Systematic Analysis of the Functions and Prognostic Values of Rna Binding Protein in Head and Neck Squamous Cell Carcinoma

song jukun (songjukun@163.com)
Guizhou provinicial people's hospital https://orcid.org/0000-0003-2542-9340
Feng Liu
Guizhou provinicial people's hospital
Bo Liu
Guizhou provinicial people's hospital
Xianlin Cheng
Guizhou provinicial people's hospital
Xinhai Yin
Guizhou provinicial people's hospital
Ke Zhao
Guizhou provinicial people's hospital
Xiaofeng Duan
Guizhou provinicial people's hospital

Primary research

Keywords: Head and Neck squamous cell carcinoma,The least absolute contraction and selection operator, RNA-binding proteins

Posted Date: June 22nd, 2020

DOI: https://doi.org/10.21203/rs.3.rs-36708/v1

Abstract

Background: Dysregulation of RNA-binding proteins (RBPs) playsan important role in controlling processes in cancer development.However, the function of RBPs has not been thoroughly and systematically documented in head and neck cancer.We aim to explore the role of RPB in the pathogenesis of HNSC.

Methods: We obtained HNSC gene expression data and corresponding clinical information from The Cancer Genome Atlas (TCGA) and the GEO databases, andidentified aberrantly expressed RBPs between tumors and normal tissues.Meanwhile, we performed a series of bioinformatics to explore the function and prognostic value of these RBPs.

Results: A total of 249 abnormally expressed RBPs were identified, including 101 down-regulated RBPs and 148 up-regulated RBPs.Using least absolute shrinkage and selection operator (LASSO) and univariate Cox regression analysis, the fifteen RPBs were identified as hub genes. With the fifteen RPBS, the prognostic prediction model was constructed.Further analysis showed that the high-risk score of the patients expressed a better survival outcome. The prediction model was validated in another HNSC dataset, and similar findings were observed.

Conclusions: Our results provide novel insights into the pathogenesis of HNSC. The fifteen RBP gene signature exhibited the predictive value of moderate HNSC prognosis, and have potential application value in clinical decision-making and individualized treatment.

Background

Head and Neck squamous cell carcinoma (HNSC) isone ofthe common types of malignancy, which is characterized by advanced diagnosis, a higher rate of lymph node metastasis, and high recurrence[1-3]. According to thepreviously published global cancer statistics report, there are approximately 400,000 oral and pharyngeal diseases, 160,000 laryngeal cancers, and 300,000 deaths worldwide each year[4-6]. The prognosis of patients with head and neck cancer stage III/IV is not ideal. The 5-year recurrence-free survival (RFS) and overall survival (OS) are 40% and 60%, respectively[7].Although the diagnosis and treatment of HNSC have achieved great progress in the past few decades, there is still a lack of effective therapies for patients. Therefore, there is an urgent need for new treatments to enrich the treatment of head and neck cancer and improve the survival rate of head and neck cancer patients, which could further understand the molecular biology basis of HNC.

RNA binding protein (RBP) is essentially a pleiotropic protein and regulates gene expression at the post-transcription level through interaction with target RNA[8, 9].Through high-throughput screening, a total of 1542 RBPs were identified in human cells, accounting for 7.5% of all protein-coding genes[8]. Studies have shown that many diseases are related to abnormal expression of RBP, including abnormal metabolism, germ cell development, muscular dystrophy,the human fragile X syndrome,nervous system diseases, and cancer[10-14]. However, only a smallfraction of RBPs have been studied intensively and found to havekey roles in cancers to date. Therefore, we collected HNSC data from the Cancer Genome Atlas (TCGA) database and conducted a systematic analysis to examine the potential molecular function and clinical significance of RPBs in HNSC.

Material And Methods

Data source and data Preprocessing

Firstly, the transcriptome sequencing data with the HTSeq-FPKMformat of 546 samples were downloaded from the TCGA database by the GDC tool(https://portal.gdc.cancer.gov/) as train group, including 502 HNSC patients and44

adjacenttumor samples.Besides, the clinical information of each patient, including age, gender, TNM stage, tumor grade, and survival time, is obtained from the GDC portal.Finally, a total of 501 samples were considered eligible in this analysis.Secondly, another HNSC dataset was also downloaded from the GEO dataset (GSE84437[15]) and serves as an external verification group.

Differentially expressed analysis

The differentiallyexpressed RPBs between adjacent tumor samples and tumorsamples were screened using the “limma” packagein R[16]. The differentially expressed RPBs of data sets with |log2 fold change| ≥0.5 and aP-value less than 0.05 was considered as the selection criteriafor subsequent analysis. The heatmap plot was also exhibited.

Functional enrichment analysis

To understand the underlying biological processes and pathwaysof aberrantly expressed RPBs, the biological functions are systematically analyzed through gene ontology (GO) enrichment, which includes three terms: molecular function, biological process, and cellular composition. The Kyoto Encyclopedia of Genes and Genomes (KEGG) is used to detect potential biological pathways for differentially expressed RBPs. Functional enrichment analysis wasconducted by “ClusterProfiler” of R packagewith the cut-off value of P less than 0.05[17].

Construction of protein-protein interaction networkand identification of sub-modules

The protein-protein interactions (PPIs) among differentiallyexpressed RBPs were detected using the String website, and the PPIs network was constructed by Cytoscape 3.6.2 software.By using the MCODE (Molecular Complex Detection) plugin in Cytoscape, we can identify sub-modules from the PPI network with a score>9 and a node count>10.Functional enrichment analysis for each module wasconducted by the “clusterprofiler” of R packagewith the cut-off value of P less than 0.05.

The screen of prognostic RPBs

Firstly, the univariate COX regression analysis was conducted to screen the prognostic RPBs with the threshold of a P value less than 0.05. Then, the least absolute shrinkage and selection operator (LASSO) was furtherused to identify the prognostic significance of candidate RBPs[18].

Construction of prognostic prediction model

With the candidate RBPs identified by LASSO, we built a prognostic prediction model. The regression coefficients from the LASSO algorithms were used to create a classification index for each sample, and used the following formula to weight the expression value of the selected miRNA: The Prognostic Index = (Exp mRNA1* Coef1) + (Exp mRNA2* Coef2) + (Exp mRNA3* Coef3) +…+ (Exp mRNAn* Coefn).

Based on the median of the prognostic index, the HNSC samples were divided into two groups: high-risk group and low-risk group. We used the Kaplan-Meier analysis and log-rank test to evaluate the difference in overall survival (OS) between the two groups. To evaluate the performance of this prognostic model, receive operating characteristic (ROC) curves were used to compare the sensitivity and specificity of the survival prediction. The ROC curves were plotted using the “survivalROC” package in R.Univariate and multivariate COX prognostic analysis was used to determine whether the established prognostic prediction model was an independent prognostic factorfor HNSC.

ValidationHub RBPs in HPA database

The protein expression of the 15 RPBsbetween HNSC and normal tissues was determined at a translational level usingimmunohistochemistry (IHC) from the Human Protein Atlasdatabase (HPA, https://www.proteinatlas.org/)[19].

Validation of the prognostic prediction model

The established prognostic prediction model was validated in another GEO (GSE65858), and the KM survival curve, ROC curve, and risk curve was plotted. And we also evaluate whether the established prognostic prediction model was an independent prognostic factorfor HNSC.

Results

Identification of Differentially ExpressedRBPs

The flowchart was shown in Figure 1. Atotal of 502 LUSC samples and 44adjacent tumor sampleswere considered eligible in the study. Differentially expressedRBPs were screened by the “limma” package of R software. In total, 249 differentially expressedRBPs were identified, including 104 downregulated RPBs and 145 up-regulated RPBs (Supplementary Table 1). The heatmap and volcano plot was exhibited in Figure 2A,B.

GO and KEGG Analysis

To uncover the potential function and molecular mechanism of differentially expressed RBPs, they were divided into two groups (upregulated and downregulated group) according to their expression levels. Functional enrichment analysis was carried out via thehypergeometric test using the ClusterProfiler R package. Upregulated differentially expressed RBPs werenotably enriched in biological processes, including metabolic process, ncRNA processing, ribonucleoprotein granule, and spliceosomal complex.The downregulated differentiallyexpressed RBPs were enriched in the mRNA processing, RNA splicing, ribosome, and translation repressor activity (Table 1, Figure 3A,B).The results of KEGG analysis for upregulated RPBs revealed that a total of nine KEGG terms was enriched, including Spliceosome, RNA transport and Ribosome biogenesis in eukaryotes, whereas thedownregulated RBPs were significantly enriched in RNA transport, mRNA surveillance pathway and RNA degradation (Table 1, Figure 3C,D).

Construction of PPI Network and identification of KeyModules

The PPI network was established using these differentially expressedRBPs in the STRING database, then a protein-protein co-expression network was visualized using Cytoscape.The PPI network for RBPs was exhibited in Figure 4A, which included 223 nodes and 1,094 edges. The two hub networks were identified via the plug-in MODE in Cytoscape. Module I included 22 genes, while Module II included 33 genes (Figure 4B,C,D). The GO and KEGG analysis in Module 1 was mainly enriched in RNA splicing, ncRNA processing, ribosome biogenesis, RNA polymerase, and NOD-like receptor signaling pathway, while the GO and KEGG analysis in Module 2were mainly enriched inmitochondrial translation, snoRNA metabolic process, nuclear RNA surveillance, Ribosome, and RNA degradation.

Selection of Prognostic-Related RBPs

Among 249 differentially expressedRBPs, 24 RPBs were regarded as prognostic-related RPBs using the univariate Cox regressionanalysis (Figure 5). Then, We used the LASSO method to build a prediction model, selected non-zero regression coefficients to identify the optimal RPBs gene set, and used the selected RPBs to build RPBs-based predictive model to predict prognosis. In LASSO logistic regression method, FifteenRPBs as optimal features were ultimately recognized, includingDNMT1, MRPL33, EZH2, PCF11, RBM24, TRMT112, DZIP1, EIF5A2, MKRN3, DARS2, PSMA6, AZGP1, LENG9, IGF2BP2,andCIRBP.The RPBs-based a predictive model was calculated as the following

formula: The risk score of eachHNSC patient was computed according to the following formula:DNMT1*(-0.21475) +MRPL33*0.177058+ EZH2*(-0.12256)+ PCF11*(-0.23465)+ RBM24*0.08662+ TRMT112*0.091039+ DZIP1*0.025207+ EIF5A2*0.147603+MKRN3*0.495024+DARS2*0.259781+PSMA6*0.161435+AZGP1* (-0.0518)+LENG9*(-0.14885)+ IGF2BP2*(-0.027825)+ CIRBP*0.00951.

To assess the predictive ability of this model, the patients were divided into high- and low-risk groups for survivalanalysis according to the median risk score. Patients in thehigh-risk group exhibited better survival rates than thosein the low-risk group (Figure 6A). The time-dependent ROC analysis was used to evaluate the prognostic performance of the 15-RBP gene signature. The AUC of theROC curve for OS was 0.67(0.62-72) at 1 years, 0.68 (0.63-0.73) at 3 years and 0.67 (0.60-0.74) at 5 years (Figure 6B). The result indicated that the model had moderate diagnostic strengthen. The expression heat map and survival status ofpatients with the 15-RBP gene biomarkers in the low- and high-risk subgroups are shown in Figure 6C-E.

Univariate and multivariate Cox regression analysis suggested that the prediction model could become an independent predictor, including age, gender, histologicalgrade,stage, TNM stage, and risk score(Figure 7, Table 3).

Validation of candidate RPBs in the HPA dataset

To further determine the expression of these hub genes in HNSC,we used immunohistochemistry results from the Human ProteinAtlas (HPA) database to show that AZGP1,DNMT1, DZIP1, IGF2BP2, LENG9, MRPL33, PSMA6, RBM24, and TRMT112were significantly increased in head and neck cancer compared withnormal tissue (Figure 8).

Validation of established prognostic RPB model

The established prognostic RPB model based on the 15 RPBs was validated in another GEO (GSE84437) dataset using the same risk assessment formula, and the result indicated that the model exhibited better predictive ability. Patients in the high-risk group suffered a lower survival probability (Figure9A). The area under the curve of the ROC curve for 1,3,5 yearreached 0.58, 0.55, and 0.58, respectively (Figure 9B). The expression heat map and survival status ofpatients with the 15-RBP gene biomarkers in the low- and high-risk subgroups are shown in Figure 9C-E.Univariate and multivariate Cox regression analysis suggested that the IRGPM could become an independent predictor, including age, gender, histologicalgrade,stage, TNM stage, and risk score(Figure 10A,B).

Discussion

RNA binding protein (RBP) is a general term for the protein that binds to RNA, which regulates the metabolic process of RNA. RBP accompanies the life of RNA, and its main role is to mediate RNA maturation, transport, positioning, and translation. One RBP may have multiple target mRNAs, and its defective expression may cause various diseases. The intricate network of interactions between RBP and its cancer-associated RNA targets will provide a better understanding of tumor biology and may reveal new cancer treatment targets[8, 20]. Recently, several studies have documented that RBPs are dysregulated in various human cancers[21-23]. However, little is known about the expression pattern and mechanism of action of RBPs in HNSC. At the work, we comprehensively analyzed the key RBPs specifically associated with theoverall survival of HNSC using bioinformatics analysis. The differentially expressedRBPs between adjacent tumor samples and tumor samples were identified. Meanwhile, a PPI network was constructed based on the RPBs, and The two hub networks were identified via the plug-in MODE in Cytoscape software. Then, we used the univariate Cox regressionanalysis to screen prognosis-related RBPs and constructed a risk model to predict HNSC prognosis based on a fifteen -RBPs signature.

The biological processes or pathways that upregulated RBPs were significantly enriched in metabolic process, ncRNA processing, ribonucleoprotein granule, and spliceosomal complex. The downregulated differentiallyexpressed RBPs were enriched in the mRNA processing, RNA splicing, ribosome, and translation repressor activity.These results were lined with previous studies. RNA binding proteins are involved in RNA metabolism[24] and play an important role in regulating RNA stability, alternative splicing, modification, localization, and translation[25]. Changing the expression of RBPs in cells may affect various physiological processes in cells, such as alternative splicing and apoptosis [26, 27]. The RNA binding domain (RBD) is a region of RBP, which has been experimentally or structurally confirmed to directly bind RNA[8].Meanwhile, a large number of studies have recently reported the role of abnormal RNA metabolism and RNA processing, alternative splicing in various diseases, including cancer development[28, 29].These results further confirm that RBP can affect the growth and metastasis of tumor cells by regulating various biological processes (such as RNA metabolism and RNA processing, RNA transport, RNA degradation).

Subsequently, the univariate Cox regressionanalysis and Lasso regression analysis were used to identifying the prognostic RPBs, and fifteen RPBs were defined as hub genes. Among these RPBs, DNMT1, MRPL33, EZH2, PCF11, RBM24, TRMT112, DZIP1, EIF5A2, MKRN3, DARS2, PSMA6, AZGP1, LENG9, IGF2BP2,andCIRBPwere considered as candidate genes. The DNMT1 was involved in the development of various tumors, including pancreatic ductal adenocarcinoma[30], breast cancer[31], and cervical cancer[32]. MRPL33 was closely related to the incidence of head and neck cancer[33], gastric cancer[34], and breast cancer[35].EZH2, PCF11, RBM24, DZIP1, EIF5A2, MKRN3, PSMA6, AZGP1, IGF2BP2,andCIRBP was also reported to involve in the development of tumor[36-39].

Conclusions

our research provides relatively novel insights into the role of RBPs in the occurrence and development of HNSC. Besides, our prognosis model shows good predictive performance in terms of survival, which may help to develop new prognostic indicators for HNSC. Also, RBP-related gene markers show a key biological background, which indicates that these RBPs can be used for clinical adjuvant therapy.Furthermore,survival analysis was also performed based on the univariate and multivariate Cox proportionalhazards regression models.

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and material

Data availability could be obtained from TCGA website.

Competing interests

The authors declare that they have no competing interests.

Funding:

The works was supported by Guizhou Science and Technology Department (Identification of Slug-regulated miRNA in oral squamous cell carcinoma and its mechanism of tumor metastasis inhibition, NO: Qiankehe LH [2014] No. 7012).

Author contributions:

J.K.S;F.L; B.L.H;X.H.Y wrote the main manuscript text;

J.K.S; K.Z;X.L.C prepared Figures 1-9;

J.K.S; X.F.Dcontributed to data analysis;

All authors reviewed the manuscript.

Acknowledgments.

None.

Reference:

1. Preciado DA, Matas A, Adams GL: **Squamous cell carcinoma of the head and neck in solid organ transplant recipients**. *Head & neck* 2002, **24**(4):319-325.
2. Patel BP, Shah PM, Rawal UM, Desai AA, Shah SV, Rawal RM, Patel PS: **Activation of MMP-2 and MMP-9 in patients with oral squamous cell carcinoma**. *Journal of surgical oncology* 2005, **90**(2):81-88.
3. Choi KK, Kim MJ, Yun PY, Lee JH, Moon HS, Lee TR, Myoung H: **Independent prognostic factors of 861 cases of oral squamous cell carcinoma in Korean adults**. *Oral oncology* 2006, **42**(2):208-217.
4. Wozniak A, Szyfter K, Szyfter W, Florek E: **[Head and neck cancer--history]**. *Przeglad lekarski* 2012, **69**(10):1079-1083.
5. Mehanna H, Paleri V, West CM, Nutting C: **Head and neck cancer--Part 1: Epidemiology, presentation, and prevention**. *Bmj* 2010, **341**:c4684.
6. Siegel RL, Miller KD, Jemal A: **Cancer statistics, 2018**. *CA: a cancer journal for clinicians* 2018, **68**(1):7-30.
7. Chi AC, Day TA, Neville BW: **Oral cavity and oropharyngeal squamous cell carcinoma--an update**. *CA Cancer J Clin* 2015, **65**(5):401-421.
8. Gerstberger S, Hafner M, Tuschl T: **A census of human RNA-binding proteins**. *Nat Rev Genet* 2014, **15**(12):829-845.
9. Velasco MX, Kosti A, Penalva LOF, Hernandez G: **The Diverse Roles of RNA-Binding Proteins in Glioma Development**. *Adv Exp Med Biol* 2019, **1157**:29-39.
10. Fredericks AM, Cygan KJ, Brown BA, Fairbrother WG: **RNA-Binding Proteins: Splicing Factors and Disease**. *Biomolecules* 2015, **5**(2):893-909.
11. Suzuki A, Niimi Y, Shinmyozu K, Zhou Z, Kiso M, Saga Y: **Dead end1 is an essential partner of NANOS2 for selective binding of target RNAs in male germ cell development**. *EMBO Rep* 2016, **17**(1):37-46.

12. Chang ET, Parekh PR, Yang Q, Nguyen DM, Carrier F: **Heterogenous ribonucleoprotein A18 (hnRNP A18) promotes tumor growth by increasing protein translation of selected transcripts in cancer cells**. *Oncotarget* 2016, **7**(9):10578-10593.
13. Chenard CA, Richard S: **New implications for the QUAKING RNA binding protein in human disease**. *J Neurosci Res* 2008, **86**(2):233-242.
14. Neelamraju Y, Gonzalez-Perez A, Bhat-Nakshatri P, Nakshatri H, Janga SC: **Mutational landscape of RNA-binding proteins in human cancers**. *RNA Biol* 2018, **15**(1):115-129.
15. Wichmann G, Rosolowski M, Krohn K, Kreuz M, Boehm A, Reiche A, Scharrer U, Halama D, Bertolini J, Bauer U *et al*: **The role of HPV RNA transcription, immune response-related gene expression and disruptive TP53 mutations in diagnostic and prognostic profiling of head and neck cancer**. *Int J Cancer* 2015, **137**(12):2846-2857.
16. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK: **limma powers differential expression analyses for RNA-sequencing and microarray studies**. *Nucleic Acids Res* 2015, **43**(7):e47.
17. Yu G, Wang LG, Han Y, He QY: **clusterProfiler: an R package for comparing biological themes among gene clusters**. *OMICS* 2012, **16**(5):284-287.
18. Wu TT, Chen YF, Hastie T, Sobel E, Lange K: **Genome-wide association analysis by lasso penalized logistic regression**. *Bioinformatics* 2009, **25**(6):714-721.
19. Thul PJ, Lindskog C: **The human protein atlas: A spatial map of the human proteome**. *Protein Sci* 2018, **27**(1):233-244.
20. Lan Y, Xiao X, He Z, Luo Y, Wu C, Li L, Song X: **Long noncoding RNA OCC-1 suppresses cell growth through destabilizing HuR protein in colorectal cancer**. *Nucleic Acids Res* 2018, **46**(11):5809-5821.
21. 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**:155-167.
22. Soni S, Anand P, Padwad YS: **MAPKAPK2: the master regulator of RNA-binding proteins modulates transcript stability and tumor progression**. *J Exp Clin Cancer Res* 2019, **38**(1):121.
23. Schultz CW, Preet R, Dhir T, Dixon DA, Brody JR: **Understanding and targeting the disease-related RNA binding protein human antigen R (HuR)**. *Wiley Interdiscip Rev RNA* 2020, **11**(3):e1581.
24. Frisone P, Pradella D, Di Matteo A, Belloni E, Ghigna C, Paronetto MP: **SAM68: Signal Transduction and RNA Metabolism in Human Cancer**. *Biomed Res Int* 2015, **2015**:528954.
25. Perron G, Jandaghi P, Solanki S, Safisamghabadi M, Storoz C, Karimzadeh M, Papadakis AI, Arseneault M, Scelo G, Banks RE *et al*: **A General Framework for Interrogation of mRNA Stability Programs Identifies RNA-Binding Proteins that Govern Cancer Transcriptomes**. *Cell Rep* 2018, **23**(6):1639-1650.
26. Wurth L: **Versatility of RNA-Binding Proteins in Cancer**. *Comp Funct Genomics* 2012, **2012**:178525.
27. Lukong KE, Chang KW, Khandjian EW, Richard S: **RNA-binding proteins in human genetic disease**. *Trends Genet* 2008, **24**(8):416-425.
28. Zhang H, Brown RD, Stenmark KR, Hu CJ: **RNA-Binding Proteins in Pulmonary Hypertension**. *Int J Mol Sci* 2020, **21**(11).
29. Zhu S, Rooney S, Michlewski G: **RNA-Targeted Therapies and High-Throughput Screening Methods**. *Int J Mol Sci* 2020, **21**(8).
30. Wong KK: **DNMT1 as a therapeutic target in pancreatic cancer: mechanisms and clinical implications**. *Cell Oncol (Dordr)* 2020.
31. Kah KW: **DNMT1: A Key Drug Target in Triple-Negative Breast Cancer**. *Semin Cancer Biol* 2020.

32. Wu L, Gong Y, Yan T, Zhang H: **LINP1 Promotes the Progression of Cervical Cancer by Scaffolding EZH2, LSD1 and DNMT1 to Inhibit the Expression of KLF2 and PRSS8**. *Biochem Cell Biol* 2020.
33. Guo T, Zambo KDA, Zamuner FT, Ou T, Hopkins C, Kelley DZ, Wulf HA, Winkler E, Erbe R, Danilova L *et al*: **Chromatin structure regulates cancer-specific alternative splicing events in primary HPV-related oropharyngeal squamous cell carcinoma**. *Epigenetics* 2020:1-13.
34. Li J, Feng D, Gao C, Zhang Y, Xu J, Wu M, Zhan X: **Isoforms S and L of MRPL33 from alternative splicing have isoformspecific roles in the chemoresponse to epirubicin in gastric cancer cells via the PI3K/AKT signaling pathway**. *Int J Oncol* 2019, **54**(5):1591-1600.
35. Liu L, Luo C, Luo Y, Chen L, Liu Y, Wang Y, Han J, Zhang Y, Wei N, Xie Z *et al*: **MRPL33 and its splicing regulator hnRNPK are required for mitochondria function and implicated in tumor progression**. *Oncogene* 2018, **37**(1):86-94.
36. Hua WF, Zhong Q, Xia TL, Chen Q, Zhang MY, Zhou AJ, Tu ZW, Qu C, Li MZ, Xia YF *et al*: **RBM24 suppresses cancer progression by upregulating miR-25 to target MALAT1 in nasopharyngeal carcinoma**. *Cell Death Dis* 2016, **7**(9):e2352.
37. Yang SW, Li L, Connelly JP, Porter SN, Kodali K, Gan H, Park JM, Tacer KF, Tillman H, Peng J *et al*: **A Cancer-Specific Ubiquitin Ligase Drives mRNA Alternative Polyadenylation by Ubiquitinating the mRNA 3' End Processing Complex**. *Mol Cell* 2020, **77**(6):1206-1221 e1207.
38. Lu M, Ge Q, Wang G, Luo Y, Wang X, Jiang W, Liu X, Wu CL, Xiao Y, Wang X: **CIRBP is a novel oncogene in human bladder cancer inducing expression of HIF-1alpha**. *Cell Death Dis* 2018, **9**(10):1046.
39. Sun D, Wang Y, Wang H, Xin Y: **The novel long non-coding RNA LATS2-AS1-001 inhibits gastric cancer progression by regulating the LATS2/YAP1 signaling pathway via binding to EZH2**. *Cancer Cell Int* 2020, **20**:204.

Tables

Table 1: Functional enrichment analysis for differentially expressed RPBs.

Differentially expressed RPBs	Terms	P-value
Down-regulated RPBs		
Biological processes	RNA splicing	P<0.001
	regulation of RNA splicing	P<0.001
	regulation of mRNA processing	P<0.001
	regulation of mRNA metabolic process	P<0.001
	regulation of mRNA splicing, via spliceosome	P<0.001
Cellular component	RNA cap binding complex	P<0.001
	ribosome	P<0.001
	cytoplasmic stress granule	P<0.001
	cytoplasmic ribonucleoprotein granule	P<0.001
	ribonucleoprotein granule	P<0.001
Molecullar function	mRNA 3'-UTR binding	P<0.001
	translation factor activity, RNA binding	P<0.001
	single-stranded RNA binding	P<0.001
	poly(A) binding	P<0.001
	poly(U) RNA binding	P<0.001
KEGG pathway	RNA transport	P<0.001
	mRNA surveillance pathway	P<0.001
	RNA degradation	P<0.001
Up-regulated RPBs		
Biological processes	RNA catabolic process	P<0.001
	defense response to virus	P<0.001
	RNA phosphodiester bond hydrolysis	P<0.001
	response to virus	P<0.001
	regulation of translation	P<0.001
Cellular component	cytoplasmic ribonucleoprotein granule	P<0.001
	ribonucleoprotein granule	P<0.001
	P-body	P<0.001
	telomerase holoenzyme complex	P<0.001
	spliceosomal complex	P<0.001

Molecullar function	catalytic activity, acting on RNA	P<0.001
	RNA helicase activity	P<0.001
	helicase activity	P<0.001
	double-stranded RNA binding	P<0.001
	ribonuclease activity	P<0.001
KEGG pathway	Spliceosome	P<0.001
	RNA degradation	P<0.001
	RNA transport	P<0.001
	mRNA surveillance pathway	P<0.001
	Measles	P<0.001

Table 2: Univariate Cox regression analysis to identify prognosis-relatedRBPs.

RPBs	HR	HR.95L	HR.95H	P-value
DNMT1	0.66792	0.506998	0.87992	0.004111
MRPL33	1.568354	1.147441	2.143669	0.004764
EZH2	0.657414	0.505662	0.854708	0.001734
PCF11	0.508046	0.332979	0.775157	0.001681
PPARGC1B	0.576362	0.379043	0.8764	0.009967
RBM24	1.191707	1.011945	1.403401	0.035527
TRMT112	1.442278	1.0616	1.959463	0.019166
NUTF2	1.453638	1.039026	2.033697	0.029005
QARS	0.669615	0.461679	0.971202	0.034516
DZIP1	1.418783	1.023762	1.966223	0.035639
EIF5A2	1.526215	1.169747	1.991313	0.001838
MKRN3	2.306887	1.462073	3.639851	0.000328
DARS2	1.402555	1.018074	1.932236	0.038495
MBNL3	0.676222	0.493429	0.926732	0.014966
PSMA6	1.550197	1.122482	2.140889	0.007781
CPEB3	0.586725	0.344384	0.999602	0.049829
RBM47	0.78889	0.626623	0.993176	0.043566
ADARB2	0.534365	0.307906	0.927381	0.025881
MKRN2	0.651642	0.425701	0.997501	0.048669
AZGP1	0.88812	0.79568	0.9913	0.034363
LENG9	0.721531	0.533861	0.975172	0.033709
SIDT1	0.624264	0.405671	0.960645	0.03215
IGF2BP2	1.208821	1.047008	1.395641	0.009696
CIRBP	0.734818	0.552169	0.977884	0.034571

Table 3: The prognostic effect of different clinical parameters.

Index	Univariate analysis				Multivariate analysis			
	HR	95%Upper CI	95%Lower CI	P-value	HR	95%Upper CI	95%Lower CI	P-value
age	1.019682	1.004522	1.035071	0.010764	1.021692	1.005341	1.038308	0.00913
gender	0.744786	0.530085	1.046448	0.089453	0.867827	0.607693	1.239316	0.435525
grade	1.218909	0.948102	1.567067	0.122532	1.041376	0.791244	1.370583	0.772366
stage	1.678967	1.330204	2.119171	1.29E-05	1.180157	0.79929	1.742511	0.404758
T	1.378889	1.164345	1.632965	0.000197	1.148962	0.895807	1.473659	0.274177
M	1.088512	0.924793	1.281214	0.307814	1.049855	0.887407	1.24204	0.570545
N	1.551352	1.306974	1.841423	5.14E-07	1.39208	1.116608	1.735513	0.003279
riskScore	1.763024	1.470252	2.114095	9.36E-10	1.718394	1.377301	2.14396	1.62E-06

Figures

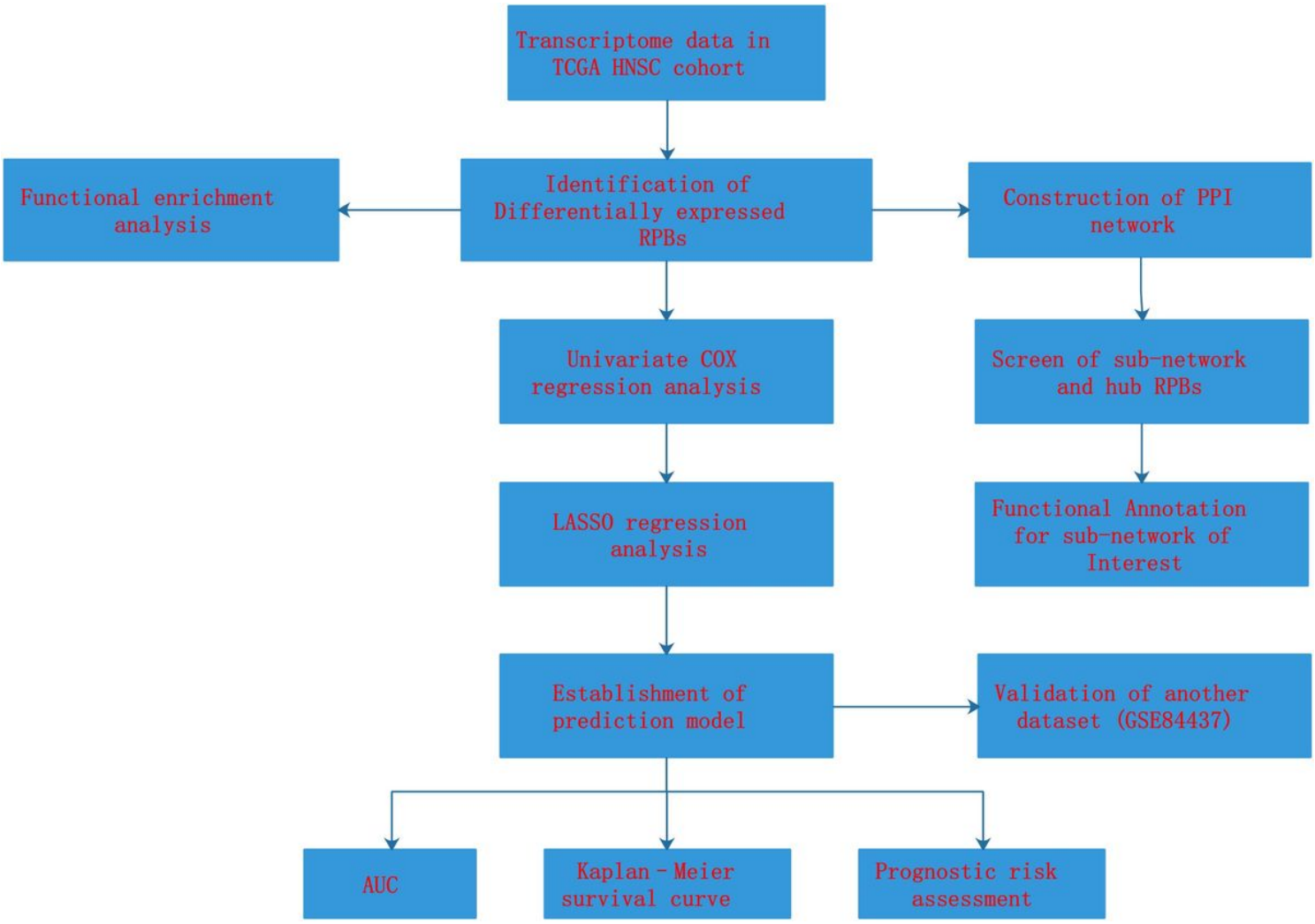

Figure 1

Flowchart delineates the study design and analysis process.

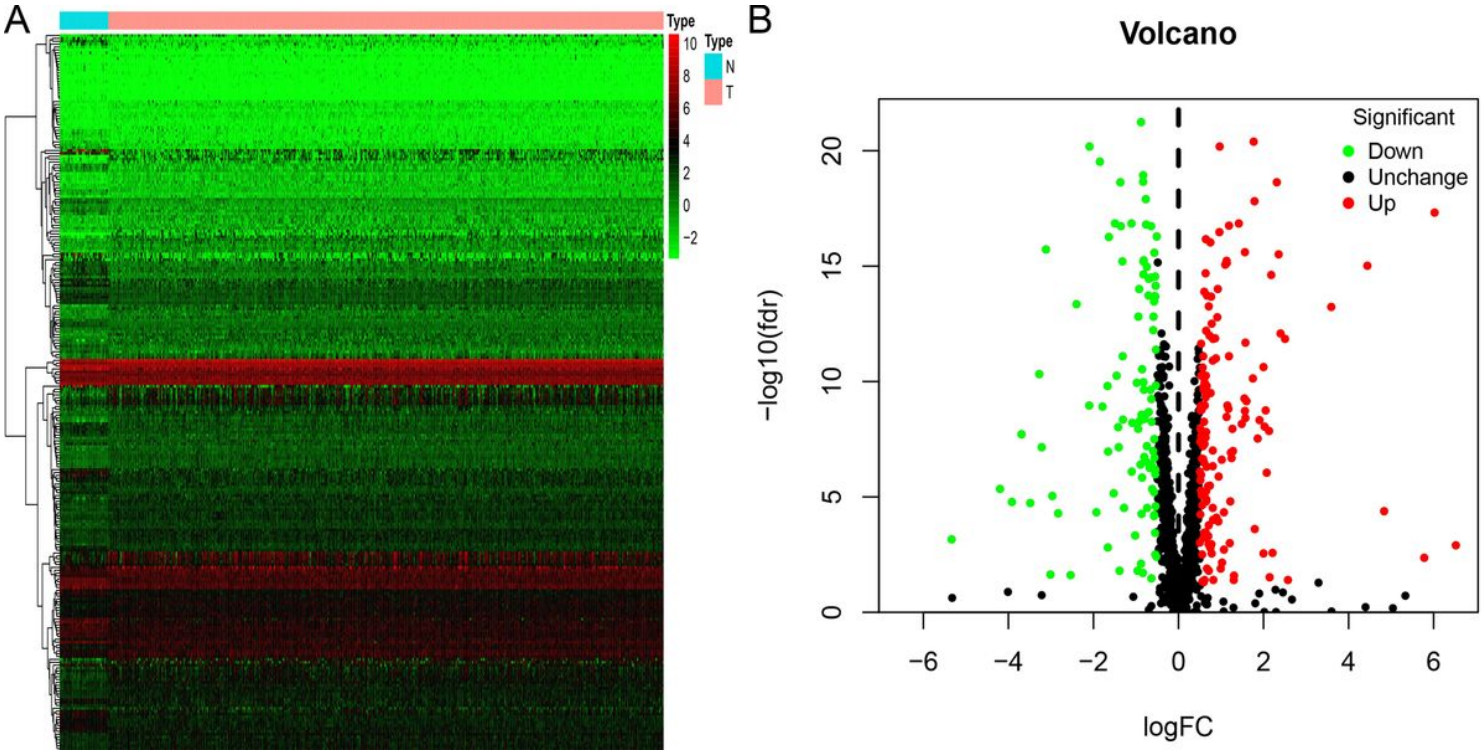

Figure 2

Identification of differentiallyexpressed RPBs between adjacent tumor samples and tumorsamples. The heatmap (A) and volcano (B) plot showed the differentiallyexpressed RPBs, a total of 249 differentially expressedRBPs were identified, including 104 downregulated RPBs and 145 up-regulated RPBs.

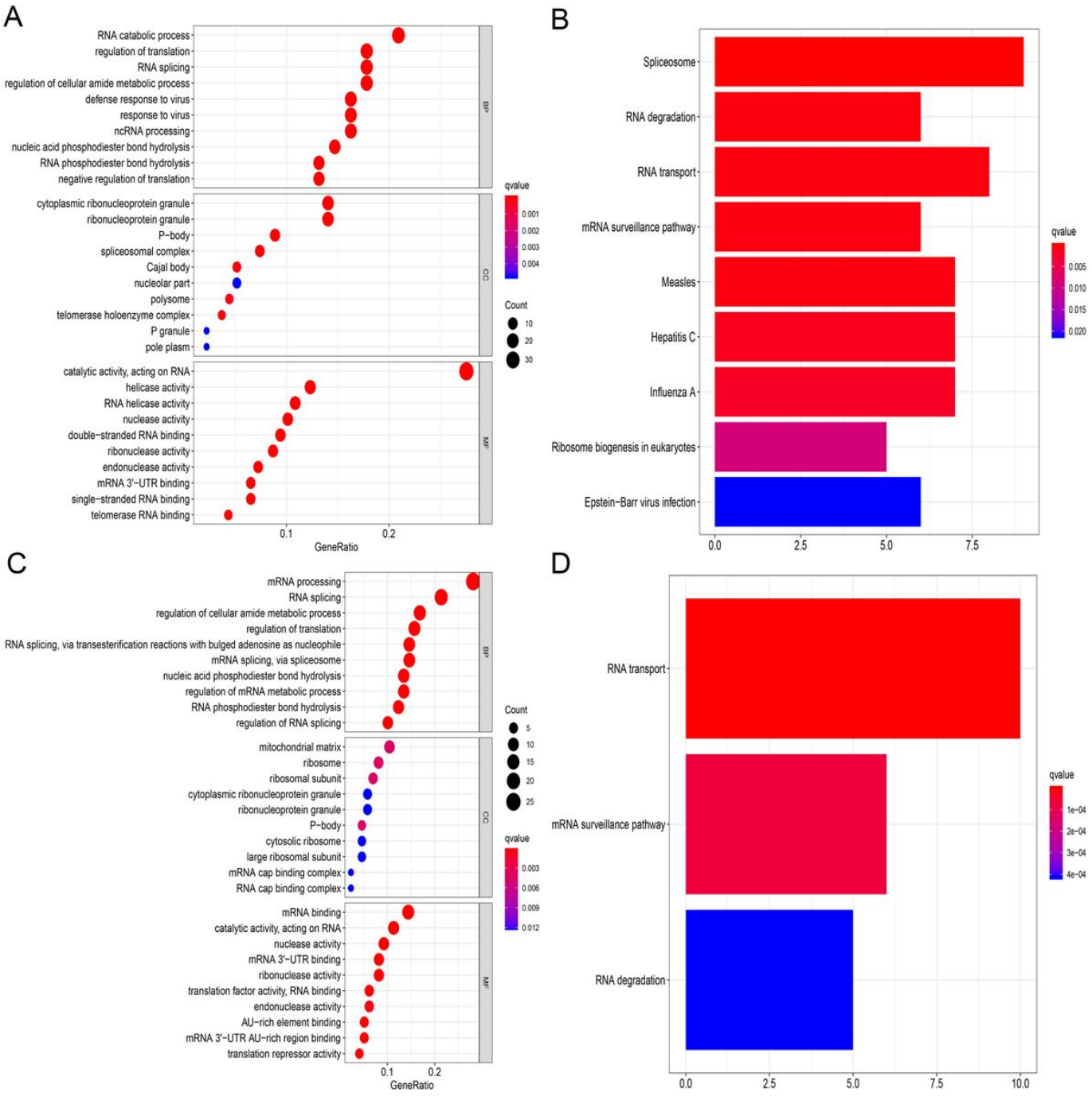

Figure 3

Functional enrichment analysis exhibited functions of the differentially expressed RPBs. A,C is for GO and KEGG analysis among upregulated RPBs; B,D is for GO and KEGG analysis among downregulated RPBs.

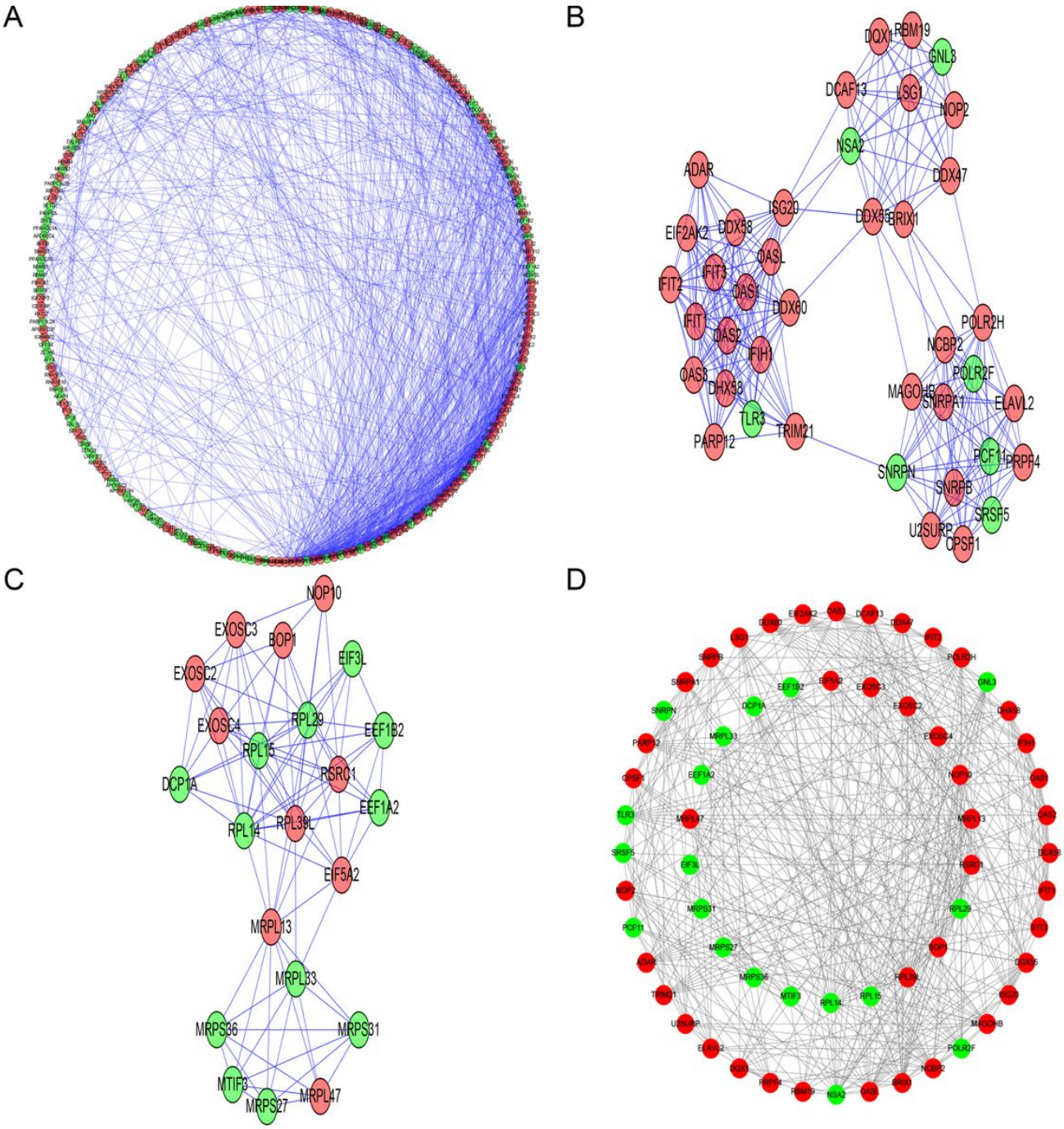

Figure 4

PPI network and module analysis. (A) PPI network for RBPs; (B) critical module I in PPI network; (C) critical module II in PPI network; (D) Circle diagram exhibited the module I (outlier) and II (inner), respectively.

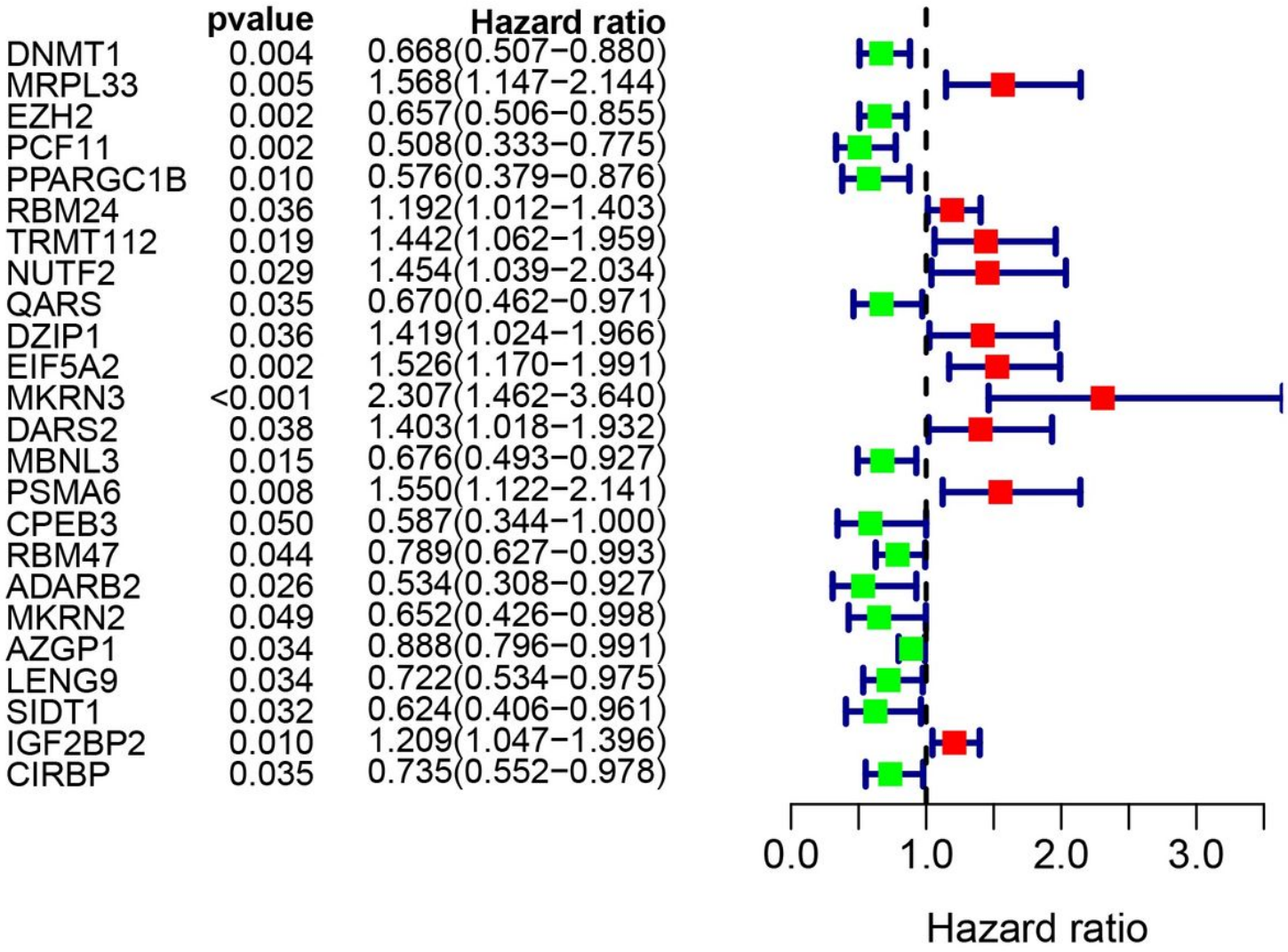

Figure 5

The forest plot of prognosis-related miRNAs between adjacent tumor samples and tumorsamples.

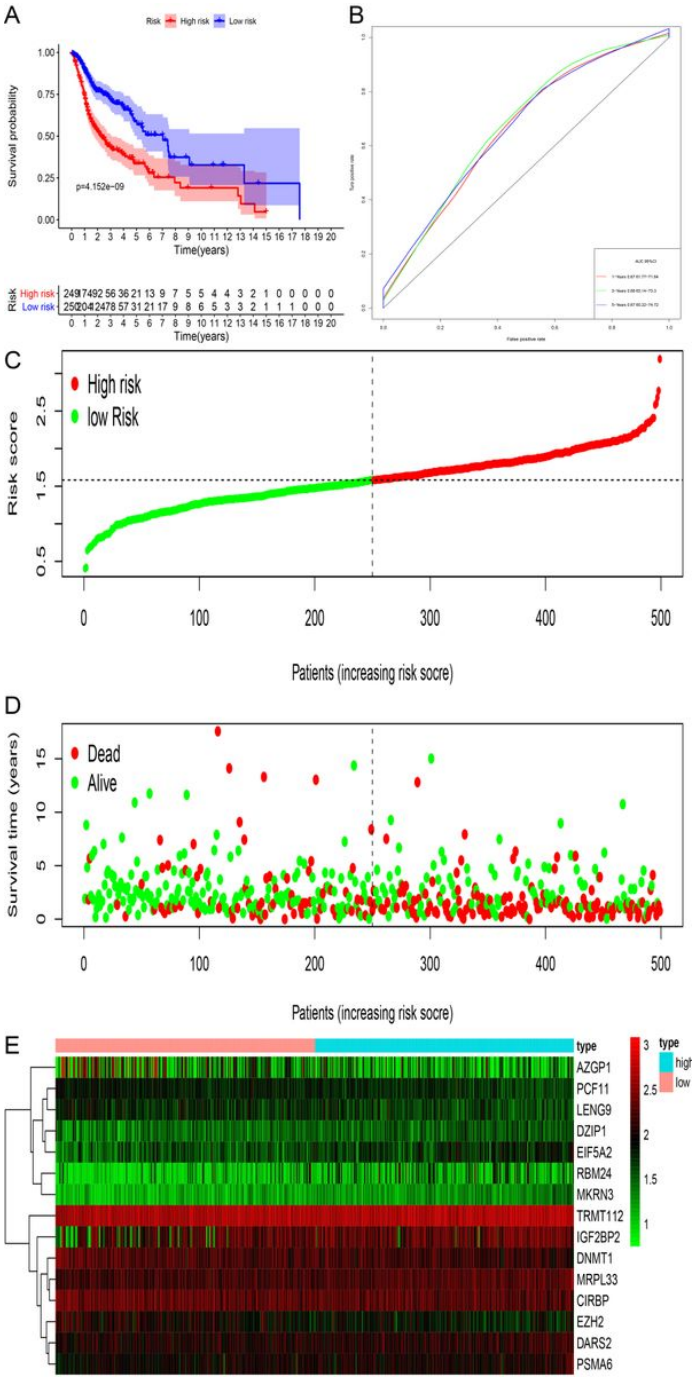

Figure 6

Risk score analysis of fifteen-gene prognostic model in TCGA HNSC cohort. (A) Survival analysis according to risk score; (B) ROC analysis; (C) Risk score of patients; (D) Heat map; (E) Survival status.

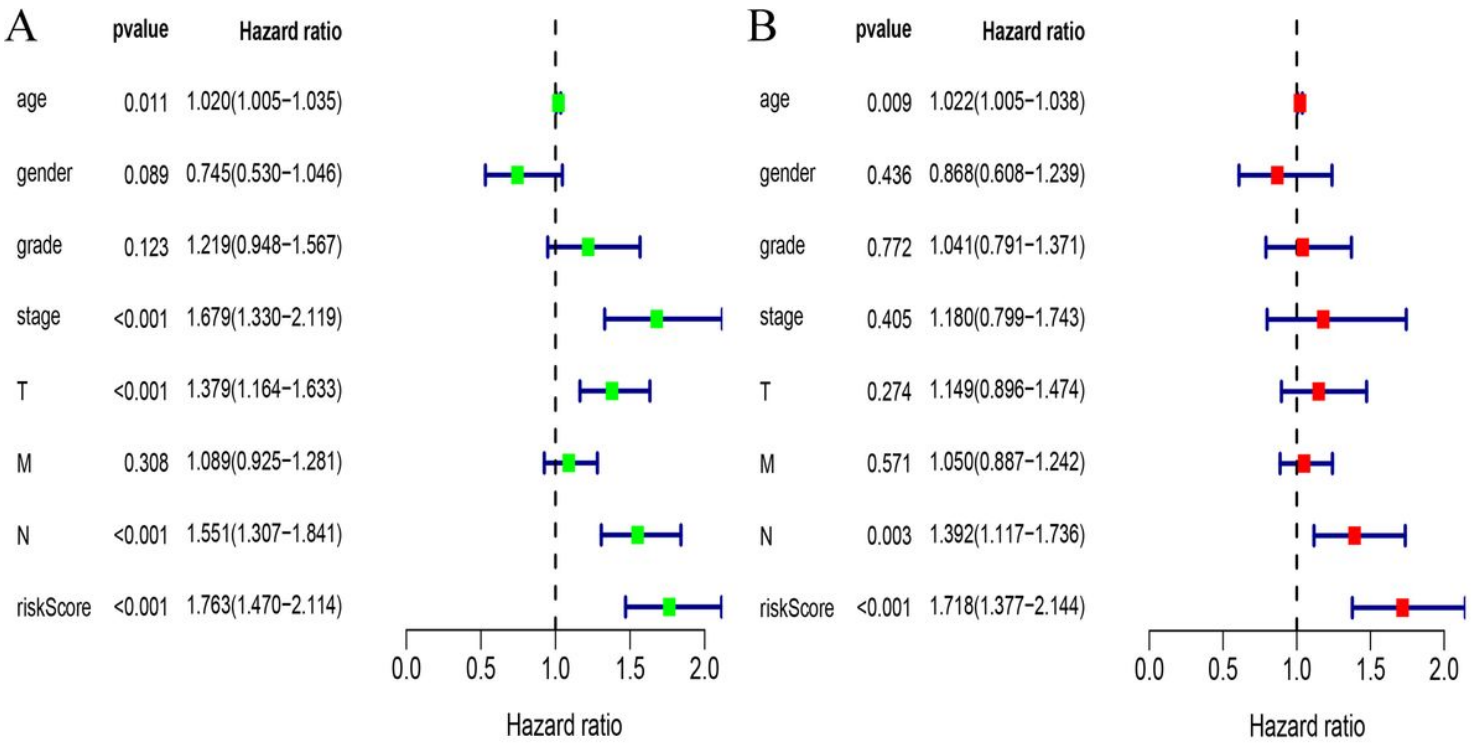

Figure 7

The prediction model could act as an independent predictor in unicovariate (A) and multivariate regression analysis (B) in TCGA HNSC cohort.

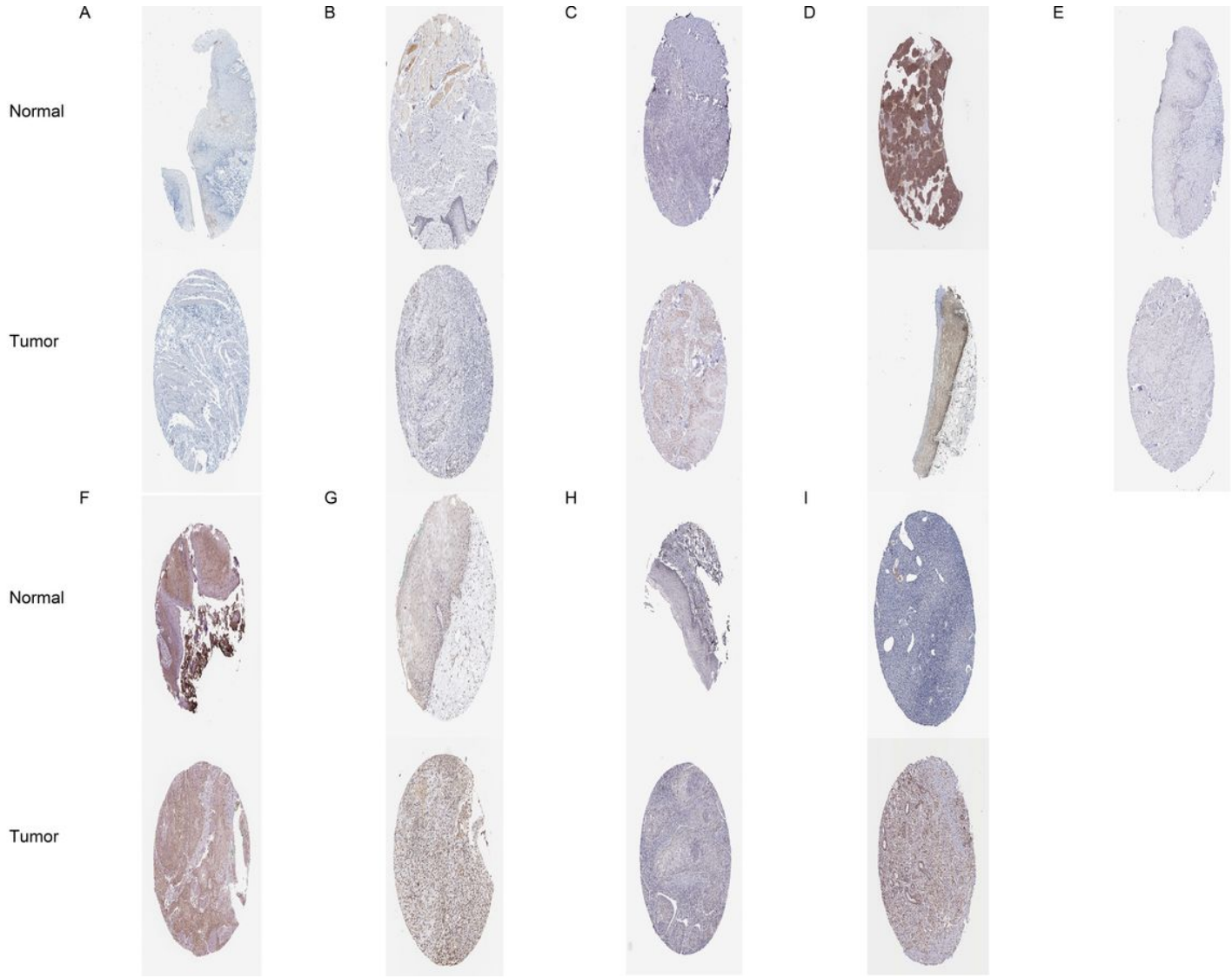

Figure 8

Validation of protein expression of hub genes in normal tissue and HNSC using the HPA database. (A) AZGP1, (B) DNMT1, (C) DZIP1, (D) IGF2BP2, (E) LENG9, (F) MRPL33, (G) PSMA6, (E) RBM24 and (D) TRMT112.

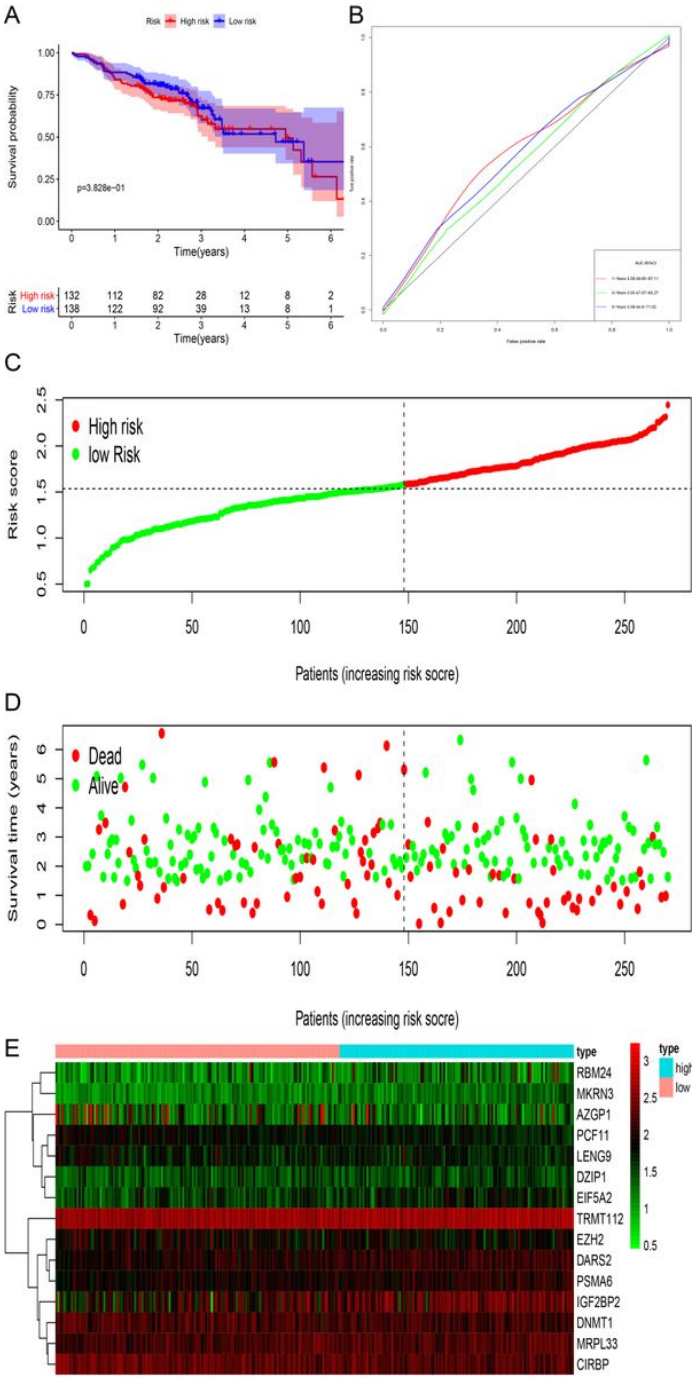

Figure 9

Risk score analysis of fifteen-gene prognostic model in GSE84437. (A) Survival analysis according to risk score; (B) ROC analysis; (C) Risk score of patients; (D) Heat map; (E) Survival status.

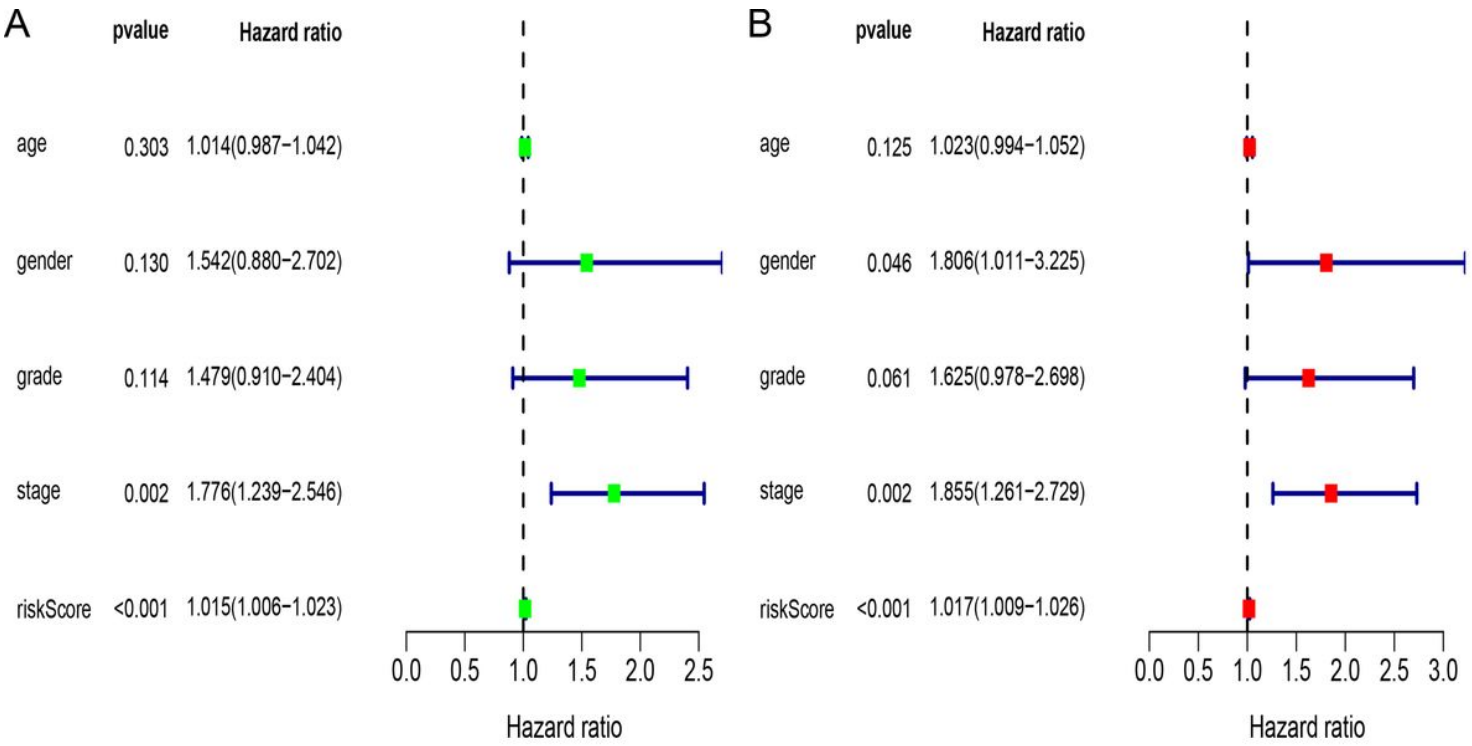

Figure 10

The prediction model could act as an independent predictor in unicovariate (A) and multivariate regression analysis (B) in GSE84437 dataset.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryTable.docx