Genome-Wide Profiling of a Prognostic RNA-Binding Protein Signature in Esophageal Cancer

Guoqing Zhang
First Affiliated Hospital of Zhengzhou University
Yan Zhang
First Affiliated Hospital of Zhengzhou University
Hang Yang
First Affiliated Hospital of Zhengzhou University
Yanan Guo
First Affiliated Hospital of Zhengzhou University
Jindong Li (13598820589@163.com)
First Affiliated Hospital of Zhengzhou University
Xiangnan Li
First Affiliated Hospital of Zhengzhou University

Article

Keywords: esophageal cancer, RNA-binding protein, prognosis, chemo-resistance, immune checkpoint inhibitor

Posted Date: March 15th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1439005/v1

Abstract

Background

the roles of RNA-binding proteins (RBPs) in esophageal cancer (EC) need elucidation.

Materials and methods

The gene expression profiles and clinical data of patients with EC were downloaded from the Xena database. Candidate genes were obtained by taking the union of RBP genes, KEGG pathway-related genes and differentially expressed RBP genes from cluster analysis. Hub genes were extracted by PPI network construction.

Result

We built a Cox proportional hazards regression model with 7 prognostic RBPs (TRMT2A, PDHA1, MPRIP, KRI1, IL17A, HSPA1A and HIST1H4J). The risk score of each patient in the internal and external dataset cohorts was calculated and then the patients were divided into two groups based on the median. There were significant differences between the RBP risk score and response to chemotherapy, with low-risk patients being more likely to achieve CR. Finally, the risk score was revealed to be significantly related to OS by univariate and multivariate analyses, and pathological stage could also be used independently to predict the prognosis of EC.

Conclusion

Our study indicated that the RBP signature could serve as a prognostic biomarker of EC and provide new insight into chemoresistance in EC patients.

Introduction

Esophageal cancer (EC), with an estimated 604,000 new cases (ranking seventh) and 544,000 deaths (ranking sixth) in 2020 worldwide, is a highly prevalent fatal cancer[1]. Esophageal adenocarcinoma (EAC) and esophageal squamous cell carcinoma (ESCC) are the 2 main types of EC. AC represents roughly two-thirds of EC cases in high-income countries, while ESCC has the leading incidence (> 90%) in certain high-risk areas in Asia (e.g., China). Despite screening and treatment progress, the mortality of EC remains high, with an average 5-year survival rate of less than 15%[2]. In recent years, even with numerous diagnostic molecular markers, clinicians have failed to achieve accurate early-stage detection of EC. Large amounts of gene expression data give clinicians an excellent opportunity to identify potential

tumor molecular markers, which may be of great clinical significance for early EC detection, diagnosis and prognosis.

RNA-binding proteins (RBPs) are a class of proteins that interact with different types of RNAs. At present, genome-wide screening in humans has identified over 1,500 RBPs[3]. Studies have shown that RBPs play a vital role in posttranscriptional regulation in life processes and can cause deleterious effects that could lead to numerous human disorders. Although RBPs are known to be involved in the initiation and development of malignant tumors, the roles of RBPs in EC development still need elucidation. No study has systematically evaluated RBP expression patterns, which may help us fully understand their roles in EC. Therefore, we downloaded EC RNA sequencing and clinical data from public databases such as The Cancer Genome Atlas (TCGA) database and the Xena database. Subsequently, we identified differentially expressed RBPs between EC and normal samples and explored their functional roles and potential clinical value. Furthermore, we established a prognostic model based on the hub RBPs, identifying them as clinically significant prognostic biomarkers of EC.

Results

3.1 Basic information

The gene expression profiles of 172 samples (161 tumors and 11 adjacent normal tissues) from EC patients were obtained from the Xena database. RBP-related genes were downloaded from the EuRBPDB database and related articles[4], and a total of 4,528 RBP-related genes were obtained. A total of 337 hub genes were obtained through DEG analysis and WGCNA. Among these genes, prognostic RBPs were identified using univariate Cox regression analysis. Subsequently, we built a LASSO-penalized Cox proportional hazards regression model with multiple prognostic RBPs. It was validated in the internal dataset and external dataset GSE72873 (n = 44). Finally, pan-cancer analysis was performed, and correlations between the RBP risk score and immunotherapy biomarkers, including TMB, mismatch repair defects, and neoantigens, were identified. The detailed workflow is shown in Fig. 1.

3.2 Identification of differentially expressed RBP genes in EC patients

In this study, we obtained 1,898 DEGs from the Xena database (Fig. 2A-B; Supplementary Tables 1). Then, we performed KEGG pathway enrichment analysis of the differentially expressed RBPs, and 62 pathways with 1,331 RBP-related genes were identified (cd-RBP-geneset1) (Fig. 2C; Supplementary Tables 2 and 3).

We found that 1,879 RBP genes can reflect the prognosis of patients with EC. Fifty-four DEGs were obtained by the log-rank test and Cox proportional hazards regression ($p < 0.05$). Hence, we performed a consensus unsupervised analysis of all samples based on these 54 RBP genes. We determined the optimal number of clusters using the consistent cumulative distribution function graph and the delta area plot (Fig. 2D-E). The final number of clusters was $k = 2$. Therefore, two clusters of patients were identified

as follows: C1 (n = 98, 40.33%) and C2 (n = 64, 26.34%). Subsequently, Kaplan–Meier analysis was used to evaluate the relationship between the clusters and prognosis. The results indicated that the C2 cluster was associated with better prognosis in terms of OS (Fig. 2f). Furthermore, 635 DEGs were obtained from cluster analysis (cd-RBP-geneset2)(Supplementary Tables 4). PCA was effectively able to distinguish the above subtypes (Fig. 2G-H).

3.3 Construction of the prognostic RBP signature

To further identify the association between RBPs and EC patients with different clinical characteristics, the expression data profiles of RBP genes, cd-RBP-geneset1 and cd-RBP-geneset2 (Fig. 3A) were transformed into a gene coexpression network using the WGCNA package in R (Fig. 3B-C). A total of 10 coexpressed modules were obtained through a one-step network construction method (Fig. 3D, E, F). Furthermore, we performed a correlation analysis between different coexpression modules and OS/OS time (Fig. 3D-E). The results indicate that the turquoise module has the most significant correlation with OS (correlation coefficient = 0.19, p = 0.0015). However, there were no modules significantly related to OS time. The distribution of the modules' average gene significance related to OS and OS time are shown in Fig. 3D and Fig. 3E, respectively. The turquoise module was selected as the most clinically significant module of OS for PPI network construction to extract hub genes through Cytoscape plugins. As a result, 337 hub genes were identified (Fig. 3G) (Supplementary Table 5).

Among these 337 hub genes, 24 prognostic RBPs (6 newly predicted) were identified using univariate Cox regression analysis (Fig. 4). Subsequently, we built a LASSO-penalized Cox proportional hazards regression model with 7 prognostic RBPs (Fig. 5A, B, C).

Risk score = TRMT2A*(-0.9916) + PDHA1*(0.9129) + MPRIP*(-1.121) + KRI1*(-0.6753) + IL17A*(1.282) + HSPA1A*(0.2014) + HIST1H4J*1.256.

3.4 Validation of the risk model

The risk model was further validated in the internal dataset and external dataset GSE72873 (n = 44).

The risk score of each patient in the internal and external dataset cohorts was calculated and then the patients were divided into two groups based on the median. In the internal dataset cohort, 80 patients each were categorized into the high-risk and low-risk groups. Similarly, in the external dataset cohort, 22 patients each were divided into the high-risk and low-risk groups. There were significant differences in survival curves between the two risk groups in the internal dataset and external dataset cohorts ($P < 0.05$) (Fig. 6C, Fig. 7C). The area under the curve (AUC) values at 1, 2 and 3 years were 0.789, 0.821 and 0.797, respectively, in the internal dataset (Fig. 6E-F). The AUC value at 3 years was 0.701 in the external dataset (Fig. 7E). The risk score distribution plot, survival status plot, and heatmap of risk gene expression are presented in Fig. 6A, B, D and Fig. 7A, B, D.

We then evaluated the model stability in different clinical subgroups, and subgroups such as age, sex, grade, stage, and T classification showed significant differences in OS between the high-risk and low-risk

groups ($p < 0.05$) (Fig. 8). In the pan-cancer analyses, except for TCGA BRCA, all other cancer types (TCGA ACC, TCGA COAD, TCGA KIRC and TCGA LGG) demonstrated significant differences in OS between the high-risk and low-risk groups (Fig. 9). We further compared our proposed model with two existing RBP-related models, prognostic model 1[5] and prognostic model 1[6], and our model achieved a higher AUC value (Supplementary Fig. 1).

3.5 Correlations between the RBP risk score and clinical features and biomarkers of ICI response

We explored the differences in RBP risk scores between different clinical features and the results show that there are significant differences between the RBP risk score and clinical features such as different sexes, N0 and N1 classification, N0 and N2 classification, stage I and stage III, stage I and stage IV, stage II and stage IV, and M0 and M1 classification (Fig. 10). We further explored the relationship between the RBP risk score and biomarkers of ICI response (such as TMB, HRD, NAL, stemness index, LOH, LST, and TAI) in cancer, and the results showed that there was a remarkable relationship between mRNAsi and the RBP risk score, while the others had no significant correlations (Fig. 11).

We then explored the differences in biomarkers for predicting ICI response (immune cell infiltration, immune score, stromal score, tumor purity, somatic mutation and CNV) between the high-risk and low-risk groups. A difference analysis of the content of various immune cells in the high- and low-risk groups revealed high levels of follicular helper T cell (Tfh cell) and resting dendritic cell infiltration in the low-risk group, while there were no significant differences in other immune cell infiltration levels (Fig. 12A). The ESTIMATE analysis showed that there were no significant differences regarding the immune score, stromal score or tumor purity (Fig. 12B, C, D). Similarly, there were no differences regarding the somatic mutation type and frequency (Fig. 12E, F). High-risk patients had significantly higher CNVs than their low-risk counterparts ($p < 0.05$) (Fig. 12G, H).

3.6 Evaluation of the RBP risk score for the prediction of response to treatment and prognosis

In terms of immunotherapy (Fig. 13A, B, C), there was no significant difference in OS between the high- and low-risk groups ($p = 0.64$). The differences in RBP risk scores between patients with different responses to treatment were not significantly different ($p > 0.05$). Additionally, the proportions of complete response (CR)/partial response (PR) and stable disease (SD)/progressive disease (PD) in the high- and low-risk groups were 25.5% and 74.5% and 20.13% and 79.87%, respectively, which were not significantly different. In terms of chemotherapy (Fig. 13D, E, F), there were significant differences between RBP risk scores and responses to chemotherapy, with low-risk patients being more inclined to achieve CR. The AUC value for the prediction of chemotherapy response was 0.796. The proportions of CR, PR, SD and PD in the high- and low-risk groups were 33.33%, 10.26%, 0% and 56.41% and 25.86%, 1.72%, 1.72% and 70.69%, respectively. We further analyzed the status of chemoresistance by the RBP risk score, and the results showed that high-risk patients tended to be resistant to cisplatin and vinblastine, while low-risk patients tended to be resistant to doxorubicin, mitomycin C and sorafenib (Fig. 13G, H, I, J, K).

3.7 Independent prognostic value of the RBP signature and clinical parameters

The risk score was found to be significantly related to OS by univariate and multivariate analyses ($p < 0.05$) (Fig. 14A, B). The results also demonstrated that pathological stage and the risk model could be used independently to predict the prognosis of EC. Furthermore, we found that the nomogram was more accurate than the risk score and pathological stage in predicting OS at 3 years. The AUCs at 3 years for the nomogram, the risk score, and pathological stage were 0.837, 0.76 and 0.762, respectively (Fig. 14C, D, E, F, G).

Discussion

EC is the most common malignant tumor of the digestive tract and has a poor prognosis. Studies have established several novel biomarkers to determine the prognosis of EC patients, such as miRNA signatures[7,8], autophagy-related signatures[9], m6A RNA methylation regulator-based signatures[10], integrated mRNA-lncRNA signatures[11], epigenetic signatures[12], lymph node metastasis-associated gene signatures[13], urinary metabolomic signatures[14], and immune-related gene signatures[15]. Although RBP-related signatures have been analyzed for a variety of cancers[16–21], related analyses for EC are rare. Hence, there is a need for a prognostic analysis of RBP-related genes in

EC patients to explore their functional roles and potential clinical significance. Therefore, in this study, we used an integrated analysis of various databases to construct a Cox proportional hazards regression model with 7 prognostic RBPs (TRMT2A, PDHA1, MPRIP, KRI1, IL17A, HSPA1A and HIST1H4J) that could be used to predict prognosis and chemoresistance in EC patients.

Notably, PDHA1 and IL17A are newly predicted RBP genes, and not all 7 prognostic RBP-related genes have been reported to be associated with prognosis in cancer. For TRMT2A, the only study has reported that TRMT2A protein expression is a biomarker of increased risk of recurrence in HER2 + breast cancer patients and may be used to predict the response to adjuvant cytotoxic chemotherapy[22]. PDHA1 has been reported to be biologically significant in several human tumors. However, for EC, only one study demonstrated that inhibition of PDHA1 gene expression in human ESCC leads to increased malignancies[23]. However, as shown in our prediction model, TRMT2A was found to be a gene that improves prognosis, while PDHA1 makes the prognosis worse, which needs further validation in EC. IL17A is a member of the IL17 cytokine family and is released by both immune and nonimmune cells (such as tumor and stromal cells) into the tumor microenvironment. Among all the IL17 cytokine family members, IL17A is the most controversial in regulating tumor immunity and has different prognostic values among various cancers. Generally, tumor-infiltrating IL-17A-producing cells (IL17A + cells) are correlated with elevated antitumor immunity[24]. A study demonstrated that IL17A deficiency reduces tumor latency and promotes metastasis in lung cancer[25], and a study showed that IL17A mRNA expression could be used as a predictive biomarker for superior response to adjuvant chemotherapy and

can indicate better patient survival in gastric cancer[26]. In pancreatic cancer, an anti-IL17A antibody can enhance the antitumor response to gemcitabine[27], which implies that IL17A may be involved in the development of drug resistance. IL17A polymorphisms were associated with the risk of various cancers, and the IL17A rs4711998 A > G polymorphism was associated with a decreased risk of EC[28]. Our study showed that IL17A can indicate poor patient survival in EC. HSPA1A was associated with unfavorable survival and poor clinicopathological features in several kinds of tumors[29–31]. Furthermore, HSPA1A was demonstrated to mediate breast cancer radioresistance[32]. However, no study has shown the effects of HSPA1A on malignant biological properties, treatment sensitivity and prognosis in EC. Importantly, MPRIP, KRI1 and HIST1H4J have not been reported to be associated with prognosis in cancer and are our newly discovered genes that may be associated with EC prognosis.

In this study, we established a seven-gene biomarker as a novel prognostic model and analyzed its ability to predict prognosis in different cohorts. Although the number of adjacent normal tissues from TCGA was relatively small (n = 11), it did not affect the reliability of the results. The prognostic performance of the model was validated by internal and external validation. Importantly, the RBP signature is also a good independent indicator of survival with adjusted clinical parameters, including age, sex, T classification, N classification, M classification, tumor grade, and stage.

We further analyzed the predictive effect of the RBP signature for chemotherapy and immunotherapy response. Interestingly, our signature can predict the efficacy of chemotherapy, while the efficacy of immunotherapy was not well predicted. Many of the chemotherapy agents used in the treatment of cancer interfere with the production of nucleic acids, RNA and DNA. Thus, chemotherapy may interfere with the binding of RBPs to RNAs. In return, the expression of RBPs may influence chemotherapy, leading to drug sensitivity or resistance. The chemotherapy agents identified (cisplatin, doxorubicin, mitomycin C, sorafenib, and vinblastine) in our study can provide new ideas and serve as a basis for future clinical drug research. However, as shown in our study, there were no significant differences in most of the biomarkers for predicting ICI responses and immune cells in the immune microenvironment. This is probably the main reason for the unpredictability of immunotherapy. Different responses to treatment and different mechanisms of resistance stem from different antitumor mechanisms. This might suggest that different signatures are needed to predict drug sensitivity/resistance and prognosis. Various immune gene signatures have been identified to predict therapeutic effects in various kinds of tumors[33–36]. Wang et al identified a prognostic immune gene signature in EC; however, this signature did not predict immunotherapy response[15]. Therefore, further studies are needed to predict the therapeutic effect of immunotherapy in EC.

In conclusion, we established a prognostic signature with 7 prognostic RBPs as biomarkers for EC patients. Our study could also contribute to providing new insight into chemoresistance in EC patients.

Materials And Methods

1. Data download

The gene expression profiles and clinical data of patients with EC were downloaded from the Xena database (https://xenabrowser.net/). A total of 161 tumors and 11 adjacent normal tissues were obtained for further analyses. RBP-related genes were downloaded from the EuRBPDB database (http://EuRBPDB.syshospital.org) and related articles[4], and a total of 4,528 RBP-related genes were obtained.

2. Identification of hub genes

Initially, we obtained differentially expressed RBP-related genes. To identify differentially expressed genes (DEGs) between EC and adjacent normal tissues, differences in expression were determined using the limma R package with a threshold of p value < 0.05. Furthermore, differentially expressed RBP-related genes were determined by intersecting DEGs with RBPs. These RBP-related genes selected as initial candidates were used to establish the prognostic RBP signature in the next step.

Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis of the differentially expressed RBPs was performed using the clusterProfiler R package to identify the KEGG pathways these genes were enriched in. A p value < 0.05 represented statistical significance. Then, KEGG pathway-related genes were identified (cd-RBP-geneset1). Cluster analysis was performed via the ConsensusClusterPlus package in R (reps = 100, pItem = 0.8, pFeature = 1, clusterAlg="km", distance="euclidean"). Survival analysis was used to compare the overall survival (OS) between categorical data. Differences in the expression of the results from cluster analysis were determined using the limma R package, with a p value < 0.05 representing statistical significance. The differentially expressed RBP genes were identified as cd-RBP-geneset2. Principal component analysis (PCA) was performed to facilitate data visualization.

Candidate genes were obtained by taking the union of RBP genes, cd-RBP-geneset1 and cd-RBP-geneset2. We used the WGCNA package in R software to build a gene coexpression network, and the minimum number of module genes was set as 50 to ensure the reliability of the results. Correlations between coexpression modules and clinical characteristics were analyzed, and the clinically most significant module was selected for protein-protein interaction (PPI) network construction to extract hub genes through Cytoscape 3.8 plugins.

3. Construction and validation of the prognostic RBP signature

We further investigated the hub genes selected for constructing the prognostic RBP-related signature. Prognostic RBPs were identified using univariate Cox regression analysis. Subsequently, we built a least absolute shrinkage and selection operator (LASSO) regression model using the glmnet package to remove the redundancy factor via 10-fold cross-validation.

Then, the model was validated in terms of prediction effect evaluation by Kaplan–Meier survival analysis and receiver operator characteristic (ROC) curves using survival and survivalROC packages in R software.

We then evaluated the model stability in different clinical subgroups by Kaplan–Meier survival analysis using a clinical gene dataset from TCGA. For external validation, the GSE72873 dataset was selected and analyzed in a similar way. We also conducted pan-cancer analysis using a clinical gene dataset from TCGA and compared our proposed model with existing RBP-related models.

4. Correlation analysis of clinical features and biomarkers of ICI response

We explored the differences in RBP risk scores between different clinical features by independent t tests. Additionally, numerous biomarkers for predicting immune checkpoint inhibitor (ICI) response have been explored in recent years, including tumor mutation burden (TMB), homologous recombination deficiency (HRD), neoantigen load (NAL), stemness index, loss of heterozygosity (LOH), large-scale state transition (LST), telomeric allelic imbalance (TAI), immune cell infiltration, immune score, stromal score, tumor purity, somatic mutation and copy number variation (CNV). We explored the relationship between RBP risk scores and biomarkers and the differences in biomarkers between the high-risk and low-risk groups. TMB was calculated with maftools package. HRD was downloaded from the Xena database (https://xenabrowser.net/). NAL was downloaded from the TSNAdb database (TSNAdb (zju.edu.cn)). The stemness index was obtained from a previous study[37]. The Pearson correlation coefficient test was used to estimate the relationship between the risk score and biomarkers. Immune cell infiltration was evaluated by the CIBERSORT algorithm. The immune score, stromal score and tumor purity were analyzed by the estimate package in R software. Somatic mutations were analyzed by the maftools package in R software. A CNV map was drawn according to the segment of copy number in EC, and CNVs were compared between the high-risk and low-risk groups. Differences between the two groups were assessed with independent t tests.

5. Prediction of response to treatment and prognosis

We evaluated the clinical significance of the RBP risk score in predicting the response to ICI treatment in the IMvigor210CoreBiologies dataset, which contained 348 clinical samples, with 298 containing immunophenotypes. We also evaluated the predictive effect of the RBP risk score for predicting responses to chemotherapy in the TCGA database. pRRophetic was employed in the sensitivity prediction of chemotherapeutic drugs (cisplatin, doxorubicin, mitomycin. C, sorafenib, and vinblastine).

6. Independent prognostic value of the prognostic RBP signature

To assess the prognostic value of the prognostic RBP signature, we applied both univariate and multivariate analyses of prognostic factors using Cox proportional hazards models. Age and risk score were treated as continuous variables. Sex, T stage, M stage, N stage, pathological stage and differentiation were treated as categorical variables. Factors with $p < 0.05$ in both univariate and multivariate analyses were identified as independent prognostic variables. Finally, we constructed a nomogram that could predict the OS probability of patients with EC.

Abbreviations

AUC,area under the curve;CNV,copy number variation; CR, complete response; DEGs,differentially expressed genes;EAC,Esophageal adenocarcinoma;EC,esophageal cancer;ESCC,esophageal squamous cell carcinoma;HRD, homologous recombination deficiency; ICI, immune checkpoint inhibitor;KEGG, Kyoto Encyclopedia of Genes and Genomes;LASSO,least absolute shrinkage and selection operator;LOH, loss of heterozygosity; LST, large-scale state transition; NAL,neoantigen load;OS, overall survival;PCA, Principal component analysis; PD, progressive disease;PPI,protein-protein interaction; PR, partial response; RBPs, RNA-binding proteins;SD, stable disease;TAI, telomeric allelic imbalance; TCGA, The Cancer Genome Atlas;Tfh cell, follicular helper T cell; TMB, tumor mutation burden;

Declarations

Author Contributions

ZGdesigned and supervised all experiments. ZG and GY carried out all experiments. YH and GY executed the data analysis and drafted the manuscript. ZG, LJ, and LX revised the manuscript. All authors reviewed the final manuscript.

Acknowledgments

The data that support the findings of this study are available on request from the corresponding author.All methods were carried out in accordance with the relevant guidelines and regulations.

Conflicts of Interest

none

Funding

This study was supported by the Key Scientific Research Projects of Institutions of Higher Learning in Henan Province (No. 21A320032).

References

1. Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians.* Feb 4 2021.
2. Xia W, Liu S, Mao Q, et al. Effect of lymph node examined count on accurate staging and survival of resected esophageal cancer. *Thoracic cancer.* May 2019;10(5):1149-1157.

3. Gerstberger S, Hafner M, Tuschl T. A census of human RNA-binding proteins. *Nature reviews. Genetics.* Dec 2014;15(12):829-845.
4. Wang Z, Tang W, Yuan J, Qiang B, Han W, Peng X. Integrated Analysis of RNA-Binding Proteins in Glioma. *Cancers.* Apr 7 2020;12(4).
5. Liu Y, Liu X, Gu Y, Lu H. A novel RNA binding protein-associated prognostic model to predict overall survival in hepatocellular carcinoma patients. *Medicine.* Jul 23 2021;100(29):e26491.
6. Man Z, Chen Y, Gao L, et al. A Prognostic Model Based on RNA Binding Protein Predicts Clinical Outcomes in Hepatocellular Carcinoma Patients. *Frontiers in oncology.* 2020;10:613102.
7. Zhao Y, Xu L, Wang X, Niu S, Chen H, Li C. A novel prognostic mRNA/miRNA signature for esophageal cancer and its immune landscape in cancer progression. *Molecular oncology.* Apr 2021;15(4):1088-1109.
8. Zhang HC, Tang KF. Clinical value of integrated-signature miRNAs in esophageal cancer. *Cancer medicine.* Aug 2017;6(8):1893-1903.
9. Du H, Xie S, Guo W, et al. Development and validation of an autophagy-related prognostic signature in esophageal cancer. *Annals of translational medicine.* Feb 2021;9(4):317.
10. Xu LC, Pan JX, Pan HD. Construction and Validation of an m6A RNA Methylation Regulators-Based Prognostic Signature for Esophageal Cancer. *Cancer management and research.* 2020;12:5385-5394.
11. Lan T, Xiao Z, Luo H, et al. Bioinformatics analysis of esophageal cancer unveils an integrated mRNA-lncRNA signature for predicting prognosis. *Oncology letters.* Feb 2020;19(2):1434-1442.
12. Jin YQ, Miao DL. Multiomic Analysis of Methylation and Transcriptome Reveals a Novel Signature in Esophageal Cancer. *Dose-response : a publication of International Hormesis Society.* Jul-Sep 2020;18(3):1559325820942075.
13. Cai W, Li Y, Huang B, Hu C. Esophageal cancer lymph node metastasis-associated gene signature optimizes overall survival prediction of esophageal cancer. *Journal of cellular biochemistry.* Jan 2019;120(1):592-600.
14. Davis VW, Schiller DE, Eurich D, Sawyer MB. Urinary metabolomic signature of esophageal cancer and Barrett's esophagus. *World journal of surgical oncology.* Dec 15 2012;10:271.
15. Wang L, Wei Q, Zhang M, et al. Identification of the prognostic value of immune gene signature and infiltrating immune cells for esophageal cancer patients. *International immunopharmacology.* Oct 2020;87:106795.
16. Zhu M, Cai J, Wu Y, Wu X, Feng L, Yin Z. A Novel RNA-Binding Protein-Based Nomogram for Predicting Survival of Patients with Gastric Cancer. *Medical science monitor : international medical journal of experimental and clinical research.* Jan 20 2021;27:e928195.
17. Yang L, Zhang R, Guo G, et al. Development and validation of a prediction model for lung adenocarcinoma based on RNA-binding protein. *Annals of translational medicine.* Mar 2021;9(6):474.

18. Xie Y, Luo X, He H, Pan T, He Y. Identification of an individualized RNA binding protein-based prognostic signature for diffuse large B-cell lymphoma. *Cancer medicine.* Apr 2021;10(8):2703-2713.
19. Wang M, Jiang F, Wei K, et al. Development and Validation of a RNA Binding Protein-Associated Prognostic Model for Hepatocellular Carcinoma. *Technology in cancer research & treatment.* Jan-Dec 2021;20:15330338211004936.
20. Tian J, Ma C, Yang L, Sun Y, Zhang Y. Prognostic Value and Immunological Characteristics of a Novel RNA Binding Protein Signature in Cutaneous Melanoma. *Frontiers in genetics.* 2021;12:723796.
21. Sheng C, Chen Z, Lei J, Zhu J, Song S. Development and Multi-Data Set Verification of an RNA Binding Protein Signature for Prognosis Prediction in Glioma. *Frontiers in medicine.* 2021;8:637803.
22. Hicks DG, Janarthanan BR, Vardarajan R, et al. The expression of TRMT2A, a novel cell cycle regulated protein, identifies a subset of breast cancer patients with HER2 over-expression that are at an increased risk of recurrence. *BMC cancer.* Mar 22 2010;10:108.
23. Liu L, Cao J, Zhao J, Li X, Suo Z, Li H. PDHA1 Gene Knockout In Human Esophageal Squamous Cancer Cells Resulted In Greater Warburg Effect And Aggressive Features In Vitro And In Vivo. *OncoTargets and therapy.* 2019;12:9899-9913.
24. Wang Z, Zhou Q, Zeng H, et al. Tumor-infiltrating IL-17A(+) cells determine favorable prognosis and adjuvant chemotherapeutic response in muscle-invasive bladder cancer. *Oncoimmunology.* 2020;9(1):1747332.
25. You R, DeMayo FJ, Liu J, et al. IL17A Regulates Tumor Latency and Metastasis in Lung Adeno and Squamous SQ.2b and AD.1 Cancer. *Cancer immunology research.* Jun 2018;6(6):645-657.
26. Wang JT, Li H, Zhang H, et al. Intratumoral IL17-producing cells infiltration correlate with antitumor immune contexture and improved response to adjuvant chemotherapy in gastric cancer. *Annals of oncology : official journal of the European Society for Medical Oncology.* Feb 1 2019;30(2):266-273.
27. Roux C, Mucciolo G, Kopecka J, Novelli F, Riganti C, Cappello P. IL17A Depletion Affects the Metabolism of Macrophages Treated with Gemcitabine. *Antioxidants (Basel, Switzerland).* Mar 10 2021;10(3).
28. Yin J, Wang L, Shi Y, et al. Interleukin 17A rs4711998 A>G polymorphism was associated with a decreased risk of esophageal cancer in a Chinese population. *Diseases of the esophagus : official journal of the International Society for Diseases of the Esophagus.* Jan 2014;27(1):87-92.
29. Oliveira D, Prahm KP, Christensen IJ, Hansen A, Høgdall CK, Høgdall EV. Gene expression profile association with poor prognosis in epithelial ovarian cancer patients. *Scientific reports.* Mar 8 2021;11(1):5438.
30. Guan Y, Zhu X, Liang J, Wei M, Huang S, Pan X. Upregulation of HSPA1A/HSPA1B/HSPA7 and Downregulation of HSPA9 Were Related to Poor Survival in Colon Cancer. *Frontiers in oncology.* 2021;11:749673.
31. Yang Z, Zhuang L, Szatmary P, et al. Upregulation of heat shock proteins (HSPA12A, HSP90B1, HSPA4, HSPA5 and HSPA6) in tumour tissues is associated with poor outcomes from HBV-related

early-stage hepatocellular carcinoma. *International journal of medical sciences.* 2015;12(3):256-263.

32. Zhang S, Wang B, Xiao H, et al. LncRNA HOTAIR enhances breast cancer radioresistance through facilitating HSPA1A expression via sequestering miR-449b-5p. *Thoracic cancer.* Jul 2020;11(7):1801-1816.
33. Hwang S, Kwon AY, Jeong JY, et al. Immune gene signatures for predicting durable clinical benefit of anti-PD-1 immunotherapy in patients with non-small cell lung cancer. *Scientific reports.* Jan 20 2020;10(1):643.
34. Ling B, Ye G, Zhao Q, Jiang Y, Liang L, Tang Q. Identification of an Immunologic Signature of Lung Adenocarcinomas Based on Genome-Wide Immune Expression Profiles. *Frontiers in molecular biosciences.* 2020;7:603701.
35. Zhang C, Zhang Z, Zhang G, et al. Clinical significance and inflammatory landscapes of a novel recurrence-associated immune signature in early-stage lung adenocarcinoma. *Cancer letters.* Jun 1 2020;479:31-41.
36. Zhuang Y, Li S, Liu C, Li G. Identification of an Individualized Immune-Related Prognostic Risk Score in Lung Squamous Cell Cancer. *Frontiers in oncology.* 2021;11:546455.
37. Malta TM, Sokolov A, Gentles AJ, et al. Machine Learning Identifies Stemness Features Associated with Oncogenic Dedifferentiation. *Cell.* Apr 5 2018;173(2):338-354 e315.

Figures

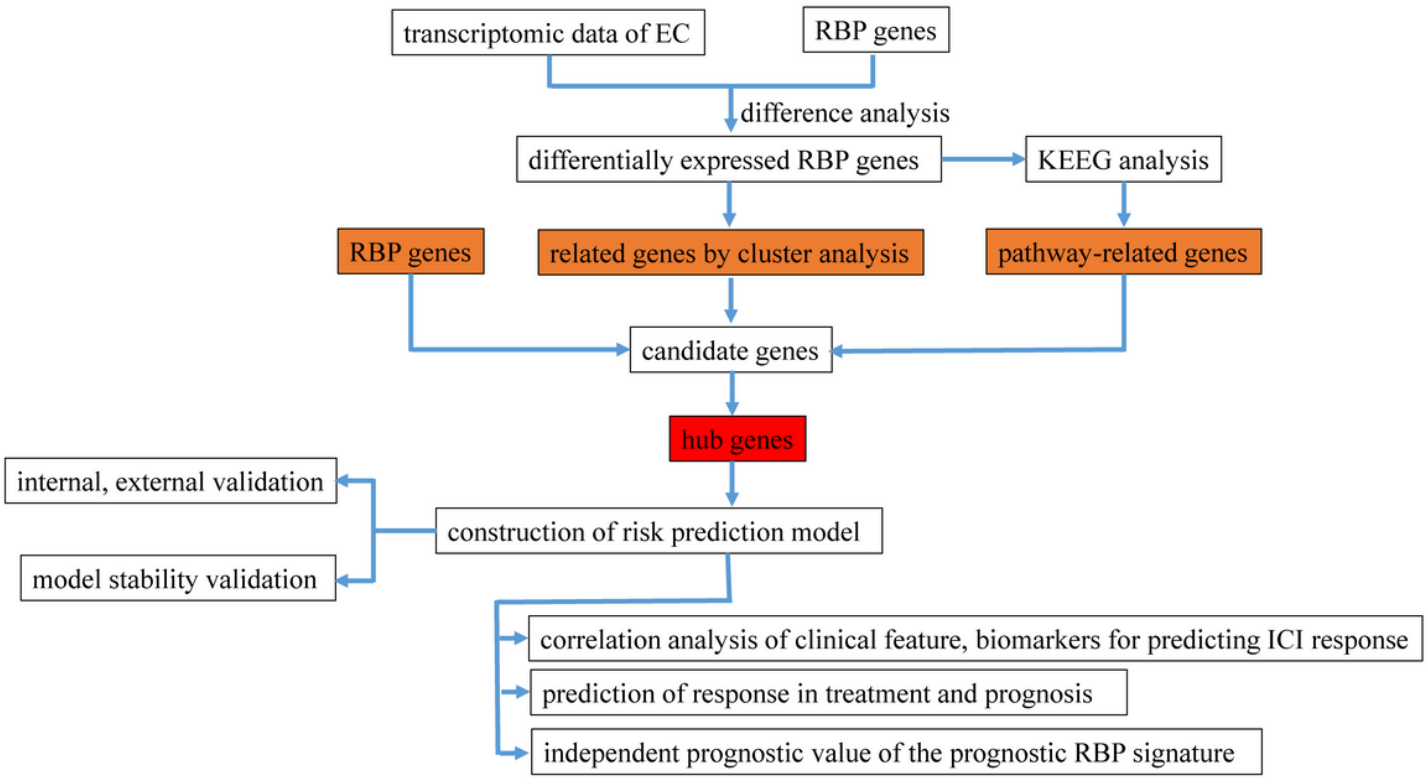

Figure 1

Diagram of the study. EC, esophageal cancer; ICIs, immune checkpoint inhibitors; KEGG, Kyoto Encyclopedia of Genes and Genomes; RBP, RNA-binding protein.

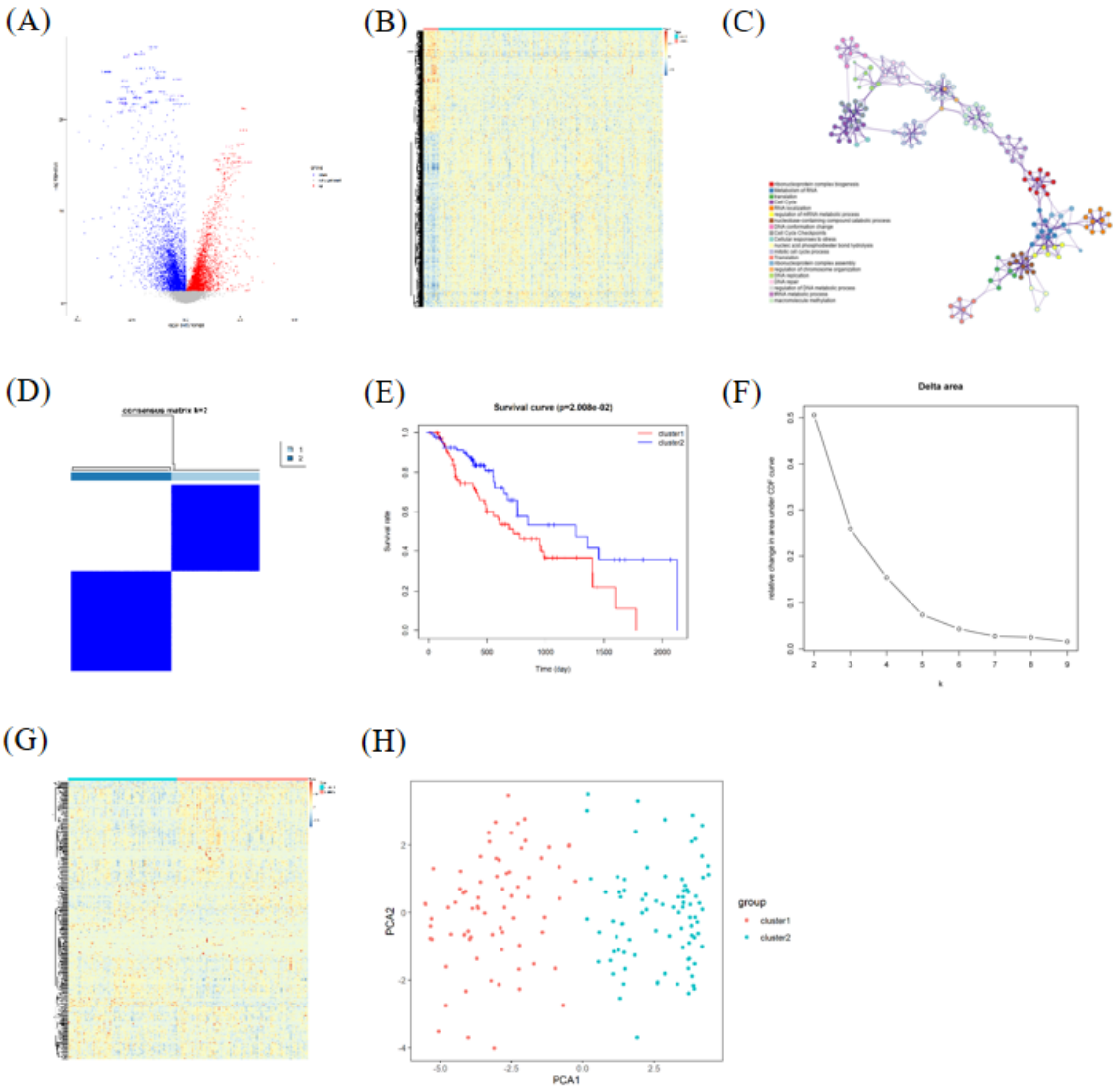

Figure 2

Identification of differentially expressed (DE) RBP-related genes. (A). Volcano plot of the DE genes. (B). Heatmap of the DE genes. (C). KEGG pathway enrichment analysis of DE RBPs. (D). Consensus clustering analysis identification of two clusters. (E) Cumulative distribution function (CDF) for k = 2–9. (F) Kaplan–Meier (K–M) curves for EC patients stratified by cluster. (G) heatmap of theprognostic RBP genes ordered by clusters. (H) Principal component analysis of different clusters.

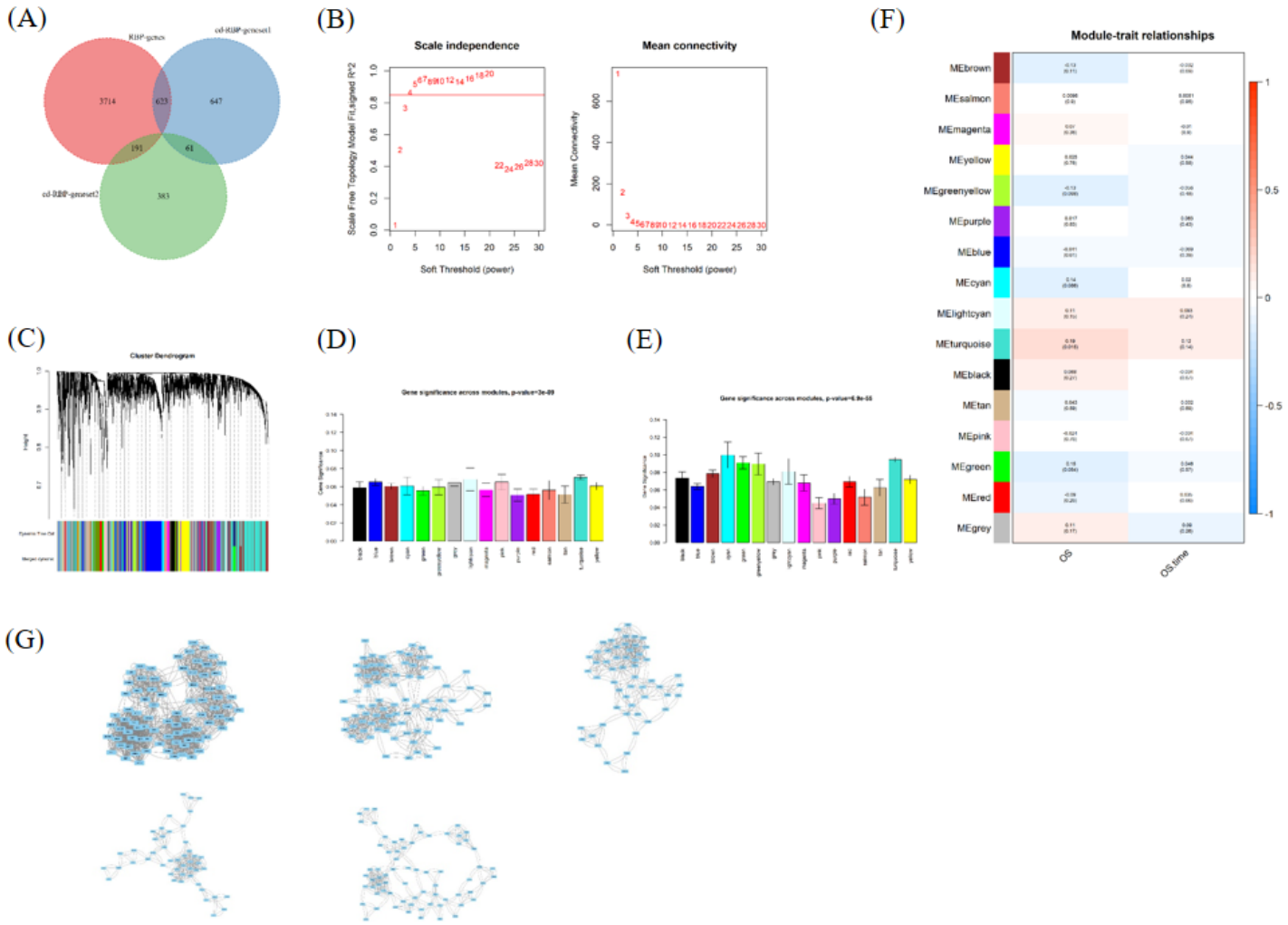

Figure 3

Construction of the prognostic RBP signature. (A)Venn diagram showing the RBP genes, cd-RBP-geneset1 and cd-RBP-geneset2. (B) The effect of different power values on the scale independence and average connectivity degree of coexpression modules of RBP-related genes. (C) The constructed coexpression modules of RBP-related genes by WGCNA software. (D)The distribution of the average gene significance related to OS in different modules is shown. (E)The distribution of the average gene significance related to OS time in different modules is shown. (F)The heatmap visualizes the correlation between the modules and the patient's clinical characteristics. (G) PPI network construction.

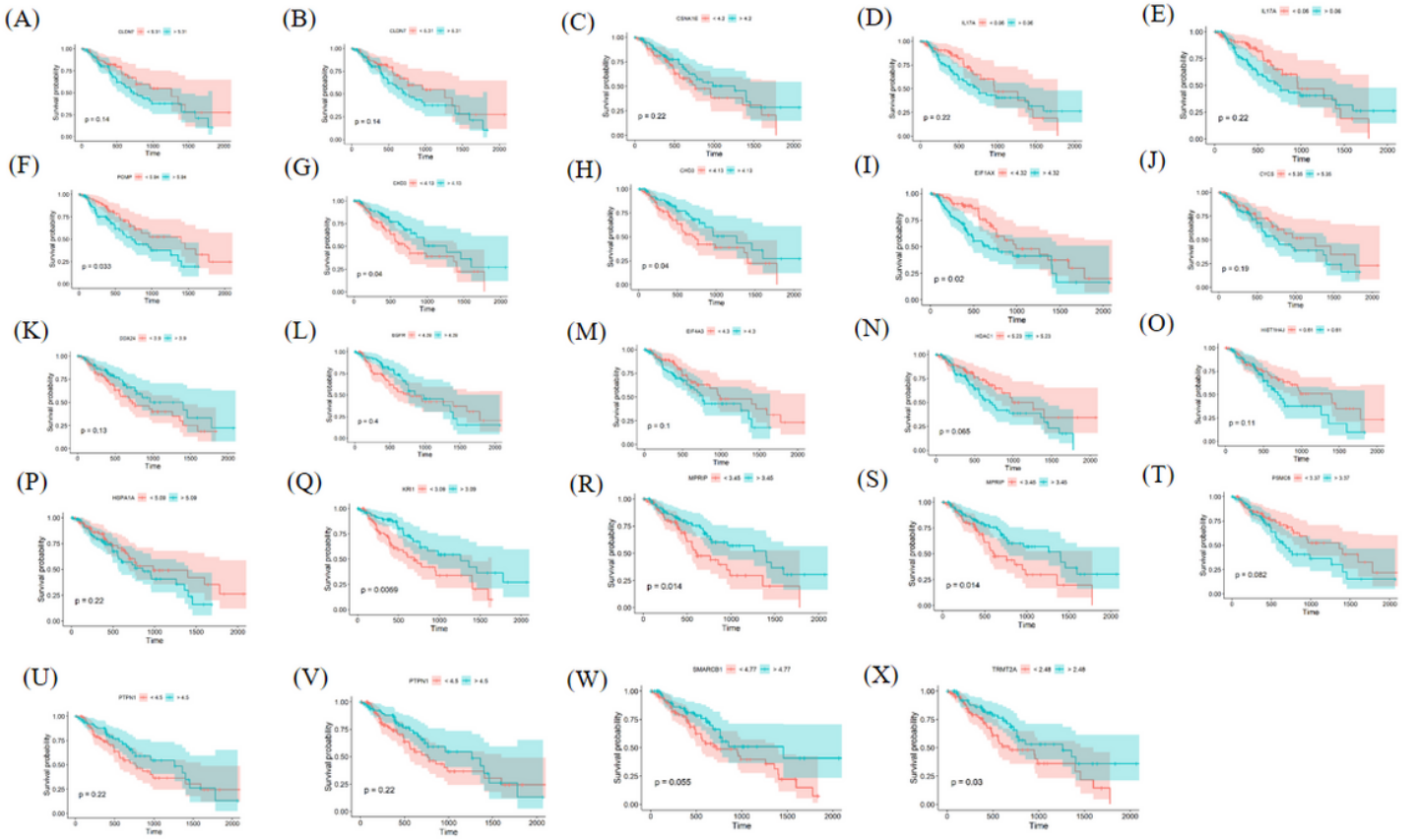

Figure 4

Prognostic value of hub genes in EC patients (Kaplan–Meier plot).

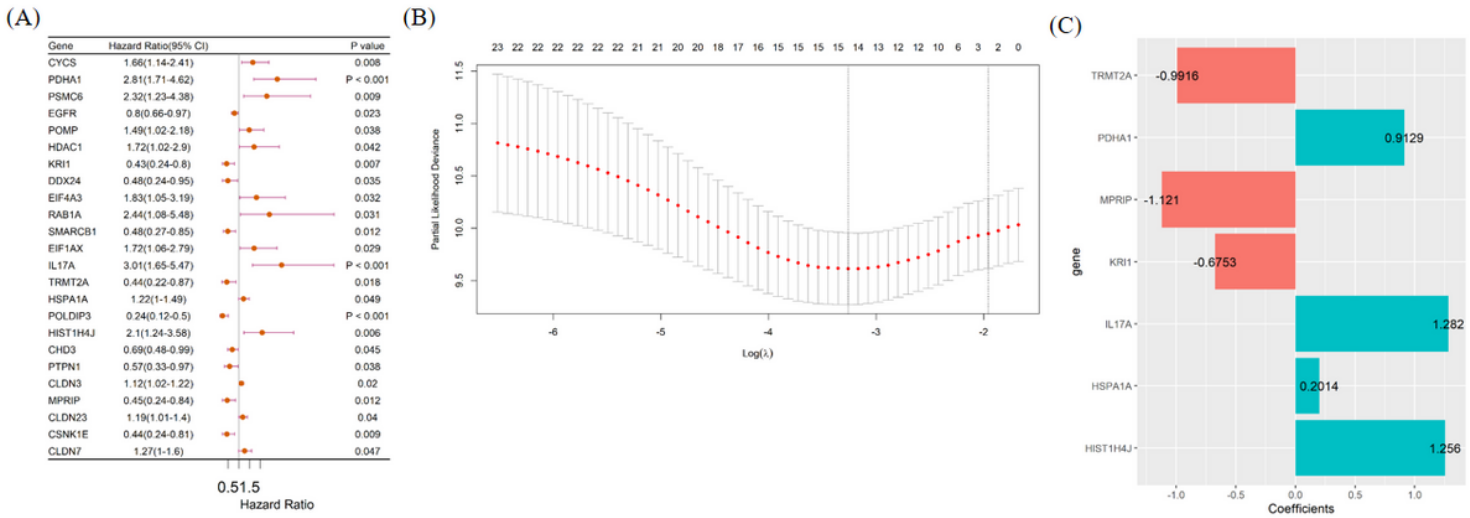

Figure 5

LASSO Cox regression analysis of RBP-related genes (A)Cox regression analysis.(B) Selection of the tuning parameter (λ) in theLASSO model via tenfold cross-validation based on minimum criteria.(C)

Identification of the prognostic risk score system with 7 prognostic RBPs.

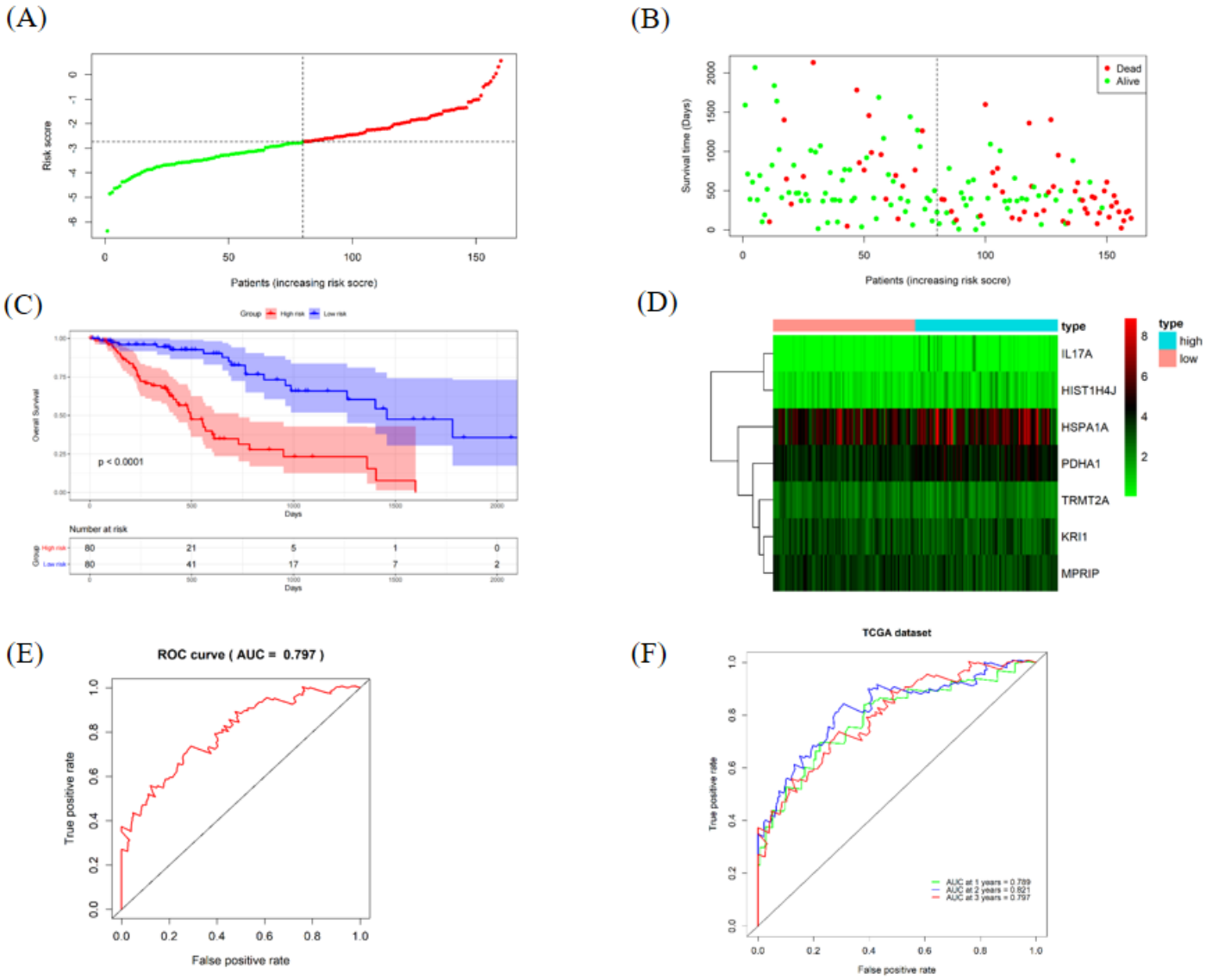

Figure 6

Internal signature validation. (A) The patients were divided into two groups: low-risk and high-risk groups. (B) As the risk scoreincreased, the survival time of patients decreased, and the number of deaths increased. (C) Kaplan–Meier analysis of ECpatients stratified by the median risk score.(D) The heatmap shows the expression profiles of the 7 RBP-related genes in the prognostic signature.(E) (F)The signature was evaluated by using the sensitivityand specificity of the ROC curve.

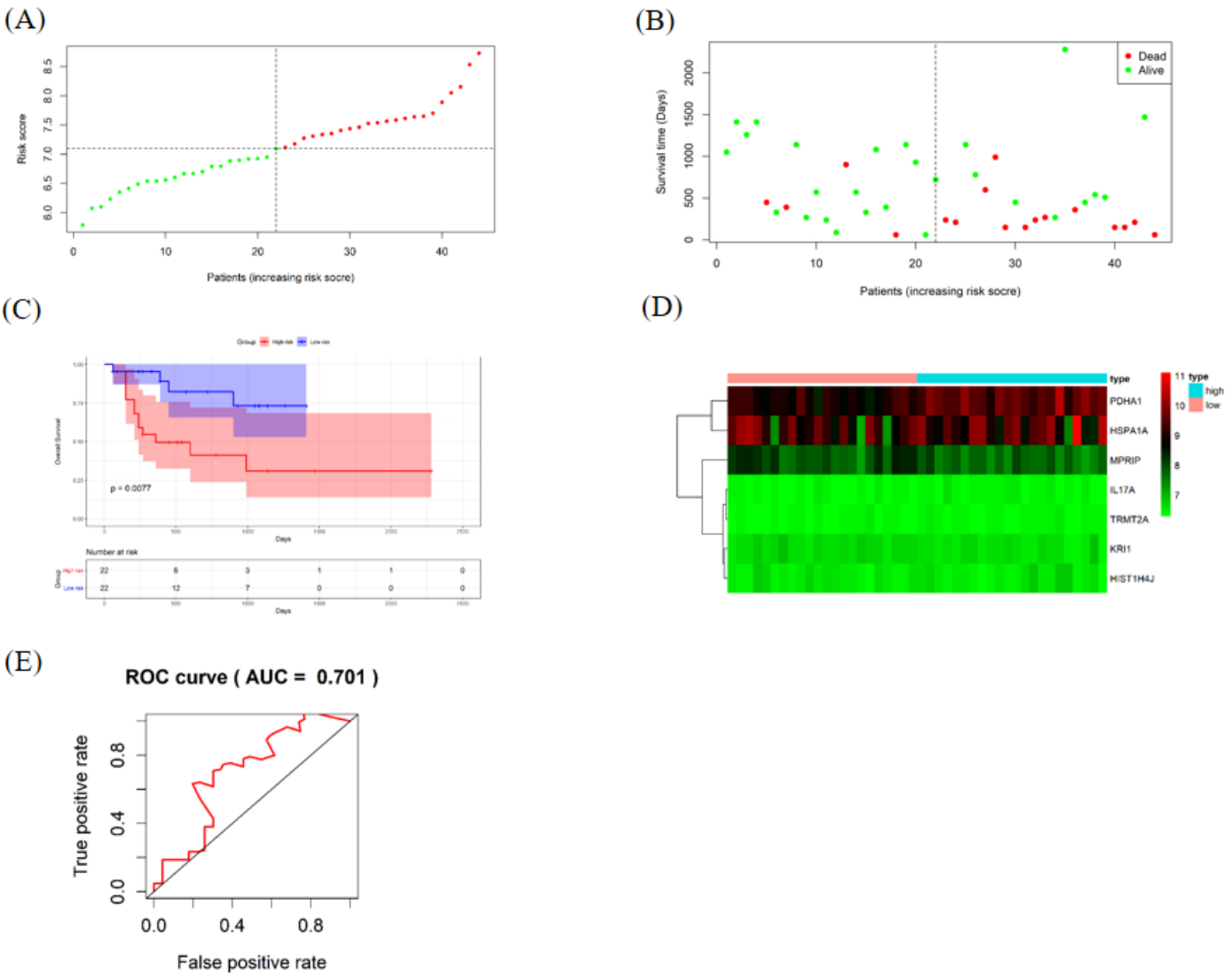

Figure 7

External signature validation. (A) The patients were divided into two groups: low-risk and high-risk groups. (B) As the risk scoreincreased, the survival time of patients decreased, and the number of deaths increased. (C) Kaplan–Meier analysis of ECpatients stratified by the median risk score.(D) The heatmap shows the expression profiles of the 7 RBP-related genes in the prognostic signature.(E)The signature was evaluated by using the sensitivityand specificity of the ROC curve.

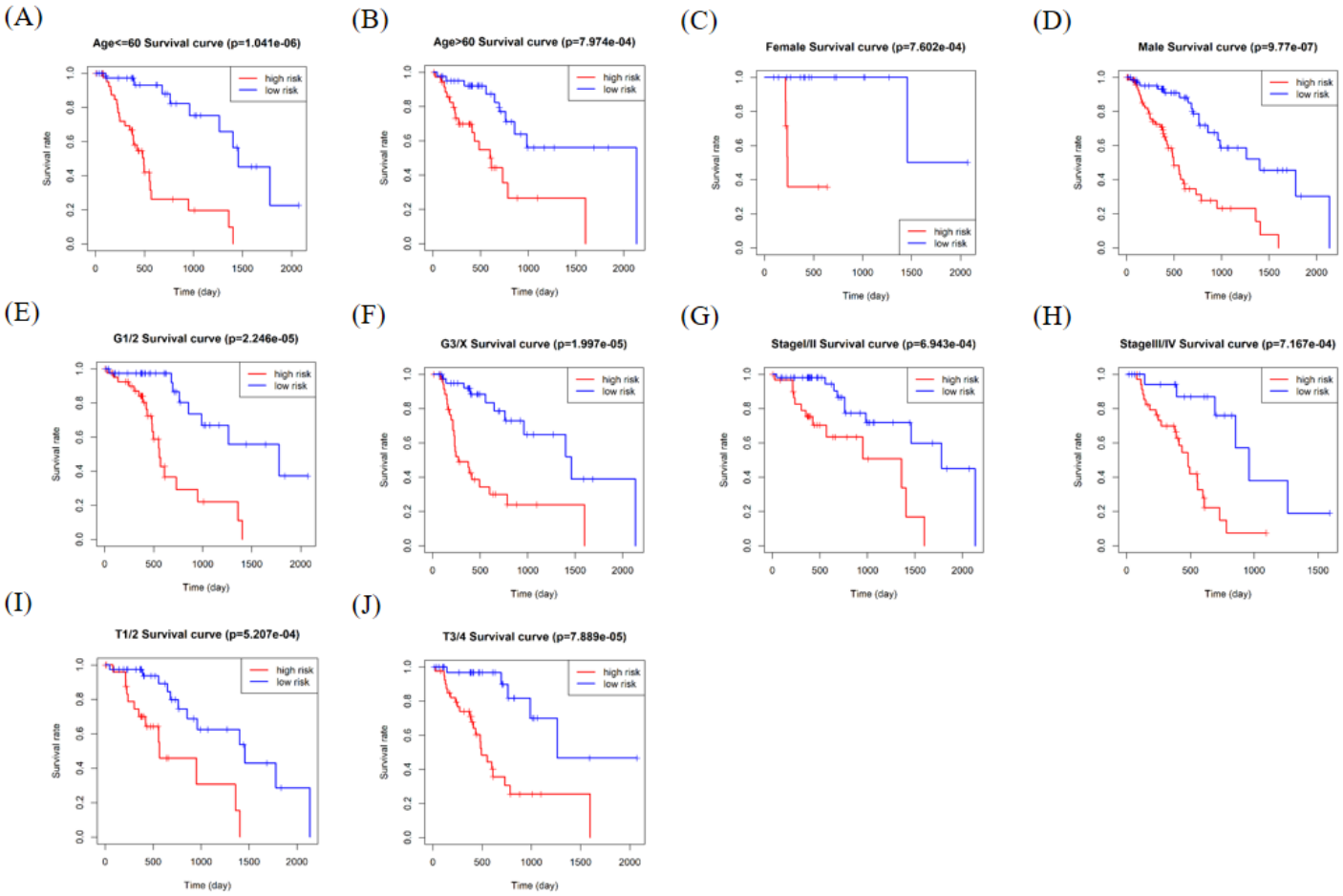

Figure 8

Verification of the prognostic value in different clinical subgroups. (A) age≤60; (B) age>60; (C) female; (D) male; (E) G1/2; (F) G3/X; (G) stage I/II; (H) stage III/IV; (I) T1/2; (J) T3/4.

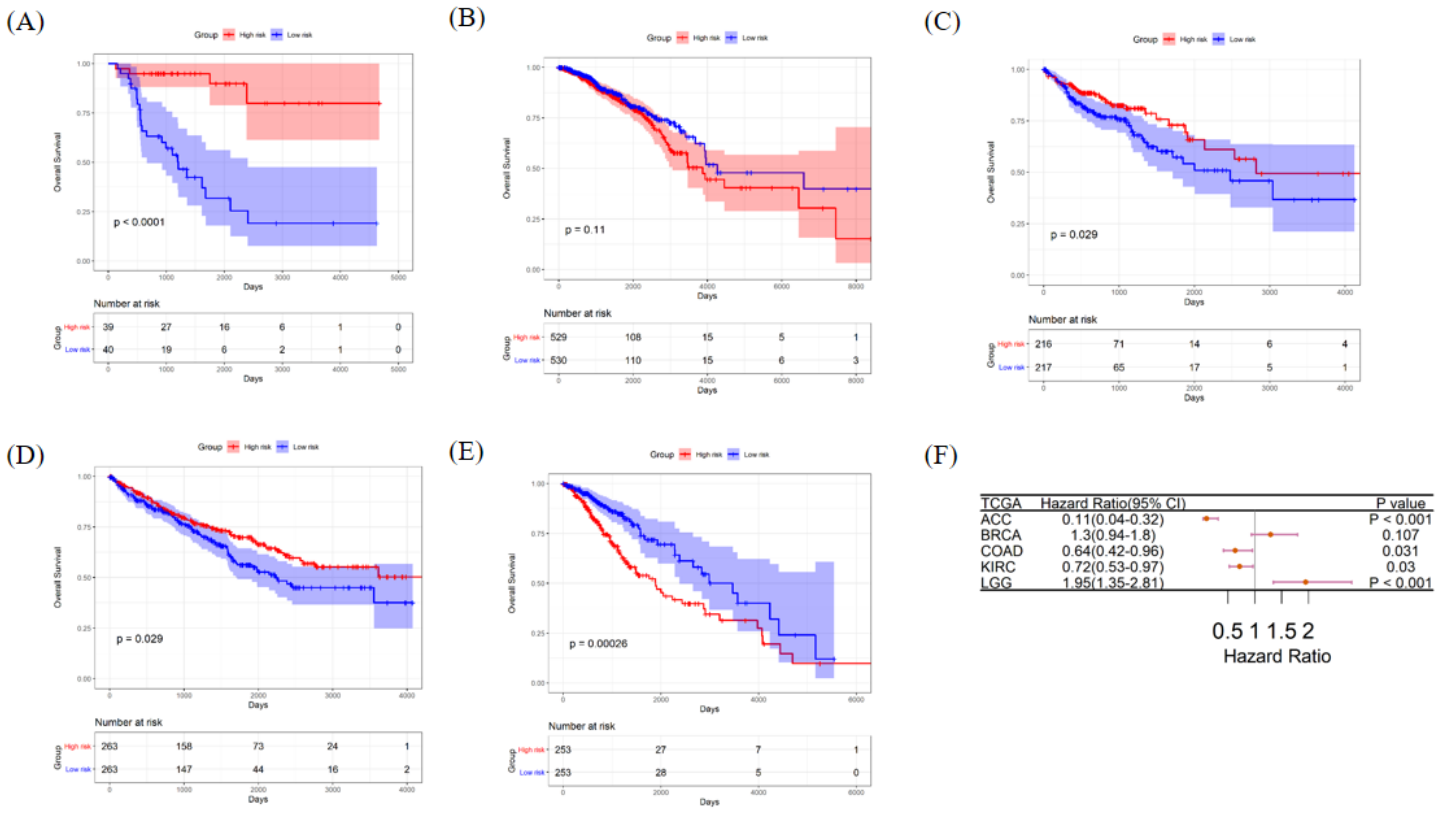

Figure 9

Pan-cancer analyses of OS between the high-risk and low-risk groups. (A) ACC; (B) BRCA; (C) COAD; (D) KIRC; (E) LGG; (F) forest plot.

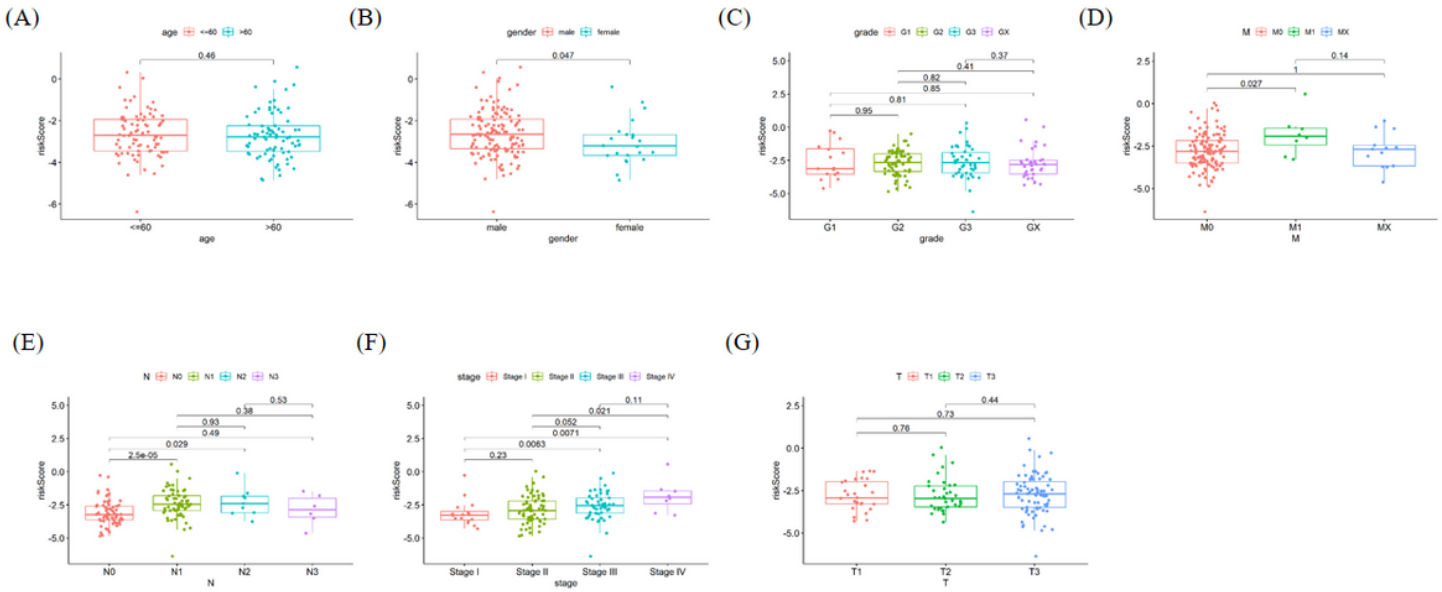

Figure 10

Differences in RBPrisk scores between different clinical features. (A) age; (B) sex; (C) grade; (D) M classification; (E) N classification; (F) stage; (G) T classification.

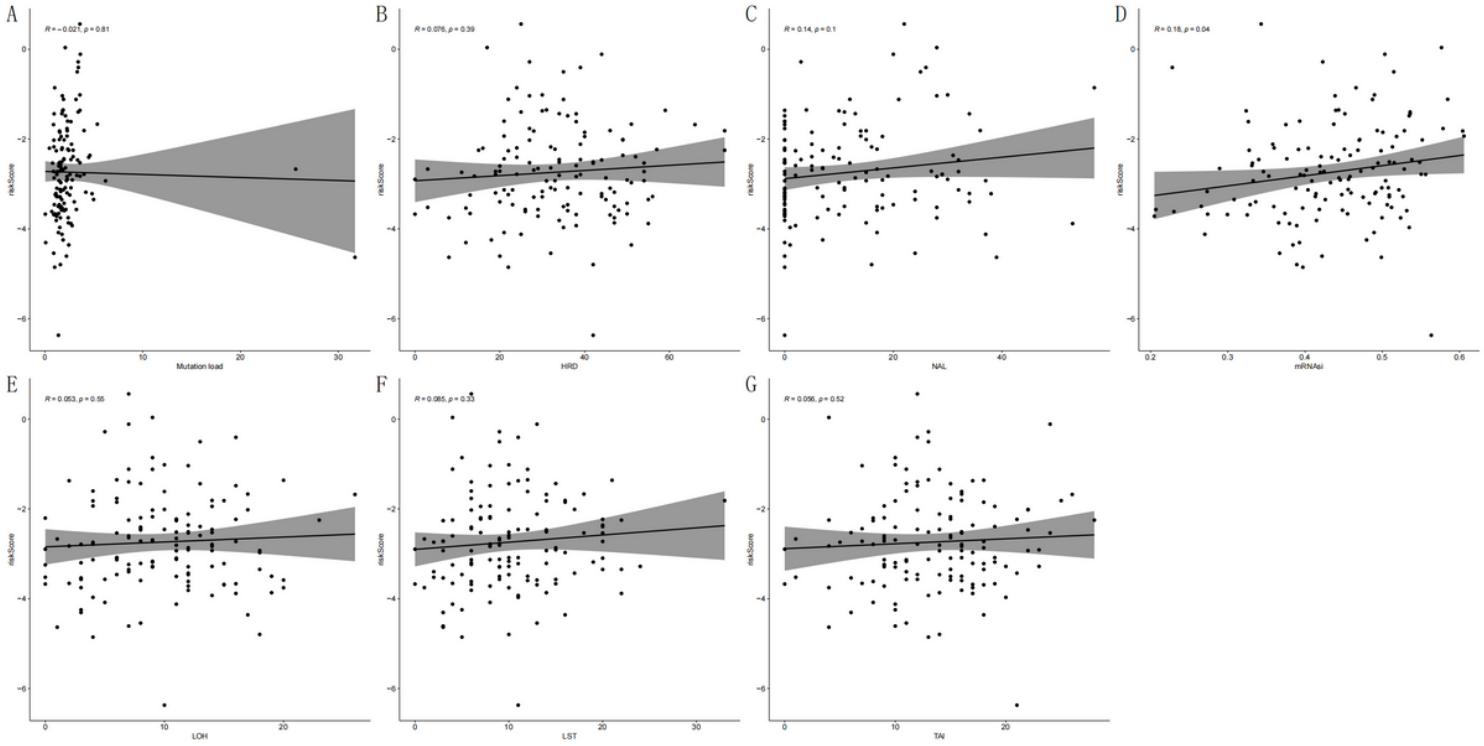

Figure 11

Relationship between the RBPrisk score and biomarkers for predicting ICI response.(A) TMB; (B) HRD; (C) NAL; (D) stemness index; (E) LOH; (F) LST; (G) TAI.

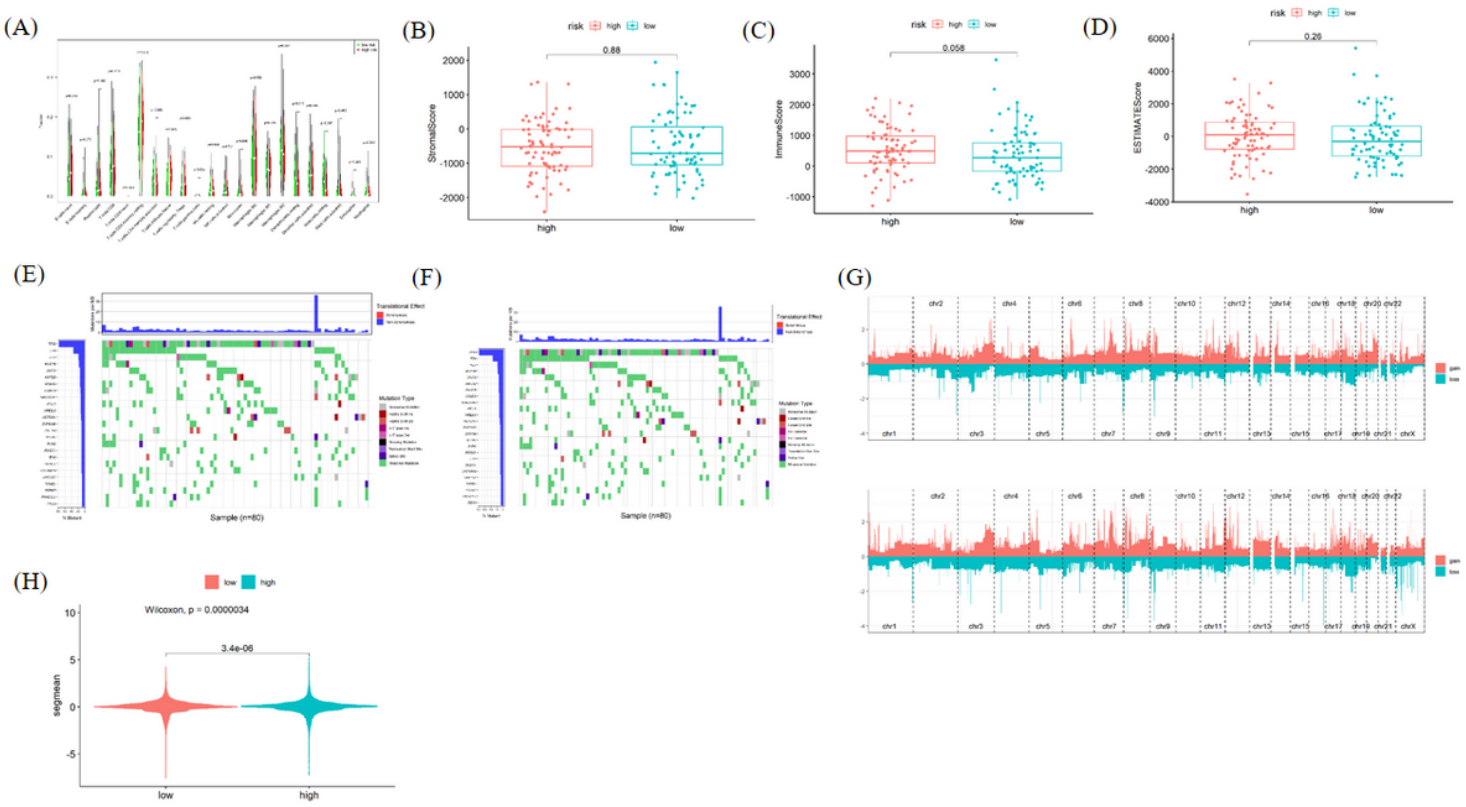

Figure 12

Differences in biomarkers for predicting ICI response between the high-risk and low-risk groups. (A)Immune cell infiltration; (B) immune score; (C) stromal score;(D) tumor purity;(E)(F)somatic mutation; (G)(H)CNV.

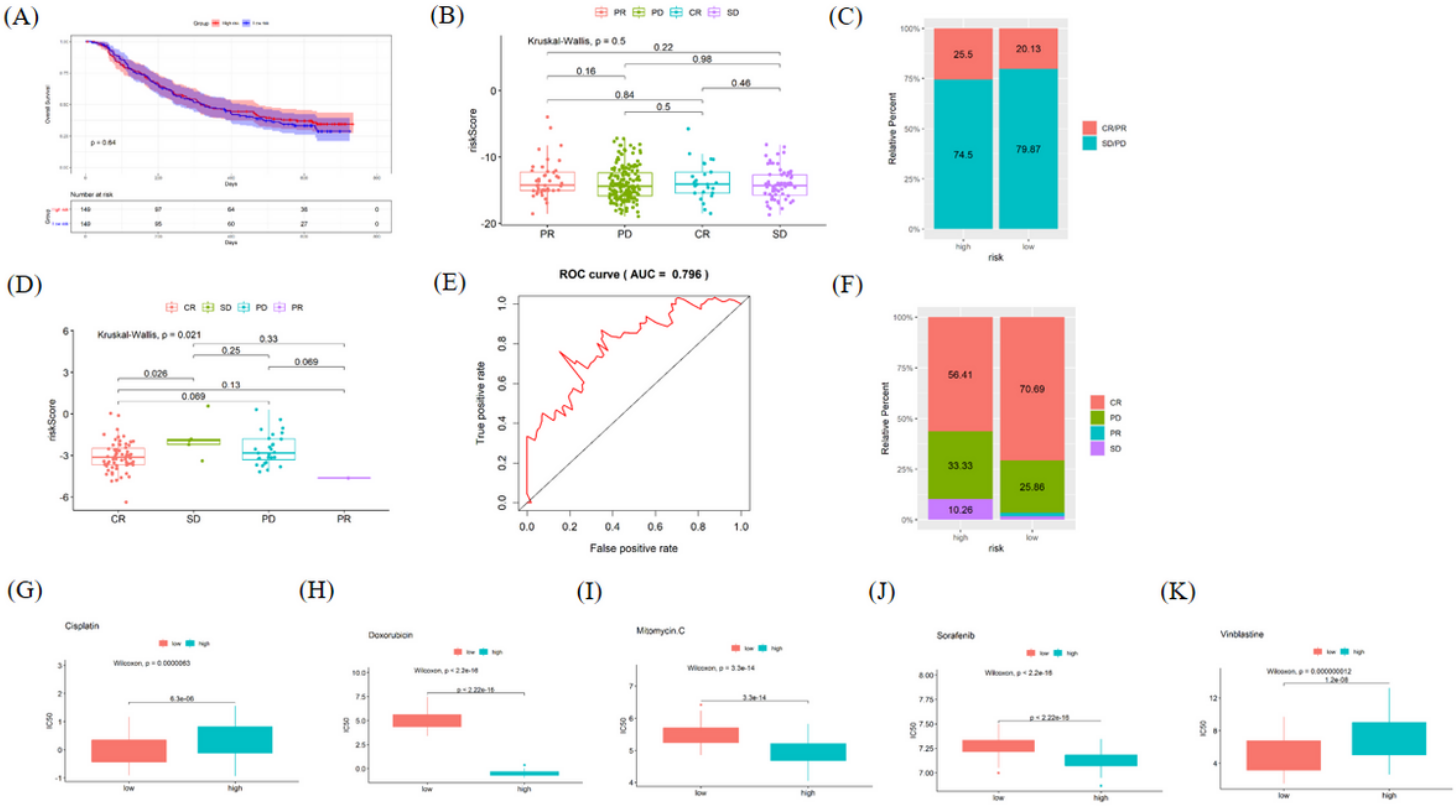

Figure 13

Evaluation of the RBP risk score for the prediction of response to treatment and prognosis. (A) Kaplan–Meier analysis stratified by the median risk score.(B) Relationship between the RBP risk score and therapeutic effect of ICIs. (C) Proportion of therapeutic effect of ICIs by the median risk score.(D) Relationship between the RBP risk score and therapeutic effect of chemotherapy.(E)ROC curve.(F) Proportion of therapeutic effect of chemotherapy by the median risk score.(G) Chemotherapy effect of cisplatin stratified by the median risk score. (H) Chemotherapy effect of doxorubicin stratified by the median risk score. (I) Chemotherapy effect of mitomycin C stratified by the median risk score.(J) Chemotherapy effect of sorafenib stratified by the median risk score.(K)Chemotherapy effect of vinblastine stratified by the median risk score.

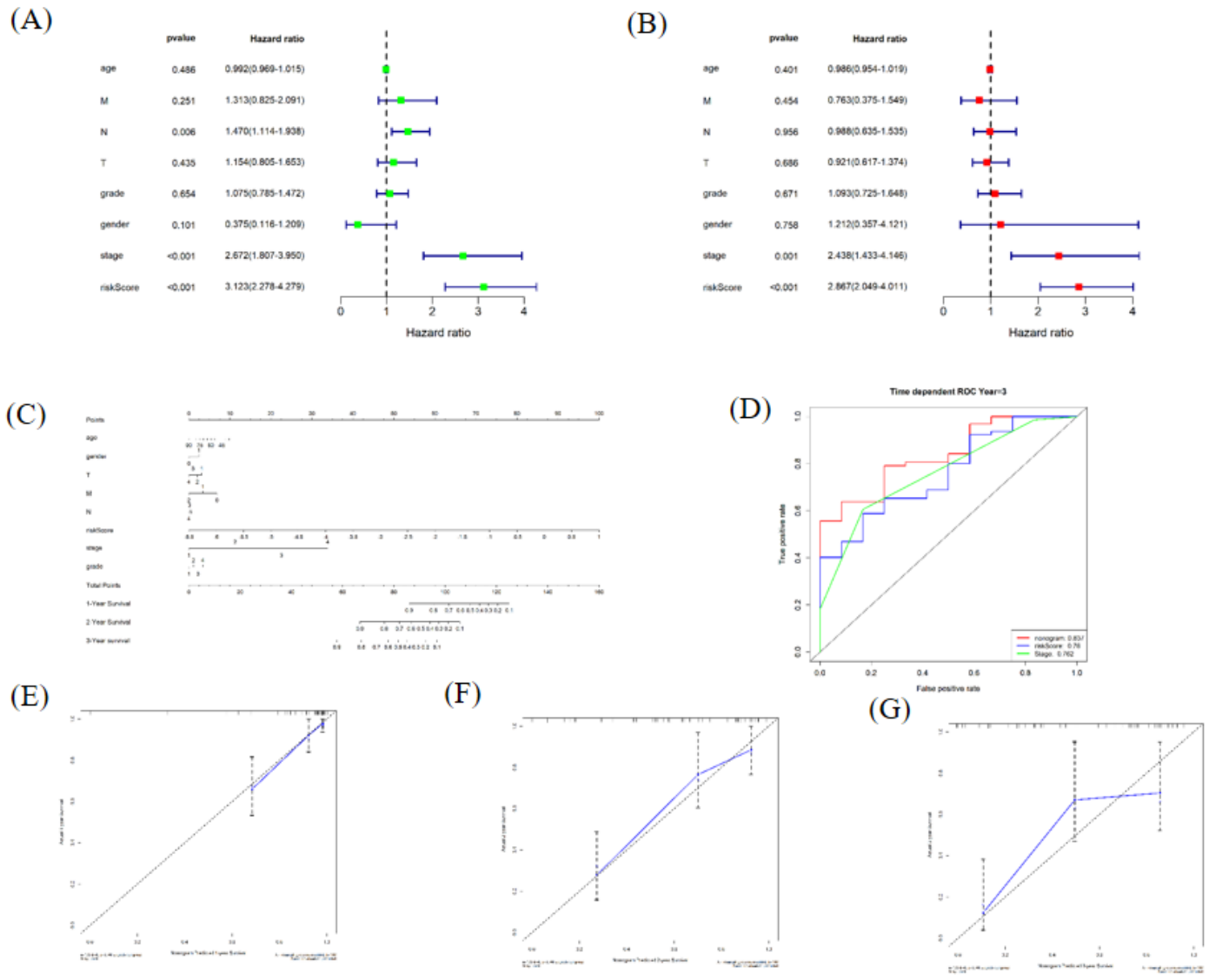

Figure 14

Independent prognostic value of the prognostic RBP signature and clinical parameters. (A) Univariate analysis. (B) Multivariate Cox analyses. (C)Nomogram for predicting the 1-, 2-, and 3-year OS of EC patients. (D)ROC curves of OS predictors. (E) (F) (G) Calibration plots of the nomogram at 1, 2, and 3 years.

Supplementary Files

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

- Supplementarylegends.docx
- Supplementarylegends.docx
- SupplementaryTable1.csv

- SupplementaryTable2.csv
- SupplementaryTable3.csv
- SupplementaryTable4.csv
- SupplementaryTable5.csv
- figureS1.tif