

# Construction of Novel Immune-Related LncRNA Signature and its Potential Prediction of Immune Status in Hepatocellular Carcinoma

**Min Deng**

Sun Yat-sen University Cancer Center <https://orcid.org/0000-0002-6775-2210>

**Jia-Bao Lin**

nanfang hospital

**Shao-Hua Li**

sun yat-sen university cancer center

**Rong-Ce Zhao**

sun yat-sen university cancer center

**Wei Wei**

sun yat-sen university cancer center

**Rong-Ping Guo** (✉ [dengmin@sysucc.org.cn](mailto:dengmin@sysucc.org.cn))

Sun Yat-sen University Cancer Center <https://orcid.org/0000-0002-7530-8902>

---

## Primary research

**Keywords:** Hepatocellular carcinoma, Immune-related lncRNA, Prognosis, Bioinformatics

**Posted Date:** May 5th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-412155/v1>

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

[Read Full License](#)

---

# Abstract

**Background:** The accuracy of the existing biomarkers in predicting the prognosis of hepatocellular carcinoma (HCC) is not satisfactory. It is necessary to explore biomarkers that can accurately predict the prognosis of HCC.

**Materials and methods:** In this study, the original transcriptome data was downloaded from The Cancer Genome Atlas (TCGA) database. Immune-related long non-coding ribonucleic acids (lncRNAs) were identified by co-expression analysis, and different expression lncRNA (DELncRNA) pairs were distinguished by univariate analysis. Besides, the least absolute shrinkage and selection operator (LASSO) penalized regression was modified. Next, the cut-off point was determined based on the area under the curve (AUC) and Akaike information criterion (AIC) values of the 5-year receiver operating characteristic curve (ROC) to establish an optimal model for identifying high-risk and low-risk groups in HCC patients. The model was then reassessed in terms of clinicopathological features, survival rate, tumor-infiltrating immune cells, immunosuppressive markers, and chemotherapy efficacy.

**Results:** A total of 1009 pairs of DELncRNA were recognized in this study, 30 of which were included in the Cox regression model for subsequent analysis. After regrouping according to the cut-off point, we can more effectively identify factors such as aggressive clinicopathological features, poor survival outcomes, specific tumor immune infiltration status, the high expression level of immunosuppressive biomarkers, and low sensitivity to chemotherapy drugs in HCC patients.

**Conclusions:** The non-specific expression level signature involved in lncRNAs shows a promising clinical value in predicting the prognosis of HCC patients.

## Background

Hepatocellular carcinoma (HCC) is the fourth most common malignant tumor globally, and its incidence is increasing annually with poor survival[1, 2]. The primary relevant risk factors associated with the development of HCC include viral hepatitis, alcoholic liver disease, nonalcoholic fatty liver disease, aflatoxin exposure, et al. [3]. In recent years, chemotherapeutics have achieved encouraging results in treating HCC, especially in the aspect of immune checkpoint inhibitors (ICIs)[4]. With the recent success of clinical trials of immunotherapy, such as Checkmate 040, Keynote-224, and IMbrave150, ICIs including nivolumab, pembrolizumab, and atezolizumab plus bevacizumab, have been approved for the treatment of HCC[5–7].

Long non-coding RNAs (lncRNAs) are non-protein-coding RNAs with a transcription length of more than 200 nucleotides[8]. Since lncRNAs are abundant, they often participate in humans' physiological processes and are closely related to the development of diseases[9–11]. In addition, lncRNAs have the ability to interact with molecules such as DNA, RNA, or protein to play the enhancement or inhibition roles[12]. Researches have reported that lncRNAs may participate in tumorigenesis through various molecular mechanisms[13, 14]. Recent work demonstrates that lncRNAs can promote the malignant

phenotypes of cancer by changing the genome or transcriptome level and varying the immune microenvironment[15]. LncRNAs can activate immune cells by expressing related genes, which leads to tumor immune cells infiltrating[16].

Tumor immune infiltration markers show prospective predictive and prognostic value in tumor diagnosis, treatment, and survival evaluation[17–19]. Moreover, lncRNAs have a close relationship with tumor immunity; the study of its combination with tumor immunity will help to establish these markers. Hong et al. constructed a model to predict the prognosis of HCC according to immune-related lncRNAs (irlncRNAs) and risk scores[20]; Wei et al. generated a marker for evaluating the prognosis of pancreatic cancer by screening nine irlncRNAs[21]; Jiang et al. established a prognostic model composed of three lncRNAs to predict the prognosis of clear cell renal cell carcinoma[21]; Wu et al. integrated clinical data and RNAs, including lncRNA, microRNA (miRNA), and mRNA, and established an immune-related signature to predict patients' survival with head and neck squamous cell carcinoma[22].

Generally speaking, the accuracy of the tumor prediction model based on the combination of two biomarkers is better than that of simple genes[23]. Up to the present, few models have studied the predictive role of lncRNAs and tumor immune-related cells in HCC. This study used a novel modeling algorithm, which does not need to involve a specific expression level, but through pairing and iteration to establish an irlncRNA signature. Subsequently, we evaluated the diagnostic effect, predictive value, tumor immune infiltration, and chemotherapy efficacy of this signature in HCC patients.

## Materials And Methods

### Transcriptome data collection and differential expression analysis

We downloaded the transcriptome data (RNAseq) corresponding to fragments per kilobase million (FPKM) from the TCGA database (<https://tcga-data.nci.nih.gov/tcga/>). The GTF files used for subsequent analysis, distinguishing mRNA and lncRNA, were extracted from the genome database, Ensembl (<http://asia.ensembl.org>). The ImmPort database (<http://www.immport.org>) was accessed to download the list of identified immune-related genes (ir-genes), serving as a screening role to filter irlncRNAs through co-expression methods. The correlation between lncRNAs and ir-genes was analyzed. The inclusion criteria of irlncRNAs were immune gene correlation coefficient  $> 0.4$  and  $p < 0.001$ . Package limma of R was performed to analyze the differential expression among irlncRNAs, and log fold change (FC)  $> 1$  and false discovery rate (FDR)  $< 0.05$  was regarded as thresholds to distinguish different expression irlncRNA (DEirlncRNA).

### DEirlncRNAs pairing

DEirlncRNAs were periodically single paired. Assuming that C was equal to lncRNA A plus lncRNA B, a 0 or 1 matrix was constructed; If the lncRNA A expression level is higher than that of lncRNA B, C is defined as 1; otherwise, C is defined as 0. After that, the constructed matrix was further screened. Suppose the expression level of lncRNA pairs is 0 or 1. In that case, it is considered that there is no

relationship between pairing and prognosis because there is no specific level of pairing to predict the survival outcome of patients correctly. When the number of lncRNA pairs with an expression of 0 or 1 accounts for more than 30% of the total pairs, it was regarded as an effective match.

### **Clinicopathological data acquisition**

The clinicopathological information of HCC patients was collected from the TCGA database. After excluding the cases with follow-up time less than 30 days and duplicate data, the adequate data were extracted.

### **Establish a risk model for evaluating the riskScore**

The single-factor analysis was the first step. Then, the lasso regression with 10-fold cross-validation was carried out, and the p-value was 0.05. Lasso regression was performed for 1000 cycles, and 1000 random stimulation were set in each cycle. The next step was to record the frequency of each pair in the Lasso regression model that was repeated 1000 times and select the pairs with a frequency more than 100 times for Cox proportional hazard regression analysis and model generation. The AUC value was calculated, and the curve was plotted. The maximum area under the curve (AUC) value is when the curve reaches the highest point, the calculation process is terminated, and the model is regarded as the best candidate. In this study, the 1-year, 3-year, and 5-year receiver operating characteristic curves (ROC) of the risk model were drawn. The formula of the riskScore for all clinical cases is as follows.

**RiskScore =  $\beta_i S_i$ .** The Akaike information criterion (AIC) value of each point of the 5-year ROC curve is used to determine the cut-off point to distinguish the high or low risk of RiskScores.

### **Risk model validation**

Kaplan-Meier analysis was used to verify the cut-off point to show the difference of survival rate between high-risk group and low-risk group. Then, the survival curve was plotted, and R was used to visualize the riskScore of each case in the model. The R packages, including glmnet, survival, survivalROC, pbapply, surfminer, and pheatmap, were utilized in these analyses. A Chi-square test was conducted to investigate the relationship between the clinicopathological features and model, which in order to validate the clinical appliance usefulness of the generated model. Then the band chart was visualized and marked as below: <0.05 was marked \*, <0.01 was marked \*\*, and <0.001 was marked \*\*\*. Wilcoxon signed-rank test was performed to analyze the differences in riskScores between various clinicopathological feature groups. Results were demonstrated by Box plots. Univariate and multivariate Cox regression analysis was performed between the riskScore and clinicopathological features to certify the possibility of an independent predictor of clinical prognosis for this model. Results were displayed by a forest plot. Packages of R, including ggpubr, survival, and pHeatmap, were used in these procedures.

### **Study on tumor-infiltrating immune cells**

The currently accepted methods for examining the immune infiltration status between samples from LIHC of the TCGA dataset were used to analyze the relationship between risk and immune cell features. These methods include XCELL, TIMER, QUANTISEQ, MCPOUNTER, EPIC, CIBERSORT-ABS, and CIBERSORT. The immune infiltrating cell content analysis between the high-risk and low-risk groups used the Wilcoxon signed-rank test. Results were shown by a box diagram. The relationship between immune infiltrating cells and riskScores was performed by Spearman correlation analysis. A lollipop chart was drawn to show the correlation coefficient of the results.  $P < 0.05$  was considered as a significant threshold. Ggplot2 packages of R conducted the procedure.

### **The noteworthy relationship between the model and the clinical therapeutics**

In order to assess the clinical application value of the model in the treatment of HCC, the IC50 of commonly used chemotherapy drugs in the TCGA project of the LIHC dataset was calculated. Anti-tumor medications such as bleomycin, doxorubicin, erlotinib, gemcitabine, methotrexate, mitomycin, paclitaxel, rapamycin, cisplatin, and sorafenib are commonly used in the treatment of various types of malignant tumors recommended by AJCC guidelines. Wilcoxon signed-rank test was performed to compare the difference of IC50 between the high-risk group and low-risk group. The results were shown in the box diagram drawn by ggplot2 and pRRophetic of R.

### **Expression analysis of immunosuppressive molecules related to ICIs**

The relationship between the model and the gene expression level associated with ICIs was examined and visualized by the ggstatsplot and violin plot.

## **Results**

### **Differential expression analysis of irlncRNAs (DEirlncRNAs)**

We downloaded transcriptome data of HCC from the TCGA database, including 50 non-tumor tissues and 374 tumor tissues. The gene transfer format (GTF) files from Ensemble were used to annotate the accessed data, and co-expression of known ir-genes and lncRNAs was analyzed. Totally, 740 irlncRNAs were identified, of which 490 were classified as DEirlncRNAs. Among the DEirlncRNAs, 16 were down-regulated, and 474 were up-regulated (**Figure 1A, 1B, Table S1, S2**).

### **DEirlncRNA pairs screening and risk model construction**

A total of 10344 valid DEirlncRNA pairs from 490 DEirlncRNAs were identified by iteration loop and 0 or 1 matrix. 1009 DEirlncRNA pairs were screened by a single factor test and modified LASSO analysis, of which 30 pairs were involved in the Cox model by the stepwise method. Results were shown in **Figure 1C, 1D, and 1E**. After that, the areas under the curve (AUC) for each receiver operating characteristic (ROC) curve of 1009 were calculated, and the curve was plotted. At the same time, we found that the maximum AUC value was obtained when the highest point equal to 0.941, then the the optimal DEirlncRNA pair was determined (**Figure 2A**). Our study also used the Akaike information criterion (AIC) values to determine the

maximum inflection point as the cut-off point of the 5-year ROC curve (**Figure 2B**). ROC curves at 1, 3, and 5 years were drawn to verify the optimality, which indicated all AUC values exceeded 0.91 (**Figure 2C**). Moreover, AUC values between the 5-year ROC curve and some common clinical parameters were also compared. (**Figure 2D**). Furthermore, 343 HCC patients' data were collected from the TCGA database, and the risk scores of all patients were calculated. Then the cut-off point was utilized to re-differentiate the high-risk group and low-risk group for verification.

### **Application of risk assessment model in the clinical evaluation**

Fifty-nine cases and 284 cases were divided into high-risk and low-risk groups according to the cut-off point. **Figure 3A** and **Figure 3B** display the riskScores and survival rate of each case. These data indicate that patients' clinical outcome in the high-risk group was inferior to that in the low-risk group. The survival of the high-risk group was poor than that of the low-risk group by Kaplan Meier analysis ( $p < 0.001$ ) (**Figure 3C**). Next, a Chi-square test was conducted to explore the relationship between the risk of HCC and clinicopathological features. The stripping diagram (**Figure 4A**) and the scatter plots examined through Wilcoxon signed-rank test displayed T classification, tumor stage, tumor grade, Child-Pugh grade, ECOG, vascular invasion, and survival status (**Figure 4B-4H**) were significantly associated with the risk. Subsequently, univariate Cox regression analysis indicated that there were statistical differences in tumor stage ( $p < 0.001$ , HR = 1.627, 95% CI [1.218 – 2.173]), vascular invasion ( $p = 0.007$ , HR = 1.737, 95% CI [1.161 – 2.599]), ECOG ( $p < 0.001$ , HR = 1.945, 95% CI [1.346 – 2.811]), and riskScore ( $p < 0.001$ , HR = 1.028, 95% CI [1.017 – 1.039]), while tumor stage ( $p = 0.018$ , HR = 1.488, 95% CI [1.072 – 2.067]), ECOG ( $p < 0.001$ , HR = 1.905, 95% CI [1.299 – 2.792]), and riskScore ( $p < 0.001$ , HR = 1.024, 95% CI [1.015– 1.034]) can be regarded as independent prognostic predictors by multivariate Cox regression analysis (**Figure 4I**).

### **Relationship between tumor-infiltrating immune cells, immune molecules and the risk model**

Since lncRNAs were initially associated with ir-gene, we explored whether this model correlates with the tumor immune microenvironment. Our study found that there is a positive correlation between the high-risk group and tumor-infiltrating immune cells, such as B cells, Neutrophils, and macrophages, while negatively correlated with CD8 + T cells, CD4 + T cells, and monocytes. Spearman correlation analysis was carried out in detail, and the results were shown as a lollipop diagram (**Figure 5A**). Since ICIs currently play a significant role in the treatment of HCC, we explored the correlation between the risk model and ICI-related biomarkers. The results showed that high risk scores were associated with the expression of CD276 ( $p < 0.001$ ), GSDME ( $p < 0.001$ ), HAVCR2 ( $p < 0.01$ ), and TNFRSF18 ( $p < 0.05$ ) (**Figure 5B-5E**). There was no statistical difference between the high-risk scores and ir-genes, such as CTLA4, PDCD1, and LAG3 (All  $p > 0.05$ , **Figure 5F-5H**).

### **Correlation analysis of risk model and chemotherapy drugs**

In addition to checkpoint blocking therapy, we also tried to explore the association between the risk and efficacy of commonly used chemotherapy drugs for HCC. We found that the high-risk score was along

with a higher half inhibitory concentration (IC<sub>50</sub>) of chemotherapy drugs for erlotinib ( $p < 0.001$ ), methotrexate ( $p = 0.0024$ ), and rapamycin ( $p = 0.0043$ ), while a lower IC<sub>50</sub> for bleomycin ( $p < 0.001$ ), doxorubicin ( $p = 0.0023$ ), gemcitabine ( $p < 0.001$ ), mitomycin ( $p < 0.001$ ), and paclitaxel ( $p = 0.011$ ), suggesting that this model can be served as a potential predictor for the sensitivity of chemotherapies (Figure 6A-6J).

## Discussion

It is necessary to improve the accuracy of prognostic markers for HCC patients. LncRNAs are closely related to normal physiological activities and the development of diseases[10, 24]. Furthermore, studies demonstrated that lncRNAs play a vital role in tumor development and anti-tumor processes[25–27]. Recent researches have focused on investigating the potential relationship between coding genes and non-coding RNAs to predict patients' prognosis with cancers[28, 29]. Regrettably, the majority of these signatures were generated by the particular expression levels of transcripts. Our research did not pay attention to the specific expression levels of lncRNAs but utilized the method of ir-gene pairing to generate a practical model with the combination of lncRNAs.

First of all, we downloaded the original information of lncRNAs from the TCGA database, and then differential co-expression analysis was performed to catalog DEir-lncRNAs. The lncRNA-pairs pairs were verified by an improved cyclic single pair method along with 0 or 1 matrix. Secondly, univariate analysis and modified LASSO penalty regression were performed to determine DEir-lncRNAs pairs, procedures including cross-validation, multiple repetitions, and random stimulation. And then, we gained the optimum model by examining each AUC value of ROC, and the optimum cut-off point was according to the AIC value of each point on the AUC to distinguish the high-risk and low-risk group in HCC cases. Finally, the model was evaluated according to various parameters, such as survival rate, clinicopathological features, tumor-infiltrating immune cells, checkpoint-associated molecules, and chemotherapeutics.

The relationship between lncRNAs and tumors has been paid more and more attention[30–32]. Deng et al. established a model to predict HCC patients' survival[33]. The method utilized in this study does not need to check the specific expression level of each lncRNA but only needs to detect the pairs with high or low expression levels. Therefore, the model is practical and straightforward in distinguishing high-risk or low-risk clinical cases. The lncRNAs included in this model are related to ir-genes, so these lncRNAs may regulate the immune microenvironment and the activation of immune cells.

Our research reveals that part of the DEir-lncRNAs included in the modeling play a vital role in the malignant phenotype of many cancers, such as MYLK – AS1[34, 35], THUMP3 – AS1[36], and DSCR8[37], especially in the development of HCC. MYLK – AS1 promotes angiogenesis and HCC progression by targeting the miR-424-5p/E2F7 axis and activating the VEGFR-2 signaling pathway[35]. THUMP3 – AS1 was associated with cell cycle and can be used as a prognostic marker in hepatitis B virus-related HCC patients[36]. Wang et al. revealed that DSCR8 promotes the progression of HCC by

activating Wnt / b-catenin signaling pathway[38]. The established model can identify new biomarkers for further tumor-related researches.

In order to achieve better accuracy and effectiveness of risk prediction, this study used the improved method of the LASSO penalty model[39]. In addition, we determined the maximum value for an optimal model by calculating each AUC value and then compared it with other clinicopathological characteristics to further improve the modeling process. The AIC value was applied to get the ideal cut-off point for model fitting instead of only using the median value to discriminate risk. After using this new method to differentiate high-risk and low-risk groups, survival outcomes and univariate and multivariate analysis of clinicopathological features were reevaluated. Moreover, the sensitivity of common chemotherapy drugs to HCC treatment was analyzed. The relationship between high-risk and low-risk groups and tumor immune infiltration, and the relationship between high-risk and low-risk groups and immune checkpoint-related genes were also studied, which indicated that this modeling algorithm has a good clinical application prospect.

The reaction of blocking immune checkpoint is closely related to tumor-infiltrating immune cells[40]. Our research used seven common recognized methods to reckon the immune infiltrating cells, which in order to investigate the relationship between risk scores and tumor-infiltrating immune cells, including XCELL[41, 42], TIMER[43, 44], QUANTISEQ[45, 46], MCPCOUNTER[47], EPIC[48], CIBERSORT-ABS, and CIBERSORT[49, 50]. Due to the defects and complexity of these algorithms, they are rarely compared with each other. Through integration analysis, our findings show that DEirIncRNA pairs have a positive correlation with tumor-infiltrating immune cells such as B cells, Neutrophils, macrophages, while negatively correlated with CD8 + T cells, CD4 + T cells, and monocytes. Wang et al. demonstrated that the immune score could predict the efficacy of immunotherapy and chemotherapy[51]. lncRNAs SATB2-AS1 can affect tumor immune cell microenvironment and inhibit colorectal cancer metastasis[34]. lncRNA-EGFR can stimulate T regulatory cell differentiation and promote immune evasion of HCC[35]. Our model suggests that the high risk was related to the sensitivity of chemotherapy drugs such as methotrexate, rapamycin, bleomycin, doxorubicin, gemcitabine, mitomycin, and paclitaxel, rather than the sensitivity of sorafenib. We believe that immunotherapy is more effective than traditional chemotherapy, mainly because immunotherapy can activate the immune function and resist tumors through active immunity. Tumoral mutations can cause a large number of neoantigens to be released, which can be recognized by T cells and cause rich immune cells to infiltrate into the tumor[52–54].

We have to admit that our study has some defects and limitations. Firstly, the research data was based on public databases. Some data are incomplete, such as some clinicopathological features and the sensitivity of drugs commonly used in the treatment of HCC, for instance, lenvatinib and oxaliplatin, have not been analyzed; Secondly, the constructed model needs external verification because the expression level of each sample is different, which may lead to the unreliability of the final model. However, this research uses various methods to verify the new modeling algorithm and optimizes and analyzes it. Although the lack of external data validation, from the analysis results, our model was acceptable. But

this study will be more convincing if there is external validation. Therefore, our team will re-collect clinicopathological data for the subsequent studies and enlarge the sample size for further verification.

## **Conclusions**

Our research shows an innovative signature established by irlncRNAs that do not need to predict the expression level of lncRNA can predict HCC patients' prognosis and may contribute to identifying those patients who could benefit from anti-tumor immunotherapy.

## **Declarations**

### **Acknowledgements**

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

### **Authors' contributions**

Conceptualization, project administration, formal analysis, software, and writing—review and editing, M.D.; Data analysis, software, figures preparation, J.-B.L.; Data curation and analysis, figures and tables preparation, S.-H.L.; Data collection and analysis, R.-C.Z; Conceived and designed this study, W.W.; Conceptualization, supervision, writing—review, and editing, R.-P.G. All authors have read and approved the final version of the manuscript.

### **Funding**

None.

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

Publicly available datasets were analyzed in this study, these can be found in The Cancer Genome Atlas (<https://portal.gdc.cancer.gov/>).

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

Not applicable.

### **Consent for publication**

Not applicable.

### **Competing interests**

The authors declare no conflicts of interest.

# Abbreviations

HCC: hepatocellular carcinoma; TCGA: The Cancer Genome Atlas; lncRNAs: long non-coding RNAs; irlncRNA: Immune-related long non-coding ribonucleic acids; DEirlncRNA: different expression irlncRNA; miRNA: microRNA; LASSO: least absolute shrinkage and selection operator; AUC: Area under the ROC curve; AIC: Akaike information criterion; ROC: Receiver operating characteristic; ICIs: immune checkpoint inhibitors; FPKM: fragments per kilobase million; ir-genes: immune-related genes; FC: fold change; FDR: false discovery rate; GTF: gene transfer format; ECOG: eastern cancer oncology group; IC50: half inhibitory concentration.

# References

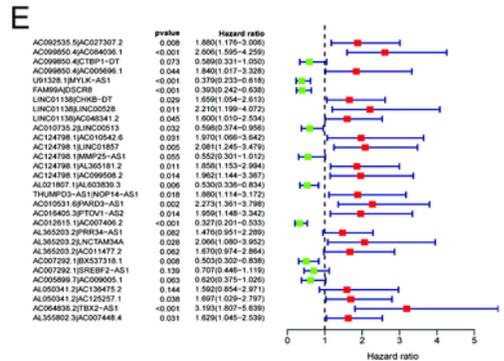
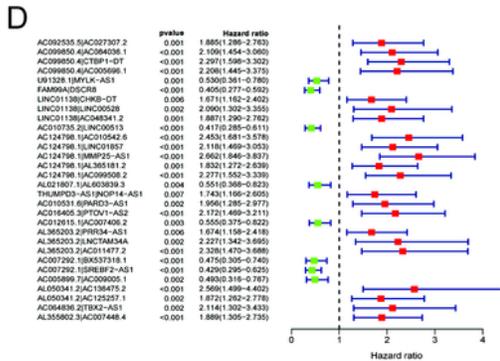
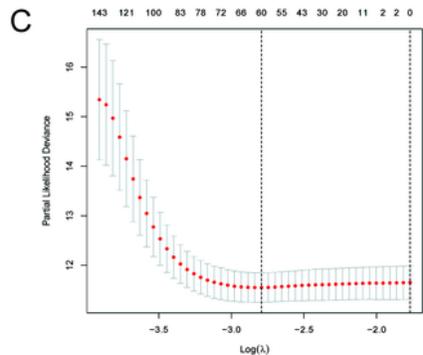
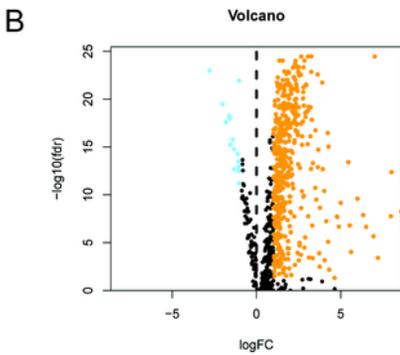
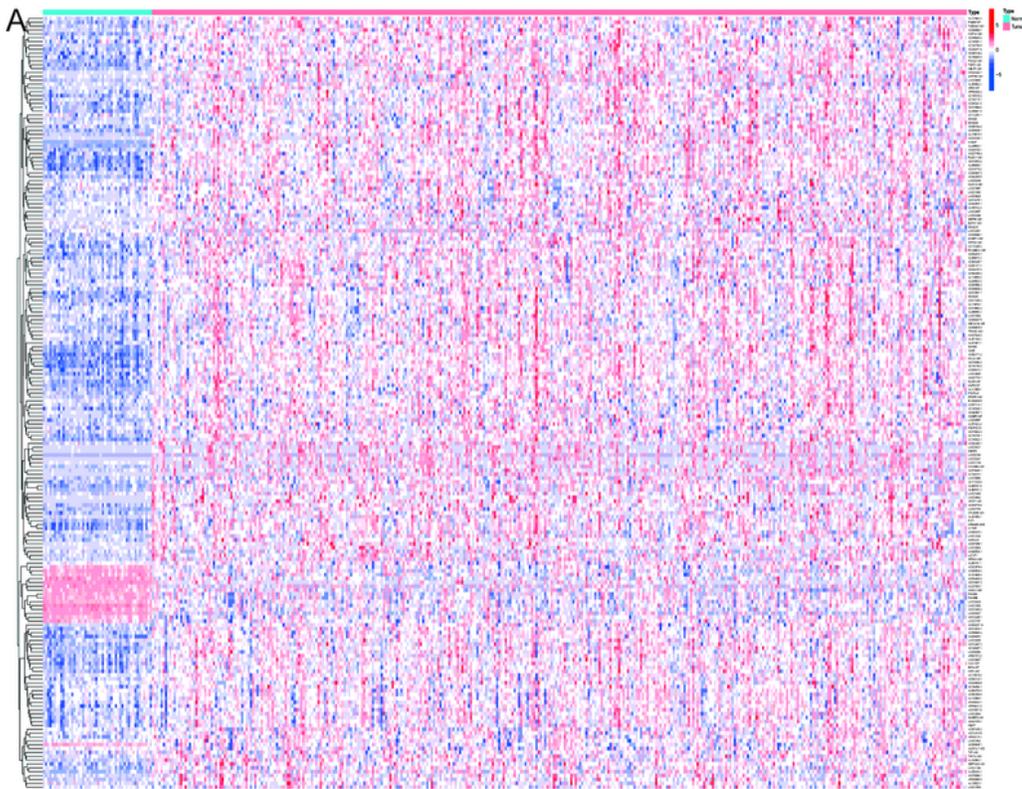
1. Villanueva A. Hepatocellular Carcinoma. *N Engl J Med*. 2019;380:1450–62.
2. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2018;68:394–424.
3. Forner A, Reig M, Bruix J. Hepatocellular carcinoma. *Lancet*. 2018.
4. Havel JJ, Chowell D, Chan TA. The evolving landscape of biomarkers for checkpoint inhibitor immunotherapy. *Nat Rev Cancer*. 2019;19:133 – 50.
5. El-Khoueiry AB, Sangro B, Yau T, Crocenzi TS, Kudo M, Hsu C, et al. Nivolumab in patients with advanced hepatocellular carcinoma (CheckMate 040): an open-label, non-comparative, phase 1/2 dose escalation and expansion trial. *Lancet*. 2017;389:2492–502.
6. Finn RS, Qin S, Ikeda M, Galle PR, Ducreux M, Kim TY, et al. Atezolizumab plus Bevacizumab in Unresectable Hepatocellular Carcinoma. *N Engl J Med*. 2020;382:1894 – 905.
7. Zhu AX, Finn RS, Edeline J, Cattan S, Ogasawara S, Palmer D, et al. Pembrolizumab in patients with advanced hepatocellular carcinoma previously treated with sorafenib (KEYNOTE-224): a non-randomised, open-label phase 2 trial. *Lancet Oncol*. 2018;19:940 – 52.
8. Qian X, Zhao J, Yeung PY, Zhang QC, Kwok CK. Revealing lncRNA Structures and Interactions by Sequencing-Based Approaches. *Trends Biochem Sci*. 2019;44:33–52.
9. Yang G, Lu X, Yuan L. LncRNA: a link between RNA and cancer. *Biochim Biophys Acta*. 2014;1839:1097 – 109.
10. Sun X, Zheng H, Sui N. Regulation mechanism of long non-coding RNA in plant response to stress. *Biochem Biophys Res Commun*. 2018;503:402-7.
11. Quinn JJ, Chang HY. Unique features of long non-coding RNA biogenesis and function. *Nat Rev Genet*. 2016;17:47–62.
12. Robinson EK, Covarrubias S, Carpenter S. The how and why of lncRNA function: An innate immune perspective. *Biochim Biophys Acta Gene Regul Mech*. 2020;1863:194419.

13. Zhao J, Lawless MW. Long noncoding RNAs and their role in the liver cancer axis. *Nat Rev Gastroenterol Hepatol.* 2013;10:703.
14. Schmitt AM, Chang HY. Long Noncoding RNAs in Cancer Pathways. *Cancer Cell.* 2016;29:452 – 63.
15. Atianand MK, Caffrey DR, Fitzgerald KA. Immunobiology of Long Noncoding RNAs. *Annu Rev Immunol.* 2017;35:177 – 98.
16. Chen YG, Satpathy AT, Chang HY. Gene regulation in the immune system by long noncoding RNAs. *Nat Immunol.* 2017;18:962 – 72.
17. Peng D, Wang L, Li H, Cai C, Tan Y, Xu B, et al. An immune infiltration signature to predict the overall survival of patients with colon cancer. *IUBMB Life.* 2019;71:1760-70.
18. Wei ZW, Wu J, Huang WB, Li J, Lu XF, Yuan YJ, et al. Immune-infiltration based signature as a novel prognostic biomarker in gastrointestinal stromal tumour. *EBioMedicine.* 2020;57:102850.
19. Zhang B, Wang Q, Fu C, Jiang C, Ma S. Exploration of the immune-related signature and immune infiltration analysis for breast ductal and lobular carcinoma. *Ann Transl Med.* 2019;7:730.
20. Hong W, Liang L, Gu Y, Qi Z, Qiu H, Yang X, et al. Immune-Related lncRNA to Construct Novel Signature and Predict the Immune Landscape of Human Hepatocellular Carcinoma. *Mol Ther Nucleic Acids.* 2020;22:937 – 47.
21. Wei C, Liang Q, Li X, Li H, Liu Y, Huang X, et al. Bioinformatics profiling utilized a nine immune-related long noncoding RNA signature as a prognostic target for pancreatic cancer. *J Cell Biochem.* 2019;120:14916-27.
22. Wu S, Dai X, Xie D. Identification and Validation of an Immune-Related RNA Signature to Predict Survival of Patients With Head and Neck Squamous Cell Carcinoma. *Front Genet.* 2019;10:1252.
23. Lv Y, Lin SY, Hu FF, Ye Z, Zhang Q, Wang Y, et al. Landscape of cancer diagnostic biomarkers from specifically expressed genes. *Brief Bioinform.* 2020;21:2175-84.
24. Zhu J, Fu H, Wu Y, Zheng X. Function of lncRNAs and approaches to lncRNA-protein interactions. *Sci China Life Sci.* 2013;56:876 – 85.
25. Li J, Meng H, Bai Y, Wang K. Regulation of lncRNA and Its Role in Cancer Metastasis. *Oncol Res.* 2016;23:205 – 17.
26. Huang Z, Zhou JK, Peng Y, He W, Huang C. The role of long noncoding RNAs in hepatocellular carcinoma. *Mol Cancer.* 2020;19:77.
27. Sanchez Calle A, Kawamura Y, Yamamoto Y, Takeshita F, Ochiya T. Emerging roles of long non-coding RNA in cancer. *Cancer Sci.* 2018;109:2093 – 100.
28. Zhu X, Tian X, Yu C, Shen C, Yan T, Hong J, et al. A long non-coding RNA signature to improve prognosis prediction of gastric cancer. *Mol Cancer.* 2016;15:60.
29. Qu L, Wang ZL, Chen Q, Li YM, He HW, Hsieh JJ, et al. Prognostic Value of a Long Non-coding RNA Signature in Localized Clear Cell Renal Cell Carcinoma. *Eur Urol.* 2018;74:756 – 63.
30. Chan JJ, Tay Y. Noncoding RNA:RNA Regulatory Networks in Cancer. *Int J Mol Sci.* 2018;19.

31. Bhan A, Soleimani M, Mandal SS. Long Noncoding RNA and Cancer: A New Paradigm. *Cancer Res.* 2017;77:3965-81.
32. Renganathan A, Felley-Bosco E. Long Noncoding RNAs in Cancer and Therapeutic Potential. *Adv Exp Med Biol.* 2017;1008:199–222.
33. Deng B, Yang M, Wang M, Liu Z. Development and validation of 9-long Non-coding RNA signature to predicting survival in hepatocellular carcinoma. *Medicine (Baltimore).* 2020;99:e20422.
34. Xu M, Xu X, Pan B, Chen X, Lin K, Zeng K, et al. LncRNA SATB2-AS1 inhibits tumor metastasis and affects the tumor immune cell microenvironment in colorectal cancer by regulating SATB2. *Mol Cancer.* 2019;18:135.
35. Jiang R, Tang J, Chen Y, Deng L, Ji J, Xie Y, et al. The long noncoding RNA Inc-EGFR stimulates T-regulatory cells differentiation thus promoting hepatocellular carcinoma immune evasion. *Nat Commun.* 2017;8:15129.
36. Zhao X, Bai Z, Li C, Sheng C, Li H. Identification of a Novel Eight-lncRNA Prognostic Signature for HBV-HCC and Analysis of Their Functions Based on Coexpression and ceRNA Networks. *Biomed Res Int.* 2020;2020:8765461.
37. You Q, Yao Y, Wu J, Cheng C, Li Y, Yuan H. YY1-induced lncRNA DSCR8 promotes the progression of ovarian cancer via miR-3192-5p/YY1 axis. *Biomed Pharmacother.* 2020;129:110339.
38. Wang Y, Sun L, Wang L, Liu Z, Li Q, Yao B, et al. Long non-coding RNA DSCR8 acts as a molecular sponge for miR-485-5p to activate Wnt/ $\beta$ -catenin signal pathway in hepatocellular carcinoma. *Cell Death Dis.* 2018;9:851.
39. Sveen A, Ågesen TH, Nesbakken A, Meling GI, Rognum TO, Liestøl K, et al. ColoGuidePro: a prognostic 7-gene expression signature for stage III colorectal cancer patients. *Clin Cancer Res.* 2012;18:6001-10.
40. Zhou G, Sprengers D, Boor P, Doukas M, Schutz H, Mancham S, et al. Antibodies Against Immune Checkpoint Molecules Restore Functions of Tumor-Infiltrating T Cells in Hepatocellular Carcinomas. *Gastroenterology.* 2017;153:1107-19.e10.
41. Aran D Cell-Type Enrichment Analysis of Bulk Transcriptomes Using xCell. *Methods Mol Biol.* 2020;2120:263 – 76.
42. Aran D, Hu Z, Butte AJ. xCell: digitally portraying the tissue cellular heterogeneity landscape. *Genome Biol.* 2017;18:220.
43. Li T, Fan J, Wang B, Traugh N, Chen Q, Liu JS, et al. TIMER: A Web Server for Comprehensive Analysis of Tumor-Infiltrating Immune Cells. *Cancer Res.* 2017;77:e108-108e110.
44. Li T, Fu J, Zeng Z, Cohen D, Li J, Chen Q, et al. TIMER2.0 for analysis of tumor-infiltrating immune cells. *Nucleic Acids Res.* 2020;48:W509-509W514.
45. Finotello F, Mayer C, Plattner C, Laschober G, Rieder D, Hackl H, et al. Molecular and pharmacological modulators of the tumor immune contexture revealed by deconvolution of RNA-seq data. *Genome Med.* 2019;11:34.

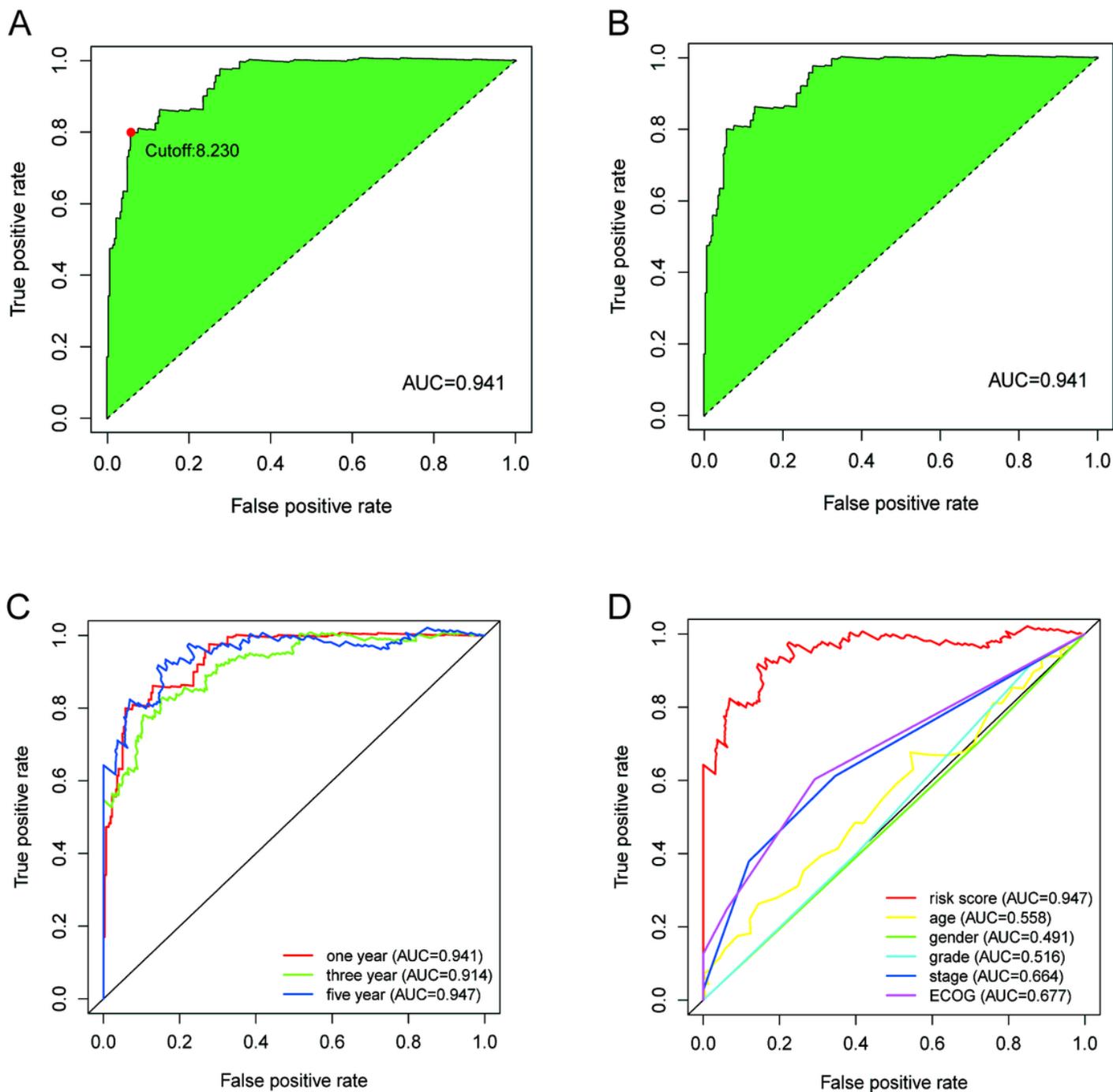
46. Plattner C, Finotello F, Rieder D. Deconvoluting tumor-infiltrating immune cells from RNA-seq data using quanTIseq. *Methods Enzymol.* 2020;636:261 – 85.
47. Dienstmann R, Villacampa G, Sveen A, Mason MJ, Niedzwiecki D, Nesbakken A, et al. Relative contribution of clinicopathological variables, genomic markers, transcriptomic subtyping and microenvironment features for outcome prediction in stage II/III colorectal cancer. *Ann Oncol.* 2019;30:1622-9.
48. 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.
49. Zhang H, Li R, Cao Y, Gu Y, Lin C, Liu X, et al. Poor Clinical Outcomes and Immuno-evasive Contexture in Intratumoral IL-10-Producing Macrophages Enriched Gastric Cancer Patients. *Ann Surg.* 2020.
50. Chen B, Khodadoust MS, Liu CL, Newman AM, Alizadeh AA. Profiling Tumor Infiltrating Immune Cells with CIBERSORT. *Methods Mol Biol.* 2018;1711:243 – 59.
51. Dai GP, Wang LP, Wen YQ, Ren XQ, Zuo SG. Identification of key genes for predicting colorectal cancer prognosis by integrated bioinformatics analysis. *Oncol Lett.* 2020;19:388 – 98.
52. Schumacher TN, Schreiber RD. Neoantigens in cancer immunotherapy. *Science.* 2015;348:69–74.
53. Chandran SS, Klebanoff CA. T cell receptor-based cancer immunotherapy: Emerging efficacy and pathways of resistance. *Immunol Rev.* 2019;290:127 – 47.
54. Yamamoto TN, Kishton RJ, Restifo NP. Developing neoantigen-targeted T cell-based treatments for solid tumors. *Nat Med.* 2019;25:1488-99.

## Figures



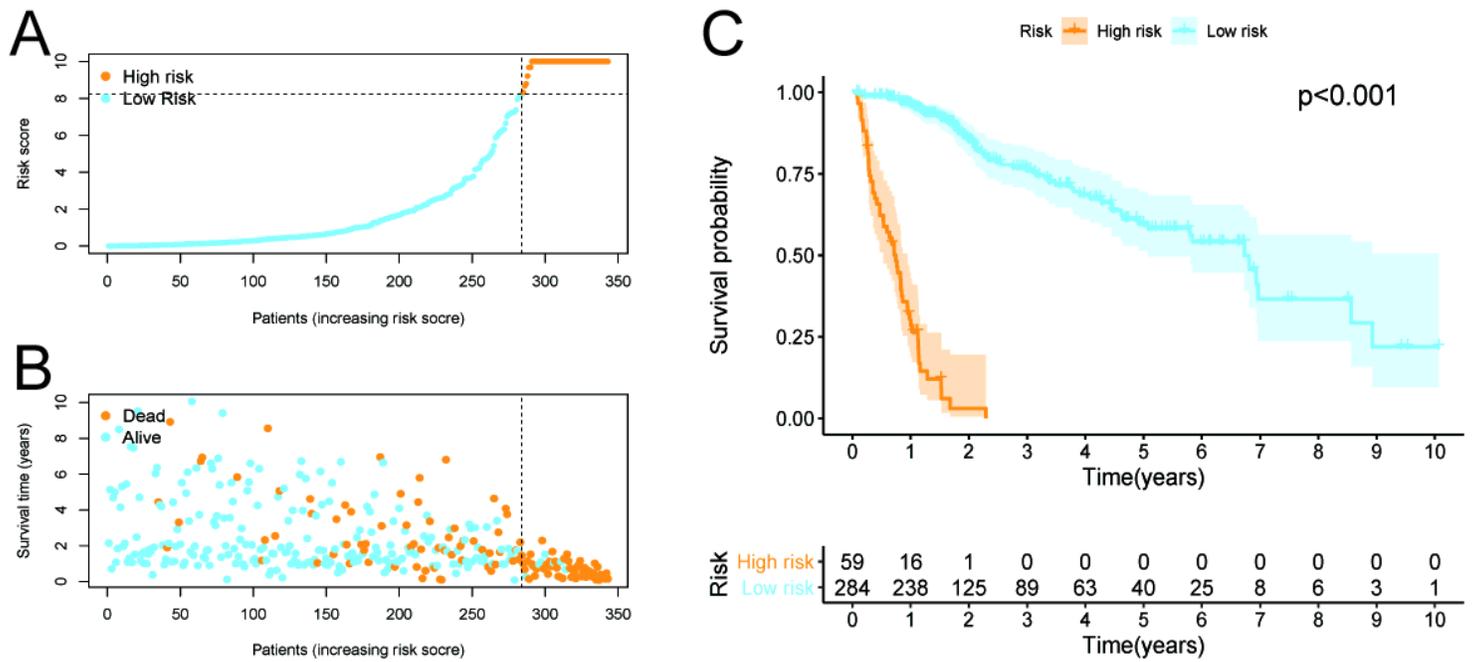
**Figure 1**

Establishment of a risk assessment model based on DEirIncRNA pairs. Differentially expressed immune-related lncRNAs (DEirIncRNAs) and volcano plot (B) were displayed. Establishment of the LASSO regression (C). 30 DEirIncRNA pairs were shown by a forest map (D and E).



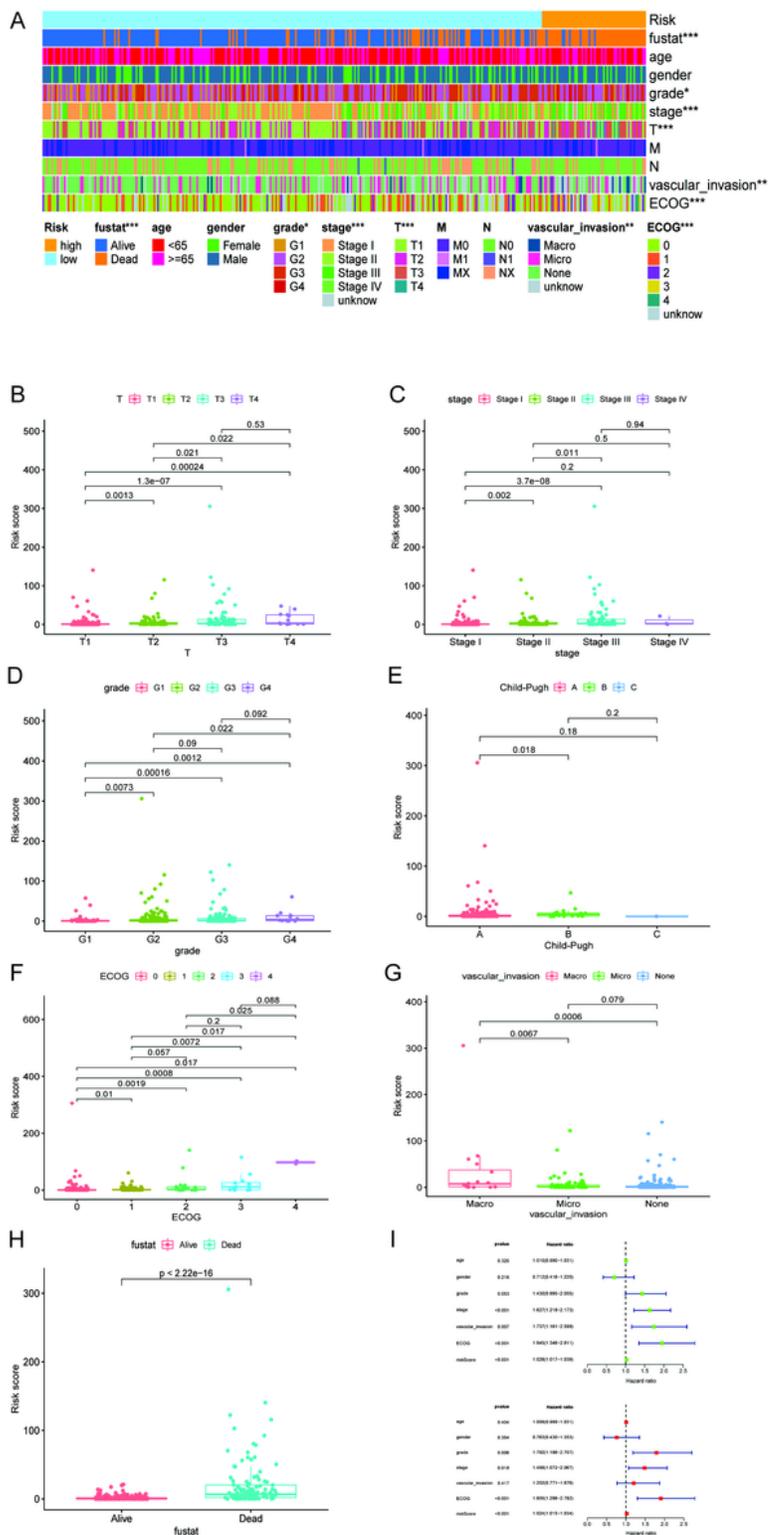
**Figure 2**

Establishment of a risk assessment model on the basis of DEIRlncRNA pairs The curve of each AUC value generated by ROCs of 1009 DEIRlncRNA pair models was drawn, and the highest point of AUC was determined; The maximum inflection point is the cut-off point acquired from AIC. (A and B). The 1-year, 3-year, and 5-year ROC of the optimal model showed that all AUC values exceeded 0.91 (C). Compared with other common clinical features, the 5-year ROC curves showed the superiority of the risk score (D).



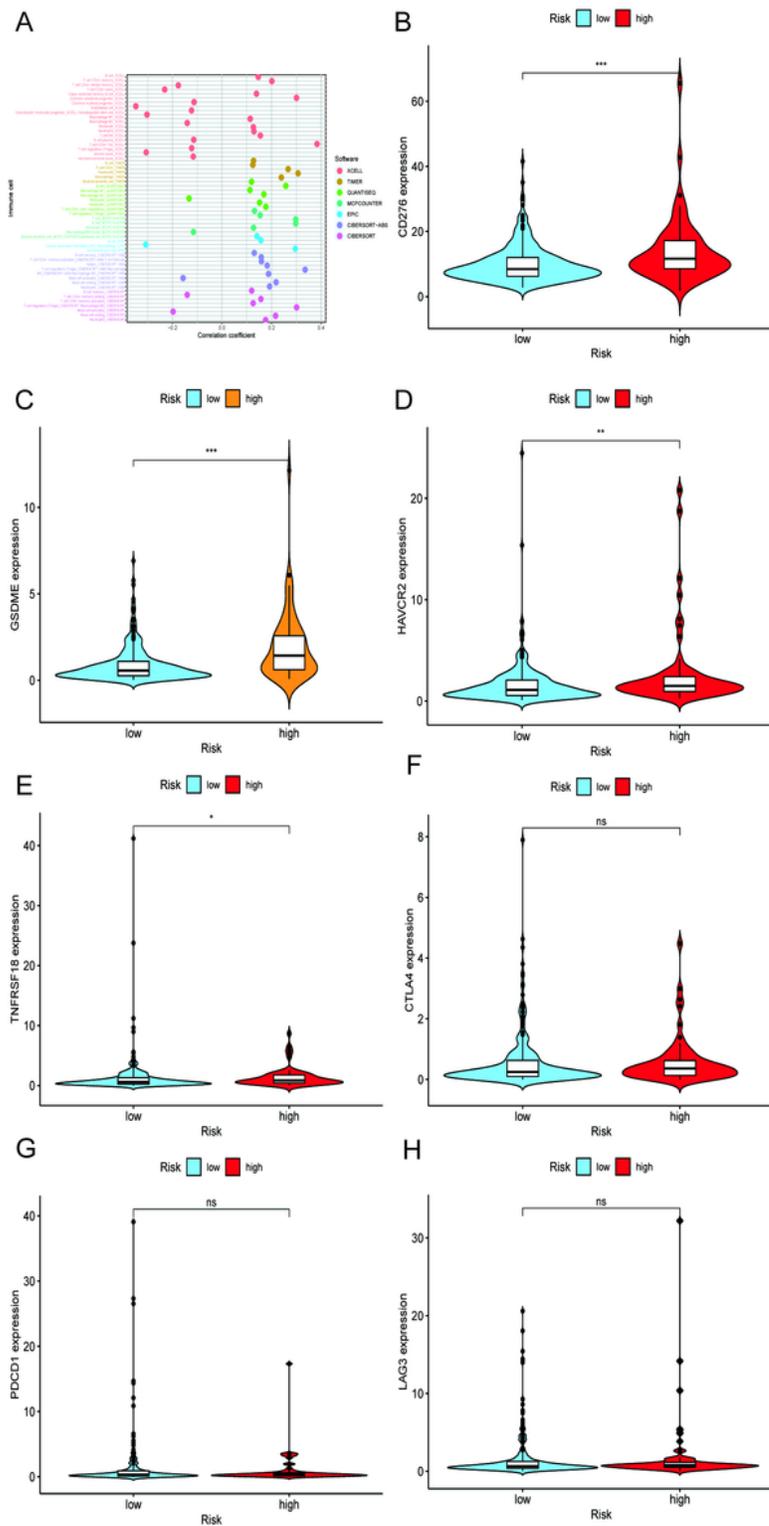
**Figure 3**

Prognostic prediction of the risk assessment model Risk scores (A) and survival outcome (B) of each case. Kaplan-Meier survival curve of high-risk group and low-risk group(C).



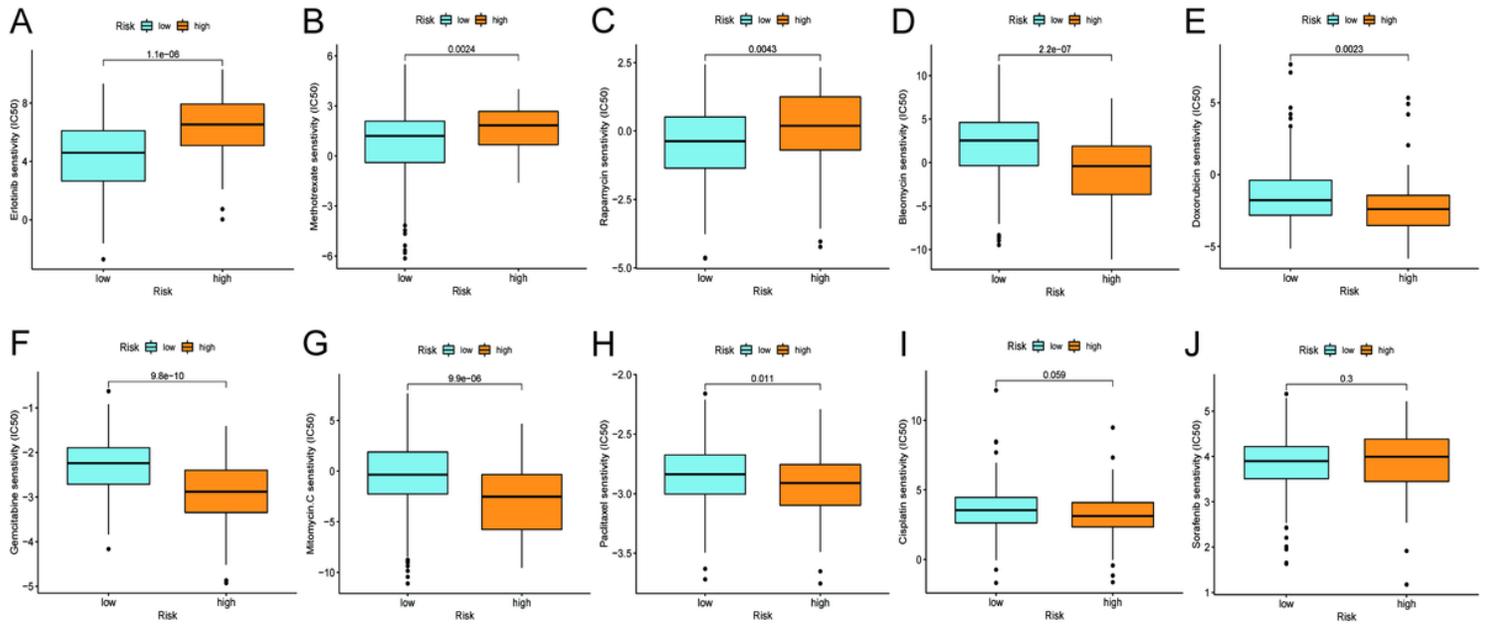
**Figure 4**

Application of risk assessment model in the clinical evaluation A strip diagram (A) and scatter plot showed that T classification (B), tumor stage (C), tumor grade (D), Child-Pugh grade (E), ECOG (F), vascular invasion (G), and survival status (H) were significantly correlated with risk score. Univariate and multivariate Cox regression analyses were performed to analyze the clinicopathological features and shown by a forest map (I).



**Figure 5**

Estimate tumor-infiltrating cells and immunosuppressive molecules through risk assessment model (A) Correlation between the high-risk group and tumor-infiltrating immune cells. High-risk scores were associated with the expression of CD276 (B), GSDME (C), HAVCR2 (D), and TNFRSF18 (E) levels. There was no statistical difference between the high-risk scores and immune-related genes, such as CTLA4 (F), PDCD1 (G), and LAG3 (H).



**Figure 6**

Relationship between risk scores and IC50 of chemotherapeutics

## Supplementary Files

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

- [supplementary.zip](#)