Integrated Analysis of the Functions and Prognostic Values of RNA Binding Proteins and Candidate Drugs in Renal Papillary Cell Carcinoma

Silin Jiang
Department of Urology, the First Affiliated Hospital of Nanjing Medical University
https://orcid.org/0000-0003-4338-0536

Xiaohan Ren
Department of Urology, the First Affiliated Hospital of Nanjing Medical University

Shouyong Liu
Department of Urology, the First Affiliated Hospital of Nanjing Medical University

Zhongwen Lu
Department of Urology, the First Affiliated Hospital of Nanjing Medical University

Aiming Xu
Department of Urology, the First Affiliated Hospital of Nanjing Medical University

Chao Qin
Department of Urology, the First Affiliated Hospital of Nanjing Medical University

Zengjun Wang (zengjunwang@njmu.edu.cn)
The First Affiliated Hospital of Nanjing Medical University https://orcid.org/0000-0002-7583-4750

Primary research

Keywords: Renal papillary cell carcinoma (KIRP), RNA binding protein (RBP), prognostic model, small molecular drugs

Posted Date: November 17th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-105341/v1

Abstract

Background: Roles of RNA binding proteins in renal papillary cell carcinoma (KIRP) remain undiscovered. We thus conducted a series of bioinformatics analyses to elucidate the associations between RBPs and prognosis of renal papillary cell carcinoma.

Methods: RNA sequencing data and clinical of KIRP were downloaded from the TCGA database. The differentially expressed RBP coding genes (DEGs) were sorted out by R software and Gene Ontology (GO) function and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses were then performed to evaluate the functional pathways. Protein-protein interaction (PPI) network of DEGs was formed through the Search Tool for the Retrieval of Interacting Genes (STRING) database and visualized by Cytoscape. Subsequently, a prognostic model was constructed by the uses of univariate Cox regression analysis, random survival forest analysis and multivariate Cox analysis. Validations of Receiver Operating Characteristic (ROC) analysis, KM analysis of overall survival and nomogram were performed as follow. Furthermore, CMap database was used to predict potential drugs.

Results: A prognostic OS-predictive model based on six RBPs (SNRPN, RRS1, INTS8, RBPMS2, IGF2BP3 and PIH1D2) was constructed. STOCK1N-28457, pyrimethamine and trapidil were determined as potential drugs according to the CMap database.

Conclusion: This study constructed a prognosis-related six-RBP signature and made a prediction of three small molecular drugs, which provided a novel insight into KIRP and assisted the development of individualized therapeutic strategies.

Background

Renal cell carcinoma (RCC), which accounts for 3% of adult malignancies, is the most mortal malignancy of urinary system [1]. RCC are mostly consist of three subtypes, renal clear cell carcinoma (KIRC), renal papillary cell carcinoma (KIRP), and renal chromophobe cell carcinoma (KICH) [2]. Renal clear cell carcinoma (KIRC) constitutes 70 percent of all RCC cases, while renal papillary cell carcinoma (KIRP) is the second common subtype of RCC and occupies 15 percent [3]. Clinically, KIRP is considered as more inert than KIRC. However, advanced cases of KIRP subtype indeed has metastatic potential and is more lethal than KIRC [4]. Therefore, a comprehensive analysis of vital genes in KIRP tumorigenesis is typically necessary to evaluate the individual prognosis, figure out therapeutic target and predict potential drugs for poor-prognosis patients.

RNA-binding proteins (RBPs) are a significant group of cellular proteins containing RNA-binding domains and have been widely considered as a key role in posttranscriptional regulation of gene expression, such as RNA shearing, transport, stability, protein translation and subcellular localization [5 6]. Recent studies have revealed links between RBPs and known cancer biomarkers [7 8]. For example, high-risk HPV E7 activates RBP RNASEH2A and PCNA expression and PCNA directs RNASEH2A activity with regard to DNA

replication. The induction of these two factors may promote DNA replication and cancer cell proliferation [9]. High expression of RBP LARP1 co-associate with BCL2 and BIK in BCL2 messenger ribonucleoprotein (mRNP) complexes in epithelial ovarian cancer (EOC) and stabilizes BCL2 while destabilize BIK, which promotes ovarian cancer cell survival and leads to adverse prognosis [10]. RBP NELFE decreases the stabilization of NDRG2 mRNA, which results in epithelial-to-mesenchymal transition through the activation of the Wnt/β-catenin signaling and the promotion the metastasis of pancreatic cancer [11]. Stated thus, defects or dysfunctional of RBP are bounded up with tumorigenesis and tumor prognosis.

As the development of high-throughput sequencing platforms, the vast amount of genomic data are extensively applied for biomarker prediction, prognosis analysis and targeted therapy [12 13]. Bioinformatics analyses provide abundant tools and specific algorithms to obtain, process and interpret biological data [14]. In this study, we conducted a series of bioinformatics analyses based on biological data downloaded from the TCGA database and finally sifted six differentially expressed RBPs associated with KIRP. Our results might provide a new direction for the understanding of progression and prognosis of KIRP.

Materials And Methods

Data Acquisition

The FPKM transcriptome profiling data of 32 normal samples and 289 KIRP samples were obtained from the TCGA database. Clinical data were downloaded from the TCGA database. The RBP genes were obtained from the published literature [15]. RBPs of KIRP were screened when the KIRP transcriptome sequencing map combined with the RBP genes.

Data processing of Differentially Expressed Genes (DEGs)

We used limma package in R software to identify different-expression genes of RBPs between the tumor group and the normal group. The identification was based on cutoffs of | log2 fold change (FC)| >0.5 and false discovery rate (FDR) < 0.05.

Functional Enrichment Analyses and Protein–Protein Interaction Network

Gene ontology (GO), which contained three terms: biological process (BP), cellular component (CC), and molecular function (MF), were applied to investigate the biological functions enrichment. The Kyoto Encyclopedia of Genes and Genomes database (KEGG) was applied to figure out potential biological pathways. All GO and KEGG enrichment analyses were conducted on R software through the clusterprofiler R package with a P-value less than 0.05. The protein–protein interactions (PPIs) among DEGs were checked using The Search Tool for the Retrieval of Interacting Genes (STRING) database [16]. The Cytoscape 3.7.2 software were then put into use for the visualization of PPI network. Subsequently,

the MCODE (Molecular Complex Detection) plug-in in Cytoscape was loaded to filter out significant modules from the PPI network with score > 5 and node counts > 5.

Prognostic Model Construction and Analyses

After establishing a combination by merging gene expression and overall survival, we conducted univariate Cox regression to select prognosis-related RBP genes ($p < 0.01$). We applied randomForestSRC package in R software to conduct the Random survival forests-variable hunting (RSFVH) algorithm to further predict the significant RBP genes from initial screened candidates. Based on these genes, a prognostic model was constructed and the rick score was calculated according to the formula as follows:

Risk Score=

N represents the number of selected genes, is the coefficient of genes in the Cox regression analysis and x equals gene expression value. P values computed by Kaplan–Meier (KM) analysis were then sorted to sift the best combination of six genes. KIRP patients from the TCGA database were divided randomly into training set and test set, and patients in either set were further categorized into high risk group and low risk group according to the median risk score. The survival R package and pROC R package were conducted to construct a ROC curve to measure the accuracy of prognosis. In addition, we plotted a nomogram to calculate the feasibility of overall survival using the nomogramEx R package.

Gene Set Enrichment Analysis

Considered that SNRPN and RRS1 were intersections between critical module 1 of PPI network and prognosis-related combination, GSEA v4.1.0 was downloaded from Broad Institute and Hallmark gene set V6.2 collection were downloaded as target set to perform an analysis of potential mechanism of actions of two genes.

Prediction of Candidate Small Molecular Drugs

We identified differentially expressed RBPs of high and low risk group by applying the limma R package. Then the Connectivity map (CMap) was conducted in order to predict small molecules as potential targeted drugs for KIRP [17].

Results

Exploration of Differentially Expressed Genes (DEGs)

The flowchart of this study is illustrated in Fig. 1. RNA sequencing data containing 32 normal samples and 289 tumor samples of KIRP and clinical data were downloaded from the TCGA database. The list of 1542 RBP genes were acquired from published literature [15]. Finally, 380 RBP-coding genes containing 129 downregulated and 251 upregulated genes met our inclusion criteria of adj $P < 0.05$ and $|log2FC|$ 0.5 (Fig. 2A). Expression of DEGs were plotted in a heatmap to visualize (Fig. 2B).

Functional Enrichment Analyses of DEGs

GO and KEGG analyses were conducted on both downregulated genes and upregulated genes using the clusterProfiler R package. The upregulated RBPs were conspicuously enriched in BPs including ncRNA metabolic process, ncRNA processing, ribonucleoprotein complex biogenesis and RNA splicing while the downregulated RBPs were enriched in regulation of translation, RNA splicing and regulation of cellular amide metabolic process. CC analyses demonstrated that the upregulated RBPs were notably enriched in ribosomal subunit, ribosome, spliceosomal complex and cytosolic ribosome whereas the downregulated RBPs were enriched in cytoplasmic ribonucleoprotein granule, ribonucleoprotein granule and spliceosomal complex. With regard to MF analyses, the results revealed that the upregulated RBPs were significantly enriched in catalytic activity, acting on RNA, ribonuclease activity, nuclease activity and mRNA 3'–UTR binding, and the downregulated RBPs were mainly enriched translation regulator activity, catalytic activity, acting on RNA, translation regulator activity, nucleic acid binding and translation factor activity, RNA binding (Fig. 3A, B). As for the KEGG analyses, we discovered that the upregulated RBPs were largely enriched in the pathways of Ribosome, Spliceosome and RNA transport (Fig. 3C). The downregulated RBPs were primarily enriched in RNA transport pathway and mRNA surveillance pathway (Fig. 3D).

PPI Network Construction and Crucial Module

PPI network was constructed by using the STRING database and visualized by applying the Cytoscape software (Fig. 4A). This PPI network comprised 346 nodes and 3164 edges. We further extract three modules with the use of the plug-in MCODE in cytoscape according to the cutoff of node counts > 5 and score > 5 (Fig. 4B). Module 1 consisted of 85 nodes and 1366 edges, module 2 comprised 13 nodes and 33 edges and module 3 included 15 nodes and 38 edges (Fig. 4C, D, E). The results of GO and KEGG for three modules revealed that genes in module 1 were mainly enriched in ribonucleoprotein complex biogenesis, cytosolic ribosome, structural constituent of ribosome and Ribosome pathway, genes in module 2 were enriched in DNA alkylation, chromatoid body, regulatory RNA binding and MicroRNAs in cancer, whereas the genes in module 3 were significantly enriched in mitochondrial translational elongation, organellar large ribosomal subunit, structural constituent of ribosome and RNA transport pathway (Table 1, 2).

Table 1
The GO functional enrichment of three crucial modules.

ONTOLOGY	ID	Description	pvalue	p.adjust	Count
Module 1					
BP	GO:0022613	ribonucleoprotein complex biogenesis	2.13E-33	1.09E-30	33
BP	GO:0000956	nuclear-transcribed mRNA catabolic process	6.85E-26	1.38E-23	22
BP	GO:0000184	nuclear-transcribed mRNA catabolic process, nonsense-mediated decay	8.14E-26	1.38E-23	19
CC	GO:0022626	cytosolic ribosome	8.69E-25	8.60E-23	18
CC	GO:0044391	ribosomal subunit	1.74E-20	8.63E-19	18
CC	GO:0022625	cytosolic large ribosomal subunit	1.27E-19	4.18E-18	13
MF	GO:0003735	structural constituent of ribosome	2.72E-19	3.02E-17	18
MF	GO:0090079	translation regulator activity, nucleic acid binding	8.11E-16	4.50E-14	13
MF	GO:0008135	translation factor activity, RNA binding	1.38E-15	5.10E-14	12
Module 2					
BP	GO:0006305	DNA alkylation	9.31E-10	2.84E-07	5
BP	GO:0006306	DNA methylation	9.31E-10	2.84E-07	5
BP	GO:0044728	DNA methylation or demethylation	3.62E-09	7.38E-07	5
CC	GO:0033391	chromatoid body	6.66E-14	2.46E-12	5
CC	GO:0036464	cytoplasmic ribonucleoprotein granule	2.32E-09	3.88E-08	6
CC	GO:0035770	ribonucleoprotein granule	3.14E-09	3.88E-08	6
MF	GO:0061980	regulatory RNA binding	9.85E-09	6.01E-07	4
MF	GO:0140098	catalytic activity, acting on RNA	4.37E-08	1.02E-06	6
MF	GO:0004521	endoribonuclease activity	5.04E-08	1.02E-06	4
Module 3					
BP	GO:0070125	mitochondrial translational elongation	4.46E-11	7.24E-09	6

ONTOLOGY	ID	Description	pvalue	p.adjust	Count
BP	GO:0070126	mitochondrial translational termination	4.78E-11	7.24E-09	6
BP	GO:0006415	translational termination	1.24E-10	1.25E-08	6
CC	GO:0000315	organellar large ribosomal subunit	4.96E-10	9.17E-09	5
CC	GO:0005762	mitochondrial large ribosomal subunit	4.96E-10	9.17E-09	5
CC	GO:0000313	organellar ribosome	4.32E-09	3.99E-08	5
MF	GO:0003735	structural constituent of ribosome	1.51E-05	0.000377	4
MF	GO:0008536	Ran GTPase binding	0.000424	0.005294	2
MF	GO:0003724	RNA helicase activity	0.001686	0.014054	2

Table 2
The KEGG functional enrichment analysis of three crucial modules.

ID	Description	pvalue	p.adjust	Count
Module 1				
hsa03010	Ribosome	2.38E-18	4.04E-17	18
hsa03040	Spliceosome	8.99E-12	7.64E-11	13
hsa03013	RNA transport	2.78E-06	1.58E-05	9
hsa03015	mRNA surveillance pathway	3.74E-05	0.000159	6
hsa03008	Ribosome biogenesis in eukaryotes	8.00E-05	0.000272	6
hsa05134	Legionellosis	0.000505	0.001431	4
hsa03020	RNA polymerase	0.001054	0.00256	3
Module 2				
hsa05206	MicroRNAs in cancer	0.001028	0.011308	3
Module 3				
hsa03013	RNA transport	3.94E-06	1.97E-05	4
hsa03040	Spliceosome	0.00496	0.009032	2
hsa03010	Ribosome	0.005419	0.009032	2

Prognostic Risk Score Model Construction and Validation

A combination of 346 RBPs from PPI network and overall survival was put into univariate Cox regression analysis to confirm prognosis-related RBPs (Fig. 5A). 60 RBPs were sorted out by the cutoff of P value < 0.01. The randomForestSRC R package was applied to carry on a Random survival forest analysis in order to distinguish RBP genes with best association with prognosis and 10 genes (EXO1, RBPMS2, PABPN1L, PIH1D2, INTS8, RRS1, CPSF4L, IGF2BP3, SNRPN and NPM3) were screened out from the 60 prognosis-related RBPs (Fig. 5B). Multivariate Cox analysis were subsequently conducted to establish a prognostic model related to overall survival and we further performed a Kaplan–Meier analysis on the 2^{10}-1 = 1023 models formed by 10 genes for the purpose of determining the best risk score model (Fig. 5C). Comparing the -log_{10}log rank p value of these 1023 models, we finally sorted out the probnostic risk score model containing six RBPs (SNRPN, RRS1, INTS8, RBPMS2, IGF2BP3 and PIH1D2). The risk score of each KIRP patient was calculated: Risk score = (0.2729931 * SNRPN) + (0.9340297 * RRS1) + (1.8014324 * INTS8) + (-0.5129049 * RBPMS2) + (1.7546410 * IGF2BP3) + (-0.4098881 * PIH1D2).

In order to evaluate the predictive capabilities of the model, we allocated KIRP patients into trainging set and test set equally, and then patients in each set were divided into high risk group and low risk group in consifderation of the median risk score. Expressions of survival status and heatmap of each set were also shown (Fig. 5D, G). Receiver operating characteristic (ROC) analyses were utilized for further estimate on the prognostic model. The AUC values in the training set was 0.9 at 1 year, 0.87 at 3 years and 0.78 at 5 years together with 0.88 at 1 year, 0.75 at 3 years and 0.69 at 5 years in the test set (Fig. 5E, H). The results revealed that patients in the high risk group tend to a significantly lower survival probability than those in low risk group (Fig. 5F, I).

In order to appraise clinical factors in prognosis, we then conducted an independent prognostis-related analysis on the training set and test set by using univariate and multivariate COX regression analyses. The univariate results revealed that, in both training set and test set, the stage and T staging could be considered as independent prognostic factors for overall survival of KIRP patients (Fig. 6A, B). In multivariate analysis, tumor stage and T staging could be considered as independent prognostic factors in the training set (Fig. 6C). Nevertheless, only tumor stage can be taken as independent prognostic factor in the test set (Fig. 6D). Finally, the nomogram were constructed with six selected genes and tumor stage to evaluate the mortality risk of 3 and 5 years (Fig. 6E). Furthermore, we plotted calibration curves which demonstrated ideal conformity between speculated outcomes and observed outcomes (Fig. 6F, G).

Gene Set Enrichment Analysis

We figured out that SNRPN and RRS1 were intersections between prognostic model and module 1 of PPI network and then made a Kaplan–Meier analysis for overall survival (Fig. 7A, B). Given the levels of SNRPN were negatively correlated with the survival while RRS1 was positively correlated with the survival, GSEA analysis was applied in the low-expression and high-expression groups. As to RRS1 high-expression group, the genes sets in hallmark collection were enriched in DNA repair, E2F targets, G2M checkpoints, MTORC1 signaling, MYC targets and unfolded protein response (Fig. 7D). As to SNRPN low-

expression group, the gene sets were enriched in mitotic spindle, E2F targets and G2M checkpoints (Fig. 7C).

Screening of Candidate Small Molecular Drugs

We applied the limma R package to identified differential-expression RBPs of different risk groups, and then 297 RBPs consisted of 261 upregulated and 36 downregulated RBPs reached the threshold of adj P < 0.05 and $|\log_2 FC| > 1$. The prediction of small molecular drugs was based on the 297 RBPs. Finally, three small molecules which were STOCK1N-28457, pyrimethamine and trapidil were cherry-picked according to the enrichment score (> 0.6), P value (< 0.05) and percent non-null (> 70) (Table 3).

Table 3
Results of CMap analysis

cmap name	mean	n	enrichment	p	specificity	percent non-null
STOCK1N-28457	-0.717	3	-0.93	0.00052	0.0051	100
pyrimethamine	-0.678	5	-0.773	0.00108	0.0067	80
Trapidil	-0.58	3	-0.889	0.00266	0.0124	100
AH-6809	0.679	2	0.934	0.00833	0	100
mercaptopurine	-0.61	2	-0.908	0.01714	0.0137	100
7-aminocephalosporanic acid	-0.389	4	-0.699	0.01717	0.0311	75
lobelanidine	0.477	4	0.689	0.01936	0.0067	75
5230742	-0.648	2	-0.897	0.0209	0.0383	100
emetine	-0.42	4	-0.677	0.02401	0.2765	75
spaglumic acid	0.598	2	0.891	0.02463	0.0397	100
exisulind	0.524	2	0.891	0.02487	0	100
loxapine	0.47	4	0.667	0.02751	0.0106	75
debrisoquine	0.441	4	0.64	0.041	0.0204	75
altizide	0.413	4	0.639	0.04132	0	75
oxybenzone	0.431	4	0.634	0.04416	0.1706	75

Discussion

Recently, RBPs were becoming progressovely more important by the profoung study conducted on its roles in various cancers and increasingly regarded as crucial factors in posttranscriptional regulation [18–

[20]. The dysfunction of posttranscriptional regulation that were related with the origination of cancer could enhance the function mutations of oncogenes and depresse mutations of tumor suppressor [21; 22]. To the best of our knowledge, this was the first study focused on the role of RBPs in the progression and prognosis of KIRP. Here, we integrated RNA sequencing data of KIRP from the TCGA database and sorted out differentially expressed RBPs between tumor and normal patients. We further coducted GO and KEGG enrichment analyses and established PPI network for these RBPs. Moreover, we constructed a OS-predictve model for predicting the prognosis of KIRP patients and performed ROC analyses to evaluate the feasibility of our model. Subsequently, GSEA was conducted to figure out the biological functions of two selected RBPs.

As for the results of biological functions and pathway enrichment analyses, DEGs were enriched in ribosome and posttranscriptional modification pathways, such as RNA splicing, RNA transport, spliceosome and translation. Large numbers of studies in recent years have reported that aberrant RNA modification and RNA metabolism were of great value in various cancers [23 24]. Alternative RNA splicing events, which were probably adjusted by RBPs, were reported prevalent in liver cancer and affected tumorigenesis in metabolism-related pathways by investigations conducted by Li, S research team [25]. In addition, the manipulation of alternative splicing was proved to be a new method to suppress tumorigenesis in glioblastoma by Mogilevsky and their colleagues [26]. Will research team discovered that spliceosome was consist of five snRNPs and numerous proteins which catalyze pre-mRNA splicing [27]. Varialbe levels of RNA and protein components influenced splice cite and research conducted by Dvinge and their partners showed that the spliceosome shaped the global transcriptome of breast cancer [28]. In ovarian cancers, Li et al. found that spliceosome could to promote proliferation and invasion by the upregulation of associate factor [29]. Previous studies have shown the relationship and possible mechanisms between RBPs and spliceosome [30 31].

Subsequently, through the application of univariate Cox regression analysis, random survival forest analysis, multivariate Cox analysis and Kaplan–Meier test, we determine six RBP-coding genes: SNRPN, RRS1, INTS8, RBPMS2, IGF2BP3 and PIH1D2. The risk score model was then constructed to predict the prognosis of patients. It was noteworthy that patients with high risk scores tend to have worse prognosis, implying that individual therapeutic schedules should be taken into consideration. The ROC curve of risk score model revealed that the six-RBP signature was comparatively reliable in predicting prognosis with the AUC values of 0.87 and 0.75 at 3 years as well as 0.78 and 0.69 at 5 years in training set and test set. A nomogram comprised independent prognostic factor and six-RBP signature was established to assist the prediction of 3-year and 5-year overall survival in clinical treatments.

In addition, we performed GSEA on SNRPN and RRS1 since they were concurrently parts of subnetwork of PPI. SNRPN (small nuclear ribonucleoprotein polypeptide N) was widely regarded as a spliceosome component [32]. As to the results of GSEA, low-expression SNRPN was enriched in E2F targets. E2F was a family of transcription factors and had various functions such as controlling the cell cycle, regulating transcription and apoptosis [33]. More importantly, E2F targets were reported to play significant roles in

several cancers. For example, Park research team announced that E2F targets were activated by EPEL to promote cell proliferation of lung cancer [34]. Meanwhile, Sun and their colleagues made a research on the roles of E2Fs in breast cancer and considered E2F4 and 6 as biomarkers, together with E2F1, 3, 5, 7 and 8 as potential targets of therapy [35]. The inhibition of E2F downregulated the ability of BRD4 binding with the promoter of miR-106b-5p and inhibited its transcription which resulted in the cellular senescence of gastric cancer cells were investigated by Dong, X team [36]. High-expression RRS1 was enriched mainly in mTORC1 signaling. The mTORC1 signaling pathway was a classical pathway connecting to tumorigenesis [37]. According to studies of He et al. and Guigon et al., the growth of pancreatic cancer and thyroid cancer were inhibited by the suppression of mTORC1 signaling [38 39]. Experiments on the the interaction mechanism of TRAF6 and p62 were carried on by Linares and their colleagues. Results revealed the importance for lung cancer cell proliferation through the activation of mTORC1 [40]. Last but not least, we made a prediction of potential small molecular drugs which may be of therapeutic benefits for KIRP patients and had a certain degree of reliability.

Overall, our study dissected the relationship between RBPs and KIRP for the first time and proposed a novel direction for exploring the tumorigenesis and prognosis of KIRP. We determined 6 RBPs which were linked with prognosis, constructed a reliable prognostic OS-predictive model and speculated three potentially useful drugs. The six RBPs could act as potential therapeutic targets of KIRP and contribute to the development of clinical treatment. Nevertheless, our study had several limitations. First, our prognostic model was only constructed on the TCGA database and lack of clinical data from the GEO database to evaluate. Meanwhile, the lack of clinical characteristics of clinical data from TCGA may decrease the credibility of our research. Moreover, our results were based on RNA sequencing, patients had the possiblity to exhibit inter-individual heterogeneity. Finally, prospective clinical studies had the necessarity to condut before our 6-RBP prognostic model put into use.

Conclusion

We applied a series of bioinformatics analysis on the aberrantly expressed RBPs, which were affiliated with tumorigenesis, invasion and prognosis, in order to investigate their potential functions, action pathways and prognostic values. Subsequently, we identified six RBPs highly associated with prognosis of KIRP and constructed a six-RBP prognostic model to predict overall survival as well as optimize the predictive ability of stage system. Moreover, we selected two important RBPs and evaluate their biological effects and made a prediction of potential drugs. To the best of our knowledge, this study represented the first study focused on prognostic values of RBPs in KIRP and provided new insight into pathogenesis and therapeutic strategies of KIRP.

Abbreviations

KIRP
renal papillary cell carcinoma

RBP
RBA binding protein
DEGs
differentially expressed genes
GO
Gene Ontology
KEGG
The Kyoto Encyclopedia of Genes and Genomes database
PPI
Protein-protein interaction
FC
Fold change
FDR
False discovery rate
ROC
Receiver operating characteristic
AUC
Area under the curves
GSEA
Gene Set Enrichment Analysis
OS
Overall survival
TCGA
The Cancer Genome Atlas

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Data Availability Statement

All datasets about TCGA are publicly available and included in the article.

Declaration of Competing Interest

The study was conducted with no existance of competing interests.

Funding

The study was supported by the National Natural Science Foundation of China (81771640).

Authors' Contributions

Zengjun Wang: conception and design of the study, funding acquisition. Silin Jiang and Xiaohan Ren: data acquisition, bioinformatics analysis, drafting and critical revision of the manuscript. Shouyong Liu and Zhongwen Lu: visualization and investigation. Aiming Xu: validation. All authors approved the final manuscript.

Acknowledgements

Not applicable.

References

1. Huang Q, Sun Y, Ma X, Gao Y, Li X, Niu Y, Zhang X, Chang C. Androgen receptor increases hematogenous metastasis yet decreases lymphatic metastasis of renal cell carcinoma. Nature communications. 2017;8:918.
2. Tabibu S, Vinod PK, Jawahar CV. Pan-Renal Cell Carcinoma classification and survival prediction from histopathology images using deep learning. Scientific reports. 2019;9:10509.
3. Al Ahmad A, Paffrath V, Clima R, Busch JF, Rabien A, Kilic E, Villegas S, Timmermann B, Attimonelli M, Jung K, et al. (2019). Papillary Renal Cell Carcinomas Rewire Glutathione Metabolism and Are Deficient in Both Anabolic Glucose Synthesis and Oxidative Phosphorylation. Cancers 11.
4. Kaldany A, Paulucci DJ, Kannappan M, Beksac AT, Anastos H, Okhawere K, Sfakianos JP, Badani KK. Clinicopathological and survival analysis of clinically advanced papillary and chromophobe renal cell carcinoma. Urol Oncol. 2019;37:727–34.
5. Burd CG, Dreyfuss G. Conserved structures and diversity of functions of RNA-binding proteins. Science. 1994;265:615–21.
6. Glisovic T, Bachorik JL, Yong J, Dreyfuss G. RNA-binding proteins and post-transcriptional gene regulation. FEBS Lett. 2008;582:1977–86.
7. Kechavarzi B, Janga SC. Dissecting the expression landscape of RNA-binding proteins in human cancers. Genome biology. 2014;15:R14.
8. Pereira B, Billaud M, Almeida R. RNA-Binding Proteins in Cancer: Old Players and New Actors. Trends in cancer. 2017;3:506–28.
9. Xu J, Liu H, Yang Y, Wang X, Liu P, Li Y, Meyers C, Banerjee NS, Wang HK, Cam M, et al. (2019). Genome-Wide Profiling of Cervical RNA-Binding Proteins Identifies Human Papillomavirus Regulation of RNASEH2A Expression by Viral E7 and E2F1. mBio 10.

10. Hopkins TG, Mura M, Al-Ashtal HA, Lahr RM, Abd-Latip N, Sweeney K, Lu H, Weir J, El-Bahrawy M, Steel JH, et al. The RNA-binding protein LARP1 is a post-transcriptional regulator of survival and tumorigenesis in ovarian cancer. Nucleic acids research. 2016;44:1227–46.
11. Han L, Zan Y, Huang C, Zhang S. NELFE promoted pancreatic cancer metastasis and the epithelial–to–mesenchymal transition by decreasing the stabilization of NDRG2 mRNA. Int J Oncol. 2019;55:1313–23.
12. Azad RK, Li J. Interpreting genomic data via entropic dissection. Nucleic acids research. 2013;41:e23.
13. Gounder MM, Nayak L, Sahebjam S, Muzikansky A, Sanchez AJ, Desideri S, Ye X, Ivy SP, Nabors LB, Prados M, et al. Evaluation of the Safety and Benefit of Phase I Oncology Trials for Patients With Primary CNS Tumors. Journal of clinical oncology: official journal of the American Society of Clinical Oncology. 2015;33:3186–92.
14. Psarros M, Heber S, Sick M, Thoppae G, Harshman K, Sick B. RACE: Remote Analysis Computation for gene Expression data. Nucleic acids research. 2005;33:W638–43.
15. Gerstberger S, Hafner M, Tuschl T. A census of human RNA-binding proteins. Nature reviews Genetics. 2014;15:829–45.
16. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, Sun Y, Jacobsen A, Sinha R, Larsson E, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal. 2013;6:pl1.
17. Lamb J, Crawford ED, Peck D, Modell JW, Blat IC, Wrobel MJ, Lerner J, Brunet JP, Subramanian A, Ross KN, et al. The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. 313. New York: Science; 2006. pp. 1929–35.
18. Soni S, Anand P, Padwad YS. MAPKAPK2: the master regulator of RNA-binding proteins modulates transcript stability and tumor progression. Journal of experimental clinical cancer research: CR. 2019;38:121.
19. Dong W, Dai ZH, Liu FC, Guo XG, Ge CM, Ding J, Liu H, Yang F. The RNA-binding protein RBM3 promotes cell proliferation in hepatocellular carcinoma by regulating circular RNA SCD-circRNA 2 production. EBioMedicine. 2019;45:155–67.
20. Singh AK, Aryal B, Zhang X, Fan Y, Price NL, Suárez Y, Fernández-Hernando C. Posttranscriptional regulation of lipid metabolism by non-coding RNAs and RNA binding proteins. Semin Cell Dev Biol. 2018;81:129–40.
21. Masuda K, Kuwano Y. Diverse roles of RNA-binding proteins in cancer traits and their implications in gastrointestinal cancers. Wiley interdisciplinary reviews RNA. 2019;10:e1520.
22. Vogelstein B, Kinzler KW. The Path to Cancer –Three Strikes and You're Out. N Engl J Med. 2015;373:1895–8.
23. Delaunay S, Frye M. RNA modifications regulating cell fate in cancer. Nat Cell Biol. 2019;21:552–9.
24. Li Y, Sahni N, Pancsa R, McGrail DJ, Xu J, Hua X, Coulombe-Huntington J, Ryan M, Tychhon B, Sudhakar D, et al. Revealing the Determinants of Widespread Alternative Splicing Perturbation in Cancer. Cell reports. 2017;21:798–812.

25. Li S, Hu Z, Zhao Y, Huang S, He X. Transcriptome-Wide Analysis Reveals the Landscape of Aberrant Alternative Splicing Events in Liver Cancer. Hepatology. 2019;69:359–75.
26. Mogilevsky M, Shimshon O, Kumar S, Mogilevsky A, Keshet E, Yavin E, Heyd F, Karni R. Modulation of MKNK2 alternative splicing by splice-switching oligonucleotides as a novel approach for glioblastoma treatment. Nucleic acids research. 2018;46:11396–404.
27. Will CL, Lührmann R. (2011). Spliceosome structure and function. Cold Spring Harbor perspectives in biology 3.
28. Dvinge H, Guenthoer J, Porter PL, Bradley RK. RNA components of the spliceosome regulate tissue- and cancer-specific alternative splicing. Genome research. 2019;29:1591–604.
29. Li Y, Guo H, Jin C, Qiu C, Gao M, Zhang L, Liu Z, Kong B. Spliceosome-associated factor CTNNBL1 promotes proliferation and invasion in ovarian cancer. Experimental cell research. 2017;357:124–34.
30. Naro C, Sette C. (2013). Phosphorylation-mediated regulation of alternative splicing in cancer. International journal of cell biology 2013, 151839.
31. Sutandy FXR, Ebersberger S, Huang L, Busch A, Bach M, Kang HS, Fallmann J, Maticzka D, Backofen R, Stadler PF, et al. In vitro iCLIP-based modeling uncovers how the splicing factor U2AF2 relies on regulation by cofactors. Genome research. 2018;28:699–713.
32. Jing J, Zhao Y, Wang C, Zhao Q, Liang Q, Wang S, Ma J. Effect of small nuclear ribonucleoprotein-associated polypeptide N on the proliferation of medulloblastoma cells. Mol Med Rep. 2015;11:3337–43.
33. Johnson DG, Schneider-Broussard R. (1998). Role of E2F in cell cycle control and cancer. Frontiers in bioscience: a journal and virtual library 3, d447-448.
34. Park SM, Choi EY, Bae DH, Sohn HA, Kim SY, Kim YJ. (2018). The LncRNA EPEL Promotes Lung Cancer Cell Proliferation Through E2F Target Activation. Cellular physiology and biochemistry: international journal of experimental cellular physiology, biochemistry, and pharmacology 45, 1270–1283.
35. Sun CC, Li SJ, Hu W, Zhang J, Zhou Q, Liu C, Li LL, Songyang YY, Zhang F, Chen ZL, et al. Comprehensive Analysis of the Expression and Prognosis for E2Fs in Human Breast Cancer. Molecular therapy: the journal of the American Society of Gene Therapy. 2019;27:1153–65.
36. Dong X, Hu X, Chen J, Hu D, Chen LF. BRD4 regulates cellular senescence in gastric cancer cells via E2F/miR-106b/p21 axis. Cell death disease. 2018;9:203.
37. Poburski D, Leovsky C, Boerner JB, Szimmtenings L, Ristow M, Glei M, Thierbach R. Insulin-IGF signaling affects cell transformation in the BALB/c 3T3 cell model. Scientific reports. 2016;6:37120.
38. He R, Yin Y, Yin W, Li Y, Zhao J, Zhang W. Prevention of pancreatic acinar cell carcinoma by Roux-en-Y Gastric Bypass Surgery. Nature communications. 2018;9:4183.
39. Guigon CJ, Fozzatti L, Lu C, Willingham MC, Cheng SY. Inhibition of mTORC1 signaling reduces tumor growth but does not prevent cancer progression in a mouse model of thyroid cancer. Carcinogenesis. 2010;31:1284–91.

40. Linares JF, Duran A, Yajima T, Pasparakis M, Moscat J, Diaz-Meco MT. K63 polyubiquitination and activation of mTOR by the p62-TRAF6 complex in nutrient-activated cells. Molecular cell. 2013;51:283–96.

Figures

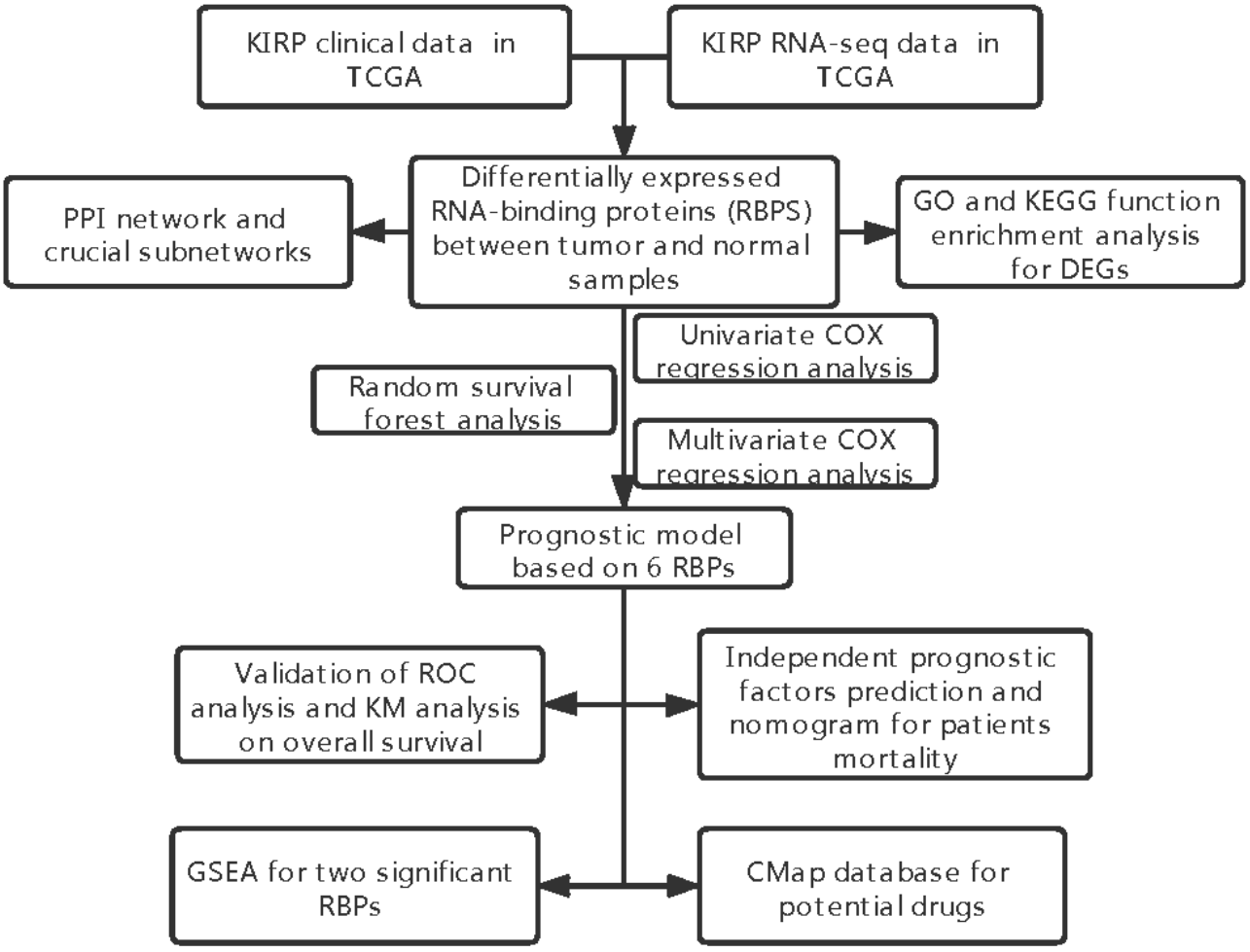

Figure 1

Flowchart of analyzation on RBPs in renal papillary cell carcinoma (KIRP).

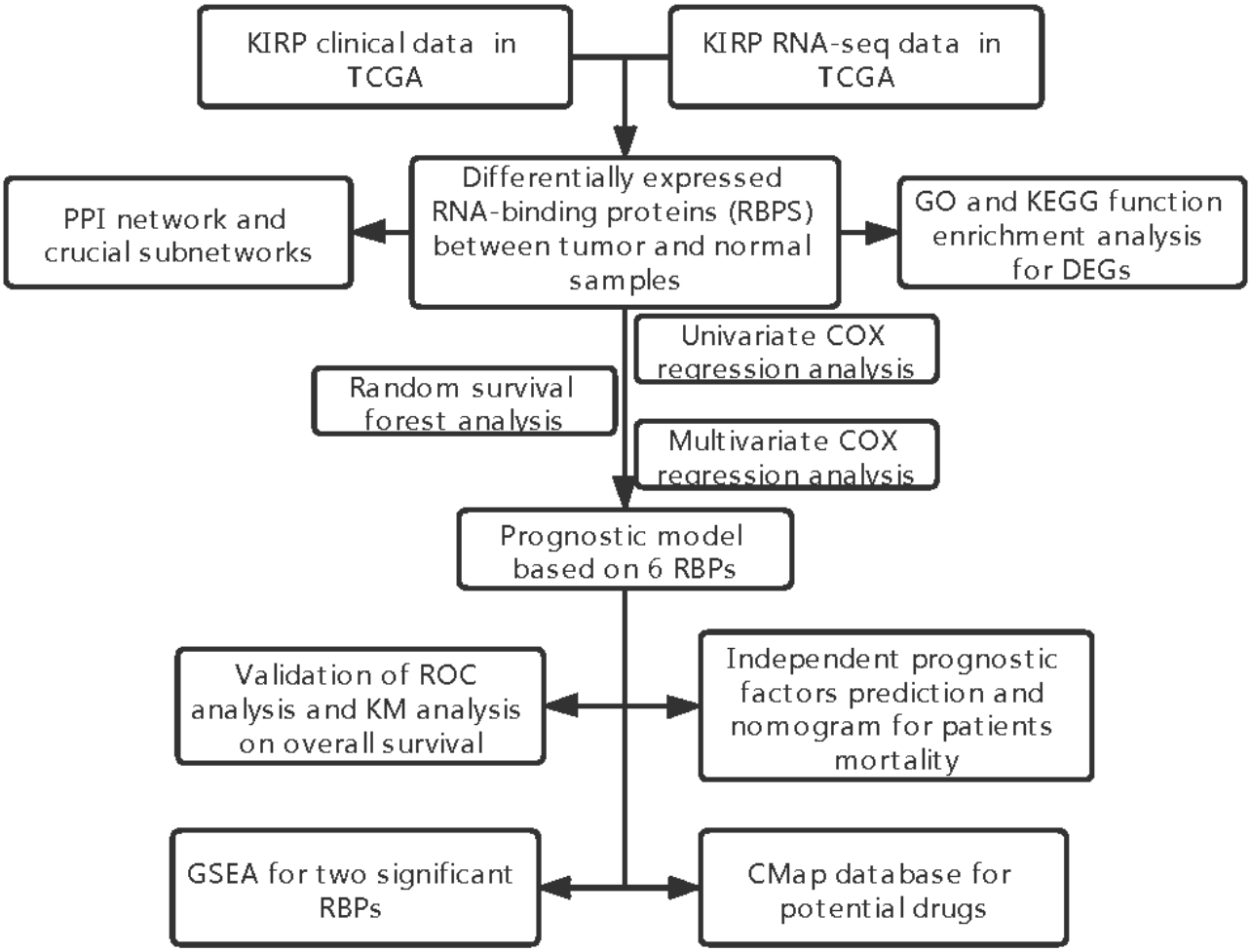

Figure 1

Flowchart of analyzation on RBPs in renal papillary cell carcinoma (KIRP).

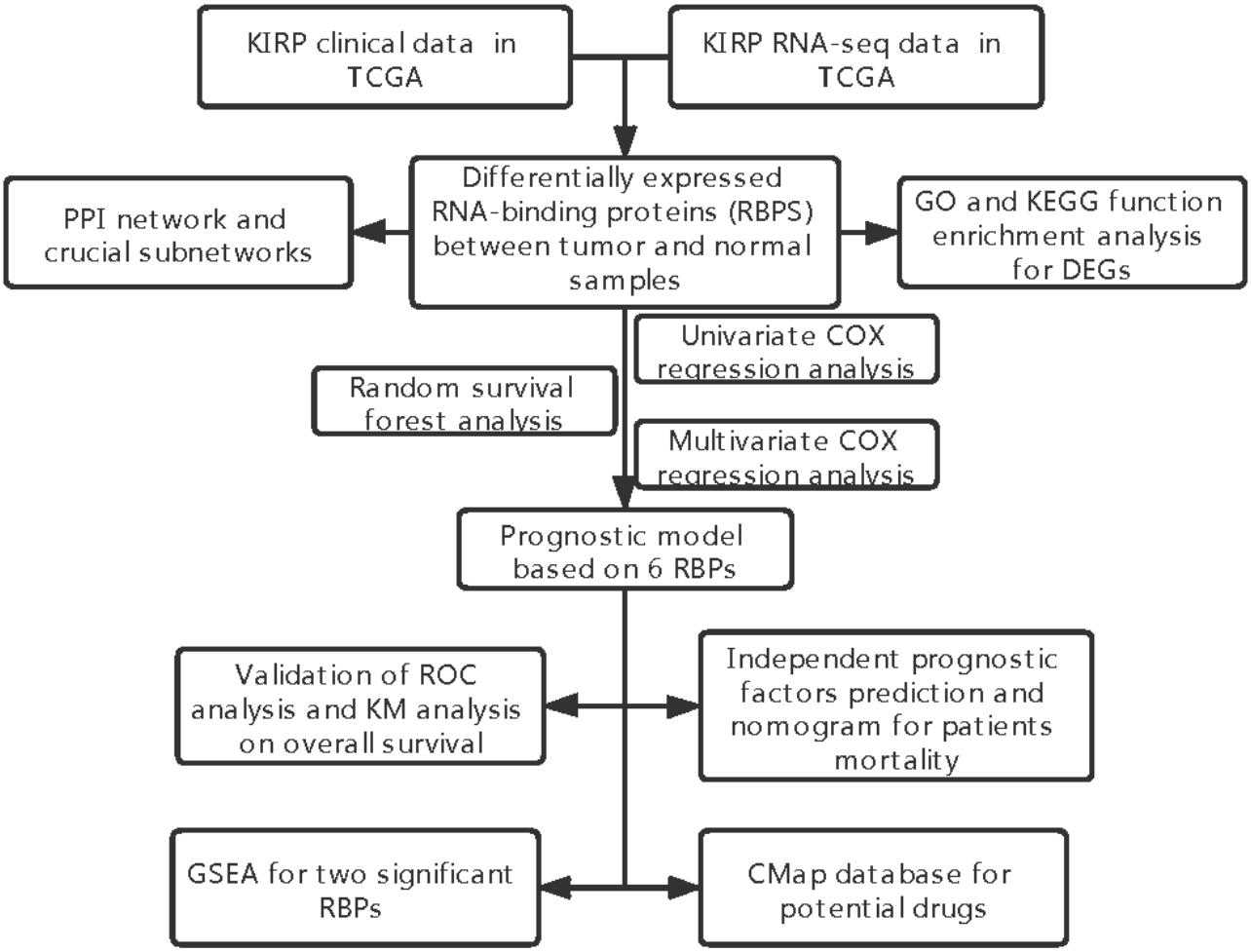

Figure 1

Flowchart of analyzation on RBPs in renal papillary cell carcinoma (KIRP).

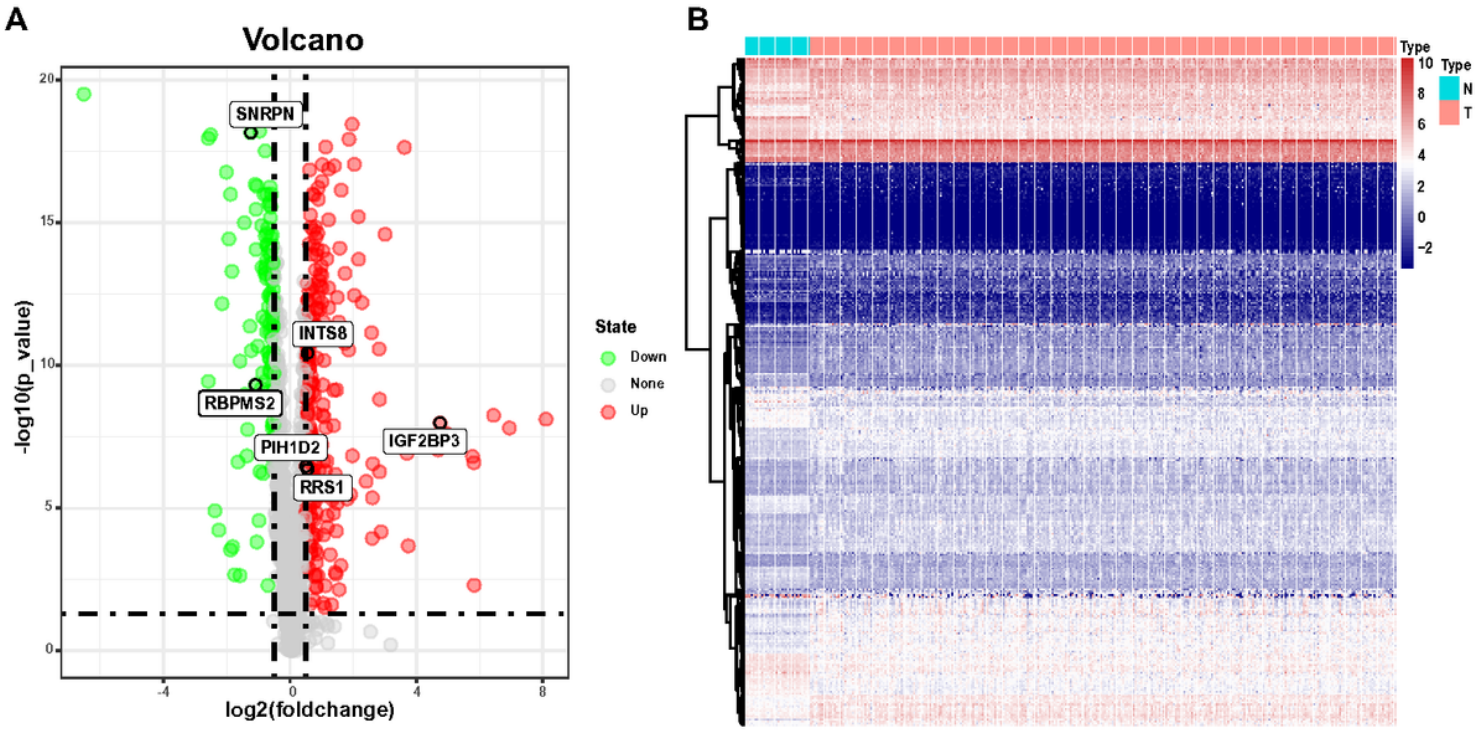

Figure 2

The differentially expressed RBPs. (A) Volcano plot of differentially expressed RBPs reaching the threshold of adj P <0.05 and |log2FC| ≥ 0.5. (B) Heatmap plot of 120 differentially RBPs expression between tumor and normal samples in KIRP.

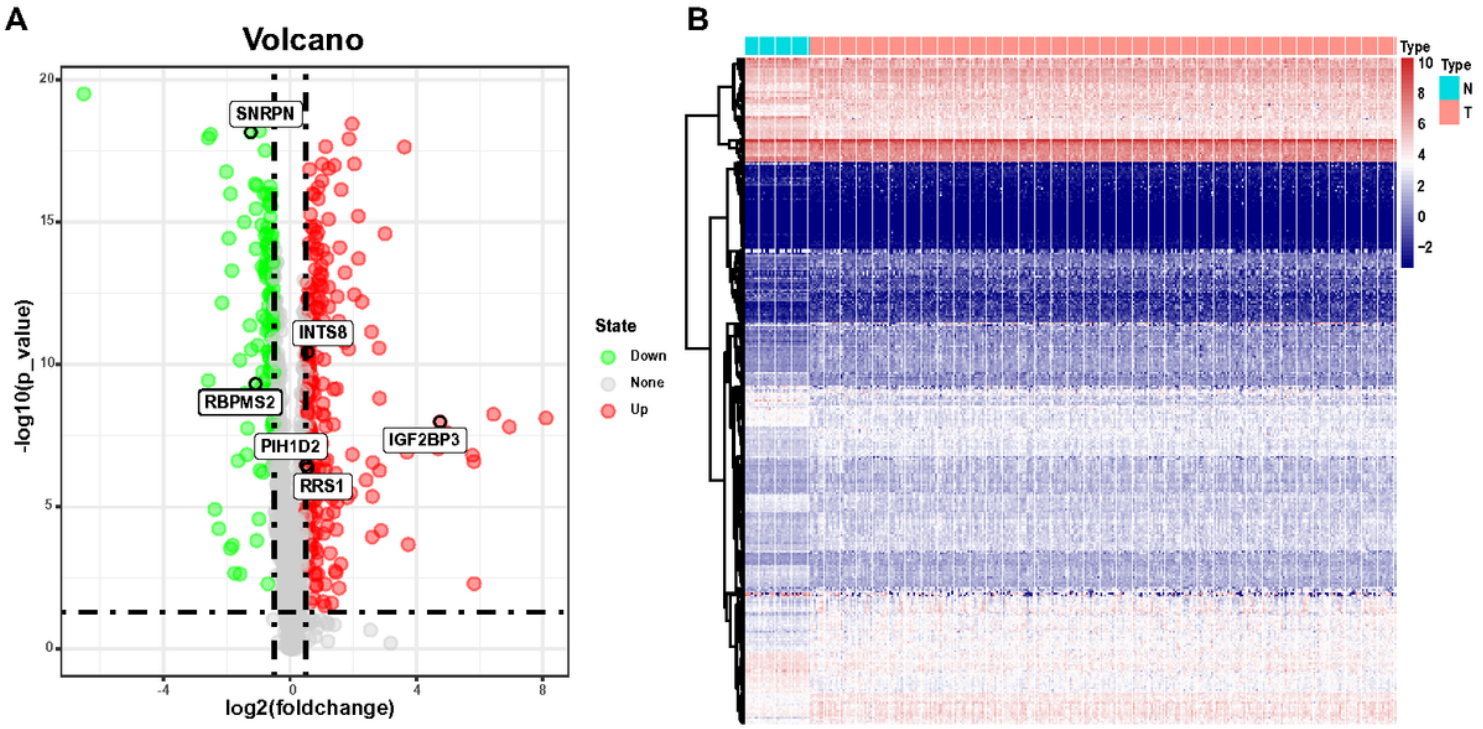

Figure 2

The differentially expressed RBPs. (A) Volcano plot of differentially expressed RBPs reaching the threshold of adj P <0.05 and |log2FC| ≥ 0.5. (B) Heatmap plot of 120 differentially RBPs expression between tumor and normal samples in KIRP.

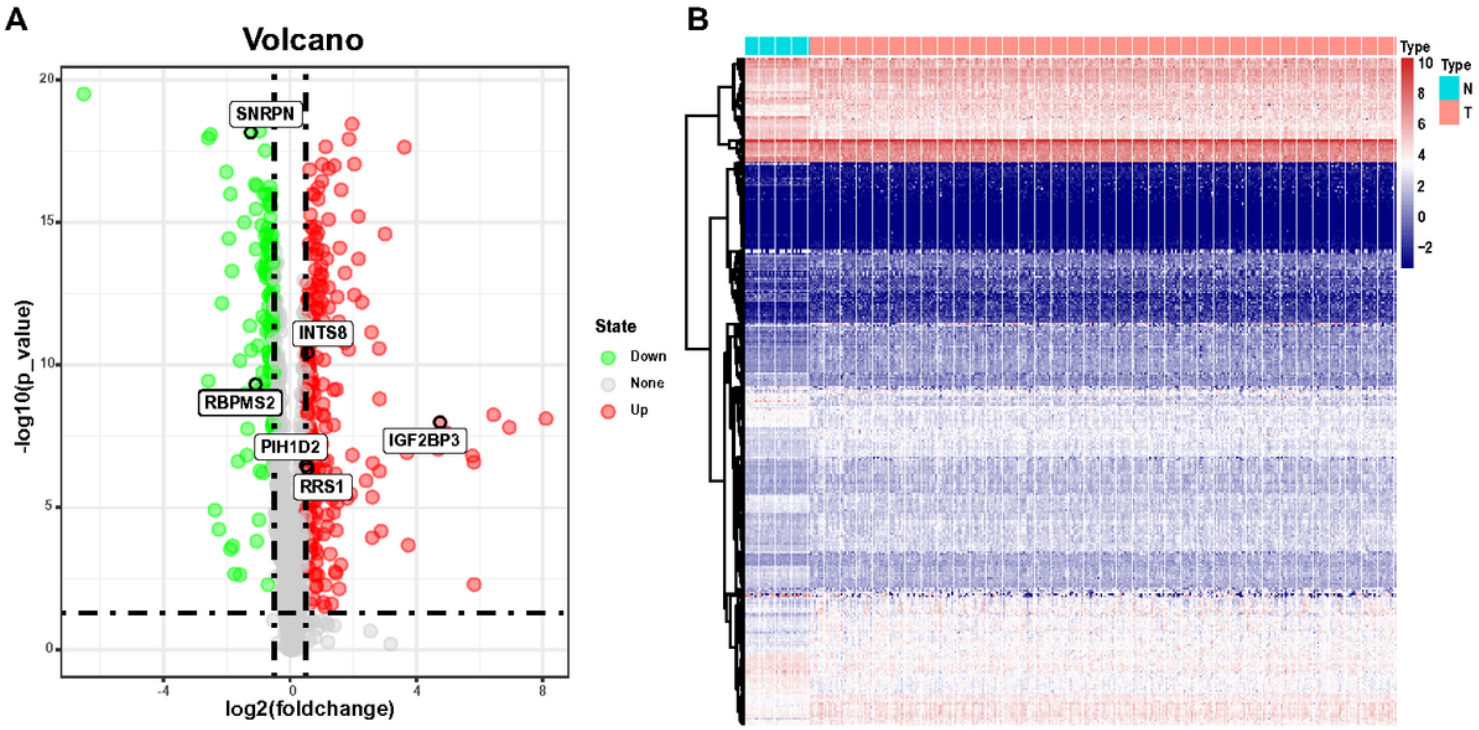

Figure 2

The differentially expressed RBPs. (A) Volcano plot of differentially expressed RBPs reaching the threshold of adj P <0.05 and |log2FC| ≥ 0.5. (B) Heatmap plot of 120 differentially RBPs expression between tumor and normal samples in KIRP.

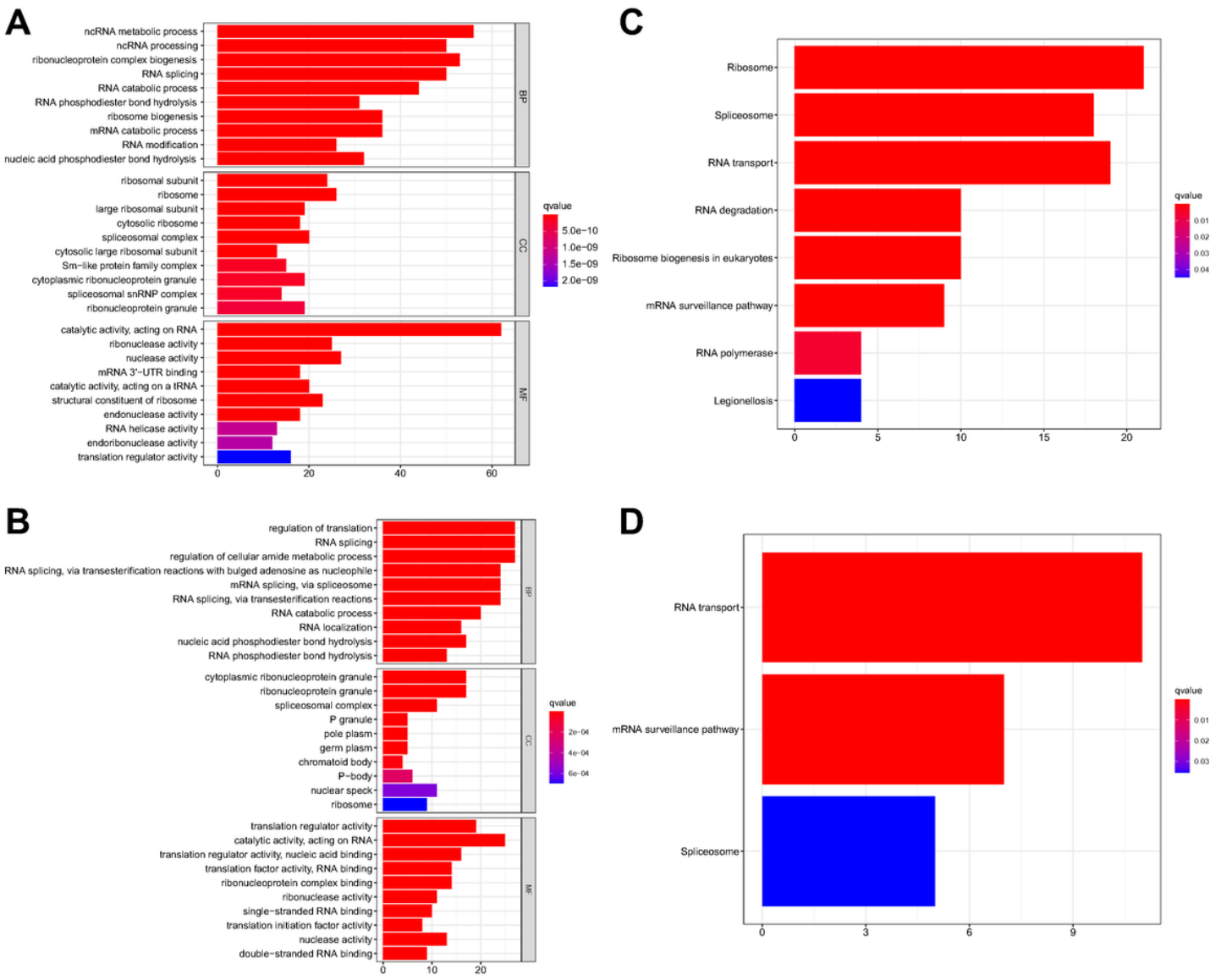

Figure 3

Functional enrichment analyses of 251 upregulated and 129 downregulated RBPs. (A) GO analysis on upregulated RBPs. (B) GO analysis on downregulated RBPs. (C) KEGG analysis on upregulated RBPs. (D) KEGG analysis on downregulated RPS.

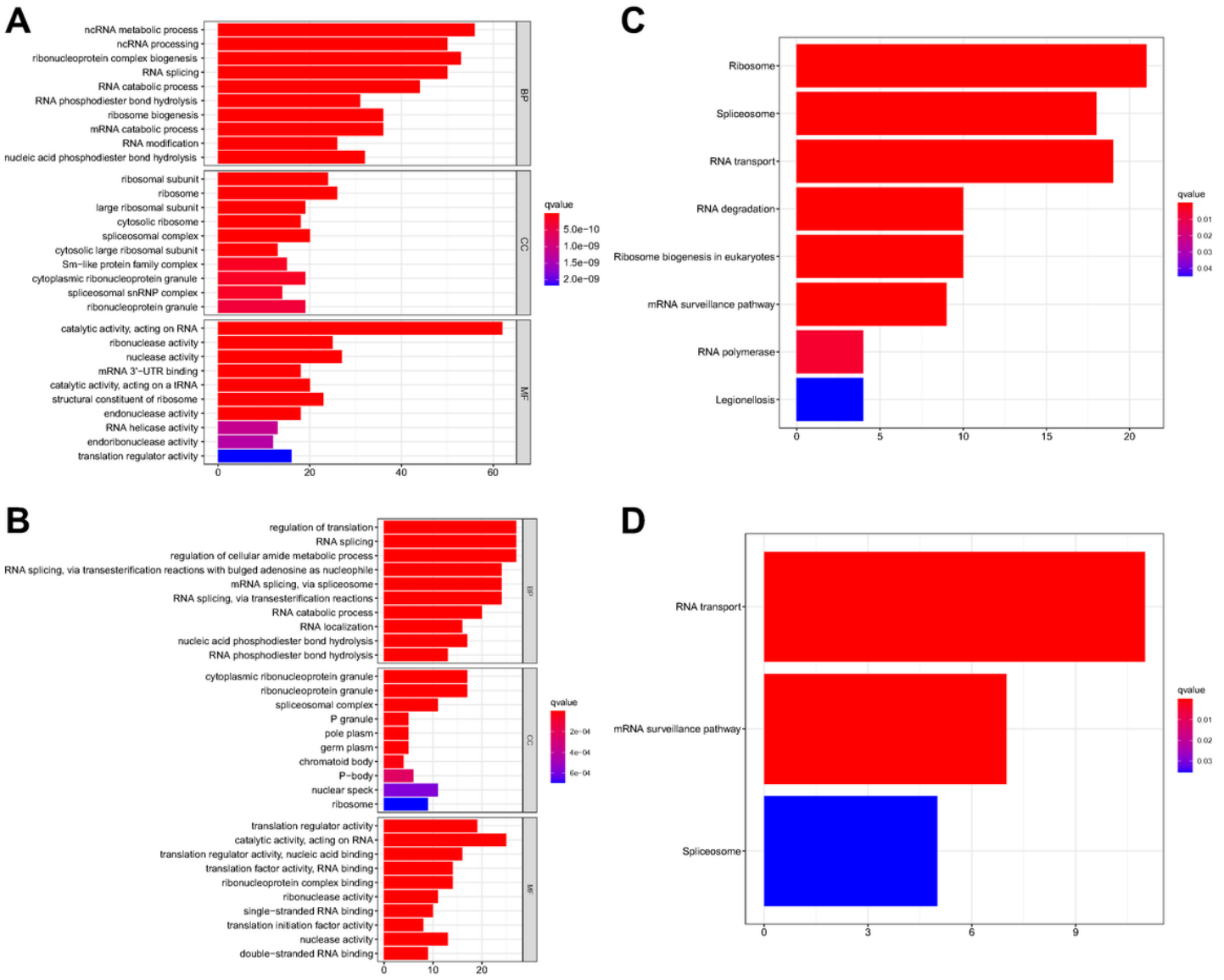

Figure 3

Functional enrichment analyses of 251 upregulated and 129 downregulated RBPs. (A) GO analysis on upregulated RBPs. (B) GO analysis on downregulated RBPs. (C) KEGG analysis on upregulated RBPs. (D) KEGG analysis on downregulated RPS.

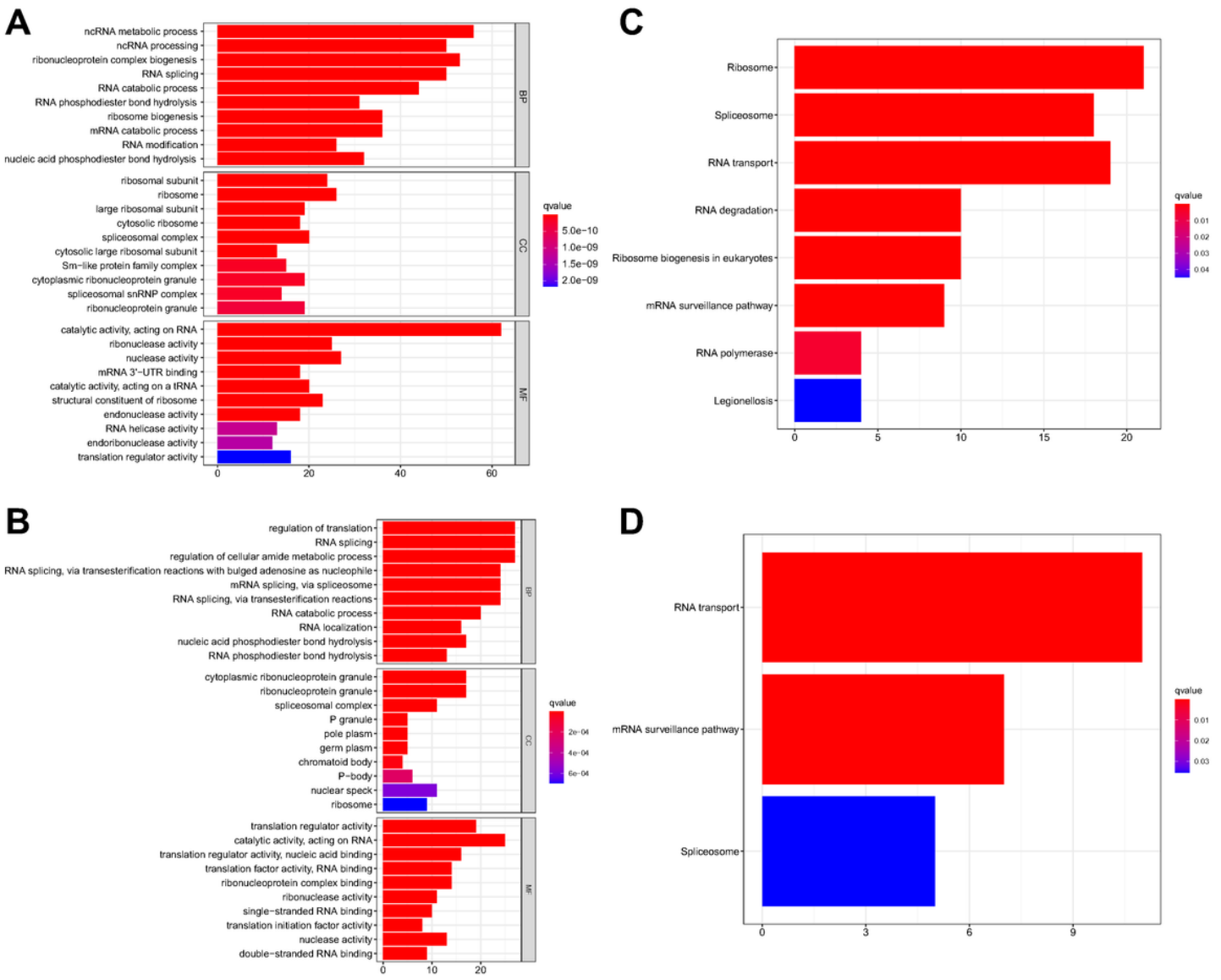

Figure 3

Functional enrichment analyses of 251 upregulated and 129 downregulated RBPs. (A) GO analysis on upregulated RBPs. (B) GO analysis on downregulated RBPs. (C) KEGG analysis on upregulated RBPs. (D) KEGG analysis on downregulated RPS.

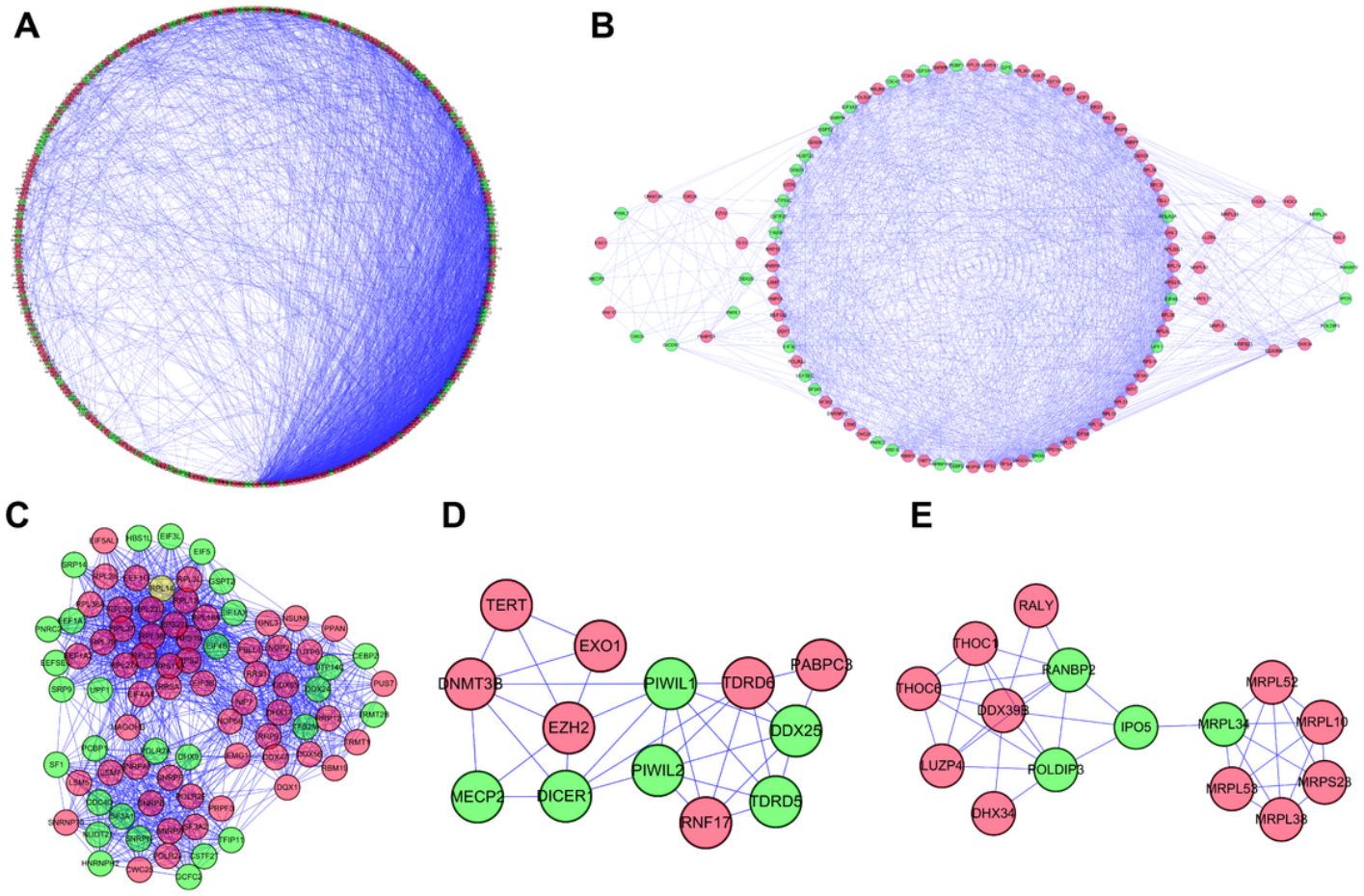

Figure 4

PPI network based on 380 differentially expressed RBPs and critical subnetworks. (A) Visualization of PPI network conducted on Cytoscape. Red nodes represent upregulated RBPs and green nodes represents downregulated RBPs. (B) Visualization of three MCODE modules. Visualization of module 1 (C), module 2 (D), and module 3 (E).

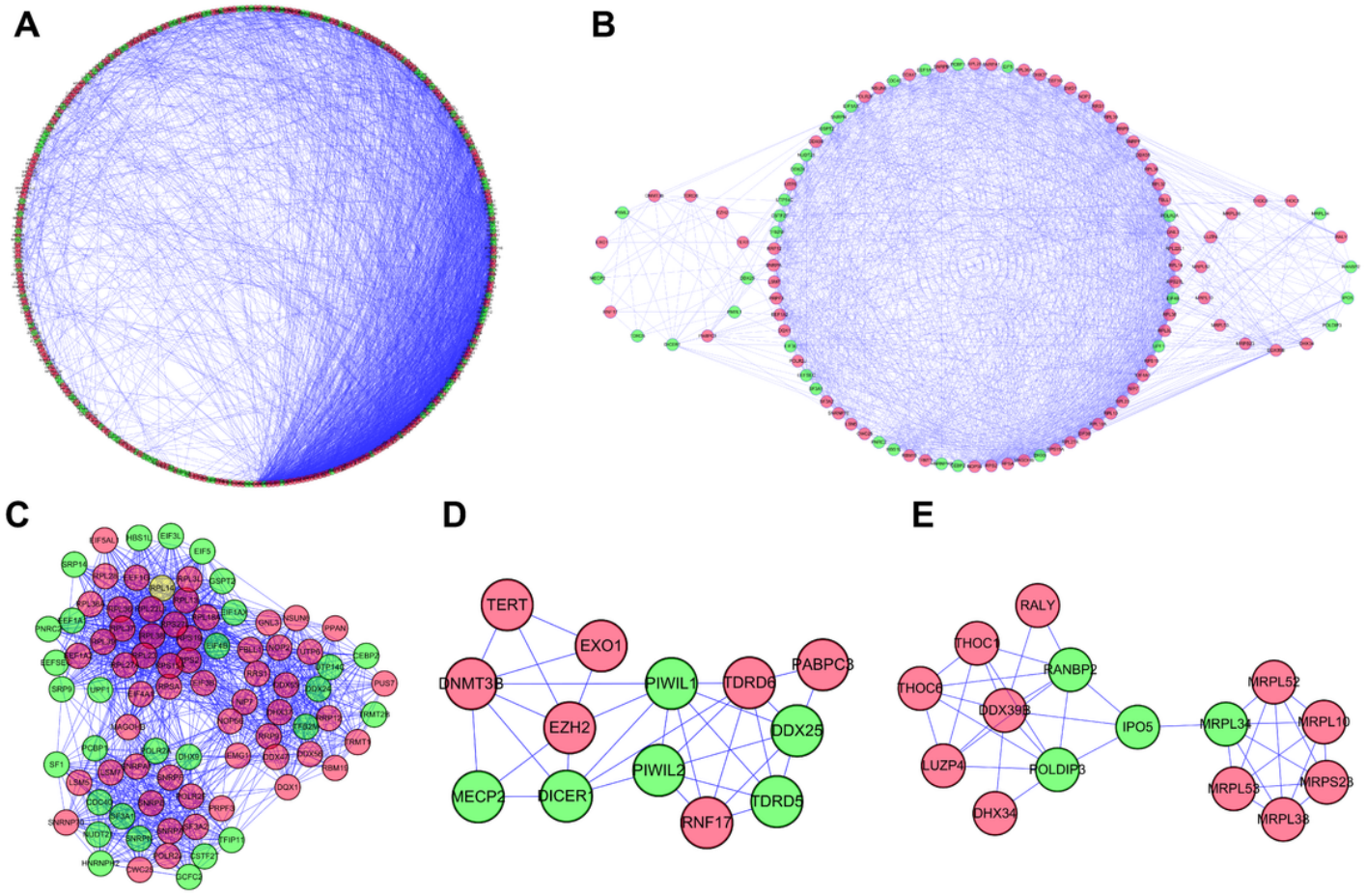

Figure 4

PPI network based on 380 differentially expressed RBPs and critical subnetworks. (A) Visualization of PPI network conducted on Cytoscape. Red nodes represent upregulated RBPs and green nodes represents downregulated RBPs. (B) Visualization of three MCODE modules. Visualization of module 1 (C), module 2 (D), and module 3 (E).

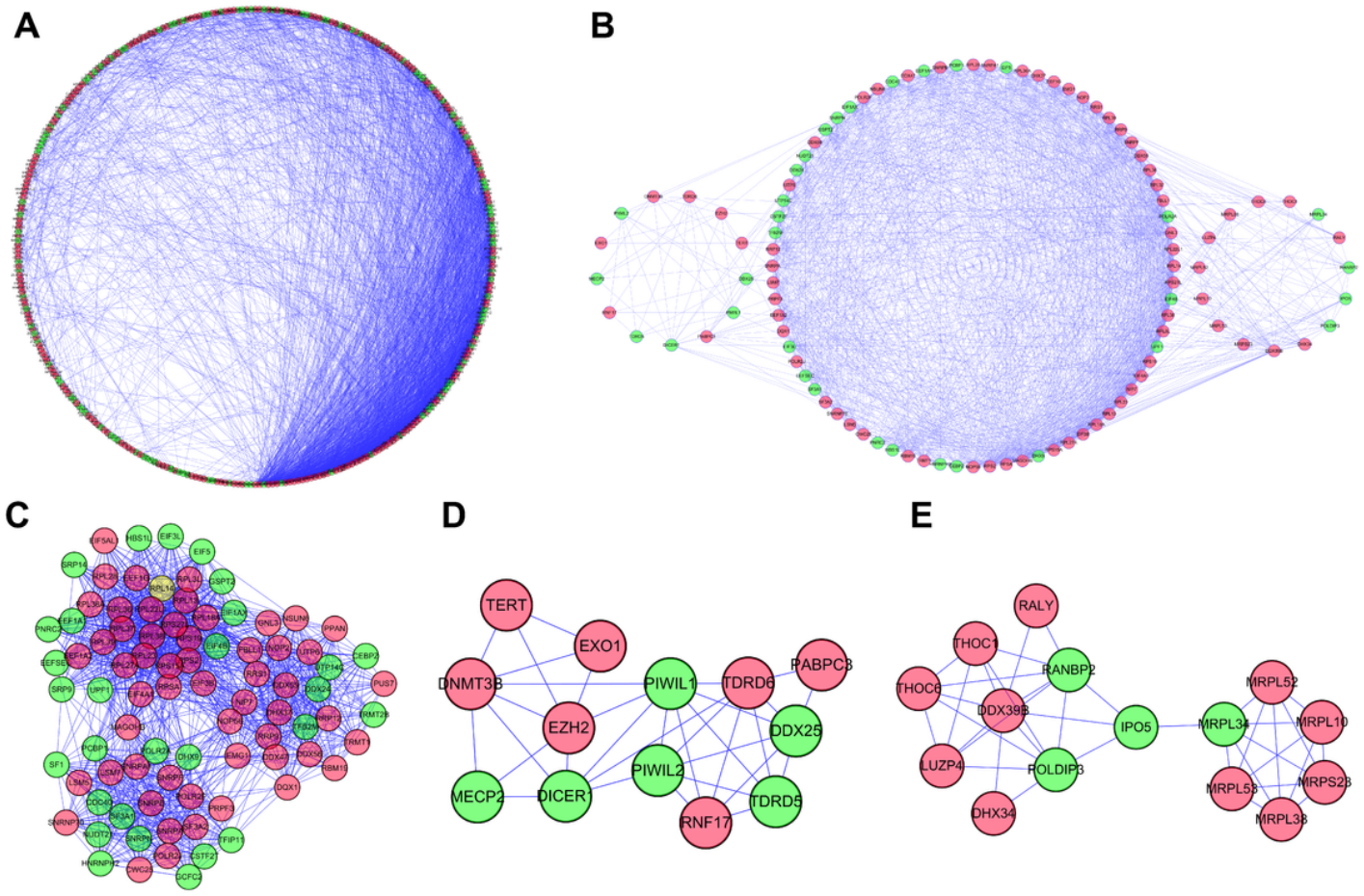

Figure 4

PPI network based on 380 differentially expressed RBPs and critical subnetworks. (A) Visualization of PPI network conducted on Cytoscape. Red nodes represent upregulated RBPs and green nodes represents downregulated RBPs. (B) Visualization of three MCODE modules. Visualization of module 1 (C), module 2 (D), and module 3 (E).

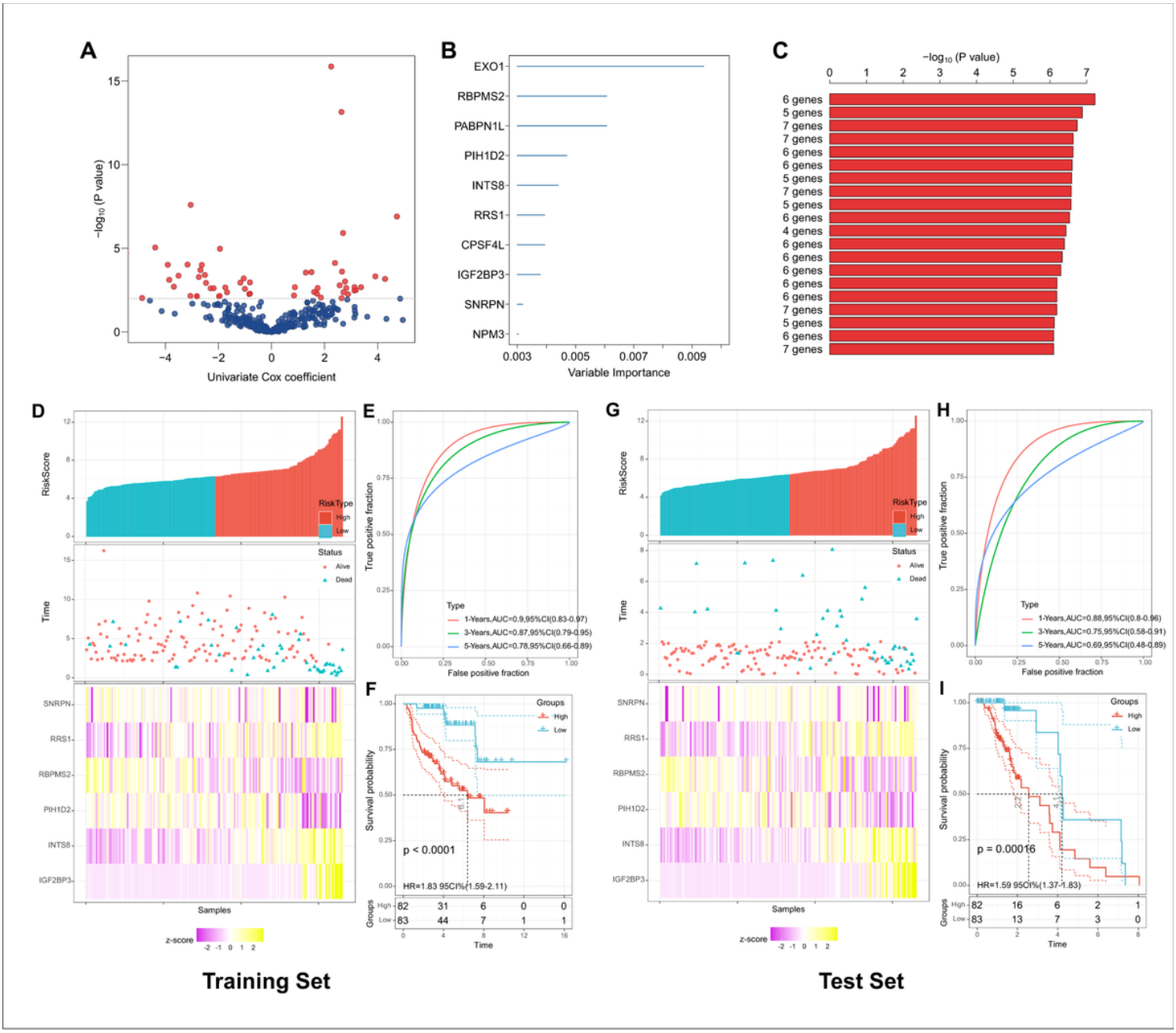

Figure 5

Identification of prognosis-related RBPs and validations of prognostic risk score model in training set and test set. (A) Volcano plot represented the prognosis-related RBPs of univariate Cox regression analysis. (B) Random survival forest analysis filtered out 10 best related genes. (C) The top 20 signatures were screened out among 1023 combinations according to the P value of Kaplan–Meier analysis. The risk score distribution, survival status distribution and heatmap of six RBPs expression in the training set (D) and test set (G). ROC analysis for predictive OS of KIRP patients at 1, 3 and 5 years in the training set (E) and test set (H). Kaplan–Meier analysis for overall survival (OS) of different risk groups in the training set (F) and test set (I).

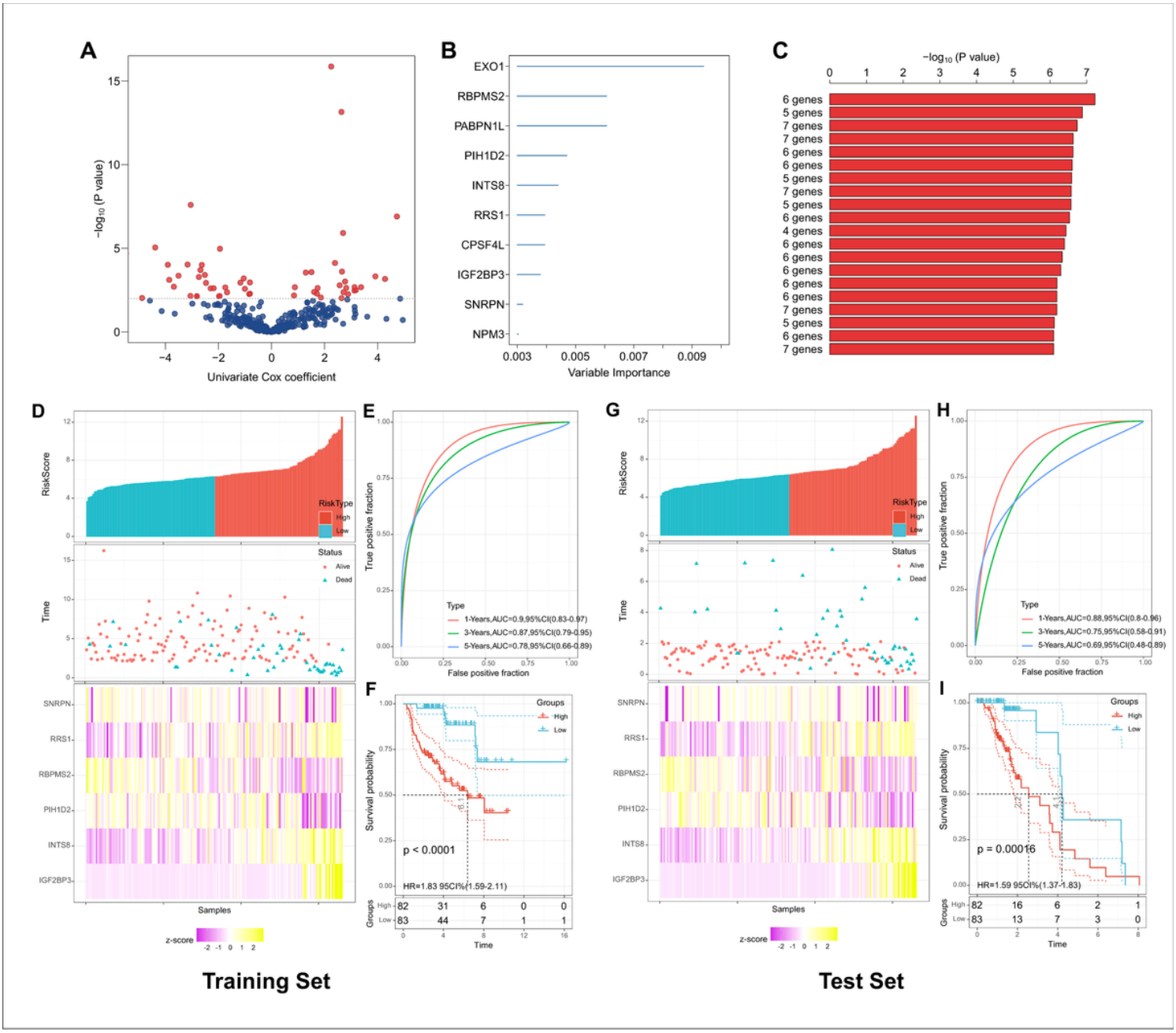

Figure 5

Identification of prognosis-related RBPs and validations of prognostic risk score model in training set and test set. (A) Volcano plot represented the prognosis-related RBPs of univariate Cox regression analysis. (B) Random survival forest analysis filtered out 10 best related genes. (C) The top 20 signatures were screened out among 1023 combinations according to the P value of Kaplan–Meier analysis. The risk score distribution, survival status distribution and heatmap of six RBPs expression in the training set (D) and test set (G). ROC analysis for predictive OS of KIRP patients at 1, 3 and 5 years in the training set (E) and test set (H). Kaplan–Meier analysis for overall survival (OS) of different risk groups in the training set (F) and test set (I).

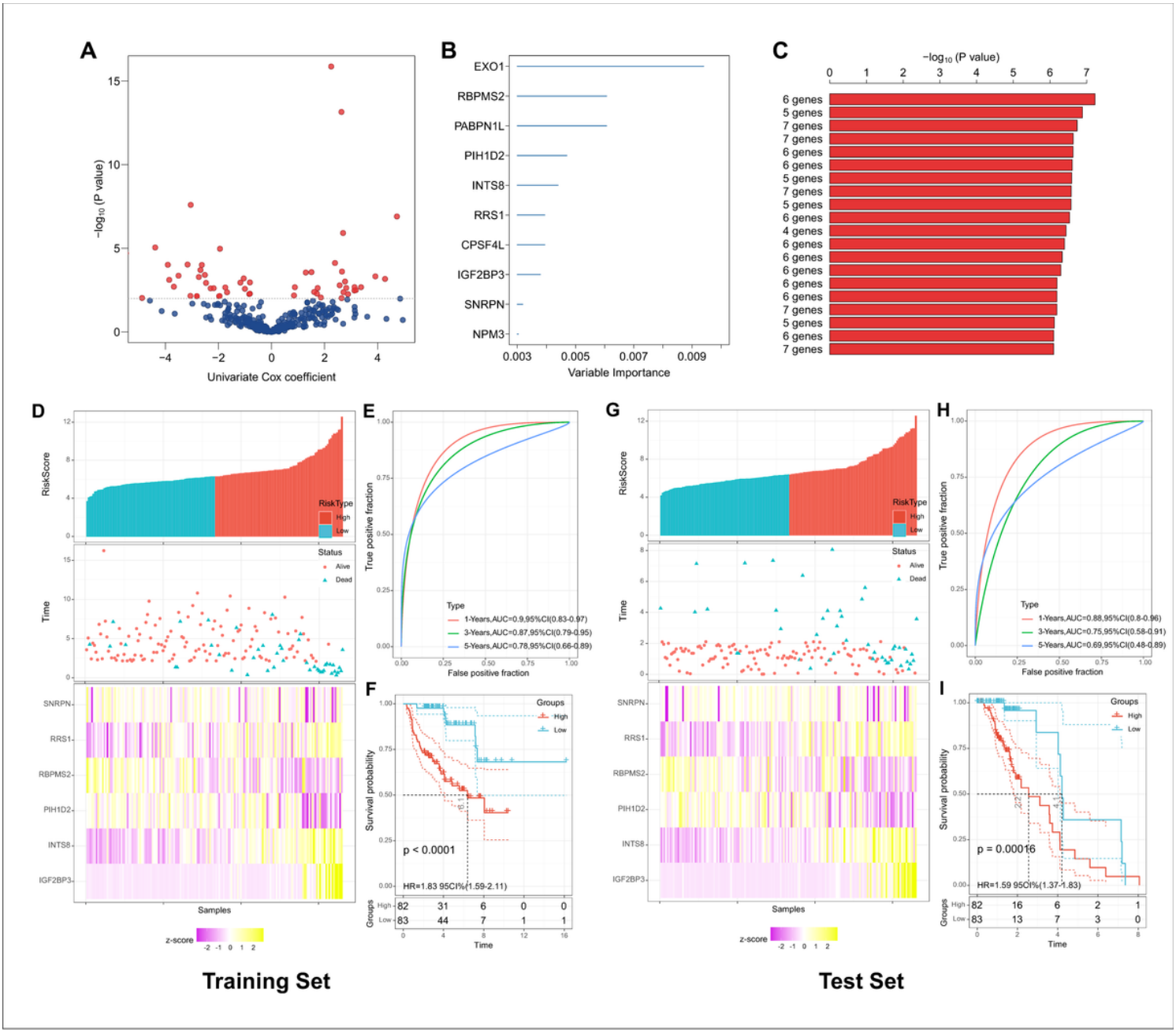

Figure 5

Identification of prognosis-related RBPs and validations of prognostic risk score model in training set and test set. (A) Volcano plot represented the prognosis-related RBPs of univariate Cox regression analysis. (B) Random survival forest analysis filtered out 10 best related genes. (C) The top 20 signatures were screened out among 1023 combinations according to the P value of Kaplan–Meier analysis. The risk score distribution, survival status distribution and heatmap of six RBPs expression in the training set (D) and test set (G). ROC analysis for predictive OS of KIRP patients at 1, 3 and 5 years in the training set (E) and test set (H). Kaplan–Meier analysis for overall survival (OS) of different risk groups in the training set (F) and test set (I).

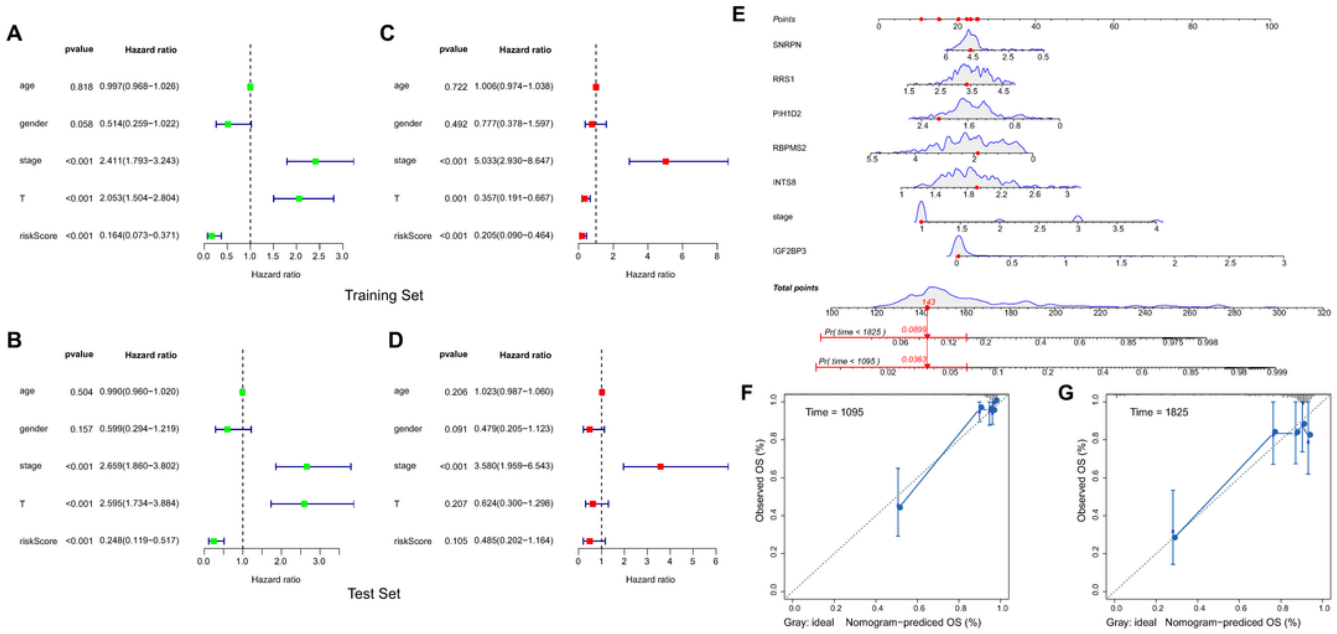

Figure 6

Identification of independent prognostic factors, nomogram for predicting the probability of patient mortality based on six RBPs as well as tumor stage and calibrations of nomogram in terms of conformity between predicted outcomes and observed outcomes at 3 and 5 years. Outcomes of univariate prognostic analysis conducted on training set (A) and test set (B). Outcomes of multivariate prognostic analysis conducted on training set (C) and test set (D). (E) Nomogram for evaluating the possibility of KIRP patients mortality at 3 and 5 years. Calibration for assessing the conformity between nomogram OS and observed OS at 3 years (F) and 5 years (G).

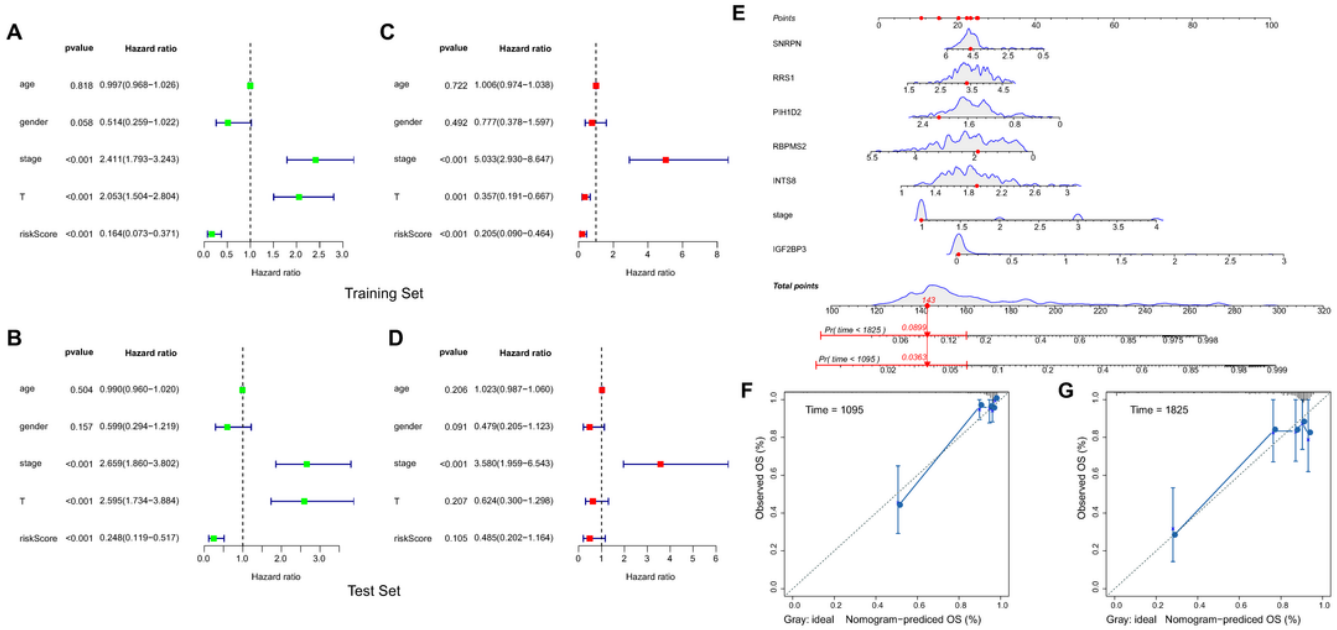

Figure 6

Identification of independent prognostic factors, nomogram for predicting the probability of patient mortality based on six RBPs as well as tumor stage and calibrations of nomogram in terms of conformity between predicted outcomes and observed outcomes at 3 and 5 years. Outcomes of univariate prognostic analysis conducted on training set (A) and test set (B). Outcomes of multivariate prognostic analysis conducted on training set (C) and test set (D). (E) Nomogram for evaluating the possibility of KIRP patients mortality at 3 and 5 years. Calibration for assessing the conformity between nomogram OS and observed OS at 3 years (F) and 5 years (G).

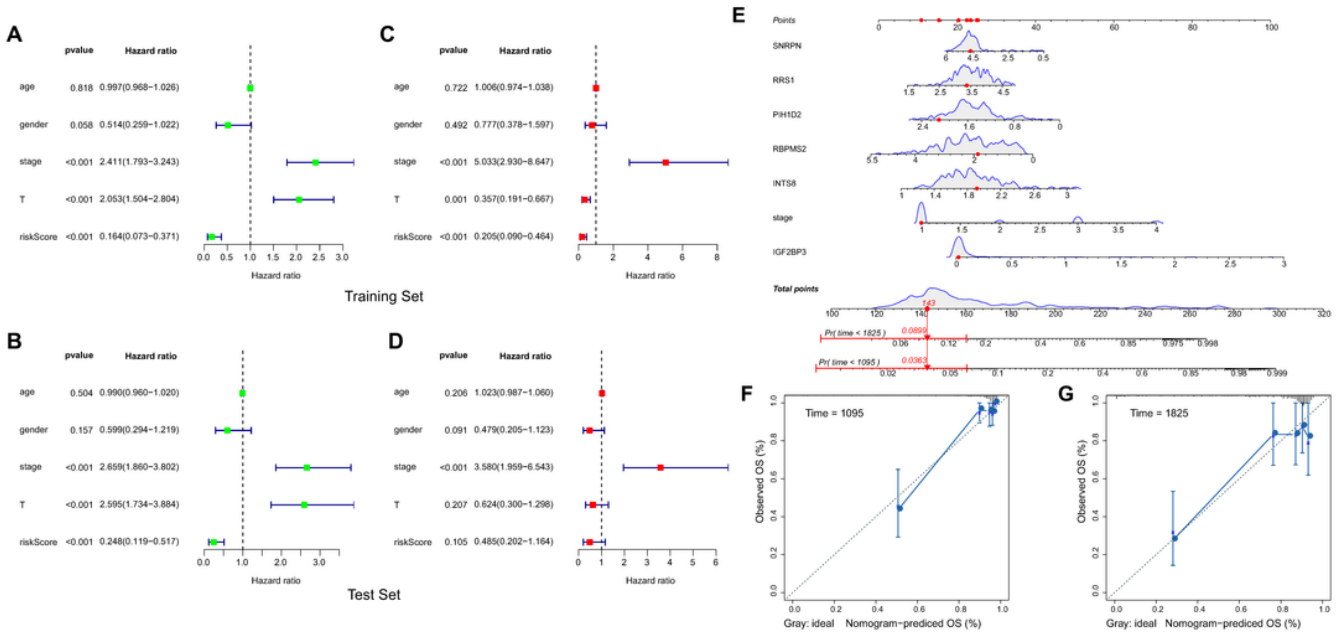

Figure 6

Identification of independent prognostic factors, nomogram for predicting the probability of patient mortality based on six RBPs as well as tumor stage and calibrations of nomogram in terms of conformity between predicted outcomes and observed outcomes at 3 and 5 years. Outcomes of univariate prognostic analysis conducted on training set (A) and test set (B). Outcomes of multivariate prognostic analysis conducted on training set (C) and test set (D). (E) Nomogram for evaluating the possibility of KIRP patients mortality at 3 and 5 years. Calibration for assessing the conformity between nomogram OS and observed OS at 3 years (F) and 5 years (G).

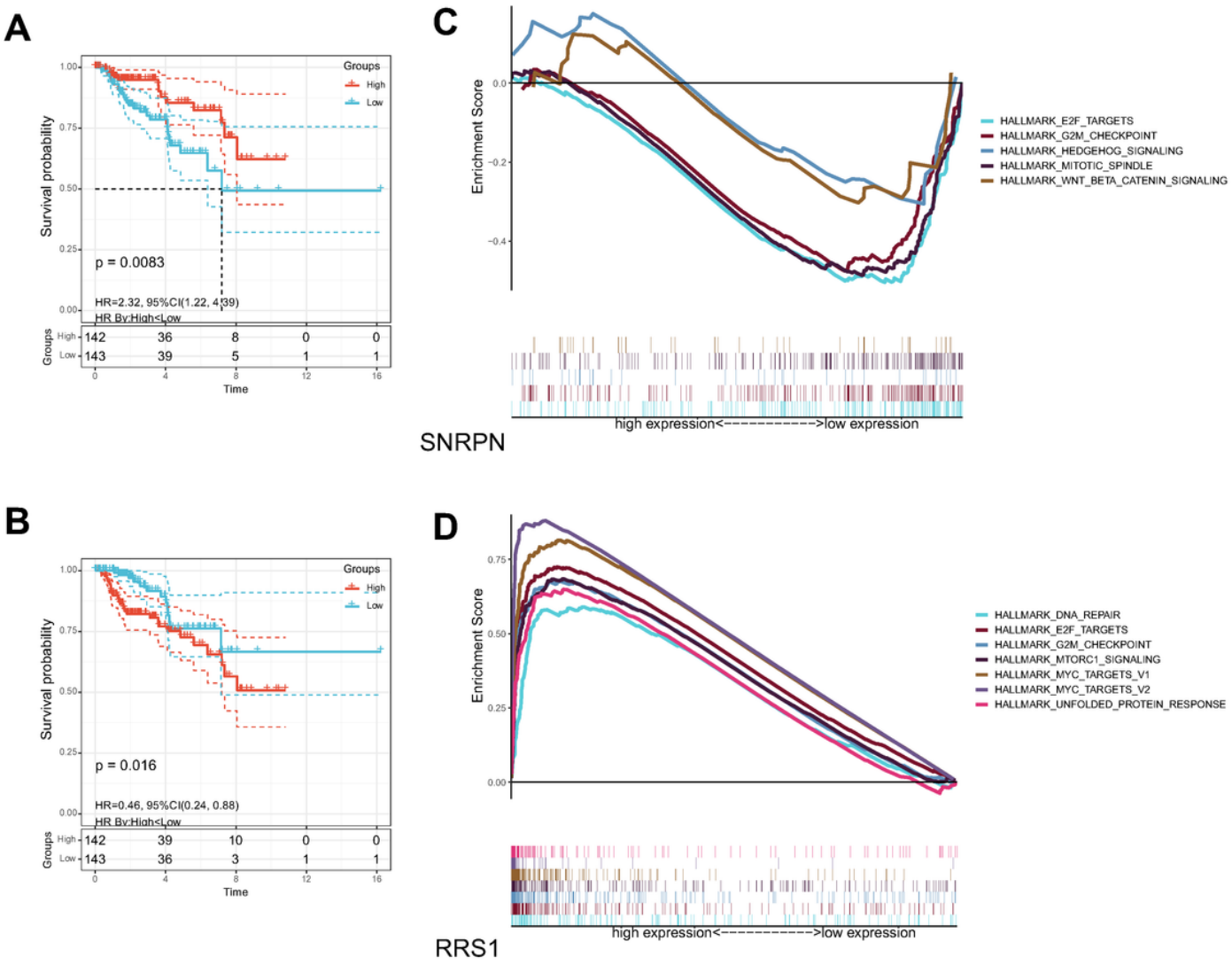

Figure 7

The survival curves and GSEA for samples with high and low expression of SNRPN and RRS1. (A) Kaplan–Meier analysis for overall survival (OS) based on the expression of SNRPN. (B) The enriched gene sets in HALLMARK collection by samples with low SNRPN expression. (C) Kaplan–Meier analysis for overall survival (OS) based on the expression of RRS1. (D) The enriched gene sets in HALLMARK collection by samples with high RRS1 expression.

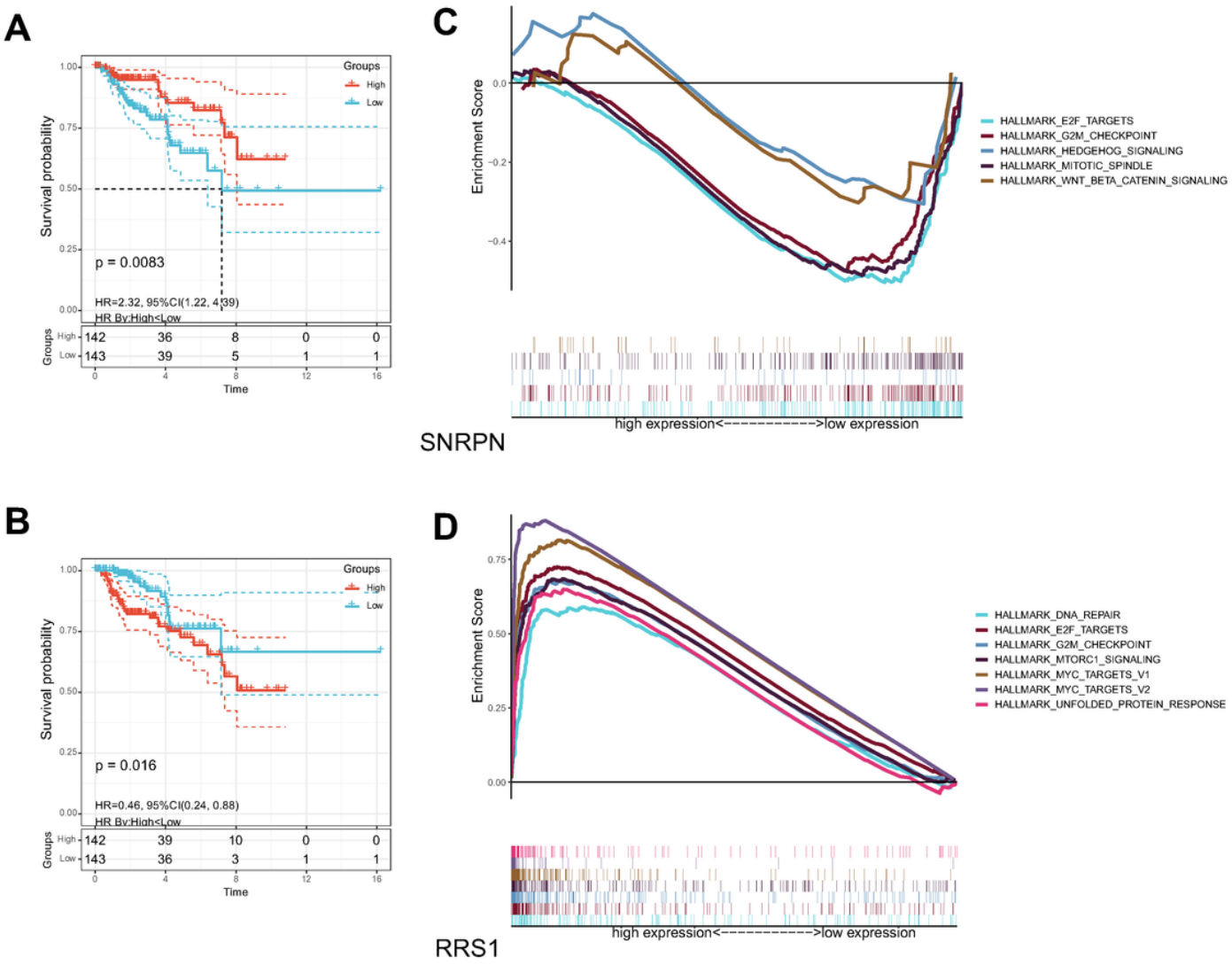

Figure 7

The survival curves and GSEA for samples with high and low expression of SNRPN and RRS1. (A) Kaplan–Meier analysis for overall survival (OS) based on the expression of SNRPN. (B) The enriched gene sets in HALLMARK collection by samples with low SNRPN expression. (C) Kaplan–Meier analysis for overall survival (OS) based on the expression of RRS1. (D) The enriched gene sets in HALLMARK collection by samples with high RRS1 expression.

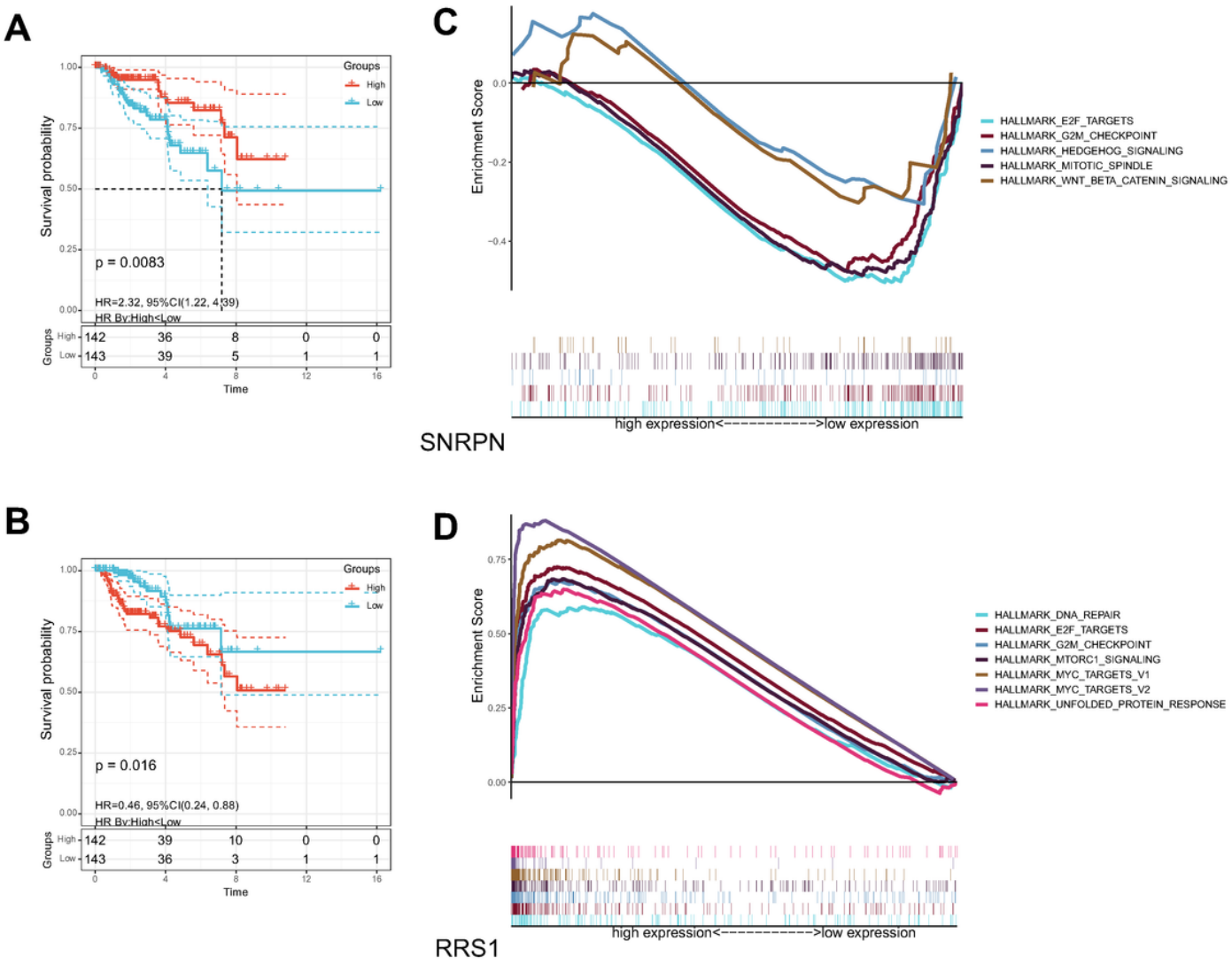

Figure 7

The survival curves and GSEA for samples with high and low expression of SNRPN and RRS1. (A) Kaplan–Meier analysis for overall survival (OS) based on the expression of SNRPN. (B) The enriched gene sets in HALLMARK collection by samples with low SNRPN expression. (C) Kaplan–Meier analysis for overall survival (OS) based on the expression of RRS1. (D) The enriched gene sets in HALLMARK collection by samples with high RRS1 expression.