

A Novel Prognostic Model In Oral Squamous Carcinoma: The Functions And Prognostic Values Of Rna Binding Proteins

Yingjuan Lu

Sun Yat-sen Memorial Hospital

Yongcong Yan

Sun Yat-Sen Memorial Hospital

Mo Liu

Sun Yat-Sen Memorial Hospital

Yancan Liang

Sun Yat-Sen Memorial Hospital

Yushan Ye

Sun Yat-Sen Memorial Hospital

Weiqi Cheng

Sun Yat-Sen Memorial Hospital

Jinsong Li

Sun Yat-Sen Memorial Hospital

Jiuyang Jiao

Sun Yat-Sen Memorial Hospital

Shaohai Chang (✉ changshaoh@126.com)

Sun Yat-sen Memorial Hospital <https://orcid.org/0000-0002-7820-3006>

Primary research

Keywords: oral squamous cell carcinoma, RNA binding proteins, prognostic model, bioinformatic tools

Posted Date: July 24th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-46791/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: The biological roles and clinical significance of RNA-binding proteins (RBPs) in oral squamous cell carcinoma (OSCC) are not fully understood. We investigated the prognostic value of RBPs in OSCC by several bioinformatic strategies.

Methods: OSCC data were obtained from a public online database, the Limma R package was used to identify differentially expressed RBPs, and functional enrichment analysis was performed to elucidate the biological functions of the above RBPs in OSCC. We performed protein-protein interaction (PPI) network and Cox regression analyses to extract prognosis-related hub RBPs. Next, we established and validated a prognostic model based on the hub RBPs by Cox regression and risk score analyses.

Results: We found that the differentially expressed RBPs were closely related to the defence response to virus and multiple RNA processes. We obtained ten prognosis-related hub RBPs (ZC3H12D, OAS2, INTS10, ACO1, PCBP4, RNASE3, PTGES3L-AARSD1, RNASE13, DDX4, and PCF11) and effectively predicted the overall survival of OSCC patients. The area under the ROC curve (AUC) of the risk score model was 0.781, suggesting that our model had good prognostic performance. Finally, we built a nomogram integrating the ten RBPs. The internal validation cohort results showed a reliable predictive capability of the nomogram for OSCC.

Conclusions: We established a novel ten-RBP-based model for OSCC that could enable precise therapeutic targets in the future.

Introduction

Oral cancer, representing a group of tumours located in the alveolar ridge, buccal cavity, mucosa, floor of the mouth, palate, tongue and other parts of the oral cavity, is the eighth most frequent cancer and accounts for an estimated 350 000 new cases and 170 000 deaths per year worldwide. Of these, oral squamous cell carcinoma (OSCC) accounts for more than 95% of oral tumours. A major proportion of global OSCC cases are diagnosed in Asia [1]. Despite some developments in the therapeutic strategies of OSCC that have been achieved over the last few years, the long-term survival rates of OSCC patients are still extremely low, and despite the promising approval of checkpoint inhibitors, precise targeted therapies remain limited in changing treatment failures owing to individual differences in genetic and epigenetic changes among different OSCC patients [2–4]. Thus, a systematic study to identify key biomarkers for diagnosis and effective targets for treatment is urgently needed for OSCC.

Nomograms are statistics-based tools that can effectively integrate different variables and are widely used in detecting prognostic factors for various cancers [5, 6]. Many models have been established for predicting the outcomes of OSCC by calculating the risk of clinicopathologic characteristics, such as age, sex, smoking status, TNM stage and therapies [7, 8]. Although these studies have enhanced our comprehension of OSCC, few have taken into account genetic features, and they are not accurate and reliable enough, as the underlying molecular mechanisms of OSCC might be considerably complex [9].

RNA-binding proteins (RBPs) are proteins that regulate the functions of RNA transcripts at multiple levels [10]. Currently, a total of 1542 RBPs have been identified, and more than 800 mRNAs have been discovered in human cells [11, 12]. These RBPs bind RNA through the globular RNA-binding domains (RBDs) and affect the splicing, stability or translation of their target RNAs [13, 14]. Considering the importance of RBPs in many cellular processes, it is unsurprising that defects in their functions affect several human diseases, including, neurological disorders, muscular atrophies and cancer [15]. Many studies have indicated that the dysregulation of RBPs is related to the biological behaviours of various cancer cells. For instance, human antigen R (HuR) is widely overexpressed in cancer cells and is known to be involved in posttranscriptional gene regulation [16]. HuR not only binds to protein-coding mRNAs but also affects lncRNAs with both positive and negative influences on their stability [17]. Serine/arginine-rich splicing factor 1 (SRSF1) is another RBP that is upregulated in glioma and handles oncogenic impacts by regulating splicing, RNA stability, and nuclear export [18]. The abnormal expression of heterogeneous nuclear ribonucleoprotein K (hnRNPK) is frequently observed in cancer and might play an important role in hepatocellular carcinoma (HCC) by changing the localization of MALAT1 [19]. Insulin-like growth factor 2 mRNA-binding protein 1 (IGF2BP1) may play an important role in regulating cell proliferation and apoptosis in HCC [20]. Mitogen-activated protein kinase-activated protein kinase-2 (MAPKAPK2) regulates transcript stability and exerts important effects in head and neck squamous cell carcinoma [21]. To sum up, these previous studies have shown that RBPs may play important roles in tumorigenesis and tumour progression. However, only a few of them have been studied recently and revealed to have critical functions in cancers.

In this study, we integrated all relevant OSCC information from The Cancer Genome Atlas (TCGA) database to perform a novel systematic analysis. To reveal the potential functions and clinical significance of RBPs in the pathogenesis of OSCC, we identified multiple differentially expressed RBPs associated with OSCC. Some of them may serve as potential biomarkers for diagnosis and prognosis.

Materials And Methods

Dataset Collection and Identification of Differentially Expressed RBPs

The RNA-sequencing data of 32 normal oral tissues and 331 OSCC tissues with corresponding clinical information were obtained from TCGA (<https://portal.gdc.cancer.gov/>). Approval of the institutional ethics committee was not necessary because the data were collected from publicly available online databases. To identify the differentially expressed genes between normal oral and OSCC tissue samples, all raw data were preprocessed by the Limma package (<http://www.bioconductor.org/packages/release/bioc/html/limma.html>), and genes with an average count value of less than one were excluded. In addition, we also identify the differently expressed RBPs with the thresholds of $|\log_2\text{-fold change (FC)}| \geq 1$ and false discovery rate (FDR) < 0.05 .

KEGG Pathway and GO Enrichment Analyses

The biological functions of these differentially expressed RBPs were systematically detected by Gene Ontology (GO) enrichment and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses. The GO analysis included terms in the cellular component (CC), molecular function (MF), and biological process (BP) categories. All enrichment analyses were carried out by WEB-based Gene Set Analysis Toolkit (WebGestalt, <http://www.webgestalt.org/>) [22]. For both analysis, P values less than 0.05 and gene numbers greater than 5 were considered statistically significant.

PPI Network Construction and Module Screening

The protein-protein interactions (PPIs) among all differently expressed RBPs were investigated in the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) database (<http://www.string-db.org/>) [23]. Cytoscape 3.8.0 software was used to further visualize the PPI network. The key modules and genes in the PPI network were selected by using the Molecular Complex Detection (MCODE) plug-in with both MCODE scores and node count numbers greater than 5 [24]. A P value < 0.05 was considered to indicate a significant difference.

Prognosis-related RBP Identification and Prognostic Model Construction

We constructed Univariate Cox regression analysis on the differentially expressed RBPs in the critical modules of the training dataset using the survival R package. Subsequently, we performed the log-rank test to extract the prognosis-related RBPs. Based on the above preliminarily screened significant candidate RBPs, a multivariate Cox proportional hazards regression model was further performed and the risk score (RS) was calculated to predict the prognostic outcomes of OSCC patients [25]. The risk score formula for each sample was as follows: $RS = \text{Expression}_{\text{gene1}} \times \beta_{\text{gene1}} + \dots + \text{Expression}_{\text{genen}} \times \beta_{\text{genen}}$ (where β is the regression coefficient derived from multivariate Cox regression). According to the median RS analysis, the OSCC patients were divided into low-risk and high-risk groups. The log-rank test was used to compare the difference in overall survival (OS) between the two groups. Besides, the survival ROC package was provided to assess the prognostic capability of the above model, and receiver operating characteristic (ROC) curve analysis was executed [26]. In addition, an internal validation cohort was used to test the predictive capability of the nomogram. Calibration curves were constructed using the rms R package to compare the observed probabilities and the predicted probabilities of the nomogram. A P value < 0.05 indicated a significant difference.

Results

Identification of Differentially Expressed RBPs in OSCC

The workflow of this study is illustrated in Figure 1. RNA-sequencing data for OSCC and corresponding clinical information were downloaded from the TCGA database. In total, 331 OSCC samples and 32 normal oral samples were analysed. The Limma R software package was applied to preprocess these data and to identify the differentially expressed RBPs [11]. A total of 1542 RBPs were analysed in this

study, and 257 met our inclusion criteria ($P < 0.05$, $|\log_2FC| > 1.0$), which consisted of 121 downregulated and 136 upregulated RBPs. The expression of these differentially expressed RBPs is shown in Figure 2.

Functional Enrichment Analysis of the Differentially Expressed RBPs

The identified RBPs were divided into two groups, the upregulation group and the downregulation group, to explore their potential functions and mechanisms. We carried out GO and pathway analyses for these differentially expressed RBPs. The results showed that the upregulated RBPs were significantly enriched in defense response to virus and regulation of mRNA metabolic process (Table 1, Figure 3A), while the downregulated RBPs were considerably enriched in mRNA processing, RNA splicing, and RNA catabolic process (Table 1, Figure 3B). Regarding the CC category, the upregulated RBPs were enriched in cytoplasmic ribonucleoprotein granule, ribonucleoprotein granule, and spliceosomal complex (Table 1, Figure 3A), while the downregulated RBPs were enriched in the ribosome, ribosomal subunit, and mitochondrial matrix (Table 1, Figure 3B). The MF analysis indicated that the upregulated differentially expressed RBPs were mainly enriched in catalytic activity acting on RNA, nuclease activity, and helicase activity (Table 1, Figure 3A), whereas the downregulated RBPs were notably enriched in mRNA binding, catalytic activity acting on RNA, and nuclease activity (Table 1, Figure 3B). Additionally, the results of KEGG pathway analysis showed that the upregulated RBPs were enriched in the RNA transport, spliceosome, and mRNA surveillance pathways (Table 1, Figure 3C), while the downregulated RBPs were mainly enriched in the RNA transport, ribosome, and mRNA surveillance pathways (Table 1, Figure 3D).

PPI Network Construction and Key Module Screening

To better understand the potential roles of these identified RBPs in OSCC, we performed a PPI network using the STRING database and Cytoscape software. This PPI network included 228 nodes and 1384 edges in total (Figure 4A). We further analysed the coexpression network to detect potential critical modules using the MCODE plug-in and identified the key modules (Figure 4B). Of these, the first essential module, module 1, contained 49 nodes and 316 edges, the second module, module 2, consisted of 8 nodes and 28 edges, module 3 consisted of 13 nodes and 43 edges, module 4 consisted of 6 nodes and 12 edges, the next module consisted of 4 nodes and 6 edges, and the last module consisted of 16 nodes and 29 edges (Supplementary Figure 1A-F). The functional analysis showed that the RBPs in these key modules were mainly enriched in mitochondrial translation, mitochondrial gene expression, and cellular protein complex disassembly. More details are provided in Supplementary Table 1.

Identification of Prognosis-related RBPs

Of the differentially expressed RBPs from the PPI network, 20 prognosis-associated candidate hub RBPs were identified by univariate Cox regression analysis (Supplementary Table 2). The candidate RBPs were analysed by multivariate Cox regression to determine their impact on patient survival time and clinical outcomes. Ten hub RBPs, ZC3H12D, OAS2, INTS10, ACO1, PCBP4, RNASE3, PTGES3L-AARSD1, RNASE13, DDX4 and PCF11, were identified as independent predictors in OSCC (Supplementary Table 3).

Prognosis-related Genetic Risk Score Model Construction and Validation

The ten hub RBPs obtained by multivariate Cox regression analysis were used to construct a prognostic model. To assess the predictive ability of this model, we further carried out a risk score analysis in the training cohort. The risk score of each patient was calculated according to the following formula: $RS = ZC3H12D * -1.505 + OAS2 * -0.006 + INTS10 * -0.135 + ACO1 * 0.108 + PCBP4 * -0.122 + RNASE3 * 2.757 + PTGES3L - AARSD1 * 2.827 + RNASE13 * -11.340 + DDX4 * 8.283 + PCF11 * -0.204$.

A total of 221 OSCC patients were divided into low-risk and high-risk groups according to the median RS. The results showed that patients in the low-risk group had significantly longer OS times than those in the high-risk group (Figure 5A). The time-dependent ROC analysis was further conducted to estimate the prognostic ability of the identified RBPs. From the chart, we can see that the area under the ROC curve (AUC) of the risk score model was 0.781 (Figure 5B), which suggested that our model has high prognostic accuracy. The expression heat map of the ten prognostic RBPs is illustrated in Figure 5C, and the RS and survival status for the low- and high-risk groups are displayed in Figure 5D. Moreover, the results from our internal validation cohort, using the same formula to assess the predictive capability of the ten-gene prognostic model, showed that patients with lower risk scores had better OS (Figure 6A-D). These results indicated that the prognostic model has significant predictive capability.

The Prognostic Value of Different Clinical Parameters

To assess the prognostic significance of different clinical features, we performed a univariate Cox regression analysis of OSCC patients in the training set and internal validation set. The results showed that stage, grade, age and RS were correlated with the OS of OSCC patients in both cohorts ($P < 0.05$) (Figure 7A, B), whereas only stage and RS were independent prognostic factors for OS in the multiple regression analysis of both cohorts ($P < 0.05$) (Figure 7C, D).

Development and Validation of a Nomogram Based on the Hub RBPs

We established a nomogram for predicting the 1-, 2- and 3-year OS based on the multivariate analysis results (Figure 8A). The points for each independent prognostic factor listed in the nomogram were summed the total points value, which is projected on the bottom scale and used to determine the 1-, 2- and 3-year OS probabilities for each patient. We plotted the calibration curves, which indicated that there was no apparent deviation from the ideal line, with optimal accordance achieved between the nomogram-predicted outcomes and the actual survival rates in both the training and validation cohorts (Figure 8B, C).

Discussion

The survival rate of OSCC patients remains low since there are limited prognostic factors that could provide effective targeted therapies and precisely predictable outcomes [9]. In recent years, the regulation of gene expression by RBPs has become the centre of attention [27]. Previous studies have demonstrated

that RBPs, which regulate oncogene expression at the posttranscriptional level, significantly affect tumorigenesis and progression, indicating that RBPs are potential cancer biomarkers [28]. In fact, several RBPs have been identified to predict OS in cancer. Zhao et al identified a novel extra-ribosomal role of RPS3 as a potential therapeutic target in HCC via the posttranscriptional regulation of SIRT1 expression [29]. Soni et al highlighted that MK2, as the master regulator of RBPs, plays an important role in the regulation of transcript stability and tumour progression, as well as the possibility of the use of MK2 as a therapeutic target in tumour management [30]. However, the role of RBPs is rarely understood in OSCC, even though there have been studies focused on the functions of RBPs in head and neck carcinoma. Berggen et al revealed that MK2 could be a potential prognostic biomarker for head and neck cancer and that MK2 pathway activation can mediate radiation resistance in head and neck squamous cell carcinoma (HNSCC) [31]. Jiang et al indicated that LINC00460 is a promising candidate prognostic predictor for HNSCC by facilitating the entry of an RBP called PRDX1 into the nucleus and promoting epithelial-mesenchymal transition (EMT) in HNSCC cells [32]. However, most of these studies analysed the functions of single RBPs, while information is now rapidly emerging on the vital functional role of the molecular network in tumour initiation and progression, indicating that we should investigate prognostic markers collectively [33]

In the present study, we identified dysregulated RBPs in OSCC based on the TCGA database and systematically analysed the related biological pathways of these RBPs. We further established a PPI network to screen the modules that may play key roles in OSCC. Subsequently, we executed Cox regression analysis and obtained ten prognosis-related RBPs. Moreover, we carried out risk score analysis of this ten-gene prognostic model in the training and validation cohorts to explore the potential biological functions and clinical significance of these ten hub RBPs. The model exhibited good predictive capability in both cohorts. Finally, we constructed and validated a nomogram to predict OSCC prognosis based on these hub RBPs. Our findings may provide novel biomarkers for indicating the prognosis of OSCC patients.

The biological function and pathway enrichment analyses of the differentially expressed RBPs showed that the upregulated RBPs were significantly enriched in the defense response to virus, response to virus, regulation of mRNA metabolic process, RNA transport, RNA degradation, spliceosome, and mRNA surveillance pathways. The downregulated RBPs were mainly enriched in the mRNA processing, RNA splicing, RNA catabolic process, RNA transport, ribosome, and mRNA surveillance pathways. Many studies have reported the role of metabolism, RNA degradation and RNA transport in OSCC [34–36]. Human papillomavirus (HPV) is closely associated with OSCC [37]. HPV infection significantly increases the number of binding sites for RBPs and subsequently upregulates a group of oncogenic genes by absorbing RBPs [38]. A previous study showed that RBPs can regulate mRNA stability or translational efficiency via ribosomes and are involved in various diseases [39]. These results suggest that RBPs can affect the biological behaviour of OSCC by interacting with components related to HPV or regulating multiple biological processes of OSCC cells.

We obtained ten hub RBPs by constructing a PPI network and performing Cox regression analysis: ZC3H12D, OAS2, INTS10, ACO1, PCBP4, RNASE3, PTGES3L-AARSD1, RNASE13, DDX4, and PCF11. Most of them are associated with tumorigenesis and the progression of cancer patients, such as OAS2 [40], INTS10 [41], ACO1 [42], PCBP4 [43], DDX4 [44] and PCF11 [45]. Of these, OAS2, an interferon (IFN)-induced antiviral enzyme, has been reported as an OSCC-related gene. Similarly, Wang et al analysed three microarray datasets and found that the expression of OAS2 was closely related to the survival rate of patients with oral cancer [46], which was consistent with our results. Interestingly, the IFN system is a central and powerful host first-line antiviral defence strategy that is inhibited by HPV 16 in HNSCC [47]. A significant level of OAS2 could activate the IFN signalling pathway and in turn produce a T-cell response that is effective in clearing HPV infection [48].

Thus, OAS2 may serve as a key regulator and may provide rational insight into immunotherapy for HPV OSCC. However, the detailed mechanisms underlying these hub RBPs in OSCC are still unexplored, and further in-depth studies to discover their potential roles would be valuable.

Overall, these results suggested that our prognostic model of the ten-RBPs has a certain value for the treatment decisions of OSCC patients. Nonetheless, our study had several limitations. First, the multivariable survival analysis contained only basic prognostic factors from the TCGA database and was unable to provide other possible clinical factors, such as the status of metastatic lesions and the HPV status of patients, which may affect the reliability of the Cox regression analysis. Second, extensive evidence indicates that OSCC is an extremely heterogeneous tumour at the genetic and molecular levels, and limited by the data of the study, the expression of all genes from the TCGA cohort was detected in a piece of OSCC tissue from one patient. In the future, we will detect the expression of the ten genes by single-cell whole-genome sequencing or quantitative RT-PCR analysis in several pieces of OSCC specimens from one patient to determine whether the ten-gene signature is a reliable and workable OS prediction marker for OSCC. In addition, we will seek cooperation with other hospitals to obtain more clinical patient cohorts for gene model validation.

Conclusion

We systemically explored the prognostic value of differentially expressed RBPs by performing a series of bioinformatics analyses in OSCC. A prognostic model of the ten hub RBPs was constructed and further validated. Our results were used to establish a novel prognostic model for OSCC and will contribute new candidate markers to aid in making therapeutic decisions in the future.

List Of Abbreviations

OSCC: Oral Squamous Cell Carcinoma

RBPs: RNA Binding Proteins

HNSCC: head and neck squamous cell carcinoma

Declarations

Ethics approval and consent to participate

Approval of the institutional ethics committee was not necessary because the data were collected from publicly available online databases.

Availability of data and materials:

The RNA-sequencing data of 331OSCC samples and 32 normal oral samples with corresponding clinical information were downloaded from the Cancer Genome Atlas (TCGA) database (<https://portal.gdc.cancer.gov/>).

Competing interests:

The authors have declared that no competing interest exists.

Fundings

This work was supported by the National Natural Science Foundation of China (NSFC; 81900959 to Y.L., 81802700 to Y.L.)

Authors' contributions:

Conception and design: Y. Lu, Y. Yan

Development of methodology: Y. Lu, Y. Yan

Analysis and interpretation of data (e.g., statistical analysis, bioinformatics analysis): Y. Lu, Y. Yan, M. Liu, Y. Liang

Writing, review, and/or revision of the manuscript: Y. Lu, Y. Yan, M. Liu, Y. Liang, Y. Ye, W. Cheng, J. Jiao, J. Li, S. Chang

Administrative, technical, or material support (i.e., reporting or organizing data): Yingjuan Lu, M. Liu, Y. Liang, J. Li, J. Jiao, S. Chang

Study supervision: J. Li, J. Jiao, S. Chang

Acknowledgements

Yingjuan Lu: conceptualization, formal analysis, investigation, funding acquisition, and writing-original draft. Yongcong Yan: conceptualization, formal analysis, investigation, visualization, and writing-original draft. Mo Liu: formal analysis, technical support, administration, and writing-original draft. Yancan Liang: formal analysis, funding acquisition, administration, and writing-original draft. Yushan Ye: writing-original draft. Weiqi Cheng: writing-original draft. Jinsong Li: methodology, resources, and supervision, and

writing-review and editing. Jiuyang Jiao: project administration, resources, supervision, and writing-review and editing. Shaohai Chang: investigation, methodology, project administration, resources, supervision, and writing-review and editing. We also appreciate American Journal Experts for their professional language polishing.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2018;68(6):394–424.
2. Colevas AD, Yom SS, Pfister DG, et al. NCCN Guidelines Insights: Head and Neck Cancers, Version 1.2018. *J Natl Compr Canc Netw.* 2018;16(5):479–90. doi:10.6004/jnccn.2018.0026.
3. Argiris A, Karamouzis MV, Raben D, Ferris RL. Head and neck cancer. *Lancet.* 2008;371(9625):1695–709.
4. Liu X, Zhao X, Gou C. Identification of novel methylated DNA marker ZNF569 for head and neck squamous cell carcinoma. *J Cancer.* 2019;10(10):2250–60.
5. Wang K, Li L, Fu L, et al. Integrated Bioinformatics Analysis the Function of RNA Binding Proteins (RBPs) and Their Prognostic Value in Breast Cancer. *Front Pharmacol.* 2019;10:140.
6. Li W, Li X, Gao LN, You CG. Integrated Analysis of the Functions and Prognostic Values of RNA Binding Proteins in Lung Squamous Cell Carcinoma. *Front Genet.* 2020;11:185.
7. Wang F, Zhang H, Wen J, et al. Nomograms forecasting long-term overall and cancer-specific survival of patients with oral squamous cell carcinoma. *Cancer Med.* 2018;7(4):943–52.
8. Chen F, Lin L, Liu F, et al. Three prognostic indexes as predictors of response to adjuvant chemoradiotherapy in patients with oral squamous cell carcinoma after radical surgery: A large-scale prospective study. *Head Neck.* 2019;41(2):301–8.
9. Mattavelli D, Lombardi D, Missale F, et al. Prognostic Nomograms in Oral Squamous Cell Carcinoma: The Negative Impact of Low Neutrophil to Lymphocyte Ratio. *Front Oncol.* 2019;9:339.
10. Nahalka J. The role of the protein-RNA recognition code in neurodegeneration. *Cell Mol Life Sci.* 2019;76(11):2043–58.
11. Gerstberger S, Hafner M, Tuschl T. A census of human RNA-binding proteins. *Nat Rev Genet.* 2014;15(12):829–45.
12. Kechavarzi B, Janga SC. Dissecting the expression landscape of RNA-binding proteins in human cancers. *Genome Biol.* 2014;15(1):R14.
13. Masuda K, Kuwano Y. Diverse roles of RNA-binding proteins in cancer traits and their implications in gastrointestinal cancers. *Wiley Interdiscip Rev RNA.* 2019;10(3):e1520.
14. Lunde BM, Moore C, Varani G. RNA-binding proteins: modular design for efficient function. *Nat Rev Mol Cell Biol.* 2007;8(6):479–90.

15. Castello A, Fischer B, Hentze MW, Preiss T. RNA-binding proteins in Mendelian disease. *Trends Genet.* 2013;29(5):318–27.
16. Wang E, Aifantis I. RNA Splicing and Cancer [published online ahead of print, 2020 May 17]. *Trends Cancer.* 2020;S2405-8033(20):30141–2.
17. Hu YP, Jin YP, Wu XS, et al. LncRNA-HGBC stabilized by HuR promotes gallbladder cancer progression by regulating miR-502-3p/SET/AKT axis. *Mol Cancer.* 2019;18(1):167.
18. Zhou X, Li X, Yu L, et al. The RNA-binding protein SRSF1 is a key cell cycle regulator via stabilizing NEAT1 in glioma. *Int J Biochem Cell Biol.* 2019;113:75–86.
19. Nguyen TM, Kabotyanski EB, Reineke LC, et al. The SINEB1 element in the long non-coding RNA Malat1 is necessary for TDP-43 proteostasis. *Nucleic Acids Res.* 2020;48(5):2621–42.
20. Gutschner T, Hämmerle M, Pazaitis N, et al. Insulin-like growth factor 2 mRNA-binding protein 1 (IGF2BP1) is an important protumorigenic factor in hepatocellular carcinoma. *Hepatology.* 2014;59(5):1900–11.
21. Soni S, Saroch MK, Chander B, Tirpude NV, Padwad YS. MAPKAPK2 plays a crucial role in the progression of head and neck squamous cell carcinoma by regulating transcript stability. *J Exp Clin Cancer Res.* 2019;38(1):175.
22. Liao Y, Wang J, Jaehnig EJ, Shi Z, Zhang B. WebGestalt 2019: gene set analysis toolkit with revamped UIs and APIs. *Nucleic Acids Res.* 2019;47(W1):W199–205.
23. Szklarczyk D, Gable AL, Lyon D, et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* 2019;47(D1):D607–13.
24. Bader GD, Hogue CW. An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinformatics.* 2003;4:2.
25. Jiang Y, Li T, Liang X, et al. Association of Adjuvant Chemotherapy With Survival in Patients With Stage II or III Gastric Cancer. *JAMA Surg.* 2017;152(7):e171087.∗.
26. Heagerty PJ, Lumley T, Pepe MS. Time-dependent ROC curves for censored survival data and a diagnostic marker. *Biometrics.* 2000;56(2):337–44.
27. Wurth L, Gebauer F. RNA-binding proteins, multifaceted translational regulators in cancer. *Biochim Biophys Acta.* 2015;1849:881–6.
28. Pereira B, Billaud M, Almeida R. RNA-binding proteins in cancer: old players and new actors. *Trends Cancer.* 2017;3:506–28.
29. Zhao L, Cao J, Hu K, et al. RNA-binding protein RPS3 contributes to hepatocarcinogenesis by post-transcriptionally up-regulating SIRT1. *Nucleic Acids Res.* 2019;47(4):2011–28.
30. 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. Published 2019 Mar 8.

31. Berggren K, Cruz SR, Hixon MD, Cowan A, Ozbun MA, Keysar S, Jimeno A, Ness SA, McCance DJ, Beswick EJ, Gan GN. Inhibition of MK2 decreases inflammatory cytokine production and tumor volumes in HPV-positive and HPV-negative models of head and neck squamous cell carcinoma. *Int J Radiat Oncol Biol Phys*. 2018;100:1372–3.
32. Jiang Y, Cao W, Wu K, et al. LncRNA LINC00460 promotes EMT in head and neck squamous cell carcinoma by facilitating peroxiredoxin-1 into the nucleus. *J Exp Clin Cancer Res*. 2019;38(1):365.
33. Li W, Gao LN, Song PP, You CG. Development and validation of a RNA binding protein-associated prognostic model for lung adenocarcinoma. *Aging*. 2020;12(4):3558–73.
34. Hu Q, Peng J, Chen X, et al. Obesity and genes related to lipid metabolism predict poor survival in oral squamous cell carcinoma. *Oral Oncol*. 2019;89:14–22.
35. Liu S, Shi L, Yang X, et al. Nuclear survivin promoted by acetylation is associated with the aggressive phenotype of oral squamous cell carcinoma. *Cell Cycle*. 2017;16(9):894–902.
36. Li M, Gao F, Yu X, et al. Promotion of ubiquitination-dependent survivin destruction contributes to xanthohumol-mediated tumor suppression and overcomes radioresistance in human oral squamous cell carcinoma. *J Exp Clin Cancer Res*. 2020;39(1):88.
37. Luo X, Donnelly CR, Gong W, et al. HPV16 drives cancer immune escape via NLRX1-mediated degradation of STING. *J Clin Invest*. 2020;130(4):1635–52.
38. Wu Y, Chen H, Chen Y, et al. HPV shapes tumor transcriptome by globally modifying the pool of RNA binding protein-binding motif. *Aging*. 2019;11(8):2430–46.
39. Uchida Y, Chiba T, Kurimoto R, Asahara H. Post-transcriptional regulation of inflammation by RNA-binding proteins via cis-elements of mRNAs. *J Biochem*. 2019;166(5):375–82.
40. Zhang Y, Yu C. Prognostic characterization of OAS1/OAS2/OAS3/OASL in breast cancer. *BMC Cancer*. 2020;20(1):575.
41. Li Y, Si L, Zhai Y, et al. Genome-wide association study identifies 8p21.3 associated with persistent hepatitis B virus infection among Chinese. *Nat Commun*. 2016;7:11664.
42. Gonzalez-Sanchez L, Cobos-Fernandez MA, Lopez-Nieva P, et al. Exploiting the passenger ACO1-deficiency arising from 9p21 deletions to kill T-cell lymphoblastic neoplasia cells [published online ahead of print, 2019 Nov 17]. *Carcinogenesis*. 2019. bgz185. doi:10.1093/carcin/bgz185.
43. Pio R, Zudaire I, Pino I, et al. Alpha CP-4, encoded by a putative tumor suppressor gene at 3p21, but not its alternative splice variant alpha CP-4a, is underexpressed in lung cancer. *Cancer Res*. 2004;64(12):4171–9.
44. Schudrowitz N, Takagi S, Wessel GM, Yajima M. Germline factor DDX4 functions in blood-derived cancer cell phenotypes. *Cancer Sci*. 2017;108(8):1612–9.
45. Ogorodnikov A, Levin M, Tattikota S, et al. Transcriptome 3'end organization by PCF11 links alternative polyadenylation to formation and neuronal differentiation of neuroblastoma. *Nat Commun*. 2018;9(1):5331.

46. Wang J, Wang Y, Kong F, et al. Identification of a six-gene prognostic signature for oral squamous cell carcinoma. *J Cell Physiol.* 2020;235(3):3056–68.
47. Tan YS, Sansanaphongpricha K, Xie Y, Donnelly CR, Luo X, Heath BR, Zhao X, Bellile E, Hu H, Chen H, et al. Mitigating SOX2-potentiated Immune Escape of Head and Neck Squamous Cell Carcinoma with a STING-inducing Nanosatellite Vaccine. *Clin Cancer Res.* 2018;24(17):4242–55.
48. Luckau S, Wehrs TP, Brandau S, Horn PA, Lindemann M. Vaccination Against Human Papilloma Viruses Leads to a Favorable Cytokine Profile of Specific T Cells. *J Immunother.* 2016;39(8):316–20.

Figures

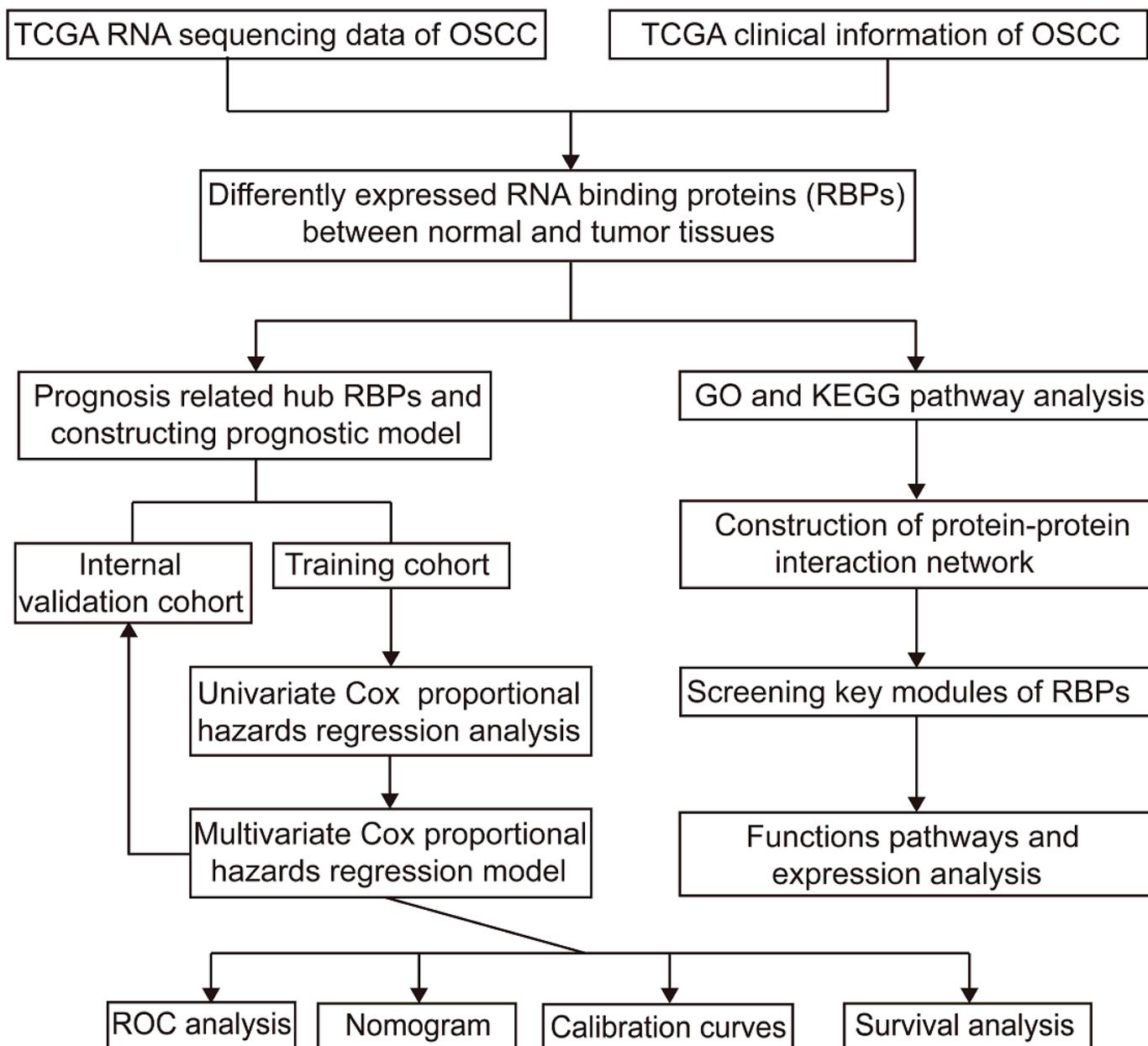


Figure 1

Study flowchart for analyzing RBPs in OSCC. OSCC, oral squamous cell carcinoma; RBPs, RNA binding proteins

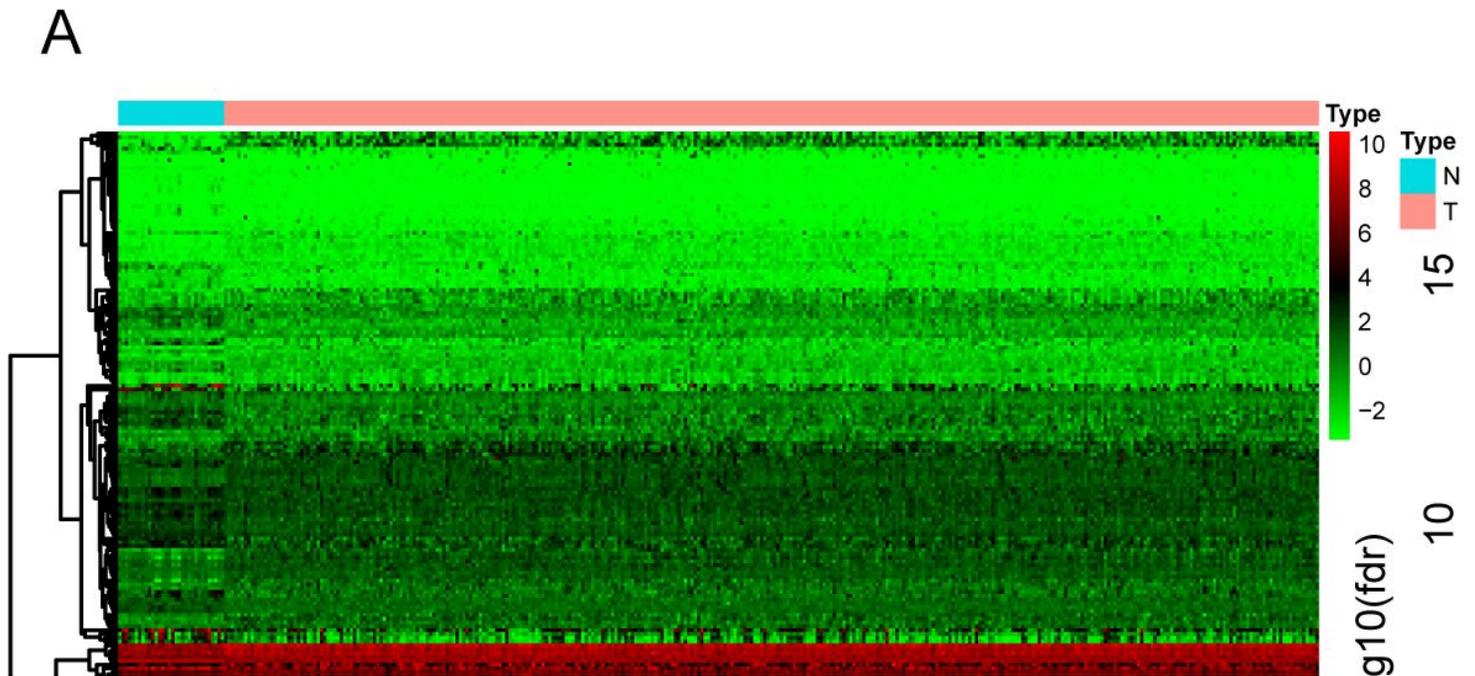


Figure 2

The differentially expressed RBPs in OSCC. (A) Hierarchical clustering of OSCC tissues and normal tissues by differentially expressed RBPs; The upper horizontal axis represents samples, and the left vertical axis represents clusters of RBPs. Red represents upregulated RBPs, and green represents downregulated RBPs. (B) Volcano plot of differentially expressed RBPs. The red dots represent upregulated RBPs, and the green dots represent downregulated RBPs. FC, Fold Change; fdr, false discovery rate.

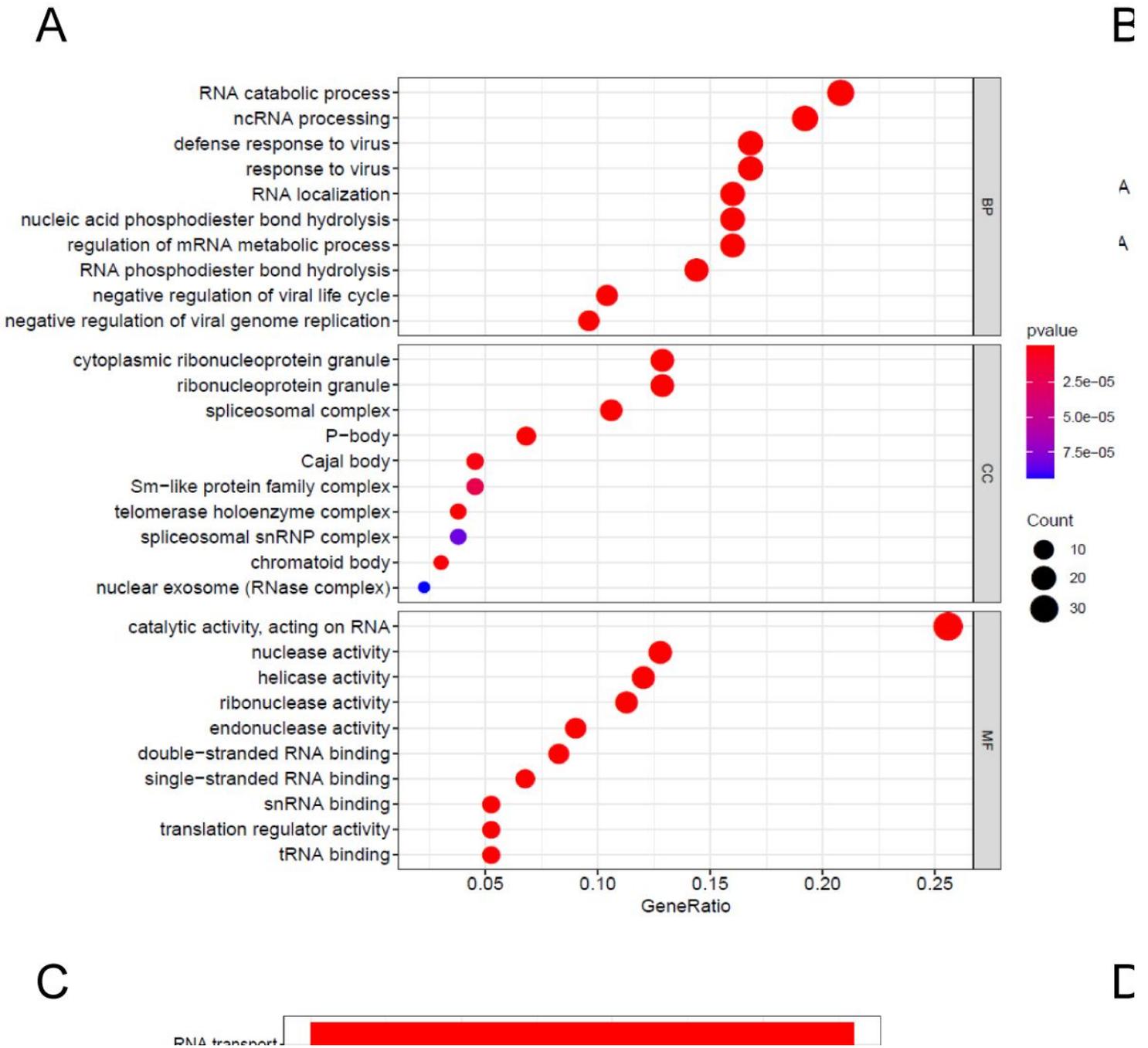


Figure 3

GO enrichment and KEGG pathway analysis of aberrantly expressed RBPs. Biological process, cellular components and molecular function enrichment for up-expressed RBPs (A) and down-expressed RBPs (B); KEGG pathway analysis for up-expressed RBPs (C) and down-expressed RBPs (D).

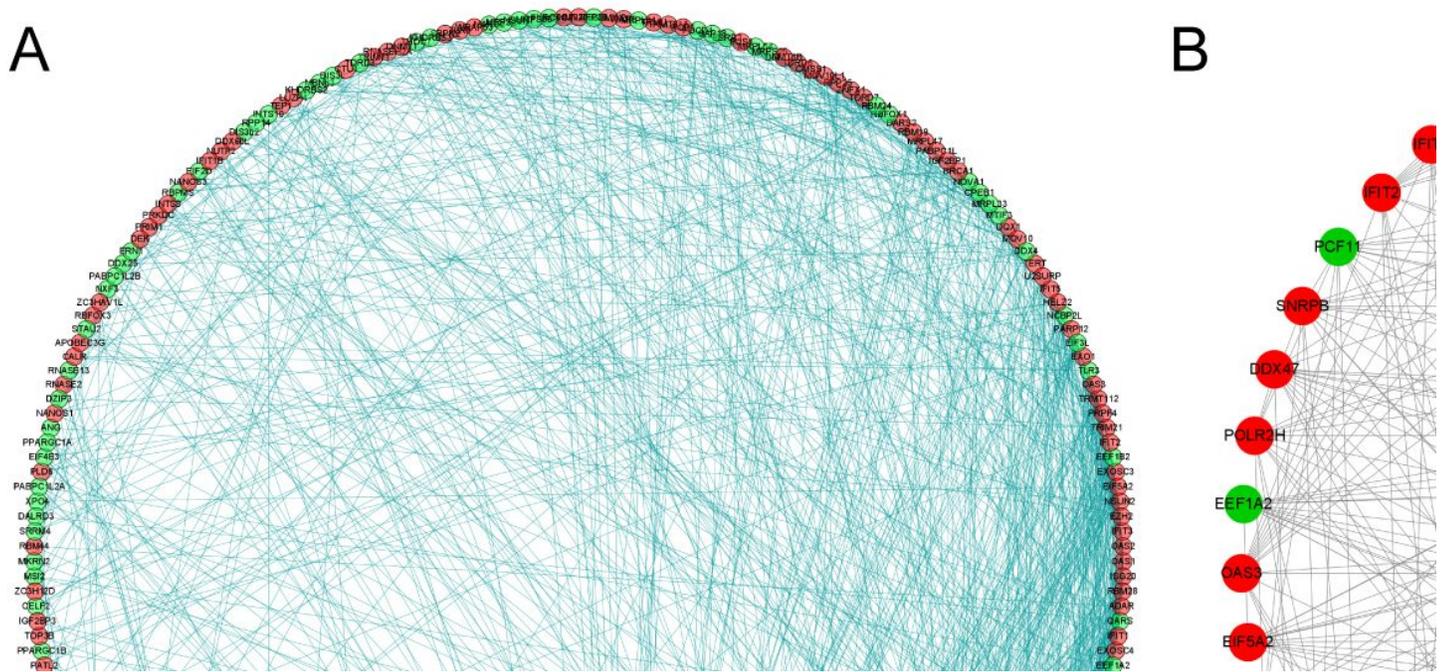


Figure 4

Protein-protein interaction network and modules analysis. (A) Protein-protein interaction network of differentially expressed RBPs; (B) Significant modules from PPI network. Green circles: down-expressed RBPs; red circles: up-expressed RBPs.

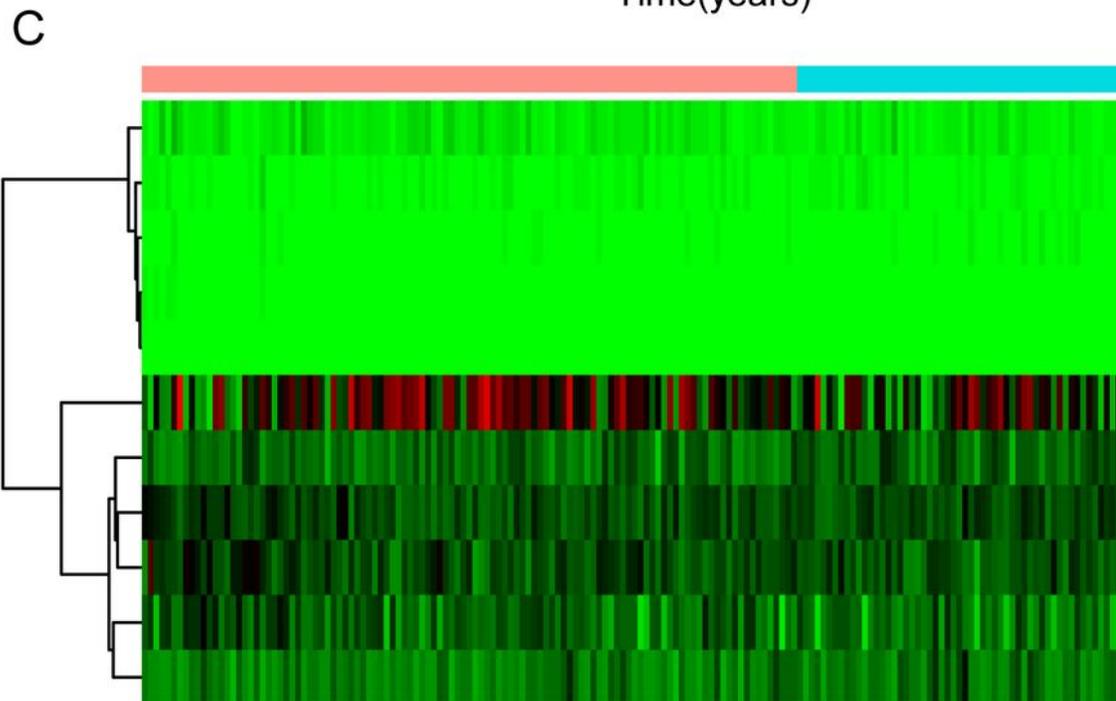
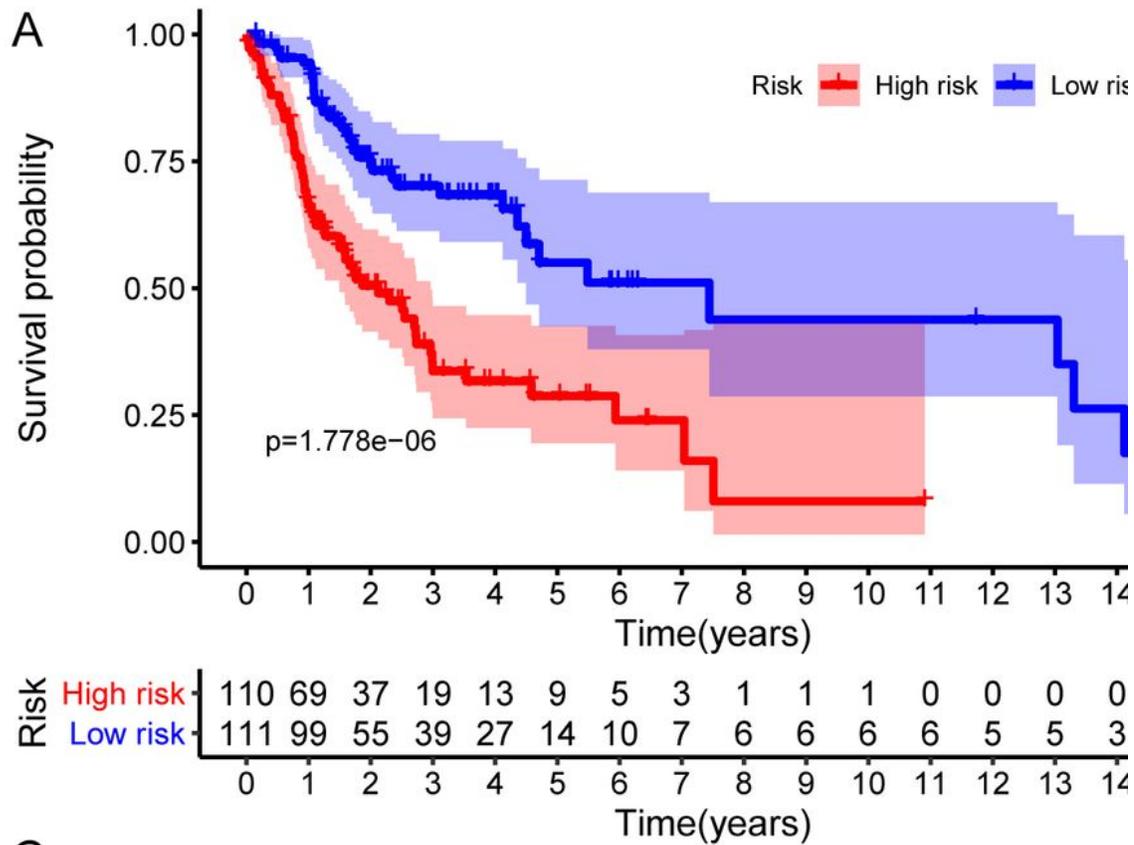


Figure 5

Risk score analysis of ten-genes prognostic model in the training cohort. (A) Survival curves for OS in low- and high-risk subgroups; (B) ROC curves for predicting OS based on risk score; (C) Expression heat map of ten prognostic RBPs in low- and high-risk subgroups; (D) Risk score distribution and survival status for low- and high-risk subgroups. OS, overall survival; ROC, receiver operating characteristic.

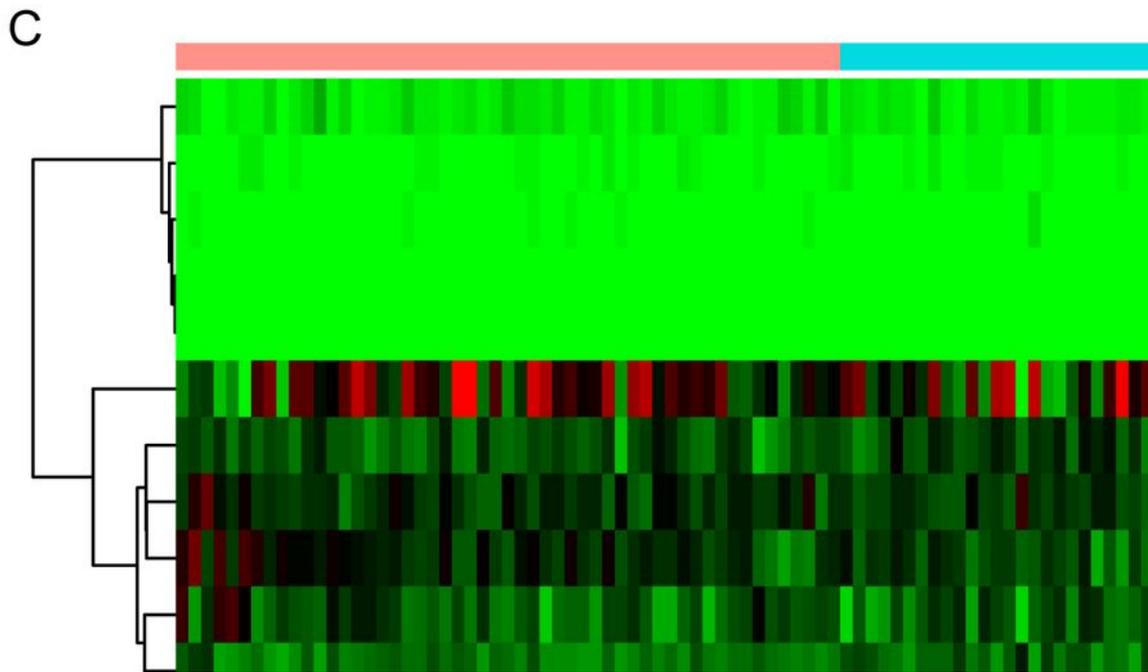
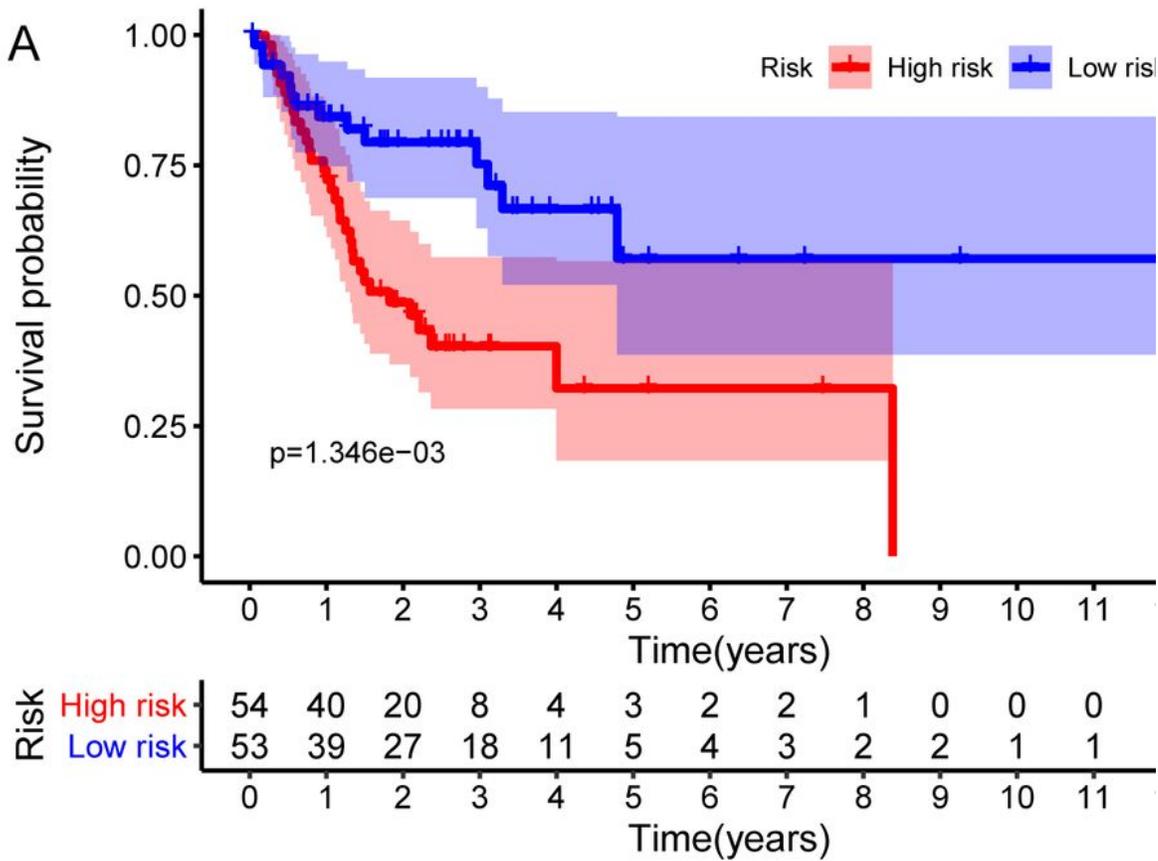


Figure 6

Risk score analysis of ten-genes prognostic model in the internal validation cohort. (A) Survival curves for OS in low- and high-risk subgroups; (B) ROC curves for predicting OS based on risk score; (C) Expression heat map of ten prognostic RBPs in low- and high-risk subgroups; (D) Risk score distribution and survival status for low- and high-risk subgroups. OS, overall survival; ROC, receiver operating characteristic.

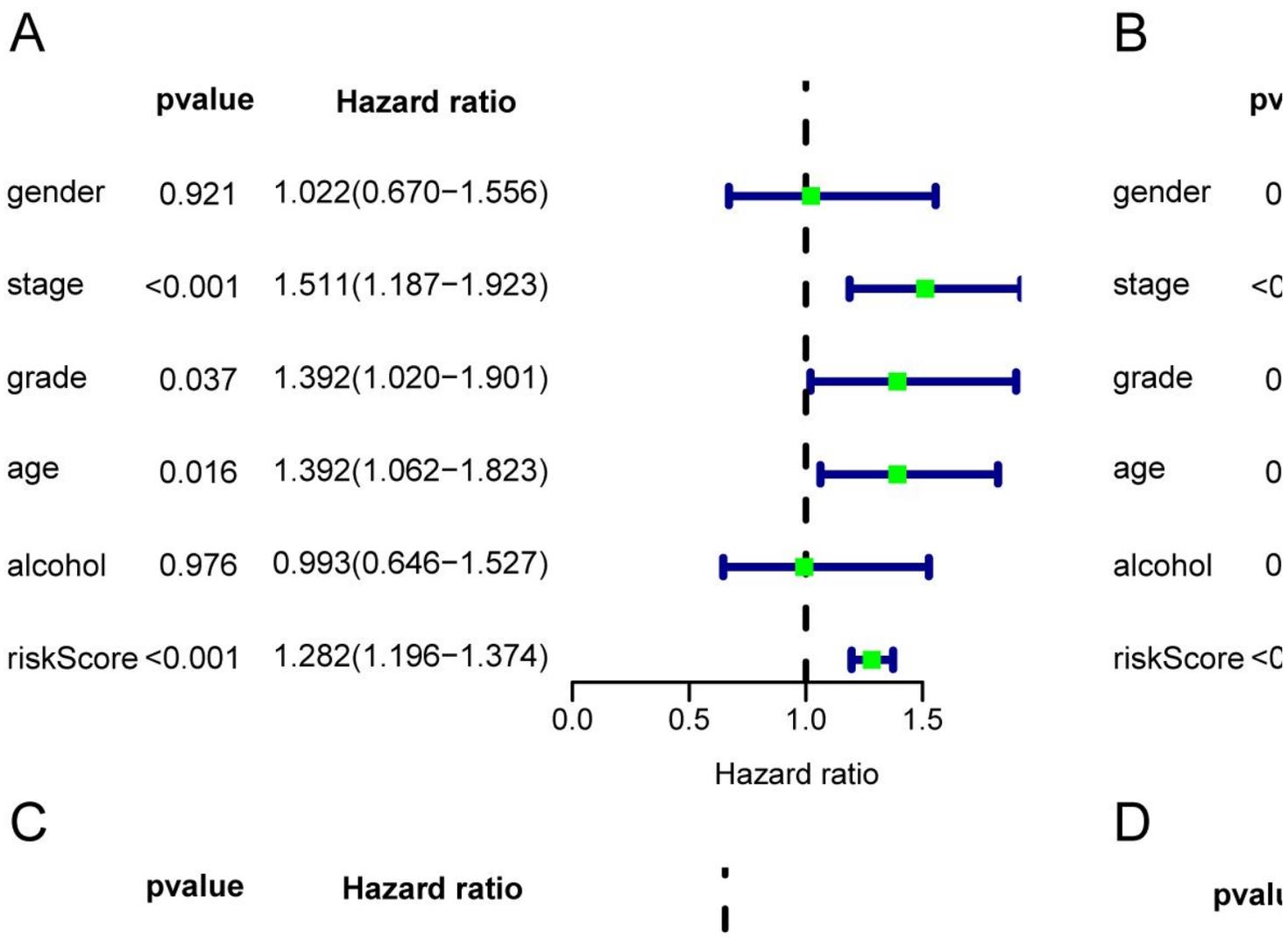


Figure 7

The prognostic value of different clinical parameters. Univariate Cox regression analysis and forest plots of the HR and 95% CIs in the training cohort (A) and internal validation cohort (C); Multivariate Cox regression analysis and forest plots of the HR and 95% CIs in the training cohort (B) and internal validation cohort (D);

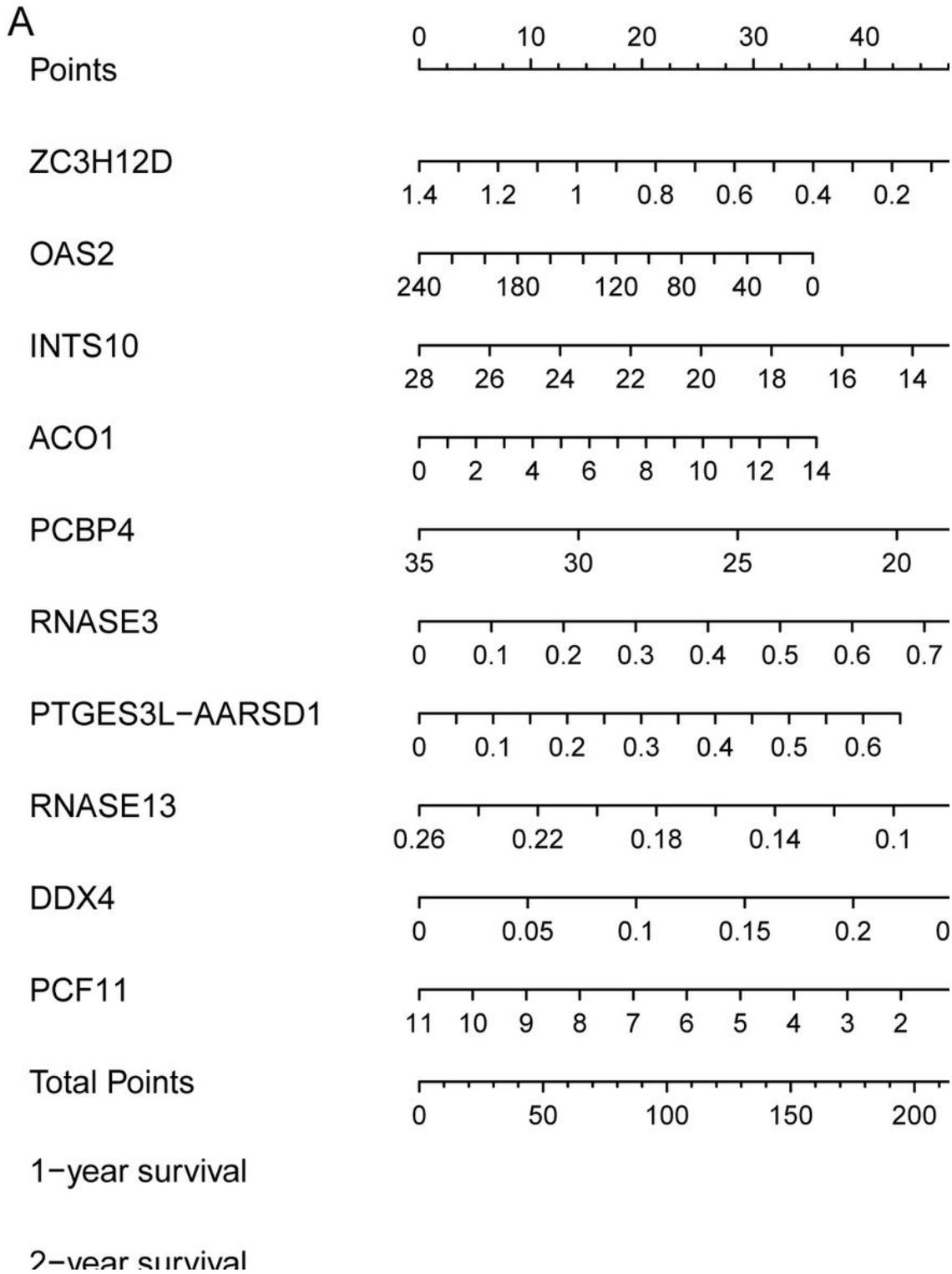


Figure 8

Prognostic nomogram based on hub RBPs and calibration plots. (A) Nomogram for predicting the 1-, 2- and 3-year OS in the training cohort. All the points identified on the top scale for each factor were added to generate a total score. The total points projected on the bottom scale were used to determine the probabilities of 1-, 2- and 3-year OS for each patient. Nomograms for predicting the 1-, 2- and 3-year OS probabilities of OSCC patients in the training cohort (B) and internal validation cohort (C).

Supplementary Files

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

- [FigureS1.tif](#)
- [SupplementaryTables.docx](#)
- [Supplementaryfigure1.docx](#)