



Identification of the Prognostic Value of Tumor Microenvironment-Related Genes in Esophageal Squamous Cell Carcinoma

OPEN ACCESS

Donglei Zhang^{1†}, Changlin Qian^{2†}, Huabing Wei^{1*} and Xiaozhe Qian^{1*}

Edited by:

Sergio Akira Uyemura,
University of São Paulo, Brazil

Reviewed by:

Shinjini Singh,
University of Texas MD Anderson
Cancer Center, United States
Ramya Sivakumar,
University of Washington,
United States

*Correspondence:

Xiaozhe Qian
qianxiaozhe@renji.com
Huabing Wei
doctorwei@163.com

[†]These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Molecular Diagnostics and
Therapeutics,
a section of the journal
Frontiers in Molecular Biosciences

Received: 01 September 2020

Accepted: 24 November 2020

Published: 14 December 2020

Citation:

Zhang D, Qian C, Wei H and Qian X
(2020) Identification of the Prognostic
Value of Tumor
Microenvironment-Related Genes in
Esophageal Squamous Cell
Carcinoma.
Front. Mol. Biosci. 7:599475.
doi: 10.3389/fmolb.2020.599475

¹ Department of Thoracic Surgery, Renji Hospital, School of Medicine, Shanghai Jiaotong University, Shanghai, China,
² Department of General Surgery, Renji Hospital, School of Medicine, Shanghai Jiaotong University, Shanghai, China

Background: Esophageal squamous cell carcinoma (ESCC) is the most prevalent histological type of esophageal cancer, but there is a lack of definite prognostic markers for this cancer.

Methods: We used the ESTIMATE algorithm to access the tumor microenvironment (TME) of ESCC cases deposited in the TCGA database, and identified TME-related prognostic genes using Cox regression analysis. A least absolute shrinkage and selector operation or LASSO algorithm was used to identify key prognostic genes. Risk scores were calculated, and a clinical predictive model was constructed to evaluate the prognostic value of TME-related genes.

Results: We found that high immune and stromal scores were significantly associated with poor overall survival ($p < 0.05$). We identified a total of 1,151 TME-related differently expression genes, among which 67 were prognosis-related genes. Through the LASSO method, 13 key prognostic genes were selected, namely, *ADAMTS16*, *LOC51089*, *CH25H*, *CORO2B*, *DLGAP1*, *GYS2*, *HAL*, *MXRA8*, *NPTX1*, *OTX1*, *RET*, *SLC24A2*, and *SPI1*, and a 13-gene risk score was constructed. A higher score was indicative of a poorer prognosis than a lower risk score (hazard ratio = 8.21, 95% confidence interval: 2.56–26.31; $P < 0.001$). The risk score was significantly correlated with immune/stromal scores and various types of infiltrating immune cells, including CD8 cells, regulatory T cells, and resting macrophages.

Conclusion: We characterized the tumor microenvironment in ESCC, and identified the key prognosis genes. The risk score based on the expression profiles of these genes is proposed as an indicator of TME status and is instrumental in predicting patient prognosis.

Keywords: esophageal squamous cell carcinoma, tumor microenvironment, prognosis, overall survival, TCGA

INTRODUCTION

Esophageal cancer is the seventh most common cause of cancer-related deaths worldwide (Fitzmaurice et al., 2018). Esophageal squamous cell carcinoma (ESCC) accounts for ~90% of all esophageal cancers, which are always in the advanced stage at the time of their first diagnosis (Rustgi and El-Serag, 2014). 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 prognosis of patients with ESCC remains unfavorable (Cunningham et al., 2006; Baba et al., 2018). The pathological subtypes of this cancer present limitations in the prediction of prognosis, as patients with similar clinical and pathological types could have totally different final outcomes (Smyth et al., 2017). Thus, it is crucial to identify prognostic biomarkers for ESCC patients and develop more effective therapies.

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 (Hegde and Chen, 2020). In addition to the genetic heterogeneity of tumor cells, the heterogeneity of immune and stromal cells also contribute to the complexity of TME; this affects the time and intensity of the anti-tumor response and becomes a major obstacle in the treatment of tumors. Even though much attention has been devoted to the role of TME in the development of cancers and their clinical outcomes (Galon et al., 2012; Pitt et al., 2016; Senbabaoglu et al., 2016), little is known about the relationship between TME and the prognosis of ESCC patients, or the role of TME-related genes in ESCC.

In the present study, we gathered information about the clinical features and RNA sequencing data of 95 ESCC tumor samples from the TCGA database and evaluated their TME profiles. We then identified the correlation between the TME profile and patient prognosis. The relevant mechanism was explored with gene expression profiling, and a TME-related gene signature model was established for predicting the prognosis of ESCC.

MATERIALS AND METHODS

Data Collection From the TCGA Database

Gene expression and clinical data of 185 esophageal cancer patients were downloaded from the TCGA database (<https://portal.gdc.cancer.gov/>). Only patients with a histological diagnosis of ESCC and who had not undergone neoadjuvant chemotherapy were included. Based on these criteria, 95 ESCC cases with gene expression data were included in our analysis. The clinical information of the patients is shown in **Table 1**.

Abbreviations: ESCC, esophageal squamous cell carcinoma; TME, tumor microenvironment; LASSO, least absolute shrinkage and selector operation; HRs, hazard ratios; DEG, differently expression gene; CD8, cluster of differentiation 8; M0, resting macrophages.

TABLE 1 | Clinical features of patients with esophageal squamous cell carcinoma in TCGA.

Clinical features	Count (%) (n = 95)
Status	
Alive	63 (66.3)
Dead	32 (33.7)
Gender	
Female	14 (14.7)
Male	81 (85.3)
Age	
≤50	21 (22.1)
>50	74 (77.9)
Grade	
G1	16 (16.8)
G2	48 (50.5)
G3	21 (22.1)
Missing	10 (10.5)
Stage	
Stage 1	7 (7.4)
Stage 2–3	82 (86.3)
Stage 4	4 (4.2)
Missing	2 (2.1)
Stage T	
T1	8 (8.4)
T2	31 (32.6)
T3	50 (52.6)
T4	4 (4.2)
Missing	2 (2.1)
Stage N	
N0	54 (56.8)
N1	29 (30.5)
N2	6 (6.3)
N3	3 (3.2)
NX	1 (1.1)
Missing	2 (2.1)
Stage M	
M0	83 (87.4)
M1	4 (4.3)
MX	5 (5.3)
Missing	3 (3.2)

Calculation of Stromal and Immune Score

The ESTIMATE algorithm was used to calculate immune and stromal scores for each tumor sample with the *estimate* R package (Yoshihara et al., 2013). The stromal, immune, and ESTIMATE scores were compared across different clinical indexes with the Wilcoxon rank-sum test (two groups) or Kruskal-Wallis *H*-test (three or more groups).

To determine the optimal cutoff value for grouping patients, maximally selected rank statistics were employed with the *maxstat* R package (Hothorn and Zeileis, 2008). Based on the cutoff value, all samples were divided into the high and low score groups.

Identification of Differentially Expressed Genes and Functional Enrichment Analysis

Linear models were used to identify differentially expressed genes (DEGs) between the two immune/stromal groups (high score group vs. low score group) using the *limma* R package (Ritchie et al., 2015). A false discovery rate (FDR) adjusted $p < 0.05$ combined with a simultaneous absolute value of >1 for logFC was set as the threshold for DEG identification. Next, upregulated and downregulated genes were identified based on the stromal and immune scores.

Functional enrichment analysis of the DEGs with the Kyoto Encyclopedia of Genes and Genomes (KEGG) and gene ontology (GO) was carried out using the *clusterProfiler* R package (Yu et al., 2012). The GO terms are classified under biological process (BP), cellular component (CC), and molecular function (MF).

Identification of Key Prognostic Genes

In order to identify key prognostic genes, we conducted a comprehensive analysis. First, we used a Cox regression analysis to estimate the association between the expression of all DEGs and the overall survival time of patients. Tumor samples were grouped into the “high expression group” and “low expression group” with the median gene expression level set as the cutoff. Through this method, 67 genes for which the $p < 0.05$ were identified as candidate prognostic genes.

Next, a least absolute shrinkage and selector operation (LASSO) algorithm was used to identify key prognostic genes with the *glmnet* R package (Friedman et al., 2010). Clinical variables, including age, sex, tumor grade, and tumor stage, were included. Lambda.min is the cutoff point at which the minimum mean cross-validated error occurs. Genes or indexes whose coefficient was not 0 at lambda.min were selected for further analysis.

External Validation of Key Prognostic Genes

The genes selected by the LASSO method were further validated in a GEO dataset (GSE44021). We extracted 73 samples of ESCC tumor tissue from the dataset, and the immune and stromal scores were calculated based on their expression profile. The correlation of the prognostic genes selected by LASSO to the immune/stromal scores was evaluated utilizing a Spearman's correlation test.

Risk Score Calculation and Model Construction

The risk scores for every tumor sample were calculated using the following formula:

$$\text{Risk score} = \frac{\sum_{i=1}^n \text{Exp}_i * \frac{\pm 1}{D_i}}{\sum_{i=1}^n \frac{1}{D_i}}$$

In the above formula, *Exp* represents the expression level of the gene, \pm is the positive or negative sign for the regression coefficient of the gene calculated with the LASSO method, and *D* represents the variance in the expression level of the gene in all the samples.

To verify that this risk score can independently predict the prognosis of patients with ESCC, multivariate Cox survival analysis was performed with sex and age as covariates.

Estimation of Immune Cell Fractions

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 (Newman et al., 2015). Normalized gene expression data were analyzed using the CIBERSORT algorithm by running 1,000 permutations. The CIBERSORT *p*-value reflects the statistical significance of the results, and a threshold <0.05 is recommended. Finally, samples with CIBERSORT $p < 0.05$ were included in the analysis of the correlation between risk scores and immune cell types.

Statistical Analysis

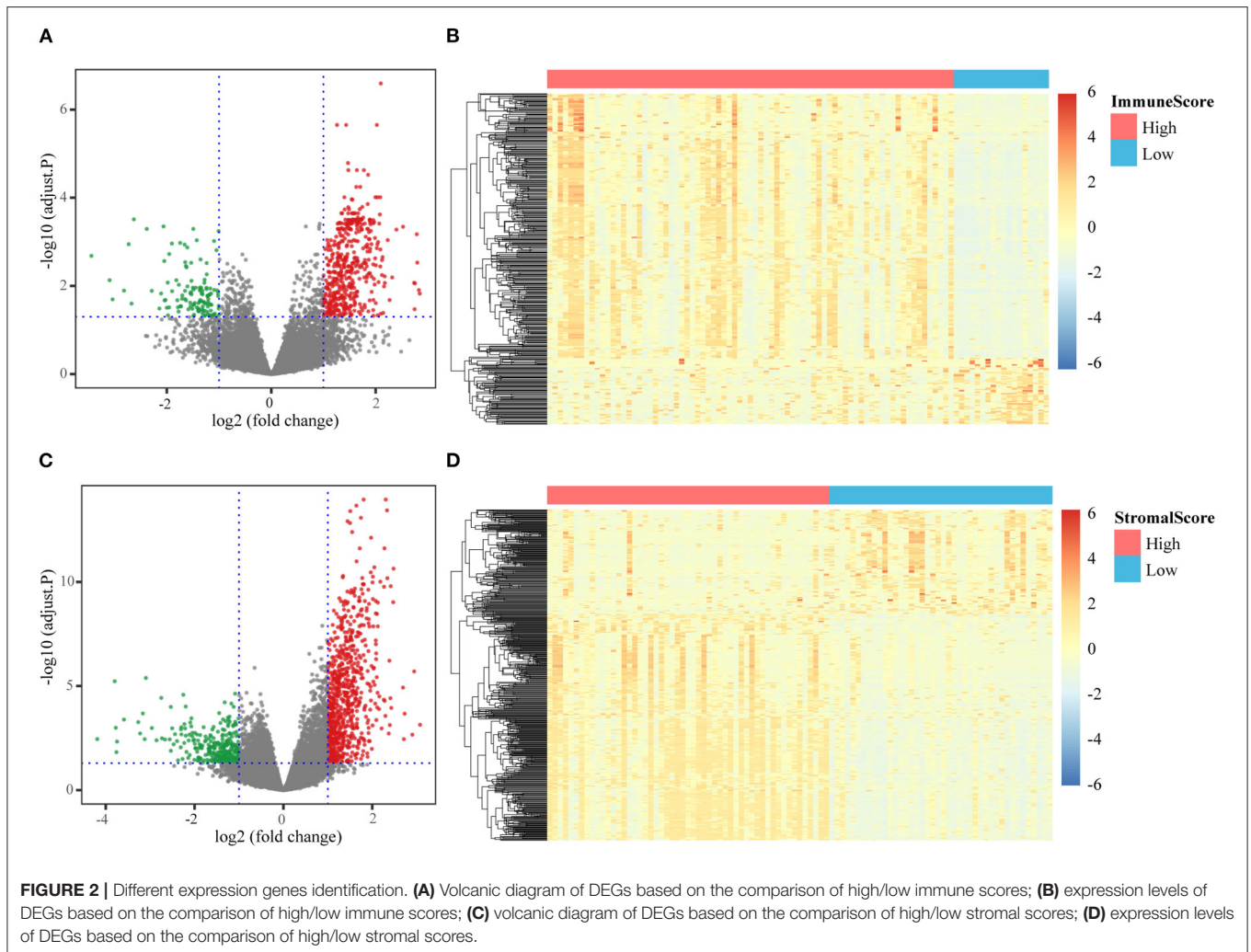
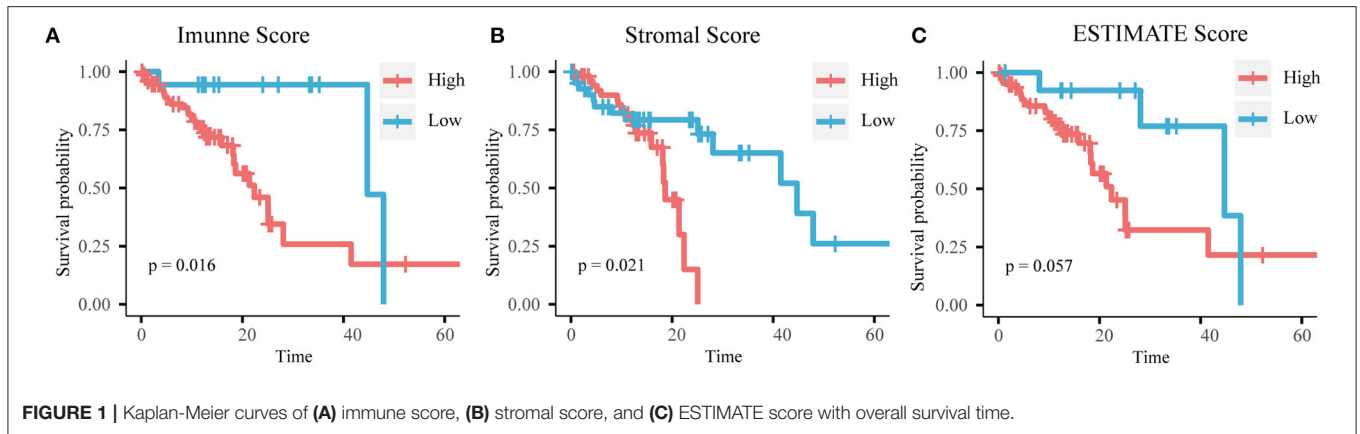
The correlation between the risk scores and immune cell infiltration or TME scores were evaluated using Spearman's correlation analysis. Survival curves were compared using the Kaplan-Meier method and the log-rank test. A ROC curve was used to calculate the AUC value for 1 and 2-year survival. Hazard ratios (HRs) and 95% confidence intervals (CIs) were calculated using Cox regression analysis. All tests were two-sided, and a $p < 0.05$ was considered to indicate significance, unless stated otherwise. All analyses were performed with R version 4.0.2.

RESULTS

Association of High Immune and Stromal Scores With Poor Prognosis

The stromal, immune, and ESTIMATE scores of all 95 ESCC tumor samples were calculated using the ESTIMATE algorithm based on their gene expression profile. The immune scores ranged from $-1,190$ to $2,705$; the stromal scores, from $-1,370$ to $1,136$; and the ESTIMATE scores, from $-2,232$ to $3,231$. There was no significant difference in the immune scores according to any of the clinical indexes, and the stromal or ESTIMATE scores did not show significant differences either (**Supplementary Figure 1**).

To determine whether there was a correlation between the stromal/immune/ESTIMATE scores and overall survival, we separated samples into high- and low-score groups according to the cutoff values determined by maximally selected rank statistics (**Supplementary Figure 2**). According to the immune scores, 77 patients were assigned to the high score group and 22 patients were assigned to the low score group. According to the stromal scores, 53 patients were assigned to the high score group and 42 patients were assigned to the low score group. According to the ESTIMATE scores, 81 patients were assigned to the high score group and 14 patients were assigned to the low score group. Kaplan-Meier survival analysis showed that a high immune score and high stromal score were associated with poor overall survival (**Figure 1**).



Identification of DEGs and Functional Analysis

Differential expression analysis was performed between the high- and low-score groups. When the threshold of $|\log_2FC|$ and FDR were set at >1 and <0.05 , respectively, 459 upregulated

genes and 124 downregulated genes were identified based on the immune scores (Figure 2A), and 767 upregulated genes and 285 downregulated genes were identified based on the stromal scores (Figure 2B). Using hierarchical clustering analysis, we found that these DEGs could be used to distinguish tumors

with high stromal/immune scores from those with low scores (Figures 2C,D).

All the DEGs were included in GO and KEGG functional enrichment analyses. The DEGs based on the immune score, were mainly enriched in BP related to immune response and immune cell activation, CC was associated with immune receptor activity, and MF terms were associated with the external side of the plasma membrane (Supplementary Figure 3A). The top KEGG pathway identified was the hematopoietic cell lineage pathway (Supplementary Figure 3C). The DEGs based on the stromal score were mainly enriched in BP terms that were strongly linked to extracellular matrix organization and extracellular structure organization, in CC terms that were associated with the collagen-containing extracellular matrix, and MF terms that were significantly related to extracellular matrix structural constituents (Supplementary Figure 3B). The top KEGG pathway identified was the cell adhesion molecule pathway (Supplementary Figure 3D).

Identification of Key Prognostic Genes Based on the Machine Learning Method

To evaluate the prognostic effect of DEGs, Cox regression analysis was performed on the expression of 1,514 DEGs and overall survival. Finally, 67 genes were found to have a significant association with overall survival (Supplementary Table 1).

We used the LASSO method to select key prognostic genes from these 67 candidate genes based on the clinical features (Supplementary Figure 4). Thirteen genes and the clinical indicators sex and age were identified (Table 2). Furthermore, the ability of these 15 features to predict survival was determined using the cv.glmnet function. ROC curve analysis revealed that the AUC value was 0.875 (Figure 3A).

We used the 13 genes (*ADAMTS16*, *LOC51089*, *CH25H*, *CORO2B*, *DLGAP1*, *GYS2*, *HAL*, *MXRA8*, *NPTX1*, *OTX1*, *RET*,

SLC24A2, and *SPI1*) selected by the LASSO algorithm for further investigation. We also analyzed the association between the expression of the 13 genes and overall survival using Cox regression analysis. Age and sex were inputted into each regression model. The high expression levels of *LOC51089* (HR = 1.34, 95% CI = 1.04–1.72), *CH25H* (HR = 1.20, 95% CI = 1.01–1.43), *CORO2B* (HR = 1.42, 95% CI = 1.13–1.78), *RET* (HR = 1.25, 95% CI = 1.04–1.50), and *SPI1* (HR = 1.41, 95% CI = 1.04–1.92) were positively correlated with overall survival, while high expression levels of *MXRA8* (HR = 0.72, 95% CI = 0.55–0.93) and *SLC24A2* (HR = 0.69, 95% CI = 0.52–0.91) were negatively correlated with overall survival (Table 2). The correlation between the 13 genes and the infiltration of immune cells was identified by Pearson's correlation analysis (Supplementary Table 2).

External Validation of Key Prognostic Genes

An external ESCC dataset from the GEO database was used to validate the correlation between the expression of the 13 key prognostic genes and the immune/stromal scores. Immune and stromal scores were calculated based on the gene expression profile, and 10 out of 13 genes were matched in this ESCC dataset. Among the 10 genes, the correlation of five genes to the immune/stromal scores was validated, including *CH25H*, *HAL*, *MXRA8*, *SLC24A2*, and *SPI1* (Supplementary Figure 5).

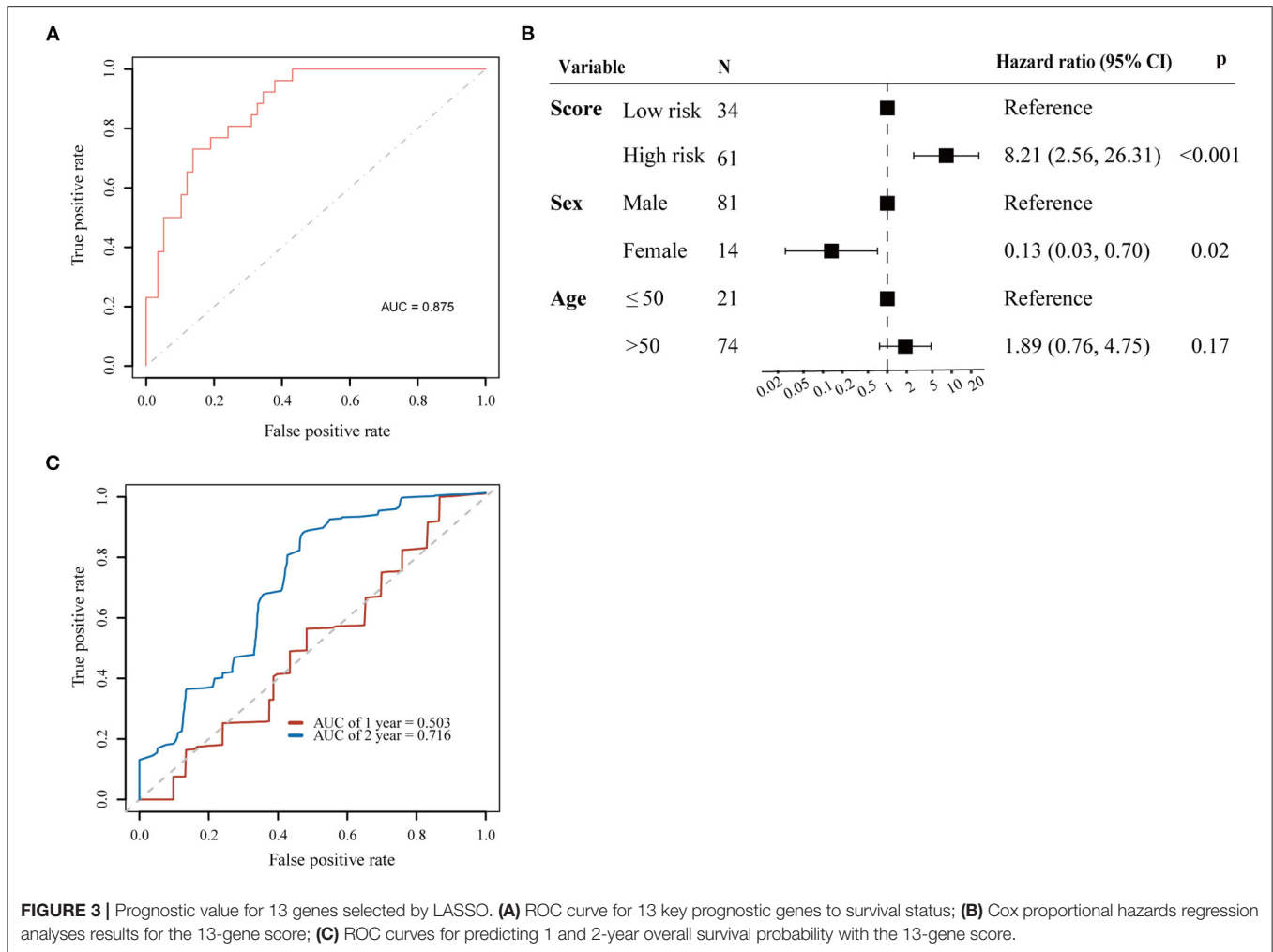
Construction of a 13-Genes Risk Score

The risk score was calculated based on the expression level (EL) of all 13 key prognostic genes. Risk score = $[(-0.22 \times \text{EL of } ADAMTS16) + (0.49 \times \text{EL of } LOC51089) + (0.26 \times \text{EL of } CH25H) + (0.42 \times \text{EL of } CORO2B) + (-0.21 \times \text{EL of } DLGAP1) + (-0.20 \times \text{EL of } GYS2) + (0.25 \times \text{EL of } HAL) + (-0.60 \times \text{EL of } MXRA8) + (0.32 \times \text{EL of } NPTX1) + (-0.53 \times \text{EL of } OTX1) +$

TABLE 2 | Thirteen key prognostic genes and two clinical indexes selected by the LASSO method.

Clinical index/gene	Description	Coefficient	Hazard ratio	95% CI	P-value
Age		0.0408			
Sex		-0.3281			
RET	Ret proto-oncogene	0.3421	1.25	1.04–1.50	0.02
LOC51089 (C1QA)	Complement C1q A chain	0.1787	1.34	1.04–1.72	0.02
CORO2B	Coronin 2B	0.2032	1.42	1.13–1.78	0.003
CH25H	Cholesterol 25-hydroxylase	0.1026	1.20	1.01–1.43	0.04
SPI1	Spi-1 proto-oncogene	0.1117	1.41	1.04–1.92	0.03
HAL	Histidine ammonia-lyase	0.0783	0.97	0.81–1.17	0.76
NPTX1	Neuronal pentraxin 1	0.0077	1.13	0.95–1.33	0.17
DLGAP1	DLG associated protein 1	-0.0174	0.95	0.81–1.11	0.53
OTX1	Orthodenticle homeobox 1	-0.0409	0.82	0.66–1.03	0.09
SLC24A2	Solute carrier family 24 member 2	-0.0341	0.69	0.52–0.91	0.008
ADAMTS16	ADAM metalloproteinase with thrombospondin type 1 motif 16	-0.1505	0.84	0.70–1.00	0.053
GYS2	Glycogen synthase 2	-0.2956	0.88	0.72–1.07	0.20
MXRA8	Matrix remodeling associated 8	-0.4201	0.72	0.56–0.93	0.01

Multivariate Cox survival analysis was performed for every gene when sex and age were used as covariates.



$(0.33 \times \text{EL of } RET) + (-0.50 \times \text{EL of } SLC24A2) + (0.83 \times \text{EL of } SPI1)/5.17$. All the EL values were logarithmically transformed.

Patients were then separated into high-risk and low-risk groups according to the cutoff point of 0.98 by maximally selected rank statistics. To evaluate the risk score model based on the LASSO algorithm, Cox regression analysis was performed for overall survival time with sex, age, and risk group as covariates. The results of multivariate Cox analysis demonstrated that the risk score could be regarded as an independent predictive factor for OS (HR = 8.21, 95% CI = 2.56–26.31; $p < 0.001$) (Figure 3B).

Next, we established a prognostic nomogram to predict 1 and 2-year OS in 95 ESCC patients. The AUC values for 1 and 2-year OS were 0.503 and 0.716, respectively (Figure 3C).

Correlation of Risk Score to TME-Score and Immune Cell Infiltration

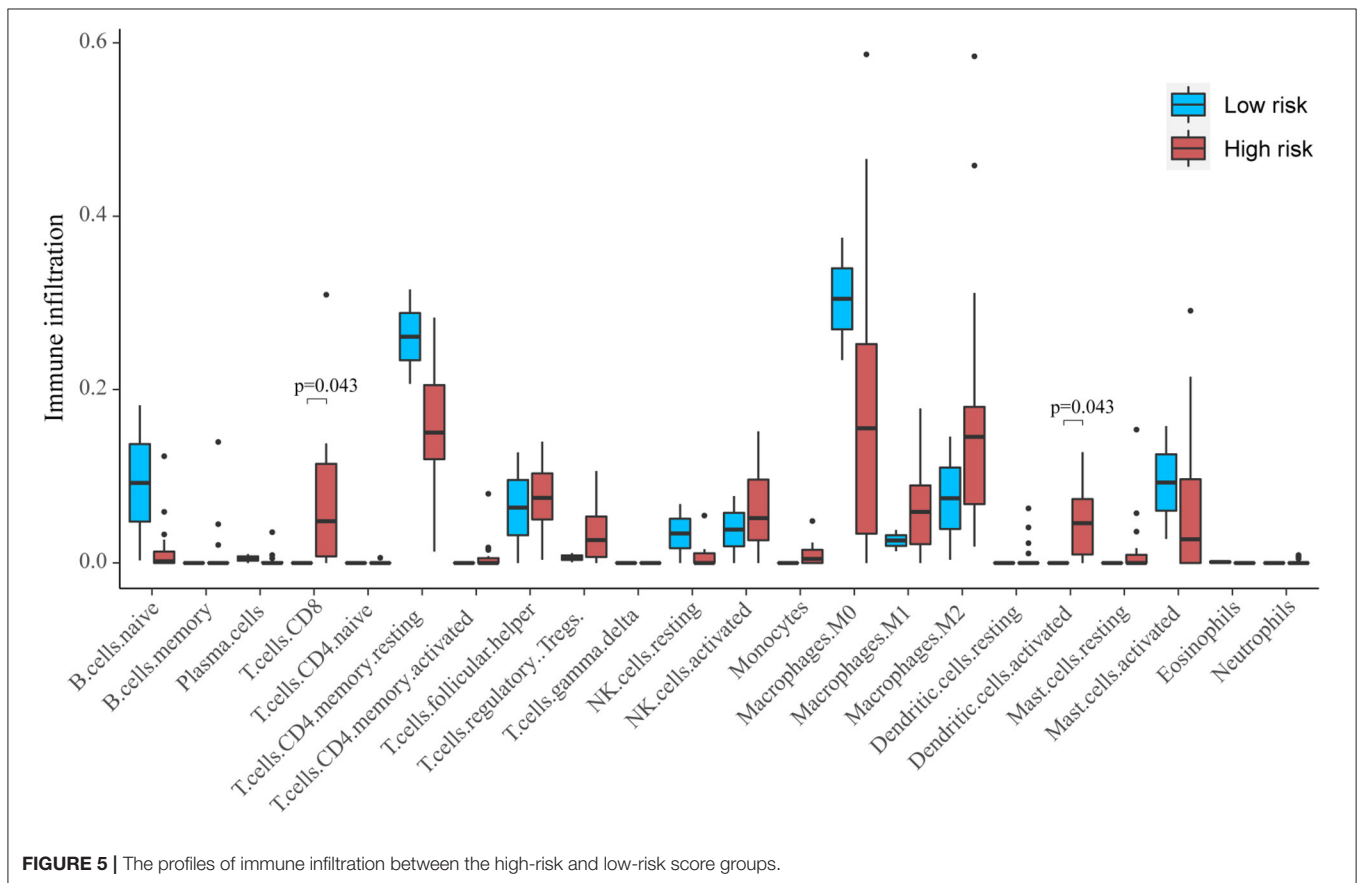
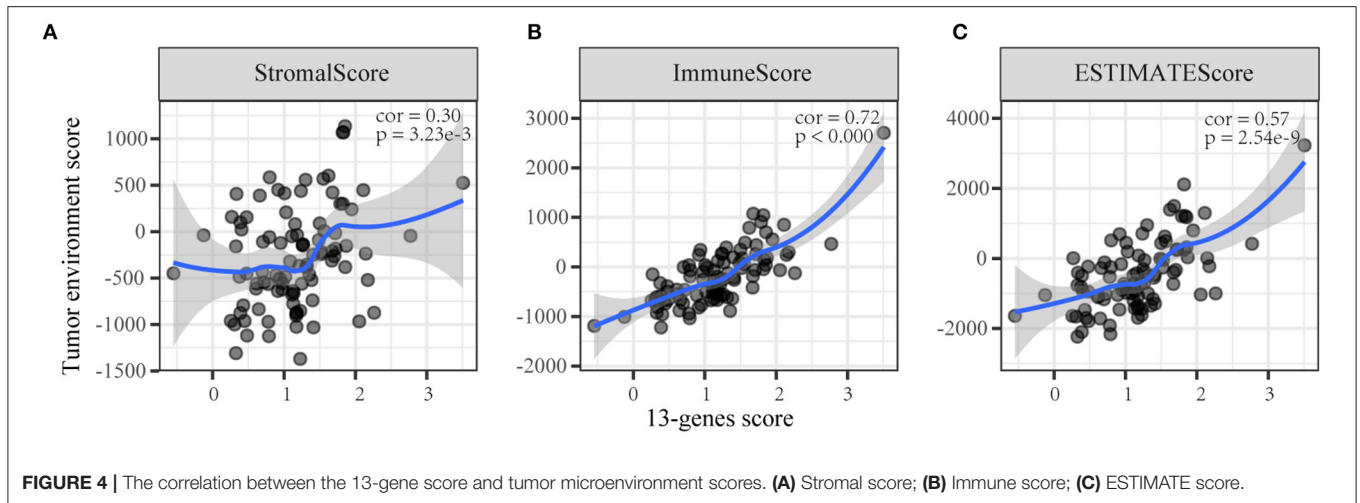
We conducted a Spearman's correlation test to evaluate the correlation between the TME score and the 13-gene risk score. The results showed that the 13-gene risk score was

significantly correlated with the immune/stromal/ESTIMATE scores (Figure 4).

Based on CIBERSORT, unmatched samples with $p > 0.05$ were removed. A total of 23 tumor samples, including 21 samples in the high-risk group and 2 samples in the low-risk group were used for further analysis. Our results showed that there were significant differences in the proportion of various immune cell fractions. The high-risk group had a higher infiltration of CD8 cells ($p = 0.043$) and dendritic cells ($p = 0.043$) (Figure 5). In addition, the correlation of the 13-gene risk score with CD8 cell, regulatory T cell, and resting macrophage (M0) fractions was confirmed using a Spearman's correlation test (Figure 6).

DISCUSSION

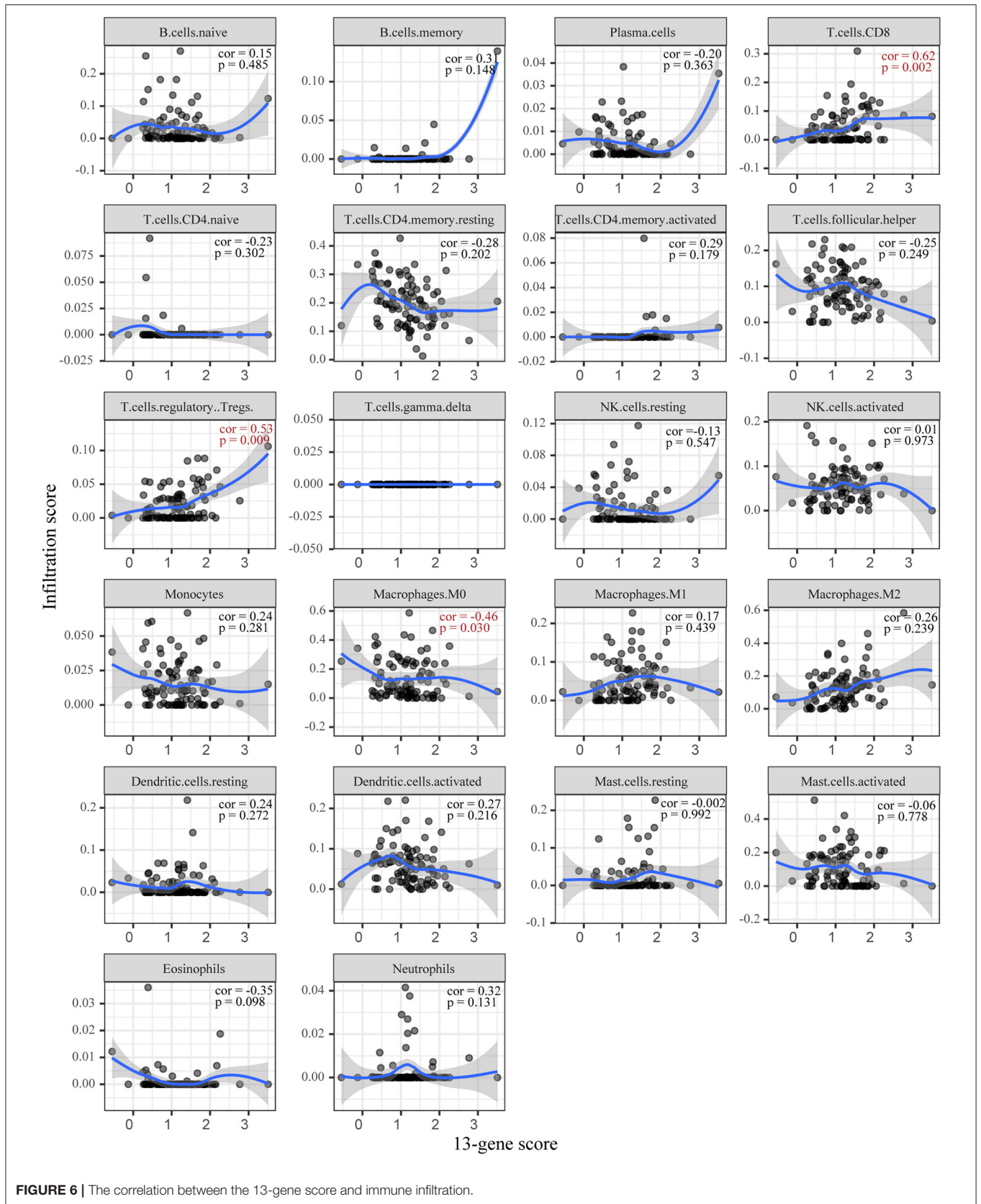
Although significant benefits of new therapeutic strategies have been reported in the past few decades, the overall survival rates of ESCC remain unsatisfactory (Cunningham et al., 2006; Baba et al., 2018). This is because ESCC is characterized by high heterogeneity that is attributable to the heterogeneity of tumor cells and the tumor environment, and as a result, the response to



therapy is highly varied among patients (Marshall and Djamgoz, 2018). Therefore, it is critical to screen for prognostic markers related to the TME of ESCC. In the present study, we identified 13 TME-related genes and developed a risk score signature for ESCC. This signature could be used to efficiently determine the overall survival time of ESCC patients.

Previous studies have established several molecular models to predict the long-term survival of patients with ESCC (Wen et al., 2014; Chen W. et al., 2020; Wang C. et al., 2020). These studies often overlook the role of TME in ESCC. The

ESTIMATE algorithm is a widely accepted and verified algorithm in studies on various cancers, such as endometrial cancer, cervical squamous cell carcinoma, and breast cancer (Priedigkeit et al., 2017; Pan et al., 2019; Chen P. et al., 2020). In this study, we calculated the stromal and immune scores of ESCC tumor tissues with the ESTIMATE algorithm in order to predict the level of the infiltration of stromal and immune cells. ESCC patients with high immune/stromal scores had a poorer prognosis than those with low scores; this indicates that the TME composition affects the clinical outcomes of ESCC patients. Furthermore, the



DEGs based on immune score were significantly enriched in the regulation of T cell activation, the regulation of lymphocyte activation, and lymphocyte differentiation, while the DEGs based on stromal score were mainly enriched in extracellular matrix organization, extracellular structure organization, and cell-substrate adhesion. These findings confirm the role of immune cells and TME heterogeneity on the clinical outcomes of ESCC.

Cox regression analysis was performed to determine the association between the expression of DEGs and survival, and 67 DEGs were identified as being associated with overall survival. Then, using the LASSO algorithm, 13 key prognostic genes and two clinical indexes were selected. These genes had independent prognostic values, and the correlation of their expression to immune/stromal scores was also validated in the external GEO dataset. These genes have previously been reported to be associated with different malignant tumors in various ways. For example, abnormally methylated CH25H has been found to be a prognostic marker for lung squamous cell carcinoma patients (Gao et al., 2019). Additionally, low CH25H levels in leukocytes from melanoma patients were correlated with poor prognosis (Ortiz et al., 2019). CORO2B was found to be involved in many biological processes of malignant transformation in BEAS-2B cells induced by cigarette smoke (Wang et al., 2019). LOC51089, also called C1QA, was found to be involved in the innate immune system and was associated with the expression of PD-L1 (Olkhov-Mitsel et al., 2020), and its abnormal expression in tumor tissues was confirmed in head and neck squamous cell carcinoma and clear cell renal cell carcinoma (Yu et al., 2019; Apanovich et al., 2020). MXRA8 is one of the predicted tumor stroma-specific markers in various cancers (Kiflemariam et al., 2015). Accordingly, kidney renal clear cell carcinoma patients with high expressions of MXRA8 had worse overall survival (Li and Xu, 2019). RET is an important proto-oncogene that can undergo oncogenic activation through both cytogenetic rearrangement and the activation of point mutations, and alterations in RET have been identified as being oncogenic in multiple malignancies (Subbiah et al., 2020). Moreover, alterations in SLC24A2, a potassium-dependent sodium-calcium exchanger, were observed in pancreatic ductal adenocarcinoma (Wang et al., 2017). Alterations in SPI1 lead to cellular proliferation and differentiation arrest, resulting in oncogenic subversion (Roos-Weil et al., 2019). Finally, high expression levels of HAL contributed to poor prognoses in breast cancer patients (Wang C. Y. et al., 2020). Our study demonstrated that all these genes may be crucial biomarkers for predicting survival outcome in ESCC. Therefore, 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 prognosis biomarkers as well as their association with immunotherapy efficacy.

It is critical to accurately predict the prognosis of patients with ESCC for the selection of appropriate treatment. In this regard, combining different independent prognostic variables into one risk score can significantly improve prognostic potential. In this study, we constructed and verified a 13-gene risk score for ESCC based on stromal/immune scores. The 13-gene model can be used as a prognostic tool independently of other clinical and pathological features. 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 TME from different aspects and has the potential to be appropriate for rational diagnosis and individualized treatment.

Several limitations of this study should be noted. First, given that the study was retrospective in nature, the risk scores need to be validated in a large cohort. Second, in this study, the associations between the 13 genes and the biological mechanisms of ESCC have not been clarified. Thus, further experiments are required to validate the exact mechanism of these 13 genes under *in vitro* and *in vivo* conditions.

CONCLUSION

In conclusion, we developed a robust TME-related gene signature for the prognostic prediction of ESCC based on samples deposited in the TCGA database. Our signature could reflect the TME features and prognosis of ESCC. These findings are the basis for more studies on the specific roles of these TME-related genes in the development and progression of ESCC.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: The Cancer Genome Atlas (<https://portal.gdc.cancer.gov/>) and Gene Expression Omnibus (GEO) Datasets (<https://www.ncbi.nlm.nih.gov/>).

AUTHOR CONTRIBUTIONS

DZ and CQ proposed the study concept, design, drafted the manuscript, collected, analyzed, and interpreted the data. XQ and HW participated in revising the manuscript. All authors contributed to the article and approved the submitted version.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmolb.2020.599475/full#supplementary-material>

REFERENCES

Apanovich, N., Peters, M., Apanovich, P., Mansorunov, D., Markova, A., Matveev, V., et al. (2020). The genes-candidates for prognostic markers of

metastasis by expression level in clear cell renal cell cancer. *Diagnostics*10:30. doi: 10.3390/diagnostics10010030

Baba, Y., Yoshida, N., Kinoshita, K., Iwatsuki, M., Yamashita, Y. I., Chikamoto, A., et al. (2018). Clinical and prognostic features of patients with esophageal

- cancer and multiple primary cancers: a retrospective single-institution study. *Ann. Surg.* 267, 478–483. doi: 10.1097/SLA.0000000000002118
- Chen, P., Yang, Y., Zhang, Y., Jiang, S., Li, X., and Wan, J. (2020). Identification of prognostic immune-related genes in the tumor microenvironment of endometrial cancer. *Aging* 12, 3371–3387. doi: 10.18632/aging.102817
- Chen, W., Li, H., Ren, J., Zheng, R., Shi, J., Li, J., et al. (2020). Selection of high-risk individuals for esophageal cancer screening: a prediction model of esophageal squamous cell carcinoma based on a multicenter screening cohort in rural China. *Int. J. Cancer.* 148, 329–339. doi: 10.1002/ijc.33208
- Cunningham, D., Allum, W. H., Stenning, S. P., Thompson, J. N., van de Velde, C. J., Nicolson, M., et al. (2006). Perioperative chemotherapy versus surgery alone for resectable gastroesophageal cancer. *N. Engl. J. Med.* 355, 11–20. doi: 10.1056/NEJMoa055531
- Fitzmaurice, C., Akinyemiju, T. F., Al Lami, F. H., Alam, T., Alizadeh-Navaei, R., Allen, C., et al. (2018). Global, regional, and national cancer incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life-years for 29 cancer groups, 1990 to 2016: a systematic analysis for the global burden of disease study. *JAMA Oncol.* 4, 1553–1568. doi: 10.1001/jamaoncol.2018.2706
- Friedman, J., Hastie, T., and Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *J Stat Softw.* 33, 1–22. doi: 10.18637/jss.v033.i01
- Galon, J., Pagès, F., Marincola, F. M., Thurin, M., Trinchieri, G., Fox, B. A., et al. (2012). The immune score as a new possible approach for the classification of cancer. *J. Transl. Med.* 10:1. doi: 10.1186/1479-5876-10-1
- Gao, C., Zhuang, J., Zhou, C., Ma, K., Zhao, M., Liu, C., et al. (2019). Prognostic value of aberrantly expressed methylation gene profiles in lung squamous cell carcinoma: a study based on the cancer genome atlas. *J. Cell. Physiol.* 234, 6519–6528. doi: 10.1002/jcp.27389
- Hegde, P. S., and Chen, D. S. (2020). Top 10 challenges in cancer immunotherapy. *Immunity* 52, 17–35. doi: 10.1016/j.immuni.2019.12.011
- Hothorn, T., and Zeileis, A. (2008). Generalized maximally selected statistics. *Biometrics* 64, 1263–1269. doi: 10.1111/j.1541-0420.2008.00995.x
- Kiflemariam, S., Ljungström, V., Pontén, F., and Sjöblom, T. (2015). Tumor vessel up-regulation of INSR revealed by single-cell expression analysis of the tyrosine kinase and phosphate in human cancers. *Am. J. Pathol.* 185, 1600–1609. doi: 10.1016/j.ajpath.2015.02.019
- Li, S., and Xu, W. (2019). Mining TCGA database for screening and identification of hub genes in kidney renal clear cell carcinoma microenvironment. *J. Cell. Biochem.* 121, 3852–3960. doi: 10.1002/jcb.29511
- Marshall, H. T., and Djamgoz, M. B. A. (2018). Immuno-oncology: emerging targets and combination therapies. *Front. Oncol.* 8:315. doi: 10.3389/fonc.2018.00315
- Newman, A. M., Liu, C. L., Green, M. R., Gentles, A. J., Feng, W., Xu, Y., et al. (2015). Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods* 12, 453–457. doi: 10.1038/nmeth.3337
- Olkhov-Mitsel, E., Hodgson, A., Liu, S. K., Vesprini, D., Bayani, J., Bartlett, J., et al. (2020). Immune gene expression profiles in high-grade urothelial carcinoma of the bladder: a NanoString study. *J. Clin. Pathol.* 1–15. doi: 10.1136/jclinpath-2020-206631
- Ortiz, A., Gui, J., Zahedi, F., Yu, P., Cho, C., Bhattacharya, S., et al. (2019). An interferon-driven oxysterol-based defense against tumor-derived extracellular vesicles. *Cancer Cell* 35, 33–45.e36. doi: 10.1016/j.ccell.2018.12.001
- Pan, X. B., Lu, Y., Huang, J. L., Long, Y., and Yao, D. S. (2019). Prognostic genes in the tumor microenvironment in cervical squamous cell carcinoma. *Aging* 11, 10154–10166. doi: 10.18632/aging.102429
- Pitt, J. M., Marabelle, A., Eggermont, A., Soria, J. C., Kroemer, G., and Zitvogel, L. (2016). Targeting the tumor microenvironment: removing obstruction to anticancer immune responses and immunotherapy. *Ann. Oncol.* 27, 1482–1492. doi: 10.1093/annonc/mdw168
- Priedigke, N., Watters, R. J., Lucas, P. C., Basudan, A., Bhargava, R., Horne, W., et al. (2017). Exome-capture RNA sequencing of decade-old breast cancers and matched decalcified bone metastases. *JCI Insight* 2:e95703. doi: 10.1172/jci.insight.95703
- Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., et al. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 43:e47. doi: 10.1093/nar/gkv007
- Roos-Weil, D., Decaudin, C., Armand, M., Della-Valle, V., Diop, M. K., Ghamlouch, H., et al. (2019). A recurrent activating missense mutation in Waldenström macroglobulinemia affects the DNA binding of the ETS transcription factor SPI1 and enhances proliferation. *Cancer Discov.* 9, 796–811. doi: 10.1158/2159-8290.CD-18-0873
- Rustgi, A. K., and El-Serag, H. B. (2014). Esophageal carcinoma. *N. Engl. J. Med.* 371, 2499–2509. doi: 10.1056/NEJMra1314530
- Senbabaoglu, Y., Gejman, R. S., Winer, A. G., Liu, M., van Allen, E. M., de Velasco, G., et al. (2016). Tumor immune microenvironment characterization in clear cell renal cell carcinoma identifies prognostic and immunotherapeutically relevant messenger RNA signatures. *Genome Biol.* 17:231. doi: 10.1186/s13059-016-1092-z
- Smyth, E. C., Lagergren, J., Fitzgerald, R. C., Lordick, F., Shah, M. A., Lagergren, P., et al. (2017). Oesophageal cancer. *Nat. Rev. Dis. Primers* 3:17048. doi: 10.1038/nrdp.2017.48
- Subbiah, V., Yang, D., Velcheti, V., Drilon, A., and Meric-Bernstam, F. (2020). State-of-the-art strategies for targeting RET-dependent cancers. *J. Clin. Oncol.* 38, 1209–1221. doi: 10.1200/JCO.19.02551
- Wang, C. Y., Chiao, C. C., Phan, N. N., Li, C. Y., Sun, Z. D., Jiang, J. Z., et al. (2020). Gene signatures and potential therapeutic targets of amino acid metabolism in estrogen receptor-positive breast cancer. *Am. J. Cancer Res.* 10, 95–113.
- Wang, C., Zhao, K., Hu, S., Huang, Y., Ma, L., Song, Y., et al. (2020). A predictive model for treatment response in patients with locally advanced esophageal squamous cell carcinoma after concurrent chemoradiotherapy: based on SUVmean and NLR. *BMC Cancer* 20:544. doi: 10.1186/s12885-020-07040-8
- Wang, J., Yu, X. F., Ouyang, N., Zhao, S., Yao, H., Guan, X., et al. (2019). MicroRNA and mRNA interaction network regulates the malignant transformation of human bronchial epithelial cells induced by cigarette smoke. *Front. Oncol.* 9:1029. doi: 10.3389/fonc.2019.01029
- Wang, L., Shao, Z., Chen, S., Shi, L., and Li, Z. (2017). A SLC24A2 gene variant uncovered in pancreatic ductal adenocarcinoma by whole exome sequencing. *Tohoku J. Exp. Med.* 241, 287–295. doi: 10.1620/tjem.241.287
- Wen, J., Yang, H., Liu, M. Z., Luo, K. J., Liu, H., Hu, Y., et al. (2014). Gene expression analysis of pretreatment biopsies predicts the pathological response of esophageal squamous cell carcinomas to neo-chemoradiotherapy. *Ann. Oncol.* 25, 1769–1774. doi: 10.1093/annonc/mdu201
- Yoshihara, K., Shahmoradgoli, M., Martínez, E., Vegesna, R., Kim, H., Torres-García, W., et al. (2013). Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat. Commun.* 4:2612. doi: 10.1038/ncomms3612
- Yu, D., Ruan, X., Huang, J., Hu, W., Chen, C., Xu, Y., et al. (2019). Comprehensive analysis of competitive endogenous RNAs network, being associated with esophageal squamous cell carcinoma and its emerging role in head and neck squamous cell carcinoma. *Front. Oncol.* 9:1474. doi: 10.3389/fonc.2019.01474
- Yu, G., Wang, L. G., Han, Y., and He, Q. Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 16, 284–287. doi: 10.1089/omi.2011.0118

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Zhang, Qian, Wei and Qian. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.