

Identification of Prognosis-Related RBPs to Reveal The Role of RNA Binding Proteins in the Progression and Prognosis of Wilms' Tumor

Xuejiao Qi (✉ uro_weixiyi@163.com)

Department of Paediatrics, Qingdao Eighth People's Hospital, No. 84 Fengshan Road, Licang District, Qingdao, 266100

Shuyu Wang

First Clinical Medical College, Nanjing Medical University, Nanjing

Yihui Dong

Department of Paediatrics, Qingdao Eighth People's Hospital, No. 84 Fengshan Road, Licang District, Qingdao, 266100

Xiaojie Lin

Department of Paediatrics, Qingdao Eighth People's Hospital, No. 84 Fengshan Road, Licang District, Qingdao, 266100

Jingqiu Chen

Department of Paediatrics, Licang District Central Hospital, Licang District, Qingdao, 266100

Research Article

Keywords: RNA binding proteins (RBPs), prognosis, risk model, Wilms' tumor

Posted Date: January 27th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-135539/v1>

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

[Read Full License](#)

Abstract

Background: Despite the various key functions of RBPs in posttranscriptional events, the mechanism of their influence on Wilms' tumor has not been well elucidated. Therefore, we constructed the research to identify several RBPs related to Wilms' tumor progression and prognosis, for the better understanding of RBPs' role in the occurrence and development of Wilms' tumor, and to provide effective reference targets for new drug development.

Methods: A total of 127 samples of different clinical characteristics including gender, race and stage were selected from TARGET to carry out our study. After the gene functional enrichment pathways, univariate Cox regression analysis and lasso regression analysis were performed to test the prognostic effect of the differentially-expressed genes and establish the prognostic index. Further Cox regression analyses were utilized to identify the independence of our model and to analyze the relationship between our model and clinical parameters. What's more, gene set enrichment analysis (GSEA) was also performed to elucidate the biological characteristics of genes involved in Wilms' tumor. $P < 0.05$ was considered to be statistically significant.

Results: 20 RBPs were statistically correlated with Wilms' tumor. After the construction of a prognostic index, patients were divided into high-and low-risk scores group. Kaplan-Meier (K-M) analyses showed that patients with high risk scores possessed poorer survival probability than patients with low risk scores in both training group and test group. Furthermore, multivariate Cox regression analysis explored the relationship between our prognostic model and clinical parameters and confirmed that our model was an independent predicted factor for Wilms' tumor.

Conclusion: Our study clarifies the application of RBPs in the prognosis of Wilms' tumor. We are confident that our risk scoring model can provide ideas for the development of new targets for broad-spectrum anticancer drugs and has great potential in clinical practice.

Introduction

Wilms' tumor (WT), the most common embryonal malignancy of the genitourinary system in children, is an embryonic kidney tumour divided into sporadic and syndrome-associated types[3]. WT can occur bilaterally, and more than 12% of WT patients have different types of congenital diseases, such as testicular insufficiency, hypospadias, hemianomegaly, and iris absence. Additionally, the proportion of congenital malformations associated with nephroblastoma accounts for 8%-17%[4]. Nephroblastoma has a high incidence in childhood and ranks second among abdominal malignancies in children, with clinical evidence that 98% of cases occur below the age of 10 years. However, with poor prognosis, easy recurrence, and difficulty to diagnose in early stage, some patients still die due to drug resistance, recurrence, and tumor metastasis despite the improvement of overall survival[5]. Conclusively, the related genes of high-risk nephroblastoma prognosis is of great significance to study.

RNA-binding proteins (RBPs) serve as the the key regulators of RNA , as except for a few RNAs that can function alone in the form of ribozymes, the majority of them are combined with proteins to form RNA-protein complexes where RBPs regulate alternative splicing, modification, transport and translation and other life activities of RNA[6]. So studying the interaction between RNA and protein is the key to exploring RNA functions, considering that modifications in RNA structure can all lead to changes in the RBP bound to it, resulting in different biological functions[2]. As a result, RBPs may be a novel promising bio-markers for cancer patients[7].

Here, we described the expression and lineage of RBPs in Wilms'tumor, and investigate the mechanisms of RBPs to the development and prognosis of Wilms'tumor.

Methods

Data collection

RNA sequence data of Wilmes' tumor was downloaded from Therapeutically Applicable Research To Generate Effective Treatments (TARGET) database (<https://ocg.cancer.gov/programs/target>). Clinical traits such as stage, gender, race and survival time were also included in our research. A total of 1542 RBPs identified by highthroughput screening in human cells were obtained from published research which represented 7.5% of all protein coding genes.

Functional enrichment pathways

With fold change>2 and P value <0.05 as a criterion, 20 differentially expressed genes were finally obtained based on the comparison between tumor tissues and normal tissues by utilizing EdgeR package in R studio. Then, for the purpose of indentifying the biological characteristics of these differentially expressed genes, these 20 genes went through functional enrichment analyses including Kyoto Encyclopedia of Genes and Genomes (KEGG) and gene ontology (GO). We utilized Database for Annotation, Visualization, and Integrated Discovery (DAVID, [https:// david.ncifcrf.gov/](https://david.ncifcrf.gov/)) to identify enriched KEGG and GO themes.

Construction of an independent prognostic model

After the identification of gene function, univariate Cox regression analysis was performed to figure out which gene expression were associated with the overall survival of children with Wilms' tumor. Finally, an independent prognostic index (PI) was established. Utilizing the median PI value as the cutoff value, patients were separated into a low- and high-risk group to carry out the further study. Kaplan-Meier (K-M) analyses were then conducted to compare the overall survival of the low- and high- risk groups and log-rank test was applied to test the differences between these two groups. Finally, receiver operating characteristic curves (ROCs) were designed to verify the prediction value of the model.

Statistical analysis

We carried out all statistical analyses using R Studio 3.6.1 (<https://www.r-project.org/>). The independent t test was used for continuous variables with normal distribution, and the Mann-Whitney U test was used for continuous variables with skewed distribution. A two-sided test was used, and a P value of <0.05 was considered statistically significant. The overall survival analyses were performed by the Kaplan-Meier method and log-rank test. The ROC curves were also created to evaluate the predictive ability of our risk signature. Univariate and multivariate Cox regression analyses were utilized to analyze the effects of the prognostic index and identify the independence of our model.

Results

1. *Validation of differentially expressed genes in Wilms' tumor*

A total of 127 samples of different clinical characteristics including gender, race and stage were selected from TARGET to carry out our study (Table 1). Using EdgeR package in R studio, we compared the differentially expressed genes between tumor tissues and normal tissues. With $P < 0.05$ and fold change > 2 as the cut off value, volcano map and heatmap were applied to visualize the differentially expressed genes (Figure 1A-B) and 20 differentially expressed RBPs were finally obtained based on the comparison between tumor tissues and normal tissues by utilizing EdgeR package in R studio.

2. *Functional annotation of these differentially expressed genes*

In order to further clarify the biological attributes of 20 differentially expressed RBPs between tumor tissues and normal tissues, Kyoto Encyclopedia of Genes and Genomes (KEGG) and gene ontology (GO) analysis were carried out. The GO analysis revealed that the most significant GO term was "Ribosome", followed by "RNA transport", "Spliceosome", "RNA degradation", "Ribosome biogenesis in eukaryotes" and "mRNA surveillance pathway". According to the KEGG pathway analysis, however, the most significantly enriched pathway was "RNA catabolic process" (Figure 2).

3. *Establishment of a prognostic signature*

To deeply explore the associations between these differentially expressed RBPs and overall survival, univariate Cox regression analysis was performed (Figure 3A). Then, lasso regression analysis was conducted to increase the robustness and select the optimal variables based on training set. Finally, we got 9 genes for the construction of our prognostic index (Figure 3B-C). After the establishment of our prognostic signature, patients were then classified into a high-risk group and a low-risk group based on the risk scores (Figure 3D). The results indicated that survival years decreased as the risk scores increased (Figure 3E). Heatmap was also utilized to visualize the different gene expression pattern between high- and low-risk groups (Figure 3F).

Based on the results of K-M analyses, we came to the conclusion that patients with high risk scores possessed poorer survival probability than patients with low risk scores in both training group and test group ($P = 2.337e-03$, $2.118e-03$, $9.921e-06$ respectively) (Figure 4A, Figure 4B, Figure 4C). The ROC curves

were also created to evaluate the predictive ability of our model. According to the ROC curve, it was proved that our prognostic index had a good sensitivity and specificity (AUC=0.765 and 0.665 for 5 years overall survival in training and validation group, respectively) (Figure 4D, Figure 4E, Figure 4F)

4. Relationship between risk score and clinical parameters

To explore whether the constructed risk signature was independent from gender, race and stage, univariate and multivariate Cox regression analyses were both performed. According to the results, only risk score worked as an independent predicted factor (Figure 5A-B). What's more, our risk score as well as clinical parameters including stage, gender and race was integrated into a nomogram to visualize the 3- and 5-year survival probability of patients with Wilms' tumor (Figure 5C, Figure 5D, Figure 5E).

We tried to further investigate the relationship between our prognostic index and clinical parameters such as gender, race and stage. Based on the results, gender and race showed no statically significant influence on the risk score ($p=0.55$ and 0.62 respectively) while stage seemed to be related with risk score ($p=0.038$) (Figure 6A-C)

5. Gene set enrichment analysis of risk scores

To explore the biological relevance of risk scores involved in progression of Wilms' tumor, we carried out a gene set enrichment analysis (GSEA) of risk scores based on the TCGA Wilms' tumor cohort. GSEA analysis indicated high risk scores was associated with PI3K_AKT_MTOR_SIGNALING, MTORC1_SIGNALING, MYC_TARGETS_V1 and G2M_CHECKPOINT pathway (Figure 7A, Figure 7B, Figure 7C, Figure7D).

Discussion

The changes in the RBPs bound to RNA are caused by the modification of structure and the change of its spatial structure, further leading to different biological functions, such as RNA splicing, mRNA stabilization and protein translation[1, 2, 6-8]. Because of their various key functions in post-transcriptional events and regulation on physiological events of cells, the changes in RBPs are related to the occurrence and development of many human diseases[2]. Although RBPs are reported to show dysregulated expression in various human cancers, little is currently known about the expression patterns and roles of RBPs in Wilms' tumor. Despite the high overall survival rate of children with Wilms' tumor, the overall prognosis of children with stage IV disease is poor[9].

Scholars have been studying RBPs recently. A applied high-throughput screening to identified 1542 RBPs affecting posttranscriptional events and regulate physiological events of cells, accounting for 7.5% of all protein coding genes[10]. Bin Zhang established a comprehensive expression pattern of RBPs across different cancer types, where RBP is always dysregulated, proving that RBPs may be a new target for the development of broad-spectrum anticancer drugs[1]. And it has been confirmed that some genes, such as P53, ki-67, etc. play an important role in the development of Wilms' tumor[11].

We construct a detailed assessment of RBPs in Wilms' tumor, based on the data from a total of 127 samples of different clinical characteristics including gender, race and stage from TCGA. After analyzing the RBPs between Wilms' tumor and normal tissues, we constructed univariate Cox regression and lasso regression analysis to increase the robustness and select the optimal variables based on training set, eventually getting 9 genes for the construction of our prognostic index. Patients were then classified into a high-risk group and a low-risk group based on the risk scores, and as expected, the results indicated that survival years decreased as the risk scores increased. Then based on the results of K-M analyses, we came to the conclusion that patients with high risk scores possessed poorer survival probability than patients with low risk scores in both training group and test group. Furthermore, we performed the ROC curve to investigate the prognostic value of the model, which proved that our prognostic indicators have good sensitivity and specificity. Univariate and multivariate Cox regression analyses also show that risk signature was independent from gender, race and stage. GSEA analysis indicated high risk scores was associated with PI3K_AKT_MTOR_SIGNALING, MTORC1_SIGNALING, MYC_TARGETS_V1 and G2M_CHECKPOINT pathway.

Cummings found that p53-mediated downregulation of Chk1 inhibit g2m checkpoint induced by DNA damage in K562 cells, leading to increased apoptosis[12]; the deregulation of the G2M checkpoint in myeloid leukaemic cell lines were also found by Higginbottom to result in loss of cell survival[13]. Zhu found that Gomisin N play an anti-hepatoma role in vitro by regulating the PI3K-AKT and mtor-ulk1 pathways[14]. But the MTORC1_SIGNALING and MYC_TARGETS_V1 pathway has rarely, if ever, been studied.

However, the limitations in our research cannot be ignored. First of all, our study sample size is small, considering which a larger queue and more abundant sequencing results were suggested. Secondly, we focus only on the level of gene expression and mutation, ignoring other events important in tumor progression, such as the amplification of copy number and the methylation of gene .Finally, the in vitro verification about the effect of RBPS on the progression and prognosis of Wilms' tumor needs to carried out. What's more, since the prognosis of adult nephroblastoma treated with a regimen for children is unsatisfactory, the effect of RBPS on adult nephroblastoma should be studied separately[15-17].

Conclusions

In summary, our study clarifies the application of RBPs in the prognosis of Wilms' tumor. The constructed RBPs risk scoring model, which is an independent factor influencing the prognosis of Wilms' tumor, can reliably predict the prognosis of Wilms' tumor. We are confident that our risk scoring model can provide ideas for the development of new targets for broad-spectrum anticancer drugs and has great potential in clinical practice.

Abbreviations

RNA-binding proteins (RBPs)

The Cancer Genome Atlas (TCGA)

gene set enrichment analysis (GSEA)

Kaplan-Meier (K-M)

Wilms'tumor (WT)

Kyoto Encyclopedia of Genes and Genomes (KEGG)

gene ontology (GO)

receiver operating characteristic curves (ROCs)

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

All data generated or analysed during this study are included in this published article and its supplementary information files. The datasets generated and analysed during the current study are available in the Therapeutically Applicable Research To Generate Effective Treatments (TARGET) database repository, <https://ocg.cancer.gov/programs/target>). We utilized Database for Annotation, Visualization, and Integrated Discovery (DAVID, <https://david.ncifcrf.gov/>) to identify enriched KEGG and GO themes. We carried out all statistical analyses using R Studio 3.6.1 (<https://www.r-project.org/>).

Competing interests

The authors declare no competing interests.

Funding

None.

Authors contributions

Xuejiao Qi designed this work. Xuejiao Qi and Shuyu Wang wrote the manuscript. Yihui Dong and Jingqiu Chen performed the bioinformatics analysis. Xiaojie Lin performed the data review. All authors have read and approved the manuscript.

Acknowledgments

None.

References

1. Zhang B, Babu KR, Lim CY, Kwok ZH, Li J, Zhou S, Yang H, Tay Y: **A comprehensive expression landscape of RNA-binding proteins (RBPs) across 16 human cancer types** %J *RNA Biology*. 2020, **17**(2):211-226.
2. Lukong KE, Chang K-w, Khandjian EW, Richard S: **RNA-binding proteins in human genetic disease**. *Trends Genet* 2008, **24**(8):416-425.
3. Pritchard-Jones K, Fleming S, Davidson D, Bickmore W, Porteous D, Gosden C, Bard J, Buckler A, Pelletier J, Housman D: **The candidate Wilms' tumour gene is involved in genitourinary development**. *Nature* 1990, **346**(6280):194-197.
4. Dumoucel S, Gauthier-Villars M, Stoppa-Lyonnet D, Parisot P, Brisse H, Philippe-Chomette P, Sarnacki S, Boccon-Gibod L, Rossignol S, Baumann C *et al*: **Malformations, Genetic Abnormalities, and Wilms Tumor**. *Pediatr Blood Cancer* 2014, **61**(1):140-144.
5. Beckwith JB, Palmer NF: **Histopathology and prognosis of Wilms tumors: results from the First National Wilms' Tumor Study**. *Cancer* 1978, **41**(5):1937-1948.
6. Kechavarzi B, Janga SC: **Dissecting the expression landscape of RNA-binding proteins in human cancers**. *Genome Biol* 2014, **15**(1).
7. Castello A, Fischer B, Hentze MW, Preiss T: **RNA-binding proteins in Mendelian disease**. *Trends Genet* 2013, **29**(5):318-327.
8. Castello A, Fischer B, Eichelbaum K, Horos R, Beckmann BM, Strein C, Davey NE, Humphreys DT, Preiss T, Steinmetz LM *et al*: **Insights into RNA Biology from an Atlas of Mammalian mRNA-Binding Proteins**. *Cell* 2012, **149**(6):1393-1406.
9. Raffensperger J: **Max Wilms and his tumor**. *J Pediatr Surg* 2015, **50**(2):356-359.
10. Gerstberger S, Hafner M, Tuschl T: **A census of human RNA-binding proteins**. *NATURE REVIEWS GENETICS* 2014, **15**(12):829-845.
11. Ooms AHAG, Gadd S, Gerhard DS, Smith MA, Auvil JMG, Meerzaman D, Chen Q-R, Hsu CH, Yan C, Nguyen C *et al*: **Significance of TP53 Mutation in Wilms Tumors with Diffuse Anaplasia: A Report from the Children's Oncology Group**. *Clin Cancer Res* 2016, **22**(22):5582-5591.
12. Cummings M, Siitonen T, Higginbottom K, Newland AC, Allen PD: **p53-mediated downregulation of Chk1 abrogates the DNA damage-induced G2M checkpoint in K562 cells, resulting in increased apoptosis**. *Br J Haematol* 2002, **116**(2):421-428.
13. Higginbottom K, Cummings M, Newland AC, Allen PD: **Etoposide-mediated deregulation of the G2M checkpoint in myeloid leukaemic cell lines results in loss of cell survival**. *Br J Haematol* 2002, **119**(4):956-964.

14. Zhu P-L, Lam DF, Li J-K, Fu X-Q, Yin C-L, Chou J-Y, Wang Y-P, Liu Y-X, Chen Y-J, Wu J-Y *et al*: **Gomisin N Exerts Anti-liver Cancer Effects and Regulates PI3K–Akt and mTOR–ULK1 Pathways in Vitro** %J *Biological and Pharmaceutical Bulletin* %J *Biological and Pharmaceutical Bulletin*. 2020, **43**(8):1267-1271.
15. Mundlos S, Pelletier J, Darveau A, Bachmann M, Winterpacht A, Zabel B: **Nuclear localization of the protein encoded by the Wilms' tumor gene WT1 in embryonic and adult tissues**. *Development* 1993, **119**(4):1329-1341.
16. Morris MR, Hesson LB, Wagner KJ, Morgan NV, Astuti D, Lees RD, Cooper WN, Lee JA, Gentle D, Macdonald F *et al*: **Multigene methylation analysis of Wilms' tumour and adult renal cell carcinoma**. *Oncogene* 2003, **22**(43):6794-6801.
17. Byrd RL, Evans AE, D'Angio GJ: **Adult Wilms tumor: effect of combined therapy on survival**. *The Journal of urology* 1982, **127**(4):648-651.

Table

Table 1. Clinical characteristics of included patients in the study

Variables	Total (n=127)	Training cohort (n=66)	Validation cohort (n=61)
Gender			
FEMALE	71	42	29
MALE	54	22	32
Race			
Black	19	6	13
Other	5	2	3
Unknow	9	4	5
White	92	52	40
Stage			
I	16	10	6
II	49	23	26
III	46	23	23
IV	14	8	6

Figures

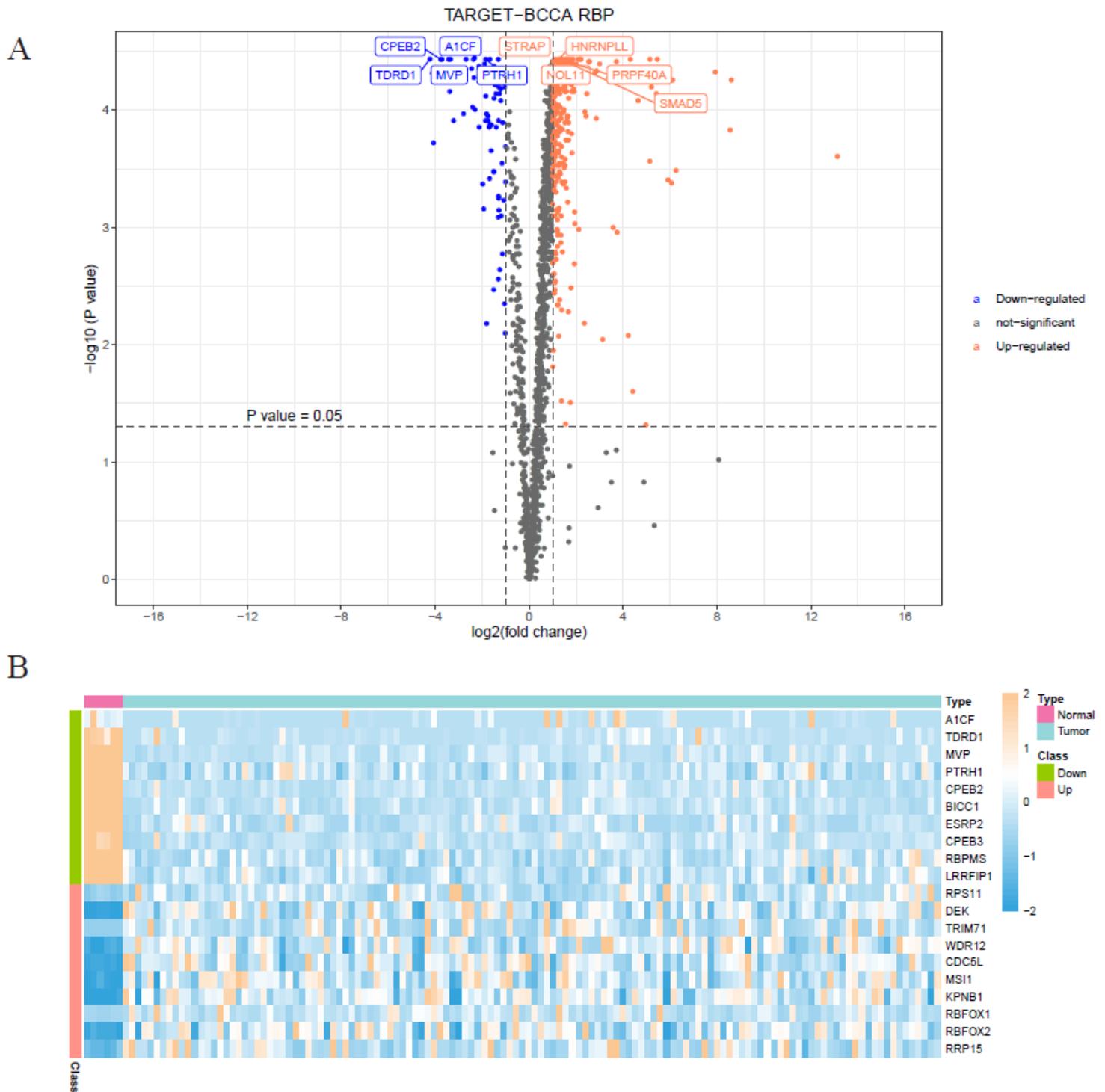


Figure 1

Identification of 20 DERBPs. (A) volcano plots of DERBPs in Wilms' tumor and normal tissues from TCGA database. (B) Heatmap plots of top 10 up-regulated and top 10 down-regulated DERBPs.

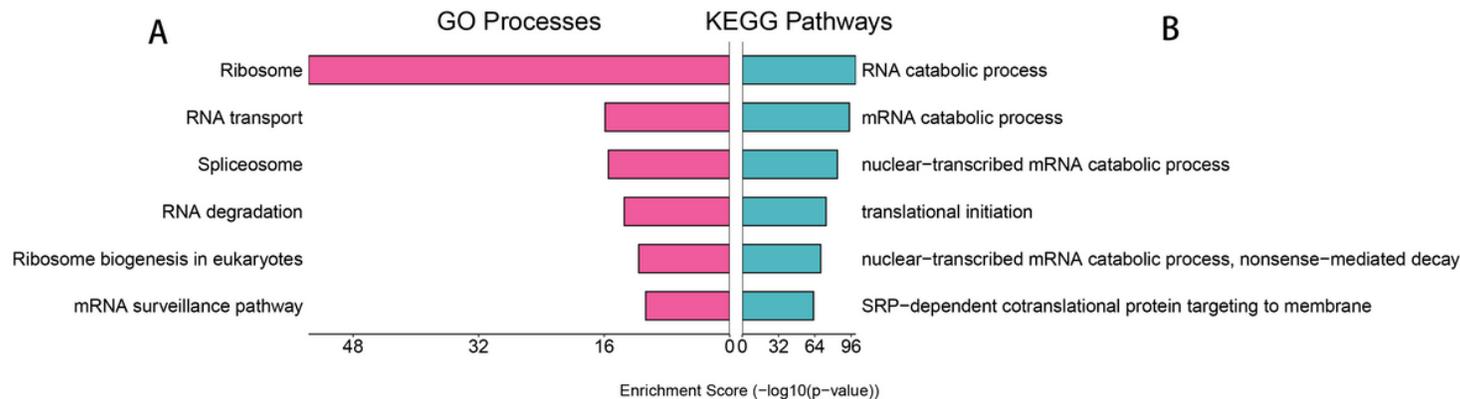


Figure 2

GO (A) and KEGG(B) enrichment analysis of 20 DERBPs.

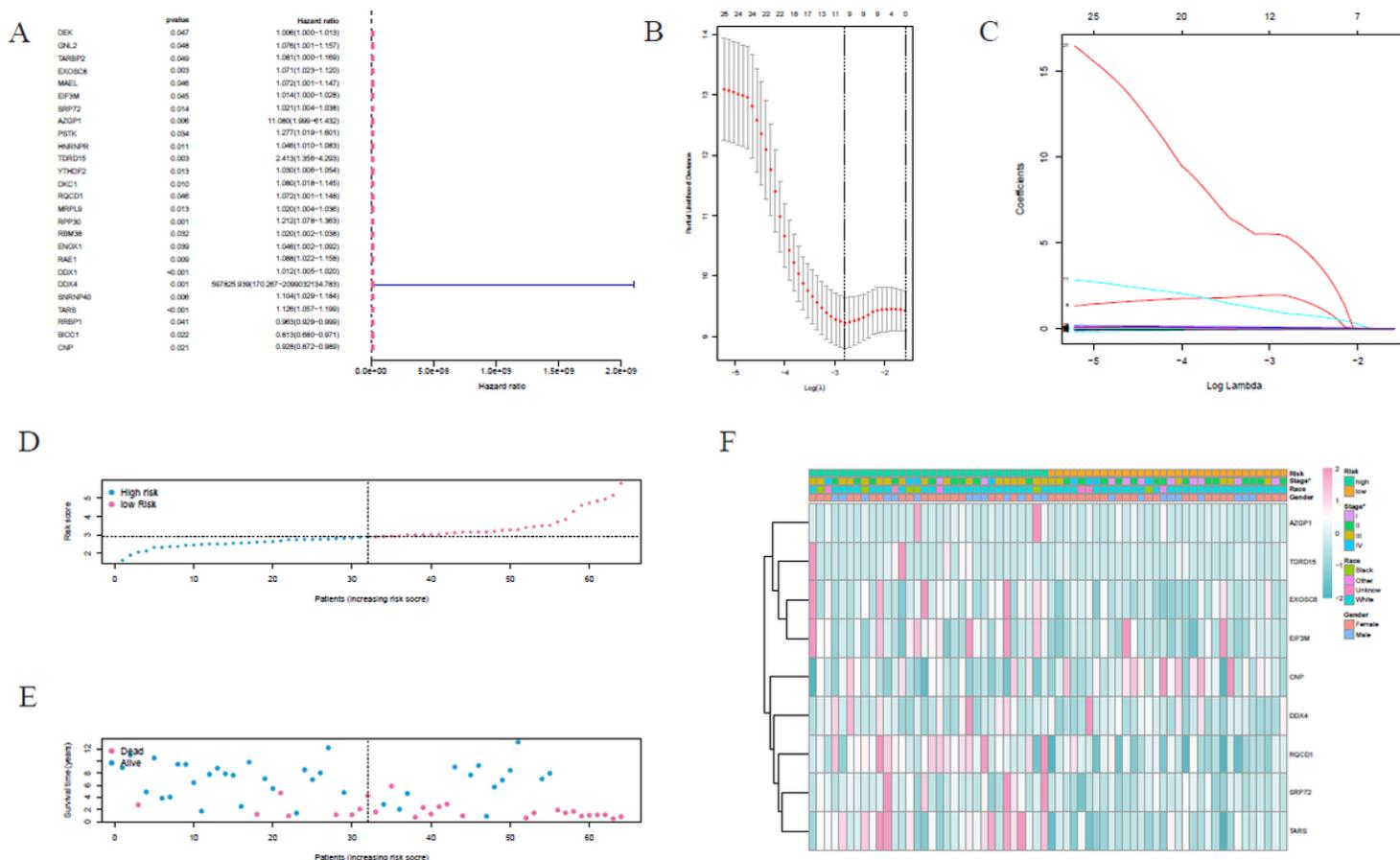


Figure 3

Univariate survival analysis and Multivariate cox regression model (A) Univariate survival analysis by cox proportional hazards models. (B) Partial Likelihood Deviance. (C) Coefficients. (D) Distribution of immune risk scores in breast cancer patients. (E) Distribution of survival status in colon cancer patients. (F) Distribution of specific risk factors in the high- and low-risk groups (divided by median value). (*P<0.05)

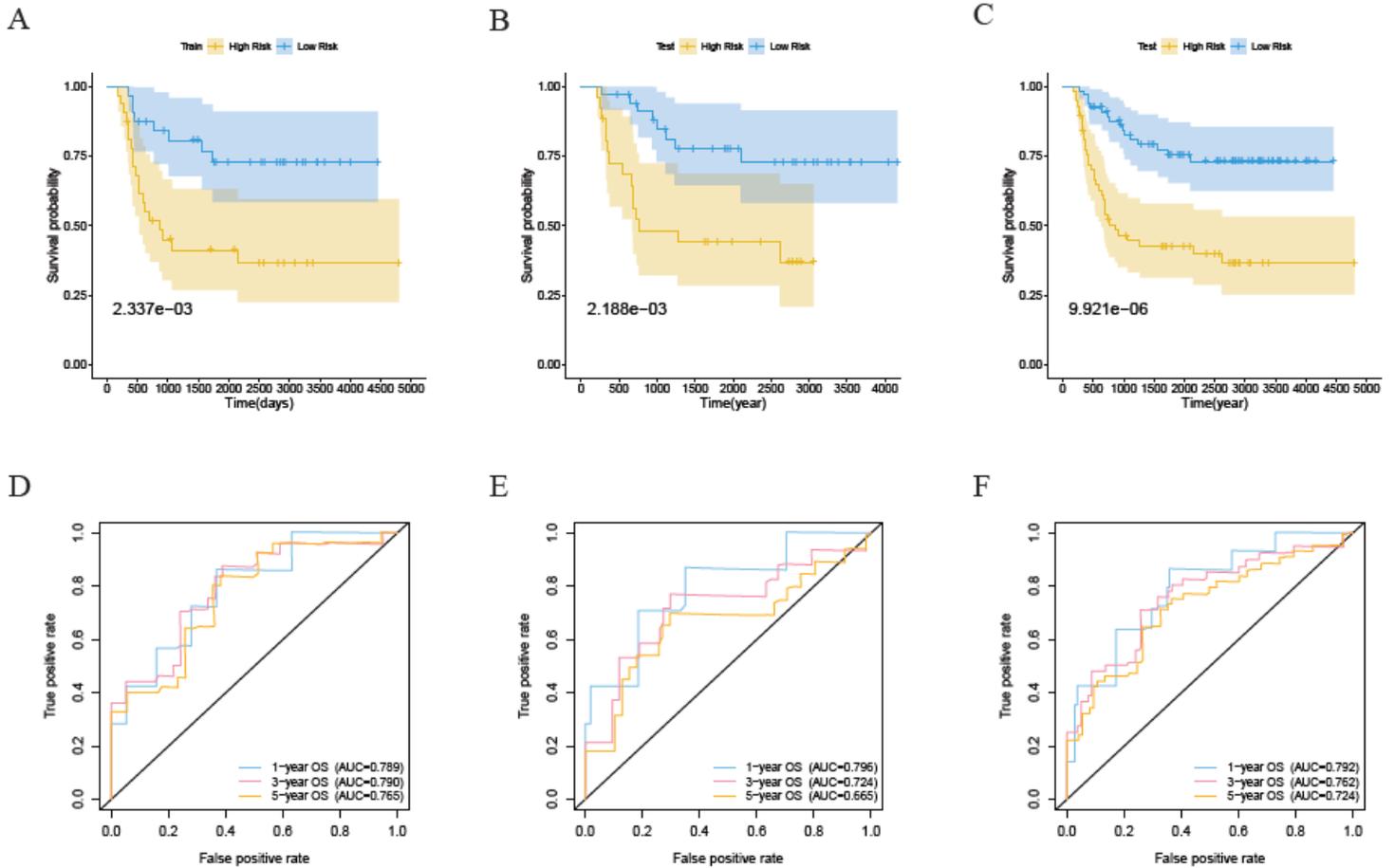


Figure 4

Kaplan-Meier curve analysis of high-risk and low-risk patients and Time dependent ROC curve analysis. (A)Kaplan-Meier curve analysis of high-risk and low-risk patients in the training cohort. (B) Kaplan-Meier curve analysis of high-risk and low-risk patients in the testing cohort.(C) Kaplan-Meier curve analysis of high-risk and low-risk patients in the entire TCGA cohort. (D)Timedependent ROC curve analysis of the training cohort. (E) Timedependent ROC curve analysis of the testing cohort. (F) Time-dependent ROC curve analysis of the entire TCGA cohort.

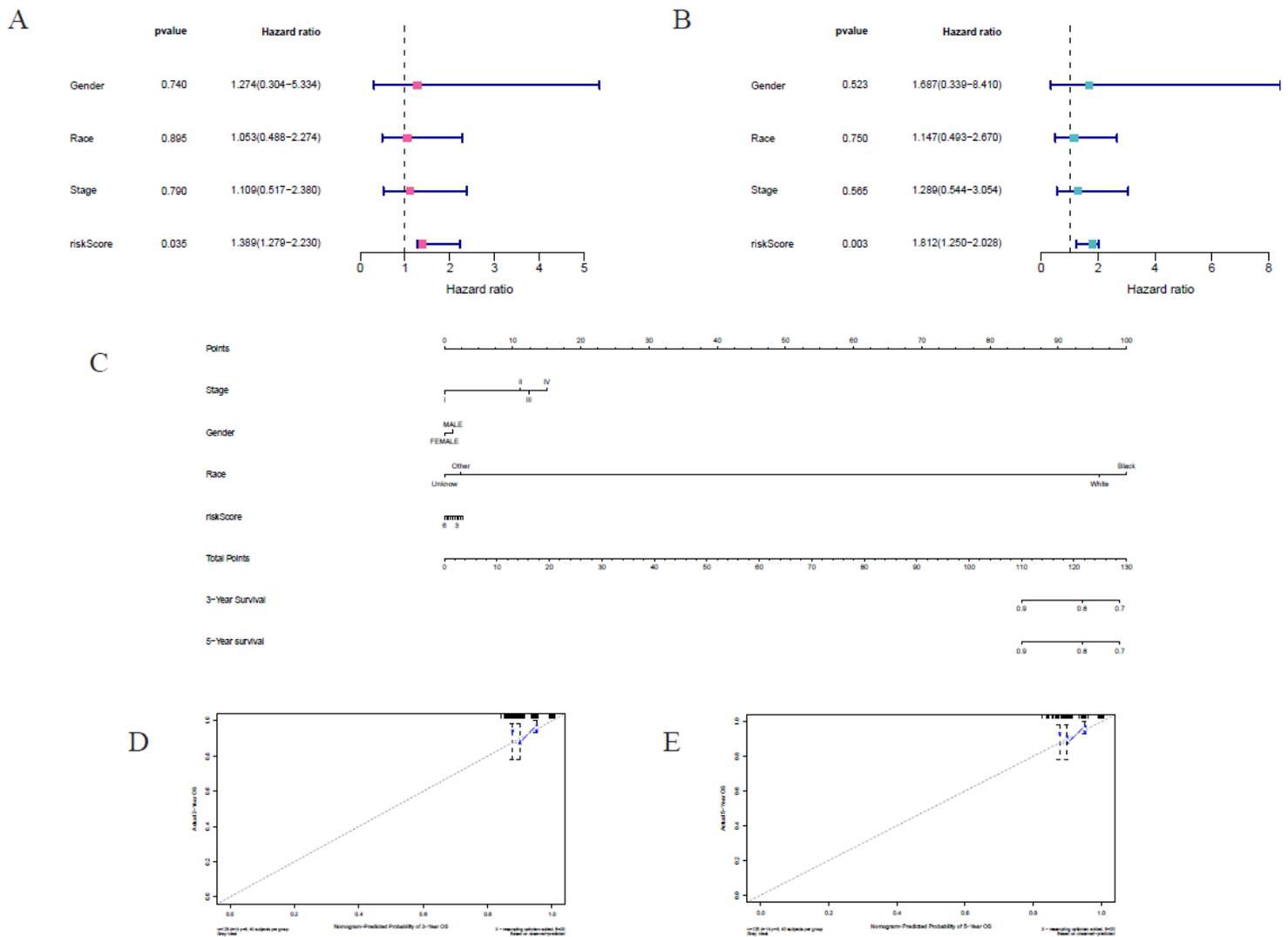
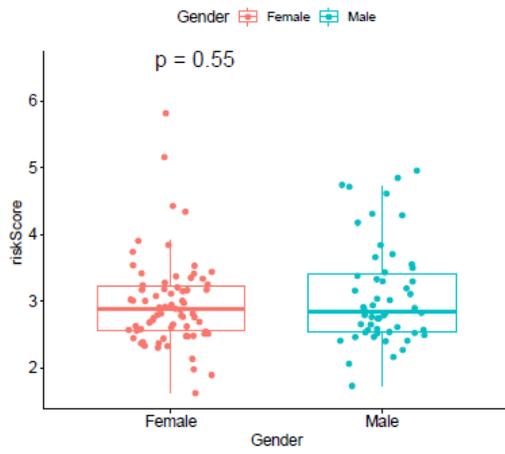


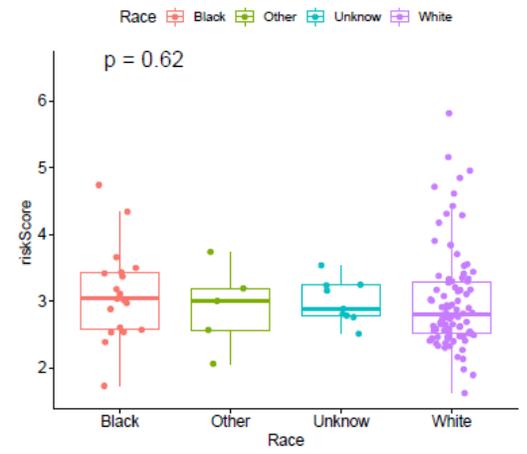
Figure 5

Cox's proportional hazard model of correlative factors in Wilms' tumor patients.(A)Univariate COX regression analysis for clinicopathological parameters affecting the overall survival.(B) Multivariate COX regression analysis for clinicopathological parameters affecting the overall survival. (C) An established nomogram to predict Wilms' tumor survival based on cox model. (D)Nomogram-Predicted Probability of 3-Year OS and Actual 3-Year OS. (E)Nomogram-Predicted Probability of 3-Year OS and Actual 3-Year OS.

A



B



C

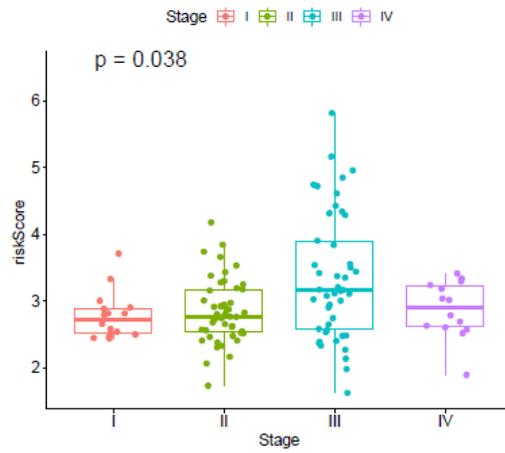


Figure 6

Correlation between RBPs risk scores and various clinical factors. (A) Gender. (B) Race. (C) Stage.

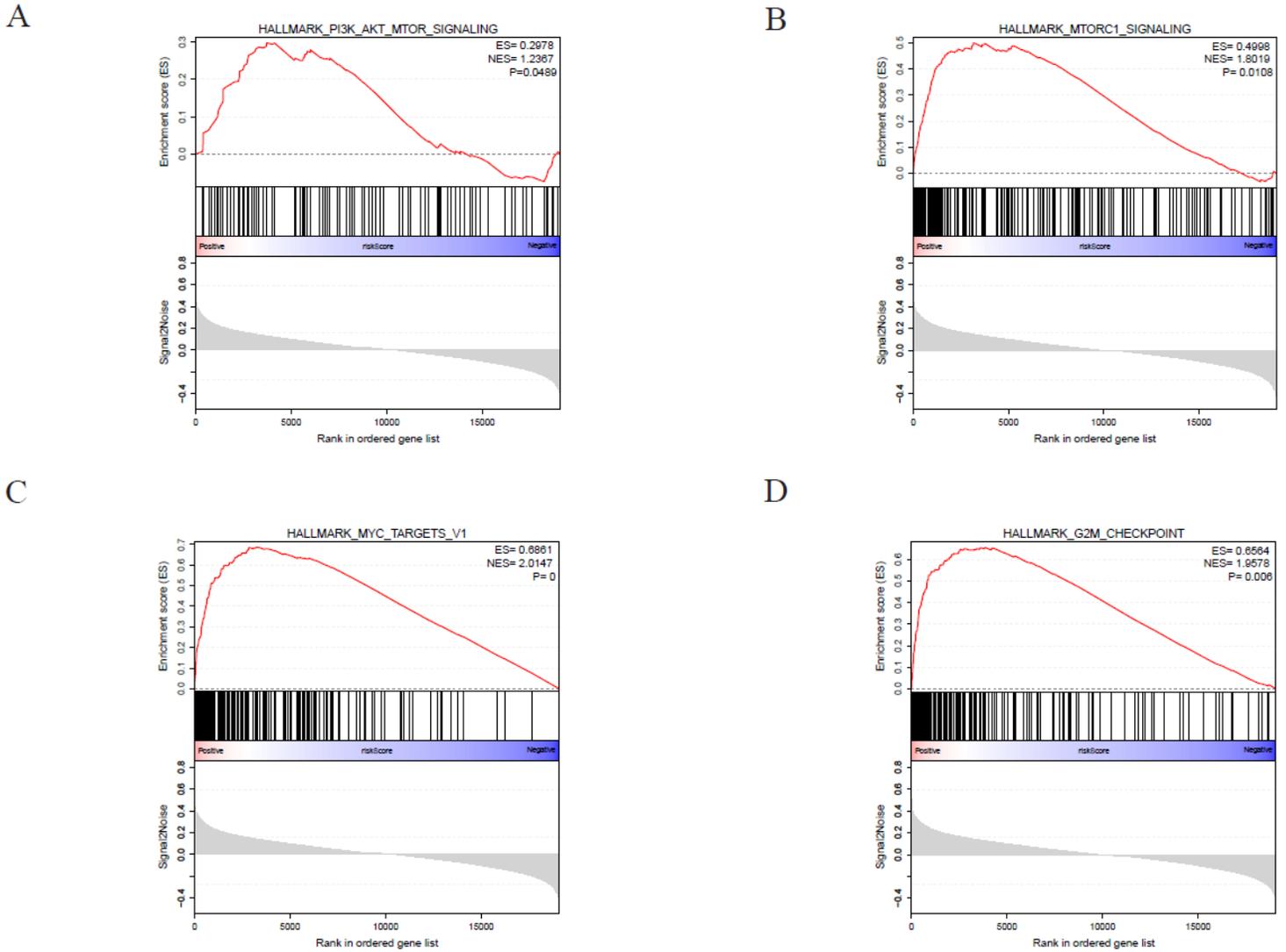


Figure 7

Gene set enrichment analysis (GSEA) of risk scores based on the TCGA Wilms' (A) HALLMARK_PI3K_AKT_MTOR_SIGNALING. (B) HALLMARK_MTORC1_SIGNALING. (C) HALLMARK_MYC_TARGETS_V1. (D) HALLMARK_G2M_CHECKPOINT