

# A Novel Signature to Predict Thyroid Cancer Prognosis and Immune Landscape Using Immune-Related LncRNA Pairs

**BO SONG**

China Medical University

**Lijun Tian**

China Medical University

**Fan Zhang**

China Medical University

**Zheyu Lin**

China Medical University

**Boshen Gong**

China Medical University

**Tingting Liu**

China Medical University

**Weiping Teng** (✉ [twp@vip.163.com](mailto:twp@vip.163.com))

China Medical University <https://orcid.org/0000-0002-6445-6192>

---

## Research

**Keywords:** TCGA, lncRNA pairs, immune infiltration, thyroid carcinoma, immune checkpoint inhibitors

**Posted Date:** November 9th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-1026831/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

**Background:** Thyroid cancer (TC) is the most common endocrine malignancy worldwide. The incidence of TC is high and increasing worldwide due to continuous improvements in diagnostic technology. TC is still often overtreated due to a lack of reliable diagnostic biomarkers. Therefore, determining accurate prognostic predictions to stratify TC patients is important.

**Methods:** Raw data were downloaded from the TCGA database, and pairwise comparisons were applied to identify differentially expressed immune-related lncRNA (DEirlncRNA) pairs. Then, we used univariate Cox regression analysis and a modified Lasso algorithm on these pairs to construct a risk assessment model for TC. Next, TC patients were assigned to high- and low-risk groups based on the optimal cutoff score of the model for the 1-year ROC curve. We evaluated the signature in terms of prognostic independence, predictive value, immune cell infiltration, ICI-related molecules and small-molecule inhibitor efficacy.

**Results:** We identified 30 DEirlncRNA pairs through Lasso regression, and 14 pairs served as the novel predictive signature. The high-risk group had a significantly poorer prognosis than the low-risk group. Cox regression analysis revealed that this immune-related signature can predict prognosis independently and reliably for TC. With the CIBERSORT algorithm, we found an association between the signature and immune cell infiltration. Additionally, several immune checkpoint inhibitor (ICI)-related molecules, such as PD-1 and PD-L1, showed a negative correlation with the high-risk group. We further found that some commonly used small-molecule inhibitors, such as sunitinib, were related to this new signature.

**Conclusions:** We constructed a prognostic immune-related lncRNA signature that can predict TC patient survival without considering the technical bias of different platforms, and this signature also sheds light on TC overall prognosis and novel clinical treatments, such as ICB therapy and small molecular inhibitors.

## Introduction

In recent decades, the incidence of thyroid cancer (TC) has been increasing steadily over the past 30 years worldwide[1]. With the development of diagnostic technology, the detection rate of TC in the world has increased year by year, ranking fifth in the incidence of malignant tumors[2], and the growing incidence of TC is raising serious public health issues worldwide. Even though most TCs have a relatively good prognosis, approximately 10% of patients with differentiated thyroid cancer (DTC) may progress to invasive disease, 5% progress to distant metastasis, and approximately 20–30% may relapse[3].

The immune system is now considered to have an important role in the elimination of cancer cells and sheds light on the mechanisms of cancer–immune evasion, contributing to tumor outgrowth[4, 5]. Although immunosurveillance prevents the development of most tumors in normal individuals, cancer cells can still deploy escape strategies such as initiating immune checkpoints[6]. Immune checkpoint inhibitors (ICIs) can restore the normal work of the immune system by suppressing the signal of "rest

from work", which is sent by immune checkpoints, and then launch an attack on tumor cells[7]. Immune checkpoint blockade (ICB) immunotherapy has shown durable responses and improved clinical outcomes for patients across most malignancies, including thyroid cancer. Growing evidence has proven that ICI-related genes such as PD-1 and PD-L1 are expressed in ATC and DTC[8]. Pembrolizumab, an anti-PD-1 antibody, may be effective for patients with TC because of the KEYNOTE-158 trial, which shows that approximately 60% of patients in this program achieved disease control [9].

In addition, several studies have demonstrated that tyrosine kinase inhibitors (TKIs) can prevent the proliferation and tumorigenicity of thyroid cancer cells. Treatment with small-molecule tyrosine kinase inhibitors, including gefitinib, pazopanib, lenvatinib, and axitinib, are approved by the FDA to treat thyroid cancer[10]. Some clinical trials were performed to evaluate the efficacy and safety of these inhibitors [11]. Nevertheless, TKIs are promising new agents for the treatment of patients with thyroid cancer.

Long noncoding ribonucleic acids (lncRNAs) belong to the family of noncoding RNAs and are functionally defined as transcripts >200 nucleotides in length with no protein-coding potential[12]. In the past few years, lncRNAs have attracted much attention due to their previously underappreciated transcriptional regulation, and they can function as both oncogenes and tumor suppressors[13]. It is clear that lncRNAs play crucial roles in the regulation of various biological and pathological behaviors of malignant tumors, especially in tumorigenesis and progression[14, 15].

Moreover, a growing body of studies has shown that the expression of lncRNAs is linked to the immune response and tumor progression. In addition, lncRNAs have an essential role in the development of immune cells and in pathogen response pathways. Notably, individual lncRNAs can act functionally through modular domains and often link protein activity to DNA or RNA targets through interactions with both [16].

Aberrant lncRNA expression has been observed in various cancers [17]. Recently, an increasing number of lncRNAs have been discovered to play an important role in TC tumorigenesis and development. For example, a study showed that the expression level of lncRNA n384546 was upregulated in TC patients and that its interference inhibited cancer cell proliferation, invasion, and migration[18]. Reliable prognostic models related to tumor immune infiltration may affect the diagnosis, evaluation and treatment decisions of tumors. LncRNAs have recently been used to establish prognostic signatures. Huang et al. [19] constructed an immune-related lncRNA signature to predict the survival outcome of patients with breast cancer. Xu et al [20] built a signature based on seven immune-related lncRNAs, which showed reliable prognostic value in hepatocellular carcinoma and may predict the outcome of immune checkpoint blockade (ICB) therapy.

In this study, we established a novel method to construct a predictive signature for TC based on immune-related lncRNA pairs, which did not need to consider the technical bias of different platforms. Next, to explore the latent role of this signature, we integrated the immune model with clinical factors of TC

patients to build a composite prognostic index, which allowed improved estimation of prognosis. We also explored the correlation between the signature and several aspects, such as immune cell infiltration, immunosuppressed biomarkers and small-molecule tyrosine kinase inhibitor efficacy. Overall, our new model might provide insight into the prognosis and clinical treatment of TC.

## Materials And Methods

### Data Download and Preprocessing and Differential Expression Analysis

The gene expression data and clinical data of TC patients were downloaded from The Cancer Genome Atlas (TCGA, <http://cancergenome.nih.gov/>) to perform comprehensive analysis. The RNA-seq profiles of 558 cases comprised 500 thyroid carcinoma samples and 58 normal thyroid samples. We also included clinical data of 505 TC patients with an overall survival time of 0 days. Next, the data were annotated to distinguish the mRNAs and lncRNAs for further analysis. We retrieved 2483 immune-related genes (IRGs) from the ImmPort database (<https://immport.niaid.nih.gov>) and performed coexpression analysis to identify the association between immune-modulating genes and all lncRNAs. lncRNAs with a correlation coefficient  $|R| \geq 0.4$  and  $p \leq 0.001$  were considered related to immune genes and were used for further analysis. Screening of the DEirlncRNAs was based on  $\log_{FC} > 1$  and  $FDR < 0.05$  using the limma R package[21].

### Pairing Immune-related lncRNAs

We performed cyclical pairwise comparisons between the DEirlncRNA expression values to generate a score for each irlncRNA pair and established a 0-or-1 matrix. If the expression level of lncRNA1 was higher than that of lncRNA2, this pair was assigned a score of 1; otherwise, the score was 0. To prevent biases and unrepeatability, we validated the expression quantities of the lncRNA pairs. After removing lncRNA pairs scoring 0 or 1 in more than 80% and less than 80% of the total pairs, the remaining pairs were considered valid matches to build the prognostic signature.

### Construction of a Novel Signature based on the DEirlncRNA Pairs

Univariate Cox regression analysis ( $p < 0.001$ ) was first implemented to identify the prognostic irlncRNA pairs with TC by using R (survival package). After that, least absolute shrink age and selection operator (LASSO) Cox regression was conducted by utilizing the Glmnet R package to reduce the number of pairs[22]. Lasso regression was carried out to acquire a well-balanced prognostic model by running 1000 cycles. Then, 14 pairs of immune-related lncRNAs were ultimately identified as our prognostic model, and the risk score of each TC patient was calculated based on the following formula:  $RiskScore = \sum_{i=1}^n \beta_i P_i$ . Here,  $n$  was the number of selected lncRNA pair,  $\beta_i$  was the coefficient of pair  $i$ , and  $P_i$  was the expression of lncRNA pair  $i$ . To validate the sensitivity and accuracy of the model, time-dependent ROC curve analysis was performed by utilizing the SurvivalROC package in R. The AUC was calculated for ROC curves, and sensitivity and specificity were calculated to assess the performance of the risk score. The optimal cutoff

score based on the 1-year ROC curve was identified to separate the patients into low-risk and high-risk groups.

### **Validation of the Risk Assessment Model**

We estimated the prognostic capability of the risk score, sex, age and tumor stage (TNM stage and clinical stage) for overall survival in terms of time-dependent AUC values. The survival difference between subgroups was evaluated by Kaplan-Meier survival analysis. The survival curve was visualized using the survival and survminer packages in R. The specific risk score value of each sample based on the signature was also used for visualization in R. We further performed principal component analysis (PCA) to assess the accuracy of the classification according to different risk scores. To predict the reliability and stability of the constructed model, Pearson's  $\chi^2$  test was used to assess the association between the signature and clinical features. The risk score differences between groups of conventional clinical features were compared by the Wilcoxon signed rank test. The analysis results were visualized by a strip chart and box diagrams. Moreover, we used univariate and multivariate Cox proportional hazards analyses of the risk score and other clinical characteristics to examine whether the model could be used as an independent variable.

### **Correlation Between the Risk Model and Immune Cell Infiltration**

The CIBERSORT algorithm was performed to estimate the tumor-infiltrating immune cell profiles of the samples in the TC dataset, followed by quality filtering, and only 509 tumor samples with  $p < 0.05$  were selected for the following analysis. To further explore whether the risk model was related to immune cell infiltration in TC, other methods, such as TIMER, XCELL, QUANTISEQ, MCP-counter, EPIC, and CIBERSORT-ABS, were also used. We applied the Wilcoxon signed-rank test to compare the tumor-infiltrating immune cells among different risk groups. A box chart was used to visualize the results. To assess the relationship between the model and the infiltrated immune cells, Spearman correlation analysis was performed. The results of the Spearman rank correlation coefficients are shown in the lollipop-style mutation diagram. We visualized the results using the ggplot2 package in R

### **Correlation Between the Risk Model and the Molecules related to ICIs**

The limma package was applied to explore the expression of ICI-related molecules among the different risk groups, and the results are shown in violin plots by using the ggpubr package in R. We also evaluated the immune score between low-risk cases and high-risk cases.

### **Analyses of the Ability of the Model in Clinical Treatment**

Small-molecule inhibitors such as gefitinib, pazopanib, lenvatinib, axitinib, AMG-706 (also known as motesanib), tipifarnib, sunitinib and sorafenib, which are now approved by the FDA, have shown great curative effects in thyroid cancer. To evaluate the ability of the signature in thyroid cancer treatment, we calculated the IC50 of these drugs in the THCA dataset. We performed the Wilcoxon signed-rank test to

analyze the differences among groups, and the results were displayed through box plots by utilizing the pRRopheti, ggpubr and ggplot2 R packages.

## Results

### Identification of Differentially Expressed irlncRNAs (DEirlncRNAs)

This study was carried out based on the flowchart shown in Figure 1A. First, RNA sequencing (RNA-seq) transcriptome data of TC were downloaded from TCGA database. A total of 505 available TC patients were included in our study, and the baseline characteristics of all the patients are shown in Table 1. A total of 1148 irlncRNAs were identified by performing a coexpression analysis (shown in Additional file 1: TableS1). Next, we set the threshold as log fold change (FC) > 1.0 or < -1.0 and false discovery rate (FDR) < 0.05 and identified 200 DEirlncRNAs between tumor and normal thyroid tissues (Figure 1B), including 117 upregulated and 83 downregulated genes (Figure 1C).

### Construction of a Novel Prognostic Signature Based on DEirlncRNA Pairs for TC

First, a total of 200 DEirlncRNAs were screened by pairwise comparison, and 13349 were constructed. With single factor Cox regression analysis, prognosis-related DE-irlncRNA pairs were further selected. Then, lasso-penalized Cox analysis was conducted to narrow the number of lncRNA pairs to 14 over 1000 repetitions (Figure 2A, B). Next, we constructed a risk assessment model that contains these 14 DEirlncRNA pairs using a risk score method and calculated the risk score of each sample. We performed receiver operating characteristic (ROC) analysis at 1 year (area under the curve (AUC)= 0.973) and identified the optimal cutoff point at 0.868 to divide 497 patients into high- and low-immune risk groups (Figure 2C, D). We further investigated the prognostic value of the signature by plotting the distributions of the risk score and survival time (Figure 2E, F). Low-risk TC patients exhibited a superior clinical outcome compared with high-risk TC patients. Then, 3-, 5-, and 10-year ROC curves were plotted and showed that the model had the ability to predict the survival outcome of TC patients with high accuracy and sensitivity (Figure 2G).

### Clinical Assessment and Evaluation of the Signature

Based on the validated optimal cutoff point, 45 TC patients with higher risk scores were included in the high-risk group, and 452 TC patients with lower risk scores were included in the low-risk group. The PCA results proved that this kind of classification was reliable (Figure 2H). Kaplan-Meier survival analysis showed that low-risk TC patients exhibited a significantly better prognosis than high-risk TC patients ( $p < 0.001$ , Figure 2I). Furthermore, we assigned THCA patients into different subgroups according to clinical characteristics, including sex (male and female), N stage (N0 and N1), stage (stage I + II and stage III + IV) and T stage (T1 + T2 and T3 + T4). KM analysis also showed that high-risk patients shared worse overall survival than low-risk patients, which indicated that the prediction of the signature was great ( $p < 0.001$ , Figure 3).

## The Signature as an Independent Prognostic Predictor

A series of chi-square tests were performed to explore the relationship between the risk of TC and common clinical features, including age, sex, TNM stage and clinical stage. The strip chart (Figure 4A) and scatter diagrams showed that age, N stage, clinical stage and survival status (Additional file 2 ) were significantly related to the risk of TC. Furthermore, we compared the AUCs of the 1-year ROC curves of the risk score to those of other clinical features and found that the risk score achieved a higher AUC value among these factors (Figure 4B). Next, univariate and multivariate Cox regression analyses were performed to validate prognostic factors associated with TC. Univariate analysis showed that age ( $p < 0.001$ , hazard ratio (HR) = 1.153, 95% confidence interval (CI) [1.095~1.213]), stage ( $p < 0.001$ , HR = 2.929, 95% CI [1.896–4.527]), T stage ( $p=0.005$ , HR = 2.384, 95% CI [1.291–4.405]), and risk score ( $p < 0.001$ , HR = 4.714, 95% CI [3.015–7.371]) were significantly associated with prognosis, whereas only age ( $p=0.040$ , HR = 1.061, 95% CI [1.003–1.122]) and risk score ( $p=0.005$ , HR = 3.409, 95% CI [1.750–6.641]) could independently predict the clinical outcome of TC patients by multivariate Cox regression analysis (Figure 4C, D, Table 2).

## Correlation of the Risk Assessment Model with Immune Cell Infiltration in TC

The TME is essential for cancer prognosis and is composed of tumor cells, stromal cells, and infiltrating immune cells. We next investigated whether the signature was associated with the TME in TC patients. The proportion of tumor-infiltrating immune subsets was determined based on the CIBERSORT algorithm, and 21 kinds of immune cell profiles were detected in the TC samples (Figure 5A, B). We found that the risk score had a significant relationship with the infiltration levels of immune cells. Correlation analysis was conducted using the Spearman correlation test, and the results are displayed in a lollipop diagram (Figure 6A). The results are listed in Additional file 3: TableS2. The high-risk samples showed a negative association with tumor-infiltrating immune cells, such as CD8+ T cells, CD4+ T cells, neutrophils, B cells, and M1 macrophages, compared to the low-risk samples, whereas they were positively associated with myeloid dendritic cells, plasma B cells, monocytes and cancer-associated fibroblasts, as analyzed by the Wilcoxon signed-rank test (Figure 6B, Additional file 4 )

## Correlation of the Signature with genes related to ICIs in TC

Immune checkpoint blockade has revolutionized cancer treatment, and ICB therapy has already been applied in thyroid cancer patients. Therefore, we employed seven key ICI-related genes, PDCD1 (also called PD-1), CD274 (also called PD-L1), PDCD1LG2 (also called PD-L2), CTLA-4, LAG3, CD74 and IDO1, to investigate whether these ICI-related molecules are related to our new signature. The results indicated that the high-risk score group showed a negative correlation with the expression of PD-1 ( $p < 0.01$ , Figure 7A), PD-L1 ( $p < 0.001$ , Figure 7B), LAG3 ( $p < 0.01$ , Figure 7C), CTLA-4 ( $p < 0.001$ , Figure 7E), PD-L2 ( $p < 0.05$ , Figure 7F) and CD74 ( $p < 0.001$ , Figure 7G), whereas IDO1 ( $p>0.05$ , Figure 7D) showed no significant difference. We also analyzed the difference in immune score between the high- and low-risk groups and found that the low-risk group had a higher immune score than the high-risk group ( $p<0.001$ , Figure 7H).

## The Signature as an indicator in TC Treatment

As a personalized medical treatment, small-molecule inhibitors have been widely applied in the clinical treatment of thyroid cancer. We next investigated the relationship between the signature and the half maximal inhibitory concentration (IC50) of some common inhibitors in the THCA dataset. As Figure 8 shows, the high-risk samples were positively related to the IC50 of gefitinib ( $p < 0.01$ , Figure 8A), sunitinib ( $p < 0.001$ , Figure 8B), and tipifarnib ( $p < 0.001$ , Figure 8C) but negatively related to axitinib ( $p < 0.001$ , Figure 8E), AMG-706 ( $p < 0.001$ , Figure 8F), and pazopanib ( $p < 0.01$ , Figure 8F). In addition, lenvatinib and sorafenib might have no relationship with the signature ( $p > 0.05$ , Figure 8D, H).

## Discussion

Thyroid cancer is the most common type of cancer of the endocrine system, and its incidence has increased almost 3-fold over the past decades. [23] Traditional clinical characteristics, such as TNM stage, can be used to predict the severity related to TC, but it is difficult to accurately estimate the risk of recurrence. [24] Thus, it is imperative to establish powerful tools that can be effectively applied to aid in the diagnosis, prognosis, and treatment of patients with TC.

Accumulating evidence shows that since lncRNAs have high tissue and cancer specificity, they might play an active role in cancer initiation, development and progression. An increasing number of studies have shown that lncRNAs promote tumor cell proliferation, invasion, metastasis, and angiogenesis and can serve as an excellent tool to modulate therapeutic decisions in cancer. Increasing evidence suggests that lncRNAs are involved in TC tumorigenesis and progression as important regulatory factors [25], thus, lncRNAs have attracted much attention as potential targets in the diagnosis and treatment evaluation of TC. For example, Liu et al. found that MALAT1 may have an oncogenic function in PTC and may thus be a potential diagnostic marker for PTC [26]. In our research, a few of the DE lncRNAs in the model, such as SMIM25, [27] LINC00900, [28] AC010980.2, [29] HAGLROS, and LINC01977 [30], have already been revealed to play roles in various cancers, especially TC, while others were found to be related to TC for the first time. Nevertheless, it is necessary to validate whether this immune-related lncRNA model could be a helpful predictive indicator in TC. Many researchers are currently focused on constructing signatures with both coding genes and noncoding RNAs, which can assess the survival status of patients with malignant carcinoma [31, 32]. Unlike most traditional risk models, our newly constructed signature involved two-lncRNA pairwise comparisons and relative ranking on the basis of gene expression entirely from the same TC patient. Although from different sequencing platforms, our prognostic model does not require gene expression data normalization. Previous studies have supported the effectiveness of this method [33, 34].

In the current study, we established an immune-related lncRNA model and evaluated its prognostic value as well as its correlation with immune cell infiltration, ICI-related genes and TKIs in TC. First, we performed differential coexpression analysis to identify DE lncRNAs based on data from TCGA. lncRNA pairs were systematically identified through pairwise comparisons in the same sample without

the need for data normalization. In addition, univariate analysis with Lasso regression analysis was performed on the pairs to validate the most suitable variables. Fourteen significant DE lncRNA pairs with maximum prognostic values were determined with multiple repeats and random stimulation. Next, we used these pairs to develop the predictive risk score model. Then, we calculated not only the 3-, 5-, and 10-year AUC values of the ROC curve but also identified the optimal cutoff point of the 1-year ROC curve to separate TC patients into high- and low-risk groups. Furthermore, Kaplan-Meier curves, time-dependent ROC curves, and Cox proportional hazards regression analysis showed that the model has independent predictive value for TC prognosis. Finally, we evaluated the relationship between this novel model and tumor-infiltrating immune cells, ICI-related molecules and small-molecule inhibitor validity.

Immune infiltrates in the tumor microenvironment (TME) play a vital role in tumor development and progression and affect the clinical outcomes of cancer patients. [35] Dysfunction of the immune status in the TME contributes to the development and progression of cancer, and this is the basis of many immunotherapy studies. Moreover, tumor immunotherapy is now considered to have an important role in the elimination of cancer cells and sheds light on the mechanisms of cancer-immune evasion, contributing to tumor outgrowth. Recent studies have suggested that lncRNAs play a central role in innate and adaptive cancer immunity regulation. [36] Immune-related lncRNAs as signatures are better at predicting prognosis than single lncRNAs. Therefore, it is necessary to explore more immune-related lncRNAs in tumors for future clinical practice. In this research, we carried out pairwise comparisons of a given set of immune-related lncRNAs and expression values. Thus, our prognostic signature could help address batch effects between different platforms and overcome the reprocessing and normalization of data.

Immune cell infiltration reflects the TME and reportedly impacts the outcome of TC progression. It is evident that immune-related lncRNAs are correlated with the development of TC. To explore the relationship between the prognostic model and immune-infiltrating cells, we applied seven commonly accepted methods, including TIMER[37, 38], CIBERSORT[39], XCELL[40], QUANTISEQ [41], MCP-counter, [42] EPIC [43] and CIBERSORT-ABS [44]. By integrating analyses, we found that the levels of Tregs, myeloid dendritic cells, monocytes, cancer-associated fibroblasts and B cell plasma in the high-risk group were higher than those in the low-risk group, while the levels of neutrophils, M1 macrophages, CD8+, CD4+ T cells, and B cells were significantly negatively correlated with the risk of signature. These results were consistent with the findings of some previous experimental studies, which aimed to determine the correlation between each cell type and the aggressiveness of TC[45-47]. For example, it has been reported that neutrophils play an antitumor role and can be beneficial to the prognosis of TC[48], which is consistent with the findings of our analysis. Our results also revealed that the abundance of Tregs was more associated with the high-risk group, which is similar to the findings of previous literature. Those studies found that the levels of Tregs in PTC were higher than those in multinodular goiter patients, and Tregs were consistently present in extraglandular invasion and lymph node metastasis[49, 50]. Monocytes have been observed to promote the occurrence and development of tumors, and their high density is closely related to thyroid tumor invasion and reduced survival, which

was also confirmed by our study[51]. These observations can be further explored for a holistic understanding of the nuances of TC microenvironment immune cell infiltration.

Tyrosine kinase inhibitors (TKIs) are an innovative personalized strategy that aim at pro-oncogenic kinases, including EGFR, MET, PDGFR, VEGFR-1, VEGFR-2, RAF, FGFR and RET. Our signature showed that the risk score was related to some of these inhibitors, such as gefitinib, sunitinib, and tipifarnib, indicating that this new model might be a new method in the assessment of efficacy to systemic therapy based on a genetic understanding in TC. Moreover, ICB immunotherapy is viewed as a promising cancer therapeutic modality for malignant tumors. The identification of PD-L1 as an immunostat blockade has led to the development of a number of cancer immunotherapies. For RAI refractory PTC patients, recent evidence has shown that overexpression of PD-L1 together with lymphocyte infiltration into the tumor TME are significantly associated with the effectiveness of ICB[52, 53]. In this study, we found that a high-risk score was negatively correlated with ICB-related genes such as PD-1, PD-L1, LAG3, CTLA-4, PD-L2 and CD74, which are commonly expressed in human cancer. In addition, some clinical trials have applied PD-1 and PD-L1 inhibitors in combination with TKIs, RAIs or chemotherapy for managing and defeating deadly TC. Our signature may provide new insight to predict which patients are more suitable for these treatments, either alone or in combination.

To the best of our knowledge, a prognostic model based on irlncRNA pairs in TC has not been reported to date. Our predictive model is based on a 0-or-1 matrix and could be applied in an individualized manner while eliminating batch bias. In addition, our signature first combined DElncRNA pairs with ICB and TKI efficacy for analysis. Various additional methods were used to support the prognostic value and feasibility of this new model.

However, this lncRNA-based prognostic signature had several limitations. First, the establishment and validation of the model was based only on the TCGA database, which might lead to selection bias. To verify the predictive values of the risk assessment model, a larger dataset and external datasets of TC should be analyzed. Second, this was a retrospectively designed study, and a prospective cohort needs to be established in future work for further verification. In addition, the calculation formulas of this prognostic signature may be too complex for clinical application.

## Conclusion

In conclusion, we constructed an immune-related lncRNA pair model without considering the technical bias of different platforms. As such, our novel model could be used without the need to eliminate batch effects and served as an independent single-sample estimate of the survival risk subgroup of TC patients. The model may provide new possibilities for translation to clinical practice for TC patients and help in distinguishing those who could benefit from ICB immunotherapy and TKI therapy.

## Declarations

## **Ethics approval and consent to participate**

TCGA belong to public databases. The patients involved in the database obtained ethical approval. Users can download relevant data for free for research and publish relevant articles. Thus, no ethical approval and patient consent were required.

## **Consent for publication**

Not applicable.

## **Competing interests**

The authors declare that they have no competing interests.

## **Availability of data and materials.**

The generated and analyzed datasets of the current research are available in TCGA database (<http://cancergenome.nih.gov/>), and the immune-related genes were obtained from the ImmPort database (<https://import.niaid.nih.gov>).

## **Funding**

This research was funded by The National Natural Science Foundation of China, Grant Number 81800700) and The Clinical Research Fund of Chinese Medical Association (Grant No. 15010010589).

## **Author Contributions**

BS is the first author of this study. WT is the corresponding author supervising this work. BS designed the study and wrote the first manuscript; BG performed the statistical analysis; WT reviewed the manuscript; LT, FZ and ZL wrote sections of the manuscript. TL performed the project administration. All authors read and approved the final manuscript.

## **Acknowledgments**

We acknowledge TCGA database for providing their platforms and contributors for uploading their meaningful datasets.

## **Author informationn**

<sup>1</sup>Department of Endocrinology and Metabolism, Institute of Endocrinology, National Health Commission Key Laboratory of Diagnosis and Treatment of Thyroid Diseases, The First Hospital of China Medical University, Shenyang, China

## **References**

1. Vaccarella S, Dal Maso L, Laversanne M, Bray F, Plummer M, Franceschi S: **The Impact of Diagnostic Changes on the Rise in Thyroid Cancer Incidence: A Population-Based Study in Selected High-Resource Countries.** *Thyroid* 2015, **25**:1127-1136.
2. Rahib L, Smith B, Aizenberg R, Rosenzweig A, Fleshman J, Matrisian L: **Projecting cancer incidence and deaths to 2030: the unexpected burden of thyroid, liver, and pancreas cancers in the United States.** *Cancer research* 2014, **74**:2913-2921.
3. Chowdhury S, Veyhl J, Jessa F, Polyakova O, Alenzi A, MacMillan C, Ralhan R, Walfish P: **Programmed death-ligand 1 overexpression is a prognostic marker for aggressive papillary thyroid cancer and its variants.** *Oncotarget* 2016, **7**:32318-32328.
4. Schreiber R, Old L, Smyth M: **Cancer immunoediting: integrating immunity's roles in cancer suppression and promotion.** *Science (New York, NY)* 2011, **331**:1565-1570.
5. Grivennikov S, Greten F, Karin M: **Immunity, inflammation, and cancer.** *Cell* 2010, **140**:883-899.
6. Galdiero MR, Varricchi G, Marone G: **The immune network in thyroid cancer.** *Oncoimmunology* 2016, **5**:e1168556.
7. Chen DS, Mellman I: **Oncology meets immunology: the cancer-immunity cycle.** *Immunity* 2013, **39**:1-10.
8. Bastman J, Serracino H, Zhu Y, Koenig M, Mateescu V, Sams S, Davies K, Raeburn C, McIntyre R, Haugen B, French J: **Tumor-Infiltrating T Cells and the PD-1 Checkpoint Pathway in Advanced Differentiated and Anaplastic Thyroid Cancer.** *The Journal of clinical endocrinology and metabolism* 2016, **101**:2863-2873.
9. Mehnert J, Varga A, Brose M, Aggarwal R, Lin C, Prawira A, de Braud F, Tamura K, Doi T, Piha-Paul S, et al: **Safety and antitumor activity of the anti-PD-1 antibody pembrolizumab in patients with advanced, PD-L1-positive papillary or follicular thyroid cancer.** *BMC cancer* 2019, **19**:196.
10. Cabanillas M, Habra M: **Lenvatinib: Role in thyroid cancer and other solid tumors.** *Cancer treatment reviews* 2016, **42**:47-55.
11. Brose MS, Nutting CM, Jarzab B, Elisei R, Siena S, Bastholt L, de la Fouchardiere C, Pacini F, Paschke R, Shong YK, et al: **Sorafenib in radioactive iodine-refractory, locally advanced or metastatic differentiated thyroid cancer: a randomised, double-blind, phase 3 trial.** *Lancet* 2014, **384**:319-328.
12. Iyer MK, Niknafs YS, Malik R, Singhal U, Sahu A, Hosono Y, Barrette TR, Prensner JR, Evans JR, Zhao S, et al: **The landscape of long noncoding RNAs in the human transcriptome.** *Nat Genet* 2015, **47**:199-208.
13. Sahu A, Singhal U, Chinnaiyan A: **Long noncoding RNAs in cancer: from function to translation.** *Trends in cancer* 2015, **1**:93-109.

14. Fang Y, Fullwood M: **Roles, Functions, and Mechanisms of Long Non-coding RNAs in Cancer.***Genomics, proteomics & bioinformatics* 2016, **14**:42-54.
15. Mercer T, Dinger M, Mattick J: **Long non-coding RNAs: insights into functions.***Nature reviews Genetics* 2009, **10**:155-159.
16. Chen Y, Satpathy A, Chang H: **Gene regulation in the immune system by long noncoding RNAs.***Nature immunology* 2017, **18**:962-972.
17. Bhan A, Soleimani M, Mandal SS: **Long Noncoding RNA and Cancer: A New Paradigm.***Cancer Res* 2017, **77**:3965-3981.
18. Feng J, Zhou Q, Yi H, Ma S, Li D, Xu Y, Wang J, Yin S: **A novel lncRNA n384546 promotes thyroid papillary cancer progression and metastasis by acting as a competing endogenous RNA of miR-145-5p to regulate AKT3.***Cell death & disease* 2019, **10**:433.
19. Huang Z, Xiao C, Zhang F, Zhou Z, Yu L, Ye C, Huang W, Li N: **A Novel Framework to Predict Breast Cancer Prognosis Using Immune-Associated LncRNAs.***Frontiers in genetics* 2020, **11**:634195.
20. Xu Q, Wang Y, Huang W: **Identification of immune-related lncRNA signature for predicting immune checkpoint blockade and prognosis in hepatocellular carcinoma.***International immunopharmacology* 2021, **92**:107333.
21. Ritchie M, Phipson B, Wu D, Hu Y, Law C, Shi W, Smyth G: **limma powers differential expression analyses for RNA-sequencing and microarray studies.***Nucleic acids research* 2015, **43**:e47.
22. Friedman J, Hastie T, Tibshirani R: **Regularization Paths for Generalized Linear Models via Coordinate Descent.***J Stat Softw* 2010, **33**:1-22.
23. Fagin JA, Wells SA, Jr.: **Biologic and Clinical Perspectives on Thyroid Cancer.***N Engl J Med* 2016, **375**:2307.
24. Haugen BR, Alexander EK, Bible KC, Doherty GM, Mandel SJ, Nikiforov YE, Pacini F, Randolph GW, Sawka AM, Schlumberger M, et al: **2015 American Thyroid Association Management Guidelines for Adult Patients with Thyroid Nodules and Differentiated Thyroid Cancer: The American Thyroid Association Guidelines Task Force on Thyroid Nodules and Differentiated Thyroid Cancer.***Thyroid* 2016, **26**:1-133.
25. Kentwell J, Gundara J, Sidhu S: **Noncoding RNAs in endocrine malignancy.***The oncologist* 2014, **19**:483-491.
26. Liu J, Dong H, Yang Y, Qian Y, Liu J, Li Z, Guan H, Chen Z, Li C, Zhang K, et al: **Upregulation of long noncoding RNA MALAT1 in papillary thyroid cancer and its diagnostic value.***Future Oncol* 2018, **14**:3015-3022.

27. Subhash S, Kalmbach N, Wegner F, Petri S, Glomb T, Dittrich-Breiholz O, Huang C, Bali KK, Kunz WS, Samii A, et al: **Transcriptome-wide Profiling of Cerebral Cavemous Malformations Patients Reveal Important Long noncoding RNA molecular signatures.***Sci Rep* 2019, **9**:18203.
28. Liu J, Wang Y, Chu Y, Xu R, Zhang D, Wang X: **Identification of a TLR-Induced Four-lncRNA Signature as a Novel Prognostic Biomarker in Esophageal Carcinoma.***Front Cell Dev Biol* 2020, **8**:649.
29. Wu L, Wen Z, Song Y, Wang L: **A novel autophagy-related lncRNA survival model for lung adenocarcinoma.***J Cell Mol Med* 2021.
30. Guo K, Chen L, Wang Y, Qian K, Zheng X, Sun W, Sun T, Wu Y, Wang Z: **Long noncoding RNA RP11-547D24.1 regulates proliferation and migration in papillary thyroid carcinoma: Identification and validation of a novel long noncoding RNA through integrated analysis of TCGA database.***Cancer Med* 2019, **8**:3105-3119.
31. Zhu X, Tian X, Yu C, Shen C, Yan T, Hong J, Wang Z, Fang JY, Chen H: **A long non-coding RNA signature to improve prognosis prediction of gastric cancer.***Mol Cancer* 2016, **15**:60.
32. Gu J, Zhang X, Miao R, Xiang X, Fu Y, Zhang J, Liu C, Qu K: **Six-long non-coding RNA signature predicts recurrence-free survival in hepatocellular carcinoma.***World journal of gastroenterology* 2019, **25**:220-232.
33. Peng PL, Zhou XY, Yi GD, Chen PF, Wang F, Dong WG: **Identification of a novel gene pairs signature in the prognosis of gastric cancer.***Cancer Med* 2018, **7**:344-350.
34. 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**:1529-1537.
35. Zhang Y, Zhang Z: **The history and advances in cancer immunotherapy: understanding the characteristics of tumor-infiltrating immune cells and their therapeutic implications.***Cellular & molecular immunology* 2020, **17**:807-821.
36. Denaro N, Merlano M, Lo Nigro C: **Long noncoding RNAs as regulators of cancer immunity.***Molecular oncology* 2019, **13**:61-73.
37. Li T, Fan J, Wang B, Traugh N, Chen Q, Liu JS, Li B, Liu XS: **TIMER: A Web Server for Comprehensive Analysis of Tumor-Infiltrating Immune Cells.***Cancer Res* 2017, **77**:e108-e110.
38. Li T, Fu J, Zeng Z, Cohen D, Li J, Chen Q, Li B, Liu XS: **TIMER2.0 for analysis of tumor-infiltrating immune cells.***Nucleic Acids Res* 2020, **48**:W509-w514.
39. Chen B, Khodadoust MS, Liu CL, Newman AM, Alizadeh AA: **Profiling Tumor Infiltrating Immune Cells with CIBERSORT.***Methods Mol Biol* 2018, **1711**:243-259.

40. Aran D: **Cell-Type Enrichment Analysis of Bulk Transcriptomes Using xCell.***Methods Mol Biol* 2020, **2120**:263-276.
41. Finotello F, Mayer C, Plattner C, Laschober G, Rieder D, Hackl H, Krogsdam A, Loncova Z, Posch W, Wilflingseder D, et al: **Molecular and pharmacological modulators of the tumor immune contexture revealed by deconvolution of RNA-seq data.***Genome Med* 2019, **11**:34.
42. Dienstmann R, Villacampa G, Sveen A, Mason MJ, Niedzwiecki D, Nesbakken A, Moreno V, Warren RS, Lothe RA, Guinney J: **Relative contribution of clinicopathological variables, genomic markers, transcriptomic subtyping and microenvironment features for outcome prediction in stage II/III colorectal cancer.***Ann Oncol* 2019, **30**:1622-1629.
43. Racle J, de Jonge K, Baumgaertner P, Speiser DE, Gfeller D: **Simultaneous enumeration of cancer and immune cell types from bulk tumor gene expression data.***Elife* 2017, **6**.
44. Tamminga M, Hiltermann TJN, Schuurin E, Timens W, Fehrmann RS, Groen HJ: **Immune microenvironment composition in non-small cell lung cancer and its association with survival.***Clin Transl Immunology* 2020, **9**:e1142.
45. Yu H, Huang X, Liu X, Jin H, Zhang G, Zhang Q, Yu J: **Regulatory T cells and plasmacytoid dendritic cells contribute to the immune escape of papillary thyroid cancer coexisting with multinodular non-toxic goiter.***Endocrine* 2013, **44**:172-181.
46. Angell TE, Lechner MG, Jang JK, Correa AJ, LoPresti JS, Epstein AL: **BRAF V600E in papillary thyroid carcinoma is associated with increased programmed death ligand 1 expression and suppressive immune cell infiltration.***Thyroid* 2014, **24**:1385-1393.
47. Hilly O, Koren R, Raz R, Rath-Wolfson L, Mizrachi A, Hamzany Y, Bachar G, Shpitzer T: **The role of s100-positive dendritic cells in the prognosis of papillary thyroid carcinoma.***Am J Clin Pathol* 2013, **139**:87-92.
48. Mao M, Huang R, Zheng J, Liang H, Huang W, Liu J, Li J: **OGDHL closely associates with tumor microenvironment and can serve as a prognostic biomarker for papillary thyroid cancer.***Cancer medicine* 2021, **10**:728-736.
49. Liu Y, Yun X, Gao M, Yu Y, Li X: **Analysis of regulatory T cells frequency in peripheral blood and tumor tissues in papillary thyroid carcinoma with and without Hashimoto's thyroiditis.***Clin Transl Oncol* 2015, **17**:274-280.
50. French JD, Weber ZJ, Fretwell DL, Said S, Klopper JP, Haugen BR: **Tumor-associated lymphocytes and increased FoxP3+ regulatory T cell frequency correlate with more aggressive papillary thyroid cancer.***J Clin Endocrinol Metab* 2010, **95**:2325-2333.

51. Zhou C, Duan D, Liu S: **Predictive Value of a Prognostic Model Based on Lymphocyte-to-Monocyte Ratio Before Radioiodine Therapy for Recurrence of Papillary Thyroid Carcinoma.***Technol Cancer Res Treat* 2021, **20**:15330338211027910.
52. Cunha L, Marcello M, Morari E, Nonogaki S, Conte F, Gerhard R, Soares F, Vassallo J, Ward L: **Differentiated thyroid carcinomas may elude the immune system by B7H1 upregulation.***Endocrine-related cancer* 2013, **20**:103-110.
53. Shi R, Qu N, Luo T, Xiang J, Liao T, Sun G, Wang Y, Wang Y, Huang C, Ji Q: **Programmed Death-Ligand 1 Expression in Papillary Thyroid Cancer and Its Correlation with Clinicopathologic Factors and Recurrence.***Thyroid : official journal of the American Thyroid Association* 2017, **27**:537-545.

## Tables

**Table 1 Clinical characteristics of the thyroid cancer patients**

Characteristic	Type	n	Proportion (%)
Age	<55	339	67.26%
	>=55	165	32.74%
Gender	FEMALE	368	73.02%
	MALE	136	26.98%
Stage	Stage I	399	79.17%
	Stage II	77	15.28%
	Stage III	22	4.36%
	Stage IV	6	1.19%
T stage	T1	144	28.57%
	T2	167	33.13%
	T3	170	33.73%
	T4	23	4.57%
M stage	M0	496	98.41%
	M1	8	1.59%
N stage	N0	272	53.97%
	N1	232	46.03%

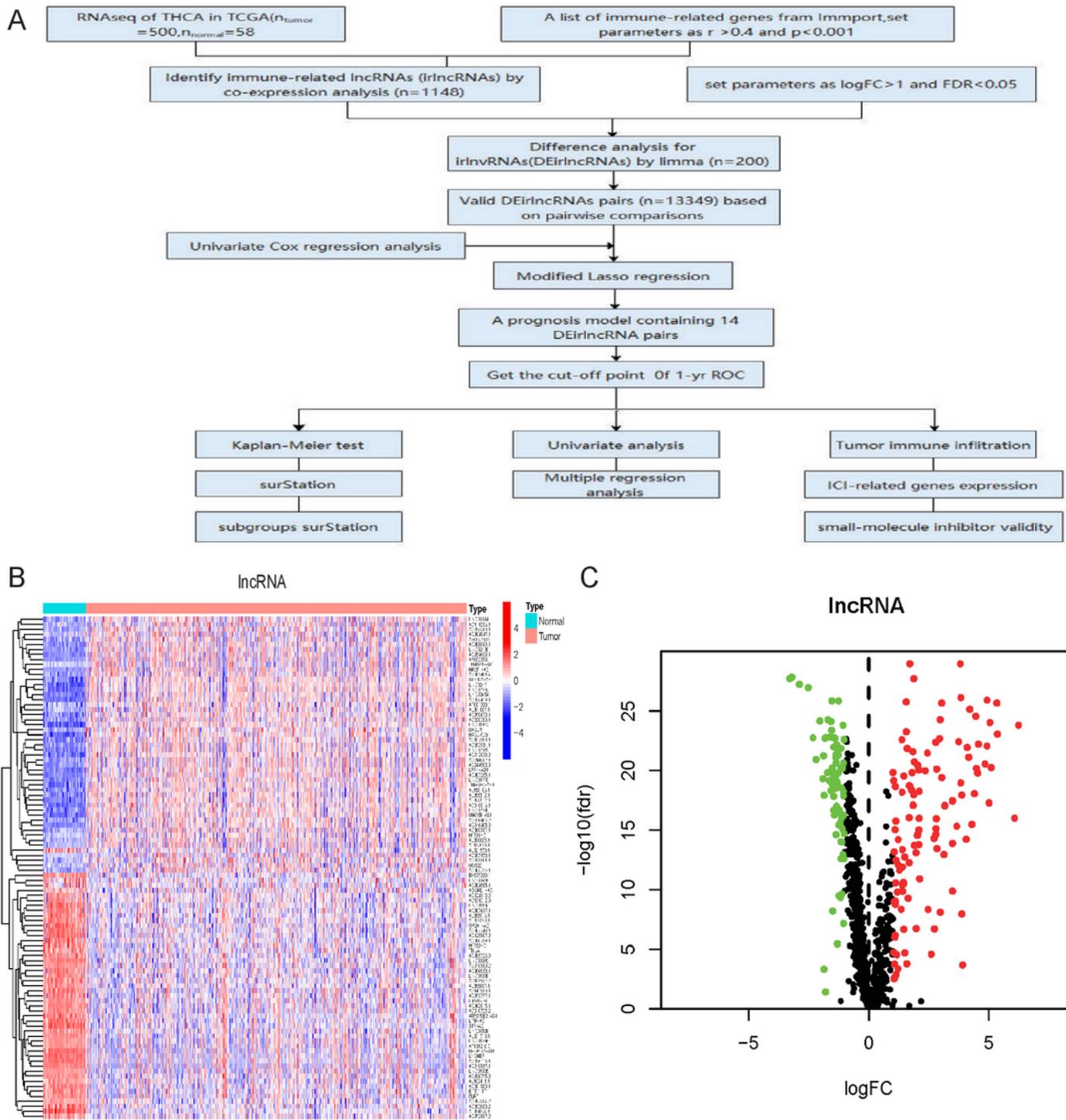
**Table 2 Independence of the signature for predicting the clinical outcomes of TC**

Id	Univariate analysis			Multivariate analysis		
	HR	95%CI	P	HR	95%CI	P
age	1.153	1.095~1.213	<0.001	1.061	1.003~1.122	0.040
gender	0.523	0.189~1.448	0.212	1.601	0.330~7.768	0.559
stage	2.929	1.896~4.527	<0.001	1.286	0.290~5.709	0.741
T	2.384	1.291~4.405	0.005	0.779	0.264~2.304	0.652
M	2.731	0.359~20.794	0.332	2.701	0.119~61.347	0.652
N	1.494	0.556~4.014	0.426	0.533	0.235~3.618	0.908
riskScore	4.714	3.015~7.371	<0.001	3.409	1.750~6.641	<0.001

(A): Univariate analysis illustrated the clinicopathological factors related to TC

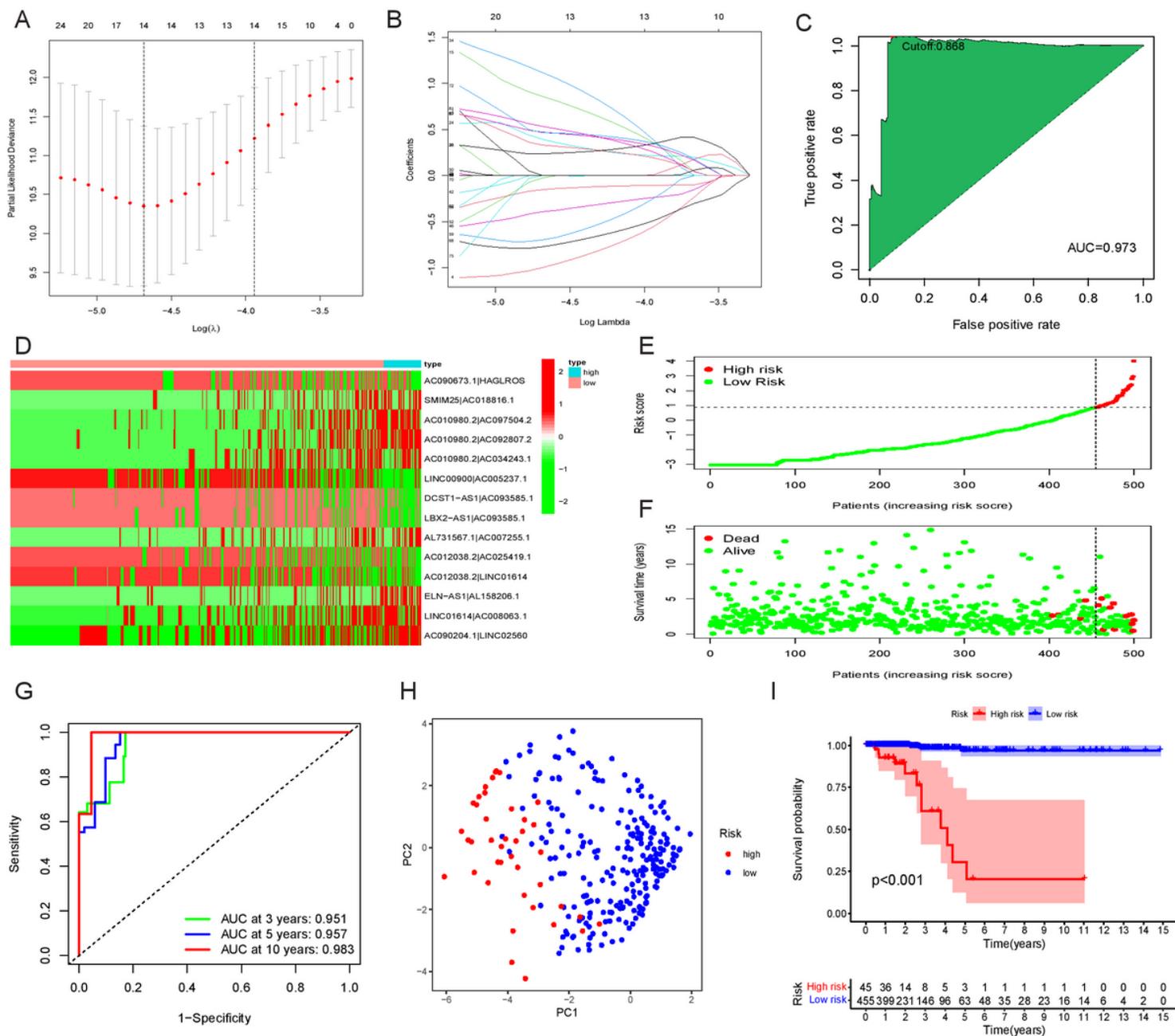
prognosis. (B): Multivariate analysis revealed the clinicopathological factors related to TC prognosis.

## Figures



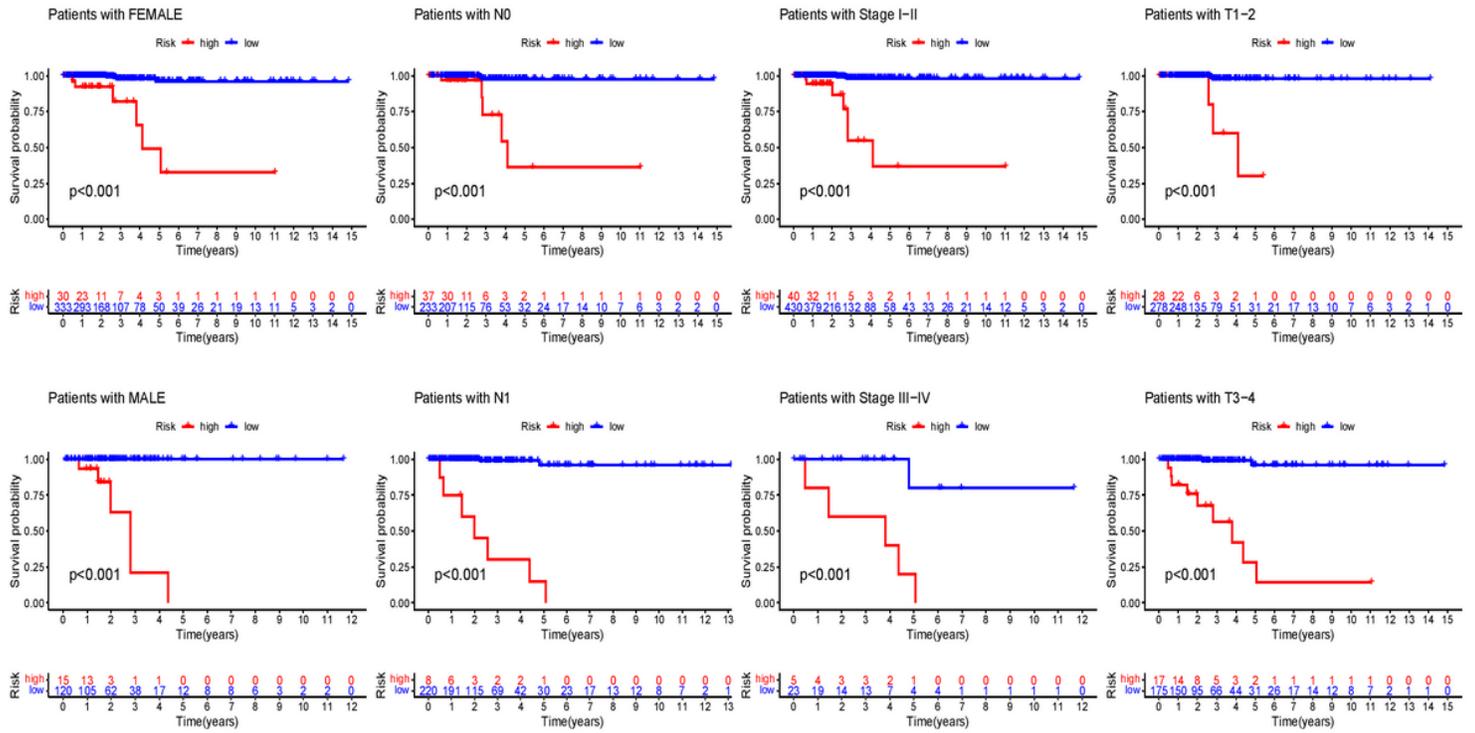
**Figure 1**

Identification of immune-related lncRNA in TC (A) Analysis workflow of this study. (B) Heat map shows the expression level of differentially expressed immune-related lncRNAs (DEirlncRNAs) in the TCGA dataset (blue, lower expression; red, higher expression). (C): Volcano plot represents the result of DEirlncRNA analysis of the THCA dataset (green: downregulated genes; black: no differentially expressed genes; red: upregulated genes)



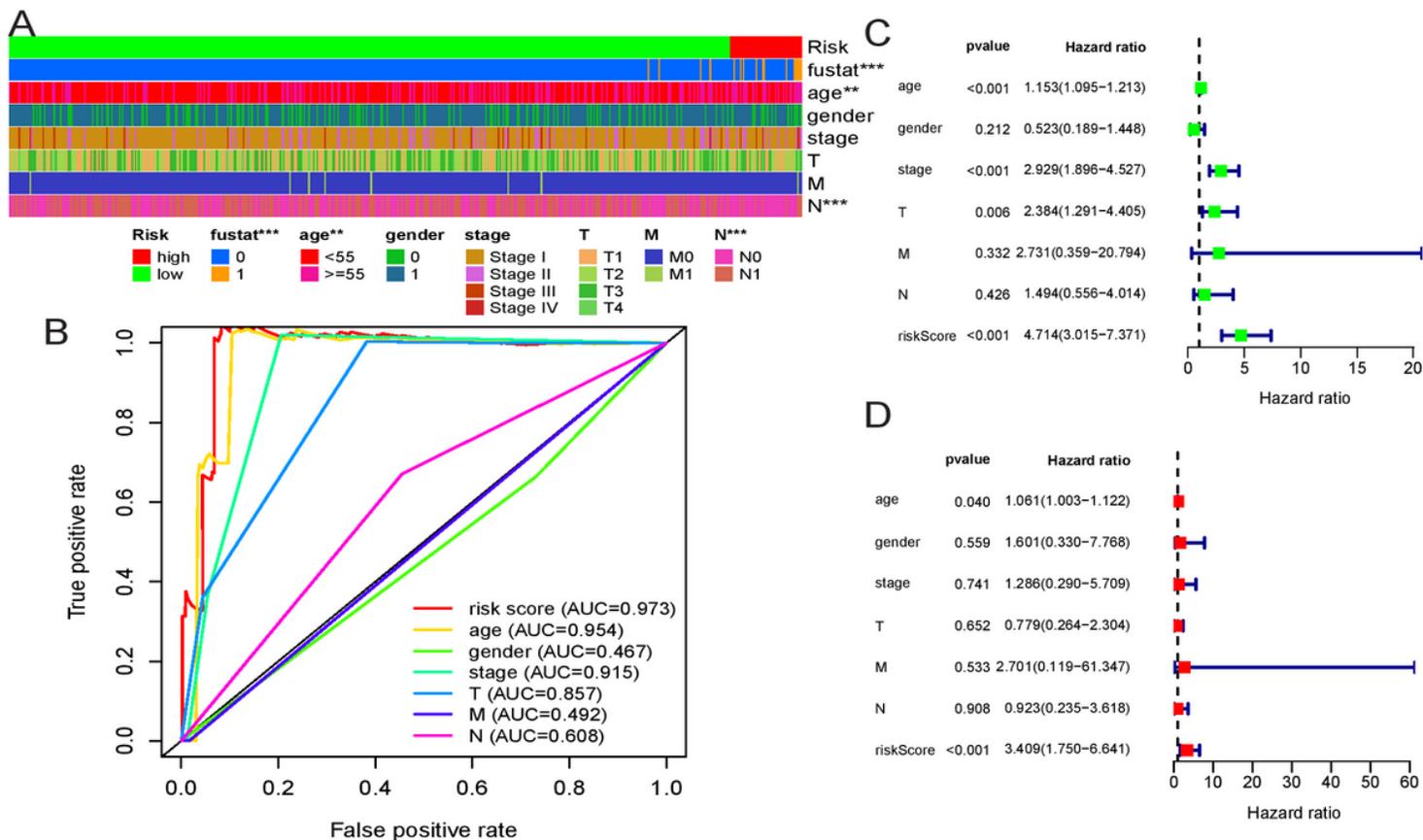
**Figure 2**

Construction of a risk assessment model and confirmation of the signature (A)(B): The results of lasso regression. (C): The cutoff value of the one-year ROC curve was used to separate the patients. (D): Fourteen DE lncRNA pairs were identified to construct the signature. (E): Distribution of lncRNA model risk score. (F): Survival status of TC patients in the subgroups. (G): ROC analysis of the risk scores for overall prognosis prediction. The 3- (red), 5- (green), and 10-year (blue) ROC curves of the model suggested that all AUC values were over 0.95 (H): PCA plot of the lncRNA model. (I): Kaplan-Meier curve presenting survival in the high-risk and low-risk sets.



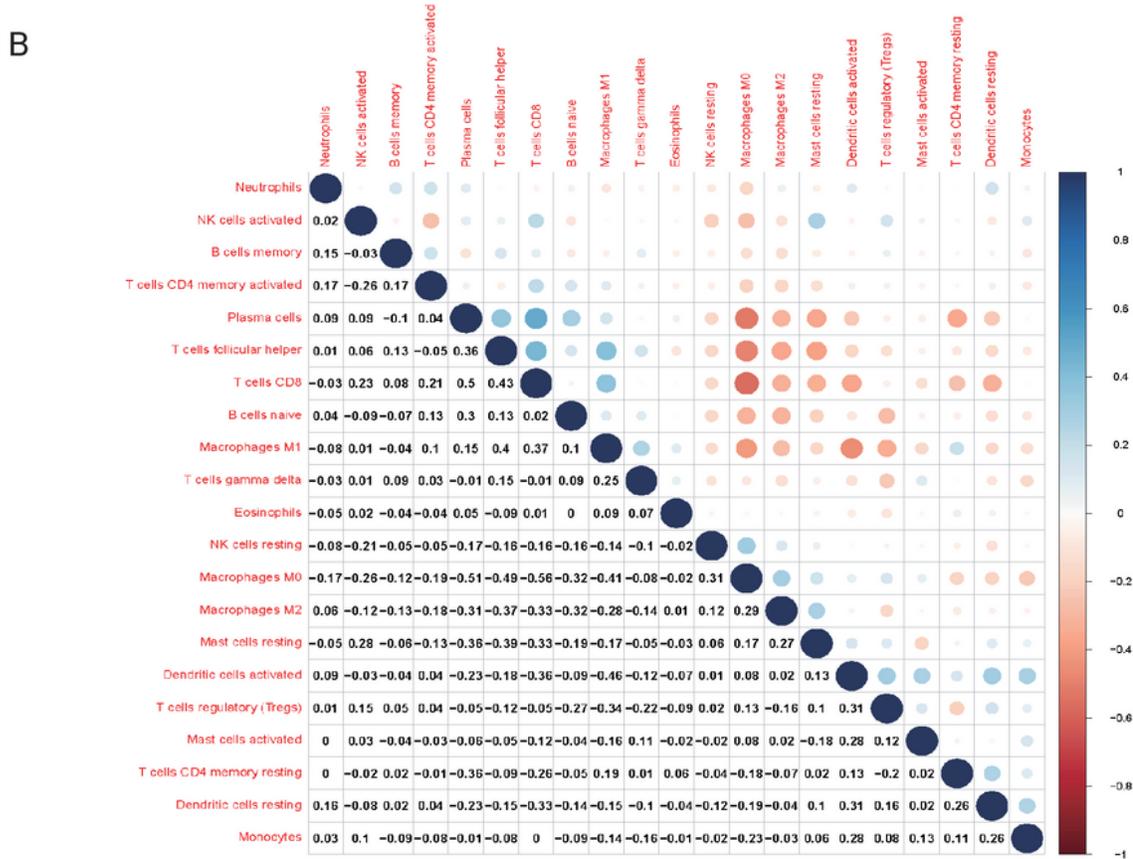
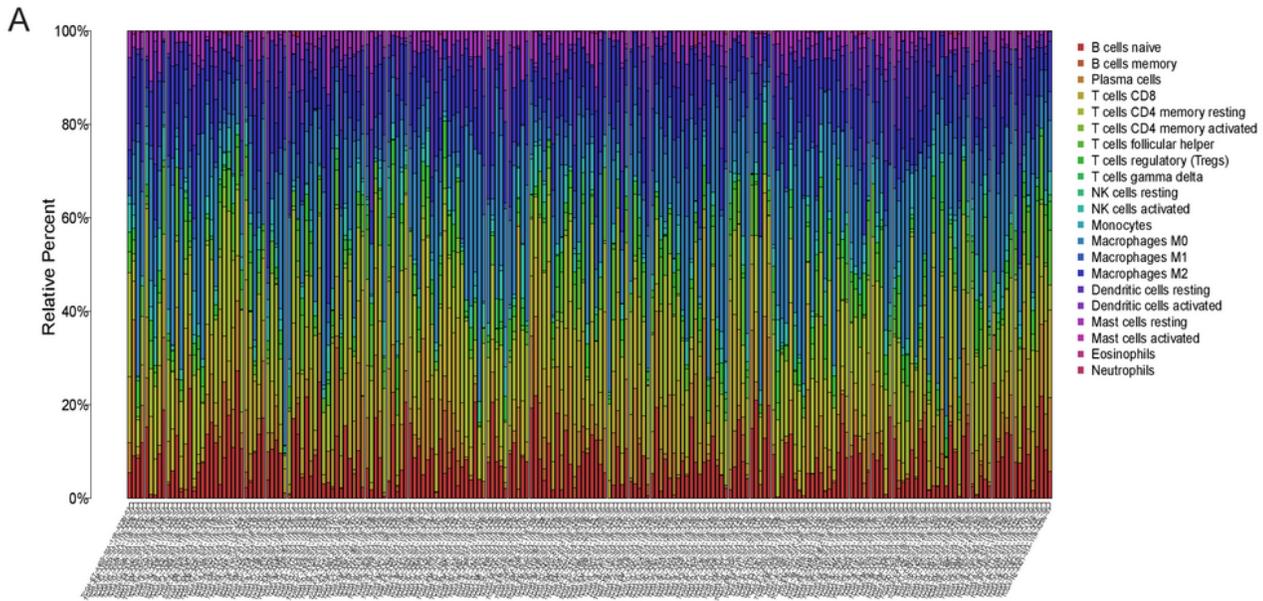
**Figure 3**

The prognostic value of the signature THCA patients were assigned into different subgroups, including sex (male and female), N stage (N0 and N1), stage (stage I + II and stage III + IV) and T stage (T1 + T2 and T3 + T4).



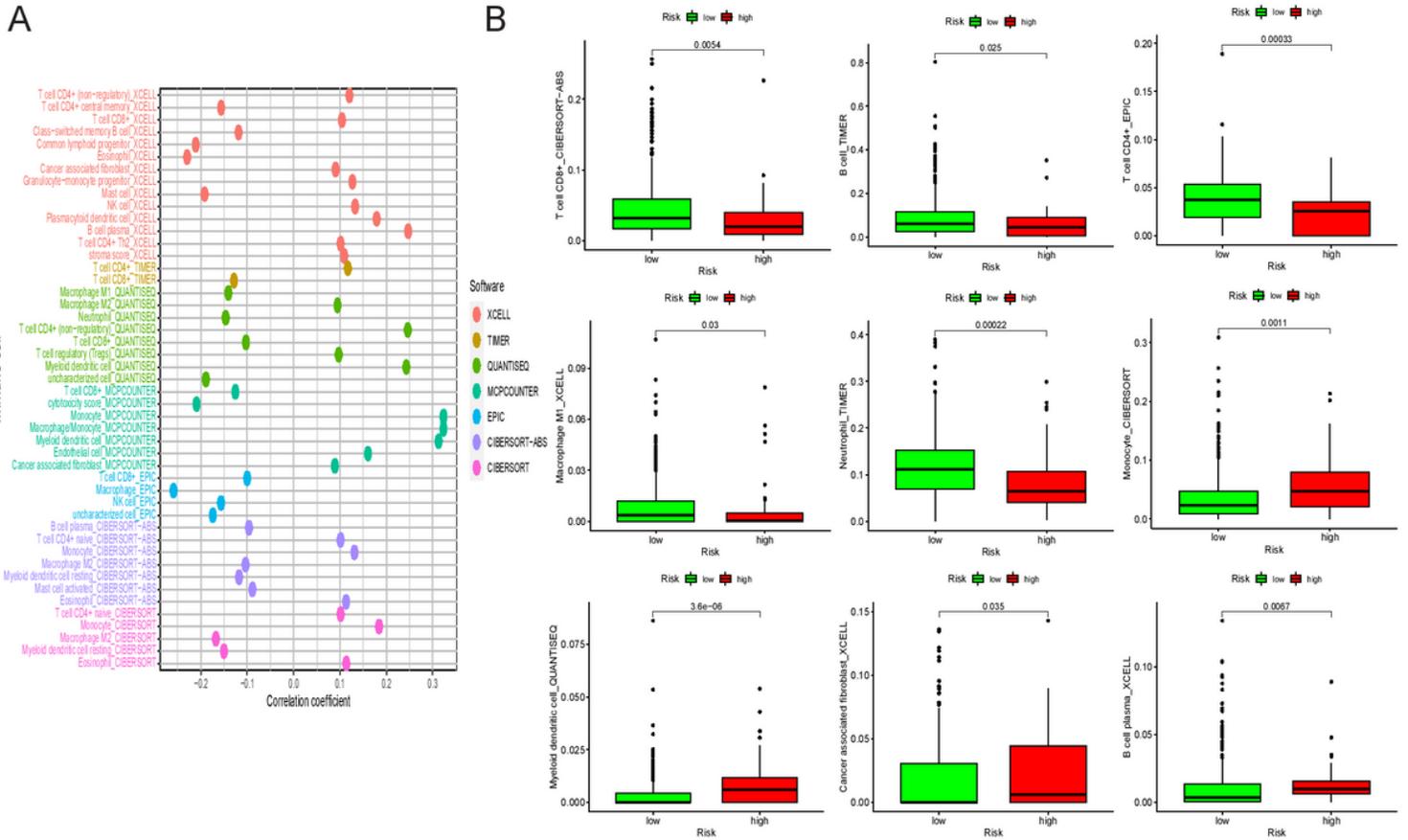
**Figure 4**

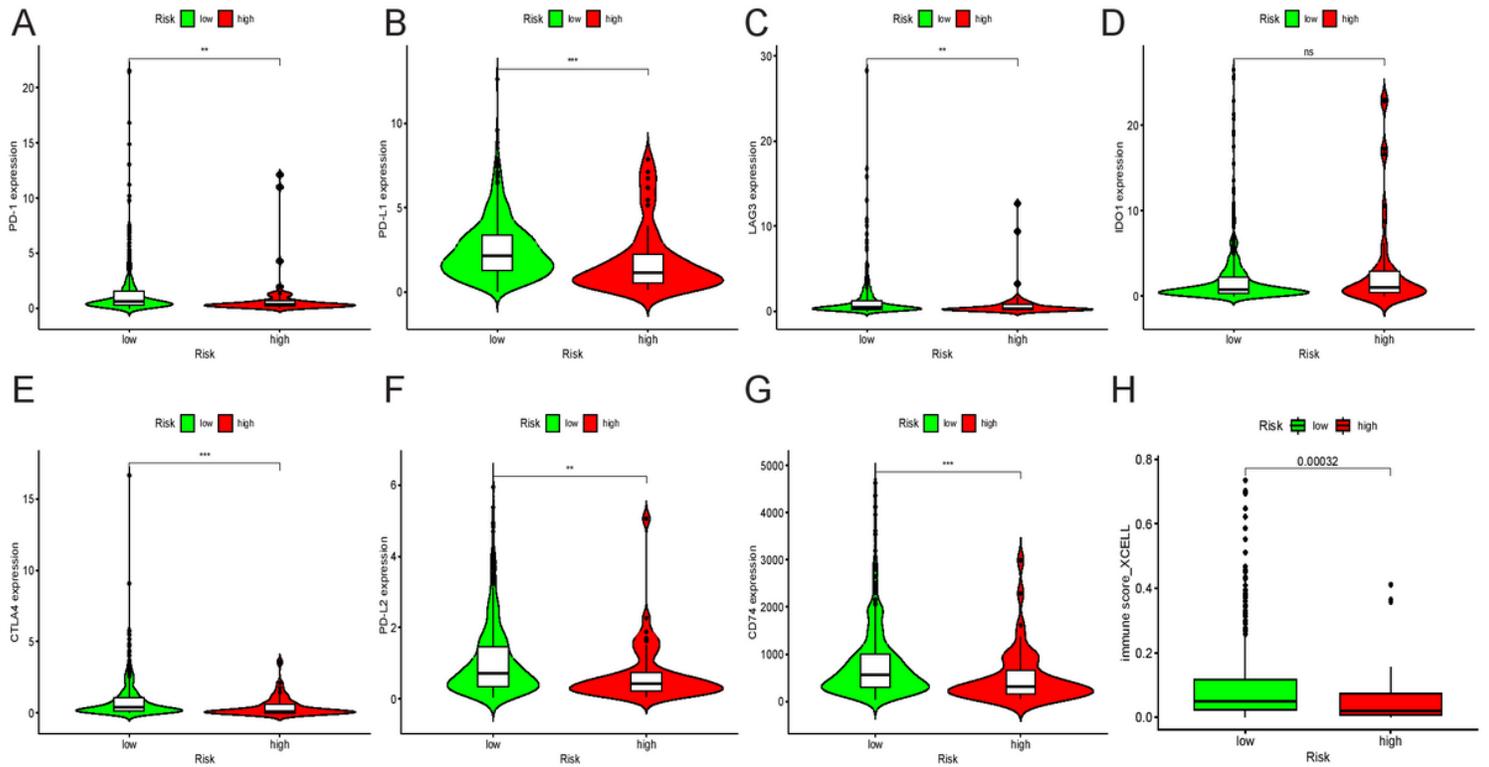
The clinical value of the signature (A) Strip chart, which was labeled as follows: <math><0.001 = \*\*\*</math>, <math><0.01 = \*\*</math>, and <math><0.05 = \*</math>. (B) Comparison of the 1-year ROC curves of the risk score with those of other clinical features showed the superiority of the risk score. (C)(D): Univariate and multivariate Cox regression analyses were applied to identify prognostic factors associated with TC.



**Figure 5**

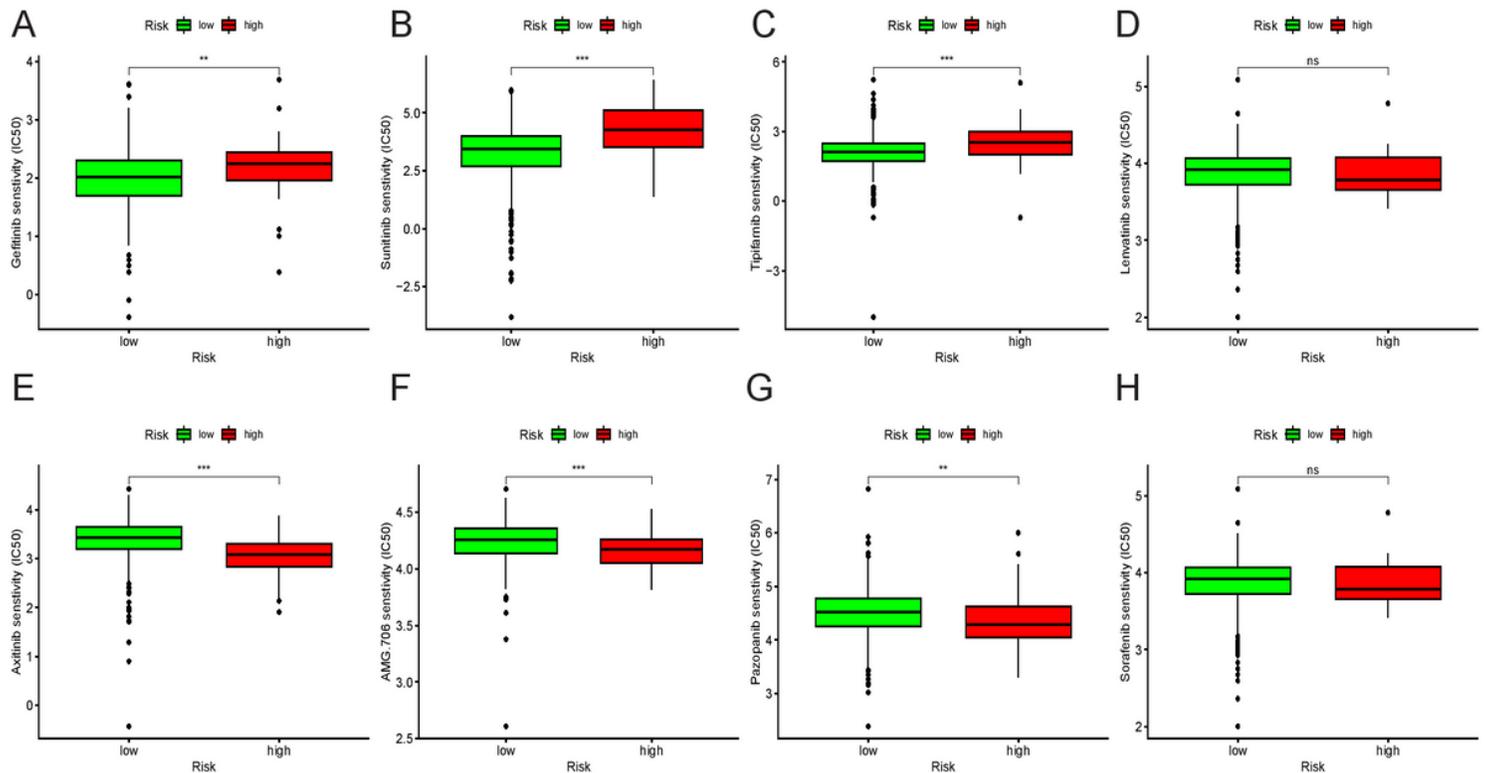
(A) Proportions of the 21 kinds of TICs explored by the CIBERSORT algorithm. (B) Correlations between 21 immune cell components.





**Figure 7**

Correlation between ICI target genes and the immune-related lncRNA signature, which was labeled as follows:  $<0.001 = ***$ ,  $<0.01 = **$ , and  $<0.05 = *$



**Figure 8**

Correlation between common TKIs and the immune-related lncRNA signature, which was labeled as follows:  $<0.001 = ***$ ,  $<0.01 = **$ , and  $<0.05 = *$ .

## Supplementary Files

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

- [SupplementalFiles.docx](#)
- [SupplementaryTable1.xls](#)