Identification and Validation of a Hypoxia-Immune-Based Prognostic mRNA Signature for Oral Squamous Cell Carcinoma

Shaohua Lv
First Affiliated Hospital of Harbin Medical University
Jianhao Li
First Affiliated Hospital of Harbin Medical University
Songlin Piao
First Affiliated Hospital of Harbin Medical University
jichen Li (001936@hrbmu.edu.cn)
Harbin Medical University https://orcid.org/0000-0001-8599-1197

Primary research

Keywords: Oral squamous cell carcinoma, hypoxia, immune, mRNA signature, Machine learning

Posted Date: June 15th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-596220/v1

Version of Record: A version of this preprint was published at Journal of Oncology on February 7th, 2022.
See the published version at https://doi.org/10.1155/2022/5286251.

Abstract

Background: Oral squamous cell carcinoma (OSCC) is a frequently encountered head and neck malignancy. Increasing evidence points towards an aberrant immune response and chronic cell hypoxia in the development of OSCC. However, there is a lack of a reliable hypoxia-immune-based gene signature that may serve to accurately prognosticate OSCC.

Methods: The mRNA expression data of OSCC patients was extracted from the TCGA database. Hypoxia status was identified using the t-distributed Stochastic Neighbor Embedding (t-SNE) algorithm. Both ESTIMATE and single-sample gene-set enrichment analysis (ssGSEA) was used for further evaluation of immune status. The DEGs in different hypoxia and immune status were determined. A Machine learning method-Least Absolute Shrinkage and Selection operator (LASSO) Cox regression analysis allowed us to select prognostically significant hypoxia- and immune-related mRNAs in order to construct prognostic gene signature to predict the overall survival (OS) of OSCC patients.

Results: A total of 773 DEGs were classified into either Hypoxia_High and Hypoxia_Low groups. Immune-associated DEG expressions were used to divide individuals into Immune_High, Immune_ Medium and Immune_Low groups. A total of 193 mRNAs which were significant in both immune function and hypoxia status were identified. With the Lasso Cox regression model, 8 signature mRNAs (FAM122C, RNF157, RANBP17, SOWAHA, KIAA1211, RIPPLY2, INSL3, and DNAH1) associated with OS were selected for further calculation of their respective risk scores. The risk score showed a significant association with age, perineural and lymphovascular invasion. In the GEO validation cohort, a better OS was observed in patients from the low-risk group in comparison to those in the high-risk group. High-risk patients also demonstrated different immune infiltration characteristics from the low-risk group. All individuals from the TCGA OSCC cohort showed similar trends in all 6 immune checkpoints, with those of the low-risk group yielding higher immune indicator scores in contrast to their high-risk counterparts.

Conclusion: The hypoxia-immune-based gene signature has prognostic potential in OSCC.

Introduction

The incidence of lip and oral cavity cancers have been on the rise, with an estimated 246,420 cases and 119,693 deaths occurring globally in 2018 [1]. Oral squamous cell carcinoma (OSCC) is the dominating subtype of oral cancer which unfortunately harbors a dismal 5-year survival rate [2]. There is still an urgent need to discover accurate prognostic biomarkers and effective drug targets for OSCC.

Solid tumors such as OSCC commonly feature cell hypoxia, which has universally been taken to indicate a negative patient outcome [3, 4]. Cell hypoxia has evolved to represent a significant feature in cancer [5]. Hypoxic conditions could trigger the migration and invasion of OSCC cells [6–8]. Angiogenesis is also a critical process in OSCC initiation and progression. Hypoxia is also able stimulate angiogenesis through activation of pro-angiogenic factors [9]. It is undeniable that cellular hypoxia contributes significantly towards the malignant properties of OSCC.

The immune system is crucial in OSCC tumorigenesis, progression, and prognosis [10–12]. Enhanced PD-1 and PD-L1 expressions were observed in OSCC patient serum and tissue samples [13]. Premalignant lesions were also found to harbor raised PD-L1 expressions [14]. Inhibiting the PD-1/PD-L1 axis is one of the most promising means to activate antitumor immunity [15]. A number of immune-related prognostic biomarkers have been identified in OSCC. Meehan K et al. had previously constructed an immune-related signature which correlated with OSCC patient prognosis [16]. At least seven immune checkpoints have been documented to be closely related to poorer patient outcomes in OSCC [17]. One of these is the CTSG gene, which is an immune-related gene that has been identified as an independent biomarker and therapeutic target for OSCC [18]. Although these immune-related prognostic biomarkers have been discovered, few studies have explored the potential benefit of a combined assessment of both immunological characteristics and the hypoxic microenvironment in OSCC [19].

Several studies have alluded towards the clinical significance of potential interactions between immune function and cellular hypoxia across various malignancies [20, 21]. However, a reliable hypoxia-immune-based prognostic gene signature has not yet been established for OSCC. We sought to fill this knowledge gap in this study. Here, we first accessed the mRNA expression profiles of OSCC patients from the TCGA database. Cell hypoxia was identified using the t-SNE algorithm while the immune status and immune-related DEGs were identified using the ESTIMATE and ssGSEA algorithms. DEGs which were significant in both immune and cellular hypoxia phenomena were identified. This pool of genes was then subjected to a Machine learning assessment to discern those which were potentially prognostic. We successfully constructed a hypoxia-immune-based gene signature that was useful in predicting the overall survival (OS) of OSCC patients.

Materials And Methods

Data collection and preprocessing

The discovery cohort included 303 OSCC individual data from the "TCGA-HNSC" project in The Cancer Genome Atlas (TCGA). The Genomic Data Commons (available at: https://portal.gdc.cancer.gov) database was then accessed to extract the corresponding gene expression data. Data was externally validated using an independent cohort. The expression profile of GSE41613 was downloaded from the Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/) based on the GPL570 Affymetrix Human Genome U133 Plus 2.0 Array platform. A total of 97 OSCC patient data were included in this validated study.

RNA-sequencing data (FPKM values) extracted from the TCGA database was translated into transcripts per kilobase million (TPM) values and normalized using log2 (TPM + 1). Raw GEO database-derived data was obtained using the RMA algorithm. Background adjustment, quantile normalization, and final summation of oligonucleotides per transcript was carried out using the median polish algorithm of the Affy software package. Finally, the different probe IDs were converted into their respective gene symbols

to annotate them, and the repeated gene expression values were averaged. All patients with included clinical information and survival data were also incorporated in this study.

Identifcation of hypoxia status and hypoxiarelated differentially expressed genes (DEGs)

The presence of cell hypoxia was determined using the t-distributed Stochastic Neighbor Embedding (t-SNE) algorithm. This non-parametric, unsupervised method was able to distinguish between patient clusters based on provided hallmarks or signatures. Based on the seven hypoxic signatures of Buffa, Elvidge, Eustace, Hu, Ragnum, Sorensen and Winter [22], a nonlinear dimensionality reduction algorithm t-SNE was used to measure the Euclidean distance of any two patients in the TCGA cohort which was then condensed into two-dimensional points. Three clusters were identified. Moreover, we obtained genes using the KEGG HIF-1 signaling pathway (https://www.kegg.jp/; ID:04066) to assess the hypoxic status. Of these genes, 14 were identified as “Increased oxygen delivery” genes and 12 were identified as “Reduced oxygen consumption” genes. Finally, the “DESeq2” software package was used to discern between DEGs significant to low- or high-hypoxic groups (P-value < 0.05 and |log2 (fold change)|>1).

Identification of immune status and immunerelated DEGs

We evaluated enrichment degrees of a total of 28 immune cells within each TCGA OSCC samples using ssGSEA method. The OSCC samples arbitrarily cohorted into high-, medium- and low-immune group using hierarchical clusters (namely “Immune_Low”, “Immune_Medium” and “Immune_High”) based on the above immune matrix. Both degrees of immunocyte infiltration degree (immune score) and stromal level (stromal score) were assessed using the Estimation of Stromal and Immune cells in Malignant Tumor tissues using Expression data (ESTIMATE) algorithm in TCGA OSCC samples to validate the above immune status grouping. The immune-related DEGs between the Immune_Low group and the Immune_High group were identified by the "DESeq2" software package using similar parameters as in the cell hypoxia analysis. The Metascape (https://metascape.org/) database was used to construct GO and KEGG enrichment pathways of the selected DEGs.

Construction and verification of prognostic signatures related to hypoxia and immune status

DEGs which appeared to be significant in both cell hypoxia and immune function were further analyzed using the Least Absolute Shrinkage and Selection operator (LASSO) and univariate Cox regression analyses. A prognostic immune-related model was constructed for prediction of the survival (OS) of OSCC cases. First, the relationship between the expression of each overlapping gene and OS was calculated using a univariate Cox regression model. The LASSO method then allowed us to select variables in the Cox regression model to identify genes with prognostic value. The standard was taken to be the optimal value of lambda. All patients were then subjected to a risk score calculation as follows:

Risk score = $\sum$ (coefficienti * expression of the signature genei)

The risk score was defined as the sum of the product of the prognostic signature gene expression levels and the corresponding LASSO model-derived coefficients. The optimal cut point was determined by a method of maximally selected rank statistics based on an individual level risk score to further refine a means to stratify patients with OSCC according to their prognosis.

Relationship between risk scores and the immune microenvironment of OSCC

Profiles of 22 immune cell infiltrates were determined in both the low- and high-risk groups using the CIBERSORT algorithm. Using gene expression data, the CIBERSORT is a deconvolution algorithm specifically designed to evaluate cellular composition in tissues. In addition, two immune checkpoints (CTLA-4 and PD-1), APM, CYT, TILs, and TIS allowed for further analysis of the relationship between the low- and high-risk groups.

The major histocompatibility complex (MHC) molecules were subjected to relative antigen presentation mechanism (APM) calculation. The degree of immune cytotoxic activity (CYT) based on granzyme A (GZMA) and perforin-1 (PRF1), both of which are related to CD8 + T cell activation, was calculated on previously reported formulas. Tumor prognosis is known to be associated with the density of T cells tumor infiltration. The proportion of tumor infiltrating lymphocytes (TILs) was calculated. Lastly, we used the tumor inflammation signature (TIS) which represents a commonly used 18-gene signature in research to quantify background adaptive immune responses which are normally inhibited in tumors. The TIS score was derived in terms of the mean of log2-FPKM gene expression of the selected marker genes.

Statistical analysis

The R version 3.6.1 and its related packages were used to carry out all data analyses. The t-SNE algorithm was performed using the “Rtsne” of the R package on the basis of nonlinear dimensionality reduction. The “estimate” package was used to determine immune scores. The “glmnet” package was used to carry out Lasso Cox regression modeling. To identify independent risk factors for survival, univariate and multivariate Cox regression analyses were employed after adjustment of covariates. The “survival” package was used to perform Cox regression model and Kaplan Meier analyses. A p value of < 0.05 suggested that the results were statistically significant.

Results

Hypoxia status and hypoxia-related DEGs in OSCC

In the TCGA cohort, cell hypoxia status was evaluated using ssGSEA analysis based on eight hypoxia signatures (Buffa, Elvidge, Eustace, Hu, Ragnum, Seigneuric, Sorensen and Winter) [22] (Fig. 1A). Based on the quantitative score of each hypoxia signature, we found that the “Seigneuric” signature had the lowest correlation with the others and was therefore not included in this study (Fig. 1B). The remaining seven gene signatures were used in the non-linear dimensionality reduction algorithm calculation of two-

dimensional points between any two patients (see methods for details). As shown in Fig. 1C, three clusters of patients were identified and each patient was assigned to their nearest cluster. A total of 64, 142, and 97 cases were classified into Cluster 1, Cluster 2 and Cluster 3. There was no significant difference in terms of patient survival among the three clusters (log-rank test, p = 0.088) (Fig. 1D). However, there was a significant difference when comparing Cluster 1 with Cluster 2 and Cluster 3 (log-rank test, p = 0.029). Clusters 2 and 3 were then combined for further comparison against Cluster 1, which appeared to possesses the most favorable overall survival outcome. Clusters 2 and 3 may differ from Cluster 1 in terms of degree of cellular hypoxia. We further explored differences of expressions of KEGG HIF-1 pathway molecules between the two new clusters (Cluster 1 vs Clusters 2 and 3). The resultant identified genes were classified as being involved in “Reduced oxygen consumption” (12 genes) or “Increased oxygen delivery” (14 genes). Of those 14 genes involved in increased oxygen delivery, 11 (78.57%) were found to be overexpressed in Clusters 2&3 compared with those in Cluster 1 (Fig. 1E). 7 of the 12 genes related to reduced oxygen consumption (58.33%) were noted to be overexpressed in Clusters 2&3 (Fig. 1F). We therefore proved that both these clusters were of different cell hypoxia status. Patients in Cluster 2&3 and Cluster 1 were then renamed to Hypoxia_High or Hypoxia_Low groups, respectively. A total of 773 DEGs were then identified using the DESeq2 package between the Hypoxia_High and Hypoxia_Low groups. Of these genes, 220 were upregulated whereas 553 were downregulated (Fig. 1G and H).

Immune status and immune-related DEGs in OSCC

Based on the proportions of 28 immune cells quantified by ssGSEA, OSCC samples in TCGA database were classified into high, medium and low immune groups (named Immune_ High,Immune_ Medium and Immune_ Low, respectively (Fig. 2A). In accordance to calculations using the ESTIMATE approach both immune and stromal scores of the Immune_ High group were elevated in contrast to those of the Immune_ Low group (Fig. 2B and C). Comparisons between high- and low-immune groups revealed 854 immune-related DEGs (Fig. 2D and E), which were subjected to Metascape tool functional enrichment analysis. The identified DEGs were predominately involved in immune response processes such as “lymphocyte activation”, “adaptive immune response”, and “leukocyte activation involved immune response” and “Immunoregulatory interactions between a Lymphoid and a non-Lymphoid cell” (Fig. 2F).

Development and validation of the risk score

A total of 193 mRNAs which were significant in both hypoxia- and immune-related genes. 11 of these mRNAs were associated with OS ($p < 0.05$) according to univariate Cox regression analysis (Fig. 3A). Of these, a total of 8 mRNA signatures were selected by the LASSO Cox regression model. (Fig. 3B and C). All the 8 mRNAs were protective against OSCC (Fig. 3D). Subsequent risk scores were derived from the following equation: Risk score = [DNAH1 expression* (-0.20877639)] + [FAM122C expression * (-0.02762894)] + [INSL3 expression* (-0.16020112)] + [KIAA1211 expression* (-0.07398287)] + [RANBP17 expression* (-0.05474305)] + [RIPPLY2 expression* (-0.13172196)] + [RNF157 expression * (-0.02933425)] + [SOWAHA expression* (-0.06916555)]. Using this formula, patients were stratified inoto

either low- or high-risk groups. The method of maximally selected rank statistics was used to determine the optimal risk score cutof (Fig. 3E). Significantly improved survival was observed in those in the low-risk group compared to their high-risk counterparts (log-rank test, $p < 0.001$) (Fig. 3F). The ROC curve showed that the classifier had relatively strong predictive power in the TCGA cohort, with a value of 0.578 for the area under the curve (AUC) (Fig. 3G). Furthermore, the TCGA cohort was also analysed for distribution of selected gene expressions, risk scores as well as survival status (Fig. 3H). Table 1 depicts all relevant clinical parameters.

Table 1
Clinical characteristics of the TCGA OSCC patients.

Characteristic	high (n = 161)	low (n = 142)	TCGA (n = 303)
Age			
≤ 62	80	81	161
> 62	81	61	142
Gender			
Female	51	48	99
Male	110	94	204
Survival status			
Living	93	100	193
Dead	68	42	110
Pathologic M			
Unknown	100	88	188
M0	61	54	115
Pathologic N			
Unknown	27	17	44
N0	59	54	113
N1	27	19	46
N2	47	51	98
N3	1	1	2
Pathologic T			
Unknown	11	7	18
T1	12	15	27
T2	50	43	93
T3	31	26	57
T4	57	51	108
Tumor stage			
Unknown	13	10	23
Stage I	9	8	17

Characteristic	high (n = 161)	low (n = 142)	TCGA (n = 303)
Stage II	27	24	51
Stage III	33	21	54
Stage IV	79	79	158

To further explore the efficacy of the constructed risk score, we performed validation tests using the GEO cohort. Using the method of maximally selected rank statistics, all OSCC patients in the GEO cohort were divided (Fig. 4A). Comparison of survival showed that improved survival rates were experienced by low-risk patients in contrast to high-risk patients (log-rank test, p = 0.00084) (Fig. 4B). In addition, we plotted the ROC curves for predicting one - and three-year survival according to the OS of patients, and the values of the area under the curve (AUC) were 0.552 and 0.605, respectively (Fig. 4C). The GEO OSCC cohort were then analyzed for distribution of selected gene expressions, risk scores as well as survival status (Fig. 4D-F).

Relationship between prognostic mRNA signature and clinical parameters

We next investigate the relationship between the risk score and clinical parameters in the TCGA database. Only OSCC samples with complete clinical information (including patient tumor stage and grade (including N and T stages), genders, age and the presence of perineural and lymphovascular invasion) were used. We found that in addition to age (Fig. 5A), perineural and lymphovascular invasion were found to correlate significantly with the hypoxia- immune mRNA signature, but not with patient tumor stage and grade (including N and T stages), genders. (Fig. 5B-H).

Both univariate and multivariate analyses were done on the TCGA dataset to further prove the significance of the constructed hypoxia- and immune-related gene signatures. Risk score, age, and pathological stage may be able to be combined to predict the prognosis of the TCGA OSCC cohort (Table 2).

Table 2
Univariate and multivariate Cox regression analysis of clinicopathological features associated with overall survival in TCGA data.

Variables	Univariate Analysis		Multivariate Analysis	
	HR (95% CI)	*P*-value	HR (95% CI)	*P*-value
Age (>62/≤62)	1.021 (1.005–1.037)	0.00942 *	1.0329 (1.0103–1.056)	0.00413 *
Gender (male/female)	1.021 (0.6894–1.511)	0.919	0.7058 (0.4183–1.191)	0.19194
Tumor grade (G4/G3/G2/G1)	1.35 (1.01–1.803)	0.0424 *	1.4226 (0.9819–2.061)	0.06242
Stage (IV/III/II/I)	1.482 (1.172–1.874)	0.00102 *	1.7513 (1.0060–3.049)	0.04758 *
Pathologic T (T4/T3/T2/T1)	1.447 (1.181–1.773)	0.000365 *	1.0449 (0.7235–1.509)	0.81483
Pathologic N (N3/N2/N1/N0)	1.411 (1.116–1.785)	0.00403 *	1.1601 (0.8620–1.561)	0.32721
Risk score	4.322 (2.096–8.912)	0.0000736 *	3.5246 (1.4818–8.383)	0.00438 *

Variability of degree of immune cell infiltration between the low- and high-risk TCGA OSCC cohorts

The degree of immune cell infiltration between low- and high-groups was evaluated using the CIBERSORT algorithm. Those in the low risk group were more likely to have higher infiltration degrees of CD8 + T cells, plasma cells, follicular helper T cells, regulatory T cells (Tregs) and naive B cells s. On the other hand, high-risk patients demonstrated higher degrees of infiltration of activated DCs, activated mast cells and neutrophils. Furthermore, the immune indicator scores and the immune checkpoint expressions in the low-risk group were all remarkably raised in comparison to those in the high-risk group patients (Fig. 6B-G).

Discussion

Both hypoxia and immune status play important roles in OSCC [23–25]. However, there is a lack of reliable and validated hypoxia- and immune-related gene signatures. Based on hypoxia-related signatures, the samples in our study were grouped into 3 clusters (Cluster 1, Cluster 2, and Cluster 3)

(Fig. 1A-C). The presence of cell hypoxia induces the expression of HIF-1α. Enhanced HIF-1α expression has previously been associated with poor prognosis and lymph nodes metastasis in OSCC patients [26]. Further evaluation of overall patient survival resulted in the combination of Clusters 2 and 3 to form the Hypoxia_High group, while those in Cluster 1 were designated as the Hypoxia_Low group (Fig. 1D). The expression of genes from two hypoxia-related gene sets, “Increased oxygen delivery” and “Reduced oxygen consumption”, were compared between Cluster 1 and Cluster 2 & 3. 11 of the 14 genes belonging to the “Increased oxygen delivery” were expressed at higher levels in Clusters 2 and 3 in contrast to Cluster 1 (Fig. 1E). On the other hand, 7 of the 12 genes associated with “Reduced oxygen consumption” were highly expressed in Cluster 1 compared to Clusters 2 and 3 (Fig. 1F). Therefore, Cluster 1 and Clusters 2 & 3 were considered as Hypoxia_Low and Hypoxia_High groups, respectively. A total of 773 DEGs were identified across both these groups (Fig. 1E and F).

We then divided these OSCC patients into three groups (Immune_High, Immune_Medium and Immune_Low) according to the abundance of immune cells (Fig. 2A). The immune score of Immune_High group was higher than that of the Immune_Medium and Immune_Low groups. The Immune_Low group has the lowest immune score amongst the three groups (Fig. 2B & C). A total of 854 immune-related genes in OSCC were determined between Immune_High and Immune_Low groups (Fig. 2E & F). All genes were found to be enriched in immune-related functions including “leukocyte activation involved immune response”, “adaptive immune response”, “lymphocyte activation” and “Immunoregulatory interactions between a Lymphoid and a non-Lymphoid cell” (Fig. 2F). These immune processes also related to the malignant properties of OSCC. Previous reports found that a higher neutrophil-to-lymphocyte ratio was a negative predictor for overall survival for patients with OSCC [27]. Dendritic cell immune response activation was able to be induced by IFN-γ inhibited OSCC growth in tumour-bearing mice [28].

To further integrate the hypoxia- and immune-related genes, the overlapping genes between the two gene sets were screened. Eight prominent mRNA signatures associated with OS from the 193 overlapping genes were selected, which were FAM122C, RNF157, RANBP17, SOWAHA, KIAA1211, RIPPLY2, INSL3, and DNAH1 (Fig. 3A-D). Among these mRNAs, KIAA1211 is known to be an oncogenic gene. Non-small cell lung cancer tissues were found to have raised KIAA1211 expressions in contrast to adjacent normal tissues. Knockdown of KIAA1211 inhibited the proliferative abilities of NSCLC cells while promoting apoptosis both *in vitro* and *in vivo* [29]. Small cell lung cancer patients with a KIAA1211 mutation possesses longer survival period than those with wild-type KIAA1211 mutations [30]. RIPPLY2 represented one of the mRNAs in a five-gene signature verified to be able to predict survival of endometrial cancer patients [31]. The tumor-promoting effect of INSL3 in cancer has also been widely addressed. The plasma level of INSL3 was found to be raised in an individual with metastatic ovarian cancer [32]. Other reports highlight the potential role of INSL3 as a marker of human testicular Leydig cell tumors [33]. INSL3 could promote tumor growth and angiogenesis in nude mice model of thyroid cancer in a manner which appeared to be related to the action of RXFP2 and the secretion of S100A4 and (pro-) cathepsin-L [34]. In pancreatic cancer patients, higher serum level of INSL3 was associated with

increased anorexia [35]. These mRNAs, which have previously been found to be of significant value in other cancers, should be further investigated for their role in OSCC.

These 8 mRNAs were used to construct a mRNA signature which may have prognostic potential in OSCC (Fig. 3E-H). This constructed risk score was then validated in a cohort from the GEO dataset (Fig. 4A-F). The risk score showed a significant association only with age, perineural and lymphovascular invasion, but not with other features including gender, T stage, N stage, and tumour stage (Fig. 5A-H). Therefore, this established gene signature may be an independent prognostic indicator.

Finally, the immune profile variability between the high- and low-risk groups were compared in the TCGA OSCC cohort. High-risk group samples showed different degree of immune cell infiltration (Fig. 6A). The low-risk group showed higher immune status than low-risk group, as evidenced by a higher amount of immune checkpoint expression including CTLA-4, PD-1, APM, CYT, and TILs (Fig. 6B-G). These checkpoints are indictors of OSCC risk and may function as therapeutic targets in OSCC. The genetic variants of CTLA-4 were associated with tobacco-related OSCC risk in the North Indian population [36]. A high number of CTLA-4 cells is associated with poor 5-year metastasis-free survival of OSCC patients [37]. Anti-PD-1 antibody is an agent which may prevent the initiation and progression of OSCC while prolonging patient survival time [38–40]. Therefore, patients in low-risk group may be more sensitive to immunotherapy.

In conclusion, hypoxia statues and immune statues both play pivotal roles in the prognostic of OSCC. Combined with hypoxia- and immune-related genes, we established a hypoxia-immune-based gene signature for prognosis of OSCC, which may serve as a reliable reference for clinical decision making.

Declarations

Ethical approval and consent to participate

Not applicable.

Consent for publication

All authors have reviewed and approved this manuscript

Availability of data and materials

TCGA Data Poral: https://portal.gdc.cancer.gov/

GEO Datasets: https://www.ncbi.nlm.nih.gov/gds/

Metascape: https://metascape.org/

Competing interests

The authors declare no conflicts of interest.

Funding

Not applicable.

Author Contributions

Jichen Li and Shaohua Lv designed the study. Shaohua Lv, Jianhao Li and Songlin Piao collected data. Shaohua Lv, Jianhao Li and Songlin Piao developed the computational model and analyzed. Shaohua Lv and Jianhao Li wrote the article. All authors reviewed the manuscript.

Acknowledgements

Not applicable.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68(6):394–424.
2. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. CA Cancer J Clin. 2020;70(1):7–30.
3. Eckert AW, Lautner MH, Schutze A, Taubert H, Schubert J, Bilkenroth U. Coexpression of hypoxia-inducible factor-1alpha and glucose transporter-1 is associated with poor prognosis in oral squamous cell carcinoma patients. Histopathology. 2011;58(7):1136–47.
4. Eckert AW, Kappler M, Schubert J, Taubert H. Correlation of expression of hypoxia-related proteins with prognosis in oral squamous cell carcinoma patients. Oral Maxillofac Surg. 2012;16(2):189–96.
5. Hubbi ME, Semenza GL. Regulation of cell proliferation by hypoxia-inducible factors. Am J Physiol Cell Physiol. 2015;309(12):C775–82.
6. Domingos PLB, Souza MG, Guimaraes TA, Santos ES, Farias LC, de Carvalho Fraga CA, Jones KM, Santos SHS, de Paula AMB, Guimaraes ALS. Hypoxia reduces the E-cadherin expression and increases OSCC cell migration regardless of the E-cadherin methylation profile. Pathol Res Pract. 2017;213(5):496–501.
7. Zhu G, Wang S, Chen J, Wang Z, Liang X, Wang X, Jiang J, Lang J, Li L. Long noncoding RNA HAS2-AS1 mediates hypoxia-induced invasiveness of oral squamous cell carcinoma. Mol Carcinog. 2017;56(10):2210–22.
8. Ishikawa T, Nakashiro K, Klosek SK, Goda H, Hara S, Uchida D, Hamakawa H. Hypoxia enhances CXCR4 expression by activating HIF-1 in oral squamous cell carcinoma. Oncol Rep. 2009;21(3):707–12.
9. Lv X, Li J, Zhang C, Hu T, Li S, He S, Yan H, Tan Y, Lei M, Wen M, et al. The role of hypoxia-inducible factors in tumor angiogenesis and cell metabolism. Genes Dis. 2017;4(1):19–24.

10. Jian SL, Hsieh HY, Liao CT, Yen TC, Nien SW, Cheng AJ, Juang JL. Galpha(1)(2) drives invasion of oral squamous cell carcinoma through up-regulation of proinflammatory cytokines. PloS one. 2013;8(6):e66133.
11. Cao MX, Zhang WL, Yu XH, Wu JS, Qiao XW, Huang MC, Wang K, Wu JB, Tang YJ, Jiang J, et al. Interplay between cancer cells and M2 macrophages is necessary for miR-550a-3-5p down-regulation-mediated HPV-positive OSCC progression. J Exp Clin Cancer Res. 2020;39(1):102.
12. Diao P, Wu Y, Li J, Zhang W, Huang R, Zhou C, Wang Y, Cheng J. Preoperative systemic immune-inflammation index predicts prognosis of patients with oral squamous cell carcinoma after curative resection. J Transl Med. 2018;16(1):365.
13. Malaspina TS, Gasparoto TH, Costa MR, de Melo EF Jr, Ikoma MR, Damante JH, Cavassani KA, Garlet GP, da Silva JS, Campanelli AP. Enhanced programmed death 1 (PD-1) and PD-1 ligand (PD-L1) expression in patients with actinic cheilitis and oral squamous cell carcinoma. Cancer Immunol Immunother. 2011;60(7):965–74.
14. Dave K, Ali A, Magalhaes M. Increased expression of PD-1 and PD-L1 in oral lesions progressing to oral squamous cell carcinoma: a pilot study. Sci Rep. 2020;10(1):9705.
15. Kondoh N, Mizuno-Kamiya M, Umemura N, Takayama E, Kawaki H, Mitsudo K, Muramatsu Y, Sumitomo S. Immunomodulatory aspects in the progression and treatment of oral malignancy. Jpn Dent Sci Rev. 2019;55(1):113–20.
16. Meehan K, Leslie C, Lucas M, Jacques A, Mirzai B, Lim J, Bulsara M, Khan Y, Wong NC, Solomon B, et al. Characterization of the immune profile of oral tongue squamous cell carcinomas with advancing disease. Cancer Med. 2020;9(13):4791–807.
17. Sievilainen M, Almahmoudi R, Al-Samadi A, Salo T, Pirinen M, Almangush A. The prognostic value of immune checkpoints in oral squamous cell carcinoma. Oral Dis. 2019;25(6):1435–45.
18. Huang GZ, Wu QQ, Zheng ZN, Shao TR, Li F, Lu XY, Ye HY, Chen GX, Song YX, Zeng WS, et al. Bioinformatics Analyses Indicate That Cathepsin G (CTSG) is a Potential Immune-Related Biomarker in Oral Squamous Cell Carcinoma (OSCC). Onco Targets Ther. 2021;14:1275–89.
19. Li L, Cao B, Liang X, Lu S, Luo H, Wang Z, Wang S, Jiang J, Lang J, Zhu G. Microenvironmental oxygen pressure orchestrates an anti- and pro-tumoral gammadelta T cell equilibrium via tumor-derived exosomes. Oncogene. 2019;38(15):2830–43.
20. Zheng S, Zou Y, Liang JY, Xiao W, Yang A, Meng T, Lu S, Luo Z, Xie X. Identification and validation of a combined hypoxia and immune index for triple-negative breast cancer. Mol Oncol. 2020;14(11):2814–33.
21. Liu Y, Wu J, Huang W, Weng S, Wang B, Chen Y, Wang H. Development and validation of a hypoxia-immune-based microenvironment gene signature for risk stratification in gastric cancer. J Transl Med. 2020;18(1):201.
22. Bhandari V, Hoey C, Liu LY, Lalonde E, Ray J, Livingstone J, Lesurf R, Shiah YJ, Vujcic T, Huang X, et al. Molecular landmarks of tumor hypoxia across cancer types. Nat Genet. 2019;51(2):308–18.

23. Gao L, Dou ZC, Ren WH, Li SM, Liang X, Zhi KQ. CircCDR1as upregulates autophagy under hypoxia to promote tumor cell survival via AKT/ERK(1/2)/mTOR signaling pathways in oral squamous cell carcinomas. Cell Death Dis. 2019;10(10):745.
24. Dong W, Chen Y, Qian N, Sima G, Zhang J, Guo Z, Wang C. SATB2 knockdown decreases hypoxia-induced autophagy and stemness in oral squamous cell carcinoma. Oncol Lett. 2020;20(1):794–802.
25. Quan H, Shan Z, Liu Z, Liu S, Yang L, Fang X, Li K, Wang B, Deng Z, Hu Y, et al. The repertoire of tumor-infiltrating lymphocytes within the microenvironment of oral squamous cell carcinoma reveals immune dysfunction. Cancer Immunol Immunother. 2020;69(3):465–76.
26. Uehara M, Sano K, Ikeda H, Nonaka M, Asahina I. Hypoxia-inducible factor 1 alpha in oral squamous cell carcinoma and its relation to prognosis. Oral Oncol. 2009;45(3):241–6.
27. Wang Y, Wang P, Andrukhov O, Wang T, Song S, Yan C, Zhang F. Meta-analysis of the prognostic value of the neutrophil-to-lymphocyte ratio in oral squamous cell carcinoma. J Oral Pathol Med. 2018;47(4):353–8.
28. Wei H, Hongya P, Linlin J, Mujiang A, Kuijie W, Duohong Z, Qingang H, Zhiyuan Z. IFN-gamma enhances the anti-tumour immune response of dendritic cells against oral squamous cell carcinoma. Arch Oral Biol. 2011;56(9):891–8.
29. Liu Z, Cao H, Shi Y, Yang R. KIAA1211 plays an oncogenic role in human non-small cell lung cancer. J Cancer. 2019;10(26):6747–53.
30. Wang L, Luan T, Zhou S, Lin J, Yang Y, Liu W, Tong X, Jiang W: **LncRNA HCP5 promotes triple negative breast cancer progression as a ceRNA to regulate BIRC3 by sponging miR-219a-5p**. *Cancer medicine* 2019.
31. Li X, Yin F, Fan Y, Cheng Y, Dong Y, Zhou J, Wang Z, Li X, Wang J. Establishment and validation of a prognostic nomogram based on a novel five-DNA methylation signature for survival in endometrial cancer patients. Cancer Med. 2021;10(2):693–708.
32. Rossato M, Barban M, Vettor R. Elevated plasma levels of the novel hormone INSL3 in a woman with metastatic ovarian cancer. Int J Biol Markers. 2007;22(2):159–60.
33. Rossato M, Tavolini IM, Calcagno A, Gardiman M, Dal Moro F, Artibani W. The novel hormone INSL3 is expressed in human testicular Leydig cell tumors: a clinical and immunohistochemical study. Urol Oncol. 2011;29(1):33–7.
34. Hombach-Klonisch S, Bialek J, Radestock Y, Truong A, Agoulnik AI, Fiebig B, Willing C, Weber E, Hoang-Vu C, Klonisch T. INSL3 has tumor-promoting activity in thyroid cancer. Int J Cancer. 2010;127(3):521–31.
35. Yeom E, Shin H, Yoo W, Jun E, Kim S, Hong SH, Kwon DW, Ryu TH, Suh JM, Kim SC, et al. Tumour-derived Dilp8/INSL3 induces cancer anorexia by regulating feeding neuropeptides via Lgr3/8 in the brain. Nat Cell Biol. 2021;23(2):172–83.
36. Bharti V, Mohanti BK, Das SN. Functional genetic variants of CTLA-4 and risk of tobacco-related oral carcinoma in high-risk North Indian population. Hum Immunol. 2013;74(3):348–52.

37. Koike K, Dehari H, Ogi K, Shimizu S, Nishiyama K, Sonoda T, Sasaki T, Sasaya T, Tsuchihashi K, Hasegawa T, et al. Prognostic value of FoxP3 and CTLA-4 expression in patients with oral squamous cell carcinoma. PloS one. 2020;15(8):e0237465.
38. Ferris RL, Blumenschein G Jr, Fayette J, Guigay J, Colevas AD, Licitra L, Harrington K, Kasper S, Vokes EE, Even C, et al. Nivolumab for Recurrent Squamous-Cell Carcinoma of the Head and Neck. N Engl J Med. 2016;375(19):1856–67.
39. Wang J, Xie T, Wang B, William WN Jr, Heymach JV, El-Naggar AK, Myers JN, Caulin C. PD-1 Blockade Prevents the Development and Progression of Carcinogen-Induced Oral Premalignant Lesions. Cancer Prev Res (Phila). 2017;10(12):684–93.
40. Monteiro de Oliveira Novaes JA, Hirz T, Guijarro I, Nilsson M, Pisegna MA, Poteete A, Barsoumian HB, Fradette JJ, Chen LN, Gibbons DL, et al: **Targeting of CD40 and PD-L1 Pathways Inhibits Progression of Oral Premalignant Lesions in a Carcinogen-induced Model of Oral Squamous Cell Carcinoma**. *Cancer Prev Res (Phila)* 2020.

Figures

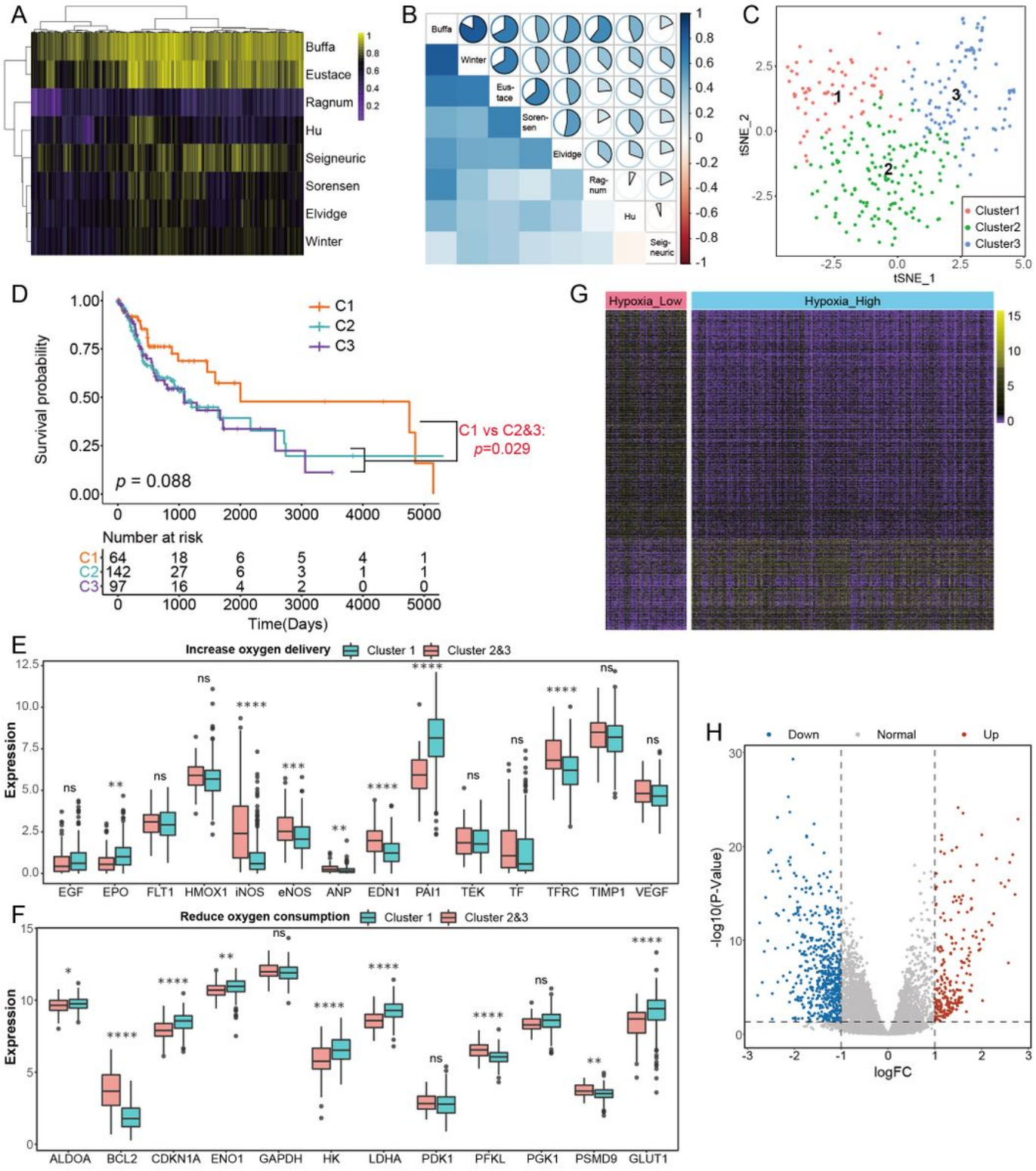

Figure 1

Identification of hypoxia status and hypoxia-associated DEGs. Quantified heatmap (A) and correlation heatmap (B) calculated by ssGSEA with respect to the eight hypoxia signature sets. (C) Dot plot of three distinct clusters determined using the t-SNE algorithm to analyze 7 hypoxia signature gene sets. (D) Overall survival as shown by a Kaplan–Meier plot for patients in three clusters. (E, F) HIF-1 KEGG pathway gene expression changes based on hypoxia status (Hypoxia_High vs. Hypoxia_Low) . Heat map

(G) and volcano plot (H) show the differentially expressed hypoxia-related genes in oral squamous cell carcinoma (OSCC). Red dots represent upregulated DEGs, blue dots represent downregulated DEGs and gray dots represent genes with no differential expression.

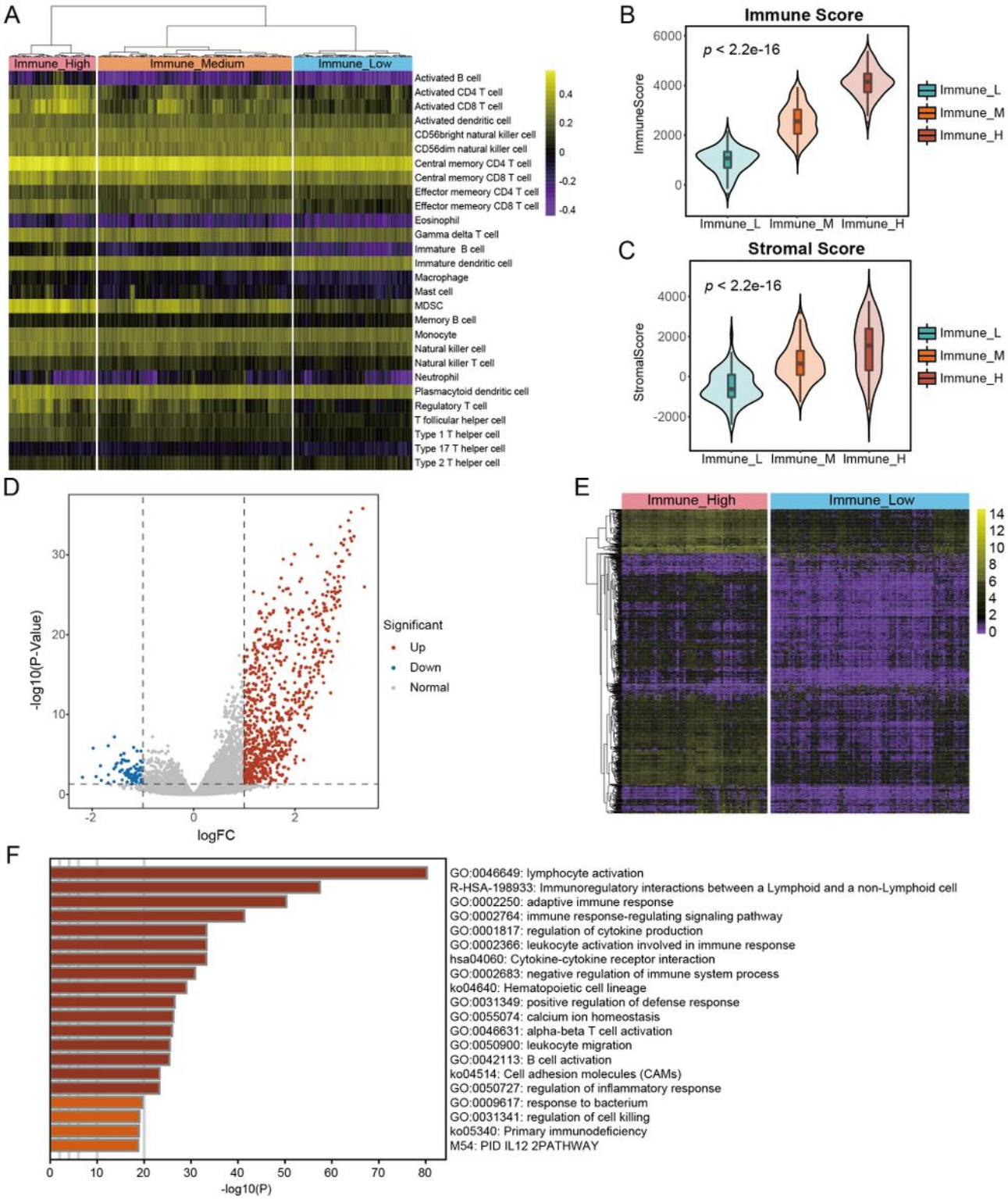

Figure 2

Identification of immune status and immune-associated DEGs. (A) Data from the TCGA database was classified into three subtypes (Immune_High group, Immune_Medium group, and Immune_Low group)

based on immune status. Comparison of Immune_High and Immune_Low groups in terms of (B) immune score and (C) stromal score. (D) Volcano plot and (E) heatmap depicting DEGs between Immune_High and Immune_Low groups. Red dots represent upregulated DEGs, blue dots represent downregulated DEGs and gray dots represent genes with no differential expression. Bar plot (F) of functional enrichment analysis of immune-related DEGs.

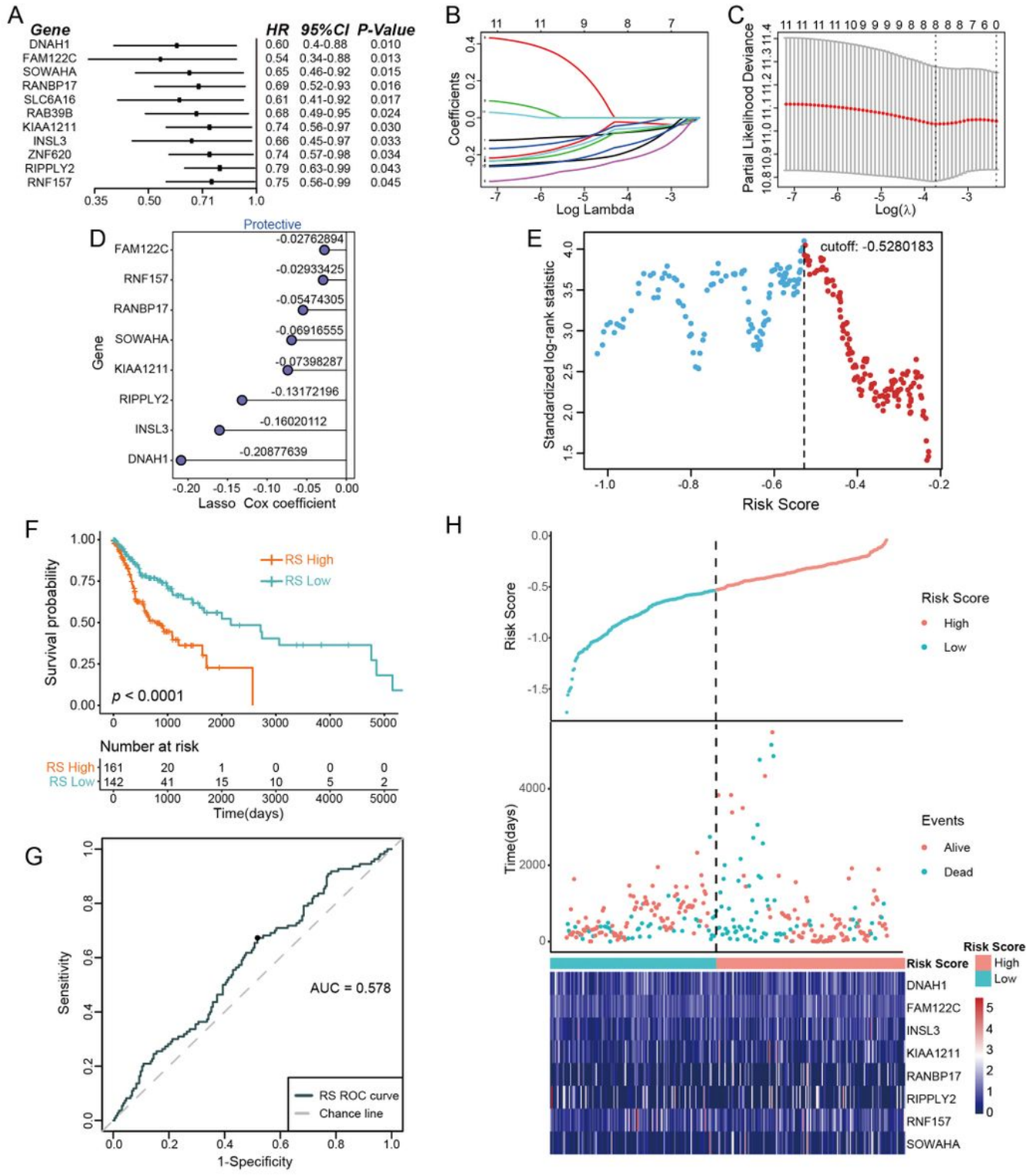

Figure 3

The construction of hypoxia-immune-based mRNA prognosis model. (A) Hazard ratios for 11 hypoxia-immune-associated prognostic mRNAs. (B) LASSO coefficient profiles. (C) The LASSO model is used to choose the tuning parameter (lambda) using a fivefold cross-validation based on minimum criteria for OS; average OS genes are represented by the upper X axis and log (lambda) is represented by the lower X axis. Partial likelihood deviance error is represented by the Y axis. (D) An ensemble of 8 genes with their respective individual coefficients. (E) Scatter plot illustrates standardized log-rank statistic value for each corresponding cutoff of hypoxia-immune-based risk score. The optimal cutoff with the maximum standard log-rank statistic is marked with a vertical dashed line. (F) Kaplan–Meier plot of overall survival for patients in low-risk and high-risk groups by hypoxia-immune-based prognosis classifier in discovery cohort. (G) The effectiveness of the risk scores are predicted using the ROC curve. (H) Expression profiles, survival status and risk score distributions of the gene signature.

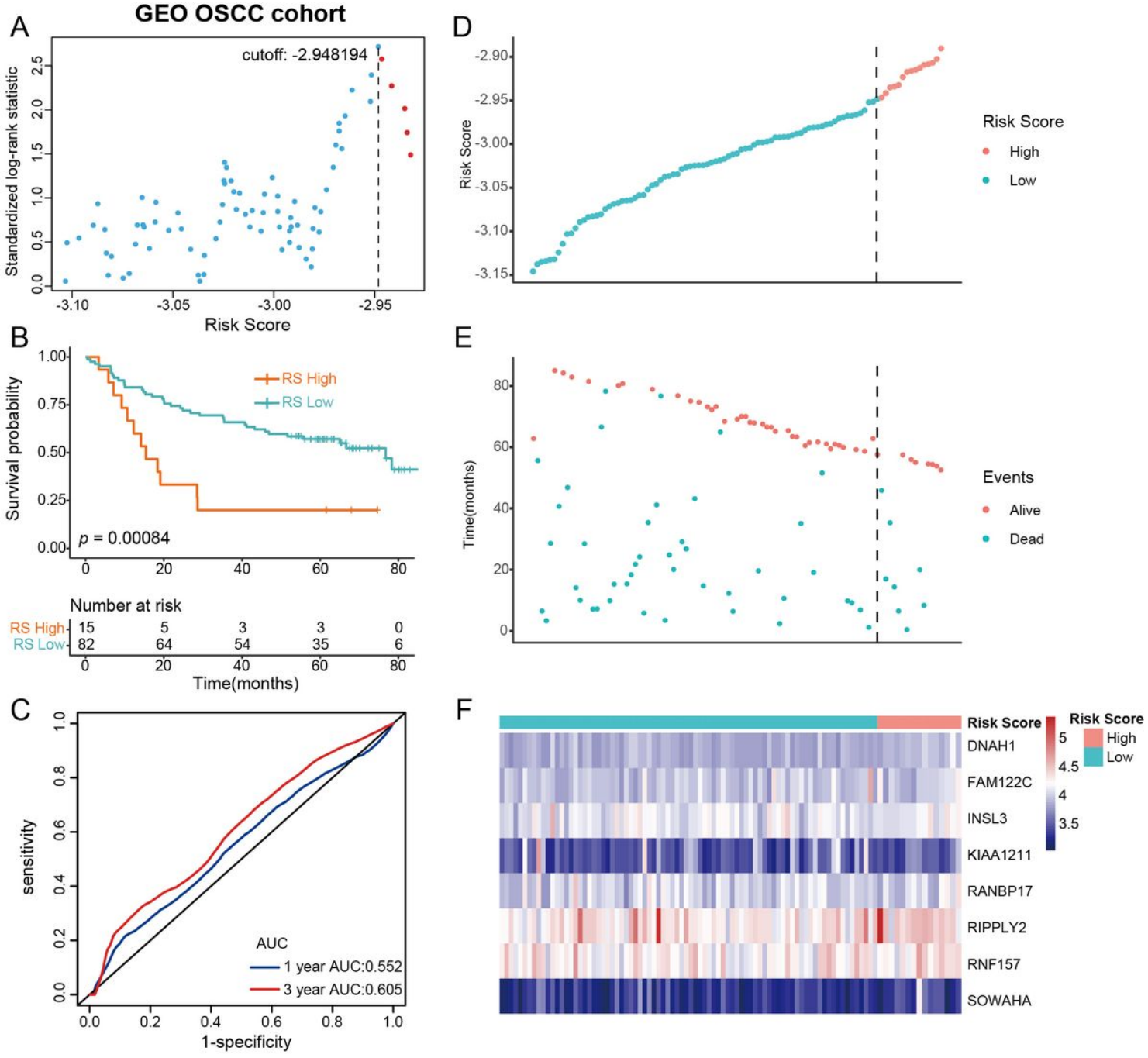

Figure 4

The GEO cohort was used to verify the hypoxia-immune-based prognosis classifier. (A) Scatter plot demonstrates the standardized log-rank statistic value for each corresponding cutoff of hypoxia-immune-based risk score. Vertical dashed lines represent the optimal cutoff with the maximum standard log-rank statistic. (B) The hypoxia-immune-based prognosis classifier is used to label patients from the GEO cohort as either low- or high-risk followed by construction of an overall survival Kaplan–Meier plot (C) 1-year and 3-year survival rate of OSCC patients in the GEO cohort as shown using ROC curves. (D) Distributions of risk score, (E) survival status and (F) expression profile of signature genes.

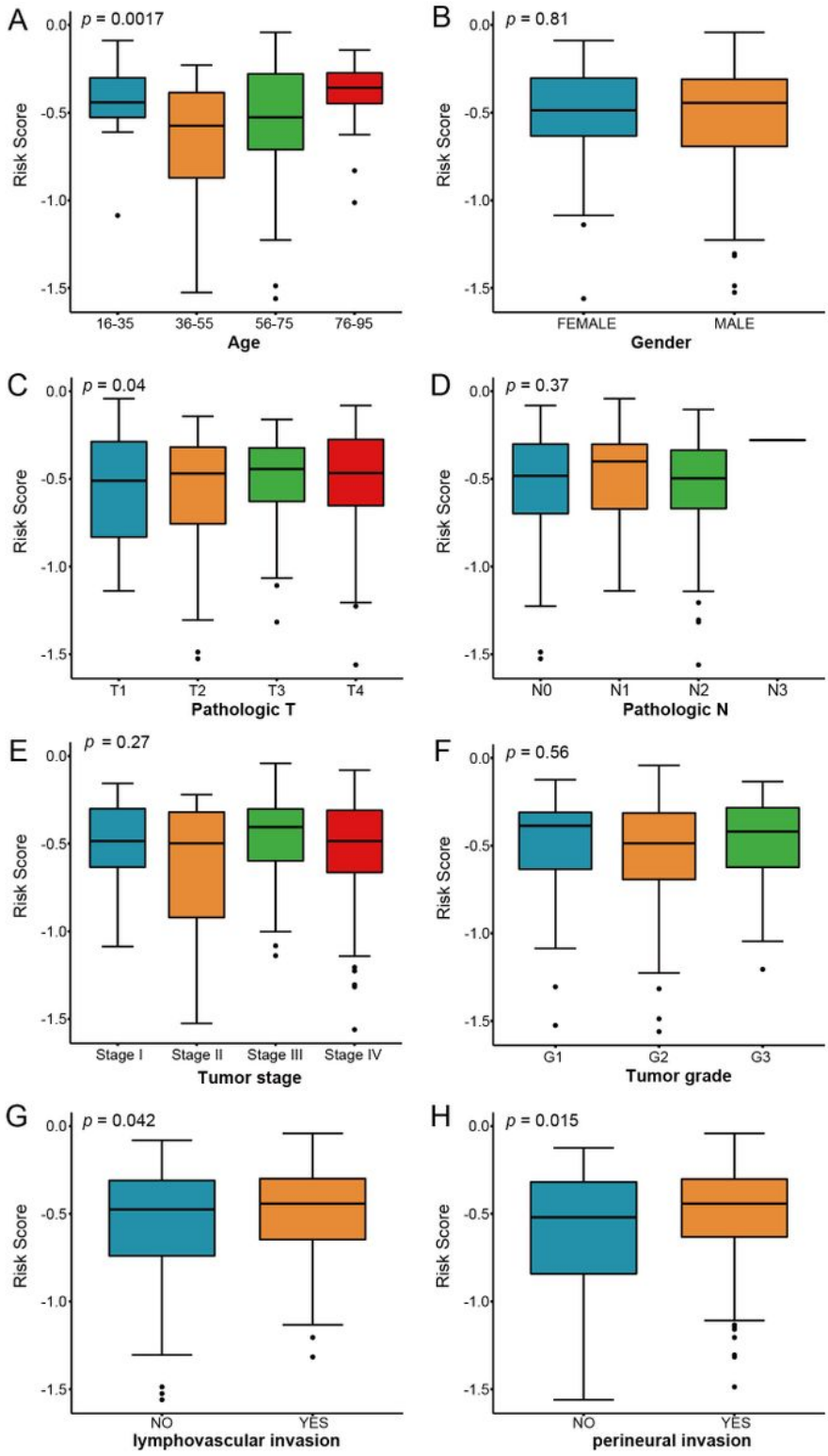

Figure 5

The TCGA cohort was used to determine the association between clinical parameters and the gene signature. Significant association was only found for (A) age, but not with (B) gender, (C) T stage, (D) N stage, (E) tumour stage, (F) tumour grade, (G) lymphovascular invasion and (H) perineural invasion.

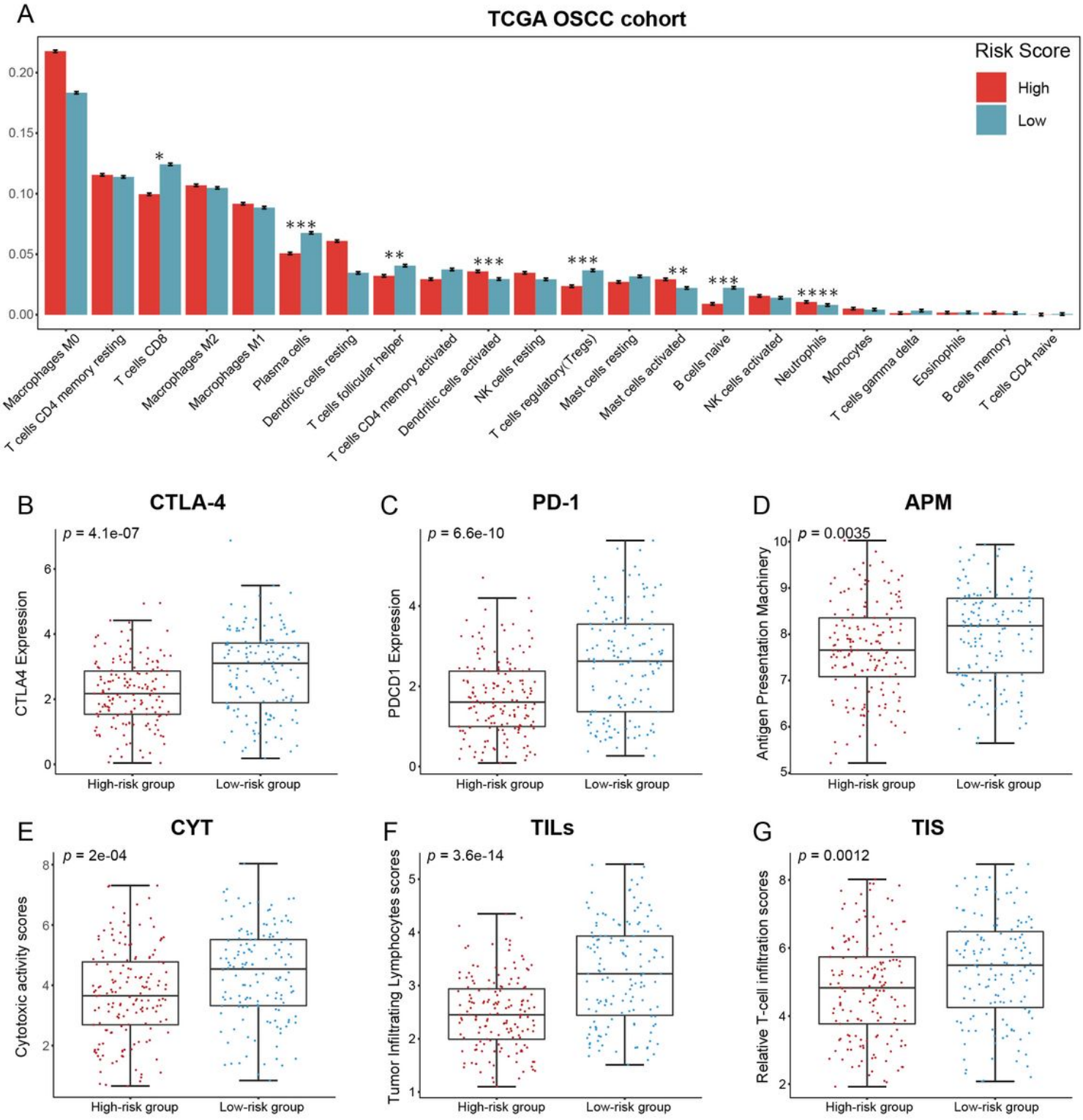

Figure 6

Analysis of immune status variability of high- and low-risk groups in the TCGA OSCC cohort. (A) The bar plot shows various immune cell infiltration between high- and low-risk patients. (B-G) Box plots illustrating markedly different immune checkpoints between high- and low-risk groups.