Original Article

Genome-wide profiling of a prognostic RNA-binding protein signature in esophageal cancer

Shaowu Sun[1#^], Junya Wang[2#], Yujie Zhang[2#], Yuetong Li[3], Yanan Guo[1], Chunyao Huang[1], Alfredo Tartarone[4], Pierlorenzo Pallante[5,6], Kaiyuan Li[1], Guoqing Zhang[1], Xue Pan[7], Xiangnan Li[1]

[1]Department of Thoracic Surgery and Lung Transplantation, First Affiliated Hospital of Zhengzhou University, Zhengzhou, China; [2]Department of Esophageal Surgery, First Affiliated Hospital of Zhengzhou University, Zhengzhou, China; [3]Undergraduate Program in Clinical Medicine, Henan University School of Medicine, Kaifeng, China; [4]Division of Medical Oncology, Department of Onco-Hematology, IRCCS-CROB, Referral Cancer Center of Basilicata, Rionero in Vulture, Italy; [5]Institute of Experimental Endocrinology and Oncology (IEOS) "G. Salvatore", National Research Council (CNR), Naples, Italy; [6]Department of Molecular Medicine and Medical Biotechnology (DMMBM), University of Naples "Federico II", Naples, Italy; [7]School of Nursing and Health, Zhengzhou University, Zhengzhou, China

Contributions: (I) Conception and design: S Sun, J Wang; (II) Administrative support: Y Zhang; (III) Provision of study materials or patients: Y Li, Y Guo; (IV) Collection and assembly of data: C Huang; (V) Data analysis and interpretation: K Li; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Xiangnan Li, MD, PhD; Guoqing Zhang, MD, PhD. Department of Thoracic Surgery and Lung Transplantation, First Affiliated Hospital of Zhengzhou University, No. 1 Jianshe Road, Zhengzhou 450052, China. Email: lxn-2000@163.com; drzhangguoqing@163.com; Xue Pan, PhD. School of Nursing and Health, Zhengzhou University, No. 100 Science Avenue, Zhengzhou 450052, China. Email: xuepan321@163.com.

Background: RNA-binding proteins (RBPs) are known to be involved in the initiation and development of malignant tumors, but the roles of RBPs in esophageal cancer (EC) remain unclear. This study aims to establish a prognostic signature based on RBPs through genome-wide analysis to predict the prognosis of EC patients and provide new insights into chemoresistance.

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 intersection of RBP genes, Kyoto Encyclopedia of Genes and Genomes pathway-related genes, and differentially expressed RBP genes from cluster analysis. Hub genes were extracted via protein-protein interaction network construction. A Cox proportional hazards regression model with seven prognostic RBPs (TRMT2A, PDHA1, MPRIP, KRI1, IL17A, HSPA1A, and HIST1H4J) was built. The risk score of each patient in internal and external dataset cohorts was calculated, and then the patients were divided into two groups based on the median value.

Results: There were significant differences in survival curves between the two risk groups in the internal and external dataset cohorts ($P<0.05$). In terms of chemotherapy, there was a significant association between RBP risk score and response to chemotherapy, with low-risk patients being more likely to achieve complete response. Finally, univariate and multivariate analyses indicated that the risk score was significantly correlated with overall survival ($P<0.05$), and pathological stage could also be used independently to predict the prognosis of EC.

Conclusions: Our study indicated that the RBP signature could serve as a prognostic biomarker of EC and provided new insights into the chemoresistance of this disease.

Keywords: Esophageal cancer (EC); RNA-binding protein (RBP); prognosis; immune checkpoint inhibitor (ICI)

^ ORCID: 0009-0003-8325-7396.

Submitted Dec 16, 2024. Accepted for publication Jan 16, 2025. Published online Feb 26, 2025.
doi: 10.21037/tcr-2024-2561
View this article at: https://dx.doi.org/10.21037/tcr-2024-2561

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 and lethal cancer (1). Esophageal adenocarcinoma (EAC) and esophageal squamous cell carcinoma (ESCC) are the two main types of EC. EAC 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 progress in screening and treatment, 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 available, clinicians have failed to achieve accurate early-stage detection of EC. Large amounts of gene expression data provide clinicians an excellent opportunity to identify potential tumor molecular markers, which may be of crucial 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). Research on RBPs in cancer has revealed their roles in regulating a wide array of biological processes associated with tumorigenesis, namely gene expression, apoptosis, epithelial-mesenchymal transition (EMT), and autophagy (4-6). In addition, the dysfunction of RBPs is also closely related to immune system disorders, neurodegenerative diseases, and other diseases of the human body (7,8). As a novel type of biomarker, RBPs have demonstrated potential to serve as therapeutic targets in various cancers, including colorectal, breast, and cervical cancer (9-11). A gene signature composed of multiple RBP-related genes can effectively predict the prognosis of certain patients with cancer (12). Although RBPs are known to be involved in the initiation and development of malignant tumors, the roles of RBPs in EC development still need to be elucidated. Moreover, 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 including The Cancer Genome Atlas (TCGA) and the Xena databases. After analysis, we identified differentially expressed RBPs between EC and normal samples and explored their functional roles and potential clinical value. Finally, we established a prognostic model based on the hub RBPs, identifying them as clinically significant prognostic biomarkers of EC. We present this article in accordance with the TRIPOD reporting checklist (available at https://tcr.amegroups.com/article/view/10.21037/tcr-2024-2561/rc).

Highlight box

Key findings

- The RNA-binding protein (RBP) signature could serve as a prognostic biomarker of esophageal cancer (EC) and provide new insights into the chemoresistance of EC in patients.

What is known and what is new?

- 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.
- We downloaded EC RNA-sequencing and clinical data from public databases. 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.

What is the implication, and what should change now?

- This study indicated that the RBP signature could serve as a prognostic biomarker of EC and provides new insights into its chemoresistance mechanisms.

Methods

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 extracted from the Eukaryotic RBP Database (EuRBPDB) (http://eurbpdb.gzsys.org.cn/) and related articles (13), and a total of 4,528 RBP-related genes were obtained.

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" package in R (The R Foundation for Statistical Computing, Vienna, Austria) with a threshold of P<0.05. Furthermore, the 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 subsequent step.

Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis of the differentially expressed RBPs was performed. "ClusterProfiler" R package was used to identify the KEGG pathways these genes were enriched in, with P<0.05 indicating statistical significance. Following this, KEGG pathway-related genes were identified as candidate-RBP-geneset1 (*cd-RBP-geneset1*). Cluster analysis was performed via the "ConsensusClusterPlus" package in R (repetitions =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 P<0.05 indicating statistical significance. The differentially expressed RBP genes were identified as candidate-RBP-geneset2 (*cd-RBP-geneset2*). Principal component analysis (PCA) was performed to facilitate data visualization.

Candidate genes were obtained via the intersection 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 at 50 to ensure the reliability of the results. Correlations between coexpression modules and clinical characteristics were analyzed, and the most clinically significant module was selected for protein-protein interaction (PPI) network construction to extract hub genes through Cytoscape 3.8 plugins.

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. We built a least absolute shrinkage and selection operator (LASSO) regression model using the "glmnet" R package to remove the redundancy factor via 10-fold cross-validation.

Subsequently, the model was validated in terms of prediction effect evaluation via Kaplan-Meier survival analysis and receiver operator characteristic (ROC) curves through the "survival" and "survivalROC" packages in R software. We then evaluated the model stability in different clinical subgroups using Kaplan-Meier survival analysis in a clinical gene dataset from TCGA. For external validation, the GSE72873 dataset was selected and analyzed in a similar manner. We also conducted pancancer analysis using a clinical gene dataset from TCGA and compared our proposed model with existing RBP-related models.

Correlation analysis of clinical features and biomarkers of immune checkpoint inhibitor (ICI) response

We explored the differences in RBP risk scores between different clinical features using independent *t*-tests. Additionally, numerous biomarkers for predicting ICI response have been explored in recent years, including tumor mutational 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 characterized 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 the R "maftools" package. HRD was downloaded from the Xena database (https://xenabrowser.net/). NAL was downloaded from the Tumor-Specific NeoAntigen database (TSNAdb) (http://biopharm.zju.edu.cn/tsnadb/). The stemness index was obtained from a previous study (14). The Pearson correlation coefficient test was used to estimate the relationship between the risk score and biomarkers. Immune cell infiltration was evaluated with the CIBERSORT algorithm. The immune score, stromal score, and tumor purity were analyzed with the "estimate" package in R, while somatic mutations were analyzed with the "maftools" package in R. 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.

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, containing 348 clinical

samples, of which, 298 containing immunophenotypes. We also evaluated the ability of the RBP risk score to predict responses to chemotherapy in TCGA database. The R package "pRRophetic" was employed in the sensitivity prediction of chemotherapeutic drugs (cisplatin, doxorubicin, mitomycin C, sorafenib, and vinblastine).

Independent prognostic value of the prognostic RBP signature

To assess the prognostic value of the 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.

Immunohistochemical (IHC) staining of hub genes

IHC staining was performed to verify differential hub gene expression between EC tumors and their adjacent normal tissues. Details of the experiment procedures are described in the Appendix 1. This study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). All patients provided written informed consent before tissues were collected, and this research was approved by the institutional Ethical Committee of the First Affiliated Hospital of Zhengzhou University (No. 2023-KY-1440-001).

Statistical analysis

All statistical analyses were performed using R software (version 4.0.3) and SPSS (version 26.0). Analysis of variance (ANOVA) or *t*-tests were used to assess differences in RNA expression levels between different subgroups. Prognostic models were constructed using univariate Cox regression and LASSO regression, and validated through Kaplan-Meier survival analysis and ROC curves. A chi-square test was used to analyze the correlation between risk-score and EC clinicopathological parameters. Survival analysis was performed using the Kaplan-Meier method. Pearson correlation analysis assessed the relationship between the RBP risk score and clinical features, as well as biomarkers of ICI response. The independent prognostic value of the RBP signature was evaluated using Cox proportional hazards models, adjusting for clinical variables. Unless otherwise specified, P values <0.05 were considered to indicate statistical significance, and all the P values were calculated as two-tailed tests.

Results

Basic information

The gene expression profiles of 172 samples (161 tumors and 11 adjacent normal tissues) from patients with EC were obtained from the Xena database. RBP-related genes were downloaded from the EuRBPDB and relevant articles (13), and a total of 4,528 RBP-related genes were obtained. A total of 337 hub genes were obtained through DEG analysis and weighted gene co-expression network analysis (WGCNA). Among these genes, prognostic RBPs were identified using univariate Cox regression analysis. Subsequently, a LASSO-penalized Cox proportional hazards regression model with multiple prognostic RBPs was built. It was validated in the internal and external [GSE72873 (n=44)] dataset. Finally, pancancer 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 *Figure 1*.

Identification of differentially expressed RBP genes in patients with EC

In this study, we obtained 1,898 DEGs from the Xena database (*Figure 2A*,*2B*; table available at https://cdn.amegroups.cn/static/public/tcr-2024-2561-1.xlsx). We then performed KEGG pathway enrichment analysis of the differentially expressed RBPs, and 62 pathways with 1,331 RBP-related genes were identified (cd-RBP-geneset1) (*Figure 2C*; tables available at https://cdn.amegroups.cn/static/public/tcr-2024-2561-2.xlsx and https://cdn.amegroups.cn/static/public/tcr-2024-2561-3.xlsx).

We found that 1,879 RBP genes could reflect the prognosis of patients with EC. Fifty-four DEGs were obtained via 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 (*Figure 2D*,*2E*). The final number of clusters was k =2. Therefore, two clusters of patients

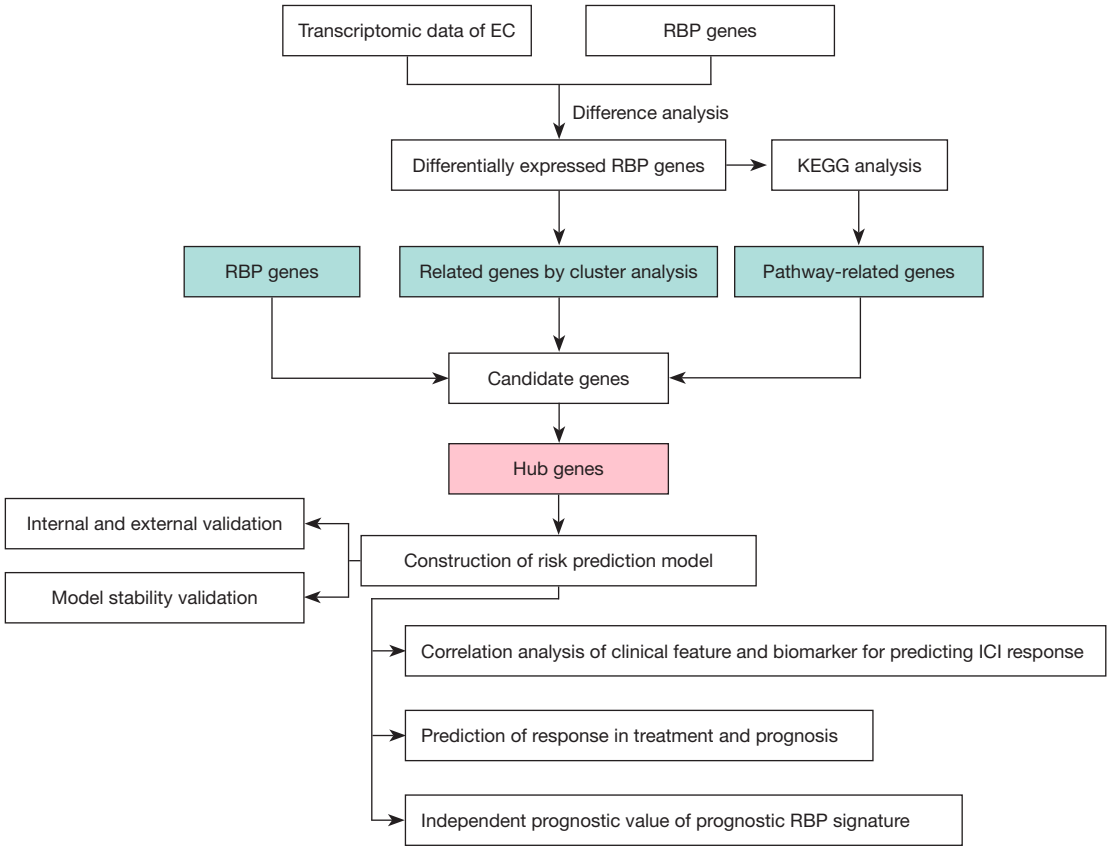

Figure 1 Flowchart of the study. EC, esophageal cancer; ICI, immune checkpoint inhibitor; KEGG, Kyoto Encyclopedia of Genes and Genomes; RBP, RNA-binding protein.

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 (*Figure 2F*). Furthermore, 635 DEGs were obtained from cluster analysis (cd-RBP-geneset2) (table available at https://cdn.amegroups.cn/static/public/tcr-2024-2561-4.xlsx). PCA was effectively able to distinguish between the above subtypes (*Figure 2G,2H*).

Construction of the prognostic RBP signature

To further identify the association between RBPs and patients with EC with different clinical characteristics, the expression data profiles of RBP genes, cd-RBP-geneset1, and cd-RBP-geneset2 (*Figure 3A*) were transformed into a gene coexpression network using the "WGCNA" package in R (*Figure 3B,3C*). A total of 16 coexpressed modules were obtained through a one-step network construction method (*Figure 3D-3F*). Furthermore, we performed a correlation analysis between different coexpression modules and OS/OS time (*Figure 3D,3E*). The results indicate that the turquoise module had the most significant correlation with OS (correlation coefficient =0.19; P=0.02) (*Figure 3F*). However, there were no modules significantly correlated with OS time. The distribution of the modules' average gene significance related to OS and OS time are shown in *Figure 3D,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 (*Figure 3G*; table available at https://cdn.amegroups.cn/static/public/tcr-2024-2561-5.xlsx).

Among these 337 hub genes, 24 prognostic RBPs (6 newly predicted) were identified using univariate Cox regression analysis (Figure S1). Subsequently, we performed multivariate Cox regression analysis (*Figure 4A*) and LASSO

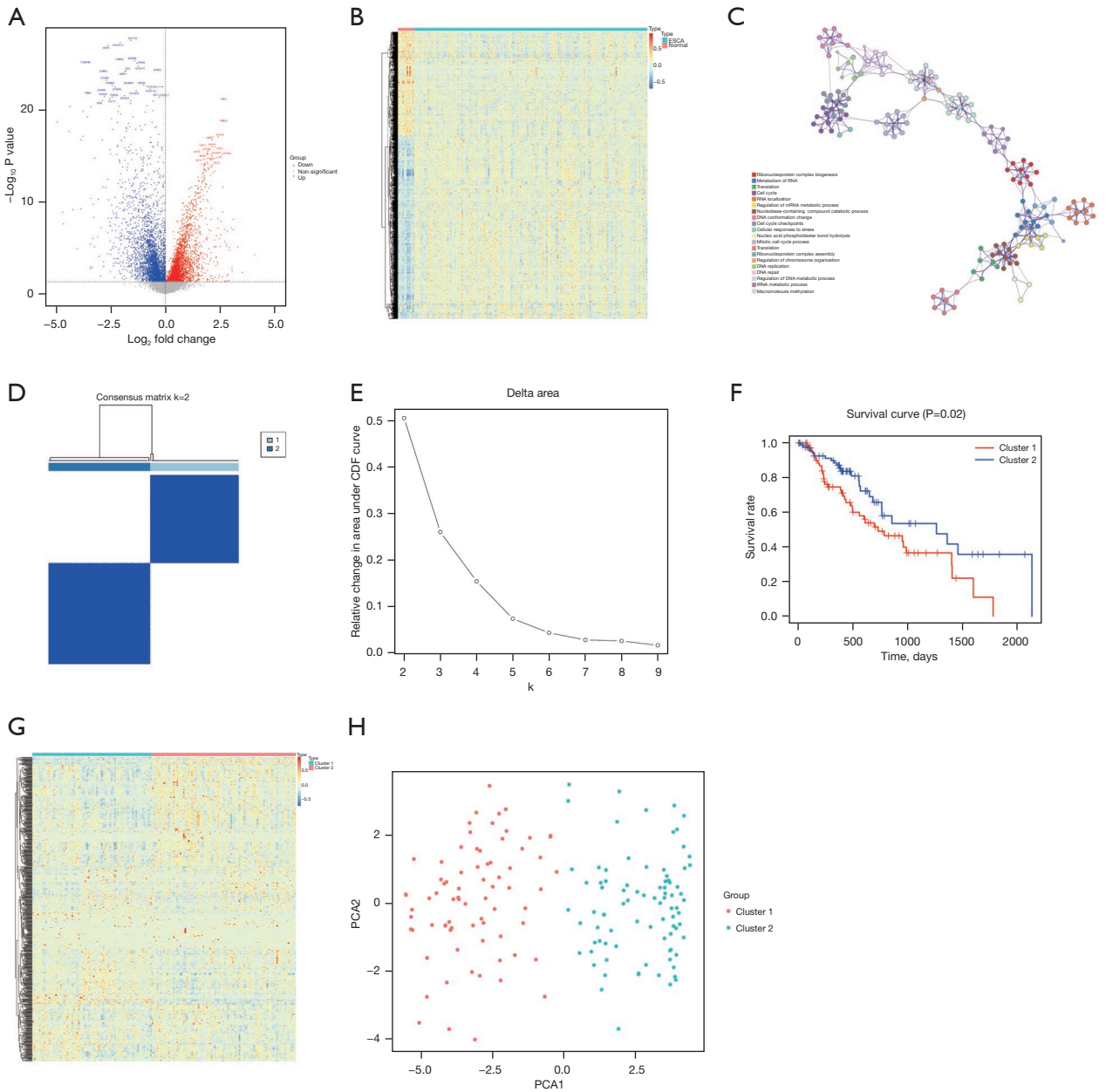

Figure 2 Identification of DE RBP-related genes. (A) Volcano plot of the DE genes. (B) Heatmap of the DE genes. (C) KEGG pathway enrichment analysis of DE RBPs. (D) Consensus clustering analysis identification of two clusters. (E) CDF for k =2–9. (F) Kaplan-Meier curves for patients with EC stratified by cluster. (G) Heatmap of the prognostic RBP genes ordered by clusters. (H) Principal component analysis of different clusters. CDF, cumulative distribution function; DE, differentially expressed; EC, esophageal cancer; ESCA, esophageal cancer; KEGG, Kyoto Encyclopedia of Genes and Genomes; RBP, RNA-binding protein.

regression analysis (*Figure 4B*) separately using these 24 genes. We then intersected the statistically significant genes from the multifactorial Cox regression with the 14 candidate genes selected by LASSO analysis, ultimately identifying 7 risk score genes. The formula for the risk score was as follows: risk score = TRMT2A × (–0.9916) + PDHA1 × 0.9129 + MPRIP × (–1.121) + KRI1 × (–0.6753) + IL17A × 1.282 + HSPA1A × 0.2014 + HIST1H4J × 1.256

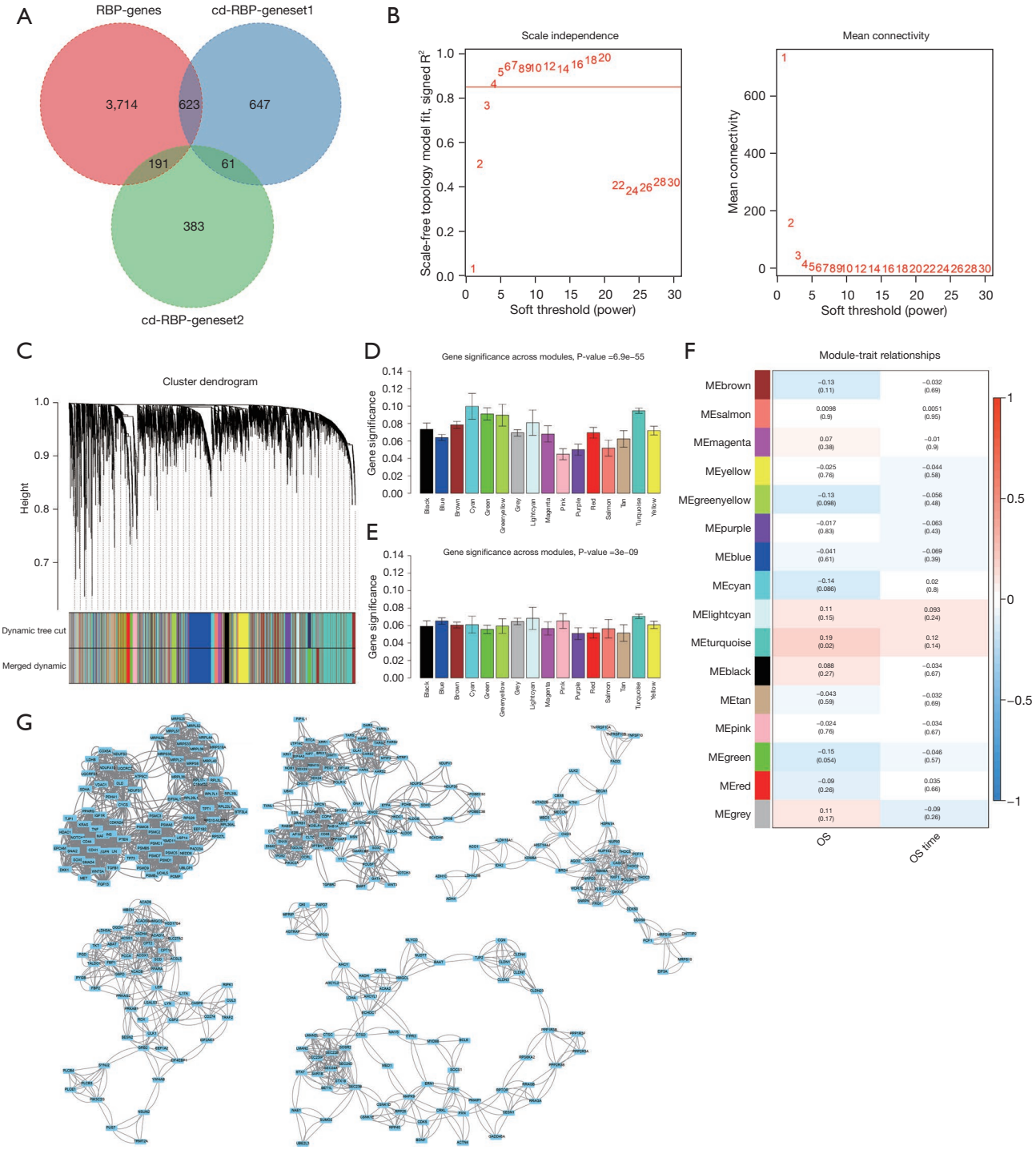

Figure 3 Construction of the prognostic RBP signature. (A) Venn diagram showing the RBP genes, cd-RBP-geneset1, and cd-RBP-geneset2. (B) The effect of different power values on the scale independence and average connectivity degree of coexpression modules of RBP-related genes. (C) The constructed coexpression modules of RBP-related genes via the "WGCNA" R software package. (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) Protein-protein interaction network construction. ME, module; OS, overall survival; RBP, RNA-binding protein; WGCNA, weighted gene co-expression network analysis.

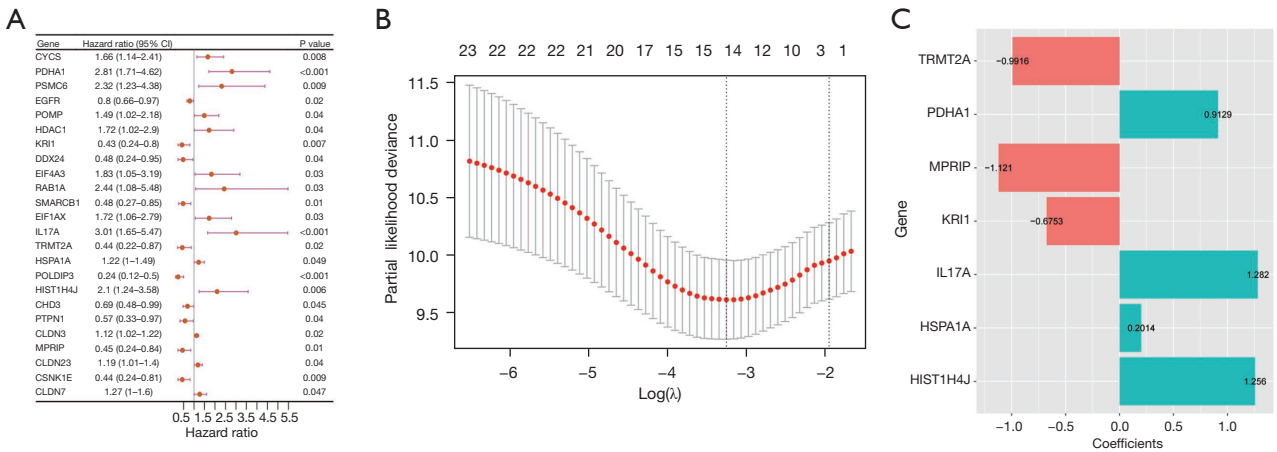

Figure 4 LASSO Cox regression analysis of RBP-related genes. (A) Cox regression analysis. (B) Selection of the tuning parameter (λ) in the LASSO model via 10-fold cross-validation based on the minimum criteria. (C) Identification of the prognostic risk score system with 7 prognostic RBPs. CI, confidence interval; LASSO, least absolute shrinkage and selection operator; RBP, RNA-binding protein.

(*Figure 4C*).

Additionally, we analyzed the expression levels of these 7 genes through IHC in EC tissue and adjacent normal tissues. And the results showed that the expression of hub genes in EC tumor tissue was higher than that in adjacent normal tissues. The H-scores in the tumor samples were significantly higher than those in their matched normal tissues (*Figure 5*).

Validation of the risk model

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

In the internal dataset cohort, the risk score of each patient was calculated, and patients were divided into two groups based on the median, with 80 patients in each of the high- and low-risk groups. The risk score distribution and survival status for this cohort are displayed in *Figure 6A,6B*. The survival curves in *Figure 6C* show that the prognosis of patients in the high-risk group was significantly poorer than that of patients in the low-risk group (P<0.001). The heatmap of risk gene expression in both groups is shown in *Figure 6D*, and the area under the curve (AUC) values at 1, 2, and 3 years were 0.789, 0.821, and 0.797, respectively, indicating the good predictive accuracy of the risk score model (*Figure 6E*). For the external dataset cohort, a similar division was made, with 22 patients in each risk group. The risk score distribution and survival status for this cohort are displayed in *Figure 7A,7B*. The survival curves also show the same trend (P=0.008) (*Figure 7C*). The heatmap of risk gene expression in both groups is shown in *Figure 7D*, and the AUC value at 3 years for this cohort was 0.701 (*Figure 7E*).

We then evaluated the model stability in different clinical subgroups, and the age, sex, grade, stage, and T classification subgroups showed significant differences in OS between the high-risk and low-risk groups (Figure S2). In the TCGA pancancer analyses, except for breast cancer (BRCA), all other cancer types [adenoid cystic carcinoma (ACC), colon adenocarcinoma/rectum carcinoma (COAD), kidney renal clear cell carcinoma (KIRC), and low-grade glioma (LGG)] showed significant differences in OS between the high- and low-risk groups (Figure S3). We further compared our proposed model with two existing RBP-related models, prognostic model 1 (15) and prognostic model 2 (16), and our model achieved a higher AUC value (Figure S4).

Correlations of RBP risk score with clinical features and biomarkers of ICI response

We explored the differences in RBP risk scores between different clinical features, and the results are shown in *Figure 8*. There were significant differences between the RBP risk score and clinical features such as sex (*Figure 8B*), M classification (M1 *vs.* M0) (*Figure 8D*), N classification (N1 and N2 *vs.* N0) (*Figure 8E*), and stage (stage III and stage IV *vs.* stage I, stage IV *vs.* stage II) (*Figure 8F*), while no statistical differences were found in age (*Figure 8A*), pathological grading (*Figure 8C*), and T classification (*Figure 8G*). We further explored the relationship between

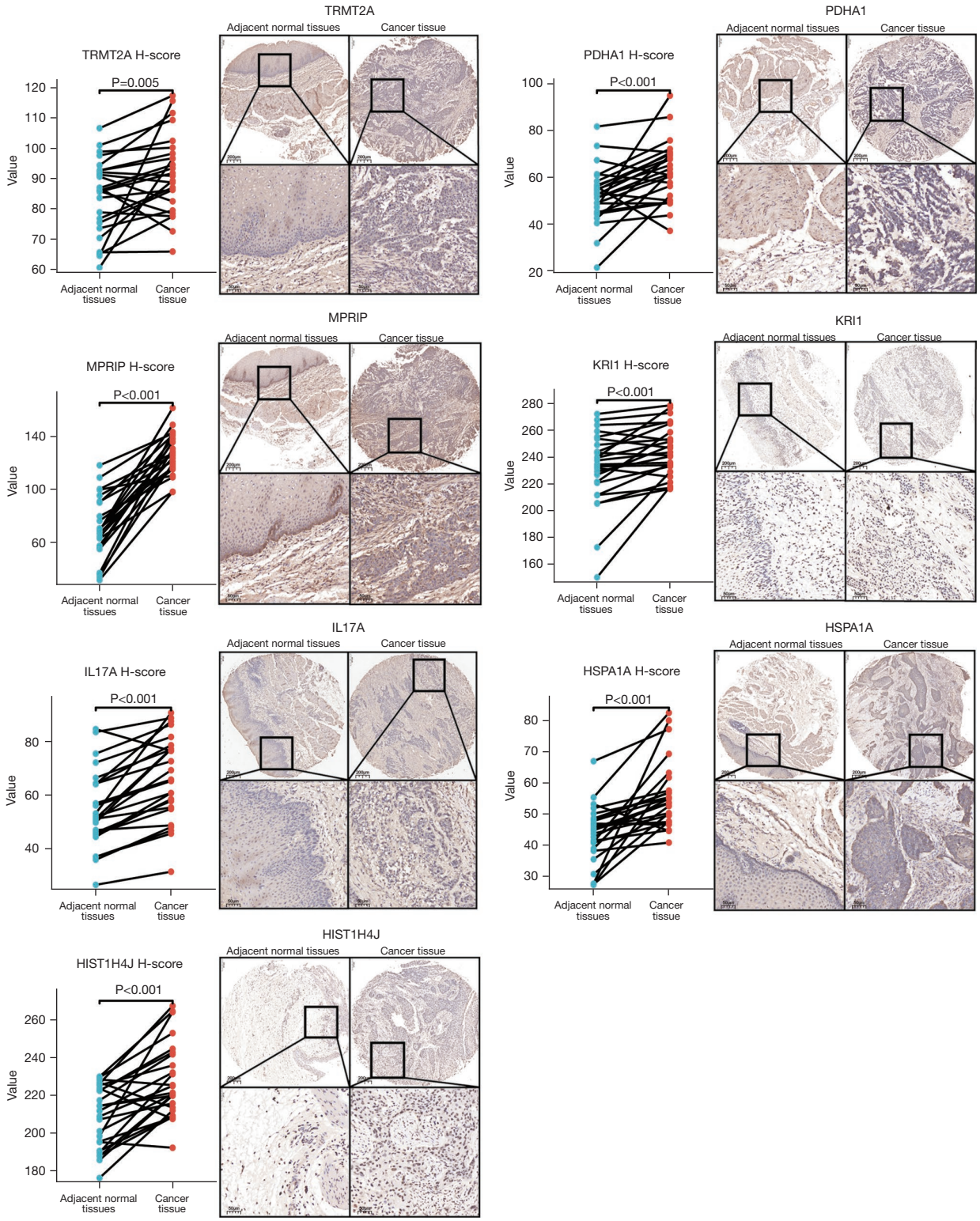

Figure 5 The expression of hub genes (*TRMT2A*, *PDHA1*, *MPRIP*, *KRI1*, *IL17A*, *HSPA1A*, *HIST1H4J*) in esophageal cancer tissue and adjacent normal tissues. Staining method: enzyme-labeled secondary antibody (Horseradish Peroxidase).

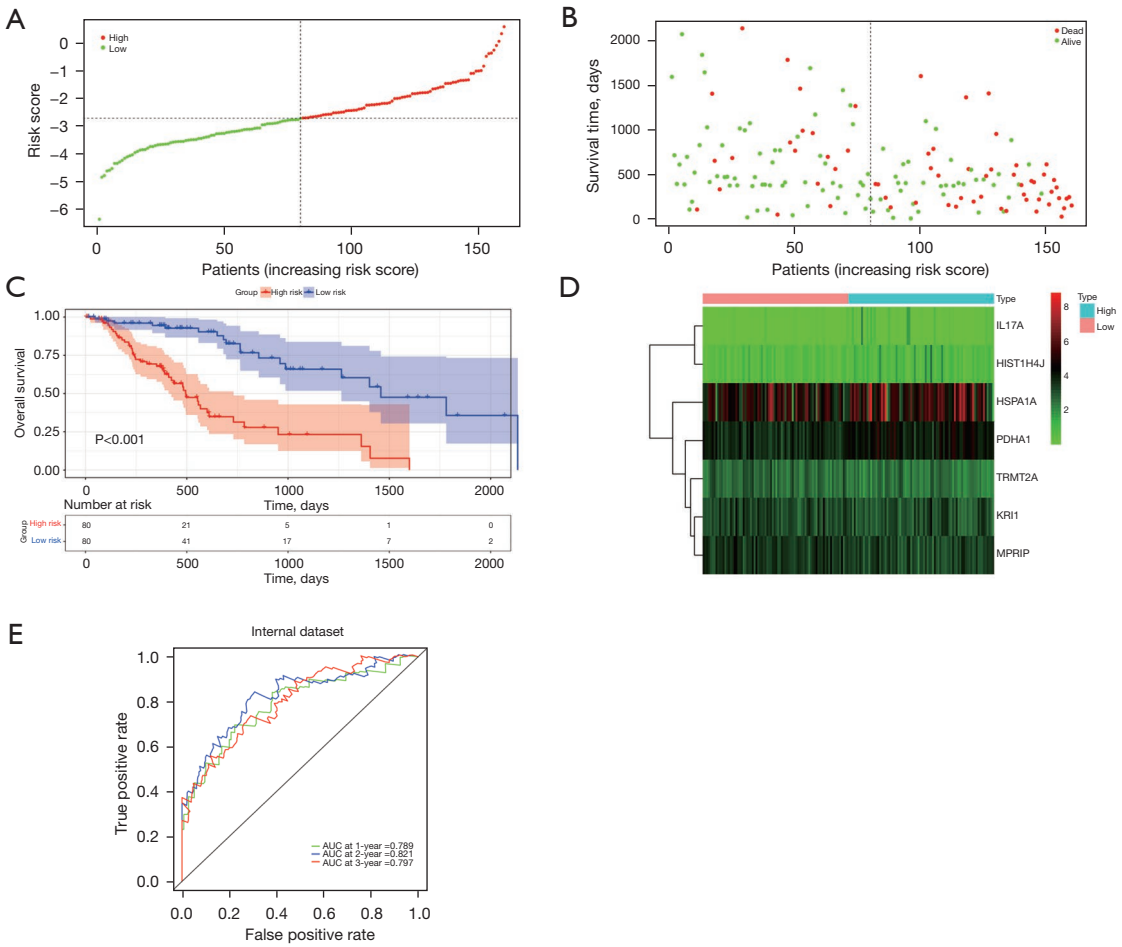

Figure 6 Internal signature validation. (A) The patients were divided into two groups: low- and high-risk groups. (B) As the risk score increased, the survival time of patients decreased, and the number of deaths increased. (C) Kaplan-Meier analysis of patients with EC stratified by the median risk score. (D) The heatmap shows the expression profiles of the seven RBP-related genes in the prognostic signature. (E) The signature was evaluated by using the sensitivity and specificity of the ROC curve. AUC, area under the curve; EC, esophageal cancer; RBP, RNA-binding protein; ROC, receiver operating characteristic.

the RBP risk score and biomarkers of ICI response (including TMB, HRD, NAL, stemness index, LOH, LST, and TAI) in cancer, and the results showed that there was a remarkable relationship between messenger RNA expression-based stemness index (mRNAsi) and the RBP risk score, while the other biomarkers had no significant correlations (Figure S5).

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 the infiltration levels of other immune cells (*Figure 9A*). The estimation of stromal and immune cells in malignant tumor tissues using expression data (ESTIMATE) analysis showed that there were no significant differences regarding the stromal score (*Figure 9B*), the immune score (*Figure 9C*) or tumor purity (*Figure 9D*). Similarly, there were no differences regarding the somatic mutation type or frequency (*Figure 9E*). High-risk patients had significantly higher CNVs than did their low-risk

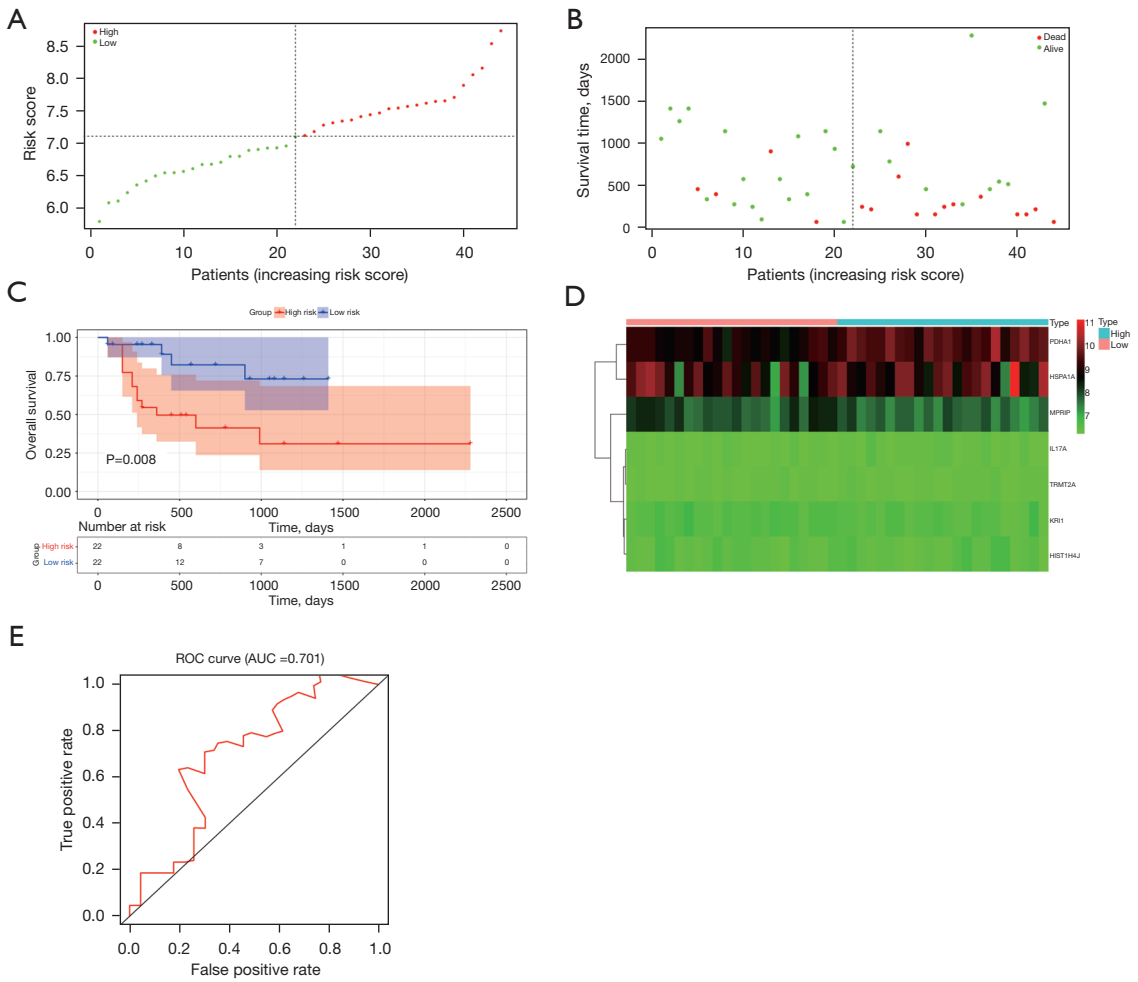

Figure 7 External signature validation. (A) The patients were divided into two groups: low- and high-risk groups. (B) As the risk score increased, the survival time of patients decreased, and the number of deaths increased. (C) Kaplan-Meier analysis of patients with EC stratified by the median risk score. (D) The heatmap shows the expression profiles of the seven RBP-related genes in the prognostic signature. (E) The signature was evaluated by using the sensitivity and specificity of the ROC curve. AUC, area under the curve; EC, esophageal cancer; RBP, RNA-binding protein; ROC, receiver operating characteristic.

counterparts (P<0.001) (*Figure 9F*).

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

In terms of immunotherapy, there was no significant difference in OS between the high- and low-risk groups (P=0.64) (*Figure 10A*). The differences in RBP risk scores between patients with different responses to treatment were not significantly different (P>0.05) (*Figure 10B*). Additionally, the proportions of complete response (CR)/partial response (PR) and stable disease (SD)/progressive disease (PD) in the high-risk group were 25.5% and 74.5%, respectively, while those in the low-risk group were 20.13% and 79.87%, respectively, and were not significantly different (*Figure 10C*).

In terms of chemotherapy, there was a significant association of RBP risk score and response to chemotherapy, with low-risk patients being more likely to achieve CR (*Figure 10D*). The AUC value for the prediction of chemotherapy response was 0.796 (*Figure 10E*) and the relative proportions of CR, PR, SD and PD were 56.41%, 0%, 10.26% and 33.33%, respectively, in the high-risk group, while they were 70.69%, 1.72%, 1.72% and 25.86%,

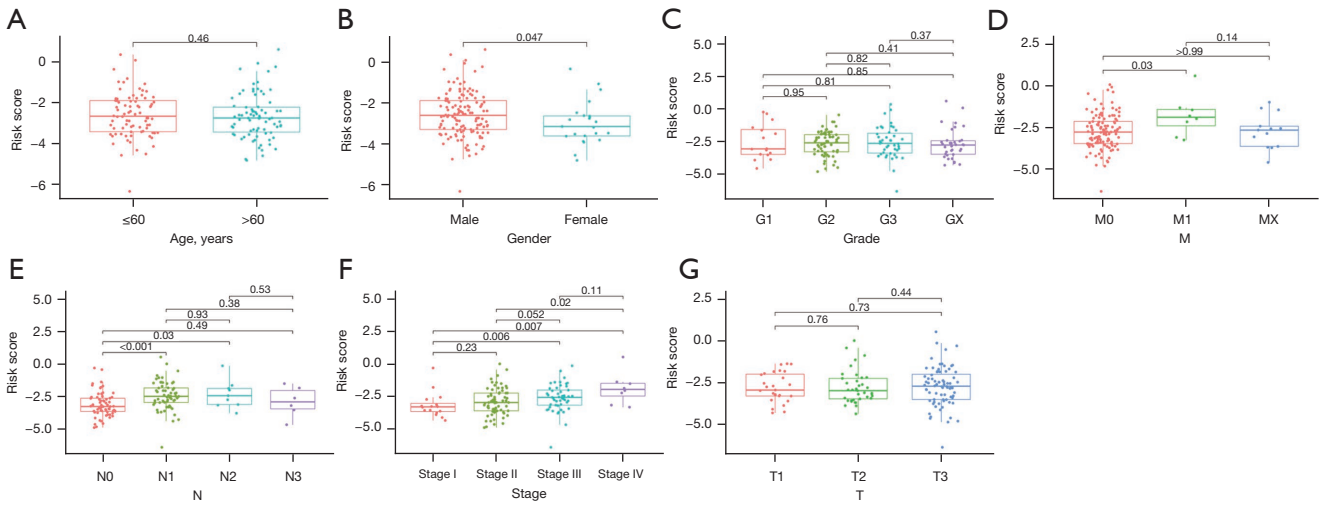

Figure 8 Differences in RBP risk scores between different clinical features. (A) Age. (B) Sex. (C) Grade. (D) M classification. (E) N classification. (F) Stage. (G) T classification. RBP, RNA-binding protein.

respectively, in the low-risk group (*Figure 10F*). We further analyzed the association of chemoresistance with the RBP risk score, and the results showed that low-risk patients tended to be more sensible to cisplatin and vinblastine, while high-risk patients tended to be more sensible to doxorubicin, mitomycin C and sorafenib (*Figure 10G*).

Independent prognostic value of the RBP signature and clinical parameters

The risk score was found to be significantly related to OS according to univariate (*Figure 11A*) and multivariate (*Figure 11B*) analyses ($P<0.05$). 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 (*Figure 11C*) was more accurate than the risk score and pathological stage in predicting OS at 3 years. In fact, the AUCs at 3 years for the nomogram, the risk score, and pathological stage were 0.837, 0.76, and 0.762, respectively (*Figure 11D*). The calibration curves indicate that the survival probabilities predicted by our model have a good fit with the actual probabilities of patient survival at 1, 2, and 3 years (*Figure 11E-11G*).

Discussion

EC is the most common malignant tumor of the digestive tract and has a poor prognosis. Studies have established several novel biomarkers signatures to determine the prognosis of patients with EC, with research directions involving microRNA (17,18), autophagy (19), DNA repair (20,21), m6A RNA methylation (22), long noncoding RNA (23), epigenetics (24), lymph node metastasis (25), urinary metabolomic (26), and immunology (27). RBP-related biomarkers have demonstrated reliable predictive capabilities for tumor risk characteristics in various cancers (28-33). In EC, existing research has focused on the diagnostic and prognostic predictive abilities of the models, with a lack of systematic investigation into the potential biological characteristics of these models (12,34). Therefore, this study constructed a novel RBP-related prognostic signature, further validated the expression of the signature through IHC, and systematically analyzed the relationship between the signature and clinical characteristics, response to ICIs, and response to immunotherapy.

Notably, not all 7 prognostic RBP-related genes have been reported to be associated with prognosis in cancer. For TRMT2A, only one study has reported that TRMT2A protein expression is a biomarker of increased risk of recurrence in patients with human epidermal growth factor receptor 2 (HER2)-positive breast cancer and may be used to predict the response to adjuvant chemotherapy (35). However, as shown in our prediction model, TRMT2A was found to be a gene associated with improved prognosis, but this needs further validation in EC. Recent studies have shown that PDHA1 plays a multifaceted role in cancer biology, including its potential as a therapeutic target, its involvement in the metabolic reprogramming of tumor cells,

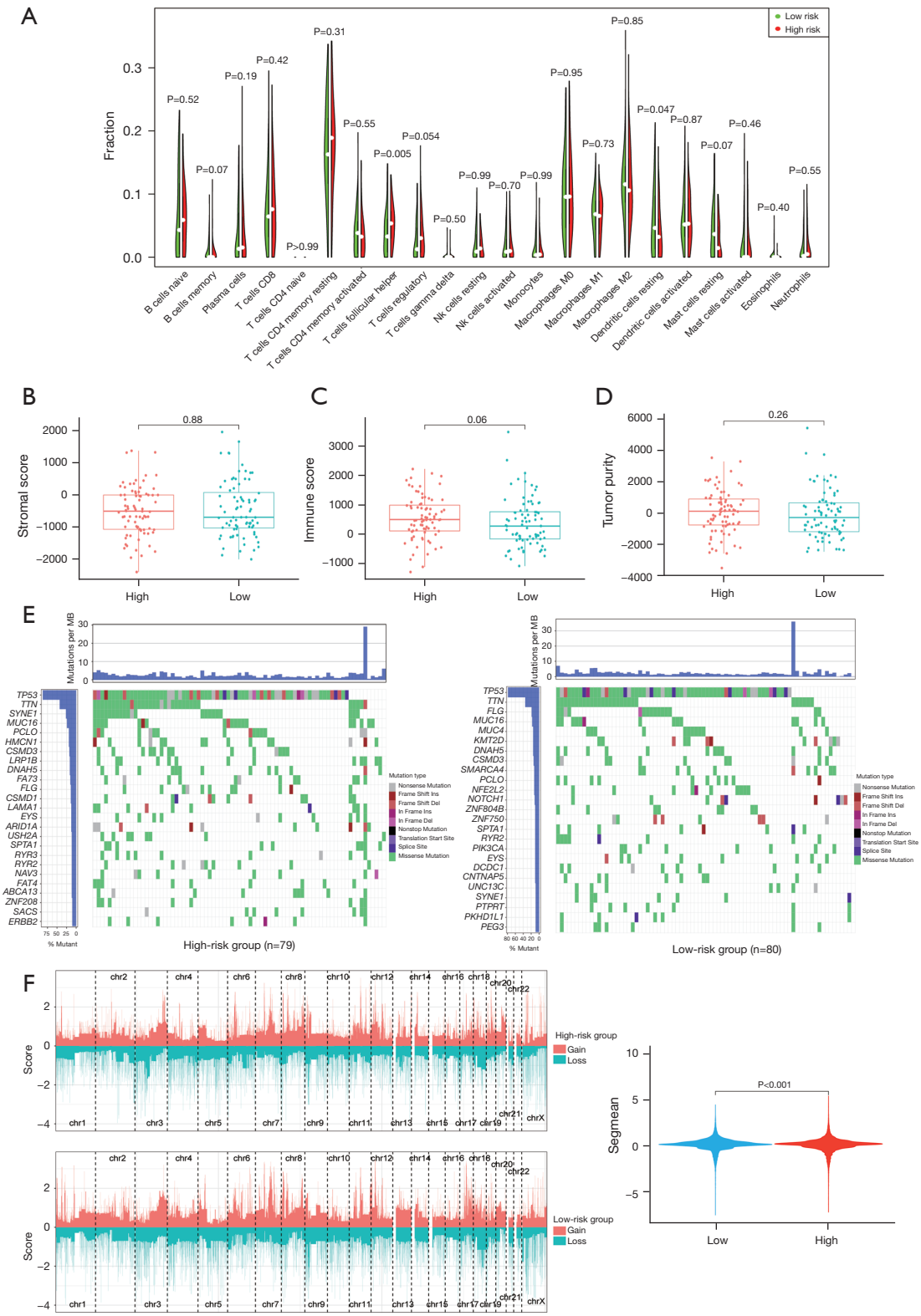

Figure 9 Differences in biomarkers for predicting ICI response between the high- and low-risk groups. (A) Immune cell infiltration. (B) Stromal score. (C) Immune score. (D) Tumor purity. (E) Somatic mutation. (F) CNV. CNV, copy number variation; ICI, immune checkpoint inhibitor; MB, Megabase.

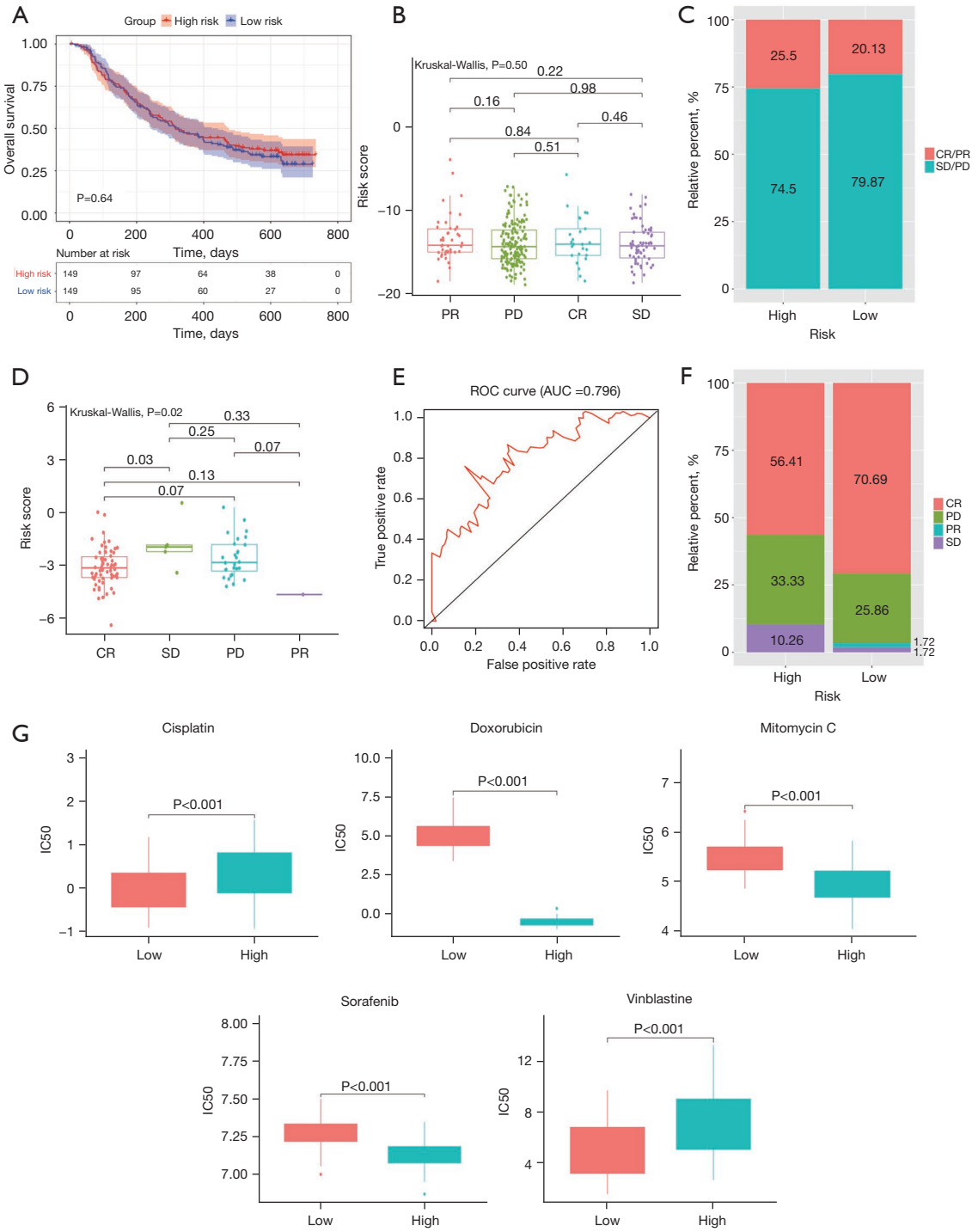

Figure 10 Evaluation of the RBP risk score for the prediction of response to treatment and prognosis. (A) Kaplan-Meier analysis stratified by the median risk score. (B) Relationship between the RBP risk score and therapeutic effect of ICIs. (C) Proportion of therapeutic effect of ICIs by the median risk score. (D) Relationship between the RBP risk score and therapeutic effect of chemotherapy. (E) ROC curve. (F) Proportion of therapeutic effect of chemotherapy stratified by the median risk score. (G) Chemotherapy effect of cisplatin, doxorubicin, mitomycin C, sorafenib, vinblastine stratified by the median risk score. AUC, area under the curve; CR, complete response; IC50, half maximal inhibitory concentration; ICI, immune checkpoint inhibitor; PD, progressive disease; PR, partial response; RBP, RNA-binding protein; ROC, receiver operating characteristic; SD, stable disease.

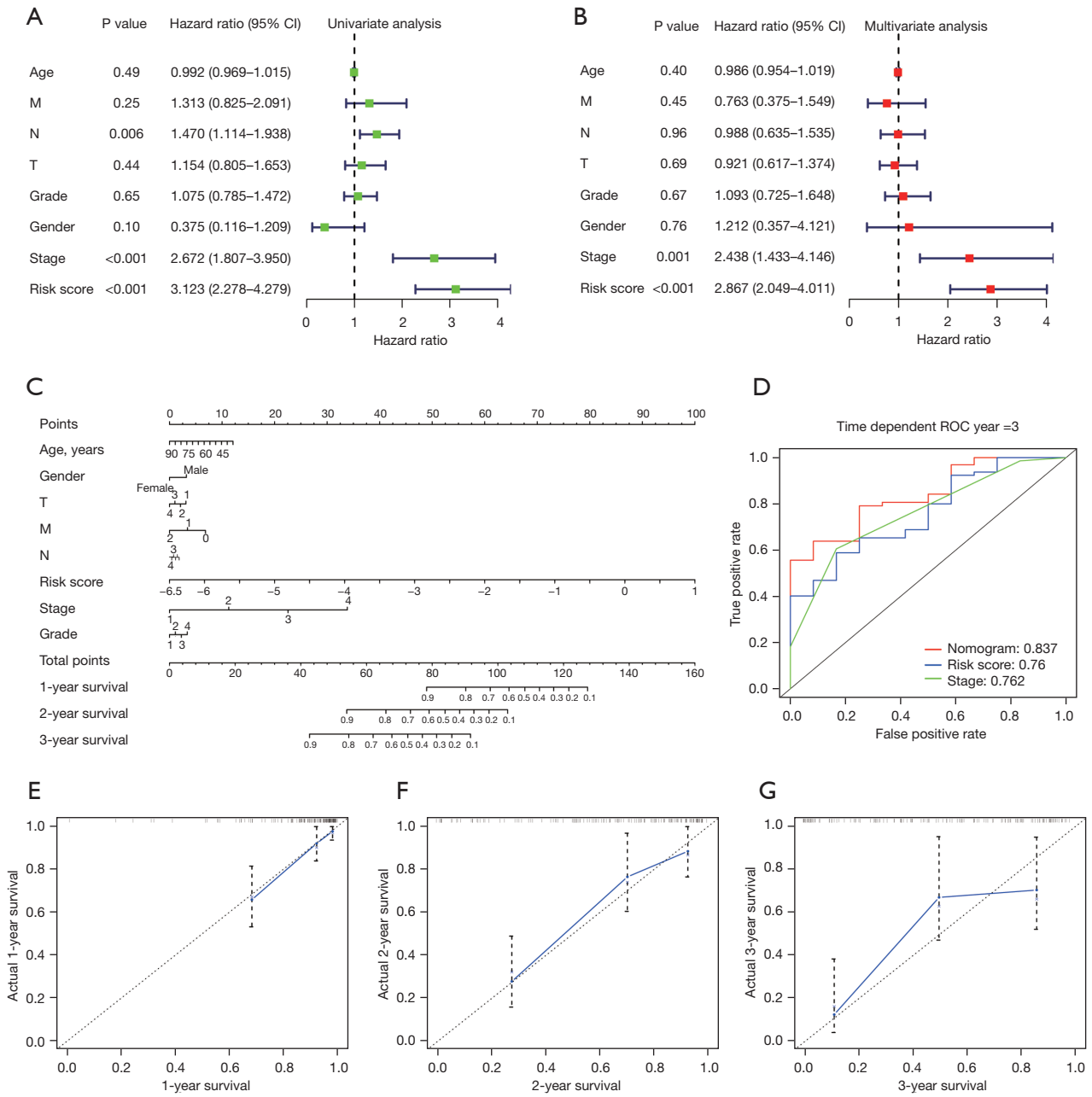

Figure 11 Independent prognostic value of the prognostic RBP signature and clinical parameters. (A) Univariate analysis. (B) Multivariate Cox analyses. (C) Nomogram for predicting the 1-, 2-, and 3-year OS of patients with EC. (D) ROC curves of OS predictors. (E-G) Calibration plots of the nomogram at 1, 2, and 3 years. EC, esophageal cancer; OS, overall survival; RBP, RNA-binding protein; ROC, receiver operating characteristic.

and its impact on immune responses and the pluripotency of stem cells. In EC, PDHA1, as a cuproptosis-related gene, can effectively predict the immune status and prognosis of tumors (36,37). IL17A is a member of the IL17 cytokine family and is released by immune, stromal and tumor cells, into the tumor microenvironment. Among all the IL17 cytokine family members, IL17A is the most controversial in terms of its role in regulating tumor immunity, and its prognostic value varies across cancer types. Generally, tumor-infiltrating IL17A-producing cells (IL17A$^+$ cells)

are associated with elevated antitumor immunity (38). A study demonstrated that IL17A deficiency reduces tumor latency and promotes metastasis in lung cancer (39), while other research 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 (40). In pancreatic cancer, an anti-IL17A antibody can enhance the antitumor response to gemcitabine (41), which implies that IL17A may be involved in the development of drug resistance. IL17A polymorphisms have been associated with the risk of various cancers, and the IL17A rs4711998 A>G polymorphism was reported to be associated with a decreased risk of EC (42). In our study, IL17A could indicate poor survival in patients with EC. Research suggests that HSPA1A is associated with unfavorable survival and poor clinicopathological features in several kinds of tumors (43-45). Furthermore, HSPA1A was demonstrated to mediate breast cancer radioresistance (46). However, no study has examined the effects that HSPA1A exerts on malignant biological properties, treatment sensitivity, or prognosis in EC. Importantly, no studies have reported MPRIP, KRI1, or HIST1H4J as being associated with prognosis in cancer, and thus our findings are the first to identify their association 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 confirmed by internal and external validation. Importantly, the RBP signature represented also a strong independent indicator of survival with adjusted clinical parameters, including age, sex, tumor grade and stage, T, N, and M classification.

We further analyzed the predictive value of the RBP signature for chemotherapy and immunotherapy response. Interestingly, our signature could 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, thus, chemotherapy may interfere with the binding of RBPs to RNAs. In turn, the expression of RBPs may influence chemotherapy, leading to drug sensitivity or resistance. The chemotherapy agents considered (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. Variability in responses to treatment and 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 a variety of tumors (47-50). Wang *et al.* identified a prognostic immune gene signature in EC; however, this signature did not predict immunotherapy response (27). Therefore, further studies are needed to predict the therapeutic effect of immunotherapy in EC. Finally, we validated the differential expression of hub genes in EC using clinical samples and IHC staining.

Conclusions

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

Acknowledgments

None.

Footnote

Reporting Checklist: The authors have completed the TRIPOD reporting checklist. Available at https://tcr.amegroups.com/article/view/10.21037/tcr-2024-2561/rc

Data Sharing Statement: Available at https://tcr.amegroups.com/article/view/10.21037/tcr-2024-2561/dss

Peer Review File: Available at https://tcr.amegroups.com/article/view/10.21037/tcr-2024-2561/prf

Funding: This study was supported by the Key Scientific Research Projects of Institutions of Higher Learning in Henan Province (No. 24A320027).

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://tcr.amegroups.com/article/view/10.21037/tcr-2024-2561/coif). The

authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. This study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). All patients provided written informed consent before tissues were collected, and this research was approved by the institutional Ethical Committee of the First Affiliated Hospital of Zhengzhou University (No. 2023-KY-1440-001).

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 Cancer J Clin 2021;71:209-49.
2. Xia W, Liu S, Mao Q, et al. Effect of lymph node examined count on accurate staging and survival of resected esophageal cancer. Thorac Cancer 2019;10:1149-57.
3. Gerstberger S, Hafner M, Tuschl T. A census of human RNA-binding proteins. Nat Rev Genet 2014;15:829-45.
4. Zhang J, Zhang G, Zhang W, et al. Loss of RBMS1 promotes anti-tumor immunity through enabling PD-L1 checkpoint blockade in triple-negative breast cancer. Cell Death Differ 2022;29:2247-61.
5. Zhu H, Chen K, Chen Y, et al. RNA-binding protein ZCCHC4 promotes human cancer chemoresistance by disrupting DNA-damage-induced apoptosis. Signal Transduct Target Ther 2022;7:240.
6. Zhao Y, Mir C, Garcia-Mayea Y, et al. RNA-binding proteins: Underestimated contributors in tumorigenesis. Semin Cancer Biol 2022;86:431-44.
7. Liu J, Cao X. RBP-RNA interactions in the control of autoimmunity and autoinflammation. Cell Res 2023;33:97-115.
8. Tao Y, Zhang Q, Wang H, et al. Alternative splicing and related RNA binding proteins in human health and disease. Signal Transduct Target Ther 2024;9:26.
9. García-Cárdenas JM, Armendáriz-Castillo I, García-Cárdenas N, et al. Data mining identifies novel RNA-binding proteins involved in colon and rectal carcinomas. Front Cell Dev Biol 2023;11:1088057.
10. Zhu J, Li Q, Wu Z, et al. Circular RNA-mediated miRNA sponge & RNA binding protein in biological modulation of breast cancer. Noncoding RNA Res 2024;9:262-76.
11. Chen X, Dong X, Li H, et al. RNA-binding proteins signature is a favorable biomarker of prognosis, immunotherapy and chemotherapy response for cervical cancer. Cancer Cell Int 2024;24:80.
12. Du H, Li Y, Pang S, et al. Development and validation of a prognostic model based on RNA binding proteins in patients with esophageal cancer. J Thorac Dis 2023;15:6178-91.
13. Wang Z, Tang W, Yuan J, et al. Integrated Analysis of RNA-Binding Proteins in Glioma. Cancers (Basel) 2020;12:892.
14. Malta TM, Sokolov A, Gentles AJ, et al. Machine Learning Identifies Stemness Features Associated with Oncogenic Dedifferentiation. Cell 2018;173:338-354.e15.
15. Liu Y, Liu X, Gu Y, et al. A novel RNA binding protein-associated prognostic model to predict overall survival in hepatocellular carcinoma patients. Medicine (Baltimore) 2021;100:e26491.
16. Man Z, Chen Y, Gao L, et al. A Prognostic Model Based on RNA Binding Protein Predicts Clinical Outcomes in Hepatocellular Carcinoma Patients. Front Oncol 2020;10:613102.
17. Zhao Y, Xu L, Wang X, et al. A novel prognostic mRNA/miRNA signature for esophageal cancer and its immune landscape in cancer progression. Mol Oncol 2021;15:1088-109.
18. Zhang HC, Tang KF. Clinical value of integrated-signature miRNAs in esophageal cancer. Cancer Med 2017;6:1893-903.
19. Du H, Xie S, Guo W, et al. Development and validation of an autophagy-related prognostic signature in esophageal cancer. Ann Transl Med 2021;9:317.
20. Du H, Wang X, Xie S, et al. Identification of a prognostic DNA repair gene signature in esophageal cancer. J Gastrointest Oncol 2024;15:829-40.
21. Wang L, Li X, Zhao L, et al. Identification of DNA-Repair-Related Five-Gene Signature to Predict Prognosis

in Patients with Esophageal Cancer. Pathol Oncol Res 2021;27:596899.
22. Xu LC, Pan JX, Pan HD. Construction and Validation of an m6A RNA Methylation Regulators-Based Prognostic Signature for Esophageal Cancer. Cancer Manag Res 2020;12:5385-94.
23. Lan T, Xiao Z, Luo H, et al. Bioinformatics analysis of esophageal cancer unveils an integrated mRNA-lncRNA signature for predicting prognosis. Oncol Lett 2020;19:1434-42.
24. Jin YQ, Miao DL. Multiomic Analysis of Methylation and Transcriptome Reveals a Novel Signature in Esophageal Cancer. Dose Response 2020;18:1559325820942075.
25. Cai W, Li Y, Huang B, et al. Esophageal cancer lymph node metastasis-associated gene signature optimizes overall survival prediction of esophageal cancer. J Cell Biochem 2019;120:592-600.
26. Davis VW, Schiller DE, Eurich D, et al. Urinary metabolomic signature of esophageal cancer and Barrett's esophagus. World J Surg Oncol 2012;10:271.
27. 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. Int Immunopharmacol 2020;87:106795.
28. Zhu M, Cai J, Wu Y, et al. A Novel RNA-Binding Protein-Based Nomogram for Predicting Survival of Patients with Gastric Cancer. Med Sci Monit 2021;27:e928195.
29. Yang L, Zhang R, Guo G, et al. Development and validation of a prediction model for lung adenocarcinoma based on RNA-binding protein. Ann Transl Med 2021;9:474.
30. Xie Y, Luo X, He H, et al. Identification of an individualized RNA binding protein-based prognostic signature for diffuse large B-cell lymphoma. Cancer Med 2021;10:2703-13.
31. Wang M, Jiang F, Wei K, et al. Development and Validation of a RNA Binding Protein-Associated Prognostic Model for Hepatocellular Carcinoma. Technol Cancer Res Treat 2021;20. doi: 10.1177/15330338211004936.
32. Tian J, Ma C, Yang L, et al. Prognostic Value and Immunological Characteristics of a Novel RNA Binding Protein Signature in Cutaneous Melanoma. Front Genet 2021;12:723796.
33. Sheng C, Chen Z, Lei J, et al. Development and Multi-Data Set Verification of an RNA Binding Protein Signature for Prognosis Prediction in Glioma. Front Med (Lausanne) 2021;8:637803.
34. Sha Y, Reyimu A, Liu W, et al. Construction and validation of a prognostic model for esophageal cancer based on prognostic-related RNA-binding protein. Medicine (Baltimore) 2024;103:e39639.
35. 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 2010;10:108.
36. Wu Z, Huang Z, Zhou X, et al. Comprehensive analysis of cuproptosis genes and cuproptosis-related genes as prognosis factors in esophageal squamous cell carcinoma. Genomics 2023;115:110732.
37. Guo P, Niu Z, Zhang D, et al. Potential impact of cuproptosis-related genes on tumor immunity in esophageal carcinoma. Aging (Albany NY) 2023;15:15535-56.
38. 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;9:1747332.
39. 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 Immunol Res 2018;6:645-57.
40. 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. Ann Oncol 2019;30:266-73.
41. Roux C, Mucciolo G, Kopecka J, et al. IL17A Depletion Affects the Metabolism of Macrophages Treated with Gemcitabine. Antioxidants (Basel) 2021;10:422.
42. 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. Dis Esophagus 2014;27:87-92.
43. Oliveira DVNP, Prahm KP, Christensen IJ, et al. Gene expression profile association with poor prognosis in epithelial ovarian cancer patients. Sci Rep 2021;11:5438.
44. Guan Y, Zhu X, Liang J, et al. Upregulation of HSPA1A/HSPA1B/HSPA7 and Downregulation of HSPA9 Were Related to Poor Survival in Colon Cancer. Front Oncol 2021;11:749673.
45. 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. Int J Med Sci 2015;12:256-63.

46. Zhang S, Wang B, Xiao H, et al. LncRNA HOTAIR enhances breast cancer radioresistance through facilitating HSPA1A expression via sequestering miR-449b-5p. Thorac Cancer 2020;11:1801-16.
47. 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. Sci Rep 2020;10:643.
48. Ling B, Ye G, Zhao Q, et al. Identification of an Immunologic Signature of Lung Adenocarcinomas Based on Genome-Wide Immune Expression Profiles. Front Mol Biosci 2020;7:603701.
49. 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 Lett 2020;479:31-41.
50. Zhuang Y, Li S, Liu C, et al. Identification of an Individualized Immune-Related Prognostic Risk Score in Lung Squamous Cell Cancer. Front Oncol 2021;11:546455.

(English Language Editor: J. Gray)

Cite this article as: Sun S, Wang J, Zhang Y, Li Y, Guo Y, Huang C, Tartarone A, Pallante P, Li K, Zhang G, Pan X, Li X. Genome-wide profiling of a prognostic RNA-binding protein signature in esophageal cancer. Transl Cancer Res 2025;14(2):1428-1446. doi: 10.21037/tcr-2024-2561