Identification of a Hypoxia-Related Gene Signature and Establishment of a Nomogram for Predicting Prognosis in Esophageal Squamous Cell Carcinoma

Wanyi Xiao
Tianjin Cancer Institute: Tianjin Tumor Hospital
Peng Tang
Tianjin Cancer Institute: Tianjin Tumor Hospital
Zhilin Sui
Tianjin Cancer Institute: Tianjin Tumor Hospital
Xianxian Wu
Cancer Hospital and Shenzhen Hospital
Yueyang Yang
Tianjin Cancer Institute: Tianjin Tumor Hospital
Ningning Zhu
Tianjin Cancer Institute: Tianjin Tumor Hospital
Youming Han
Tianjin Cancer Institute: Tianjin Tumor Hospital
Lei Gong
Tianjin Cancer Institute: Tianjin Tumor Hospital
Zhentao Yu (yuzhentao@tjmuch.com)
Tianjin Cancer Institute: Tianjin Tumor Hospital https://orcid.org/0000-0002-5785-8492
Hongdian Zhang
Tianjin Cancer Institute: Tianjin Tumor Hospital

Primary research

Keywords: esophageal squamous cell carcinoma, hypoxia, hypoxia-related genes, nomogram, prognosis

Posted Date: November 2nd, 2021

DOI: https://doi.org/10.21203/rs.3.rs-1008316/v1

Abstract

Background

Esophageal squamous cell carcinoma (ESCC), the major subtype of esophageal cancer in China, has a dismal prognosis. Tumor hypoxia, a typical characteristic of many solid tumors, leads to tumor invasion and poor prognosis. We aimed to develop a hypoxia-based gene signature to assist in the diagnosis and prognosis prediction of ESCC.

Methods

We integrated gene expression files from the Gene Expression Omnibus (GEO) and The Cancer Genome Atlas (TCGA) databases to identify potential hypoxia-related genes (HRGs). Euclidean-based consensus clustering was performed on the GSE53625 ESCC samples. Immune cell infiltration, differential expression, and Kaplan-Meier survival analyses were performed. Candidate modules were screened using weighted gene coexpression network analysis and then intersected with differentially expressed genes from clustered subgroups to construct the gene signatures that were verified in internal and external datasets. A nomogram was developed with risk scores and clinicopathological variables. Finally, immunohistochemistry was used to detect the protein expression level of relevant genes. Additionally, the relationships between risk scores and the tumor microenvironment were explored.

Results

A total of 5 potential differentially expressed HRGs and 6 critical HRGs were identified. A prognostic model was constructed by screening key HRGs, and higher risk scores were associated with poorer prognosis. The nomogram based on risk scores predicted the 1-, 3-, and 5-year survival rates well, thus determining prognostic risk stratification. The p53, Wnt, and hypoxia signaling pathways may be some of the regulatory mechanisms of hypoxia associated with the tumor microenvironment. Furthermore, the high expression of BGN and low expression of IL-18 in ESCC tissues was verified and was associated with a shorter survival time.

Conclusions

Our study systematically explored HRGs and determined the prognostic value of a 6-hypoxia gene signature. A prognostic model was developed and validated based on this, providing potential prognostic predictors and therapeutic targets for ESCC.

Background

Esophageal squamous cell carcinoma (ESCC) is the predominant type of esophageal cancer in China, accounting for more than 90% of pathological types[1, 2]. Due to the heterogeneous biological characteristics of ESCC[3], there are various mechanisms that promote tumor cell migration and invasion leading to tumor progression, which is one of the important reasons for the poor prognosis of patients.

Rapid growth and massive angiogenesis in many tumors increase metabolic demand and oxygen consumption[4, 5], resulting in hypoxia. The hypoxia has received widespread attention as a hallmark of malignant tumor microenvironment. Hypoxia is closely related to abnormal biological behaviors, leading to aberrant gene expression changes[6], anti-apoptosis[7], suboptimal treatment outcomes[8] and ultimately to poor prognosis as

tumor cells adapt to hypoxia[9]. It has been found that ESCC cancer tissues are invariably hypoxic. Several studies have demonstrated that hypoxia is involved in the migration and aggressive development of ESCC and is associated with increased malignancy and poor prognosis[10]. And hypoxia plays a key role in metastasis and proliferation. Extensive reviews have reported subsequent proteomic changes that exacerbate tumor progression and metastasis[11, 12]. Therefore, focusing on hypoxia-inducible factors and related markers may provide a more efficient way to identify new biomarkers for ESCC patients.

With advances in bioinformatics and high-throughput technologies[13], analyses that target the regulation networks, functional pathways key molecules between genes and phenotypes gene expression levels are improving the current understanding of the molecular mechanisms underlying the biological behaviors. Since tumor hypoxia cannot be predicted based on clinical size, stage or differentiation, molecular biomarkers capable of assessing the hypoxic status of ESCC are needed. Such biomarkers can be used to detect tumor hypoxia at an early stage, assess the response to treatment, predict the prognosis of ESCC patients, and select appropriate personalized treatment regimens for patients.

It has been reported that hypoxia and the tumor immune microenvironment also interact with each other[14]. The adaptation of tumor cells to the hypoxic environment leads not only to aberrant gene expression but also to the formation of the tumor microenvironment (TME) and tumor immune microenvironment. Exploring the effects of hypoxia on both and the potential remodeling effects will help reveal the complex interactions among them and provide new clues for the treatment of ESCC.

In this study, we aimed to identify and validate hypoxia-related genes (HRGs) in ESCC and their relevance to patient prognosis and the TME. Then, we aimed to develop a prognostic prediction nomogram based on the HRG signature by investigating the expression pattern of HRGs to assist in the prognostic risk prediction of patients with esophageal cancer and to facilitate personalized treatment.

Methods

Data acquisition and processing

We integrated all public RNA-seq data from the Gene Expression Omnibus (GEO) database and The Cancer Genome Atlas (TCGA) database, including the GSE53625 and TCGA ESCC datasets, to identify potential hypoxia-related genes in ESCC. Clinical information for these samples, including follow-up data of 179 ESCC and 179 paired normal control samples, was downloaded from the GSE53625 dataset based on GPL18109. Additionally, the RNA-FPKM data and clinical data of 82 ESCC samples were retrieved for subsequent external validation analysis using the TCGA data portal. Corresponding normal samples included 11 TCGA paracancerous samples and 1445 samples of 54 noncancer sites in GTEx. The expression file of genes with the median value was calculated when more than one expression file had matched patients.

Tissue microarrays (TMAs) made by Shanghai Outdo Biotech Co., Ltd. were used for immunohistochemical staining (IHC). Samples from 232 ESCC tissues and 58 adjacent normal esophageal tissues between 2009 and 2010 at Tianjin Medical University Cancer Institute and Hospital were obtained with reliable information on survival. Patients who had received neoadjuvant chemotherapy or radiotherapy and those with cancer types other than ESCC were excluded from the study. Specimens were taken from the center of the tumor. Paired normal tissues were taken from surgically dissected tissues ～5 cm away from the tumor. Among the patients, there were

192 males and 40 females with a median age of 68 years old. All patients were followed up until September 2016 with a median survival of 31 months. Enrolled patients authorized the collection and use of his or her material and signed written informed consent forms in advance. Ethical approval was obtained from the Research Ethics Committee of Tianjin Medical University Cancer Institute and Hospital.

Extraction of HRGs

The Cancer Single-Cell State Atlas (CancerSEA)[15] is a multifunctional website designed to comprehensively explore the different functional states of cancer at the single-cell level, covering 41,900 single cancer cells from 25 cancer types in 14 functional states, including hypoxia. The list of 112 hypoxia genes was retrieved by enquiring “hypoxia” in the gene pool of all cancer types. The Molecular Signatures Database (MSigDB)[16], a gene set database, defines many existing gene sets as gene collections from the perspective of location, function, metabolic pathways, target binding, etc. The “HALLMARK-HYPOXIA” gene set containing 200 named genes was derived from sets of effector features representing well-defined biological states or processes. A total of 270 unique elements were subsequently obtained by intersecting the abovementioned hypoxia gene sets from the two sources and were identified as key genes involved in hypoxia activity.

Differential analysis of gene expression

Differential gene expression was determined using the “limma” package in R software, accounting for the nonindependence of samples from the same participant using limma’s duplicateCorrelation. Multiple comparisons were adjusted using the Benjamini–Hochberg false discovery rate (FDR). An FDR < 0.05 and a log2 |fold change| > 2 were deemed as cutoff values for differentially expressed genes (DEGs).

Identification of hypoxia subgroups associated with prognosis by consensus clustering

Consensus clustering is an algorithm used to determine the optimal number of clusters in a gene microarray dataset. To identify distinct subgroups of ESCC for optimal classification purposes, Euclidean-based consensus clustering was performed on the GSE53625 dataset by the K-means algorithm using the ConsensusClusterPlus R package. The clustering process was performed 500 times, which was user-specified, with each iteration containing 80% of the samples and the number of clusters set to 2–10. Each algorithm was run, and the consensus values and the stability of the clustering results were assessed by applying the given clustering method to random subsets of data.

After executing ConsensusClusterPlus, the graphical output results included heatmaps of the consensus matrices, which displayed the clustering results, consensus cumulative distribution function (CDF) plots, and Delta area plots, which allowed us to determine an approximate number of clusters. The optimal number of clusters is determined by the CDF plots and the Delta plots of the relative change in the area under the curve of the cumulative distribution function of the consensus scores. The optimal number of clusters is usually chosen as the value of K corresponding to the last inflection point of the CDF and the smallest slope of Delta.

Weighted gene coexpression network analysis (WGCNA)[17]

WGCNA was used to construct the gene coexpression network to identify required gene modules by performing significant association analysis with the phenotype of interest. The GSE53625 gene expression file for 179 ESCC samples was used as a data source to construct a scale-free network by calculating the connection strength

between genes using the R package “WGCNA”. Scale-free R^2 ranging from 0 to 1 was used to determine a scale-free topology model. To minimize the effects of noise and spurious associations, the adjacency matrix was transformed into a topological overlap matrix (TOM), which was used to form modules by dynamic tree cut. Here, we set the minimal module size to 20 and the cut height to 0.25. Screening was done to identify candidate HRGs for next-step analysis, which included functional enrichment analysis of modules (to see if their functional characteristics are consistent with the purpose of the study), correlation analysis between modules and clinical phenotypes (to determine the modules with the highest correlation with the trait of interest) and extraction of key information from modules (to find the core genes of modules and to predict gene function).

Construction and evaluation of the prognosis prediction model of ESCC

The stepAIC algorithm running in the R “MASS” package was used to construct a prognostic model that incorporated the best variables to obtain the optimal model. The algorithm was run based on the combination of expression profiles that intersected from selected modules and DEGs and prognostic information from 179 ESCC samples. The risk score of each sample was calculated according to the expression levels of the samples and then divided into two groups, namely, high and low scores, according to the median value after sorting. Half of the 179 ESCC samples of GSE53625 were randomly taken as the training group to construct the model. While the other half (internal validation set) and all of the GSE53625 datasets were used as testing datasets to assess the robustness of the prognostic model. To assess the accuracy and predictive ability of the model, the same coefficient as the training set was used for the external validation dataset, that is, TCGA. Kaplan-Meier curve analysis was further conducted to evaluate the relationship between the risk score and overall survival. The area under the curve (AUC) of the receiver operating characteristic (ROC) curves was calculated and compared to examine the classifier performance using the “time ROC R” package.

Determining the classification features using Cox proportional risk regression models

In this study, survival risk scores were combined with clinical features for multivariate Cox regression analysis, and significant prognostic variables obtained from the resulting multivariate Cox regression model were then introduced into the final nomogram model. The nomogram may be helpful for calculating a certain survival probability based on total points by summing the assigned scores of independent factors. Calibration curves were plotted to assess the consistency between the predicted survival probabilities and the actual survival proportions at 1, 3, and 5 years, and higher overlap with the 45-degree standard curve indicated better predictive agreement within the model. In this study, we used the “rms”, “foreign” as well as “survival” packages in R software for nomogram construction and calibration curve plotting.

Construction of the protein-protein interaction (PPI) network and functional analysis

We further used the STRING website (http://string-db.org/) to explore the protein interaction relationship of hypoxia-related DEGs based on the PPI network. To investigate the underlying function of differentially expressed HRGs, we used the “clusterProfiler” R package for Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis and for Gene Ontology (GO) analysis with respect to three domains: cellular component (CC), biological process (BF), and molecular function (MF).

Implementation of gene set enrichment analysis (GSEA)

To further explore biological signaling pathways between differentially activated consensus subgroups and the high/low-risk groups, GSEA was conducted using GSEA software employing GSE53625 data. P < 0.05 and a Q value less than 0.25 were considered to denote significant enrichment.

The ESTIMATE and CIBERSORT algorithms

The stromal, immune and ESTIMATE score for each patient was calculated with the ESTIMATE algorithm through the R “estimate” package. The fraction of 22 immune cell types for each sample was assessed through cell type identification by estimating the relative subsets of RNA transcripts (CIBERSORT; https://cibersort.stanford.edu/).

Immunohistochemical analysis based on TMA data

TMAs were used for IHC to examine the protein expression level of selected HRGs. In brief, the tissue sections were deparaffinized (70℃, 2 hours), rehydrated and subjected to antigen repair with heated antigen retrieval solution (10 mmol/l sodium citrate buffer, pH 6.0) (100℃, 10 minutes). The activity of endogenous peroxidase was blocked with 0.3% H2O2 and 5% goat serum. Then, the sections were incubated with primary antibodies (1:50, BGN, Proteintech; 1:100, IL-18, ORIGENE) overnight at 4°C and then incubated with a biotinylated secondary antibody for 20 minutes at room temperature based on the reagent instructions. Diaminobenzidine (DAB, Zhongshan Inc.) was used as a chromogen and produced a brown color, and then samples were counterstained with hematoxylin. The stained slides were scanned using an automatic slice scanning system.

Two experienced pathologists who were blinded to the clinical data scored the staining results of each case. Staining was assessed independently according to the intensity of staining (no staining, 0; weak, 1; moderate, 2; and strong, 3) and the percentage of positively stained tumor cells (0%, 0; 1 to 30% positive, 1; 31% to 70% positive, 2; 71-100% positive, 3). A total score of 0 to 9 was obtained by multiplying the results of the staining intensity and staining percentage scores, and tissues with a total score of 0 to 3 were considered to have low expression, while tissues with a score of 3 to 9 were considered to have high expression.

Statistical analysis

Bioinformatics analysis and statistical analysis were conducted using R (version 4.0.2). Comparisons between two groups were presented via the Wilcoxon rank-sum test and chi–squared test, while multiple comparisons were assessed via the Kruskal–Wallis test. The cutoff point of each subgroup was identified by the survminer package in R. Kaplan–Meier curves for overall survival (OS) analysis are presented between different subgroups, followed by the log-rank test. ROC curves for 1-, 3- and 5-year survival were delineated to evaluate the predictive efficacy of the risk score. The *P*-values were corrected by Bonferroni’s test. A two-sided $P < 0.05$ was considered statistically significant.

Results

Identification of differentially expressed hypoxia genes associated with OS

After data preprocessing, a total of 3166 DEGs in the GEO set and 4808 DEGs in TCGA were obtained. These are illustrated in the volcano plots in Figure 1A and 1B and heatmaps in Supplementary Figure 1. We then intersected the DEGs with the list of 270 HRGs obtained from the collection of the CancerSEA database and MSigDB database to obtain differentially expressed hypoxia genes. As shown in the Venn diagrams in Figure 1C, 30

unique elements (15 upregulated and 15 downregulated) were identified and consistently differentially expressed in the GEO training set and TCGA validation set (Supplementary Table 1).

A subsequent PPI network was constructed by uploading the aforementioned 30 target genes to the STRING online database and visualized using Cytoscape to elucidate their interactions (Figure 1D). In the functional analysis of DEGs, the top 6 enriched GO annotations and KEGG pathways in each category were visualized intuitively as Circos plots (Figure 1E). The expression data of 5 real hub genes extracted from univariate analysis ($P < 0.05$) were used to assess the impact of these variables on the survival outcome and were visualized with forest plots. As shown in Figure 1F, SLC2A1, PGM2, SULT2B1, and CA9 all had a lower risk of death except BGN, suggesting that BGN may be expressed at higher levels in tumor tissues than in normal tissues.

Consensus clustering to identify distinct subgroups and intercluster prognosis analysis

A total of 179 tumor samples from the GSE53625 dataset were consistently clustered based on expression similarity, and the correlation between hypoxic conditions and clinical phenotypes was investigated. After executing ConsensusClusterPlus, we obtained the cluster consensus and item-consensus results. Considering the minimum descent slope of the CDF and Delta area plots (Figure 2A-B), although K = 2 is not the last inflection point that minimizes the area under the CDF curve, the consensus matrix helps us to determine the clearest division when divided into two clusters, named C1 and C2. Compared to those of the other categorical numbers (Supplementary Figure 2), the consensus matrix graph corresponding to K = 2 showed that the distribution of 2 blue blocks on the diagonal along the white background was well defined (Figure 2C).

The results of Kaplan-Meier survival analysis revealed significant differences in prognosis among cluster 1 and cluster 2 ($P = 0.014$). As shown in Figure 2D, the samples in cluster 2 had better performance for overall survival than those in cluster 1. Sorting the samples by cluster produced the gene expression heatmap shown in Supplementary Figure 3, which indicated the composition and quantity of clustering. Gene expression patterns differed among the subgroups, suggesting the credibility of the 2 structural clusters. A heatmap corresponding to the dendrogram in Figure 2E, annotated by grade, T stage, N stage, stage, sex, and 5 real hub genes as mentioned previously, shows the distribution of these features based on cluster 1 or 2 intuitively. This result demonstrated the heterogeneity between the two clusters.

Identifying the correlations between the obtained clusters and immune infiltration

We calculated the levels of 22 immune cell types in each sample and compared their differences in the C1 and C2 subgroups (Figure 2F). The results showed that 11 types of immune cells, including CD8 T cells, resting memory CD4 T cells, follicular helper T cells, regulatory T cells (Tregs), activated NK cells, monocytes, M0 macrophages, activated dendritic cells and activated mast cells, had significantly different infiltration levels in different subgroups. Then, we performed GSEA to explore the underlying regulatory mechanism that contributed to the differences in the tumor immune microenvironment between the 2 clusters, and the results showed that they were mainly related to the activation of P53 signaling pathway molecules (Figure 2G). These results indicated that each cluster had unique gene expression and pathway characteristics, and hypoxic cells may potentially regulate the tumor immune microenvironment between the two clusters through the P53 signaling pathway.

Identification of modules associated with hypoxia by WGCNA

After sample clustering to detect outliers, the WGCNA was then restricted to 178 patients from GSE53625. Different power values (1–20) were analyzed, and the best power value of β = 6 (scale-free R^2= 0.95) was chosen for soft thresholding for subsequent coexpression scale-free network construction (Supplementary Figure 4A-C). Eight gene modules were obtained and visualized for the next analysis (Figure 3A). Module-feature relationship analysis revealed that the green module was significantly associated with consensus subgroups (C1, r = -0.46, *P* = 1e-10; C2, r=0.46, *P* = 1e-10, Supplementary Figure 4D-E), suggesting that the module is suitable for identifying the hub genes associated with C1/C2. Therefore, the green module was chosen for further analysis.

Constructing and evaluating a hypoxia-related prognosis signature

The results of differential expression analysis among the obtained clusters are shown in Figure 3B. After further screening for the most significant hub genes, 259 differentially expressed HRGs (Figure 3C) resulted from the intersection of 443 green module genes and 731 DEGs from consensus clustering analysis. The hypoxia gene signature with the best prognostic performance was constructed using the stepAIC algorithm based on the expression level of the above genes, identifying the risk score by using the 6 most relevant genes, including PNPLA1, CARD18, IL-18, SLC37A2, ADAMTS18 and FAM83C (Figure 4). The risk score was calculated as follows: risk score = (0.066265*PNPLA1 expression) + (-0.149270*CARD18 expression) + (-0.183367*IL18 expression) + (-0.037724079*SLC37A2 expression) + (0.119388782*ADAMTS18 expression) + (-0.031834954*FAM83C expression). The samples were assigned a risk score and ordered to determine whether the expression level varied systematically with the risk score (Figure 4A-B). The expression levels of the 6 genes distinctly decreased as the risk score increased, except for ADAMTS18 (Figure 4C). After the risk curves were plotted based on the risk score versus patient survival status and divided into the high- and low-risk groups based on the median risk score, a higher percentage of patient deaths was associated with high-risk patients, and a higher percentage of long-term survival was associated with low-risk patients (*P* = 0.0063). Furthermore, the prognoses differed significantly between the two groups, as shown in Figure 4D. The results of ROC curve analysis using the risk score calculated for each sample are shown in Figure 4E. The AUCs for 1-, 3-, and 5-year prognostic prediction were 0.71, 0.68 and 0.71, respectively, indicating the relatively excellent predictive efficacy of the model.

Then, the established prognostic signature was further validated in the test group, including the internal validation set and all GSE53625 and TCGA datasets. Expression level profiles for the 6 selected genes were obtained from 3 testing group samples, and the risk scores were calculated using the abovementioned method for each patient. Consistent with the above results, sorting the samples by risk score produced the heatmap shown in Figure 5A-C and Supplementary Figures 5A-C and 6A-C, which indicated that the risk score increased as the expression levels decreased, except for ADAMTS18. The Kaplan-Meier survival curves (Figure 5D) showed that this risk model could effectively distinguish the survival of the high- and low-risk groups in the validation set (*P* = 0.0014 for the internal validation set, *P* < 0.001 for all datasets, and *P* = 0.043 for the TCGA dataset). In the external validation group, the AUCs of the 1-, 3-, and 5-year overall survival predictions for the risk scores were 0.64, 0.78, and 0.79, respectively (Figure 5E). The efficiency results of the risk score classification for prognosis prediction at 1, 3, and 5 years are presented. These results were consistent with those obtained from the training dataset, demonstrating the relatively excellent predictive accuracy and stability of our risk score model.

Evaluating the independent role of the prognostic signature and building a predictive nomogram for OS prediction

To verify the use of these candidate prognostic genes as independent biomarkers, univariate and multivariate Cox regression analyses were used to evaluate whether the predictive value of the model-incorporated prognostic

signature was affected by other clinical factors (Figure 6A-B). The results indicated that age (P = 0.010, HR = 1.011), stage, location and risk score (P < 0.001, HR = 1.869) were independent prognostic factors for OS. Taking these results of the univariate and multivariate Cox analyses into consideration, the risk score integrated with age and stage was finally chosen together to construct a nomogram model, as presented in Figure 6C. The survival probability for the individuals at 1, 3, and 5 years was obtained through the function conversion relationship of the total scores. The calibration plot of the nomogram (Figure 6D) showed better consistency between the predicted OS outcomes and actual observations, indicating a good predictive performance of the hypoxia-related prognosis model.

Correlation analysis of the risk score with clinicopathological features and immune infiltration

The characteristics of the clinicopathological features in the high- and low-risk groups are shown by heat maps (Figure 7A), which were closely correlated with the clinical phenotypes. We then analyzed the relationship between the risk score and immune cell infiltration as well as the tumor microenvironment in ESCC. The results of the correlation analysis between the immune-related score and the risk score (Figure 7B, Supplementary Table 2) showed that the stromal and immune scores were comparable between the high- and low-risk groups. However, the estimated score was higher in the high-risk group than in the low-risk group, and the difference was statistically significant (P = 0.03). Differences in the infiltration of 22 immune cell subtypes in the high- and low-risk groups are shown in Figure 7C. GSEA between the high- and low-risk groups showed that the most significantly enriched pathway was the P53 signaling pathway (Figure 7D). Notably, we found consistent results for the Wnt pathway and hypoxia pathway (Figure 7E and 7F). The results reconfirmed that esophageal cancer cells in a hypoxic state could affect the tumor immune microenvironment through an underlying potential regulatory mechanism.

Validation of the expression of selected HRGs

To verify the accuracy of the abovementioned HRGs obtained from previous analyses, we further detected the protein expression levels of BGN and IL-18 according to previous publications and antibody availability. The clinical details of the 232 patients involved are presented in Table 1. The results of IHC-based staining of TMA showed that BGN was highly expressed in these ESCC samples and mainly localized to the cytoplasm of cancer cells (Figure 8A-D). IL-18 is normally expressed in cancerous tissues, with a significantly higher percentage of expression in normal tissues than in tumor tissues (Figure 8E-H). Moreover, the survival analysis suggested that patients with higher BGN expression were predicted to have poorer survival (Figure 8I). In contrast, patients with high expression of IL-18 showed better overall survival (Figure 8J). The chi–squared test indicated that high BGN expression was notably associated with tumor size, tumor invasion, and lymph node metastasis (P < 0.05 for all) of patients with ESCC (Table 1). For IL-18, no significant correlation between the IL-18 level and clinicopathological factor except lymph node metastasis (P < 0.001), was observed.

Discussion

ESCC is an aggressive malignancy with heterogeneity and genomic diversity, resulting in variable prognostic outcomes[18]. Therefore, it is necessary to identify novel and effective prognostic indicators to help oncologists assess the prognosis of patients with ESCC. Exploring mechanisms of ESCC progression can be beneficial for prognosis prediction. An abundance of molecular biomarkers have been reported[19], and genetic diagnosis and treatment have become more effective tools with clinical application. Since the predictive power of single

indicators is limited and influenced by confounders, this study used a comprehensive analysis of multiple datasets to develop a hypoxia-related gene signature and to construct a prognostic prediction model with good performance in predicting the OS of ESCC patients.

Hypoxia is one of the features of rapid tumor growth. Adaptation of tumor cells to a hypoxic environment leads to increased aggressiveness and a treatment-resistant tumor phenotype through physiological and genomic mechanisms, contributing to a poor prognosis in various cancers[5, 20]. The relationship between hypoxia and ESCC has also been demonstrated in many studies[21], which, based on high-throughput technologies and bioinformatics methodologies, have revealed some specific genomic alterations in an attempt to clarify the underlying mechanisms for the purpose of improving prognosis. Hypoxia induces upregulation of the expression of some genes, including CA9, GLUT1, VEGF, ADM and AK3[22, 23]. Therefore, we also used publicly available databases to discover more HRGs.

Thus, we overlapped the list of hypoxic signature genes downloaded from the TCGA and GSE53625 datasets, which helped to identify 30 genes most associated with hypoxia status. After analysis, five potential HRGs were identified: SLC2A1, PGM2, SULT2B1, CA9, and BGN. However, the expression of these genes needs further validation.

The gene expression signature assimilates expression information from multiple hypoxia-dependent genes, which can convey more information about hypoxia than measuring only a single genetic factor, and can function as a suitable hypoxia and prognosis biomarker[24]. A 5-gene prognostic signature based on the hypoxic pathway of cervical cancer was deemed to be an independent prognostic factor and had potential for application as a clinical biomarker[25]. Similarly, a seven-gene prognostic signature in melanoma[26] was proposed by Shou Y et al. Although the heterogeneous gene expression of 15 previously described hypoxia genes was confirmed in ESCC and gastric cancer, the results obtained from 95 tumor paraffin samples may have limitations[11]. We used high-throughput sequencing data from public databases and a combination of the above bioinformatic analysis methods to construct the first nomogram of ESCC OS that encompasses both clinical attributes and the effect of selected HRGs quantified by a risk score system.

Moreover, samples were successfully divided into two clusters to maximize the clustering of genes and samples with common features. To narrow the field and improve the predictive efficiency, the DEGs between two clusters were matched up with the screened green module genes to identify a set of genes most likely to be associated with hypoxia. Specifically, a hypoxia model with a 6-gene signature was used to predict the prognosis of ESCC patients. It was established upon a combination of PNPLA1, CARD18, IL-18, SLC37A2, ADAMTS18 and FAM83C gene mRNA expression and verified in both internal and external testing datasets. Upregulation of BGN was confirmed again. IL-18 was thought to be a protective factor with the opposite expression pattern. Survival analysis indicated that patients with high risk scores had poor prognoses. The AUCs for 1-, 3-, and 5-year prognostic prediction indicated the relatively excellent predictive efficacy of the model. The TCGA data were used as an external validation set, and the constructed risk score model was quite robust based on the results of survival analysis and the AUC. The risk score in both univariate and multivariate Cox regression models was a prognostic predictor and was identified as an independent prognostic factor. Then, a nomogram was developed and demonstrated the reliability of the predictions by the high degree of coincidence with the calibration curves. We developed the first prognostic prediction model based on the genetic risk profile of hypoxia, with excellent prognostic power and generalizability.

We investigated critical features of the tumor hypoxic environment in hopes of providing clues for clinical diagnosis and immunotherapy. Functional enrichment analysis indicated that these genes are enriched in pathways such as the P53, Wnt, and hypoxia signaling pathways. Among them, P53 is one of the major apoptotic signaling pathways that is regulated in a hypoxia-dependent or nonhypoxia-dependent manner through HIF-1, a major transcription factor activated by hypoxia[27]. These enriched KEGG pathways revealed that molecular alterations in the high-risk group were closely associated with the hypoxic microenvironment of ESCC. To investigate the tumor immunity relevance of the gene signature, the associations of the gene signature with tumor purity and the presence of infiltrating stromal/immune cells in tumor tissues were evaluated. The estimated score was significantly higher in the high-risk group than in the low-risk group. The complex interaction between tumor cells and immune cells in hypoxic environments remains to be further explored.

Combining the 5 hub genes selected earlier, a list of the 11 most relevant genes was obtained. Based on the results of previous studies and the availability of antibodies, two of them were selected for experimental validation of these bioinformatics analyses. In addition, IHC staining results showed that BGN was highly expressed in clinical samples versus adjacent normal tissues, which was consistent with previous studies. Higher BGN expression was notably associated with larger tumor size, deeper invasion, more metastatic lymph nodes and poorer prognosis. IL-18 may serve as a possible tumor protective factor and higher levels of expression was associated with less lymph node metastasis and longer survival time. In our study, taken together, this model can be used to assess the prognostic risk of ESCC patients. In general, the gene signature performed well at predicting the OS of ESCC.

BGN, a member of the family of leucine-rich small proteoglycans (SLRPs), has been detected tumor-associated upregulation in esophageal[28, 29], pancreatic[30], gastric[31], colorectal[32], prostate[33] and ovarian cancer[34] and was correlated with tumor progression and poor prognosis. This is consistent with the findings of the present study obtained by analyzing ESCC samples from public databases. IL-18, the protein encoded by this gene, is a proinflammatory and immunomodulatory cytokine and a member of the IL-1 family. IL-18 shows antitumor activity in different experimental tumor models through activation of NK and/or T cell responses and has a protective effect against rectal cancer development. According to previous studies, IL-18 activation may be a double-edged sword that promotes tumor development and progression, enhances antitumor immunity and limits tumor growth[35]. In our study, it is suggested that this gene has a tumor suppressive effect. These findings are in line with data from studies on various other cancer types, especially ESCC[36].

The following are limitations of this study: First, from the perspective of data sources, potential selection bias could not be ruled out due to transcriptomic and corresponding clinical data obtained from the TCGA and GEO databases. The majority of these patients were white, African American, or Latino, which may result in inconsistent RNA-seq results and poor reproducibility of our findings in patients of other races. Second, the datasets included in this study were all from public databases, and certain limitations and preferences could not be excluded as these were retrospective analyses. Third, our prediction score was developed based on microarray data from the GEO database, and it is necessary to use external datasets with complete clinical information and gene expression information for validation. Further in vitro and in vivo studies are needed to investigate the mechanisms of the key genes involved in prognosis prediction.

Conclusions

In conclusion, our study identified a 6-gene signature and a prognostic nomogram incorporating the gene signature and clinical prognostic factors associated with tumor prognosis. The constituents of the gene signature may serve as potential prognostic predictors and provide therapeutic targets with future clinical applications for ESCC. The prognostic nomogram may reliably facilitate individualized treatment and medical decision making.

Abbreviations

ESCC: Esophageal squamous cell carcinoma; GEO: Gene Expression Omnibus; TCGA: The Cancer Genome Atlas; HRGs: hypoxia-related genes; TME: tumor microenvironment; TMAs: Tissue microarrays; IHC: immunohistochemical staining; CancerSEA: Cancer Single-Cell State Atlas; MSigDB: Molecular Signatures Database; FDR: false discovery rate; DEGs: differentially expressed genes; CDF: cumulative distribution function; WGCNA: Weighted gene coexpression network analysis; TOM: topological overlap matrix; AUC: Area under the curve; ROC: Receiver operating characteristic; PPI: protein-protein interaction; KEGG: Kyoto Encyclopedia of Genes and Genomes; GO: Gene Ontology; CC: : cellular component; BF: biological process; MF: molecular function; GSEA: gene set enrichment analysis; OS: overall survival; Tregs: regulatory T cells.

Declarations

Author contribution

(I) Conception and design: Xiao WY, Tang P, Yu ZT and Zhang HD; (II) Administrative support: Gong L, Yu ZT and Zhang HD; (III) Provision of study materials or patients: Xiao WY, Sui ZL, Wu XX; (IV) Collection and assembly of data: Yang YY, Zhu Ningning, Han YM; (V) Data analysis and interpretation: Xiao WY, Tang P, Zhang HD, Gong L; (VI) Manuscript writing: Xiao WY; (VII) Final approval of manuscript: All authors.

Funding

National Natural Science Foundation of China (Grant/Award Number: 82002551 and 81772619), Bethune Charity Foundation (Grant Number: HZB-20190528-11 and HZB-20190528-18), Science and Technology Project of Tianjin Municipal Health Commission (KJ20119), Clinical Trial Project of Tianjin Medical University (2017kylc006).

Availability of data and materials

Deidentified data and related documents will be made available upon request.

Ethics approval and consent to participate

The project has been approved by the ethics committee of Tianjin Medical University Cancer Institute and Hospital and each participant has signed written informed consent.

Consent for publication

Not applicable

Competing interests

The authors declare that they have no confict of interest.

References

1. Zeng H, Zheng R, Guo Y, Zhang S, Zou X, Wang N, Zhang L, Tang J, Chen J, Wei K *et al*: **Cancer survival in China, 2003-2005: a population-based study**. *Int J Cancer* 2015, **136**(8):1921-1930.
2. He Z, Ke Y: **Precision screening for esophageal squamous cell carcinoma in China**. *Chin J Cancer Res* 2020, **32**(6):673-682.
3. Zhang J, Wang JL, Zhang CY, Ma YF, Zhao R, Wang YY: **The prognostic role of FZD6 in esophageal squamous cell carcinoma patients**. *Clin Transl Oncol* 2020, **22**(7):1172-1179.
4. Peerlings J, Van De Voorde L, Mitea C, Larue R, Yaromina A, Sandeleanu S, Spiegelberg L, Dubois L, Lambin P, Mottaghy FM: **Hypoxia and hypoxia response-associated molecular markers in esophageal cancer: A systematic review**. *Methods* 2017, **130**:51-62.
5. Jing X, Yang F, Shao C, Wei K, Xie M, Shen H, Shu Y: **Role of hypoxia in cancer therapy by regulating the tumor microenvironment**. *Mol Cancer* 2019, **18**(1):157.
6. Chen F, Chu L, Li J, Shi Y, Xu B, Gu J, Yao X, Tian M, Yang X, Sun X: **Hypoxia induced changes in miRNAs and their target mRNAs in extracellular vesicles of esophageal squamous cancer cells**. *Thorac Cancer* 2020, **11**(3):570-580.
7. Wang Y, Xie Z, Lu H: **Significance of halofuginone in esophageal squamous carcinoma cell apoptosis through HIF-1alpha-FOXO3a pathway**. *Life Sci* 2020, **257**:118104.
8. Macedo-Silva C, Miranda-Goncalves V, Lameirinhas A, Lencart J, Pereira A, Lobo J, Guimaraes R, Martins AT, Henrique R, Bravo I *et al*: **JmjC-KDMs KDM3A and KDM6B modulate radioresistance under hypoxic conditions in esophageal squamous cell carcinoma**. *Cell Death Dis* 2020, **11**(12):1068.
9. Vaupel P, Mayer A: **Hypoxia in cancer: significance and impact on clinical outcome**. *Cancer Metastasis Rev* 2007, **26**(2):225-239.
10. Muz B, de la Puente P, Azab F, Azab AK: **The role of hypoxia in cancer progression, angiogenesis, metastasis, and resistance to therapy**. *Hypoxia (Auckl)* 2015, **3**:83-92.
11. Winther M, Alsner J, Tramm T, Nordsmark M: **Hypoxia-regulated gene expression and prognosis in loco-regional gastroesophageal cancer**. *Acta Oncol* 2013, **52**(7):1327-1335.
12. Mao Y, Wang Y, Dong L, Zhang Y, Zhang Y, Wang C, Zhang Q, Yang S, Cao L, Zhang X *et al*: **Hypoxic exosomes facilitate angiogenesis and metastasis in esophageal squamous cell carcinoma through altering the phenotype and transcriptome of endothelial cells**. *J Exp Clin Cancer Res* 2019, **38**(1):389.
13. Liu M, An H, Zhang Y, Sun W, Cheng S, Wang R, Wang X, Feng L: **Molecular analysis of Chinese oesophageal squamous cell carcinoma identifies novel subtypes associated with distinct clinical outcomes**. *EBioMedicine* 2020, **57**:102831.
14. van Uden P, Kenneth NS, Rocha S: **Regulation of hypoxia-inducible factor-1alpha by NF-kappaB**. *Biochem J* 2008, **412**(3):477-484.
15. Yuan H, Yan M, Zhang G, Liu W, Deng C, Liao G, Xu L, Luo T, Yan H, Long Z *et al*: **CancerSEA: a cancer single-cell state atlas**. *Nucleic Acids Res* 2019, **47**(D1):D900-D908.
16. Liberzon A, Birger C, Thorvaldsdottir H, Ghandi M, Mesirov JP, Tamayo P: **The Molecular Signatures Database (MSigDB) hallmark gene set collection**. *Cell Syst* 2015, **1**(6):417-425.

17. Langfelder P, Horvath S: **WGCNA: an R package for weighted correlation network analysis**. *BMC Bioinformatics* 2008, **9**:559.
18. Li R, Li P, Xing W, Qiu H: **Heterogeneous genomic aberrations in esophageal squamous cell carcinoma: a review**. *American journal of translational research* 2020, **12**(5):1553-1568.
19. Shao N, Han Y, Song L, Song W: **Clinical significance of hypoxia-inducible factor 1alpha, and its correlation with p53 and vascular endothelial growth factor expression in resectable esophageal squamous cell carcinoma**. *J Cancer Res Ther* 2020, **16**(2):269-275.
20. Akanji MA, Rotimi D, Adeyemi OS: **Hypoxia-Inducible Factors as an Alternative Source of Treatment Strategy for Cancer**. *Oxid Med Cell Longev* 2019, **2019**:8547846.
21. Hu X, Lin J, Jiang M, He X, Wang K, Wang W, Hu C, Shen Z, He Z, Lin H *et al*: **HIF-1alpha Promotes the Metastasis of Esophageal Squamous Cell Carcinoma by Targeting SP1**. *J Cancer* 2020, **11**(1):229-240.
22. Harris BH, Barberis A, West CM, Buffa FM: **Gene Expression Signatures as Biomarkers of Tumour Hypoxia**. *Clin Oncol (R Coll Radiol)* 2015, **27**(10):547-560.
23. Yang L, West CM: **Hypoxia gene expression signatures as predictive biomarkers for personalising radiotherapy**. *The British journal of radiology* 2019, **92**(1093):20180036.
24. Moon EJ, Brizel DM, Chi JT, Dewhirst MW: **The potential role of intrinsic hypoxia markers as prognostic variables in cancer**. *Antioxid Redox Signal* 2007, **9**(8):1237-1294.
25. Yang Y, Li Y, Qi R, Zhang L: **Constructe a novel 5 hypoxia genes signature for cervical cancer**. *Cancer Cell Int* 2021, **21**(1):345.
26. Shou Y, Yang L, Yang Y, Zhu X, Li F, Xu J: **Identification of Signatures of Prognosis Prediction for Melanoma Using a Hypoxia Score**. *Front Genet* 2020, **11**:570530.
27. Sermeus A, Michiels C: **Reciprocal influence of the p53 and the hypoxic pathways**. *Cell Death Dis* 2011, **2**:e164.
28. Zhu YH, Yang F, Zhang SS, Zeng TT, Xie X, Guan XY: **High expression of biglycan is associated with poor prognosis in patients with esophageal squamous cell carcinoma**. *International journal of clinical and experimental pathology* 2013, **6**(11):2497-2505.
29. Tang L, Chen Y, Peng X, Zhou Y, Jiang H, Wang G, Zhuang W: **Identification and Validation of Potential Pathogenic Genes and Prognostic Markers in ESCC by Integrated Bioinformatics Analysis**. *Front Genet* 2020, **11**:521004.
30. Yang L, Cui R, Li Y, Liang K, Ni M, Gu Y: **Hypoxia-Induced TGFBI as a Serum Biomarker for Laboratory Diagnosis and Prognosis in Patients with Pancreatic Ductal Adenocarcinoma**. *Lab Med* 2020, **51**(4):352-361.
31. Liu Z, Liu S, Guo J, Sun L, Wang S, Wang Y, Qiu W, Lv J: **Identification and Analysis of Key Genes Driving Gastric Cancer Through Bioinformatics**. *Genet Test Mol Biomarkers* 2021, **25**(1):1-11.
32. Zhou XG, Huang XL, Liang SY, Tang SM, Wu SK, Huang TT, Mo ZN, Wang QY: **Identifying miRNA and gene modules of colon cancer associated with pathological stage by weighted gene co-expression network analysis**. *Onco Targets Ther* 2018, **11**:2815-2830.
33. Jacobsen F, Kraft J, Schroeder C, Hube-Magg C, Kluth M, Lang DS, Simon R, Sauter G, Izbicki JR, Clauditz TS *et al*: **Up-regulation of Biglycan is Associated with Poor Prognosis and PTEN Deletion in Patients with Prostate Cancer**. *Neoplasia* 2017, **19**(9):707-715.

34. Pan X, Ma X: **A Novel Six-Gene Signature for Prognosis Prediction in Ovarian Cancer**. *Front Genet* 2020, **11**:1006.
35. Saetang J, Chonpathompikunlert P, Sretrirutchai S, Roongsawang N, Kayasut K, Voravuthikunchai SP, Sukketsiri W, Tipmanee V, Sangkhathat S: **Anti-cancer effect of engineered recombinant interleukin 18**. *Adv Clin Exp Med* 2020, **29**(10):1135-1143.
36. Tsuboi K, Miyazaki T, Nakajima M, Fukai Y, Masuda N, Manda R, Fukuchi M, Kato H, Kuwano H: **Serum interleukin-12 and interleukin-18 levels as a tumor marker in patients with esophageal carcinoma**. *Cancer Lett* 2004, **205**(2):207-214.

Tables

Table 1 Associations of BGN and IL-18 expression with clinicopathological variables of ESCC patients from the TMA dataset.

Clinicopathological variables	n	BGN expression		x^2	*P* value	IL18 expression		x^2	*P* value
		Low level	High level			Low level	High level		
Gender				1.924	0.165			3.513	0.061
Male	192	70 (36.5%)	122 (63.5%)			108 (56.3%)	84 (43.8%)		
Female	40	10 (25%)	30 (75%)			16 (40.0%)	24 (60.0%)		
Age□years□				0.837	0.360			0.595	0.440
≤66	114	36 (31.6%)	78 (68.4%)			58 (50.9%)	56 (49.1%)		
□66	118	44 (37.3%)	74 (62.7%)			66 (55.9%)	52 (44.1%)		
Tumor size				4.210	**0.040**			0.038	0.845
≤3.5	109	45 (41.3%)	64 (58.7%)			59 (54.1%)	50 (45.9%)		
□3.5	123	35 (28.5%)	88 (71.5%)			65 (52.8%)	58 (47.2%)		
Tumor Location				1.088	0.581			1.710	0.425
Upper	11	5 (45.5%)	6 (54.5%)			4 (36.4%)	7 (63.6%)		
Middle	171	56 (32.7%)	115 (67.3%)			91 (53.2%)	80 (46.8%)		
Lower	50	19 (38.0%)	31 (62.0%)			29 (58.0%)	21 (42.0%)		
Histological grade				3.312	0.191			13.439	0.179
I	6	4 (66.7%)	2 (33.3%)			1 (16.7%)	5 (83.3%)		
II	169	59 (34.9%)	110 (65.1%)			91 (53.8%)	78 (46.2%)		
III	57	17 (29.8%)	40 (70.2%)			32 (56.1%)	25 (43.9%)		
Tumor invasion depth				22.483	**0.000**			0.353	0.553
T1 - T2	56	34 (60.7%)	22 (39.3%)			28 (50.0%)	28 (50.0%)		
T3 - T4	176	46 (26.1%)	130 (73.9%)			96 (54.5%)	80 (45.5%)		
Lymph node metastasis				4.768	**0.029**			14.552	**0.000**

N0	128	52 (40.6%)	76 (59.4%)	54 (42.2%)	74 (57.8%)
N1 - N3	104	28 (26.9%)	76 (73.1%)	70 (67.3%)	34 (32.7%)

Abbreviations: BGN, biglycan; IL-18, interleukin 18; ESCC, esophageal squamous cell carcinoma; TMA, TMA tissue microarray.

Figures

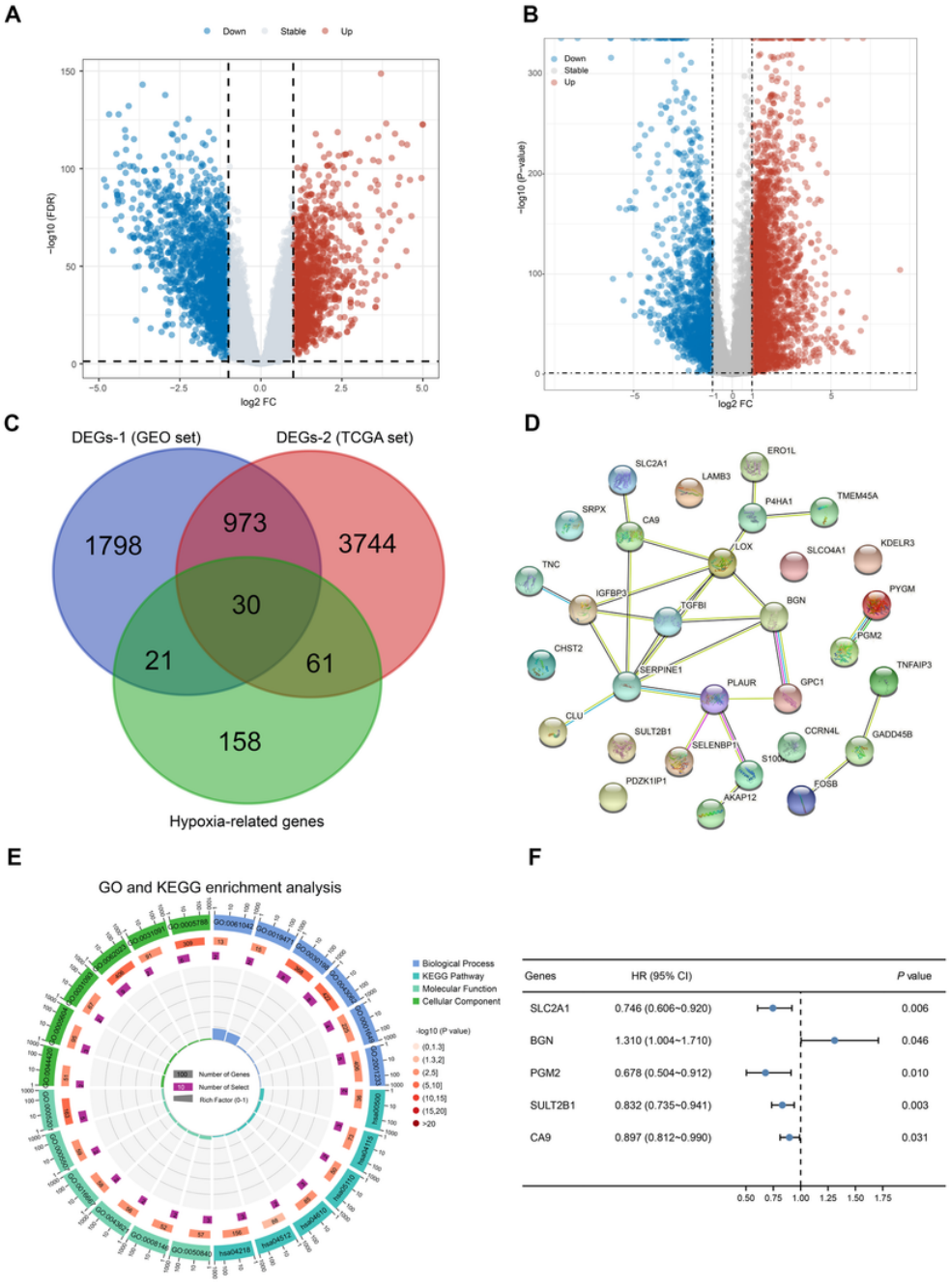

Figure 1

Identification of potential HRGs and related functional analysis. Differential expression analysis in ESCC (A-B). Volcano plots showing DEGs in (A) GES53625 and (B) TCGA-ESCC. (C) Venn diagrams of the 30 overlapping differentially expressed HRGs from three datasets. (D) PPI network of the differentially expressed HRGs constructed in STRING and visualized by Cytoscape. (E) The top 10 significant terms from the GO and KEGG analyses of differentially expressed HRGs. Abbreviations: HRGs, hypoxia related genes; ESCC, esophageal squamous cell carcinoma; DEGs, differentially expressed genes; PPI, protein–protein interaction; GO, Gene ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes.

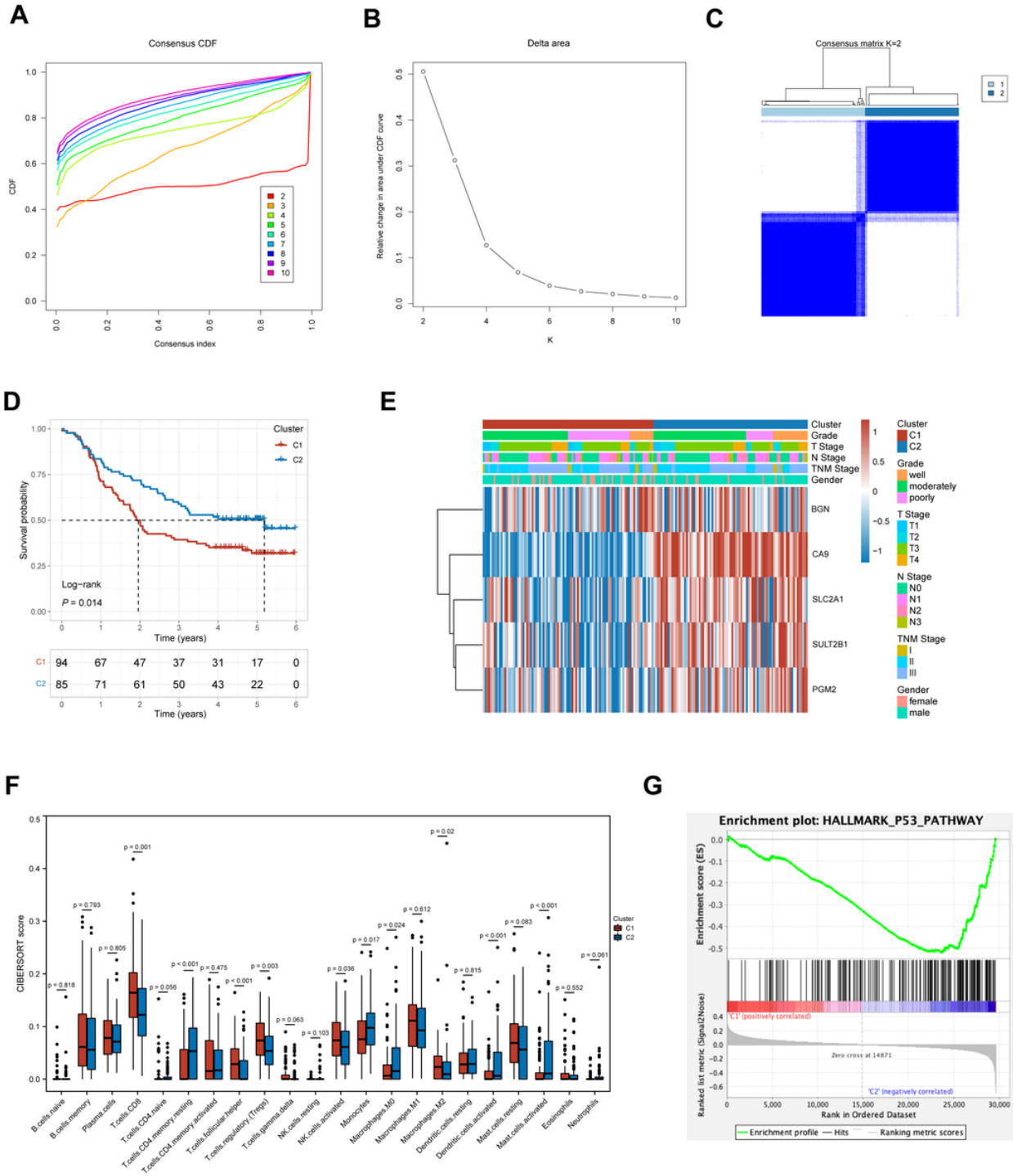

Figure 2

Analysis of subgroups with clinicopathological characteristics and immune infiltration. (A) Cumulative distribution map of clustering consistency. (B) Clustering Delta area map. (C) Display of the clustering results corresponding to K = 2. (D) The results of Kaplan-Meier survival analysis of different clusters. (E) A heatmap corresponding to the dendrogram annotated by grade, T stage, N stage, stage, sex, and 5 real hub genes. (F) The infiltration of 22 immune cell subtypes in different clusters. (G) GSEA was used to demonstrate the correlation between genes from different clusters and the KEGG enriched pathways. Abbreviations: GSEA, gene set enrichment analysis; KEGG, Kyoto Encyclopedia of Genes and Genomes.

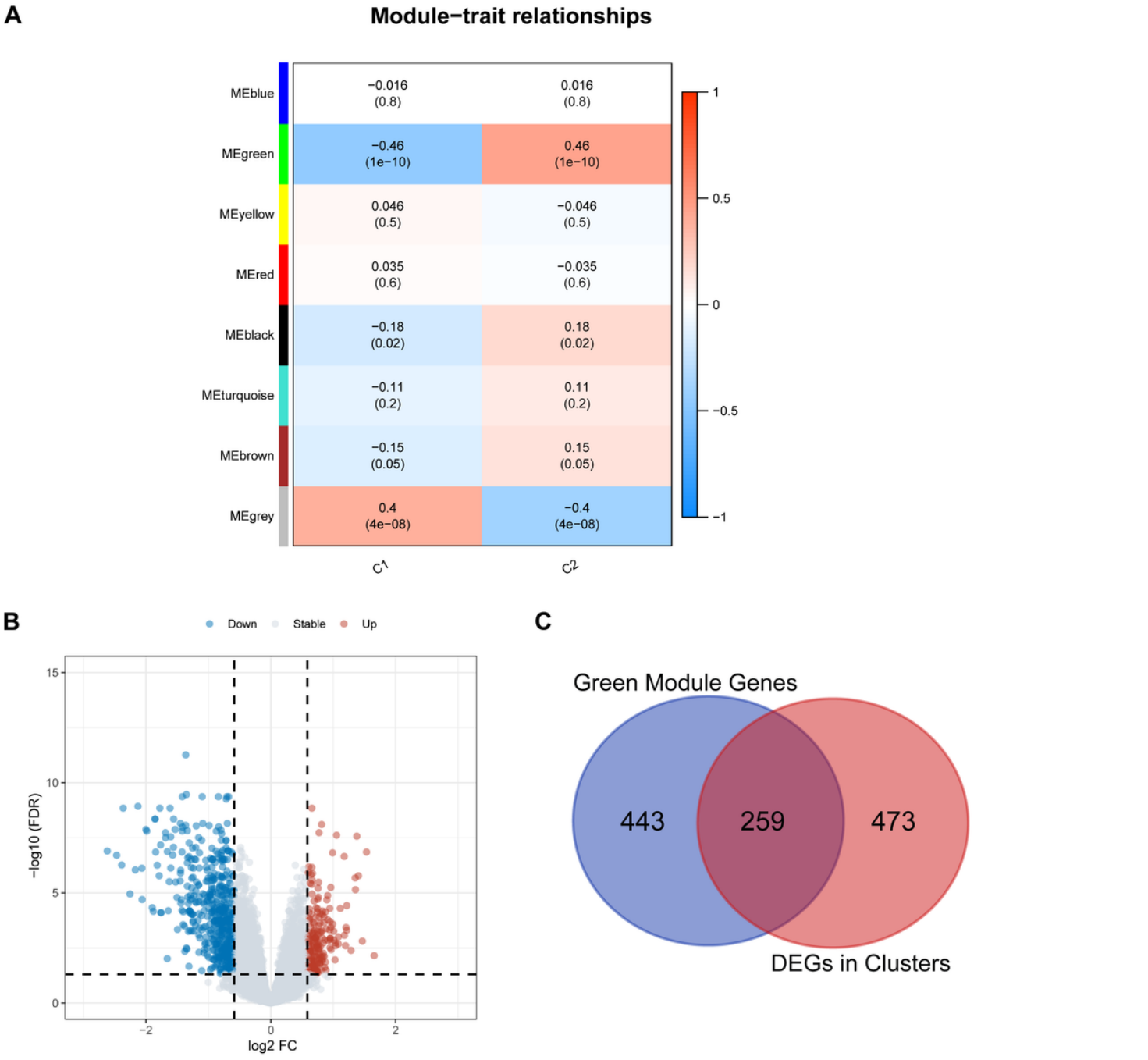

Figure 3

Screening of key HRGs. (A) The relationship of module features with consensus subgroups was assessed by eight gene modules obtained from WGCNA. (B) Volcano plots showing the DEGs in two clusters. (C) Venn diagrams of the 259 differentially expressed HRGs from intersection of green module genes and consensus clustering DEGs. Abbreviations: HRGs, hypoxia related genes; DEGs, differentially expressed genes; WGCNA, the weighted gene co-expression network analysis.

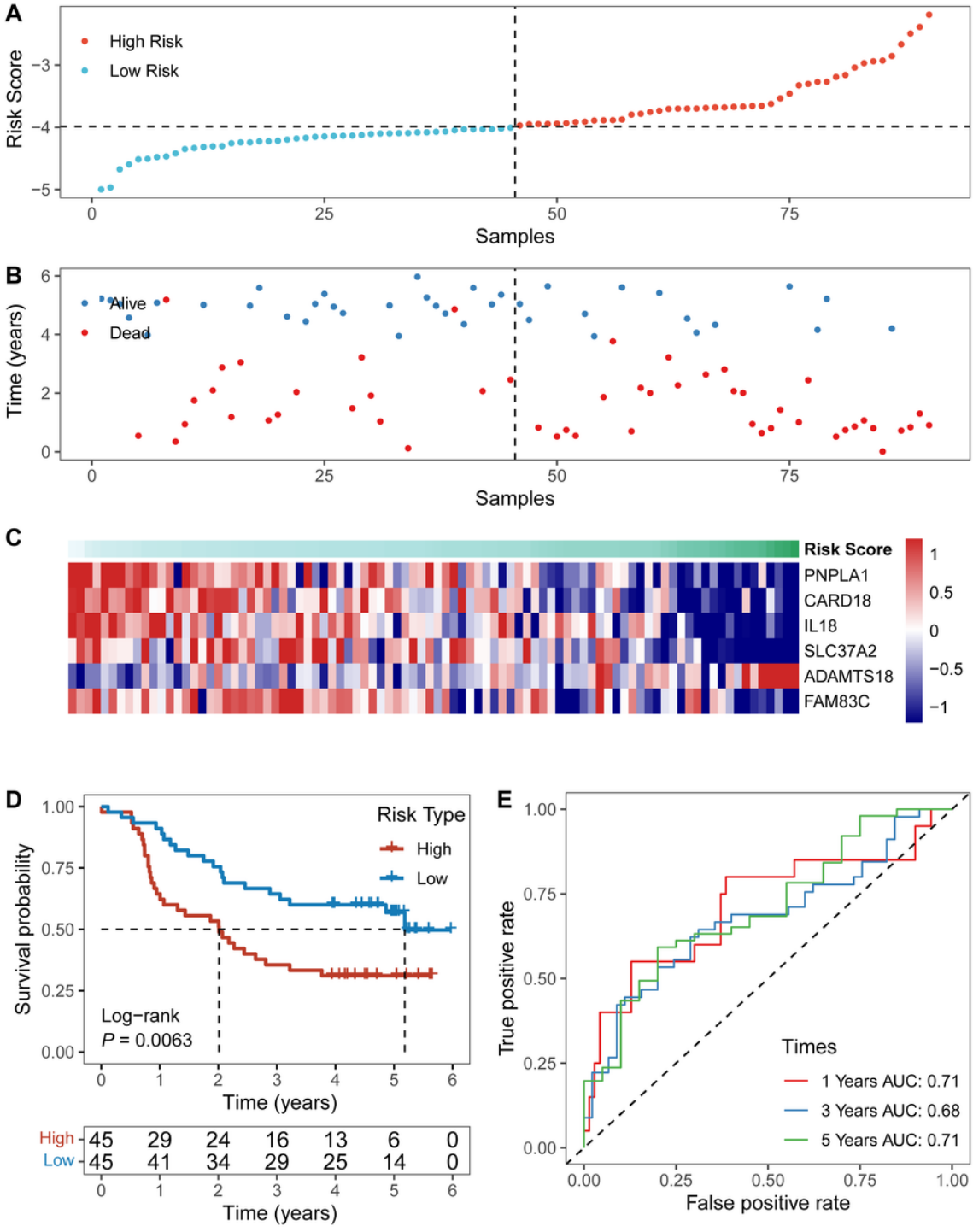

Figure 4

Constructing the prognostic hypoxia gene features in the training set. The samples were assigned a risk score and ordered to determine whether the expression level (A) and survival time (B) varied systematically with the risk score. (C) Expression levels of the 6 HRGs based on the risk score. (D) Survival curve distribution of the risk score.

(E) ROC curves and AUCs of risk score classifications. Abbreviations: HRGs, hypoxia related genes; ROC, receiver operating characteristic; AUC, area under the curve.

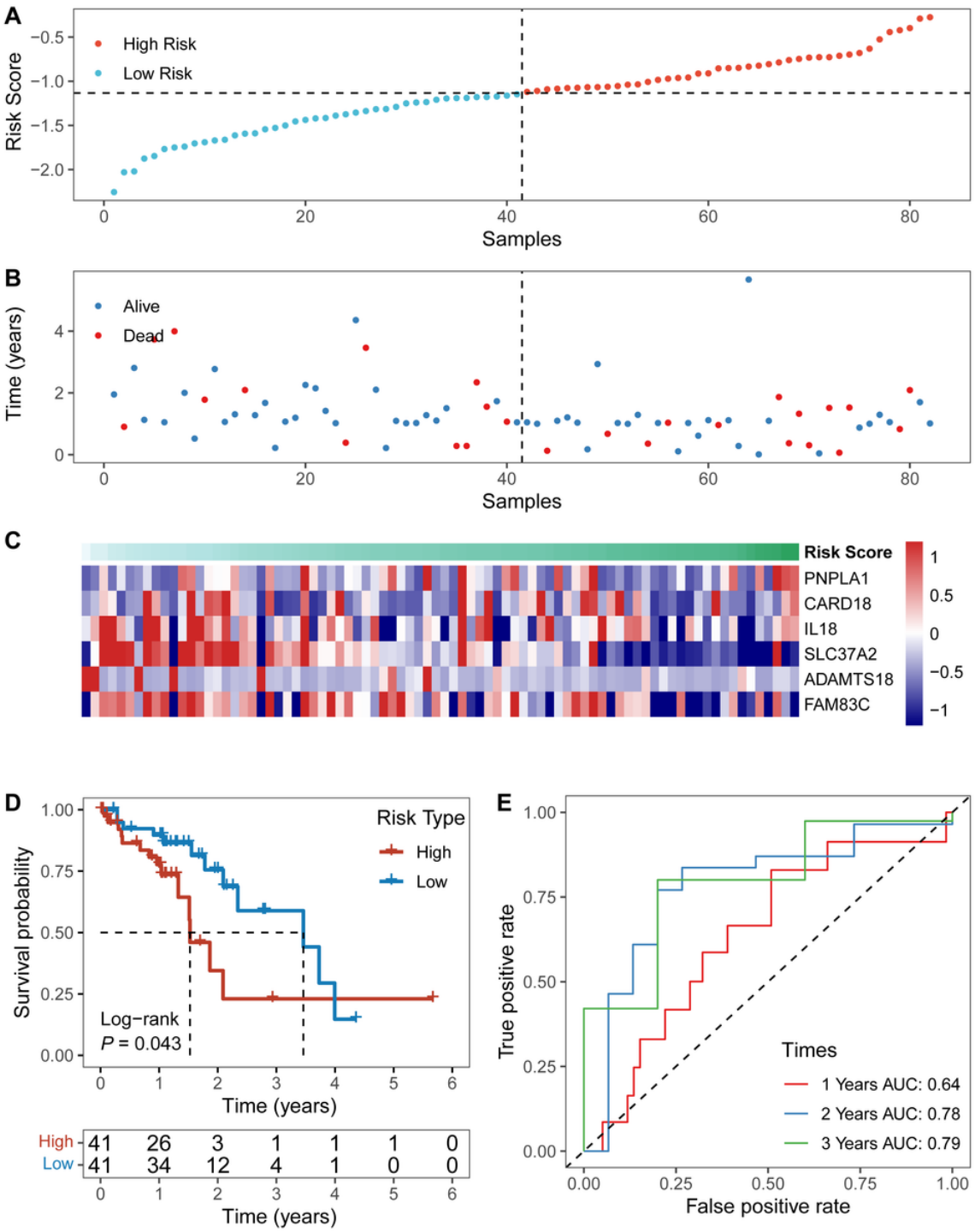

Figure 5

Validating the prognostic hypoxia gene features in the external TCGA dataset. The samples were assigned a risk score and ordered to determine whether the expression level (A) and survival time (B) varied systematically with the risk score. (C) Expression levels of the 6 HRGs based on risk scores. (D) Survival curve distribution of the risk score. (E) ROC curves and AUCs of risk score classifications. Abbreviations: HRGs, hypoxia related genes.

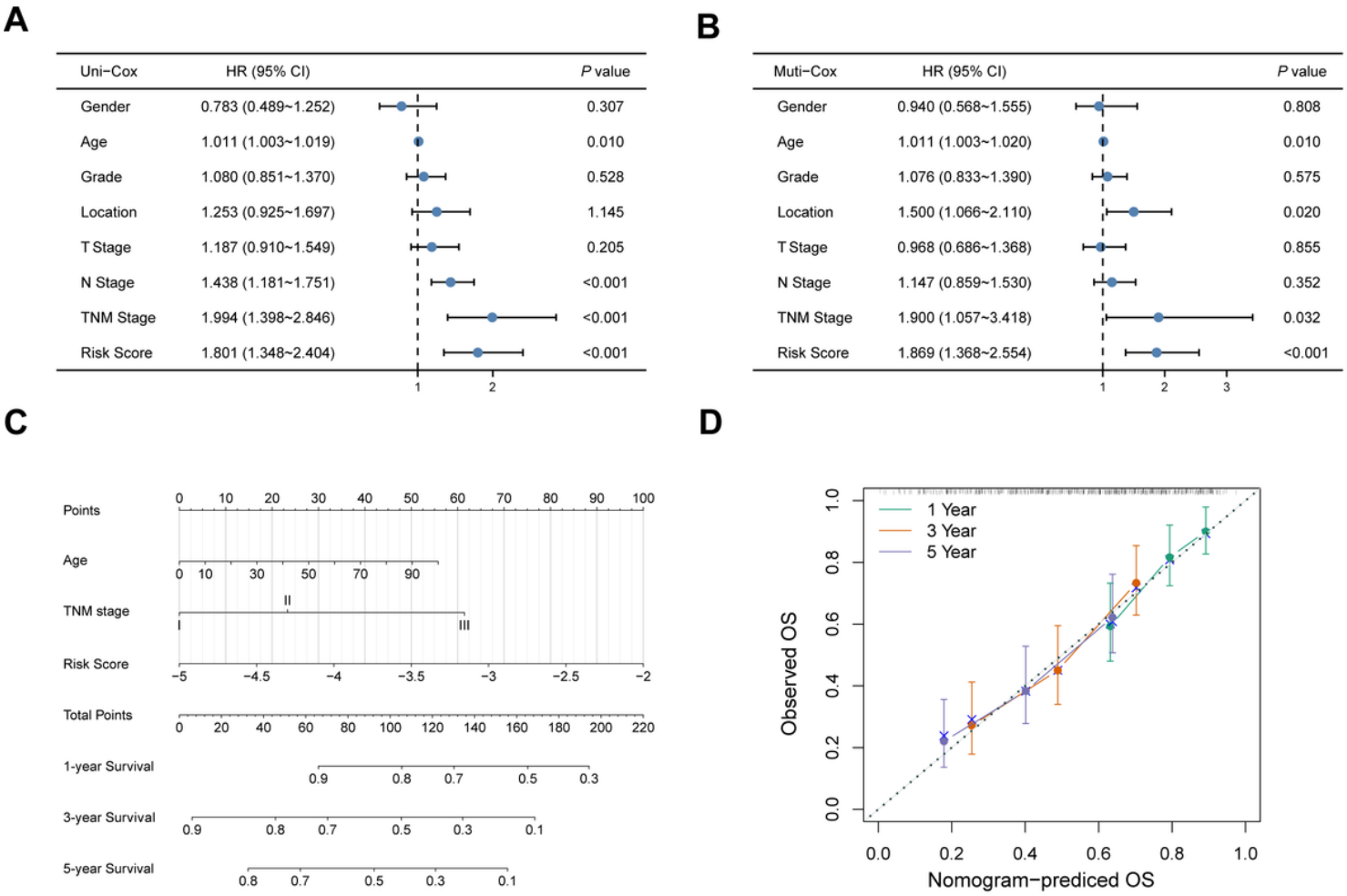

Figure 6

Evaluating the independent role of the prognostic signature and building a predictive nomogram. Univariate (A) and multivariate (B) Cox regression analyses were used to evaluate the predictive value of the model-incorporated prognostic signature. (C) Integrating the risk score with age and stage to construct a nomogram model. (D) The calibration curve of the nomogram.

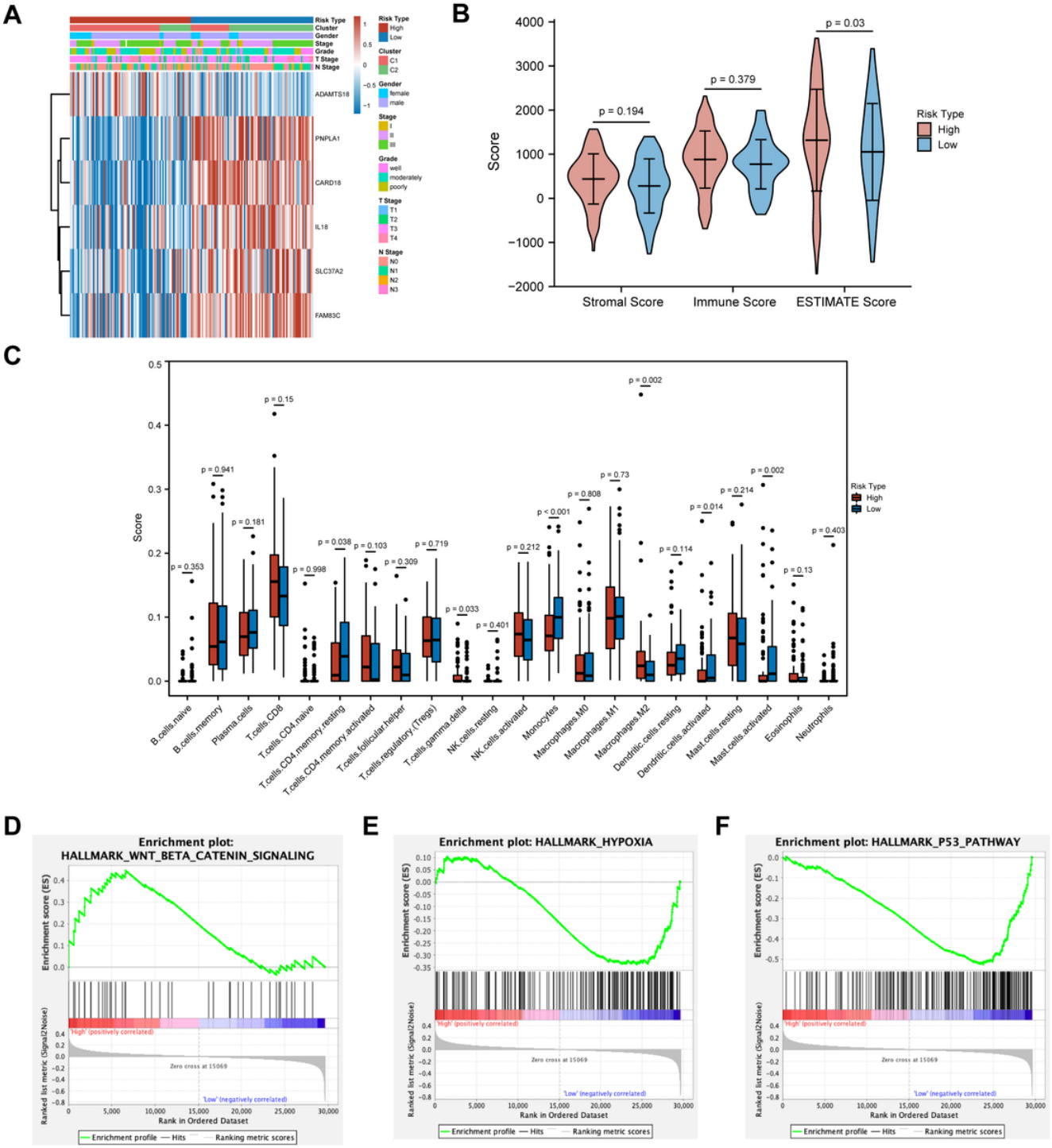

Figure 7

Analysis of the correlation between the risk score and immune infiltration. (A) Expression and clinical features of the 5 prognostic genes in the high- and low-risk groups of the training set. (B) The results of the correlation analysis between the immune-related score and the risk score. (C) The infiltration of 22 immune cell subtypes in the high- and low-risk groups. (D-F) GSEA was used to demonstrate the correlation between HRG expression and the KEGG enriched pathways. Abbreviations: GSEA, gene set enrichment analysis; KEGG, Kyoto Encyclopedia of Genes and Genomes.

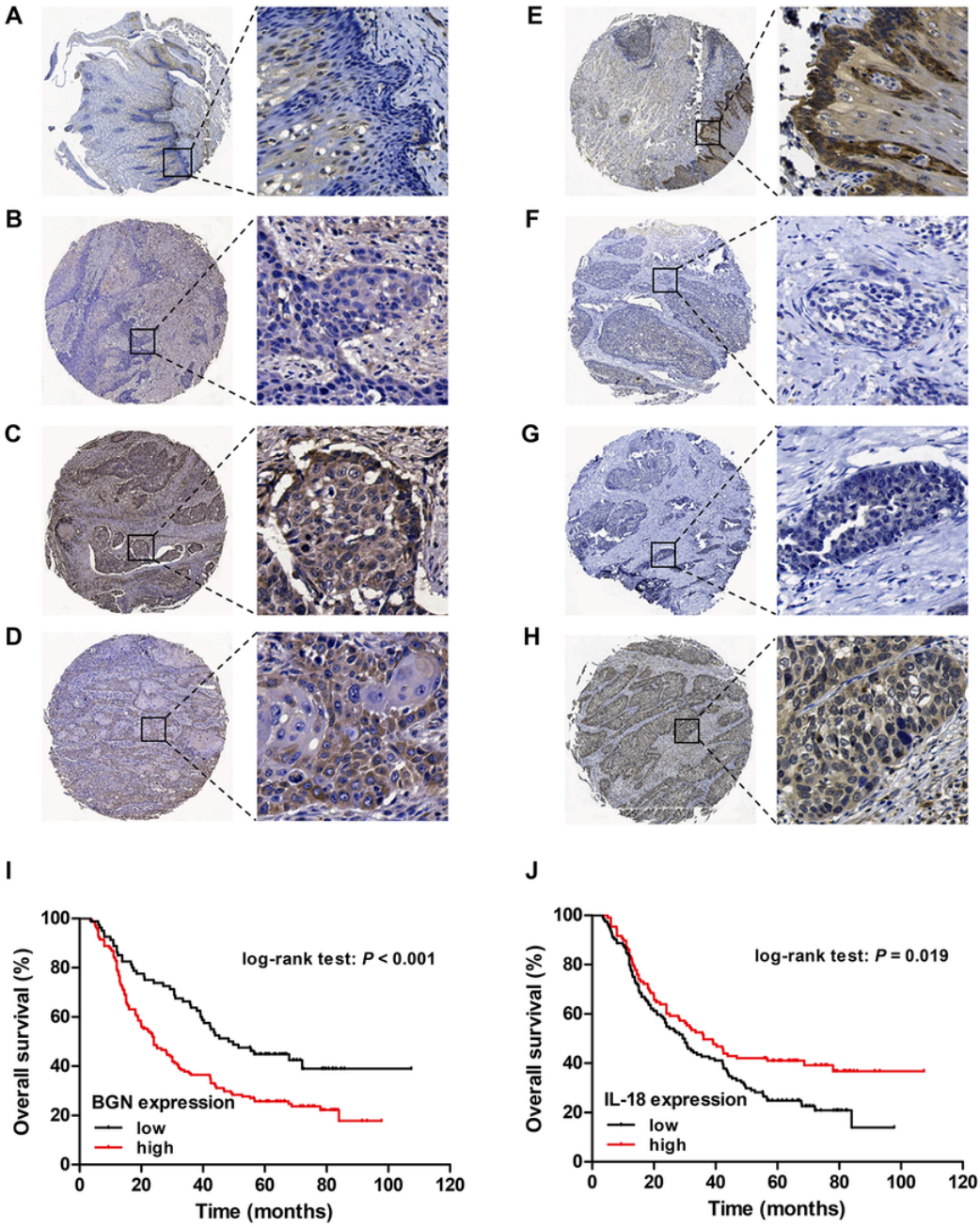

Figure 8

The expression of HRGs protein in ESCC and normal tissues was detected in TMA. Representative image of IHC staining for low expression of BGN in adjacent non-tumor tissue (A) and low expression (B), high expression case 1 (C), and high expression case 2 (D) of BGN in human primary tumor tissue, stained 5x, inset 40x magnification. Representative image of IHC staining for high expression of IL-18 in adjacent non-tumor tissue (E) and low expression case 1 (F), low expression case 2 (G) and high expression (H) of IL-18 in human primary tumor tissue, stained 5x, inset 40x magnification; (I) Kaplan-Meier curve of BGN expression in the TMA dataset. (J) Kaplan-Meier curve of IL-18 expression in the TMA dataset. Abbreviations: HRGs, hypoxia related genes; ESCC, esophageal squamous cell carcinoma; TMA, tissue microarray; BGN, biglycan; IL-18, interleukin 18.

Supplementary Files

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

- S.Fig.1.tif
- S.Fig.2.tif
- S.Fig.3.tif
- S.Fig.4.tif
- S.Fig.5.tif
- S.Fig.6.tif
- S.Table.1.pdf
- S.Table.2.pdf