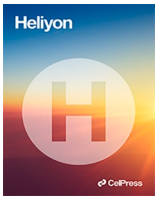
Heliyon 10 (2024) e38438

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Research article

Pyroptosis-related genes features on prediction of the prognosis in liver cancer: An integrated analysis of bulk and single-cell RNA sequencing

Zhihao Zhang [a], Feng Liu [b,**], Xin Lan [c], Fuhai Wang [b], Jiahao Sun [d], Honglong Wei [b,*]

[a] *Department of General Surgery, Traditional Chinese medical hospital of Huangdao District, Qingdao, 266001, China*
[b] *Department of Hepatobiliary Surgery, The First Affiliated Hospital of Shandong First Medical University, Jinan, 271016, China*
[c] *Department of Nephrology, Traditional Chinese Medical Hospital of Huangdao District, Qingdao, 266001, China*
[d] *Department of General Surgery, Qingdao Hiser Hospital Affiliated of Qingdao University (Qingdao Traditional Chinese Medicine Hospital), Qingdao, 266033, China*

ARTICLE INFO

Keywords:
Liver cancer
Biological markers
Pyroptosis
Prognosis features
Survival analysis

ABSTRACT

Objective: This study explores the impact of pyroptosis-related genes (PRG) on the prognosis of liver cancer (LC).

Methods: 421 samples (371 tumor samples and 50 normal samples) from the Cancer Genome Atlas (TCGA) were included in this study. GSE14520 dataset (data of RNA expression and relevant clinicopathological features), GSE125449 dataset (single-cell data in LC) and HCCDB18 dataset (validation on the reliability of the model) were downloaded as appropriate. Download the PRG and its corresponding pathway information from the gene set enrichment analysis (GSEA) website. The consensus clustering was performed by ConsensusClusterPlus package. Differentially expressed genes (DEGs) were identified using limma package, and prognostic features were constructed using un/multivariate and Lasso Cox regression. Pathway enrichment analysis was conducted by ssGSEA method. Receiver Operating Characteristic and the survival analysis were conducted by timeROC and Survminer packages. The Seurat package was used for single-cell RNA sequencing (scRNA-seq) analysis. For cellular validation, following the quantification on the key genes via reverse-transcription quantitative PCR, the Transwell and scratch assays were applied to evaluate the *in-vitro* invasion and migration of LC cells Huh-7.

Results: 12 prognosis-related genes were identified to be related to the progression of LC. Three subtypes including C1, C2 and C3 were categorized using the 12 prognosis-related genes and PRGs significantly related to the prognosis of LC patients. The worst and best prognosis was seen in C3 subtype and C2 subtype, respectively. Hallmark pathway enrichment analysis has shown the concurrent immunoactivation and immune escape in C3 subtype. A RiskScore model was constructed using 8 key genes (KPNA2, UCK2, FTCD, CBX2, RAB32, HMMR, S100A9 and ANXA10) from the DEGs of three subtypes. The RiskScore system as an independent prognostic factor dividing the patients into high and low risk groups, and patients of the high-risk group had poor prognosis in both test set and validation set. A nomogram model combining the risk score had the extreme higher benefit. Further, 6 subclusters were identified from scRNA-seq analysis,

* Corresponding author.
** Corresponding author.
E-mail addresses: liufsd@163.com (F. Liu), weihl2012@126.com (H. Wei).

https://doi.org/10.1016/j.heliyon.2024.e38438
Received 11 July 2024; Received in revised form 4 September 2024; Accepted 24 September 2024
Available online 25 September 2024

where the highest PYROPTOSIS score was seen in Monocytic-Macrophages. The quantification on the key genes has suggested the high expressions of KPNA2, UCK2, CBX2, RAB32, HMMR and S100A9 and the low expressions of FTCD and ANXA10 in LC cells Huh-7. Particularly, UCK2 knockdown evidently diminished the number of invaded and migrated LC cells *in vitro*.
Conclusion: The risk model associated with pyproptosis is crucial for the tumor immunity of LC and may serve as a prognostic indicator for patients suffering from LC. Our findings will offer new perspectives for immunotherapies targeting LC.

1. Introduction

Worldwide, liver cancer (LC) ranks as the 7th most common type of cancer and serves as the 4th leading cause of cancer-related deaths. The highest death rates associated with this disease have been observed in individuals from the East Asian region. Major risk factors for liver cancer include chronic infections with hepatitis B and C viruses, persistent exposure to aflatoxins, as well as alcohol consumption and smoking [1–3]. Estimates indicate that by the year 2025, over one million people will be affected by liver cancer annually [4]. Despite the continuous rise in the incidence of liver cancer, treatment options remain remarkably limited, with patients facing very low chances of survival (5-year survival rate: 18 % only) [5]. In recent times, the field of LC has seen a growing interest in anti-tumor immunotherapy [6]. Therapies that are guided by biomarkers and immune checkpoint inhibitors are emerging as promising methods for treating LC [7]. Nevertheless, the issue of drug resistance in tumors remains a significant challenge within LC therapy [8]. Therefore, it is critical to explore trustworthy new biomarkers and risk assessment models to forecast the prognosis of patients with TLC and evaluate the efficacy of anti-cancer treatment approaches.

Pyroptosis is a programmed cell death process triggered by some inflammasomes, which can lead to cell swelling and membrane rupture and excite a robust inflammation response associated with innate immunity [9,10]. Recent evidences have stressed the dual role pyroptosis plays in the occurrence and development of tumors [11]. Pyroptosis can serve as a novel strategy to eliminate cancer via triggering pyroptotic cell death and activating intense anti-tumor immunity. In addition, pyroptosis is capable of promoting tumor development by inducing excessive and persistent local inflammation [11]. Emerging discoveries have gradually expanded our understanding on the role of pyroptosis in liver diseases like LC, along with some relevant mechanisms of action [12,13]. Besides, a prior study has screened some pyroptosis-related genes (PRGs) in LC and preliminarily constructed a relevant survival nomogram based on these genes [14]. In light of this, our current study aims to further expand the results on PRGs and their involvement in predicting the prognosis of LC patients.

With the rapid development of high-throughput sequencing, a number of methods have been already applied to define the biomarkers of the disease and some notable achievements have been achieved in predicting the prognosis of cancer [15]. While it comes to liver physiology and disease biology, it should not be neglected that in addition to transcriptome analysis (which enables the study of gene expression dynamics in human liver disease), single-cell RNA sequencing (scRNA-seq) has both enabled the study of transcriptional activity at the single cell level and paved the way for the identification of previously unknown cell types or cell subtypes in normal and diseased liver [16]. Great achievements in LC have been made integrating bulk RNA sequencing and scRNA-seq [17,18]. Of note, while linking RNA sequencing with LC, a novel association of PRG signature has been unveiled to be associated with the prognosis of hepatocellular carcinoma (the most prevalent form of LC) [19]. The aim of this study, accordingly, is to identify predictive PRGs in LC and then implement a validation across multiple datasets to demonstrate the robustness of the model. Hereafter, our research extends to the single-cell level and investigates the occurrence of pyroptosis in different cell types. The relevant results are presented as follows.

2. Methods

2.1. Data collection

RNA sequencing data and the clinicopathological information, including 421 samples (371 tumor samples and 50 normal samples), were downloaded from the Cancer Genome Atlas (TCGA, https://portal.gdc.cancer.gov/). A testing dataset (GSE14520) comprising RNA expression data, paired clinical information, and follow-up details in LC, as well as the single-cell dataset GSE125449, were obtained from the Gene Expression Omnibus (GEO, https://www.ncbi.nlm.nih.gov/geo/). Besides, the HCCDB18 dataset (http://lifeome.net/database/hccdb/home.html) was downloaded to further validate the reliability of the model. Those samples with missing information on the survival time and status were excluded during the integration analysis.

Furthermore, the gene sets for REACTOME_PYROPTOSIS and the complete set of Hallmark pathways were acquired from the Gene Set Enrichment Analysis (GSEA).

2.2. Consensus clustering

The ConsensusClusterPlus package was applied to construct a consensus matrix and classify the samples [20]. The parameters were set as follows: clusterAlg = “km” (indicating k-means clustering), distance = “euclidean” (using Euclidean distance metric), number of resampling iterations = 500, and subsampling rate of each resampling iteration at 80 %.

2.3. Differential expressed genes (DEGs) identification

The limma package was applied to identify the DEGs in LC with the thresholds of adjusted p-value <0.05 and |log2FC| > 1 [21].

2.4. Risk model establishment

Utilizing TCGA samples as the training set, we employed a univariate Cox analysis to select key genes significantly associated with prognosis ($p < 0.05$). The limma package was employed for univariate Cox proportional hazard regression on the DEGs, with a threshold of $p < 0.05$ for filtering. Subsequently, the glmnet package was utilized for Least Absolute Shrinkage and Selection Operator (LASSO) Cox regression to further compress the selected genes [22], incorporating a 10-fold cross-validation to mitigate the collinearity of the risk model. The risk score for each patient was calculated using the formula:

$$\text{Risk Score} = \Sigma \beta i \times \text{Expi}$$

(β represents the Cox regression coefficient for the gene)

Next, we calculated the risk score for each sample based on its expression level and used the "surv_cutpoint" function in the survminer package to determine the optimal threshold for the classification of high and low risk groups. To validate the robustness of the clinically predictive model based on the risk-associated gene signature, the following data cohorts, including HCCDB18 and GSE14520, were applied as needed.

2.5. Construction and validation of prognostic nomograms

Both univariate and multivariate Cox regression analyses were employed to assess the independent prognostic value of the risk score. Receiver Operating Characteristic (ROC) curves with the area under the curve (AUC) value was plotted by the "timeROC" package to evaluate the predictive capabilities of age, gender, AJCC stage, Grade, and the risk score [23]. Subsequently, the prognostic model and clinical factors identified through multivariate regression analysis were integrated to construct prognostic nomograms. Calibration curves were generated to assess the deviation between the nomograms and the ideal model. The classification performance of the model was further evaluated through the calculation of ROC curves, Decision Curve Analysis (DCA) curves, and the construction of calibration curves.

2.6. Single cell filtering, dimensionality reduction, clustering and annotation

Single-cell analysis was implemented on 9 LC samples from the GSE125449 dataset. Filtering was firstly applied to the single-cell data by setting a minimum expression requirement of each gene in at least three cells, ensuring each cell expressed at least 200 genes. The PercentageFeatureSet function in Seurat package was subsequently applied to calculate the proportions of mitochondrial and rRNA, making sure that the number of genes expressed in each cell genes was between 200 and 3000, with mitochondrial content below 10 % and a minimum unique molecular identifier (UMI) count greater than 200 [24]. Standardization and identification of highly variable genes were performed. Subsequently, the ScaleData function was employed to scale all genes, followed by PCA dimensionality reduction to identify anchor points with dim = 20. Batch correction was then carried out using the Harmony algorithm to harmonize the data across batches.

2.7. Cell culture and transfection

In this study, the following cell lines, including normal hepatocyte cell line THLE-2 (SCSP-5068) and LC cell line Huh7 (SCSP-526), were all tested via Short Tandem Repeat (STR) profiling analysis and ordered from Cell Bank, Chinese Academy of Science (Shanghai, China). BEGM culture media (CC-3170, Lonza, Basel, Switzerland) blended with 5 ng/mL epidermal growth factor (130-093-825, Miltenyi Biotec, Inc., Auburn, CA), 70 ng/mL phosphoethanolamine (P612818, Aladdin, Shanghai, China) and 10 % bovine calf serum (C0234, Beyotime Institute, Shanghai, China) were applied for the culture of THLE-2 cells, and Huh7 cells were grown in Dulbecco's modified Eagle's medium (11965065, Gibco, Grand Island, NY) with 10 % bovine calf serum. All cells were incubated in the CO_2 incubator at the conditions of 37 °C, 5 % CO_2.

For the transfection via liposome, the corresponding small interfering RNAs for UCK2 knockdown (target sequence: 5′-GCCTTTGACAATGAACTCATTCT-3′) or the control were synthesized and purchased from GenePharma (Shanghai, China), which were then transfected into Huh7 cells with the help of lipofectamine 2000 transfection reagent (11668027, Invitrogen, Carlsbad, CA) at 37 °C for 48 h. These stably transfected cells were hereafter harvested for the invasion and migration assays, respectively.

2.8. Reverse-transcription quantitative PCR

RNAiso Plus total RNA extraction kit (#9108, Takara Bio, Shiga, Japan) was applied to extract the total cellular RNA from normal hepatocyte cell line THLE-2 and Huh7 as per the manuals. The corresponding complementary DNA (cDNA) was synthesized with a corresponding assay kit (#6210B, Takara Bio, Japan) and subjected to the PCR analysis in a relevant PCR assay kit (#RR036B, Takara Bio, Japan) in ABI7500 thermocycler (ThermoFisher Scientific, Waltham, MA). The relative mRNA expression level was finally

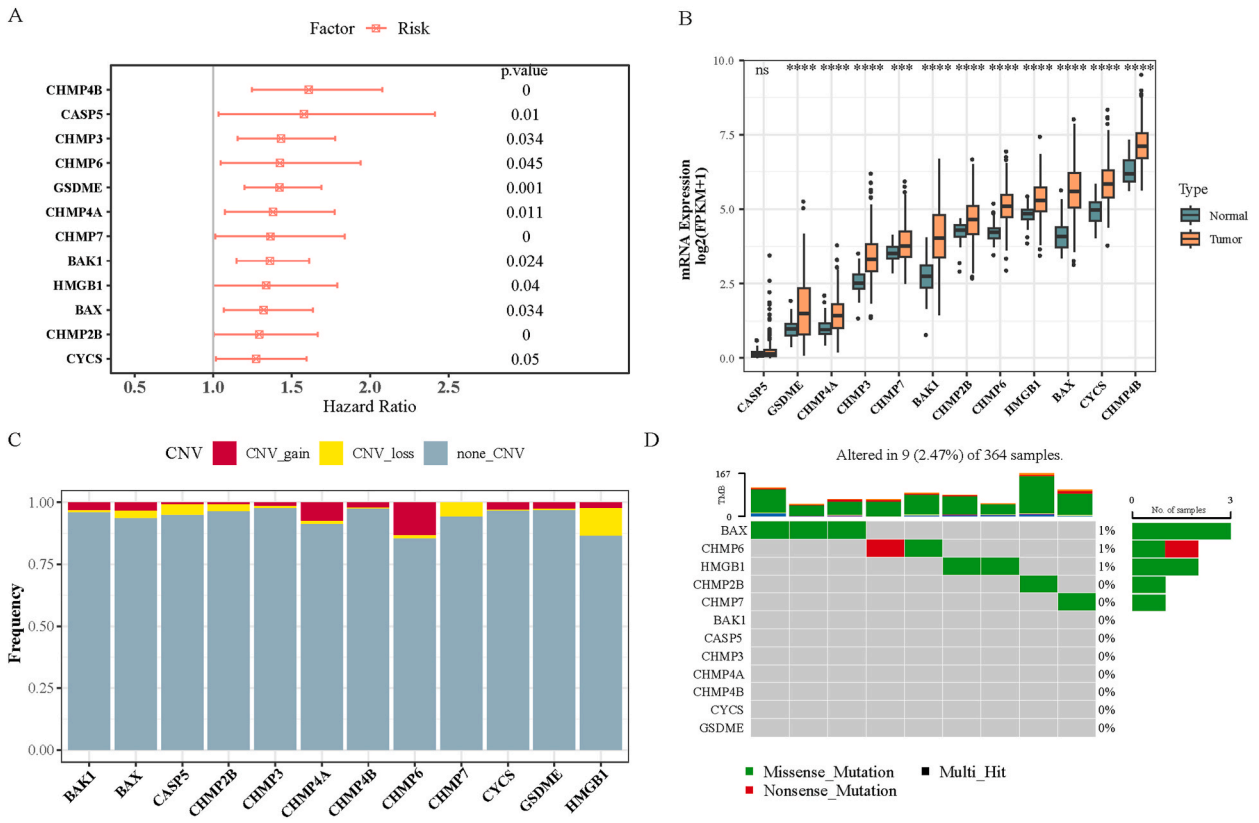


Fig. 1. SNV and CNV analysis of pyroptosis-related genes in liver cancer. A: Forest plot of pyroptosis-related genes significantly related to the prognosis in liver cancer. B: Box plot on the expression levels of pyroptosis-related genes significantly related to the prognosis in liver cancer and normal tissues. C: Changes on the copy number in pyroptosis-related genes significantly related to the prognosis. D: Somatic cell mutation frequency of pyroptosis-related genes significantly related to the prognosis.

calculated via the method $2^{-\Delta\Delta ct}$ and normalized to the housekeeping control GAPDH [25]. The primers applied in this study were all listed in Supplementary Table 1.

2.9. Transwell cell invasion assay

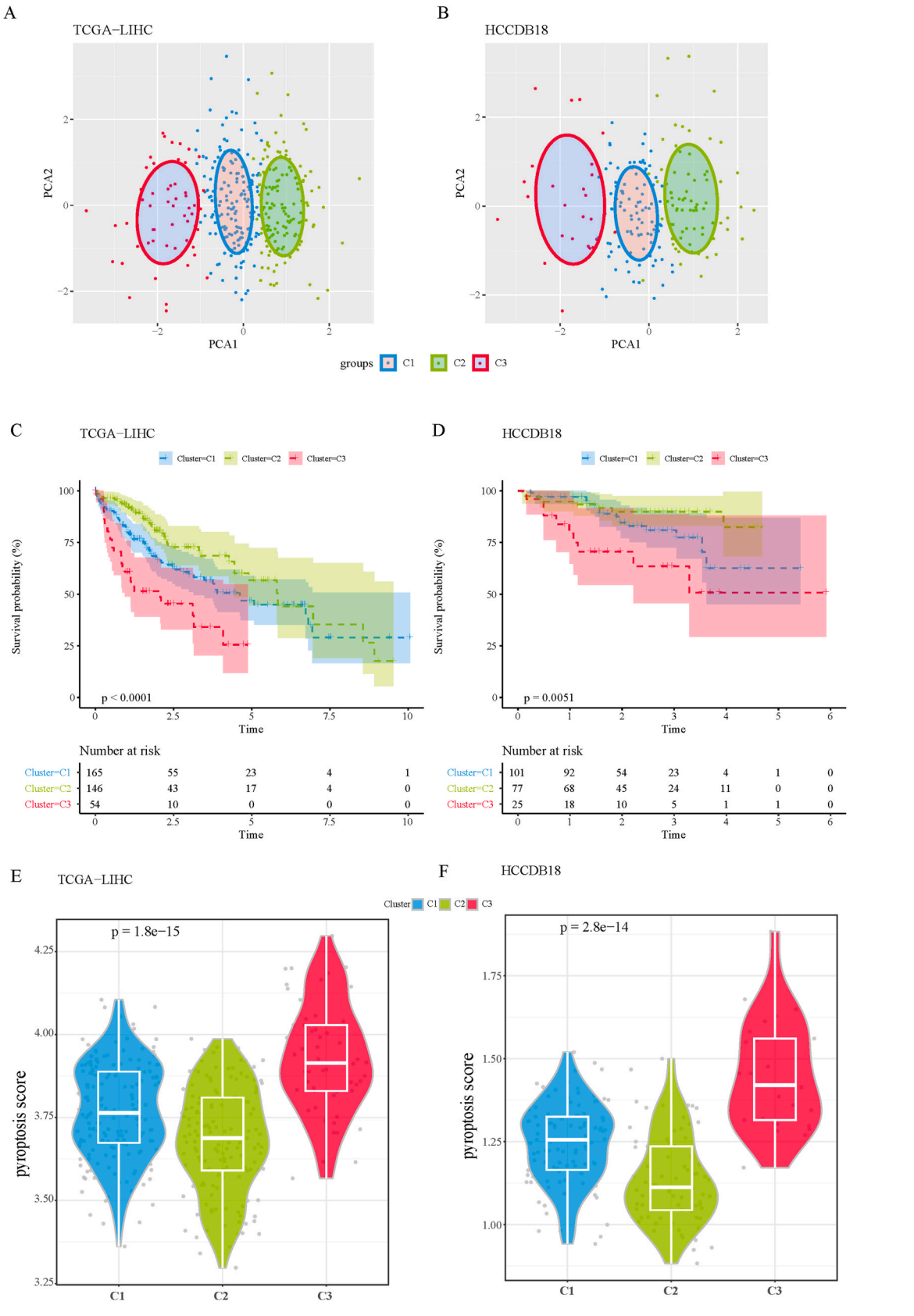
A total number of 1×10^5 Huh7 cells were added to the upper chambers of the transwell inserts (3422, Corning, Inc., Corning, NY) coated with matrix gel (C0371, Beyotime Institute, China) and added with 200 μL cell culture medium deprived of the serum. The bottom chamber was accordingly added with 600 μL complete medium with 10 % serum as the chemoattractant. After 48 h, the cells invaded to the bottom chamber were serially fixed in 4 % paraformaldehyde (P0099, Beyotime Institute, China) for 15 min and stained in 0.1 % crystal violet staining solution (C0121, Beyotime Institute, China) for 10 min at ambient temperature. An inverted light microscope (DP27, Olympus Corp., Tokyo, Japan) was applied to observe the invaded cells and then the number of cells was calculated.

2.10. Scratch cell migration assay

Huh7 cells were seeded in a 6-well plate at 1×10^5 cells per well and the wound on the surface was made when cell grew >90 % confluence. Then these cells were grown in the serum-deprived culture media for 48 h, and the wound width was visualized and imaged at 0 and 48 h under an inverted light microscope. The wound closure (%) was accordingly calculated based on the width observed.

2.11. Statistical analyses

All statistical analyses were realized in R software (version 3.6.0) and GraphPad Prism software (version 8.0.2). The Wilcoxon rank-sum test or Student's t-test was employed to assess differences between two groups of continuous variables. The differences among three groups of continuous variables were evaluated via Kruskal-Wallis test. The Chi-square test was applied to calculate differences in categorical variables among different groups. The log-rank test was used to compare survival times among patients in different groups



(caption on next page)

Fig. 2. Molecular subtyping of liver cancer based on the pyroptosis-related genes
A-B: PCA analysis on the subtypes of liver cancer according to the data of TCGA and HCCDB18. C–D: The survival of patients with different molecular subtypes of liver cancer (C1, C2 and C3) based on the data of TCGA and HCCDB18, and the results were shown in the Kaplan-Meier curves. E–F: Pyroptosis score on the different molecular subtypes of liver cancer (C1, C2 and C3) using the data from TCGA and HCCDB18.

based on the Kaplan-Meier curve. A significance level at $p < 0.05$ was indicative of the statistical significance.

3. Results

3.1. SNV and CNV analysis of PRGs in LC

In the beginning, a univariate COX analysis was performed on the 27 genes of REACTOME_PYROPTOSIS in GSEA, revealing that there were 12 genes significantly associated with prognosis, which were also revealed as the risk factors ($p < 0.05$, Fig. 1A). The expression analysis on these 12 genes expressed in LC and adjacent tissues revealed the significantly higher expression of 11 genes in LC tissue, excepting for CASP5 (Fig. 1B). Then, it was discovered that these genes had no apparent amplification or deletion of gene copies of copy number variations (Fig. 1C). Notably, the CHMP6 gene exhibited a substantial copy number amplification compared with other genes and these genes had a low mutation frequency. The mutation frequencies of BAX, CHMP6, and HMGB1 are 1 %, while those for the remaining genes are below 1 % (Fig. 1D). Based on these analyses, it can be concluded that these PRGs may have a certain relationship with the progression of LC.

3.2. Molecular subtyping of LC based on the PRGs

Then, we explored the relationship between the above PRGs and LC. Based on the above 12 PRGs and prognosis-related genes, a consensus clustering was performed in TCGA cohort with the optimal clustering number from the cumulative distribution function (CDF). According to the results on the CDF Delta area curve (Supplementary Fig. 1), the samples were perfectly classified into three molecular subtypes, a discovery seen in both TCGA sample and HCCDB18 dataset.

To verify the rationality of the classification, PCA analysis was performed using 12 genes, and the obvious boundary was seen in the three molecular subtypes, according to the data of TCGA and HCCDB18 datasets (Fig. 2A and B). Next, we analyzed the survival difference of these three subtypes, revealing the worst prognosis in the C3 subtype and the best prognosis in the C2 subtype (Fig. 2C and D).

Meanwhile, we analyzed the expression of cellular pyroptosis pathway scores in different molecular subtypes. It was clear that the trends of prognosis and cellular pyroptosis scores were consistent in TCGA and HCCDB18, with the worst prognosis and the highest cellular pyroptosis score observed in the C3 subtype (Fig. 2E and F). These discoveries hinted the possible effect of the pyroptosis pathway on the prognosis of LC patients.

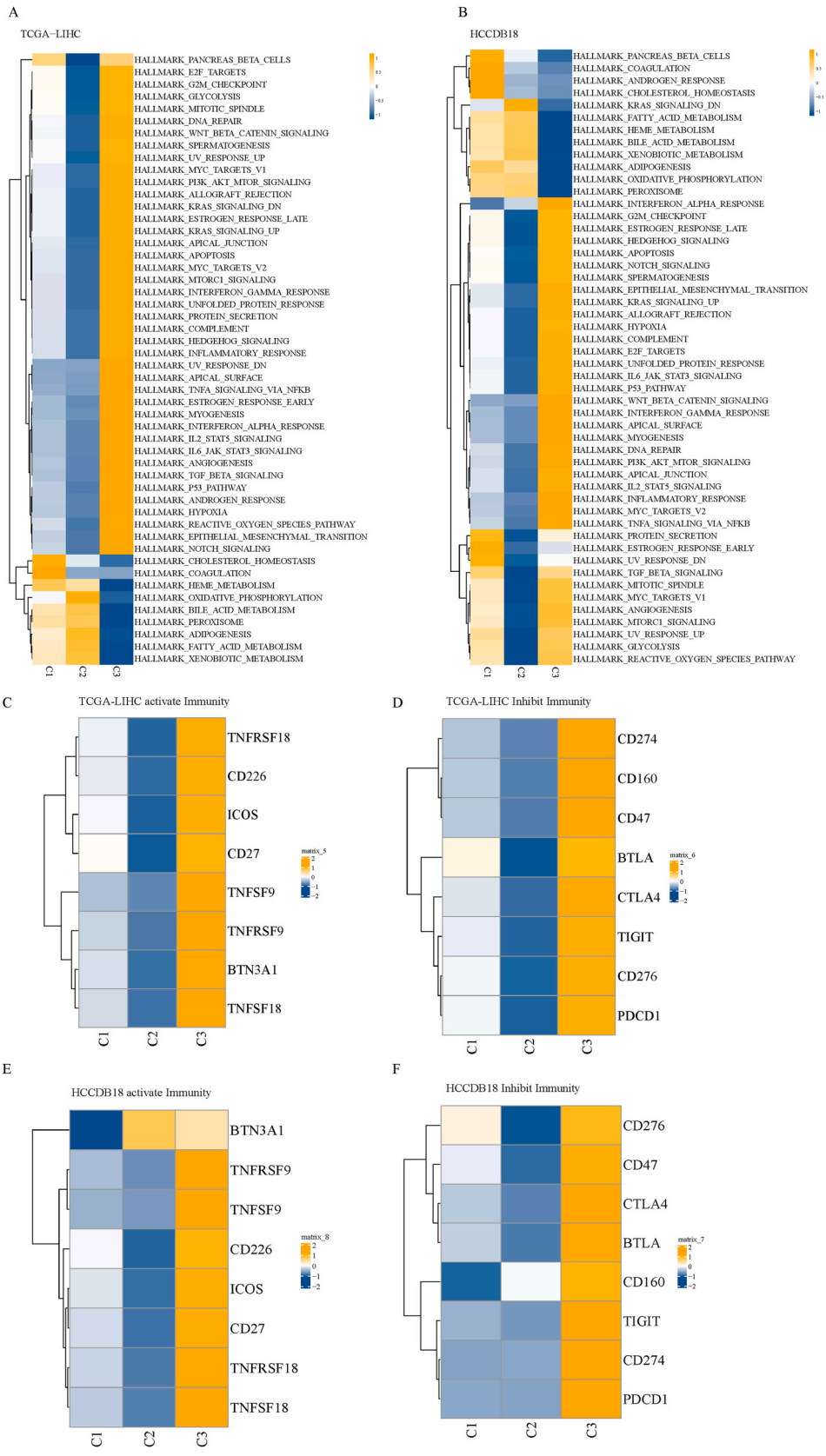
3.3. Changes in immune pathways across different molecular subtypes of LC

Furthermore, we analyzed the enriched Hallmark pathways, those related to immunity and metabolism in particular, across different molecular subtypes of LC by ssGSEA. Heatmaps were accordingly generated to illustrate the enrichment status, and subsequent comparisons of immune differences within subtypes revealed compelling results. In both the TCGA and HCCDB18 datasets, the C3 subtype of LC with the poorest prognosis exhibited the enrichment across the majority of pathways, encompassing some known to promote cancer progression and metabolic pathways conducive to cancer development, such as Glycolysis (Fig. 3A and B).

Unexpectedly, we also observed the enrichment of the C3 subtype in the activation of immune-related pathways, which was contrary to our initial expectations. This leads us to speculate whether the immunoactivation might be concurrent with the immune escape, which allows tumor cells to survive and progress. To test this hypothesis, we tested the expression of immunoactivation-related genes and found that these genes were highly expressed in the C3 subtype compared to the other two subtypes, consistent with the results on the enriched pathway. Subsequently, the expressions of immune checkpoint genes were determined and higher expressions of these genes were also seen in the C3 subtype (Fig. 3C–F). Although the C3 subtype shows signs of both immune activation and immune escape, this phenomenon is not contradictory. We hypothesize that tumor cells of the C3 subtype may manipulate the tumor microenvironment through transient immune activation, which attracts immune cell infiltration, but at the same time protect themselves from killing by the immune system through immune escape mechanisms (e.g., increased expression of immune checkpoint molecules).

3.4. Construction of prognostic models

The DEGs analysis was performed among these 3 molecular subtypes of LC, and the corresponding results were illustrated in volcano plot (Fig. 4A). Hereafter, the univariate and LASSO Cox and stepwise regression analysis of DEGs was performed to identify the candidate model genes (Fig. 4B and C). Following this, a stepwise multivariable regression analysis was employed to further narrow down the genes and calculated the regression coefficient. Ultimately, 8 genes were identified as key prognostic factors, and the RiskScore was calculated as follows (Fig. 4D): RiskScore= (−0.254 × expression level of KPNA2) + (0.298 × expression level of UCK2)



(caption on next page)

Fig. 3. Changes in immune pathways across different molecular subtypes of liver cancer
A-B: The enrichment of Hallmark pathway in different liver cancer subtypes based on the data cohort from TCGA and HCCDB18. C–F: Expression of immunoactivation-related or immune escape-related genes in different liver cancer subtypes based on the data from TCGA and HCCDB18 data cohort.

+ (−0.084 × expression level of FTCD) + (0.17 × expression level of CBX2) + (0.18 × expression level of RAB32) + (0.309 × expression level of HMMR) + (0.12 × expression level of S100A9) + (−0.101 × expression level of ANXA10).

We computed the individual risk scores based on the expression levels of gene and divided the samples into high and low risk subgroups based on the optimal cutoff point of Riskscore determined via surv_cutpoint function in the survminer package. Subsequently, the ROC analysis was conducted on the RiskScore for prognostic classification employing the R package timeROC, and the predictive efficiency for 1-, 3-, and 5-year survival in all TCGA samples was also accessed. A notably high AUC for each time point was demonstrated in the model, with higher AUC indicative of the better prediction accuracy and the robust prognostic performance. Besides, the poorer outcomes were seen in the high-risk subgroup (Fig. 4E–H). Finally, we analyzed the proportions of survival and mortality in the high and low Risk subgroups, from which we confirmed the higher proportion in the low-risk subgroup (Fig. 4I).

3.5. Validation on the accuracy of the prognostic model

To validate the accuracy of the model, we assessed the prognostic classification performance on external validation sets (HCCDB18 and GSE14520). It is evident that the model exhibits high robustness and accuracy, with poorer prognosis showing in the high-risk group (Fig. 5A–B, D-E). Further, we analyzed the distribution of mortality and survival within the high and low Risk groups in HCCDB18 and GSE14520 cohorts, and a higher proportion of mortality was seen in the high-risk group of both cohorts (Fig. 5C–F).

3.6. Manifestation of RiskScore on clinicopathological features in different molecular subtypes of LC

In this phase, based on the TCGA dataset, we compared the differences in RiskScore and PYROPTOSIS scores among our identified molecular subtypes of LC. The corresponding results revealed that all samples in the C3 subtype, with the worst prognosis, belonged to the high-risk group (Fig. 6A). Notably, the trends on PYROPTOSIS scores were consistent with those of RiskScore, and the PYROPTOSIS score was highest in C3 subtype and lowest in C2 subtype (Fig. 6B). Also, we compared the differences in PYROPTOSIS scores between these two groups in the TCGA dataset, and the significantly lower scores were seen in the low-risk group (Fig. 6C). By examining the distribution of RiskScore among clinical-pathological features in the TCGA dataset, significant differences in RiskScore were clear in different AJCC Stages and Grades (Fig. 6D and E).

3.7. Improving effects of RiskScore on prognostic modeling and survival prediction by incorporating clinicopathological features

Here, we conducted univariate and multivariate Cox regression analyses based on RiskScore and clinicopathological features in the TCGA-LIHC cohort. The results revealed that AJCC stage and RiskScore were significant prognostic factors (Fig. 7A–B). Then, to quantify the risk assessment and survival probability of LC patients, a nomogram was constructed combining RiskScore and AJCC stage (Fig. 7C). Notably, it was indicated in the model that RiskScore had the most substantial impact on the prediction of survival rate. Furthermore, we evaluated the predictive accuracy of the model using a calibration curve (Fig. 7D), where the calibration points for 1, 3, and 5 years closely aligned with the standard curve, suggesting excellent predictive performance of the nomogram. Additionally, we assessed the reliability of the model using DCA, and a significant benefit of the nomogram over the extreme curves was observed (Fig. 7E).

3.8. Single-cell analysis

In the end, the single-cell analysis was implemented on 9 LC samples from the GSE125449 dataset. After standardization, dimensionality reduction, and clustering, 6 major cell subtypes were identified (Fig. 8A), comprising a total of 3913 cells. The signature genes for each cell subtype were also delineated (Fig. 8B), like Endothelial, Epithelial cell, Fibroblasts, NK/T cells, Monocytic-Macrophages, Plasma cells.

The results on the expressions of 8 genes identified during the construction of risk model in these cells were illustrated in Fig. 8C, with the majority of genes expressed in epithelial cells and S100A9 and RAB32 expressed in Monocytic-Macrophages. Next, we calculated the mean gene expression for each cell group using the AverageExpression function and computed pyroptosis scores for each cell subtype using the gsva package. The differences in scores among different cells were compared, revealing the highest pyroptosis score in Monocytic-Macrophages (Fig. 8D). In a word, it could be deduced that the particular occurrence of cell pyroptosis in Monocytic-Macrophages may contribute to the progression of LC.

3.9. In-vitro cellular validation

To make a preliminary investigation on the involvement of the key genes of RiskScore analysis in LC, the levels of these key genes in LC cells Huh7 and normal hepatocyte THLE-2 were quantified and compared. Accordingly, the higher expressions of KPNA2, UCK2,

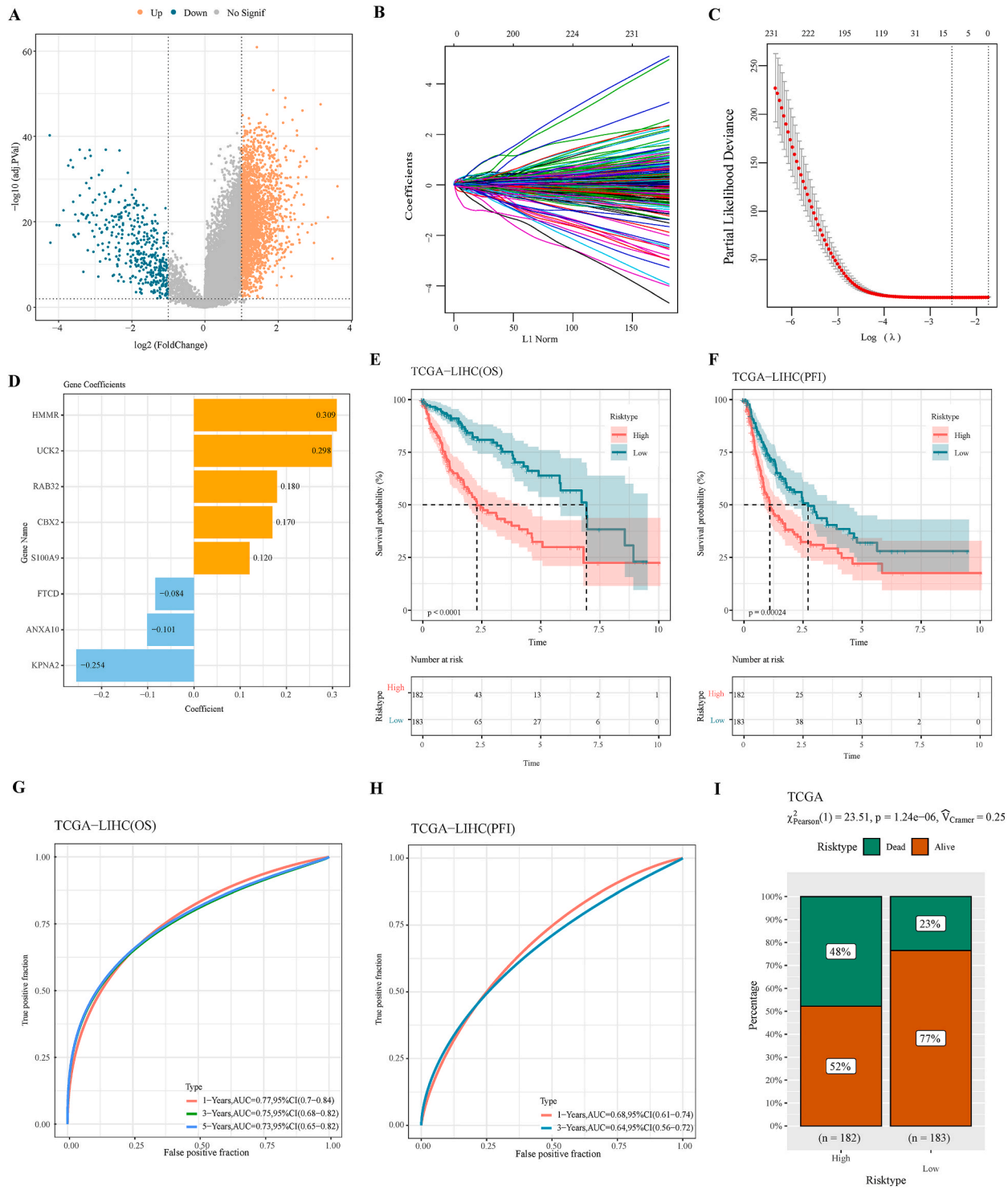


Fig. 4. Construction of prognostic models

A: Volcano plot showing differentially expressed genes in different molecular subtypes of liver cancer. B: Trajectory of each independent variable with lambda trajectory of each independent variable. C: Confidence intervals under lambda. D: Risk coefficients corresponding to modeled genes from multifactorial screening. E–F: Kaplan-Merve survival curves of OS and PFI in the high and low RiskScore subgroups of the TCGA data cohort. G–H: ROC curves for OS and PFI in the high and low RiskScore subgroups of the TCGA data cohort. I: Percentage of mortality and survival in the high and low RiskScore subgroups of the TCGA data cohort.

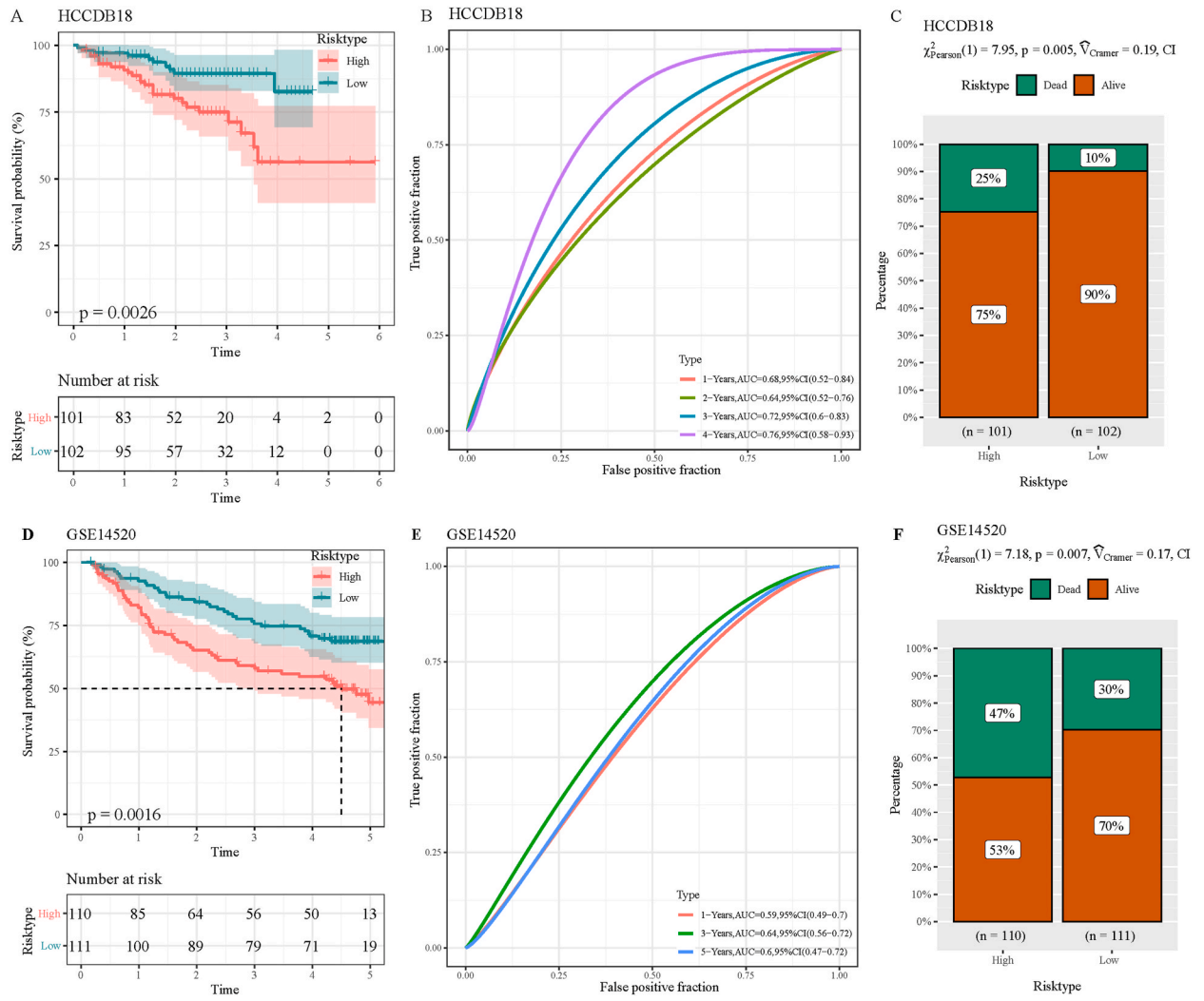


Fig. 5. Validation on the accuracy of the prognostic model
A-B Kaplan-Merve survival curves and ROC curves for RiskScore in the HCCDB18 cohort. C: Percentage of mortality and survival in the high and low risk subgroups of the HCCDB18G cohort. D–E: Kaplan-Merve survival curves and ROC curves for RiskScore in the GSE14250 cohort. F: Percentage of mortality and survival in the high and low risk groups of the GSE14250 cohort.

CBX2, RAB32, HMMR and S100A9 and the lower expressions of FTCD and ANXA10 were seen in LC cells Huh7 as compared to those in normal hepatocyte THLE-2 ($p < 0.05$, Fig. 9A). In the subsequent transwell cell invasion and scratch cell migration assays, it was additionally noted that the silencing of UCK2 via specific small interfering RNA could visibly diminish the number of invaded and migrated cells at 48 h ($p < 0.01$, Fig. 9B–C).

4. Discussion

Our current study, in the beginning, examined the genes obtained from GSEA and confirmed 12 of these genes as the risk factors in LC, which, along with those PRGs significantly related to the prognosis of LC, were then applied for the consensus clustering. Three molecular subtypes of LC, accordingly, have been identified, with a worst prognosis and a highest pyroptosis score seen in C3 subtype of LC. The relevant heatmaps showing the enriched pathways of these subtypes have demonstrated the weird concurrence of both immunoactivation and immune escape in C3 subtype of LC, which could be possibly interpreted by the abnormally high expressions of genes relevant to these two procedures. Such a discovery, so far as we are concerned, was seen in glioma, where cuproptosis, another form of cell death, can facilitate immunoactivation but promotes immune escape T. Hereafter, with the aim to construct a prognostic model, 8 PRGs have been sorted identified, including KPNA2, UCK2, CBX2, RAB32, HMMR, S100A9, FTCD and ANXA10. Then these PRGs were applied to compute the risk score which divided LC patients in TCGA data cohort and external validation sets (HCCDB18 and GSE14520) into high-risk and low-risk groups. Further exploration has suggested the overall survival (OS) of low-risk group was

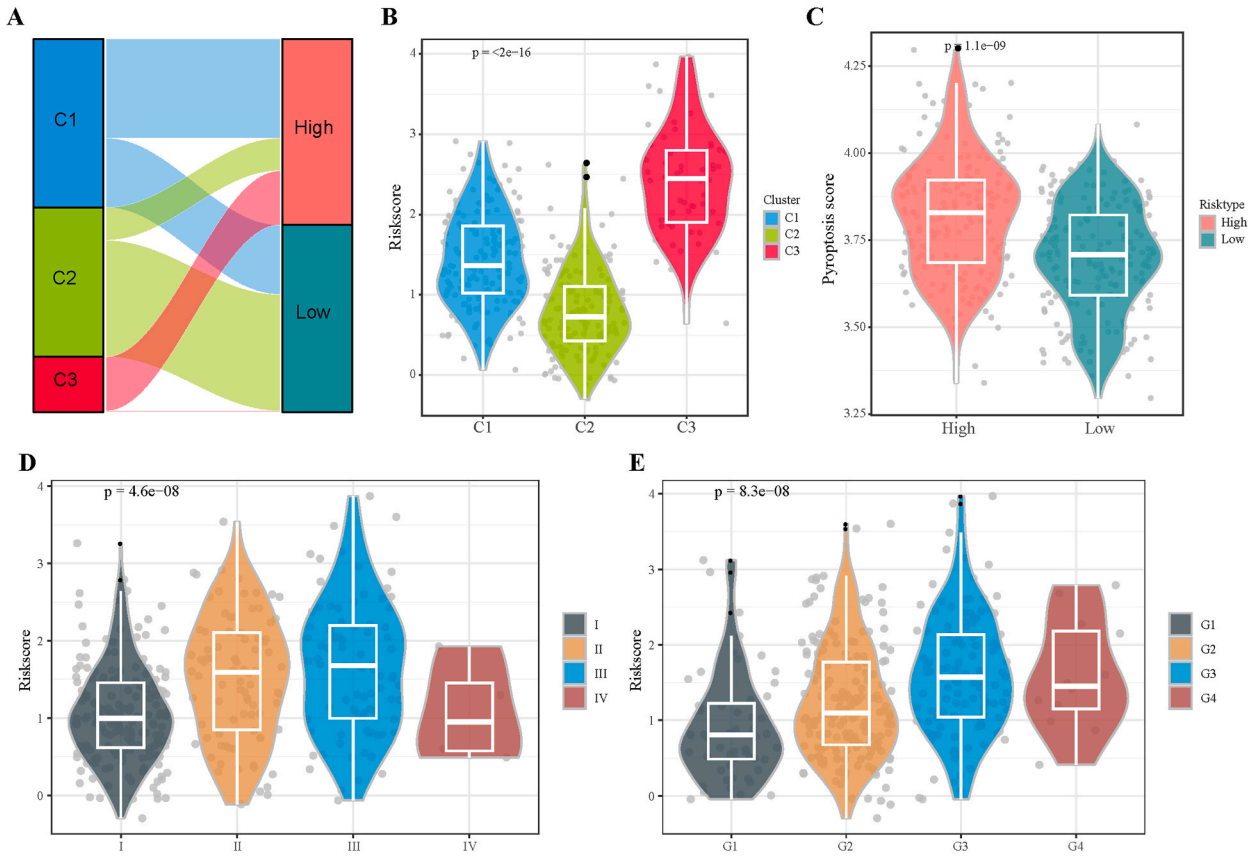


Fig. 6. Manifestation of RiskScore on clinicopathological features in different molecular subtypes of liver cancer. A: The distribution of high and low risk groups in the subtypes of liver cancer in TCGA cohort. B: The pyroptosis scores in the subtypes of liver cancer in TCGA cohort. C: The pyroptosis scores in the high and low risk subgroups. D: The distribution of RiskScore among different AJCCStage in TCGA cohort. E: The distribution of RiskScore among different grades in TCGA cohort.

better than that of the high-risk group, a trend consistent with the previous study [26]. Additionally, we have confirmed the manifestation of risk score in both different subtypes of LC divided in our study (as indicated by the fact that the C3 subtype of LC was categorized into the high-risk group) and different clinicopathological features, AJCC stage and tumor grade, for instance. Besides, we further validated the performance of our RiskScore in improving the prognosis and carried out a scRNA-seq to reveal the potential involvement of S100A9 and RAB32 as well as monocytic-macrophages in pyroptosis of LC.

Numerous studies have indicated the association of cell pyroptosis with anti-cancer immunity, which has altered the strategies for anti-tumor therapy in cancer immunology and tumorigenesis [27]. Besides, both pyroptosis and PRGs 3 serve as potential therapeutic targets for various cancers [28]. For instance, a system analysis based on PRGs carried out by Yan et al. has identified Gasdermin C as a novel therapeutic target for pancreatic adenocarcinoma [29]. Another study in gastric cancer has established a 5 gene signature based on 63 PRGs which plays a profound role in the prognosis of gastric cancer [30]. Besides, 35 PRGs have been identified to participate in tumor immunity and have the potential to predict the prognosis of hepatocellular carcinoma, [26]. In our current study, 8 PRGs have been identified, including KPNA2, UCK2, CBX2, RAB32, HMMR, S100A9, FTCD and ANXA10, and their differential expression in LC was observed. In other words, the higher expressions of KPNA2, UCK2, CBX2, RAB32, HMMR and S100A9 and the lower expressions of FTCD, ANXA10 were seen in LC cells Huh7 as compared to those in normal hepatocyte THLE-2. Further transwell cell invasion and scratch cell migration assays results have demonstrated the anti-LC effects of UCK2 silencing on LC cells, as exemplified by the reduced number of invaded and migrated LC cells. KPNA2 is a member of the nuclear transport protein family overexpressed in a variety of cancers like hepatocellular carcinoma, as revealed in a single-cell transcriptome analysis [31,32]. UCK2 is a rate-limiting enzyme overexpressed in several solid and hematological cancers, which is associated with the aggressiveness of hepatocellular carcinoma [33]. CBX2 is a polycomb protein which is closely related to tumorigenesis and tumor development, which can drive a stem cell-like phenotype in hepatocellular carcinoma [34,35]. Rab32 belongs to the Ras-like small GTPase superfamily and acts as an oncogene in hepatocellular carcinoma, with the potential to promote tumor cell growth and invasion [36,37]. HMMR, also known as CD168, has been reported play diverse functional roles in various cancers, including lung, liver, bladder, and gastric cancers [38]. S100A9 is primarily expressed in neutrophils and monocytes, playing a crucial role in regulating various inflammatory responses and inflammation-related diseases [39]. Increasing evidences have further proven the involvement of S100A9 in tumors like liver cancer

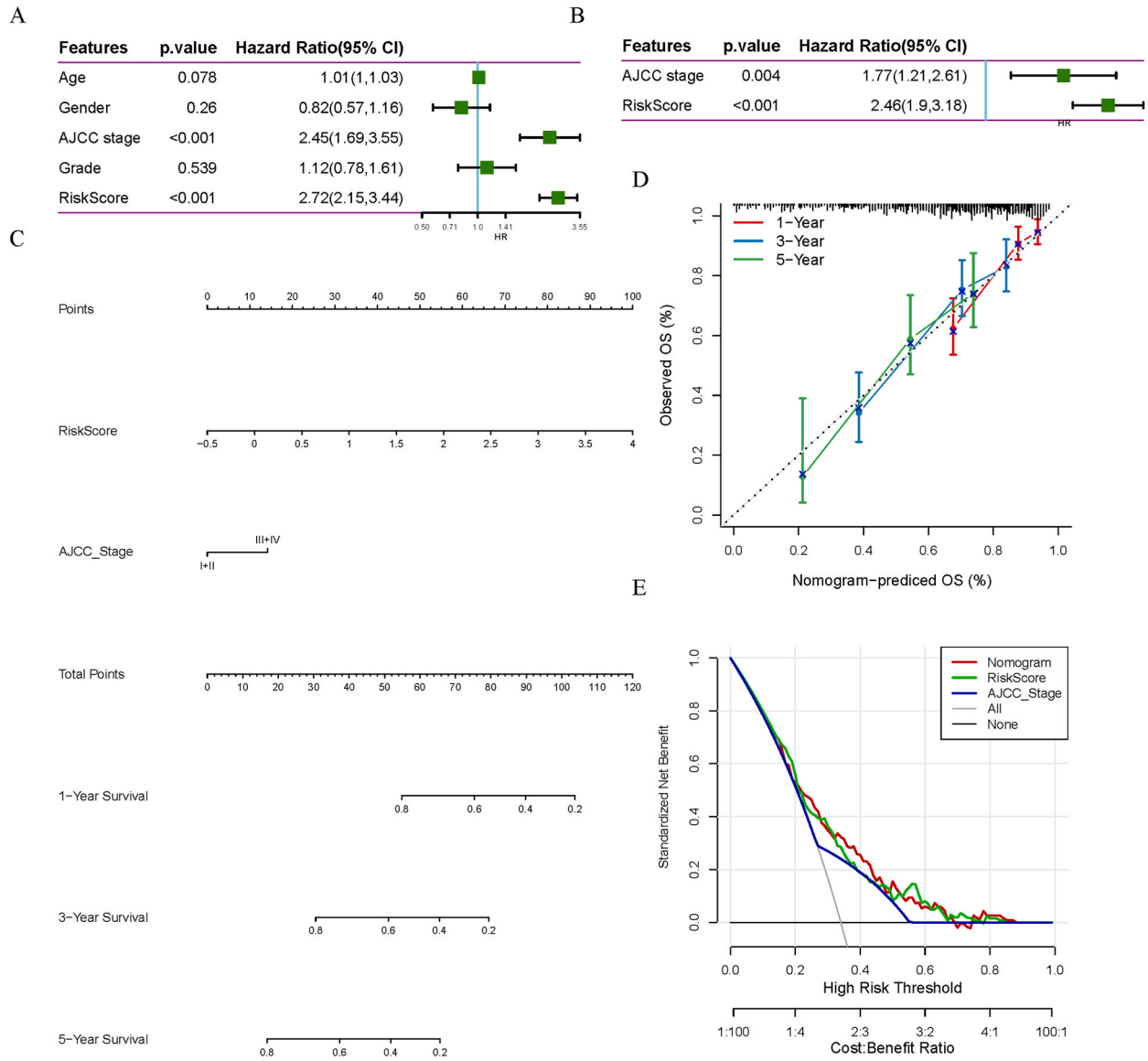


Fig. 7. Improving effects of RiskScore on prognostic modeling and survival prediction by incorporating clinicopathological features. A: Univariate Cox analysis on RiskScore and the clinicopathological features. B: Multivariate Cox analysis on RiskScore and the clinicopathological features. C: Constructed nomogram combining RiskScore and AJCC stage to quantify the risk assessment and survival probability of liver cancer patients. D–E: The calibration curve and the decision curve of the monogram.

[40,41]. FTCD exhibits selective high expression in normal liver and hepatic loss of FTCD contributes to lipid accumulation and hepatocarcinogenesis [42,43]. ANXA10 is a member of the annexin protein family located on chromosome 4q33 and a bioinformatics analysis and experimental validation have confirmed its role as a prognostic biomarker and suppressor of hepatocellular carcinoma [43]. In addition to their involvement in HCC, some other studies have proven the involvement of some of these 8 PRGs in pyroptosis with the peculiar mechanisms implicated [44,45]. To our surprise, another existing study linking pyroptosis and hepatocellular carcinoma has further demonstrated the potential of CBX2 as one of the genes used to construct the pyroptosis-related prognostic model [46]. Likewise, in our current study, CBX2 was listed as one of the PRGs applied to construct the RiskScore (constructed using these 8 PRGs). Also, the scRNA-seq result has revealed the potential involvement of S100A9 and RAB32 as well as monocytic-macrophages in pyroptosis of LC. An in-vitro study by Yang et al. has already proven the promoting effects of S100A9 on the polarization of M2 macrophages so as to drive LC progression [41]. Nonetheless, whether S100A9 and RAB32 are involved in modulating macrophage pyroptosis in LC has not been explored, which, we believe, will be the direction of our future studies. Conducting additional research to explore the foundational connection between S100A9, RAB32, and the prognosis of patients with LC is essential.

In addition to the one mentioned above, there are some other inescapable limitations of this study that should be addressed. For

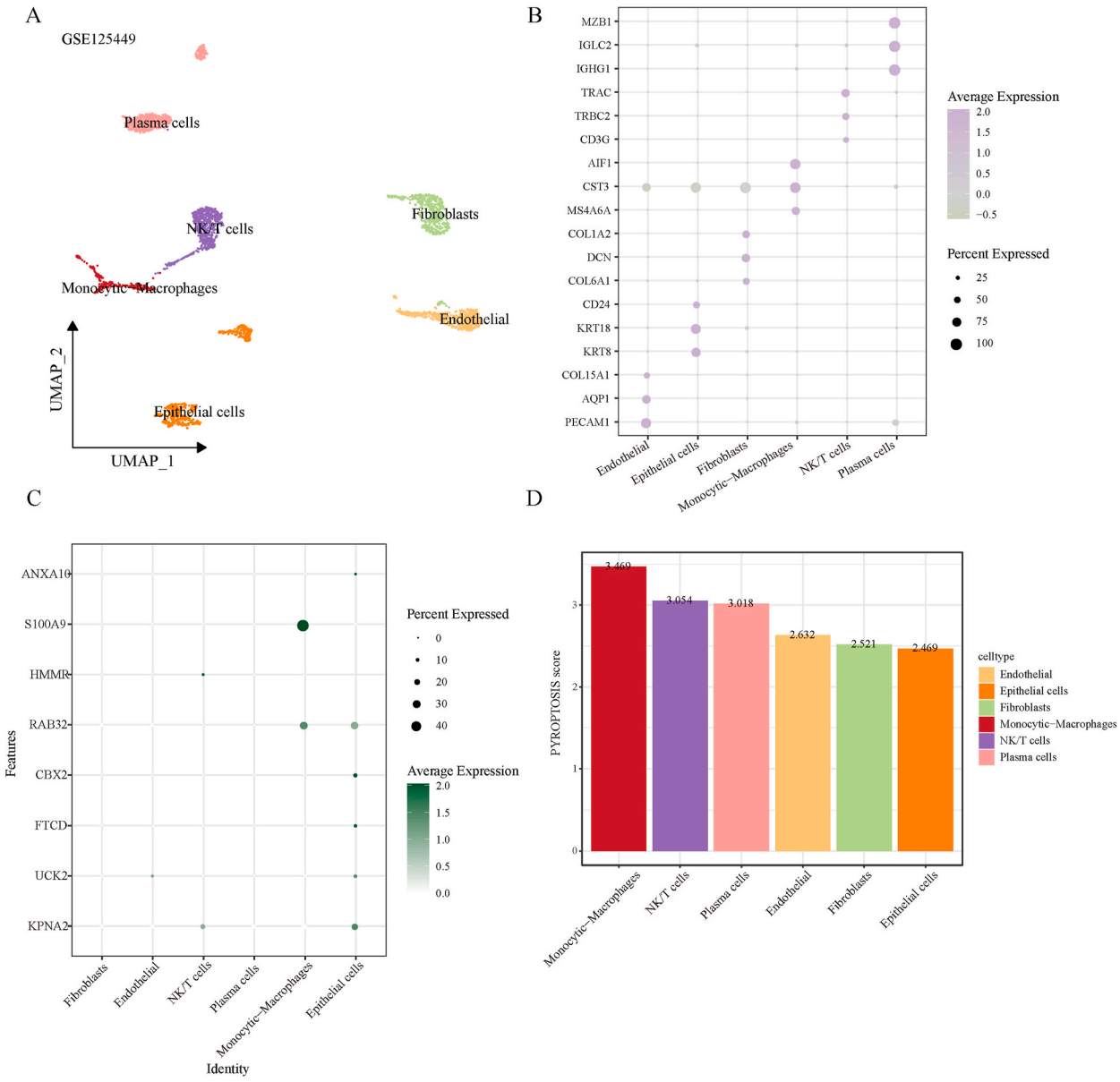


Fig. 8. Single-cell analysis. A: UMAP plot for cell annotation based on the samples from the dataset GSE125449. B: Specific marker genes expressions in different cell subtypes (endothelial, epithelial cells, fibroblasts, monocytic-macrophages, NK/T cells and plasma cells). C: Expression levels of genes identified during risk model construction in different cell subtypes (endothelial, epithelial cells, fibroblasts, monocytic-macrophages, NK/T cells and plasma cells). D: Pyroptosis scores of different cell subtypes (endothelial, epithelial cells, fibroblasts, monocytic-macrophages, NK/T cells and plasma cells).

instance, the sample size for scRNA-seq is relatively small, which only includes one dataset GSE125449. Whether the results can be reproduced to other LC-related datasets remains a mystery. Secondly, all these results are generated solely from bioinformatics and accordingly, relevant verification will be implemented to consolidate the results.

5. Conclusion

Extensive research has elucidated the role of PRGs in LC and other tumors. This study employs comprehensive and systematic bioinformatics analysis to reveal the significant involvement of PRGs in LC, where 8 key PRGs have been identified. These findings offer new insights into the potential molecular mechanisms of HCC, shedding light on the role of cell pyroptosis and paving the way for further exploration of more effective immunotherapies. Meanwhile, subtype analysis based on PRGs may offer additional personalized treatment options for LC patients.

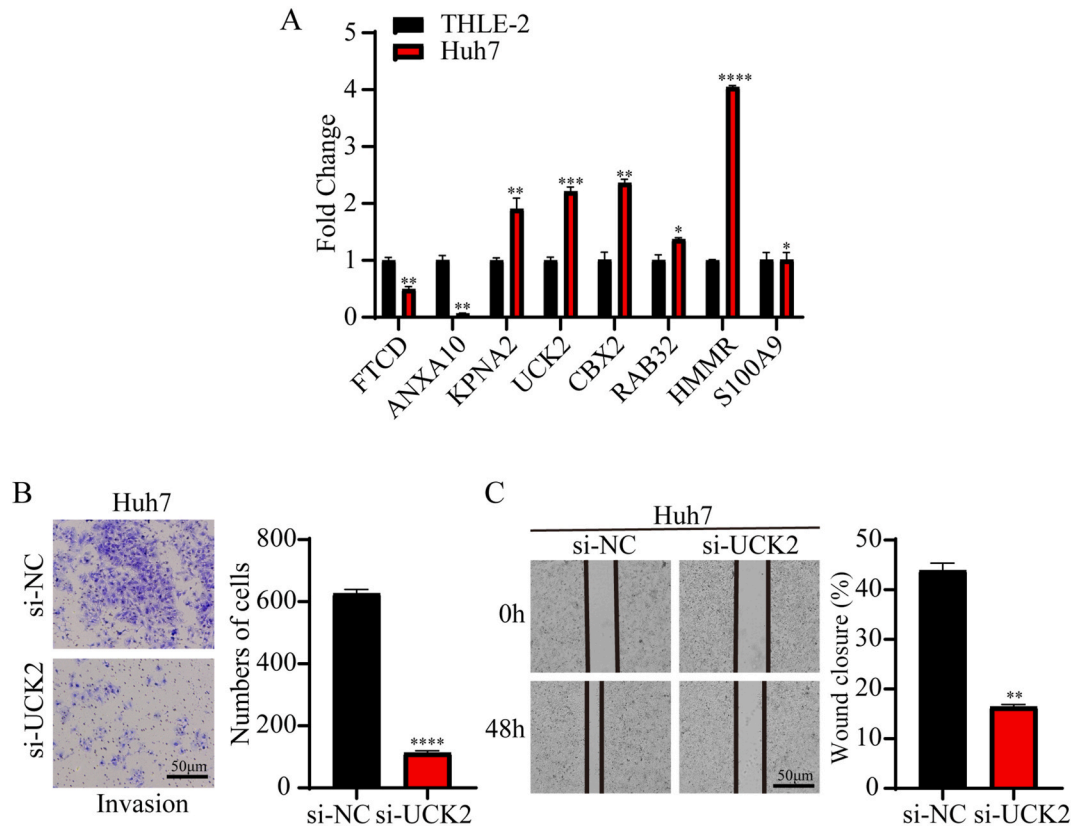


Fig. 9. In-vitro cellular validation. A: Quantified mRNA level of 8 key PRGs (KPNA2, UCK2, FTCD, CBX2, RAB32, HMMR, S100A9 and ANXA10) in normal hepatocyte THLE-2 and LC cell line Huh7 via reverse-transcription quantitative PCR. B: Transwell cell invasion assay showing the effects of UCK2 knockdown on the invasion of LC cells Huh7 at 48 h. C: Scratch cell migration assay depicting the effects of UCK2 silencing on the migration of LC cells Huh-7 cells at 0 and 48 h.

Data availability statement

The datasets generated and/or analyzed during the current study are available in the [GSE14520] repository [https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc= GSE14520], and [GSE125449] repository, [https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc= GSE125449].

Funding

This study is supported by The Youth Science Foundation Grant Program of Shandong First Medical University (No.202201-091).

Ethics approval statement

Not applicable.

Patient consent statement

Not applicable.

CRediT authorship contribution statement

Zhihao Zhang: Writing – review & editing, Writing – original draft, Software, Project administration, Funding acquisition, Data curation, Conceptualization. **Feng Liu:** Writing – review & editing, Writing – original draft, Software, Resources, Project administration, Methodology, Formal analysis, Conceptualization. **Xin Lan:** Visualization, Supervision, Project administration, Methodology, Investigation. **Fuhai Wang:** Visualization, Software, Resources, Methodology, Formal analysis. **Jiahao Sun:** Visualization, Supervision, Resources, Project administration, Methodology, Investigation. **Honglong Wei:** Writing – original draft, Visualization,

Supervision, Project administration, Investigation, Data curation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

None.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e38438.

Abbreviation

liver cancer (LC)
pyroptosis-related genes (PRGs)
single-cell RNA sequencing (scRNA-seq)
The Cancer Genome Atlas (TCGA)
Gene Expression Omnibus (GEO)
Gene Set Enrichment Analysis (GSEA)
Differential expressed genes (DEGs)
Receiver Operating Characteristic (ROC)
the area under the curve (AUC)
Decision Curve Analysis (DCA)
unique molecular identifier (UMI)
complementary DNA (cDNA)
cumulative distribution function (CDF)
overall survival (OS)
Regulated cell death (RCD)
programmed cell death (PCD)

References

[1] D. Anwanwan, S.K. Singh, S. Singh, V. Saikam, R. Singh, Challenges in liver cancer and possible treatment approaches, Biochim. Biophys. Acta, Rev. Cancer 1873 (1) (2020) 188314.
[2] L-x Wang, A-r Kong, H. Dong, Early diagnosis and prognosis of hepatocellular carcinoma based on a ceRNA array, Oncologie 25 (3) (2023) 245–255.
[3] W. Su, F. Su, H. Wang, Y. Wang, L. Ye, P. Zhu, et al., NMR-Based metabolomic techniques identify the anticancer effects of three polyphyllins in HepG2 cells, Curr. Pharmaceut. Anal. 18 (4) (2022) 415–426.
[4] J.M. Llovet, R.K. Kelley, A. Villanueva, A.G. Singal, E. Pikarsky, S. Roayaie, et al., Hepatocellular carcinoma, Nat. Rev. Dis. Prim. 7 (1) (2021) 6.
[5] R.L. Siegel, K.D. Miller, N.S. Wagle, A. Jemal, Cancer statistics, 2023, Ca - Cancer J. Clin. 73 (1) (2023) 17–48.
[6] Y. Zheng, Y. Li, J. Feng, J. Li, J. Ji, L. Wu, et al., Cellular based immunotherapy for primary liver cancer, J. Exp. Clin. Cancer Res. : CR 40 (1) (2021) 250.
[7] Q. Gao, H. Zhu, L. Dong, W. Shi, R. Chen, Z. Song, et al., Integrated proteogenomic characterization of HBV-related hepatocellular carcinoma, Cell 179 (2) (2019), 561-77.e22.
[8] M.H. Bao, C.C. Wong, Hypoxia, metabolic reprogramming, and drug resistance in liver cancer, Cells 10 (7) (2021).
[9] B. Chai, J. Qiu, W. Pan, Z. Ma, Pyroptosis-related noncoding RNAs and cancer involvement, Oncologie 25 (4) (2023) 335–343.
[10] Y. Fang, S. Tian, Y. Pan, W. Li, Q. Wang, Y. Tang, et al., Pyroptosis: a new frontier in cancer, Biomedicine & pharmacotherapy = Biomedecine & pharmacotherapie 121 (2020) 109595.
[11] Z. Rao, Y. Zhu, P. Yang, Z. Chen, Y. Xia, C. Qiao, et al., Pyroptosis in inflammatory diseases and cancer, Theranostics. 12 (9) (2022) 4310–4329.
[12] C. Stoess, Y.K. Choi, J. Onyuru, H. Friess, H.M. Hoffman, D. Hartmann, et al., Cell death in liver disease and liver Surgery, Biomedicines 12 (3) (2024).
[13] Y. Li, R. Zhao, Z. Xiu, X. Yang, Y. Zhu, J. Han, et al., Neobavaisoflavone induces pyroptosis of liver cancer cells via Tom20 sensing the activated ROS signal, Phytomedicine : international journal of phytotherapy and phytopharmacology 116 (2023) 154869.
[14] H. Jun, Z. ZeXin, Screening of pyroptosis-related genes influencing the therapeutic effect of dehydroabietic acid in liver cancer and construction of a survival nomogram, Biochem. Biophys. Res. Commun. 585 (2021) 103–110.
[15] H. Chi, S. Zhao, J. Yang, X. Gao, G. Peng, J. Zhang, et al., T-cell exhaustion signatures characterize the immune landscape and predict HCC prognosis via integrating single-cell RNA-seq and bulk RNA-sequencing, Front. Immunol. 14 (2023) 1137025.
[16] A. Saviano, N.C. Henderson, T.F. Baumert, Single-cell genomics and spatial transcriptomics: discovery of novel cell states and cellular interactions in liver physiology and disease biology, J. Hepatol. 73 (5) (2020) 1219–1230.
[17] R. Gupta, Y. Schrooders, D. Hauser, M. van Herwijnen, W. Albrecht, Braak B. Ter, et al., Comparing in vitro human liver models to in vivo human liver using RNA-Seq, Arch. Toxicol. 95 (2) (2021) 573–589.
[18] A.K. Swain, P. Pandey, R. Sera, P. Yadav, Single-cell transcriptome analysis identifies novel biomarkers involved in major liver cancer subtypes, Funct. Integr. Genom. 23 (3) (2023) 235.

[19] Y. Li, Y. Li, X. Zhang, X. Duan, H. Feng, Z. Yu, et al., A novel association of pyroptosis-related gene signature with the prognosis of hepatocellular carcinoma, Front. Oncol. 12 (2022) 986827.
[20] M.D. Wilkerson, D.N. Hayes, ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking, Bioinformatics 26 (12) (2010) 1572–1573.
[21] M.E. Ritchie, B. Phipson, D. Wu, Y. Hu, C.W. Law, W. Shi, et al., Limma powers differential expression analyses for RNA-sequencing and microarray studies, Nucleic Acids Res. 43 (7) (2015) e47.
[22] J. Friedman, T. Hastie, R. Tibshirani, Regularization paths for generalized linear models via coordinate descent, J. Stat. Software 33 (1) (2010) 1–22.
[23] P. Blanche, J.F. Dartigues, H. Jacqmin-Gadda, Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks, Stat. Med. 32 (30) (2013) 5381–5397.
[24] R. Satija, J.A. Farrell, D. Gennert, A.F. Schier, A. Regev, Spatial reconstruction of single-cell gene expression data, Nat. Biotechnol. 33 (5) (2015) 495–502.
[25] K.J. Livak, T.D. Schmittgen, Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method, Methods (San Diego, Calif) 25 (4) (2001) 402–408.
[26] M. Deng, S. Sun, R. Zhao, R. Guan, Z. Zhang, S. Li, et al., The pyroptosis-related gene signature predicts prognosis and indicates immune activity in hepatocellular carcinoma, Mol. Med. 28 (1) (2022) 16.
[27] R. Tang, J. Xu, B. Zhang, J. Liu, C. Liang, J. Hua, et al., Ferroptosis, necroptosis, and pyroptosis in anticancer immunity, J. Hematol. Oncol. 13 (1) (2020) 110.
[28] S.K. Hsu, C.Y. Li, I.L. Lin, W.J. Syue, Y.F. Chen, K.C. Cheng, et al., Inflammation-related pyroptosis, a novel programmed cell death pathway, and its crosstalk with immune therapy in cancer treatment, Theranostics 11 (18) (2021) 8813–8835.
[29] C. Yan, Y. Niu, F. Li, W. Zhao, L. Ma, System analysis based on the pyroptosis-related genes identifies GSDMC as a novel therapy target for pancreatic adenocarcinoma, J. Transl. Med. 20 (1) (2022) 455.
[30] S.H. Guan, X.Y. Wang, P. Shang, Q.C. Du, M.Z. Li, X. Xing, et al., Pyroptosis-related genes play a significant role in the prognosis of gastric cancer, World journal of clinical cases 10 (24) (2022) 8490–8505.
[31] Y. Han, X. Wang, The emerging roles of KPNA2 in cancer, Life Sci. 241 (2020) 117140.
[32] C. Han, J. Chen, J. Huang, R. Zhu, J. Zeng, H. Yu, et al., Single-cell transcriptome analysis reveals the metabolic changes and the prognostic value of malignant hepatocyte subpopulations and predict new therapeutic agents for hepatocellular carcinoma, Front. Oncol. 13 (2023) 1104262.
[33] Y. Fu, X.D. Wei, L. Guo, K. Wu, J. Le, Y. Ma, et al., The metabolic and non-metabolic roles of UCK2 in tumor progression, Front. Oncol. 12 (2022) 904887.
[34] M. Jangal, B. Lebeau, M. Witcher, Beyond EZH2: is the polycomb protein CBX2 an emerging target for anti-cancer therapy? Expert Opin. Ther. Targets 23 (7) (2019) 565–578.
[35] Q. Meng, Q. Zhou, X. Chen, J. Chen, Prognostic hub gene CBX2 drives a cancer stem cell-like phenotype in HCC revealed by multi-omics and multi-cohorts, Aging 15 (22) (2023) 12817–12851.
[36] E. McGrath, D. Waschbüsch, B.M. Baker, A.R. Khan, LRRK2 binds to the Rab32 subfamily in a GTP-dependent manner via its armadillo domain, Small GTPases 12 (2) (2021) 133–146.
[37] Q. Li, J. Wang, Y. Wan, D. Chen, Depletion of Rab32 decreases intracellular lipid accumulation and induces lipolysis through enhancing ATGL expression in hepatocytes, Biochem. Biophys. Res. Commun. 471 (4) (2016) 492–496.
[38] Q. Wang, G. Wu, L. Fu, Z. Li, Y. Wu, T. Zhu, et al., Tumor-promoting roles of HMMR in lung adenocarcinoma, Mutation research 826 (2023) 111811.
[39] Y. Chen, Y. Ouyang, Z. Li, X. Wang, J. Ma, S100A8 and S100A9 in cancer, Biochim. Biophys. Acta, Rev. Cancer 1878 (3) (2023) 188891.
[40] M. Valiente, J.M. Sepúlveda, A. Pérez, Emerging targets for cancer treatment: S100A9/RAGE, ESMO open 8 (1) (2023) 100751.
[41] S. Wang, Y. Zhou, R. Yu, J. Ling, B. Li, C. Yang, et al., Loss of hepatic FTCD promotes lipid accumulation and hepatocarcinogenesis by upregulating PPARγ and SREBP2, JHEP reports : innovation in hepatology 5 (10) (2023) 100843.
[42] B. Li, T. Xu, C. Liu, G. Meng, Y. Sun, L. Qian, et al., Liver-enriched genes are associated with the prognosis of patients with hepatocellular carcinoma, Sci. Rep. 8 (1) (2018) 11197.
[43] C. Zhang, L. Peng, H. Gu, J. Wang, Y. Wang, Z. Xu, ANXA10 is a prognostic biomarker and suppressor of hepatocellular carcinoma: a bioinformatics analysis and experimental validation, Sci. Rep. 13 (1) (2023) 1583.
[44] C. Gong, J. Ma, Y. Deng, Q. Liu, Z. Zhan, H. Gan, et al., S100A9(-/-) alleviates LPS-induced acute lung injury by regulating M1 macrophage polarization and inhibiting pyroptosis via the TLR4/MyD88/NFκB signaling axis, Biomedicine & pharmacotherapy = Biomedecine & pharmacotherapie 172 (2024) 116233.
[45] J. Wang, P. Chen, G. Han, Y. Zhou, X. Xiang, M. Bian, et al., Rab32 facilitates Schwann cell pyroptosis in rats following peripheral nerve injury by elevating ROS levels, J. Transl. Med. 22 (1) (2024) 194.
[46] J. Ding, X. He, W. Luo, W. Zhou, R. Chen, G. Cao, et al., Development and validation of a pyroptosis-related signature for predicting prognosis in hepatocellular carcinoma, Front. Genet. 13 (2022) 801419.