Original Article

Prognostic model construction and immune microenvironment analysis of esophageal cancer based on gene expression data and microRNA target genes

Bingbing Gu[1,2]^, Shuai Zhang[1,2], Zhe Fan[1,2], Jiajing Che[1,2], Shuting Li[1,2], Yunfei Li[1,2], Keyu Pan[1,2], Jiali Lv[1,2], Cheng Wang[1,2], Tao Zhang[1,2], Jialin Wang[1,3]

[1]Department of Biostatistics, School of Public Health, Cheeloo College of Medicine, Shandong University, Jinan, China; [2]Institute for Medical Dataology, Cheeloo College of Medicine, Shandong University, Jinan, China; [3]Shandong Cancer Hospital and Institute, Shandong First Medical University and Shandong Academy of Medical Sciences, Jinan, China

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

Correspondence to: Tao Zhang, MD, PhD. Department of Biostatistics, School of Public Health, Cheeloo College of Medicine, Shandong University, PO Box 100, 44 Wenhuaxi Rd., Jinan 250012, China. Email: taozhang@sdu.edu.cn; Jialin Wang. Shandong Cancer Hospital and Institute, Shandong First Medical University and Shandong Academy of Medical Sciences, 440 Jiyan Road, Jinan 250117, China. Email: wangjialin6681@gmail.com.

Background: Accumulating evidence suggests that microRNA-target genes are closely related to tumorigenesis and progression. This study aims to screen the intersection of differentially expressed mRNAs (DEmRNAs) and the target genes of differentially expressed microRNAs (DEmiRNAs), and to construct a prognostic gene model of esophageal cancer (EC).

Methods: Gene expression, microRNA expression, somatic mutation, and clinical information data of EC from The Cancer Genome Atlas (TCGA) database were used. The intersection of DEmRNAs and the target genes of DEmiRNAs predicted by the Targetscan database and microRNA Data Integration Portal (mirDIP) database were screened. The screened genes were used to construct a prognostic model of EC. Then, the molecular and immune signatures of these genes were explored. Finally, the GSE53625 dataset from the Gene Expression Omnibus (GEO) database was further used as a validation cohort to confirm the prognostic value of the genes.

Results: Six genes on the grounds of the intersection of DEmiRNAs target genes and DEmRNAs were identified as prognostic genes, including *ARHGAP11A*, *H1.4*, *HMGB3*, *LRIG1*, *PRR11*, and *COL4A1*. Based on the median risk score calculated for these genes, EC patients were divided into a high-risk group (n=72) and a low-risk group (n=72). Survival analysis showed that the high-risk group had a significantly shorter survival time than the low-risk group (TCGA and GEO, P<0.001). The nomogram evaluation showed high reliability in predicting the 1-year, 2-year, and 3-year survival probability of EC patients. Compared to low-risk group, higher expression level of M2 macrophages was found in high-risk group of EC patient (P<0.05), while *STAT3* checkpoints showed attenuated expression level in high-risk group.

Conclusions: A panel of differential genes was identified as potential EC prognostic biomarkers and showed great clinical significance in EC prognosis.

Keywords: Esophageal cancer (EC); prognostic model; immune microenvironment; expression data; microRNA target genes

^ ORCID: 0000-0002-9240-703X.

Submitted Nov 10, 2022. Accepted for publication Mar 30, 2023. Published online Apr 13, 2023.
doi: 10.21037/tcr-22-2588
View this article at: https://dx.doi.org/10.21037/tcr-22-2588

Introduction

Esophageal cancer (EC) ranks the seventh common cancer in terms of incidence (572,000 new cases) and the sixth most common cause of death from cancer worldwide in 2,018 (509,000 deaths) (1). Due to the insidious onset, highly invasive properties, and rapid progress, most EC patients are diagnosed at an advanced stage. The prognosis of EC patients is poor (2). Consequently, there are still in urgent need for more reliable and non-invasive prognostic biomarkers to guide EC personalized treatment.

Recent studies has proved dysregulation of specific microRNAs could contribute to metaplastic and neoplastic processes in the oesophageal mucosa. MicroRNAs are short, non-coding RNAs regulating gene expression. Li *et al.* found microRNAs-target gene played an important role in cancer initiation and progression (3). Although the contribution of microRNAs to EC development has been extensively studied, little has been done on microRNA targeting genes in EC. Previous studies showed tumor microenvironment (TME) may play a crucial role in the tumorigenesis and progression of EC. The TME includes a complex collection of components to keep a dynamic balance exists between the pro- and anti-tumor factors within the TME, which profoundly influences the prognosis of patients with cancer (4-7).

This study aimed to identify a panel of potential prognostic genes according to the intersection of differentially expressed mRNAs (DEmRNAs) and the target genes of differentially expressed microRNAs (DEmiRNAs), and to explore the prognostic value of these genes. We further clarified the immunological roles of prognostic genes among EC patients, therefore providing a comprehensive insights into esophageal tumor development. We present this article/case in accordance with the TRIPOD reporting checklist (available at https://tcr.amegroups.com/article/view/10.21037/tcr-22-2588/rc).

Highlight box

Key findings

- A prognostic model consisting of six genes (*ARHGAP11A*, *H1.4*, *HMGB3*, *LRIG1*, *PRR11*, and *COL4A1*) is a strong predictor of esophageal cancer.

What is known and what is new?

- Dysregulation of specific microRNAs could contribute to metaplastic and neoplastic processes in the oesophageal mucosa and tumor microenvironment (TME) may play a crucial role in the tumorigenesis and progression of esophageal cancer.
- This study combined intersection of differentially expressed mRNAs (DEmRNAs) and the target genes of differentially expressed microRNAs (DEmiRNAs) to construct a prognostic model and found that the new model correlated with prognosis with esophageal cancer.

What is the implication, and what should change now?

- This study highlights the importance of microRNA target genes in clinical practice and implies more accurate diagnosis and prognostic prediction for EC patients.

Methods

Data downloading and processing

Gene expression data included 160 tumor samples and 11 normal samples, microRNA data included 185 tumor samples and 13 normal samples, 366 somatic mutation data and corresponding clinical data were collected from The Cancer Genome Atlas (TCGA) dataset (8) as the training cohort. Furthermore, The GSE53625 dataset from the Gene Expression Omnibus (GEO) database as the validation cohort was downloaded (9). Data with incomplete survival information were excluded. The training dataset comprised 144 EC patients while the validation dataset comprised 179 EC patients.

The edgeR (10) package was employed to identify DEmiRNAs or DEmRNAs. The criteria of DEmiRNAs or DEmRNAs were false discovery rate (FDR) q value <0.05 and $\log_2$ |fold change (FC)| >1.5.

Target genes of DEmiRNAs were selected by TargetScan (http://www.targetscan.org/vert_72/) and microRNA Data Integration Portal (mirDIP) databases (http://ophid.utoronto.ca/mirDIP/index.jsp). Further, a panel of prognostic genes were identified by taking the intersection of DEmRNAs and target genes of DEmiRNAs. To better understand the deep molecular mechanism behind these potential prognostic biomarkers, we conducted pathway enrichment analysis based on Gene Ontology (GO) database (11) and Kyoto Encyclopedia of Genes and Genomes (KEGG) database (12).

Table 1 The clinical characteristics and sample size for TCGA and GEO datasets

Variable	TCGA		GEO	
	Alive (n=90)	Dead (n=54)	Alive (n=106)	Dead (n=73)
Age, years, mean (SD)	62.20 (12.3)	62.83 (11.3)	57.44 (8.5)	60.66 (9.2)
Follow-time, years, mean (SD)	1.46 (1.1)	1.60 (1.3)	5.02 (0.5)	1.64 (1.1)
Male, n (%)	72 (80.0)	50 (92.6)	22 (20.8)	11 (15.1)
Stage, n (%)				
Stage I	13 (14.4)	3 (5.6)	0 (0.0)	0 (0.0)
Stage II	47 (52.2)	21 (38.9)	36 (34.0)	41 (56.2)
Stage III	28 (31.1)	20 (37.0)	67 (63.2)	25 (34.2)
Stage IV	0 (0.0)	8 (14.8)	3 (2.8)	7 (9.6)
Missing	2 (2.2)	2 (3.7)	0 (0.0)	0 (0.0)

TCGA, The Cancer Genome Atlas; GEO, Gene Expression Omnibus.

Construction and validation of a prognostic model

Least absolute shrinkage and selection operator (LASSO) Cox regression analysis was used to select differential genes to construct final prognostic model using the R package "glmnet" (13). Then, Cox regression analysis was performed to construct the best prognostic model and make proportional hazards assumption with R package "survival" and "MASS". The prognostic genes eventually screened in the training dataset were used to construct the prognostic model, and then verified in the validation dataset. According to the prognostic model formula, the risk score of each patient was calculated, and then divided EC patients into high-risk and low-risk groups according to the median risk value. Subsequently, we constructed a prognostic model using genes, tumor node metastasis (TNM) stage, and risk group, and validation in the validation dataset; 1-, 2-, and 3-year receiver operating characteristic (ROC) curves were drawn using the R package "timeROC" (14). Then, according to the genes, TNM stage, and risk group and other prognostic factors, the nomogram was constructed to predict overall survival (OS).

Analysis of immune status of EC patients

Moreover, immune cell infiltration, gene checkpoint gene expression, and significantly mutated gene identification were also assessed between the high-risk group and low-risk group. The software CIBERSORT (15) was used to deconvolve the expression matrix of human immune cell subtypes to calculate the relative proportions of 22 immune cells in EC samples. The EC immune checkpoint genes were searched in the literature and found that the most studied EC immune checkpoint inhibitors were *CTLA4*, *PDCD1*, *LAG3*, *TIGIT*, *IDO1*, *TDO2*, *CD19*, *STAT3*, *CD47* and *WT1* (16-24). The differences between the high risk group and the low risk group of these immune checkpoint genes were explored. The R package "maftools" was used to analyze variant genes for somatic mutation data. Somatic mutation data for different risk groups were used to detect mutation types and single nucleotide variants (SNV). Differential mutant genes (DMG) were identified based on Fisher's exact test, $P<0.05$ (25).

Statistical analysis

Baseline characteristics between different groups were analyzed using Student's *t*-test or Wilcoxon rank sum test as appropriate for continuous variables and χ^2 test for categorical variables. All statistical analyses were carried out with R (4.1.2).

Ethical statement

The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

Results

TCGA and GEO datasets and patients characteristics

Baseline clinical characteristics of included EC patients were presented in *Table 1*, including 144 samples from

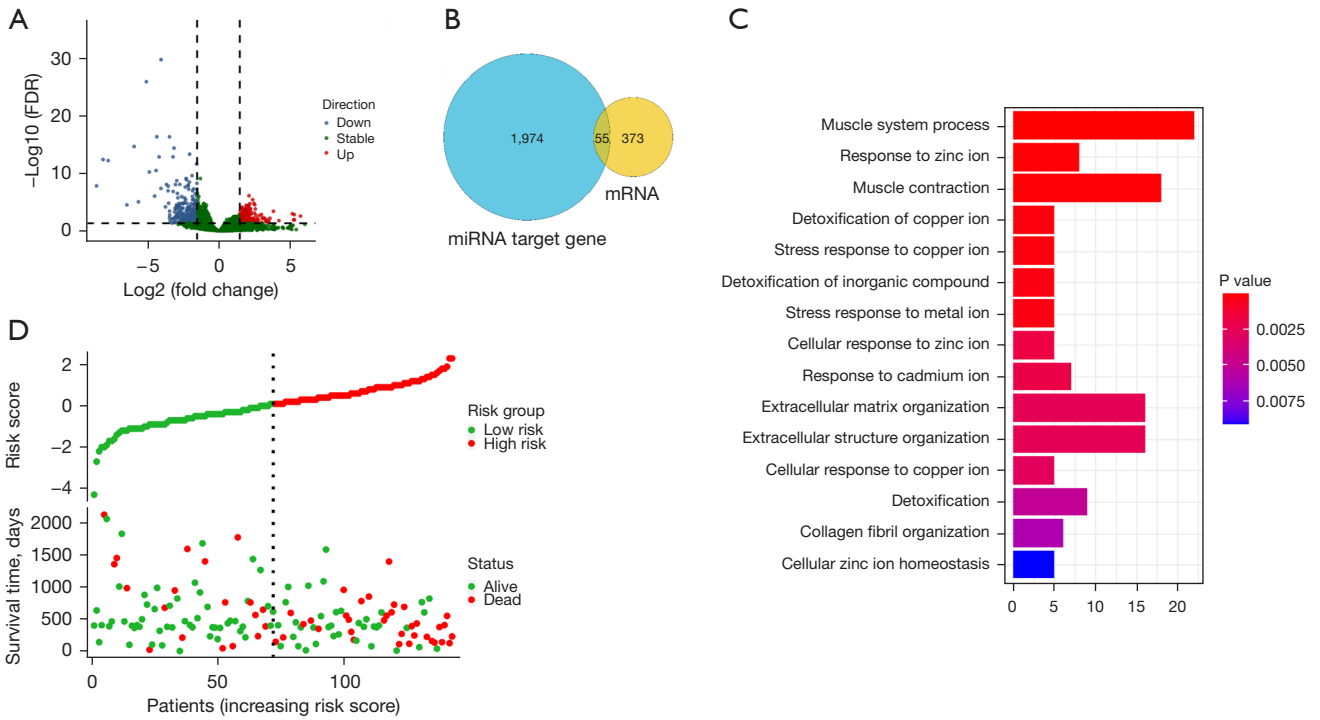

Figure 1 Results of differential gene analysis. (A) Volcano plot showing the differentially up-regulated (red nodes) and down-regulated genes (blue nodes). (B) Venn diagram of DEmRNAs and DEmiRNAs target genes. (C) Enrichment analysis of the common genes between DEmRNAs and DEmiRNAs target genes. (D) Scatterplots illustrate the distribution of risk scores and survival status of EC patients. FDR, false discovery rate; DEmRNAs, differentially expressed mRNAs; DEmiRNAs, differentially expressed microRNAs; EC, esophageal cancer.

TCGA database as training set and 179 samples from GEO database as validation set. EC patients in training set were more likely to be older males than patients in validation set (P<0.05).

Identification of differentially expressed genes and enrichment analysis

In total, 428 genes were identified as DEmRNAs, including 177 up-regulated genes and 251 down-regulated genes in gene expression data (*Figure 1A*); 2,029 target genes of DEmiRNAs were identified by TargetScan and mirDIP databases. There are 55 common genes between DEmRNAs and DEmiRNAs target genes (*Figure 1B*). GO and KEGG enrichment analysis was performed for DEmRNAs, DEmiRNAs and 55 common genes, and the results showed that DEmRNAs were mainly enriched in mitotic nuclear division, nuclear division and chromosome segregation (Figure S1A), DEmiRNAs was mainly enriched in muscle system process, response to zinc ion and muscle contract ion (Figure S1B), and 55 common genes were mainly enriched in muscle system process, response to zinc ion and muscle contraction and other KEGG pathways (*Figure 1C*).

Construction and validation of a risk score prognostic model

Differential genes were screened by LASSO Cox analysis, including *ARHGAP11A*, *CELF2*, *H1.4*, *H2AC8*, *HMGB3*, *ZNF367*, *LRIG1*, *PRR11*, and *COL4A1*. Then, these differential genes were included in multivariate Cox stepwise regression analysis, and genes with P<0.05 were selected to construct a prognostic model. The prognostic genes were *ARHGAP11A*, *H1.4*, *HMGB3*, *LRIG1*, *PRR11*, and *COL4A1*. The prognostic genes were validated in the validation cohort. A risk score for each patient was calculated and then divided samples from the training and validation cohorts into high-risk group and low-risk group based on the median, respectively. The results (*Figure 1D*)

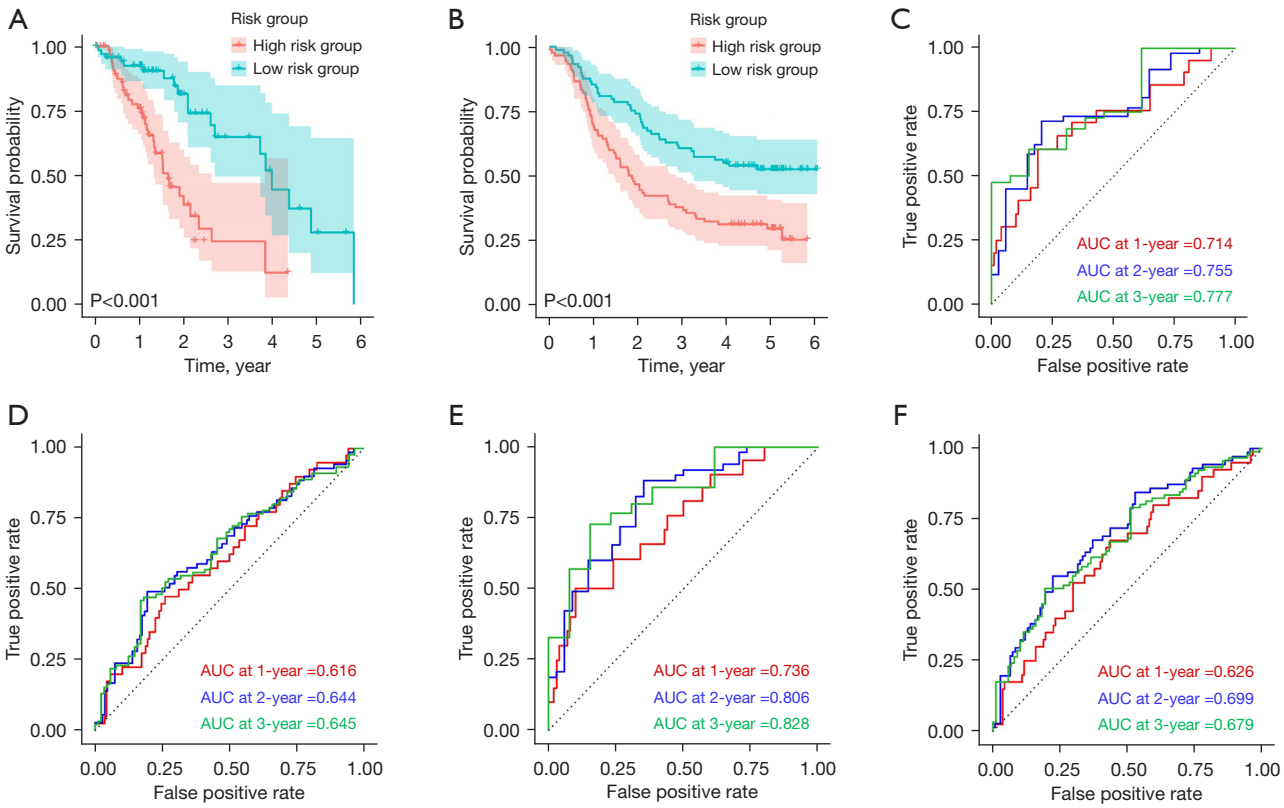

Figure 2 Construction and validation prognostic models. (A) Kaplan-Meier survival analysis with the 6 genes and TNM stage risk score as a stratification factor in the TCGA cohort. (B) Kaplan-Meier survival analysis with the 6 genes and TNM stage risk score as a stratification factor in the GEO cohort. (C) ROC curves of the risk score for predicting 1-, 2-, and 3-year survival in the TCGA cohort. (D) ROC curves of the risk score for predicting 1-, 2-, and 3-year survival in the GEO cohort. (E) ROC curves of the risk score combined with clinical factors for predicting 1-, 2-, and 3-year survival in the TCGA cohort. (F) ROC curves of the risk score combined with clinical factors for predicting 1-, 2-, and 3-year survival in the GEO cohort. AUC, area under the curve; TNM, tumor node metastasis; TCGA, The Cancer Genome Atlas; GEO, Gene Expression Omnibus; ROC, receiver operating characteristic.

showed that the prognosis of the high-risk group was significantly poorer than that of the low-risk group in the training cohort. Compared with the low-risk group in *Figure 2A*,*2B*, patients in the high-risk group had significantly shorter OS in both the training and validation cohorts (all P<0.001) datasets.

This study constructed a risk score prediction model using the above prognostic genes, and the 1-, 2-, and 3-year areas under the curve (AUC) for EC risk scores in the training cohort were 0.714, 0.755, and 0.777, respectively (*Figure 2C*). The AUC in the validation cohort were 0.616, 0.644, and 0.645, respectively. These findings further indicated these prognostic genes had great potential in predicting future prognosis for EC patients (*Figure 2D*). Subsequently, used eight independent prognostic factors including prognostic genes, TNM stage, and risk group to constructed a risk prediction model. For prognosis prediction for the training cohort, the performance of the models with 1-, 2-, and 3-year AUC equaled 0.736, 0.806, and 0.828 (*Figure 2E*). For validated prognostic predictions, the performance of the models with 1-, 2-, and 3-year AUC was 0.626, 0.699, and 0.679, respectively (*Figure 2F*).

Independent prognostic role of the prognostic gene signature

A multivariate Cox regression model consisting of age, TNM stage, gender, and risk group was used to confirm whether the risk score was an independent prognostic indicator (*Figure 3A*). TNM stage and risk group were found

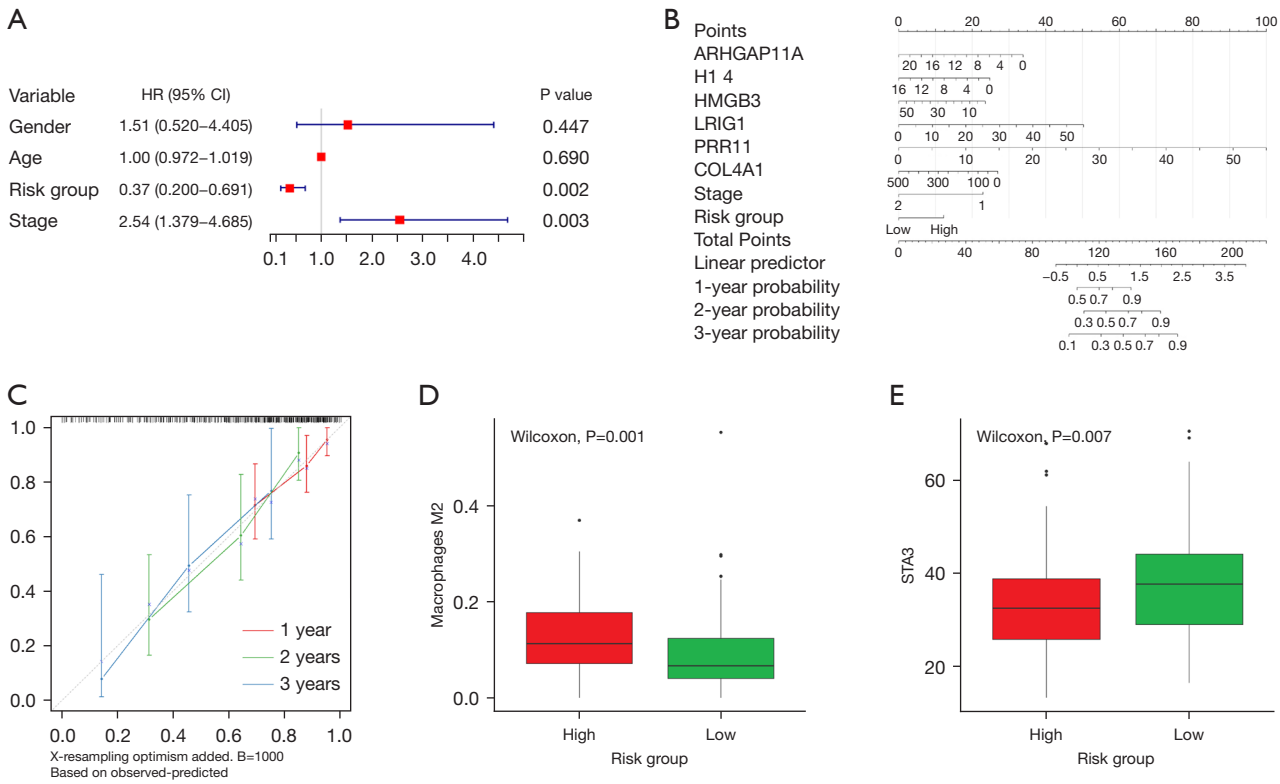

Figure 3 Cox regression analysis and immune infiltration analysis. (A) Multivariate Cox regression analysis of forest plots. (B) Nomogram to predict the probability of 1, 2, and 3 years OS in EC patients. (C) Normative curve for predicting the probability of 1, 2, and 3 years OS in EC patients. The X axis represents nomogram predicted survival, and the Y axis represents the actual survival. (D) Comparisons of the immune cell members between high-risk group and low-risk group. (E) Significant immune checkpoint genes for EC patients with risk score as a grouping factor. HR, hazard ratio; CI, confidence interval; OS, overall survival; EC, esophageal cancer.

to remain significantly associated with OS [hazard ratio (HR) =0.37, 95% CI: 0.20–0.69, P=0.002; HR =2.54, 95% CI: 1.38–4.69, P=0.003]. This study constructed a nomogram to predict patient OS by eight independent prognostic factors, including prognostic genes, TNM stage, and risk group (*Figure 3B*,*3C*). The calibration curve clarified an excellent calibration performance of this prognostic model.

Immune status of EC patients with different risk groups

The CIBERSORT method was used to estimate the immune infiltration differences between 22 immune cells in EC patients in the high-risk group and low-risk group. The correlation between different immune cell infiltration rates was weaker (Figure S2). The expression level of Macrophage M2 in the high-risk group was significantly higher than that in the low-risk group (P=0.001) (*Figure 3D*), the difference of other immune cells was not statistically significant (Figure S3). The relationship between different risk groups and key immune checkpoints (*CTLA4*, *PDCD1*, *LAG3*, *TIGIT*, *IDO1*, *TDO2*, *CD19*, *STAT3*, *CD47*, *WT1*) was analyzed. The expression level of *STAT3* immune checkpoint was significantly higher in low-risk group than in high-risk group (P<0.05) (*Figure 3E*). Combining these results, the study found that the different risk groups of EC can accurately indicate the immune level.

Figure 4A,*4B* showed a panoramic of the 15 genes with the highest mutation frequency in the low-risk and high-risk groups for each sample. *TP53* and *TTN* occupy the top two places in the two cohorts, indicating that these two genes were mainly involved in the occurrence and progression of tumors. As shown in *Figure 5A*,*5B*, the number of genes that significantly co-occur in pairs in the high-risk group is less than that in the low-risk group, while the number of genes with mutually exclusive mutations is greater than that in the low-risk group.

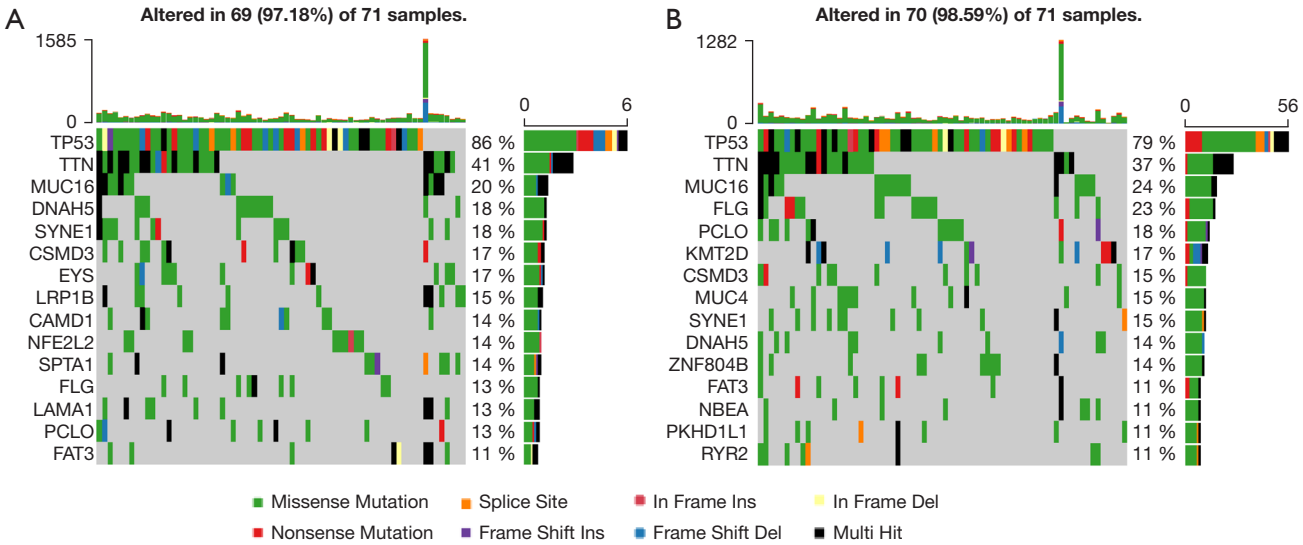

Figure 4 Waterfall plot shows the mutation distribution of the top 15 most frequently mutated genes. The central panel shows the types of mutations in each EC sample. The upper panel shows the mutation frequency of each EC sample. The bar plots on the right side show the frequency and mutation type of genes mutated in the low-risk and high-risk cohort. The bottom panel is the legend for mutation types. (A) High risk group. (B) Low risk group. EC, esophageal cancer.

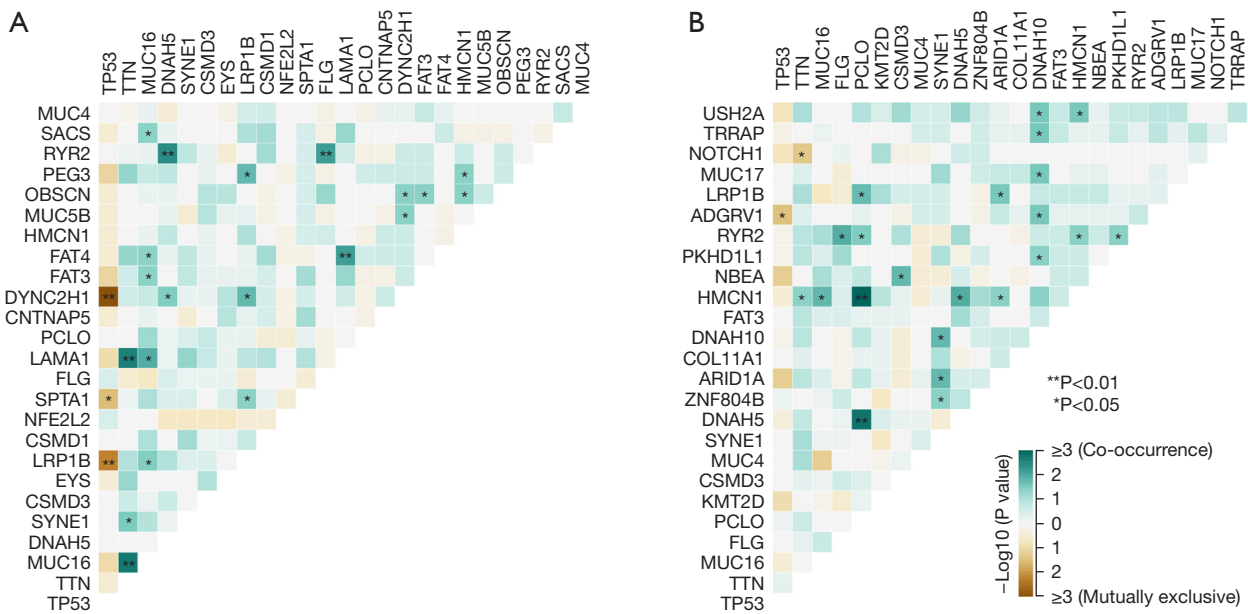

Figure 5 The heatmap illustrates the mutually co-occurring and exclusive mutations of the top 25 frequently mutated genes. The color and symbol in each cell represent the statistical significance of the exclusivity or co-occurrence for each pair of genes. (A) High risk group. (B) Low risk group. Symbols indicated statistical significance for the Mann-Whitney U test: *P<0.05; **P<0.01.

Discussion

In this study, the intersection of DEmRNAs and target genes of DEmiRNAs was applied to identify EC prognostic genes, including *ARHGAP11A*, *H1.4*, *HMGB3*, *LRIG1*, *PRR11*, and *COL4A1*. Six prognostic genes were used to build EC prognostic model in training set, then GEO database was applied to validate their prognostic value among EC patients. According to the median risk score, included EC patients were divided into high-risk or low-risk group. Then, in the immune microenvironment, the expression level of M2 in macrophages in the high-risk group was observably higher than that in the low-risk group. For the immune checkpoint, only the *STAT3* immune checkpoint in EC patients was significant in the low-risk group compared with the high-risk group. Finally, among the somatic mutation genes of EC, the mutation frequency of the top 15 genes, *TP53* and *TTN* were in the top 2 in the high-risk group and the low-risk group were found.

Most genes included in the prognostic model of this study have been reported to be associated with many different malignancies. For example, Lawson *et al.* found that *ARHGAP11A* is highly expressed in a subtype of basal-like breast cancer (BLBC). Previous study showed decreased migration velocity of BLBC without *ARHGAP11A*, indicating their oncogenic potential could be evaluated by their ability to induce cellular transformation (26). *RhoGAPs* is a newly discovered class of cancer biomarkers. *ArhGAP11A*, a member of the *RhoGAP* family, induces cell cycle arrest and apoptosis by binding tumor suppressor p53. *ArhGAP11A* directly binds p53 and promotes its transcriptional function through a Rho-independent mechanism. The expression level of *ArhGAP11A* was proved to have impact on cell proliferation and apoptosis (27,28).

HMGB3 is a member of the high mobility group box (*HMGB*) family, which is overexpressed in gastric and bladder cancers. And in colorectal cancer, breast cancer is up-regulated at both mRNA and protein levels (29-31). In addition, *LRIG1* is associated with survival in patients with various cancers (32). The *LRIG1* gene is located on chromosome 3p14.3, and deletion of this region is associated with a variety of tumors. The degree of low expression of *LRIG1* increases with the increase of tumor grade, and the susceptibility of glioma is related to the *LRIG1* gene (33). *LRIG1* functions as a thyroid tumor suppressor (34).

Previous studies have explored the prognostic value of differential genes among EC patients, this study further combined the target genes of DEmiRNAs to construct prognostic model. Both ROC analysis and calibration curve showed this prognostic model in the study had better discrimination and calibration performance. In addition, we also explored the difference of TME between high-risk group and low-risk group. In addition, we used the external data GEO dataset as a validation cohort. However, several limitations of this study need to be considered. Firstly, the prognostic model has only been validated in the GEO dataset and further validation in other independent large sample cohorts is required to ensure the reliability of our model. Secondly, for immune checkpoints, we only found a part of immune checkpoints, and many others were not included in the analysis. Thirdly, in this study, a small sample size was included. Lastly, these findings were from population based study, further experiments are needed to clarify deep molecular mechanism underlying the association between TME and EC.

Conclusions

The use of DEmRNA and DEmiRNA target genes to make a more accurate diagnosis and prognosis prediction of EC patients, so as to formulate individualized treatment plans on this basis, for the prognosis of patients bring greater benefits.

Acknowledgments

We are very grateful to those who participated in our research. Without their dedication and cooperation, our research could not begin. We are very grateful to all patients and staff who participated and contributed to TCGA and GEO for all their support.

Funding: This work was supported by grants from the National Natural Science Foundation of China (Nos. 82222064, 81973147, 81573246), Health Science and Technology Development Program of Shandong Province (No. 202005010068), National Key Research and Development Program (No. 2022YFC2010100), and Shandong University Distinguished Young Scholars. The funders had no role in study design, data collection, and analysis, decision to publish, or preparation of the manuscript.

Footnote

Reporting Checklist: The authors have completed the

TRIPOD reporting checklist. Available at https://tcr.amegroups.com/article/view/10.21037/tcr-22-2588/rc

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://tcr.amegroups.com/article/view/10.21037/tcr-22-2588/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. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

References

1. Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018;68:394-424.
2. Watanabe M, Otake R, Kozuki R, et al. Recent progress in multidisciplinary treatment for patients with esophageal cancer. Surg Today 2020;50:12-20.
3. Li X, Yu X, He Y, et al. Integrated Analysis of MicroRNA (miRNA) and mRNA Profiles Reveals Reduced Correlation between MicroRNA and Target Gene in Cancer. Biomed Res Int 2018;2018:1972606.
4. Zhang Y, Zhu M, Mo J, et al. Tumor microenvironment characterization in esophageal cancer identifies prognostic relevant immune cell subtypes and gene signatures. Aging (Albany NY) 2021;13:26118-36.
5. Pang J, Pan H, Yang C, et al. Prognostic Value of Immune-Related Multi-IncRNA Signatures Associated With Tumor Microenvironment in Esophageal Cancer. Front Genet 2021;12:722601.
6. Li C, Zhou W, Zhu J, et al. Identification of an Immune-Related Gene Signature Associated with Prognosis and Tumor Microenvironment in Esophageal Cancer. Biomed Res Int 2022;2022:7413535.
7. Huai Q, Guo W, Han L, et al. Identification of prognostic genes and tumor-infiltrating immune cells in the tumor microenvironment of esophageal squamous cell carcinoma and esophageal adenocarcinoma. Transl Cancer Res 2021;10:1787-803.
8. Tomczak K, Czerwińska P, Wiznerowicz M. The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. Contemp Oncol (Pozn) 2015;19:A68-77.
9. Barrett T, Wilhite SE, Ledoux P, et al. NCBI GEO: archive for functional genomics data sets--update. Nucleic Acids Res 2013;41:D991-5.
10. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 2010;26:139-40.
11. Huang Q, Wu LY, Wang Y, et al. GOMA: functional enrichment analysis tool based on GO modules. Chin J Cancer 2013;32:195-204.
12. Wanggou S, Feng C, Xie Y, et al. Sample Level Enrichment Analysis of KEGG Pathways Identifies Clinically Relevant Subtypes of Glioblastoma. J Cancer 2016;7:1701-10.
13. Engebretsen S, Bohlin J. Statistical predictions with glmnet. Clin Epigenetics 2019;11:123.
14. George B, Seals S, Aban I. Survival analysis and regression models. J Nucl Cardiol 2014;21:686-94.
15. Newman AM, Liu CL, Green MR, et al. Robust enumeration of cell subsets from tissue expression profiles. Nat Methods 2015;12:453-7.
16. Hao S, Huang G, Feng J, et al. Non-NF2 mutations have a key effect on inhibitory immune checkpoints and tumor pathogenesis in skull base meningiomas. J Neurooncol 2019;144:11-20.
17. Pardoll DM. The blockade of immune checkpoints in cancer immunotherapy. Nat Rev Cancer 2012;12:252-64.
18. Dyck L, Mills KHG. Immune checkpoints and their inhibition in cancer and infectious diseases. Eur J Immunol 2017;47:765-79.
19. Johnson DE, O'Keefe RA, Grandis JR. Targeting the IL-6/JAK/STAT3 signalling axis in cancer. Nat Rev Clin Oncol 2018;15:234-48.
20. Feng M, Jiang W, Kim BYS, et al. Phagocytosis checkpoints as new targets for cancer immunotherapy. Nat Rev Cancer 2019;19:568-86.
21. Ruffo E, Wu RC, Bruno TC, et al. Lymphocyte-activation gene 3 (LAG3): The next immune checkpoint receptor. Semin Immunol 2019;42:101305.

22. Ogasawara M, Miyashita M, Yamagishi Y, et al. Immunotherapy employing dendritic cell vaccination for patients with advanced or relapsed esophageal cancer. Ther Apher Dial 2020;24:482-91.
23. Harjunpää H, Guillerey C. TIGIT as an emerging immune checkpoint. Clin Exp Immunol 2020;200:108-19.
24. Zhai L, Ladomersky E, Lenzen A, et al. IDO1 in cancer: a Gemini of immune checkpoints. Cell Mol Immunol 2018;15:447-57.
25. Mayakonda A, Lin DC, Assenov Y, et al. Maftools: efficient and comprehensive analysis of somatic variants in cancer. Genome Res 2018;28:1747-56.
26. Lawson CD, Fan C, Mitin N, et al. Rho GTPase Transcriptome Analysis Reveals Oncogenic Roles for Rho GTPase-Activating Proteins in Basal-like Breast Cancers. Cancer Res 2016;76:3826-37.
27. Kagawa Y, Matsumoto S, Kamioka Y, et al. Cell cycle-dependent Rho GTPase activity dynamically regulates cancer cell motility and invasion in vivo. PLoS One 2013;8:e83629.
28. Xu J, Zhou X, Wang J, et al. RhoGAPs attenuate cell proliferation by direct interaction with p53 tetramerization domain. Cell Rep 2013;3:1526-38.
29. Zhang Z, Chang Y, Zhang J, et al. HMGB3 promotes growth and migration in colorectal cancer by regulating WNT/β-catenin pathway. PLoS One 2017;12:e0179741.
30. Gong Y, Cao Y, Song L, et al. HMGB3 characterization in gastric cancer. Genet Mol Res 2013;12:6032-9.
31. Tang HR, Luo XQ, Xu G, et al. High mobility group-box 3 overexpression is associated with poor prognosis of resected gastric adenocarcinoma. World J Gastroenterol 2012;18:7319-26.
32. Yu Q, Li Y, Peng S, et al. Exosomal-mediated transfer of OIP5-AS1 enhanced cell chemoresistance to trastuzumab in breast cancer via up-regulating HMGB3 by sponging miR-381-3p. Open Med (Wars) 2021;16:512-25.
33. Lindquist D, Alsina FC, Herdenberg C, et al. LRIG1 negatively regulates RET mutants and is downregulated in thyroid cancer. Int J Oncol 2018;52:1189-97.
34. Mao F, Holmlund C, Faraz M, et al. Lrig1 is a haploinsufficient tumor suppressor gene in malignant glioma. Oncogenesis 2018;7:13.

Cite this article as: Gu B, Zhang S, Fan Z, Che J, Li S, Li Y, Pan K, Lv J, Wang C, Zhang T, Wang J. Prognostic model construction and immune microenvironment analysis of esophageal cancer based on gene expression data and microRNA target genes. Transl Cancer Res 2023;12(5):1165-1174. doi: 10.21037/tcr-22-2588

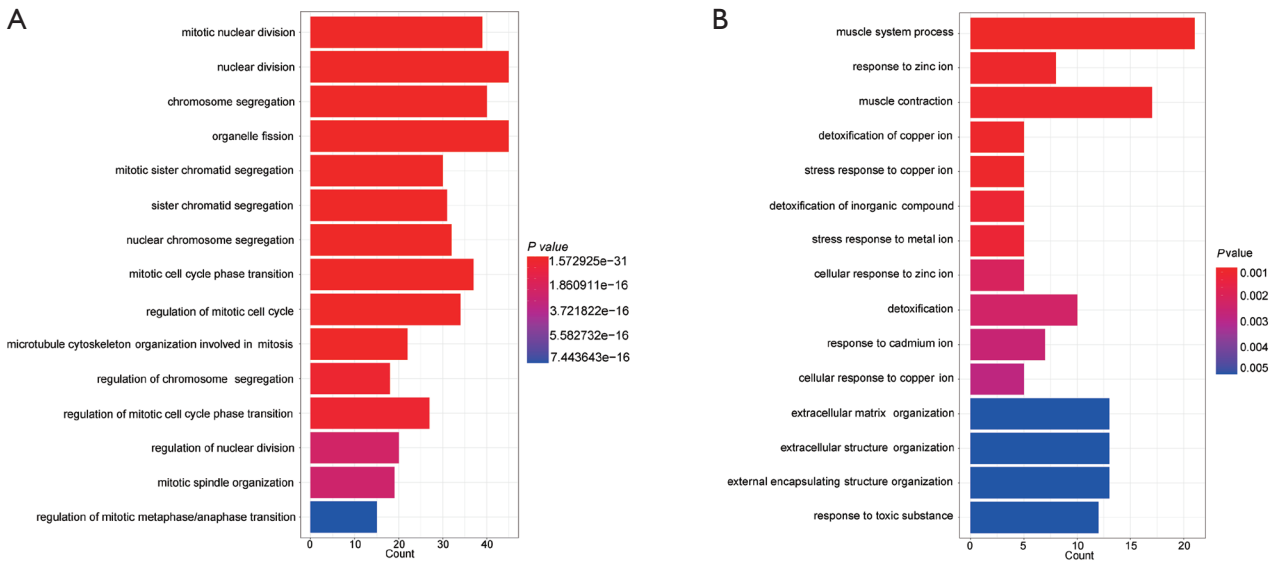

Figure S1 Enrichment analysis of mRNA differentially genes. (A) mRNA differentially up-regulated genes. (B) mRNA differentially down-regulated genes.

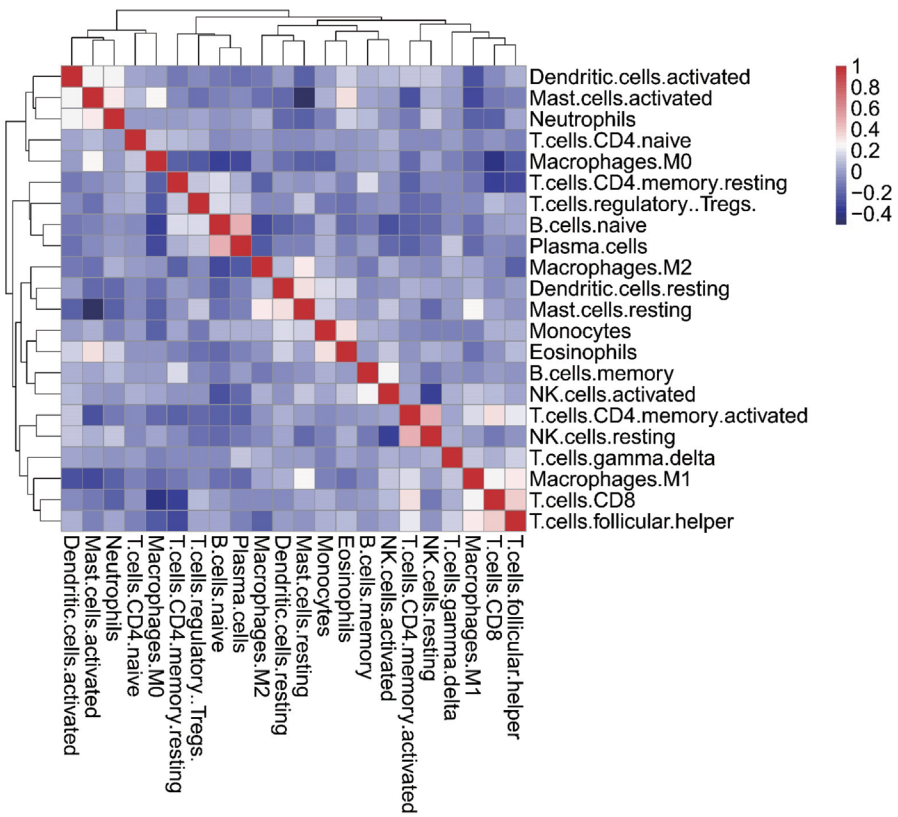

Figure S2 The correlation between different immune cell infiltration rates. NK, natural killer.

https://dx.doi.org/10.21037/tcr-22-2588

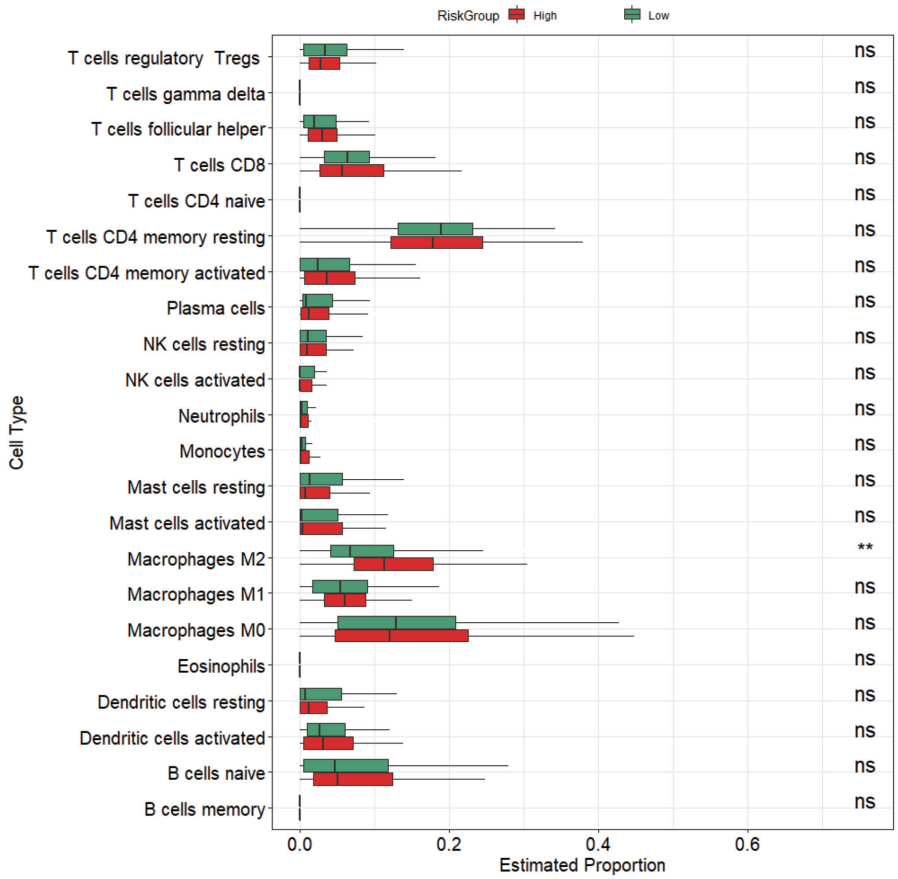

Figure S3 The proportion of immune infiltrating cells in different subgroups: ns, P>0.05; **, P≤0.001. NK, natural killer; ns, no significance.

https://dx.doi.org/10.21037/tcr-22-2588