# Development of a miR-3648-Targeted Immune-Related 4-Gene Signature as a Prognostic Biomarker in Esophageal Adenocarcinoma Based on ehe WGCNA Algorithm

**Donglei Zhang**
Shanghai Jiao Tong University School of Medicine Affiliated Renji Hospital
**Hang Yin**
Shanghai Jiao Tong University School of Medicine Affiliated Renji Hospital
**Ping Xu**
Shanghai Jiao Tong University School of Medicine Affiliated Renji Hospital
**Xiaozhe Qian** ( qianxiaozhe@renji.com )
Shanghai Jiao Tong University School of Medicine Affiliated Renji Hospital https://orcid.org/0000-0002-5160-8572

**Primary research**

**Keywords:** Esophageal adenocarcinoma, microRNA, prognostic signature, immune cell infiltration, weighted gene coexpression network analysis.

**Posted Date:** June 2nd, 2021

**DOI:** https://doi.org/10.21203/rs.3.rs-549122/v1

# Abstract

**Background**: Esophageal adenocarcinoma (EA) has a poor prognosis and is a typical immunogenic malignant tumor. Abnormal expression of miR-3648 has been reported in EA, but its value in prognosis prediction and immune cell infiltration imbalance mediation is still unknown. We aimed to mine immune-related genes (IRGs) targeted by miR-3648 and construct a multigene signature to improve the prognostic prediction of EA.

**Methods**: The gene expression data of EA tumor or normal tissues from The Cancer Genome Atlas (TCGA) database and GTEx database were downloaded. Weighted gene coexpression network analysis (WGCNA), the CIBERSORT algorithm and Cox regression analysis were applied to identify IRGs and to construct a prognostic signature and nomogram.

**Results**: miR-3648 was obviously highly expressed in EA tumor tissues and was correlated with patient survival time [hazard ratio (HR) = 1.28, 95% confidence interval (CI): 1.09-1.49, p = 0.002]. A total of 70 miR-3648-targeted genes related to immune cell infiltration were identified, and a novel 4-gene signature (C10orf55, DLL4, PANX2 and NKAIN1) was established. The prognostic signature-based risk score has superior capability to predict overall survival (AUC = 0.740 for 1 year; AUC = 0.717 for 3 years; AUC = 0.622 for 5 years). A higher score was indicative of a poorer prognosis than a lower score (HR = 1.69, 95% CI: 1.08-2.64, p = 0.20, adjusted by TNM stage).

**Conclusion**: miR-3648 might play a crucial role in the progression of EA. The novel miR-3648-targeted immune-related 4-gene signature is expected to become a potential prognostic marker in EA.


# Introduction

Esophageal adenocarcinoma (EA) is a highly lethal carcinoma. [1] Patients with EA are often diagnosed at an advanced stage, which results in less than half of the patients being eligible for potentially curative treatment at diagnosis. [2] Although recent advances in therapeutic approaches for gastroesophageal tumors have significantly improved the curative resection rates, as well as the disease-free and overall survival rates, the 5-year overall survival rate for patients with advanced EA is lower than 20%. [3] There are significant prognostic discrepancies in patients with the same TNM stage. [4] To improve risk prediction, it is necessary to identify objective biomarkers to predict response to EA or to predict survival.

The tumor microenvironment (TME) is composed of stromal cells, immune cells, extracellular matrix molecules, and inflammatory mediators. Tumor cells could promote immune escape by forming an immunosuppressive microenvironment. [5] The heterogeneity of immune cells also affects the time and intensity of the antitumor response and becomes a major obstacle to the treatment of tumors.[6] There is still a lack of systematic and comprehensive research on the correlation between immune infiltration and EA. Understanding the cancer-related immune landscape is vital to the treatment and prognosis of EA.

microRNA (miRNA) is an important gene regulatory substance that affects gene expression by pairing with the 3'-UTR of the target gene. Approximately one-third of all human genes are directly regulated by miRNAs. miRNAs regulate many important cellular processes such as cell proliferation, migration and apoptosis. [7] Several cancer-associated miRNAs have now been identified, of which miRNA-3648 (miR-3648) is one of the most established and broad-acting miRNAs.[8–10] miR-3648 is abnormally expressed in many cancers, including esophageal carcinoma. [10] Dysregulation of miR-3648 can promote cell proliferation and induce the invasion and metastasis of cancer cells. [9, 11] However, the impact of miR-3648 on clinical characteristics in EA patients is less well explored, as is the correlation between miR-3648 and the immune microenvironment. The in-depth exploration of the immune regulatory network driven by miR-3648 has the potential to provide suggestions for the prognostic improvement of EA.

In the present study, we utilized weighted gene coexpression network analysis (WGCNA) to construct a coexpression network of differentially expressed target genes of miR-3648. The infiltrated immune cells were identified with the CIBERORT algorithm, and immune-related genes in the coexpression network were determined. Finally, Cox analysis was performed to construct a multigene prognostic signature, and a nomogram was established to improve patient survival prediction capacity.

# Methods And Materials

### Data source and clinical information

We downloaded EA-related miRNA-seq, mRNA-seq expression profiles and clinical information from the TCGA database (https://cancergenome.nih.gov/). miRNA data were extracted from 87 tumor tissues and 13 normal esophageal tissues. We also downloaded the mRNA-seq expression profiles of 54 extra normal esophageal tissues as controls from the GTEx database (https://www.gtexportal.org/). Finally, the expression data of 19835 genes in 87 tumor samples and 67 normal samples were summarized as mRNA data.

### Differential expression analysis and target gene prediction

Linear models were used to identify differentially expressed miRNAs/mRNAs in tumor/normal groups using the *limma* R package. [12] A false discovery rate (FDR)-adjusted P value of <0.05 combined with a simultaneous absolute value of >1 for logFC was set as the threshold for differentially expressed miRNA/mRNA identification.

The correlation between miRNAs and the overall survival (OS) of the patient was further estimated by a Cox regression analysis. The diagnostic and prognostic capability of significant miRNAs was measured by the area under the curve (AUC) of receiver operating characteristic (ROC) analysis.

Subsequently, 12 databases (Microt4, miRWalk, mir-bridge, miRanda, miRDB, miRMap, Pictar2, PITA, MiRNAMap, RNAhybrid, RNA22 and TargetScan) in miRWalk2.0 (http://zmf.umm.uni-

heidelberg.de/apps/zmf/mirwalk2/) were employed to determine the target genes of miR-3648. The genes that could be found in no less than 3 databases were identified as the target genes of miR-3648.

Then, the cross-linked genes between the target genes of miR-3648 and the differentially expressed genes were determined as the differentially expressed target genes (DETGs) of miR-3648.

### Tumor-infiltrating immune cell analysis

CIBERSORT (https://cibersort.stanford.edu/) and leucocyte signature matrix 22 (LM22) were used to quantify the proportions of different immune cell types in the ESCC samples from the TCGA database. [13] Normalized gene expression data were analyzed using the CIBERSORT algorithm by running 1000 permutations. The CIBERSORT p value reflects the statistical significance of the results, and a threshold less than 0.05 is recommended.

### Weighted gene coexpression network analysis

The R package “weighted gene coexpression network analysis” (WGCNA) [14] was utilized to identify modules of highly correlated DETGs as previously described. [15] Briefly, the DETG expression matrix was used to define gene coexpression similarity based on Pearson’s correlation coefficient of paired genes, followed by an adjacency matrix transformed from coexpression similarity using the power function. Here, with a soft threshold $\beta = 4$ and with $R^2 > 0.9$, the coexpression network distribution exhibited an essentially scale-free topology. Next, a topological overlap matrix (TOM) is built based on the adjacency matrix A. The TOM-based dissimilarity subsequently led to distinct modules defined as clusters of densely interconnected genes. A dynamic hybrid tree-cutting algorithm is further used to build a hierarchical clustering tree to divide modules.

The infiltration fractions of immune cells in tumor tissues and normal tissues were employed as clinical traits. Pearson’s correlation test was then used to evaluate the module correlations with clinical traits. Among these clinical traits, the genes of the most significant module were determined to be immune-related genes (IRGs) for subsequent analysis.

### Identification of Key Prognostic Genes

To identify key prognostic genes, we conducted a Cox regression analysis to estimate the association between the expression of the genes of the most significant module and overall survival time of patients. Through this method, 4 genes for which P was <0.05 were identified as candidate prognostic genes.

### Construction of prognostic signature

We selected genes with excellent prognostic performance and diagnostic capability to construct a prognostic signature. The risk score (RS) for each tumor sample was calculated using the following formula.

$$RS = \sum_{i=1}^{n} \mathrm{Coef}_i \times Exp_i$$

In the above formula, n represents the number of genes included in the prognostic signature. $Exp_i$ represents the expression level of gene i. $Coef_i$ represents the Cox regression coefficient of gene i. The prognostic independence of the risk score under the interference of multiple clinical characteristics was analyzed through univariate and multivariate Cox regression analysis.

### Establishment of nomogram

Then, the independent prognostic characteristics were included in the construction of the nomogram. Meanwhile, the consistency index of actual survival and predicted survival was employed to assess the survival prediction capability of the nomogram. The prediction results and observation results in the calibration curve were visualized to measure the prediction performance of the nomogram.

### Statistical analysis

The Wilcoxon test was applied to estimate the difference between groups. ROC analysis and Kaplan–Meier (K-M) curve analysis were performed to evaluate the prognostic performance and diagnostic capacity. The construction of the nomogram employed the R package “rms v6.0-2”. All R program packages for statistical analysis were implemented on the “R v4.0.4” platform. $P < 0.05$ was considered statistically significant.

## Results

### Recognition of differentially expressed microRNAs associated with prognosis

A total of 78 differentially expressed miRNAs were identified in TCGA EA samples. (see Additional file 1: Fig. S1) Not surprisingly, miR-3648 was obviously more highly expressed in tumor tissue than in normal tissues ($p < 0.001$; Fig. 1A) and showed excellent discrimination between tumor tissue and normal tissue (AUC = 0.935, p = 4.03e-15; Fig. 1B). We then applied Cox regression analysis to evaluate the impact of miRNAs on patient survival. In addition to miR-3648, 4 other miRNAs (miR-550a-1, miR-550a-2, miR-3682 and miR-503) associated with the prognosis of EA were identified. (Table 1) The ROC curves for the prediction of 1-year survival time are shown in Fig. 1C. Among these miRNAs, miR-3648 had the highest prediction capability (AUC = 0.770). Fig. 1D shows the survival rate of patients with miR-3648 estimated by the Kaplan-Meier method. A total of 12289 potential target genes of miR-3648 were identified based on 12 databases of miRWalk2.0. The 608 genes existing in no less than 3 databases were precisely established as the target genes of miR-3648.

### Differentially expressed target genes of miR-3648-5p

A total of 5746 differentially expressed genes (DEGs, satisfying $p < 0.05$ and absolute (log2FC) > 1) were identified in the gene differential expression analysis. These genes included 2817 upregulated genes and 2929 downregulated genes (Fig. 2A). Furthermore, we cross-linked target genes and differentially expressed genes to identify 170 differentially expressed target genes (DETGs) (Fig. 2B).

### Tumor-immune microenvironment fluctuation

We obtained the infiltration fraction of 22 immune cells based on the CIBERSORT deconvolution algorithm. The arithmetic means of the immune cell infiltration fraction of 87 tumor samples and 67 normal samples were taken as their infiltration proportion in the tumor group and the control group. The infiltration proportion of each immune cell subtype represents the ratio of its number to the number of all 22 immune cells in the immune microenvironment. M0 macrophages, M1 macrophages, activated dendritic cells, activated mast cells, activated memory CD4 T cells, follicular helper T cells and regulatory Tregs were the most common immune cell populations in EA tumor tissues, while resting dendritic cells, resting mast cells, monocytes, activated NK cells, resting memory CD4 T cells and naïve CD4 T cells were the most common immune cell populations in normal tissues. (Fig. 3)

### Identification of immune-related genes

We constructed a cp-expression network of 170 DETGs based on the WGCNA algorithm. Pick $\beta = 4$ to satisfy the scale-free network law. The two gene modules and 37 oligogenes were identified under the parameter settings of minModuleSize = 20 and mergeCutHeight = 0.25. (Fig. 3A) The two gene modules are the turquoise module and the blue module. The turquoise module contains 70 genes, and the blue module contains 63 genes. Subsequently, the infiltration fractions of 13 immune cells in different samples were used as clinical traits to obtain their correlation with gene modules. (Fig. 4B) The turquoise module is highlighted because it has the most significant negative correlation with the infiltration of T-cell regulatory Tregs (cor = -0.48, p = 3e-6). Additionally, activated mast cells and activated NK cells were negatively correlated with the turquoise module. The infiltration of resting dendritic cells is also closely related to the turquoise module. Therefore, the 70 genes of the turquoise module were identified as immune-related genes (IRGs).

### Construction and validation of 4-gene signature

The 4 genes that were closely related to OS were identified based on univariate Cox regression analysis of 70 IRGs. (Table 2) Then, the risk score (RS) of patients was calculated based on the linear combination of the expression values of these four genes multiplied by the Cox coefficients. The formula for calculating the risk score is:

RS = 0.29 * expression of C10orf55 + 0.39 * expression of DLL4 + 0.15 * expression of PANX2 + 0.16 * expression of NKAIN1.

All patients were divided into high-risk or low-risk groups by using the median RS as the cutoff. The median of RS in TCGA cohort was 9.04. The 87 patients were separated into 43 high-risk patients and 44

low-risk patients. Low-risk patients had longer survival times than high-risk patients (p = 0.0012; Fig. 5A). The AUCs of the risk score for 1-year, 3-year and 5-year survival prediction were 0.740, 0.717, and 0.622, respectively (Fig. 5B).

### Independent prognosis analysis of 4-gene signature

The prognostic performance of the risk score was obtained via univariate and multivariate Cox regression analysis of the risk score and multiple clinical characteristics (sex, age, TNM stage and classification). The univariate Cox regression analysis results of the TCGA cohort revealed that improved risk score (HR = 1.76, p < 0.001) and TNM stage (stage III/IV, HR = 3.63, p = 0.002) were risk factors for EA deterioration. (Fig. 5C) Therefore, the TNM stage was used as a covariate in multivariate Cox regression analysis to evaluate the prognostic performance of the risk score. The results of multivariate Cox analysis demonstrated that the risk score could be regarded as an independent predictive factor for OS (HR = 1.69, 95% CI = 1.08–2.64; p = 0.020) (Fig. 5D).

### Nomogram analysis

The TNM stage and the risk score are listed as candidate indicators for nomogram construction due to their excellent prognostic ability. The nomograms established with the risk score (C-index = 0.661) and TNM stage (C-index = 0.630) showed good predictive potential in the TCGA cohort. Finally, we combined the risk score and TNM stage to establish an optimal nomogram for survival prediction (C-index = 0.698).

## Discussion

EA has a poor prognosis and is a typical immunogenic malignant tumor. Immune regulation plays a crucial role in the progression of esophageal cancer. Because of the high heterogeneity of the tumor microenvironment (TME) in EA, the response to therapy is highly varied among patients. Therefore, it is critical to screen prognostic markers related to the immune environment of ESCC. In this study, we found that miR-3648 was significantly upregulated in EA tumor tissues compared with normal tissues and identified 4 miR-3648-targeted genes related to the TME. We than developed a risk score signature for EA. This signature could be used as a potential prognostic biomarker in EA patients.

As a generally recognized carcinogen, miR-3648 is abnormally expressed in many cancers. [9, 10] A previous study suggested that miR-3648 can promote the proliferation of prostate cancer cells. [11] Upregulated miR-3648 was reported to induce the invasion and metastasis of human bladder cancer. [9] Consistent with previous findings, miR-3648 has a sharply high expression in tumor tissues compared to normal tissues in EA and has good clinical performance. Furthermore, the high expression of miR-3648 was related to the short survival time of EA patients. This finding suggests that miR-3648 might play a crucial role in the progression of EA.

The imbalance of the immune microenvironment mediated by miR-3648 in EA is still unknown. To comprehend the complexity of the EA microenvironment, we constructed a coexpression network using a

gene expression matrix to predict the level of immune cell infiltration. We found that M0 macrophages, M1 macrophages, activated dendritic cells, activated mast cells, activated memory CD4 T cells, follicular helper T cells and regulatory Tregs were the most common immune cell populations in EA tumor tissues. The immune-related genes targeted by miR-3648 most related to T cell regulatory Tregs were further identified. Tregs are an immunosuppressive subset of $CD4^+$ T cells and can suppress anticancer immunity, thus promoting tumor development and progression.[16] The correlation of miR-3648 with the tumor immune environment may be a prominent mechanism in the pathological process of EA. Therefore, mining the IRGs driven by miR-3648 and building an immune-related prognostic signature has great potential to identify high-risk patients and guide EA treatment.

We utilized the WGCNA algorithm to identify 70 IRGs based on the 170 DETGs of miR-424-5p and 13 DIICs. Cox regression analysis was performed to determine the association between the expression of IRGs and survival, and 4 genes were identified as being associated with overall survival. These genes have previously been reported to be associated with different malignant tumors in various ways. For example, C10orf55 has been found to be a potential biomarker to predict relapse of acute myeloid leukemia. [17] DLL4 plays a key role in tumor angiogenesis, functioning as a negative regulator of tumor angiogenesis and upregulated in the tumor vasculature. [18, 19] A previous study indicated that DLL4 could be an independent prognostic factor for predicting the overall survival of patients with nonsmall cell lung cancer and was correlated with immunocyte infiltration. [20] PANX2 is a channel-forming glycoprotein and is associated with many common diseases, including cancers. [21] PANX2 was reported as a growth regulator in glioma cells [22] and was associated with the survival of clear cell renal cell carcinoma patients [23]. NKAIN1 was overexpressed in urinary sediments from patients with prostate cancer and could be a promising biomarker for the early diagnosis of prostate cancer. [24] Our findings demonstrated that all these genes may be crucial biomarkers for predicting survival outcome in EA. Both *in vitro* and *in vivo* experiments are necessary to validate the expression of these genes and their roles in tumor cell proliferation, metastasis and invasion. Further clinical studies are also required to determine whether these genes are independent prognostic biomarkers as well as their association with immunotherapy efficacy.

We constructed an IRG-based 4-gene (C10orf55, DLL4, PANX2 and NKAIN1) prognostic signature. The prognostic signature and nomogram show extremely high accuracy in identifying high-risk patients and predicting prognosis. The novel immune-related prognostic signatures could be used as potential prognostic biomarkers of EA. Based on this signature, we could conveniently monitor the infiltration of immune cells and further reduce the degree of immune response. Thus, this signature could reflect these changes in the TME from different aspects and has the potential to be appropriate for rational diagnosis and individualized treatment.

Although the 4-gene signature achieves an independent prognosis in EA, our study still has certain limitations. First, the study included only the DETGs of miR-3648, and the prognostic signature does not represent the genome-wide transcription profile of EA. Second, given that the study was retrospective in nature, the risk score needs to be validated in a large cohort. Third, in this study, the associations between

the 4 genes and the biological mechanisms of EA were not clarified. Finally, as a retrospective study, our results should be verified in prospective multicenter clinical trials.

# Conclusions

In summary, we conducted WGCNA and the deconvolution algorithm of CIBERSORT to evaluate fluctuations in the immune microenvironment and identified 70 miR-3648-targeted genes related to immune cell infiltration. Subsequently, a novel 4-gene signature (C10orf55, DLL4, PANX2 and NKAIN1) with good prognostic independence was established. A nomogram combining risk score and TNM stage was determined for survival prediction with great accuracy. The novel immune-related 4-gene signature is expected to become a potential prognostic marker of EA.

# Declarations


### Acknowledgements

We want to thank acknowledge the database available to us for this study


### Authors' contributions

Zhang, Yin and Qian proposed the study concept, design, and drafted the manuscript. Zhang and Yin collected, analyzed, and interpreted the data. Qian and Xu participated in revising the manuscript. All authors contributed to the article and approved the submitted version.


### Funding

Not applicable.


### Availability of data and materials

The dataset supporting the conclusions of this article is available in The Cancer Genome Atlas (https://portal.gdc.cancer.gov/).

### Ethics approval and consent to participate

This study was approved by Ethics Committee of Renji Hospital.

### Consent for publication

All authors have seen and agreed to publish.

### Competing interests

The authors declare no potential conflicts of interest.

**Author details**

[1] Department of Thoracic Surgery, Renji Hospital, School of Medicine, Shanghai Jiaotong University, Shanghai 200127, China. [2] Division of Gastroenterology and Hepatology, Key Laboratory of Gastroenterology and Hepatology, Ministry of Health, Renji Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai Institute of Digestive Disease, Shanghai, China.

# References

1. Siegel RL, Miller KD, Fuchs HE, and Jemal A, *Cancer Statistics, 2021.* CA Cancer J Clin, 2021. **71**(1): p. 7-33.
2. Smyth EC, Lagergren J, Fitzgerald RC, Lordick F, Shah MA, Lagergren P, et al., *Oesophageal cancer.* Nat Rev Dis Primers, 2017. **3**: p. 17048.
3. Cohen DJ and Leichman L, *Controversies in the treatment of local and locally advanced gastric and esophageal cancers.* J Clin Oncol, 2015. **33**(16): p. 1754-9.
4. Markar SR, Gronnier C, Pasquer A, Duhamel A, Behal H, Thréreaux J, et al., *Discrepancy Between Clinical and Pathologic Nodal Status of Esophageal Cancer and Impact on Prognosis and Therapeutic Strategy.* Ann Surg Oncol, 2017. **24**(13): p. 3911-3920.
5. Hegde PS and Chen DS, *Top 10 Challenges in Cancer Immunotherapy.* Immunity, 2020. **52**(1): p. 17-35.
6. Tamborero D, Rubio-Perez C, Muiños F, Sabarinathan R, Piulats JM, Muntasell A, et al., *A Pan-cancer Landscape of Interactions between Solid Tumors and Infiltrating Immune Cell Populations.* Clin Cancer Res, 2018. **24**(15): p. 3717-3728.
7. Fabian MR and Sonenberg N, *The mechanics of miRNA-mediated gene silencing: a look under the hood of miRISC.* Nat Struct Mol Biol, 2012. **19**(6): p. 586-93.
8. Rashid F, Awan HM, Shah A, Chen L, and Shan G, *Induction of miR-3648 Upon ER Stress and Its Regulatory Role in Cell Proliferation.* Int J Mol Sci, 2017. **18**(7).
9. Sun W, Li S, Yu Y, Jin H, Xie Q, Hua X, et al., *MicroRNA-3648 Is Upregulated to Suppress TCF21, Resulting in Promotion of Invasion and Metastasis of Human Bladder Cancer.* Mol Ther Nucleic Acids, 2019. **16**: p. 519-530.
10. Hu Y, Dingerdissen H, Gupta S, Kahsay R, Shanker V, Wan Q, et al., *Identification of key differentially expressed MicroRNAs in cancer patients through pan-cancer analysis.* Comput Biol Med, 2018. **103**: p. 183-197.
11. Xing R, *miR-3648 Promotes Prostate Cancer Cell Proliferation by Inhibiting Adenomatous Polyposis Coli 2.* J Nanosci Nanotechnol, 2019. **19**(12): p. 7526-7531.
12. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al., *limma powers differential expression analyses for RNA-sequencing and microarray studies.* Nucleic Acids Res, 2015. **43**(7): p. e47.

13. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al., *Robust enumeration of cell subsets from tissue expression profiles.* Nat Methods, 2015. **12**(5): p. 453-7.
14. Langfelder P and Horvath S, *WGCNA: an R package for weighted correlation network analysis.* BMC Bioinformatics, 2008. **9**: p. 559.
15. Yang Z, Zi Q, Xu K, Wang C, and Chi Q, *Development of a macrophages-related 4-gene signature and nomogram for the overall survival prediction of hepatocellular carcinoma based on WGCNA and LASSO algorithm.* Int Immunopharmacol, 2021. **90**: p. 107238.
16. Nishikawa H and Sakaguchi S, *Regulatory T cells in tumor immunity.* Int J Cancer, 2010. **127**(4): p. 759-67.
17. Walker CJ, Mrózek K, Ozer HG, Nicolet D, Kohlschmidt J, Papaioannou D, et al., *Gene expression signature predicts relapse in adult patients with cytogenetically normal acute myeloid leukemia.* Blood Adv, 2021. **5**(5): p. 1474-1482.
18. Liu Z, Fan F, Wang A, Zheng S, and Lu Y, *Dll4-Notch signaling in regulation of tumor angiogenesis.* J Cancer Res Clin Oncol, 2014. **140**(4): p. 525-36.
19. Perez-Fidalgo JA, Ortega B, Simon S, Samartzis EP, and Boussios S, *NOTCH signalling in ovarian cancer angiogenesis.* Ann Transl Med, 2020. **8**(24): p. 1705.
20. Li R, Liu X, Zhou XJ, Chen X, Li JP, Yin YH, et al., *Identification and validation of the prognostic value of immune-related genes in non-small cell lung cancer.* Am J Transl Res, 2020. **12**(9): p. 5844-5865.
21. Penuela S, Harland L, Simek J, and Laird DW, *Pannexin channels and their links to human disease.* Biochem J, 2014. **461**(3): p. 371-81.
22. Lai CP, Bechberger JF, and Naus CC, *Pannexin2 as a novel growth regulator in C6 glioma cells.* Oncogene, 2009. **28**(49): p. 4402-8.
23. Kim KM, Hussein UK, Bae JS, Park SH, Kwon KS, Ha SH, et al., *The Expression Patterns of FAM83H and PANX2 Are Associated With Shorter Survival of Clear Cell Renal Cell Carcinoma Patients.* Front Oncol, 2019. **9**: p. 14.
24. Leyten GH, Hessels D, Smit FP, Jannink SA, de Jong H, Melchers WJ, et al., *Identification of a Candidate Gene Panel for the Early Diagnosis of Prostate Cancer.* Clin Cancer Res, 2015. **21**(13): p. 3061-70.

# Tables

Table 1 Cox analysis of miRNAs.

| miRNA | Coefficient | HR (95% CI for HR) | P value |
|---|---|---|---|
| miR-550a-1 | 0.37 | 1.4 (1.0-2.0) | 0.030 |
| miR-550a-2 | 0.30 | 1.4 (1.0-1.8) | 0.043 |
| miR-3682 | 0.31 | 1.4 (1.0-1.9) | 0.043 |
| miR-503 | 0.35 | 1.4 (1.0-2.0) | 0.040 |
| miR-3648 | 0.24 | 1.3 (1.1-1.5) | 0.002 |

HR, hazard ratio; CI, confidence interval.

Table 2 Cox analysis of 4 genes in the signature.

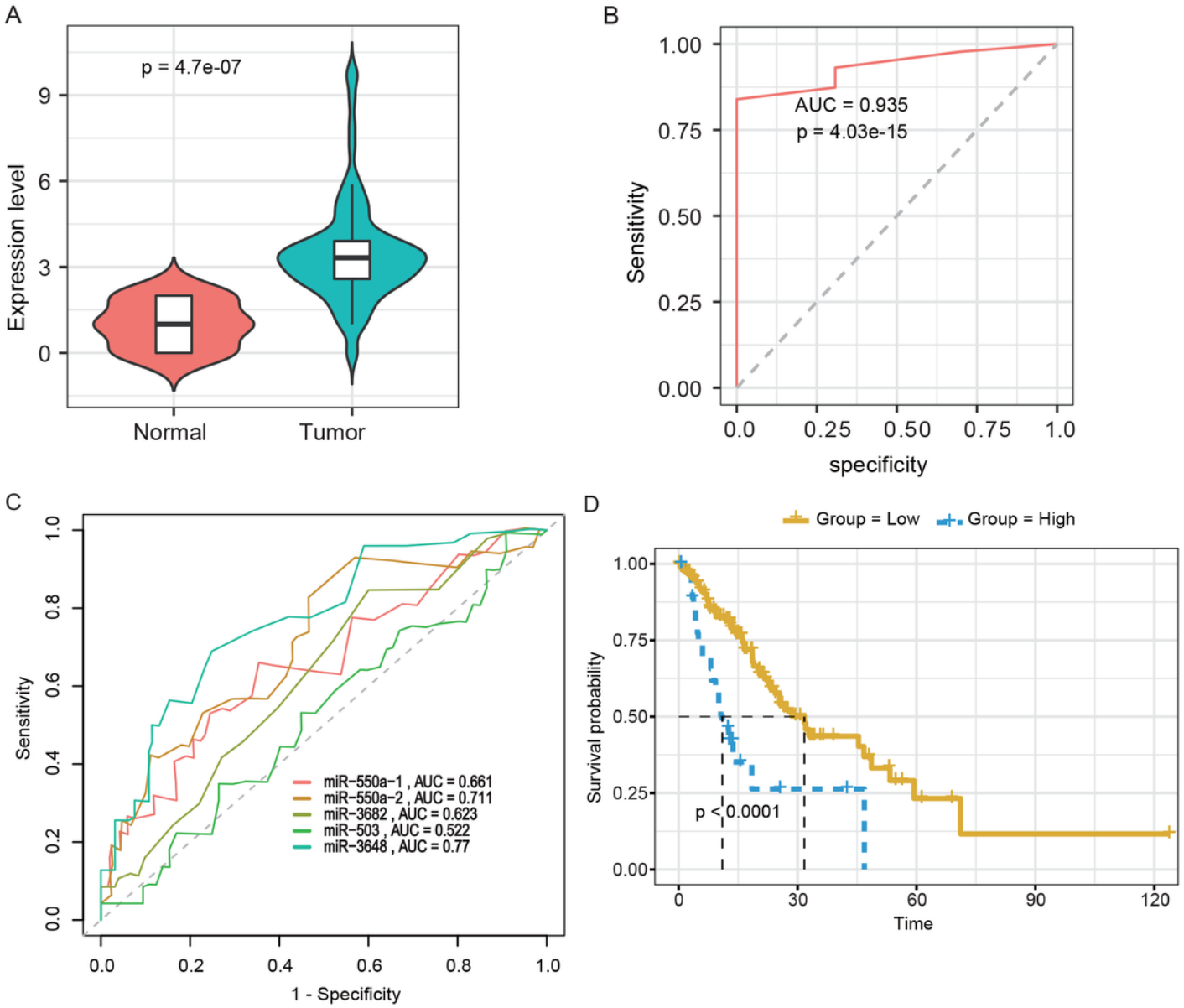
| Gene symbol | Gene full name | Coefficient | HR (95% CI for HR) | P value |
|---|---|---|---|---|
| C10orf55 | Chromosome 10 putative open reading frame 55 | 0.29 | 1.3 (1.1-1.7) | 0.017 |
| DLL4 | Delta-like canonical Notch ligand 4 | 0.39 | 1.5 (1.1-2.0) | 0.018 |
| PANX2 | Pannexin 2 | 0.15 | 1.2 (1.0-1.3) | 0.030 |
| NKAIN1 | Na+/K+ transporting ATPase interacting 1 | 0.16 | 1.2 (1.0-1.4) | 0.046 |

HR, hazard ratio; CI, confidence interval.

# Supplementary

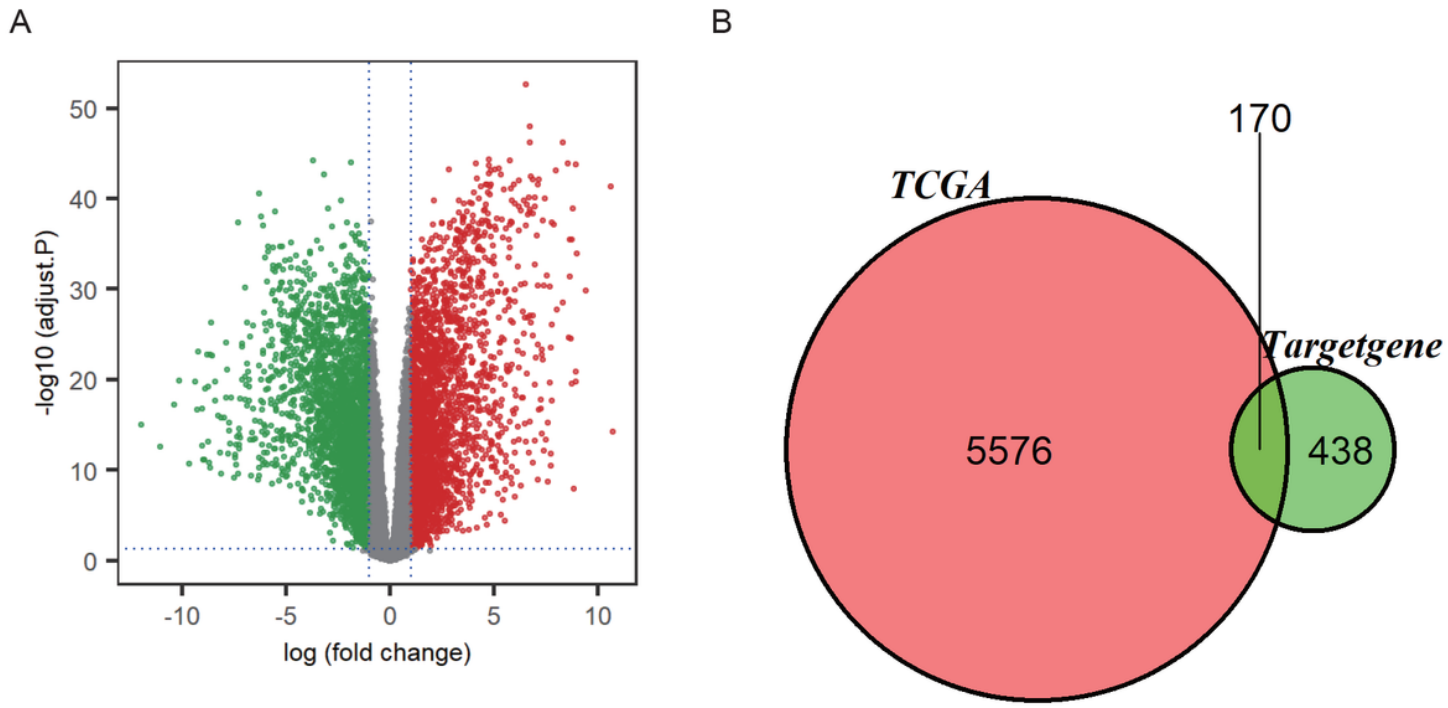
Additional file 1: Fig. S1 is not available with this version

# Figures



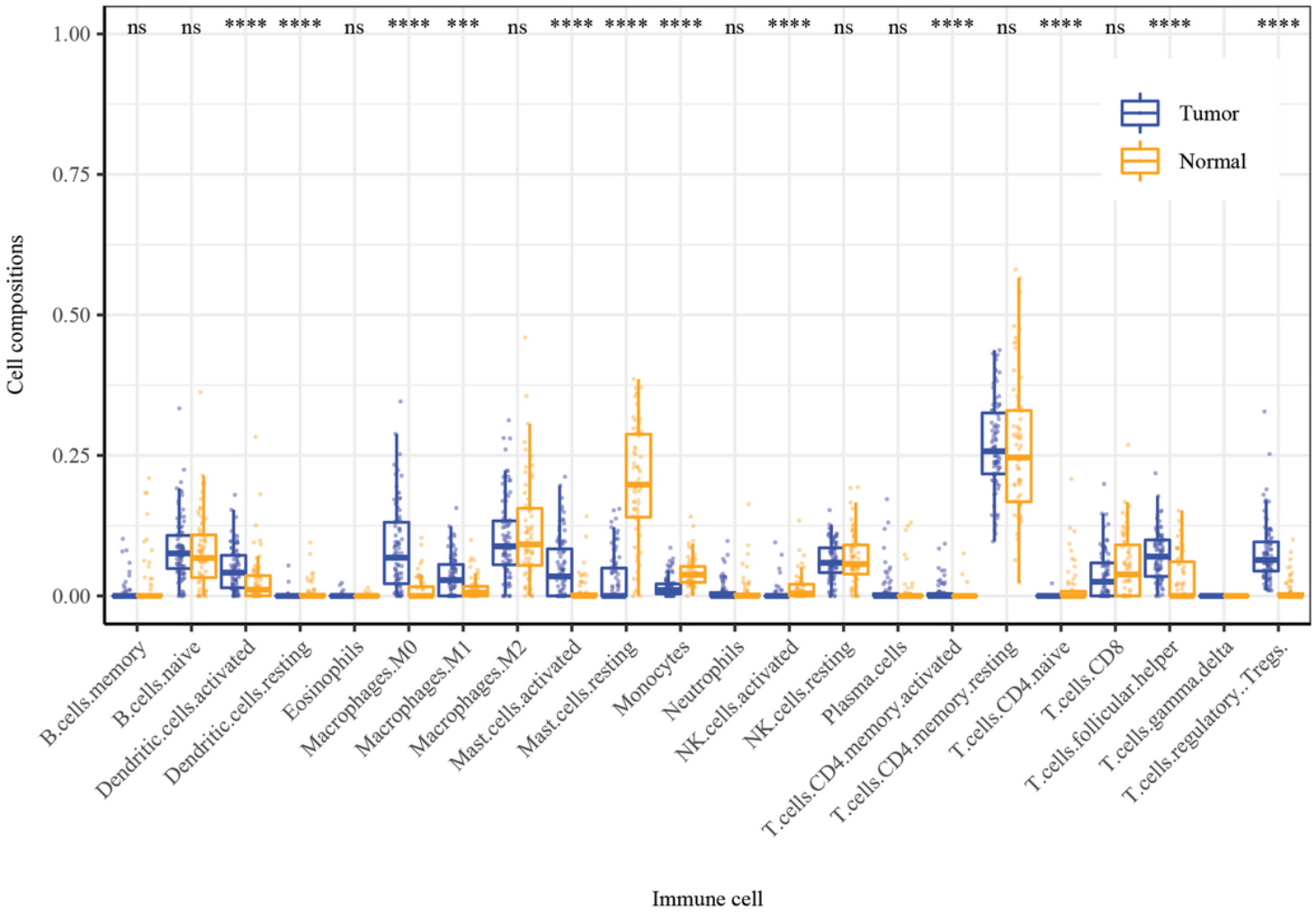

## Figure 1

Clinical performance of miRNAs. (A) Expression level of miR-3648 in EA tumor/normal tissues; (B) ROC curve of miR-3648 in diagnosis; (C) ROC curves for miRNAs to predict 1-year survival; (D) Kaplan-Meier curves of high/low expressed groups of miR-3648 with overall survival time.



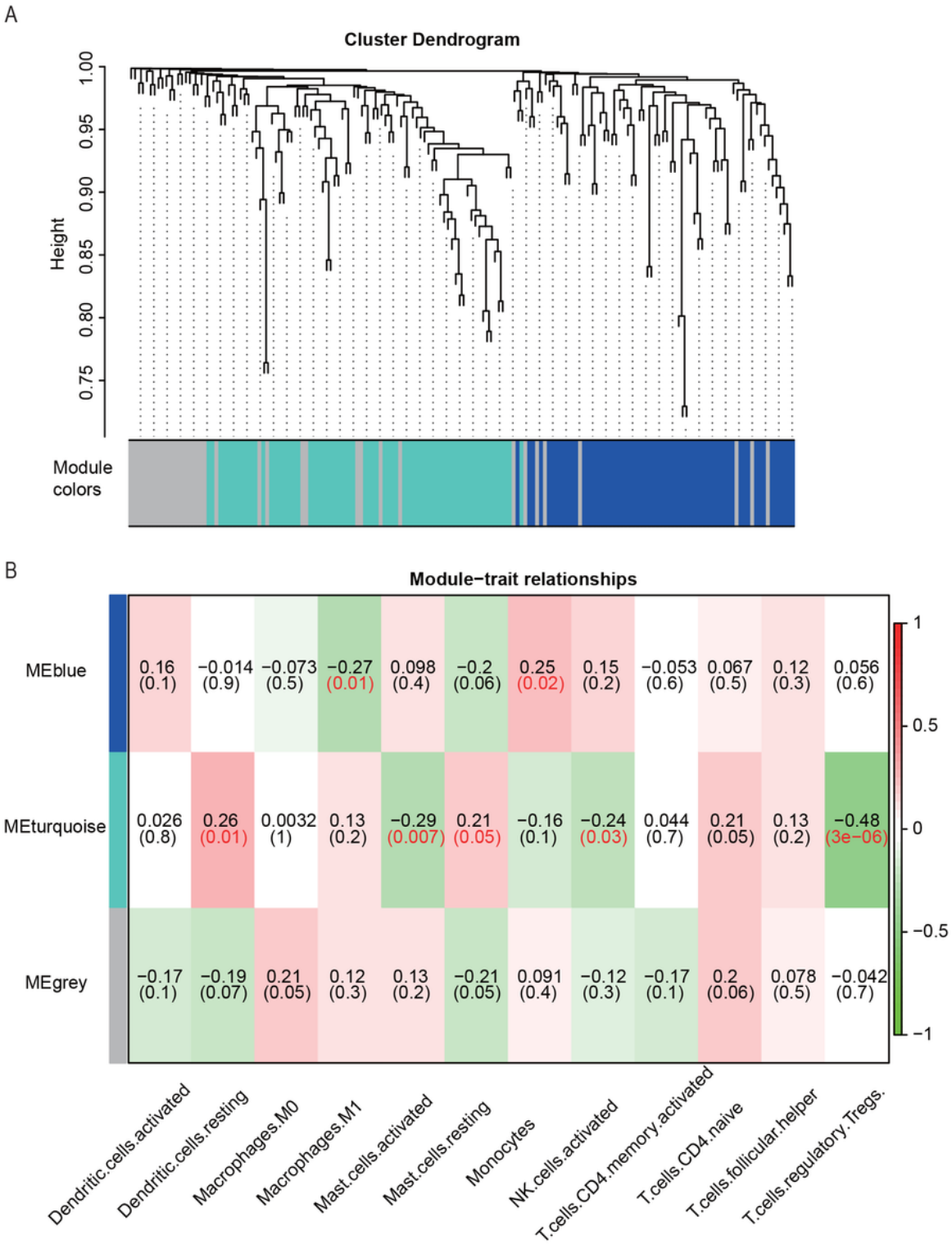

## Figure 2

miR-3648-targeted gene identification. (A) Volcano diagram of differentially expressed genes of esophageal adenocarcinoma; (B) Venn diagram of miR-3648-targeted genes in differentially expressed genes;



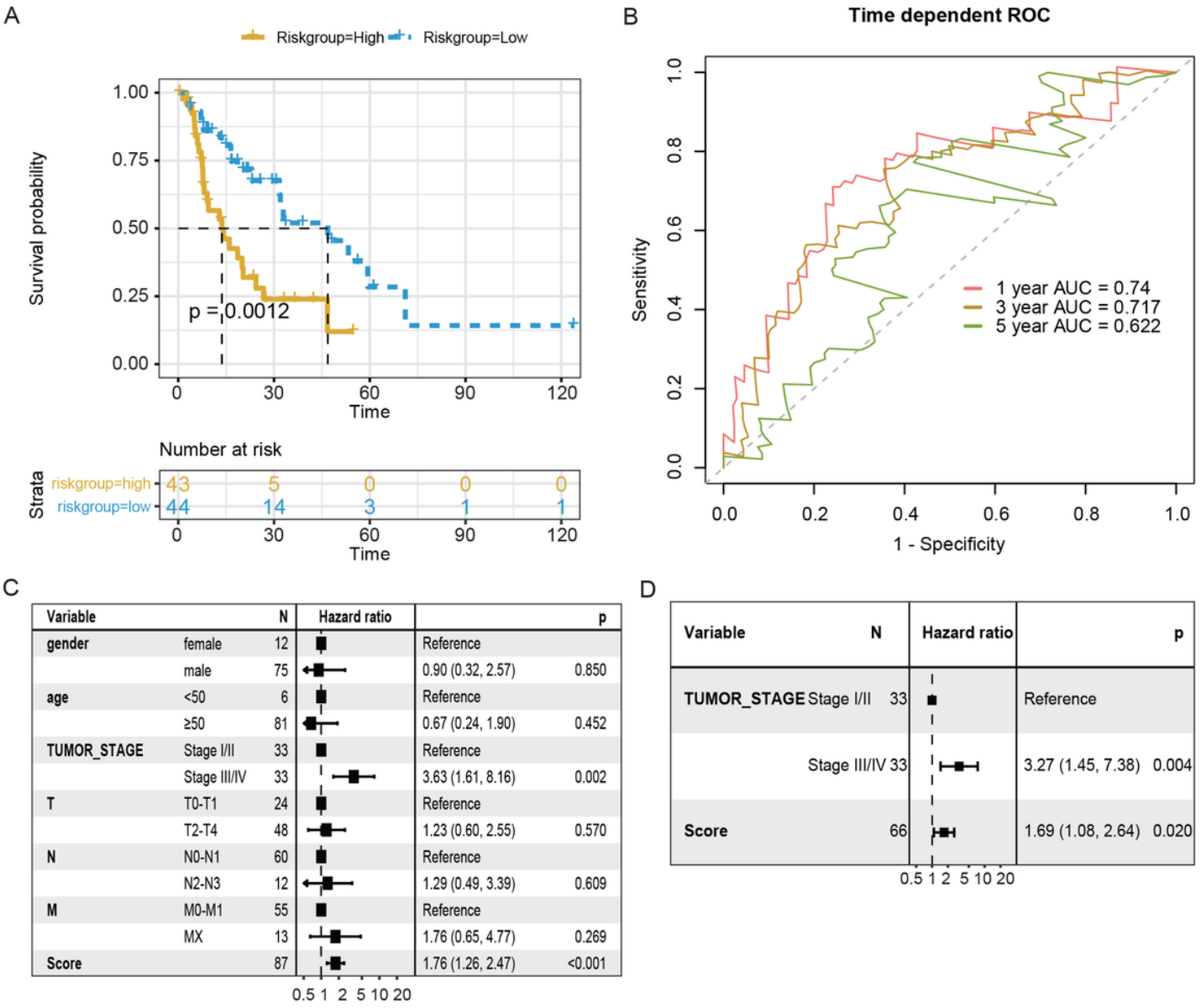

## Figure 3

The profiles of immune infiltration of esophageal adenocarcinoma.




## Figure 4

Weighted gene coexpression network analysis (WGCNA) of the DETGs of miR-3648. (A) 2 modules and 37 oligogenes were clustered. Each branch of the tree diagram represents genes, and genes clustered into the same module are assigned the same module color. The oligogenes are summarized in the gray module. (B) Correlation analysis between modules and immune infiltration. P values < 0.05 are colored red.




## Figure 5

Prognostic value for risk score. (A) Kaplan-Meier curves of high/low risk score groups with overall survival time; (B) ROC curves for the risk score to predict 1-, 3-, and 5-year survival; (C) Univariate and (D) multivariate Cox regression analysis of the risk score and clinical characteristics.

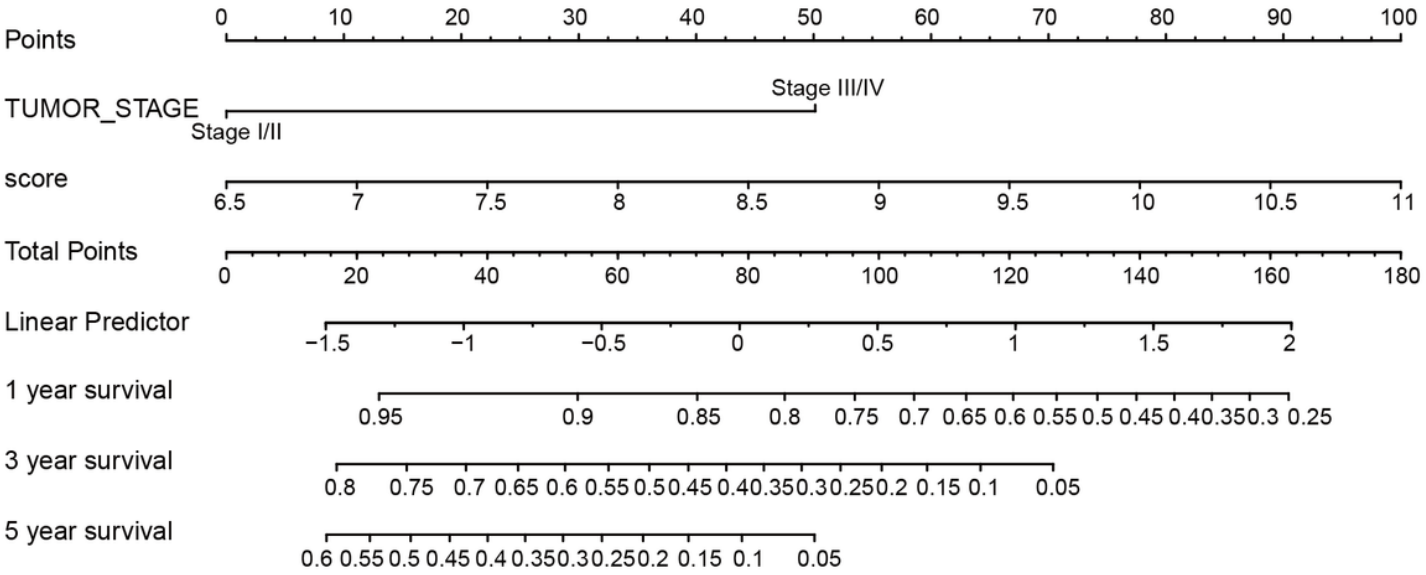



**Figure 6**

Nomogram for survival prediction.

# Supplementary Files

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

- TCGAvsGTExcounts.txt
- Patientsinformation.csv
- ESCA.miRseqrawcounts.txt