell lung cancer cells identifies components of the RNA splicing machinery as targets for anticancer treatment. Mol Oncol 2017; 11: 534-551.
[50] Tapak L, Saidijam M, Sadeghifar M, Poorolajal J and Mahjub H. Competing risks data analysis with high-dimensional covariates: an application in bladder cancer. Genomics Proteomics Bioinformatics 2015; 13: 169-176.
[51] Valles I, Pajares MJ, Segura V, Guruceaga E, Gomez-Roman J, Blanco D, Tamura A, Montuenga LM and Pio R. Identification of novel deregulated RNA metabolism-related genes in non-small cell lung cancer. PLoS One 2012; 7: e42086.
[52] Xu WJ, Huang HX, Yu L and Cao LH. Meta-analysis of gene expression profiles indicates genes in spliceosome pathway are up-regulated in hepatocellular carcinoma (HCC). Med Oncol 2015; 32: 96.
[53] Chung FH, Lee HH and Lee HC. ToP: a trend-of-disease-progression procedure works well for identifying cancer genes from multi-state cohort gene expression data for human colorectal cancer. PLoS One 2013; 8: e65683.
[54] Liu JF, Ma TF, Gao MZ, Liu YL, Liu J, Wang SC, Xie YK, Wen Q, Wang L, Cheng J, Liu SX, Zou J, Wu J, Li WM and Xie HP. Proteomic characterization of proliferation inhibition of well-differentiated laryngeal squamous cell carcinoma cells under below-background radiation in a deep underground environment. Front Public Health 2020; 8: 584964.
[55] Sun X, Zhang HW, Luo LH, Zhong KZ, Ma YS, Fan LL, Fu D and Wan LJ. Comparative proteomic profiling identifies potential prognostic factors for human clear cell renal cell carcinoma. Oncol Rep 2016; 36: 3131-3138.
[56] Lan YL, Lou JC, Hu JL, Yu ZK, Lyu W and Zhang B. Downregulation of SNRPG induces cell cycle arrest and sensitizes human glioblastoma cells to temozolomide by targeting Myc through a p53-dependent signaling pathway. Cancer Biol Med 2020; 17: 112-131.
[57] Lv J, He Y, Li LL and Wang ZH. Alternative splicing events and splicing factors are prognostic in adrenocortical carcinoma. Front Genet 2020; 11: 918.
[58] Wang LW, Zhou W, Li H, Yang H and Shan NC. Clinical significance, cellular function, and potential molecular pathways of CCT7 in endometrial cancer. Front Oncol 2020; 10: 1468.

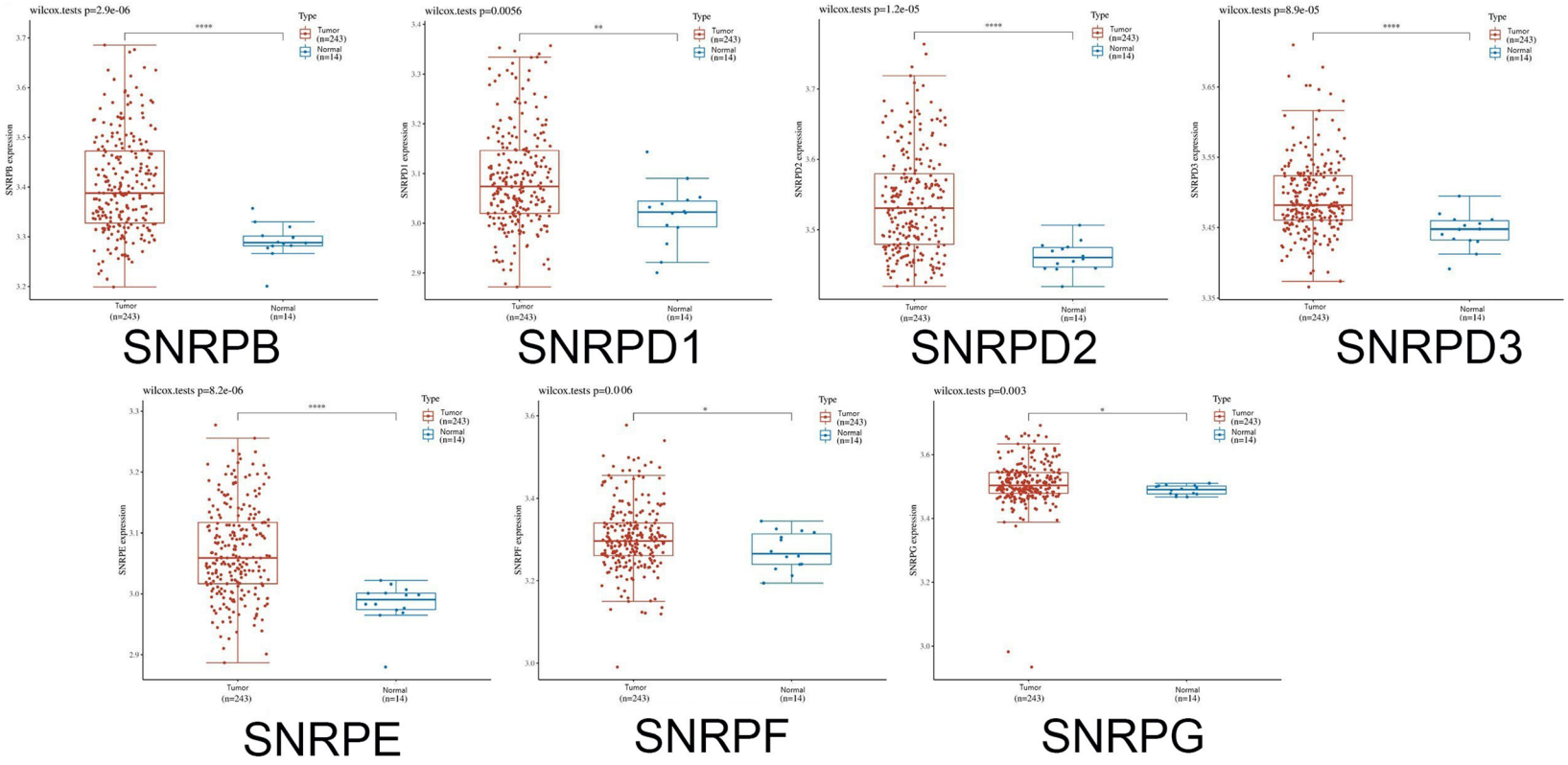


Figure S1. mRNA expression levels of snRNPs in HCC samples and normal liver samples (GEO). *P<0.05, **P<0.01, ***P<0.001, ****P<0.0001.

Table S1. Staining index of 24 patients with HCC in the IHC dataset

Sample	Age	Tissue	Staining index (SNRPB)	Staining index (SNRPD1)	Staining index (SNRPD2)	Staining index (SNRPD3)	Staining index (SNRPE)	Staining index (SNRPF)	Staining index (SNRPG)
1a	35	Tumor	6	6	6	6	5	5	5
1b	35	Normal	0	1	1	0	0	0	1
2a	51	Tumor	5	5	7	6	6	6	6
2b	51	Normal	0	2	0	0	1	1	2
3a	54	Tumor	7	5	6	6	6	6	5
3b	54	Normal	2	2	1	1	2	1	2
4a	42	Tumor	6	6	7	5	7	7	6
4b	42	Normal	1	0	0	2	1	0	0
5a	80	Tumor	7	6	7	6	6	5	6
5b	80	Normal	1	1	1	0	1	0	1
6a	36	Tumor	7	6	5	6	7	6	6
6b	36	Normal	1	2	ND	0	1	0	1
7a	57	Tumor	6	5	6	7	7	7	5
7b	57	Normal	1	2	2	1	1	2	1
8a	68	Tumor	7	7	6	7	6	7	6
8b	68	Normal	0	2	0	0	0	1	2
9a	74	Tumor	7	7	7	7	7	6	5
9b	74	Normal	2	1	1	3	2	2	4
10a	55	Tumor	6	7	6	6	7	6	7
10b	55	Normal	ND	1	0	5	5	5	6
11a	37	Tumor	7	6	7	7	7	6	2
11b	37	Normal	3	2	2	1	2	0	2
12a	34	Tumor	6	6	6	7	6	7	4
12b	34	Normal	1	2	1	2	2	0	1
13a	59	Tumor	6	5	7	6	7	6	7
13b	59	Normal	2	1	0	1	2	0	2
14a	49	Tumor	4	6	5	4	5	7	6
14b	49	Normal	2	3	1	ND	1	2	2
15a	47	Tumor	7	7	6	7	6	5	6
15b	47	Normal	1	1	1	2	ND	0	3
16a	44	Tumor	6	4	5	6	7	7	7
16b	44	Normal	ND	3	0	0	ND	0	2
17a	43	Tumor	7	6	ND	5	6	5	1
17b	43	Normal	0	1	4	1	1	2	7
18a	46	Tumor	6	6	7	7	7	6	5
18b	46	Normal	3	2	2	0	2	1	1
19a	49	Tumor	6	6	6	5	4	1	0
19b	49	Normal	1	0	0	0	3	1	ND
20a	41	Tumor	7	7	6	6	7	4	6
20b	41	Normal	2	1	2	1	1	ND	2
21a	47	Tumor	6	4	5	7	6	6	5
21b	47	Normal	0	ND	1	1	1	2	1
22a	43	Tumor	6	5	6	5	5	7	6
22b	43	Normal	1	0	2	3	2	1	ND
23a	65	Tumor	5	7	7	6	6	5	6
23b	65	Normal	4	1	0	2	1	1	2
24a	47	Tumor	6	5	6	5	7	6	5
24b	47	Normal	2	0	3	2	2	3	1

ND: Not detected.

Table S2. Basic characteristics of the 374 HCC patients

Variables	HCC patients (n=374)
Gender (male/female)	253/121
Age (years)	
≤60	177 (47.5%)
>60	196 (52.5%)
Weight (kg)	
≤70	184 (53.2%)
>70	162 (46.8%)
Albumin (g/dl)	
<3.5	69 (23%)
≥3.5	231 (77%)
AFP (ng/ml)	
≤400	215 (76.8%)
>400	65 (23.2%)
Fibrosis ishak score	
0	75 (34.9%)
1/2	31 (14.4%)
3/4	28 (13%)
5/6	81 (37.7%)
Child-Pugh grade	
A	219 (90.9%)
B	21 (8.7%)
C	1 (0.4%)
Histologic grade	
G1	55 (14.9%)
G2	178 (48.2%)
G3	124 (33.6%)
G4	12 (3.3%)
Pathologic stage	
Stage I	173 (49.4%)
Stage II	87 (24.9%)
Stage III	85 (24.3%)
Stage IV	5 (1.4%)
T stage	
T1	183 (49.3%)
T2	95 (25.6%)
T3	80 (21.6%)
T4	13 (3.5%)
N stage	
N0	254 (98.4%)
N1	4 (1.6%)
M stage	
M0	268 (98.5%)
M1	4 (1.5%)

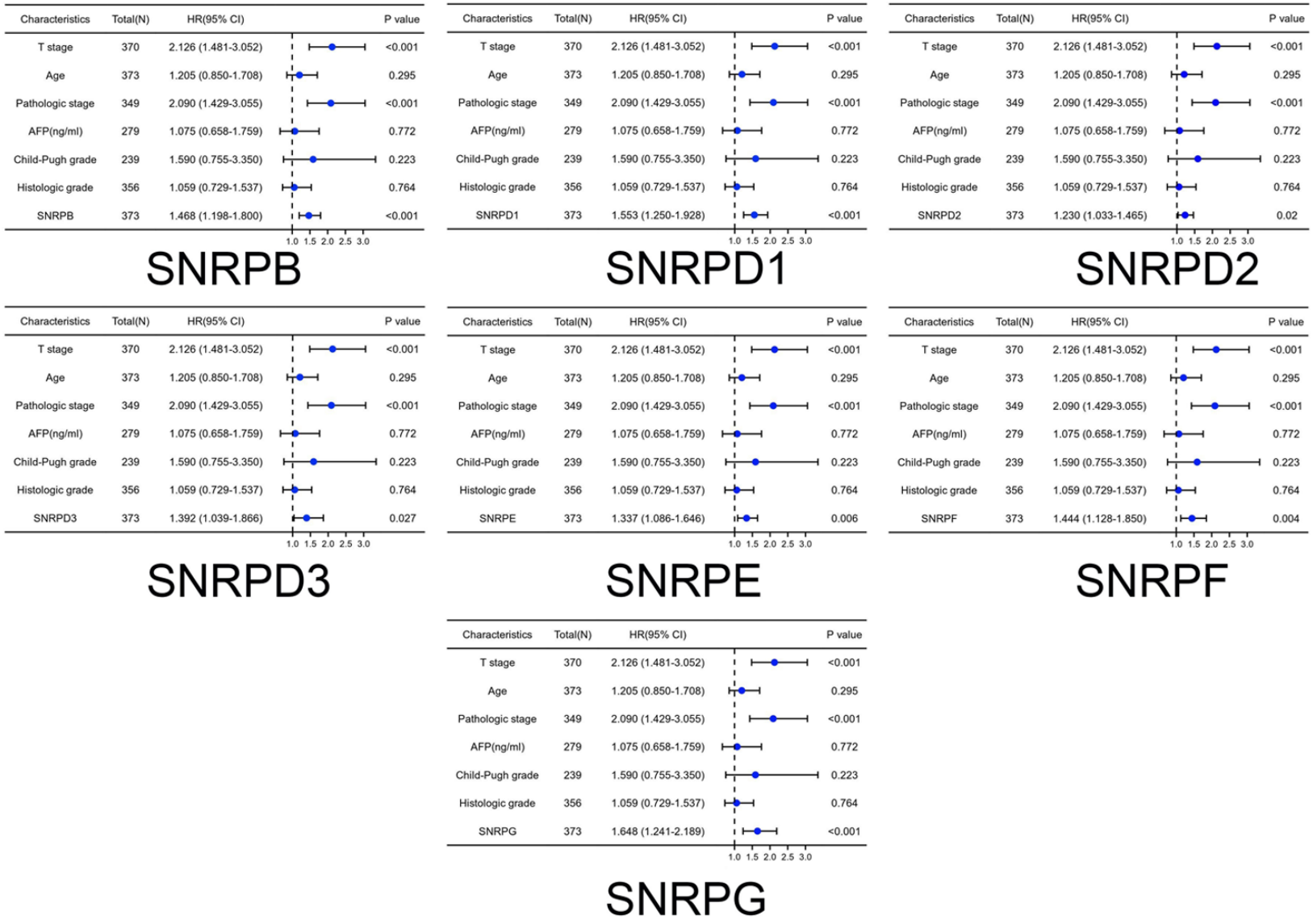


Figure S2. Univariate Cox regression analysis of snRNP expression for OS of HCC patients from the TCGA dataset.

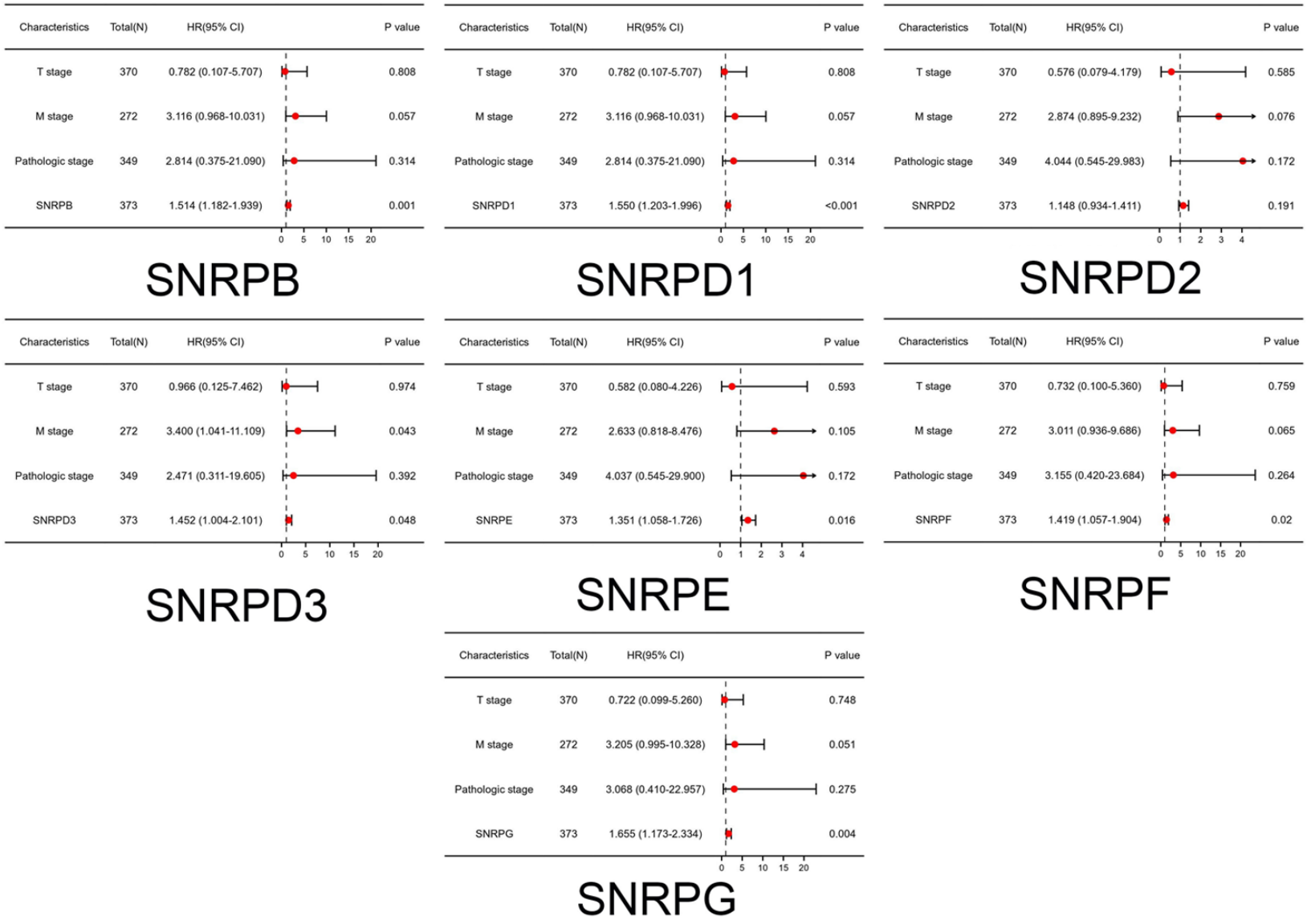


Figure S3. Multivariate Cox regression analysis of snRNP expression for OS of HCC patients from the TCGA dataset.

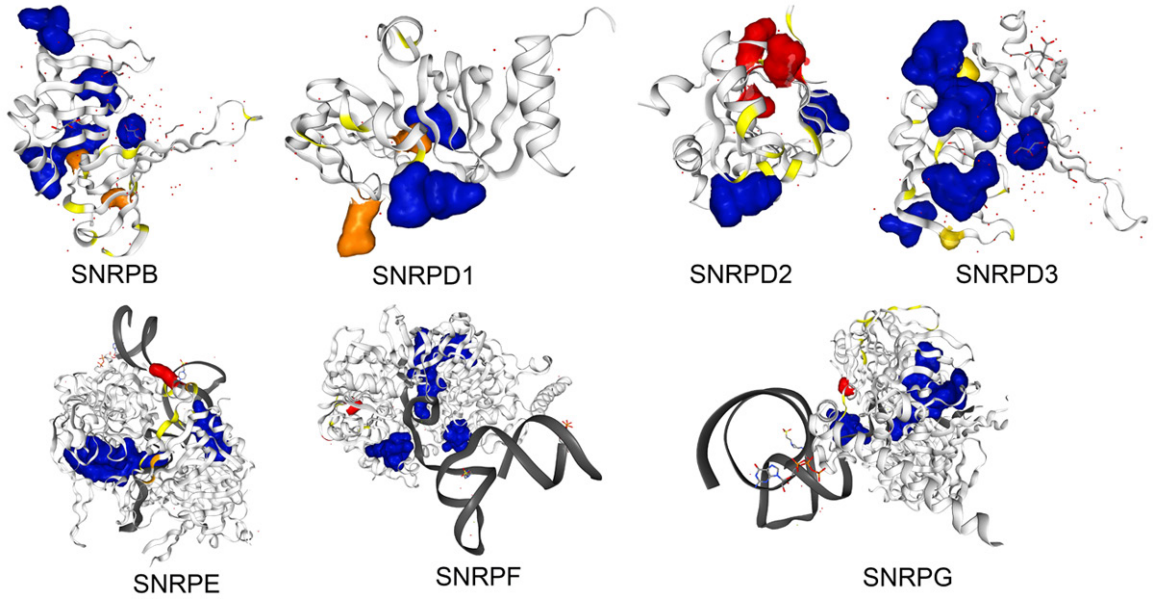


Figure S4. 3D structure of seven snRNPs (COSMIC v94). Blue groups indicate predicted binding sites that are predicted targets for drugs.

Table S3. The specific enriched terms for GO and KEGG enrichment analysis between snRNPs and their 50 most frequently altered neighboring genes (DAVID 6.8)

Pathway	GeneRatio	*P*-value	-log10 (*P*-value)	Count	Class
RNA splicing	24/138	2.58E-11	10.5883802940367	24	BP
RNA splicing, via transesterification reactions with bulged adenosine as nucleophile	23/138	3.89E-12	11.4100503986742	23	BP
mRNA splicing, via spliceosome	23/138	3.89E-12	11.4100503986742	23	BP
RNA splicing, via transesterification reactions	23/138	3.89E-12	11.4100503986742	23	BP
spliceosomal complex	19/142	1.10E-14	13.9586073148417	19	CC
mitotic nuclear division	18/138	1.02E-12	11.991399828238	18	BP
catalytic step 2 spliceosome	15/142	1.10E-14	13.9586073148417	15	CC
spliceosomal snRNP complex	13/142	2.06E-11	10.6861327796308	13	CC
U12-type spliceosomal complex	9/142	1.59E-11	10.7986028756795	9	CC
U2-type catalytic step 2 spliceosome	9/142	2.67E-11	10.5734887386354	9	CC
ribonucleoprotein complex binding	7/138	0.010712703	1.9701009353423	7	MF
cyclin-dependent protein serine/threonine kinase regulator activity	5/138	0.010051082	1.99778718388172	5	MF
DNA helicase activity	5/138	0.014418963	1.84106597259663	5	MF
single-stranded DNA-dependent ATP-dependent DNA helicase activity	3/138	0.014418963	1.84106597259663	3	MF
3'-5' DNA helicase activity	3/138	0.014418963	1.84106597259663	3	MF
Spliceosome	17/81	7.75E-14	13.1106982974936	17	KEGG
Cell cycle	14/81	1.47E-11	10.8326826652518	14	KEGG
Human T-cell leukemia virus 1 infection	10/81	6.11E-05	4.21395878975744	10	KEGG
Oocyte meiosis	7/81	0.000293793	3.53195855605287	7	KEGG
DNA replication	4/81	0.000433406	3.36310508067458	4	KEGG

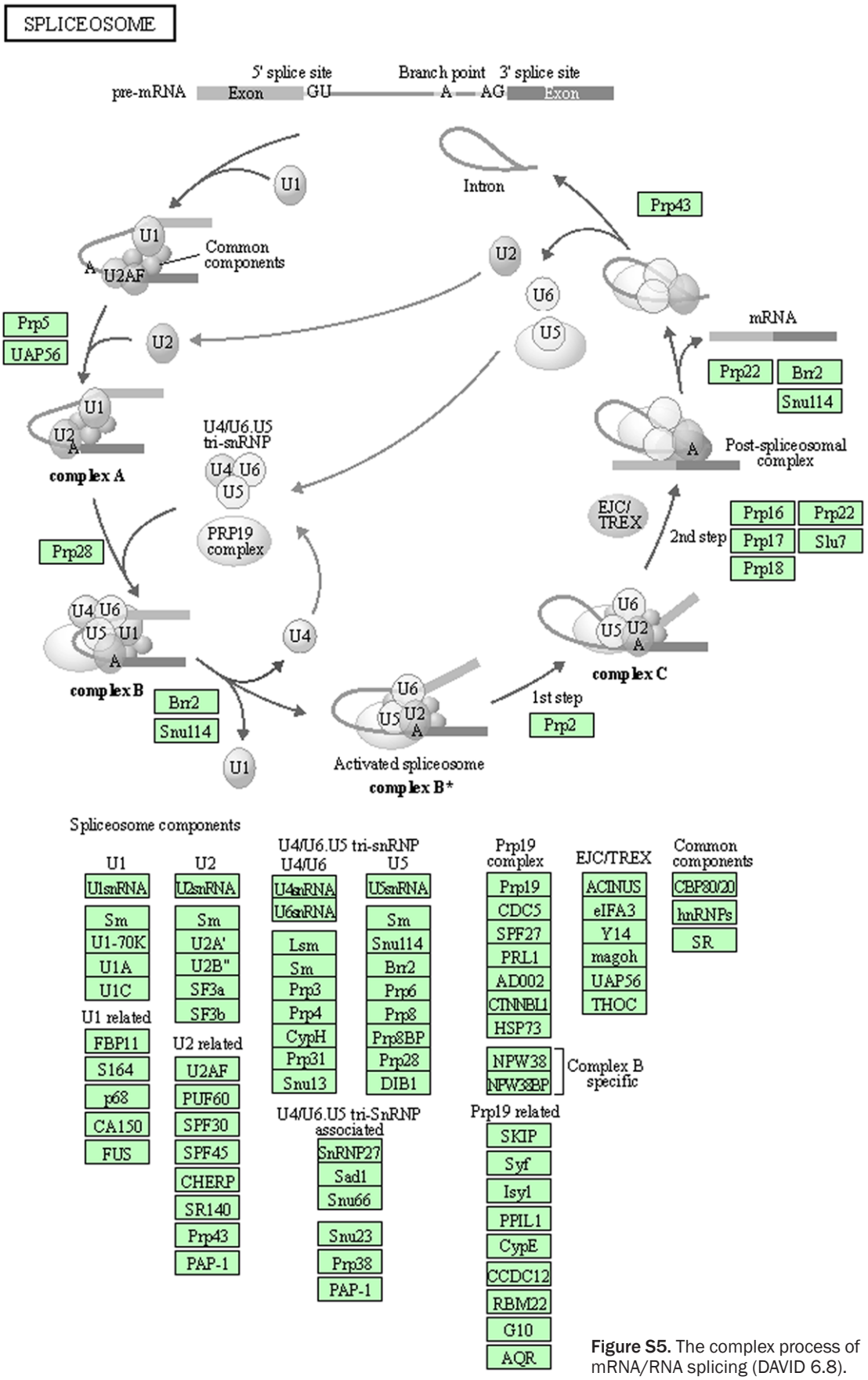


Figure S5. The complex process of mRNA/RNA splicing (DAVID 6.8).