

RESEARCH

Open Access



# Evaluate the immune-related eRNA models and signature score to predict the response to immunotherapy in thyroid carcinoma

Pu Wu, Jinyuan Shi, Zhiyuan Wang, Wei Sun and Hao Zhang\*

## Abstract

**Background:** The functional alterations of eRNAs have been reported to be correlated with tumorigenesis. However, the roles of eRNAs in thyroid cancer (THCA) remain still unclear. This study aimed to construct an immune-related eRNA prognostic signature that could effectively predict the survival and prognosis for THCA.

**Methods:** The Weighted Gene Co-Expression Network Analysis (WGCNA) was performed to identify THCA-specific immune-related hub genes and immune-related eRNAs were obtained using Pearson correlation analysis. Univariate and least absolute shrinkage and selection operator (LASSO) Cox regression were conducted to construct an immune-related eRNA prognostic signature in training cohort, and the predictive capability was verified in test cohort and entire cohort. Kaplan–Meier analysis, principal component analysis (PCA), receiver operating characteristic (ROC) curves, and nomogram were used to validate the risk signature. Furthermore, CIBERSORT, ESTIMATE and ssGSEA were analyzed to explore the tumor immune microenvironment (TIME) of the risk signature, and the response of potential immunotherapeutic were also discussed.

**Results:** A total of 125 immune-related eRNAs were obtained and 16 immune-related eRNAs were significantly correlated with overall survival (OS). A 9-immune-related eRNA prognostic signature was constructed, and the risk score was identified as an independent predictor. High-risk groups were associated with a poorer OS. Immune microenvironment analysis indicated that low risk score was correlated with higher immuneScore, high immune cell infiltration, and the better response of immunotherapy. Additionally, we also detected 9 immune-related eRNA expression levels in sixty-two matched tumorous and non-tumorous tissues using qRT-PCR analysis.

**Conclusion:** Our immune-related eRNA risk signature that was an independent prognostic factor was strongly correlated with the immune microenvironment and may be promising for the clinical prediction of prognosis and immunotherapeutic responses in THCA patients.

**Keywords:** eRNAs, Tumor immune microenvironment, THCA, Prognostic signature, Immune checkpoints

## Introduction

The incidence of thyroid cancer (THCA) is increasing steadily in the past few decades [1]. THCA is the most common malignancy of the endocrine system, it can be

classified into four main histological types: well-differentiated papillary thyroid carcinoma (PTC), follicular thyroid cancer (FTC), medullary thyroid cancer (MTC) and undifferentiated or anaplastic thyroid cancer (ATC) [2]. PTC is the most predominant subtype and accounts for more than 85% among all the THCA cases. Although the majority of PTC patients usually have an excellent overall prognosis after surgery and radioactive iodine therapy, regional neck lymph node metastasis (LNM) in advanced

\*Correspondence: haozhang@cmu.edu.cn

Department of Thyroid Surgery, The First Hospital of China Medical University, Shenyang, China



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

PTC is associated with local recurrence and distant metastasis, and this has been reported as an independent factor of a poor prognosis [3, 4]. Therefore, it is urgent to explore and identify sensitive prognostic biomarkers for THCA to facilitate rational individualized treatment.

The tumor microenvironment (TME) consists of the stromal and immune cells. The constant interactions between tumor cells and the TME play crucial roles in tumor initiation, progression, metastasis, as well as response to therapies [5]. The TME has attracted considerable attention and interest for their potential to consider to be a therapeutic target in various cancers for anti-tumorigenesis treatment [6]. The immune system plays critical roles in tumor formation and progression. A growing body of evidence has indicated that immune-related genes (IRGs) have increasingly been recognized as biomarkers to predict cancer patient prognosis [7]. Presently, some IRG signatures had been constructed to serve as efficiently predictive and prognostic models among hepatocellular carcinoma (HCC) [8, 9], cervical cancer [10], head and neck squamous cell carcinoma (HNSCC) [11, 12], gastric cancer (GC) [13], neuroblastoma (NBL) [14], prostatic adenocarcinoma (PRAD) patients [15].

Long non-coding RNAs (lncRNAs) have gained substantial attention due to their multi-faceted ability to regulate gene expression [16]. Recent studies showed that enhancers were found to be transcriptionally active and generate and transcribe noncoding RNAs, which were known as enhancer RNAs (eRNAs). eRNAs played a fundamental role in tumor initiation, development, and treatment. eRNAs involved in various cancer signaling pathways through regulating their target genes. Oncogene-induce eRNAs can directly contribute to tumorigenesis. In contrast, tumor suppressors can also induce eRNAs to suppress tumors. eRNAs have emerged as important regulators of the immune response, and played a central role in controlling immune-related functions [17]. Furthermore, eRNAs could also regulate clinically actionable genes and immune checkpoints, which indicated the potentially clinical utility of eRNAs in cancer therapy. Unfortunately, there are limited studies focusing on immune-related eRNAs and the potential prognostic value of immune-related eRNAs in THCA remains unclear; therefore, effective prognostic biomarkers for THCA are urgently needed.

In this study, we aimed to identify immune-related eRNAs using Pearson correlation analysis and construct immune-related eRNAs prognostic signature to systematically explore the prognostic value of the risk signature in TCGA patients from TCGA database. We then further investigated the associations between prognostic signature and clinicopathological factors, tumor immune microenvironment (TIME), immunotherapy responses.

Furthermore, a nomogram was established to predict the OS of THCA patients. Our study may provide a new insight to predict the potential response for immunotherapy for THCA patients.

## Materials and methods

### Data extraction

The RNA sequencing transcriptome data (58 normal samples and 510 THCA samples) and related clinical information of THCA were downloaded from The Cancer Genome Atlas (TCGA) database and Genotype-Tissue Expression Project (GTEx) database. The clinical characteristics of all THCA patients were listed in Table 1. The immune-related genes (IRGs) were downloaded from the Immunology Database and Analysis Portal (ImmPort, <https://www.immport.org/>) and InnateDB (<https://www.innatedb.com/>) databases.

### Differentially expressed immune-related genes (DE-IRGs)

The differentially expressed genes (DEGs) were identified using “limma” package in the R software (version 4.0.4) between tumor samples and normal specimens and visualized by the heatmap and volcano plot according to the screening criterion of  $|\log_2 \text{fold change (FC)}| > 1.0$  and

**Table 1** Patients’ clinical characteristics of training, test and entire cohorts

Variables	Total cohort %	Training cohort %	Test cohort %	p-value
Age				
< =60	389 (77.49)	190 (75.7)	199 (79.28)	0.3926
> 60	113 (22.51)	61 (24.3)	52 (20.72)	
Gender				
Female	367 (73.11)	188 (74.9)	179 (71.31)	0.4207
Male	135 (26.89)	63 (25.1)	72 (28.69)	
Stage				
Stage I-II	333 (66.33)	158 (62.95)	175 (69.72)	0.1003
Stage III-IV	167 (33.27)	93 (37.05)	74 (29.48)	
Unknow	2 (0.4)	0 (0)	2 (0.8)	
T stage				
T1-2	307 (61.16)	154 (61.35)	153 (60.96)	1
T3-4	193 (38.45)	96 (38.25)	97 (38.65)	
Unknow	2 (0.4)	1 (0.4)	1 (0.4)	
M stage				
M0	282 (56.18)	139 (55.38)	143 (56.97)	1
M1	9 (1.79)	4 (1.59)	5 (1.99)	
Unknow	211 (42.03)	108 (43.03)	103 (41.04)	
N stage				
N0	229 (45.62)	111 (44.22)	118 (47.01)	0.7806
N1	223 (44.42)	112 (44.62)	111 (44.22)	
Unknow	50 (9.96)	28 (11.16)	22 (8.76)	

false discovery rate (FDR) value < 0.05 in the TCGA dataset18. Differentially expressed immune-related genes (DE-IRGs) were extracted from DEGs. The visual Venn diagram was constructed by the online tool to show the intersection of the DEGs and IRGs.

#### Weighted gene co-expression network analysis (WGCNA)

WGCNA was performed to identify THCA-specific immune genes related to the co-expression modules using “WGCNA” R package based on the DE-IRGs [19]. The  $A_{ij} = |S_{ij}|^\beta$  ( $A_{ij}$ : adjacency matrix between gene  $i$  and gene  $j$ ,  $S_{ij}$ : similarity matrix made by Pearson’s correlation coefficient of all pairs of genes, and  $\beta$ : soft thresholding value) was used to show the weighted adjacency matrix with a scale-free co-expression network and then transformed into a topological overlap matrix (TOM), and gene modules were identified [20]. We calculated the module eigengene (ME) of each module to identify the most significant module. Finally, the module that was highly correlated with THCA was selected for subsequently analysis.

#### Identification of immune-related eRNAs in THCA

The eRNAs were predicted according to the PreSTIGE algorithm as previously described [21, 22]. The expression of 1565 eRNAs among 1584 eRNAs were retrieved from the TCGA-THCA samples. The relationship between the eRNAs and IRGs was explored to identify immune-related eRNAs using correlation analysis ( $|R| > 0.4$ ,  $P < 0.001$ ).

#### Construction and validation of the immune-related eRNAs prognostic signature

All patients were randomly divided into a training cohort ( $n=251$ ) and a validation cohort ( $n=251$ ) at 1:1 ratio. The training cohort was used for constructing the immune-related eRNAs prognostic signature, and the predictive performance was verified in the test cohort and total cohort. The Univariate Cox proportional hazard regression analysis was conducted to identify the immune-related eRNAs with prognostic value in the training cohort, and the filter  $p$ -value was set at 0.05. The prognostic model was subsequently established using the Least Absolute Shrinkage and Selection Operator (LASSO) penalized Cox proportional hazards regression with “glmnet” R package [23]. The model was determined by penalty parameter ( $\lambda$ ) with ten-fold cross-validation following the minimum criteria. The risk score of each THCA patient was calculated by the following formula: Risk score =  $\beta_A$ \* Expression level of Gene A +  $\beta_B$ \* Expression level of Gene B + ... +  $\beta_N$ \* Expression level of Gene N, ( $\beta$ : regression coefficient) [24]. Patients were separated into high and low-risk group based on

the median risk score. The principal component analysis (PCA) was performed for two risk groups using “scatterplot3d” R package [25]. Kaplan–Meier survival curves were constructed between two risk groups using “survival” R package and the time-dependent receiver operational feature curves (ROC) were carried out with the “survival ROC” R package to verify the sensitivity and specificity of the signature in all cohorts [26]. Univariate and multivariate cox regression analyses were used to determine the independent prognostic factor for OS in all cohorts.

#### Development of prognostic nomogram

The nomogram was constructed based on the risk score and clinical features (age, gender, and pathological stage) to predict the survival risk of THCA patients using “rms” R package. The calibration curves were used for comparing the consistency between the predicted and actual survival to evaluate the predictive probability of the nomogram.

#### Functional enrichment analysis

The DEGs between the high- and low-risk groups were identified according to the filtering criteria ( $|\log_2FC| \geq 1$  and  $FDR < 0.05$ ). The Gene Ontology (GO) analysis was conducted to explore the potential biological processes based on these DEGs using the “clusterProfiler” R package. Gene Set Enrichment Analysis (GSEA) was performed to analyze the difference of the immune response between different risk groups using GSEA 4.1.0. A NOM  $p$ -value < 0.05 was defined as statistically significant.

#### Assessment of immune cell infiltration in THCA

The infiltration levels of immune cells in THCA were estimated based on the gene expression profiles using CIBERSORT algorithm [27]. The differential infiltrating levels of immune cells in high- and low-risk groups were evaluated by the Wilcoxon rank-sum test. Furthermore, the tumor microenvironment score (tumor purity, immune score, stromal score and estimate score) was calculated using the “ESTIMATE” R package [28].

To determine the association between the risk score and immune status, single-sample gene set enrichment analysis (ssGSEA) was performed to calculate the infiltration abundance of 16 immune cells and the activity of 13 immune-related pathways by utilizing the “GSVA” R package [29].

#### Immunophenoscore analysis

Immunophenoscore (IPS) was estimated based on the expression of the four determining components of immunogenicity, including effector cells, immunosuppressive cells, major histocompatibility complex (MHC)

molecules, and immunomodulators, which can well predict the response to immune checkpoint inhibitors (ICIs). IPS was calculated with a scale ranging from 0 to 10 according to representative cell-type gene expression Z-scores. The IPS of THCA patients were downloaded from The Cancer Immunome Atlas (TCIA) (<https://tcia.at/home>) [30].

### Quantitative real-time PCR

Sixty-two pairs of PTC tumorous and adjacent normal tissue specimens were collected from the First Affiliated Hospital of China Medical University. The clinicopathological characteristics of 62 THCA patients from our hospital were displayed in Table 2. Total RNA was extracted from tissue samples using RNAiso (Takara, Dalian, China), then RNA was reverse transcribed into cDNA with the QuantiTect Reverse Transcription Kit (Takara, Shiga, Japan). Quantitative Real-Time PCR (qRT-PCR) analyses were performed with SYBR-Green (Takara, Shiga, Japan) to validate mRNA expression level, and the level of GAPDH served as an internal control. The relative expression level was calculated based on the comparative Ct ( $2^{-\Delta\Delta C_t}$ ) method. The primers' sequences are listed in Additional file 3: Table S1.

**Table 2** The clinicopathological factors in 62 PTC patients

Characteristics	Samples(N = 62)	Percentage (%)
Age		
≤ 60	54	87.1
> 60	8	12.9
Gender		
Female	45	72.6
Male	17	27.4
Tumor size		
< 2 cm	43	69.4
≥ 2 cm	19	30.6
Extrathyroidal invasion		
Yes	7	11.3
No	55	88.7
Multicentricity		
Yes	23	37.1
No	39	62.9
Stage		
Stage I-II	56	90.3
Stage III-IV	6	9.7
T		
T1-2	41	66.1
T3-4	21	33.9
N		
N0	18	29.0
N1	44	71.0

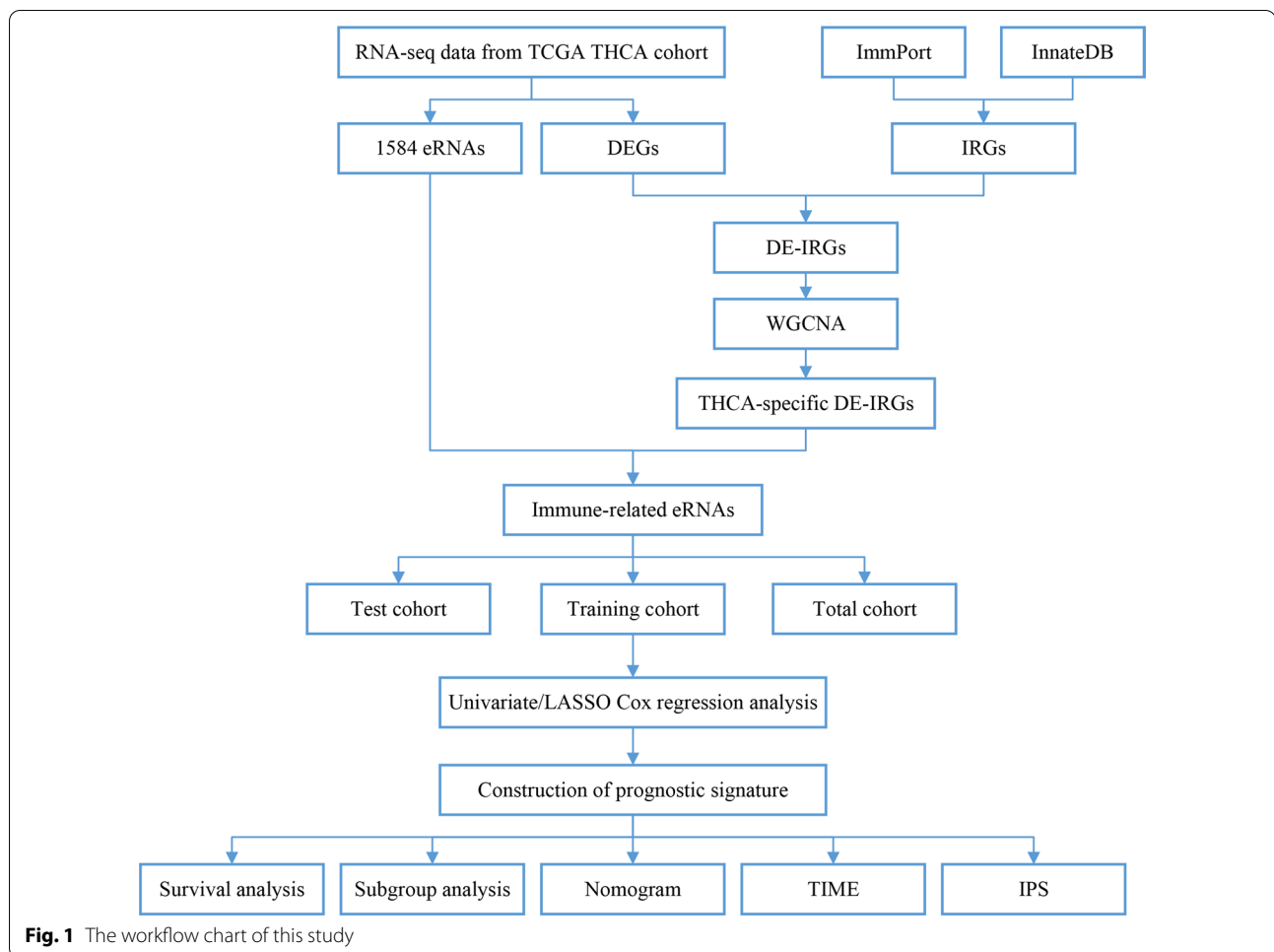
## Results

### Screening of THCA-specific DE-IRGs between normal and tumor tissues

The framework of the analytical process in our study was shown in Fig 1. A total of 3451 DEGs were identified based on the screening criteria of  $|\log_2(\text{Fold Change})| > 1$  and  $\text{FDR} < 0.05$ , including of 1699 downregulated genes and 1752 upregulated genes (Fig. 2A, C). Additionally, 362 DE-IRGs were extracted from these DEGs (Fig. 2D). Among them, 172 DE-IRGs were downregulated while 190 genes were upregulated in THCA samples (Fig. 2B, E). Subsequently, WGCNA was performed to construct a weighted co-expression network based on the DE-IRGs expression matrix. We used the soft-thresholding power of  $\beta = 11$  to achieve a scale-free network in the present study. Two gene modules (turquoise and grey modules) were identified. Among two modules, grey module ( $r = 0.62$ ,  $p = 1e-62$ ) showed the strongest correlation with THCA tissues (Fig. 3A). Therefore, we regarded the grey module as THCA-specific module and 271 DE-IRGs were identified for subsequent analysis.

### Construction of immune-related eRNAs risk signature

The expression matrixes of 1565 eRNAs among 1584 eRNAs and 271 DE-IRGs in grey module were extracted from the TCGA-THCA samples. A total of 125 immune-related eRNAs were obtained between 1565 eRNAs and 271 DE-IRGs using Pearson correlation analysis with  $|R| > 0.4$ ,  $p < 0.001$ . The Univariate Cox regression analysis was conducted in the training cohort to explore the prognostic value of 125 immune-related eRNAs. Sixteen eRNAs were significantly correlated with OS in the training cohort ( $p < 0.05$ ) (Fig. 3B). The expression levels of 16 prognostic eRNAs were presented in heatmap (Fig. 3C). The LASSO Cox regression analysis was performed to build the prognostic signature in the training set based on the 16 prognostic genes. According to the minimum criteria, a 9-gene risk signature consisting of FAAHP1, TP73-AS1, WDFY3-AS2, LINC01184, AL365259.1, TMEM184A, AC007255.1, IQANK1 and AC084375.1 was constructed (Fig. 3D, E). The immune-related eRNAs co-expression network was constructed based on the 9 genes (Fig. 3F). The risk score was calculated according to the following formula: risk score =  $(0.1087^* \text{expression level of FAAHP1}) + (0.0749^* \text{expression level of TP73-AS1}) + (0.3578^* \text{expression level of WDFY3-AS2}) + (0.3184^* \text{expression level of LINC01184}) + (-0.4084^* \text{expression level of AL365259.1}) + (0.1747^* \text{expression level of TMEM184A}) + (-0.0476^* \text{expression level of AC007255.1}) + (0.3280^* \text{expression level of IQANK1}) + (-0.1432^* \text{expression level of AC084375.1})$ . Risk score of each patient in the training cohort was

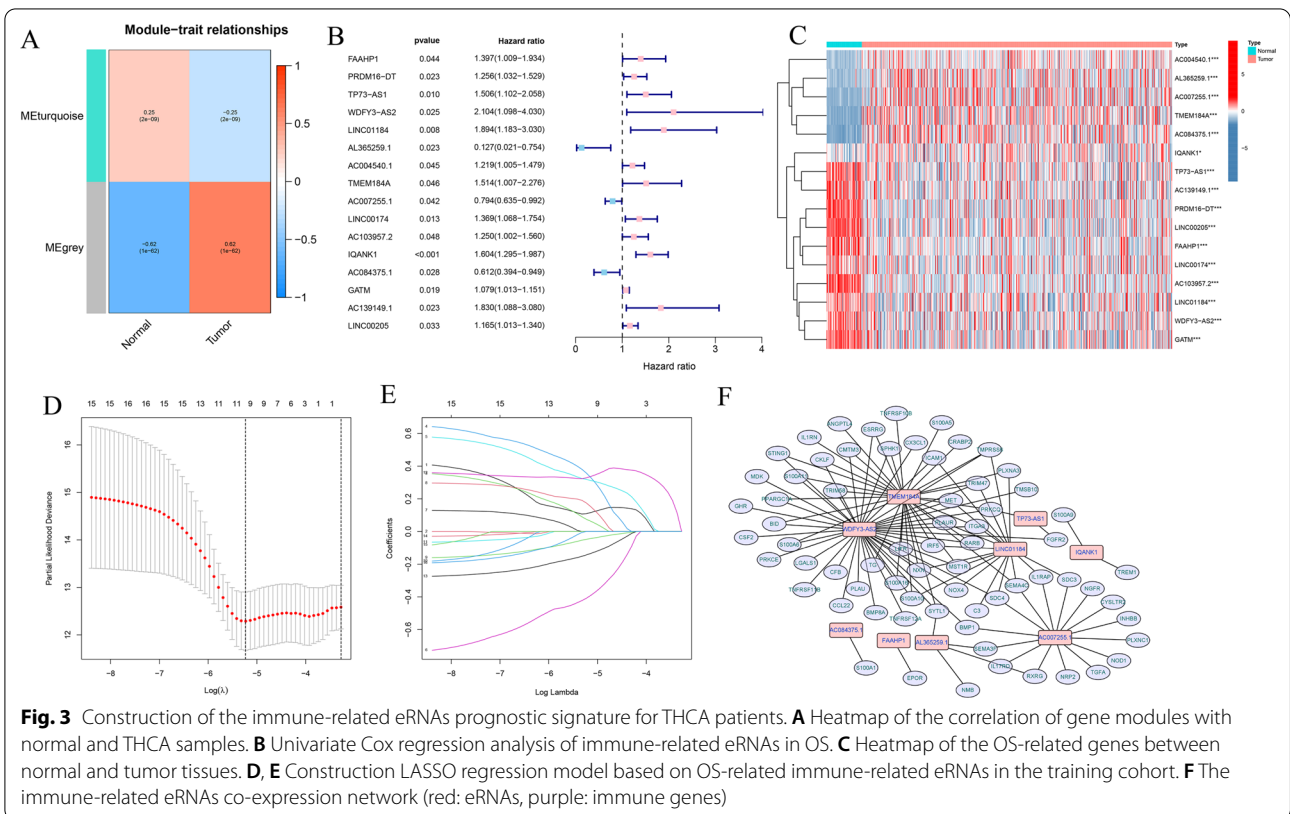
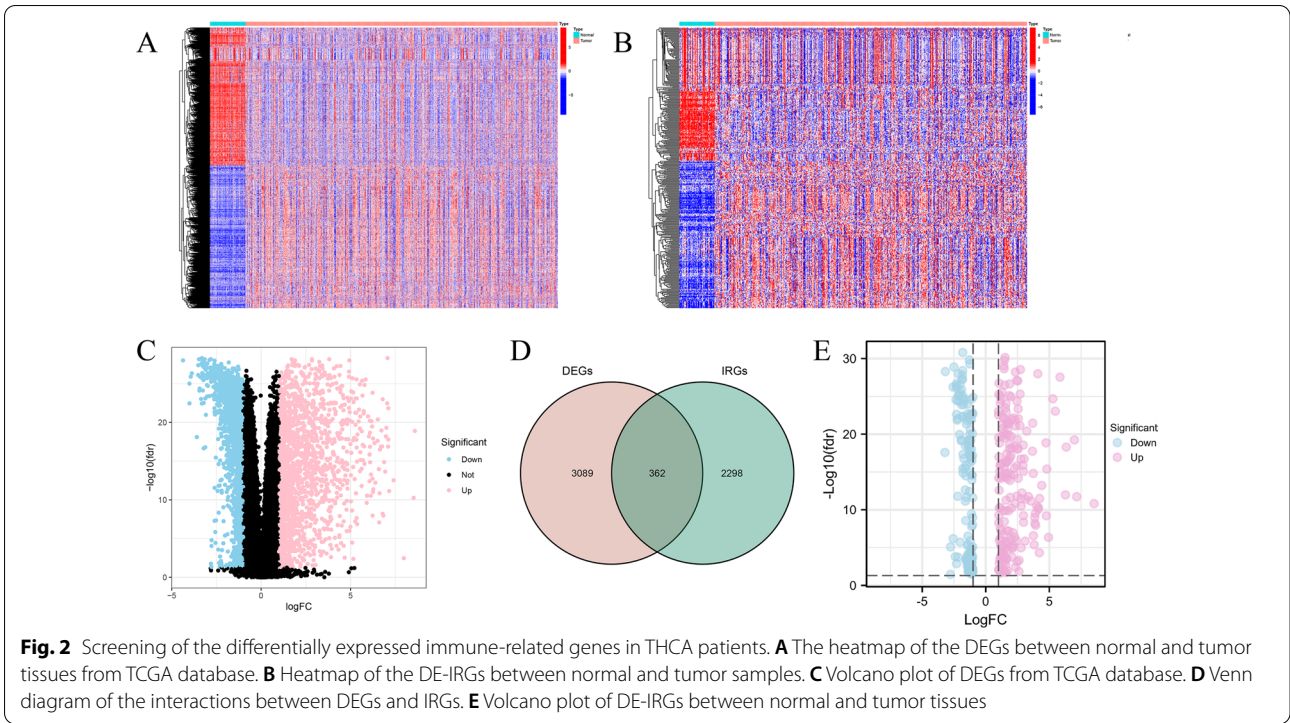


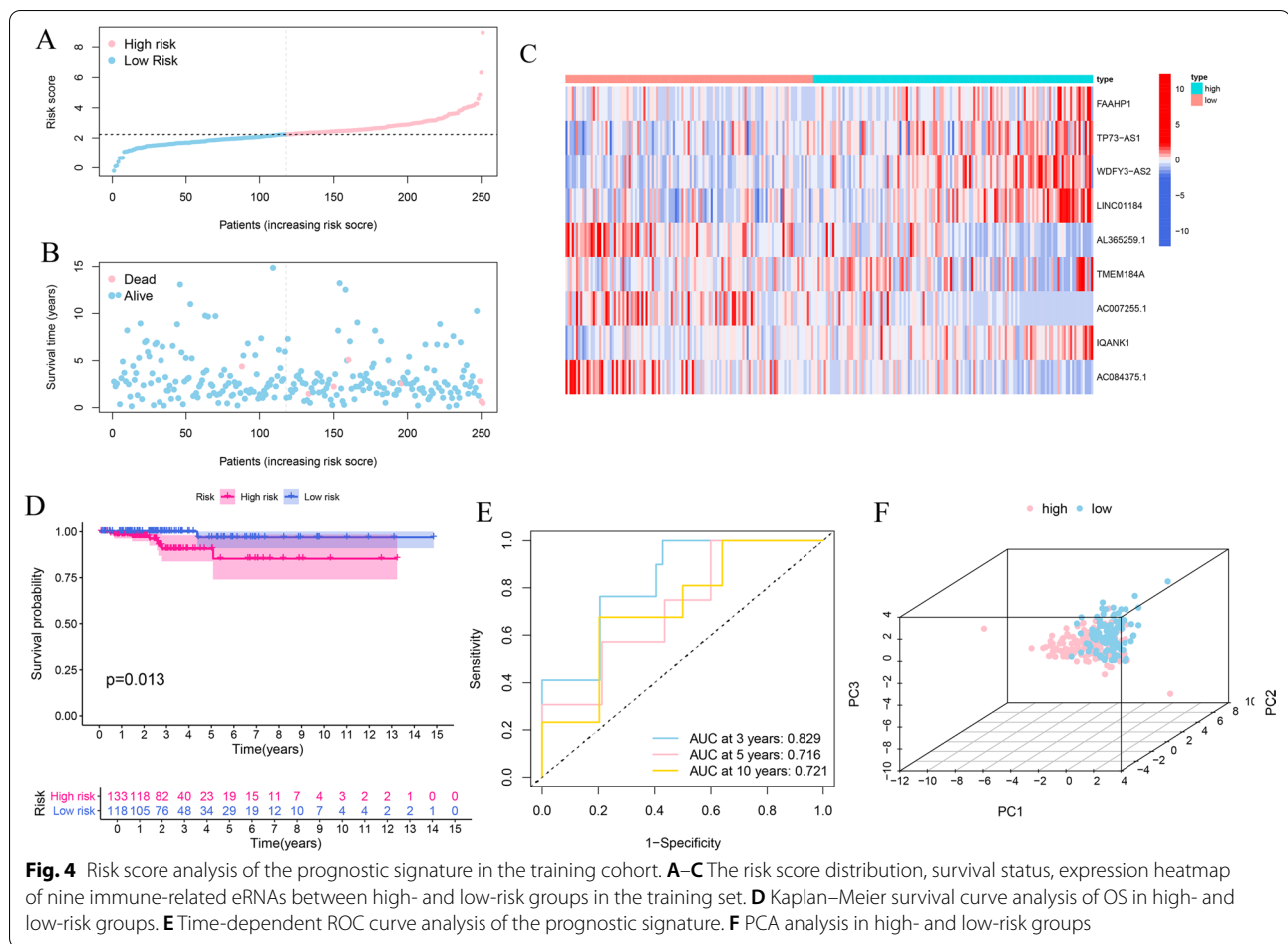
calculated and then patients were separated into high and low-risk subgroups according to the median risk score (Fig. 4A). Patients with high-risk group had more deaths and a shorter survival time than those in low-risk group (Fig. 4B). The heatmap revealed the expression patterns of 9 eRNAs between two different risk subgroups (Fig. 4C). The PCA analysis indicated that high and low-risk patients were well separated into two clusters (Fig. 4F). Kaplan-Meier survival curve indicated that high-risk group had a significantly poorer OS than low-risk groups ( $p = 0.013$ ) (Fig. 4D). ROC analysis was conducted to evaluate the predictive accuracy of the risk score and the areas under the ROC curve (AUC) were 0.829 for 3 year, 0.716 for 5 year, and 0.721 for 10 year (Fig. 4E).

Furthermore, the predictive capability of the risk signature was validated using test cohort and total cohort. The risk score of each patient was calculated according to the same formula as the training cohort. Patients in test cohort and total cohort were also classified into high- and low-risk groups using the same median score

in the training set. The OS of the patients with high-risk score was lower than that of the low-risk groups in the test cohort ( $p = 0.019$ ) (Fig. 5D). The AUC values of the ROC curve were 0.764 at 3 year, 0.885 at 5 year, and 0.889 at 10 year (Fig. 5E). We ranked the risk scores of patients in the test set and analyzed their distribution, survival status, and the expressions heatmap of nine biomarkers between high- and low-risk groups (Fig. 5A–C). The PCA plot indicated satisfactory separation in different risk subgroups (Fig. 5F).

The risk score distribution, survival status and the expression heatmap of nine eRNAs in the total cohort were displayed in Fig. 6A–C. The PCA analysis showed that risk genes could separate two risk groups (Fig. 6F). The result of the OS suggested that high-risk groups had a poorer prognosis compared with low-risk groups (Fig. 6D). The ROC analysis showed that the risk signature exhibited a reliable predictive capability (AUC = 0.813 in 3 year, 0.819 in 5 year and 0.824 in 10 year) (Fig. 6E). These results indicated that the risk signature was a reliable index.





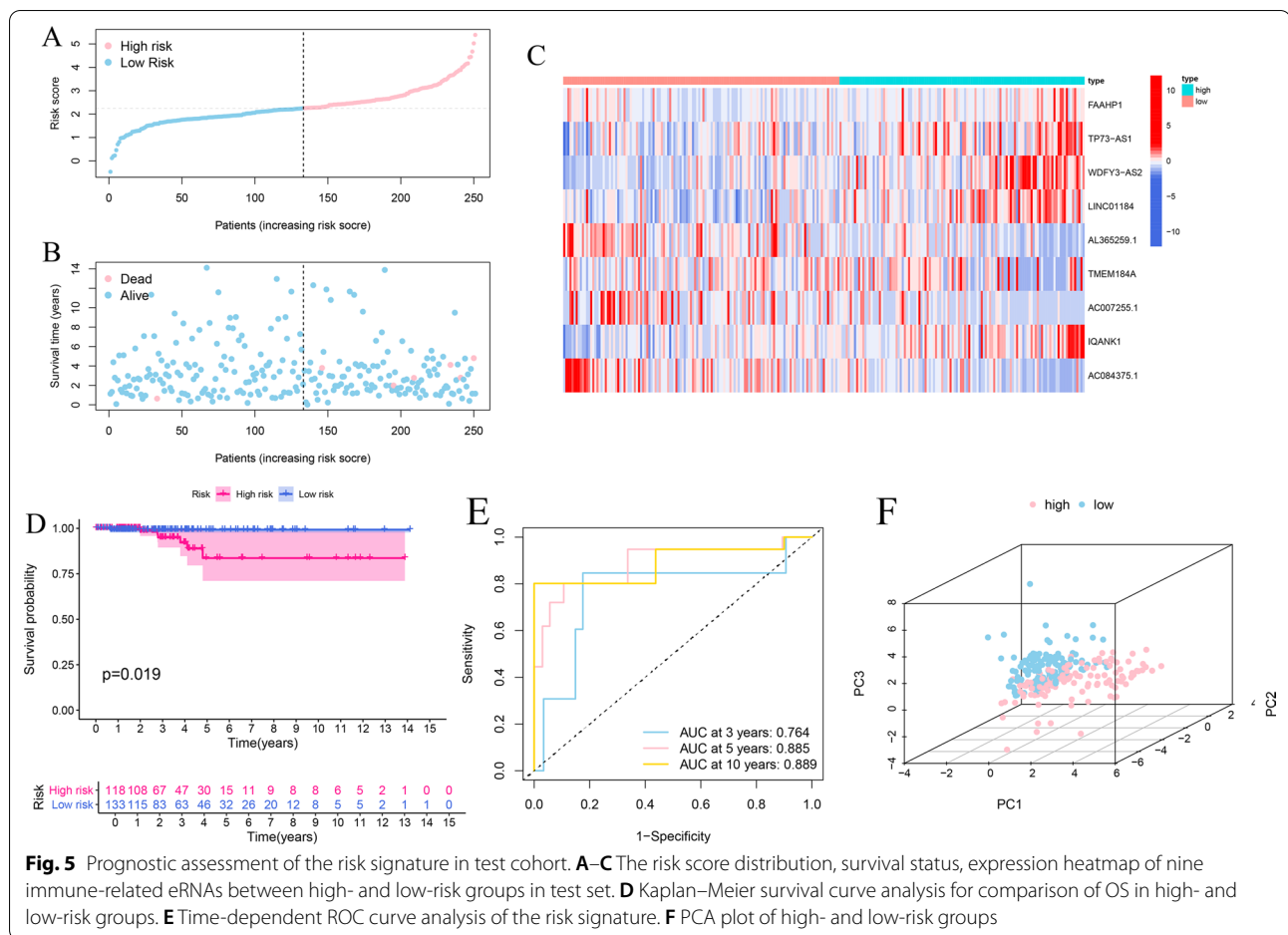
### Clinical value of risk signature

The possible relationships between risk signature and clinicopathological features were explored and the heatmap showed the distributions of clinical characteristics (survival status, age, gender, clinical stage and TNM stage) in different risk subgroups. The risk score was significantly associated with survival status, age and N stage ( $p < 0.05$ ) (Fig. 7A). In order to examine the predictive effects of the risk signature, all patients with THCA were separated into different subgroups according to different clinical characteristics, including age, gender, pathological tumor stage and TNM stage. The K-M survival curve suggested that high risk score predicted poorer prognosis in age  $> 60$  ( $p = 0.002$ ), gender ( $p = 0.003$  in female), stage ( $p = 0.029$  in stage I–II and  $p = 0.002$  in stage III–IV). Similar results were also observed in T1–2 ( $P = 0.013$ ) and T3–4 groups ( $P = 0.012$ ), stage N0 ( $P = 0.034$ ) and stage N1 groups ( $P = 0.023$ ), and M0 stage ( $P = 0.022$ ) (Fig. 7E). Furthermore, ROC analysis was conducted to evaluate the predictive performance based on the risk score and clinicopathological characteristics. The AUCs of the risk score (0.813), age (0.941) and stage

(0.713) at 3 year, risk score (0.820), age (0.896) and stage (0.773) at 5 year, and risk score (0.826), age (0.932) and stage (0.835) at 10 year, indicating a better predictive capability (Fig. 7B–D). More accurate prediction power was observed after combining risk score with other clinical features.

### Independent prognostic value of the risk signature

Univariate and multivariate Cox proportional hazards regression analyses were performed in training cohort, test cohort and total cohort to evaluate the independent prognostic value of risk score with clinical factors, including age, gender and pathological stage for THCA patients. Univariate Cox regression analysis showed that age and risk score were significantly associated with OS [training cohort: hazard ratio (HR) = 3.447, 95% confidence interval (CI) = 1.943–6.115,  $P < 0.001$ ; test cohort: HR = 2.957, 95% CI = 1.500–5.828,  $P = 0.002$ ; total cohort: HR = 3.376, 95% CI = 2.237–5.095,  $P < 0.001$ ] (Fig. 8A, C, E). After including other confounding variables in multivariate Cox regression analysis, the risk score was further identified as an independent prognostic



factor [training set: HR (95% CI) = 2.163 (1.266–3.696),  $P = 0.005$ ; test set: HR (95% CI) = 2.749 (1.427–5.295),  $P = 0.003$ ; total set: HR (95% CI) = 2.360 (1.613–3.452),  $P < 0.001$ ] (Fig. 8B, D, F).

Subsequently, we developed a prognostic nomogram to provide a quantitative analysis tool for predicting prognosis in patients with THCA based on the risk score and clinicopathological features (age, gender and stage) (Fig. 8G). The calibration curves of 3 year, 5 year and 10 year OS demonstrated an ideal consistency in predictive and actual survivals (Fig. 8H–J).

### Functional enrichment analysis of DEGs between different risk groups

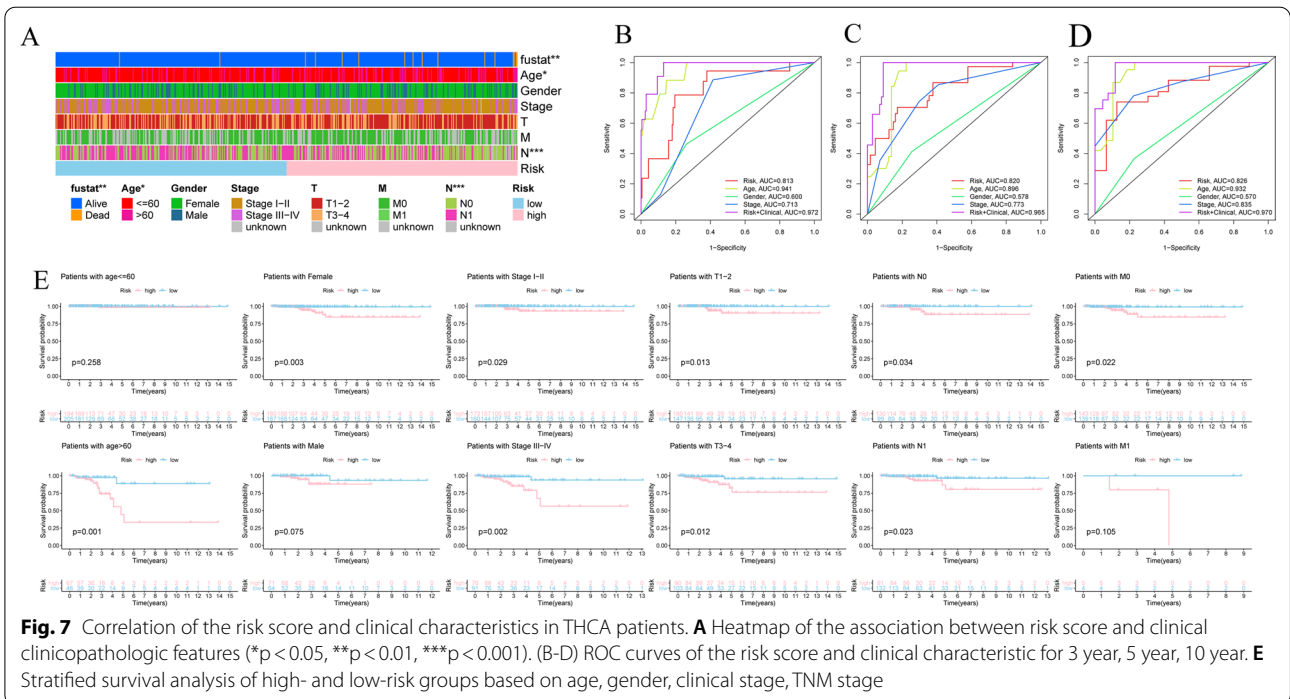
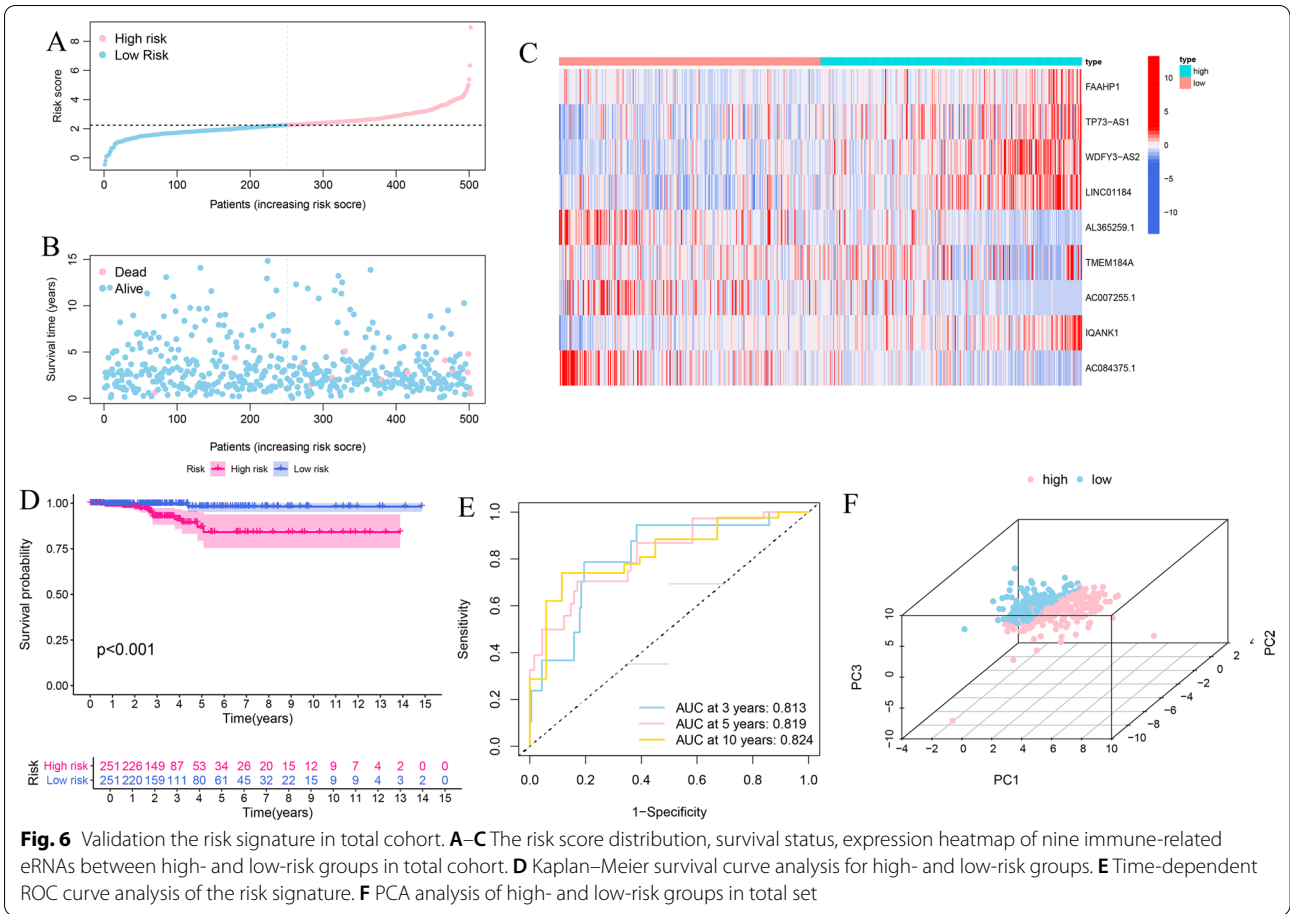
To further investigate the differences in gene functions between different risk subgroups, we identified 355 DEGs (201 downregulated genes and 154 upregulated genes in high-risk group) between high- and low-risk groups (Additional file 4: Table S2). The result of GO functional analysis suggested that the DEGs were mainly enriched in humoral immune response, immunoglobulin complex, and antigen binding (Fig. 9A). Additionally, GSEA

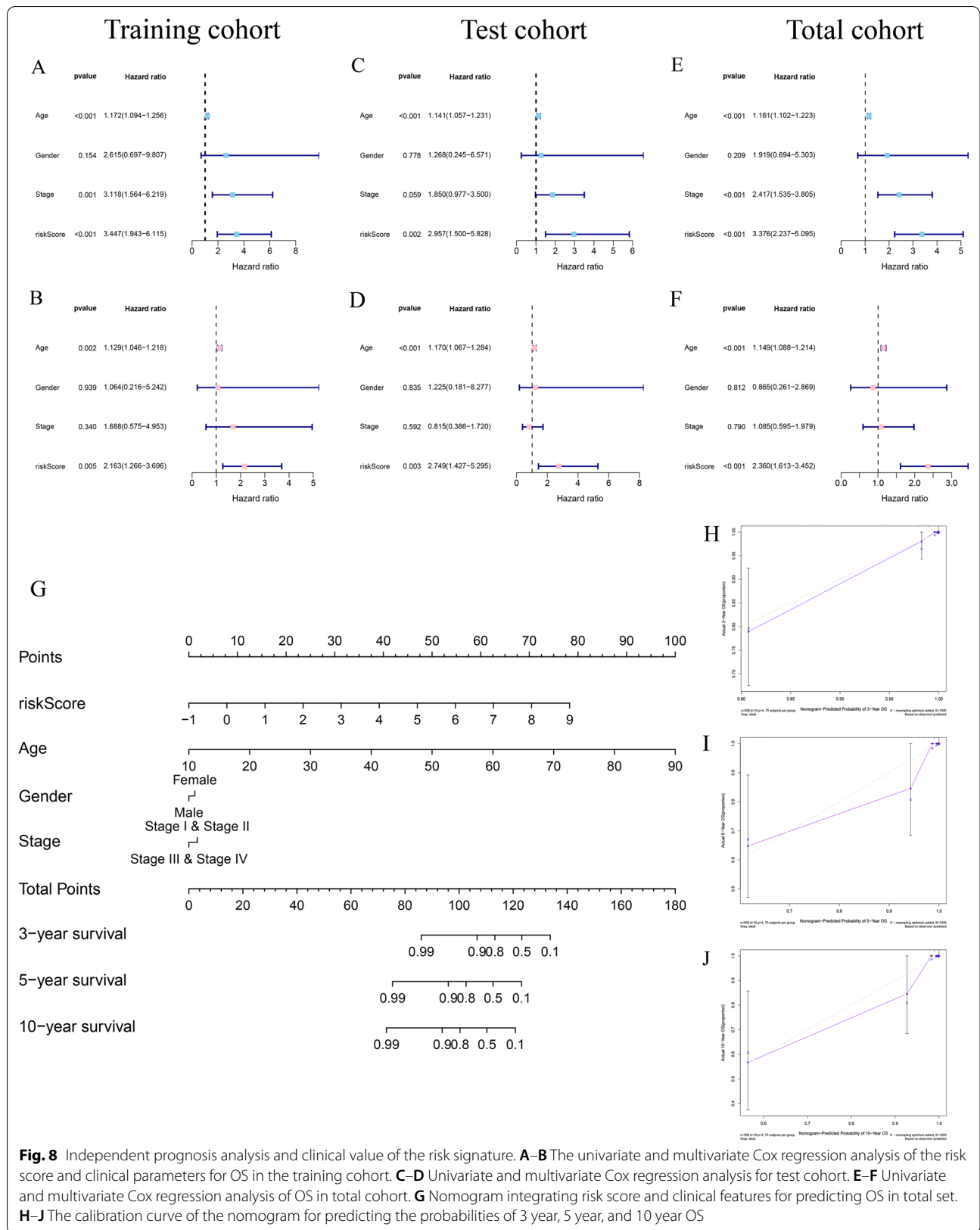
indicated that humoral immune response, regulation of humoral immune response, and positive regulation of humoral immune response were significantly involved in low-risk group (Fig. 9B).

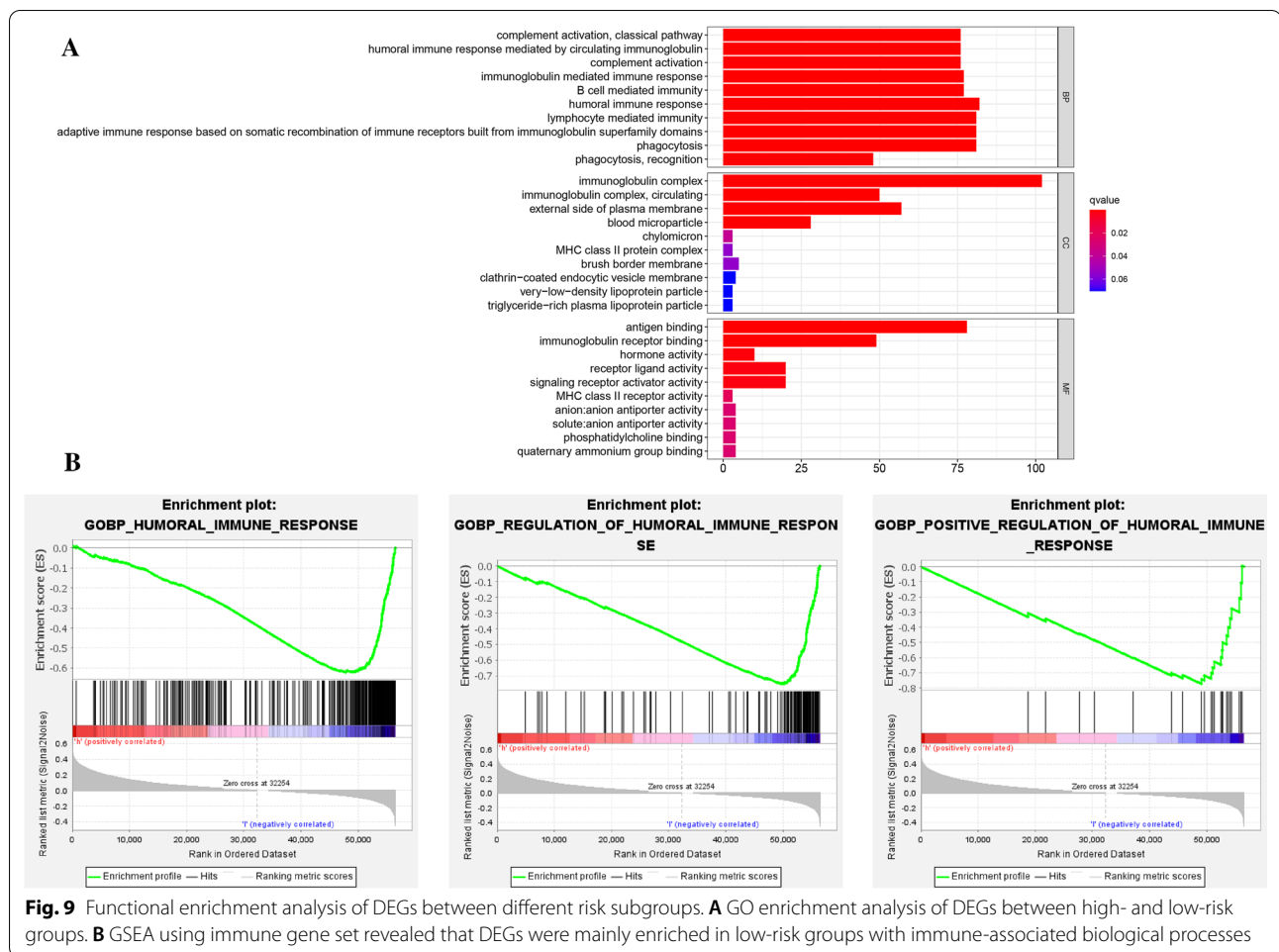
### Comparison of the tumor immune microenvironment in different risk subgroups

To explore the association between risk signature and tumor immune microenvironment (TIME), we calculated the relative proportion of each kind of immune cells among THCA patients based on the RNA-sequencing data using CIBERSORT algorithm. The proportions of 21 kinds of tumor-infiltrating immune cells in the low- and high-risk groups were displayed in Additional file 1: Figure S1A. The bar plot indicated that the proportions of dendritic cells resting and T cells CD4 memory activated were significantly higher in low-risk samples compared to high-risk patients ( $p < 0.05$ ) (Additional file 1: Figure S1B). The infiltrating ratios of Plasma cells, T cells CD8, NK cells activated, Monocytes, Mast cells resting, Macrophages M0, Macrophages M1, Macrophages M2, Dendritic cells resting, and Dendritic cells activated were







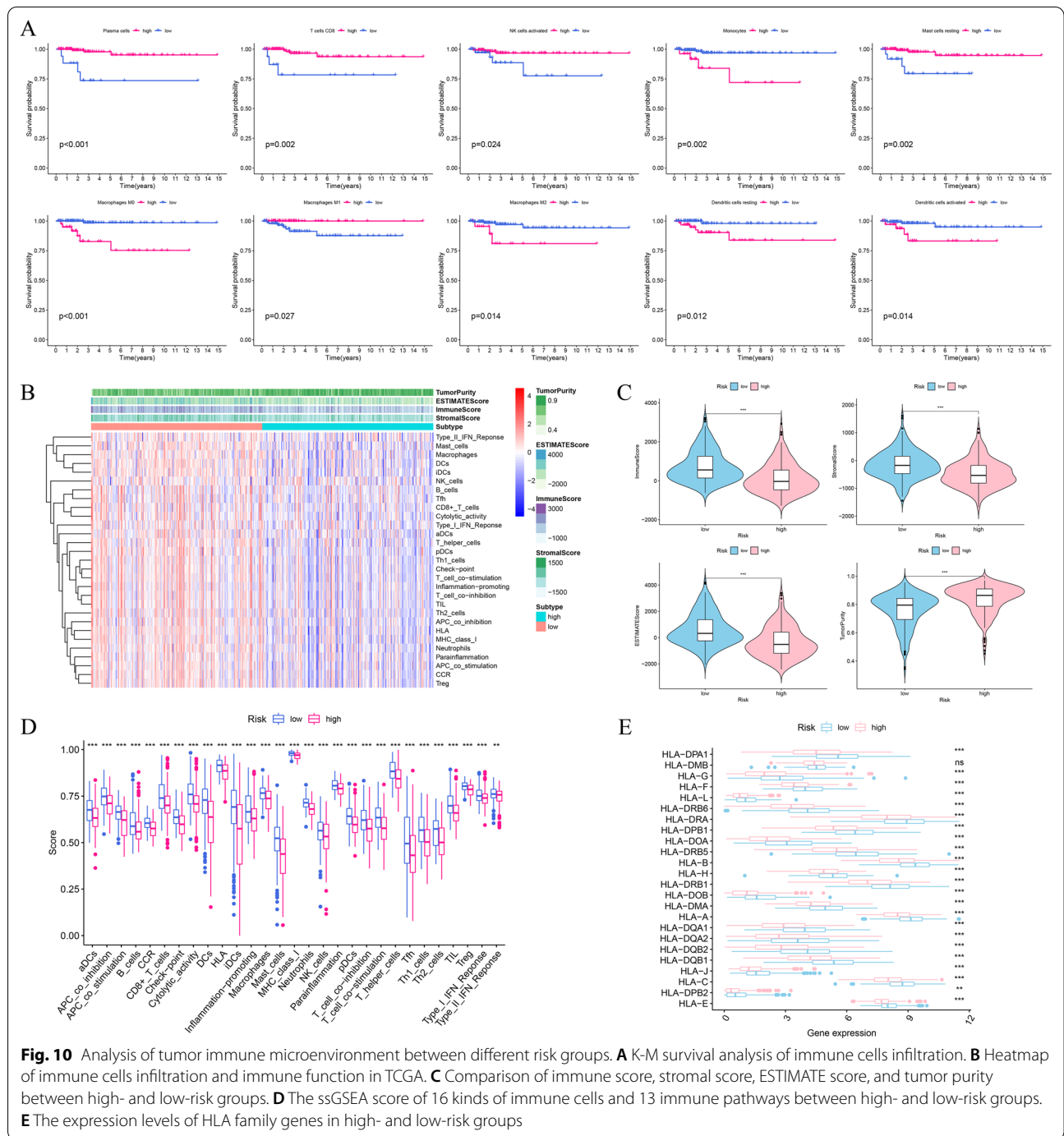


significantly associated with OS ( $p < 0.05$ ) (Fig. 10A). The higher infiltrating abundance of Monocytes, Macrophages M0, Macrophages M2, Dendritic cells resting, and Dendritic cells activated were tend to have a poorer OS, while high levels of Plasma cells, T cells CD8, NK cells activated, Mast cells resting, and Macrophages M1 were correlated with better OS.

We also calculated the immune score, stromal score, ESTIMATE score, and tumor purity in each THCA sample to investigate the differences of the TIME between high- and low-risk groups based on ESTIMATE algorithm. We performed ssGSEA to quantify the enrichment scores of 16 kinds of immune cells and the activity of 13 types of immune-related signaling pathways in THCA samples. The heatmap showed 29 immune-related gene sets for THCA patients and correlation between different risk subgroups and immune score, stromal score, ESTIMATE score, and tumor purity (Fig. 10B). The immune score, stromal score and ESTIMATE score were significantly higher in low-risk groups than in high-risk groups ( $p < 0.001$ ), while the tumor purity was lower in low-risk

groups ( $p < 0.001$ ) (Fig. 10C). Moreover, 29 immune-related gene sets were all significantly upregulated in low-risk groups ( $p < 0.01$ ) (Fig. 10D). Furthermore, the expression levels of the HLA family genes were higher in the low-risk groups except for HLA-DMB ( $p < 0.01$ ) (Fig. 10E).

The correlation between IPS and prognostic risk signature in THCA was explored to predict the patients' response to immune checkpoint inhibitors (ICIs). The IPS, IPS-CTLA4, IPS-PD1/PD-L1/PD-L2, and IPS-PD1/PD-L1/PD-L2 + CTLA4 scores were markedly higher in low-risk subgroups ( $p < 0.001$ ) (Fig. 11A). The expression levels of PD1, PD-L1, PD-L2, CTLA4, TIGIT, TIM-3, BTLA, and LAG3 were significantly higher in low-risk groups ( $p < 0.001$ ) (Fig. 11B). Furthermore, we also compared the difference in expression levels of cytokines between two different risk subgroups. There were significant differences in the expression levels of interleukin 1 beta (IL-1 $\beta$ ), IL-2, IL-6, IL-10, IL-18, tumor necrosis factor (TNF), granzyme A (GZMA), and GZMB in high- and low-risk groups ( $p < 0.001$ ) (Fig. 11C). These results

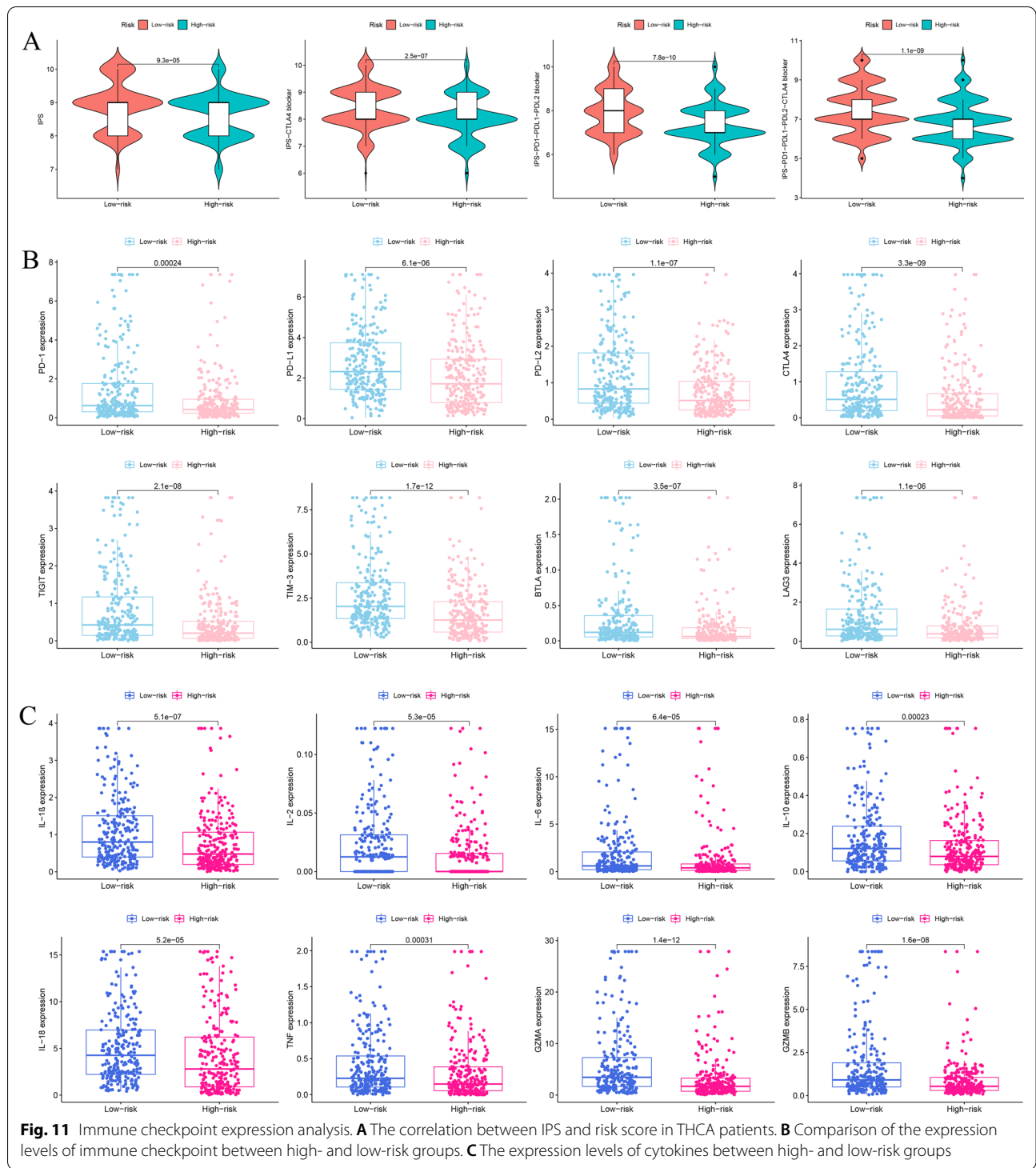


indicated that low-risk groups patients appeared to have a better opportunity for ICI treatment.

**Validation expression of risk eRNAs in THCA tissues**

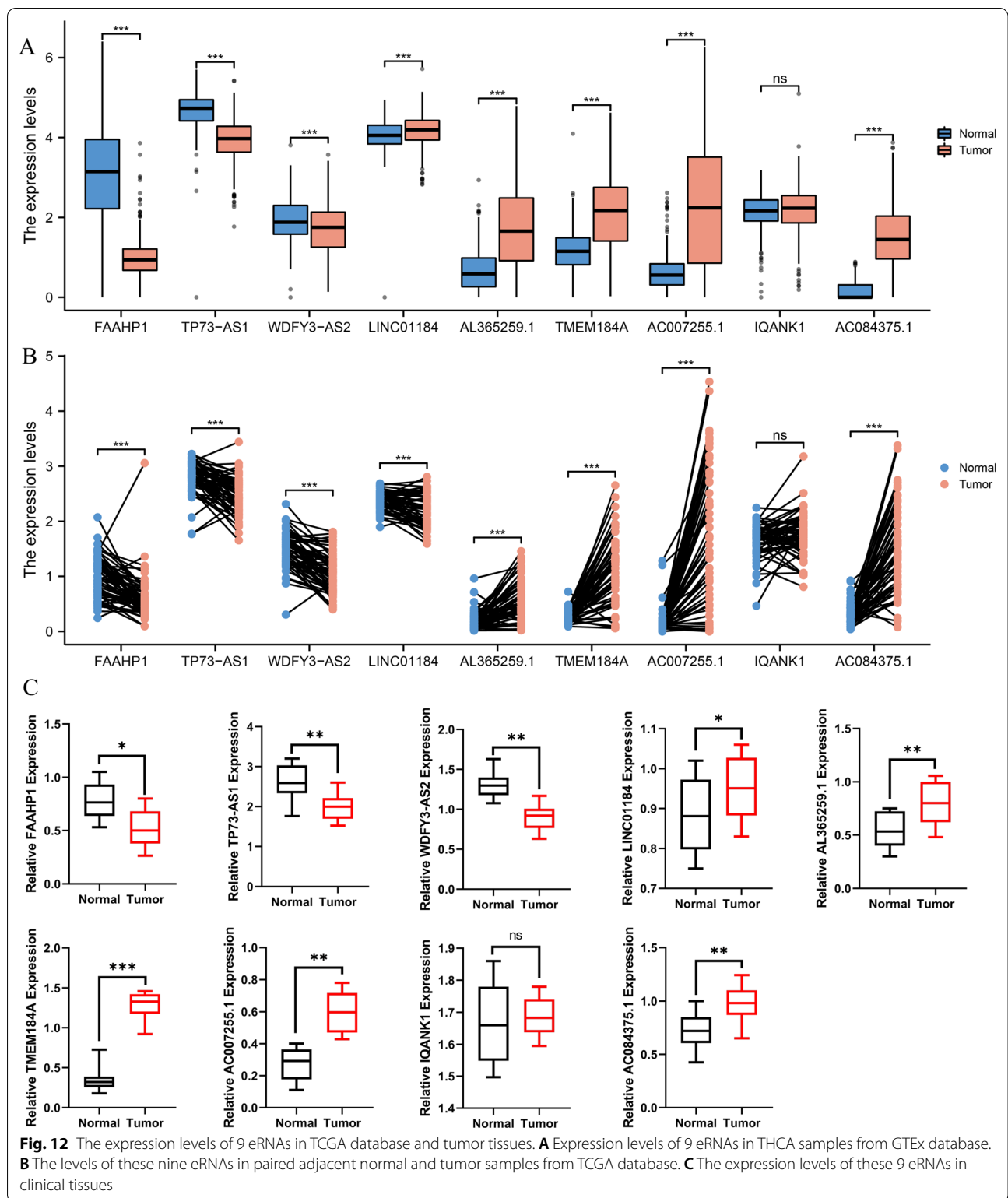
The expressions of FAAHP1, TP73-AS1, and WDFY3-AS2 were significantly down-expressed, while LINC01184, AL365259.1, TMEM184A,

AC007255.1, and AC084375.1 expression were significantly increased in THCA tissues compared with normal samples in TCGA and GTEx databases ( $p < 0.001$ ) (Fig. 12A–B). Nevertheless, IQANK1 expression showed no statistical difference. The FAAHP1, TP73-AS1, WDFY3-AS2, LINC01184, AL365259.1, TMEM184A, AC007255.1, and AC084375.1 indicated



well diagnostic accuracy with the area under the ROC curve (AUC) > 0.7 (Additional file 2: Figure S2). To further validate the expression level of 9 prognostic eRNAs, we analyzed the differential expression levels

in normal samples and THCA tissues based on TCGA database as well as in 62 pairs of thyroid cancer tissues and adjacent normal tissues by qRT-PCR (Fig. 12C). The expression levels were consistent with the results of bioinformatic analysis.



## Discussion

THCA is the most common form of endocrine system tumor, and its incidence has rapidly increased over the past few decades. PTC is the most predominant pathological subtype among all the THCA cases. Most PTC patients usually have a favorable prognosis when surgical treatment, radioactive iodine therapy, and thyroid-stimulating hormone (TSH) suppression treatment are implemented. Nonetheless, more than 10% of PTC patients may suffer local recurrence and distant metastasis after the initial treatment. Therefore, it is essential to explore new potential molecular targets for clinical therapy in order to improve patient outcomes.

TME consists of immune cells and non-immune stromal cells, and immune system was found to make a crucial contribution to cancer development and progression. Cancer immunotherapy has made tremendous progress for some cancer types in recent years.

Predictive or prognostic biomarkers related to the TIME have great promise for identifying novel molecular therapeutic targets and improving cancer patients clinical management of immunotherapy [31].

Long non-coding RNAs (lncRNAs) have emerged as enormous amount and diverse functions in the past decade and have been reported to play important roles in a variety of biological processes, including cell proliferation, death, and tumor growth [32]. Enhancer was described as distal regulatory DNA that regulate the transcription of target genes by interacting with promoters of target genes. Studies suggested that enhancers can transcribe non-coding RNAs, which was been defined as enhancer RNAs (eRNAs) [33]. eRNAs were limited to around 500–2000 bp and had shorter half-lives [34]. Extensive evidences suggested significant roles of eRNAs in tumorigenesis and they involved in various cancer signaling pathways through regulating their target genes. The activation of oncogenic signaling pathways in human cancers often enhanced enhancer activation and production of eRNAs. CELF2 is highly expressed in stomach adenocarcinoma [35]. APH1A is highly expressed in grade-3 hepatocellular carcinoma (HCC). EN1 is highly expressed in breast cancer (BRAC), and ESRI can increase eRNA transcription in BRAC [36]. TAOK1 is associated with overall survival in clear cell renal cell carcinoma [22, 37]. SCRIB was differentially expressed among lung adenocarcinoma patients [38]. Tumor suppressors-induced eRNAs may implicated in tumor suppression, while oncogene-induced eRNAs can directly promote tumorigenesis. Therefore, eRNAs were closely correlated to malignancy formation and progression.

In our study, we identified 3451 DEGs from 510 THCA samples and 58 normal samples in TCGA cohort, and 362 DE-IRGs (172 downregulated 190 upregulated

genes) were extracted. WGCNA was performed to identify THCA-specific immune-related hub genes based on 362 DE-IRGs, and two gene modules (turquoise and grey modules) were obtained. Finally, the grey module was regarded as THCA-specific module. Pearson correlation analysis was used to evaluate immune-related eRNAs. Subsequently, 9-immune-eRNAs prognostic signature (FAAHP1, TP73-AS1, WDFY3-AS2, LINC01184, AL365259.1, TMEM184A, AC007255.1, IQANK1 and AC084375.1) was constructed based on univariate Cox regression analysis and LASSO Cox regression analysis, which was an independent prognostic factor for OS. TP73-AS1, also known as KIAA0495, is abnormally expressed in many cancers [39]. Previous studies indicated that TP73-AS1 could be a key role in regulating HCC cells proliferation and its expression level was associated with poor prognosis of HCC patients [40]. Besides, Wang et al. showed TP73-AS1 also interfered the metastasis and proliferation of ovarian cancer [41]. Of note, we firstly found the expression level of TP73-AS1 is higher in normal thyroid cancer, suggesting that it might have different mechanism in regulating tumor progression comparing other cancers. Increasingly evidences have implicated that lncRNAs participated in the process of cell growth, invasion. Studies showed that overexpressed WDFY3-AS2 suppressed the proliferation, invasion, and epithelial-to-mesenchymal transition (EMT) in ovarian cancer [42]. Furthermore, WDFY3-AS2 also could promote cisplatin resistance by the expression of miR-139-5p/SDC4 in ovarian cancer, which may provide a promising drug target to drug resistance [43]. WDFY3-AS2 participated in the development and progression of oesophageal squamous cell carcinoma (ESCC) by regulating miR-2355-5p/SOCS2 axis, which suggested that WDFY3-AS2 might be an underlying predictor and novel therapeutic target for ESCC patients [44]. Other studies suggested that the potential value of long noncoding RNA WDFY3-AS2 might be novel prognostic biomarker for lung adenocarcinoma (LUAD) [45], glioma [46], and esophageal cancer (EC) [47]. AC007255.1, an immune-related prognostic eRNA, was up-regulated in EC tissues and high expression indicated a poorer prognosis. It was closely related to immune response and infiltration levels of immune cells, such as B cell, dendritic cell and neutrophil [48]. Sui et al. demonstrated that LINC01184 was highly expressed in colorectal cancer, and it could affect the the proliferation and invasion of colorectal cancer cells through the linc01184-miR-331-HER2-p-Akt/ERK1/2 pathway [49]. Subsequently, the risk score was calculated and separated all patients into high- and low-risk subgroups based on median risk score, and Kaplan–Meier survival curves indicated that the high-risk groups had poorer clinical results than that of the low-risk

groups. Univariate and multivariate Cox regression analyses showed the risk score was an independent prognostic factor. Additionally, GO functional analysis suggested that the DEGs between different risk groups were mainly enriched in humoral immune response, immunoglobulin complex, and antigen binding. GSEA showed that humoral immune response, regulation of humoral immune response, and positive regulation of humoral immune response were significantly enriched in low-risk group. To further estimate the TIME of the prognostic signature, ESTIMATE and CIBERSORT were performed to estimate the immune score, stromal score, and tumor purity in THCA sample. More dendritic cells resting and T cells CD4 memory activated were significantly infiltrated in TIME of low-risk group. In addition, patients in the low-risk groups had higher immune score, stromal score and ESTIMATE score than in high-risk group. We applied ssGSEA to assess the immune status of the risk signature, the results suggested that 29 immune-related cells were significantly up-regulated in low-risk group. At the same time, the expression levels of cytokines and immunosuppressor molecules (PD1, PD-L1, PD-L2, CTLA4, TIGIT, TIM-3, BTLA, and LAG3) were significantly higher in low-risk groups, implying more tumor immunogenicity in the low-risk group. The good response of ICIs might be one of the reasons for the good clinical outcome in the low-risk group. Therefore, patients with low-risk scores might be more likely to benefit from ICI treatment.

The present study has some limitations. Our data was retrieved from TCGA public database instead of our own cohort, the predictive power of the prognostic signature should be validated using an external validation cohort, such as GEO cohort. Nevertheless, information on overall survival of THCA patients was unfortunately largely lacking in the GEO databases. Moreover, some basic experiments should be performed to further validate our bioinformatics analysis results in the future. The efficacy of immunotherapy in THCA patients was needed to be validated in large clinical trials.

## Conclusion

Taken together, we are the first to construct an immune-related eRNAs prognostic signature to efficiently predict the survival and prognosis with high specificity for THCA patients. Significant differences were founded between prognostic signature and TIME, all patients with different risk levels exhibited different response to immunotherapy. This study could provide a novel insight into a potentially novel prognostic prediction and offer opportunity for individualized immunotherapy of THCA patients in future studies.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12935-022-02722-8>.

**Additional file 1: Figure S1.** Tumor-infiltrating immune cells in different risk subgroups. **A** The abundance of 21 immune cells between high- and low-risk groups. **B** Differences in fractions of tumor-infiltrating immune cells between two risk subgroups.

**Additional file 2: Figure S2.** The diagnostic values of 9 eRNAs for THCA patients.

**Additional file 3: Table S1.** Premier sequences for qRT-PCR analysis.

**Additional file 4: Table S2.** The differentially expressed genes between high and low-risk group.

## Acknowledgements

We acknowledge the Cancer Genome Atlas (TCGA) database for sharing their meaningful datasets.

## Author contributions

PW performed the data analysis and wrote the manuscript. JS and ZW performed the experiments. WS and HZ reviewed and revised the manuscript. All authors read and approved the final manuscript.

## Funding

This work was supported by the National Natural Science Foundation of China (81902726), the Natural Science Foundation of Education Bureau of Liaoning Province (QNZR2020002), Natural Science Foundation of Education Bureau of Liaoning Province (QNZR2020009), the Natural Science Foundation of Liaoning Province (2020-MS-143), and the Natural Science Foundation of Liaoning Province (2020-MS-186), the Natural Science Foundation of Liaoning Province (2021-MS-193) and Science and Technology Project of Shenyang City (21-173-9-31).

## Availability of data and materials

Gene expression profiles, clinical information of THCA in this study are available from the public database (TCGA, <https://portal.gdc.cancer.gov/>) and Genotype-Tissue Expression Project (GTEx) database (<https://www.gtexportal.org/>). The immune related genes are acquired from the Immunology Database and Analysis Portal database (ImmPort, <https://immport.niaid.nih.gov/>) and InnateDB (<https://www.innatedb.com/>) databases. The IPS values are downloaded from The Cancer Immunome Atlas (TCIA, <https://tcia.at/home>).

## Declarations

### Ethics approval and consent to participate

The research was approved by the Institutional Research Ethics Committees of the First Affiliated Hospital of China Medical University. Informed consent for publication was obtained from all patients for collection of tissue samples prior to the surgery.

### Consent for publication

Not applicable.

### Competing interests

The authors declare that they have no competing interests.

Received: 23 March 2022 Accepted: 17 September 2022  
Published online: 10 October 2022

## References

- Filetti S, et al. Thyroid cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Ann Oncol.* 2019;30:1856–83. <https://doi.org/10.1093/annonc/mdz400>.



2. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. *CA Cancer J Clin*. 2020;70:7–30. <https://doi.org/10.3322/caac.21590>.
3. Ahn JH, et al. A prospective randomized controlled trial to assess the efficacy and safety of prophylactic central compartment lymph node dissection in papillary thyroid carcinoma. *Surgery*. 2022;171:182–9. <https://doi.org/10.1016/j.surg.2021.03.071>.
4. Agrawal N, et al. Indications and extent of central neck dissection for papillary thyroid cancer: an American head and neck society consensus statement. *Head Neck*. 2017;39:1269–79. <https://doi.org/10.1002/hed.24715>.
5. Xiao Y, Yu D. Tumor microenvironment as a therapeutic target in cancer. *Pharmacol Ther*. 2021;221: 107753. <https://doi.org/10.1016/j.pharmthera.2020.107753>.
6. Lin J, Kato M, Nagata K, Okuwaki M. Efficient DNA binding of NF- $\kappa$ B requires the chaperone-like function of NPM1. *Nucleic Acids Res*. 2017;45:3707–23. <https://doi.org/10.1093/nar/gkw1285>.
7. Jia D, et al. Mining TCGA database for genes of prognostic value in glioblastoma microenvironment. *Aging*. 2018;10:592–605. <https://doi.org/10.18632/aging.101415>.
8. Peng Y, et al. Identification of a prognostic and therapeutic immune signature associated with hepatocellular carcinoma. *Cancer Cell Int*. 2021;21:98. <https://doi.org/10.1186/s12935-021-01792-4>.
9. Dai Y, et al. An immune-related gene signature for predicting survival and immunotherapy efficacy in hepatocellular carcinoma. *Cancer Immunol Immunother*. 2021;70:967–79. <https://doi.org/10.1007/s00262-020-02743-0>.
10. Chen Y, et al. Development and validation of a prognostic signature based on immune genes in cervical cancer. *Front Oncol*. 2021;11: 616530. <https://doi.org/10.3389/fonc.2021.616530>.
11. Zhang Y, et al. A novel immune-related prognostic signature in head and neck squamous cell carcinoma. *Front Genet*. 2021;12: 570336. <https://doi.org/10.3389/fgene.2021.570336>.
12. Qiu Y, et al. Development and validation of a robust immune prognostic signature for head and neck squamous cell carcinoma. *Front Oncol*. 2020;10:1502. <https://doi.org/10.3389/fonc.2020.01502>.
13. Guan X, Xu ZY, Chen R, Qin JJ, Cheng XD. Identification of an immune gene-associated prognostic signature and its association with a poor prognosis in gastric cancer patients. *Front Oncol*. 2020;10: 629909. <https://doi.org/10.3389/fonc.2020.629909>.
14. Jin W, et al. Exploration of the molecular characteristics of the tumor-immune interaction and the development of an individualized immune prognostic signature for neuroblastoma. *J Cell Physiol*. 2021;236:294–308. <https://doi.org/10.1002/jcp.29842>.
15. Zhao HB, et al. Novel immune-related signature for risk stratification and prognosis in prostatic adenocarcinoma. *Cancer Sci*. 2021;112:4365–76. <https://doi.org/10.1111/cas.15062>.
16. Rondina MT, et al. Longitudinal RNA-seq analysis of the repeatability of gene expression and splicing in human platelets identifies a platelet *seip* splice QTL. *Circ Res*. 2020;126:501–16. <https://doi.org/10.1161/CIRCRESAHA.119.315215>.
17. Tian W, et al. A novel prognostic tool for glioma based on enhancer RNA-regulated immune genes. *Front Cell Dev Biol*. 2021;9: 798445. <https://doi.org/10.3389/fcell.2021.798445>.
18. Ritchie ME, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43: e47. <https://doi.org/10.1093/nar/gkv007>.
19. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*. 2008;9:559. <https://doi.org/10.1186/1471-2105-9-559>.
20. Zhu W, Ru L, Ma Z. Identification of a novel four-gene signature correlated with the prognosis of patients with hepatocellular carcinoma: a comprehensive analysis. *Front Oncol*. 2021;11: 626654. <https://doi.org/10.3389/fonc.2021.626654>.
21. Murakami S, Gadad SS, Kraus WL. A PreSTIGEous use of lncRNAs to predict enhancers. *Cell Cycle*. 2015;14:1619–20. <https://doi.org/10.1080/15384101.2015.1032650>.
22. Cai S, Hu X, Chen R, Zhang Y. Identification and validation of an immune-related eRNA prognostic signature for hepatocellular carcinoma. *Front Genet*. 2021;12: 657051. <https://doi.org/10.3389/fgene.2021.657051>.
23. Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *J Stat Softw*. 2010;33:1–22.
24. Wu P, Shi J, Sun W, Zhang H. Identification and validation of a pyroptosis-related prognostic signature for thyroid cancer. *Cancer Cell Int*. 2021;21:523. <https://doi.org/10.1186/s12935-021-02231-0>.
25. Fu XW, Song CQ. Identification and validation of pyroptosis-related gene signature to predict prognosis and reveal immune infiltration in hepatocellular carcinoma. *Front Cell Dev Biol*. 2021;9: 748039. <https://doi.org/10.3389/fcell.2021.748039>.
26. Heagerty PJ, Lumley T, Pepe MS. Time-dependent ROC curves for censored survival data and a diagnostic marker. *Biometrics*. 2000;56:337–44. <https://doi.org/10.1111/j.0006-341x.2000.00337.x>.
27. Newman AM, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods*. 2015;12:453–7. <https://doi.org/10.1038/nmeth.3337>.
28. Yoshihara K, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat Commun*. 2013;4:2612. <https://doi.org/10.1038/ncomms3612>.
29. Rooney MS, Shukla SA, Wu CJ, Getz G, Hacohen N. Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell*. 2015;160:48–61. <https://doi.org/10.1016/j.cell.2014.12.033>.
30. Charoentong P, et al. Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade. *Cell Rep*. 2017;18:248–62. <https://doi.org/10.1016/j.celrep.2016.12.019>.
31. Li B, Cui Y, Diehn M, Li R. Development and validation of an individualized immune prognostic signature in early-stage nonsquamous non-small cell lung cancer. *JAMA Oncol*. 2017;3:1529–37. <https://doi.org/10.1001/jamaoncol.2017.1609>.
32. Niknafs YS, et al. The lncRNA landscape of breast cancer reveals a role for DSCAM-AS1 in breast cancer progression. *Nat Commun*. 2016;7:12791. <https://doi.org/10.1038/ncomms12791>.
33. Zhang Z, et al. Transcriptional landscape and clinical utility of enhancer RNAs for eRNA-targeted therapy in cancer. *Nat Commun*. 2019;10:4562. <https://doi.org/10.1038/s41467-019-12543-5>.
34. Liu Y, et al. Current advances on the important roles of enhancer RNAs in gene regulation and cancer. *Biomed Res Int*. 2018;2018:2405351. <https://doi.org/10.1155/2018/2405351>.
35. Barretina J, et al. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature*. 2012;483:603–7. <https://doi.org/10.1038/nature11003>.
36. Beltran AS, Graves LM, Blancafort P. Novel role of Engrailed 1 as a prosurvival transcription factor in basal-like breast cancer and engineering of interference peptides block its oncogenic function. *Oncogene*. 2014;33:4767–77. <https://doi.org/10.1038/onc.2013.422>.
37. Harvey KF, Zhang X, Thomas DM. The hippo pathway and human cancer. *Nat Rev Cancer*. 2013;13:246–57. <https://doi.org/10.1038/nrc3458>.
38. Li P, et al. Enhancer RNA SLIT2 inhibits bone metastasis of breast cancer through regulating P38 MAPK/c-Fos signaling pathway. *Front Oncol*. 2021;11: 743840. <https://doi.org/10.3389/fonc.2021.743840>.
39. Chu F, Xue L, Miao H. Long noncoding RNA TP73-AS1 in human cancers. *Clin Chim Acta*. 2020;500:104–8. <https://doi.org/10.1016/j.cca.2019.09.024>.
40. Li S, et al. The long non-coding RNA TP73-AS1 modulates HCC cell proliferation through miR-200a-dependent HMGB1/RAGE regulation. *J Exp Clin Cancer Res*. 2017;36:51. <https://doi.org/10.1186/s13046-017-0519-z>.
41. Wang X, Yang B, She Y, Ye Y. The lncRNA TP73-AS1 promotes ovarian cancer cell proliferation and metastasis via modulation of MMP2 and MMP9. *J Cell Biochem*. 2018;119:7790–9. <https://doi.org/10.1002/jcb.27158>.
42. Li W, et al. Long noncoding RNA WDFY3-AS2 suppresses tumor progression by acting as a competing endogenous RNA of microRNA-18a in ovarian cancer. *J Cell Physiol*. 2020;235:1141–54. <https://doi.org/10.1002/jcp.29028>.
43. Wu Y, Wang T, Xia L, Zhang M. lncRNA WDFY3-AS2 promotes cisplatin resistance and the cancer stem cell in ovarian cancer by regulating hsa-miR-139-5p/SDC4 axis. *Cancer Cell Int*. 2021;21:284. <https://doi.org/10.1186/s12935-021-01993-x>.
44. Zhang Q, et al. lncRNA WDFY3-AS2 suppresses proliferation and invasion in oesophageal squamous cell carcinoma by regulating miR-2355-5p/SOCS2 axis. *J Cell Mol Med*. 2020;24:8206–20. <https://doi.org/10.1111/jcmm.15488>.
45. Ren P, Hong X, Chang L, Xing L, Zhang H. USF1-induced overexpression of long noncoding RNA WDFY3-AS2 promotes lung adenocarcinoma

progression via targeting miR-491-5p/ZNF703 axis. *Mol Carcinog.* 2020;59:875–85. <https://doi.org/10.1002/mc.23181>.

46. Wu F, et al. Expression profile analysis of antisense long non-coding RNA identifies WDFY3-AS2 as a prognostic biomarker in diffuse glioma. *Cancer Cell Int.* 2018;18:107. <https://doi.org/10.1186/s12935-018-0603-2>.
47. Kong Q, et al. Long Noncoding RNA WDFY3-AS2 represses the progression of esophageal cancer through miR-18a/PTEN axis. *J Oncol.* 2021;2021:9951010. <https://doi.org/10.1155/2021/9951010>.
48. Wang Q, Yu X, Yang N, Xu L, Zhou Y. LncRNA AC007255.1, an immune-related prognostic enhancer RNA in esophageal cancer. *PeerJ.* 2021. <https://doi.org/10.7717/peerj.11698>.
49. Sui Y-X, Zhao D-L, Yu Y, Wang L-C, Rizzo R. The role, function, and mechanism of long intergenic noncoding RNA1184 (linc01184) in colorectal cancer. *Dis Markers.* 2021;1–11:2021. <https://doi.org/10.1155/2021/8897906>.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

