Molecular Therapy
Nucleic Acids
Original Article

Immune-Related lncRNA to Construct Novel Signature and Predict the Immune Landscape of Human Hepatocellular Carcinoma

Weifeng Hong,[1,6] Li Liang,[2,6] Yujun Gu,[3] Zhenhua Qi,[4] Haibo Qiu,[5] Xiaosong Yang,[4] Weian Zeng,[4] Liheng Ma,[1] and Jingdun Xie[4]

[1]Department of Medical Imaging, The First Affiliated Hospital of Guangdong Pharmaceutical University, Guangzhou, Guangdong 510000, China; [2]Departments of Medical Oncology, Zhongshan Hospital of Fudan University, Shanghai 200032, China; [3]Department of Ultrasonic Medicine, The First Affiliated Hospital of Sun Yat-Sen University, Guangzhou, Guangdong 510000, China; [4]Department of Anesthesiology, Sun Yat-Sen University Cancer Center, State Key Laboratory of Oncology in Southern China, Collaborative Innovation for Cancer Medicine, Guangzhou, Guangdong 510000, China; [5]Department of Gastric and Pancreatic Surgery, Sun Yat-Sen University Cancer Center, State Key Laboratory of Oncology in Southern China, Collaborative Innovation for Cancer Medicine, Guangzhou, Guangdong 510000, China

The signature composed of immune-related long noncoding ribonucleic acids (irlncRNAs) with no requirement of specific expression level seems to be valuable in predicting the survival of patients with hepatocellular carcinoma (HCC). Here, we retrieved raw transcriptome data from The Cancer Genome Atlas (TCGA), identified irlncRNAs by co-expression analysis, and recognized differently expressed irlncRNA (DEirlncRNA) pairs using univariate analysis. In addition, we modified Lasso penalized regression. Then, we compared the areas under curve, counted the Akaike information criterion (AIC) values of 5-year receiver operating characteristic curve, and identified the cut-off point to set up an optimal model for distinguishing the high- or low-disease-risk groups among patients with HCC. We then reevaluated them from the viewpoints of survival, clinic-pathological characteristics, tumor-infiltrating immune cells, chemotherapeutics efficacy, and immunosuppressed biomarkers. 36 DEirlncRNA pairs were identified, 12 of which were included in a Cox regression model. After regrouping the patients by the cut-off point, we could more effectively differentiate between them based on unfavorable survival outcome, aggressive clinic-pathological characteristics, specific tumor immune infiltration status, low chemotherapeutics sensitivity, and highly expressed immunosuppressed biomarkers. The signature established by paring irlncRNA regardless of expression levels showed a promising clinical prediction value.

INTRODUCTION

The incidence and mortality of hepatocellular carcinoma (HCC) combined with intrahepatic cholangiocarcinoma are estimated to be 42,810 and 30,160, respectively, in the United States in 2020.[1] In Europe, there were 0.82 and 0.77 million estimated new liver cancer cases and deaths in 2018,[2] whereas the number among Chinese patients were 0.46 and 0.31 million in 2015, respectively.[3] The risk factors contributing to HCC development include viral infections by hepatitis B virus, hepatitis C virus, cirrhosis by any causes such as alcoholic hepatitis or primary biliary cirrhosis, metabolism disorders, and intake of aflatoxin-contaminated food.[4–7] Nevertheless, considerable progress has been made in medicinal therapy for HCC, especially for immune checkpoint inhibitors (ICIs).

Nivolumab is approved to be administrated to patients with HCC who have progressed on or after standard sorafenib therapy, as supported by the CheckMate 040 trial, which showed that the disease control rates reached approximately 60%.[8] Pembrolizumab, another anti-PD-1 antibody, has also been approved for patients with HCC undergoing sorafenib therapy because of the KEYNOTE-240 trial, which showed that more than two-thirds of patients exhibited response to treatment.[9,10] Other clinical trials such as CheckMate 459 to evaluate the response to ICIs for HCC are in progress.

Long noncoding RNAs (lncRNAs), which do not code for proteins, refer to a type of RNA the transcript length of which is more than 200 nucleotides. lncRNAs are abundant and account for approximately 80% of the human transcriptome. They regulate 70% of the gene expression in humans and do not act in a universal manner because they can interact with DNA, RNA, and proteins and exhibit enhancement or inhibition effects. Both normal development and tumorigenesis such as HCC involve the regulation of lncRNAs.[11,12]

Received 1 August 2020; accepted 6 October 2020;
https://doi.org/10.1016/j.omtn.2020.10.002.

[6]These authors contributed equally to this work.

Correspondence: Jingdun Xie, Department of Anesthesiology, Sun Yat-Sen University Cancer Center, State Key Laboratory of Oncology in Southern China, Collaborative Innovation for Cancer Medicine, Guangzhou, Guangdong 510000, China.
E-mail: xiejd6@mail.sysu.edu.cn
Correspondence: Li Liang, Departments of Medical Oncology, Zhongshan Hospital of Fudan University, Shanghai 200032, China.
E-mail: liang.li@zs-hospital.sh.cn

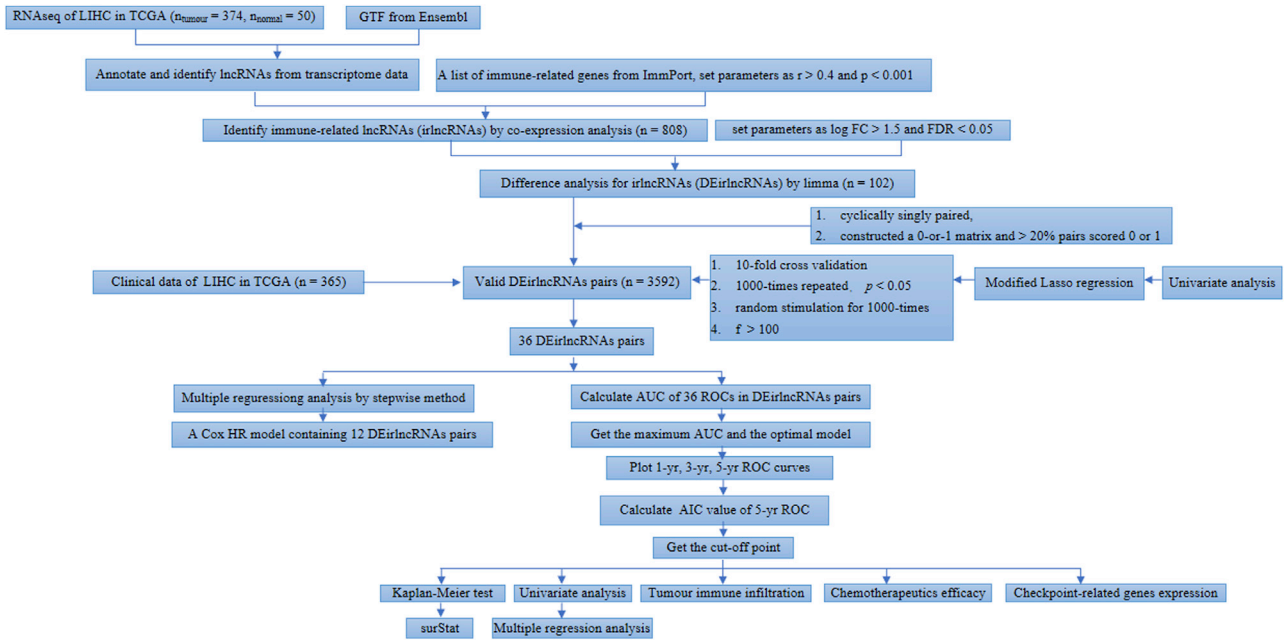


Figure 1. Flow Chart of This Study

Recent evidence has suggested that lncRNAs contribute to the malignant phenotypes of cancer not only through genomic or transcriptomic alterations but also by altering the immune microenvironment[13] because lncRNAs direct the expression of genes related to the activation of immune cells, thus resulting in tumor immune-cell infiltration.[14]

Signatures focusing on the tumor immune infiltration show promising predictive and prognostic value in diagnosis, evaluation, and treatment of cancer.[15–17] Moreover, lncRNAs contribute to the establishment of these signatures. Wei et al.[18] identified nine immune-related lncRNAs (irlncRNAs) and constructed a signature to evaluate the prognosis of patients with pancreatic cancer. Wu et al.[19] integrated mRNA, microRNA (miRNA), lncRNA, and clinical data and set up an immune-related signature that could identify a superior survival outcome among patients with head and neck squamous cell carcinoma. Jiang et al.[20] used a new modeling method and set up a three irlncRNA signatures to indicate the prognosis for clear cell renal cell carcinoma. Song et al.[21] built a signature based on eight lncRNAs and validated the immune-related genes *in situ* for bladder cancer.

In contrast, two-biomarker combinations are superior than simple genes in terms of the accuracy of a diagnostic model for cancers.[22] Few models have examined the role of lncRNA in this setting before. In this study, we utilized a novel modeling algorithm, paring, and iteration to construct an irlncRNA signature, which did not require any specific expression levels. Then, we estimated its predictive value among patients with HCC, as well as its diagnostic effectiveness, chemotherapeutic efficacy, and tumor immune infiltration.

RESULTS

Identification of Differentially Expressed irlncRNAs (DEirlncRNAs)

The process flow of this study is shown in Figure 1. First, we retrieved the transcriptome profiling data of hepatoma from the liver hepatocellular carcinoma (LIHC) project of the The Cancer Genome Atlas (TCGA) database; 50 normal and 374 tumor samples were included. Next, the data were annotated according to gene transfer format (GTF) files from Ensembl, and a co-expression analysis was performed between known ir-genes and lncRNAs. A total of 808 irlncRNAs were identified (shown in Table S1), and 102 were distinguished as DEirlncRNAs (Figure 2A), among which 75 were upregulated and 27 were downregulated (Figure 2B).

Establishment of DEirlncRNA Pairs and a Risk Assessment Model

Using an iteration loop and a 0-or-1 matrix screening among 102 DEirlncRNAs, 3,592 valid DEirlncRNA pairs were identified. After a single factor test followed by modified Lasso regression analysis, 36 DEirlncRNA pairs were extracted, 12 of which were included in a Cox proportional hazards model by the stepwise method (Figure 2C). Next, we calculated the areas under curve (AUCs) for each receiver operating characteristic (ROC) curve of 36 pairs, drew the curved line, and found the highest point referring to 0.904 (Figure 3A) to confirm the most ideal DEirlncRNA pair for this maximum AUC value (Figure 3B). To validate the optimality, we not only plotted the 1-, 3-, and 5-year ROC curves, which suggested that all AUC values were over 0.85 (Figure 3C) but also compared the 5-year ROC curves with other clinical

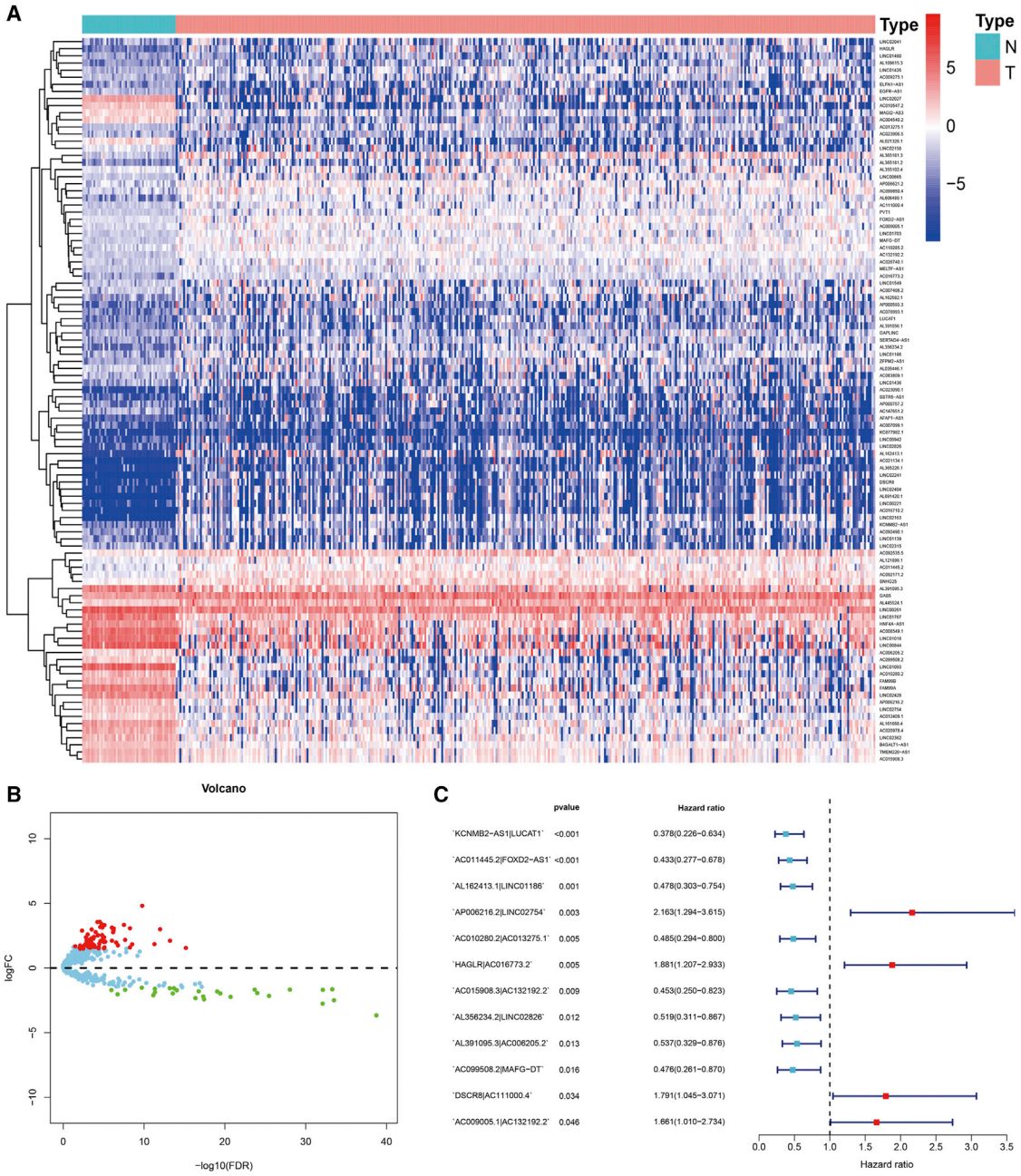


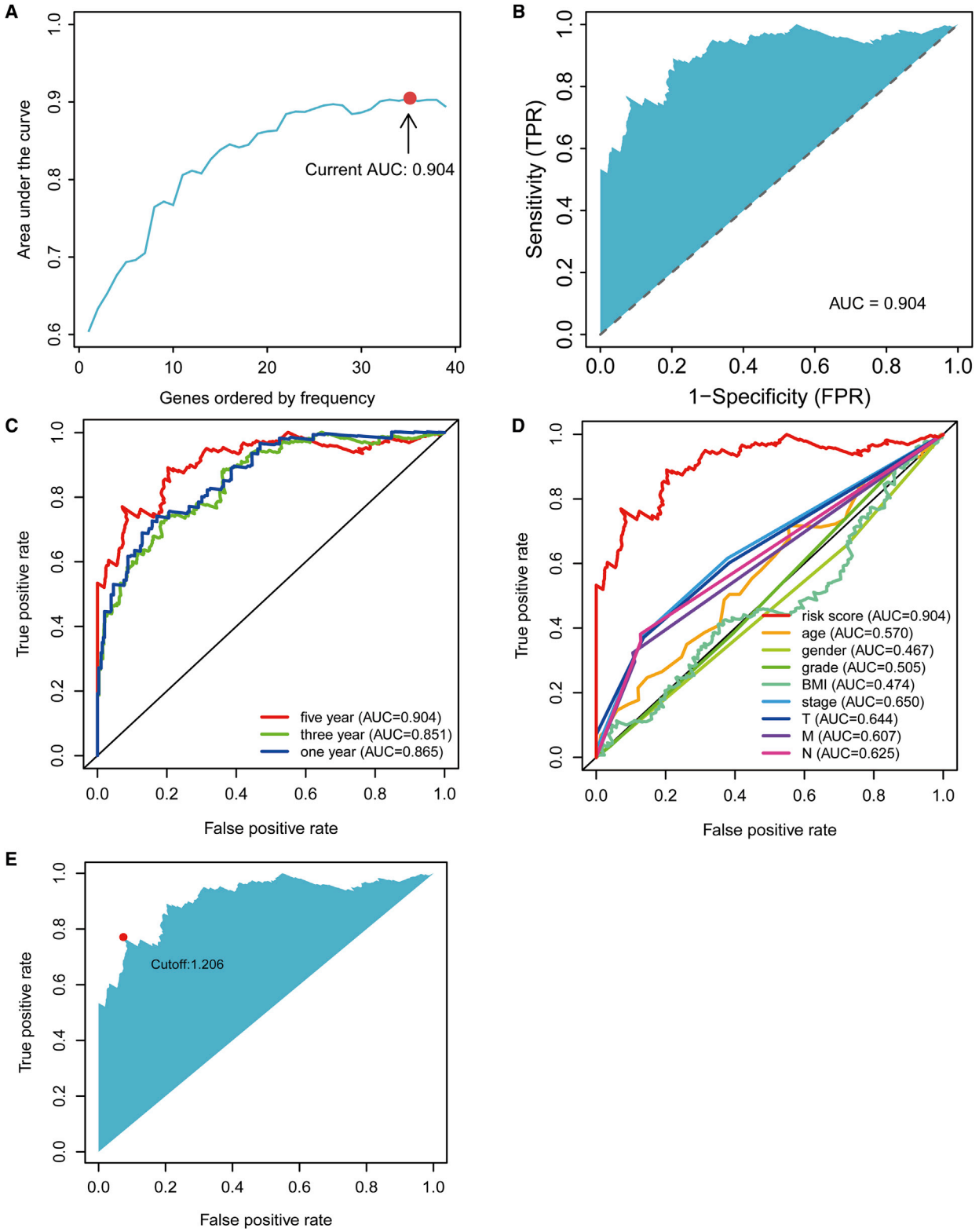
Figure 2. Establishment of a Risk Assessment Model using DEirlncRNA Pairs

Identification of differentially expressed immune-related lncRNAs (DEirlncRNAs) using TCGA datasets and annotation by Ensembl. (A and B) The heatmap (A) and volcano plot (B) are shown. (C) A forest map showed 12 DEirlncRNA pairs identified by Cox proportional hazard regression in the stepwise method.

characteristics (Figure 3D). We also recognized the maximum inflection point as the cut-off point on the 5-year ROC curve using the Akaike information criterion (AIC) values (Figure 3E). We collected data of 365 acceptable cases of patients with HCC from TCGA and calculated the risk scores for all of them. We used the identified cut-off point to re-distinguish high- and low-risk groups in the cohort for validation.

Clinical Evaluation by Risk Assessment Model

According to the cut-off point confirmed previously, 157 cases were classified into the high-risk group and 208 into the low-risk group. RiskScores and survival of each case are shown in Figure 4A and Figure 4B. These figures suggest that the clinical outcome of patients in the low-risk group was superior to that of patients in the high-risk group. Kaplan-Meier analysis showed that patients in the low-risk



(legend on next page)

group exhibited a longer survival time than those in the high-risk group ($p < 0.0001$) (Figure 4C). Then, we performed a series of chi-square tests to investigate the relationship between the risk of hepatocellular carcinoma and clinicopathological characteristics. The strip chart (Figure 5A) and consequent scatter diagrams obtained by the Wilcoxon signed-rank test showed that T stage (Figure 5B), M stage (Figure 5C), tumor grade (Figure 5D), clinical stage (Figure 5E), and survival status (Figure 5F) were significantly related to the risk. Next, we demonstrated that clinical stage ($p < 0.001$, HR = 1.717, 95% CI [1.3866–2.127]), T stage ($p < 0.001$, HR = 1.689, 95% CI [1.379–2.069]), and riskScore ($p < 0.001$, HR = 1.079, 95% CI [1.064–1.094]) showed statistical differences by univariate Cox regression analysis (Figure 5G), whereas only riskScore ($p < 0.001$, HR = 1.075, 95% CI [1.058–1.092]) presented as an independent prognostic predictor by multivariate Cox regression analysis (Figure 5H). The detailed values of univariate and multivariate Cox regression analyses are shown in Table S2.

Estimation of Tumor-Infiltrating Immune Cells and Immunosuppressive Molecules with Risk Assessment Model

Because lncRNAs and immune-related genes were initially connected, we consequently investigated whether the model was related to the tumor immune microenvironment. We revealed that the high-risk group was more positively associated with tumor-infiltrating immune cells such as macrophages, monocytes, and $CD8^+$ T cells, whereas they were negatively associated with fibroblasts and $CD4^+$ T cells, as revealed by the Wilcoxon signed-rank test (see Figure S1 and Table S3). A detailed Spearman correlation analysis was conducted, and the resulting diagram exhibited a lollipop shape, as shown in Figure 6A. The results are listed in Table S4. Because ICIs are administered for treating hepatocellular carcinoma in clinical practice, we investigated whether the risk model was related to ICI-related biomarkers and discovered that high risk scores were positively correlated with high expression of CTLA4 ($p < 0.05$, Figure 6B), HAVCR2 ($p < 0.001$, Figure 6C), LAG3 ($p > 0.05$, Figure 6D), and PDCD1 ($p > 0.05$, Figure 6E), whereas the latter two showed no statistical differences.

Analysis of the Correlation between the Risk Model and Chemotherapeutics

Besides checkpoint blockades therapy, we attempted to identify associations between risk and the efficacy of common chemotherapeutics in treating hepatocellular carcinoma in the TCGA project of the LIHC dataset. We showed that a high risk score was associated with a lower half inhibitory centration (IC_{50}) of chemotherapeutics such as doxorubicin ($p = 0.026$), mitomycin ($p < 0.001$), and cisplatin ($p < 0.001$), whereas it was associated with a higher IC_{50} for vinblastine ($p < 0.01$), which indicated that the model acted as a potential predictor for chemosensitivity (Figure 6F; Table S5).

DISCUSSION

Recent studies have focused on setting up signatures with both coding genes and noncoding RNAs to evaluate the prognosis of patients with malignant tumor.[23,24] However, most of these signatures are based on quantifying the expression levels of transcripts. In this study, we were inspired by the strategy of immune-related gene pairing and attempted to construct a reasonable model with two-lncRNA combinations and did not adopt their exact expression levels in the signature for the first time.[25]

First, we retrieved raw data of lncRNAs from TCGA, performed a differential co-expression analysis to classify DEirlncRNAs, and validated lncRNA-pairs using an improved method of cyclically single pairing along with a 0-or-1 matrix. Second, we performed univariate analysis combined with a modified Lasso penalized regression including procedures of cross validation, multi-times repeat, and random stimulation to determine DEirlncRNAs pairs. Third, we calculated not only each AUC value of ROC to get the optimal model but also counted the AIC value of each point on the AUC to distinguish the optimal cut-off point to differentiate the high- or low-risk-group among patients with HCC. Fourth, we evaluated this novel model under various clinical settings including survival, clinical pathological characteristics, tumor-infiltrating immune cells, chemotherapy, and checkpoint related biomarkers.

Deng et al.[26] analyzed the expression levels of nine lncRNAs to set up a signature for predicting the survival of patients with HCC. Generally, lncRNAs with high abundance possess significant biological function. Our algorithm suggested that we could initially identify DEirlncRNAs and establish the most significant irlncRNA pair. Thus, pairs with higher or lower expression only had to be detected instead of examining specific expression values of every lncRNA. This novel model had an advantage of clinical practicability to distinguish the high or low risk for clinical cases. Because lncRNAs are associated with immune-related genes, these lncRNAs were potentially involved in regulating the remodeling of the immune microenvironment or the activation of immune cells. Zhang et al.[27] investigated irlncRNAs from the LIHC project of TCGA and validated them in GSE76427 from GEO databases. In this study, some of the DEirlncRNAs in the process of modeling that have been already identified play an important role in malignant phenotypes of various cancer types, such as LUCAT1,[28–30] FOXD2-AS1,[31–33] HAGLR,[34–36] and DSCR8,[37] especially for HCC, while others were revealed for the first time. For instance, Jiao et al.[38] reported that lncRNA LUCAT1 exhibits clinical value and suggested an unfavorable survival outcome for patients with liver cancer. Hu revealed that FOXD2-AS1 accelerates HCC tumorigenesis by upregulating MAPK31 expression,[39] and Wang

Figure 3. Establishment of a Risk Assessment Model by DEirlncRNA Pairs

(A) Plot a curve of every AUC value generated by ROCs of 36 DEirlncRNA pair models and to identify the highest point of the AUC. (B) The ROC of the optimal DEirlncRNA pair models was related to the maximum AUC. (C) The 1-, 3-, and 5-year ROC of the optimal model suggested that all AUC values were over 0.85. (D) A comparison of 5-year ROC curves with other common clinical characteristics showed the superiority of the riskScore. (E) RiskScore for 365 patients with HCC; the maximum inflection point is the cut-off point obtained by the AIC.

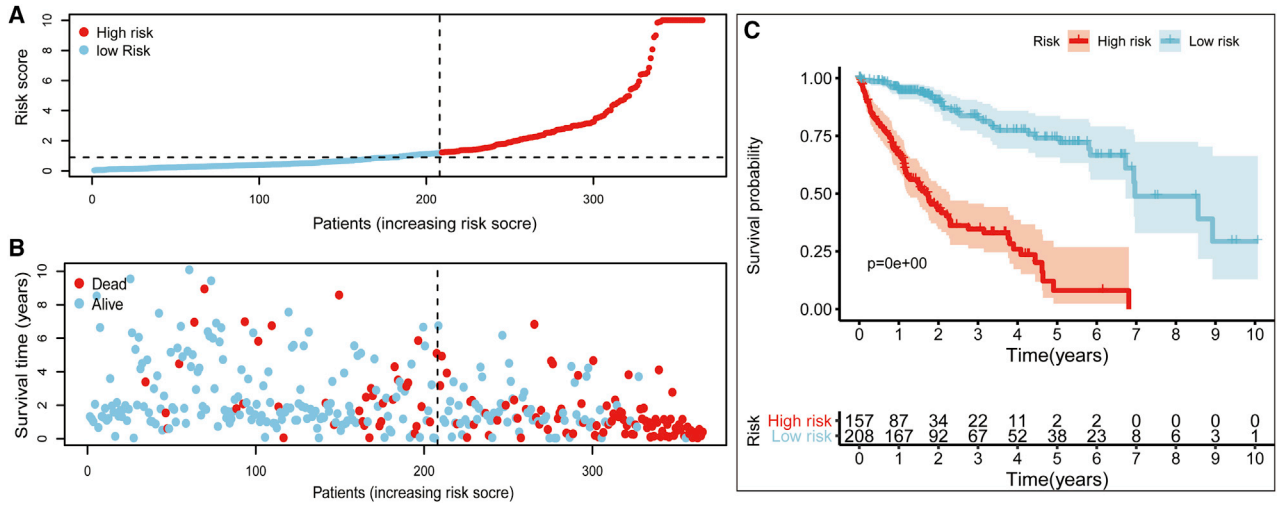


Figure 4. Risk Assessment Model for Prognosis Prediction

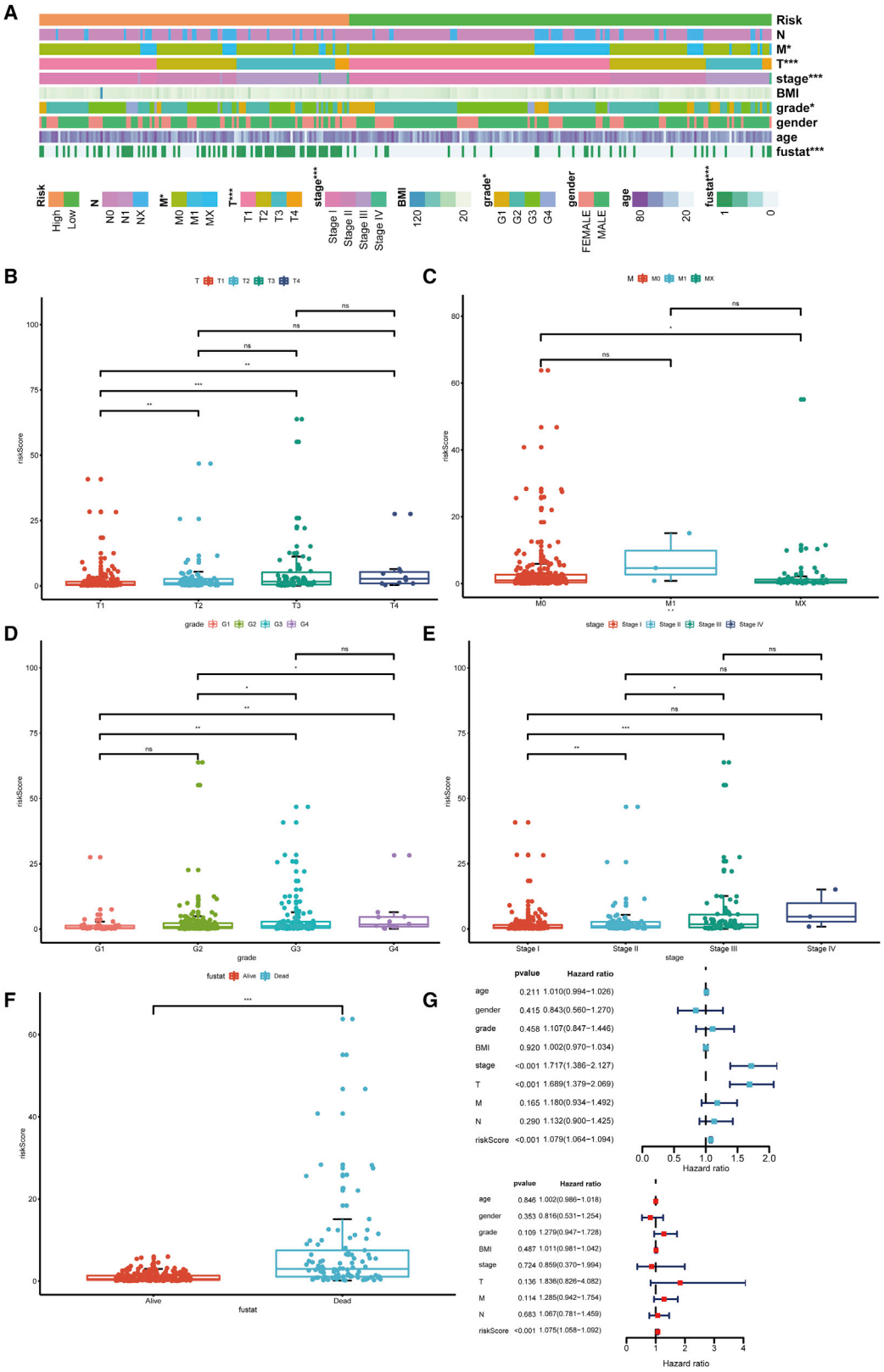
(A and B) Risk scores (A) and survival outcome (B) of each case are shown. (C) Patients in the low-risk group experienced a longer survival time tested by the Kaplan-Meier test.

et al.[40] disclosed that DSCR8 promotes HCC progression by activating Wnt/β-catenin cascades. Therefore, the proposed model can identify novel biomarkers for further research.

To improve the accuracy and efficacy of prediction on risk, we used a modified method of Lasso penalized modeling proposed by Sveen et al.[41] It showed that incorporating the factors into a Cox regression model based on the rank of occurrence frequency in the process of repeating instead of the intersection of occurrence, as the frequency suggesting the impact of the factor in the model. Moreover, we further improved the procedure of modeling as follows: we calculated every AUC value to identify the maximum value for an optimal model followed by comparison with other clinical parameters. We also used the AIC values to get an optimal cut-off point for model fitting instead of distinguishing the risk just by the median value. After distinguishing the high and low risk group by this novel upgradation, we reevaluated the survival outcome, performed univariate and multivariate analyses of clinicopathological characteristics, and analyzed the efficacy of chemotherapeutics for HCC treatment, tumor immune infiltration, and biomarkers related to checkpoint inhibitors, which implied that this modeling algorithm worked well.

Intertumoral tumor-infiltrating immune cells affect the response to anti-checkpoint blockades. Patients with more $CD8^+$ T cell infiltration experienced a superior treatment response from pembrolizumab than those with less infiltration, which was indicated by the KEYNOTE-001 study.[42] To explore the relationship between risk scores and tumor-infiltrating immune cells, we used seven common acceptable methods to estimate the immunoinfiltrating cell, including TIMER,[43,44] CIBERSORT,[45,46] XCELL,[47,48] QUANTISEQ,[49,50] MCPcounter,[51] EPIC,[52] and CIBERSORT-ABS.[53] Because of the complexity and defects of these algorithms, inter-comparison was rarely performed. By integrating analyzing, our results revealed that DEirlncRNA pairs were more positively related to tumor-infiltrating immune cells such as $CD8^+$ T cells, macrophages, monocytes, and myeloid dendritic cells. Wang et al.[54] reported that immune scores based on immunogenomic analysis could indicate therapeutic benefits for chemotherapy and immunotherapy. Yue et al.[55] suggested that gene signatures related to stromal infiltration were associated with metastasis of ovarian cancer, whereas Wu et al.[56] suggested that LINC00665 influenced the infiltration level of macrophages, dendritic cells, and inhibited regulatory T cells to prevent T cell exhaustion. Our model suggested that high risk was associated with sensitivity to chemotherapeutics such as doxorubicin, vinblastine, mitomycin, and cisplatin, instead of that to the multiple kinase inhibitor—sorafenib. We assumed that immunotherapy benefits better than chemotherapy through eliminating cancer cells, generating more neoantigens, and suppressing tumor development. Although signatures were correlated with checkpoint-related biomarkers such as CTLA4 and HAVCR2 rather than LAG3 and PDCD1, we believed that specific biomarkers instead of signatures should be identified and validated because of differences between various immune cells and immune-related phenotypes.[57] Furthermore, the abundant neoantigens caused by tumoral mutation, numerous T cells that recognize the neoantigens, and rich intertumoral infiltration by immune cells are critically important.

However, we recognized some shortcomings and limitations in this study. For example, the raw dataset for initial analysis was relatively insufficient because it was simply downloaded from TCGA. We failed to retrieve datasets simultaneously supporting information of lncRNA expression levels, clinicopathological characteristics, and survival outcomes for patients with HCC. Constructed models require external validation because of the difference in expression levels of each sample, which may make the final model unreliable. We exploratively constructed a 0-or-1 matrix to screen all lncRNA-pairs to minimize the sample errors due to variations of expression.



(legend on next page)

In addition, various methods were used to confirm this new modeling algorithm, which was optimal and utilized in this study. Based on these results, we assumed that our model was acceptable in spite of the lack of external data validation. However, external validation by other clinical datasets would be beneficial. Hence, in future work, we will recollect clinical samples and expand the sample size for further verification, the evaluation of which will be time-consuming.

In conclusion, this study demonstrated that a novel signature constructed by irlncRNAs that did not require prediction of lncRNA expression levels could predict prognosis for patients with HCC and might help in distinguishing those who could benefit from anti-tumor immunotherapy.

MATERIALS AND METHODS

Retrieval of Transcriptome Data, Preparation, and Differentially Expressed Analysis

Transcriptome profiling (*RNaseq*) data harmonized to fragments per-kilobase million (FPKM) of hepatoma from TCGA (https://tcga-data.nci.nih.gov/tcga/) for the LIHC project were downloaded. GTF files were downloaded from Ensembl (http://asia.ensembl.org) for annotation to distinguish the mRNAs and lncRNAs for subsequent analysis. A list of recognized immune-related genes (ir-genes) was downloaded from the ImmPort database (http://www.immport.org) and was used to screen irlncRNAs by a co-expression strategy. Correlation analysis was performed between ir-genes and all lncRNAs. Those with immune gene correlation coefficients more than 0.4 and p value less than 0.001 were considered as irlncRNAs. To identify the DEirlncRNA, we used *R* package *limma* for differential expression analysis among irlncRNAs. The thresholds were set as log fold change (FC) >1.5 along with false discovery rate (FDR) <0.05.

Pairing DEirlncRNAs

The DEirlncRNAs were cyclically singly paired, and a 0-or-1 matrix was constructed assuming C is equal to lncRNA A plus lncRNA B; C is defined as 1 if the expression level of lncRNA A is higher than lncRNA B, otherwise C is defined as 0. Then, the constructed 0-or-1 matrix was further screened. No relationship was considered between pairs and prognosis if the expression quantity of lncRNA pairs was 0 or 1 because pairs without a certain rank could not properly predict patient survival outcome. When the amounts of lncRNA-pairs of which expression quantity was 0 or 1 accounted for more than 20% of total pairs, it was considered a valid match.

Obtaining Clinical Data of Patients

Clinical data of patients with LIHC were retrieved from the LIHC project of TCGA. The valid data were extracted by excluding data with a follow up time of 0 days, as well as repeated data.

Establishment of a Risk Model to Evaluate the riskScore

A single factor analysis was first performed, following which, Lasso regression was performed with 10-fold cross validation and a p value of 0.05. The Lasso regression was run for 1,000 cycles, and for each cycle, a random stimulation was set up for 1,000 times. Next, the frequency of each pair in the 1,000-times-repeated Lasso regression model was recorded and pairs with frequency more than 100 times were selected for Cox proportional hazard regression analysis, as well as for construction of the model. The AUC value of each model was also calculated and was drawn as a curve. If the curve reached the highest point, indicating the maximum AUC value, the calculation procedure was terminated while the model was taken as the optimal candidate. The 1-, 3-, and 5-year ROC curves of the model were plotted. The following formula was used to calculate the riskScore with the constructed risk model for all clinical cases: $\text{RiskScore} = \sum_{i=1}^{k} \beta i Si$. The AIC values of every points of the 5-year ROC curve were evaluated to identify the maximum inflection point, which was considered as the cut-off point to distinguished high or low risk of RiskScores.

Validation of the Constructed Risk Model

To validate this cut-off point, we performed Kaplan-Meier analysis to show the survival difference of patients in the high- or low-risk group, and the survival curve was used for visualization. The specific riskScore values of each sample in the model were also visualized by *R* tools. The *R* packages utilized in these steps included *survival*, *glmnet*, *pbapply*, *survivalROC*, *survminer*, and *pHeatmap*.

To verify the clinical application value of the constructed model, we performed the chi-square test to analyze the relationship between the model and clinicopathological characteristics. The band diagram was used for visualization and was labeled as follows: <0.001 = ***, <0.01 = ** *, and <0.05 = *. Wilcoxon signed-rank test was used to calculate the riskScore differences among different groups of these clinicopathological characteristics. The box diagram was used to show the analysis results. To confirm whether the model can be used as an independent clinical prognostic predictor, we performed univariate and multivariate Cox regression analyses between the riskScore and clinicopathological characteristics. A forest map was used to demonstrate the results. The *R* packages utilized in these operations were survival, *pHeatmap*, and *ggupbr*.

Investigation of Tumor-Infiltrating Immune Cells

To analyze the relationship between the risk and immune-cell characteristics, we considered the currently acknowledged methods to calculate the immune infiltration statues among the samples from the TCGA project of the LIHC dataset including TIMER,

Figure 5. Clinical Evaluation by the Risk Assessment Model

(A–F) A strip chart (A) along with the scatter diagram showed that (B) T stage, (C) M stage, (D) tumor grade, (E) clinical stage, and (F) survival status were significantly associated with the riskScore. (G) A univariate Cox hazard ratio analysis demonstrated that clinical stage ($p < 0.001$, HR = 1.717, 95% CI [1.3866–2.127]), T stage ($p < 0.001$, HR = 1.689, 95% CI [1.379–2.069]), and riskScore ($p < 0.001$, HR = 1.079, 95% CI [1.064–1.094]) were statistically different. (H) Only riskScore ($p < 0.001$, HR = 1.075, 95% CI [1.058–1.092]) presented as an independent prognostic predictor by multivariate Cox regression.

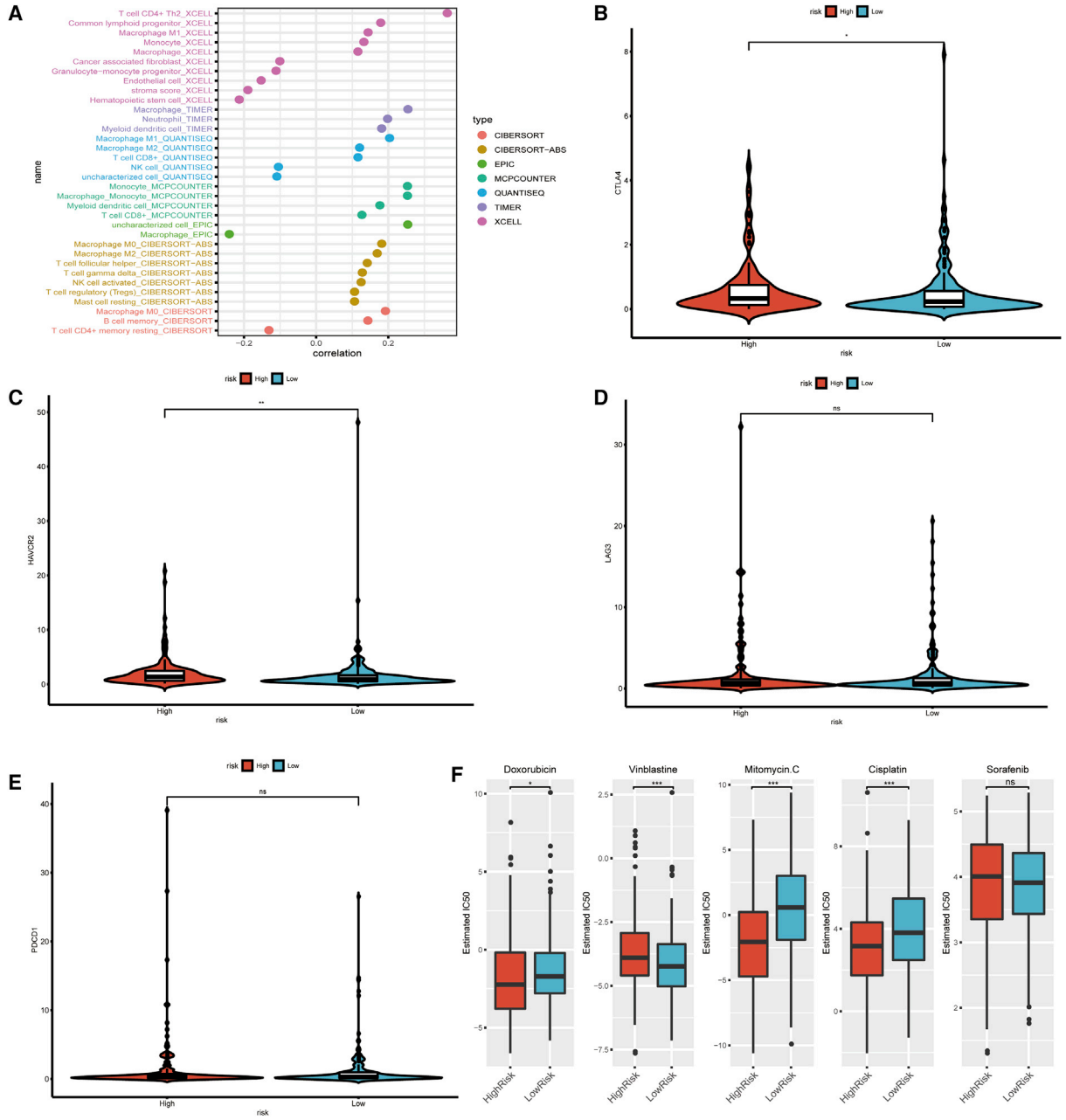


Figure 6. Estimation of Tumor-Infiltrating Cells and Immunosuppressed Molecules by the Risk Assessment Model

(A) Patients in the high-risk group were more positively associated with tumor-infiltrating immune cells such as macrophages, monocytes, and $CD8^+$ T cells, whereas they were negatively associated with fibroblasts and $CD4^+$ T cells, as shown by Spearman correlation analysis. (B–E) High risk scores were positively correlated with upregulated (B) CTLA4, (C) HAVCR2, (D) LAG3, and (E) PDCD1 levels, whereas the latter two showed no statistical difference in patients with HCC. (F) The model acted as a potential predictor for chemosensitivity as high risk scores were related to a lower IC_{50} for chemotherapeutics such as doxorubicin, mitomycin, and cisplatin, whereas they were related to a higher IC_{50} for vinblastine.

CIBERSORT, XCELL, QUANTISEQ, MCPcounter, EPIC, and CIBERSORT. The differences in immune infiltrating cell content explored by these methods between high- and low-risk groups of the constructed model were analyzed by Wilcoxon signed-rank test; the results are shown in a box chart. Spearman correlation analysis was performed to analyze the relationship between the riskScore values and the immune infiltrated cells. The correlation coefficients of the results were shown in a lollipop diagram. The significance threshold was set as $p < 0.05$. The procedure was performed using R *ggplot2* packages.

Exploration of the Significance of the Model in the Clinical Treatment

To evaluate the model in the clinic for hepatocellular carcinoma treatment, we calculated the IC_{50} of common administrating chemotherapeutic drugs in the TCGA project of the LIHC dataset. Antitumor drugs such as adriamycin, mitomycin, vinblastine, cisplatin, and sorafenib are recommended for liver cancer treatment by AJCC guidelines. The difference in the IC_{50} between the high and low risk groups was compared by Wilcoxon signed-rank test and the results are shown as box drawings obtained using with pRRophetic and ggplot2 of R.

Analyses of the Immunosuppressive Molecules Expressing Related to ICIs

To study the relationship between the model and the expression level of genes related to ICIs, we performed ggstatsplot package and violin plot visualization.

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.omtn.2020.10.002.

AUTHOR CONTRIBUTIONS

W.H. and L.L. conducted the formal analysis and wrote the original draft; W.H., L.L., and J.X. performed the project administration; Y.G. and Z.Q. participated in software; H.Q. conducted data curation; X.Y., W.Z., and L.M. contributed to writing, reviewing, and editing the article; and J.X. provided funding acquisition. All authors read and approved the final submitted manuscript.

CONFLICTS OF INTEREST

The authors declare no competing interests.

ACKNOWLEDGMENTS

This study was supported by the GuangDong Natural Science Foundation for Distinguished Young Scholars (grant number 2019B151502010), National Natural Science Foundation of China (grant number 81870878), and GuangDong Natural Science Foundation (2018A030313092).

REFERENCES

1. Siegel, R.L., Miller, K.D., and Jemal, A. (2020). Cancer statistics, 2020. CA Cancer J. Clin. *70*, 7–30.
2. Ferlay, J., Colombet, M., Soerjomataram, I., Dyba, T., Randi, G., Bettio, M., Gavin, A., Visser, O., and Bray, F. (2018). Cancer incidence and mortality patterns in Europe: Estimates for 40 countries and 25 major cancers in 2018. Eur. J. Cancer *103*, 356–387.
3. Chen, W., Zheng, R., Baade, P.D., Zhang, S., Zeng, H., Bray, F., Jemal, A., Yu, X.Q., and He, J. (2016). Cancer statistics in China, 2015. CA Cancer J. Clin. *66*, 115–132.
4. Kulik, L., and El-Serag, H.B. (2019). Epidemiology and Management of Hepatocellular Carcinoma. Gastroenterology *156*, 477–491.
5. Singal, A.G., Lampertico, P., and Nahon, P. (2020). Epidemiology and surveillance for hepatocellular carcinoma: New trends. J. Hepatol. *72*, 250–261.
6. McGlynn, K.A., Petrick, J.L., and El-Serag, H.B. (2020). Epidemiology of Hepatocellular Carcinoma. Hepatology.
7. Mak, L.Y., Wong, D.K., Pollicino, T., Raimondo, G., Hollinger, F.B., and Yuen, M.F. (2020). Occult hepatitis B infection and hepatocellular carcinoma: Epidemiology, virology, hepatocarcinogenesis and clinical significance. J. Hepatol. *73*, 952–964.
8. El-Khoueiry, A.B., Sangro, B., Yau, T., Crocenzi, T.S., Kudo, M., Hsu, C., Kim, T.Y., Choo, S.P., Trojan, J., Welling, T.H.R., et al. (2017). Nivolumab in patients with advanced hepatocellular carcinoma (CheckMate 040): an open-label, non-comparative, phase 1/2 dose escalation and expansion trial. Lancet *389*, 2492–2502.
9. Zhu, A.X., Finn, R.S., Edeline, J., Cattan, S., Ogasawara, S., Palmer, D., Verslype, C., Zagonel, V., Fartoux, L., Vogel, A., et al.; KEYNOTE-224 investigators (2018). Pembrolizumab in patients with advanced hepatocellular carcinoma previously treated with sorafenib (KEYNOTE-224): a non-randomised, open-label phase 2 trial. Lancet Oncol. *19*, 940–952.
10. Finn, R.S., Ryoo, B.Y., Merle, P., Kudo, M., Bouattour, M., Lim, H.Y., Breder, V., Edeline, J., Chao, Y., Ogasawara, S., et al.; KEYNOTE-240 investigators (2020). Pembrolizumab As Second-Line Therapy in Patients With Advanced Hepatocellular Carcinoma in KEYNOTE-240: A Randomized, Double-Blind, Phase III Trial. J. Clin. Oncol. *38*, 193–202.
11. Schmitt, A.M., and Chang, H.Y. (2016). Long Noncoding RNAs in Cancer Pathways. Cancer Cell *29*, 452–463.
12. Zhao, J., and Lawless, M.W. (2013). Long noncoding RNAs and their role in the liver cancer axis. Nat. Rev. Gastroenterol. Hepatol. *10*, 703.
13. Atianand, M.K., Caffrey, D.R., and Fitzgerald, K.A. (2017). Immunobiology of Long Noncoding RNAs. Annu. Rev. Immunol. *35*, 177–198.
14. Chen, Y.G., Satpathy, A.T., and Chang, H.Y. (2017). Gene regulation in the immune system by long noncoding RNAs. Nat. Immunol. *18*, 962–972.
15. Peng, D., Wang, L., Li, H., Cai, C., Tan, Y., Xu, B., and Le, H. (2019). An immune infiltration signature to predict the overall survival of patients with colon cancer. IUBMB Life *71*, 1760–1770.
16. Zhang, B., Wang, Q., Fu, C., Jiang, C., and Ma, S. (2019). Exploration of the immune-related signature and immune infiltration analysis for breast ductal and lobular carcinoma. Ann. Transl. Med. *7*, 730.
17. Wei, Z.W., Wu, J., Huang, W.B., Li, J., Lu, X.F., Yuan, Y.J., Xiong, W.J., Zhang, X.H., Wang, W., He, Y.L., and Zhang, C.H. (2020). Immune-infiltration based signature as a novel prognostic biomarker in gastrointestinal stromal tumour. EBioMedicine *57*, 102850.
18. Wei, C., Liang, Q., Li, X., Li, H., Liu, Y., Huang, X., Chen, X., Guo, Y., and Li, J. (2019). Bioinformatics profiling utilized a nine immune-related long noncoding RNA signature as a prognostic target for pancreatic cancer. J. Cell. Biochem. *120*, 14916–14927.
19. Wu, S., Dai, X., and Xie, D. (2019). Identification and Validation of an Immune-Related RNA Signature to Predict Survival of Patients With Head and Neck Squamous Cell Carcinoma. Front. Genet. *10*, 1252.
20. Jiang, Y., Gou, X., Wei, Z., Tan, J., Yu, H., Zhou, X., and Li, X. (2020). Bioinformatics profiling integrating a three immune-related long non-coding RNA signature as a prognostic model for clear cell renal cell carcinoma. Cancer Cell Int. *20*, 166.
21. Song, Y., Jin, D., Chen, J., Luo, Z., Chen, G., Yang, Y., and Liu, X. (2020). Identification of an immune-relatedlLong non-coding RNA signature and nomogram as prognostic target for muscle-invasive bladder cancer. Aging (Albany NY) *12*, 12051–12073.
22. Lv, Y., Lin, S.-Y., Hu, F.-F., Ye, Z., Zhang, Q., Wang, Y., and Guo, A.-Y. (2019). Landscape of cancer diagnostic biomarkers from specifically expressed genes. Brief. Bioinform. Published online December 8, 2019. https://doi.org/10.1093/bib/bbz131.
23. Zhu, X., Tian, X., Yu, C., Shen, C., Yan, T., Hong, J., Wang, Z., Fang, J.Y., and Chen, H. (2016). A long non-coding RNA signature to improve prognosis prediction of gastric cancer. Mol. Cancer *15*, 60.
24. Qu, L., Wang, Z.L., Chen, Q., Li, Y.M., He, H.W., Hsieh, J.J., Xue, S., Wu, Z.J., Liu, B., Tang, H., et al. (2018). Prognostic Value of a Long Non-coding RNA Signature in Localized Clear Cell Renal Cell Carcinoma. Eur. Urol. *74*, 756–763.
25. Gruss, I., Twardowski, J.P., and Cierpisz, M. (2019). The Effects of Locality and Host Plant on the Body Size of *Aeolothrips intermedius* (Thysanoptera: Aeolothripidae) in the Southwest of Poland. Insects *10*, 266.

26. Deng, B., Yang, M., Wang, M., and Liu, Z. (2020). Development and validation of 9-long Non-coding RNA signature to predicting survival in hepatocellular carcinoma. Medicine (Baltimore) *99*, e20422.

27. Zhang, Y., Zhang, L., Xu, Y., Wu, X., Zhou, Y., and Mo, J. (2020). Immune-related long noncoding RNA signature for predicting survival and immune checkpoint blockade in hepatocellular carcinoma. J. Cell. Physiol. *235*, 9304–9316.

28. Luzón-Toro, B., Fernández, R.M., Martos-Martínez, J.M., Rubio-Manzanares-Dorado, M., Antiñolo, G., and Borrego, S. (2019). LncRNA LUCAT1 as a novel prognostic biomarker for patients with papillary thyroid cancer. Sci. Rep. *9*, 14374.

29. Huan, L., Guo, T., Wu, Y., Xu, L., Huang, S., Xu, Y., Liang, L., and He, X. (2020). Hypoxia induced LUCAT1/PTBP1 axis modulates cancer cell viability and chemotherapy response. Mol. Cancer *19*, 11.

30. Campbell, A., Mohl, J.E., Gutierrez, D.A., Varela-Ramirez, A., and Boland, T. (2020). Thermal Bioprinting Causes Ample Alterations of Expression of LUCAT1, IL6, CCL26, and NRN1L Genes and Massive Phosphorylation of Critical Oncogenic Drug Resistance Pathways in Breast Cancer Cells. Front. Bioeng. Biotechnol. *8*, 82.

31. Xu, T.P., Wang, W.Y., Ma, P., Shuai, Y., Zhao, K., Wang, Y.F., Li, W., Xia, R., Chen, W.M., Zhang, E.B., and Shu, Y.Q. (2018). Upregulation of the long noncoding RNA FOXD2-AS1 promotes carcinogenesis by epigenetically silencing EphB3 through EZH2 and LSD1, and predicts poor prognosis in gastric cancer. Oncogene *37*, 5020–5036.

32. Li, R., Chen, S., Zhan, J., Li, X., Liu, W., Sheng, X., Lu, Z., Zhong, R., Chen, L., Luo, X., et al. (2020). Long noncoding RNA FOXD2-AS1 enhances chemotherapeutic resistance of laryngeal squamous cell carcinoma via STAT3 activation. Cell Death Dis. *11*, 41.

33. Zhou, G., Huang, Z., Meng, Y., Jin, T., Liang, Y., and Zhang, B. (2020). Upregulation of long non-coding RNA FOXD2-AS1 promotes progression and predicts poor prognosis in tongue squamous cell carcinoma. J. Oral Pathol. Med. Published online June 12, 2020. https://doi.org/10.1111/jop.13074.

34. Lu, C., Ma, J., and Cai, D. (2017). Increased HAGLR expression promotes non-small cell lung cancer proliferation and invasion via enhanced de novo lipogenesis. Tumour Biol. *39*, 1010428317697574.

35. Guo, X., Chen, Z., Zhao, L., Cheng, D., Song, W., and Zhang, X. (2019). Long non-coding RNA-HAGLR suppressed tumor growth of lung adenocarcinoma through epigenetically silencing E2F1. Exp. Cell Res. *382*, 111461.

36. Yang, C., Shen, S., Zheng, X., Ye, K., Sun, Y., Lu, Y., and Ge, H. (2019). Long noncoding RNA HAGLR acts as a microRNA-143-5p sponge to regulate epithelial-mesenchymal transition and metastatic potential in esophageal cancer by regulating LAMP3. FASEB J. *33*, 10490–10504.

37. You, Q., Yao, Y., Wu, J., Cheng, C., Li, Y., and Yuan, H. (2020). YY1-induced lncRNA DSCR8 promotes the progression of ovarian cancer via miR-3192-5p/YY1 axis. Biomed. Pharmacother. *129*, 110339.

38. Jiao, Y., Li, Y., Ji, B., Cai, H., and Liu, Y. (2019). Clinical Value of lncRNA LUCAT1 Expression in Liver Cancer and its Potential Pathways. J. Gastrointestin. Liver Dis. *28*, 439–447.

39. Hu, W., Feng, H., Xu, X., Huang, X., Huang, X., Chen, W., Hao, L., and Xia, W. (2020). Long noncoding RNA FOXD2-AS1 aggravates hepatocellular carcinoma tumorigenesis by regulating the miR-206/MAP3K1 axis. Cancer Med. *9*, 5620–5631.

40. Wang, Y., Sun, L., Wang, L., Liu, Z., Li, Q., Yao, B., Wang, C., Chen, T., Tu, K., and Liu, Q. (2018). Long non-coding RNA DSCR8 acts as a molecular sponge for miR-485-5p to activate Wnt/β-catenin signal pathway in hepatocellular carcinoma. Cell Death Dis. *9*, 851.

41. Sveen, A., Ågesen, T.H., Nesbakken, A., Meling, G.I., Rognum, T.O., Liestøl, K., Skotheim, R.I., and Lothe, R.A. (2012). ColoGuidePro: a prognostic 7-gene expression signature for stage III colorectal cancer patients. Clin. Cancer Res. *18*, 6001–6010.

42. Garon, E.B., Hellmann, M.D., Rizvi, N.A., Carcereny, E., Leighl, N.B., Ahn, M.J., Eder, J.P., Balmanoukian, A.S., Aggarwal, C., Horn, L., et al. (2019). Five-Year Overall Survival for Patients With Advanced Non–Small-Cell Lung Cancer Treated With Pembrolizumab: Results From the Phase I KEYNOTE-001 Study. J. Clin. Oncol. *37*, 2518–2527.

43. Li, T., Fan, J., Wang, B., Traugh, N., Chen, Q., Liu, J.S., Li, B., and Liu, X.S. (2017). TIMER: A Web Server for Comprehensive Analysis of Tumor-Infiltrating Immune Cells. Cancer Res. *77*, e108–e110.

44. Li, T., Fu, J., Zeng, Z., Cohen, D., Li, J., Chen, Q., Li, B., and Liu, X.S. (2020). TIMER2.0 for analysis of tumor-infiltrating immune cells. Nucleic Acids Res. *48* (W1), W509–W514.

45. Chen, B., Khodadoust, M.S., Liu, C.L., Newman, A.M., and Alizadeh, A.A. (2018). Profiling Tumor Infiltrating Immune Cells with CIBERSORT. Methods Mol. Biol. *1711*, 243–259.

46. Zhang, H., Li, R., Cao, Y., Gu, Y., Lin, C., Liu, X., Lv, K., He, X., Fang, H., Jin, K., et al. (2020). Poor Clinical Outcomes and Immunoevasive Contexture in Intratumoral IL-10-Producing Macrophages Enriched Gastric Cancer Patients. Ann Surg. Published online June 11, 2020. https://doi.org/10.1097/SLA.0000000000004037.

47. Aran, D. (2020). Cell-Type Enrichment Analysis of Bulk Transcriptomes Using xCell. Methods Mol. Biol. *2120*, 263–276.

48. Aran, D., Hu, Z., and Butte, A.J. (2017). xCell: digitally portraying the tissue cellular heterogeneity landscape. Genome Biol. *18*, 220.

49. Finotello, F., Mayer, C., Plattner, C., Laschober, G., Rieder, D., Hackl, H., Krogsdam, A., Loncova, Z., Posch, W., Wilflingseder, D., et al. (2019). Molecular and pharmacological modulators of the tumor immune contexture revealed by deconvolution of RNA-seq data. Genome Med. *11*, 34.

50. Plattner, C., Finotello, F., and Rieder, D. (2020). Deconvoluting tumor-infiltrating immune cells from RNA-seq data using quanTIseq. Methods Enzymol. *636*, 261–285.

51. Dienstmann, R., Villacampa, G., Sveen, A., Mason, M.J., Niedzwiecki, D., Nesbakken, A., Moreno, V., Warren, R.S., Lothe, R.A., and Guinney, J. (2019). Relative contribution of clinicopathological variables, genomic markers, transcriptomic subtyping and microenvironment features for outcome prediction in stage II/III colorectal cancer. Ann. Oncol. *30*, 1622–1629.

52. Racle, J., de Jonge, K., Baumgaertner, P., Speiser, D.E., and Gfeller, D. (2017). Simultaneous enumeration of cancer and immune cell types from bulk tumor gene expression data. eLife *6*, e26476.

53. Tamminga, M., Hiltermann, T.J.N., Schuuring, E., Timens, W., Fehrmann, R.S., and Groen, H.J. (2020). Immune microenvironment composition in non-small cell lung cancer and its association with survival. Clin. Transl. Immunology *9*, e1142.

54. Dai, G.P., Wang, L.P., Wen, Y.Q., Ren, X.Q., and Zuo, S.G. (2020). Identification of key genes for predicting colorectal cancer prognosis by integrated bioinformatics analysis. Oncol. Lett. *19*, 388–398.

55. Yue, H., Wang, J., Chen, R., Hou, X., Li, J., and Lu, X. (2019). Gene signature characteristic of elevated stromal infiltration and activation is associated with increased risk of hematogenous and lymphatic metastasis in serous ovarian cancer. BMC Cancer *19*, 1266.

56. Wu, M., Shang, X., Sun, Y., Wu, J., and Liu, G. (2020). Integrated analysis of lymphocyte infiltration-associated lncRNA for ovarian cancer via TCGA, GTEx and GEO datasets. PeerJ *8*, e8961.

57. Zhang, B., Wu, Q., Li, B., Wang, D., Wang, L., and Zhou, Y.L. (2020). m^6A regulator-mediated methylation modification patterns and tumor microenvironment infiltration characterization in gastric cancer. Mol. Cancer *19*, 53.

OMTN, Volume *22*

Supplemental Information

Immune-Related lncRNA to Construct Novel Signature and Predict the Immune Landscape of Human Hepatocellular Carcinoma

Weifeng Hong, Li Liang, Yujun Gu, Zhenhua Qi, Haibo Qiu, Xiaosong Yang, Weian Zeng, Liheng Ma, and Jingdun Xie

Figure S1. The representative results of the evaluation of tumor infiltrating immune cells with risk assessment model.

Table S1. A total of 808 immune-related lncRNAs were identified by co-expression analysis (submitted as a separate Excel file).

Table S2. The detail values of univariate and multivariate cox regression analysis.

id	Univariate analysis			Multivariate analysis		
	HR	95% CI	P	HR	95% CI	P
age	1.01	0.99 ~1.03	0.211	1.00	0.99 ~1.02	0.846
gender	0.84	0.56 ~1.27	0.415	0.82	0.53 ~1.25	0.353
grade	1.11	0.85 ~1.45	0.458	1.28	0.95 ~1.73	0.109
BMI	1.00	0.97 ~1.03	0.920	1.01	0.98 ~1.04	0.487
stage	1.72	1.39 ~2.13	0.000	0.86	0.37 ~1.99	0.724
T	1.69	1.38 ~2.07	0.000	1.84	0.83 ~4.08	0.136
M	1.18	0.93 ~1.49	0.165	1.29	0.94 ~1.75	0.114
N	1.13	0.90 ~1.42	0.290	1.07	0.78 ~1.46	0.683
riskScore	1.08	1.06 ~1.09	0.000	1.07	1.06 ~1.09	0.000

Table S3. The p value of comparing tumour infiltrating immune cells and risk sore.

gene	pVal	gene	pVal
Macrophage_EPIC	0.0000	Myeloid dendritic cell resting_CIBERSORT-ABS	0.1448
Hematopoietic stem cell_XCELL	0.0000	B cell plasma_CIBERSORT-ABS	0.1491
T cell CD4+ Th2_XCELL	0.0000	B cell_TIMER	0.1576
stroma score_XCELL	0.0000	T cell CD4+ central memory_XCELL	0.1594
Macrophage M0_CIBERSORT	0.0000	Myeloid dendritic cell activated_CIBERSORT-ABS	0.1645
Macrophage M0_CIBERSORT-ABS	0.0000	Myeloid dendritic cell activated_CIBERSORT	0.1653
uncharacterized cell_EPIC	0.0000	B cell_XCELL	0.1784
Endothelial cell_XCELL	0.0000	NK cell_XCELL	0.1878
Granulocyte-monocyte progenitor_XCELL	0.0000	Endothelial cell_MCPCOUNTER	0.1943
Common lymphoid progenitor_XCELL	0.0000	T cell CD4+ naive_XCELL	0.2013
T cell CD8+ naive_XCELL	0.0001	Myeloid dendritic cell resting_CIBERSORT	0.2080
microenvironment score_XCELL	0.0002	B cell plasma_CIBERSORT	0.2159
Monocyte_QUANTISEQ	0.0002	cytotoxicity score_MCPCOUNTER	0.2217
Monocyte_MCPCOUNTER	0.0006	T cell CD4+ (non-regulatory)_QUANTISEQ	0.2225
Macrophage_Monocyte_MCPCOUNTER	0.0006	Cancer associated fibroblast_XCELL	0.2226
Neutrophil_TIMER	0.0006	Macrophage M2_CIBERSORT	0.2282
T cell CD8+ central memory_XCELL	0.0006	Macrophage M1_XCELL	0.2295
Macrophage M1_QUANTISEQ	0.0009	T cell CD4+ naive_CIBERSORT	0.2385
Macrophage_TIMER	0.0009	T cell CD4+ naive_CIBERSORT-ABS	0.2418
Cancer associated fibroblast_EPIC	0.0017	Myeloid dendritic cell_MCPCOUNTER	0.2433
T cell CD4+_TIMER	0.0018	Mast cell_XCELL	0.2638
Myeloid dendritic cell_TIMER	0.0022	T cell CD8+_QUANTISEQ	0.2886
T cell CD8+_CIBERSORT	0.0026	T cell follicular helper_CIBERSORT	0.2923
Common myeloid progenitor_XCELL	0.0036	NK cell activated_CIBERSORT	0.2931
Monocyte_CIBERSORT	0.0050	Neutrophil_XCELL	0.2948
T cell CD4+ memory_XCELL	0.0091	Class-switched memory B cell_XCELL	0.3058
Mast cell resting_CIBERSORT-ABS	0.0092	B cell naive_XCELL	0.3126
Mast cell resting_CIBERSORT	0.0109	B cell_EPIC	0.3144
Mast cell activated_CIBERSORT	0.0109	T cell CD4+ Th1_XCELL	0.3271
Neutrophil_CIBERSORT-ABS	0.0110	Myeloid dendritic cell_QUANTISEQ	0.3822
T cell regulatory (Tregs)_CIBERSORT-ABS	0.0130	B cell naive_CIBERSORT	0.4136
Neutrophil_CIBERSORT	0.0149	T cell CD8+_MCPCOUNTER	0.4202
Mast cell activated_CIBERSORT-ABS	0.0155	Macrophage M2_QUANTISEQ	0.4722
Macrophage M2_XCELL	0.0161	T cell CD8+ effector memory_XCELL	0.4723
T cell CD4+ memory resting_CIBERSORT	0.0182	NK cell activated_CIBERSORT-ABS	0.4736
B cell plasma_XCELL	0.0263	T cell regulatory (Tregs)_XCELL	0.4944
Plasmacytoid dendritic cell_XCELL	0.0279	Macrophage M1_CIBERSORT	0.4969
T cell CD4+ effector memory_XCELL	0.0281	Macrophage M1_CIBERSORT-ABS	0.5200
T cell CD8+_XCELL	0.0292	Myeloid dendritic cell_XCELL	0.5488
T cell regulatory (Tregs)_CIBERSORT	0.0313	T cell CD4+_EPIC	0.5665

NK cell_MCPCOUNTER	0.0331	B cell naive_CIBERSORT-ABS	0.6310
T cell NK_XCELL	0.0333	Monocyte_XCELL	0.6611
T cell regulatory (Tregs)_QUANTISEQ	0.0375	T cell gamma delta_CIBERSORT	0.7033
Macrophage M2_CIBERSORT-ABS	0.0388	T cell gamma delta_CIBERSORT-ABS	0.7143
Eosinophil_XCELL	0.0398	Eosinophil_CIBERSORT	0.7445
Neutrophil_QUANTISEQ	0.0424	Eosinophil_CIBERSORT-ABS	0.7475
Endothelial cell_EPIC	0.0455	T cell CD4+ memory resting_CIBERSORT-ABS	0.7831
T cell CD4+ memory activated_CIBERSORT-ABS	0.0460	B cell memory_CIBERSORT	0.7858
T cell CD8+_TIMER	0.0471	NK cell resting_CIBERSORT	0.7866
T cell CD4+ memory activated_CIBERSORT	0.0502	Macrophage_XCELL	0.7871
Neutrophil_MCPCOUNTER	0.0521	Myeloid dendritic cell activated_XCELL	0.7963
T cell follicular helper_CIBERSORT-ABS	0.0681	B cell memory_XCELL	0.7972
Monocyte_CIBERSORT-ABS	0.0910	B cell memory_CIBERSORT-ABS	0.7988
T cell_MCPCOUNTER	0.0946	NK cell resting_CIBERSORT-ABS	0.8092
T cell CD8+_CIBERSORT-ABS	0.1098	immune score_XCELL	0.8620
B cell_QUANTISEQ	0.1197	T cell CD8+_EPIC	0.9047
NK cell_QUANTISEQ	0.1210	T cell gamma delta_XCELL	0.9296
T cell CD4+ (non-regulatory)_XCELL	0.1241	NK cell_EPIC	0.9377
uncharacterized cell_QUANTISEQ	0.1294	B cell_MCPCOUNTER	0.9477
Cancer associated fibroblast_MCPCOUNTER	0.1307		

Table S4. The detail comparison results of correlation ship between tumour infiltrating immune cells and risk sore.

symbol	type	correlation	pValue
B cell memory_CIBERSORT	CIBERSORT	0.1407	0.0071
Macrophage M0_CIBERSORT	CIBERSORT	0.1894	0.0003
T cell CD4+ memory resting_CIBERSORT	CIBERSORT	-0.1336	0.0106
Macrophage M0_CIBERSORT-ABS	CIBERSORT-ABS	0.1791	0.0006
Macrophage M2_CIBERSORT-ABS	CIBERSORT-ABS	0.1665	0.0014
Mast cell resting_CIBERSORT-ABS	CIBERSORT-ABS	0.1037	0.0476
NK cell activated_CIBERSORT-ABS	CIBERSORT-ABS	0.1220	0.0197
T cell follicular helper_CIBERSORT-ABS	CIBERSORT-ABS	0.1390	0.0078
T cell gamma delta_CIBERSORT-ABS	CIBERSORT-ABS	0.1253	0.0166
T cell regulatory (Tregs)_CIBERSORT-ABS	CIBERSORT-ABS	0.1038	0.0474
Macrophage_EPIC	EPIC	-0.2435	0.0000
uncharacterized cell_EPIC	EPIC	0.2512	0.0000
Macrophage_Monocyte_MCPCOUNTER	MCPCOUNTER	0.2504	0.0000
Monocyte_MCPCOUNTER	MCPCOUNTER	0.2504	0.0000
Myeloid dendritic cell_MCPCOUNTER	MCPCOUNTER	0.1739	0.0009
T cell CD8+_MCPCOUNTER	MCPCOUNTER	0.1242	0.0176
Macrophage M1_QUANTISEQ	QUANTISEQ	0.2009	0.0001
Macrophage M2_QUANTISEQ	QUANTISEQ	0.1177	0.0245
NK cell_QUANTISEQ	QUANTISEQ	-0.1073	0.0404
T cell CD8+_QUANTISEQ	QUANTISEQ	0.1132	0.0307
uncharacterized cell_QUANTISEQ	QUANTISEQ	-0.1116	0.0331
Macrophage_TIMER	TIMER	0.2517	0.0000
Myeloid dendritic cell_TIMER	TIMER	0.1790	0.0006
Neutrophil_TIMER	TIMER	0.1954	0.0002
Cancer associated fibroblast_XCELL	XCELL	-0.1034	0.0484
Common lymphoid progenitor_XCELL	XCELL	0.1763	0.0007
Endothelial cell_XCELL	XCELL	-0.1553	0.0029
Granulocyte-monocyte progenitor_XCELL	XCELL	-0.1139	0.0295
Hematopoietic stem cell_XCELL	XCELL	-0.2158	0.0000
Macrophage M1_XCELL	XCELL	0.1413	0.0069
Macrophage_XCELL	XCELL	0.1130	0.0309
Monocyte_XCELL	XCELL	0.1296	0.0132
stroma score_XCELL	XCELL	-0.1921	0.0002
T cell CD4+ Th2_XCELL	XCELL	0.3608	0.0000

Table S5. The p value of comparing chemotherapeutics sensitivity and risk sore.

Drug	p-value
Doxorubicin	0.026
Vinblastine	0.0043
Mitomycin.C	1.40E-09
Cisplatin	3.60E-05
Sorafenib	0.24