

Prognostic signature for hepatocellular carcinoma based on 4 pyroptosis-related genes

Yuanbin Zhong (✉ 71981232@qq.com)

Primary research

Keywords: hepatocellular carcinoma, pyroptosis, risk scores

Posted Date: October 15th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-970746/v1>

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

[Read Full License](#)

Abstract

Background

Hepatocellular carcinoma (HCC) is a cancer with a poor prognosis. Many recent studies have suggested that pyroptosis is important in tumour progression. However, the role of pyroptosis-related genes (PRGs) in HCC remains unclear.

Materials and methods

We identified differentially expressed PRGs in tumours versus normal tissues. Through univariate, LASSO, and multivariate Cox regression analyses, a prognostic PRG signature was established. The signature effectiveness was evaluated by time-dependent receiver operating characteristic (ROC) curve and Kaplan–Meier (KM) survival analysis. The signature was validated in the ICGC (LIRI-JP) cohort. In addition, single-sample gene enrichment analysis (ssGSEA) showed the infiltration of major immune cell types and the activity of common immune pathways in different subgroups.

Results

Twenty-nine pyroptosis-related DEGs from The Cancer Genome Atlas-Liver Hepatocellular Carcinoma (TCGA-LIHC) dataset were detected, and four genes (*CTSV*, *CXCL8*, *MKI67* and *PRF1*) among them were selected to construct a prognostic signature. Then, the patients were divided into high- and low-risk groups. The pyroptosis-related signature was significantly associated with overall survival (OS). In addition, the patients in the high-risk group had lower levels of immune infiltration.

Conclusion

The prognostic signature for HCC based on 4 pyroptosis-related genes has reliable prognostic and predictive value for HCC patients.

Background

Hepatocellular carcinoma (HCC) is one of the most common cancers, and its outlook is very poor. Worldwide, it is a leading cancer type in 11 countries and regions and is the third leading cause of cancer-related death in 23 countries after lung cancer and colorectal cancer. According to GLOBCAN 2020, 830,180 new deaths from HCC occur every year[1].In most patients, surgery is only an option for those in the early stages of liver cancer (less than stage IIIB and Child-Pugh stage A)[2]. Patients with advanced liver cancer often take conservative treatments and have a poor prognosis. However, the early clinical manifestations of HCC are often atypical, and it is difficult to diagnose HCC only from the clinical manifestations. AFP is a common serum molecular marker used to diagnose liver cancer. However,

normal AFP levels may be present in approximately 30% of patients with HCC[3]. Accordingly, the establishment of a multigene classifier to predict HCC outcomes is important. Such a classifier could have important reference value for predicting the survival and treatment outcomes of patients with liver cancer.

Pyroptosis is a form of regulated cell death (RCD), and the Nomenclature Committee on Cell Death recommended that it be recognized as a type of RCD that critically depends on the formation of plasma membrane pores by members of the gasdermin (GSDM) protein family, often (but not always) as a consequence of inflammatory caspase activation[4]. GSDM pores formed in the plasma membrane eventually lead to cell lysis, and a large number of damage-related molecular patterns (DAMPs), such as ATP, IL-1 β , and HMGB1, are released at different times and stages[5], resulting in inflammation. Studies have shown that certain inflammatory factors, such as IL-1 β , play a role in promoting tumour invasion. Solid tumours that have been shown to upregulate IL-1 β include breast, colon, lung, and head and neck cancers and melanoma, and patients with these IL-1 β -upregulated tumours generally show a poor prognosis[6]. In many tumours, IL-1 β may increase tumour cell metastasis by promoting angiogenesis[7, 8]. However, IL-1 β has also been shown to inhibit cancer. Haabeth *et al.* demonstrated that IL-1 β can drive tumour-specific Th1 response activation to inhibit the progression of B-cell myeloma and lymphoma[9, 10]. However, there are few studies on pyroptosis in HCC.

Here, we conducted a systematic study of pyroptosis-related genes (PRGs) in HCC to determine differences in the expression of these genes between tumour patients and nontumour patients, to assess the correlation between their expression and patient prognosis and to study the effects of these PRGs on the tumour-immune microenvironment.

Results

CC analysis to explore patient differences

To effectively distinguish liver cancer patients by PRGs, we conducted CC analysis on all HCC patients in the TCGA cohort. The cluster variable K was increased from 2 to 10; when K =2, the correlation within subgroups was the highest, while the correlation between groups was the lowest. However, the samples could not be well distinguished (Supplementary Fig. 1a). In addition, when KM survival curves were generated for the two subgroups, we found no significant difference between the two groups (Supplementary Fig. 1b).

Identification of differentially expressed PRGs between tumour and normal tissue

Overall, 184 genes related to pyroptosis were retrieved from the GeneCards database, and 29 differentially expressed genes with $|Log_2 FC| > 1$ were screened by the "limma" package (Fig. 1a,b). $P < 0.05$ was considered to indicate statistically significant DEGs. Of the DEGs, 8 PRGs were upregulated in liver cancer tissues, and 21 PRGs were downregulated in tumour tissues.

To further explore the association between these PRGs, protein-protein interaction (PPI) analysis was performed, setting the minimum required interaction score at 0.4 (medium confidence). Cytoscape software (Version 3.7.2) was employed for visualization (Fig. 1c), and we found that IL1B interacts with many genes at the protein level. In addition, we mapped the correlation network containing these 29 genes and explored their relationships at the gene level (Fig. 1d).

Construction and development of a prognostic signature for HCC

Through univariate Cox regression analysis, we found 10 DEGs related to the prognosis of HCC among 29 PRGs (Fig. 2a). High expression of *ANXA2*, *CTSV*, *CXCL8*, *IL13RA2*, *MKI67*, *TLR2* and *TREM2* may promote survival, while high expression of *CD14*, *MST1* and *NLRP6* may adversely affect survival. LASSO regression analysis further narrowed ten genes down to seven (*ANXA2*, *CTSV*, *CXCL8*, *MKI67*, *NLRP6*, *NLRP6*, and *PRF1*) (Fig. 2b, c). Finally, we obtained 4 key prognosis-related PRGs—*CTSV*, *CXCL8*, *MKI67* and *PRF1*—by stepwise multivariate regression analysis (Fig. 3d, e). The prognostic PRG signature was constructed as follows: risk score = (0.249* *CTSV* expression) +(0.137* *CXCL8* expression) +(0.264* *MKI67* expression) +(-0.224* *PRF1* expression). *CTSV*, *CXCL8* and *MKI67* were high-risk factors (HR>1), and *PRF1* was a protective factor (HR<1). After calculating the risk scores of all samples, the median was used to divide the samples into two subgroups (high and low risk), with a cut-off value of 0.384. Patients in the high- and low-risk groups were well distinguished by PCA (Fig. 3a), and the high-risk score group had a higher mortality rate and shorter survival time (Fig. 3b, c). By plotting ROC curves, we found that risk scores also had certain significance in differentiating patients' half-year, 1-year, 2-year, and 3-year survival (Fig. 3d). The area under the ROC curve (AUC) was 0.691 at 6 months, 0.725 at 1 year, 0.694 at 2 years and 0.681 at 3 years. In addition, multivariate Cox regression proved that our prognostic signature was an independent prognostic factor for HCC (Fig. 3e).

External database validation of the PRG-related prognostic signature in HCC

TIMER database analysis confirmed that *CTSV* and *MKI67* were highly expressed in tumour tissues, while *CXCL8* and *PRF1* were expressed at low levels in tumour tissues (Fig. 4a), which was consistent with our findings. We investigated the genetic changes in the *CTSV*, *CXCL8*, *MKI67* and *PRF1* genes. The mutations of these four genes in TCGA HCC samples were assessed via the cBioPortal database. *MKI67* had the highest frequency of genetic alterations (3%), and missense mutation was the most common alteration (Fig. 4b).

Construction of a nomogram

Univariate and multivariate Cox regression analyses showed that the expression of *CTSV*, *CXCL8*, *MKI67* and *PRF1* was an independent prognostic factor in liver cancer patients. On this basis, we constructed a nomogram to predict the patient's probability of survival at one, two, and three years. The expression

levels of four genes (*CTSV*, *CXCL8*, *MKI67* and *PRF1*) were included as indicators to establish the risk score (Fig. 5a). In addition, a calibration plot was drawn and indicated that our nomogram had good predictive performance (Fig. 5b, c, d).

Validation of the PRG-related prognostic signature for HCC

To verify the availability of our PRG-related model, we downloaded liver cancer datasets from the ICGC database to verify the risk score equation. We used a cut-off value of 0.384 for the training set to divide the ICGC database samples into high- and low-risk subgroups. Consistent with the results from the TCGA database, patients in the high- and low-risk groups could still be distinguished by PCA (Fig. 6a). Patients with high risk scores had worse prognoses (Fig. 6b), and the OS of high-risk patients was significantly reduced compared with that of low-risk patients ($P = 6.501e-03$) (Fig. 6c). The AUCs of the prognostic signature at six months, one year, two years and three years were 0.673, 0.731, 0.681 and 0.705, respectively (Fig. 6d). Multivariate Cox regression proved that the prognostic signature was an independent prognostic factor for HCC (Fig. 6e). All of the above evidence indicates that PRGs are of certain clinical value in our prognostic signature of HCC.

Gene enrichment analysis

To further explore the differences in gene function and pathways between the high- and low-risk subgroups, the "limma" package was used to identify DEGs between the two subgroups of TCGA patients. The differential criteria were an absolute value of $\log FC > 1$ and $P < 0.05$. In total, 45 genes related to the risk score were found. Of these, 43 genes were upregulated in the high-risk group, and 2 were downregulated. Then, KEGG and GO enrichment analysis was conducted on these genes (Fig. 7a, b), and the results showed that most of the high risk-related pathways were cell cycle, cell metabolism and immune-related pathways.

Analysis of the immune microenvironment in patient subgroups

To explore the differences in immune infiltration and the activity of immune-related pathways in high- to low-risk subgroups, single-sample gene enrichment analysis (ssGSEA) was used to calculate the immune score of 16 kinds of immune cells in each sample. The results showed that the high-risk group usually had a lower level of immune infiltration. In particular, B cells, CD8+ T cells, macrophages, mast cells, neutrophils, NK cells, pDCs (plasmacytoid dendritic cells), T helper cells, and TILs (tumour-infiltrating lymphocytes) (Fig. 8a) were found at lower levels in this group. The activity of immune-related pathways was also generally downregulated in the high-risk group (Fig. 8b).

Discussion

HCC is a serious health hazard worldwide. Some studies have shown that pyroptosis plays an important role in the occurrence and development of tumours[11, 12], but the role of PRGs in liver cancer remains unclear. In this study, we first assessed 29 PRGs that were differentially expressed in tumour and normal samples and conducted PPI network and correlation analysis to explore the associations of these genes in HCC. To further explore the prognostic value of these PRGs, we established a risk profile of four genes by univariate Cox regression, LASSO-Cox regression, and multivariate Cox regression analyses, and the risk score equation was verified to be clinically significant in the ICGC database. A nomogram was constructed to estimate the patient's one-, two-, and three-year survival probability. Enrichment analysis showed that the DEGs between the high- and low-risk groups were involved in immune-related pathways. We conducted ssGSEA and found that the level of immune infiltration and the activity of related pathways were decreased significantly in the high-risk group.

Among the four key PRGs obtained, cathepsin V (*CTSV*) promotes the proliferation and invasion of tumour cells in breast cancer[13] and colorectal cancer[14], but its role in liver cancer is unclear. *CXCL8/IL-8* is a well-known chemokine that mediates cancer cell motility, invasion and metastasis by promoting epithelial-mesenchymal transition (*EMT*)[15], and IL-8 has been found to increase cancer proliferation in vitro[16]. In Zhao *et al.*'s study[17], *MKI67* was confirmed to be related to the prognosis of liver cancer. Perforin plays a dominant role in the CD8+ T cell-mediated lysis of HCV-replicating human hepatoma cells. Our results indicate that these four key genes are independent factors affecting the prognosis of HCC, and the signature including all these genes has good predictive performance in terms of prognosis and recurrence.

In conclusion, our study suggests that the prognosis of HCC is closely related to pyroptosis. We constructed a promising PRG prognostic model to predict the prognosis of HCC patients. It will be a helpful reference for clinical and treatment decision-making. At present, there are few studies on pyroptosis, especially its mechanism in HCC. The four genes in the signature may play an important role in pyroptosis in HCC. Our study preliminarily confirmed the prognostic value of these PRGs in HCC; however, how they play a role in HCC needs further exploration.

Conclusion

In summary, we identified pyroptosis-related DEGs between HCC and normal tissues and provided a signature to evaluate the prognosis of HCC based on pyroptosis-related genes.

Materials And Methods

Data collection

The mRNA expression profiles (mRNA-seq) and corresponding clinical data of patients were obtained from The Cancer Genome Atlas (TCGA) database. We downloaded TCGA-liver hepatocellular cancer (LIHC) sample information after log2 (FPKM +1) processing from the Xena platform

(<https://xena.ucsc.edu/>)[18]. In total, 273 primary tumour solid tissue samples and 197 normal samples were included as the external validation dataset from the International Cancer Genome Consortium (ICGC), and 184 PRGs were retrieved from the GeneCards database (25th August 2021. <https://www.genecards.org/>) with the keyword “pyroptosis”. The above three databases are publicly available and accessible.

Consensus clustering (CC) analysis

Unsupervised class discovery is a highly useful technique in cancer research, where intrinsic groups sharing biological characteristics may exist but are unknown. The CC method provides quantitative and visual stability evidence for estimating the number of unsupervised classes in a dataset[19]. We used the “ConsensusClusterPlus” R package to duplicate tumour samples in TCGA 1000 times to find the optimal classification cluster of tumour samples, and approximately 80% of the samples were selected in each iteration.

Search for differentially expressed genes (DEGs) in the TCGA-LIHC dataset

We used the “limma” R package to obtain DEGs[20]. DEGs with $|log_2\text{-fold change (FC)}| \geq 1$ and adj P-value (FDR) < 0.05 were used for subsequent analysis.

Construction of the prognostic PRG signature

Univariate Cox regression, least absolute shrinkage and selection operator (LASSO)-Cox regression and multivariate Cox regression analyses were used to identify the genes associated with pyroptosis and establish a prognostic PRG signature for HCC. In the univariate Cox regression analysis, $P < 0.05$ was considered to indicate statistical significance, indicating that the gene is associated with overall survival (OS). The LASSO Cox regression model was then utilized to narrow down the candidate genes by using 10-fold cross validation via the “glmnet” R package[21]. Then, we carried out multivariate Cox regression analysis and established the risk scoring equation according to the standardized regression coefficients of the factors.

$$\text{Risk Score} = \sum_i (\text{Coef}_i \cdot \text{Expi})$$

To evaluate the predictive ability of the prognostic PRG signature, we used the median risk score to divide the sample into high- and low-risk groups. Kaplan–Meier (KM) survival curves were drawn with the “survival” R package, and the P value was calculated. In addition, time-dependent receiver operating characteristic (ROC) curve analyses were performed to determine how accurately the signature predicted patient survival at six months, one year, two years, and three years. The risk curve was also used to verify the relationship between the risk score and patient survival status. Principal component analysis (PCA) integrated the risk score and patient survival information through dimensionality reduction and presented relevant information in a low-dimensional manner, which could verify the signature's ability to distinguish

between high- and low-risk groups[22]. In addition, we performed a similar process with data from the ICGC database, and the prognostic PRG signature was further validated in this independent cohort.

Validation of prognostic genes in an external database

The mRNA expression levels of four marker genes were verified with the Tumour Immune Estimation Resource (TIMER) database (<https://cistrome.shinyapps.io/timer/>). In addition, we used the cBioPortal database (<https://www.cbioportal.org/>) to investigate the genetic alterations of the related genes. The above two databases are publicly available and accessible.

Construction of the nomogram and evaluation of its predictive value

The expression levels of the four marker genes in the risk scoring equation were included in the nomogram to assess the 1-year, 2-year and 3-year survival rates of the patients, and calibration curves were drawn to assess the prognostic performance of the nomogram.

Validation of the independence of the risk scores from other clinical variables

The clinical information, including sex, age, tumour grade, tumour stage and risk scores, of patients providing the samples was included in the univariate Cox regression analysis, and bilateral P values < 0.05 were considered to be statistically significant. The clinical features were also incorporated into the multivariate Cox regression analysis to determine whether the prognostic model prediction ability was independent of the conventional clinical features. Hazard ratios (HRs) and 95% confidence intervals (CIs) were calculated.

Gene enrichment analysis

We next investigated the biological processes involving the DEGs between the high- and low-risk patient groups. DEGs with P < 0.05 and absolute value of log2 FC > 1 were selected for enrichment analysis. For the Kyoto Encyclopedia of Genes and Genomes (KEGG) and gene ontology (GO) analyses, P < 0.05 and false discovery rate (q) < 1 were considered indicators of statistical significance.

Immune infiltration analysis

The "GSVA" package was used to calculate the immune score of the samples, explore the infiltration of 16 major immune cell types and investigate the activity of 13 common immune pathways in different subgroups in the high- and low-risk groups.

Abbreviations

HCC: Hepatocellular carcinoma

RCD: Pyroptosis is a form of regulated cell death

PRG: Pyroptosis-related genes

mRNA-seq: mRNA expression profiles

TCGA: The Cancer Genome Atlas

ICGC: International Cancer Genome Consortium

CC: Consensus clustering

DEG: Differentially expressed gene

LASSO: Least absolute shrinkage and selection operator

OS: Overall survival

KM: Kaplan–Meier

ROC: Receiver operating characteristic

PCA: Principal component analysis

TIMER: Tumour Immune Estimation Resource

HR: Hazard ratio

AUC: Area under the ROC curve

CI: Confidence interval

KEGG: Kyoto Encyclopedia of Genes and Genomes

GO: Gene ontology

PPI: Protein-protein interaction

ssGSEA: Single-sample gene enrichment analysis

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and materials

The datasets analysed in this research were downloaded from The Cancer Genome Atlas (TCGA, <https://portal.gdc.cancer.gov/repository>), the UCSC Xena platform (<https://xenabrowser.net/>) and the International Cancer Genome Consortium (ICGC, <https://dcc.icgc.org/releases/current/Projects/LIRI-JP>).

Competing interests

The authors declare that they have no competing interests.

Funding

This study was supported by the Natural Science Foundation of Jiangxi Province (20192BAB205059) and the National Natural Science Foundation of China (82060112).

Authors' contributions

YZ and LD designed the study. SD and JG carried out the data extraction and statistical analysis. WL, YZ, YD, CW and HH drafted the manuscript. HX, SG, SL, ZL, FX modified the manuscript. All authors contributed to the article and approved the submitted version.

Acknowledgements

We would like to acknowledge TCGA, the Xena platform and the ICGC network for providing data.

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F: **Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries.** *CA Cancer J Clin* 2021, **71**(3):209-249.
2. Recio-Boiles A, Babiker HM: **Liver Cancer.** In: *StatPearls*. Treasure Island (FL): StatPearls Publishing
3. Marengo A, Rosso C, Bugianesi E: **Liver Cancer: Connections with Obesity, Fatty Liver, and Cirrhosis.** *Annu Rev Med* 2016, **67**:103-117.
4. Galluzzi L, Vitale I, Aaronson SA, Abrams JM, Adam D, Agostinis P, Alnemri ES, Altucci L, Amelio I, Andrews DW *et al.* **Molecular mechanisms of cell death: recommendations of the Nomenclature Committee on Cell Death 2018.** *Cell Death Differ* 2018, **25**(3):486-541.

5. Tsuchiya K: **Switching from Apoptosis to Pyroptosis: Gasdermin-Elicited Inflammation and Antitumor Immunity.** *Int J Mol Sci* 2021, **22**(1).
6. Apte RN, Krelin Y, Song X, Dotan S, Recih E, Elkabets M, Carmi Y, Dvorkin T, White RM, Gayvoronsky L et al: **Effects of micro-environment- and malignant cell-derived interleukin-1 in carcinogenesis, tumour invasiveness and tumour-host interactions.** *Eur J Cancer* 2006, **42**(6):751-759.
7. Nakao S, Kuwano T, Tsutsumi-Miyahara C, Ueda S-i, Kimura YN, Hamano S, Sonoda K-h, Saijo Y, Nukiwa T, Strieter RM et al: **Infiltration of COX-2-expressing macrophages is a prerequisite for IL-1 beta-induced neovascularization and tumor growth.** *J Clin Invest* 2005, **115**(11):2979-2991.
8. Saijo Y, Tanaka M, Miki M, Usui K, Suzuki T, Maemondo M, Hong X, Tazawa R, Kikuchi T, Matsushima K et al: **Proinflammatory cytokine IL-1 beta promotes tumor growth of Lewis lung carcinoma by induction of angiogenic factors: in vivo analysis of tumor-stromal interaction.** *J Immunol* 2002, **169**(1):469-475.
9. Haabeth OAW, Lorvik KB, Hammarström C, Donaldson IM, Haraldsen G, Bogen B, Corthay A: **Inflammation driven by tumour-specific Th1 cells protects against B-cell cancer.** *Nat Commun* 2011, **2**:240.
10. Haabeth OAW, Lorvik KB, Yagita H, Bogen B, Corthay A: **Interleukin-1 is required for cancer eradication mediated by tumor-specific Th1 cells.** *Oncoimmunology* 2016, **5**(1):e1039763.
11. Xia X, Wang X, Cheng Z, Qin W, Lei L, Jiang J, Hu J: **The role of pyroptosis in cancer: pro-cancer or pro-"host"?** *Cell Death Dis* 2019, **10**(9):650.
12. Zhang Z, Zhang Y, Xia S, Kong Q, Li S, Liu X, Junqueira C, Meza-Sosa KF, Mok TMY, Ansara J et al: **Gasdermin E suppresses tumour growth by activating anti-tumour immunity.** *Nature* 2020, **579**(7799):415-420.
13. Sereesongsaeng N, McDowell SH, Burrows JF, Scott CJ, Burden RE: **Cathepsin V suppresses GATA3 protein expression in luminal A breast cancer.** *Breast Cancer Res* 2020, **22**(1):139.
14. Wang C-H, Wang L-K, Wu C-C, Chen M-L, Kuo C-Y, Shyu R-Y, Tsai F-M: **Cathepsin V Mediates the Tazarotene-induced Gene 1-induced Reduction in Invasion in Colorectal Cancer Cells.** *Cell Biochem Biophys* 2020, **78**(4):483-494.
15. Fernando RI, Castillo MD, Litzinger M, Hamilton DH, Palena C: **IL-8 signaling plays a critical role in the epithelial-mesenchymal transition of human carcinoma cells.** *Cancer Res* 2011, **71**(15):5296-5306.
16. Khazali AS, Clark AM, Wells A: **Inflammatory cytokine IL-8/CXCL8 promotes tumour escape from hepatocyte-induced dormancy.** *Br J Cancer* 2018, **118**(4):566-576.

17. Yao Z, Dong Y, Wu G, Zhang Q, Yang D, Yu J-H, Wang W-P: **Preoperative diagnosis and prediction of hepatocellular carcinoma: Radiomics analysis based on multi-modal ultrasound images.** *BMC Cancer* 2018, **18**(1):1089.
18. Goldman MJ, Craft B, Hastie M, Repečka K, McDade F, Kamath A, Banerjee A, Luo Y, Rogers D, Brooks AN *et al*: **Visualizing and interpreting cancer genomics data via the Xena platform.** *Nat Biotechnol* 2020, **38**(6):675-678.
19. Wilkerson MD, Hayes DN: **ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking.** *Bioinformatics* 2010, **26**(12):1572-1573.
20. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK: **limma powers differential expression analyses for RNA-sequencing and microarray studies.** *Nucleic Acids Res* 2015, **43**(7):e47.
21. Tibshirani R: **The lasso method for variable selection in the Cox model.** *Stat Med* 1997, **16**(4):385-395.
22. Jolliffe IT, Cadima J: **Principal component analysis: a review and recent developments.** *Philos Trans A Math Phys Eng Sci* 2016, **374**(2065):20150202.

Figures

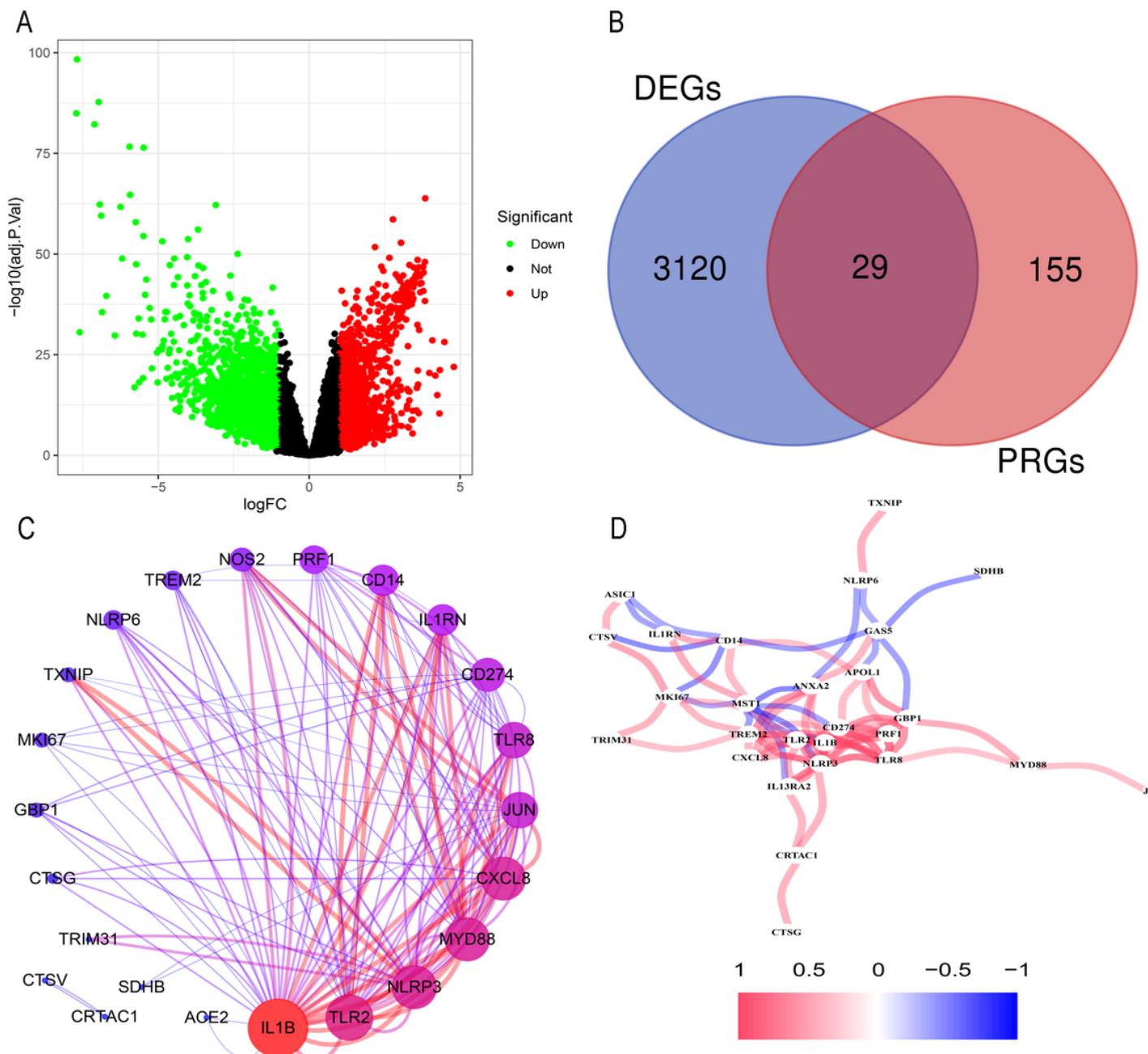
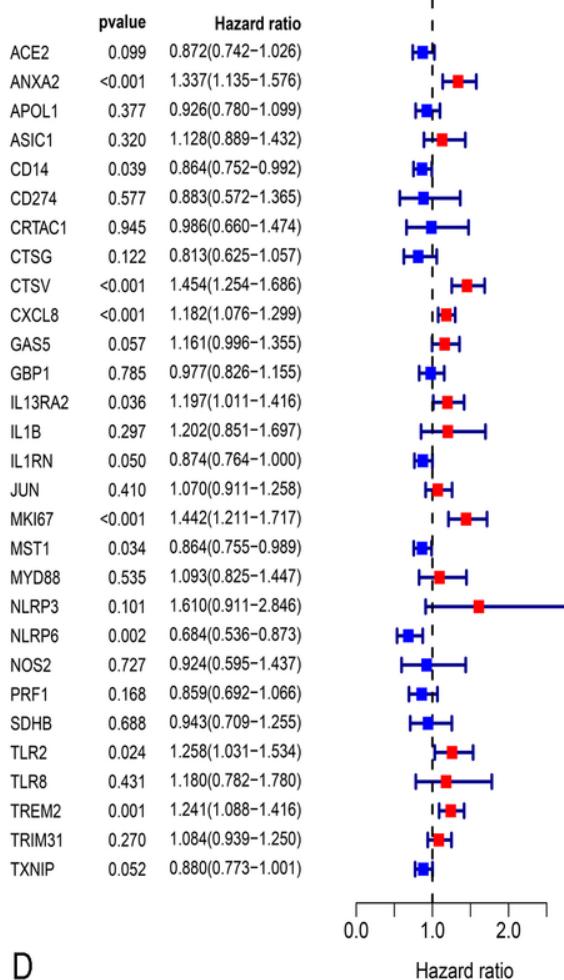


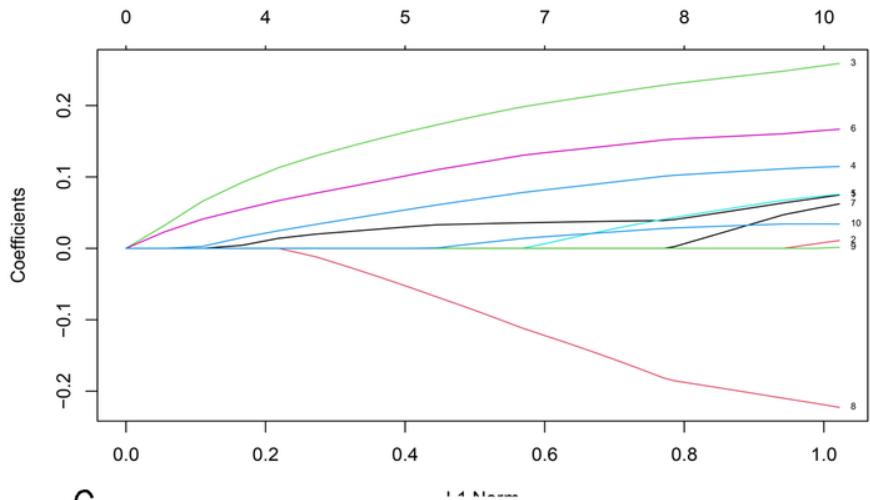
Figure 1

Searching for PRGs and links between them. Volcano plot of all DEGs between HCC and normal samples. Red: upregulated DEGs; Black: nonsignificant genes; Green: downregulated DEGs (a). Overall, 29 PRGs were found in DEGs (b). PPI analysis and the correlation network were drawn to explore the relationship between 29 genes (c, d).

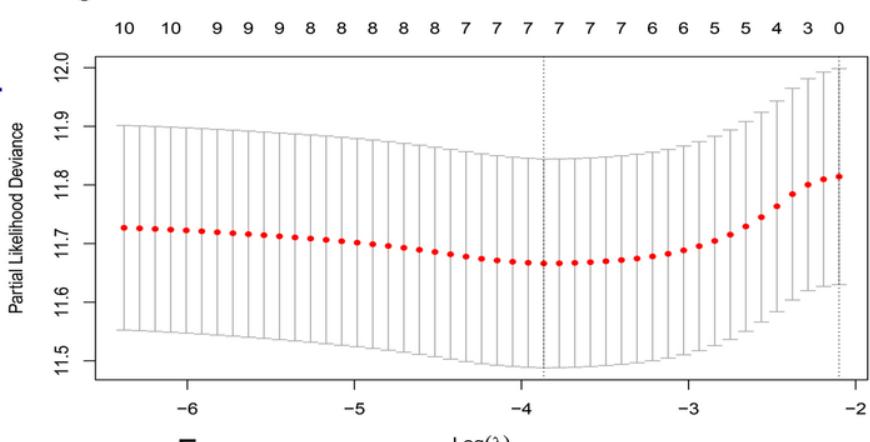
A



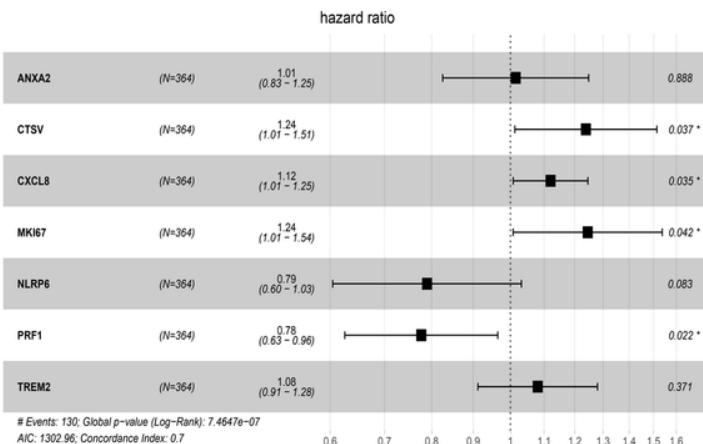
B



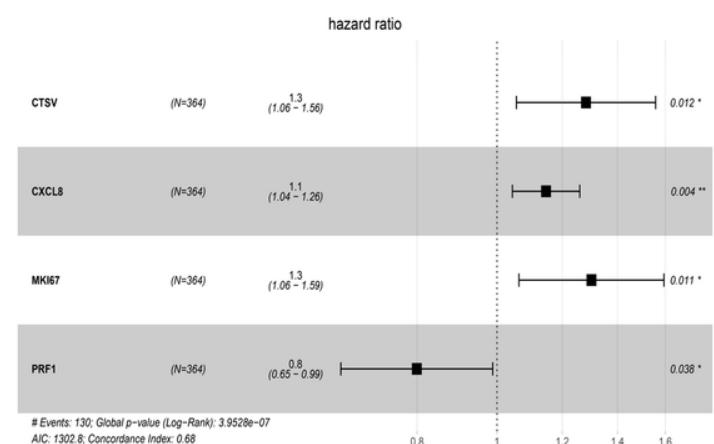
C



D



E

**Figure 2**

Construction and development of a prognostic signature for HCC. Univariate Cox regression was used to identify genes associated with OS (a). LASSO regression of the 10 OS-related genes (b,c). The prognostic signature was established based on multivariate Cox regression (d, e).

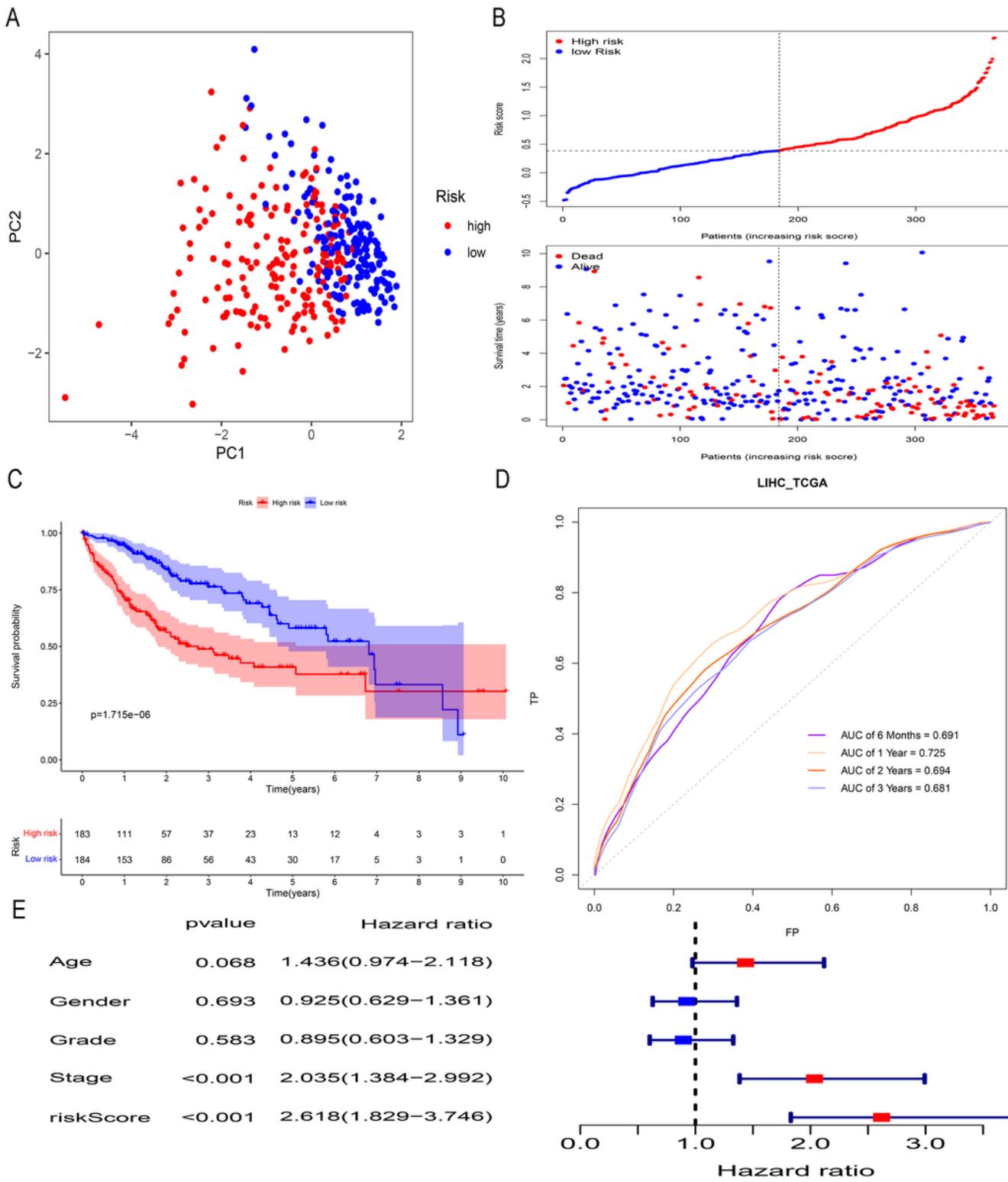


Figure 3

Internal validation of the TCGA cohort. PCA plot for samples based on the risk scores (a). Risk score and survival status distribution of HCC samples from the TCGA database (b). KM curves of samples in the high- and low-risk groups (c). ROC curves demonstrated the predictive efficiency of the risk score (d). Multivariate Cox regression showed that the risk score was an independent prognostic factor in the TCGA database (e).

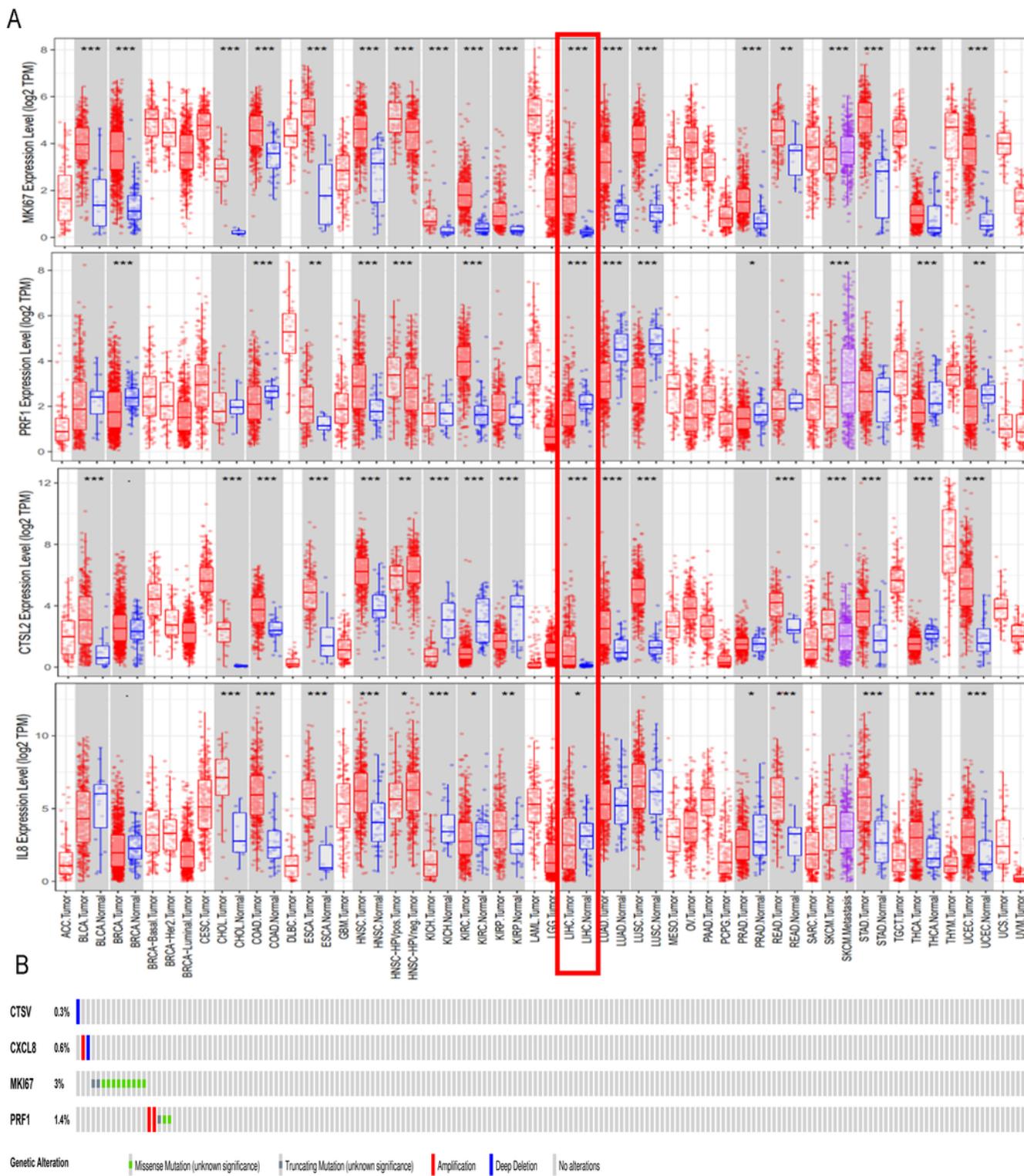


Figure 4

External database validation of the PRG-related prognostic signature in HCC. Human CTSV, CXCL8, MK67 and PRF1 levels in different tumour types from TCGA were determined using TIMER (a). The mutations of these four genes in TCGA HCC samples were assessed via the cBioPortal database (b).

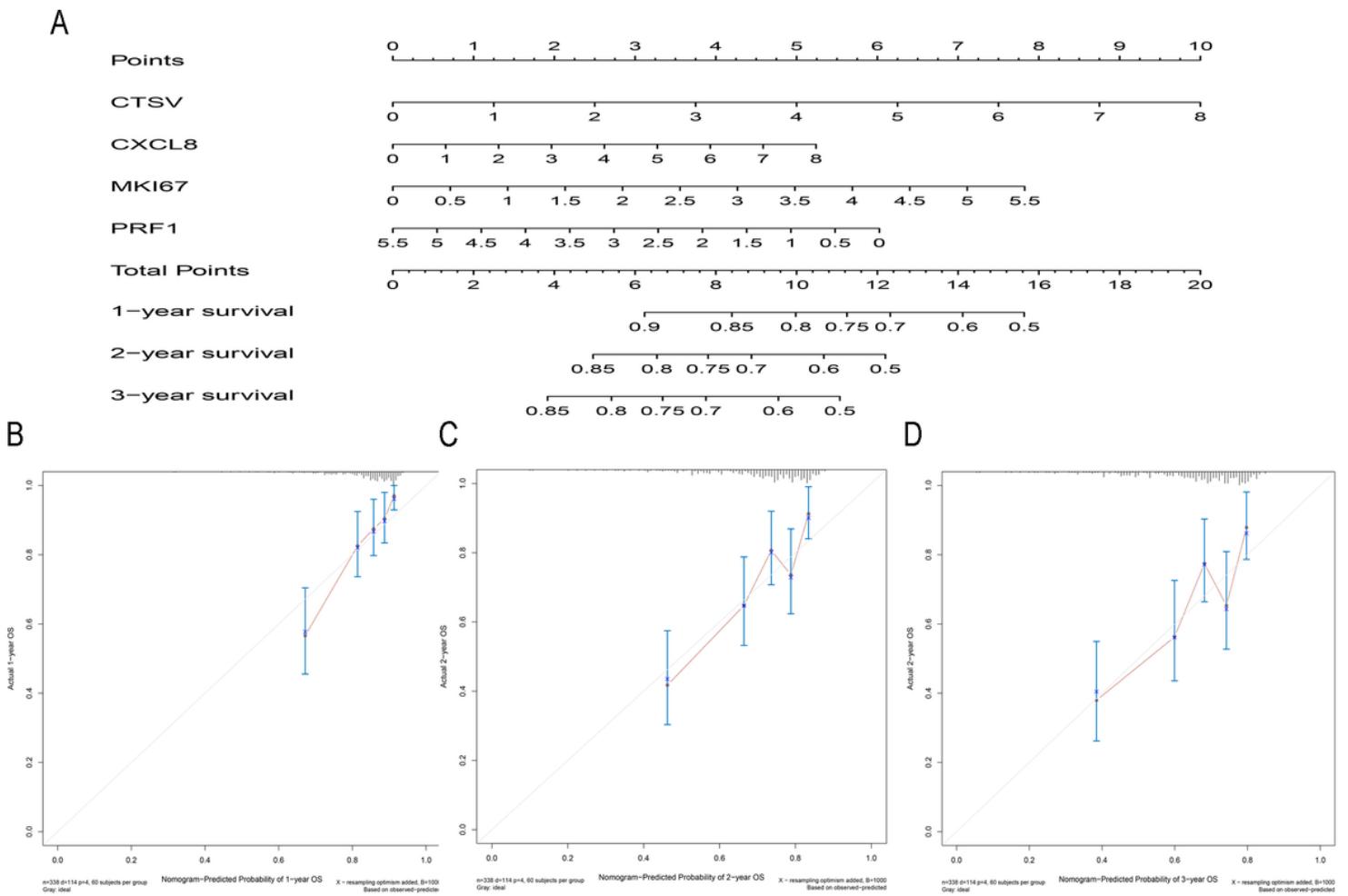


Figure 5

Construction and verification of a nomogram. The nomogram was built based on the expression of 4 genes in the training cohort (a). The calibration plots showed good predictive performance for OS at 1, 2, and 3 years (b, c, d).

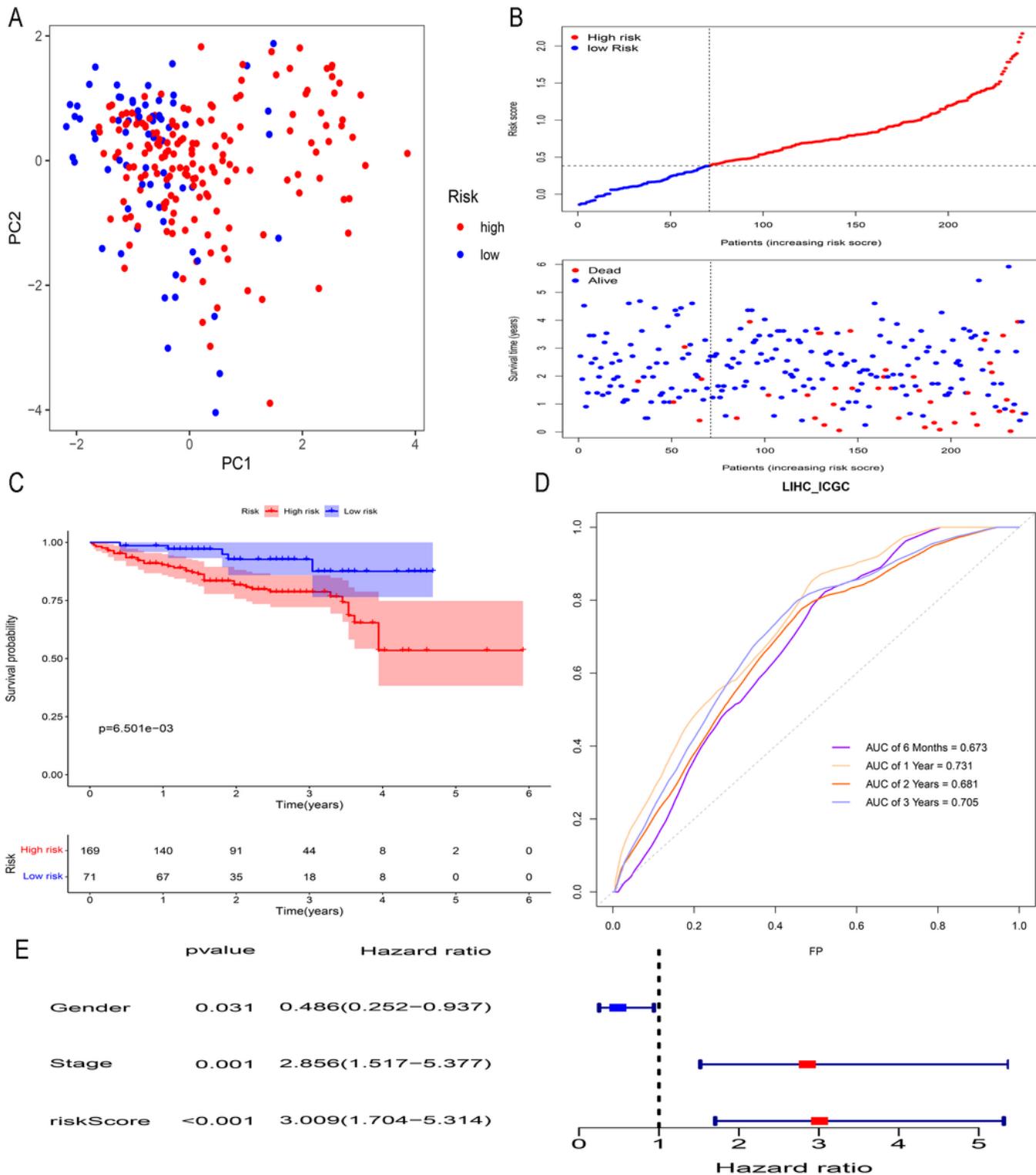


Figure 6

External validation of the ICGC cohort. PCA was performed by dividing HCC samples in the ICGC database into high- and low-risk groups using the cut-off values in TCGA (a). Samples with a high risk score in ICGC HCC samples had a worse prognosis (b,c). The ROC curve verifies the accuracy of the prognostic signature (d). Multivariate Cox regression showed that the risk score was an independent prognostic factor in the ICGC database (e).

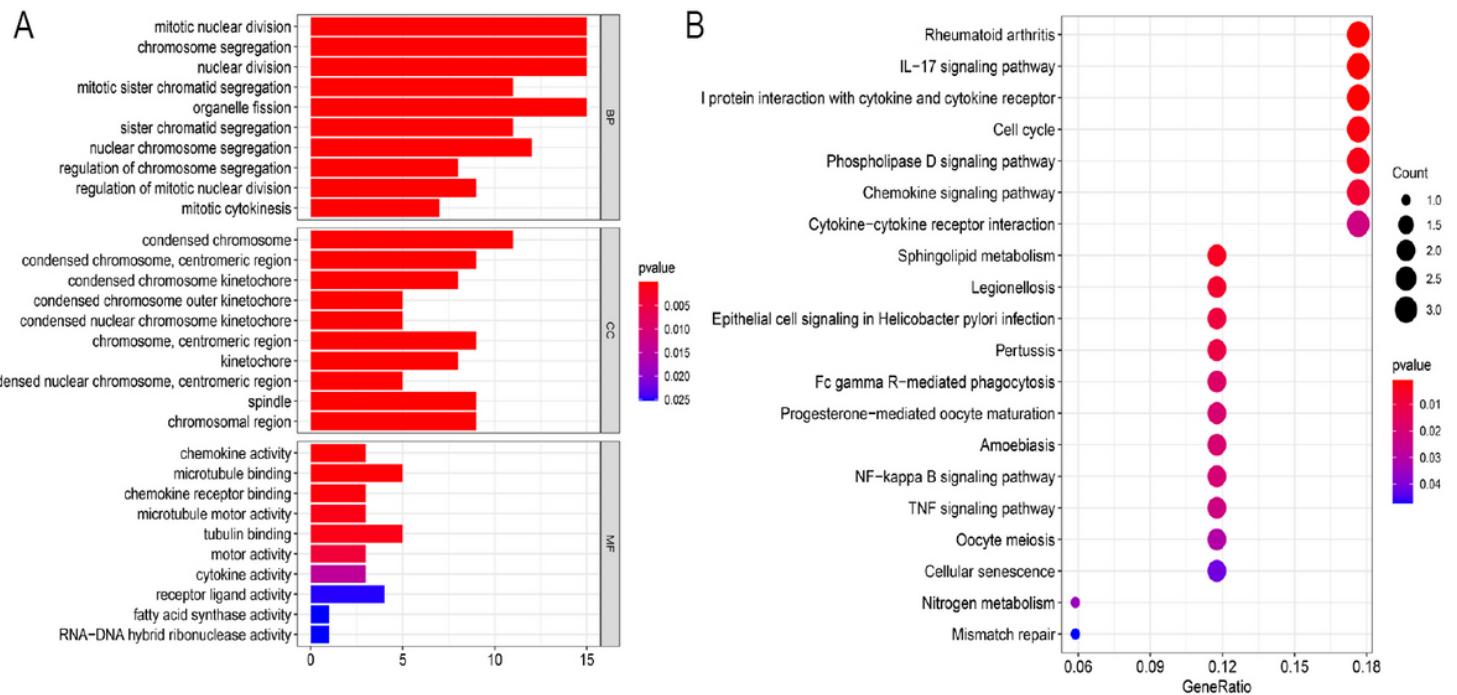


Figure 7

Functional analysis based on the DEGs between the two risk groups in the TCGA cohort. GO enrichment analysis, including biological process analysis, cellular component analysis and molecular function (a). KEGG pathway analysis (b).

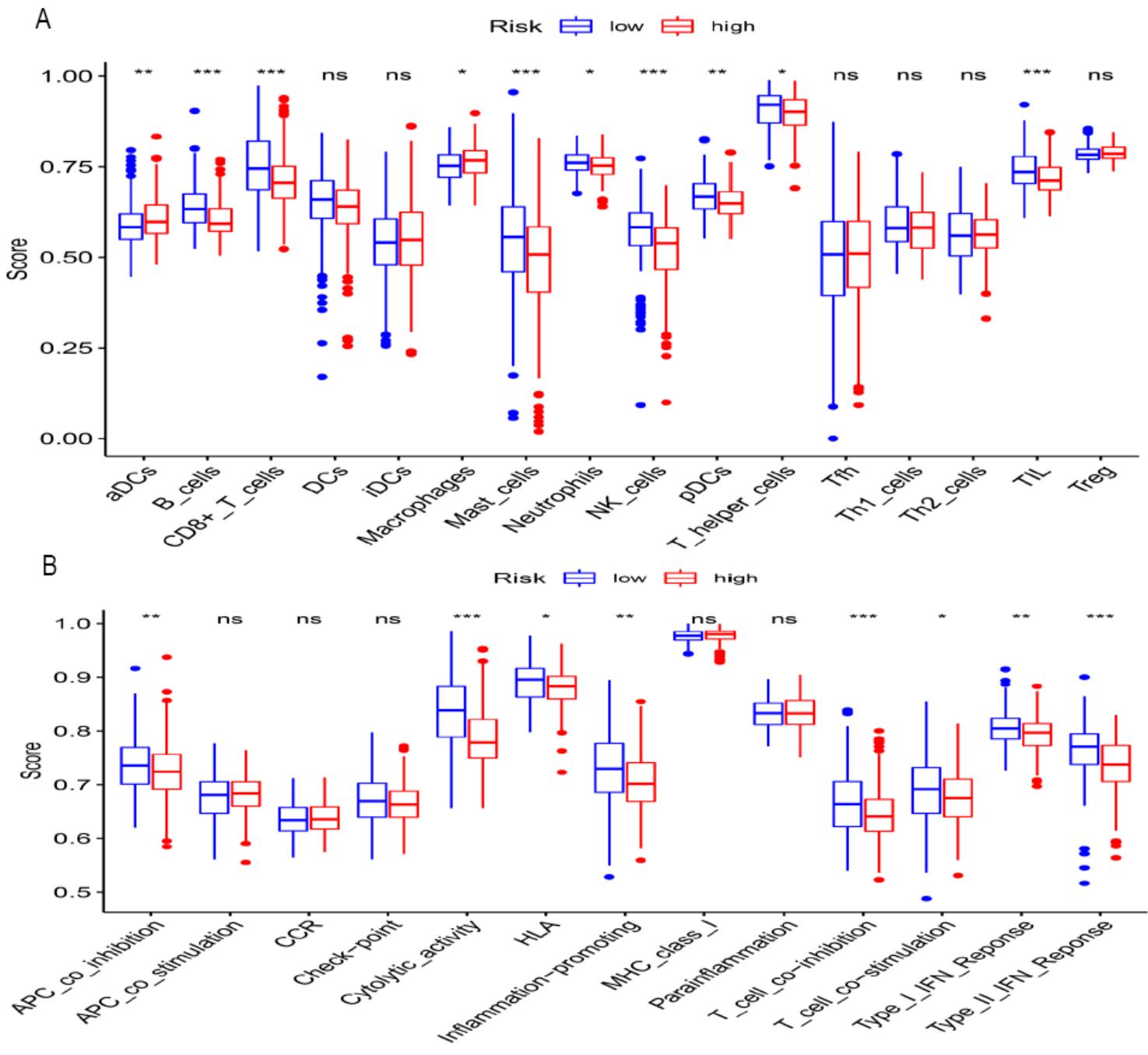


Figure 8

Analysis of the immune microenvironment in patient subgroups. Comparison of the enrichment scores of 16 types of immune cells and 13 immune-related pathways between the low- and high-risk groups in the TCGA cohort.

Supplementary Files

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

- [Supplementarymaterial.docx](#)